

Electron-beam energy reconstruction for neutrino oscillation measurements

<https://doi.org/10.1038/s41586-021-04046-5>

Received: 29 June 2020

Accepted: 22 September 2021

Published online: 24 November 2021

 Check for updates

M. Khachatryan^{1,56}, A. Papadopoulou^{2,56}, A. Ashkenazi^{2,56}, F. Hauenstein^{1,2}, L. B. Weinstein¹, O. Hen², E. Piasetzky³, the CLAS Collaboration* & e4v Collaboration*

Neutrinos exist in one of three types or ‘flavours’—electron, muon and tau neutrinos—and oscillate from one flavour to another when propagating through space. This phenomena is one of the few that cannot be described using the standard model of particle physics (reviewed in ref. ¹), and so its experimental study can provide new insight into the nature of our Universe (reviewed in ref. ²). Neutrinos oscillate as a function of their propagation distance (L) divided by their energy (E). Therefore, experiments extract oscillation parameters by measuring their energy distribution at different locations. As accelerator-based oscillation experiments cannot directly measure E , the interpretation of these experiments relies heavily on phenomenological models of neutrino–nucleus interactions to infer E . Here we exploit the similarity of electron–nucleus and neutrino–nucleus interactions, and use electron scattering data with known beam energies to test energy reconstruction methods and interaction models. We find that even in simple interactions where no pions are detected, only a small fraction of events reconstruct to the correct incident energy. More importantly, widely used interaction models reproduce the reconstructed energy distribution only qualitatively and the quality of the reproduction varies strongly with beam energy. This shows both the need and the pathway to improve current models to meet the requirements of next-generation, high-precision experiments such as Hyper-Kamiokande (Japan)³ and DUNE (USA)⁴.

The three types of neutrinos are described in two different bases: flavour and mass. The weak nuclear interaction of neutrinos with other particles is described using flavour (ν_e , ν_μ and ν_τ), whereas their propagation through space is described using mass. Each flavour state is a linear combination of the three mass states (ν_1 , ν_2 and ν_3)¹.

In the simpler case of two neutrino flavours, the oscillation probability from ν_μ to ν_e is given by:⁵

$$P_{\nu_\mu \rightarrow \nu_e}(E, L) \approx \sin^2(2\theta) \sin^2\left(\frac{\Delta m^2 L}{4E}\right), \quad (1)$$

where $\Delta m^2 = m_{\nu_1}^2 - m_{\nu_2}^2$ is the neutrino mass difference squared that determines the oscillation wavelength as a function of L/E , and θ is the neutrino mixing angle that determines the oscillation amplitude.

Accelerator-based measurements produce beams that predominantly contain either ν_μ or $\bar{\nu}_\mu$. At a distance L from the neutrino production point some ν_μ will oscillate to ν_e , resulting in fluxes of approximately:

$$\Phi_e(E, L) \propto P_{\nu_\mu \rightarrow \nu_e}(E, L) \Phi_\mu(E, 0), \quad (2)$$

where the proportionality constant depends on the experiment geometry.

$\nu_\mu \rightarrow \nu_e$ oscillations are thus observed by measuring the neutrino fluxes $\Phi_e(E, L)$ and $\Phi_\mu(E, L)$ as a function of energy or distance. The three-flavour

oscillation equations are similar but include additional terms. Charge-parity (CP) symmetry violation in the leptonic sector would add a phase (δ_{CP}) to the three-flavour oscillation with an opposite sign for neutrinos and anti-neutrinos ($\bar{\nu}$)^{6,7} (see Methods for details).

Experimentally, the neutrino flux is extracted from the measured neutrino interaction rate with atomic nuclei in neutrino detectors. This interaction rate is given by:

$$N_\alpha(E_{\text{rec}}, L) \propto \sum_i \int \Phi_\alpha(E, L) \sigma_i(E) f_{\sigma_i}(E, E_{\text{rec}}) dE, \quad (3)$$

where $\sigma_i(E)$ is the neutrino interaction cross-section for process i (for example, quasi-elastic (QE) scattering, resonance production, and so on), α is the neutrino flavour, and E_{rec} is the neutrino energy reconstructed from the measured angles, momenta and/or energies of the detected particles.

$f_{\sigma_i}(E, E_{\text{rec}})$ is a smearing matrix relating the real (E) and reconstructed (E_{rec}) neutrino energies. E_{rec} differs from E owing to both experimental effects (for example, detector resolutions, inefficiencies, backgrounds) and also nuclear interaction effects (for example, nucleon (proton or neutron) motion, meson currents, nucleon reinteraction). Although experimental effects are generally understood, nuclear effects are irreducible and must be accounted for using theoretical models, typically implemented in neutrino event generators.

¹Old Dominion University, Norfolk, VA, USA. ²Massachusetts Institute of Technology, Cambridge, MA, USA. ³Tel Aviv University, Tel Aviv, Israel. ⁵⁶These authors contributed equally: M. Khachatryan and A. Papadopoulou. *Lists of authors and their affiliations appear at the end of the paper. [✉]e-mail: adishka@mit.edu

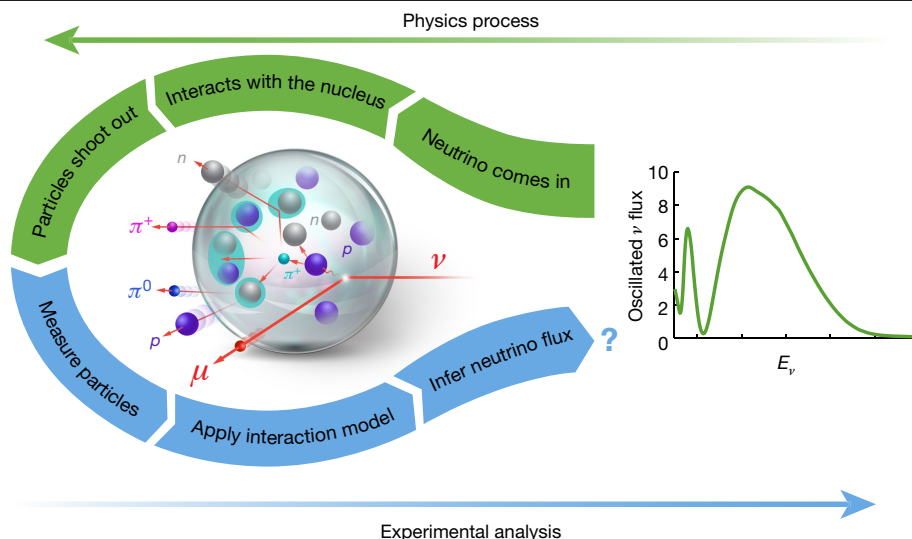


Fig. 1 | Neutrino oscillations and energy spectra measurements. Neutrino energy spectra reconstruction depends on our ability to model the interaction of neutrinos with nuclei and the propagation of particles through the nucleus. This flow chart shows the process, starting with an oscillated far-detector

incident-energy spectrum (green), differentiating the physical neutrino interactions (green arrows) from the experimental analysis (blue arrows), and ending up with an inferred incident-energy spectrum that provides an estimate for the real spectrum.

The precision to which oscillation parameters can be determined experimentally therefore depends on our ability to extract $\Phi_\alpha(E, L)$ from $N_\alpha(E_{\text{rec}}, L)$, see Fig. 1. This is largely determined by the accuracy of the theoretical models used to calculate $\sigma_i(E)$ and $f_{\sigma_i}(E, E_{\text{rec}})$. The models currently used have many free parameters that are poorly constrained and are ‘tuned’ by each neutrino experiment. Current oscillation experiments report substantial systematic uncertainties owing to these interaction models^{7–10} and simulations show that energy reconstruction errors can lead to large biases in extracting δ_{CP} at DUNE¹¹. There is a robust theoretical effort to improve these models^{12–14}.

Because there are no mono-energetic high-energy neutrino beams, these models cannot be tested for individual neutrino energies. Instead, experiments tune models of $\sigma_i(E)$ and $f_{\sigma_i}(E, E_{\text{rec}})$ to reproduce their near-detector data, where the unoscillated flux $\Phi(E, 0)$ is calculated from hadronic reaction rates^{15–17}.

Although highly informative, such integrated constraints are insufficient to ensure that the models are correct for each value of E . Thus, for precision measurements using a broad-energy neutrino beam, the degree to which the near-detector data alone can constrain models is unclear, since the neutrino flux can be very different at the far detector, owing to oscillations.

Here we report a measurement of $f_{\sigma_i}(E, E_{\text{rec}})$ for mono-energetic electron–nucleus scattering, and use it to test interaction models widely used by neutrino oscillation analyses. Both types of leptons, e and ν , interact similarly with nuclei, via a vector current; whereas neutrinos have an additional axial-vector current. The nuclear ground state is the same in both cases and many of the nuclear reaction effects are similar (see Methods for details). Therefore, any model of neutrino interactions (vector + axial-vector) should also be able to reproduce electron (vector) interactions. The data presented here can therefore test neutrino–nucleus interaction models to be used in analysis of neutrino oscillation measurements. Although previous work has compared these interaction models with inclusive electron scattering, (e, e') ,^{18,19} here we compare semi-exclusive electron scattering data (data with one or more detected hadrons) with these interaction models.

We examined events with one detected electron, one proton and zero pions ($1p0\pi$), which were expected to be dominated by well understood QE scattering (that is, by scattering of the lepton from a single moving nucleon in the nucleus). Even these simpler events reconstructed to the

correct energy less than 50% of the time, and the models used strongly overestimated the amount of mis-reconstructed events that are due to non-QE processes at the higher incident energies. This highlights a major shortcoming in our current understanding of neutrino interactions which, if not corrected, could limit the exploitation of the full potential of next-generation, high-precision oscillation experiments, namely DUNE and Hyper-Kamiokande.

Electron data selection

The experiment measured electron scattering from ^4He , ^{12}C and ^{56}Fe nuclei at beam energies of 1.159, 2.257 and 4.453 GeV, and detected the scattered electron and knocked-out particles over a wide range of angles and momenta in the CEBAF Large Acceptance Spectrometer (CLAS)²⁰ at the Thomas Jefferson National Accelerator Facility (Jefferson Lab). We detected electrons with energy $E_e \geq 0.4$, 0.55 and 1.1 GeV for $E_{\text{beam}} = 1.159$, 2.257 and 4.453 GeV, respectively, and angles $15^\circ \leq \theta_e \leq 45^\circ$; hadrons with momenta above 150 to 300 MeV/c and $10^\circ \leq \theta_h \leq 140^\circ$; and photons with energy $E_\gamma \geq 300$ MeV. These hadron detection thresholds are similar to those of neutrino detectors²¹; however, neutrino detectors have full angular coverage and lower lepton energy thresholds. See Methods for details concerning all aspects of this section.

The incident energies used here span the range of typical accelerator-based neutrino beams (Extended Data Fig. 1). The carbon data are relevant for scintillator-based experiments such as MINERvA and NOvA²² and similar to the oxygen in water-based Cherenkov detectors such as Super-Kamiokande²³ and Hyper-Kamiokande⁶⁰. The iron is similar to the argon in the liquid argon time projection chambers of MicroBoone²³, the Fermilab short-baseline oscillation program²⁴ and DUNE²⁵.

We selected events with one electron and zero pions or photons from π^0 decay above threshold. We did this to maximize the contribution of well understood events where the incident lepton scattered quasi-elastically from a single nucleon in the nucleus, as is done in many neutrino oscillation analyses^{1,26}.

Electrons, unlike neutrinos, radiate bremsstrahlung photons in the electric field of the nucleus. We vetoed events where the photons from scattered-electron radiation were detected in CLAS.

We subtracted from our data contributions from events where unwanted pions or photons were produced but not detected owing

to the incomplete CLAS angular acceptance (approximately 50%). We used events with a detected unwanted particle (for example, pion or extra proton), and for each event we constructed a ‘simulation’ where we rotated the unwanted-particle momentum around the (known) momentum transfer direction many times to determine the probability P of detecting similar events. We then subtracted those events from our dataset.

This produced an $(e, e')_{0\pi}$ dataset where events included any number of detected or undetected protons and neutrons as well as charged pions and photons below the CLAS detection threshold. We also separately examined the subset of events with exactly one detected proton—that is, $(e, e'p)_{1p0\pi}$ —subtracting contributions from events with additional undetected protons above threshold.

We divided the yield by the integrated beam charge and target thickness to get a normalized yield. We corrected the data for the CLAS acceptance and for bremsstrahlung radiation to determine the cross-section.

The electron–nucleon cross-section is much more forward peaked than the neutrino cross-section. We accounted for that by weighting each event by Q^4 .

We considered several major sources of systematic uncertainties, including the angular dependence of the pion-production cross-section (for the undetected-pion subtraction), the effects of fiducial cuts on undetected particle subtraction, photon identification cuts, the sector-to-sector variation of the cross-section, and the CLAS acceptance corrections. The normalization uncertainty was about 3% and total point-to-point systematic uncertainties ranged from 7% to 25%, with the largest uncertainties for the smallest cross-sections.

From neutrino to electron scattering

We compared our mono-energetic electron data to predictions of the GENIE²⁷ simulation, which is used by most neutrino experiments in the USA and has an electron-scattering version (e -GENIE) that was recently overhauled to be consistent with the neutrino version¹⁹. GENIE includes QE lepton scattering, interactions of the lepton with two nucleons exchanging a meson (meson exchange currents or MEC, often referred to as ‘2p2h’), resonance production in nuclei (RES), and ‘deep inelastic scattering’ (DIS, which also includes all non-resonant meson production), as well as rescattering (final-state interactions) of the outgoing hadrons. We compared two GENIE ‘tunes’: G2018, which reproduces measured neutrino-inclusive²⁸ and electron-inclusive cross-sections; and SuSAv2, which uses modern, theoretically inspired, recently implemented QE and MEC models (see Methods for details).

We generated events using e -GENIE, propagated the events through CLAS fiducial cuts and acceptance maps to determine which particles were detected, and smeared the momenta of these particles based on the known CLAS resolution. We then analysed the resulting simulated events using the same code as the data and compared the two (see Methods for details).

The inclusive electron–nucleus and neutrino–nucleus event distributions generated by e -GENIE (weighted by Q^4) and GENIE are very similar¹⁹. This bolsters the relevance of our electron study to neutrino interactions.

Incident energy reconstruction

There are two general approaches for reconstructing the incident neutrino energy, based on the particle detection capabilities of the neutrino detector.

Water Cherenkov detectors measure only charged leptons and pions. If the neutrino scattered quasi-elastically from a stationary nucleon in the nucleus, its energy can be reconstructed from the measured lepton as:

$$E_{QE} = \frac{2M_N \varepsilon + 2M_N E_l - m_l^2}{2(M_N - E_l + k_l \cos \theta_l)}, \quad (4)$$

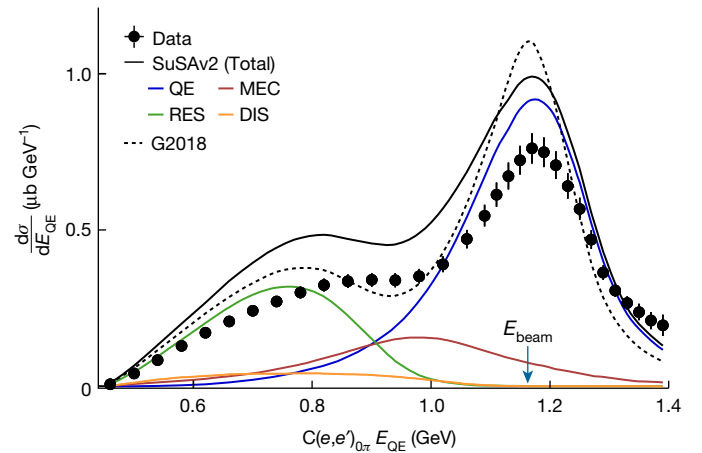


Fig. 2 | Quasi-elastic reconstructed energy. The 1.159-GeV $C(e, e')_{0\pi}$ cross-section plotted as a function of the reconstructed energy E_{QE} for data (black points), GENIE SuSAv2 (solid black curve) and GENIE G2018 (dotted black curve). The coloured lines show the contributions of different processes to the GENIE SuSAv2 cross-section: QE (blue), MEC (red), RES (green) and DIS (orange). Error bars show the 68% (1σ) confidence limits for the statistical and point-to-point systematic uncertainties added in quadrature. Error bars are not shown when they are smaller than the size of the data point. Normalization uncertainty of 3% not shown. μb , μbarn .

where ε is the average nucleon separation energy, M_N is the nucleon mass, and m_l , E_l , k_l and θ_l are, respectively, the mass, energy, momentum and angle of the scattered lepton.

Figure 2 shows the E_{QE} distribution for 1.159-GeV $C(e, e')_{0\pi}$ events, which are most relevant for Tokai-to-Kamioka (T2K) and Hyper-Kamiokande. We observe a broad peak centred at the real beam energy with a large tail extending to lower energies. The peak is Doppler-broadened by the motion of the nucleons in the nucleus. The tail is caused by non-QE reactions that pass the $(e, e')_{0\pi}$ selection. The tail is cut off at the lowest energies by the CLAS minimum detected electron energy of 0.4 GeV.

The SuSAv2 e -GENIE peak has the correct width, but is somewhat greater in magnitude than the data and overestimates the tail by about 25%. The G2018 e -GENIE peak also exceeds the data, but is too narrow, with a Gaussian width of $\sigma = 76$ MeV, compared to 89 MeV for the data. This is due to inexact modelling of the nuclear ground state momentum distribution. The tail dips below the data at around 0.9 GeV, and is larger than the data at lower reconstructed energies. Neither model describes the data quantitatively well.

Tracking detectors measure all charged particles above their detection thresholds. The calorimetric incident neutrino energy is then the sum of all the detected particle energies:

$$E_{\text{cal}} = \sum (E_i + \varepsilon_i) \quad (5)$$

where E_i are the detected nucleon kinetic energies and the lepton and meson total energies and ε is the average nucleon separation energy for the detected protons.

Figure 3 shows the cross-section as a function of E_{cal} for 1.159-, 2.257- and 4.453-GeV $C(e, e')_{1p0\pi}$ events and 2.257- and 4.453-GeV $\text{Fe}(e, e')_{1p0\pi}$ events. All spectra show a sharp peak at the real beam energy, followed by a large tail at lower energies. For carbon, only 30–40% of the events reconstruct to within 5% of the real beam energy; see Extended Data Table 1. For iron this fraction is only 20–25%, highlighting the crucial need to well model the low-energy tail of these distributions. e -GENIE overpredicts the fraction of events in the peak at 1.159 GeV and greatly underpredicts it at 4.453 GeV; see Extended Data Fig. 2a.

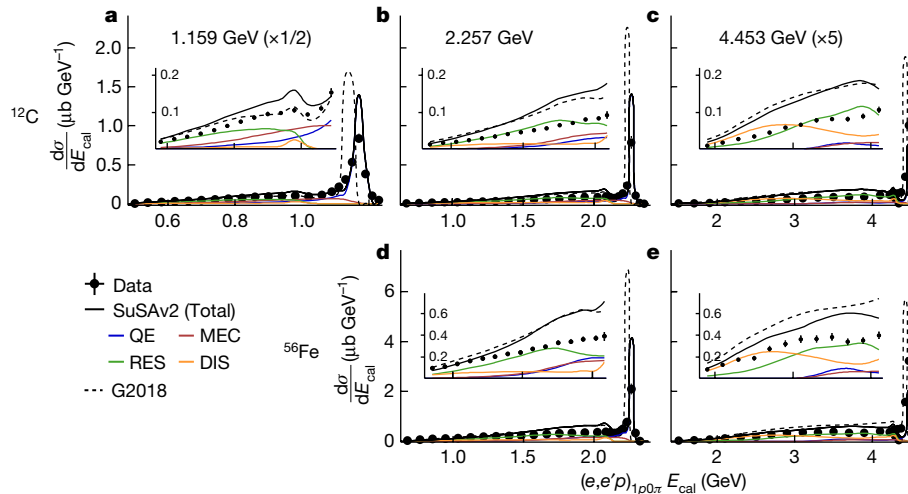


Fig. 3 | Calorimetric reconstructed energy. The $A(e, e'p)_{1p0\pi}$ cross-section plotted as a function of the reconstructed calorimetric energy E_{cal} for data (black points), SuSAv2 (black solid curve) and G2018 (black dotted curve). Different panels show results for different beam energy and target nucleus combinations: **a–c**, Carbon target at 1.159 GeV (**a**), 2.257 GeV (**b**) and 4.453 GeV

(**c**). **d, e**, Iron target at 2.257 GeV (**d**) and 4.453 GeV (**e**) incident beam.

The 1.159-GeV yields have been scaled by 1/2 and the 4.453-GeV yields have been scaled by 5 to have the same vertical scale. The insets show the cross-sections with the same horizontal scale and an expanded vertical scale. Line colours, error bars and uncertainties are the same as in Fig. 1.

e -GENIE using SuSAv2 greatly overpredicts the peak cross-section at 1.159 and 2.257 GeV, and substantially underestimates the peak cross-section at 4.453 GeV; see Extended Data Fig. 2b. e -GENIE using the older G2018 models overestimates the peak cross-section at all three incident energies. It also reconstructs the peak position (that is, the incident energy) to be 10, 25 and 36 MeV too low for ^4He , C and Fe, respectively, at all three beam energies. This is due to an error in the G2018 QE models.

This beam-energy dependence of the data–GENIE discrepancy could have significant implications for neutrino flux reconstruction.

At 1.159 GeV, e -GENIE using SuSAv2 slightly overpredicts the low-energy tail and e -GENIE using G2018 is reasonably close. Both models greatly overpredict the low-energy tail at the higher beam energies (see Fig. 3, insets). The tail seems to be dominated by resonance production (plus DIS at 4.453 GeV) that did not result in the production of other charged particles above detection threshold. This overprediction is also seen in inclusive electron scattering from the proton and deuteron, and thus appears to be due to the electron–nucleon interaction, rather than the result of nuclear modelling¹⁹ (see Methods for details).

SuSAv2 describes the peak cross-section equally well for C and for Fe, whereas G2018 overestimates the peak cross-section more for Fe than for C. Both models predict a greater peak fraction (relative to the data) for Fe than for C, particularly at 2.2 GeV; see Extended Data Fig. 2 and Extended Data Table 1.

Although the $(e, e')_{0\pi}$ QE reconstruction of equation (4) gives a much broader peak at the true beam energy than the calorimetric energy E_{cal} owing to the effects of nucleon motion (see Extended Data Fig. 4), it has the same tail of lower-energy events for the same $(e, e')_{1p0\pi}$ dataset.

Transverse variables and model tuning

Neutrino experiments use ‘single transverse variables’ (STVs) to enhance their sensitivity to different aspects of the reaction mechanism. These STVs are independent of the incident neutrino energy^{29–31},

$$\mathbf{P}_T = \mathbf{P}_T^{e'} + \mathbf{P}_T^p \quad (6)$$

$$\delta\alpha_T = \arccos \frac{\mathbf{P}_T^{e'} \cdot \mathbf{P}_T}{P_T^{e'} P_T} \quad (7)$$

$$\delta\phi_T = \arccos \frac{\mathbf{P}_T^{e'} \cdot \mathbf{P}_T^p}{P_T^{e'} P_T^p} \quad (8)$$

where $\mathbf{P}_T^{e'}$ and $\mathbf{P}_T^p$ are the three-momenta of the detected lepton and proton perpendicular to the direction of the incident lepton, respectively, $P_T^{e'}$ and P_T^p are their respective magnitudes, and $P_T = |\mathbf{P}_T|$. Purely QE events without final-state interactions will have small P_T , consistent with the motion of the struck nucleon. Events with small P_T should thus reconstruct to the correct incident energy. $\delta\alpha_T$ measures the angle between $\mathbf{P}_T$ and the transverse momentum transfer ($\mathbf{q}_T = -\mathbf{P}_T^{e'}$) in the transverse plane and is isotropic in the absence of final-state interactions. $\delta\phi_T$ measures the opening angle between the detected proton momentum and the transverse momentum transfer and is forward-peaked. $\delta\phi_T$ is intended to probe regions where MEC/ $2p2h$ events dominate^{29–31}.

The P_T distribution for 2.257-GeV C($e, e'p$)_{1p0 π} is shown in Fig. 4 (and the other targets and energies are shown in Extended Data Fig. 5). Both data and e -GENIE peak at relatively low momenta, as expected, and both have a large tail extending out to 1 GeV/c and containing about half of the measured events. The high- P_T tail is predominantly due to resonance production that did not result in an additional pion or nucleon above the detection threshold. e -GENIE using SuSAv2 reproduces the shape of the data moderately well, suggesting adequate reaction modelling, including the contribution of non-QE processes such as resonance production.

As expected, both data and e -GENIE/SuSAv2 events with $P_T < 200$ MeV/c almost all reconstruct to the correct incident energy. However, events with $P_T \geq 400$ MeV/c do not reconstruct to the correct energy and are poorly reproduced by e -GENIE.

This disagreement indicates that including high- P_T data in oscillation analyses could bias the extracted parameters. High- P_T data account for 25–50% of the measured events, and so care must be taken to improve the models implemented in GENIE, so that they can reproduce the high- P_T data. This will be especially true at the higher incident neutrino energies expected for DUNE.

The $\delta\alpha_T$ distributions become progressively less isotropic at higher energies and heavier targets, indicating the increasing importance of FSI and of non-QE reaction mechanisms. GENIE agrees best with data at the lowest beam energy. At the higher beam energies GENIE describes the relatively flat smaller angles much better than the back-angle peak. GENIE also describes the lowest-energy $\delta\phi_T$ distribution. At higher

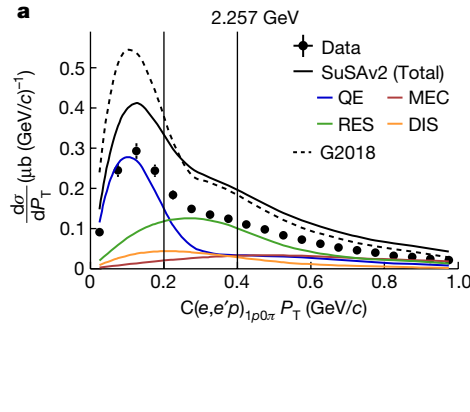


Fig. 4 | Reconstructed energies and perpendicular momenta. **a**, The 2.257-GeV $C(e, e'p)_{1p0\pi}$ cross-section plotted P_T for data (black points), SuSAv2 (black solid line) and G2018 (black dashed line). The vertical lines at 200 MeV/c and 400 MeV/c

energies, GENIE overestimates the height of the forward peak (see Extended Data Fig. 6).

Summary

In conclusion, we have used Jefferson Lab CLAS electron–nucleus scattering data with known incident energies to perform a test of our ability to reconstruct incident neutrino energies from measured lepton–nucleus collisions. This is a crucial step in the neutrino oscillation experiment analysis chain.

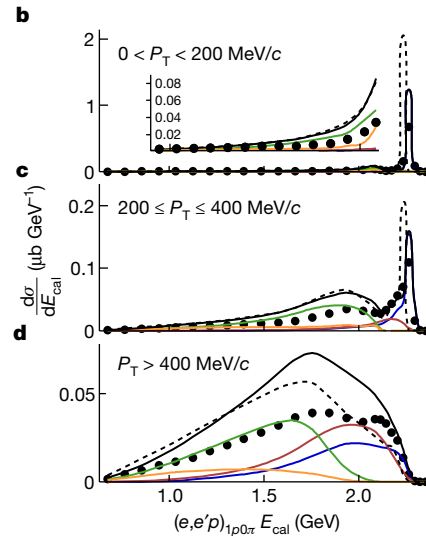
Most of the $1p0\pi$ events do not reconstruct to the correct incident energy. The interaction model describes the size but not the exact shape of the low-energy tail for QE energy reconstruction at 1.159 GeV, despite reproducing differential inclusive electron–scattering cross-sections. The same interaction model greatly overestimates the low-energy tail for calorimetric energy reconstruction at 2.257 and 4.453 GeV electrons. The more modern SuSAv2 model describes the QE events better than the older G2018 model, but there are still substantial discrepancies.

As we enter a precision era of neutrino studies, it is critical to improve the neutrino models to the same level of accuracy and precision. Electron scattering offers a way to test aspects of these models that complements theoretical calculations and neutrino near-detector data.

Combining the neutrino energy reconstruction studies presented here with the standard near-detector neutrino data analyses could greatly reduce the systematic modelling uncertainties of next-generation oscillation experiments. This could be complemented by efforts such as DUNE-PRISM, which would provide quasi-monochromatic neutrino beams by making linear combinations of off-axis neutrino fluxes. DUNE-PRISM would be sensitive to axial currents and would have very different systematic uncertainties than electron measurements. Future experiments with the improved CLAS12 spectrometer (see Methods for details) will extend the electron measurements to more nuclei, much smaller scattering angles, and to a wider range of momentum transfers. All of these will be needed to reduce the systematic modelling uncertainties of next-generation oscillation experiments.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information,



separate the three bins in P_T . **b–d**, The cross-section plotted versus E_{cal} for: $P_T < 200$ MeV/c (**b**), $200 \text{ MeV/c} \leq P_T \leq 400 \text{ MeV/c}$ (**c**) and $P_T > 400 \text{ MeV/c}$ (**d**). Line colours, error bars and uncertainties are the same as in Fig. 2.

acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-021-04046-5>.

- Particle Data Group. Review of particle physics. *Phys. Rev. D* **98**, 030001 (2018).
- Mohapatra, R. N. et al. Theory of neutrinos: a white paper. *Rep. Prog. Phys.* **70**, 1757–1867 (2007).
- Hyper-Kamiokande Proto-Collaboration. Hyper-Kamiokande design report. Preprint at <https://arxiv.org/abs/1805.04163> (2018).
- DUNE Collaboration. The DUNE Far Detector interim design report. Volume 1: physics, technology and strategies. Preprint at <https://arxiv.org/abs/1807.10334> (2018).
- Gonzalez-Garcia, M. C. & Nir, Y. Neutrino masses and mixing: evidence and implications. *Rev. Mod. Phys.* **75**, 345–402 (2003).
- Fukugita, M. & Yanagida, T. Baryogenesis without grand unification. *Phys. Lett. B* **174**, 45–47 (1986).
- The T2K Collaboration. Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations. *Nature* **580**, 339–344 (2020).
- T2K Collaboration. Search for CP violation in neutrino and antineutrino oscillations by the T2K Experiment with 2.2×10^{21} protons on target. *Phys. Rev. Lett.* **121**, 171802 (2018).
- Alvarez-Ruso, L. et al. NuSTEC White Paper: status and challenges of neutrino nucleus scattering. *Prog. Part. Nucl. Phys.* **100**, 1–68 (2018).
- NOvA Collaboration. New constraints on oscillation parameters from ν_e appearance and ν_μ disappearance in the NOvA experiment. *Phys. Rev. D* **98**, 032012 (2018).
- Ankowski, A. et al. Missing energy and the measurement of the CP-violating phase in neutrino oscillations. *Phys. Rev. D* **92**, 091301 (2015).
- Rocco, N. Ab initio calculations of lepton–nucleus scattering. *Front. Phys.* **8**, 00116 (2020).
- Dolan, S., Megias, G. D. & Bolognesi, S. Implementation of the SuSAv2-meson exchange current 1p1h and 2p2h models in GENIE and analysis of nuclear effects in T2K measurements. *Phys. Rev. D* **101**, 033003 (2020).
- Rocco, N., Lovato, A. & Benhar, O. Unified description of electron–nucleus scattering within the spectral function formalism. *Phys. Rev. Lett.* **116**, 192501 (2016).
- MINERvA Collaboration. Neutrino flux predictions for the NuMI beam. *Phys. Rev. D* **94**, 092005 (2016).
- Maan, K. K. on behalf of the NOvA Collaboration. Constraints on the neutrino flux in NOvA using the near-detector data. In *Proc. Sci. 38th Intl Conf. High Energy Physics (ICHEP2016)* Vol. 282 931 (Sissa, 2016).
- T2K Collaboration. T2K near detector constraints for oscillation results. In *18th Intl Workshop on Neutrino Factories and Future Neutrino Facilities Search 2* (2017); <https://arxiv.org/abs/1701.02559>
- Ankowski, A. & Friedland, A. Assessing the accuracy of the GENIE event generator with electron–scattering data. *Phys. Rev. D* **102**, 053001 (2020).
- Papadopolou, A. et al. Inclusive electron scattering and the GENIE neutrino event generator. *Phys. Rev. D* **103**, 113003 (2021).
- Mecking, B. A. et al. The CEBAF large acceptance spectrometer (CLAS). *Nucl. Instrum. Meth. A* **503**, 513–553 (2003).
- MINERvA Collaboration. Direct measurement of nuclear dependence of charged current quasielastic-like neutrino interactions using MINERvA. *Phys. Rev. Lett.* **119**, 082001 (2017).
- Aliaga, L. et al. Design, calibration, and performance of the MINERvA detector. *Nucl. Instrum. Meth. A* **743** 130–159 (2014).
- Acciarri, R. et al. Design and construction of the MicroBooNE detector. *J. Instrum.* **12**, P02017 (2017).

24. The ICARUS-WA104 Collaboration, The LAr1-ND Collaboration, The MicroBooNE Collaboration & additional Fermilab contributors. A proposal for a three detector short-baseline neutrino oscillation program in the Fermilab Booster Neutrino Beam. Preprint at <https://arxiv.org/abs/1503.01520>

25. The DUNE Collaboration. Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) conceptual design report. Volume 2: The physics program for DUNE at LBNF. Preprint at <https://arxiv.org/abs/1512.06148> (2016).

26. Katori, T. & Martini, M. Neutrino–nucleus cross sections for oscillation experiments. *J. Phys. G* **45**, 013001 (2018).

27. Andreopoulos, C. et al. The GENIE neutrino Monte Carlo generator. *Nucl. Instrum. Meth. A* **614**, 87–104 (2010).

28. MicroBooNE Collaboration. First measurement of inclusive muon neutrino charged current differential cross sections on argon at $E_\nu \sim 0.8$ GeV with the MicroBooNE detector. *Phys. Rev. Lett.* **123**, 131801 (2019).

29. Lu, X.-G. et al. Measurement of nuclear effects in neutrino interactions with minimal dependence on neutrino energy. *Phys. Rev. C* **94**, 015503 (2016).

30. The T2K Collaboration. Characterization of nuclear effects in muon–neutrino scattering on hydrocarbon with a measurement of final-state kinematics and correlations in charged-current pionless interactions at T2K. *Phys. Rev. D* **98**, 032003 (2018).

31. MINERvA Collaboration. Measurement of final-state correlations in neutrino muon–proton mesonless production on hydrocarbon at $\langle E_\nu \rangle = 3$ GeV. *Phys. Rev. Lett.* **121**, 022504 (2018).

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021

the CLAS Collaboration

M. Khachatryan^{1,53}, A. Papadopolou^{2,53}, A. Ashkenazi², F. Hauenstein^{1,2}, A. Nambrath², A. Hrnjic², L. B. Weinstein¹, O. Hen², E. Piasetzky³, M. Betancourt⁴, S. Dytman⁵, K. Mahn⁶, P. Coloma^{7,8}, S. Adhikari⁹, M. J. Amarian¹, Giovanni Angelini¹⁰, H. Atac¹¹, L. Barion¹², M. Battaglieri^{13,14}, I. Bedlinskiy¹⁵, A. Beck², F. Benmokhtar¹⁶, A. Bianconi^{17,18}, A. S. Biselli^{19,20}, F. Bossu²¹, S. Boiarinov¹³, W. J. Briscoe¹⁰, W. K. Brooks^{22,23}, D. Bulumulla¹, V. D. Burkert¹³, D. S. Carman¹³, J. C. Carvajal⁹, A. Celentano¹⁴, P. Chatagnon²³, V. Chesnokov²⁴, T. Chetry²⁵, G. Ciullo^{12,26}, L. Clark²⁷, B. A. Clary²⁸, E. O. Cohen³, P. L. Cole^{29,30}, M. Contalbrigo¹², V. Crede³¹, R. Cruz-Torres², A. D’Angelo^{32,33}, N. Dashyan³⁴, R. De Vita¹⁴, M. Defurne²¹, A. Denniston², A. Deur¹³, S. Diehl²⁸, C. Djalali^{35,36}, M. Duer³, R. Dupre²³, H. Egiyan¹³, M. Ehrhart³⁷, A. El Alaoui²², L. El Fassi²⁵, L. Elouadrhiri¹³, P. Eugenio³, R. Fersch³⁸, A. Filippi³⁹, T. A. Forest³⁰, G. Gavalian^{13,40}, G. P. Gilfoyle⁴¹, K. L. Giovanetti⁴², F. X. Girod¹³, D. I. Glazier²⁷, E. Golovatch²⁴, R. W. Gothe³⁶, K. A. Griffioen⁴³, M. Guidal²³, L. Guo^{9,13}, H. Hakobyan^{22,34}, M. Hattawy¹, T. B. Hayward⁴³, D. Heddle^{13,38}, K. Hicks³⁵, A. Hobart²³, M. Holtrop³⁰, Y. Ilieva^{10,36}, D. G. Ireland²⁷, B. S. Ishkhanov²⁴, E. L. Isupov²⁴, H. S. Jo⁴⁴, K. Joo³⁸, D. Keller⁴⁵, A. Khanal⁹, M. Khandaker⁴⁶, C. W. Kim¹⁰, W. Kim⁴⁴, I. Korover², V. Kubarovskiy^{13,47}, S. E. Kuhn¹, L. Lanza³², M. Leali^{17,18}, P. Lenisa^{12,26}, K. Livingston²⁷, I. J. D. MacGregor²⁷, D. Marchand²³, N. Markov¹³,

L. Marsicano¹⁴, V. Mascagna¹⁷, B. McKinnon²⁷, S. May-Tal Beck², T. Mineeva²², M. Mirazita⁴⁸, V. Mokeev^{13,24}, A. Movsisyan¹², C. Munoz Camacho²³, P. Nadel-Turonski¹³, K. Neupane³⁶, S. Niccolai²³, G. Niculescu⁴², M. Osipenko¹⁴, A. I. Ostrovidov³¹, M. Paolone¹¹, L. L. Pappalardo^{12,26}, R. Paremuzyan⁴⁰, K. Park⁴⁴, E. Pasyuk¹³, O. Pogorelko¹⁵, J. Poudel¹, Y. Prok^{1,45}, D. Protopopescu²⁷, J. Pybus², M. Ripani¹⁴, B. Raue⁹, J. Ritman⁴⁹, A. Rizzo^{32,33}, G. Rosner²⁷, P. Rossi^{15,48}, F. Sabati²⁰, C. Salgado⁴⁶, A. Schmidt¹⁰, R. A. Schumacher²⁰, E. P. Segarra², Y. G. Sharabian¹³, U. Shrestha³⁵, Iu. Skorodumina^{24,36}, D. Sokhan²⁷, O. Soto⁴⁸, N. Sparveris¹¹, S. Stepanyan¹¹³, I. I. Strakovsky¹⁰, S. Strauch^{10,36}, N. Tyler³⁶, R. Tyson²⁷, M. Ungaro^{13,47}, L. Venturelli^{17,18}, H. Voskanyan³⁴, E. Voutier²³, D. Watts⁵⁰, X. Wei¹³, M. H. Wood^{36,51}, N. Zachariou⁵⁰, J. Zhang⁴⁵ & Z. W. Zhao⁵²

e4v Collaboration

M. Khachatryan^{1,56}, A. Papadopolou^{2,56}, A. Ashkenazi², F. Hauenstein^{1,2}, A. Nambrath², A. Hrnjic², L. B. Weinstein¹, O. Hen², E. Piasetzky³, M. Betancourt⁴, S. Dytman⁵, K. Mahn⁶, P. Coloma^{7,8}, S. Dolan⁵³, G. D. Megias⁵⁴ & S. Gardiner⁵⁵

⁴Fermi National Accelerator Laboratory, Batavia, IL, USA. ⁵University of Pittsburgh, Pittsburgh, PA, USA. ⁶Michigan State University, East Lansing, MI, USA. ⁷Instituto de Física Corpuscular, Universidad de Valencia and CSIC, Valencia, Spain. ⁸Universidad Autonoma de Madrid, Madrid, Spain. ⁹Florida International University, Miami, Florida, USA. ¹⁰The George Washington University, Washington, DC, USA. ¹¹Temple University, Philadelphia, PA, USA. ¹²INFN, Ferrara, Italy. ¹³Thomas Jefferson National Accelerator Facility, Newport News, VA, USA. ¹⁴INFN, Genova, Italy. ¹⁵National Research Centre Kurchatov Institute - ITEP, Moscow, Russia. ¹⁶Duquesne University, Pittsburgh, PA, USA. ¹⁷Università degli Studi di Brescia, Brescia, Italy. ¹⁸INFN, Pavia, Italy. ¹⁹Fairfield University, Fairfield, CT, USA. ²⁰Carnegie Mellon University, Pittsburgh, PA, USA. ²¹University of Paris-Saclay, Gif-sur-Yvette, France. ²²Universidad Tecnica Federico Santa Maria, Valpara, Chile. ²³University Paris-Saclay, Orsay, France. ²⁴Lomonosov Moscow State University, Moscow, Russia. ²⁵Mississippi State University, Mississippi State, MS, USA. ²⁶University of Ferrara, Ferrara, Italy. ²⁷University of Glasgow, Glasgow, United Kingdom. ²⁸University of Connecticut, Storrs, CT, USA. ²⁹Lamar University, Beaumont, TX, USA. ³⁰Idaho State University, Pocatello, ID, USA. ³¹Florida State University, Tallahassee, FL, USA. ³²INFN, Rome, Italy. ³³Università di Roma Tor Vergata, Rome, Italy. ³⁴Yerevan Physics Institute, Yerevan, Armenia. ³⁵Ohio University, Athens, OH, USA. ³⁶University of South Carolina, Columbia, SC, USA. ³⁷Argonne National Laboratory, Argonne, IL, USA. ³⁸Christopher Newport University, Newport News, VA, USA. ³⁹INFN, Sezione di Torino, Torino, Italy. ⁴⁰University of New Hampshire, Durham, NH, USA. ⁴¹University of Richmond, Richmond, VA, USA. ⁴²James Madison University, Harrisonburg, VA, USA. ⁴³College of William and Mary, Williamsburg, VA, USA. ⁴⁴Kyungpook National University, Daegu, Republic of Korea. ⁴⁵University of Virginia, Charlottesville, VA, USA. ⁴⁶Norfolk State University, Norfolk, VA, USA. ⁴⁷Rensselaer Polytechnic Institute, Troy, NY, USA. ⁴⁸INFN, Laboratori Nazionali di Frascati, Frascati, Italy. ⁴⁹Institute für Kernphysik (Juelich), Juelich, Germany. ⁵⁰University of York, York, UK. ⁵¹Canisius College, Buffalo, NY, USA. ⁵²Duke University, Durham, NC, USA. ⁵³CERN, European Organization for Nuclear Research, Geneva, Switzerland. ⁵⁴Research Center for Cosmic Neutrinos, Institute for Cosmic Ray Research, University of Tokyo, Kashiwa, Chiba, Japan. ⁵⁵Fermi National Accelerator Laboratory, Batavia, IL, USA.

Methods

Three-flavour neutrino oscillation

Equation (1) gives the neutrino oscillation probability for the simplified case of only two types of neutrino. The full three-flavour probability for $\nu_\mu \rightarrow \nu_e$ oscillation (in vacuum) is given by^{32–34}:

$$P_{\nu_\mu \rightarrow \nu_e}(E, L) \approx A \sin^2 \frac{\Delta m_{13}^2 L}{4E} - B \cos \left(\frac{\Delta m_{13}^2 L}{4E} + \delta_{CP} \right) \sin \frac{\Delta m_{13}^2 L}{4E}, \quad (9)$$

where $\Delta m_{13}^2 = m_{\nu_1}^2 - m_{\nu_3}^2$ is the neutrino mass difference squared, which determines the oscillation wavelength as a function of L/E , and δ_{CP} is a phase that might break CP symmetry. The coefficients A and B depend primarily on the neutrino oscillation mixing angles, $A = \sin^2 \theta_{23} \sin^2 2\theta_{13}$ and $B = -\frac{\sin 2\theta_{12} \sin 2\theta_{23}}{2 \sin \theta_{13}} \sin \frac{\Delta m_{21}^2 L}{4E} \sin^2 2\theta_{13}$. The different flavour neutrinos (labelled ν_e , ν_μ and ν_τ) are linear combinations of the different mass neutrinos labelled 1, 2 and 3.

Lepton cross-sections

Neutrinos and electrons interact with atomic nuclei by exchanging intermediate vector bosons, a massive $W^\pm$ or Z for the neutrino and a massless photon for the electron. Electrons interact via a vector current ($j_{EM}^\mu = \bar{u} \gamma^\mu u$) and neutrinos interact via vector and axial-vector ($j_{CC}^\mu = \bar{u} \gamma^\mu (1 - \gamma^5) u \frac{-ig_W}{2\sqrt{2}}$) currents, where u and $\bar{u}$ are representations of spin-1/2 particle and anti-particles, respectively.

Fundamental considerations (the number of independent momenta, Lorentz invariance, parity conservation, and current conservation) give an inclusive (e, e') electron–nucleon elastic scattering cross-section that depends on only two structure functions:

$$\frac{d^2\sigma^e}{dx dQ^2} = \frac{4\pi\alpha^2}{Q^4} \left[\frac{1-y}{x} F_2^e(x, Q^2) + y^2 F_1^e(x, Q^2) \right]. \quad (10)$$

Here F_1^e and F_2^e are the standard electromagnetic vector structure functions, $Q^2 = \mathbf{q}^2 - \nu^2$ is the squared momentum transfer and $\mathbf{q}$ and ν are the three-momentum and energy transfers, $x = Q^2/(2m\nu)$ is the Bjorken scaling variable, m is the nucleon mass, $y = \nu/E_e$ is the electron fractional energy loss, and α is the fine structure constant. This formula shows the simplest case where $Q^2 \gg m^2$.

The corresponding inclusive charged current (CC) ($\nu, \ell^\pm$) neutrino–nucleon elastic cross-section (where $\ell^\pm$ is the outgoing charged lepton) has a similar form. The vector part of the current is subject to the same fundamental considerations as above, but the axial-vector part of the current does not conserve parity. This leads to a third, axial, structure function:

$$\frac{d^2\sigma^\nu}{dx dQ^2} = \frac{G_F^2}{2\pi} \left[\frac{1-y}{x} F_2^\nu(x, Q^2) + y^2 F_1^\nu(x, Q^2) - y(1-y/2) F_3^\nu(x, Q^2) \right]. \quad (11)$$

Here F_1^ν and F_2^ν are the parity-conserving neutrino–nucleus (vector) structure functions, F_3^ν is the axial structure function, and G_F is the Fermi constant. The vector form factors (F_1^ν and F_2^ν) have both vector–vector and axial–axial contributions.

The strength of the interaction is very different: $4\pi\alpha^2/Q^4$ for electrons versus $G_F^2/(2\pi)$ for neutrinos, where the factor of $1/Q^4$ in the electron cross-section is due to the photon propagator. When compensated for the factor of $1/Q^4$, the shapes of the electron–nucleus and neutrino–nucleus cross-sections, as calculated in GENIE, are very similar¹⁹.

Nuclear medium effects such as nucleon motion, binding energy, two-body currents and final-state interactions will be similar or identical.

To calculate the electron–nucleus cross-section in GENIE, we changed the interaction strength, set the axial structure functions to zero, and set the axial–axial parts of the vector structure functions to zero.

Thus, models of the neutrino–nucleus cross-section must be able to describe the more limited electron–nucleus cross-section.

Experimental set-up and particle identification

CLAS used a toroidal magnetic field with six sectors of drift chambers, scintillation counters, Cherenkov counters and electromagnetic calorimeters to identify electrons, pions, protons and photons, and to reconstruct their trajectories²⁰. The six sectors functioned as six independent spectrometers (see Extended Data Fig. 9a).

We used the e2a data, which was measured in 1999 and was used in many published analyses^{35–40}. We measured the momentum and charge of the outgoing charged particles from their measured positions in the drift chambers and the curvature of their trajectories in the magnetic field. We identified electrons by requiring that the track originated in the target, produced a time-correlated signal in the Cherenkov counter, and deposited enough energy in the electromagnetic calorimeter. We identified charged pions and protons by requiring that the track originated in the target and that the measured time of flight agreed (to within $\pm$ three times the standard deviation of the detector resolution) with that calculated from the particle's momentum and assumed mass. We identified photons by requiring a signal in the electromagnetic calorimeter that implied a velocity greater than about 0.96c (see ref. ⁴¹ for details).

We corrected low-momentum protons for energy losses traversing the target and detector material. We used the CLAS GEANT Monte Carlo simulation to simulate the proton energy loss in CLAS as a function of proton momentum. The maximum correction was about 20 MeV/c for a proton momentum of 300 MeV/c. The correction was negligible for protons with momenta greater than 600 MeV/c.

We detected protons with momenta $p_p \geq 300$ MeV/c and angles $\theta_p \geq 10^\circ$, charged pions with momenta $p_\pi \geq 150$ MeV/c and angles $\theta_{\pi^+} \geq 10^\circ$ and $\theta_{\pi^-} \geq 22^\circ$, and photons with energy $E_\gamma \geq 300$ MeV and $8^\circ \leq \theta_\gamma \leq 45^\circ$. We applied separate fiducial cuts for electrons, π^- , positive particles and photons to select momentum-dependent regions of CLAS where the detection efficiency was constant and close to one. We also determined the minimum electron angle (as a function of electron momentum p) for each beam energy as:

$$\theta_e^{1.1} \geq 17^\circ + \frac{7^\circ}{p [\text{GeV}]}, \quad (12)$$

$$\theta_e^{2.2} \geq 16^\circ + \frac{10.5^\circ}{p [\text{GeV}]}, \quad (13)$$

$$\theta_e^{4.4} \geq 13.5^\circ + \frac{15^\circ}{p [\text{GeV}]}, \quad (14)$$

and the minimum π^- angle as

$$\theta_{\pi^-}^{1.1} \geq 17^\circ + \frac{4^\circ}{p [\text{GeV}]},$$

and

$$\theta_{\pi^-}^{2.2,4.4} \geq 25^\circ + \frac{7^\circ}{p [\text{GeV}]},$$

for $p_{\pi^-} < 0.35$ GeV/c and

$$\theta_{\pi^-}^{2.2,4.4} \geq 16^\circ + \frac{10^\circ}{p [\text{GeV}]}$$

for $p_{\pi^-} \geq 0.35$ GeV/c. The minimum π^+ and proton angle was $\theta > 12^\circ$ for all datasets and momenta.

We measured the delivered beam charge using the CLAS Faraday Cup.

Energy calibration

The beam energy was equal to the injector energy plus the pass number times the linac energy. The three-pass beam energy was measured using the Hall A arc measurement and the four-pass energy was measured using the Hall C arc measurement. These gave a central linac energy of 1.0979 GeV and Hall B one-, two- and four-pass beam energies of 1.159, 2.257 and 4.453 GeV, respectively. We assigned an uncertainty of 2×10^{-3} to these energies, on the basis of the difference between the Hall A and Hall C measurements.

We used elastic electron scattering from hydrogen to correct the electron momentum as a function of angle for uncertainties in the CLAS magnetic field and in CLAS tracking chamber locations. These corrections also greatly narrowed the elastic peak width. Typical correction factors were less than 1%. We checked the momentum correction factors at lower scattered electron energies using the $H(e, e'\pi^+)X$ and $^3\text{He}(e, e'pp)X$ reactions and found that they gave the correct missing mass for the undetected neutron (see Extended Data Fig. 9).

Nucleon removal energies

The average nucleon removal energy, ε , used in reconstructing the incident energies in equations (4), (5) was determined from the data. We used $\varepsilon = M_A - M_{A-1} - m_p + \Delta\varepsilon$, where $M_A - M_{A-1} - m_p$ is the difference in the binding energies for knocking a proton out of nucleus A. We adjusted $\Delta\varepsilon$ so that the peaks in the E_{cal} spectrum for low- P_T events reconstructed to the correct beam energy. We found $\Delta\varepsilon = 5$ and 11 MeV for ^{12}C and ^{56}Fe , respectively, which are consistent with average excitation energies from single-nucleon knockout from nuclei.

Subtraction of undetected pions and photons

Because the CLAS geometrical coverage is incomplete, we needed to subtract for undetected pions and photons to achieve a true 0π event sample. We assumed that the photons came from either radiation by the outgoing electron approximately parallel to its motion or from π^0 decay. We identified the radiated photons by requiring that they be detected within $\delta\phi \leq 30^\circ$ and $\delta\theta \leq 40^\circ$ of the scattered electron and removed them from the dataset.

We determined the undetected pion or photon contribution from the events with detected pions or photons. We assumed that the pion-production cross-section was independent of $\phi_{q\pi}$, the angle between the electron-scattering plane (the plane containing the incident and scattered electrons and the virtual photon) and the hadron plane (the plane containing the virtual photon and pion). For each detected $(e, e'\pi)$ event, we rotated the pion around the momentum transfer direction $\mathbf{q}$ randomly many times. For each rotation we determined if the particle would have been detected, that is, if the particle was within the fiducial region of the detector. If it was, we used acceptance maps to determine the probability that it would have been detected. The particle acceptance is then $A_\pi = N_{\text{det}}/N_{\text{rot}}$, where N_{rot} is the number of rotations and N_{det} is the number of times the pion would have been detected. Then the corresponding number of undetected $(e, e'\pi)$ events for that detected $(e, e'\pi)$ event is $(N_{\text{rot}} - N_{\text{det}})/N_{\text{det}}$. We used that as a weight to subtract for the undetected pion events.

For example, if one specific $(e, e'\pi)$ event would have been detected 250 times out of 1,000 rotations, then we inferred that for each event we detected, there were three more that we did not detect. We calculated the reconstructed energy (and other appropriate variables) for that event and subtracted it from the reconstructed energy spectrum (and other corresponding distributions) with a weight of three. We did this separately for π^+ , π^- and photons.

To subtract the undetected $(e, e'p\pi)$ and $(e, e'p\gamma)$ events to get the $(e, e'p)_{0\pi}$ sample, we rotated the proton and pion (or photon) together around $\mathbf{q}$ and determined the number of detected proton-only events N_{det}^p and the number of detected proton and pion events $N_{\text{det}}^{p\pi}$. We used $(N_{\text{det}}^p - N_{\text{det}}^{p\pi})/N_{\text{det}}^{p\pi}$ as a weight to subtract for the undetected $(e, e'p\pi)$ and $(e, e'p\gamma)$ events. We

also subtracted the $(e, e'p)$ event sample for extra protons in the same way to get a true $1p0\pi$ sample. The proton and pion multiplicity plots are shown in Extended Data Fig. 3. *e*-GENIE dramatically overpredicts the number of events with large proton and pion multiplicities.

We also accounted for the effects of, for example, events with two detected pions (or photons). When we rotated these events, each rotated event could have been detected as a 2π event, a 1π event or a 0π event. If it appeared as a 0π event, we subtracted its contribution from the various 0π spectra as described above. If it appeared as a 1π event, we included it in the set of 1π events with the appropriate negative weight. We then treated it as a regular 1π event, which we then rotated and added to the 0π dataset as described above. (We note that some of the detected 1π events are actually 2π events with an undetected pion. When we account for the effects of these events, we are left with fewer true 1π events. This reduces the contamination of the 1π events in the 0π channel.)

In practice, we started with the highest multiplicity events, for example, $(e, e'3\pi)$, and subtracted their contributions to each of the detected lower multiplicity channels—for example, $(e, e'2\pi)$, $(e, e'\pi)$ and $(e, e'0\pi)$. We then worked recursively, rotating higher multiplicity events to determine and subtract their contributions to the lower multiplicity channels, and then considering each of the lower multiplicity channels in turn.

We considered event multiplicities up to three pions and photons (total) for (e, e') and up to three protons, pions and photons (total) for $(e, e'p)$, where the subtraction converged. The effects of the subtraction (and its convergence) can be seen in Extended Data Fig. 7. The number of events with an undetected $\pi^\pm$ or photon is about equal to the number of events with a detected $\pi^\pm$ or photon, consistent with the approximately 50% CLAS geometrical acceptance. The effect of including two $\pi^\pm$ or photon events is much less than that of the one $\pi^\pm$ or photon events and the effect of including three $\pi^\pm$ or photon events is negligible.

We tested the subtraction method by applying it to GENIE-simulated data. The resulting subtracted spectra agreed reasonably with the true $1p0\pi$ spectra.

Cross-section determination

We determined the inclusive 37.5° 1.159-GeV $C(e, e')$ cross-section (see Extended Data Fig. 3b):

$$\frac{d\sigma}{d\Omega d\omega} = \frac{N_e}{\Delta\Omega N_i N_t}$$

where N_e is the number of detected electrons in sector 1 within $36^\circ \leq \theta_e \leq 39^\circ$ and a 12° range in ϕ_e , $\Delta\Omega = \sin\theta_e d\theta_e d\phi_e = 6.68$ milliradians, N_i is the number of incident electrons and $N_t = 0.179 \text{ g cm}^{-2} = 8.95 \times 10^{-9}$ nuclei per μbarn . The measured cross-sections are in reasonable agreement with the GENIE predictions and are also consistent with the SLAC measurements at lower and higher energies⁴².

We determined the cross-section as a function of reconstructed energy (for particles above the minimum angles, see equation (14)) in several steps. We first weighted all events by a factor of Q^4 to account for the major difference in electron–nucleus and neutrino–nucleus scattering.

1. Determine the number of weighted $(e, e')_{0\pi}$ or $(e, e'p)_{1p0\pi}$ events, corrected for events with undetected pions, photons and (if appropriate) extra protons, as a function of reconstructed energy.

2. Divide the number of events by the number of target nuclei per area and the number of incident beam electrons to get the normalized yield.

3. Correct for electron radiation by multiplying the resulting spectra by the ratio of *e*-GENIE without electron radiation divided by *e*-GENIE with electron radiation (see Extended Data Fig. 8g–i). This includes a multiplicative factor to account for the effects of internal radiation.

4. Correct for electron and proton acceptance and other detector effects using *e*-GENIE. The acceptance correction factor is the ratio of the number of true $(e, e'p)_{1p0\pi}$ events without detector effects to the number of true $(e, e'p)_{1p0\pi}$ events with detector effects. The detector effects included momentum resolution, fiducial (acceptance) cuts

and efficiency maps. The fiducial cuts determine the useful areas of the detector as a function of particle momenta and angles, and the efficiency maps describe the efficiency of the detector as a function of particle momenta and angles. This factor corrects the effective electron and proton solid angles to almost 4π . It excludes all electrons, pions and protons below their minimum angles (see equation (14)).

5. Finally, divide by the reconstructed energy bin width.

We calculated the acceptance correction factor using both G2018 and SuSAv2. We used the bin-by-bin average of the two as the acceptance correction factor and the bin-by-bin difference divided by $\sqrt{12}$ as the uncertainty (see Extended Data Fig. 8a–f.)

GENIE simulations

We generated events with the electron-scattering version of GENIE (*e*-GENIE), one of the standard neutrino event generators. *e*-GENIE has been substantially modified recently to fix known issues, to use reaction mechanisms as close to those of *v*-GENIE as possible (version v3.00.06), and to include the effects of the real photon part of electron bremsstrahlung¹⁹. We used CLAS acceptance maps to determine the probability that each particle was detected and smeared the momenta of the particles with an effective CLAS resolution (we used electrons and proton momentum resolutions of 0.5% and 1%, respectively, for the 2.257- and 4.453-GeV data and 1.5% and 3% for the 1.159-GeV data, which was taken with a lower torus magnetic field). We then analysed the events in the same way as the data.

We used *e*-GENIE with the newly implemented SuSAv2 models for electron–nucleon QE scattering and for MEC. The SuSAv2 QE model is based on the superscaling exhibited by inclusive electron scattering data and uses relativistic mean field theory to describe both the nuclear initial state and the inclusive cross-section modifications due to the outgoing nucleon reinteracting with the residual nucleus^{13,43,44}. This inclusive model is then joined with a modified version of GENIE's local Fermi gas model in a factorized way to produce predictions for the outgoing nucleon. The SuSAv2 MEC model describes the lepton interaction with a transverse electromagnetic isovector MEC (which excites two-particle two-hole (2p2h) states) in the framework of ref. ⁴⁵. Because it uses superscaling to describe a wide range of nuclei with a single scaling function, SuSAv2 was never intended to model nuclei lighter than ¹²C.

We compared that with the older *e*-GENIE version (tune G18_10a_02_11a, referred to here as G2018) which used the local Fermi gas model; the Rosenbluth cross-section for electron–nucleon QE scattering and the empirical Dytman model⁴⁶ of MEC (meson exchange currents or 2p2h currents), which describes it as a Gaussian distribution located between the QE and first resonance peaks⁴⁷.

Both the SuSAv2 and G2018 tunes used the Berger–Sehgal model⁴⁸ of electron–nucleon resonance production (RES), which includes cross-sections of 16 resonances calculated in the Feynman–Kislinger–Ravndal (FKR) model⁴⁹, without interference between them. There are slight differences in the tuning parameters for the two tunes.

Deep inelastic scattering (DIS) in both tunes is modelled using Bodek and Yang⁵⁰. Hadronization is modelled using an approach that transitions gradually between the AGKY model⁵¹ and the PYTHIA 6 model⁵². At low values of the hadronic invariant mass W , the Bodek–Yang differential cross-section is scaled by tunable parameters that depend on the multiplicity of hadrons in the final state⁵³.

e-GENIE used two models of the final-state interactions of outgoing protons and pions with the residual nucleus. G2018 used the Intranuke package^{54,55} with hA2018, an empirical data-driven method, using the cross-section of pions and nucleons with nuclei as a function of energy up to 1.2 GeV, and the CEM03⁵⁶ calculation normalized to low-energy data for higher energies. SuSAv2 used hN, a full intra-nuclear cascade calculation of the interactions of pions, kaons, photons and nucleons with nuclei. In the hN model, each outgoing particle can interact successively with any or all the nucleons it encounters on its path leaving the nucleus, and any particles created in those interactions can also

subsequently reinteract. The ability of the two models to describe hadron–nucleus data is very similar.

e-GENIE includes radiative corrections based on the formalism of ref. ⁵⁷ to account for radiation of a real photon by the electron either before or after scattering, for radiation and reabsorption of a virtual photon (vertex correction), and for virtual pair production by the exchanged photon (vacuum polarization). This includes radiated real photons up to 15% of the incident electron energy. The vertex correction and vacuum polarization corrections were of the order of 6%.

Although many of the reaction mechanisms in GENIE have been tuned to approximately reproduce inclusive electron scattering cross-sections, the hadron-production part of the models have rarely, if ever, been compared to precision data.

This version of *e*-GENIE¹⁹ and *v*-GENIE²⁷ described inclusive electron scattering and charged current neutrino scattering cross-sections moderately well. The SuSAv2 models described the data around the QE peak and in the dip region between the QE peak and the first resonance peak much better than the G2018 tune. The data were less well described at energy transfers at and above the first resonance peak (the region where RES and DIS scattering dominate).

The difference between data and GENIE at smaller reconstructed energies attributed to the RES and DIS processes might be due to one of two options: (a) the GENIE RES and DIS processes were tuned to *v*-deuteron data only and thus the axial + vector response is better constrained than the vector part; or (b) the distinction between RES and DIS is artificial and some resonance contributions might be double counted in the two processes. There is a suggestion that GENIE also overpredicts resonance production in neutrino–nucleus interactions (see figures 13–16 in ref. ⁵⁸), lending support to the second hypothesis.

Systematic uncertainties

We considered several major sources of systematic uncertainties, including the angular dependence of the pion-production cross-section (for the undetected-pion subtraction), the effects of fiducial cuts on undetected particle subtraction, photon identification cuts, the sector-to-sector variation of the data to *e*-GENIE ratio, the model dependence of the acceptance correction, and uncertainties in the normalization measurement.

When we rotated events containing pions around the momentum transfer vector, we assumed that the cross-section did not depend on $\phi_{q\pi}$. We tested the $\phi_{q\pi}$ independence of the pion-production cross-section by weighting the subtraction using the measured $\phi_{q\pi}$ -dependent $H(e, e'\pi\pi)$ cross-sections of ref. ⁵⁹. This changed the subtracted spectra by about 1% and was included as a systematic uncertainty.

The subtraction of events with undetected pions depends on the CLAS acceptance for such particles. The final spectrum should be independent of the CLAS pion acceptance. We estimated the effect of varying the CLAS acceptance on the undetected particle subtraction by comparing the results using the nominal fiducial cuts and using fiducial cuts with the ϕ acceptance in each CLAS sector reduced by 6° or about 10–20%. This changed the resulting subtracted spectra by about 1% at 1.159 and 2.257 GeV and by 4% at 4.453 GeV. This difference was included as a point-to-point systematic uncertainty.

We also varied the photon identification cuts. We identified photons as neutral particle hits in the calorimeter with a velocity greater than two standard deviations (3σ at 1.159 GeV) below the mean of the photon velocity peak (at $v = c$). We varied this limit by $\pm 0.25\sigma$. This gave an uncertainty in the resulting subtracted spectra of 0.1%, 0.5% and 2% at 1.159, 2.257 and 4.453 GeV, respectively.

CLAS has six almost-identical sectors (see Extended Data Fig. 9a). The primary difference among the sectors is the distribution of dead detector channels. We attempted to account for these dead channels in our fiducial cuts (where we cut away dead or partially dead regions of the detector) and in our acceptance maps, where we measured the effect of the dead detectors on the particle detection efficiency and

applied that efficiency to the particles generated in the *e*-GENIE Monte Carlo simulation. If our fiducial cuts and acceptance maps completely accounted for the effect of the dead and inefficient detector channels, then the ratio of data to *e*-GENIE should be the same for all six sectors.

We discarded sectors with anomalous data to *e*-GENIE ratios and used the variance of the ratios for the remaining sectors as a measure of the uncertainty in the measured normalized yields. This gave a point-to-point systematic uncertainty of 6%.

We calculated the acceptance correction factors using both G2018 and SuSAv2. We shifted the G2018 results so that the energy reconstruction peaks lined up at the correct beam energy. We used the bin-by-bin average of the two correction factors as the acceptance correction factor and the bin-by-bin difference divided by $\sqrt{12}$ as the uncertainty. We averaged the uncertainty over the entire peak to avoid large uncertainties due to small misalignments. See Extended Data Fig. 8a–f.

The overall normalization was determined using inclusive 4.4-GeV $H(e, e')$ measurements. The measured and simulated $H(e, e')$ cross-sections agreed to within an uncertainty of 3%, which we use as a normalization uncertainty⁶⁰.

We added the statistical uncertainty and the point-to-point systematic uncertainties in quadrature and displayed them on the data points.

Future CLAS12 experiments

Jefferson Lab Experiment E12-17-006, ‘Electrons for Neutrinos: Addressing Critical Neutrino–Nucleus Issues’ (scientific rating: A) will take further data on more targets with a greater kinematical range using the upgraded CLAS12 detector. The approved experiment includes measurements on ⁴He, C, Ar and Sn with 1-, 2- and 4-GeV electron beams, as well as measurements on O with 1- and 2-GeV electron beams. The 1- and 2-GeV measurements will be performed with a minimum electron scattering angle of 5°, compared to a minimum CLAS angle of about 20°. This will extend the measurements down to the much lower momentum transfers typical of some neutrino experiments. It will therefore allow comparison with the lower beam-energy data of T2K and Hyper-Kamiokande. The first part of the experiment is scheduled to run in the second half of 2021.

Data availability

The raw data from this experiment are archived in the Jefferson Lab’s mass storage silo under the CLAS E2 run-period dataset. Access to these data can be facilitated by contacting either the corresponding authors or the Jefferson Lab computing centre at helpdesk@jlab.org.

32. Freund, M. Analytic approximations for three neutrino oscillation parameters and probabilities in matter. *Phys. Rev. D* **64**, 053003 (2001).
33. Cervera, A. et al. Golden measurements at a neutrino factory. *Nucl. Phys. B* **579**, 17–55 (2000).
34. Cervera, A. et al. Erratum to: “Golden measurements at a neutrino factory”: [Nucl. Phys. B 579 (2000) 17]. *Nucl. Phys. B* **593**, 731–732 (2001).
35. Osipenko, M. et al. Measurement of the nucleon structure function F_2 in the nuclear medium and evaluation of its moments. *Nucl. Phys. A* **845**, 1–32 (2010).
36. CLAS Collaboration. Measurement of two- and three-nucleon short-range correlation probabilities in nuclei. *Phys. Rev. Lett.* **96**, 082501 (2006).
37. CLAS Collaboration. Survey of A_{LT} asymmetries in semi-exclusive electron scattering on ⁴He and ¹²C. *Nucl. Phys. A* **748**, 357–373 (2005).
38. CLAS Collaboration. Proton source size measurements in the $eA \rightarrow e'ppX$ reaction. *Phys. Rev. Lett.* **93**, 192301 (2004).
39. CLAS Collaboration. Two-nucleon momentum distributions measured in ³He($e, e'pp$)n. *Phys. Rev. Lett.* **92**, 052303 (2004).
40. CLAS Collaboration. Observation of nuclear scaling in the $A(e, e')$ reaction at $x_B > 1$. *Phys. Rev. C* **68**, 014313 (2003).
41. Hen, O. et al. Momentum sharing in imbalanced Fermi systems. *Science* **346**, 614–617 (2014).
42. Seablock, R. M. et al. Electroexcitation of the $\Delta(1232)$ in nuclei. *Phys. Rev. Lett.* **62**, 1350–1353 (1989).
43. González-Jiménez, R. et al. Extensions of superscaling from relativistic mean field theory: the SuSAv2 model. *Phys. Rev. C* **90**, 035501 (2014).
44. Megias, G. D. et al. Inclusive electron scattering within the SuSAv2 meson-exchange current approach. *Phys. Rev. D* **94**, 013012 (2016).
45. De Pace, A., Nardi, M., Alberico, W. M., Donnelly, T. W. & Molinari, A. The 2p–2h electromagnetic response in the quasielastic peak and beyond. *Nucl. Phys. A* **726**, 303–326 (2003).

46. Katori, T. Meson exchange current (MEC) models in neutrino interaction generators. In *NuInt12: 8th Intl Workshop On Neutrino–Nucleus Interactions in the Few-GeV Region* (eds. Da Motta, H. et al.) <https://doi.org/10.1063/1.4919465> (AIP, 2015).
47. GENIE Collaboration. Neutrino–nucleon cross-section model tuning in GENIE v3. *Phys. Rev. D* **104**, 072009 (2021).
48. Berger, Ch. & Sehgal, L. M. Lepton mass effects in single pion production by neutrinos. *Phys. Rev. D* **76**, 113004 (2007).
49. Feynman, R. P., Kislinger, M. & Ravndal, F. Current matrix elements from a relativistic quark model. *Phys. Rev. D* **3**, 2706 (1971).
50. Bodek, A. & Yang, U. K. Higher twist, E_p scaling, and effective LO PDFs for lepton scattering in the few GeV region. *J. Phys. G* **29**, 1899–1905 (2003).
51. Yang, T., Andreopoulos, C., Gallagher, H., Hofmann, K. & Kehayias, P. A hadronization model for few-GeV neutrino interactions. *Eur. Phys. J. C* **63**, 1–10 (2009).
52. Sjostrand, T., Mrenna, S. & Skands, P. Z. PYTHIA 6.4 physics and manual. *J. High Energy Phys.* **2006**, 026 (2006).
53. Andreopoulos, C. et al. The GENIE neutrino Monte Carlo generator: physics and user manual. Preprint at <https://arxiv.org/abs/1510.05494> (2015).
54. Dytman, S. A. & Meyer, A. S. Final state interactions in GENIE. *AIP Conf. Proc.* **1405**, 213 (2011).
55. Merenyi, R. et al. Determination of pion intranuclear rescattering rates in ν_μ -Ne versus ν_μ -D interactions for the atmospheric ν flux. *Phys. Rev. D* **45**, 743–751 (1992).
56. Mashnik, S. G., Sierk, A. J., Gudima, K. K. & Baznat, M. I. CEM03 and LAQGSM03—new modeling tools for nuclear applications. *Phys. Conf. Ser.* **41**, <https://doi.org/10.1088/1742-6596/41/1/037> (2006).
57. Mo, L. W. & Tsai, Y.-S. Radiative corrections to elastic and inelastic ep and μp scattering. *Rev. Mod. Phys.* **41**, 205–235 (1969).
58. The T2K Collaboration. Measurement of the muon neutrino charged-current single π^+ production on hydrocarbon using the T2K off-axis near detector ND280. *Phys. Rev. D* **101**, 012007 (2020).
59. The T2K Collaboration. Exclusive n^0p electroproduction off protons in the resonance region at photon virtualities $0.4 \text{ GeV}^2 \leq Q^2 \leq 1 \text{ GeV}^2$. *Phys. Rev. C* **101**, 015208 (2020).
60. Osipenko, M. A *Kinematically Complete Measurement of the Proton Structure Function F_2 in the Resonance Region and Evaluation of its Moments*. PhD thesis, Moscow State Univ. (2002); https://www.jlab.org/Hall-B/general/thesis/Osipenko_thesis.ps.gz
61. The T2K Collaboration. Improved constraints on neutrino mixing from the T2K experiment with 3.13×10^{21} protons on target. *Phys. Rev. D* **103**, 112008 (2021).
62. DUNE Collaboration. Long-baseline neutrino oscillation physics potential of the DUNE experiment. *Eur. Phys. J. C* **80**, 978 (2020).

Acknowledgements We acknowledge the efforts of the staff of the Accelerator and Physics Divisions at Jefferson Lab that made this experiment possible. We thank L. Pickering for useful discussions. The analysis presented here was carried out as part of the Jefferson Lab Hall B Data-Mining project supported by the US Department of Energy (DOE). The research was supported also by DOE, the US National Science Foundation, the Israel Science Foundation, the Chilean Comisión Nacional de Investigación Científica y Tecnológica, the French Centre national de la recherche scientifique and Commissariat à l’Energie Atomique et aux Energies Alternatives, the French–American Cultural Exchange, the Italian Istituto Nazionale di Fisica Nucleare, the National Research Foundation of Korea, and the UK Science and Technology Facilities Council. P. Coloma acknowledges support from project PROMETEO/2019/083. This project has been supported by the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 674896 (Elusives, H2020-MSCA-ITN- 2015-674896). G.D.M. acknowledges support from the Spanish Ministerio de Economía y Competitividad and ERDF (European Regional Development Fund) under contract FIS2017-88410-P, by the University of Tokyo ICRs Inter-University Research Program FY2020 & 2021, refs no. A07 and A06; by the Junta de Andalucía (grant no. FQM160); and by the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 839481. This document was prepared by the $e4\nu$ Collaboration using the resources of the Fermi National Accelerator Laboratory (Fermilab), a US Department of Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under contract no. DE-AC02-07CH11359. Jefferson Science Associates operates the Thomas Jefferson National Accelerator Facility for the DOE, Office of Science, Office of Nuclear Physics under contract DE-AC05-06OR23177. The raw data from this experiment are archived in Jefferson Lab’s mass storage silo.

Author contributions The CEBAF Large Acceptance Spectrometer was designed and constructed by the CLAS Collaboration and Jefferson Lab. Data acquisition, processing and calibration, Monte Carlo simulations of the detector and data analyses were performed by a large number of CLAS Collaboration members, who also discussed and approved the scientific results. The analysis presented here was performed by M. Khachatryan, A.P., A.A., A. Hrnjic and A.N. with guidance from A.A., F.H., O.H., E. Piasetzky and L.B.W., and was reviewed by the CLAS Collaboration. S. Dytman, M. Betancourt and K.M. provided expertise on neutrino scattering. S. Dytman, G.M., S. Dolan and S.G. helped develop *e*-GENIE. P. Coloma performed a simulation of the DUNE sensitivity to the oscillation parameters, and determined the impact of our results on the fit.

Competing interests The authors declare no competing interests.

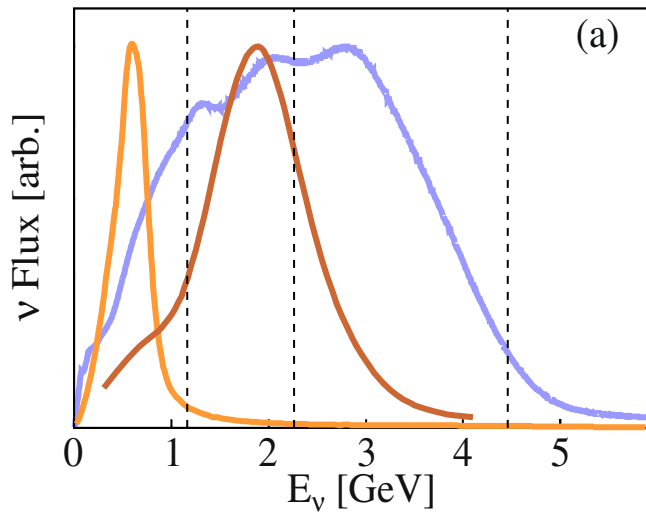
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-021-04046-5>.

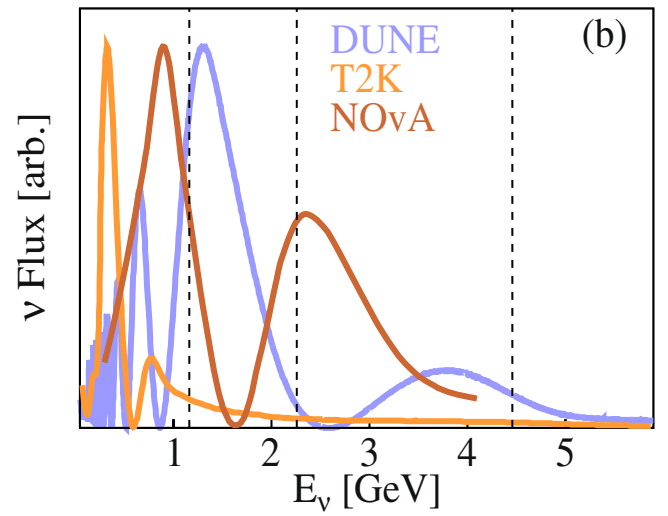
Correspondence and requests for materials should be addressed to A. Ashkenazi.

Peer review information *Nature* thanks Tingjun Yang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

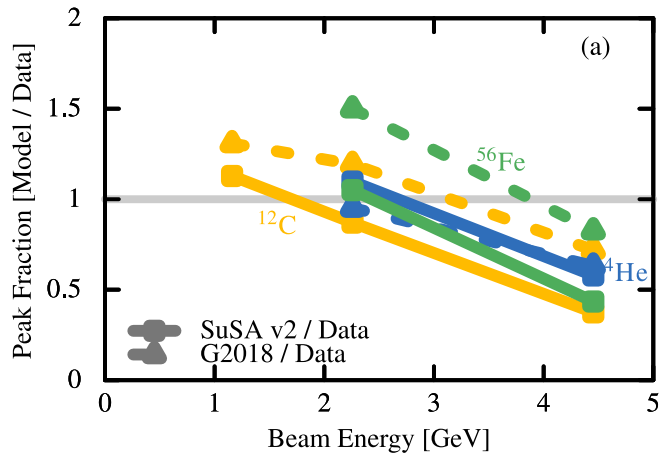
Reprints and permissions information is available at <http://www.nature.com/reprints>.



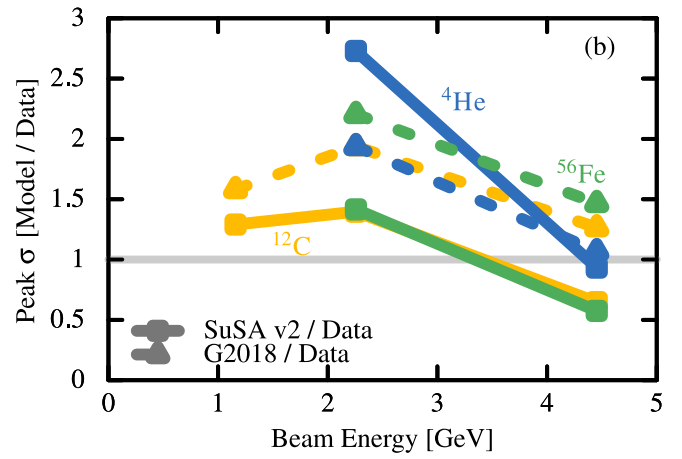
Extended Data Fig. 1 | Energy distributions of different ν_μ beams. Left, Before oscillation at the near detector; and right, after oscillation at the far detector^{61,62}. The vertical lines show the three electron beam energies of this



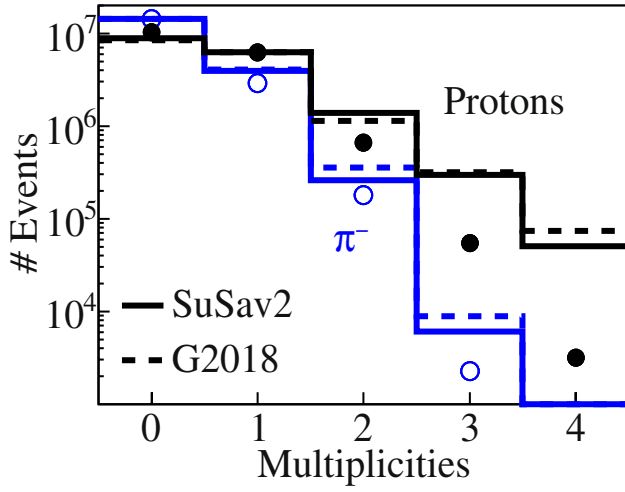
measurement. The NOvA far-detector beam flux is calculated using the near-detector flux and the neutrino oscillation parameters from the Particle Data Group. arb., arbitrary units.



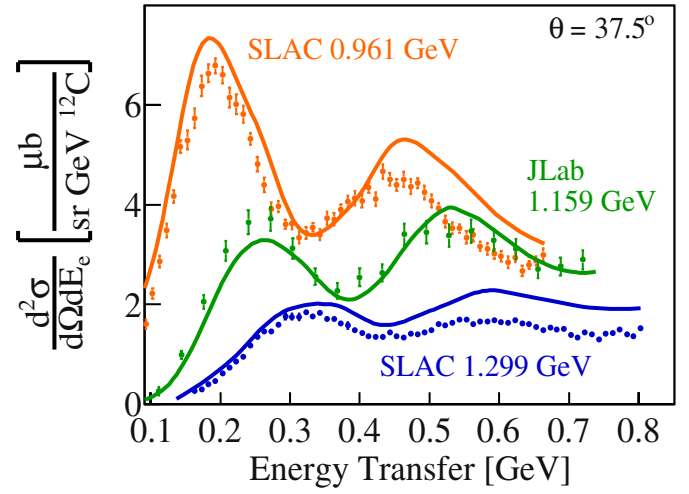
Extended Data Fig. 2 | Peak energy reconstruction fraction and width. Left, The ratio of *e*-GENIE to data for the fraction of the weighted cross-section that reconstructs to the correct incident energy, plotted versus incident energy; and right, the *e*-GENIE–data weighted cross-section ratio for events that reconstruct to the correct incident energy, plotted versus incident energy. The triangles and dashed lines indicate the G2018/data ratios and the squares and



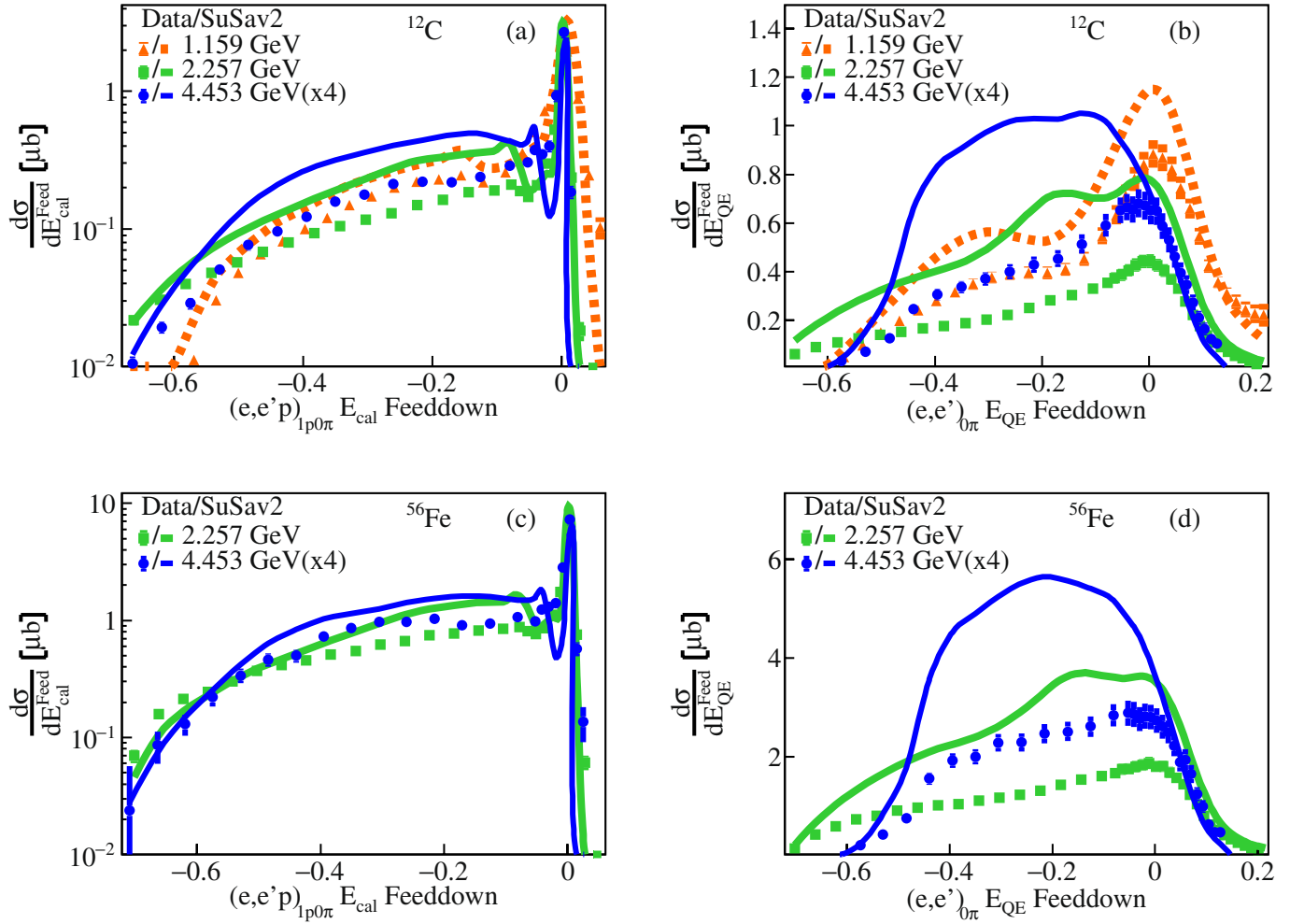
solid lines indicate the SuSAv2/data ratios. SuSAv2 is not intended to model nuclei lighter than ^{12}C . Yellow shows the carbon, blue shows helium and green shows iron. Error bars show the 68% (1σ) confidence limits for the statistical and point-to-point systematic uncertainties added in quadrature. Error bars are not shown when they are smaller than the size of the data point. Normalization uncertainties of 3% not shown.



Extended Data Fig. 3 | Particle multiplicities and inclusive cross-section extraction. Left, The proton (black) and charged pion (blue) multiplicities for data (points), SuSav2 (solid histogram) and G2018 (dashed histogram) for 2.257-GeV carbon. Right, Comparison between the inclusive $C(e, e')$ cross-sections measured at 37.5° for data (points) and SuSav2 (lines) for the

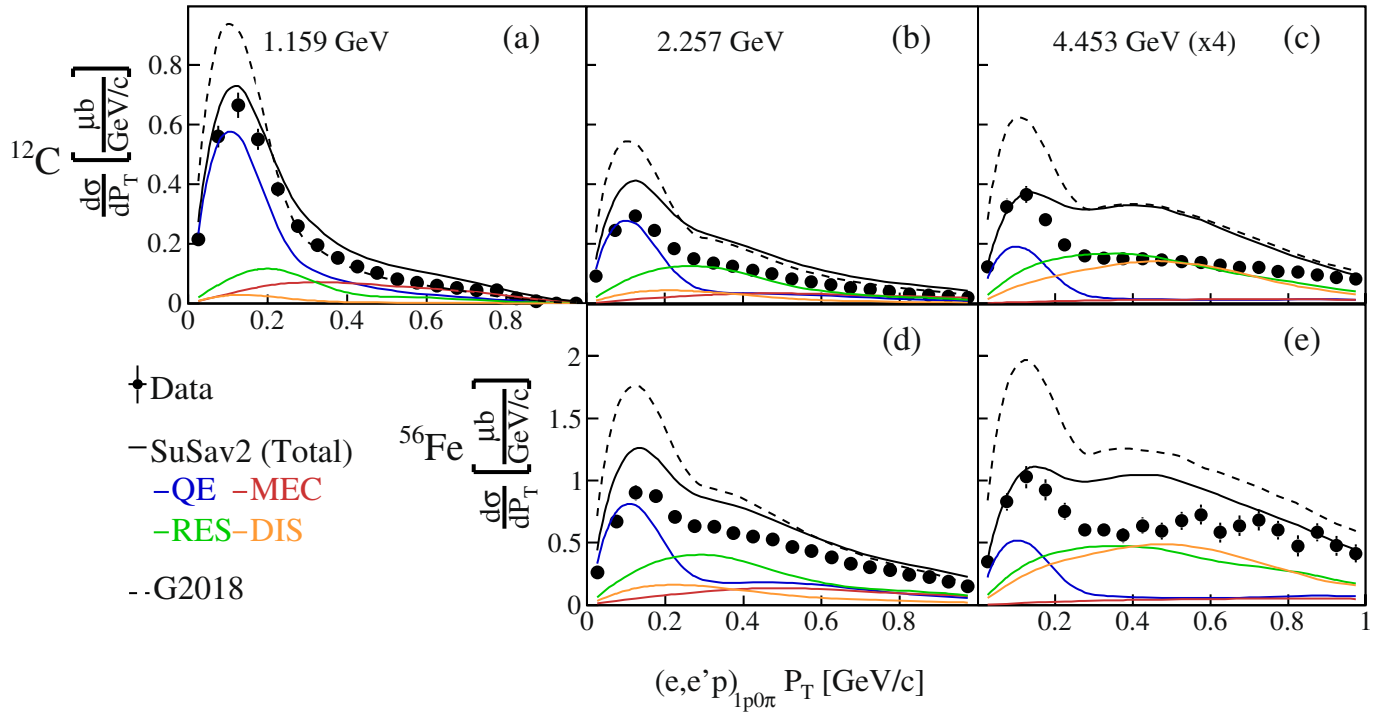


0.961- and 1.299-GeV SLAC data⁴² and our 1.159-GeV CLAS data. Error bars show the 68% (1σ) confidence limits for the statistical and point-to-point systematic uncertainties added in quadrature. Error bars are not shown when they are smaller than the size of the data point. Normalization uncertainties of 3% not shown.



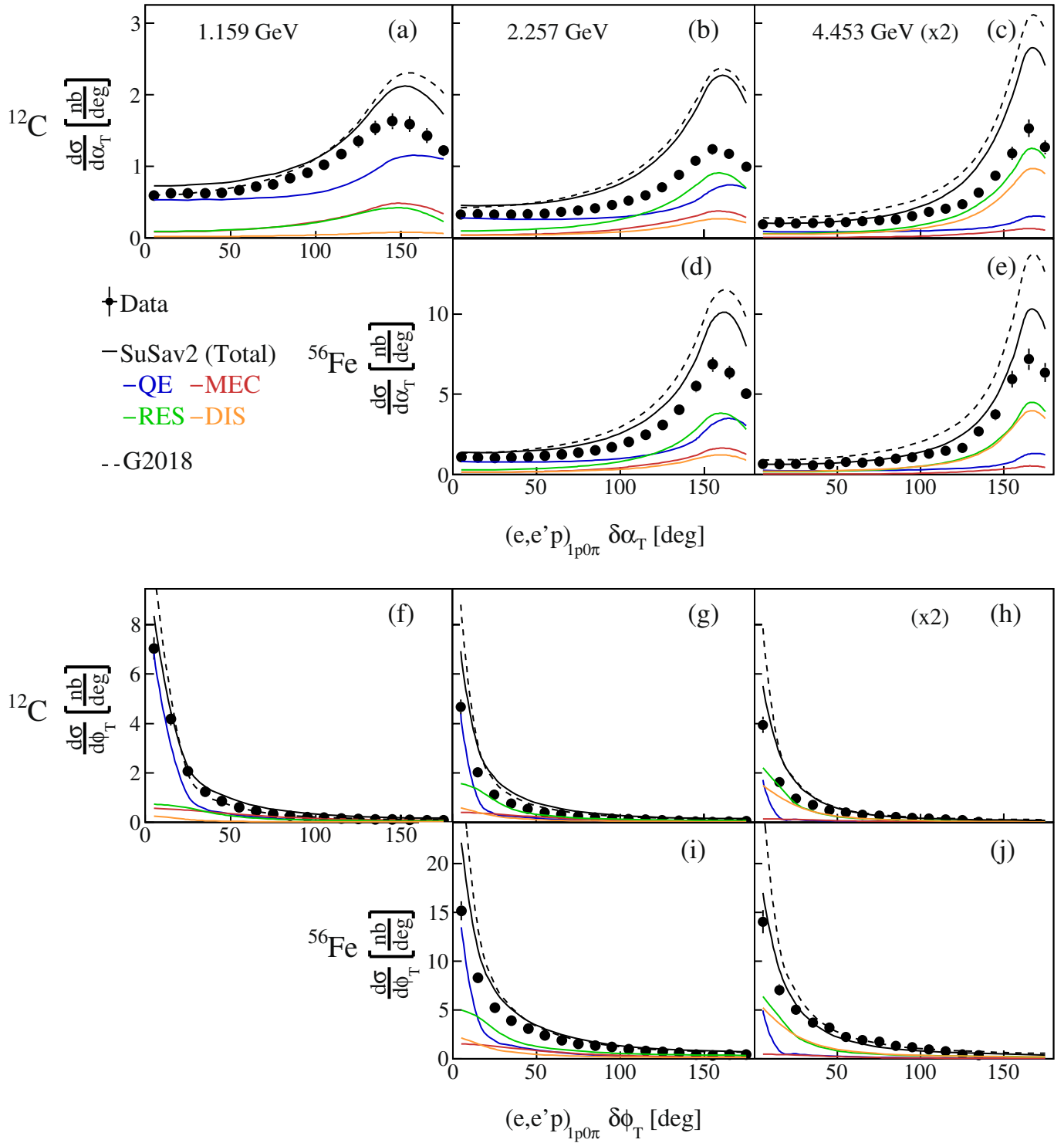
Extended Data Fig. 4 | Energy feed-down cross-sections. a–d, $(E_{\text{rec}} - E_{\text{true}})/E_{\text{true}}$ for data (points) and SuSav2 (lines) for 1.159 GeV (red triangles and dotted lines), 2.257 GeV (green squares and solid lines) and 4.453 GeV (blue dots and solid lines) for $C E_{\text{cal}}$ (a), $C E_{\text{QE}}$ (b), $\text{Fe } E_{\text{cal}}$ (c), and $\text{Fe } E_{\text{QE}}$ (d). The plots are area-normalized and each bin has been scaled by the bin width. Error bars show

the 68% (1σ) confidence limits for the statistical and point-to-point systematic uncertainties added in quadrature. Error bars are not shown when they are smaller than the size of the data point. Normalization uncertainties of 3% not shown.



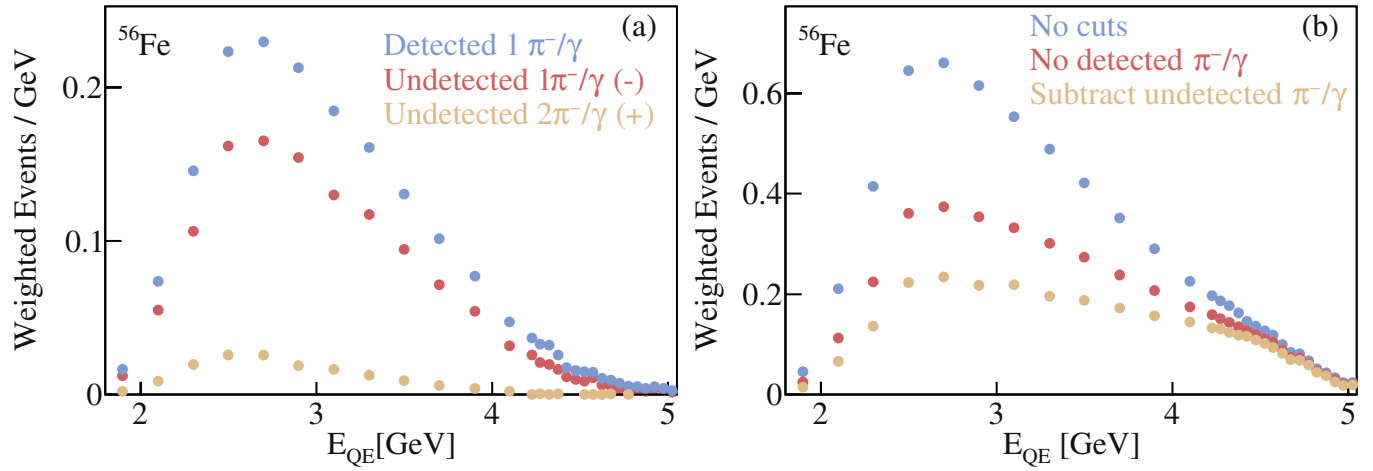
Extended Data Fig. 5 | Transverse missing-momentum-dependent differential cross-section. The cross-section plotted versus transverse missing momentum P_T for data (black points), SuSav2 (black solid curve) and G2018 (black dotted curve). Different panels show results for different beam energy and target nucleus combinations: **a–c**, Carbon target at 1.159 GeV (**a**), 2.257 GeV (**b**) and 4.453 GeV (**c**). **d, e**, Iron target at 2.257 GeV (**d**) and 4.453 GeV (**e**). The 4.453-GeV yields have been scaled by four to have the same vertical

scale. Coloured lines show the contributions of different processes to the SuSav2 GENIE simulation: QE (blue), MEC (red), RES (green) and DIS (orange). Error bars show the 68% (1σ) confidence limits for the statistical and point-to-point systematic uncertainties added in quadrature. Error bars are not shown when they are smaller than the size of the data point. Normalization uncertainties of 3% not shown.



Extended Data Fig. 6 | $\delta\alpha_T$ -dependent differential cross-section. a–j. The cross-section plotted versus $\delta\alpha_T$ (a–e) and versus $\delta\phi_T$ (f–j) for data (black points), SuSav2 (black solid curve) and G2018 (black dotted curve). Different panels show results for different beam energy and target nucleus combinations: a–c, Carbon target at 1.159 GeV (a), 2.257 GeV (b) and 4.453 GeV (c). d, e, Iron target at 2.257 GeV (d) and 4.453 GeV (e). The 4.453-GeV yields have been scaled by two to have the same vertical scale. Coloured lines show the

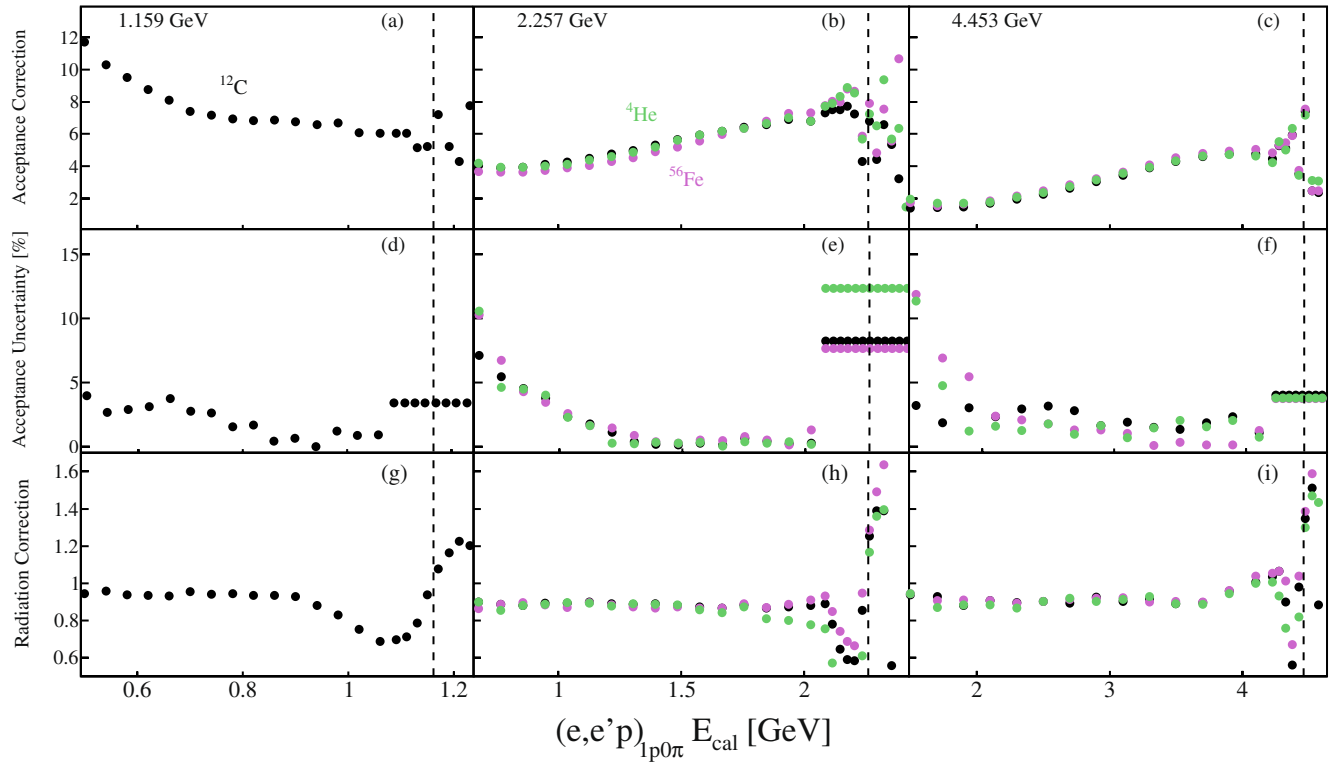
contributions of different processes to the SuSav2 GENIE simulation: QE (blue), MEC (red), RES (green) and DIS (orange). Error bars show the 68% (1σ) confidence limits for the statistical and point-to-point systematic uncertainties added in quadrature. Error bars are not shown when they are smaller than the size of the data point. Normalization uncertainties of 3% not shown.



Extended Data Fig. 7 | The effect of undetected pion subtraction.

The number of weighted events as a function of reconstructed energy E_{QE} for 4.453-GeV $^{56}\text{Fe}(e, e')$ events for: left, events with a detected $\pi^\pm$ or photon (blue), events with one (red) or two (light brown) undetected $\pi^\pm$ or photons; and right, all $(e, e'X)$ events with detected or undetected $\pi^\pm$ or photon (blue), (e, e') events

with no detected $\pi^\pm$ or photon (red), and (e, e') events after subtraction for undetected $\pi^\pm$ or photon (light brown). The uncertainties are statistical only and are shown at the 1σ or 68% confidence level. Error bars are not shown when they are smaller than the size of the data point.

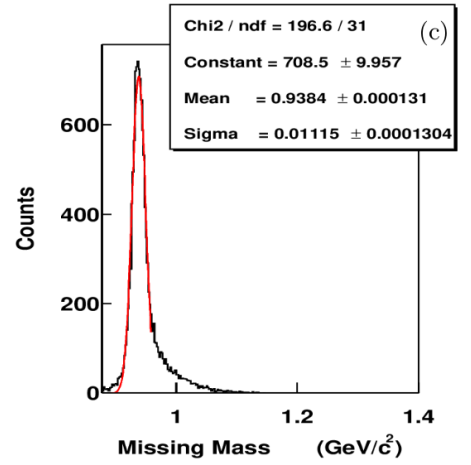
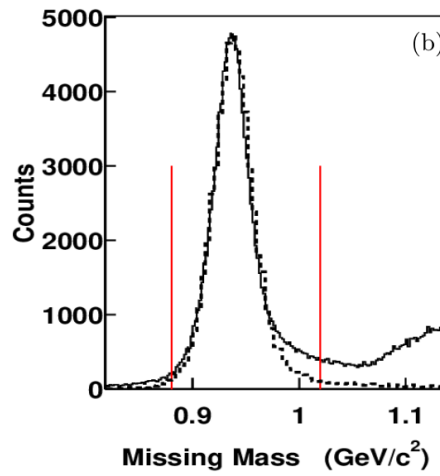
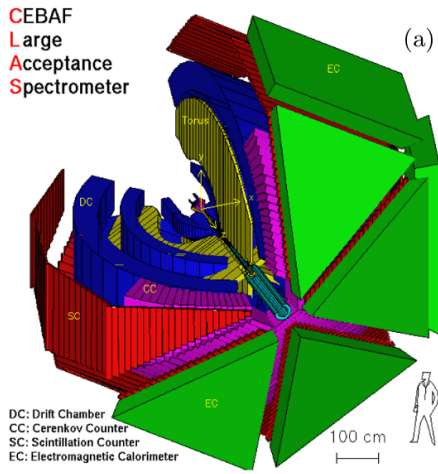


Extended Data Fig. 8 | Acceptance and radiation corrections.

a–c, Acceptance correction factors; **d–f**, acceptance correction factor uncertainties; and **g–i**, electron radiation correction factors plotted versus E_{cal} for the three incident beam energies. Results for carbon are shown in black,

helium in green and iron in magenta. The left column (**a**, **d**, **g**) shows the 1.159-GeV results, the middle column (**b**, **e**, **h**) shows the 2.257-GeV results and the right column (**c**, **f**, **i**) shows the 4.453-GeV results.

CEBAF
Large
Acceptance
Spectrometer



Extended Data Fig. 9 | CLAS detector and its calibration performance.

a. Cutaway drawing of CLAS showing the sector structure and the different detectors. Yellow, toroidal magnet; blue, drift chambers; magenta, Cherenkov counter; red, scintillation counters (time of flight); green, electromagnetic calorimeter. The beam enters from the upper left and the target is in the center

of CLAS. CLAS detector image reproduced with permission of the CLAS Collaboration. **b.** The 2.257-GeV $^3\text{He}(e, e'pp)X$ missing mass for data (solid histogram) and simulation (dashed histogram). **c.** The $\text{H}(e, e'\pi^+)X$ missing mass for data (black) and fit to data (red).

Extended Data Table 1 | $(e, e'p)_{1p0\pi}$ events reconstructed to the correct beam energy

		1.159 GeV		2.257 GeV		4.453 GeV	
		Peak	Peak	Peak	Peak	Peak	Peak
		Fraction	Sum [μ b]	Fraction	Sum [μ b]	Fraction	Sum [μ b]
^4He	Data	-	-	41	0.48	38	0.15
	SuSAv2	-	-	45	1.31	22	0.14
	G2018	-	-	39	0.93	24	0.16
^{12}C	Data	39	4.13	31	1.26	32	0.34
	SuSAv2	44	5.33	27	1.76	12	0.20
	G2018	51	6.53	37	2.44	23	0.43
^{56}Fe	Data	-	-	20	3.73	23	1.01
	SuSAv2	-	-	21	5.28	10	0.58
	G2018	-	-	30	8.22	19	1.48

‘Peak fraction’ refers to the fraction of events reconstructed to the correct beam energy and ‘Peak sum’ refers to the integrated weighted cross-section (as shown in Fig. 3) reconstructed to the correct beam energy. The peak integration windows are $1.1 \leq E_{\text{cal}} \leq 1.22$ GeV, $2.19 \leq E_{\text{cal}} \leq 2.34$ GeV and $4.35 \leq E_{\text{cal}} \leq 4.60$ GeV for the three incident beam energies, respectively. Uncertainties are shown graphically in Extended Data Fig. 2. SuSAv2 is not intended to model nuclei lighter than ^{12}C .