Very light sterile neutrinos at NOvA and T2K

André de Gouvêa[1], Giancarlo Jusino Sánchez[1], and Kevin J. Kelly[2,3]

[1]*Northwestern University, Department of Physics and Astronomy, 2145 Sheridan Road, Evanston, Illinois 60208, USA*
[2]*Theoretical Physics Department, Fermilab, P.O. Box 500, Batavia, Illinois 60510, USA*
[3]*Theoretical Physics Department, CERN, Esplande des Particules, 1211 Geneva 23, Switzerland*

Over the last several years, our understanding of neutrino oscillations has developed significantly due to the long-baseline measurements of muon-neutrino disappearance and muon-to-electron-neutrino appearance at the T2K and NOvA experiments. However, when interpreted under the standard-three-massive-neutrinos paradigm, a tension has emerged between the two experiments' data. Here, we examine whether this tension can be alleviated when a fourth, very light neutrino is added to the picture. Specifically, we focus on the scenario in which this new neutrino has a mass similar to, or even lighter than, the three mostly active neutrinos that have been identified to date. We find that, for some regions of parameter space, the four-neutrino framework is favored over the three-neutrino one with moderate ($\lesssim 2\sigma$) significance. Interpreting these results, we provide a future outlook for near-term and long-term experiments if this four-neutrino framework is indeed true.

I. INTRODUCTION

Long-baseline neutrino oscillation experiments aim at studying the phenomenon of neutrino oscillations by taking advantage of the known neutrino oscillation lengths, proportional to (the inverse of) the mass-squared differences $\Delta m_{21}^2 \equiv m_2^2 - m_1^2$ or $\Delta m_{31}^2 \equiv m_3^2 - m_1^2$, where $m_{1,2,3}$ are the masses of the neutrino mass eigenstates $\nu_{1,2,3}$, respectively. The neutrino masses are labeled such that $m_2^2 > m_1^2$ and $|\Delta m_{31}^2| > \Delta m_{21}^2$. With this definition, the sign of Δm_{31}^2 is an observable and captures the neutrino-mass ordering: normal ordering (NO) when Δm_{31}^2 is positive, inverted ordering (IO) when Δm_{31}^2 is negative.

Among the objectives of long-baseline experiments is testing the standard-three-massive-neutrinos paradigm, which states that there are three neutrino mass eigenstates and that these interact via neutral-current and charged-current weak interactions. As far as the charged-current weak interactions are concerned, three orthogonal linear combinations of $\nu_{1,2,3}$ couple to the W-boson and the charged leptons ℓ_α ($\alpha = e, \mu, \tau$). In more detail, $\nu_\alpha = U_{\alpha i} \nu_i$ ($i = 1, 2, 3$) couples to ℓ_α and the W-boson, and $U_{\alpha i}$ are the elements of the unitary leptonic mixing matrix. On the other hand, assuming the standard-three-massive-neutrinos paradigm is correct, long-baseline experiments are capable of measuring, sometimes with great precision, the neutrino oscillation parameters—the parameters which define $U_{\alpha i}$ and the mass-squared differences. One way to test the standard-three-massive-neutrinos paradigm is to assume it is correct; measure the oscillation parameters using different oscillation processes or different experimental setups; and compare the results. If different measurements of the same quantity disagree at a high confidence level, we would claim the underlying formalism—in this case the standard three-massive-neutrinos paradigm—is deficient.

Among the current generation of long-baseline experiments are the Tokai to Kamioka experiment (T2K) [1,2], in Japan, and the NuMI Off-axis ν_e Appearance (NOvA) experiment [3,4], in the United States. They are sensitive to several of the neutrino oscillation parameters, including some that are, at present, virtually unknown: the neutrino mass-ordering and the CP-odd parameter δ_{CP} that governs whether and how much CP-invariance is violated in the lepton sector. Data from T2K and NOvA have been analyzed assuming the standard-three-massive-neutrinos paradigm and have led to interesting measurements of the oscillation parameters. Just as interesting, perhaps, is the fact that there is some tension between T2K and NOvA data.

The tension, which was first demonstrated by Refs. [5,6], has been quantified and examined critically in the three-neutrino framework by various authors [7–10]. In a little more detail, both T2K and NOvA measure electronlike and

muonlike events associated to a pion decay-in-flight neutrino source ($\pi \to \mu\nu_\mu$). Measurements are performed at both near and far detectors and the detectors are exposed to both "neutrino" and "antineutrino" beams. With all this information, they can infer the ν_μ and $\bar{\nu}_\mu$ survival probabilities $P(\nu_\mu \to \nu_\mu)$ and $P(\bar{\nu}_\mu \to \bar{\nu}_\mu)$, respectively, and the ν_e and $\bar{\nu}_e$ appearance probabilities $P(\nu_\mu \to \nu_e)$ and $P(\bar{\nu}_\mu \to \bar{\nu}_e)$, respectively. At T2K, typical neutrino energies are around 600 MeV and the baseline is 295 km. Typical NOvA energies are around 2 GeV and the baseline is 810 km.

Assuming the standard-three-massive-neutrinos paradigm, the T2K and NOvA disappearance data are consistent but the appearance data, for both neutrinos and antineutrinos, are in disagreement when analyzed within the NO. If the mass ordering is NO, T2K prefers δ_{CP} values close to $3\pi/2$.[1] In contrast, when analyzed under the NO, NOvA data have no strong preference for any particular value of δ_{CP}, however, they disfavor the combination of δ_{CP} and the mixing angle $\sin^2 \theta_{23}$ preferred by T2K at roughly 2σ confidence. This tension may be addressed by instead considering the IO, where both experiments prefer $\delta_{CP} \approx 3\pi/2$ [2,4,7]. However, global fits to all neutrino oscillation data [8–10] (particularly including those from reactor antineutrino experiments [11–13]), as well as comparisons with cosmological observations [14,15], prefer NO at $\sim 2 - 3\sigma$, leaving the T2K-NOvA tension unaddressed.

Whether the tension can be alleviated by the presence of physics beyond the standard-three-massive-neutrinos paradigm has also been the subject of intense exploration (see, for example, Refs. [16–22]). Here, we would like to explore, in some detail, whether the tension between T2K and NOvA can be interpreted as evidence for new light neutrino states. This issue has been discussed before [18], assuming the new neutrino state ν_4 with mass m_4 is relatively heavy: $|\Delta m_{41}^2| \gg |\Delta m_{31}^2|$. Instead, here we concentrate on $|\Delta m_{41}^2|$ values that are $\mathcal{O}(|\Delta m_{31}^2|)$ or smaller, down to $\mathcal{O}(\Delta m_{21}^2)$, and explore the full parameter space associated with the fourth neutrino. In Sec. II, we describe the four-neutrino oscillation formalism of interest. We also discuss how the existence of a light fourth neutrino may help alleviate the T2K–NOvA tension. In Sec. III we present our simulations of NOvA and T2K data and discuss how these are used, in Sec. IV, to compare the standard-three-massive-neutrinos paradigm and the fourth-neutrino hypothesis. We present some concluding remarks in Sec. V. Some results are included in appendices: Appendix A includes detailed numerical results from our analyses, Appendix B presents an alternate, extremely light sterile neutrino analysis, and Appendix C discusses some

Monte Carlo studies of T2K, NOvA, and their combination in light of the sterile neutrino analyses.

II. FOUR-FLAVOR NEUTRINO OSCILLATIONS

We assume there are four neutrino mass eigenstates $\nu_{1,2,3,4}$, and that these are related to the four interaction eigenstates $\nu_{e,\mu,\tau}$ and ν_s (where we assume the ν_s state does not participate in the weak interactions) via a 4×4 unitary mixing matrix:

$$U = R(\theta_{34})R(\theta_{24}, \delta_{24})R(\theta_{14}, \delta_{14})R(\theta_{23})R(\theta_{13}, \delta_{13})R(\theta_{12}),$$
$$(2.1)$$

where R are 4×4 rotation matrices in the ij-plane associated with a rotation angle θ_{ij}. The nontrivial entries of the different R in Eq. (2.1) are given by

$$R(\theta_{ij}) = \begin{pmatrix} c_{ij} & s_{ij} \\ -s_{ij} & c_{ij} \end{pmatrix} \quad R(\theta_{ij}, \delta_{ij}) = \begin{pmatrix} c_{ij} & s_{ij}e^{-\delta_{ij}} \\ -s_{ij}e^{\delta_{ij}} & c_{ij} \end{pmatrix},$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. This extension to the standard-three-massive-neutrinos paradigm includes one more independent mass-squared difference and five new mixing parameters: three mixing angles ($\theta_{14}, \theta_{24}, \theta_{34}$) and two complex phases (δ_{14}, δ_{24}).

The 4×4 mixing matrix is defined in such a way that, in the limit $\theta_{14}, \theta_{24}, \theta_{34} \to 0$, $\nu_4 = \nu_s$ and $\nu_{1,2,3}$ are linear superpositions of only the active states $\nu_{e,\mu,\tau}$. In this limit, we recover the standard-three-massive-neutrinos paradigm. We will be interested in the case where θ_{14}, θ_{24}, θ_{34} are relatively small and will refer to $\nu_{1,2,3}$ as the mostly active states. The mostly active states will be defined in the usual way, including the ordering of their masses, which is either "normal" (NO) or "inverted" (IO), as discussed in Sec. I. With this in mind, we define

$$\Delta m_{4l}^2 \equiv \begin{cases} m_4^2 - m_1^2, & \text{if } m_1 < m_3 \text{ (NO)} \\ m_4^2 - m_3^2, & \text{if } m_3 < m_1 \text{ (IO)} \end{cases}. \quad (2.2)$$

In order to allow for all different relevant orderings of the four masses, we allow for both the NO and IO of the mostly active states and for both positive and negative values of Δm_{4l}^2. The four qualitatively different mass orderings are depicted in Fig. 1. As far as the magnitude of Δm_{4l}^2, we will restrict our analyses to $(10^{-5} < |\Delta m_{4l}^2| < 10^{-1})$ eV^2. Inside this range, we expect nontrivial oscillation effects to manifest themselves in the far detectors of T2K and NOvA but not in the corresponding near detectors. When $|\Delta m_{4l}^2|$ is smaller than 10^{-5} eV^2, the new oscillation length associated to Δm_{4l}^2 is too long and outside the reach of T2K and NOvA. Instead, when $|\Delta m_{4l}^2|$ is larger than 10^{-1} eV^2, we expect very fast oscillations in the far detectors of T2K and NOvA and nontrivial effects in the corresponding near

---

[1]We will use the convention that CP-violating phases are defined over $[0, 2\pi]$.
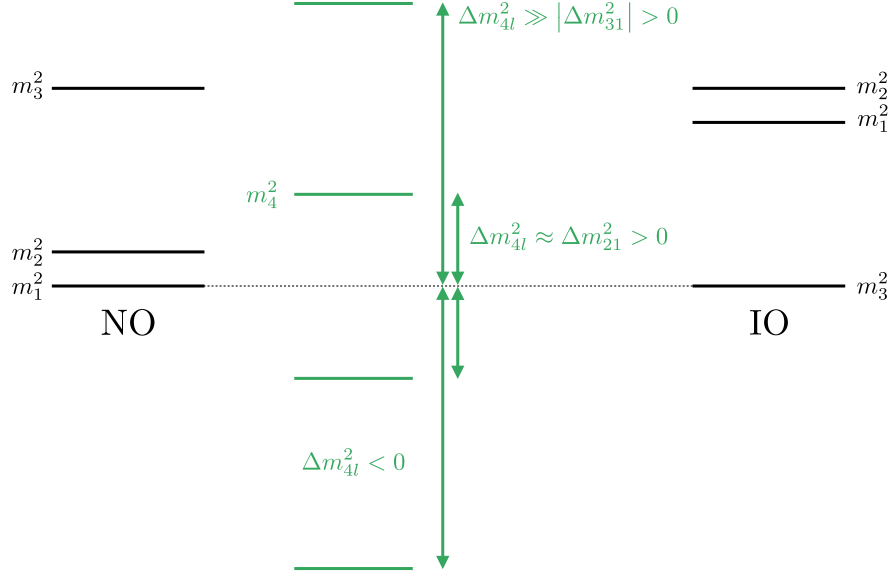
FIG. 1.   Definition, including the sign convention, of Δm^2_{4l} given the NO or IO for the mostly active states.

detectors. This region of parameter space was explored in Ref. [18].

The active neutrinos interact with the medium as they propagate from the source to the far detector. These interactions modify the equations that govern the flavor evolution of the neutrino states via effective potentials for forward charged-current (CC) and neutral-current (NC) scattering. The neutrino flavor evolution equation can be written as a Schrödinger-like equation with an effective Hamiltonian given by, in the flavor basis, $H_F = 1/(2E_\nu)(U\mathbf{M}^2 U^\dagger + \mathbf{A})$, where

$$\mathbf{M}^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \Delta m^2_{21} & 0 & 0 \\ 0 & 0 & \Delta m^2_{31} & 0 \\ 0 & 0 & 0 & \Delta m^2_{41} \end{pmatrix},$$

$$\mathbf{A} = \begin{pmatrix} 2E_\nu V_{\mathrm{CC}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2E_\nu V_{\mathrm{NC}} \end{pmatrix}. \qquad (2.3)$$

For neutrinos, $V_{\mathrm{CC}} = -2V_{\mathrm{NC}} = 3.8 \times 10^{-5}\,(\mathrm{eV}^2/\mathrm{GeV})\rho[\frac{\mathrm{g}}{\mathrm{cm}^3}]$ are the CC and NC matter potentials, respectively. The CC and NC matter potentials arise from active-neutrino interactions with electrons and all matter particles, respectively—because the NC interaction is identical for all active neutrinos but absent for the sterile neutrino, it appears with a negative sign in the "sterile-sterile" element of Eq. (2.3), $\mathbf{A}_{ss} = -2E_\nu V_{\mathrm{NC}}$. For antineutrinos, the matter potentials have the opposite sign. ρ is the density—assumed to be constant—of the medium, assumed to be neutral. When neutrality is assumed, V_{NC} is half as large as

V_{CC} and negative. For the NOvA and T2K experiments, we fix the baselines to be $L_{\mathrm{NOvA}} = 810\,\mathrm{km}$ and $L_{\mathrm{T2K}} = 295\,\mathrm{km}$, respectively, while the near-far detector average matter densities are taken to be, respectively, $\rho_{\mathrm{NOvA}} = 2.8\,\mathrm{g/cm}^3$ [4] and $\rho_{\mathrm{T2K}} = 2.6\,\mathrm{g/cm}^3$ [2].

We are interested in improving the quality of fits to the combined T2K and NOvA data, particularly in the ν_e and $\bar{\nu}_e$ appearance channels. If allowing for a fourth neutrino, with $|\Delta m^2_{4l}| \approx 10^{-2}\,\mathrm{eV}^2$, then the oscillation probability $P(\nu_\mu \to \nu_e)$ will change significantly for the baseline lengths and energies of interest for the two experiments. To try to understand this potential fit improvement, we show in Fig. 2 the oscillation probabilities for (anti) neutrinos at T2K and NOvA under the best-fit three-neutrino hypothesis, with parameters given in the 3ν IO column of Table I and obtained with the analyses described below. Similarly, we also present the best-fit oscillation probabilities at these energies under the four-neutrino hypothesis for the best fit we obtain, with oscillation parameters also given in the 4ν IO column of Table I. We also allow $\sin^2\theta_{34}$ to vary beyond the nominal value of 0.56 in Table I, in order to demonstrate that it has a noticeable impact (despite not appearing in the vacuum-calculated oscillation probabilities), especially at NOvA. Upon inspection of Fig. 2, we highlight two main reasons that the effects at T2K and NOvA may be different. One is that the dominant values of L/E, keeping in mind that both beams have a narrow energy profile, are not identical for the two experiments. This means that for relatively "fast" Δm^2_{4l} the value of the new oscillation phase will not be the same for the two experiments. The other is that the matter effects are more pronounced at NOvA relative to T2K. These allow the effective oscillation frequencies and mixing parameters to be distinct at the two experimental setups.
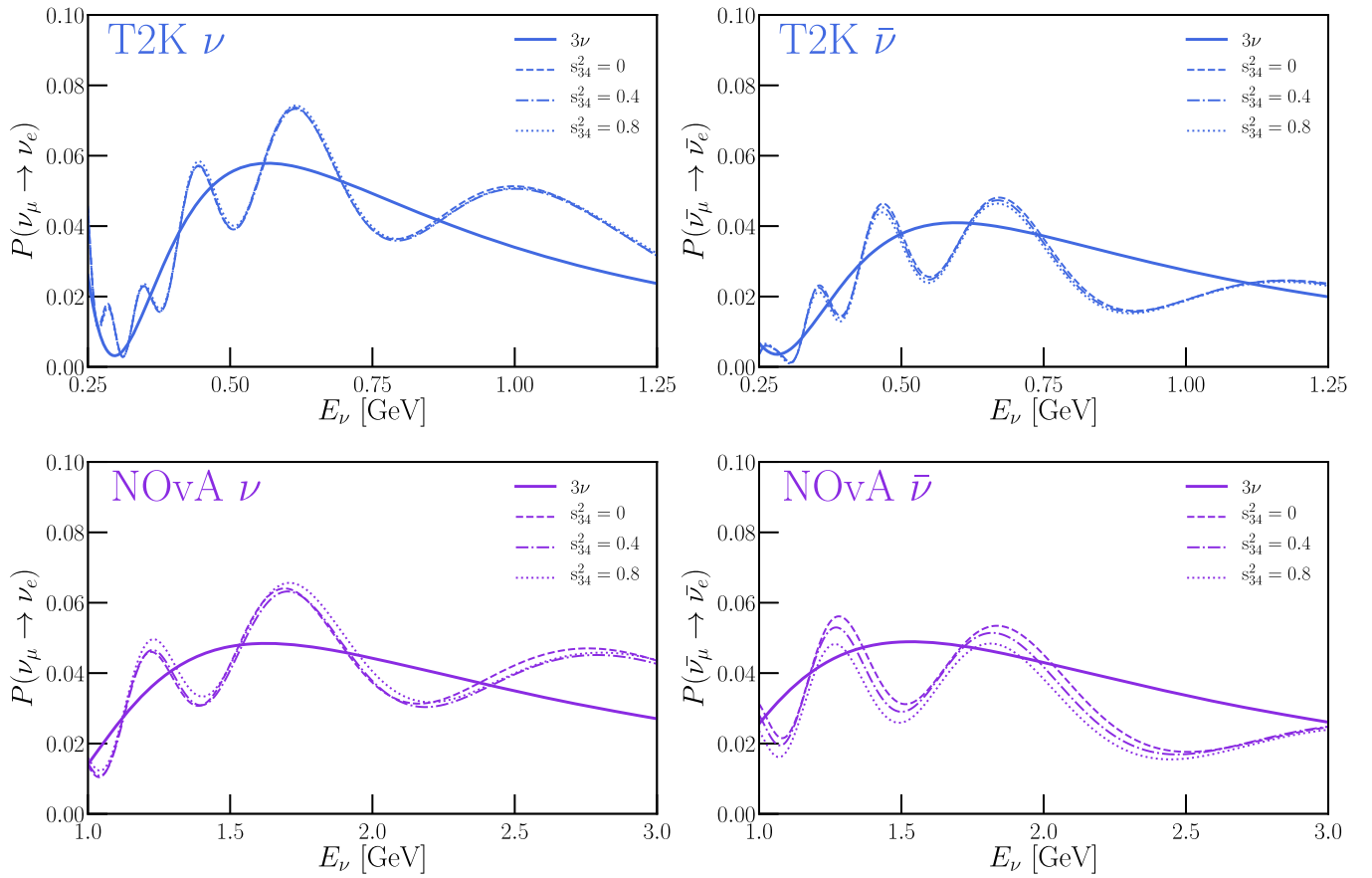
FIG. 2. Appearance oscillation probabilities at T2K (top, blue) and NOvA (bottom, purple) comparing three-neutrino oscillation probabilities (solid lines, parameters from Table I, column 2 "3ν IO") against four-neutrino ones (nonsolid lines, parameters from Table I, column 4 "4ν IO"). Left panels show probabilities for neutrino oscillation, whereas right ones show antineutrino oscillation. For the four-neutrino probabilities, three choices of $\sin^2 \theta_{34}$ are used for illustrative purposes: dashed/dot-dashed/dotted lines correspond to $\sin^2 \theta_{34} = 0/0.4/0.8$.

In vacuum, $P(\nu_\mu \to \nu_e)$ does not depend on θ_{34}; this is not the case in matter. An easy way to see this is to express the propagation Hamiltonian in the mass basis. In the absence of matter effects, the dependency on the mixing parameters is encoded in the initial and final interaction eigenstates and since neither ν_e nor ν_μ, when expressed as linear superpositions of the mass eigenstates, depend on θ_{34}, then neither can $P(\nu_\mu \to \nu_e)$. Instead, when the matter effects are present, the matter potential in the mass basis depends on θ_{34}. Hence we expect $P(\nu_\mu \to \nu_e)$ to also depend on θ_{34} as long as matter effects are relevant. The dependency on θ_{34} can be seen in Fig. 2. As expected, it is rather small at T2K and larger at NOvA, where matter effects are relatively more pronounced.

Another nontrivial aspect of four-neutrino oscillations for T2K and NOvA is the sterile-sterile component of the matter potential, $\mathbf{A}_{ss}$. As discussed above, for a sterile neutrino, $\mathbf{A}_{ss} = -2E_\nu V_{NC}$, but a fourth active neutrino would have $\mathbf{A}_{ss} = 0$, or equivalently, $V_{NC} = 0$. Figure 3 depicts the ratio of the appearance probabilities in matter relative to what those would be in vacuum. We show this

comparison for the three-neutrino hypothesis (solid lines) as well as the four-neutrino one, where we compare the two hypotheses, $V_{NC} = -1/2V_{CC}$ (a sterile neutrino) and 0 (an active one). Here, we see that the "sterileness" of the fourth neutrino has an observable impact on the oscillation probabilities at both T2K and especially NOvA. We emphasize however that all of the analyses below make the more theoretically robust assumption that the fourth neutrino is sterile and $V_{NC} = -1/2V_{CC}$.

III. SIMULATING DATA FROM NOvA AND T2K

As discussed earlier, both NOvA and T2K operate with beams with a flux of predominantly ν_μ ($\bar{\nu}_\mu$) when operating in (anti)neutrino mode. Both experiments' far detectors are designed to study the disappearance of ν_μ and $\bar{\nu}_\mu$, as well as the appearance of ν_e and $\bar{\nu}_e$. Using the most recent publications from NOvA [4] and T2K [2], and building off the simulations of Refs. [7,23,24], we perform simulations to determine the expected event rates in the disappearance and appearance channels of
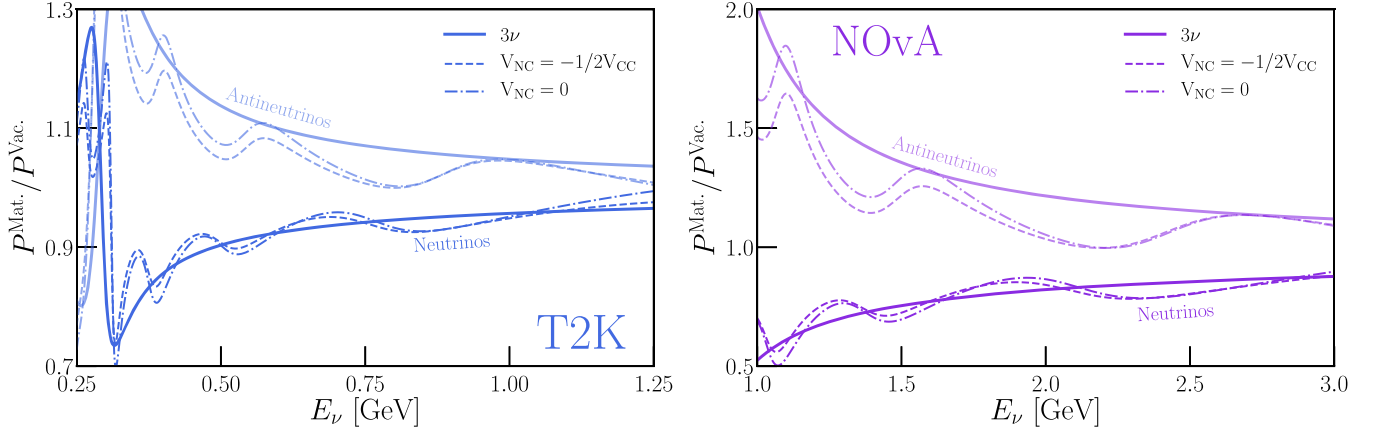
FIG. 3. Ratio of appearance oscillation probabilities in matter to those in vacuum at T2K (left) and NOvA (right). Solid lines correspond to the three-neutrino oscillation probabilities. Dashed and dot-dashed lines correspond to a fourth neutrino that is sterile or active, respectively. Parameters are taken from columns 2 and 4 from Table I corresponding to the three-neutrino and four-neutrino cases, respectively.

both experiments given a set of three- or four-neutrino oscillation parameters. We then compare these expected event rates against the experiments' published event rates and construct a test statistic using Poissonian bin expectations.

In the remainder of this section, we briefly explain the process by which we simulate the expected event rates, as well as the number of data points for each experiment that enter our test statistic. To center our discussion, we will rely on several benchmark sets of oscillation parameters with which we calculate the expected observables at NOvA and T2K. We adopt two benchmark sets each for the 3ν and 4ν assumptions, listed in Table I, allowing for the mostly active neutrinos to follow either the normal (NO) or

TABLE I. Oscillation parameters assumed when depicting oscillation probabilities and expected event rates. The four columns correspond to the three-neutrino (3ν) and four-neutrino (4ν) hypotheses, as well as whether the three mostly active neutrinos follow the normal (NO) or inverted (IO) mass ordering.

Parameter	3ν NO	3ν IO	4ν NO	4ν IO
$\sin^2\theta_{12}$	0.307	0.307	0.321	0.314
$\sin^2\theta_{13}$	0.022	0.022	0.023	0.023
$\sin^2\theta_{23}$	0.57	0.57	0.43	0.45
$\Delta m_{21}^2/10^{-5}$ eV^2	7.53	7.53	7.53	7.53
$\Delta m_{31}^2/10^{-3}$ eV^2	2.51	-2.41	2.49	-2.39
δ_{CP}	3.66	4.71	4.09	4.46
$\sin^2\theta_{14}$	...	...	0.043	0.021
$\sin^2\theta_{24}$	...	...	0.060	0.053
$\sin^2\theta_{34}$	...	...	0.37	0.56
$\Delta m_{41}^2/$eV^2	...	...	1.1×10^{-2}	-1.1×10^{-2}
δ_{14}	...	...	0.01	4.88
δ_{24}	...	...	1.82	5.89

inverted (IO) orderings. As we will discuss in Sec. IV, these parameters are the best-fit points obtained by our fit to the *combination* of T2K and NOvA under the different hypotheses.

A. NOvA

Our simulation of NOvA, designed to match the results of Ref. [4], includes the disappearance channels of neutrino and antineutrino mode (19 bins each, with neutrino energies ranging from 0 to 5 GeV) as well as event rate measurements of the appearance channels,[2] totaling 40 data points. This simulation corresponds to a total exposure of 13.6×10^{20} (12.5×10^{20}) protons on target (POT) in (anti) neutrino mode.

Figure 4 shows the expected event rates in NOvA for neutrino mode ν_μ disappearance (left), antineutrino mode $\bar\nu_\mu$ disappearance[3] (center), and a joint comparison of neutrino (x-axis) and antineutrino (y-axis) mode $\nu_\mu \to \nu_e$ (or $\bar\nu_\mu \to \bar\nu_e$) appearance (right panel). We compare the NOvA benchmark oscillation predictions, using the parameters in Table I (purple histograms/curves[4] for NO, green for IO, and dark curves for 3ν, faint ones for 4ν), to the observed event rates from the experiment (black). Error bars here are only statistical. In the left and center panels,

---

[2] For simplicity, we sum the expected event rate for the entire neutrino energy range and compare it against the observed 82 (33) appearance events of operation in (anti)neutrino mode. We have compared this approach to one that includes the spectral information of the appearance channels and find no qualitative impact on our results.

[3] In contrast to Ref. [4], our disappearance channel panels depict the event rate per bin as opposed to event rate per unit energy, causing our higher-energy bins (with larger bin width) to appear exaggerated.

[4] Where the faint curves are not visible in the left/center panels, the four-neutrino hypothesis predicts the same rate as the three-neutrino one(s).
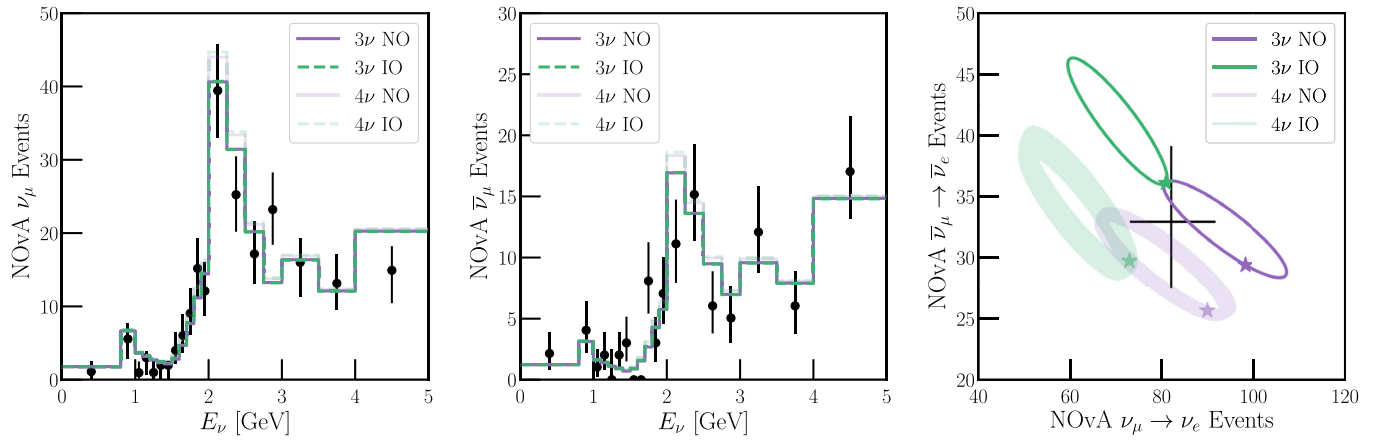
FIG. 4. Expected and observed event rates in NOvA's ν_μ disappearance (left), $\bar{\nu}_\mu$ disappearance (center), and $\nu_e/\bar{\nu}_e$ appearance (right) channels. We compare the prediction under the 3ν (solid/dashed lines) and 4ν (faint lines/regions) hypotheses, with parameters from Table I, with the observed data (black). Purple curves correspond to the mostly active neutrinos following the normal mass ordering (NO), where green ones correspond to the inverted mass ordering (IO). In the right panel, the *CP*-violating phases are allowed to vary in the predicted rates. Data points from Ref. [4].
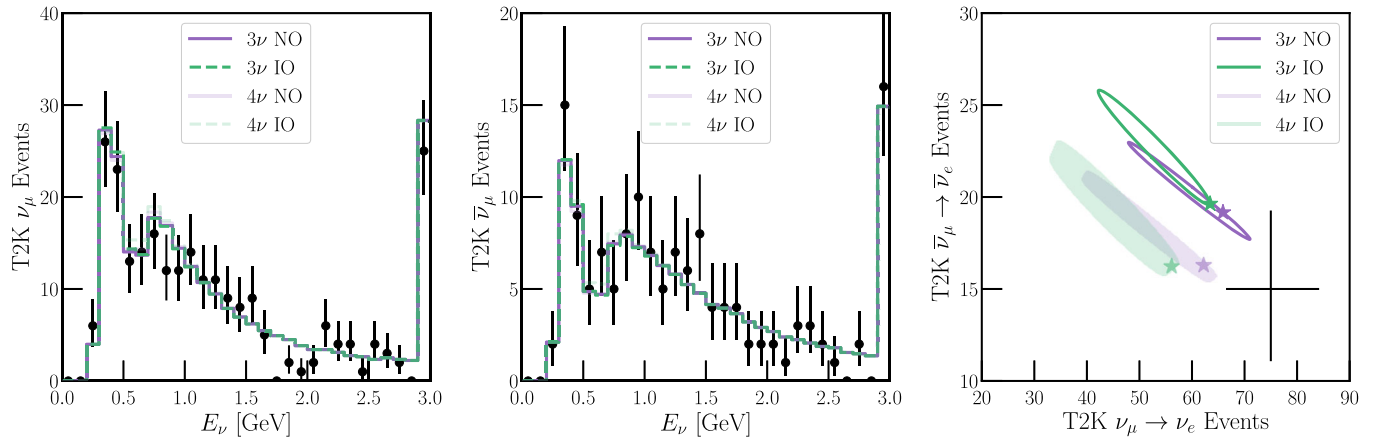
FIG. 5. Expected and observed event rates in T2K's ν_μ disappearance (left), $\bar{\nu}_\mu$ disappearance (center), and $\nu_e/\bar{\nu}_e$ appearance (right) channels. We compare the prediction under the 3ν (solid/dashed lines) and 4ν (faint lines/regions) hypotheses, with parameters from Table I, with the observed data (black). Purple curves correspond to the mostly active neutrinos following the normal mass ordering (NO), where green ones correspond to the inverted mass ordering (IO). In the right panel, the *CP*-violating phases are allowed to vary in the predicted rates. Data points are from Ref. [2].

all oscillation parameters are fixed according to Table I. In contrast, the right panel allows δ_{CP} to vary for the 3ν curves, and all three *CP*-violating phases to vary in the 4ν case. This allows for a set of ellipses in this bi-event parameter space instead of a single one. In the right panel, stars indicate the predicted event rates when the *CP*-violating phases are fixed to their values in Table I.

B. T2K

We simulate T2K in much the same spirit as NOvA, with the goal of matching the results presented in Ref. [2]. In the case of T2K, the disappearance channels each consist of 30 bins–100 MeV in width from 0 to 2.9 GeV, and one bin

corresponding to neutrino energies above 2.9 GeV. For the appearance channel, we take advantage of the expected neutrino-energy spectrum with bins of 125 MeV width from 0 to 1.25 GeV in each channel.[5] This yields 80 data points in our T2K analysis. Our T2K simulation corresponds to an exposure of 14.94×10^{20} (16.35×10^{20}) POT in (anti)neutrino mode operation.

Similar to Fig. 4, we show in Fig. 5 our expected event rates in the different T2K channels—the left panel is for ν_μ

---

[5]Refs. [23,24], however, have demonstrated that total-rate measurements of T2K's appearance channel result in similar parameter estimation to the collaboration's results.

TABLE II. Best-fit parameters of our analyses of T2K, NOvA, and a combined analysis of the two under the three-neutrino hypothesis. We determine the best-fit point under the normal (NO) and inverted (IO) mass-ordering hypotheses, as well as the overall preference for the NO over IO, $\Delta\chi^2_{\rm NO,IO}$, for each analysis. In each, a prior on $\sin^2(2\theta_{13})$ from Daya Bay is included, and $\sin^2\theta_{12} = 0.307$ and $\Delta m^2_{21} = 7.53 \times 10^{-5}$ eV^2 are fixed to their best-fit points from other experimental results.

3ν		$\sin^2\theta_{13}$	$\sin^2\theta_{23}$	$\Delta m^2_{31}/10^{-3}$ eV^2	$\delta_{\rm CP}$	χ^2	$\Delta\chi^2_{\rm NO,IO}$
T2K	NO	0.022	0.56	2.52	4.58	66.82	1.48
	IO	0.022	0.56	−2.41	4.71	68.19	
NOvA	NO	0.022	0.58	2.52	2.34	43.40	0.14
	IO	0.022	0.57	−2.41	4.78	43.55	
Joint	NO	0.022	0.57	2.51	3.67	115.58	−3.76
	IO	0.022	0.57	−2.41	4.72	111.82	

disappearance, center for $\bar{\nu}_\mu$ disappearance, and the right panel is the combined ν_e and $\bar{\nu}_e$ appearance. For clarity of display, we sum the total expected event rates in the ν_e and $\bar{\nu}_e$ channels in the right panel. Here, the oscillation parameters correspond to those given in Table I and, in the right panel, the *CP*-violating phases are allowed to vary.

C. Test statistic

We take the expected and observed event rates in NOvA (40 data points), T2K (80), or a combination of them (120) and construct a test statistic using Poisson statistics for the log-likelihood (matching a χ^2 function in the limit of large event rates):

$$\chi^2 = \sum_{i\in{\rm bins}} -2\left(-\lambda_i + x_i + x_i \log\left(\frac{\lambda_i}{x_i}\right)\right), \quad (3.1)$$

where λ_i (x_i) represents the expected (observed) event rate in bin i for a given experiment/channel.

We will be interested in several pieces of information from the test statistic in Eq. (3.1). When performing parameter estimations, we will use contours of $\Delta\chi^2$ about its minimum to represent preferred regions/intervals of parameter space. When comparing best-fit points under different hypotheses, i.e., comparing preference for the 4ν scenario over the 3ν one, we will compare the minimum χ^2 when varying over oscillation parameters, taking into account the number of degrees of freedom in such a fit.

D. Analysis and priors

The main focus of this work is on the long-baseline experiments NOvA and T2K, which are sensitive to oscillation effects associated with mass-squared differences of order of 10^{-3} eV^2. On the other hand, the solar mass-squared difference has been well-measured by solar neutrino [25,26] and reactor antineutrino [27] experiments to be $\Delta m^2_{21} = 7.53 \times 10^{-5}$ eV^2 while the associated mixing angle is measured to be $\sin^2\theta_{12} = 0.307$, both at the few percent level. Due to the lack of sensitivity to these

quantities at NOvA/T2K, we fix them[6] in our analyses. While NOvA and T2K are sensitive to $\sin^2\theta_{13}$ through their appearance channels, their measurement capability is significantly weaker than that of Daya Bay [11], RENO [12], and Double Chooz [13] reactor antineutrino experiments. In our fits, we include Daya Bay's measurement as a Gaussian prior on the quantity $4|U_{e3}|^2(1 - |U_{e3}|^2) = 0.0856 \pm 0.0029$, which is $\sin^2(2\theta_{13})$ when considering the three-neutrino hypothesis [11].

IV. RESULTS

This section details the results of our analyses. First, in Sec. IV A, we summarize the results of fits of our NOvA and T2K simulations and their combination under the three-neutrino hypothesis. Then, Sec. IV B discusses the results of these fits under the four-neutrino hypothesis, including a comparison of the three-neutrino and four-neutrino hypotheses.

A. Three-neutrino results

Our first three-neutrino analysis is focused on finding the best-fit points of each experimental analysis (T2K, NOvA, and a combined fit). For this, we perform two fits for each experiment/combination, one assuming that neutrinos follow the normal mass ordering (NO, $\Delta m^2_{31} > 0$) and one assuming that they follow the inverted one (IO, $\Delta m^2_{31} < 0$). Recent results have demonstrated that, under the three-neutrino hypothesis, T2K and NOvA each exhibit mild preference for the NO over the IO, but their combination has a mild preference for the IO [7–10]. When combined with all reactor antineutrino data and other experimental results, the global preference is for the NO at relatively low significance.

We find a result consistent with these previous results, summarized in Table II. As in all of our analyses, Δm^2_{21} and

---

[6]Specifically, we fix the matrix-element-squared $|U_{e2}|^2$, which is equal to $\sin^2\theta_{12} \cos^2\theta_{13} \cos^2\theta_{14}$ in the four-neutrino framework, to its best-fit value of 0.300. This causes $\sin^2\theta_{12}$ to vary for large θ_{14}.
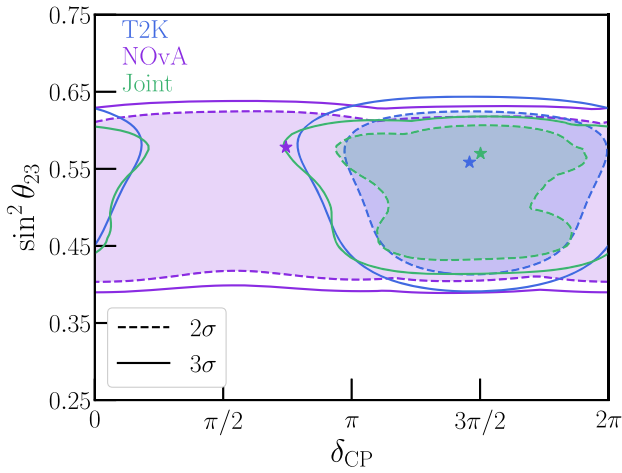
FIG. 6. Parameter estimation of δ_{CP} and $\sin^2\theta_{23}$ from T2K (blue), NOvA (purple), and their combination (green) at 2σ (dashed lines) and 3σ (solid lines) C.L.

$\sin^2\theta_{12}$ are fixed, and a prior is included from the results of Daya Bay on $\sin^2(2\theta_{13})$. We present both the overall test statistic at this best-fit point for each analysis as well as the preference for the NO over the IO in the right-most column (positive values indicate preference for NO, negative for IO). We note here that all of the best-fit χ^2 obtained are comparable to (and in the case of T2K and the joint fit, less than) the number of degrees of freedom, implying that these are all good fits to their respective datasets. Finally, we see that the joint-fit χ^2 under the NO hypothesis is around five units of χ^2 larger than the sum of the two individual fits whereas, under the IO hypothesis, it is roughly the same— this highlights the so-called NOvA/T2K tension, where the results disagree under the NO hypothesis but not under the IO one. The values from the "Joint" fit in Table II correspond to the benchmark values we adopted in the three-neutrino case in Table I.

We also perform a parameter estimation under the three-neutrino hypothesis, both to prepare our expectations for the four-neutrino analyses and to validate our results compared against the official results of the experimental collaborations. The free/fixed parameters and test statistic are identical to those when determining the best-fit points. For simplicity, we perform an analysis of the parameters $\sin^2\theta_{13}$, $\sin^2\theta_{23}$, Δm^2_{31}, and δ_{CP} and marginalize over $\sin^2\theta_{13}$ and Δm^2_{31} (including both the NO and IO hypotheses), and present the joint measurement of $\sin^2\theta_{23}$ and δ_{CP}.

Figure 6 presents the results of this analysis at 2σ (dashed, filled contours) and 3σ (solid lines) C.L. (confidence limit) for T2K (blue), NOvA (purple), and the joint fit (green). Stars of each color represent the best-fit points obtained in Table II. Once the mass ordering is marginalized, NOvA has no sensitivity to δ_{CP}, and constrains $\sin^2\theta_{23}$ to be between roughly 0.37 and 0.65 at 3σ C.L.

In the NO, NOvA can take on nearly any value of δ_{CP}, however it disfavors the combination $\delta_{CP} = 3\pi/2$, $\sin^2\theta_{23} > 1/2$ at relatively high significance. Under the IO, NOvA prefers this combination. Regardless of the mass ordering, T2K prefers $\delta_{CP} = 3\pi/2$ and constrains $\sin^2\theta_{23}$ to be in a similar range as NOvA. When the two are combined, the preferred regions are very similar to those obtained in the fit to T2K data alone.

B. Four-neutrino results

We begin our four-neutrino analyses by repeating the process that led to Table II—we determine the best-fit points under the four-neutrino hypothesis for T2K, NOvA, and their combination. Now that we are considering four-neutrino oscillations, we allow for all four mass orderings discussed in Sec. II (see Fig. 1). This amounts to dividing the analysis based on the signs of Δm^2_{31} and Δm^2_{4l}, where l represents m_1 in the NO and m_3 in the IO, the lightest of the mostly active neutrinos.

Table III summarizes these twelve analyses (four each for NOvA, T2K, and their Joint fit), giving the best-fit parameters as well as the overall χ^2 of each fit in the four-neutrino hypothesis. Near the bottom we give the preferred ordering of masses from each experiment/combination— T2K and the Joint fit both prefer $m_4 < m_3 < m_1 < m_2$, where NOvA prefers $m_1 < m_2 < m_3 < m_4$. The preference for the sign of Δm^2_{4l} is small in all cases—individual fit results for all four mass orderings and all three experimental combinations are provided for completeness in Appendix A. When allowing for a fourth neutrino, neither T2K nor NOvA have a strong preference for the sign of Δm^2_{31}. T2K prefers $\Delta m^2_{31} < 0$ at $\Delta\chi^2 = 0.1$, where NOvA prefers $\Delta m^2_{31} > 0$ at $\Delta\chi^2 = 0.02$. However, the combined fit prefers $\Delta m^2_{31} < 0$ at $\Delta\chi^2 = 4.6$ an even stronger preference for negative Δm^2_{31} than when data are analyzed under the three-neutrino hypothesis.

The bottom row of Table III presents the improvement in each experimental analysis (as well as the combined one) compared to the results of the three-neutrino analysis. We find that the fits to both the T2K[7] and NOvA data improve by roughly five units in χ^2, and the combined fit improves by nearly nine units. However, we note two very important caveats here:

(1) The results of the three-neutrino fit in Table II demonstrate that, relative to the number of degrees of freedom, good fits have been achieved. So, when comparing the three-neutrino fit—four parameters— to the four-neutrino one—ten parameters—one must take into account the fact that this minimization is being performed over an additional six parameters.

---

[7]This result is consistent with what the T2K collaboration reported in Ref. [28], which found an improvement of $\Delta\chi^2 = 4.7$.

TABLE III. Best-fit parameters of the four-neutrino analyses of T2K, NOvA, and their combination. We allow for all possible orderings of the neutrino mass eigenstates, hence Δm_{31}^2 and Δm_{4l}^2 can each be negative. In each analysis, a prior on $|U_{e3}|^2(1-|U_{e3}|^2)$ from Daya Bay is included, and $|U_{e2}|^2 = 0.300$ and $\Delta m_{21}^2 = 7.53 \times 10^{-5}$ eV^2 are fixed to their best-fit points from other experimental results.

4ν	T2K	NOvA	Joint
$\sin^2 \theta_{13}$	0.024	0.022	0.023
$\sin^2 \theta_{23}$	0.43	0.44	0.43
$\Delta m_{31}^2/10^{-3}$ eV^2	-2.39	2.43	-2.39
δ_{CP}	4.41	0.00	4.46
$\sin^2 \theta_{14}$	7.8×10^{-2}	6.9×10^{-3}	4.3×10^{-2}
$\sin^2 \theta_{24}$	4.1×10^{-2}	1.2×10^{-1}	6.0×10^{-2}
$\sin^2 \theta_{34}$	0.78	0.29	0.37
$\Delta m_{4l}^2/$eV^2	-8.5×10^{-3}	1.0×10^{-2}	-8.5×10^{-3}
δ_{14}	1.82	3.51	4.88
δ_{24}	2.64	3.15	5.89
$\chi_{4\nu}^2$	61.95	38.10	102.83
Ordering	$m_4 < m_3$ $< m_1 < m_2$	$m_1 < m_2$ $< m_3 < m_4$	$m_4 < m_3$ $< m_1 < m_2$
$\chi_{3\nu}^2 - \chi_{4\nu}^2$	4.87	5.30	8.99

(2) When determining the statistical significance, the comparison of $\chi_{3\nu}^2 - \chi_{4\nu}^2$ must be scrutinized to see whether these test statistics follow a χ^2 distribution. We have performed some basic Monte Carlo studies of our T2K and NOvA simulations (see Appendix C) and found that, when statistical fluctuations are considered, one will often find best-fit points with $\Delta m_{4l}^2 \approx 10^{-2}$ eV^2 that improve each experiment's fit by a couple of units of χ^2. This is likely driven by the sizes of the energy bins (around 100 MeV) used in the T2K and NOvA analyses—at T2K/NOvA baselines/energies, a new oscillation driven by a mass-squared splitting of 10^{-2} eV^2 will evolve significantly[8] over the span of a single bin. This new fast oscillation can "absorb" individual bins' statistical fluctuations and lead to an artificial improvement in the test statistic. This is validated by the results of Ref. [28], which found that an improvement of $\Delta \chi^2 = 4.7$ at T2K (between the three-neutrino and four-neutrino hypotheses) corresponds to only $\sim 1.0\sigma$ preference for a fourth neutrino, in contrast with the preference derived assuming Wilks' theorem [29] holds, $\sim 1.7\sigma$.

---

[8]For this Δm^2, the argument of the term $\sin^2(\Delta m^2 L/4E_\nu)$ that enters the oscillation probabilities changes by an appreciable fraction of π.

When considering the results of Table III (and that the best-fit points are close to $|\Delta m_{4l}^2| \approx 10^{-2}$ eV^2) in light of these two caveats, we find that, while a very light sterile neutrino improves the "tension" between T2K and NOvA, there is not strong evidence in favor of a four-neutrino hypothesis over the three-neutrino one.

In order to determine whether the sterile neutrino solution to the NOvA/T2K tension persists in light of caveat 2 above, we also perform an alternate analysis in Appendix B where we restrict $\Delta m_{21}^2 \lesssim |\Delta m_{4l}^2| < 10^{-3}$ eV^2. This allows us to avoid fast oscillations in the T2K/NOvA far detectors and any statistical pathologies that may arise. We find that there remains a preference for four neutrinos over three neutrinos at a level of $\Delta \chi^2 = 4.1$. While this is smaller than what we observed for $|\Delta m_{4l}^2| \approx 10^{-2}$ eV^2, it is nevertheless comparable to the preference for nonstandard interactions as a solution to this tension found in Refs. [16,19] at the level of $\Delta \chi^2 \approx 4.4 - 4.5$.

We generalize this best-fit procedure by, instead of minimizing over all parameters (including Δm_{4l}^2), scanning over Δm_{4l}^2 values. We again allow for both positive and negative values of this new mass-squared difference and for both the normal and inverted mass orderings for the three mostly active states. Figure 7 presents the results of this approach. The top panels (blue lines) show the results for T2K, middle panels (purple) for NOvA, and bottom panels (green) for the combined analysis. In each row, the left (right) panel corresponds to negative (positive) values of Δm_{4l}^2. Dark (light) lines in each case correspond to the NO (IO) among the mostly active neutrinos. Dashed lines in each panel indicate the best-fit χ^2 under the three-neutrino hypothesis presented in Table II. Stars indicate the overall best-fit point of each analysis (when considering all different mass orderings), and lines are made bold if they constitute the minimum χ^2 for a given experimental analysis for all of these choices of mass orderings.

The findings of Table III (and the corresponding tables in Appendix A) are borne out in Fig. 7, showing that the fits prefer $|\Delta m_{4l}^2| \sim 10^{-2}$ eV^2 in all cases, with moderate improvements relative to the three-neutrino fits. Above, we discussed the possibility that this preference has to do with the energy resolution and binning of the experiments and the statistical significance when interpreting confidence levels from $\Delta \chi^2$ may be overstated. If we restrict ourselves to $|\Delta m_{4l}^2| \lesssim 10^{-3}$ eV^2 to avoid this concern, we still find moderate preference for a fourth neutrino—see Appendix B for further discussion.

Moving on from best-fit determinations, we now construct constraints on the new parameters, specifically $\sin^2 \theta_{24}$ and Δm_{4l}^2 (the ones to which these experiments have the greatest sensitivity). In order to present constraints at a particular confidence level and compare against other literature results, we assume for this exercise
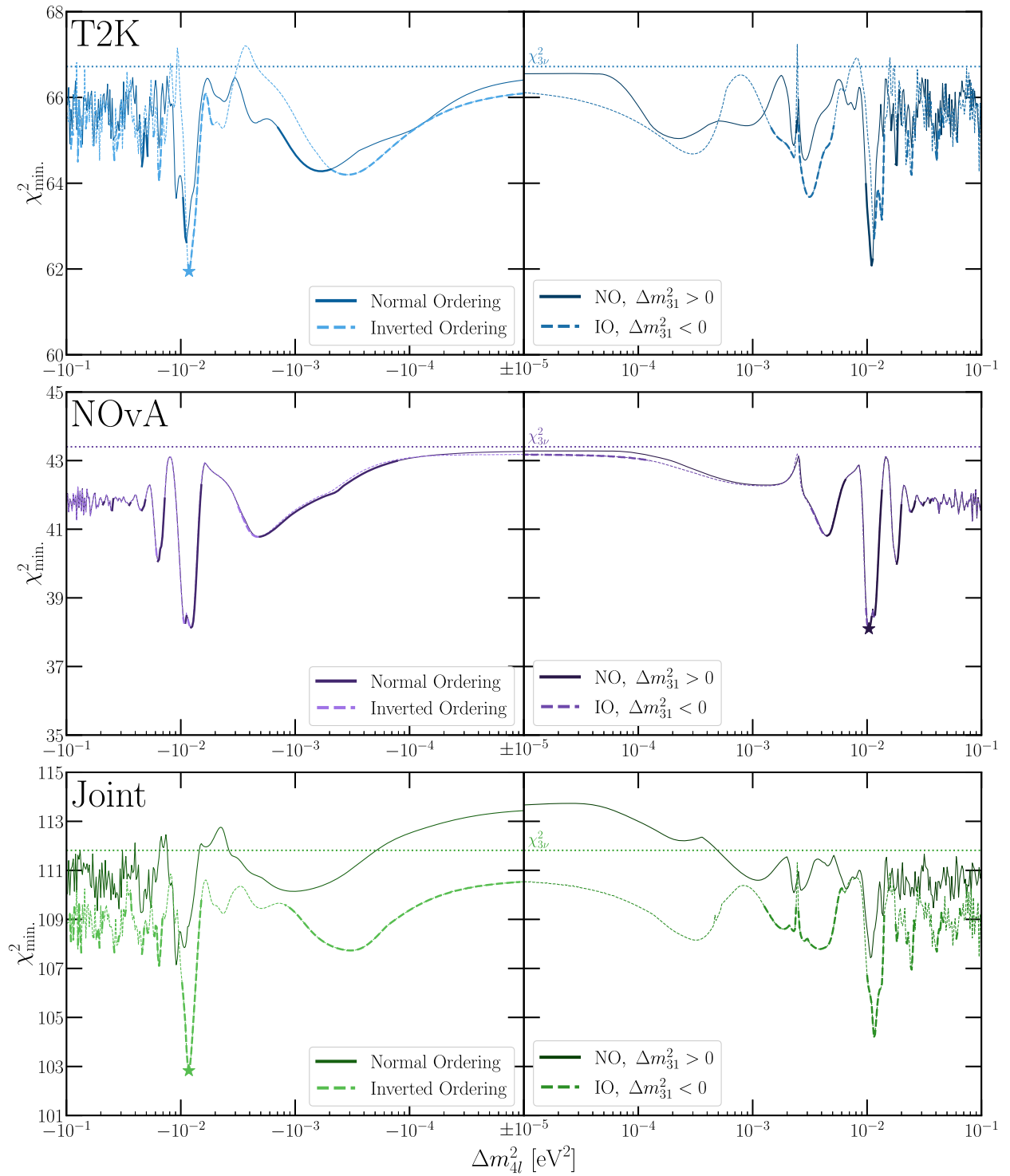
FIG. 7. Best-fit χ^2 obtained using our analysis of T2K (top, blue), NOvA (middle, purple), and a joint fit of the two (bottom, green) as a function of different values of Δm_{4l}^2. Different tones within each panel indicate different mass orderings (the signs of Δm_{31}^2 and Δm_{4l}^2). The minimization has been performed across all other oscillation parameters except for θ_{12} and Δm_{21}^2, which are fixed.

that Wilks' theorem holds [29]. After marginalizing over the remaining oscillation parameters (still fixing $|U_{e2}|^2$ and Δm_{21}^2), we present 2σ C.L. constraints from T2K (blue) and NOvA (purple) in Fig. 8. In generating these

constraints, we have marginalized over the signs of both Δm_{31}^2 and Δm_{4l}^2. Colored stars indicate the best-fit point in $(\sin^2\theta_{24}, |\Delta m_{4l}^2|)$ of the given fits. In Fig. 8 we also compare against the 90% C.L. constraint from the
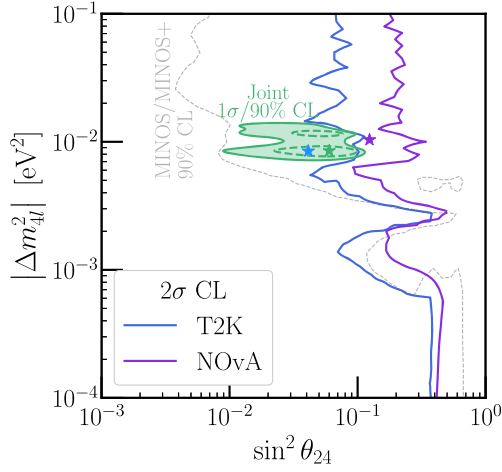
FIG. 8. Constraints on $\sin^2\theta_{24}$ versus Δm^2_{4l} at 2σ C.L. from T2K (blue) and NOvA (purple) after marginalizing over all other parameters (except for $|U_{e2}|^2$ and Δm^2_{21}, which are fixed and a prior from Daya Bay on $|U_{e3}|^2$—see text), including the signs of Δm^2_{31} and Δm^2_{4l}. The green region indicates the preferred region from a combined analysis at 1σ (dashed) and 90% (solid) C.L., and the gray, dashed line shows the 90% C.L. constraint from MINOS/MINOS+ [30]. All confidence levels presented here are derived assuming Wilks' theorem holds.

MINOS/MINOS+ experiment [30] as a faint gray line.[9] Finally, we also present in green the preferred region at $1\sigma/90\%$ C.L.[10] ($\Delta\chi^2 = 2.3$, 4.61 assuming Wilks' theorem for two parameters) by our combined T2K and NOvA analysis. This result is in tension with that of the MINOS/MINOS+ result, however, our preferred region has not been Feldman-Cousins corrected, and the results would likely agree if a higher confidence level were assumed. T2K has reported constraints in the $\sin^2\theta_{24}$ versus Δm^2_{41} parameter space in Ref. [28]—we find comparable results here despite the simplified assumptions we have made in our analysis and the slightly larger dataset considered in this work.

While Fig. 8 compares constraints and preferred regions in the parameter space $\sin^2\theta_{24}$ versus $|\Delta m^2_{4l}|$, it is also important to consider the parameters that have been marginalized in this construction. For concreteness, we focus on the preferred region (green) from the combined T2K/NOvA analysis that we have performed. The best-fit point, at $|\Delta m^2_{4l}| = 8.5 \times 10^{-3}$ eV², corresponds to mixing angles

$$\{\sin^2\theta_{14}, \sin^2\theta_{24}, \sin^2\theta_{34}\} = \{4.3\times10^{-2}, 6.0\times10^{-2}, 0.37\},$$

$$(4.1)$$

or mixing-matrix elements

$$\{|U_{e4}|^2, |U_{\mu4}|^2, |U_{\tau4}|^2\} = \{4.3\times10^{-2}, 5.7\times10^{-2}, 0.33\}.$$

$$(4.2)$$

For these low values of $|\Delta m^2_{4l}|$, the strongest constraints on $|U_{e4}|^2$ come from reactor antineutrino oscillation experiments such as Daya Bay [31] and Bugey-3 [32]. A combined analysis [33] constrains $\sin^2\theta_{14} \lesssim 4 \times 10^{-3}$ at 90% C.L., in significant tension with the value found in Eq. (4.1).

Constraints on $|U_{\tau4}|^2$ are more difficult to extract, as they often arise in tandem with $|U_{\mu4}|^2$ and depend strongly on Δm^2_{41} [34]. While specific constraints in this region of $|\Delta m^2_{4l}|$ have not been explicitly derived, $|U_{\tau4}|^2 = 0.33$ is possibly in tension with existing results from neutrino experiments. T2K, which analyzed its neutral-current data in addition to the datasets considered here, has constrained $|U_{\tau4}|^2 \lesssim 0.5$ for both $\Delta m^2_{41} = 3 \times 10^{-3}$ eV² and 0.1 eV² at 90% C.L. [28]. Atmospheric neutrino experiments, including Super-Kamiokande [35] and IceCube [36] have constrained $|U_{\tau4}|^2 \lesssim 0.2$ at high confidence, however, these analyses are restricted to $\Delta m^2_{41} \gtrsim 0.1$ eV² where the fourth-neutrino-driven oscillations are averaged out. A more thorough investigation of this 10^{-2} eV² regime would prove useful if this hint persists in future NOvA/T2K data.

Lastly, we also note that for such large $|U_{\alpha4}|^2$ and small m_4, this fourth neutrino would thermalize with the SM in the early universe and remain relativistic and in thermal equilibrium throughout much of the universe's evolution. Such thermalized light species are in tension with cosmological observations surrounding the cosmic microwave background and big bang nucleosynthesis—we refer the reader to Ref. [37] for further discussion on these effects. Regardless, if such a light sterile neutrino truly is behind any tension between NOvA and T2K, additional work is necessary to resolve tension between this terrestrial solution and cosmological observations.

When discussing Fig. 7, we considered the possibility of analyzing only the region $|\Delta m^2_{4l}| \lesssim 10^{-3}$ eV², in part to avoid concerns regarding energy resolution and bin widths. We noted that in that region, a solution to the NOvA/T2K tension persists with a preference of $\Delta\chi^2 \approx 4.1$. This regime has the added benefit that constraints from MINOS/MINOS+ (as seen in Fig. 8), Daya Bay/Bugey-3/others, and Super-Kamiokande/IceCube are considerably weaker. Such an *extremely light* sterile neutrino, as we discuss in Appendix B, with $|\Delta m^2_{4l}| \approx 7 \times 10^{-4}$ eV² should be paid particular attention as more data

---

[9]This result assumed Δm^2_{31} and Δm^2_{41} to both be positive, however, due to the lack of mass-ordering sensitivity at MINOS, the result likely does not depend strongly on this choice.

[10]We choose 90% C.L. for clarity (the 2σ C.L. region spans the entire range of $|\Delta m^2_{4l}|$ of the figure and a comparable region of $\sin^2\theta_{24}$) and for a direct comparison against the MINOS/MINOS+ result.

from T2K and NOvA are unveiled, especially if any tension between the two persists.

T2K and NOvA will continue collecting data—if a very light sterile neutrino does in fact exist with $|\Delta m_{4l}^2| \approx 10^{-2}$ eV^2, more data will continue to shed light and potentially lead to a discovery. In the next generation, the Deep Underground Neutrino Experiment (DUNE) [38] and Hyper-Kamiokande (HK) [39] experiments will have sensitivity to light sterile neutrinos in the same region of $|\Delta m_{4l}|^2$ given that they operate in a similar L/E_ν as NOvA and T2K. The two experiments, and any combined analysis, will have excellent sensitivity to test this solution to the T2K/NOvA tension [40,41].

V. CONCLUDING REMARKS

As more data from neutrino oscillation experiments are collected, we are able to test the standard-three-massive-neutrinos paradigm with better precision. Concurrently, there is always the possibility that disagreements arise, especially when data from multiple experiments are analyzed. In these instances, exploring different explanations of such tensions is invaluable, whether they are related to statistical fluctuations, deeper systematic issues, or new physics beyond the standard-three-massive-neutrinos paradigm.

Such a tension has been noted when comparing the latest data from the Tokai to Kamioka (T2K) and NuMI Off-axis ν_e Appearance (NOvA) experiments. These measure the (dis)appearance of ν_e (ν_μ) in a ν_μ beam at relatively long baselines. When analyzed under the three-neutrino hypothesis, their results disagree at around the 90% confidence level. Previous studies of combination T2K and NOvA data have highlighted that this tension is reduced when, for instance, the inverted neutrino mass ordering is considered instead of the normal ordering [7–10], or when additional, beyond-the-Standard-Model neutrino/matter interactions are included in the analyses [16,19].

We have demonstrated here that an alternative approach can remedy this tension—the addition of a fourth, very light, sterile neutrino. This very light new neutrino would be associated to a mass-squared difference, relative to the lightest mostly active neutrino, of order 10^{-2} eV^2. We have studied the four-neutrino hypothesis when applied to the T2K and NOvA data independently, as well as their combination. For the combined data, we find that the four-neutrino hypothesis is preferred over the three-neutrino one at the level of $\Delta\chi^2 \approx 9$. When interpreting this in terms of statistical significance, two difficulties arise: first, the number of additional parameters in the four-neutrino hypothesis relative to the three-neutrino one (six additional parameters). Second, the oscillations associated with a new mass-squared difference on the order of 10^{-2} eV^2 are significant within individual bins in these long-baseline experiments, which leads to an artificial

preference for sterile neutrinos due to statistical fluctuations.

Due to the second challenge, in order to avoid relatively fast oscillations, we also explored an alternative extremely light sterile neutrino analysis where the fourth neutrino is fixed to be associated to a mass-squared difference smaller (in magnitude) than 10^{-3} eV^2. In this context, we find moderate improvement relative to the three-neutrino hypothesis, at the level of $\Delta\chi^2 \approx 4$. While this is less significant, it is comparable to the improvement offered by nonstandard neutrino interactions and merits further investigation.

NOvA and T2K are still collecting and analyzing data. As they progress, the experiments and combined analyses thereof will allow for deeper testing of these different, interesting regimes of four-neutrino oscillations with a very light or extremely light fourth neutrino. If they confirm the existence of such a new, light fermion state, then future experiments (including the spiritual successors DUNE and Hyper-Kamiokande) will be able to probe the new particle's properties with even greater precision.

APPENDIX A: DETAILED FIT RESULTS

In Sec. IV, we provided best-fit points of our analyses of T2K, NOvA, and their combination under the three- and four-neutrino hypotheses. When discussing the best-fit points under the four-neutrino hypothesis (Table III), we showed the results of the analysis (i.e., which signs of Δm_{31}^2 and Δm_{4l}^2) that provided the best overall fit to each experimental dataset. In this appendix, we provide the results to all four fits for each experiment/combination. Table IV does so for our analyses of T2K and NOvA data separately, and Table V does so for their combination.

APPENDIX B: ALTERNATIVE ANALYSES WITH VERY SMALL MASS-SQUARED DIFFERENCE

We find, in Sec. IV, a solution to the NOvA/T2K tension with a new, light sterile neutrino with a mass-squared difference $|\Delta m_{4l}^2| \approx 10^{-2}$ eV^2. However, there are technical challenges associated with this relatively large mass-squared difference for the NOvA/T2K analyses, also as discussed in Sec. IV. For those reasons, we choose to pursue a different version of the analyses from the main

TABLE IV. Best-fit 4ν parameters of our four T2K (left) and NOvA (right) analyses. See Sec. IV B for more detail.

4ν	T2K			
	NO		IO	
	$\Delta m^2_{4l}>0$	<0	>0	<0
$\sin^2\theta_{13}$	0.024	0.024	0.024	0.024
$\sin^2\theta_{23}$	0.44	0.44	0.44	0.43
$\Delta m^2_{31}/10^{-3}\,\mathrm{eV}^2$	2.49	2.48	-2.38	-2.39
δ_{CP}	4.94	4.89	4.45	4.42
$\sin^2\theta_{14}$	7.1×10^{-2}	7.8×10^{-2}	8.0×10^{-2}	7.8×10^{-2}
$\sin^2\theta_{24}$	4.2×10^{-2}	4.0×10^{-2}	4.1×10^{-2}	4.1×10^{-2}
$\sin^2\theta_{34}$	5.2×10^{-2}	5.2×10^{-2}	5.6×10^{-1}	7.8×10^{-1}
$\Delta m^2_{4l}/\mathrm{eV}^2$	1.1×10^{-2}	-9.0×10^{-3}	1.1×10^{-2}	-8.5×10^{-3}
δ_{14}	3.51	3.14	2.08	1.83
δ_{24}	6.10	5.89	2.72	2.64
$\chi^2_{4\nu}$	62.07	62.63	62.80	61.95
Best-fit		$m_4<m_3<m_1<m_2$		
$\chi^2_{3\nu}-\chi^2_{4\nu}$		4.87		

	NOvA			
	NO		IO	
$\Delta m^2_{4l}>0$	<0	>0	<0	
0.022	0.022	0.022	0.022	
0.44	0.62	0.59	0.41	
2.43	2.44	-2.32	-2.35	
0.00	5.22	3.19	4.58	
6.9×10^{-3}	1.6×10^{-2}	8.9×10^{-3}	1.4×10^{-2}	
1.2×10^{-1}	1.2×10^{-1}	1.3×10^{-1}	1.1×10^{-1}	
0.29	0.79	0.34	0.69	
1.0×10^{-2}	-8.0×10^{-3}	1.0×10^{-2}	-8.1×10^{-3}	
3.51	4.07	4.81	4.69	
3.15	3.21	0.12	0.15	
38.10	38.14	38.13	38.16	
	$m_1<m_2<m_3<m_4$			
	5.30			

text, this time restricting ourselves to $|\Delta m^2_{4l}| \leq 10^{-3}\,\mathrm{eV}^2$. As with the analyses in the main text, we fix Δm^2_{21} to its best-fit value ($7.53 \times 10^{-5}\,\mathrm{eV}^2$).

First, we illustrate how the oscillation probabilities $P(\nu_\mu \to \nu_e)$ and $P(\bar{\nu}_\mu \to \bar{\nu}_e)$ at T2K/NOvA energies and baselines behave for a very light sterile neutrino, similar to the discussion in Sec. II (see Fig. 2). Instead of a relatively large $|\Delta m^2_{4l}| \approx 10^{-2}\,\mathrm{eV}^2$, Fig. 9 depicts the impact of a new mass-squared difference $\Delta m^2_{4l} = -3.4 \times 10^{-4}\,\mathrm{eV}^2$ (and an inverted mass ordering for the three mostly active neutrinos). The remaining oscillation parameters we use are

TABLE V. Best-fit 4ν parameters of our four combined T2K+NOvA analyses. See Sec. IV B for more detail.

4ν	Combined T2K and NOvA			
	NO		IO	
	$\Delta m^2_{4l}>0$	<0	>0	<0
$\sin^2\theta_{13}$	0.023	0.025	0.023	0.023
$\sin^2\theta_{23}$	0.45	0.45	0.44	0.43
$\Delta m^2_{31}/10^{-3}\,\mathrm{eV}^2$	2.49	2.51	-2.36	-2.39
δ_{CP}	4.09	3.88	1.72	4.47
$\sin^2\theta_{14}$	2.1×10^{-2}	1.1×10^{-1}	3.4×10^{-2}	4.3×10^{-2}
$\sin^2\theta_{24}$	5.3×10^{-2}	3.3×10^{-2}	5.3×10^{-2}	6.0×10^{-2}
$\sin^2\theta_{34}$	0.56	0.21	1.1×10^{-2}	0.37
$\Delta m^2_{4l}/\mathrm{eV}^2$	1.1×10^{-2}	-1.1×10^{-2}	1.2×10^{-2}	-8.5×10^{-3}
δ_{14}	0.01	0.03	6.09	4.88
δ_{24}	1.82	1.18	0.53	5.89
$\chi^2_{4\nu}$	107.41	107.62	104.27	102.83
Best-fit		$m_4<m_3<m_1<m_2$		
$\chi^2_{3\nu}-\chi^2_{4\nu}$		8.99		

from the "Joint" column in Table VI, corresponding to the best-fit parameters of the combined T2K and NOvA analysis when the new mass-squared difference is restricted to be $|\Delta m^2_{4l}| \leq 10^{-3}\,\mathrm{eV}^2$.

The top panels of Fig. 9 show oscillation probabilities at T2K, and the bottom panels at NOvA; the left (right) panels correspond to neutrino (antineutrino) oscillations. As with Fig. 2, we allow $\sin^2\theta_{34}$ to vary to demonstrate its nontrivial impact on these oscillation probabilities—the dashed/dot-dashed/dotted lines correspond to $\sin^2\theta_{34} = 0$, 0.4, 0.8, respectively. Compared with Fig. 2, here the "new" oscillation length driven by $\Delta m^2_{21} < |\Delta m^2_{4l}| < |\Delta m^2_{31}|$ is relatively long as a function of the neutrino energy, leading at zeroth order to an overall shift in normalization relative to the three-neutrino oscillation probabilities. Across the energies of interest for T2K and NOvA, this leads to larger values of $P(\nu_\mu \to \nu_e)$ and smaller values of $P(\bar{\nu}_\mu \to \bar{\nu}_e)$. As in Fig. 2, the impact of nonzero $\sin^2\theta_{34}$ is more prevalent for NOvA, with its longer baseline, than for T2K. Figure 10 depicts the impact of matter effects for this relatively smaller value of Δm^2_{4l} and is to be compared to Fig. 3.

The best-fit points obtained from this low-Δm^2_{4l} fit to T2K data, NOvA data, and the combined datasets are listed in Table VI. As in the result discussed in the main text, NOvA favors NO for the mostly active states while T2K and the Joint fits favor the IO for the mostly active states. All fits point to m_4 as the lightest neutrino mass. The improvement relative to the three-neutrinos scenario is largest for the Joint fit—a little over four units of χ^2—but rather modest. In summary, the data do not significantly
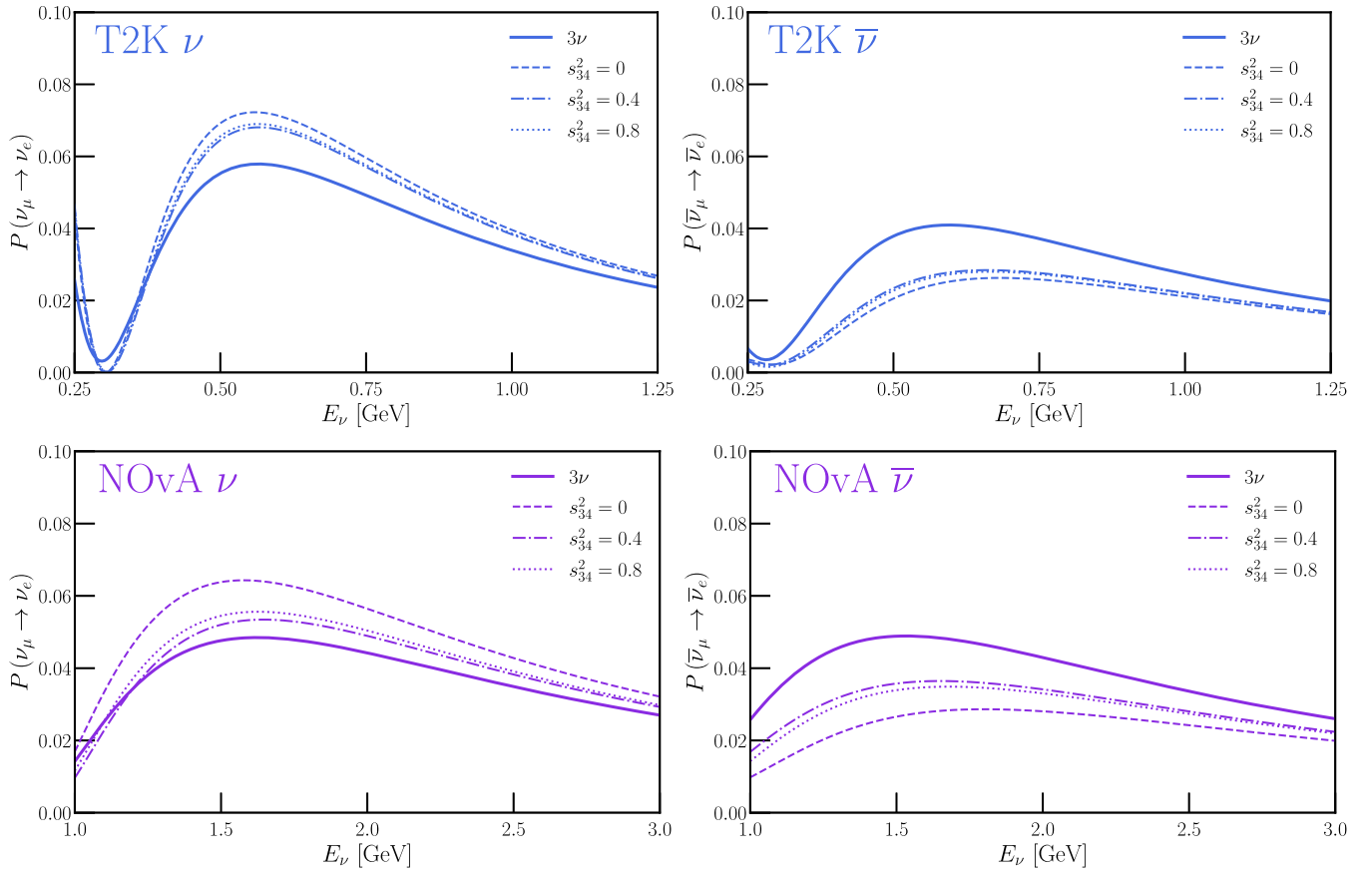
FIG. 9.   Oscillation probabilities at T2K (top) and NOvA (bottom) comparing three-neutrino oscillation probabilities (solid lines, parameters from Table I) against four-neutrino ones (nonsolid lines, parameters from the "Joint" column in Table VI). Left panels show probabilities for neutrino oscillation, whereas right ones show antineutrino oscillation. For the four-neutrino probabilities, three choices of $\sin^2\theta_{34}$ are used for demonstration: dashed/dot-dashed/dotted lines correspond to $\sin^2\theta_{34} = 0, 0.4, 0.8$.

TABLE VI.   Best-fit parameters of our 4ν analyses when restricted to $|\Delta m^2_{41}| \leq 10^{-3}$ eV^2. Other details identical to Table III.

4ν	T2K	NOvA	Joint
$\sin^2\theta_{13}$	0.025	0.022	0.026
$\sin^2\theta_{23}$	0.41	0.63	0.53
$\Delta m^2_{31}/10^{-3}$ eV^2	-2.37	2.44	-2.39
δ_{CP}	4.05	2.98	4.21
$\sin^2\theta_{14}$	0.13	6.2×10^{-3}	0.14
$\sin^2\theta_{24}$	8.2×10^{-2}	6.1×10^{-2}	7.6×10^{-2}
$\sin^2\theta_{34}$	0.63	0.79	0.48
$\Delta m^2_{41}/$eV^2	-3.5×10^{-4}	-1.0×10^{-3}	-3.4×10^{-4}
δ_{14}	4.66	2.77	5.34
δ_{24}	5.04	3.21	5.39
$\chi^2_{4\nu}$	64.20	41.50	5.39
Ordering	$m_4 < m_3 <$ $m_1 < m_2$	$m_4 < m_1 <$ $m_2 < m_3$	$m_4 < m_3 <$ $m_1 < m_2$
$\chi^2_{3\nu} - \chi^2_{4\mathrm{nu}}$	2.62	1.90	4.11

favor the four-neutrino hypothesis over the three-neutrino one.

Figure 11 depicts the region of the $|\Delta m^2_{41}| \times \sin^2\theta_{24}$ parameter space that is allowed by the combination of T2K and NOvA data at the one-sigma level, including all possible four-neutrino mass orderings (see Fig. 1) and assuming $|\Delta m^2_{41}|$ is less than 10^{-3} eV^2, along with the 2σ constraints from NOvA (purple) and T2K (blue). The stars indicate the best-fit points and the dashed line existing bounds from MINOS/MINOS+. Unlike the result discussed in the main text, here the best fit point is not in tension with existing neutrino oscillation bounds thanks to the more limited sensitivity of MINOS/MINOS+ and reactor antineutrino experiments to new mass-squared differences less than 10^{-3} eV^2.

Like the results discussed in the main text, here, the best-fit points in Table VI all prefer large values of $\sin^2\theta_{34}$, i.e., they suggest that ν_4 has an $\mathcal{O}(1)\nu_\tau$ component. As discussed in Sec. IV, while large $\sin^2\theta_{34}$ are excluded by existing data, relevant constraints were obtained only for relatively large $|\Delta m^2_{41}| \gtrsim 0.1$ eV^2.
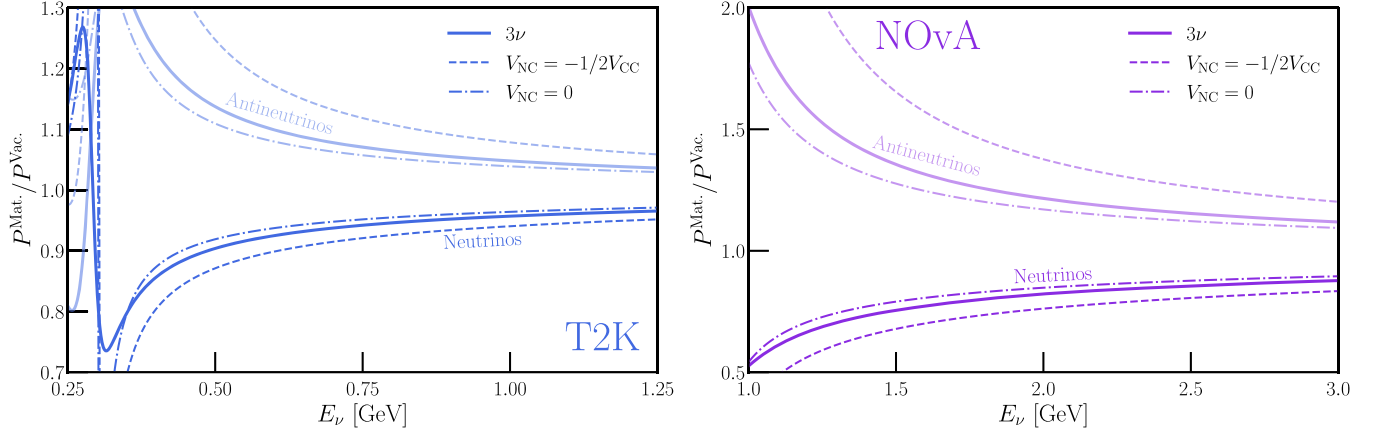
FIG. 10. Ratio of oscillation probabilities, similar to Fig. 3, considering an extremely light sterile neutrino with $\Delta m_{4l}^2 = -3.4 \times 10^{-4}$ eV^2 and oscillation parameters as given in Table VI.

APPENDIX C: TEST STATISTIC STUDIES AND PSEUDOEXPERIMENTS

Section IV demonstrated that all three fits, those to the T2K and NOvA data individually as well as their combination, prefer the four-neutrino hypothesis over the three-neutrino one to some degree of confidence. This is expected, as the three-neutrino hypothesis is a subset of the four-neutrino one—what is more difficult to predict is the level at which this preference is found. Specifically, we found that the best-fit-point to the data under the four-neutrino hypothesis compared to that of the three-neutrino hypothesis for T2K, NOvA, and the joint fit exhibited a preference at the level of $\Delta\chi^2 = 4.87$, 5.30, and 8.99, respectively. Also in Sec. IV, we discussed the fact that these three fits tend to favor $|\Delta m_{4l}^2| \approx 10^{-2}$ eV^2 and opined on whether this is a coincidence due to the binning used by T2K and NOvA or a real, physical effect.

In this appendix, we attempt to quantify some of these observed challenges—how significant these preferences are,
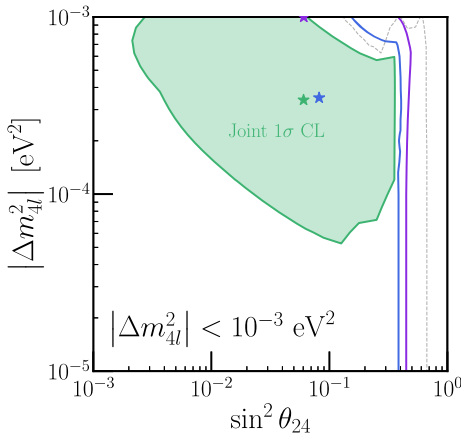

FIG. 11. Similar to Fig. 8 but under the analysis assumption that $|\Delta m_{4l}^2| < 10^{-3}$ eV^2.

and whether the preferred new mass-squared splitting is spurious. To do so, we perform a number of pseudoexperiments corresponding to each analysis. We simulate data for each experiment assuming the three-neutrino hypothesis is true, assuming $\sin^2\theta_{12} = 0.304$, $\sin^2\theta_{13} = 0.0212$, $\sin^2\theta_{23} = 0.532$, $\Delta m_{21}^2 = 7.53 \times 10^{-5}$ eV^2, $\Delta m_{31}^2 = 2.45 \times 10^{-3}$ eV^2, and $\delta_{CP} = 4.39$ (given as reference values in Ref. [2]). For each pseudoexperiment, we include Poissonian fluctuations on the expected data according to this hypothesis. Then, using the same analysis strategies as in the main text, we obtain the best-fit points and χ^2 values for the three-neutrino and four-neutrino hypotheses.

The normalized distribution of $\Delta\chi^2 \equiv \chi_{3\nu}^2 - \chi_{4\nu}^2$ is shown in Fig. 12. We show the histograms obtained by performing pseudoexperiments of the three different analyses in solid, colored lines, compared against the $\Delta\chi^2$ obtained when analyzing the data as vertical, dashed lines. We also display the χ^2 distribution assuming six degrees of freedom (corresponding to the difference between the number of parameters in the two analysis hypotheses) as a gray line, which seems to track the distribution of the joint-fit pseudoexperiments well. As a result of this procedure, we can determine the statistical significances of the three preferences—the p-values of the observed data at T2K, NOvA, and their combination are 0.53, 0.21, and 0.22, respectively. These values correspond to preference for the four-neutrino hypothesis at the level of 0.58σ, 1.26σ, and 1.22σ—none of which corresponds to a significant preference.

Finally, we determine whether the best-fit points obtained when analyzing data, all with $|\Delta m_{4l}^2| \approx 10^{-2}$ eV^2 are expected when including Poissonian fluctuations of simulated three-neutrino data. We determine, for each pseudoexperiment, the best-fit values of $\sin^2\theta_{24}$ and $|\Delta m_{4l}^2|$ obtained when analyzing the pseudodata under the four-neutrino hypothesis, displaying the distributions of these best-fit values in Fig. 13. Here, the dark regions indicate where the fits prefer the combination of parameters
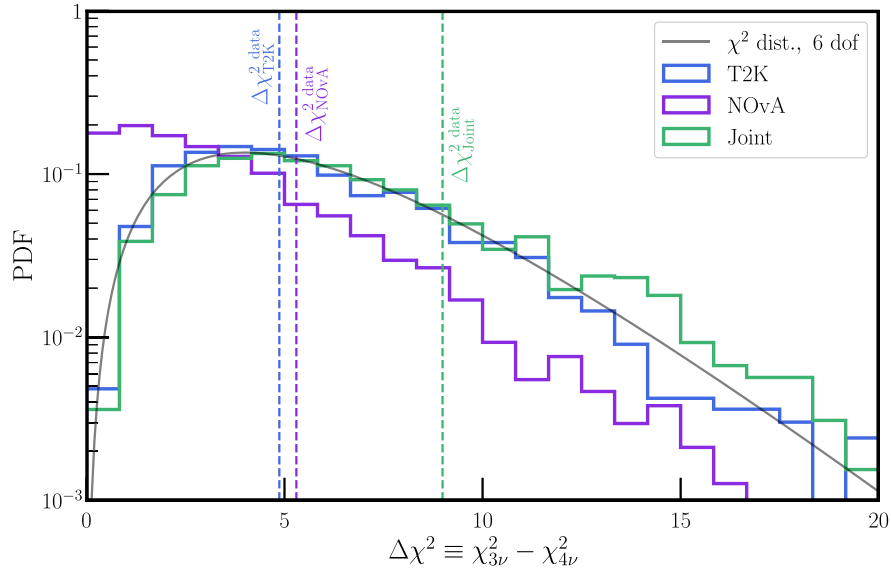
FIG. 12. Preference for the four-neutrino hypothesis over the three-neutrino one as indicated by pseudoexperiments simulating T2K (blue), NOvA (purple), and their combination (green). We also display the PDF of the chi-squared distribution assuming six degrees of freedom (gray), as well as the preferences indicated when analyzing the actual datasets (dashed lines).
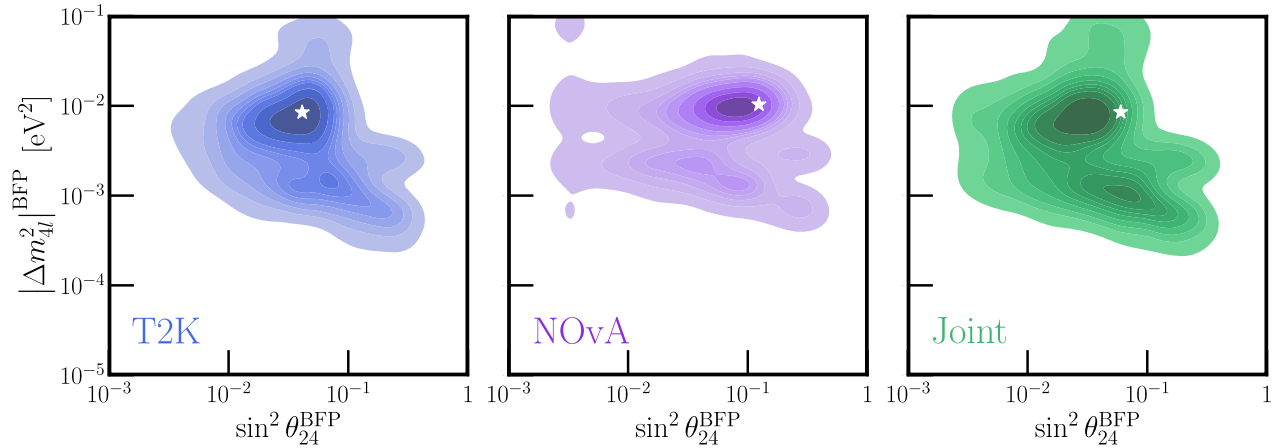

FIG. 13. Best-fit values of $\sin^2 \theta_{24}$ and $|\Delta m_{4l}^2|$ obtained when performing pseudoexperiments of T2K (left), NOvA (center), and their combination (right). White stars in each panel indicate the best-fit values of these parameters when analyzing the corresponding dataset. In the colored distributions, darker (lighter) colors indicate regions where the fit prefers the values more (less) frequently.

most frequently, and the white stars show the best-fit parameters obtained in each analysis from Sec. IV. For all three analyses, the best-fit obtained when analyzing the data is nearly exactly consistent with the most likely points obtained by these procedures. This indicates that such fit values of $|\Delta m_{4l}^2|$ are to be expected due to the construction of the test statistic and the experimental particulars, furthering the evidence that the results obtained in the main text are due to statistical fluctuations instead of the actual presence of a fourth, very light neutrino.

[1] K. Abe *et al.* (T2K Collaboration), Indication of Electron Neutrino Appearance from an Accelerator-Produced Off-Axis Muon Neutrino Beam, Phys. Rev. Lett. **107,** 041801 (2011).

[2] K. Abe *et al.* (T2K Collaboration), Improved constraints on neutrino mixing from the T2K experiment with 3.13×10^{21} protons on target, Phys. Rev. D **103,** 112008 (2021).

[3] P. Adamson *et al.* (NOvA Collaboration), First Measurement of Electron Neutrino Appearance in NOvA, Phys. Rev. Lett. **116,** 151806 (2016).

[4] M. A. Acero *et al.* (NOvA Collaboration), An improved measurement of neutrino oscillation parameters by the NOvA experiment, Phys. Rev. D **106,** 032004 (2022).

[5] P. Dunne, Latest neutrino oscillation results from T2K, 10.5281/zenodo.3959558(2020).

[6] A. Himmel, New oscillation results from the NOvA experiment, 10.5281/zenodo.3959581(2020).

[7] K. J. Kelly, P. A. N. Machado, S. J. Parke, Y. F. Perez-Gonzalez, and R. Z. Funchal, Neutrino mass ordering in light of recent data, Phys. Rev. D **103,** 013004 (2021).

[8] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, T. Schwetz, and A. Zhou, The fate of hints: Updated global analysis of three-flavor neutrino oscillations, J. High Energy Phys. 09 (2020) 178.

[9] P. F. de Salas, D. V. Forero, S. Gariazzo, P. Martínez-Miravé, O. Mena, C. A. Ternes, M. Tórtola, and J. W. F. Valle, 2020 global reassessment of the neutrino oscillation picture, J. High Energy Phys. 02 (2021) 071.

[10] F. Capozzi, E. Di Valentino, E. Lisi, A. Marrone, A. Melchiorri, and A. Palazzo, Unfinished fabric of the three neutrino paradigm, Phys. Rev. D **104,** 083031 (2021).

[11] D. Adey *et al.* (Daya Bay Collaboration), Measurement of the Electron Antineutrino Oscillation with 1958 Days of Operation at Daya Bay, Phys. Rev. Lett. **121,** 241805 (2018).

[12] G. Bak *et al.* (RENO Collaboration), Measurement of Reactor Antineutrino Oscillation Amplitude and Frequency at RENO, Phys. Rev. Lett. **121,** 201801 (2018).

[13] H. de Kerret *et al.* (Double Chooz Collaboration), Double Chooz θ_{13} measurement via total neutron capture detection, Nat. Phys. **16,** 558 (2020).

[14] R. Jimenez, C. Pena-Garay, K. Short, F. Simpson, and L. Verde, Neutrino masses and mass hierarchy: Evidence for the normal hierarchy, arXiv:2203.14247.

[15] S. Gariazzo *et al.*, Neutrino mass and mass ordering: No conclusive evidence for normal ordering, arXiv:2205.02195.

[16] P. B. Denton, J. Gehrlein, and R. Pestes, *CP*-Violating Neutrino Nonstandard Interactions in Long-Baseline-Accelerator Data, Phys. Rev. Lett. **126,** 051801 (2021).

[17] L. S. Miranda, P. Pasquini, U. Rahaman, and S. Razzaque, Searching for non-unitary neutrino oscillations in the present T2K and NOvA data, Eur. Phys. J. C **81,** 444 (2021).

[18] S. S. Chatterjee and A. Palazzo, Interpretation of NOvA and T2K data in the presence of a light sterile neutrino, arXiv: 2005.10338.

[19] S. S. Chatterjee and A. Palazzo, Nonstandard Neutrino Interactions as a Solution to the *NOvA* and T2K Discrepancy, Phys. Rev. Lett. **126,** 051802 (2021).

[20] D. V. Forero, C. Giunti, C. A. Ternes, and M. Tortola, Nonunitary neutrino mixing in short and long-baseline experiments, Phys. Rev. D **104,** 075030 (2021).

[21] U. Rahaman, Looking for Lorentz invariance violation (LIV) in the latest long baseline accelerator neutrino oscillation data, Eur. Phys. J. C **81,** 792 (2021).

[22] U. Rahaman, S. Razzaque, and S. U. Sankar, A review of the tension between the T2K and NOvA appearance data and hints to new physics, Universe **8,** 109 (2022).

[23] S. A. R. Ellis, K. J. Kelly, and S. W. Li, Leptonic unitarity triangles, Phys. Rev. D **102,** 115027 (2020).

[24] S. A. R. Ellis, K. J. Kelly, and S. W. Li, Current and future neutrino oscillation constraints on leptonic unitarity, J. High Energy Phys. 12 (2020) 068.

[25] K. Abe *et al.* (Super-Kamiokande Collaboration), Solar neutrino measurements in Super-Kamiokande-IV, Phys. Rev. D **94,** 052010 (2016).

[26] Y. Nakajima, Recent results and future prospects from Super-Kamiokande, 10.5281/zenodo.3959640 (2020).

[27] A. Gando *et al.* (KamLAND Collaboration), Reactor on-off antineutrino measurement with KamLAND, Phys. Rev. D **88,** 033001 (2013).

[28] K. Abe *et al.* (T2K Collaboration), Search for light sterile neutrinos with the T2K far detector Super-Kamiokande at a baseline of 295 km, Phys. Rev. D **99,** 071103 (2019).

[29] S. S. Wilks, The large-sample distribution of the likelihood ratio for testing composite hypotheses, Ann. Math. Stat. **9,** 60 (1938).

[30] P. Adamson *et al.* (MINOS+ Collaboration), Search for Sterile Neutrinos in MINOS and MINOS+ Using a Two-Detector Fit, Phys. Rev. Lett. **122,** 091803 (2019).

[31] F. P. An *et al.* (Daya Bay Collaboration), Improved Search for a Light Sterile Neutrino with the Full Configuration of the Daya Bay Experiment, Phys. Rev. Lett. **117,** 151802 (2016).

[32] Y. Declais *et al.*, Search for neutrino oscillations at 15-meters, 40-meters, and 95-meters from a nuclear power reactor at Bugey, Nucl. Phys. **B434,** 503 (1995).

[33] P. Adamson *et al.* (MINOS+, Daya Bay Collaborations), Improved Constraints on Sterile Neutrino Mixing from Disappearance Searches in the MINOS, MINOS+, Daya Bay, and Bugey-3 Experiments, Phys. Rev. Lett. **125,** 071801 (2020).

[34] M. Dentler, A. Hernández-Cabezudo, J. Kopp, P. A. N. Machado, M. Maltoni, I. Martinez-Soler, and T. Schwetz, Updated global analysis of neutrino oscillations in the presence of eV-scale sterile neutrinos, J. High Energy Phys. 08 (2018) 010.

[35] K. Abe *et al.* (Super-Kamiokande Collaboration), Limits on sterile neutrino mixing using atmospheric neutrinos in Super-Kamiokande, Phys. Rev. D **91,** 052019 (2015).

[36] M. G. Aartsen *et al.* (IceCube Collaboration), Search for sterile neutrino mixing using three years of IceCube Deep-Core data, Phys. Rev. D **95,** 112002 (2017).

[37] S. Gariazzo, P. F. de Salas, and S. Pastor, Thermalisation of sterile neutrinos in the early Universe in the $3 + 1$ scheme with full mixing matrix, J. Cosmol. Astropart. Phys. 07 (2019) 014.

[38] B. Abi *et al.* (DUNE Collaboration), Deep Underground Neutrino Experiment (DUNE), far detector technical design report, volume II: DUNE physics, arXiv:2002.03005.

[39] K. Abe *et al.* (Hyper-Kamiokande Collaboration), Hyper-Kamiokande design report, arXiv:1805.04163.

[40] J. M. Berryman, A. de Gouvêa, K. J. Kelly, and A. Kobach, Sterile neutrino at the Deep Underground Neutrino Experiment, Phys. Rev. D **92**, 073012 (2015).

[41] K. J. Kelly, Searches for new physics at the Hyper-Kamiokande experiment, Phys. Rev. D **95**, 115009 (2017).