PHYSICAL REVIEW D **110,** 012005 (2024)

Expanding neutrino oscillation parameter measurements in NOvA using a Bayesian approach

M. A. Acero,[2] B. Acharya,[32] P. Adamson,[13] N. Anfimov,[27] A. Antoshkin,[27] E. Arrieta-Diaz,[28] L. Asquith,[40] A. Aurisano,[7] A. Back,[21] N. Balashov,[27] P. Baldi,[26] B. A. Bambah,[18] A. Bat,[3,12] K. Bays,[31,20] R. Bernstein,[13] T. J. C. Bezerra,[40] V. Bhatnagar,[34] D. Bhattarai,[32] B. Bhuyan,[16] J. Bian,[26,31] A. C. Booth,[36,40] R. Bowles,[21] B. Brahma,[19] C. Bromberg,[29] N. Buchanan,[9] A. Butkevich,[23] S. Calvez,[9] T. J. Carroll,[42,48] E. Catano-Mur,[47] J. P. Cesar,[42] A. Chatla,[18] S. Chaudhary,[16] R. Chirco,[20] B. C. Choudhary,[11] A. Christensen,[9] T. E. Coan,[39] A. Cooleybeck,[48] L. Cremonesi,[36] G. S. Davies,[32] P. F. Derwent,[13] P. Ding,[13] Z. Djurcic,[1] M. Dolce,[43] D. Doyle,[9] D. Dueñas Tonguino,[7] E. C. Dukes,[45] A. Dye,[32] R. Ehrlich,[45] M. Elkins,[25] E. Ewart,[21] P. Filip,[24] J. Franc,[10] M. J. Frank,[37] H. R. Gallagher,[43] F. Gao,[35] A. Giri,[19] R. A. Gomes,[15] M. C. Goodman,[1] M. Groh,[9,21] R. Group,[45] A. Habig,[30] F. Hakl,[22] J. Hartnell,[40] R. Hatcher,[13] M. He,[17] K. Heller,[31] V. Hewes,[7] A. Himmel,[13] B. Jargowsky,[26] J. Jarosz,[9] F. Jediny,[10] C. Johnson,[9] M. Judah,[9,35] I. Kakorin,[27] D. M. Kaplan,[20] A. Kalitkina,[27] J. Kleykamp,[32] O. Klimov,[27] L. W. Koerner,[17] L. Kolupaeva,[27] R. Kralik,[40] A. Kumar,[34] C. D. Kuruppu,[38] V. Kus,[10] T. Lackey,[13,21] K. Lang,[42] P. Lasorak,[40] J. Lesmeister,[17] A. Lister,[48] J. Liu,[26] J. A. Lock,[40] M. Lokajicek,[24] M. MacMahon,[44] S. Magill,[1] W. A. Mann,[43] M. T. Manoharan,[8] M. Manrique Plata,[21] M. L. Marshak,[31] M. Martinez-Casales,[25] V. Matveev,[23] B. Mehta,[34] M. D. Messier,[21] H. Meyer,[46] T. Miao,[13] V. Mikola,[44] W. H. Miller,[31] S. Mishra,[4] S. R. Mishra,[38] R. Mohanta,[18] A. Moren,[30] A. Morozova,[27] W. Mu,[13] L. Mualem,[5] M. Muether,[46] K. Mulder,[44] D. Myers,[42] D. Naples,[35] A. Nath,[16] S. Nelleri,[8] J. K. Nelson,[47] R. Nichol,[44] E. Niner,[13] A. Norman,[13] A. Norrick,[13] T. Nosek,[6] H. Oh,[7] A. Olshevskiy,[27] T. Olson,[17] A. Pal,[33] J. Paley,[13] L. Panda,[33] R. B. Patterson,[5] G. Pawloski,[31] O. Petrova,[27] R. Petti,[38] R. K. Plunkett,[13] L. R. Prais,[32] A. Rafique,[1] V. Raj,[5] M. Rajaoalisoa,[7] B. Ramson,[13] B. Rebel,[13,48] P. Roy,[46] O. Samoylov,[27] M. C. Sanchez,[14,25] S. Sánchez Falero,[25] P. Shanahan,[13] P. Sharma,[34] A. Sheshukov,[27] S. Shukla,[4] D. K. Singha,[18] W. Shorrock,[40] I. Singh,[11] P. Singh,[36,11] V. Singh,[4] E. Smith,[21] J. Smolik,[10] P. Snopok,[20] N. Solomey,[46] A. Sousa,[7] K. Soustruznik,[6] M. Strait,[31] L. Suter,[13] A. Sutton,[14,25,45] K. Sutton,[5] S. Swain,[33] C. Sweeney,[44] A. Sztuc,[44] B. Tapia Oregui,[42] P. Tas,[6] T. Thakore,[7] J. Thomas,[44,48] E. Tiras,[12,25] Y. Torun,[20] J. Trokan-Tenorio,[47] J. Urheim,[21] P. Vahle,[47] Z. Vallari,[5] K. J. Vockerodt,[36] T. Vrba,[10] M. Wallbank,[7] T. K. Warburton,[25] M. Wetstein,[25] D. Whittington,[41] D. A. Wickremasinghe,[13] T. Wieber,[31] J. Wolcott,[43] M. Wrobel,[9] S. Wu,[31] W. Wu,[26] Y. Xiao,[26] B. Yaeggy,[7] A. Yankelevich,[26] K. Yonehara,[13] Y. Yu,[20] S. Zadorozhnyy,[23] J. Zalesak,[24] and R. Zwaska[13]

(NOvA collaboration)

[1]*Argonne National Laboratory, Argonne, Illinois 60439, USA*
[2]*Universidad del Atlantico, Carrera 30 No. 8-49, Puerto Colombia, Atlantico, Colombia*
[3]*Bandirma Onyedi Eylül University, Faculty of Engineering and Natural Sciences, Engineering Sciences Department, 10200, Bandirma, Balıkesir, Turkey*
[4]*Department of Physics, Institute of Science, Banaras Hindu University, Varanasi, 221 005, India*
[5]*California Institute of Technology, Pasadena, California 91125, USA*
[6]*Charles University, Faculty of Mathematics and Physics, Institute of Particle and Nuclear Physics, Prague, Czech Republic*
[7]*Department of Physics, University of Cincinnati, Cincinnati, Ohio 45221, USA*
[8]*Department of Physics, Cochin University of Science and Technology, Kochi 682 022, India*
[9]*Department of Physics, Colorado State University, Fort Collins, Colorado 80523-1875, USA*
[10]*Czech Technical University in Prague, Brehova 7, 115 19 Prague 1, Czech Republic*
[11]*Department of Physics and Astrophysics, University of Delhi, Delhi 110007, India*
[12]*Department of Physics, Erciyes University, Kayseri 38030, Turkey*
[13]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[14]*Florida State University, Tallahassee, Florida 32306, USA*
[15]*Instituto de Física, Universidade Federal de Goiás, Goiânia, Goiás, 74690-900, Brazil*
[16]*Department of Physics, IIT Guwahati, Guwahati, 781 039, India*
[17]*Department of Physics, University of Houston, Houston, Texas 77204, USA*
[18]*School of Physics, University of Hyderabad, Hyderabad, 500 046, India*
[19]*Department of Physics, IIT Hyderabad, Hyderabad, 502 205, India*
[20]*Illinois Institute of Technology, Chicago Illinois 60616, USA*
[21]*Indiana University, Bloomington, Indiana 47405, USA*
[22]*Institute of Computer Science, The Czech Academy of Sciences, 182 07 Prague, Czech Republic*

2470-0010/2024/110(1)/012005(21)

Published by the American Physical Society

[23]*Institute for Nuclear Research of Russia, Academy of Sciences 7a, 60th October Anniversary prospect,*
Moscow 117312, Russia
[24]*Institute of Physics, The Czech Academy of Sciences, 182 21 Prague, Czech Republic*
[25]*Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, USA*
[26]*Department of Physics and Astronomy, University of California at Irvine, Irvine, California 92697, USA*
[27]*Joint Institute for Nuclear Research, Dubna, Moscow region 141980, Russia*
[28]*Universidad del Magdalena, Carrera 32 No 22-08 Santa Marta, Colombia*
[29]*Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA*
[30]*Department of Physics and Astronomy, University of Minnesota Duluth, Duluth, Minnesota 55812, USA*
[31]*School of Physics and Astronomy, University of Minnesota Twin Cities,*
Minneapolis, Minnesota 55455, USA
[32]*University of Mississippi, University, Mississippi 38677, USA*
[33]*National Institute of Science Education and Research, Khurda, 752050, Odisha, India*
[34]*Department of Physics, Panjab University, Chandigarh, 160 014, India*
[35]*Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA*
[36]*Particle Physics Research Centre, Department of Physics and Astronomy,*
Queen Mary University of London, London E1 4NS, United Kingdom
[37]*Department of Physics, University of South Alabama, Mobile, Alabama 36688, USA*
[38]*Department of Physics and Astronomy, University of South Carolina,*
Columbia, South Carolina 29208, USA
[39]*Department of Physics, Southern Methodist University, Dallas, Texas 75275, USA*
[40]*Department of Physics and Astronomy, University of Sussex,*
Falmer, Brighton BN1 9QH, United Kingdom
[41]*Department of Physics, Syracuse University, Syracuse New York 13210, USA*
[42]*Department of Physics, University of Texas at Austin, Austin, Texas 78712, USA*
[43]*Department of Physics and Astronomy, Tufts University, Medford, Massachusetts 02155, USA*
[44]*Physics and Astronomy Department, University College London,*
Gower Street, London WC1E 6BT, United Kingdom
[45]*Department of Physics, University of Virginia, Charlottesville, Virginia 22904, USA*
[46]*Department of Mathematics, Statistics, and Physics, Wichita State University,*
Wichita, Kansas 67260, USA
[47]*Department of Physics, William & Mary, Williamsburg, Virginia 23187, USA*
[48]*Department of Physics, University of Wisconsin-Madison,*
Madison, Wisconsin 53706, USA

(Received 13 November 2023; accepted 24 May 2024; published 10 July 2024)

NOvA is a long-baseline neutrino oscillation experiment that measures oscillations in charged-current $\nu_\mu \to \nu_\mu$ (disappearance) and $\nu_\mu \to \nu_e$ (appearance) channels, and their antineutrino counterparts, using neutrinos of energies around 2 GeV over a distance of 810 km. In this work we reanalyze the dataset first examined in our previous paper [Phys. Rev. D **106**, 032004 (2022)] using an alternative statistical approach based on Bayesian Markov chain Monte Carlo. We measure oscillation parameters consistent with the previous results. We also extend our inferences to include the first NOvA measurements of the reactor mixing angle θ_{13}, where we find $0.071 \le \sin^2 2\theta_{13} \le 0.107$, and the Jarlskog invariant, where we observe no significant preference for the CP-conserving value $J = 0$ over values favoring CP violation. We use these results to examine the effects of constraints from short-baseline measurements of θ_{13} using antineutrinos from nuclear reactors when making NOvA measurements of θ_{23}. Our long-baseline measurement of θ_{13} is shown to be consistent with the reactor measurements, supporting the general applicability and robustness of the Pontecorvo-Maki-Nakagawa-Sakata framework for neutrino oscillations.

DOI: 10.1103/PhysRevD.110.012005

I. INTRODUCTION

In the three-flavor neutrino oscillation paradigm, transitions among the three flavor eigenstates ν_e, ν_μ, ν_τ are governed by the matrix elements between these states and

---

the mass eigenstates ν_1, ν_2, ν_3.[1] Initial constraints on the elements of this Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix were obtained by numerous experiments using neutrinos with a variety of energy spectra over various baselines [1–9]. Decomposing the PMNS matrix yields a set of rotationlike "mixing angles" θ_{12}, θ_{13}, θ_{23}, and a phase δ_{CP}. Contemporary neutrino oscillation experiments seek to make precision measurements of these parameters, as well as the differences between the squared mass eigenvalues ($\Delta m^2_{ij} \equiv m^2_i - m^2_j$). These results have fundamental implications for models of neutrino mass and lepton flavor [10–14], as well as models of baryogenesis via charge-parity (CP) symmetry violation [15–19]. Where multiple experiments can access the same parameters using different neutrino flavors or energies, the overall validity of the three-neutrino framework can also be tested.

Long-baseline (LBL) accelerator neutrino oscillation experiments measure oscillations in $\nu_\mu \to \nu_\mu$ (disappearance) and $\nu_\mu \to \nu_e$ (appearance) channels. These channels constrain the mixing angles θ_{13}, θ_{23}, the mass-squared splitting Δm^2_{32}, and the CP-violating phase δ_{CP}. Current measurements of θ_{23} [20–23] are consistent with maximal mixing ($\theta_{23} = \pi/4$), which would suggest a μ-τ symmetry in the flavors' mixing into the ν_3 mass eigenstate; a nonmaximal value such as $\theta_{23} < \pi/4$ (lower octant, LO) or $\theta_{23} > \pi/4$ (upper octant, UO) would indicate a preferential coupling of ν_τ or ν_μ, respectively, with ν_3. Current experimental uncertainties on θ_{23} are the largest among the mixing angles [24]. LBL oscillation measurements, where neutrinos traverse significant quantities of matter, are also impacted by the coherent forward scattering of ν_es on electrons in the Earth [25]. This modifies the oscillation probabilities $\mathcal{P}(\nu_\mu \to \nu_e)$ and $\mathcal{P}(\bar{\nu}_\mu \to \bar{\nu}_e)$ with opposite signs. The direction of the resulting change in rate depends on whether ν_3 is the heaviest neutrino state (normal ordering, NO) or the lightest (inverted ordering, IO). Current observations from LBL (and other) experiments [20,22,26,27] prefer the NO, though the strength of the NO hypothesis depends on which measurements are included [28]. A similar story holds for δ_{CP}, where LBL experiments provide the only constraints [28].

NOvA is a long-baseline neutrino oscillation experiment that observes the ν_μ disappearance and ν_e appearance channels using neutrinos of energies around 2 GeV over a distance of 810 km. Previously [29–35], we have presented NOvA constraints on Δm^2_{32}, $\sin^2\theta_{23}$, and δ_{CP} using a classical frequentist approach. However, the Feldman-Cousins technique that is required in order to obtain correct frequentist confidence regions for these variables [20,36] poses challenges when confronted with highly degenerate sets of parameters. It also does not allow for *post hoc* transformations of the variables considered in the analysis. In this work we present a new analysis of the dataset from Ref. [35], based on Bayesian Markov chain Monte Carlo (MCMC), which enables us to extend our inferences to include θ_{13} and the Jarlskog invariant J, for which $J \neq 0$ unambiguously indicates CP violation. We use these results to examine the implications of assuming short-baseline, nuclear-reactor antineutrino constraints on θ_{13} when making measurements of other oscillation parameters in NOvA.

We also investigate the consistency of the PMNS framework by comparing the constraint from reactor experiments with our long-baseline measurement of θ_{13}.

II. THREE-FLAVOR NEUTRINO OSCILLATIONS IN NOVA

In this paper, we reanalyze data collected from an exposure of 13.6×10^{20} 14 kton-equivalent protons on target (POT) in the neutrino-enriched beam mode and 12.5×10^{20} POT in the analogous antineutrino mode. The dataset, simulations, reconstruction, and estimation of systematic uncertainties remain unchanged for this analysis. A brief overview of these components follows, with detailed descriptions available in Ref. [35]. Extensive discussion of the new analysis method, its implementations, and the resulting inferences are presented in Sec. III.

A. The NOvA experiment

NOvA observes oscillations using neutrinos from the Neutrinos at the Main Injector beamline [37] at Fermilab using two functionally identical tracking calorimeter detectors that differ primarily in size. The 0.3 kton near detector (ND), the smaller of the detectors, is located 1 km from the neutrino production target, 100 m underground. The far detector (FD), by contrast, is 14 kton and is located on the surface at Ash River, Minnesota, 810 km from the target. Both detectors are built from rectangular cells, made from polyvinyl chloride and of 3.9×6.6 cm^2 cross sectional area, with 3.9 m (ND) or 15.5 m (FD) length. These are arranged in alternating horizontal and vertical planes and filled with a mineral-oil-based liquid scintillator. A stack of alternating active planes and steel plates is placed downstream of the remainder of the ND to range out muons while a small rock overburden is placed above the FD to aid in rejecting the cosmic-ray background. The detectors are placed 14.6 mrad from the central axis of the neutrino beam to receive a narrow-band neutrino flux predominantly between 1 and 3 GeV.

B. Simulation and selection

We use GEANT4 (v10.4) [38,39] to simulate the production of hadrons from interactions of the primary proton beam with the target as well as their transport through the beam optics. These simulations are reweighted using the

---

[1] The same is true for their antineutrino counterparts $\bar{\nu}_e$, $\bar{\nu}_\mu$, $\bar{\nu}_\tau$ and $\bar{\nu}_1$, $\bar{\nu}_2$, $\bar{\nu}_3$. Throughout this paper the symbol ν will refer to both neutrinos and antineutrinos unless otherwise specified.

Package to Predict the FluX [40] to include constraints from external hadron production data [41–59]. Simulated interactions of neutrinos that arise from decays of those hadrons are generated using GENIE 3.0.6 [60,61] and modified by corrections we derive from NOvA ND and external data. In particular, NOvA ND data are used to produce a NOvA-specific tune of the IFIC València 2p2h model [62,63] that describes charged-current neutrino scattering from correlated pairs of nucleons. The NOvA tune also modifies final-state interactions (simulated with the GENIE hN full intranuclear cascade model) using pion-nucleus scattering data [64–70]. The outgoing final-state particles are propagated through the detector using GEANT4 and a custom NOvA readout simulation [71].

We use groups of spatially and temporally proximate cells with activity above threshold to apply basic data quality and containment selection cuts to events in both data and simulation. Within these groups we assign vertices and reconstruct likely particle trajectories. Ultimately we divide the events into ν_μ charged-current (CC), ν_eCC, neutral-current, or cosmogenic background categories using NOvA's convolutional neural network (CNN)-based classifier [72]. Boosted decision trees (BDTs) are used to further reject cosmic backgrounds in the FD samples. Both sets of tags are utilized together to create ν_μ CC and ν_e CC candidate samples. Fully contained ν_e candidates at the FD are further divided into high and low purity samples based on the CNN score in order to enhance the signal-to-background rejection capability of the fit. To improve the statistical strength of the fit, we recover an additional sample of "peripheral" events that fail the containment or cosmic rejection BDT but pass stricter particle ID requirements. We estimate neutrino energy for ν_μ CC events using the muon track length and the total deposited calorimetric energy of the hadronic system. The energy for ν_e CC events is estimated using a function of calorimetric energy that takes as input the energy of the event's reconstructed trajectories divided into electromagnetic and hadronic components, as identified by a separate CNN-based classifier [73].

C. Near-to-far extrapolation and systematics

Predictions for the neutrino event rates at the FD are constrained using the high-statistics neutrino interactions measured in the ND. These measurements are used to devise corrections to the ND prediction, which we propagate to the FD by adjusting for the differing efficiency and flux between the detectors using simulations. We call this process "extrapolation." We apply oscillations to this data-driven prediction when comparing to FD data during inference (Sec. III B below).

Corrections to the signal spectra for both ν_μ disappearance and ν_e appearance channels arise from the ν_μ spectra at the ND. The near-to-far extrapolation for the ν_μ disappearance samples at the FD is performed in quartiles of hadronic energy fraction ($f_{\mathrm{had}} = E_{\mathrm{had}}/E_\nu$, with E_ν being the reconstructed neutrino energy and E_{had} the reconstructed hadronic energy).

Extrapolating in the f_{had} bins has the effect of grouping events that share similar hadronic system characteristics so that compatible events are constrained together in the FD sample, despite the detectors' somewhat different acceptances. Performing the oscillation fit in these bins enhances sensitivity to the oscillation dip as a function of E_ν since we achieve a finer energy resolution for bins with smaller f_{had}. We further subdivide the near-to-far extrapolation into three bins of reconstructed transverse momentum of the outgoing charged lepton p_T for both appearance and disappearance channels. Like the f_{had} subdivision, this allows us to better match the constraints from the much smaller ND to the FD, in this case adjusting for the differing containment of events with leptons that emerge at large angles relative to the beam direction (i.e., large p_T). The predictions in p_T bins are summed prior to the oscillation fit. We constrain the small beam backgrounds in the ν_e appearance channels with a similar procedure based on ND ν_e candidates after first decomposing them into neutral-current, ν_μ CC, and intrinsic ν_e categories using data-driven constraints. ($\bar{\nu}_e$ beam backgrounds are all constrained together rather than being decomposed this way.) Cosmic backgrounds for all FD samples are determined from dedicated FD cosmic data samples. The remaining minor backgrounds are estimated from simulation. More detail on these procedures can be found in our previous paper [35].

We evaluate the impact of systematic uncertainties on the analysis by repredicting the sample spectra described above with altered parameters in the simulation. Uncertainties in the neutrino flux and interaction model are treated using event reweighting. Uncertainties in the detector calibration and custom modeling of light in the detectors, on the other hand, must be fully resimulated. After suppressing those that result in negligible changes to our spectra, 67 sets of systematically shifted simulations remain from these techniques, one set for each uncertainty. Each variation in each set is extrapolated to the FD using ND data via the same process as above, which constrains the impact of the uncertainties on the predicted spectra. We use the extrapolated variations to obtain 67 parametrized interpolations of the uncertainties' constrained effect on the observables, and 67 free parameters corresponding to the latter are what are marginalized during the oscillation inference in Sec. III.

Figure 1 shows the reconstructed energy spectra of the data observed at the FD using the ν_e and ν_μ CC selections described in Sec. II B, separated by the neutrino-enriched and antineutrino-enriched beam modes. Overlaid on these plots are bands of FD predictions produced according to the extrapolation procedure above. These illustrate the spectra predicted using the 68.3%, 95.4%, and 99.7% highest probability values of the systematics just described and of the relevant oscillation parameters, determined using the

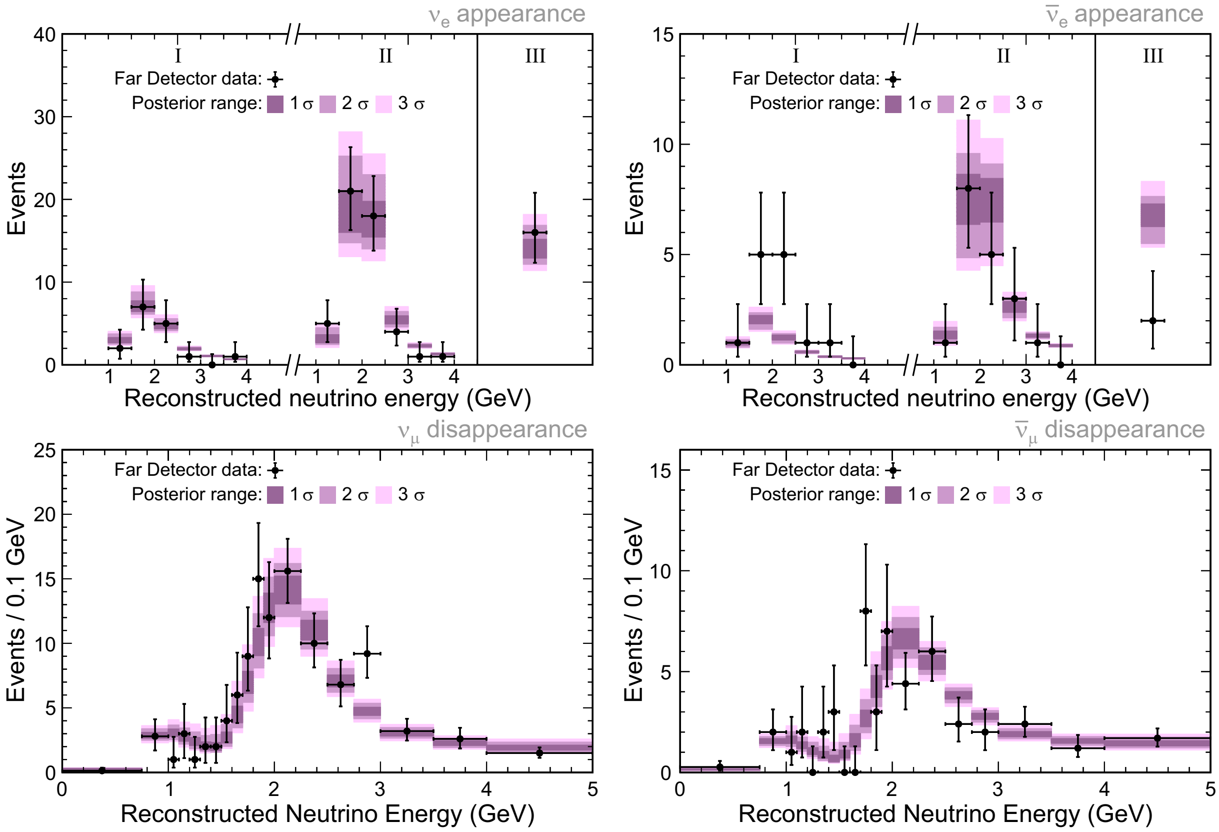


FIG. 1. Reconstructed neutrino energy distribution of selected data events (black crosses) in FD ν_eCC samples (top) and FD ν_μCC samples (bottom) in neutrino-enriched beam mode (left) and antineutrino-enriched beam mode (right). The colored bands correspond to the range of 1σ (darkest), 2σ, and 3σ (lightest) of the extrapolated FD spectra produced using the combinations of the oscillation and systematic parameters sampled by our MCMC algorithms, illustrating the posterior distributions resulting from our data fit (described further in Sec. III B). The FD ν_e samples are divided into bins of low (I) and high (II) particle ID confidence as well as the peripheral (III) sample discussed in Sec. II B. The four E_{frac} ν_μ subsamples have been combined together in each of the lower two plots.

MCMC algorithms that will be detailed in the next section. For economy, here and in the following these ranges are labeled with the conventional shorthand in terms of Gaussian standard deviations σ, using the corresponding z scores of 1σ, 2σ, and 3σ, respectively.

A Poisson likelihood [24] computed over the bins between the FD data and ND-constrained predictions such as those shown here forms the likelihood component of the posterior computed in Bayes' theorem below.

III. OSCILLATION PARAMETER INFERENCES USING MARKOV CHAIN MONTE CARLO

We derive posterior probability density distributions for relevant oscillation parameters using Bayes' theorem [74]. Marginalizing away the nuisance parameters of our model, which include the tens of systematic uncertainties described in Sec. II C, is a challenging problem because it requires an integral over many dimensions. We therefore turn to a Monte Carlo method for computing the posterior: MCMC. In MCMC, we draw a collection of sequential samples from the posterior with a frequency proportional to the posterior probability density. Histograms that approximate the posterior shape (with accuracy governed by the sample count) may be computed in any variable(s) of interest using these samples. In so doing, any dimensions not explicitly summed are implicitly marginalized.[2] In our implementations, we obtain MCMC samples that draw values from the oscillation parameter space (Δm^2_{32}, $\sin^2\theta_{13}, \sin^2\theta_{23}, \delta_{CP}$) and the 67 aforementioned parameters corresponding to our systematic uncertainties, for a total of 71 degrees of freedom.

Numerous algorithms for obtaining MCMC samples exist; we have implemented two for this analysis. The conclusions obtained from them agree with one another. Though descriptions of both methods are readily found in the literature, there are certain implementation choices that must be made for each, which we discuss in Sec. III A.

---

[2] For an accessible introduction to MCMC methods, the reader is referred to Ref. [75]. An exhaustive treatment may be found at Ref. [76].

Sec. III B lays out our resulting inferences on the parameters.

A. NOvA MCMC implementations

1. *"ARIA":* MR^{2}T^2 *algorithm*

The traditional MCMC algorithm, the Metropolis-Rosenbluth-Rosenbluth-Teller-Teller (MR^{2}T^2) method,[3] is straightforward. We call our implementation ARIA, in honor of Arianna Rosenbluth, who first implemented the method in machine code [80]. Proceeding from an initial seed in the parameter space, subsequent samples are selected by proposing a jump to a new set of coordinates, and accepting or rejecting that proposal according to an acceptance rule [24]. This process is repeated until a sufficient number of samples have been collected. There is no explicit stopping criterion. The method also does not specify the distribution to be used in the proposal algorithm. In our implementation, we use a common choice, which is of a multivariate Gaussian. We determine the characteristic length scales and correlations of the Gaussian empirically in order to optimize sampling efficiency (see Appendix A). Our ARIA results below have 10^5 effective samples (see Appendix B).

2. *"Stan": Hamiltonian MCMC method*

Though the MR^{2}T^2 method proposes samples quickly, they are typically highly autocorrelated. Its sampling proposals can also be inefficient if the posterior has both sharply concentrated and broader regions. Other MCMC methods have been developed to address these shortcomings, including one called "Hamiltonian" MCMC inference (HMCMC). We implemented a C++ interface to the Stan modeling platform [81] to obtain HMCMC samples.

The main difference between HMCMC and MR^{2}T^2 is how proposals are generated. Rather than proposing randomly, HMCMC views the posterior surface as a topographical one that can be explored by a fictitious particle. Samples correspond to this particle's trajectories under the influence of a gravitational potential whose gradient aligns with the direction of higher posterior density. Endowing the particle with an initial momentum that counterbalances the centripetal force from gravitation produces stable trajectories that traverse the highest density region of posterior space [82]. HMCMC does this by numerically integrating Hamilton's equations for the fictitious particle system with its position $\vec{q}$ (which correspond to the parameters of interest) and momentum $\vec{p}$ coordinates, and a Hamiltonian $H = -\log(\text{posterior})$. This approach produces samples that are nearly uncorrelated at the expense of additional computing cycles to compute the gradient of the posterior. The 7×10^7 Stan samples we obtained may be compared to the 10^5 effective samples mentioned above for ARIA. We find Stan's default choices of the sampling distribution for the pseudoparticle kinetic energies and the integration stopping condition to be sufficient for our needs (see Appendix C).

Its topographical nature means that unlike MR^{2}T^2, HMCMC is ill suited to parameters that assume only one of a discrete set of values, which would manifest as discontinuities in the trajectories considered. This presents a difficulty in neutrino oscillation parameter inference, where the absolute value of Δm^2_{32} is known with relatively good precision, but its sign (which determines the neutrino mass ordering) remains an important unknown. While it is possible to allow HMCMC to explore the entire range of Δm^2_{32}, we find in practice that this results in poor exploration, as few trajectories manage to "jump" across the wide disfavored region $|\Delta m^2_{32}| \lesssim 2 \times 10^{-3}\ \text{eV}^2$. Instead, at the end of each trajectory determination, we introduce a separate MR^{2}T^2-like step which considers the possibility of changing the sign of Δm^2_{32} according to the prior probability chosen for it (50%, i.e., uniform prior; see Sec. III A 3 below). If the acceptance ratio between both mass orderings satisfies the MR^{2}T^2 criteria, the proposed sign is retained; if not, it is reverted to its previous value.[4]

3. *Choices of prior*

In Bayesian inference, the posterior probabilities are influenced by the choice of prior probability densities, which encode assumptions made about the parameters before the data is examined. If the data used for a measurement is sufficient, its constraint on the posterior will typically overwhelm the prior, rendering the prior choice unimportant. However, when data is sparse, or when the prior vanishes in regions of the parameter space, the choice of prior may affect the result. The priors we choose differ according to the parameters being considered.

Parameters of interest: $|\Delta m^2_{32}|$, $\text{sgn}(\Delta m^2_{32})$, $\sin^2\theta_{23}$. We prefer to use "uninformed" priors, which do not favor any particular value, for the physics parameters we intend to measure directly. In practice, this usually amounts to a prior uniform in the variable in question (which, in the special case of the binary parameter $\text{sgn}(\Delta m^2_{32})$, corresponds to 50% probability for each of the two options). However, uniformity is not preserved under a change of variable: for instance, a prior uniform in θ_{23} is not uniform in the

---

[3] Historically this method was known as the "Metropolis" (or "Metropolis-Hastings") method, referring to the first author of the seminal papers [77,78]. We follow more recent convention and hereafter refer to it by all of the authors' names. See also Ref. [79].

[4] Though explicitly treating the mass ordering in this way is not required by the ARIA method, since it does not require continuity between samples, we find that it significantly improves the sampler performance. Thus we use it for ARIA as well.

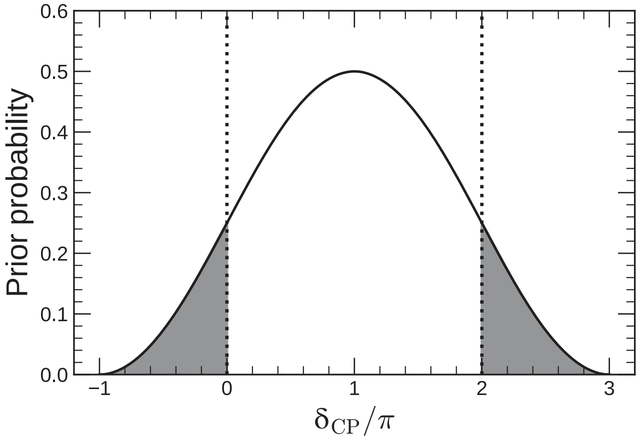


FIG. 2. Special prior distribution from Eq. (1) used for δ_{CP}. The sum of the prior for any point within the region between the dotted lines with the corresponding point in the gray shaded areas that differs from it by ±2π is always $\frac{1}{2}$.

measured variable $\sin^2\theta_{23}$. In our results below, we have studied the impact of priors uniform both in a particular variable and relevant functions of it, and we report when the prior choice significantly affects the results.

Parameter of interest: δ_{CP}. While δ_{CP} is intended to have a uniform prior as well, it receives additional special treatment. Its cyclical nature as the phase of a complex number results in an infinite set of values having identical consistency with the data. This can cause MCMC samplers never to converge on a single value for δ_{CP} and consequently to sample much more slowly. To combat this problem, we developed a novel special prior over δ_{CP}:

$$\Pi(\delta_{CP}) = \begin{cases} \frac{1}{2}\sin^2\left(\frac{1}{4}(\delta_{CP}+\pi)\right), & -1 \le \delta_{CP}/\pi \le 3 \\ 0, & \text{otherwise} \end{cases}. \tag{1}$$

This function is illustrated in Fig. 2. This prior forces the samplers to remain near a single phase of δ_{CP}, $0 \le \delta_{CP}/\pi \le 2$, as the prior vanishes outside of $[-1, 3]$. Because it makes the transition to the vanishing regions in a differentiable manner, this prior is suitable for use with HMCMC. Moreover, the sum of the prior's value at every point $0 \le \delta_{CP}/\pi \le 2$ with that of all the points outside that interval that share the same phase is a value that does not depend on δ_{CP}:

$$\sum_{n=-\infty}^{\infty} \Pi(\delta_{CP} + 2\pi n) = \frac{1}{2}. \tag{2}$$

This implies that the prior behaves identically to a uniform prior when used in conjunction with the oscillation probability. For the rest of this paper, whenever we refer to a "uniform" prior in δ_{CP}, we mean this prior.

When in the subsequent sections we study the effect of choosing a prior uniform in δ_{CP} vs $\sin(\delta_{CP})$, we reweight the MCMC samples obtained with the prior above, taking the Jacobian factor $\partial(\sin(\delta_{CP})) = \cos(\delta_{CP})$ as the weight.

Parameter of interest: $\sin^2 2\theta_{13}$. As discussed further in Sec. III B below, we consider two separate cases for $\sin^2 2\theta_{13}$: one where its prior is treated identically to that of $\sin^2\theta_{23}$ (see Sec. III A 3), and one where measurements from short-baseline reactor antineutrino oscillations are applied as a constraint. In the latter case, we impose a Gaussian prior with standard deviation obtained from the 2019 world average of reactor measurements [83], analogous to the treatment in our most recent frequentist result [35]: $\sin^2 2\theta_{13} = 0.085 \pm 0.003$.

Systematic uncertainties. Our implementation varies systematic uncertainty parameters in units of their standard deviation from the model's nominal value. The response of the prediction as a function of standard deviation from nominal for each systematic uncertainty is the same as in the frequentist analysis [84]. The model being fitted uses predictions constructed using the extrapolation described in Sec. II C, which has the effect of implicitly applying the relevant constraints from the ND. As this method of employing the ND information does not rely on fitting model parameters to the ND data—rather, it effectively reduces the systematic uncertainty variations to consist of only the components not shared by the ND and FD—the systematic uncertainty parameters remain *a priori* uncorrelated in our MCMC sampling. Thus, for all systematic uncertainties we use a unit Gaussian distribution, uncorrelated from all other parameters, as their prior.

B. Oscillation inferences: Results and discussion

In this section we describe our inferences regarding the PMNS neutrino oscillation parameters, given the data, model, and the Bayesian methodology described above. Two separate sets of results are obtained. The first uses a Gaussian prior on $\sin^2 2\theta_{13}$, imposing a constraint from reactor antineutrino experiments (Secs. III B 2 and III B 3). The second uses a prior uniform in $\sin^2 2\theta_{13}$, yielding results constrained only by the NOvA data (Sec. III B 4). Samples obtained from ARIA and Stan produce essentially identical distributions in all the variables considered (Appendix D).

1. *Goodness of fit*

To evaluate the goodness of the fit for this Bayesian analysis we use posterior predictive p values (PPP) [85]. In a PPP test, the coordinates from each MCMC sample are used to form a prediction for the observed spectra. A Poisson χ^2 statistic is computed between each prediction and the data, which we denote as χ^2_{data}. A second prediction is made for each sample by applying Poisson fluctuations to the prediction above, and a second $\chi^2_{\text{pseudodata}}$ calculated between this pseudodata and the original prediction.

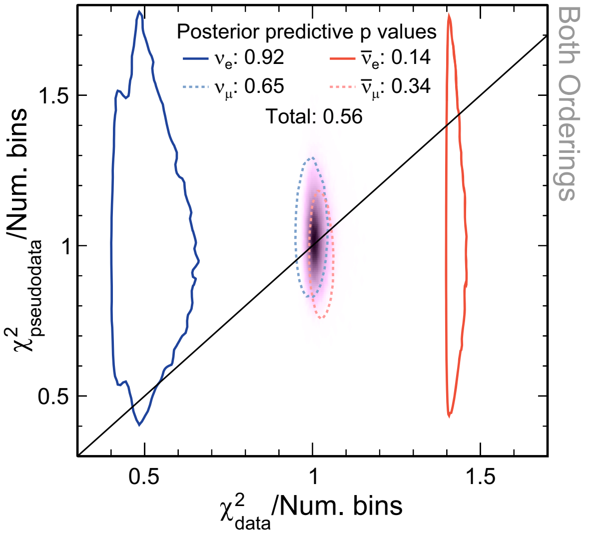


FIG. 3. Posterior predictive p values from real data MCMC samples, with θ_{13} constraint from reactor experiments applied. The purple distribution is a scatterplot of the binned χ^2 computed between the model and real data spectra (x axis), against a similar χ^2 between the model and pseudodata spectra (y axis), both divided by the number of degrees of freedom in the fit, computed for each MCMC sample. The dashed and solid contours show 1σ intervals from the same posterior-predictive distributions calculated only for ν_e (dark blue, left solid), $\bar{\nu}_e$ (red, right solid), ν_μ (light blue, left dashed), and $\bar{\nu}_\mu$ (light red, right dashed) data samples. The posterior predictive p values shown in the legend are the fraction of each distribution that lies above the diagonal $\chi^2_{\text{data}} = \chi^2_{\text{pseudodata}}$ line (black).

Because the data spectra are unchanged in this process, χ^2_{data} incorporates only variations in the oscillation parameters and systematic uncertainties. Conversely, since the pseudodata distributions have Poisson fluctuations applied but use the same oscillation parameters and systematic uncertainty pulls as the base model for each MCMC sample, $\chi^2_{\text{pseudodata}}$ treats only statistical uncertainties.

The distribution of these $(\chi^2_{\text{data}}, \chi^2_{\text{pseudodata}})$ pairs for the entire ensemble of spectra considered in Fig. 1 is shown as the purple shading in Fig. 3. The PPP then consists of the fraction of points in this ensemble that lie above the $\chi^2_{\text{data}} = \chi^2_{\text{pseudodata}}$ line. In the limit of infinite MCMC samples, a model that perfectly describes the data apart from statistical variations will produce a PPP of 0.5. We observe that the shaded distribution in Fig. 3 is distributed evenly around 1 unit of χ^2 per bin in both axes, and the PPP we obtain is 0.56. Both of these imply that the model is a good representation of the data.

We also find good p values for the ν_μ and $\bar{\nu}_\mu$ samples considered independently, whose 1σ credible regions (those enclosing 68.3% of the posterior, as with the ranges for the spectra in Fig. 1) are indicated by the dashed light blue and light red colored contours in Fig. 3, respectively. By contrast, the effect of fluctuations in the smaller-statistics samples can be seen more readily in the analogous ν_e (solid dark blue) and $\bar{\nu}_e$ (solid red) contours. Both contours are much larger than their ν_μ counterparts, particularly along the y axis, which corresponds to the dimension where statistical uncertainties are considered. In the ν_e contour, the offset downwards from unity along the x axis, and the corresponding shift above the diagonal along the y axis, together suggest a set of fluctuations that are relatively close to the Asimov (unfluctuated) model prediction. Each member of the ensemble of pseudodata spectra thus typically has larger χ^2 relative to the Asimov than that of the data. This relative closeness of the ν_e prediction to the data can also be seen in the top left panel of Fig. 1. The unusual shape of this contour arises because the PPP distribution for ν_e consists of two modes superimposed upon one another, corresponding to the high-probability regions within the NO that will be discussed below in Fig. 7. On the other hand, the situation is reversed for the $\bar{\nu}_e$ spectra, where the data fluctuations seen in Fig. 1 result in larger deviations from the Asimov prediction than the bulk of the pseudodata spectra, and the contour consequently shifts downward from the diagonal. The fact that combining all four subsamples together produces a PPP that is closer to 0.5 than any of them are individually is good evidence that the deviations from PPP of 0.5 in the various subsamples are predominantly driven by statistical, rather than systematic, effects.[5] Moreover, studies where Poisson fluctuations were applied to the model to produce fake data spectra, which were then subjected to the PPP computation process, indicated that both large and small PPP values such as those we observe here do naturally occur in the subsamples. We conclude therefore that our set of MCMC samples reflect PMNS parameters with a good description of the physics exhibited in our data.

2. *PMNS parameter measurements*

We produced Bayesian credible regions for the PMNS neutrino oscillation parameters using the data spectra and the MCMC samplers described in Secs. III A and II C, respectively. These credible regions together with the posterior probability distributions are shown in Figs. 4 and 7. In both cases the constraint on θ_{13} from the reactors used is the 2019 PDG's combination of extant measurements [83] and is applied in the form of a Gaussian prior on $\sin^2 2\theta_{13}$, as explained in Sec. III A 3. In the figures in this section, credible intervals that show the normal and inverted mass orderings separately are created by first making one shared credible interval that spans both. The separate panels then display the relevant regions from this shared interval that apply to the specified ordering. By constructing them in this manner, we ensure the NO and IO

---

[5] The reader may find further information on interpretations of posterior predictive p values in Refs. [86,87].

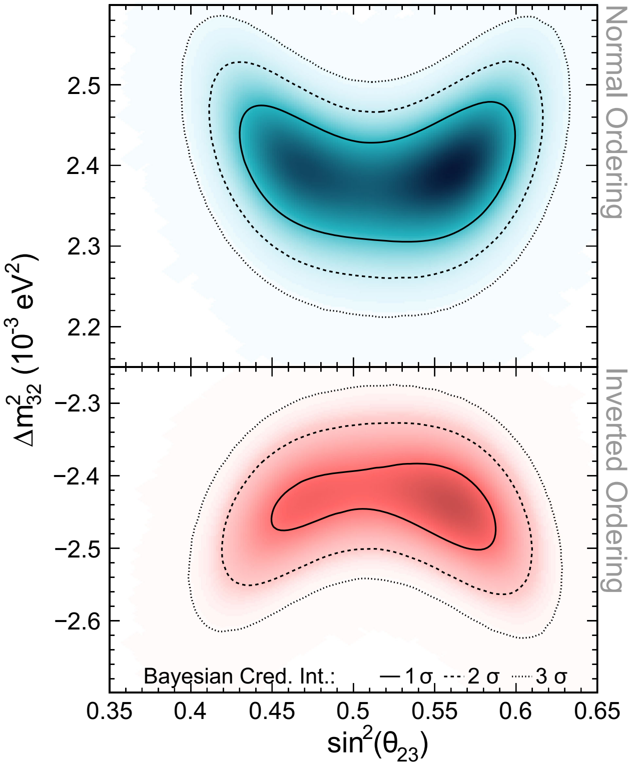


FIG. 4. Binned posterior probability densities (shaded) for $\sin^2\theta_{23}$-Δm^2_{32}, marginalized over both mass orderings and plotted separately for the normal (top) and the inverted (bottom) mass orderings (marginalization over the mass orderings is explained at the beginning of Sec. III B 2). Contours indicate regions enclosing 1σ, 2σ, and 3σ of the posterior probability.

intervals share a highest-posterior density point and posterior probability distribution, preserving any NOvA preference towards one of the mass orderings in the credible region limits. Similarly, the distributions and intervals labeled "both orderings" are created from MCMC samples over all values for Δm^2_{32}, summing together the normal and inverted ordering posteriors before extracting the credible intervals. Versions of these distributions with an alternate marginalization scheme that considers each ordering independently may be found in the Supplemental Material [88].

Table I shows the highest posterior probability density (HPD) points together with the 1σ credible intervals. These points are given for all the PMNS oscillation parameters of interest, split into both, normal, and inverted mass orderings (using the same methodology regarding the mass ordering as for the figures). For some of the parameters the 1σ region spans disjoint areas; we denote this with a union symbol $\cup$. These high posterior probability regions are in generally good agreement with the frequentist analysis of the same dataset [35].

TABLE I. HPD points together with the 1σ Bayesian credible interval limits for the PMNS parameters of interest, marginalized over all the mass ordering (MO) hypotheses. Marginalization over the mass orderings is explained at the beginning of Sec. III B 2. In these results a Gaussian prior corresponding to the reactor constraint on $\sin^2 2\theta_{13}$ (see Sec. III A 3) is applied.

	MO	HPD	1σ
δ_{CP} (Prior uniform in δ_{CP})	Both	0.91π	$[0.02\pi, 0.31\pi] \cup [0.68\pi, 1.67\pi]$
	Normal	0.89π	$[0.54\pi, 1.07\pi] \cup [1.99\pi, 0.48\pi]$
	Inverted	1.44π	$[1.26\pi, 1.65\pi]$
$\sin^2\theta_{23}$	Both	0.56	$[0.45, 0.49] \cup [0.52, 0.59]$
	Normal	0.56	$[0.44, 0.59]$
	Inverted	0.56	$[0.55, 0.57]$
Δm^2_{32} ($\times 10^{-3}$ eV^2)	Normal	2.39	$[2.32, 2.46]$
	Inverted	-2.44	$[-2.47, -2.41]$

Figure 4 shows the $\sin^2\theta_{23}$-Δm^2_{32} plane, where the denser MCMC samples (darker color) and larger credible regions in the upper panel relative to the lower indicate a mild preference for the normal ordering. This conclusion holds in the presence of the entire systematic uncertainty model discussed in Sec. II C. However, the preferred regions in each mass ordering depend in a nontrivial way on the systematic uncertainties, as illustrated in Fig. 5. The most significant effect is in Δm^2_{32}, where the systematic effects not only broaden the preferred region but also shift the most probable value to larger absolute magnitudes. The most important of the uncertainties contributing to this

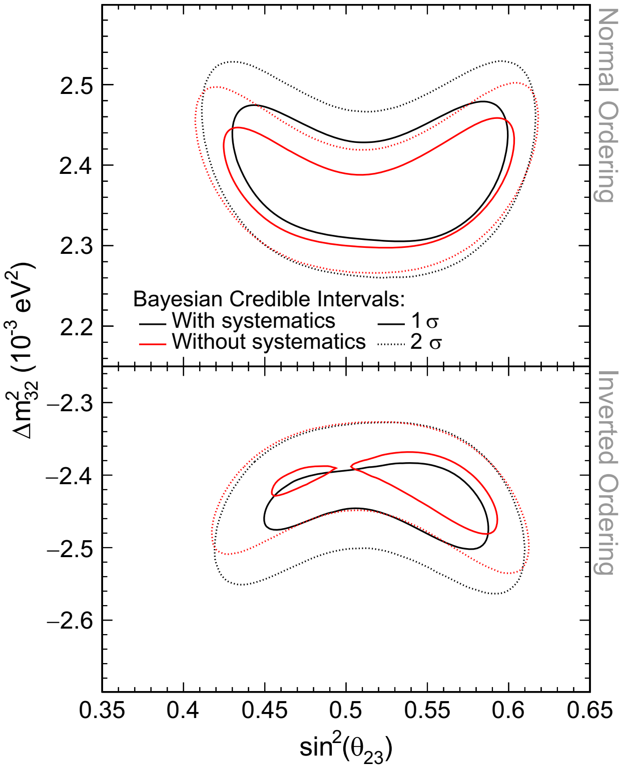


FIG. 5. Credible interval comparisons for $\sin^2\theta_{23}$-Δm^2_{32} when sampling with only statistical uncertainties (red) and with all the NOvA systematic parameters (black). The external constraint on θ_{13} from the reactor experiments was applied in these fits.

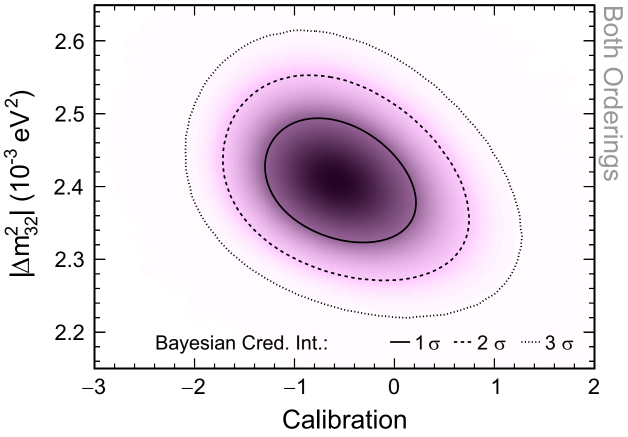


FIG. 6. Binned posterior probability density (shaded) with 1, 2, and 3σ credible intervals in $|\Delta m^2_{32}|$ and the NOvA absolute calibration systematic uncertainty, where the horizontal axis is measured in units of standard deviations from the nominal value. The external constraint on θ_{13} from reactor experiments was applied in this fit.

movement is in the absolute calibration of the calorimetric energy scale. Because this directly affects reconstructed neutrino energies, it shifts the expected number of events in the trough of the ν_μ and $\bar{\nu}_\mu$ disappearance spectra and is thus anticorrelated with Δm^2_{32} (correlation coefficient −0.29), as can be seen in Fig. 6. Shifting Δm^2_{32} in this way also moves $\sin^2\theta_{23}$ closer to maximal disappearance ($\sin^2\theta_{23}^{\mathrm{MD}} \approx 0.51$; the value depends slightly on $\sin^2\theta_{13}$). Because the ν_μ disappearance spectra are essentially identical for values reflected across the maximal disappearance line, the credible regions are nearly symmetric around it; the degeneracy is broken only by the ν_e appearance spectra, which have less statistical power. Thus, the credible regions for smaller $|\sin^2\theta_{23} - \sin^2\theta_{23}^{\mathrm{MD}}|$ appear narrower, even though the sensitivity is unchanged. This effect is responsible for the apparently slightly tighter constraint on $\sin^2\theta_{23}$ observed in Fig. 5 under the effect of systematic uncertainies.

The situation is different in the δ_{CP}-$\sin^2\theta_{23}$ plane, which is shown in Fig. 7. Here, we observe preferences for *CP*-nonconserving values of δ_{CP} (i.e., nonintegral values of δ_{CP}/π) in both normal and inverted orderings, and for the upper octant ($\sin^2\theta_{23} > 0.5$) of θ_{23} (reflected also in Fig. 4), though the *CP*-conserving points are only weakly disfavored in NO. However, in contrast to the conclusions for $\sin^2\theta_{23}$-Δm^2_{32}, these inferences are minimally affected by the presence of systematic uncertainties, as Fig. 8 makes clear—even if the resolution does degrade slightly and the credible regions grow. Here the minimal impact of systematic uncertainties owes primarily to the smaller statistics of the ν_e and especially $\bar{\nu}_e$ appearance samples that drive the sensitivity to these parameters, which results in statistical uncertainties dominating the uncertainty budget.

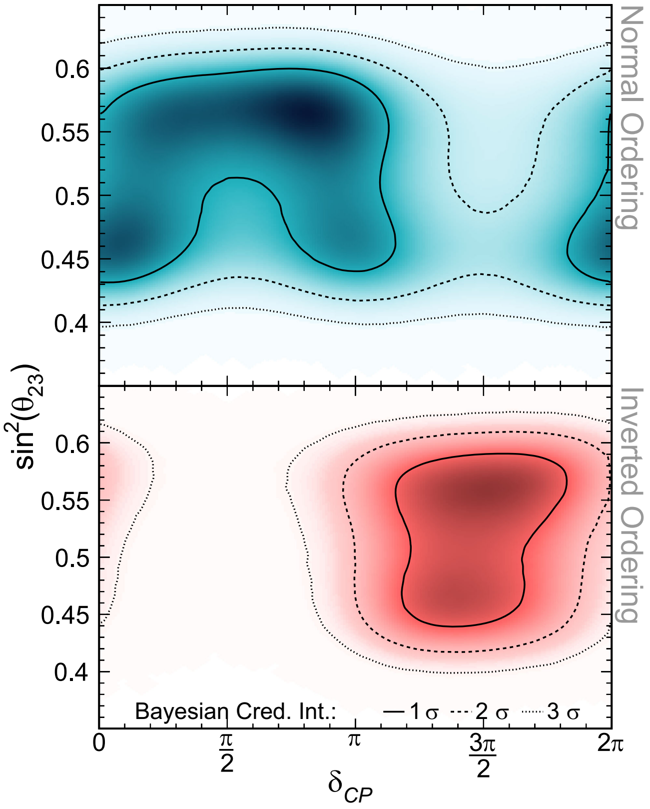


FIG. 7. Binned posterior probability density (shaded) with 1, 2, and 3σ credible intervals for δ_{CP}-$\sin^2\theta_{23}$, marginalized over both mass orderings for the normal mass ordering (top, blue) and the inverted mass ordering (bottom, red). The external constraint on θ_{13} from reactor experiments was applied.

To more quantitatively assess the mass ordering and octant preferences, we give the posterior probabilities inferred for each combination of hypotheses in Table II. We also express them in a less prior-dependent way using Bayes factors. In both cases the evidence for either option is weak.[6] The analogous Gaussian *p* values also correspond to significances of less than 1σ. The interpretations of the octant and the mass ordering hypothesis preferences are in good agreement with the 2020 frequentist analysis of the same dataset, which used profiling with Feldman-Cousins corrections instead of marginalization [20,35]. This general agreement is also true for the parameters' intervals, with small differences expected given the differing statistical methods used. To examine the *CP*-conservation situation more comprehensively, taking the whole PMNS matrix into consideration, we will use the Jarlskog invariant measure, explored in the next section.

---

[6] The reader unfamiliar with the interpretation of Bayes factors is referred to the standard treatments of Jeffreys [89] or Kass and Raftery [90].

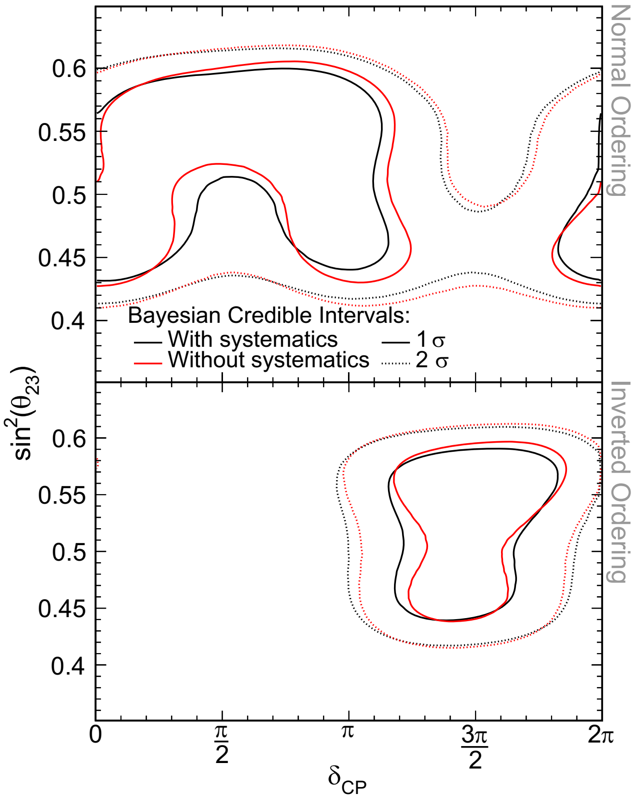


FIG. 8. Credible interval comparisons for δ_{CP}-$\sin^2\theta_{23}$ between statistical-only fits and fits including all the NOvA systematic parameters, with external constraint on θ_{13} from reactor experiments.

3. CP violation—Jarlskog invariant

The Jarlskog invariant is a measure of the strength of charge-parity violation that is independent of how the mixing matrix is parametrized. The neutrino-mixing form of the Jarlskog [91] parallels its development in the quark sector [92]. Under the three-neutrino-flavor assumption, its definition is

$$J \equiv \cos(\theta_{12})\cos^2(\theta_{13})\cos(\theta_{23})\sin(\theta_{12}) \times \sin(\theta_{13})\sin(\theta_{23})\sin(\delta_{CP}), \tag{3}$$

where if any of the factors is zero, the invariant vanishes and the neutrino mixing matrix is *CP* conserving. Nonzero values of J, on the other hand, indicate *CP* violation. We produce this measurement by taking the MCMC chain with all the oscillation parameters' values and calculating the Jarlskog invariant at each MCMC step. As NOvA is not sensitive to the value of θ_{12}, for J we use a Gaussian prior analogous to the one used for the reactor constraint (Sec. III A 3) but whose range consists of the PDG's 2019 average of solar and LBL reactor neutrino measurements: $\sin^2\theta_{12} = 0.307 \pm 0.013$ [83], the same value used in our previous results.

As described in Sec. III A 3, we use uninformed priors for the oscillation parameters where possible, meaning that the priors are uniform in the variable that is being shown. The trigonometric dependence of J on the oscillation parameters shown in Eq. (3) makes clear that constructing a prior uniform in J would result in nonuniform distributions over its constituent variables. We studied numerous priors in the oscillation parameters not constrained by external data and found that for any reasonable choice of prior for θ_{23} or Δm^2_{32}—that is, one that does not vanish or diverge across the range of values allowed by other contemporary experiments—the posterior in J is essentially unchanged, due to the strong constraints afforded by the NOvA data. Thus the only degree of freedom where the prior is of major concern is δ_{CP}. Since the Jarlskog invariant is written in terms of $\sin\delta_{CP}$, the natural formulation is in terms of a prior uniform in $\sin\delta_{CP}$. At the same time, there are theoretical considerations that suggest a prior uniform in δ_{CP} may be more appropriate [93]. We therefore consider both a prior uniform in $\sin\delta_{CP}$ and one uniform in δ_{CP}.

Figure 9 shows the inferred values of the Jarlskog invariant extracted from a fit to the NOvA data with the external constraint on θ_{13} from the reactor experiments, marginalized over the normal (top) and inverted (bottom) mass orderings. The regions most favored by the data tend towards *CP* violation in the NO, although the 1σ intervals do include *CP* conservation when the prior uniform in $\sin\delta_{CP}$ is used. The IO's preferred regions are more markedly distant from *CP* conservation, particularly for the prior uniform in δ_{CP}. As expected, these trends mirror those observed when considering *CP*-conserving and *CP*-violating values of δ_{CP} in Sec. III B 2.

A more quantitative way of measuring the NOvA preference for *CP* conservation with the Jarlskog invariant is through Bayes factors. These can be calculated with the Savage-Dickey density ratio method, which computes Bayes factors for point hypotheses [94,95]. Using this method we can calculate the Bayes factor for the *CP*-conserving value $J = 0$, nested under the unconstrained hypothesis where J can take any value. Bayes factors always depend formally on the choice of prior, but in many circumstances (such as those discussed above), if two hypotheses being compared use the same prior, it will

TABLE II. Bayes factors (posterior probabilities) in all the θ_{23} octant and mass ordering hypotheses combinations. A weak preference towards the normal mass ordering and upper octant is observed. Probabilities summed across rows or columns may differ slightly from the totals due to rounding. In these results the reactor constraint on $\sin^2 2\theta_{13}$ is applied.

	Normal Ordering	Inverted Ordering	Total
Upper Octant	0.71 (0.42)	0.26 (0.21)	1.67 (0.63)
Lower Octant	0.35 (0.26)	0.13 (0.12)	0.60 (0.38)
Total	2.08 (0.68)	0.48 (0.33)	(1.0)

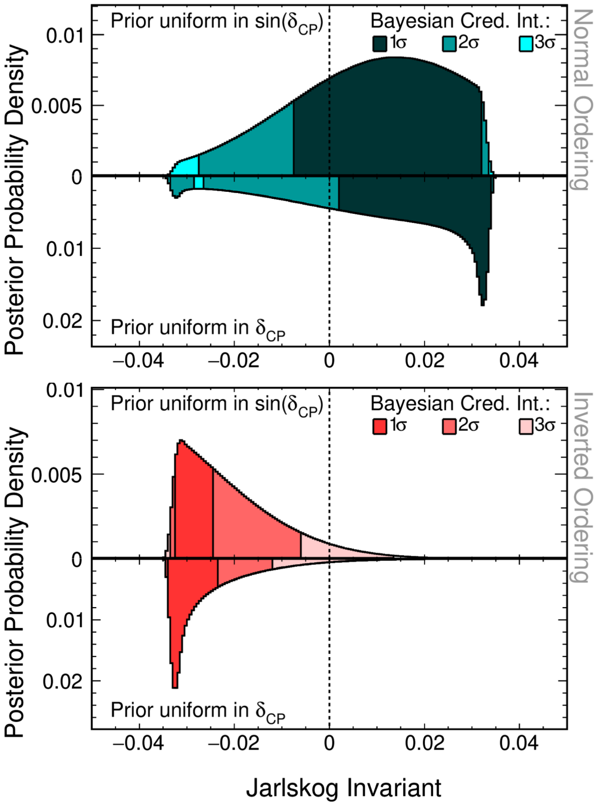


FIG. 9. Posterior probability density for the Jarlskog invariant, marginalized over both mass orderings and plotted separately for the NO (top plot) and IO (bottom plot). The top half of each panel shows the posterior with a prior uniform in $\sin\delta_{CP}$, while the bottom half uses a prior uniform in δ_{CP}. A CP-conserving line is drawn at $J = 0$. The external constraint in θ_{13} from the reactor experiments is used.

cancel. This is not the case in the Savage-Dickey method. Therefore, we computed the Bayes factor for $J = 0$ under several hundred combinations of priors for θ_{23}, Δm^2_{32}, and δ_{CP}. The combinations used in the previous discussion (uniform in θ_{23}, $|\Delta m^2_{32}|$, and $\mathrm{sgn}(\Delta m^2_{32})$, plus the two variants in δ_{CP}) produced Bayes factors that were among the most conservative that we found (though it is possible to engineer priors with, for example, smaller ranges where they are uniform to produce even more conservative Bayes factors). We therefore took these as representative Bayes factors and inverted them to obtain the associated Bayes factors for CP violation over CP conservation, $J \neq 0$.

Table III shows the Bayes factors for CP violation against CP conservation for normal, inverted, and both mass orderings, calculated for priors uniform in $\sin\delta_{CP}$ or δ_{CP}. All of these probabilities point towards a preference for CP nonconservation, although the preferences are minimal regardless of the δ_{CP} prior or the mass ordering (and an analogous frequentist p value would indicate significances less than 1σ, apart from the inverted ordering, where they range 1.1–1.2σ[7]). However, we reiterate that the inherent prior dependence of the Savage-Dickey method means these values can only be treated as representative of a class of possible interpretations of the evidence. Future neutrino oscillation experiments, in which the evidence for the ordering is expected to be much stronger, will likely wish to study this problem further.

TABLE III. Bayes factors for preference of CP violation over CP conservation, extracted using the Savage-Dickey method at $J = 0$. Priors uniform in δ_{CP} and $\sin\delta_{CP}$ are both shown. The preferences are given for the normal (NO), inverted (IO), and both (BO) mass orderings.

Prior	NO	IO	BO
Uniform $\sin\delta_{CP}$	1.2	3.4	1.5
Uniform δ_{CP}	1.0	3.8	1.6

4. Using only NOvA constraints on θ_{13}

NOvA's oscillation measurements simultaneously constrain a combination of mixing angles (θ_{23}, θ_{13}), the mass-squared splitting Δm^2_{32}, the CP phase δ_{CP}, and the neutrino mass ordering. Therefore, applying a strong external 1D constraint on θ_{13} from reactor antineutrino experiments—and thereby reducing the available solution space—increases the sensitivity to other parameters. This is especially true of the octant of θ_{23}, as we will show below. However, using an uninformed (uniform) prior in $\sin^2 2\theta_{13}$ enables us to directly compare a NOvA-only measurement against that of the reactor experiments. In so doing we can examine the robustness of the PMNS description across short-baseline reactor antineutrino experiments and long-baseline accelerator-based oscillation experiments. We also may study whether our preferences change in the presence of an external constraint on the data. In the following, we use recomputed posteriors that assumed a uniform prior over $\sin^2 2\theta_{13}$. The PPP goodness-of-fit metric from Sec. III B 1, when recomputed for the associated posterior spectra, was found not to change more than 0.01 for any of the samples shown in Fig. 3, suggesting that the data is well modeled without the constraint.

Without an external constraint on $\sin^2 2\theta_{13}$, the relevant degrees of freedom present a complex space with many intercorrelations among the parameters and numerous possible solutions. We can use the type of

---

[7] We emphasize that these significances cannot be read off of Fig. 9. The probability density shown there is marginalized independently at each value of J. However, the binary hypothesis test $J = 0$ vs $J \neq 0$ requires a simultaneous marginalization across the whole $J \neq 0$ space. For the latter, a point-hypothesis treatment, such as the Savage-Dickey formalism we use here, is necessary.

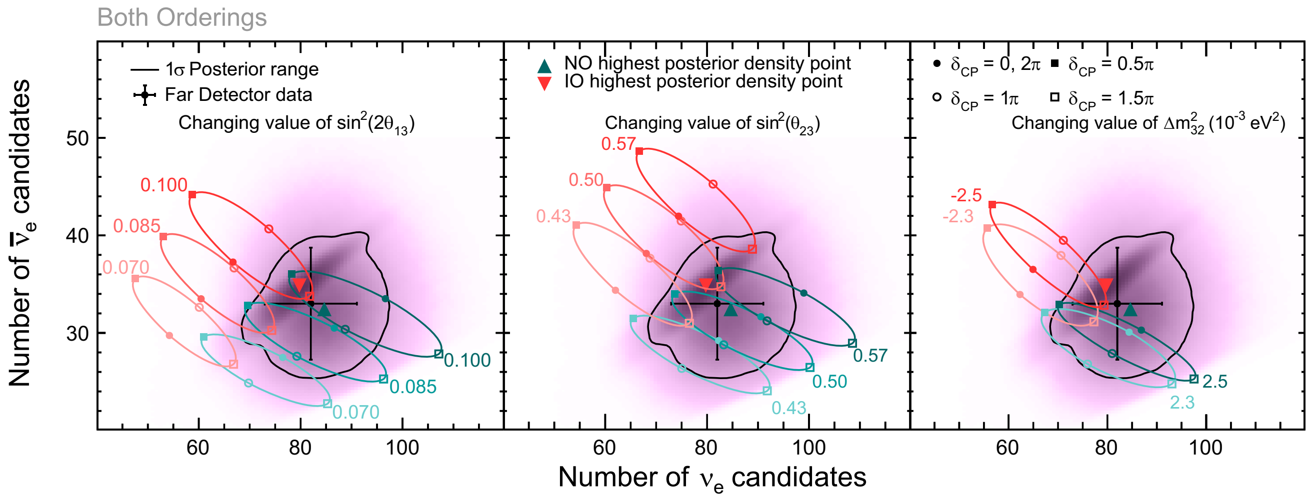
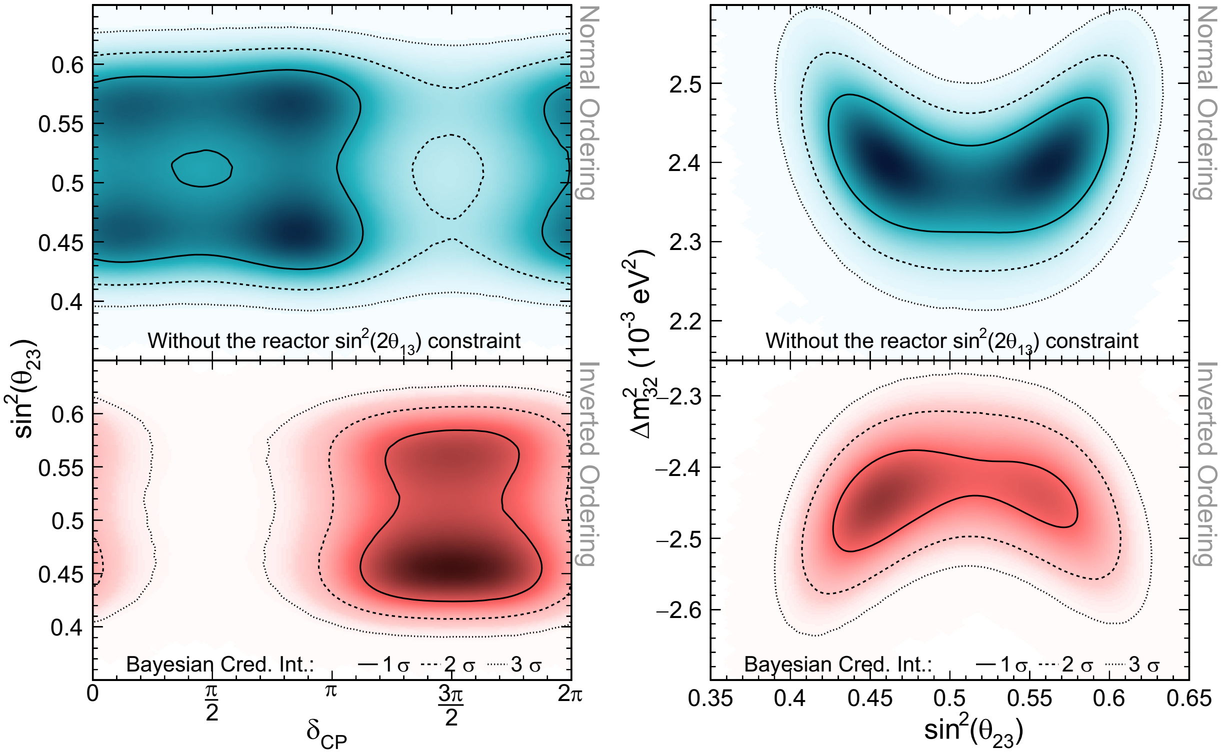


FIG. 10. Posterior probability density (purple shading)—that is, the probability that a given (ν_e, $\bar{\nu}_e$) pair of counts is the true underlying value in nature (see text)—compared to the measurement and associated expected statistical variance (black cross) of the total number of $\bar{\nu}_e$ candidates vs ν_e candidates. The black solid contours indicate the regions enclosing 68.3% of the posterior. Each panel overlays a set of ellipses corresponding to the predicted event counts over the range of possible values of δ_{CP}, with the parameter notated nearby held at that value (blue for NO, red for IO, with varying shading depending on the parameter value) and other parameter values as given in Table IV. The triangular markers show the highest posterior density in the ν_e-$\bar{\nu}_e$ space for when restricting to NO (blue) and IO (red) hypotheses. The posteriors shown here employ a uniform prior over $\sin^2 2\theta_{13}$.

posterior distributions shown in Fig. 1 to explore these possibilities: for example, by computing the total number of predicted ν_e and $\bar{\nu}_e$ candidates for the parameters of each MCMC sample and comparing the distribution of these predictions in (ν_e, $\bar{\nu}_e$) candidate space to what we observe in the data. This is shown in Fig. 10. To guide our intuition, we overlay ellipses corresponding to the ranges of predicted number of events for the possible values of δ_{CP} at fixed values of the other parameters, subject to constraint from our data in that the HPD value in each ordering is used for parameters not explicitly varied. Though not all dimensions of the way our observed ν_e and ν_μ candidate energy spectra interact with the oscillation parameters can be represented in this presentation, we can still note several important features. First, the highest-density (highest-probability) region lies essentially equally along the overlap region between NO and IO ellipses in all panels, meaning we will find good solutions for either, given an appropriate value of δ_{CP}. However, the NO ellipses subtend more of the posterior region, which will result in a slight overall preference for NO. Second, it is clear from the left panel that the value of $\sin^2 2\theta_{13}$ preferred by the reactor average, $\sin^2 2\theta_{13} = 0.085$, is compatible with the highest-probability region for NOvA. Third, comparing the left and center panels, there is obvious degeneracy between $\sin^2 2\theta_{13}$ and $\sin^2\theta_{23}$, since independently varying them produces similar effects on the predictions. And finally, the right panel demonstrates that the value of Δm^2_{32} plays a minor (though not negligible) role in the appearance channel; its primary constraints arise from the ν_μ disappearance measurement. These features will be examined more quantitatively in the distributions that follow. Table IV shows the highest posterior probability points together with the 1σ credible intervals for all the PMNS oscillation parameters of interest, extracted from the fit with a prior uniform in $\sin^2 2\theta_{13}$. Similarly to Table I, the values are split into both, normal, and inverted mass orderings with any disjoint 1σ regions denoted with a union symbol ∪. As compared to the results of the fit

TABLE IV. The HPD points together with the 1σ Bayesian credible interval limits for the PMNS parameters of interest marginalized over all the MO hypotheses. Marginalization over the mass orderings is explained at the beginning of Sec. III B 2. Extracted from a fit with prior uniform in $\sin^2 2\theta_{13}$.

	MO	HPD	1σ
δ_{CP}	Both	1.53π	$[0.67\pi, 1.88\pi] \cup [0.06\pi, 0.12\pi]$
(Prior uniform	Normal	0.85π	$[0.48\pi, 1.04\pi] \cup [1.99\pi, 0.41\pi]$
in δ_{CP})	Inverted	1.52π	$[1.25\pi, 1.75\pi]$
$\sin^2 2\theta_{13}$	Both	0.087	[0.071, 0.107]
	Normal	0.084	[0.065, 0.108]
	Inverted	0.094	[0.083, 0.106]
$\sin^2\theta_{23}$	Both	0.46	$[0.43, 0.50] \cup [0.53, 0.58]$
	Normal	0.46	[0.43, 0.59]
	Inverted	0.46	[0.44, 0.48]
Δm^2_{32}	Normal	2.39	[2.33, 2.46]
($\times 10^{-3}$ eV^2)	Inverted	−2.44	[−2.48, −2.40]



(a) δ_{CP}–$\sin^2\theta_{23}$ (b) $\sin^2\theta_{23}$–Δm^2_{32}

FIG. 11. Posterior probability density (shaded) with 1, 2, and 3σ credible intervals for δ_{CP}-$\sin^2\theta_{23}$ (left) and $\sin^2\theta_{23}$-Δm^2_{32} (right), marginalized over both mass orderings, plotted separately for NO (top, blue) and IO (bottom, red). Contours extracted from a fit with prior uniform in $\sin^2 2\theta_{13}$.

with the external constraint from reactor experiments, the central values for δ_{CP} and $\sin^2\theta_{23}$ shown here are the most different. This is not unexpected since there is also a degeneracy between these two parameters, and both δ_{CP} and $\sin^2\theta_{23}$ have multiple areas of high probability.

Figure 11 shows our preferred regions in δ_{CP}-$\sin^2\theta_{23}$ and $\sin^2\theta_{23}$-Δm^2_{32} spaces without the external θ_{13} constraint. Comparing with Figs. 4 and 7, it is evident that removing the constraint diminishes the NOvA sensitivity especially to the octant of θ_{23} (there is now near symmetry across $\sin^2\theta_{23} = 0.5$). Although the octant preference is weakened, the central values of the PMNS parameters $\sin^2\theta_{23}$, Δm^2_{32}, and especially δ_{CP} do not otherwise see any significant change. This is as expected from Fig. 10. The results without the reactor constraint are therefore fully compatible with the standard NOvA results with the external constraint on $\sin^2 2\theta_{13}$ applied. We also observe that although the sensitivity is reduced without the constraint, and although that for a given value of δ_{CP} there is always a combination of parameters within either mass ordering compatible with the data, certain δ_{CP}-$\sin^2\theta_{23}$-ordering combinations are still excluded with reasonable confidence, such as $(\delta_{CP} = \frac{\pi}{2}, \mathrm{IO})$ and $(\delta_{CP} = \frac{3\pi}{2}, \sin^2\theta_{23} = 0.5, \mathrm{NO})$. Combinations such as these produce strong asymmetries in the $(\nu_e, \bar{\nu}_e)$ counts, and as Fig. 10 makes clear, such parameter values lie in regions outside the posterior point cloud, and are thus disfavored. Table V examines the posterior probabilities for each of the octant and mass-ordering hypotheses more quantitatively. Compared to Table II, we note a similar weakened preference for the upper octant of θ_{23} for the fit without the reactor constraint.

TABLE V. Bayes factors (posterior probabilities) for all the θ_{23} octant and mass ordering hypotheses, with a marginal preference towards the normal mass ordering and upper octant. Numbers extracted from a fit with a uniform prior in $\sin^2 2\theta_{13}$. Probabilities summed across rows or columns may differ slightly from the totals due to rounding.

	Normal Ordering	Inverted Ordering	Total
Upper Octant	0.53 (0.35)	0.20 (0.17)	1.05 (0.51)
Lower Octant	0.39 (0.28)	0.42 (0.20)	0.95 (0.49)
Total	1.70 (0.63)	0.59 (0.37)	(1.0)

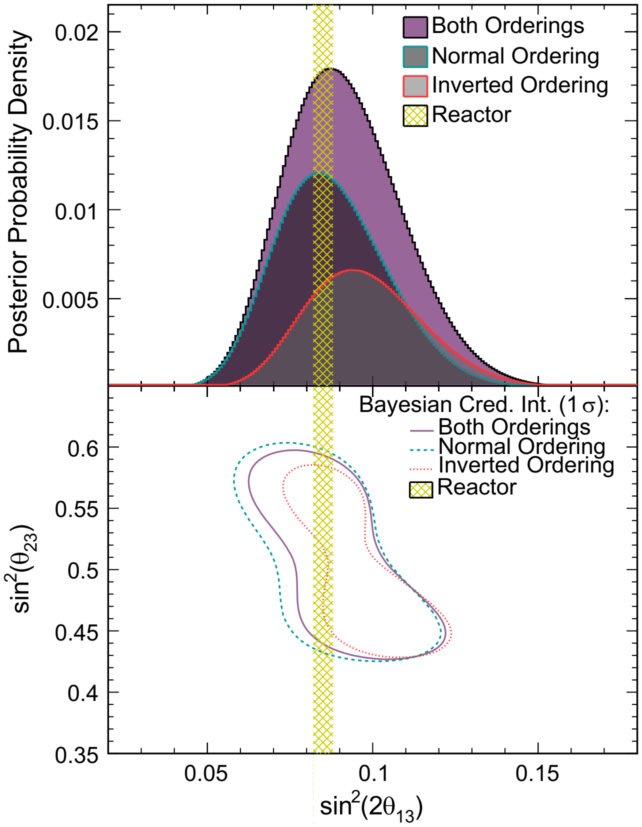


FIG. 12. Comparison of the 2D $\sin^2 2\theta_{13}$-$\sin^2\theta_{23}$ credible intervals (bottom) and 1D $\sin^2 2\theta_{13}$ posterior probability densities (top) between both (purple), normal (light-blue), and inverted (red) mass orderings with a prior uniform in $\sin^2 2\theta_{13}$ and $\sin^2\theta_{23}$. The reactor experiments' 1σ interval in $\sin^2 2\theta_{13}$ from the PDG 2019 [83] is shown in yellow hatched bar.

In Fig. 12 we study the impact of applying the reactor constraint on our mass-ordering inference. We first observe that the NOvA and reactor measurements for $\sin^2 2\theta_{13}$ are in good agreement, as evidenced by the overlap of the yellow bar (indicating the reactor 1σ range) with the peak of the "both orderings" posterior in the upper plot. While we know of no widely accepted measure for quantifying such consistency, we combined this "both orderings" posterior with a Gaussian probability distribution centered at $\sin^2 2\theta_{13} = 0.085$ and of width 0.003 (representing the reactor constraint) using the technique of conflation [96,97]. The result is identical to the reactor distribution we began with, which (as the examples in Ref. [97] show) does not occur if the probability densities being conflated agree poorly. This accords with the fact that the PPP values did not meaningfully improve when removing the reactor constraint, as noted previously. When we examine the orderings separately, we find that the normal ordering contains more posterior probability, as expected from Fig. 10 and as reflected in the larger NO contour in the lower panel here. Moreover, as we see in the marginal posterior distribution shown in the top panel, the posterior restricted to IO prefers generally larger values of $\sin^2 2\theta_{13}$. The associated correlation with $\sin^2\theta_{23}$ also pushes θ_{23} into the lower octant, as seen in the lower panel; this point will be developed further momentarily. Because the value of $\sin^2 2\theta_{13}$ measured by the reactor experiments is more consistent with the NOvA NO posterior, this results in the slightly stronger preference for the NO in Table II as compared to Table V. However, the difference is small, which indicates that the (mild) NO preference observed in the data is largely independent of the reactor constraint.

The degeneracy in the measurement between $\sin^2\theta_{23}$ and $\sin^2 2\theta_{13}$ we noted in Fig. 10 can be studied directly by examining their joint posterior probability distribution, marginalized over all other parameters and the mass ordering. We show this in Fig. 13. The central panel exhibits a clear anticorrelation between the octant of θ_{23} and the value of $\sin^2 2\theta_{13}$, which is expected since both parameters enter at leading order in the $\nu_\mu \to \nu_e$ and $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillation probabilities. Here the overlap of the reactor measurements (again indicated by the yellow hatched bar) with our marginal posterior for $\sin^2 2\theta_{13}$ (right panel) favors the upper octant over the lower octant when we constrain the results to specifically the upper or lower octant of θ_{23}. Thus, we see that the preference for the upper octant of θ_{23} in Table II is an emergent behavior that arises from the application of the reactor constraint. Similar changes in the strength of the octant preference when the reactor constraint is applied have been noted in results from T2K [26] and simulation studies for DUNE [98], though to our knowledge this is the first time the underlying θ_{13}-θ_{23} anticorrelation has been examined in detail with a data result. We also note that though the marginal posterior for $\sin^2\theta_{23}$, in the top panel, shows a higher posterior density in the lower octant, the total posterior probability integrated across the upper octant is slightly larger than the corresponding lower octant probability; but as Table V makes clear, this preference is entirely insignificant.

We emphasize that reactor neutrino experiments and accelerator neutrino experiments measure the PMNS $\sin^2 2\theta_{13}$ by examining different sectors of neutrino oscillations, over a wide range of baselines. Reactor neutrino experiments measure the $\bar{\nu}_e$ survival probability $\mathcal{P}(\bar{\nu}_e \to \bar{\nu}_e)$ with low-energy (few-MeV) $\bar{\nu}_e$s over a short (few-km) baseline. Conversely, accelerator neutrino experiments simultaneously measure ν_e appearance in a few-GeV ν_μ beam, and $\bar{\nu}_e$ appearance in a $\bar{\nu}_\mu$ beam, both over a long (hundreds-of-km) baseline. In long-baseline measurements $\mathcal{P}(\nu_\mu \to \nu_\mu)$, $\mathcal{P}(\bar{\nu}_\mu \to \bar{\nu}_\mu)$, $\mathcal{P}(\nu_\mu \to \nu_e)$, and $\mathcal{P}(\bar{\nu}_\mu \to \bar{\nu}_e)$ are all exploited to constrain the PMNS oscillation parameters, including $\sin^2 2\theta_{13}$. Thus, the consistency observed between long- and short-baseline

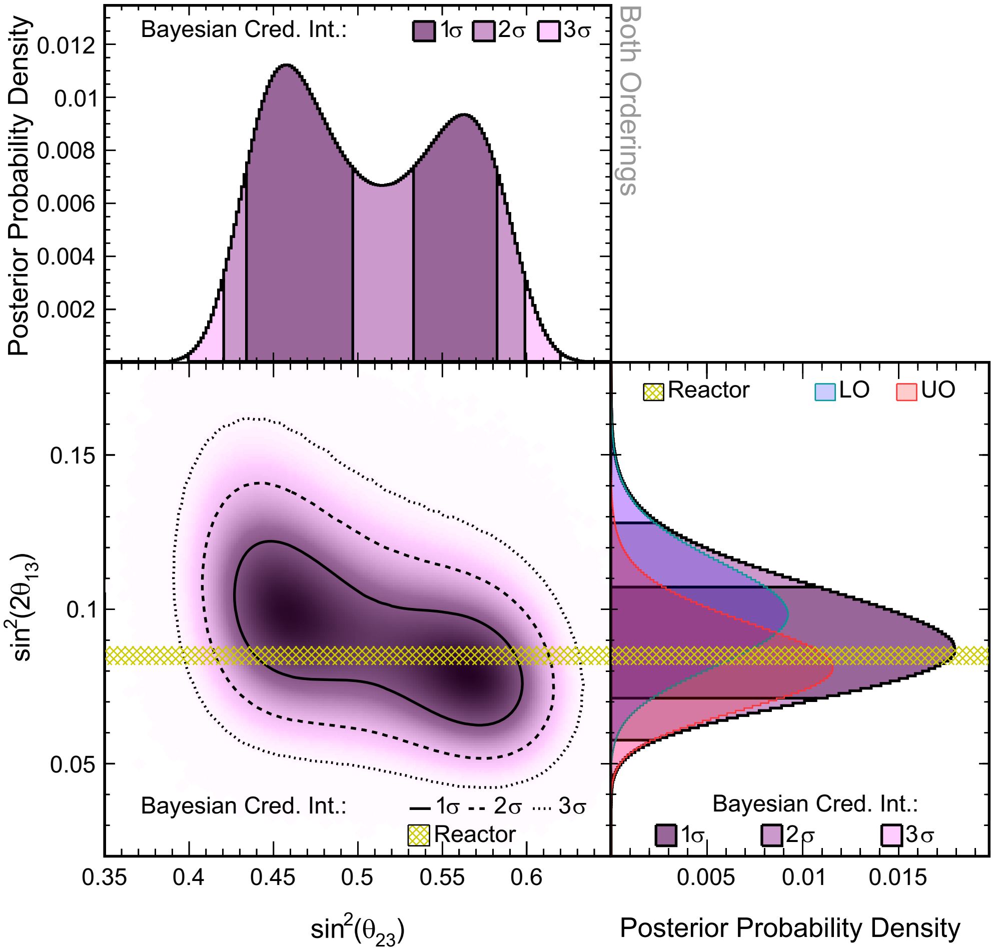


FIG. 13. $\sin^2\theta_{23} - \sin^2 2\theta_{13}$ posterior probability densities with 1, 2, and 3σ credible intervals, marginalized over both mass orderings. The posterior density was extracted from a fit with a prior uniform in $\sin^2 2\theta_{13}$. The reactor experiments' 1σ interval in $\sin^2 2\theta_{13}$ from the PDG 2019 [83] is shown in the yellow hatched bar. The right panel shows the posterior probability for $\sin^2 2\theta_{13}$, with its contributions from the upper octant (UO, transparent, red outline) and lower octant (LO, transparent, blue outline) of θ_{23}.

measurements lends support to the PMNS interpretation of neutrino oscillations.

IV. CONCLUSIONS

We have refit the NOvA dataset of 13.6×10^{20} POT in neutrino beam mode and 12.5×10^{20} POT in antineutrino mode using a Bayesian statistical approach. NOvA data continues to be consistent with maximal mixing for $\sin^2\theta_{23}$ and the regions of the $\sin^2\theta_{23}$-Δm^2_{32} and δ_{CP}-$\sin^2\theta_{23}$ spaces preferred by this analysis are consistent with those of the previous analysis done using a frequentist fit.

With the introduction of the Bayesian analysis, we also expand the neutrino oscillation parameters measured by NOvA. We report for the first time NOvA measurements that do not require constraints on $\sin^2 2\theta_{13}$ from reactor antineutrino oscillations. Moreover, we are also able to include new results for $\sin^2 2\theta_{13}$ and the Jarlskog invariant J; these were impractical to produce under the preceding frequentist method due to the necessity of Feldman-Cousins corrections. The inferences on J provide a parametrization-independent measurement of CP violation and indicate that NOvA data has a weak preference for CP violation, which becomes slightly more pronounced when assuming the inverted mass ordering.

As NOvA measures a convolution of mixing angles, δ_{CP}, and mass ordering in the oscillation probabilities, removing the external constraint on θ_{13} reduces our constraining power to determine the octant of θ_{23}, mass ordering, and δ_{CP}. While our sensitivity is reduced without the external constraint (particularly for the octant), we note that the conclusions arising from the analysis remain unchanged.

This measurement of $\sin^2 2\theta_{13}$ using electron neutrinos and antineutrinos with energies in the GeV range and propagating for hundreds of kilometers is fully consistent with measurements performed using few-MeV electron antineutrinos from nuclear reactors propagating for a few kilometers. The consistency of results using the PMNS framework across a broad regime of conditions bolsters its applicability. More stringent tests of CP violation and the consistency of our $\sin^2 2\theta_{13}$ measurement with those of reactors will be possible with increased statistics in upcoming NOvA measurements.

ACKNOWLEDGMENTS

This document was prepared by the NOvA collaboration using the resources of the Fermi National Accelerator Laboratory (Fermilab), a U.S. 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. This work was supported by the U.S. Department of Energy; the U.S. National Science Foundation; the Department of Science and Technology, India; the European Research Council; the MSMT CR, GA UK, Czech Republic; the RAS, MSHE, and RFBR, Russia; CNPq and FAPEG, Brazil; UKRI, STFC and the Royal Society, United Kingdom; and the state and University of Minnesota. We are grateful for the contributions of the staffs of the University of Minnesota at the Ash River Laboratory, and of Fermilab. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising.

APPENDIX A: DETERMINING STEP SIZES FOR ARIA

As noted in Sec. III A 1, an $\mathrm{MR}^2\mathrm{T}^2$ chain is derived sample by sample using a repeated two-step procedure:

1. Proposal: The coordinates of a potential new sample are selected from a probability distribution centered on the current sample (or initial seed).
2. Acceptance: The proposal selected above is either accepted or rejected according to the rule of detailed balance, i.e., that every step in the chain be exactly reversible.

If accepted, then the proposed coordinates become the next sample. If rejected, then the previous sample is repeated to become the next sample.

The $\mathrm{MR}^2\mathrm{T}^2$ algorithm does not specify the distribution to be used in step 1 above, however. In our implementation, we use the most common choice, a multivariate Gaussian:

$$g(\vec{x}'|\vec{x}) = (2\pi)^{-\frac{N}{2}}(\det\Sigma)^{-\frac{1}{2}}\exp\left(-\frac{1}{2}(\vec{x}'-\vec{x})^T\Sigma^{-1}(\vec{x}'-\vec{x})\right), \tag{A1}$$

where $\vec{x}$ represents the current sample coordinates, $\vec{x}'$ the proposed next coordinates, and N the dimensionality of the coordinate space. The matrix Σ imposes a length scale on the "distance" between successive samples, and (especially when it is diagonal) its elements are usually called the "step sizes" of the sampling for each degree of freedom. The ideal asymptotic fraction of samples accepted in step 2, α, is 23.4% under a wide range of circumstances [99,100]. Though this figure is strictly true only for $N \to \infty$, it has been shown to hold approximately even for parameter counts as low as $N = 5$ [101]. Because the outcome of step 2 is related to the proposals generated in step 1, we tuned the values of Σ to arrive at $\alpha = 23.4\%$.

Our overall heuristic in the tuning procedure is to maintain step sizes that yield similar autocorrelations (defined rigorously below) across all the parameters. This results in the most efficient exploration of the parameter space [76]. We first optimized the step sizes for the parameters of interest, θ_{13}, θ_{23}, $|\Delta m^2_{32}|$, and δ_{CP}. We constructed a chain that sampled only those parameters using a unit matrix for Σ. We computed α for this chain and scaled the relevant elements of Σ in order to arrive at a tolerable preliminary acceptance rate of about 20%. We then computed the k-lag autocorrelation for each parameter θ, which measures the average correlation between MCMC sample n and sample $n+k$ across all n [102]:

$$r_k = \frac{\sum_{n=1}^{N-k}(\theta_n - \bar{\theta})(\theta_{n+k} - \bar{\theta})}{\sum_{n=1}^{N}(\theta_n - \bar{\theta})^2}, \tag{A2}$$

where θ_n refers to the value of parameter θ at step n, and $\bar{\theta}$ is its mean value. These autocorrelations are shown in Fig. 14. Using these r_k, we further adjusted the elements of Σ so the oscillation parameters would have similar autocorrelations.

To optimize the step sizes for the nuisance parameters (systematic uncertainties), we constructed a chain sampling only those parameters, again beginning with Σ entries of unity for them. As with the oscillation parameters, we adjusted Σ to ensure that none of the nuisance parameters had significantly different autocorrelations from the others. We then constructed a new, much longer chain, and subsequently computed a covariance matrix over the nuisance parameters from it. The Cholesky decomposition of this matrix, L, was used in the next step.

A final chain, this time sampling both oscillation and nuisance parameters, was constructed using the adjusted

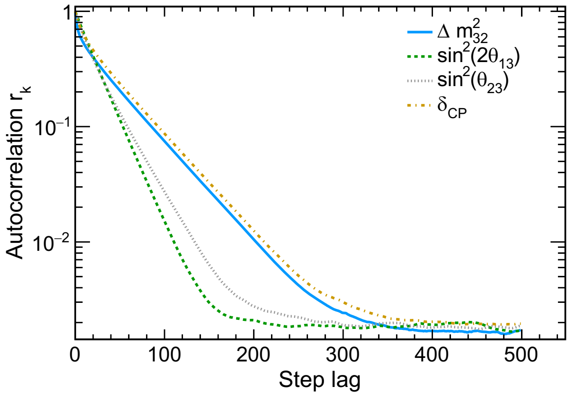


FIG. 14. Lag autocorrelation [see Eq. (A2)] computed using a $\mathrm{MR}^2\mathrm{T}^2$ chain sampling only oscillation parameters.

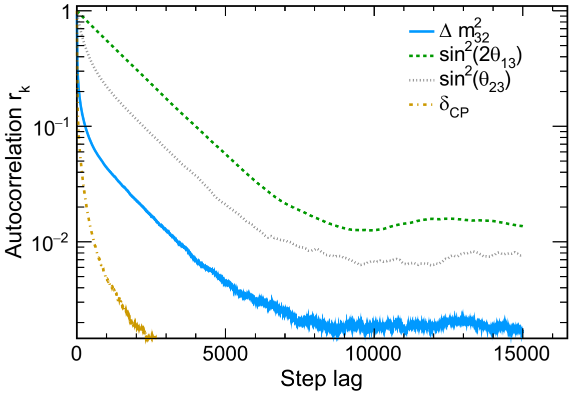


FIG. 15. Lag autocorrelation [see Eq. (A2)] computed using a $\mathrm{MR}^2\mathrm{T}^2$ chain after step size tuning with all parameters included.

Σ entries for both. While sampling, the proposed values for the nuisance parameters were multiplied by the decomposed covariance scaled by a tunable factor, βL. Using this chain, we recomputed the autocorrelations for all the parameters. The elements of Σ were readjusted to obtain similar autocorrelations from the oscillation parameters, now in the presence of the nuisance parameters. We also adjusted β to yield similar autocorrelations to those of the oscillation parameters. A final global scale was applied to Σ and β to finally arrive at $\alpha = 23.4\%$. The autocorrelations for the oscillation parameters at the end of the tuning procedure are shown in Fig. 15.

APPENDIX B: DETERMINING WARMUP FRACTION AND EFFECTIVE SAMPLE SIZE FOR ARIA

The property of sample proportionality to posterior density in the $\mathrm{MR}^2\mathrm{T}^2$ method is only guaranteed as asymptotic behavior. Therefore, it is usually necessary to discard some number of samples N_0 at the beginning of the Markov chain while the chain "burns in" or "warms up." Moreover, it is usually impossible to find a choice of step sizes Σ in Eq. (A1) that proposes samples that are both fully uncorrelated with the previous one and whose acceptance probabilities are high enough to not impose severe computing requirements. Chains can, in principle, be "thinned" by discarding all but the kth sample to reduce autocorrelations. Though we do not actually thin the chains, the resulting fraction $N_{\mathrm{eff}} = \frac{1}{k}(N - N_0)$, the "number of effective samples," still corresponds to the statistical power of the chain. The plateau in Fig. 15 of around or less than 1% represents the asymptotic behavior of the autocorrelations with the step size tuning procedure we use. Therefore we interpret our effective sample size as being computed as above with $k = 10^4$.

Because autocorrelations in our analysis are relatively long (see Fig. 15) and ARIA runs fairly quickly, we produced very long chains of 5×10^6 samples each. Thus, when we compared our posterior densities using $N_0 = 0$, $N_0 = \{1,3,5\} \times 10^3$, $N_0 = \{1,3,5\} \times 10^4$, and $N_0 = \{1,3\} \times 10^5$, we found them all to be indistinguishable. Thus, we did not find any need to discard warmup samples from our chains.

APPENDIX C: ALGORITHM CHOICES FOR HMCMC

As noted in Sec. III, HMCMC generates proposals by numerically integrating a Hamiltonian for a fictitious particle, whose potential arises from treating the log-posterior in analogy to gravity:

$$\frac{d\vec{q}}{dt} = \frac{\partial H}{\partial \vec{p}} = \frac{\partial T}{\partial \vec{p}}$$
$$\frac{d\vec{p}}{dt} = -\frac{\partial H}{\partial \vec{q}} = -\frac{\partial T}{\partial \vec{q}} - \frac{\partial V}{\partial \vec{q}} \tag{C1}$$

where T and V are the kinetic and potential energies of the system, respectively.

There are two ingredients of HMCMC left unspecified by the method. In both cases Stan's default choices were found to be suitable for our needs. The first is the distribution of kinetic energies from which T in Eq. (C1) is chosen. Stan's default is the Euclidean-Gaussian kinetic energy distribution:

$$T(\vec{q}, \vec{p}) = \frac{1}{2}\vec{p}^{\mathrm{T}} M^{-1} \vec{p} + \log|M| + \mathrm{const.} \tag{C2}$$

Here the mass matrix M (analogous to the effect of mass in gravitation) is a parameter that is automatically inferred by Stan during its warm-up sampling by iterative adjustments based on a running covariance over the samples. The second implementation choice is how long the integrator is allowed to run for each particular trajectory. Stan uses an algorithm called No-U-Turns [103], which is a heuristic method that halts integration when two trajectories, extending in each direction from the starting point along the initial momentum, begin to converge towards one another. An upper limit of integrator steps is also supplied as a parameter to Stan; in this analysis, we find that all our trajectories end within 2^{11} steps.

APPENDIX D: EQUIVALENCE OF ARIA AND STAN RESULTS

We extracted the posterior distributions for all of the results shown in this paper using both the ARIA and Stan samplers described in Sec. III. The posteriors were in all cases nearly indistinguishable, with tiny differences that occasionally caused the boundaries of credible intervals to shift by single bins. This is illustrated in Figs. 16 and 17.

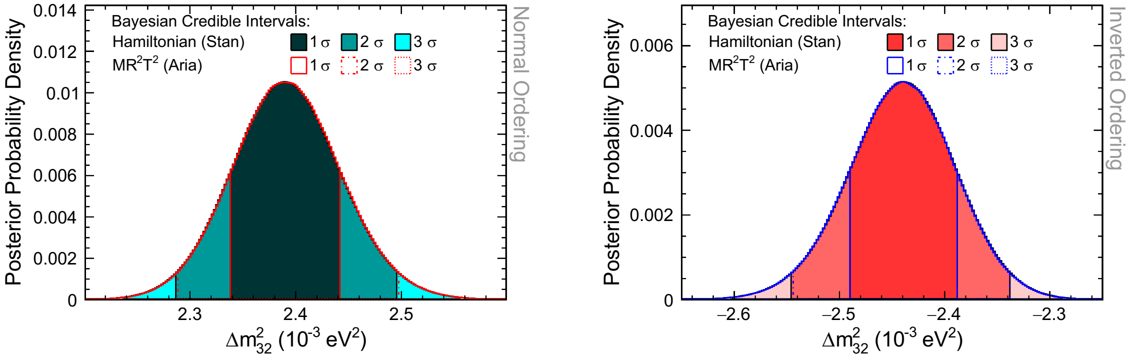


FIG. 16. Posterior probability density in Δm^2_{32} for normal ordering (left) and inverted ordering (right). Samples from Stan are shown as the shaded distributions (1σ, 2σ, 3σ regions indicated by darkest to lightest fill). Samples from ARIA are shown as the red (normal ordering) or blue (inverted ordering) overlaid lines (solid, dashed, dotted lines represent the boundaries of the 1σ, 2σ, 3σ regions, respectively).

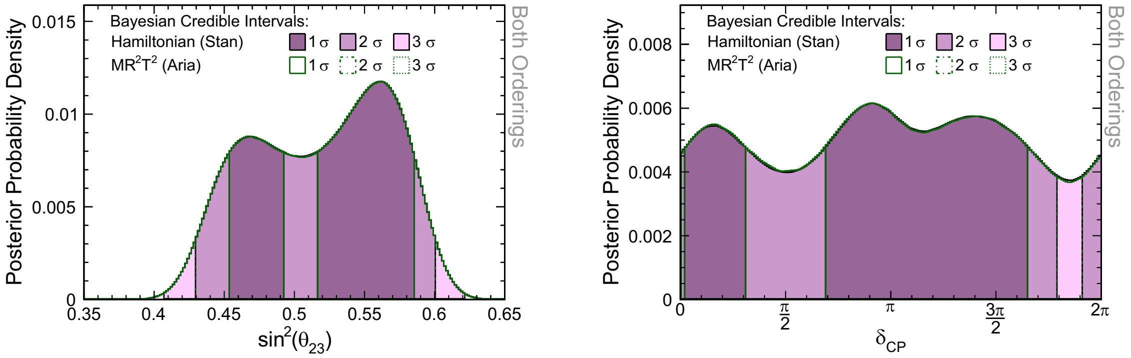


FIG. 17. Posterior probability density in $\sin^2\theta_{23}$ (left) and δ_{CP} (right), marginalized over both mass orderings. Line styles as in Fig. 16.

---

[1] Y. Fukuda *et al.* (Super-Kamiokande Collaboration), Phys. Rev. Lett. **81**, 1562 (1998).

[2] S. Fukuda *et al.* (Super-Kamiokande Collaboration), Phys. Lett. B **539**, 179 (2002).

[3] Q. R. Ahmad *et al.* (SNO Collaboration), Phys. Rev. Lett. **89**, 011301 (2002).

[4] K. Eguchi *et al.* (KamLAND Collaboration), Phys. Rev. Lett. **90**, 021802 (2003).

[5] D. G. Michael *et al.* (MINOS Collaboration), Phys. Rev. Lett. **97**, 191801 (2006).

[6] K. Abe *et al.* (T2K Collaboration), Phys. Rev. Lett. **107**, 041801 (2011).

[7] Y. Abe *et al.* (Double Chooz Collaboration), Phys. Rev. Lett. **108**, 131801 (2012).

[8] F. P. An *et al.* (Daya Bay Collaboration), Phys. Rev. Lett. **108**, 171803 (2012).

[9] J. K. Ahn *et al.* (RENO Collaboration), Phys. Rev. Lett. **108**, 191802 (2012).

[10] R. N. Mohapatra and A. Y. Smirnov, Annu. Rev. Nucl. Part. Sci. **56**, 569 (2006).

[11] H. Nunokawa, S. J. Parke, and J. W. F. Valle, Prog. Part. Nucl. Phys. **60**, 338 (2008).

[12] G. Altarelli and F. Feruglio, Rev. Mod. Phys. **82**, 2701 (2010).

[13] S. F. King, J. Phys. G **42**, 123001 (2015).

[14] S. T. Petcov, Eur. Phys. J. C **78**, 709 (2018).

[15] M. Fukugita and T. Yanagida, Phys. Lett. B **174**, 45 (1986).

[16] W. Buchmuller and M. Plumacher, Phys. Lett. B **389**, 73 (1996).

[17] W. Buchmuller, P. Di Bari, and M. Plumacher, Ann. Phys. (Amsterdam) **315**, 305 (2005).

[18] W. Buchmuller, R. D. Peccei, and T. Yanagida, Annu. Rev. Nucl. Part. Sci. **55**, 311 (2005).

[19] A. Pilaftsis, Phys. Rev. D **56**, 5431 (1997).

[20] M. A. Acero *et al.* (NOvA Collaboration), arXiv:2207.14353 [Phys. Rev. D (to be published)].

[21] K. Abe *et al.* (T2K Collaboration), Phys. Rev. D **108**, 072011 (2023).

[22] M. Jiang *et al.* (Super-Kamiokande Collaboration), Prog. Theor. Exp. Phys. **2019**, 053F01 (2019).

[23] R. Abbasi *et al.* (IceCube Collaboration), Phys. Rev. D **108**, 012014 (2023).

[24] R. L. Workman *et al.* (Particle Data Group), Prog. Theor. Exp. Phys. **2022**, 083C01 (2022).

[25] L. Wolfenstein, Phys. Rev. D **17**, 2369 (1978).

[26] K. Abe *et al.* (T2K Collaboration), Eur. Phys. J. C **83**, 782 (2023).

[27] P. Adamson *et al.* (MINOS Collaboration), Phys. Rev. Lett. **110**, 171801 (2013).

[28] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, T. Schwetz, and A. Zhou, J. High Energy Phys. 09 (2020) 178.

[29] P. Adamson *et al.* (NOvA Collaboration), Phys. Rev. D **93**, 051104 (2016).

[30] P. Adamson *et al.* (NOvA Collaboration), Phys. Rev. Lett. **116**, 151806 (2016).

[31] P. Adamson *et al.* (NOvA Collaboration), Phys. Rev. Lett. **118**, 151802 (2017).

[32] P. Adamson *et al.* (NOvA Collaboration), Phys. Rev. Lett. **118**, 231801 (2017).

[33] M. A. Acero *et al.* (NOvA Collaboration), Phys. Rev. D **98**, 032012 (2018).

[34] M. A. Acero *et al.* (NOvA Collaboration), Phys. Rev. Lett. **123**, 151803 (2019).

[35] M. A. Acero *et al.* (NOvA Collaboration), Phys. Rev. D **106**, 032004 (2022).

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

[37] P. Adamson *et al.*, Nucl. Instrum. Methods Phys Res., Sect. A **806**, 279 (2016).

[38] S. Agostinelli *et al.* (GEANT4 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **506**, 250 (2003).

[39] Geant4 Collaboration, geant4-data.web.cern.ch, https://geant4-data.web.cern.ch/ReleaseNotes/ReleaseNotes4.10.4.html (2017).

[40] L. Aliaga *et al.* (MINERvA Collaboration), Phys. Rev. D **94**, 092005 (2016); **95**, 039903(A) (2017).

[41] J. M. Paley *et al.* (MIPP Collaboration), Phys. Rev. D **90**, 032001 (2014).

[42] C. Alt *et al.* (NA49 Collaboration), Eur. Phys. J. C **49**, 897 (2007).

[43] N. Abgrall *et al.* (NA61/SHINE Collaboration), Phys. Rev. C **84**, 034604 (2011).

[44] D. S. Barton *et al.*, Phys. Rev. D **27**, 2580 (1983).

[45] S. M. Seun, Measurement of $\pi - K$ ratios from the NuMI target, Ph.D. thesis, Harvard University, 2007, 10.2172/935004.

[46] G. M. Tinti, Sterile neutrino oscillations in MINOS and hadron production in pC collisions, Ph.D. thesis, Oxford University, 2010.

[47] A. V. Lebedev, Ratio of pion kaon production in proton carbon interactions, Ph.D. thesis, Harvard University, 2007, 10.2172/948174.

[48] B. Baatar *et al.* (NA49 Collaboration), Eur. Phys. J. C **73**, 2364 (2013).

[49] P. Skubic *et al.*, Phys. Rev. D **18**, 3115 (1978).

[50] S. P. Denisov, S. V. Donskov, Yu. P. Gorin, R. N. Krasnokutsky, A. I. Petrukhin, Yu. D. Prokoshkin, and D. A. Stoyanova, Nucl. Phys. **B61**, 62 (1973).

[51] A. S. Carroll *et al.*, Phys. Lett. **80B**, 319 (1979).

[52] K. Abe *et al.* (T2K Collaboration), Phys. Rev. D **87**, 012001 (2013); **87**, 019902(A) (2013).

[53] T. K. Gaisser, G. B. Yodh, V. D. Barger, and F. Halzen, in *14th International Cosmic Ray Conference (ICRC 1975) Munich, Germany, 1975* (Max-Planck Inst. for Extraterrestriche Physik, Munich, 1975), pp. 2161–2166.

[54] J. W. Cronin, R. Cool, and A. Abashian, Phys. Rev. **107**, 1121 (1957).

[55] J. V. Allaby *et al.* (IHEP-CERN Collaboration), Phys. Lett. **30B**, 500 (1969).

[56] M. J. Longo and B. J. Moyer, Phys. Rev. **125**, 701 (1962).

[57] B. M. Bobchenko *et al.*, Yad. Fiz. **30**, 1553 (1979) [Sov. J. Nucl. Phys. **30**, 805 (1979)].

[58] V. B. Fedorov, Yu. G. Grishuk, M. V. Kosov, G. A. Leksin, N. A. Pivnyuk, S. V. Shevchenko, V. L. Stolin, A. V. Vlasov, and L. S. Vorobev, Yad. Fiz. **27**, 413 (1978) [Sov. J. Nucl. Phys. **27**, 222 (1978)].

[59] R. J. Abrams, R. L. Cool, G. Giacomelli, T. F. Kycia, B. A. Leontic, K. K. Li, and D. N. MichaelPhys. Rev. D **1**, 1917 (1970)]

[60] C. Andreopoulos *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **614**, 87 (2010).

[61] C. Andreopoulos, C. Barry, S. Dytman, H. Gallagher, T. Golan, R. Hatcher, G. Perdue, and J. Yarba, arXiv:1510.05494.

[62] J. Nieves, I. Ruiz Simo, and M. J. Vicente Vacas, Phys. Rev. C **83**, 045501 (2011).

[63] R. Gran, J. Nieves, F. Sanchez, and M. J. Vicente Vacas, Phys. Rev. D **88**, 113007 (2013).

[64] B. W. Allardyce *et al.*, Nucl. Phys. **A209**, 1 (1973).

[65] A. Saunders, S. Hoeibraten, J. J. Kraushaar, B. J. Kriss, R. J. Peterson, R. A. Ristinen, J. T. Brack, G. Hofman, E. F. Gibson, and C. L. Morris, Phys. Rev. C **53**, 1745 (1996).

[66] O. Meirav, E. Friedman, R. R. Johnson, R. Olszewski, and P. Weber, Phys. Rev. C **40**, 843 (1989).

[67] S. M. Levenson *et al.*, Phys. Rev. C **28**, 326 (1983).

[68] D. Ashery, I. Navon, G. Azuelos, H. K. Walter, H. J. Pfeiffer, and F. W. Schleputz, Phys. Rev. C **23**, 2173 (1981).

[69] D. Ashery *et al.*, Phys. Rev. C **30**, 946 (1984).

[70] E. S. Pinzon Guerra *et al.* (DUET Collaboration), Phys. Rev. C **95**, 045203 (2017).

[71] A. Aurisano, C. Backhouse, R. Hatcher, N. Mayer, J. Musser, R. Patterson, R. Schroeter, and A. Sousa (NOvA Collaboration), J. Phys. Conf. Ser. **664**, 072002 (2015).

[72] A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M. D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa, and P. Vahle, J. Instrum. **11**, P09001 (2016).

[73] F. Psihas, E. Niner, M. Groh, R. Murphy, A. Aurisano, A. Himmel, K. Lang, M. D. Messier, A. Radovic, and A. Sousa, Phys. Rev. D **100**, 073005 (2019).

[74] Mr. Bayes and Mr. Price, Philos. Trans. (1683-1775) **53**, 370 (1763).

[75] J. S. Speagle, arXiv:1909.12313.

[76] S. Brooks, A. Gelman, G. Jones, and X.-L. Meng, Handbook of Markov Chain Monte Carlo (Chapman & Hall/CRC, Boca Raton, FL, 2011), pp. 1–592.

[77] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[78] W. K. Hastings, Biometrika **57**, 97 (1970).

[79] J. E. Gubernatis, Phys. Plasmas **12**, 057303 (2005).

[80] M. Rosenbluth, AIP Conf. Proc. **690**, 22.

[81] Stan Development Team, Stan Modeling Language Users Guide and Reference Manual, version 2.26.1, Stan Development Team (2021).

[82] M. Betancourt, arXiv:1701.02434.

[83] M. Tanabashi *et al.* (Particle Data Group), Phys. Rev. D **98**, 030001 (2018) and 2019 update.

[84] N. Nayak, A joint measurement of ν_μ-disappearance and ν_e-appearance in the NuMI beam using the NOvA experiment, Ph.D. thesis, UC, Irvine (main), UC, Irvine, 2021.

[85] A. Gelman, X.-L. Meng, and H. Stern, Statistica Sinica **6**, 733 (1996).

[86] A. Gelman, Y. Goegebeur, F. Tuerlinckx, and I. Van Mechelen, J. R. Stat. Soc. Ser. C **49**, 247 (2002).

[87] A. Gelman, Electron. J. Stat. **7**, 2595 (2013).

[88] See the Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevD.110.012005 for one- and two-dimensional posterior density distributions under all combinations of usage of reactor constraint and marginalization scheme.

[89] H. Jeffreys, *Theory of Probability*, 3rd ed. (Oxford University Press, Oxford, England, 1961).

[90] R. E. Kass and A. E. Raftery, J. Am. Stat. Assoc. **90**, 773 (1995).

[91] P. B. Denton and S. J. Parke, Phys. Rev. D **100**, 053004 (2019).

[92] C. Jarlskog, Phys. Rev. Lett. **55**, 1039 (1985).

[93] P. B. Denton and R. Pestes, J. High Energy Phys. 05 (2021) 139.

[94] J. M. Dickey, Ann. Math. Stat. **42**, 204 (1971).

[95] J. Mulder, E. J. Wagenmakers, and M. Marsman, arXiv:2004.09899.

[96] T. P. Hill, Trans. Am. Math. Soc. **363**, 3351 (2011).

[97] R. F. Fox, T. P. Hill, and J. Miller, Chaos **21**, 033102 (2011).

[98] B. Abi *et al.* (DUNE Collaboration), Eur. Phys. J. C **80**, 978 (2020).

[99] A. Gelman, W. R. Gilks, and G. O. Roberts, Ann. Appl. Probab. **7**, 110 (1997).

[100] G. O. Roberts, Stochastics and Stochastic Rep. **62**, 275 (1998).

[101] G. O. Roberts and J. S. Rosenthal, Stat. Sci. **16**, 351 (2001).

[102] *NIST/SEMATECH e-Handbook of Statistical Methods*, edited by C. Croarkin and P. Tobias (2013), 10.18434/M32189.

[103] M. D. Hoffman and A. Gelman, J. Mach. Learn. Res. **15**, 1593 (2014).