

# Atmospheric neutrino oscillations: The full picture

Philipp Eller<sup>\*</sup>

Physics Department, *Technical University of Munich*, James-Frank-Str. 1,  
85748 Garching bei München, Germany



(Received 8 June 2026; accepted 21 August 2026; published 9 September 2026)

We present the first combined oscillation analysis of recent atmospheric neutrino datasets, featuring data from Super-Kamiokande, IceCube-DeepCore, and KM3NeT/ORCA together with reactor data from Daya Bay. Such combinations have long been considered infeasible outside experimental collaborations; we demonstrate that a unified physics model can simultaneously describe all datasets with no significant parameter tensions. Fitting 839 048 events across 1536 bins with 91 parameters, our combined analysis yields competitive measurements of the neutrino mixing parameters, and prefers the normal over the inverted mass ordering at  $3\sigma$  significance.

DOI: [10.1103/v4hl-pbpn](https://doi.org/10.1103/v4hl-pbpn)

## I. INTRODUCTION

Atmospheric neutrinos have played a pivotal role in the study of neutrino oscillations from early on. The Super-K collaboration famously reported  $5\sigma$  evidence against the no-oscillation hypothesis in 1998 [1]. The oscillation interpretation was then confirmed in a complementary channel by the SNO collaboration using solar neutrinos in 2002 [2].

Since then, a multitude of experiments have measured the three oscillation angles and the (absolute) values of the mass splitting to high precision. All in all, good agreement with the three-flavor oscillation picture is observed, with a few notable exceptions known as “anomalies.”

What remains still unknown, however, are the presence of  $CP$ -violation and the ordering of the neutrino masses (NMO). From solar oscillations, it was established that the second mass eigenstate is heavier than the first ( $m_1 < m_2$ ) [2,3], but it is not yet clear whether the third mass eigenstate is the heaviest [ $m_1 < m_2 < m_3$ ; normal ordering (NO)], or the lightest [ $m_3 < m_1 < m_2$ ; inverted ordering (IO)].

These two questions have been studied using accelerator neutrinos, most recently by analyses from the T2K collaboration [4], favoring a value of  $\delta_{CP}$  of around  $3/2\pi$  and a Bayes factor of 1.7 preferring the NO over the IO (or a  $\Delta\chi^2_{IO-NO} \approx 2.1$ ), and the NO $\nu$ A collaboration [5] also reported a similar  $1.4\sigma$  preference of the NO over the

IO, however, their data prefer a  $\delta_{CP}$  value of  $0.87\pi$  assuming NO.

It was realized, as for example discussed in [6–8], that this tension in the  $\delta_{CP}$  phase is problematic when combining the two datasets. In fact, since the tension in  $\delta_{CP}$  is not present under IO, a combined fit does prefer IO over NO (Bayes factor 1.3), even though both individual fits prefer NO over IO. This tension between the two datasets was also confirmed by an official result in a joint analysis of both collaborations [9].

Another experiment with sensitivity to the NMO that has already started taking data is JUNO, which recently reported its first oscillation results [10]. However, only solar parameters were constrained, and more data is needed to make statements on the NMO.

Independent sensitivity to the NMO comes from matter effects on Earth-crossing atmospheric neutrinos. Super-K reports a preference of NO vs IO with a  $\Delta\chi^2_{IO-NO} = 5.69$  [11]. An effort combining Super-K with T2K reports similar findings [12]. IceCube has not made any official statement about the NMO since 2019, when they reported a mild preference of NO over IO at  $\Delta\chi^2_{IO-NO} = 0.738$  using 3 years of data [13]. A recent PhD thesis, however, reports  $\Delta\chi^2_{IO-NO} = 4.398$  using 9.3 years of IceCube data [14]. KM3NeT observes a similarly mild preference of NO over IO in their first oscillation analysis of  $\Delta\chi^2_{IO-NO} = 0.62$  [15].

The question is: can these neutrino experiments be combined into a consistent picture to measure oscillation parameters and make a combined statement on the NMO?

Such combinations of public datasets have been the tale of multiple “global fitter” groups for many years, producing very successful global results on oscillation parameters and beyond the Standard Model physics. The sore spot in these combinations is the lack of atmospheric data included.

<sup>\*</sup>Contact author: philipp.eller@tum.de

Some groups had expressed inability to reproduce the Super-K results, as discussed in [16], with the last inclusion of Super-K in a combined fit in 2014 [17]. And only after that point in time have IceCube data been introduced into the analysis in 2018 [18], by which time Super-K had already been removed. In Ref. [19], the authors did succeed in including data from DeepCore and ANTARES in their global fit, but similar problems with Super-K were reported. Other groups have reported general difficulties, for example, in [8] “For atmospheric neutrinos, currently involving hundreds of bins, dozens of systematic uncertainties, and refined statistical separation of event classes by flavor proxies, the construction of  $\chi^2$  maps based on public information has become eventually unfeasible outside the experimental collaborations.” We understood this statement as a challenge rather than a warning.

At the time of writing, we did not find any combined analysis featuring more than one recent atmospheric dataset, and hence we present here the first of its kind, combining the data of Super-K, IceCube and KM3NeT, together with the Daya Bay reactor data, in a fully joint fit at the data level, including a careful modeling of the underlying physics with 91 fit parameters, fitting 839 048 events distributed across 1536 bins simultaneously. We find that a combined analysis can simultaneously describe all data well, and no large parameter tensions were observed. We report our measured values and uncertainties on the mixing parameters  $\delta_{CP} = 3.78^{+0.89}_{-0.884}$ ,  $\theta_{13} = 0.149^{+0.00281}_{-0.00274}$ ,  $\theta_{23} = 0.785^{+0.0318}_{-0.0407}$ , and  $\Delta m_{31}^2 = 2.51^{+0.0463}_{-0.0441} \times 10^{-3} \text{ eV}^2$ , and find a preference of NO vs IO at  $\chi^2_{\text{IO-NO}} = 9.11$ , corresponding to  $3\sigma$ . We also disfavor a  $\delta_{CP}$  value of zero at  $\Delta\chi^2 = 8.06$ . Figure 1 and Table I give an overview of these main results.

## II. DATASETS

We introduce the publicly available datasets used for our analysis in the following, including a discussion of detector-specific uncertainties. The physics models (atmospheric neutrino flux, oscillation probabilities, and interaction cross sections) and analysis methods will be discussed later in Sec. III. In Fig. 6, the standard 90% CL 2D contours in the atmospheric mixing parameters plane  $\Delta m_{31}^2 - \sin^2(\theta_{23})$  can be found, comparing the values obtained from our implementations with the official reference curves. A similar validation plot for our Daya Bay implementation can be found in [20].

### A. Super-K

We use Super-K phases I–V, which comprise 6511 days (484 kton-years) of data. The official oscillation analysis of the sample was presented in [11], and more details can also be found in [21]. The analysis constrains the atmospheric mixing parameters  $\Delta m_{31}^2$  and  $\theta_{23}$ , the leptonic  $CP$  violating phase  $\delta_{CP}$ , and the reactor angle  $\theta_{13}$ . The analysis reported a preference for the normal ordering at the  $\approx 2\sigma$  level.

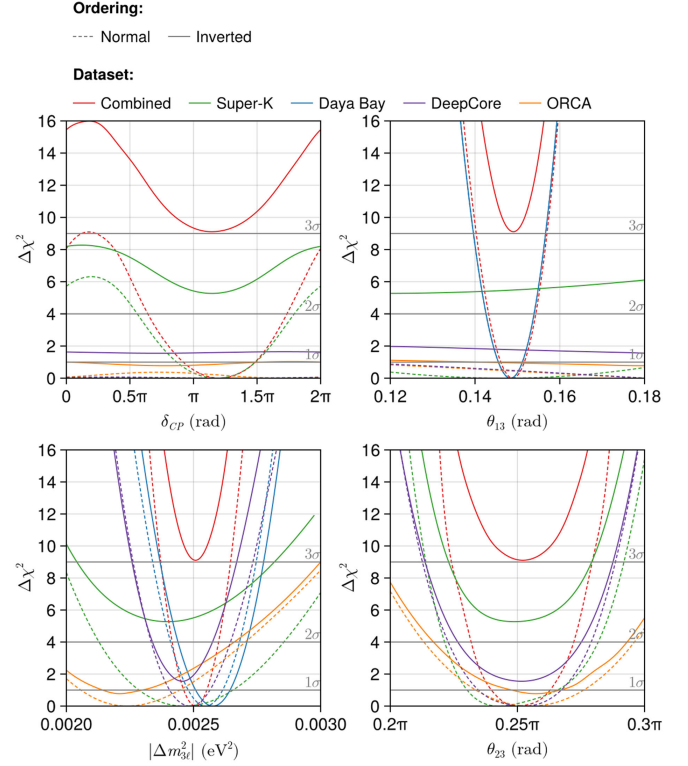


FIG. 1.  $\chi^2$  profiles for the atmospheric mixing parameters  $|\Delta m_{3\ell}^2|$  and  $\theta_{23}$ , the leptonic  $CP$  violating phase  $\delta_{CP}$ , and the reactor angle  $\theta_{13}$ . Shown are separately the NO and IO profiles for the five analyses. Note that  $\Delta m_{3\ell}^2$  corresponds to  $\Delta m_{31}^2$  under NO and  $\Delta m_{32}^2$  under IO.

The Super-K collaboration provides a public dataset [22] accompanying the publication, which we use here. The dataset provides expected and observed counts in all 930 analysis bins (expected is split up into  $\nu_e$ ,  $\bar{\nu}_e$ ,  $\nu_\mu$ ,  $\bar{\nu}_\mu$ ,  $\nu_\tau + \bar{\nu}_\tau$  charged current (CC) channels, and one flavor agnostic neutral current (NC) channel) for the best-fit NO and IO oscillation parameters, but without systematics applied. The lack of systematic detector uncertainties is a limitation of the data release, and we approximate these by 38 separate fit parameters that handle bin migrations, selection efficiencies, normalizations, and energy scale uncertainties. The data release does not provide per-phase information, which means we cannot introduce per-phase uncertainties, as is done in the official analysis. Another limitation of the dataset is that, for each analysis bin and channel, only summary statistics (mean, variance, and the [2.3, 5.9, 50, 84.1, 97.7]% quantiles) for the bin’s underlying true energy ( $E_\nu$ ) and true zenith ( $\theta$ ) distributions are provided. From this, we reconstruct a two-dimensional probability distribution over the  $(E_\nu, \cos\theta)$  space by fitting a mixture of two log-normal distributions over the energy, and a von Mises-Fisher distribution over  $\cos\theta$ . These distributions model the detector resolution, which effectively acts as a smearing matrix from true to reconstructed variables. To capture the energy-dependent

TABLE I. Best-fit points and 68% CL error bars for the mixing parameters (NO), reported for each of the five analyses separately.

| Parameter                                       | Dataset         | Best fit     | +1 $\sigma$    | -1 $\sigma$    |
|-------------------------------------------------|-----------------|--------------|----------------|----------------|
| $\delta_{CP}$ (rad)                             | Daya Bay        | ...          | ...            | ...            |
|                                                 | Super-K         | 3.72         | 0.989          | 0.961          |
|                                                 | DeepCore        | ...          | ...            | ...            |
|                                                 | ORCA            | ...          | ...            | ...            |
|                                                 | <b>Combined</b> | <b>3.78</b>  | <b>0.89</b>    | <b>0.884</b>   |
| $\theta_{13}$ (rad)                             | Daya Bay        | 0.148        | 0.00286        | 0.00281        |
|                                                 | Super-K         | 0.147        | 0.046          | 0.0411         |
|                                                 | DeepCore        | ...          | ...            | ...            |
|                                                 | ORCA            | ...          | ...            | ...            |
|                                                 | <b>Combined</b> | <b>0.149</b> | <b>0.00281</b> | <b>0.00274</b> |
| $\Delta m_{31}^2$ ( $10^{-3}$ eV <sup>2</sup> ) | Daya Bay        | 2.55         | 0.0682         | 0.0672         |
|                                                 | Super-K         | 2.46         | 0.178          | 0.192          |
|                                                 | DeepCore        | 2.49         | 0.0803         | 0.0782         |
|                                                 | ORCA            | 2.24         | 0.182          | 0.207          |
|                                                 | <b>Combined</b> | <b>2.51</b>  | <b>0.0463</b>  | <b>0.0441</b>  |
| $\theta_{23}$ (rad)                             | Daya Bay        | ...          | ...            | ...            |
|                                                 | Super-K         | 0.762        | 0.0388         | 0.0728         |
|                                                 | DeepCore        | 0.793        | 0.056          | 0.0463         |
|                                                 | ORCA            | 0.801        | 0.0672         | 0.0661         |
|                                                 | <b>Combined</b> | <b>0.785</b> | <b>0.0318</b>  | <b>0.0407</b>  |

correlation in the angular reconstruction resolution, we include an exponential energy dependence of the von Mises-Fisher  $\kappa$  parameter, as motivated in [23], as an additional per-bin fit parameter in the construction of the smearing matrices.

Since the released data provides separate expectations for NO and IO, this lets us test our reweighting procedure at nominal (i.e., without systematics applied) to check the smearing matrices. We get very good agreement for almost all bins (an average per-bin  $\Delta\chi^2$  of 0.0012), except for some very low momentum bins: the FC SK I-III sub-GeV 1-ring  $e$ -like 0-decay electrons  $\log P = [2.0, 2.6]$  (20 bins), the SK IV-V FC sub-GeV 1-ring  $\bar{\nu}_e$ -like 0-neutron  $\log P = [2.0, 2.6]$  (20 bins), and the SK IV-V FC sub-GeV 1-ring  $\bar{\nu}_e$ -like 1-neutron  $\log P = [2, 2.4]$  (10 bins). For these 50 bins, we get a per-bin  $\Delta\chi^2$  of 0.017, which we consider too large. We suspect this is due to edge effects. These 50 offending bins do not contribute to the NMO signal (only 0.1% of the total NMO difference is in these bins), and due to the observed mismodeling, we exclude them from our analysis. This reduces the 930 bins of the original sample to 880 bins used here. If the actual resolution function were provided in a future data release, this situation could be avoided.

Figure 2 shows the Super-K data used together with the best-fitting model predictions (including systematic uncertainties). The overall goodness-of-fit when fitting Super-K alone is 15.9%, and the preference for NO over IO is  $\Delta\chi^2_{\text{IO-NO}}$  is 5.27 in excellent agreement with the official result of 5.23 (see Table II).

## B. DeepCore

We use the IceCube-DeepCore analysis presented in [24], which analyzes 7.5 years of effective lifetime of atmospheric neutrino data for the atmospheric mixing parameters  $\Delta m_{31}^2$  and  $\theta_{23}$ .

The collaboration provides a public data release [25], which we use here. The dataset provides Monte Carlo (MC) simulation and observed event counts in 200 analysis bins, split into ten reconstructed energy bins spanning approximately 6.3 to 160 GeV, ten reconstructed zenith bins covering  $\cos\theta \in [-1, 0.1]$ , and two particle identification categories (mixed and tracks). The MC is provided separately for  $\nu_e$  CC,  $\nu_\mu$  CC,  $\nu_\tau$  CC, and NC interactions for both neutrinos and antineutrinos, along with an atmospheric muon background contribution. Expected event counts are obtained by reweighting the MC to the oscillated atmospheric neutrino flux.

Detector systematic uncertainties are contained in the data release. These are modeled using precomputed linear response functions encoding the bin-by-bin sensitivities to bulk ice absorption, bulk ice scattering, overall digital optical module optical efficiency, and hole ice properties. These uncertainties are tabulated as a function of  $|\Delta m_{31}^2|$  and interpolated linearly at each fit point to capture the mass-splitting dependence of the detection efficiency. Together with overall normalization (effective area scale) and atmospheric muon normalization, this results in seven detector-specific systematic parameters.

Figure 3 shows the DeepCore data together with the best-fitting model predictions. The overall goodness-of-fit when



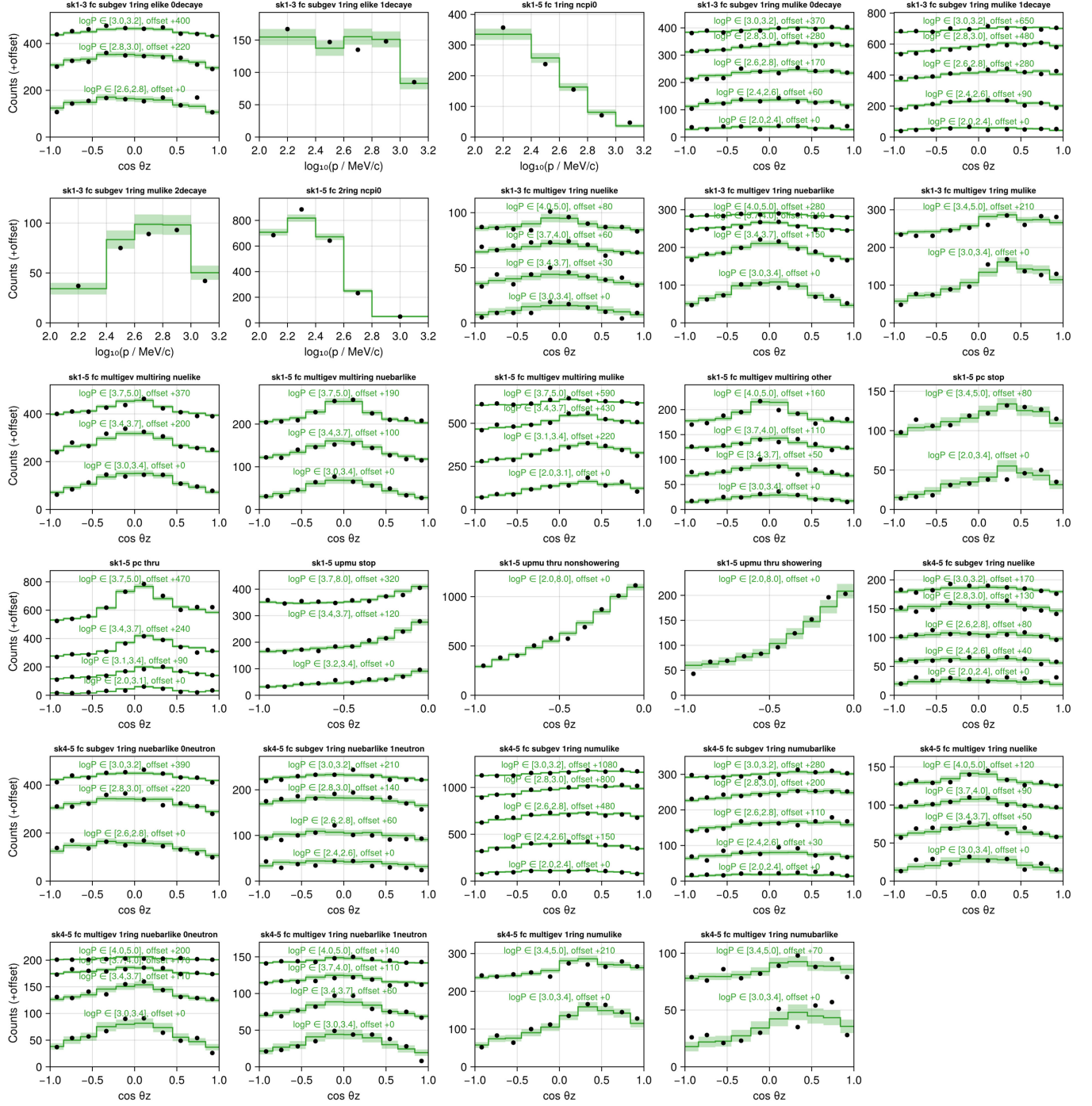


FIG. 2. All Super-K data points (black) used in this analysis, overlaid with our model predictions (colorful bands) at the global best-fit point. The width of the bands represents the  $1\sigma$  statistical uncertainty. The various reconstructed momentum ( $\log P$ ) bins are offset with constants provided in the plots to improve visibility.

fitting DeepCore alone is 28.4% (see Table II), in good agreement with the 26.1% reported in [24].

Our preference for NO over IO is  $\Delta\chi^2_{\text{IO-NO}} = 1.55$ , while the collaboration does not provide this number. The PhD thesis [14] reports a  $\Delta\chi^2_{\text{IO-NO}}$  of 4.398 using a much larger DeepCore event sample with 9.3 years of data and 150 257 events [26], which is not available publicly.

### C. ORCA

KM3NeT has published their first oscillation results using 510 days of data collected from a detector configuration with six lines (ORCA6) [15], resulting in a 433 kton-years exposure. The data was used both to measure atmospheric mixing parameters and analyze differences between the two mass orderings. A dataset

TABLE II. Summary of the datasets used in the five fits. Reported are also Pearson  $\chi^2$ , including prior penalty terms. The goodness-of-fit p values are obtained from the  $\chi^2$  distribution. The “Individual Fit” column refers to fitting one dataset at a time, while the “Combined Fit” uses the same best-fit parameter values from the combined fit for all datasets. The last column shows the  $\chi^2$  difference between the best-fitting NO and IO expectations.

| Dataset  | Bins | Events | Individual fit |         | Combined fit |         | $\Delta\chi^2_{\text{IO-NO}}$ |
|----------|------|--------|----------------|---------|--------------|---------|-------------------------------|
|          |      |        | $\chi^2$       | p value | $\chi^2$     | p value |                               |
| Daya Bay | 26   | 752954 | 23.93          | 0.580   | 24.46        | 0.550   | 0.00                          |
| Super-K  | 880  | 58352  | 921.86         | 0.159   | 925.06       | 0.142   | 5.27                          |
| DeepCore | 200  | 21914  | 210.95         | 0.284   | 215.80       | 0.211   | 1.55                          |
| ORCA     | 430  | 5828   | 475.47         | 0.064   | 500.06       | 0.011   | 0.78                          |
| Combined | 1536 | 839048 | ...            | ...     | 1648.22      | 0.023   | 9.11                          |

was made available with the publication [27], which we use here.

The dataset provides binned response matrices to map from  $(E_\nu, \cos\theta_{\text{true}})$  to  $(E_{\text{reco}}, \cos\theta_{\text{reco}})$  per flavor, and separately for NC and CC interactions. Furthermore, the sample is divided into three event categories: “high-purity tracks,” “low-purity tracks,” and “showers.” The dataset uses a very coarse grid of ten logarithmically spaced points in true energy per decade. At these points, we reweight the spectra to account for changes in atmospheric flux, oscillations, and interaction cross-sections. To better capture the physics, we oversample  $E_\nu$  to 50 points per decade, matching what we use for Super-K and DeepCore. The data release does not include any parametrized systematic detector uncertainties, and the official analysis uses only an energy-scale uncertainty, which we also model. Otherwise, we implement the same five normalization uncertainties as

in the official analysis, resulting in six detector-specific parameters.

The official goodness of fit p value is reported to be 0.012 [15], while we get 0.064 when fitting ORCA alone. This increased goodness of fit is largely due to the additional systematic uncertainties that we consider compared to the KM3NeT analysis (see also Table II). When running with a reduced set of systematic uncertainties, similar to those used by the KM3NeT collaboration, our goodness of fit p value reduces to 0.022. The collaboration also reports a  $\Delta\chi^2$  between NO and IO of 0.62, while we observe a compatible value of 0.78. Figure 4 shows the ORCA data with our best-fit predictions overlaid.

### D. Daya Bay

Daya Bay was a reactor neutrino experiment that measured the reactor angle  $\theta_{13}$  to world-leading precision and also constrained the atmospheric mass splitting [28]. The analysis is based on an event sample collected during 3158 days of operation, resulting in 752 954 candidate

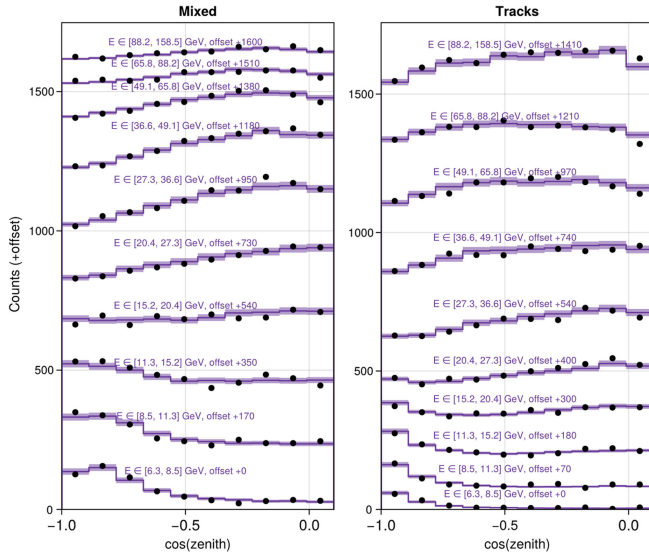


FIG. 3. All DeepCore data points (black) used in this analysis, overlaid with our model predictions (colorful bands) at the global best-fit point. The width of the bands represents the  $1\sigma$  statistical uncertainty. The various reconstructed energy bins are offset with constants provided in the plots to improve visibility.

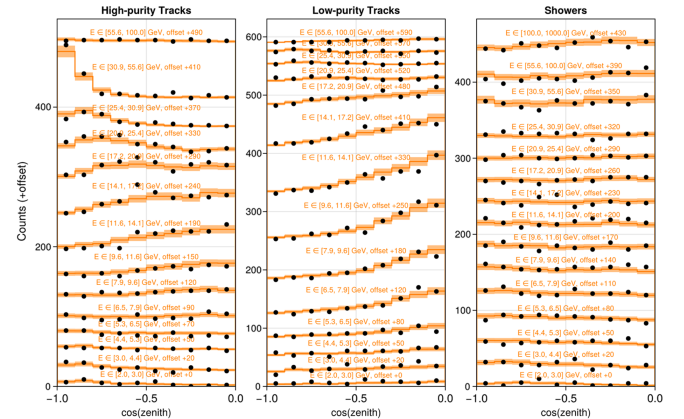


FIG. 4. All ORCA data points (black) used in this analysis, overlaid with our model predictions (colorful bands) at the global best-fit point. The width of the bands represents the  $1\sigma$  statistical uncertainty. The various reconstructed energy bins are offset with constants provided in the plots to improve visibility.

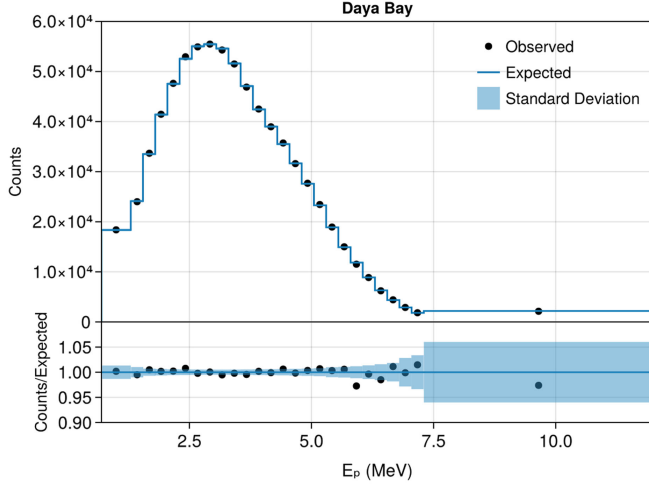


FIG. 5. All Daya Bay data points (black) used in this analysis, overlaid with our model predictions (colorful bands) at the global best-fit point. The width of the bands represents the  $1\sigma$  statistical and systematic uncertainty (from the covariance matrix).

neutrino events collected in the far detectors. The publication contains a data release as Supplemental Material, which we use here. The data release provides a full covariance matrix to account for systematic uncertainties and their correlations between bins. This means we do not need any parameters other than the oscillation ones for Daya Bay in our fit. This treatment is justified, since the reactor neutrinos are fully complementary to the atmospheric neutrino datasets. The source of neutrinos is different, the neutrino energies are much lower (MeV), and the interactions happen via inverse beta decay. This means that there are no systematic uncertainties shared with the atmospheric experiments.

Figure 5 shows the used data points in the 26 bins and our best-fitting expectation. We get an excellent  $\chi^2$  goodness of fit of 58%. There is no preference ( $\Delta\chi^2 < 10^{-4}$ ) for any mass ordering, as expected. See also Table II.

### III. ANALYSIS

The analysis models the expected event rates from atmospheric neutrino fluxes, oscillation probabilities, and interaction cross sections. These effects are used to reweight the datasets provided by the collaborations. In addition, detector-specific uncertainties (as discussed in Sec. II) are included, and a joint likelihood over the datasets is used, including a penalty term for priors on nuisance parameters, in a frequentist analysis.

For the nominal predictions of the atmospheric neutrino fluxes, we rely on the calculations in [29], which provide tabulated values for several locations [30]. We use the Kamioka site for Super-K, the South Pole site for DeepCore, and the Frejus site for ORCA. The modeling of systematic uncertainties concerning flux ratios ( $\nu_e/\bar{\nu}_e$ ,  $\nu_\mu/\bar{\nu}_\mu$ , and  $\nu_e/\nu_\mu$ ) in three energy bands (sub-GeV,

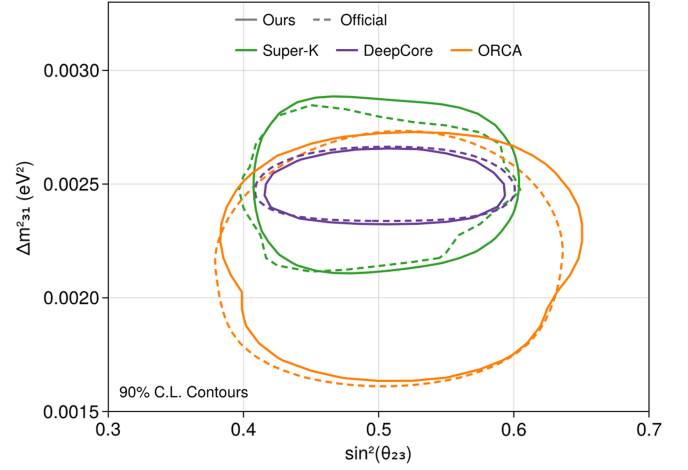


FIG. 6. Validation plot for the three atmospheric datasets included. Shown are the 90% CL contours for the atmospheric mixing parameters  $\Delta m_{31}^2$  and  $\sin^2(\theta_{23})$  for our implementation (solid lines) and the official results (dashed lines), for normal ordering.

1–10 GeV, and above), and angular asymmetries (up down and up horizontal) follows [31], and an additional uncertainty  $\Delta_\gamma$  modifying the spectral shape  $E^{-(\gamma+\Delta_\gamma)}$ . This totals to 12 parameters modifying the atmospheric flux included in our model. The overall flux normalization, which is typically used by collaborations to account for other effects as well, is left unconstrained in the analysis. In ORCA and DeepCore, this is handled via a fit parameter for the overall normalization for each. In Super-K, it is handled via an overall norm, and two separate, energy-dependent parameters scaling the  $<1$  GeV and the  $>1$  GeV regions (see [21] for details).

We use the PREM model [32] to compute matter densities across four concentric Earth layers (inner core, outer core, mantle, crust). In addition, we add a 20 km offset for the production height in the atmosphere. A fit parameter changes the electron density in the Earth around the nominal PREM values with a 6.8% uncertainty.

To compute oscillation probabilities, we use the method presented in [33], which allows for correct averaging over energies and baselines in the presence of matter effects. We use this to average effects finer than our computational grid (50 points, log-spaced per decade of  $E_\nu$ ), and to account for an uncertainty in the atmospheric production height of  $\pm 10$  km.

The neutrino interactions in the detectors were simulated with NEUT 5.4.0 for Super-K, while GENIE v2 (G00\_00A) [34] was used in DeepCore and ORCA. In a combined analysis, however, we need to treat the datasets on a common ground; therefore, we reweight the DeepCore and ORCA spectra from GENIE to NEUT expectations. An equally valid option would be to keep ORCA and DeepCore at the GENIE reference values, and reweight Super-K instead. We tested that the choice of this reference



does not significantly affect the results presented in this study (changes of  $\lesssim 0.2\sigma$ ).

For cross-section uncertainties, we evaluated the per-process differences between the predictions of NEUT 5.4.0 and GENIE 3 with different tunes (G18\_02A, G18\_10A, and G21\_11A). This results in a normalization, a shape, and a  $\nu/\bar{\nu}$  ratio uncertainty per interaction process. For charged-current (CC) interactions, we differentiate between quasi-elastic (1p1h), meson-exchange current (2p2h), resonant pion production (1pi), deep-inelastic scattering, and all other processes (other). In addition, for quasielastic scattering, we add a parameter controlling the  $\nu_e/\nu_\mu$  ratio, and introduce a separate normalization for sub-GeV events. Finally, we have an additional 25% uncertainty on the  $\nu_\tau$  CC cross section. For neutral current interactions, we only consider an inclusive reweighting. This results in 21 cross-section specific nuisance parameters included in our fit.

All data releases separate between flavors, CC and NC interactions, and  $\nu$  vs  $\bar{\nu}$  neutrinos (only in Super-K, NC and  $\nu_\tau$  CC do not provide  $\nu$  vs  $\bar{\nu}$  separation), which we use to reweight the differential cross sections  $d\sigma/dE$  separately. In the DeepCore data release, each MC event also contains the information of the simulated CC interaction process, which we use for per-process reweighting.

Data analysis is implemented in Julia, and the code is publicly available under the MIT open-source license [35,36]. It interfaces with BAT.jl [37] for running statistical inference. The code is fully forward- and reverse-mode autodifferentiable, which allows for the computation of partial derivatives used here for gradient descent optimization, the Fisher information, and Jacobians for a change of variables. The same framework has previously been used for the results shown in [20,38–40].

The analysis of Daya Bay data uses a  $\chi^2$  test statistic including the provided covariance matrix to account for systematic uncertainties. All other datasets are fitted with a Poisson likelihood and added prior pull terms,

$$\chi^2 = \sum_{i \in \text{bins}} -2 \log \left( \frac{e^{-E_i} E_i^{O_i}}{O_i!} \right) + \sum_{j \in \text{priors}} \frac{(\nu_j - \mu_j)^2}{\sigma_j^2} \quad (1)$$

where  $E_i$  and  $O_i$  are the expected and observed number of events in each bin  $i$ , and the second sum over  $j$  acts as a prior penalty for nuisance parameters  $\nu_j$  that have a normal prior with mean  $\mu_j$  and standard deviation  $\sigma_j$ .

We minimize the  $\chi^2$  over all nuisance parameters  $\nu_j$  to find maximum likelihood estimators (MLEs)  $\hat{\nu}_j$  using gradient descent. To constrain the mixing parameters, we compute the profile likelihood ratio over fixed grids of parameter values while fitting all other parameters. Setting all parameters to their global MLEs for both orderings gives the difference between IO and NO, denoted  $\Delta\chi^2_{\text{IO-NO}}$ .

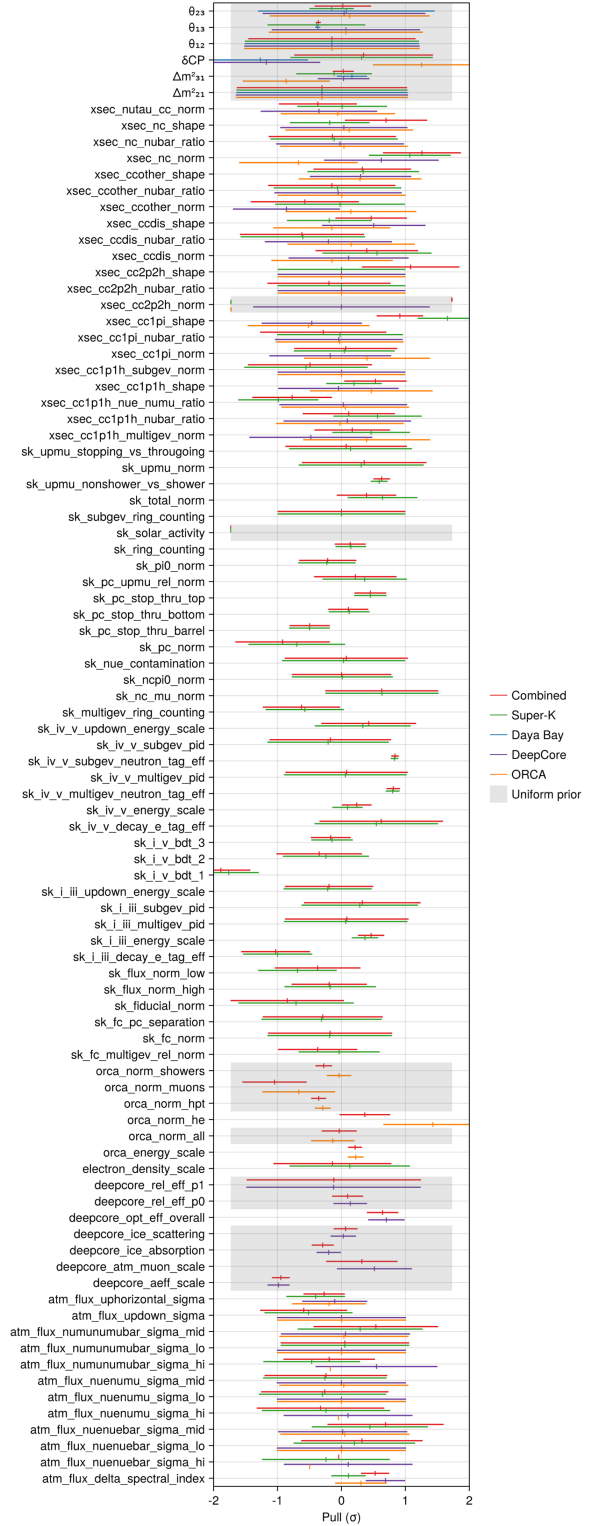


FIG. 7. Best-fit parameter values and approximate uncertainty (see text) for the five different fits. The values are shown as pulls in the prior space. Uniform priors span  $\pm\sqrt{12}/2$ , indicated by gray boxes; otherwise, normal priors are used.

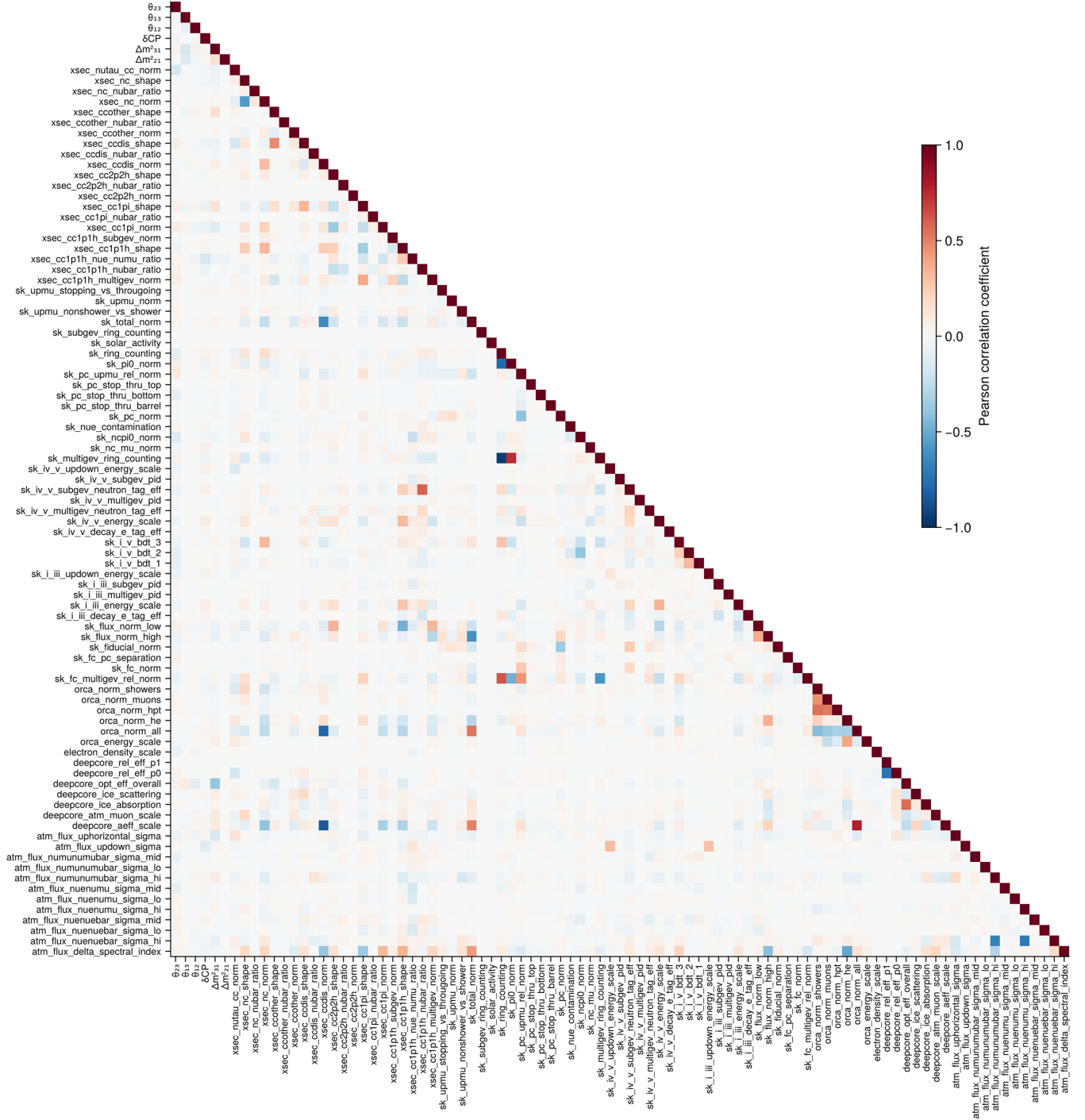


FIG. 8. Pearson correlation coefficients of all fit parameters in the combined fit, obtained from inverting the Hessian (see text).

#### IV. RESULTS

We perform five separate analyses: fitting each experiment individually, and once in a combined fit. The combined fit needs to find parameter values that can accommodate all data at the same time. This means that one set of oscillation, flux, and cross-section parameters needs to be able to explain all four datasets, with a total of 839 048 events divided into 1536 bins simultaneously.

Because of the reduced degree of freedom in a combined fit, the goodness-of-fit per dataset will strictly degrade compared to that from the individual fits. The question is: will the combined fit still be able to explain the data, or will large tensions in shared parameters be observed?

Table II gives the breakdown of  $\chi^2$  and goodness-of-fit  $p$  values. The goodness-of-fit computed over individual datasets degrades from the individual to the combined fit



only minimally, from 58% to 55% in the case of Daya Bay, as expected, since no nuisance parameters are shared. For Super-K, the number goes from 15.9% to 14.2%, and for DeepCore from 28.4% to 21.1%, which represent only modest decreases, and excellent goodness-of-fit was achieved in the combined case. For ORCA, the goodness-of-fit was already not excellent in the individual fit (6.4%), and in the combined case, it is at 1.1%. This number is consistent with the official 1.2% reported by the collaboration in [15]. The low ORCA p value also causes the combined p value across all datasets to be 2.3%. This is a limitation of the ORCA dataset, and not our combined fit, as can be seen from removing the ORCA data points from the global goodness of fit computation, which increases the p value to 14.1%.

Figure 7 provides further insight into the fit parameters under the five analyses. Shown are the best-fit values with uncertainties estimated via a Laplace approximation, where the posterior covariance is approximated by the inverse of the Fisher information matrix evaluated at the best-fit point. These uncertainties are approximate, assume Gaussian posteriors, and should not be directly compared with a proper profile-likelihood analysis. In particular, parameters with uniform priors whose best-fit values lie near the prior boundary are not well described by a Gaussian approximation and may have severely underestimated uncertainties (e.g., “xsec\_cc2p2h\_norm,” “sk\_solar\_activity”). The plot demonstrates that there are no large differences in the fit parameters across the five analyses, indicating that the model can describe all datasets with the same underlying physics.

Figure 8 provides further insight into the model parameters, showing the Pearson correlation coefficient (also obtained from the inverse Fisher information) for the combined fit. Some expected correlations are clearly visible, for example, the ORCA normalization parameters are strongly correlated with one another. So are the deep-core detector uncertainties. Furthermore, correlations can be seen, for example, between each experiment’s overall normalization and cross-section normalizations.

Figure 1 shows the profile likelihood contours (converted to  $\Delta\chi^2$ ) obtained from the five fits for the mixing parameters. As expected, Super-K is the only dataset with sensitivity to  $\delta_{CP}$ , and Daya Bay constrains  $\theta_{13}$ . All four experiments are sensitive to the atmospheric mass splitting. ORCA prefers a somewhat lower value, but has large uncertainties, the other experiments all fit around  $2.5 \times 10^{-3} \text{ eV}^2$ , and none of the fits are in tension. The atmospheric mixing angle is constrained by the three atmospheric datasets, all of which prefer maximal, or close to maximal, mixing. The best-fit numbers and uncertainties for the five analyses are reported in Table I.

A preference for normal over inverted ordering is present in all three atmospheric datasets. In the combined analysis including Daya Bay, this preference is at a  $\Delta\chi^2_{\text{IO-NO}} = 9.11$ .

Assuming normality, this corresponds to approximately  $3\sigma$  preference for the normal ordering [41].

## V. CONCLUSIONS

We presented the first comprehensive combined oscillation analysis of multiple atmospheric neutrino datasets. Our combination includes data from Super-K, IceCube, and KM3NeT, paired with Daya Bay, and demonstrates the feasibility of such an analysis. Our study demonstrates that the data can be combined with a unified physics model, and no large parameter tensions were observed. Atmospheric data offer a powerful way to measure mixing parameters and to discern the neutrino mass orderings. We found consistent mixing parameter values when fitting all datasets individually and jointly. We disfavor the inverted neutrino mass ordering at  $\Delta\chi^2 = 9.11$ .

The three detectors used to collect the atmospheric samples had recently undergone extensive hardware upgrades: Super-K is taking data with Gadolinium loading [42], IceCube has successfully installed the IceCube-Upgrade [43] featuring five new strings, and ORCA by now comprises 38 detector lines. This means exciting new analyses and results on atmospheric neutrino oscillations are on the horizon. In a combined fit, either based on public data as presented here or in an intercollaboration effort, a decisive answer to the open question of the neutrino mass ordering seems attainable in the near future [44].

When combining atmospheric with additional reactor data from JUNO, rather than only Daya Bay, another boost in NMO sensitivity is expected [45,46], further advocating for combined analyses that include atmospheric neutrinos.

A more immediate and practical improvement would also come from a more complete Super-K data release, including per-phase information, detector resolution functions, and a description of detector uncertainties. Similarly, updated and improved public data release by IceCube and KM3NeT would benefit combined analyses.

## ACKNOWLEDGMENTS

We would like to thank O. Schulz, T. Wester, D. Schultheiß, T. Kozynets, and A. Zander for useful discussions. We thank the Super-K, Daya Bay, IceCube, and KM3NeT collaborations for providing public data releases. P.E. is a member of the IceCube collaboration and an observer of the KM3NeT collaboration. This work has been supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy—EXC-2094/2–390783311, and the SFB 1258–283604770.

## DATA AVAILABILITY

Data analysis is implemented in Julia, and the code is publicly available under the MIT open-source license [47].

- [1] Y. Fukuda *et al.* (Super-Kamiokande Collaboration), Evidence for oscillation of atmospheric neutrinos, *Phys. Rev. Lett.* **81**, 1562 (1998).
- [2] Q. R. Ahmad *et al.* (SNO Collaboration), Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory, *Phys. Rev. Lett.* **89**, 011301 (2002).
- [3] K. Eguchi *et al.* (KamLAND Collaboration), First results from KamLAND: Evidence for reactor anti-neutrino disappearance, *Phys. Rev. Lett.* **90**, 021802 (2003).
- [4] K. Abe *et al.* (T2K Collaboration), Results from the T2K experiment on neutrino mixing including a new far detector  $\mu$ -like sample, *Phys. Rev. Lett.* **135**, 261801 (2025).
- [5] S. Abubakar *et al.* (NOvA Collaboration), Precision measurement of neutrino oscillation parameters with 10 years of data from the NOvA experiment, *Phys. Rev. Lett.* **136**, 011802 (2026).
- [6] 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.
- [7] P. F. de Salas, D. V. Forero, S. Gariazzo, P. Martínez-Miravé, O. Mena, C. A. Ternes, M. Tortola, and J. W. F. Valle, 2020 global reassessment of the neutrino oscillation picture, *J. High Energy Phys.* **02** (2021) 071.
- [8] F. Capozzi, W. Giarè, E. Lisi, A. Marrone, A. Melchiorri, and A. Palazzo, Neutrino masses and mixing: Entering the era of subpercent precision, *Phys. Rev. D* **111**, 093006 (2025).
- [9] S. Abubakar *et al.* (T2K Collaboration and NOvA Collaboration), Joint neutrino oscillation analysis from the T2K and NOvA experiments, *Nature (London)* **646**, 818 (2025).
- [10] A. Abusleme *et al.* (JUNO Collaboration), Measurement of reactor neutrino oscillation with the first JUNO data, *Nature (London)* **654**, 343 (2026).
- [11] T. Wester *et al.* (Super-Kamiokande Collaboration), Atmospheric neutrino oscillation analysis with neutron tagging and an expanded fiducial volume in Super-Kamiokande I–V, *Phys. Rev. D* **109**, 072014 (2024).
- [12] K. Abe *et al.* (T2K Collaboration and Super-Kamiokande Collaboration), First joint oscillation analysis of super-Kamiokande atmospheric and T2K accelerator neutrino data, *Phys. Rev. Lett.* **134**, 011801 (2025).
- [13] M. G. Aartsen *et al.* (IceCube Collaboration), Development of an analysis to probe the neutrino mass ordering with atmospheric neutrinos using three years of IceCube DeepCore data, *Eur. Phys. J. C* **80**, 9 (2020).
- [14] M. V. P. Rodriguez, Measurement of the neutrino mass ordering with 9.28 years of icecube deepcore data, Ph.D. thesis, University of Wisconsin–Madison, 2024.
- [15] S. Aiello *et al.* (KM3NeT Collaboration), Measurement of neutrino oscillation parameters with the first six detection units of KM3NeT/ORCA, *J. High Energy Phys.* **10** (2024) 206; **10** (2025) 41.
- [16] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, I. Martinez-Soler, and T. Schwetz, Updated fit to three neutrino mixing: Exploring the accelerator-reactor complementarity, *J. High Energy Phys.* **01** (2017) 087.
- [17] M. C. Gonzalez-Garcia, M. Maltoni, and T. Schwetz, Updated fit to three neutrino mixing: Status of leptonic  $CP$  violation, *J. High Energy Phys.* **11** (2014) 052.
- [18] I. Esteban, M. C. Gonzalez-Garcia, A. Hernandez-Cabezudo, M. Maltoni, and T. Schwetz, Global analysis of three-flavour neutrino oscillations: Synergies and tensions in the determination of  $\theta_{23}$ ,  $\delta_{CP}$ , and the mass ordering, *J. High Energy Phys.* **01** (2019) 106.
- [19] P. F. de Salas, D. V. Forero, C. A. Ternes, M. Tortola, and J. W. F. Valle, Status of neutrino oscillations 2018:  $3\sigma$  hint for normal mass ordering and improved  $CP$  sensitivity, *Phys. Lett. B* **782**, 633 (2018).
- [20] T. Kozynets, P. Eller, A. Zander, M. Ettengruber, and D. J. Koskinen, Constraints on non-unitary neutrino mixing in light of atmospheric and reactor neutrino data, *J. High Energy Phys.* **05** (2025) 130.
- [21] T. Wester, Discerning the Neutrino Mass Ordering using Atmospheric Neutrinos in Super-Kamiokande I–V, Ph.D. thesis, Boston University, Boston University, 2023.
- [22] Super-Kamiokande Collaboration, Data release: Atmospheric neutrino oscillation analysis with neutron tagging and an expanded fiducial volume in Super-Kamiokande I–V, [10.5281/zenodo.8401262](https://zenodo.org/record/8401262) (2023).
- [23] C. Jesús-Valls, Parametrizing the reconstruction performance of Super-Kamiokande in the Sub-GeV to TeV neutrino energy range, [arXiv:2505.20812](https://arxiv.org/abs/2505.20812).
- [24] R. Abbasi *et al.* (IceCube Collaboration), Measurement of atmospheric neutrino mixing with improved IceCube DeepCore calibration and data processing, *Phys. Rev. D* **108**, 012014 (2023).
- [25] IceCube Collaboration, Replication Data for: Measurement of atmospheric neutrino mixing with improved IceCube DeepCore calibration and data processing, [10.7910/DVN/B4RITM](https://doi.org/10.7910/DVN/B4RITM) (2025).
- [26] R. Abbasi *et al.* ((IceCube Collaboration) and IceCube Collaboration), Measurement of atmospheric neutrino oscillation parameters using convolutional neural networks with 9.3 years of data in IceCube DeepCore, *Phys. Rev. Lett.* **134**, 091801 (2025).
- [27] KM3NeT Collaboration, Oscillation analysis ORCA433kt-y / ORCA 6, [10.5072/FK2/Y0UXVW](https://arxiv.org/abs/10.5072/FK2/Y0UXVW) (2025).
- [28] F. P. An *et al.* (Daya Bay Collaboration), Precision measurement of reactor antineutrino oscillation at kilometer-scale baselines by Daya Bay, *Phys. Rev. Lett.* **130**, 161802 (2023).
- [29] M. Honda, M. Sajjad Athar, T. Kajita, K. Kasahara, and S. Midorikawa, Atmospheric neutrino flux calculation using the NRLMSISE-00 atmospheric model, *Phys. Rev. D* **92**, 023004 (2015).
- [30] M. Honda, M. Sajjad Athar, T. Kajita, K. Kasahara, and S. Midorikawa, Atmospheric neutrino flux tables for one-year-average (HAKKM, 2014), <http://www-rcn.icrr.u-tokyo.ac.jp/mhonda/public/nflx2014> (2014).
- [31] G. D. Barr, T. K. Gaisser, S. Robbins, and T. Stanev, Uncertainties in atmospheric neutrino fluxes, *Phys. Rev. D* **74**, 094009 (2006).
- [32] A. M. Dziewonski and D. L. Anderson, Preliminary reference earth model, *Phys. Earth Planet. Inter.* **25**, 297 (1981).

- [33] M. Maltoni, From ray to spray: Augmenting amplitudes and taming fast oscillations in fully numerical neutrino codes, *J. High Energy Phys.* **11** (2023) 033.
- [34] C. Andreopoulos, C. Barry, S. Dytman, H. Gallagher, T. Golan, R. Hatcher, G. Perdue, and J. Yarba, The GENIE Neutrino Monte Carlo Generator: Physics and User Manual, *Nucl. Instrum. Methods Phys. Res., Sect. A* **614**, 84 (2015).
- [35] P. Eller *et al.*, Newtrinos.jl (2026).
- [36] P. Eller and D. Schultheiß, Newtrinos.jl: A Julia package for global analysis of neutrino data, [arXiv:2608.12940](https://arxiv.org/abs/2608.12940).
- [37] O. Schulz, F. Beaujean, A. Caldwell, C. Grunwald, V. Hafych, K. Kröninger, S. L. Cagnina, L. Röhrig, and L. Shtembari, Bat.jl: A julia-based tool for bayesian inference, *SN Comput. Sci.* **2**, 210 (2021).
- [38] M. Ettengruber, A. Zander, and P. Eller, Testing the number of neutrino species with a global fit of neutrino data, *Phys. Rev. D* **109**, 095016 (2024).
- [39] P. Eller, M. Ettengruber, and A. Zander, Neutrino data analysis of extra-dimensional theories with massive bulk fields, *Phys. Rev. D* **112**, 055009 (2025).
- [40] S. Lonardi, M. Ettengruber, and P. Eller, Searching for the  $N$ -naturalness tower of neutrinos, [arXiv:2607.14243](https://arxiv.org/abs/2607.14243).
- [41] M. Blennow, P. Coloma, P. Huber, and T. Schwetz, Quantifying the sensitivity of oscillation experiments to the neutrino mass ordering, *J. High Energy Phys.* **03** (2014) 028.
- [42] K. Abe *et al.* (Super-Kamiokande Collaboration), Second gadolinium loading to Super-Kamiokande, *Nucl. Instrum. Methods Phys. Res., Sect. A* **1065**, 169480 (2024).
- [43] R. Abbasi *et al.* (IceCube Collaboration), Physics potential of the IceCube Upgrade for atmospheric neutrino oscillations, *Phys. Rev. D* **113**, 072009 (2026).
- [44] C. A. Argüelles, P. Fernández, I. Martínez-Soler, and M. Jin, Measuring oscillations with a million atmospheric neutrinos, *Phys. Rev. X* **13**, 041055 (2023).
- [45] M. G. Aartsen *et al.* (IceCube-Gen2 Collaboration), Combined sensitivity to the neutrino mass ordering with JUNO, the IceCube Upgrade, and PINGU, *Phys. Rev. D* **101**, 032006 (2020).
- [46] S. Aiello *et al.* (KM3NeT Collaboration and JUNO Collaboration), Combined sensitivity of JUNO and KM3NeT/ORCA to the neutrino mass ordering, *J. High Energy Phys.* **03** (2022) 055.
- [47] <https://github.com/Newtrinos-org/Newtrinos.jl>.