Eur. Phys. J. C (2020) 80:9
https://doi.org/10.1140/epjc/s10052-019-7555-0

THE EUROPEAN PHYSICAL JOURNAL C

Regular Article - Experimental Physics

Development of an analysis to probe the neutrino mass ordering with atmospheric neutrinos using three years of IceCube DeepCore data

IceCube Collaboration*

M. G. Aartsen[16], M. Ackermann[53], J. Adams[16], J. A. Aguilar[12], M. Ahlers[20], M. Ahrens[45], C. Alispach[25], K. Andeen[35], T. Anderson[50], I. Ansseau[12], G. Anton[24], C. Argüelles[14], J. Auffenberg[1], S. Axani[14], P. Backes[1], H. Bagherpour[16], X. Bai[42], A. Barbano[25], S. W. Barwick[27], V. Baum[33], R. Bay[8], J. J. Beatty[18,19], K.-H. Becker[52], J. Becker Tjus[11], S. BenZvi[44], D. Berley[17], E. Bernardini[53], D. Z. Besson[28], G. Binder[8,9], D. Bindig[52], E. Blaufuss[17], S. Blot[53], C. Bohm[45], M. Börner[21], S. Böser[33], O. Botner[51], E. Bourbeau[20], J. Bourbeau[32], F. Bradascio[53], J. Braun[32], H.-P. Bretz[53], S. Bron[25], J. Brostean-Kaiser[53], A. Burgman[51], R. S. Busse[32], T. Carver[25], C. Chen[6], E. Cheung[17], D. Chirkin[32], K. Clark[29], L. Classen[37], G. H. Collin[14], J. M. Conrad[14], P. Coppin[13], P. Correa[13], D. F. Cowen[49,50], R. Cross[44], P. Dave[6], J. P. A. M. de André[22], C. De Clercq[13], J. J. DeLaunay[50], H. Dembinski[38], K. Deoskar[45], S. De Ridder[26], P. Desiati[32], K. D. de Vries[13], G. de Wasseige[13], M. de With[10], T. DeYoung[22], A. Diaz[14], J. C. Díaz-Vélez[32], H. Dujmovic[47], M. Dunkman[50], E. Dvorak[42], B. Eberhardt[32], T. Ehrhardt[33], B. Eichmann[11], P. Eller[50], J. J. Evans[34,a], P. A. Evenson[38], S. Fahey[32], A. R. Fazely[7], J. Felde[17], K. Filimonov[8], C. Finley[45], A. Franckowiak[53], E. Friedman[17], A. Fritz[33], T. K. Gaisser[38], J. Gallagher[31], E. Ganster[1], S. Garrappa[53], L. Gerhardt[9], K. Ghorbani[32], T. Glauch[36], T. Glüsenkamp[24], A. Goldschmidt[9], J. G. Gonzalez[38], D. Grant[22], Z. Griffith[32], M. Günder[1], M. Gündüz[11], C. Haack[1], A. Hallgren[51], L. Halve[1], F. Halzen[32], K. Hanson[32], D. Hebecker[10], D. Heereman[12], K. Helbing[52], R. Hellauer[17], F. Henningsen[36], S. Hickford[52], J. Hignight[22], G. C. Hill[2], K. D. Hoffman[17], R. Hoffmann[52], T. Hoinka[21], B. Hokanson-Fasig[32], K. Hoshina[32,b], F. Huang[50], M. Huber[36], K. Hultqvist[45], M. Hünnefeld[21], R. Hussain[32], S. In[47], N. Iovine[12], A. Ishihara[15], E. Jacobi[53], G. S. Japaridze[5], M. Jeong[47], K. Jero[32], B. J. P. Jones[4], W. Kang[47], A. Kappes[37], D. Kappesser[33], T. Karg[53], M. Karl[36], A. Karle[32], U. Katz[24], M. Kauer[32], J. L. Kelley[32], A. Kheirandish[32], J. Kim[47], T. Kintscher[53], J. Kiryluk[46], T. Kittler[24], S. R. Klein[8,9], R. Koirala[38], H. Kolanoski[10], L. Köpke[33], C. Kopper[22], S. Kopper[48], D. J. Koskinen[20], M. Kowalski[10,53], K. Krings[36], G. Krückl[33], N. Kulacz[23], S. Kunwar[53], N. Kurahashi[41], A. Kyriacou[2], M. Labare[26], J. L. Lanfranchi[50], M. J. Larson[17], F. Lauber[52], J. P. Lazar[32], K. Leonard[32], M. Leuermann[1], Q. R. Liu[32], E. Lohfink[33], C. J. Lozano Mariscal[37], L. Lu[15], F. Lucarelli[25], J. Lünemann[13], W. Luszczak[32], J. Madsen[43], G. Maggi[13], K. B. M. Mahn[22], Y. Makino[15], K. Mallot[32], S. Mancina[32], I. C. Mariş[12], R. Maruyama[39], K. Mase[15], R. Maunu[17], K. Meagher[12], M. Medici[20], A. Medina[18], M. Meier[21], S. Meighen-Berger[36], T. Menne[21], G. Merino[32], T. Meures[12], S. Miarecki[8,9], J. Micallef[22], G. Momenté[33], T. Montaruli[25], R. W. Moore[23], M. Moulai[14], R. Nagai[15], R. Nahnhauer[53], P. Nakarmi[48], U. Naumann[52], G. Neer[22], H. Niederhausen[46], S. C. Nowicki[23], D. R. Nygren[9], A. Obertacke Pollmann[52], A. Olivas[17], A. O'Murchadha[12], E. O'Sullivan[45], T. Palczewski[8,9], H. Pandya[38], D. V. Pankova[50], N. Park[32], P. Peiffer[33], C. Pérez de los Heros[51], D. Pieloth[21], E. Pinat[12], A. Pizzuto[32], M. Plum[35], P. B. Price[8], G. T. Przybylski[9], C. Raab[12], A. Raissi[16], M. Rameez[20], L. Rauch[53], K. Rawlins[3], I. C. Rea[36], R. Reimann[1], B. Relethford[41], G. Renzi[12], E. Resconi[36], W. Rhode[21], M. Richman[41], S. Robertson[9], M. Rongen[1], C. Rott[47], T. Ruhe[21], D. Ryckbosch[26], D. Rysewyk[22], I. Safa[32], S. E. Sanchez Herrera[23], A. Sandrock[21], J. Sandroos[33], M. Santander[48], S. Sarkar[40], S. Sarkar[23], K. Satalecka[53], M. Schaufel[1], P. Schlunder[21], T. Schmidt[17], A. Schneider[32], J. Schneider[24], L. Schumacher[1], S. Sclafani[41], D. Seckel[38], S. Seunarine[43], M. Silva[32], R. Snihur[32], J. Soedingrekso[21], D. Soldin[38], S. Söldner-Rembold[34], M. Song[17], G. M. Spiczak[43], C. Spiering[53], J. Stachurska[53], M. Stamatikos[18], T. Stanev[38], A. Stasik[53], R. Stein[53], J. Stettner[1], A. Steuer[33], T. Stezelberger[9], R. G. Stokstad[9], A. Stößl[15], N. L. Strotjohann[53], T. Stuttard[20], G. W. Sullivan[17], M. Sutherland[18], I. Taboada[6], F. Tenholt[11], S. Ter-Antonyan[7], A. Terliuk[53], S. Tilav[38], L. Tomankova[11], C. Tönnis[47], S. Toscano[13], D. Tosi[32], M. Tselengidou[24], C. F. Tung[6], A. Turcati[36], R. Turcotte[1], C. F. Turley[50], B. Ty[32], E. Unger[51], M. A. Unland Elorrieta[37], M. Usner[53], J. Vandenbroucke[32], W. Van Driessche[26], D. van Eijk[32], N. van Eijndhoven[13], S. Vanheule[26], J. van Santen[53], M. Vraeghe[26], C. Walck[45],

A. Wallace[2], M. Wallraff[1], N. Wandkowsky[32], T. B. Watson[4], C. Weaver[23], M. J. Weiss[50], J. Weldert[33], C. Wendt[32], J. Werthebach[32], S. Westerhoff[32], B. J. Whelan[2], N. Whitehorn[30], K. Wiebe[33], C. H. Wiebusch[1], L. Wille[32], D. R. Williams[48], L. Wills[41], M. Wolf[36], J. Wood[32], T. R. Wood[23], K. Woschnagg[8], G. Wrede[24], S. Wren[34], D. L. Xu[32], X. W. Xu[7], Y. Xu[46], J. P. Yanez[23], G. Yodh[27], S. Yoshida[15], T. Yuan[32]

[1] III. Physikalisches Institut, RWTH Aachen University, 52056 Aachen, Germany
[2] Department of Physics, University of Adelaide, Adelaide, 5005, Australia
[3] Department of Physics and Astronomy, University of Alaska Anchorage, 3211 Providence Dr., Anchorage, AK 99508, USA
[4] Department of Physics, University of Texas at Arlington, 502 Yates St., Science Hall Rm 108, Box 19059, Arlington, TX 76019, USA
[5] CTSPS, Clark-Atlanta University, Atlanta, GA 30314, USA
[6] School of Physics and Center for Relativistic Astrophysics, Georgia Institute of Technology, Atlanta, GA 30332, USA
[7] Department of Physics, Southern University, Baton Rouge, LA 70813, USA
[8] Department of Physics, University of California, Berkeley, CA 94720, USA
[9] Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[10] Institut für Physik, Humboldt-Universität zu Berlin, 12489 Berlin, Germany
[11] Fakultät für Physik & Astronomie, Ruhr-Universität Bochum, 44780 Bochum, Germany
[12] Université Libre de Bruxelles, Science Faculty CP230, 1050 Brussels, Belgium
[13] Vrije Universiteit Brussel (VUB), Dienst ELEM, 1050 Brussels, Belgium
[14] Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[15] Department of Physics and Institute for Global Prominent Research, Chiba University, Chiba 263-8522, Japan
[16] Department of Physics and Astronomy, University of Canterbury, Private Bag 4800, Christchurch, New Zealand
[17] Department of Physics, University of Maryland, College Park, MD 20742, USA
[18] Department of Physics and Center for Cosmology and Astro-Particle Physics, Ohio State University, Columbus, OH 43210, USA
[19] Department of Astronomy, Ohio State University, Columbus, OH 43210, USA
[20] Niels Bohr Institute, University of Copenhagen, 2100 Copenhagen, Denmark
[21] Department of Physics, TU Dortmund University, 44221 Dortmund, Germany
[22] Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA
[23] Department of Physics, University of Alberta, Edmonton, AB T6G 2E1, Canada
[24] Erlangen Centre for Astroparticle Physics, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany
[25] Département de physique nucléaire et corpusculaire, Université de Genève, 1211 Geneva, Switzerland
[26] Department of Physics and Astronomy, University of Gent, 9000 Gent, Belgium
[27] Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA
[28] Department of Physics and Astronomy, University of Kansas, Lawrence, KS 66045, USA
[29] SNOLAB, 1039 Regional Road 24, Creighton Mine 9, Lively, ON P3Y 1N2, Canada
[30] Department of Physics and Astronomy, UCLA, Los Angeles, CA 90095, USA
[31] Department of Astronomy, University of Wisconsin, Madison, WI 53706, USA
[32] Department of Physics and Wisconsin IceCube Particle Astrophysics Center, University of Wisconsin, Madison, WI 53706, USA
[33] Institute of Physics, University of Mainz, Staudinger Weg 7, 55099 Mainz, Germany
[34] School of Physics and Astronomy, The University of Manchester, Oxford Road, Manchester M13 9PL, UK
[35] Department of Physics, Marquette University, Milwaukee, WI 53201, USA
[36] Physik-department, Technische Universität München, 85748 Garching, Germany
[37] Institut für Kernphysik, Westfälische Wilhelms-Universität Münster, 48149 Münster, Germany
[38] Bartol Research Institute and Department of Physics and Astronomy, University of Delaware, Newark, DE 19716, USA
[39] Department of Physics, Yale University, New Haven, CT 06520, USA
[40] Department of Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK
[41] Department of Physics, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, USA
[42] Physics Department, South Dakota School of Mines and Technology, Rapid City, SD 57701, USA
[43] Department of Physics, University of Wisconsin, River Falls, WI 54022, USA
[44] Department of Physics and Astronomy, University of Rochester, Rochester, NY 14627, USA
[45] Oskar Klein Centre and Department of Physics, Stockholm University, 10691 Stockholm, Sweden
[46] Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794-3800, USA
[47] Department of Physics, Sungkyunkwan University, Suwon 440-746, Korea
[48] Department of Physics and Astronomy, University of Alabama, Tuscaloosa, AL 35487, USA
[49] Department of Astronomy and Astrophysics, Pennsylvania State University, University Park, PA 16802, USA
[50] Department of Physics, Pennsylvania State University, University Park, PA 16802, USA
[51] Department of Physics and Astronomy, Uppsala University, Box 516, 75120 Uppsala, Sweden
[52] Department of Physics, University of Wuppertal, 42119 Wuppertal, Germany
[53] DESY, 15738 Zeuthen, Germany

Received: 20 February 2019 / Accepted: 11 December 2019 / Published online: 4 January 2020

Abstract The Neutrino Mass Ordering (NMO) remains one of the outstanding questions in the field of neutrino physics. One strategy to measure the NMO is to observe matter effects in the oscillation pattern of atmospheric neutrinos above ~ 1 GeV, as proposed for several next-generation neutrino experiments. Moreover, the existing IceCube DeepCore detector can already explore this type of measurement. We present the development and application of two independent analyses to search for the signature of the NMO with three years of DeepCore data. These analyses include a full treatment of systematic uncertainties and a statistically-rigorous method to determine the significance for the NMO from a fit to the data. Both analyses show that the dataset is fully compatible with both mass orderings. For the more sensitive analysis, we observe a preference for normal ordering with a p-value of $p_{\mathrm{IO}} = 15.3\%$ and $\mathrm{CL_s} = 53.3\%$ for the inverted ordering hypothesis, while the experimental results from both analyses are consistent within their uncertainties. Since the result is independent of the value of δ_{CP} and obtained from energies $E_\nu \gtrsim 5$ GeV, it is complementary to recent results from long-baseline experiments. These analyses set the groundwork for the future of this measurement with more capable detectors, such as the IceCube Upgrade and the proposed PINGU detector.

1 Introduction

The question of the Neutrino Mass Ordering (NMO) is one of the main drivers of the field of neutrino oscillation physics. The NMO describes the ordering of the three neutrino mass eigenstates m_1, m_2, and m_3. The two possible scenarios depend on the sign of $\Delta m^2_{31} = m^2_3 - m^2_1$, often referred to as the atmospheric *mass splitting*, where negative values are known as *Inverted Ordering* (IO) and positive values as *Normal Ordering* (NO).

The three neutrino mass states do not correspond directly to the three neutrino flavor states ν_e, ν_μ, and ν_τ. Instead, each mass state is a superposition of the flavour states, with the mixing described by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix U [1–3], such that

$$\nu_\alpha = \sum_{i=1}^{3} U_{\alpha,i}\,\nu_i\,, \tag{1}$$

where $\alpha \in \{e, \mu, \tau\}$ labels the flavor states and $i \in \{1, 2, 3\}$ labels the mass states. By convention, ν_1 is the state containing the most electron flavor, and ν_3 is the state containing the least.

The mixing matrix U can be parameterized by a CP-violating phase δ_{CP} and three mixing angles θ_{12}, θ_{13}, and θ_{23}. In the case of Majorana neutrinos, two additional phases are included, which are of no relevance for this work. Since U is non-diagonal, flavor changes are observed depending on the energy and propagation distance of a neutrino, which are commonly known as neutrino oscillations. The oscillations are described by the mass splittings, mixing angles, and the CP-violating phase [4].

For propagation through dense matter, the neutrino oscillations are modulated by interactions with electrons, which give rise to matter effects [5] such as the so-called *MSW effect* and *parametric enhancement* [6–10]. Depending on the NMO, these modulations arise mainly in the neutrino (NO) or anti-neutrino channel (IO) [11]. In measurements of solar neutrino oscillations, they were used to determine the ordering of the neutrino states ν_1 and ν_2 by finding $m_2 > m_1$. Moreover, these modulations can be observed for atmospheric neutrinos that undergo matter effects during their propagation through the Earth. In contrast to long-baseline accelerator experiments, the signature observed in IceCube is largely independent of the value of δ_{CP}, which allows for a complementary measurement of the NMO at higher energies, using atmospheric neutrinos [12].

Atmospheric neutrinos are produced in the Earth's atmosphere by interactions of cosmic rays with the nucleons of the air, generating mesons. These mesons decay generating electron and muon (anti-)neutrinos, which propagate through the Earth and can eventually be detected by an underground neutrino detector, such as IceCube [13]. The baseline of propagation through Earth can be inferred by measuring the incoming zenith angle of the neutrino. The highest-energy oscillation maximum arises at $E_\nu \sim 25$ GeV for vertically up-going neutrinos, moving to lower energy at shorter baselines towards the horizon. For energies above a few GeV, the oscillations are mostly driven by the parameters θ_{23} and Δm^2_{31}, which are therefore referred to as *atmospheric oscillation parameters* [4], while for vacuum only oscillations the value of θ_{13} is too small for any detectable effect. Considering matter effects, however, the effective value of θ_{13} under the right conditions can become sizeable, resulting in oscillation with electron flavors as shown in Fig. 2.

In atmospheric oscillations, the impact of the presence of matter arises mainly below $E_\nu \sim 15$ GeV. The strength of these matter effects depends on the Earth's matter profile, which we take as given by the Preliminary Reference Earth Model (PREM), shown in Fig. 1 [14].

The oscillation probabilities for muon-flavored atmospheric neutrinos and anti-neutrinos to be found in the flavor state $\alpha \in \{e, \mu, \tau\}$ for a given zenith angle θ_ν, and neutrino energy E_ν, are shown in Fig. 2. They are calculated

* e-mail: analysis@icecube.wisc.edu
[a] e-mail: justin.evans@manchester.ac.uk
[b] Earthquake Research Institute, University of Tokyo, Bunkyo, Tokyo 113-0032, Japan

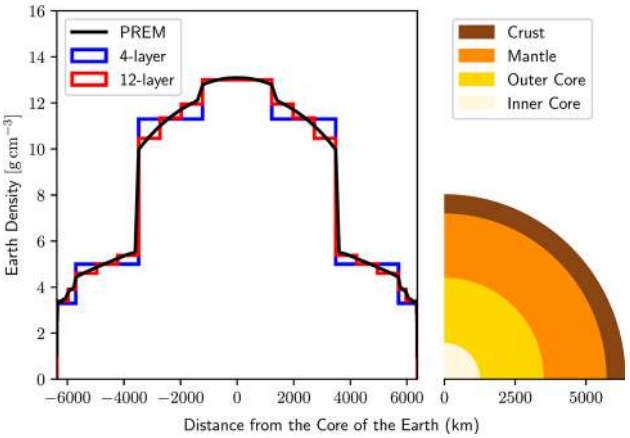

Fig. 1 Earth density profile, according to the *Preliminary Reference Earth Model* (PREM) and its approximation by 4- and 12-layers of constant density (commonly called PREM4 and PREM12, respectively) [14]

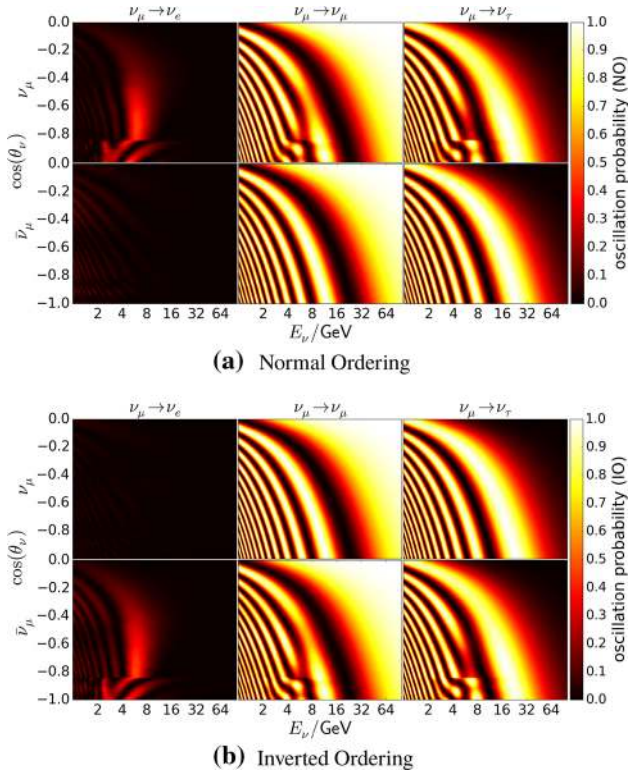

Fig. 2 Oscillation probabilities for an atmospheric ν_μ or $\overline{\nu}_\mu$ upon reaching the IceCube detector, as a function of the cosine of the zenith angle, θ_ν, and the energy, E_ν, of the neutrino, for the NO (**a**) and the IO (**b**) hypotheses. The probabilities are shown for the neutrino appearing as each of the three possible flavors, with the neutrino and anti-neutrino cases shown as the top and bottom rows in each panel. The dominant mixing of ν_μ and ν_τ is clearly visible, while the ν_e flavor is mostly decoupled, except for a small contribution from matter effects below $E_\nu \sim 15\,\text{GeV}$

with the *PROB3++* [15] package and the PREM12 approximation (cf. Fig. 1), which are consistently used throughout this work. Due to the Earth's geometry and its core-mantle structure, the visible modulations of atmospheric neutrino oscillations feature a clear zenith-dependence.

Note that the oscillation patterns for neutrinos and anti-neutrinos flip between the two orderings. Thus, the NMO can be determined by finding the enhancement in transition probabilities from matter effects either in the neutrino channel (NO) or anti-neutrino channel (IO). For detectors insensitive to distinguishing neutrinos from anti-neutrinos on an event-by-event level, the NMO still leads to a visible net-effect in the amplitude of the observed matter effects, because the atmospheric fluxes and the cross sections for neutrinos and anti-neutrinos differ [16,17]. These differences mean that atmospheric neutrinos are measured at higher rates than the corresponding anti-neutrinos. Due to this rate difference, the strength of observed matter effects in a combined sample of neutrinos and anti-neutrinos is increased in case of NO and decreased in case of IO, which is the main signature targeted in this work.

The determination of the NMO has important implications for searches for neutrinoless double-β decay, where the entire mass region allowed in the case of IO is in reach of the next generation of experiments [18,19]. The NMO must also be determined as part of the search for CP-violation in the lepton sector, where the sensitivity to δ_{CP} depends strongly on the ordering [20,21]. Therefore, a measurement of the NMO is targeted by several future long-baseline, reactor, and atmospheric neutrino experiments, such as DUNE [22], JUNO [23], PINGU [16,24], ORCA [25], and Hyper-Kamiokande [26]. Moreover, current neutrino experiments such as T2K [27], NOvA [28], and Super-Kamiokande [29] provide first indications of the NMO. Combining the results from several experiments, recent global fits prefer Normal over Inverted Ordering at $\sim 2-3.5\,\sigma$ with a small preference for the upper octant (i.e. $\sin^2(\theta_{23}) > 0.5$) [30–33].

2 The IceCube neutrino observatory

The *IceCube Neutrino Observatory* [13] is a $\sim 1\,\text{km}^3$ neutrino detector at the Geographic South Pole, optimized for detecting atmospheric and astrophysical neutrinos above $E_\nu \sim 100\,\text{GeV}$. It consists of 86 strings running through the ice vertically from the surface almost to the bedrock, carrying a total of 5160 Digital Optical Modules (DOMs) at depths between 1450 and 2450 m [34]. Each DOM houses a 10" photomultiplier tube and digitizing electronics, surrounded by a glass sphere [13,35,36].

In the center of the detector, some of these strings form a more densely instrumented volume called *DeepCore* [37]. It consists of 8 strings with an increased vertical density of DOMs with higher quantum-efficiency, surrounding one Ice-

Table 1 Overview of the main differences between the two NMO analyses in terms of the total number of observed events, the selection strategy, the reconstruction likelihood, the reconstructed energy range, the number of analysis bins (given as number of E_ν^{reco}, $\vartheta_\nu^{\text{reco}}$, PID bins), the background (atmospheric muon) description, the template generation, and the estimated fractions of the data sample from each contribution

	Data events	Selection strategy	Recon. likelihood	Energy range	Analysis binning	Background description	Template generation	Estimated contributions [%] $\text{CC}\nu_e/\text{CC}\nu_\mu/\text{CC}\nu_\tau/\text{NC}/\mu/$ noise
$\mathcal{A}$	43 214	High statistics	Hit-based	4–90 GeV	10, 10, 3	Simulation	KDEs	21.7 / 58.4 / 6.2 / 8.8 / 4.8 / 0.1
$\mathcal{B}$	23 053	Quality events	Charge-based	5–80 GeV	10, 5, 2	Data	Histograms	29.4 / 58.0 / 2.0 / 10.4 / 0.2 / –

Cube string. Due to the denser instrumentation and the higher quantum-efficiency DOMs, the DeepCore infill has a lower energy threshold than the surrounding IceCube array. The corresponding detection efficiency of DeepCore increases steeply between $\sim$ 3 GeV and $\sim$ 10 GeV and flattens for higher energies [13,37].

Neutrinos are detected by the Cherenkov emissions of their charged secondary particles, which are generated by Charged Current (CC) and Neutral Current (NC) interactions with the nucleons of the ice. In the case of CC muon-neutrino interactions, a hadronic cascade is initiated at the primary vertex, combined with an outgoing muon. The muon can propagate large distances through the detector, leading to an elongated shape of the energy deposition and thus of the Cherenkov light emission. Such events are called *track-like* signatures. In contrast, CC electron-neutrino, NC, and the majority of CC tau-neutrino interactions, do not produce a muon that can travel large distances. Instead, they initiate an electromagnetic and/or hadronic cascade that develops over a distance of a few meters. The light emission of this cascade is considerably smeared around the Cherenkov angle of the shower direction. Such events are called *cascade-like*. At low energies below a few tens of GeV, the separation of track- and cascade-like events becomes increasingly difficult, due to the short muon track and the coarse detector granularity. For oscillation measurements with DeepCore, this separation of track-like and cascade-like events is used to partially distinguish neutrino flavors [37].

For the analyses presented here, we use the Honda 3D atmospheric neutrino simulation [38], and the GENIE neutrino interaction generator [39] with KNO [40] and PYTHIA [41]. For quasielastic and resonance events, the axial masses are set to $M_A^{\text{qe}} = 0.99\,\text{GeV}$ and $M_A^{\text{res}} = 1.12\,\text{GeV}$, respectively. Simulation of the atmospheric muon background uses CORSIKA [42], with the Polygonato-Hörandel model of the muon energy spectrum [43]. Muons are propagated through the ice using PROPOSAL [44]; the propagation of all other particles is based on GEANT4 [45,46]. Cherenkov photons are propagated throught the ice using a GPU-based code [47]. More details of the simulation can be found in [48].

3 Data samples and reconstruction

In this work, two independent likelihood analyses are used to extract information about the NMO from DeepCore data. They are henceforth labelled Analysis $\mathcal{A}$ and $\mathcal{B}$, and the main differences between the two analyses are summarized in Table 1. Analysis $\mathcal{A}$ is designed to optimize the sensitivity to the NMO with DeepCore and considered the main result of this work, while Analysis $\mathcal{B}$ is designed to resemble the proposed PINGU analysis from [24], using only events that are fully-contained in the DeepCore detector, and is used as a confirmatory result here. Further details about Analyses $\mathcal{A}$ and $\mathcal{B}$ can be found in [49] and [50], respectively. The use of two independent analyses with partially complementary data sets gives great confidence in the quantitative conclusions of the analysis presented here and the treatment and impact of the systematic uncertainties.

The analyses are based on DeepCore data taken between May 2012 and April 2014, comprising a total livetime of 1006 (1022) days for Analysis $\mathcal{A}$ ($\mathcal{B}$). The difference in livetime arises from slightly different criteria on the stability of data acquisition. The data is run through two largely independent processing chains, where both samples are acquired by filtering the data in several successive steps of selection. These steps include the application of selection criteria on well-understood variables, as well as machine-learning methods, namely Boosted Decision Trees [51]. The selections are aiming for a reduction of the background of atmospheric muons and triggered noise, while maintaining a large fraction of well-reconstructed, low-energy neutrino events below $E_\nu \sim 100\,\text{GeV}$. Both samples are described in more detail in [48]. Compared to [48], the samples used in this work differ by the following modifications:

First, events with a reconstructed vertex outside the detector that enter from below are not vetoed in Analysis $\mathcal{A}$ using the lower part of the DeepCore detector, as it is done for downgoing and horizontal events using the surrounding IceCube detector. This increases the statistics at the expense of a reduced energy resolution for these uncontained events, especially at high energies. The loss in energy resolution is due to the unobserved fraction of deposited energy outside the

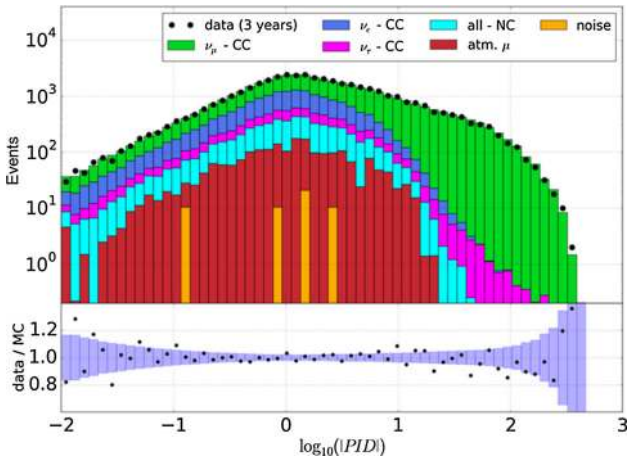

Fig. 3 The distribution of the particle identification variable for Analysis $\mathcal{A}$. The blue band on the data/MC ratio is the statistical uncertainty

detector volume. Second, the range of reconstructed energies considered is extended for both analyses compared to [48], from 56 to 90 GeV (80 GeV) for Analysis $\mathcal{A}$ ($\mathcal{B}$), allowing us to constrain nuisance parameters outside the strongest oscillation region. Third, both analyses use exclusively upgoing events (i.e. $\cos(\theta_\nu^{\text{reco}}) < 0$) to reduce the background from atmospheric muons.

The final samples are reconstructed with the same algorithm for Analyses $\mathcal{A}$ and $\mathcal{B}$ [48,49]. It is based on a likelihood function that links the number and the arrival times of the observed Cherenkov photons in all DOMs to a physics hypothesis. The physics hypothesis is given by the position and time of the interaction vertex, the neutrino direction, and the neutrino energy, which are the parameters of the likelihood optimization. The reconstruction is run separately for a starting track and a cascade-only hypothesis, where the starting track hypothesis features a cascade at the primary vertex with an additional parameter L for the length of an outgoing muon track. Since the track hypothesis allows for fitting the track length to $L = 0$, the 7-dimensional cascade-only-hypothesis is nested within the 8-dimensional track-hypothesis. The log-likelihood difference between track and cascade-only hypothesis is used as the flavor-separating variable, called *Particle Identification* (PID). Besides the reconstructed neutrino zenith angle θ_ν^{reco} and neutrino energy E_ν^{reco}, the PID is used as a third observable entering the likelihood analyses described in Sect. 4. The distribution of the PID variable for Analysis $\mathcal{A}$ is shown in Fig. 3.

In the reconstruction, the optimized likelihood function differs between the two analyses: For Analysis $\mathcal{B}$, the reconstruction likelihood is defined using the observed charge binned in time for each DOM as a proxy for the observed number of Cherenkov photons. Since some deviations were found between data and Monte Carlo in charge-related quantities, the likelihood was reformulated in a charge-independent way for Analysis $\mathcal{A}$, such that the charge amplitude information was removed and the only information used is whether a DOM is hit or not hit in multiple bins of time. In terms of the resolutions in reconstructed zenith angle θ_ν^{reco} and neutrino energy E_ν, the impact of the likelihood reformulation was found to be small. Moreover for Analysis $\mathcal{B}$, the impact of the charge mismatch is estimated to be small in comparison to the statistical uncertainty on the observed NMO.

After the data selection, the number of events in Sample $\mathcal{A}$ exceed the number of events in Sample $\mathcal{B}$ by a factor of 1.87, while providing similar resolutions in energy and zenith angle.

Note that for Analysis $\mathcal{B}$, the atmospheric muon background is estimated from data in an off-signal region, while for Analysis $\mathcal{A}$, it is obtained from Monte Carlo simulations (cf. Table 1). As a result, there is no *a priori* Monte Carlo prediction for the atmospheric muon contamination in Sample $\mathcal{B}$. However, the fraction of atmospheric muons is fitted in the analysis as discussed in Sect. 6. The contamination of triggered noise was found to be only $\lesssim 0.1\%$ for both samples. It is included into the likelihood fit for Analysis $\mathcal{A}$, while it is neglected for Analysis $\mathcal{B}$.

The final samples consist of CC muon neutrino, CC electron neutrino, CC tau neutrino, NC, and atmospheric muon events. These different components are called *contributions* in the following and are simulated separately in Monte Carlo except for the atmospheric muon contribution used in Analysis $\mathcal{B}$ that is parametrized from an off-signal data region.

The estimated fraction of the data samples from each contribution is shown in Table 1. These fractions are calculated using the best-fit values for all systematic parameters, discussed in Sect. 4.

4 Analyses

Both Analyses $\mathcal{A}$ and $\mathcal{B}$ use a binned likelihood method to determine the NMO by observing the signature from Fig. 2 in reconstructed variables. Since a separation of all flavors cannot be done with DeepCore, the PID is used to distinguish track- and cascade-like events, while neutrino energy and zenith angle are obtained from the reconstruction described in Sect. 3.

For both analyses, the binning is summarized in Table 1. For Analysis $\mathcal{B}$ only two PID bins are used to separate track- and cascade-like events, analogously to [16], while Analysis $\mathcal{A}$ uses three PID bins. This is motivated by the weak separation power at low energies, where the confidence in the separation can be taken into account by including an additional, intermediate PID bin. The binning in neutrino energy

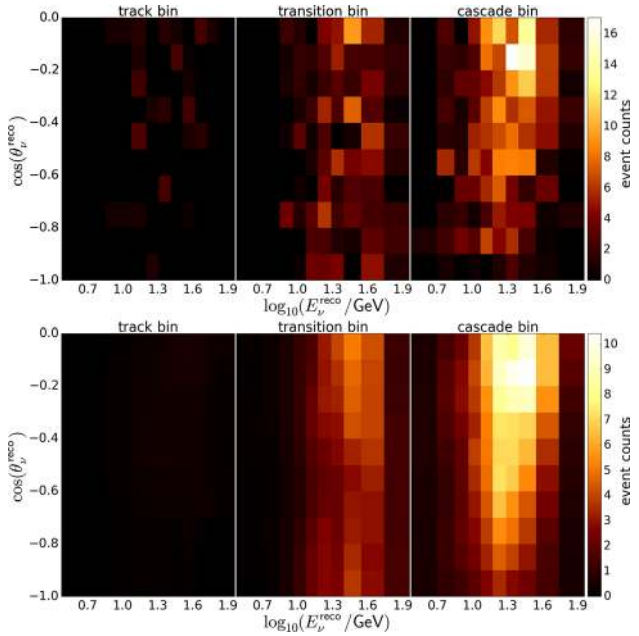

Fig. 4 Comparison of Monte Carlo template for atmospheric muons in Analysis $\mathcal{A}$, generated as histogram (top) and KDE (bottom): the latter one is used in the analysis, due to the reduced impact of limited Monte Carlo statistics

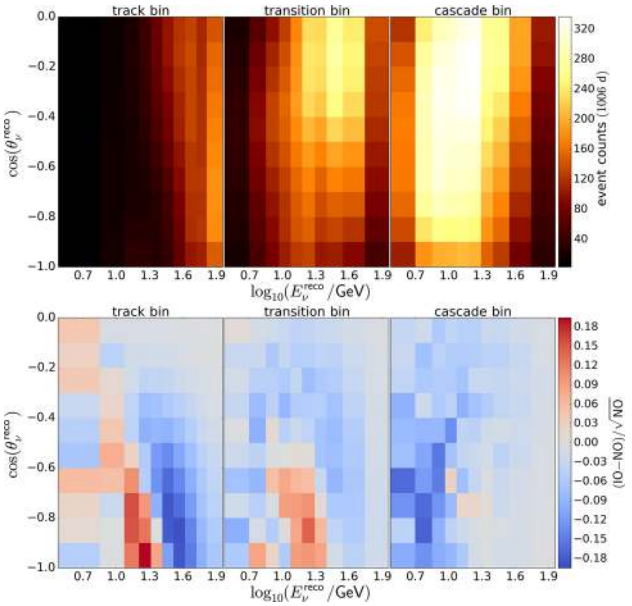

Fig. 5 Top: the distribution in PID, zenith angle, and neutrino energy for Analysis $\mathcal{A}$ that enters the likelihood calculation; bottom: corresponding signature of the NMO, given as expected pull on the bin content in case IO is observed but NO is tested, using Poissonian statistics

and zenith angle is chosen to be uniform in $\log_{10}(E_\nu^{\mathrm{reco}})$ and $\cos(\theta_\nu^{\mathrm{reco}})$ for Analysis $\mathcal{B}$. For Analysis $\mathcal{A}$, it is also uniform in $\cos(\theta_\nu^{\mathrm{reco}})$, while it is optimized in $\log_{10}(E_\nu^{\mathrm{reco}})$ to roughly follow the available statistics and maintain a large number of bins in the most interesting region at $E_\nu \sim 10\,\mathrm{GeV}$.

In Analysis $\mathcal{B}$, the binning is used to generate Monte Carlo distributions, called *templates*, in E_ν^{reco}, $\theta_\nu^{\mathrm{reco}}$, and PID for each contribution to the data sample, using histograms. In contrast, Analysis $\mathcal{A}$ applies an adaptive *Kernel Density Estimation* (KDE) method to produce these templates, which smooths the fluctuations from limited Monte Carlo statistics. These uncertainties arise mainly from the atmospheric muon template, where the available Monte Carlo statistics are similar to those from experimental data, due to the time-intensive simulation of atmospheric muons.

The KDE method is analogous to the one used in [52] and based on [53]. However, the method from [53] is extended by reflecting the KDE at the boundaries of the binned parameter space and integrating the resulting distribution to obtain a prediction for the bin content [54]. For the atmospheric muons, this is illustrated in Fig. 4, where the Monte Carlo template for atmospheric muons is generated with histograms (top) and the above mentioned KDE method (bottom). In the case of histograms, the fluctuations in the bin content, arising from limited Monte Carlo statistics, are clearly visible.

The uncertainties on the KDE prediction are estimated using *bootstrapping* for every contribution from Sect. 3 separately [55]. For each contribution, which consists of N MC events, events are drawn randomly from this sample, replacing the event each time so that it can be drawn again, until N events have been drawn. This new sample of N events is called a bootstrapped sub-sample, and from this a new KDE template is generated. This process is repeated several times and the uncertainty on each bin content in the original KDE template is estimated from the resulting distribution of bin contents in the bootstrapped samples.

For Analysis $\mathcal{A}$, the three-dimensional template obtained from the combination of all Monte Carlo contributions is shown in Fig. 5. Additionally, the expected pulls on each bin are shown in the case that the true ordering is inverted but the NO hypothesis is tested. This is used as a metric to visualize the signature of the NMO [16]. As can be seen in Fig. 5, the expected pulls between NO and IO are small, which already indicates the low sensitivity due to the limited resolution and statistics of DeepCore at energies $E_\nu \lesssim 15\,\mathrm{GeV}$.

Using these distributions, likelihoods are defined for both analyses. For Analysis $\mathcal{A}$, the negative log-likelihood LLH is given by

$$\mathrm{LLH} = \left[-\sum_{i\in\{\mathrm{bins}\}} \ln\left(\frac{p_i^{\mathrm{tot}}(N_i^{\mathcal{A}}, \mu_i^{\mathcal{A}}, \sigma_{\mu_i}^{\mathcal{A}})}{p_i^{\mathrm{tot}}(N_i^{\mathcal{A}}, N_i^{\mathcal{A}}, \sigma_{\mu_i}^{\mathcal{A}})} \right) \right] + \frac{1}{2}S, \quad (2)$$

where the term S is common to the likelihood of both analyses and will be defined after discussing the other terms. The term $p_i^{\mathrm{tot}}(N_i^{\mathcal{A}}, \mu_i^{\mathcal{A}}, \sigma_{\mu_i}^{\mathcal{A}})$ gives the probability of observing $N_i^{\mathcal{A}}$ events in bin i, if $\mu_i^{\mathcal{A}}$ events are expected. It is obtained by

a convolution of a Poissonian distribution and a narrow log-normal probability density function that describes the uncertainty $\sigma_{\mu_i}^{\mathcal{A}}$ on the Monte Carlo prediction μ_i. The uncertainty $\sigma_{\mu_i}^{\mathcal{A}}$ is obtained from a quadratic combination of the individual template uncertainties for every contribution, obtained from bootstraping.

Due to the KDE method used in Analysis $\mathcal{A}$, the dominant template uncertainties in the description of atmospheric muons are strongly reduced, such that the uncertainties on the total template are typically $\sim$ 10% of the Poissonian error expected from data fluctuations. Thus, for Analysis $\mathcal{A}$ these template uncertainties contribute only marginally to the following results.

For Analysis $\mathcal{B}$, the likelihood is adapted from [56], where a χ^2-value is calculated by quadratically combining the Poissonian error on the predicted bin content $\mu_i^{\mathcal{B}}$ with the uncertainty $\sigma_{\mu_i}^{\mathcal{B}}$ on the combined template of all contributions. It is given by

$$\chi^2 = 2\mathrm{LLH} = \sum_{i\in\{\mathrm{bins}\}} \frac{\left(N_i^{\mathcal{B}} - \mu_i^{\mathcal{B}}\right)^2}{\mu_i^{\mathcal{B}} + (\sigma_{\mu_i}^{\mathcal{B}})^2} + S, \tag{3}$$

where the labels are analogous to Eq. (2). Here, the uncertainties $\sigma_{\mu_i}^{\mathcal{B}}$ on the templates are estimated from the statistical error due to limited Monte Carlo and an uncertainty on the atmospheric muon template, estimated from off-signal data.

The dominant systematic uncertainties are included in both likelihood fits using nuisance parameters. These nuisance parameters comprise uncertainties in the atmospheric neutrino flux, the atmospheric oscillation parameters, the neutrino-nucleon cross sections, and the detector response. All systematic parameters are allowed to vary simultaneously and independently in the fit; we assume there are no correlations between the pulls on the various parameters. The parameters are listed in Table 2. To account for external constraints on these systematic parameters, Gaussian priors are included into the likelihood by the term S,

$$S = \sum_{s\in\{\mathrm{sys}\}} \left(\frac{s - s_0}{\sigma_s}\right)^2, \tag{4}$$

where the sum runs over all systematic parameters. For each parameter, the tested value s is compared to the expected baseline value s_0 with respect to its estimated uncertainty σ_s. The baseline value s_0 and width σ_s of each prior are identical for both analyses, and are stated in Table 2; the central value and the width are motivated by the provided references where possible. As indicated in Table 2, the prior for some parameters was removed in Analysis $\mathcal{B}$. Due to the small sensitivity to the NMO, the prior assumption was found to imply a preference on the NMO in case the true parameter value differs from the baseline value, which is avoided by removing the corresponding priors from the likelihood. Thus, no external knowledge is included on these parameters, allowing for larger deviations from the baseline value.

The parameters N_ν, N_{ν_e}, N_{NC}, and N_μ are used to vary the normalizations of the different contributions from Table 1. Thus, they account for uncertainties in interaction cross sections, the total neutrino and muon fluxes, the ν_e/ν_μ production ratio, and detection efficiencies.

Additional uncertainties on the neutrino fluxes predicted in [38] are modelled by the parameters γ_ν, $\sigma_\nu^{\mathrm{zenith}}$, and $\Delta(\nu/\bar{\nu})$. Here, γ_ν incorporates uncertainties in the neutrino energy spectrum, arising from flux, and cross section uncertainties, according to a reweighting of Monte Carlo events $\propto (E_\nu/\mathrm{GeV})^{\gamma_\nu}$, while $\sigma_\nu^{\mathrm{zenith}}$ and $\Delta(\nu/\bar{\nu})$ incorporate the dominant uncertainties from [58] in an *ad hoc* parametrization. The uncertainties on the production of atmospheric muons arising from the spectrum and compositions of the cosmic ray primary flux are represented by the parameter γ_μ. Note that γ_μ is only included as an uncertainty for Analysis $\mathcal{A}$, since the atmospheric muon template in Analysis $\mathcal{B}$ is estimated from data.

Uncertainties in neutrino-nucleon interactions are represented by the parameters M_A^{res} and M_A^{qe}, which model the axial mass of resonant and quasi-elastic interactions. Note that uncertainties on the cross section for deep inelastic scattering were also parametrized, but found to be negligible and therefore not included into the likelihood fit.

Detector uncertainties are modelled by the parameters ϵ_{opt}, $\epsilon_{\mathrm{lateral}}$, and $\epsilon_{\mathrm{head-on}}$, which describe the optical detection efficiency of the DOMs. The value of ϵ_{opt} gives the total detection efficiency per photon, relative to the baseline scenario. In contrast, the parameters $\epsilon_{\mathrm{lateral}}$ and $\epsilon_{\mathrm{head-on}}$ describe the dependence of the photon detection efficiency on the inclination angle of the incoming photon. Here, $\epsilon_{\mathrm{lateral}}$ changes the slope of the acceptance curve, while $\epsilon_{\mathrm{head-on}}$ controls the acceptance of very vertically upgoing photons independently. Besides actual uncertainties in the DOMs' detection efficiency, these parameters incorporate uncertainties with respect to the optical properties of the ice in the refrozen drill holes that surround the DOMs.

All of the systematic parameters mentioned above are described in more detail in [48]. Besides the parameters included in the fit, additional uncertainties have been investigated and tested for their possible effect on the analysis [49,50]. These parameters are the normalizations of sub-dominant experimental backgrounds (detector noise and event pile-up from coincident atmospheric muons), additional uncertainties on the optical properties of the ice, the oscillation parameters (θ_{12}, θ_{13}, Δm^2_{21}, and δ_{CP}), and Bjorken-x dependent uncertainties in the cross section for deep-inelastic neutrino-nucleon scattering. Two types of tests were performed to determine the impact of these parameters. In the first test, a parameter is injected into a MC fake dataset, shifted from its nominal value by $\pm$ 1σ in the case of

Table 2 Systematics treated as nuisance parameters in the likelihood analysis, including normalization (N), detector response (D), oscillation (O), flux (F), and neutrino-nucleon interaction (I) uncertainties. These parameters are discussed in more detail in [48]. The table gives the baseline value and, if the parameter is used with a prior in the likelihood, the standard deviation of the Gaussian prior, as well as the experimental best-fit values for both analyses and ordering hypotheses

Label	Type	Description of parameter	Baseline ± Prior	Analysis $\mathcal{A}$		Analysis $\mathcal{B}$	
				NO	IO	NO	IO
N_ν	N, F	Normalization of total neutrino template	1[ae]	0.83	0.84	0.98	0.99
N_{ν_e}	N, F	Normalization of ν_e flux before oscillations	1 ± 0.05[ad]	1.00	1.00	1.37	1.38
N_{NC}	N, I	Normalization of NC events	1 ± 0.2[a]	0.74	0.75	0.99	0.99
N_μ	N, F	Normalization of atmos. muon events	1[ae]	1.35	1.34	0.2%[c]	0.2%[c]
ϵ_{opt}	D	Overall optical efficiency [13]	1 ± 0.1[ad]	1.00	1.00	0.92	0.92
$\epsilon_{\mathrm{lateral}}$	D	Lateral dependence of optical efficiency [13]	0 ± 1[b]	0.68	0.68	− 0.46	− 0.46
$\epsilon_{\mathrm{head-on}}$	D	Head-on optical efficiency [13]	0[be]	−1.01	−1.01	− 2.00	− 1.92
$\Delta m^2_{31}/(10^{-3}\,\mathrm{eV}^2)$	O	Atmospheric mass-splitting	2.5(NO)/ − 2.43(IO)[e]	2.626	−2.511	2.462	− 2.348
$\sin^2(\theta_{23})$	O	Atmospheric neutrino mixing angle	0.455[e]	0.476	0.485	0.558	0.539
γ_ν	F	Neutrino spectral index unc. [38]	0.0 ± 0.1[d]	0.073	0.071	− 0.025	− 0.027
γ_μ	F	Atmospheric muon spectrum unc. [50,57]	0.0 ± 1.0[b]	0.04	0.04	–	–
$\sigma_\nu^{\mathrm{zenith}}$	F	Zenith-dependent unc. in $\nu/\bar{\nu}$ flux [58]	0.0 ± 1.0[bd]	−0.12	−0.11	− 0.86	− 0.89
$\Delta(\nu/\bar{\nu})$	F	Energy-dependent unc. in $\nu/\bar{\nu}$ ratio [58]	0.0 ± 1.0[b]	−1.03	−1.02	0.05	0.07
$M_A^{\mathrm{res}}/\mathrm{GeV}$	I	Axial mass unc. of resonant events [39]	1.12 ± 0.22	1.091	1.095	1.003	0.999
$M_A^{\mathrm{qe}}/\mathrm{GeV}$	I	Axial mass unc. of quasi-elastic events [39]	0.99 ± 0.25	0.862	0.867	0.881	0.888

[a] Relative to the nominal value of this parameter
[b] Parametrized with respect to the value and the uncertainty obtained from the provided reference
[c] Given as fraction of the total sample, since no Monte Carlo prediction exists to compare to
[d] No prior used for likelihood in Analysis $\mathcal{B}$
[e] Parameter allowed to vary freely (no prior) in both analyses

a detector systematic and by $\pm\ 3\sigma$ in the case of an oscillation parameter. This MC fake data is fit using the same MC set, but with the parameter in question fixed to its unshifted nominal value to assess whether the uncertainty in the systematic parameter can bias the measured ordering hypothesis. This test is repeated for MC fake datasets generated with both mass orderings. For none of these systematic or oscillation parameters is a bias observed in the measured preference for the mass ordering of more than 0.05σ. For the case of δ_{CP}, the value of $\delta_{\mathrm{CP}} = 180°$ is being used as the 'nominal' value for the MC fake dataset, with the value of $\delta_{\mathrm{CP}} = 270°$ injected into the MC used in the fit; a negligible (less than 0.01σ) bias in the ordering preference is observed. In the second test, the same shifted values as above are injected into the MC fake dataset, but now the parameter under test is allowed to vary in the fit. This allows us to determine if the inclusion of these parameters into the fit causes any loss of sensitivity to the mass ordering. None of the parameters in question cause a loss of sensitivity of more than 0.05σ; the inclusion of δ_{CP} reduces the sensitivity by less than 0.03σ. Since all the parameters described in this paragraph are shown to have no impact on the mass ordering sensitivity, or the potential to cause a bias, they have been set to their nominal values (in the case of oscillation parameters, to the NuFit [32] best-fit values), and are not included in the final fit in order to minimise the computing time required for the multi-parameter minimisation.

For Analysis $\mathcal{A}$ ($\mathcal{B}$), the negative log-likelihood from Eqs. (2) (3) is optimized. To do this, $\mathrm{LLH} \equiv -0.5\chi^2$ is used as the negative log-likelihood for Analysis $\mathcal{B}$. During this optimization, the first and the second octant in θ_{23} are fitted separately for both orderings, allowing all the parameters listed in Table 2 to vary, and the fit optimizing the LLH is taken as the best-fit for this ordering. The resulting difference, $2\Delta\mathrm{LLH}_{\mathrm{NO-IO}} \equiv \Delta\chi^2_{\mathrm{NO-IO}}$, between the NO and IO hypotheses is then calculated for both analyses.

Finally, $2\Delta\mathrm{LLH}_{\mathrm{NO-IO}}$ ($\chi^2_{\mathrm{NO-IO}}$) is used as a *test-statistic* (TS) in Sect. 6 for Analyses $\mathcal{A}$ ($\mathcal{B}$) to derive the experimental result from the fit to the data.

5 Sensitivity to the neutrino mass ordering

The determination of the Neutrino Mass Ordering is a binary hypothesis test, which requires the test of two non-nested hypotheses. This is different from most other applications in particle physics, where a general hypothesis $\mathcal{H}_G$ is tested against a specific one, $\mathcal{H}_S$, in the sense that the specific hypothesis is obtained for a certain realization of the parameters of $\mathcal{H}_G$. For such nested hypotheses, *Wilks' Theorem* is commonly used to derive sensitivities and to estimate limits on fitted parameters [59]. In contrast, Wilks' Theorem does not apply to the determination of the Neutrino Mass Ordering, since the discrete choice of Normal or Inverted Ordering is not related to the fixing of degrees of freedom [60].

Due to the subtleties involved in the statistical treatment and since a determination of the NMO is expected within the next decade, the correct method to quantify the preference is object of many discussions [60–62]. Here, two methods are used to estimate the sensitivity, which are described in the following.

The first method is a statistically rigorous analysis of the resulting likelihood values, using the obtained value of $2\Delta\mathrm{LLH}_{\mathrm{NO-IO}}$ as a TS. It derives the resulting sensitivity, given by the expected confidence in the determination of the NMO, from a frequentist coverage test. To do this, the data is fit with both ordering hypotheses giving a value for the TS and two sets of best-fit systematic parameters, η^{NO} and η^{IO}. These fits are called *fiducial fits* (FD) in the following.

From these parameters, the resulting best-fit templates are generated for NO and IO. Then, these templates are used to generate *Pseudo-Experiments* or *Pseudo-Trials* (PT)s by adding Poissonian fluctuations on the bin-contents, as expected in a real-world experiment; in this analysis, which has a sensitivity dominated by the statistical uncertainty, it is unnecessary to fluctuate each PT according to the systematic uncertainties. Afterwards, each PT is fitted with both ordering hypotheses, resulting in a new value for the $\mathrm{TS} = \Delta\chi^2_{\mathrm{NO-IO}} = 2\Delta\mathrm{LLH}_{\mathrm{NO-IO}}$. From these PTs, two distributions of the TS are obtained for the two sets of injected parameters η^{NO} and η^{IO}.

This process of creating PTs for η^{NO} and η^{IO} and fitting them with both hypotheses is repeated several times to estimate a TS distribution for each of the ordering hypotheses. The TS distributions for NO and IO are then used to estimate the analysis sensitivity, i.e. the expected p-values for the exclusion of each hypothesis. To do this, the fraction of PTs for NO (IO) that is to the right (left) of the median of the IO (NO) distribution is taken as the expected p-value for the exclusion of the NO (IO) hypothesis, if IO (NO) is the true ordering. This is sketched in Fig. 6 for two generic distributions.

The frequentist method is summarized as a flow-chart in Fig. 7. Note that this procedure is similar to the treatment of data, described in Sect. 6, where the experimental fit is used as fiducial fit to produce PTs. Unfortunately, the frequentist method is computationally very expensive. Thus, for performing more detailed parameter studies, a second, faster method is used.

The second method for deriving sensitivities is an *Asimov approach* adapted from [60]. Instead of generating PTs, the total MC template, with no Poissonian fluctuations, is fitted directly for both hypotheses. In the following, we refer to this MC template as the generated-ordering (GO) hypothesis, $\mathcal{H}_{\mathrm{GO}}$, where the GO can be either NO or IO. This is then fitted under assumptions of both hypotheses, NO and IO, where the

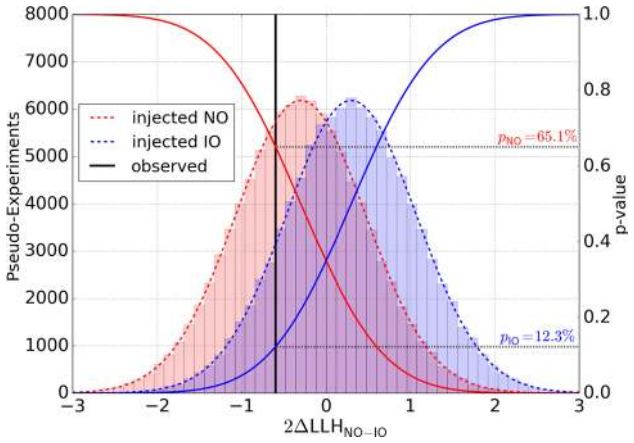

Fig. 6 Sketch of the frequentist method, using idealized distributions to illustrate the concepts. The red (blue) distribution will be derived from from PTs assuming the $\mathcal{H}_{NO}$ ($\mathcal{H}_{IO}$) hypothesis. The black vertical line represents a hypothetically observed value of ΔLLH_{NO-IO}. The resulting p-values (right, vertical axis) for the hypotheses are derived from the cumulative density distributions, marked as red (blue) solid lines for NO (IO)

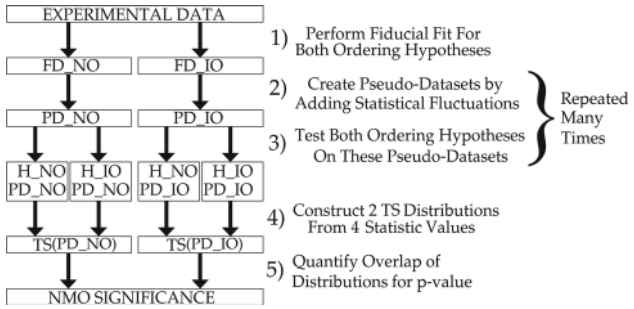

Fig. 7 Flow-chart representing the procedure of the frequentist method used to derive p-values for the NO and IO hypotheses. Abbreviations, as defined in the text, are FD (fiducial fits), PD (pseudo-dataset), H (hypothesis), and TS (test statistic)

hypothesis used in the fit is called the fitted-ordering (FO) hypothesis, $\mathcal{H}_{FO}$. The negative log-likelihood value obtained from the fit is $\overline{LLH}_{FO}(\mathcal{H}_{GO}) = 0$ if $\mathcal{H}_{FO} = \mathcal{H}_{GO}$ and $\overline{LLH}_{FO}(\mathcal{H}_{GO}) > 0$ otherwise, where the bars indicate that the values were obtained by injecting the template of the GO directly.

The resulting value of $2\overline{\Delta LLH}_{NO-IO}$ is assumed to be representative for the behavior obtained using PTs. The sensitivity to the generated ordering, n_σ^{GO}, in terms of one-sided Gaussian standard deviations is

$$n_\sigma^{GO} = \frac{\overline{\Delta LLH}_{NO-IO}(\mathcal{H}_{GO}) - \overline{\Delta LLH}_{NO-IO}(\mathcal{H}_{\widetilde{GO}})}{\sqrt{2\overline{\Delta LLH}_{NO-IO}(\mathcal{H}_{\widetilde{GO}})}}, \quad (5)$$

where $\widetilde{GO} \in \{IO, NO\}$ is the opposite hypothesis to GO, generated with the best-fit set of systematic parameters $\eta^{\widetilde{GO}} \in \{\eta^{IO}, \eta^{NO}\}$ corresponding to the set $\eta^{GO} \in$

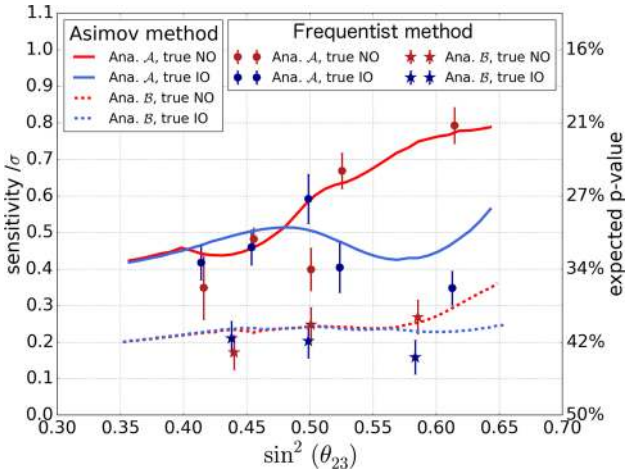

Fig. 8 Sensitivities of Analyses $\mathcal{A}$ and $\mathcal{B}$ to the NMO in terms of one-sided Gaussian sigmas (left vertical axis) and p-values (right vertical axis) derived by the Asimov-method (lines), and validated at certain values of $\sin^2(\theta_{23})$ using the frequentist method (markers). The statistical errors on the frequentist points arise from the finite number of PTs used due to the computationally intensive nature of the frequentist method

$\{\eta^{NO}, \eta^{IO}\}$ used for $\mathcal{H}_{GO}$. Note that the sensitivity $n_\sigma^{\widetilde{GO}}$ describes the expected p-value for the exclusion of the $\widetilde{GO}$ hypothesis in the case that the true ordering is the GO [60].

The choice of one- instead of two-sided Gaussian standard deviations is motivated by the fact that an experiment with no sensitivity to the NMO, i.e. if the two distributions for NO and IO in Fig. 6 were identical, would lead to a 50% chance of obtaining the correct ordering by random chance. This should not be misinterpreted as sensitivity and thus should give $n_\sigma^{NO,\,IO} = 0$, which is the case for one-sided but not two-sided Gaussians.

The resulting sensitivities for both methods are shown in Fig. 8, as a function of the true value of $\sin^2(\theta_{23})$. The blue and red lines indicate the result from the Asimov method for Analysis $\mathcal{A}$ (solid lines) and Analysis $\mathcal{B}$ (dashed lines). The sensitivities are validated at certain values of $\sin^2(\theta_{23})$ using the frequentist method, as indicated by the circle ($\mathcal{A}$) and star ($\mathcal{B}$) markers, where the uncertainties arise from the finite number of PTs.

As visible in Fig. 8, the resulting sensitivity is $< 1\sigma$ for both orderings and analyses. Moreover, Analysis $\mathcal{A}$ is more sensitive to the NMO than Analysis $\mathcal{B}$, which is due to the increased statistics, the additional bins in PID, energy and zenith, and the reduced impact from limited Monte Carlo statistics, due to the usage of KDEs in the generation of Monte Carlo templates.

Note that a characteristic shape is found for the $\sin^2(\theta_{23})$-dependence of $n_\sigma^{NO,\,IO}$, which is different for the NO and IO hypotheses. The observed features are similar to those found for the PINGU sensitivity in [16]. They arise from the interplay of the two independent octant fits for $\overline{LLH}_{GO}$ and

$\overline{\text{LLH}}_{\widetilde{\text{GO}}}$, used to calculate the values of $\overline{\Delta\text{LLH}}_{\text{NO-IO}}(\mathcal{H}_{\text{GO}})$ and $\overline{\Delta\text{LLH}}_{\text{NO-IO}}(\mathcal{H}_{\widetilde{\text{GO}}})$ in Eq. (5), where the preferred octant is not necessarily the true one in the case that $\widetilde{\text{GO}}$ is fitted. As a result, the behavior of $n_\sigma^{\text{NO, IO}}$ changes each time the octant is flipped for one of the two negative log-likelihood differences (ΔLLH) in Eq. (5).

The observed sensitivities for the Asimov method agree roughly with the PTs, while perfect agreement is not expected due to several approximations used in the Asimov-method [60]. However, the Asimov method is used as an estimator for the true sensitivity.

Note that for some observed values in Fig. 6 the p-values for both hypotheses can be small, in case the observed data agrees with neither the NO nor IO hypotheses. For example, this could be the case for $\Delta\text{LLH}_{\text{NO-IO}} > 2$ or $\Delta\text{LLH}_{\text{NO-IO}} < -2$, which is in the tail of both distributions in Fig. 6. In this case, the small p-value might lead to the wrong impression that the data clearly favors the alternative over the null hypothesis. To properly account for this, the p-values are combined into a CL_S-value,

$$\text{CL}_S^{\mathcal{A}/\mathcal{B}}(\mathcal{H}_{\text{TO}}) = \frac{p_{\mathcal{A}/\mathcal{B}}(\mathcal{H}_{\text{TO}})}{1 - p_{\mathcal{A}/\mathcal{B}}(\mathcal{H}_{\widetilde{\text{TO}}})}, \tag{6}$$

where TO is the tested ordering and $\widetilde{\text{TO}}$ is the opposite ordering hypotheses. This equation is taken from [63] where a more detailed discussion of its derivation can be found. Its value is limited to $\text{CL}_S \in [0, 1]$, where $\text{CL}_S \approx 1$ indicates no preference for one hypothesis over the other and $\text{CL}_S \approx 0$ indicates a strong disfavoring of the given hypothesis. The CL_S value can be interpreted as *confidence* in the result with a confidence level of $1 - \text{CL}_S$. More illustratively, the CL_S value describes how much less likely the observed value would occur under the disfavored hypothesis, compared to the favored one.

Finally, potential improvements of the sensitivity are tested for Analysis $\mathcal{A}$. By fixing individual and combinations of systematic parameters in the Asimov fit, the absolute gain in sensitivity from an improved understanding of systematic uncertainties is found to be small, except for the oscillation parameters. This is due to the weak NMO signature, which barely pulls the systematic parameters and thus is only weakly affected by fixing them. Instead, it is found that the sensitivity could be improved in the future by additional data statistics and improvements on the event reconstruction, which reduce the smearing-out of the NMO signature due to the low resolution in energy, zenith, and PID at the lowest energies [49].

6 Results

For both analyses, the experimental data is fitted with the likelihood method, described in Sect. 4. The data, along with the best-fit predictions, are shown for Analysis $\mathcal{B}$ in Fig. 9. The resulting best-fit values for all systematic parameters are shown in Table 2. The observed pulls are within the expected ranges for all parameters, taking statistical fluctuations and the uncertainties of the true value of each parameter into account. The corresponding values of the metric for the NO (IO) hypothesis are 2LLH = 293.38 (294.12) for Analysis $\mathcal{A}$ and $\chi^2 = 107.82$ (107.50) for Analysis $\mathcal{B}$. The metric is used as a *goodness-of-fit* estimator for the agreement of data and Monte Carlo by comparing these values to the expectation from PTs. The resulting p-values for Analyses $\mathcal{A}$ and $\mathcal{B}$ are $p_{\text{gof}}^{\mathcal{A}} = 43.5\%$ and $p_{\text{gof}}^{\mathcal{B}} = 11.0\%$, indicating the data to be well-described by the MC templates.

For Analyses $\mathcal{A}$ and $\mathcal{B}$, the observed values of the test-statistic are $2\Delta\text{LLH}_{\text{NO-IO}} = -0.738$ and $\Delta\chi^2_{\text{NO-IO}} = 0.3196$. Thus, the fits for the main result ($\mathcal{A}$) and the confirmatory result ($\mathcal{B}$) prefer NO and IO, respectively, while both results are compatible within their statistical uncertainties, i.e. both results have a test statistic within one unit of zero.

To estimate the corresponding p-values, PTs are generated with the best-fit parameters η^{NO} and η^{IO} from Table 2; for each PT, both ordering hypotheses are fitted. The resulting distributions of $\text{TS} = 2\Delta\text{LLH}_{\text{NO-IO}}$ and $\text{TS} = \chi^2_{\text{NO-IO}}$ are shown in Fig. 10. The experimentally observed value is indicated by the solid, vertical black line, indicating the preference for Normal over Inverted Ordering in Analysis $\mathcal{A}$ and Inverted over Normal Ordering in Analysis $\mathcal{B}$.

The resulting p- and CL_S-values for the main result are

$$p_{\mathcal{A}}(\mathcal{H}_{\text{NO}}) = 71.1\% \ (\text{CL}_S^{\mathcal{A}}(\mathcal{H}_{\text{NO}}) = 83.0\%), \tag{7}$$

$$p_{\mathcal{A}}(\mathcal{H}_{\text{IO}}) = 15.7\% \ (\text{CL}_S^{\mathcal{A}}(\mathcal{H}_{\text{IO}}) = 53.3\%), \tag{8}$$

while for the confirmatory result we find

$$p_{\mathcal{B}}(\mathcal{H}_{\text{NO}}) = 11.4\% \ (\text{CL}_S^{\mathcal{B}}(\mathcal{H}_{\text{NO}}) = 73.5\%), \tag{9}$$

$$p_{\mathcal{B}}(\mathcal{H}_{\text{IO}}) = 84.5\% \ (\text{CL}_S^{\mathcal{B}}(\mathcal{H}_{\text{IO}}) = 95.4\%). \tag{10}$$

In addition to testing the NMO with PTs, the likelihood is scanned across $\sin^2(\theta_{23})$ for the more sensitive Analysis $\mathcal{A}$ and both ordering hypotheses. The resulting scan is shown in Fig. 11, where the LLH is shown with respect to its global minimum. The vertical offset between the NO and IO curves indicates the preference for NO over IO, which is visible at all values of $\sin^2(\theta_{23})$. The observed minimum is in the lower octant, near $\sin^2(\theta_{23}) = 0.455$, for both orderings, while maximal mixing is separated from the best-fit point by only $2\Delta\text{LLH}_{\text{NO-IO}} = 0.128$ (0.681) for NO (IO). As a result, the preference for the lower octant is small, such that a substantial range of $\sin^2(\theta_{23}) > 0.5$ is still compatible with the observed data for NO and IO.

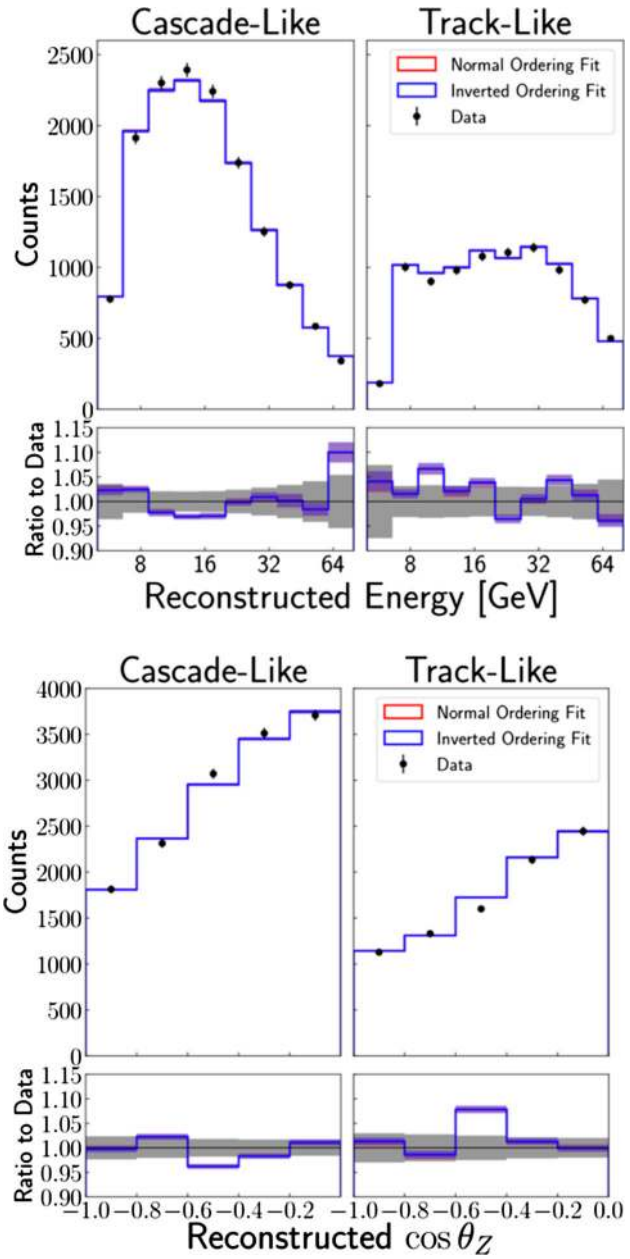

Fig. 9 The energy and zenith-angle distributions of the data from Analysis $\mathcal{B}$. Also shown are the best-fit simulations for both orderings, where the red and blue lines fall almost on top of each other

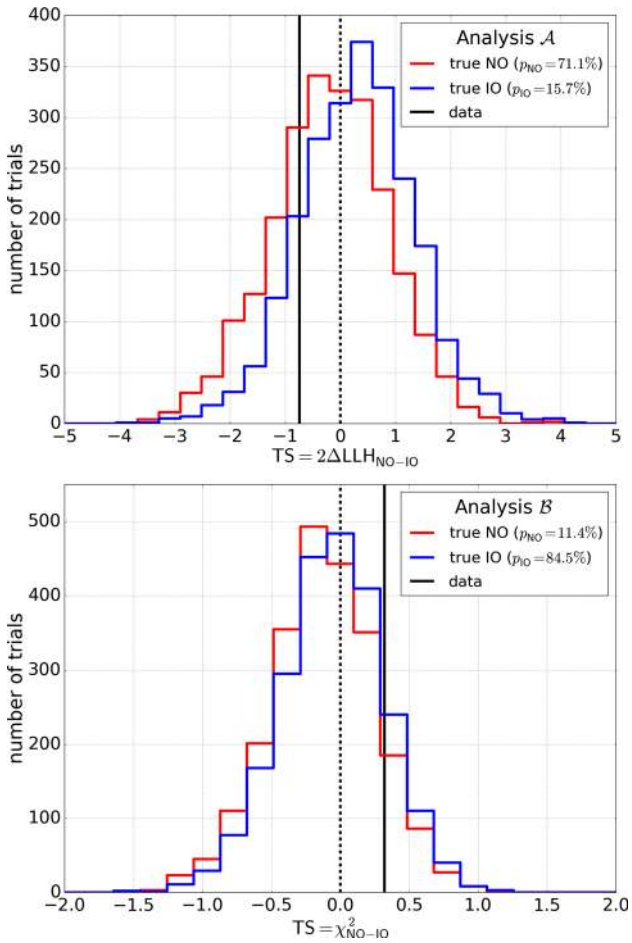

Fig. 10 Distribution of the TS from PTs, generated with the best-fit systematic parameters η^{NO} and η^{IO} from Table 2 for Analysis $\mathcal{A}$ (top) and Analysis $\mathcal{B}$ (bottom). The red and blue distributions are obtained for the NO and IO hypotheses, respectively, while the black, solid vertical line shows the observed value in data, giving the p-values for the NO and IO hypotheses stated in the legends

Note that the preference for NO over IO in Analysis $\mathcal{A}$ already indicates an observed preference for matter effects in data (cf. Sect. 1), i.e. a preference for matter effects over vacuum oscillations. To quantify this preference, the fit is repeated assuming vacuum oscillations. The resulting log-likelihood difference between matter effects (MA) and vacuum oscillations (VA) is $\Delta LLH_{MA-VA} = -0.869\ (-0.500)$ in case NO (IO) is assumed. Thus, matter effects are preferred over vacuum oscillations, independent of the assumption on the NMO. The p-values and CL_S-values that quantify the

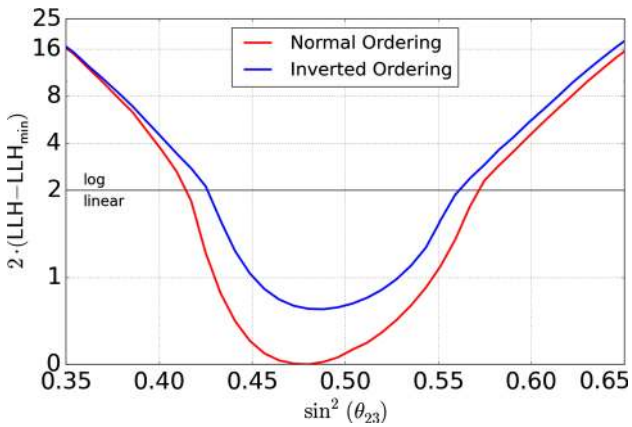

Fig. 11 The negative log-likelihood (LLH) as a function of $\sin^2(\theta_{23})$ for Analysis $\mathcal{A}$, relative to the global minimum LLH_{min}. The preference for NO over IO is visible over all the range of $\sin^2(\theta_{23})$ with the best-fit for both orderings being in the lower octant ($\sin^2(\theta_{23}) < 0.5$)

preference for matter effects (Mat) and vacuum oscillations (Vac) are

$$p(\mathcal{H}_{\mathrm{Mat}}|\mathcal{H}_{\mathrm{NO}}) = 62.3\%, \quad \mathrm{CL}_S(\mathcal{H}_{\mathrm{Mat}}|\mathcal{H}_{\mathrm{NO}}) = 71.0\%, \tag{11}$$

$$p(\mathcal{H}_{\mathrm{Vac}}|\mathcal{H}_{\mathrm{NO}}) = 12.3\%, \quad \mathrm{CL}_S(\mathcal{H}_{\mathrm{Vac}}|\mathcal{H}_{\mathrm{NO}}) = 32.6\%, \tag{12}$$

$$p(\mathcal{H}_{\mathrm{Mat}}|\mathcal{H}_{\mathrm{IO}}) = 53.2\%, \quad \mathrm{CL}_S(\mathcal{H}_{\mathrm{Mat}}|\mathcal{H}_{\mathrm{IO}}) = 68.4\%, \tag{13}$$

$$p(\mathcal{H}_{\mathrm{Vac}}|\mathcal{H}_{\mathrm{IO}}) = 22.2\%, \quad \mathrm{CL}_S(\mathcal{H}_{\mathrm{Vac}}|\mathcal{H}_{\mathrm{IO}}) = 47.4\%. \tag{14}$$

7 Conclusion

We have developed two independent likelihood analyses to demonstrate the extraction of the neutrino mass ordering from atmospheric neutrino data. We have applied these analyses to three years of IceCube DeepCore data. The first analysis aims for an optimized sensitivity with DeepCore, the second for an analysis chain as similar as possible to the proposed NMO analysis with PINGU [16]. The sensitivities were estimated with two independent methods. For the more sensitive, main analysis, the sensitivity was found to be $\sim 0.45 - 0.65\,\sigma$ (one-sided Gaussian), within the most interesting region close to maximum mixing ($\sin^2(\theta_{23}) \in [0.45, 0.55]$) for both orderings, while for the confirmatory analysis, the sensitivity was found to be $\sim 50\%$ smaller.

Due to the weak signature of the NMO in DeepCore, the sensitivity is found to be mostly unaffected by improvements in the understanding of systematic uncertainties. Instead, a future gain in sensitivity might come from additional statistics or potential improvements in the resolution of the event reconstruction.

The analyses presented here find the data to be fully compatible with both mass orderings. The main analysis observes a preference for NO over IO at $2\Delta\mathrm{LLH}_{\mathrm{NO-IO}} = -0.738$, which corresponds to a p-value of 15.3% ($\mathrm{CL_s} = 53.3\%$) for the IO hypothesis, based on the presented frequentist method. This result is in line with recently reported preferences for the NO by Super-Kamiokande [29], T2K [27], NOνA [28], MINOS [64], and recent global best fits [31,32]. However, it complements these results due to the higher energy range used for determining the NMO ($E_\nu \gtrsim 5\,\mathrm{GeV}$) and the fact that it is independent of the value of δ_{CP}. Finally, the data indicates a preference for matter effects over vacuum oscillations, independent of the assumption on the NMO.

The study presented here allows us to consider what future steps will allow a determination of the NMO with atmospheric neutrino data. Given the statistically-limited nature of this result, it is clear that a reduction of systematic uncertainties is not a priority, and we have performed studies to show that even the most optimistic reduction of systematic uncertainties can achieve at most a 10% improvement in the NMO sensitivity of this dataset [49]. The same study also showed that a removal of backgrounds (atmospheric muons and triggered noise) delivers at most a 5% improvement in sensitivity. In the coming years, a factor of four more statistics is expected from DeepCore (including both additional data and expected data-selection improvements), and this can result in a factor of two improvement in sensitivity. A more significant improvement that can be made is in the measurement resolutions: our studies [49] show that a 50% improvement in resolution on both neutrino direction and $\log_{10}(E_\nu)$ would produce a factor of two improvement in the sensitivity of this dataset. To achieve an NMO determination in a reasonable timescale, a final necessary improvement is a lowering of the neutrino energy threshold; this, along with the improved resolutions, can be achieved by the PINGU concept [16,24] that reduces the energy threshold to below 10 GeV to enable a 3σ determination of the NMO for even the least optimistic values of the oscillation parameters.

Besides the experimental result, the presented analyses provide a *proof-of-concept* for determining the NMO from matter effects in atmospheric neutrino oscillations with the IceCube Upgrade [65] or PINGU [16]. They test the full analysis chain by means of real DeepCore data and validate the understanding and treatment of systematic uncertainties, which are largely consistent with those that will be encountered by future IceCube extensions.

Acknowledgements USA – U.S. National Science Foundation-Office of Polar Programs, U.S. National Science Foundation-Physics Division, Wisconsin Alumni Research Foundation, Center for High Throughput Computing (CHTC) at the University of Wisconsin-Madison, Open Science Grid (OSG), Extreme Science and Engineering Discovery Environment (XSEDE), U.S. Department of Energy-National Energy Research Scientific Computing Center, Particle astrophysics research computing center at the University of Maryland, Institute for Cyber-Enabled Research at Michigan State University, and Astroparticle physics computational facility at Marquette University; Belgium – Funds for Scientific Research (FRS-FNRS and FWO), FWO Odysseus and Big Science programmes, and Belgian Federal Science Policy Office (Belspo); Germany – Bundesministerium für Bildung und Forschung (BMBF), Deutsche Forschungsgemeinschaft (DFG), Helmholtz Alliance for Astroparticle Physics (HAP), Initiative and Networking Fund of the Helmholtz Association, Deutsches Elektronen Synchrotron (DESY), and High Performance Computing cluster of the RWTH Aachen; Sweden – Swedish Research Council, Swedish Polar Research Secretariat, Swedish National Infrastructure for Computing (SNIC), and Knut and Alice Wallenberg Foundation; Australia – Australian Research Council; Canada – Natural Sciences and Engineering Research Council of Canada, Calcul Québec, Compute Ontario, Canada Foundation for Innovation, WestGrid, and Compute Canada; Denmark – Villum Fonden, Danish National Research Foundation (DNRF), Carlsberg Foundation; New Zealand – Marsden Fund; Japan – Japan Society for Promotion of Science (JSPS) and Institute for Global Prominent Research (IGPR) of Chiba University; Korea – National Research Foundation of Korea (NRF); Switzerland – Swiss National Science

Foundation (SNSF); UK – Science and Technology Facilities Council (STFC), part of UK Research and Innovation. The IceCube collaboration acknowledges the significant contributions to this manuscript from Martin Leuermann and Steven Wren.

Data Availability Statement This manuscript has associated data in a data repository. [Authors' comment: Data are publicly released on a regular basis by IceCube at https://icecube.wisc.edu/science/data/access/. The data used in this publication will be made available at this URL.]

Funded by SCOAP[3].

References

1. B. Pontecorvo, Sov. Phys. JETP **26**, 984 (1968)
2. V.N. Gribov, B. Pontecorvo, Phys. Lett. B **28**, 493 (1969)
3. Z. Maki, M. Nakagawa, S. Sakata, Prog. Theor. Phys. **28**, 870 (1962)
4. J.P. Yañez, A. Kouchner, Adv. High Energy Phys. **2015**, 271968 (2015). arXiv:1509.08404
5. L. Wolfenstein, Phys. Rev. D **17**, 2369 (1978)
6. S.P. Mikheyev, AYu. Smirnov, Sov. J. Nucl. Phys. **42**, 913 (1985)
7. S.P. Mikheyev, AYu. Smirnov, Il Nuovo Cimento C **9**, 17 (1986)
8. E. Akhmedov, Nuclear Physics B **538**, 25 (1999). arXiv:9805272
9. S.T. Petcov, Phys. Lett. B **434**, 321 (1998). arXiv:hep-ph/9805262
10. M.V. Chizhov, S.T. Petcov, Phys. Rev. D **63**, 073003 (2001). arXiv:hep-ph/9903424
11. O. Mena, I. Mocioiu, S. Razzaque, Phys. Rev. D **78**, 093003 (2008). arXiv:0803.3044
12. X. Qian, P. Vogel, Prog. Part. Nucl. Phys. **83**, 1 (2015). arXiv:1505.01891
13. M .G. Aartsen et al., [IceCube], J. Instrum. **12**, 03012 (2017a). arXiv:1612.05093
14. A. Dziewonski, D. Anderson, Phys. Earth Planet. Interiors **25**, 297 (1981)
15. R. Wendell et al., http://www.phy.duke.edu/~raw22/public/
16. M. G. Aartsen et al., [IceCube PINGU], Letter of Intent: The Precision IceCube Next Generation Upgrade (PINGU) (2014). arXiv:1401.2046v2.
17. E.K. Akhmedov, S. Razzaque, AYu. Smirnov, J. High Energy Phys. **2013**, 82 (2013). arXiv:1205.7071
18. C. Patrignani et al., Particle Data Group. Chin. Phys. **40**, 100001 (2016)
19. A. Giuliani, The Mid and Long Term Future of Neutrinoless Double Beta Decay, talk at XXVIII international conference on neutrino physics and astrophysics, 4–9 June 2018, Heidelberg, Germany (2018). https://doi.org/10.5281/zenodo.1286915
20. C. Hagedorn, R.N. Mohapatra, E. Molinaro, C.C. Nishi, S.T. Petcov, Int. J. Mod. Phys. A **33**, 1842006 (2018). arXiv:1711.02866
21. G. Branco, R.G. Felipe, F. Joaquim, Rev. Mod. Phys. **84**, 515 (2012). arXiv:1111.5332
22. B. Abi et al., [DUNE], The DUNE Far Detector Interim Design Report Volume 1: Physics, Technology and Strategies (2018). arXiv:1807.10334
23. F. An et al., [JUNO], J. Phys. **G 43**, 030401 (2016). arXiv:1507.05613
24. M .G. Aartsen et al., [IceCube-Gen2], J. Phys. **G 44**, 054006 (2017). arXiv:1607.02671
25. S. Adrián-Martínez et al., [KM3NeT], J. Phys. **G 43**, 084001 (2016). arXiv:1601.07459
26. K. Abe et al., [Hyper-Kamiokande], PTEP **2018**, 063C01 (2018). arXiv:1611.06118
27. K. Abe et al., [T2K], Phys. Rev. Lett. **121**, 171802 (2018). arXiv:1807.07891
28. M .A. Acero et al., [NOvA], Phys. Rev. **D 98**, 032012 (2018). arXiv:1806.00096
29. Y. Hayato [Super-Kamiokande], Atmospheric Neutrino Results from Super-Kamiokande, talk at XXVIII International Conference on Neutrino Physics and Astrophysics, 4–9 June 2018, Heidelberg, Germany, (2018). https://doi.org/10.5281/zenodo.1298567
30. F. Capozzi, E. Lisi, A. Marrone, A. Palazzo, Prog. Part. Nucl. Phys. **102**, 48 (2018)
31. I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, I. Martinez-Soler, T. Schwetz, NuFIT 3.2 (2018). http://www.nu-fit.org
32. I. Esteban, M.C. Gonzalez-Garcia, M. Maltoni, I. Martinez-Soler, T. Schwetz, J. High Energy Phys. **2017**, 87 (2017). arXiv:1611.01514
33. P.F. de Salas, D.V. Forero, C.A. Ternes, M. Tortola, J.W.F. Valle, Phys. Lett. B **782**, 633 (2018). arXiv:1708.01186
34. M.G. Aartsen et al., IceCube. Nucl. Instrum. Meth. **A711**, 73 (2013). arXiv:1301.5361
35. R. Abbasi et al., IceCube. Nucl. Instrum. Meth. **A618**, 139 (2010). arXiv:1002.2442
36. R. Abbasi et al., IceCube. Nucl. Instrum. Meth. **A601**, 294 (2009). arXiv:0810.4930
37. R. Abbasi et al., IceCube. Astropart. Phys. **35**, 615 (2012). arXiv:1109.6096
38. M. Honda, M. Sajjad Athar, T. Kajita, K. Kasahara, S. Midorikawa, Phys. Rev. **D92**, 023004 (2015). arXiv:1502.03916
39. C. Andreopoulos et al., Nucl. Instrum. Meth. A **614**, 87 (2010). arXiv:0905.2517
40. Z. Koba, H.B. Nielsen, P. Olesen, Nucl. Phys. B **40**, 317 (1972)
41. T. Sjostrand, S. Mrenna, P.Z. Skands, JHEP **05**, 026 (2006). arXiv:0603175
42. D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz, T. Thouw, Tech. Rep. FZKA 6019 (1998). http://inspirehep.net/record/469835/files/FZKA6019.pdf
43. J.R. Hörandel, Astropart. Phys. **19**, 193 (2003)
44. J .H. Koehne, K. Frantzen, M. Schmitz, T. Fuchs, W. Rhode, D. Chirkin, J. Becker Tjus, Comput. Phys. Commun **184**, 2070 (2013)
45. S. Agostinelli et al., GEANT4. Nucl. Instrum. Meth. **A506**, 250 (2003)
46. L. Radel, C. Wiebusch, Astropart. Phys. **38**, 53 (2012). arXiv:1206.5530
47. C. Kopper et al., https://github.com/claudiok/clsim
48. M .G. Aartsen et al., [IceCube], Phys. Rev. **D99**, 032007 (2019). arXiv:1901.05366
49. M. Leuermann, Doctoral thesis, RWTH Aachen University (2018). https://publications.rwth-aachen.de/record/751704
50. S. Wren, Doctoral thesis, University of Manchester (2018). https://www.research.manchester.ac.uk/portal/files/66046203/FULL_TEXT.PDF
51. Y. Freund, R.E. Schapire, J. Comput. Syst. Sci. **55**, 119 (1997)
52. M .G. Aartsen et al., [IceCube], Astrophys. J. **833**, 3 (2016). arXiv:1607.08006
53. B. Wang, X. Wang, Bandwidth selection for weighted kernel density estimation (2007). arXiv:0709.1616

54. D.B.H. Cline, J.D. Hart, Statistics **22**, 69 (1991)
55. B. Efron, R. Tibshirani, Stat. Sci. **1**, 54 (1986)
56. M .G. Aartsen et al., [IceCube], Phys. Rev. Lett. **120**, 071801 (2018). arXiv:1707.07081
57. J.J. Evans, D. Garcia-Gamez, S.D. Porzio, S. Söldner-Rembold, S. Wren, Phys. Rev. D **95**, 023012 (2017). arXiv:1612.03219
58. G.D. Barr, T.K. Gaisser, S. Robbins, T. Stanev, Phys. Rev. D **74**, 094009 (2006). arXiv:0611266
59. S.S. Wilks, Ann. Math. Stat. **9**, 60 (1938)
60. E. Ciuffoli, J. Evslin, X. Zhang, J. High Energy Phys. **1401**, 095 (2014). arXiv:1305.5150
61. M. Blennow, P. Coloma, P. Huber, T. Schwetz, J. High Energy Phys. **1403**, 028 (2014). arXiv:1311.1822
62. X. Qian, A. Tan, W. Wang, J.J. Ling, R.D. McKeown, C. Zhang, Phys. Rev. D **86**, 113011 (2012). arXiv:1210.3651
63. X. Qian, A. Tan, J.J. Ling, Y. Nakajima, C. Zhang, Nucl. Instrum. Meth. A **827**, 63 (2016). arXiv:1407.5052
64. A. Aurisano [MINOS/MINOS+], talk at XXVIII International Conference on Neutrino Physics and Astrophysics, 4–9 June 2018, Heidelberg, Germany, (2018). https://doi.org/10.5281/zenodo.1286760
65. U. Katz, talk at XXVIII International Conference on Neutrino Physics and Astrophysics, 4–9 June 2018, Heidelberg, Germany, (2018). https://doi.org/10.5281/zenodo.1287686