Constraints on Oscillation Parameters from $\nu_e$ Appearance and $\nu_\mu$ Disappearance in NOvA

P. Adamson,11 L. Aliaga,11 D. Ambrose,26 N. Anfinov,22 A. Autoshkin,22 E. Arrieta-Diaz,31 K. Augsten,9 A. Aurisano,6 C. Backhouse,4 M. Baird,33,17 B. A. Bambah,15 K. Bays,4 B. Behera,16 S. Bending,37 R. Bernstein,11 V. Bhatnagar,27 B. Bhuyan,13 J. Bian,20,26 T. Blackburn,29 A. Bolshakov,22 C. Bromberg,24 J. Brown,26 G. Brunetti,11 N. Buchanan,8 A. Butkevich,18 V. Bychkov,26 M. Campbell,37 E. Catano-Mur,19 S. Childress,11 B. C. Choudhury,10 B. Chowdhury,29 T. E. Coan,31 J. A. B. Coelho,36 M. Colo,40 J. Cooper,11 L. Corwin,30 L. Cremonesi,37 D. Cronin-Hennessy,26 G. S. Davies,17 J. P. Davies,33 P. F. Derwent,11 R. Dharmapalan,1 P. Ding,11 Z. Djuric,1 E. C. Dukes,38 H. Duyang,29 S. Edalayath,7 R. Ehrlich,38 G. J. Feldman,14 M. J. Frank,28,38 M. Gabrielyan,26 H. R. Gallagher,36 S. Germani,37 T. Ghosh,12 A. Girli,36 R. A. Gomes,12 M. C. Goodman,1 V. Grichine,23 R. Group,38 D. Grover,3 B. Guo,29 A. Habig,25 J. Hartnell,33 R. Hatcher,11 A. Hatzikoutelis,34 K. Heller,26 A. Himmel,11 A. Holin,37 J. Hylen,11 F. Jediny,9 M. Judah,8 G. K. Kafka,14 D. Kalra,27 S. M. S. Kasahara,26 S. Kasetti,15 R. Keloth,7 L. Kolupaeva,22 S. Kotelnikov,23 I. Kourbanis,11 A. Kreymer,11 A. Kumar,27 S. Kurbankov,38 K. Lang,35 W. M. Lee,11 S. Lin,8 J. Liu,40 M. Lokajícek,2 J. Lozier,4 S. Luchuk,18 K. Maan,27 S. Magill,1 W. A. Mann,36 M. L. Marshak,26 K. Materna,11 V. Matveev,18 D. P. Méndez,33 M. D. Messier,17 H. Meyer,39 T. Miao,11 W. H. Miller,26 S. R. Mishra,29 R. Mohanta,15 A. Moren,25 L. Mualem,4 M. Muether,29 S. Mufson,17 R. Murphy,17 J. Musser,17 J. K. Nelson,40 R. Nichol,37 E. Niner,17,11 A. Norman,11 T. Nosek,5 Y. Oksuzian,38 A. Olsheskiy,22 T. Olson,36 J. Paley,11 P. Pandey,10 R. B. Patterson,4 G. Pawloski,26 D. Pershey,4 O. Petra,22 R. Petti,29 S. Phan-Budd,41 R. K. Plunkett,11 R. Poling,26 B. Potukuchi,21 C. Principato,38 F. Pshas,17 A. Radovic,40 R. A. Rameika,11 B. Rebel,11 B. Reed,30 D. Rocco,26 P. Rojas,8 V. Ryabov,23 K. Sachdev,11,26 P. Sail,35 O. Samoylov,22 M. C. Sanchez,19 R. Schroeter,14 J. Sepulveda-Quiroz,19 P. Shanahan,11 A. Sheshukov,22 J. Singh,27 J. Singh,21 P. Singh,10 V. Singh,3 J. Smolik,9 N. Solomey,39 E. Song,38 A. Sousa,6 K. Soustruznik,5 M. Strait,26 L. Suter,11 R. L. Talaga,1 M. C. Tamsett,25 P. Tas,5 R. B. Thayyuthallil,7 J. Thomas,37 X. Tian,29 S. C. Tognini,12 J. Tripathi,27 A. Tsaris,11 J. Urheim,17 P. Vahle,40 J. Vasil,17 L. Vinton,33 A. Vold,26 T. Vrba,9 B. Wang,31 M. Wetsste,19 D. Whittington,17 S. G. Wojcicki,32 J. Wolcott,36 N. Yadav,13 S. Yang,6 J. Zalesak,2 B. Zamorano,33 and R. Zwaska11

(The NOvA Collaboration)

1Argonne National Laboratory, Argonne, Illinois 60439, USA
2Institute of Physics, The Czech Academy of Sciences, 182 21 Prague, Czech Republic
3Department of Physics, Institute of Science, Banaras Hindu University, Varanasi, 221 005, India
4California Institute of Technology, Pasadena, California 91125, USA
5Charles University, Faculty of Mathematics and Physics, Institute of Nuclear Physics, Prague, Czech Republic
6Department of Physics, University of Cincinnati, Cincinnati, Ohio 45221, USA
7Department of Physics, Cochin University of Science and Technology, Kochi 682 022, India
8Department of Physics, Colorado State University, Fort Collins, CO 80523-1875, USA
9Czech Technical University in Prague, Brehova 7, 115 19 Prague 1, Czech Republic
10Department of Physics and Astrophysics, University of Delhi, Delhi 110007, India
11Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA
12Instituto de Física, Universidade Federal de Goiás, Goiânia, Goiás, 74690-900, Brazil
13Department of Physics, IIT Guwahati, Guwahati, 781039, India
14Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA
15Department of Physics, School of Physics, University of Hyderabad, Hyderabad, 500 046, India
16Department of Physics, IIT Hyderabad, Hyderabad, 502 205, India
17Indiana University, Bloomington, Indiana 47405, USA
18Inst. for Nuclear Research of Russia, Academy of Sciences 7a, 60th October Anniversary prospect, Moscow 117312, Russia
19Department of Physics and Astronomy, University of California at Irvine, Irvine, California 92697, USA
20Department of Physics and Astronomy, University of California at Irvine, Irvine, California 92697, USA
21Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, USA
22Joint Institute for Nuclear Research, Dubna, Moscow region 141980, Russia
23Nuclear Physics Department, Lebedev Institute, Leningrad Prospect 53, 119991 Moscow, Russia
24Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA
25Department of Physics and Astronomy, University of Minnesota Duluth, Duluth, Minnesota 55812, USA
26School of Physics and Astronomy, University of Minnesota Twin Cities, Minneapolis, Minnesota 55455, USA
27Department of Physics, Panjab University, Chandigarh, 106 014, India
28Department of Physics, University of South Alabama, Mobile, Alabama 36688, USA
29Department of Physics and Astronomy, University of South Carolina, South Carolina 29208, USA
This Letter reports updated results on the rate of $\nu_\mu \rightarrow \nu_e$ transitions in the NOvA experiment [1] and constraints on oscillation parameters from the first combined fit of $\nu_e$ appearance and $\bar{\nu}_\mu$ disappearance data. The measurement, also probed by MINOS [2] and T2K [3] experiments, is sensitive to three unknowns in neutrino physics: the octant of $\theta_{23}$ (whether $\theta_{23}$ is less than, equal to, or greater than $\pi/4$), the neutrino mass hierarchy, and the amount of CP violation in the lepton sector. At the baseline and neutrino energy range of the NOvA experiment the probability for $\nu_\mu$ to oscillate to $\nu_e$ is primarily proportional to the combination $\sin^2 \theta_{23} \sin^2 2\theta_{13}$. The disappearance of muon neutrinos is sensitive to the mixing angle $\theta_{23}$ which is relatively weakly constrained to be near-maximal ($\sin^2 \theta_{23} \approx 0.5$) [2–4]. Reactor neutrino measurements tightly constrain $\sin^2 2\theta_{13}$ at 0.085 ± 0.005 [5–7]. The coherent forward scattering of the neutrino beam with electrons in the Earth enhances the electron neutrino appearance probability in the case of normal mass hierarchy (NH), where $\Delta m^2_{32} > 0$, and suppresses it for inverted mass hierarchy (IH), where $\Delta m^2_{32} < 0$. The possible violation of CP symmetry in the lepton sector is parameterized by $\delta_{CP}$. CP-conserving oscillations occur if $\delta_{CP} = 0$ or $\pi$, while $\nu_e$ appearance is enhanced around $\delta_{CP} = 3\pi/2$, and suppressed around $\delta_{CP} = \pi/2$. At NOvA’s energy and baseline the impact of these three factors on the $\nu_e$ appearance probability are of similar magnitudes, which can lead to degeneracies between them, particularly when analyzing oscillations in neutrinos alone. For antineutrinos, the mass hierarchy and CP phase have the opposite effect on the oscillation probability, while increasing values of $\sin^2 \theta_{23}$ increase the appearance probabilities for $\nu_e$ and $\bar{\nu}_e$ alike.

NOvA [8] observes neutrinos produced in Fermilab’s NuMI [9] beamline in two detectors. The Far Detector (FD) is located on the surface, 14.6 mrad off the central beam axis, 810 km from the neutrino parent production source. The Near Detector (ND) is located 100 m underground, 1 km from the source and measures the neutrino beam spectrum before oscillations occur. It is positioned to maximize the overlap between the neutrino energy spectra observed at the two detectors. At these locations, the beam is peaked around 2 GeV with neutrino energies mainly in the 1 to 3 GeV range. According to simulations, the neutrino beam at the ND is predominantly $\nu_\mu$, with 1.8% $\bar{\nu}_\mu$ and 0.7% $\nu_e +\bar{\nu}_e$ components for neutrino energies between 1 and 3 GeV.

The two functionally equivalent detectors [1, 4, 8, 10] are constructed from planes of extruded PVC cells [11]. The cells have a rectangular cross section measuring 3.9 cm by 6.6 cm and are 15.5 m (3.9 m) long in the FD (ND). Planes alternate the long cell dimension between vertical and horizontal orientations perpendicular to the beam. Each cell is filled with liquid scintillator [12]. Light is collected by a loop of wavelength-shifting fiber inside the cell. The fiber ends terminate on a single pixel of an avalanche photodiode (APD) [13]. The FD (ND) has a total active mass of 14 kt (193 t). In the fiducial region, the detectors are 62% scintillator by mass.

The data analyzed were collected between February 6, 2014 and May 2, 2016. The exposure is equivalent to $6.05 \times 10^{20}$ protons-on-target (POT) collected in the full detector and corresponds to more than double the exposure used in previous results [1, 4]. The fiducial mass for the full detector is 10.3 kt. The average neutrino beam power increased from 250 kW to 560 kW during the data-taking period.

Measuring electron-neutrino appearance requires identification of charged-current (CC) interactions of $\nu_e$ and understanding the various backgrounds that are also selected at the FD. The signature of $\nu_e$ CC interactions in the NOvA detectors is an electromagnetic shower plus...
any associated hadronic recoil energy. The largest background arises from Neutral Current (NC) interactions of beam neutrinos that produce \( \pi^0 \) which decay to photons that mimic the signature of an electron. The intrinsic \( \nu_e \) component of the NuMI beam represents an irreducible background to this search. Charged current interactions of \( \nu_\mu \) with a short muon track and a hadronic shower with some electromagnetic activity comprise a smaller background. Other small backgrounds include cosmic ray induced events, particularly where a photon or a neutron enters from the sides of the detector and charged-current interactions of \( \nu_e \), which mostly occur above 3 GeV.

For this analysis a new \( \nu_e \) CC classifier was developed to select a signal sample with improved purity and efficiency. The Convolutional Visual Network (CVN) [14] is a convolutional neural network and was designed using deep learning techniques from the field of computer vision [15, 16]. Recorded hits in the detectors are formed into clusters by grouping hits in time and space to isolate individual interactions [17, 18]. The CVN classifier takes the hits from these clusters, without any further reconstruction, as input and applies a series of trained linear operations to extract complex, abstract classifying features from the image. A multilayer perceptron [19, 20] at the end of the network uses these features to create the classifier output. Training is conducted using a mixture of simulated FD \( \nu_\mu \) CC, \( \nu_e \) CC, \( \nu_\tau \) CC, and NC events as well as a sample of FD cosmic data.

The NOvA simulation chain uses FLUKA [21], GEANT4 [22], FLUGG [23], GENIE [24] and a custom detector simulation [25] to model neutrino production in the beamline and subsequent interaction in the detector. Neutrino scattering off substructure in the nucleus is added to the simulation using an empirical model of multinucleon excitations and long range correlations [26–29]. The implementation of this model in the NOvA simulation is tuned to match an observed excess of events in data over simulation in bins of reconstructed three-momentum transfer [30]. Additionally, the rate of non-resonant single pion production in charged-current interactions is effectively reduced by 50%, motivated by a recent reanalysis of \( \nu_\mu \)-deuteron pion-production data [31, 32].

For the purpose of energy reconstruction and event containment, the event cluster is further reconstructed to determine particle paths. A Hough transform is applied to the cluster to identify global features, characterized as Hough lines[33]. The intersections of these lines seed an algorithm to produce a three dimensional vertex for the cluster [34]. In both the horizontal and vertical detector views hits are grouped into prongs radiating from the vertex [35, 36]. Prongs are then matched between the views based on energy deposition characteristics.

The energy responses of the detectors are calibrated using minimum ionizing energy deposits in a region 1 to 2 meters from the end of tracks corresponding to stopping cosmic ray muons. To reconstruct the electron neutrino candidate energy, the prong with the most calorimetric energy is assumed to be an electromagnetic shower caused by the outgoing electron. The remaining energy deposits in the event are attributed to the hadronic recoil system. The reconstructed \( \nu_e \) energy is taken as a quadratic function of the electromagnetic and hadronic calorimetric responses. The function is a parameterization of the simulated true electron neutrino energy in relation to these quantities, and yields an energy resolution of ~7% in both detectors.

To suppress the cosmic ray induced background in the FD, selected events are required to be in a 12\,\mu m window centered on the 10\,\mu m beam spill. A large fraction of cosmic events deposit energy close to the detector edges and are removed due to containment requirements. Requiring a small reconstructed transverse momentum fraction with respect to the beam direction rejects cosmic events with angles too steep to be consistent with a NuMI beam event. The cosmic background rejection criteria are tuned using neutrino beam simulation and a large sample of cosmic data recorded asynchronously with the neutrino beam.

The maximum of the \( \nu_e \) appearance signal is expected just below the peak neutrino energy at NOvA. Restricting the energy range of selected events to 1-3 GeV removes a large fraction of the NC and cosmic backgrounds which are predominately of lower reconstructed energy, and intrinsic \( \nu_e \) CC events which dominate at higher energies. We similarly constrain the length of the longest track and number of hits in an event to remove clear muon tracks or poorly reconstructed events. Other than containment requirements, the \( \nu_e \) CC selection criteria in the ND are very similar to those in the FD.

The selection criteria are chosen to maximize the figure of merit defined as \( S / \sqrt{S + B} \), where \( S \) and \( B \) are the number of signal and background events, respectively. The final \( \nu_e \) selection criteria select contained appearance signal with 73.5% efficiency and 75.5% purity, representing a gain in sensitivity of 30% compared to the \( \nu_e \) classifiers used in the previously reported results [1]. These criteria also reject 97.6% of the NC and 99.0% of the \( \nu_\mu \) CC beam backgrounds. The cosmic ray backgrounds are suppressed by seven orders of magnitude, and only 0.53 ± 0.14 cosmic events are estimated to be selected in the final \( \nu_e \) appearance sample based on the performance of \( \nu_e \) selection criteria on cosmic data. Of the beam backgrounds that pass all \( \nu_e \) selection, 91% contain some form of energetic electromagnetic shower. To further improve the statistical power of this analysis, events selected in the FD are split into three \( \nu_e \) classifier bins, containing signal \( \nu_e \) CC events with low, medium and high purity. The analysis is performed in four energy bins between 1 and 3 GeV for each of the three classifier bins.

The ND has negligible \( \nu_e \) appearance signal, and is used to estimate the beam neutrino induced background
FIG. 1: Reconstructed energy of events selected in the ND data and simulation by the $\nu_e$ CC selection criteria in the three $\nu_e$ classifier (CVN) bins. The left-most panel is the lowest purity classifier bin, while the right-most has the highest purity.

rates to the appearance measurement. According to simulation, the kinematics of the events that pass the $\nu_e$ CC selection criteria in the ND are representative of and adequately cover those selected in the FD. Figure 1 shows that there is an overall $\sim 10\%$ excess of data over simulation in the $\nu_e$ CC selected events in the ND. Since the NC, $\nu_\mu$ CC and beam $\nu_e$ CC background components are affected differently by oscillations, the total background selected in the ND data is broken down into these components which are then used to estimate the corresponding components in the FD.

Both the $\nu_\mu$ and intrinsic $\nu_e$ components of the beam peak arise primarily from pions decaying through the process ($\pi^+ \rightarrow \mu^+ + \nu_\mu$), as well as the subsequent muon decay ($\mu^+ \rightarrow e^+ + \bar{\nu}_\mu + \nu_e$). At higher energies they originate from kaon decays. The pion and kaon hadron yields can be derived from the low and high-energy $\nu_\mu$ CC rate in the ND data and are used to correct the $\nu_e$ CC rate in the simulation. Pion yields are adjusted in bins of transverse and longitudinal pion momentum, while the kaon yield is simply scaled. From this method, it is inferred that the kaon yield is higher by $17\%$ and the pion yield lower by $3\%$ than predicted by the simulation. This results in an overall $1\%$ increase in the estimated intrinsic $\nu_e$ CC background rate in the 1 to 3 GeV range in the ND.

Some of the $\nu_\mu$ CC interactions that are a background to the $\nu_e$ CC selection have a muon hidden in the shower associated with the hadronic recoil. In these events, the time-delayed electron from muon decay (Michel electron) may often be found. The hadronic recoil system also produces this signature due to the presence of charged pions that decay to muons. However, on average, $\nu_\mu$ CC interactions have one more Michel electron than $\nu_e$ CC and NC interactions. The $\nu_\mu$ CC and NC background components are varied in each bin of energy and $\nu_e$ classifier to obtain the best match to the distribution of the number of Michel electron candidates in data. The intrinsic $\nu_e$ CC background component is held fixed at the value obtained from the pion and kaon yield analysis. This method leads to an integrated increase of $17.7\%$ and $10.4\%$ in the $\nu_\mu$ CC and NC background rates relative to those predicted by the ND simulation. These corrections derived from the ND data account for the $10\%$ discrepancy with simulation and are applied to the background spectra in the FD simulation in the analysis bins. The spectra are then weighted by the appropriate 3-flavor oscillation probability to obtain the final estimates of the beam backgrounds in the FD. After applying these data-driven constraints, the predicted background composition in the FD for this analysis is $45.3\%$ NC, $38\%$ intrinsic $\nu_e$ CC, $8.4\%$ $\nu_\mu$ CC, $1.8\%$ $\nu_\tau$ CC, and $6.5\%$ cosmic events.

The $\nu_e$ appearance signal expected in the FD is also constrained by the observed neutrino beam spectrum in the ND. A sample of $\nu_\mu$ candidates are selected in the ND data using the latest $\nu_\mu$ selection criteria as described in [30], and the underlying true energy spectrum is derived from a reconstructed to true energy migration matrix. The spectrum of true $\nu_e$ CC signal events selected in the FD simulation is corrected by the ratio of the $\nu_\mu$ CC true energy spectrum derived from ND data to the simulated $\nu_\mu$ CC spectrum. The adjusted FD signal spectrum is weighted by the $\nu_e$ appearance probability and mapped back to the reconstructed energy spectrum for the final estimate of $\nu_e$ appearance signal. This extrapolation is carried out for the energy spectra in all three $\nu_e$ classifier bins. Figure 2 shows the variation in the number of FD events predicted as a function of the assumed oscillation parameters.

The ND data are also used to verify the simulated $\nu_e$ CC selection efficiency. For events that pass the $\nu_\mu$ CC selection criteria in ND data and simulation, the energy deposits along the reconstructed track of the candidate muon are removed [37]. An electron with the same energy and direction is simulated in its place to construct $\nu_e$ CC-like interactions in both data and simulation. The event is reconstructed again with the electron shower embedded in it and the $\nu_e$ selection cuts are applied. The efficiency of the $\nu_e$ CC selection criteria in the ND between data and simulation for identifying neutrino events with inserted electrons matches to within $1\%$.

Systematic uncertainties are evaluated by reweighting or generating new simulated event samples modified to account for each uncertainty in the ND and FD. The full analysis, including background component estimation in the ND data and extrapolation to FD, is performed with these systematically shifted simulation samples to predict the altered signal and background spectra at the FD. Calibration and normalization are the leading sources of systematic uncertainty for background and signal, respectively. Other sources of systematic uncertainty
considered include neutrino flux, modeling of neutrino interactions and detector response. The overall effect of the uncertainties summed in quadrature on the total event count is 5.0% (10.5%) on the signal (background). The statistical uncertainties of 20.1% (34.9%) on the signal (background) therefore dominate.

After the event selection criteria and analysis procedures were finalized, inspection of the FD data revealed 33 $\nu_e$ candidates, of which $8.2 \pm 0.8$ (syst.) events are predicted to be background [42]. Figure 3 shows a comparison of the event distribution with the expectations at the best fit point as a function of the classifier variable and reconstructed neutrino energy.

To extract oscillation parameters, the $\nu_e$ CC energy spectrum in bins of event classifier is fit simultaneously with the FD $\nu_\mu$ CC energy spectrum [30]. The NOvA $\nu_\mu$ disappearance result constrains $\sin^2 \theta_{23}$ around degenerate best fit points of 0.404 and 0.624. The likelihood between the observed spectra and the Poisson expectation in each bin is computed as a function of the oscillation parameters $|\Delta m_{32}^2|$, $\theta_{23}$, $\theta_{13}$, $\delta_{CP}$, and the mass hierarchy. Each source of systematic uncertainty is incorporated into the fit as a nuisance parameter, which varies the predicted FD spectrum according to the shifts determined from systematically shifted samples. Where systematic uncertainties are common between the two data sets, the nuisance parameters associated with the effect are correlated appropriately. Gaussian penalty terms are applied to represent the estimates of the 1$\sigma$ ranges of these parameters, and the knowledge of $\sin^2 2\theta_{13} = 0.085 \pm 0.005$ from reactor experiments [38].

Figure 4 shows the regions of $(\sin^2 \theta_{23}, \delta_{CP})$ space allowed at various confidence levels. The likelihood surface is profiled over the parameters $|\Delta m_{32}^2|$ and $\theta_{13}$ while the solar parameters $\Delta m_{21}^2$ and $\theta_{12}$ are held fixed. The sig-
significances are derived using the Feldman-Cousins unified approach [39] to account for the statistical effects of low event count and physical boundaries.

Figure 5 shows the significance at which values of $\delta_{CP}$ are disfavored for each hierarchy and octant combination. The value of $\sin^2 \theta_{23}$ is profiled within the specified octant. There are two degenerate best fit points, both in the normal hierarchy, $\sin^2 \theta_{23} = 0.404$, $\delta_{CP} = 1.48\pi$ and $\sin^2 \theta_{23} = 0.623$, $\delta_{CP} = 0.74\pi$. The inverted hierarchy predicts fewer events than are observed for all values of $\delta_{CP}$ and both octants. The best-fit point in the inverted hierarchy occurs near $\delta_{CP} = 3\pi/2$ and is 0.46 $\sigma$ from the global best-fit points. The inverted mass hierarchy in the lower octant is disfavored at greater than 93% C.L. for all values of $\delta_{CP}$, and excluded at greater than 3 $\sigma$ significance outside the range $0.97\pi < \delta_{CP} < 1.94\pi$. The T2K collaboration has recently published results based on their observation of $\nu_\mu$ ($\bar{\nu}_\mu$) disappearance and $\nu_e$ ($\bar{\nu}_e$) appearance [40]. While their data favor a near-maximal value of $\theta_{23}$, they disfavor CP conservation at 90% C.L., with a weak preference for normal mass hierarchy. These observations are broadly consistent with the NOvA result.

To conclude, in the first combined fit of the NOvA $\nu_e$ appearance and $\nu_\mu$ disappearance data, the inverted mass hierarchy with $\theta_{23}$ in the lower octant is disfavored at greater than 93% C.L. for all values of $\delta_{CP}$. Future data-taking in antineutrino mode, where the impact of the mass hierarchy and CP phase are reversed with respect to their effect on neutrinos, will help resolve the remaining degeneracies in the parameters.

This work was supported by the US Department of Energy; the US National Science Foundation; the Department of Science and Technology, India; the European Research Council; the MSMT CR, GA UK, Czech Republic; the RAS, RMES, and RFBR, Russia; CNPq and FAPEG, Brazil; and the State and University of Minnesota. We are grateful for the contributions of the staffs at the University of Minnesota module assembly facility and Ash River Laboratory, Argonne National Laboratory, and Fermilab. Fermilab is operated by Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the US DOE.

* Deceased.

1. P. Adamson et al., Phys. Rev. Lett. 116, 151806 (2016).
2. P. Adamson et al., Phys. Rev. Lett. 112, 191801 (2014).
3. K. Abe et al., Phys. Rev. D 91, 072010 (2015).
4. P. Adamson et al., Phys. Rev. D 93, 051104 (2016).
5. F. P. An et al., Phys. Rev. Lett. 115, 111802 (2015).
6. S. B. Kim Nucl. Phys. B 908, 94 (2016).
7. Y. Abe et al., JHEP 1601, 163 (2016).
8. NOvA Technical Design Report, FERMILAB-DESIGN-2007-01.
9. P. Adamson et al., Nucl. Instrum. Meth. A 806, 279 (2016); NuMi Technical Design Handbook, FERMILAB-DESIGN-1998-01.
10. S. Magill, J. Phys.: Conf. Ser. 404, 012035 (2012); P. Border, et al., Nucl. Instrum. Meth. A 463, 194 (2001).
11. R. L. Talaga, et al., FERMILAB-PUB-15-049-ND-PPD.
12. S. Mufson et al., Nucl. Instrum. Meth. A 799, 1 (2015).
13. The NOvA APD is a custom variant of the Hamamatsu S8850, http://www.hamamatsu.com/us/en/product/alpha/S/4112/S8850-02/index.html.
14. A. Aurisano and A. Radovic and D. Rocco et al., JINST 11 (2016).
15. C. Szegedy et al., arXiv:1409.4842, (2015).
16. Rumelhart, D. E. and Hinton, G. E. and Williams, R. J., Nature 323, 533-536 (1986).
17. M. Baird, Ph.D. Thesis, Indiana University (2015).
18. M. Ester et al., Proc. of 2nd International Conference on Knowledge Discovery and Knowledge Engineering and Knowledge Management, 226 (1996).
19. F. Rosenblatt, Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. Spartan Books, 1962.
20. R. Reed and R. Marks, Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks. A Bradford book. MIT Press, 1999.
21. T. Bohlen et al., Nucl. Data Sheets 120, 211 (2014); A. Ferrari et al., Tech. Rep. CERN-2005-010 (2005).
22. S. Agostinelli et al., Nucl. Instrum. Meth. A 506, 250 (2003); J. Allison et al., IEEE Trans. Nucl. Sci. 53, 270 (2006).
23. M. Campanella et al., Tech. Rep. CERN-ATL-SOFT-99-004 (1999).
24. C. Andreopoulos et al., Nucl. Instrum. Meth. A 614, 87 (2010); C. Andreopoulos et al., arXiv:1510.05494.
25. A. Aurisano et al., J. Phys.: Conf. Ser. 664, 072002 (2015).
26. M. Martini and M. Ericson, Phys. Rev. C 87, 065501 (2013), 1303.7199.
27. R. Gran, J. Nieves, F. Sanchez, and M. J. Vicente Vacas,
Phys. Rev. D 88, 113007 (2013), 1307.8105.

[28] G.D. Megias et al., Phys. Rev. D 91, 073004 (2015), 1412.1822.

[29] O. Lalakulich, K. Gallmeister, and U. Mosel, Phys. Rev. C 86, 014614 (2012), 1203.2935, [Erratum: Phys. Rev.C90,no.2,029902(2014)].

[30] P. Adamson et al., Phys. Rev. Lett. 118, 151802 (2017).

[31] C. Wilkinson, P. Rodrigues, S. Cartwright, L. Thompson, and K. McFarland, Phys. Rev. D 90, 112017 (2014), 1411.4482.

[32] P. Rodrigues, C. Wilkinson, and K. McFarland, Eur. Phys. J. C 76, 474 (2016).

[33] L. Fernandes and M. Oliveira, Patt. Rec. 41, 299 (2008).

[34] M. Gyulassy and M. Harlander, Comput. Phys. Commun. 66, 31 (1991); M. Ohlsson, and C. Peterson, Comput. Phys. Commun. 71, 77 (1992); M. Ohlsson, Comput. Phys. Commun. 77, 19 (1993); R Frühwirth and A. Strandlie, Comput. Phys. Commun. 120, 197 (1999).

[35] R. Krishnapuram and J. M. Keller, IEEE Trans. Fuzzy Syst. 1, 98 (1993); M. S. Yang, and K. L. Wu, Pattern Recognition 39, 5 (2006).

[36] E. Niner, Ph.D. Thesis, Indiana University (2015).

[37] K. Sachdev, Ph.D. Thesis, University of Minnesota (2015).

[38] K. A. Olive et al. (Particle Data Group), Chin. Phys. C 38, 090001 (2014) and 2015 update.

[39] G. J. Feldman and R. D. Cousins, Phys. Rev. D 57, 3873 (1998).

[40] K. Abe et al., Phys. Rev. Lett. 118, 151801 (2017).

[41] Laske G. Bassin, C. and G. Masters. EOS Trans AGU, F897:81, 2000.

[42] The backgrounds are computed at the best fit oscillation parameters: $\sin^2 \theta_{23} = 0.40$, $\sin^2 \theta_{13} = 0.085$, $\Delta m_{32}^2 = 2.67 \times 10^{-3} \text{eV}^2$ and $\delta_{CP} = 1.48\pi$. The matter density, computed for the average depth of the NuMI beam in the earth crust for the NOvA baseline of 810 km using the CRUST2.0 [41] model, is 2.84 g/cm$^3$. 
