UNIVERSITY OF COPENHAGEN

MSc in Physics

Sensitivity to the Neutrino Mass Ordering in the IceCube Upgrade

Amalie Beate Albrechtsen

Supervised by D. Jason Koskinen

May 2023

Amalie Beate Albrechtsen

Sensitivity to the Neutrino Mass Ordering in the IceCube Upgrade

MSc in Physics, May 22nd, 2023

Supervisor: D. Jason Koskinen

University of Copenhagen

Faculty of Science

Niels Bohr Institute

Blegdamsvej 17

2100 Copenhagen Ø

"*We warn the reader that there is no universal convention for the term 'confidence level'.*"

The Review of Particle Properties, 1986
Quoted by R. J. Barlow 1989

"*Niels Bohr supposedly said that if quantum mechanics did not make you dizzy then you did not really understand it. I think that the same can be said about statistical inference!*"

Robert D. Cousins, 1994

Preface

In working with this thesis it has become clear to me that some care must be taken in the colors chosen to present the findings. The common choice of red for Normal Ordering and blue for Inverted Ordering has guided me to find a color palette that not only captures the physics but also highlights the beauty of the data involved. The colors in this thesis are therefore inspired by Pantone's New York Fashion Week Spring/Summer 2023 color palette [3].

Acknowledgements

I would like to extend my sincere gratitude to my supervisor Jason Koskinen. I have immensely enjoyed stepping into the world of IceCube neutrino physics once again, and it has been a pleasure to work with you. I never expected that I would come to think of myself as an expert, or even just knowledgeable, on statistics, but somehow in working with you I have grown accustomed to be confident in my skills. Thank you for always pushing for understanding, and thank you for always being understanding.

My deepest and most heartfelt thank you should also go to the entire IceCube group at NBI. You lovely, warm, funny, and crazy people almost made me consider wanting a PhD. To Tom and Tania without whom no one would ever get anything done, to Markus for his excellent questions, to Kathrine for always making me feel welcome, to Jorge, Moust, and Linea for much needed coffee breaks and office gossip, and to Clotilde and James for their support and presence in my life even after having left.

To the people around me that bring me joy and meaning.
Thank you for everything.

Copenhagen, May 2023

Author's Contribution

The work presented in this thesis is a product of work done by the author and other members of the IceCube Collaboration. For clarity, the contributions of the author will be stated explicitly here.

The statistical considerations of Chapter 4 are based on previous work in regards to the Neutrino Mass Ordering, but the discussion of the conceptual meaning of the fundamental approaches in relation to the analysis at hand and the presented correction to the Gaussian approximation has been made by the author, alongside the investigation into the θ_{23} octant dependency. The statistical method for two discrete non-nested hypotheses presented in Chapter 4 was implemented in the analysis software PISA used by IceCube, in collaboration with Tom S. Stuttard (NBI). The obtained sensitivities are based on simulation of the IceCube Upgrade Detector by James V. Mead (prev. NBI) and an event-selection developed by Jorge Prado (NBI), Kayla L. DeHolton, and Jan Weldert (PennState) but computed by the author.

Abstract

The ordering of the neutrino mass states is one of the remaining features of the Standard Model not yet explained. Insight into whether the third mass state is the heaviest (Normal Ordering) or the lightest (Inverted Ordering) will, among other things, impact the possibility of determining the charge-parity violating phase δ_{CP}. Matter effects to atmospheric neutrino oscillations below $\sim 15\,\mathrm{GeV}$ depend on the Neutrino Mass Ordering (NMO), and such low energy effects are expected to be measurable by the upcoming extension to the IceCube detector, the IceCube Upgrade.

This thesis studies the expected sensitivity to the NMO for the IceCube Upgrade planned to be deployed in the Antarctic summer 2025-26. The non-nested nature of the ordering requires statistics beyond the standard methods, and different frequentist approaches are investigated. From data simulated according to the expected Upgrade geometry and efficiency, median sensitivities in a Gaussian approximation for both orderings are reported within the expected global fit range of θ_{23}. From pseudo-experiments, these sensitivities are confirmed at one value of θ_{23}. Only one of the sensitivity approaches studied properly accounts for a degeneracy between the NMO and the octant of the atmospheric mixing angle θ_{23}, and the commonly used value of $\sqrt{2LLH}$ is found to overestimate the sensitivity by to 1σ. It is found that even though, a frequentist approach cannot yield a determination of the neutrino mass ordering, the IceCube Upgrade is expected to be able to reject the wrong ordering at 3σ within three years, when detector uncertainties are not considered. The findings suggest that future measurements from the IceCube Upgrade will significantly improve the current global sensitivity of the NMO.

Contents

1 Introduction

Particle physics is engaged with the most fundamental building blocks of our universe, from the Higgs mechanism providing mass to particles, to the electromagnetic interaction enabling us to live in a digital and electrified world. The search for *explanations* has always been the cornerstone of science and physics, and as the knowledge we have built has accumulated we have been able to push the boundaries of our understanding ever further.

At the frontier of particle physics, we find the problems for which we do not yet have a satisfactory explanation, a substantial set of which are somehow linked to the properties of the *neutrino*. Initially proposed to explain non-conservation of energy in beta-decays, the neutrino later proved to exhibit a multitude of new and non-intuitive physics. In the Standard Model, describing particle physics, the presence of neutral leptons, like the neutrinos, is not prohibited, but such leptons would *have to be* massless to be fully described by the theory. The ability for neutrinos to mix has however proven that they are massive, with masses constrained by cosmological measurements to be *tiny* compared to all other known matter [4].

Neutrinos exhibit what is known as *flavor oscillations* where a neutrino produced as a certain type (flavor) may interact as a different type at a later time [5]. This is possible because the neutrino states which are allowed to *interact* with other particles and the neutrino states that describe the time *evolution* of the neutrinos are not identical. Precise measurements of the structure of these oscillations provide a unique insight into the properties of the fundamental neutrino states. The differences between the masses of the neutrino states (ν_1, ν_2, ν_3) are well determined in *size*, as the *splitting* of the evolution states affects the phase of the resulting oscillations. Only the sign of the smaller mass difference Δm^2_{12} has however been measured by the time of this project [6]. Whether the larger mass squared difference Δm^2_{32} is positive or negative determines the *ordering* of the mass states which is said to be Normal for $\Delta m^2_{32} > 0$ and Inverted for $\Delta m^2_{32} < 0$. This is known as the Neutrino Mass Ordering (NMO) and is the subject of the analysis presented in this project.

Determination of the NMO will provide crucial information to the efforts of explaining *how* neutrinos get their masses and is also linked to the possible *charge-parity* violation in the lepton sector, with implications for the explanation of why we live in a matter

and not *anti*-matter dominated universe. The NMO signal has been a sub-leading contribution to the current generation of experiments and no strong preference for either the Inverted (IO) or Normal (NO) ordering has been found. The next generation of neutrino experiments will however provide the energy-resolution required for a measurement of the NMO.

One such next-generation detector is the IceCube Upgrade [7]. Planned for deployment in the Antarctic summer 2025-26, the Upgrade is an extension of the current IceCube detector located at the South Pole [8]. Spanning a volume of 1 cubic-kilometer of ice, the IceCube detector measures atmospheric neutrinos created all around the globe and provides measurements of neutrinos with energies from $\mathcal{O}(\mathrm{GeV})$ to $\mathcal{O}(\mathrm{PeV})$, and path lenghts from the thickness of the atmosphere to the diameter of Earth. With the Upgrade, the full detector is expected to achieve energy resolution in the GeV-range required for measurements of the Neutrino Mass Ordering which manifests itself as modulations to the oscillation patterns experienced either by neutrinos or anti-neutrinos that have traversed the dense Earth. This project will explore to what level the IceCube Upgrade will be able to separate the two mass orderings.

The true mass ordering can *either* be NO or IO, demanding special care in the considerations of the statistics employed for its determination, must be taken. Leaving the well-known comforts of Wilk's Theorem [9] brings new and yet unresolved issues, as the impact of parameters describing the oscillation patterns can heavily impact the signature of the two orderings in the detector.

Even though the Upgrade Detector will not be deployed for a few more years, studies into its potential are crucial at this time. Such studies will both set the expectations for future measurements and guide the efforts in the coming time to areas of impact on the NMO analysis. In this project, the theoretical background for measuring the signature of the NMO is described, including the theory of neutrino interactions and oscillation patterns in matter. The IceCube Upgrade and its ability to measure atmospheric neutrinos, and the impact of the NMO on such a measurement, will then be presented including the analysis tools involved in the simulation of events. From there, the statistics of an NMO analysis is discussed and a method accounting for the *composite and non-nested* [10] nature of the hypotheses is described. Applying these methods to simulated data of the IceCube Upgrade, sensitivities are obtained by applying a Gaussian approximation and this approximation's validity is tested at one point in the parameter space of the analysis. Different statistical interpretations of the sensitivity are compared and the impact of the atmospheric oscillation angle θ_{23} is investigated. The results will be discussed in relation to current sensitivities to the NMO from IceCube and other experiments.

2 Particle Physics Theory

Neutrinos are the lightest and most un-intuitive of the elementary particles. They exhibit strange quantum mechanical behavior and they interact with other constituents of the Universe only to the very slightest degree. They are therefore naturally of the greatest interest to particle physicists of today. Simply the fact that neutrinos have mass slams the door wide open into physics beyond the standard model, which is otherwise one of the most successful physical theories. Because neutrinos are strange and interesting in ways, that can cover full careers, this project will only deal with one small corner of the picture that is neutrino physics, namely the Neutrino Mass Ordering (NMO).

In order to fully appreciate the subtleties that go into being able to measure the NMO, this chapter will deal with the basic principles regarding neutrino interactions, neutrino oscillations and importantly, how the Earth's density allows us to make statements about the masses of the neutrinos. Because that is in essence what the ordering is - measuring the masses of three different particles. These particles are the *mass* states of the neutrinos and their ordering can either be *Normal* or *Inverted*, depending on whether the third mass state is the heaviest or the lightest, respectively.

2.1 The Standard Model

The Standard Model very successfully describes the building blocks of our Universe with fundamental symmetries and symmetry-breakings, without however providing *quite* satisfactory explanations for everything.

The particles this project evolves around, the neutrinos, are the neutral part of the lepton doublets also containing the charged leptons e, μ, τ. Though electrically neutral, the neutrinos have weak isospin and weak hypercharge enabling them to interact with the force-carriers of the weak interaction, the $W^{\pm}$ and Z bosons, and thereby also with both charged leptons and quarks [5]. From the observation of neutrino oscillations it is known that neutrinos have non-zero masses, because oscillations would otherwise not be possible. The mass of the neutrinos are however not explained by the Standard Model and neither is it predicted whether neutrinos are their own

anti-particles and therefore *Majorana* particles [11]. Even though the Standard Model does not fully explain the neutrino it does provide the frame from which we can study the nature of these evasive, light, and oscillating particles.

2.1.1 The Weak Interaction and Chirality

Neutrinos can interact only via the weak force, which is mediated by the massive $W^{\pm}$ and Z bosons. Even though the weak interaction is called *weak*, the dimensionless coupling constant α_w is actually larger than the electromagnetic one, but the large masses of the bosons, compared to the massless photon, causes weak interaction decay rates to be suppressed by a factor of q^4/m_w^4 [5].

The weak interaction does not conserve parity and therefore has a different mathematical structure than both Quantum Electro Dynamics (QED) and Quantum Chromo Dynamics (QCD), which are both parity conserving interactions. The parity violating nature of the weak interaction was found by observing the angular distribution of electrons from nuclear β-decay, which showed a suppression of electrons in one direction over another. The form of the interaction vertex can be found to be of the type vector minus axial (V-A) and includes the left-handed *chiral* projection operator $P_L = \frac{1}{2}(1 - \gamma^5)$. All particles in the Standard Model can be described as four-vector spinors, which can be expressed in the base of chirality - coinciding with helicity in the ultrarelativistic limit. Helicity is the projection of a particle's spin onto the direction of its momentum, but unfortunately such an intuitive expression of chirality is only valid for ultra-relativistic particles. The V-A structure of the weak interaction allows *only* left-handed (LH) and right-handed (RH) chiral states of a particle to take part in the interaction, and the weak interaction is therefore said to be maximally parity violating.

This "handedness" of the weak interaction has some interesting effects on the interactions of neutrinos. Since neutrinos are almost massless, they are ultrarelativistic and helicity and chiral states are identical [6]. Since angular momentum is conserved in *all* particle interactions, this greatly limits the number of allowed neutrino interactions. One such example is the production of neutrinos from cosmic rays consisting mainly of $\pi^{\pm}$ decays. The decay, $\pi^+ \rightarrow l^+ + \nu_l$, can, because of the relative masses of the pion and the charged leptons, only produce electrons or muons and not the heavier tau-leptons. Because the weak-coupling constant is the same for all lepton-flavors, one would expect the ratio of decays to electrons and muons to be one. This is not the case as the weak decay will always produce a left-handed chiral, and therefore also helicity, neutrino. Because the pion is a spin-0 particle, the spin of the charged lepton *must* be in the opposite direction to the neutrino's and therefore the anti-lepton

should also be in a LH helicity state. Since the weak interaction only allows for LH chiral particles and RH chiral anti-particles, this means that the charged lepton has to be a RH chiral state projected onto the LH helicity state. As helicity and chirality coincide in the ultra-relativistic limit, the amount of LH helicity in a RH chiral state is larger for less relativistic and more massive particles. This explains why the branching ratio of the decay rates is measured to be $\frac{\Gamma(\pi^+ \to e^+ \nu_e)}{\Gamma(\pi^+ \to \mu^+ \nu_\mu)} = 1.2 \times 10^{-4}$ [12, 13].

2.1.2 Neutrino Scattering

Neutrino interactions involving a $W^\pm$ boson is known as Charged Current (CC) interactions as they can change the "flavor charge" of the involved particles, whereas interactions involving a Z boson are known as Neutral Current (NC) interactions as they do not exchange any charge.

For the case of IceCube, the target material is ice, and depending on the energy of the incoming neutrino, different interactions will dominate the cross-section. For high energy neutrinos ($E_\nu > 10\,\mathrm{GeV}$), the dominant interaction is Deep Inelastic Scattering (DIS), below $100\,\mathrm{MeV}$ the cross-section is dominated by nuclear processes, and in the intermediate region multiple processes contribute [14]. For oscillation measurements with IceCube, we are interested in energies from $1\,\mathrm{GeV}$ to $300\,\mathrm{GeV}$ and especially for an NMO study with signal in the low end of the energy range, as we will see in the next section, not only DIS but also other neutrino interaction types should be considered. In Figure 2.1 the total neutrino charged current cross section is shown together with contributions from the three largest contributing processes, Deep Inelastic Scattering (DIS), Resonance (RES) and Quasi-Elastic scattering (QE).

In a deep inelastic scattering process the neutrino has enough energy to break up the nucleon, resulting in a hadronic cascade. The underlying interaction is neutrino-quark scattering either by $\nu_l d \to l^- u$ or $\nu_l \bar{u} \to l^- \bar{d}$ or equivalently for anti-neutrinos by $\bar{\nu}_l u \to l^+ d$ or $\bar{\nu}_l \bar{d} \to l^+ \bar{u}$. From chirality arguments analogous to the case of pion-decay, neutrinos will have higher cross-sections with matter than anti-neutrinos, as anti-neutrino-quark scattering will have angular suppression due to the total spin of the initial particles being non-zero. Similarly for neutrino-anti-quark scattering, anti-neutrinos will have a larger cross section than neutrinos. Because nucleons not only consists of valence-quarks (u, d), but also sea-quarks, which can just as easily be anti-quarks as quarks, the ratio between the cross-sections of neutrino and anti-neutrinos is not $1/3$ as would be expected from angular considerations but is instead measured to be $\sim 1/2$ [5]. We will see later, how this is important for the NMO signal in IceCube, as the NMO signal is an effect in either the neutrino or anti-neutrino sector.

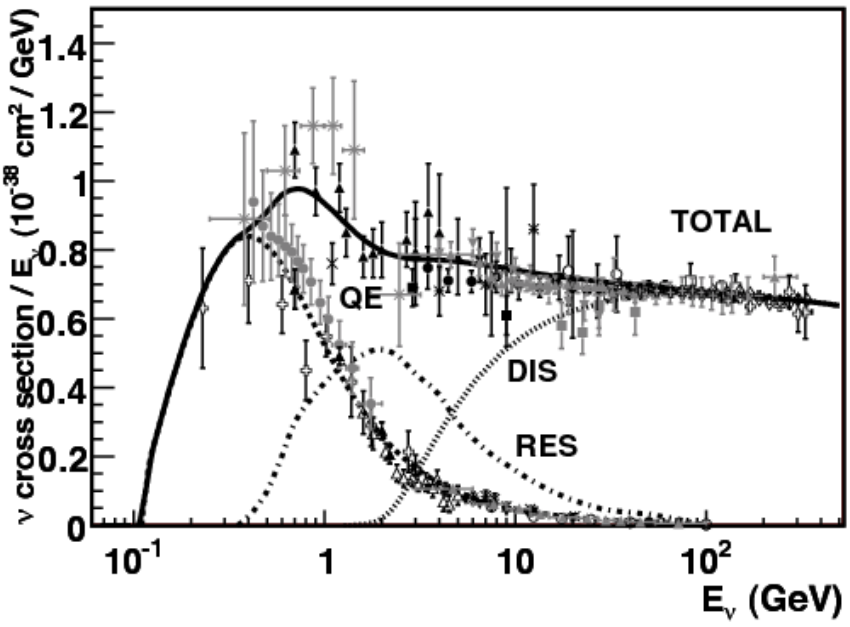

FIGURE 2.1.
Total neutrino per nucleon CC cross sections scaled to neutrino energy as a function of energy and cross sections for different contributing processes, quasi-elasic (QE), resonance (RES) and deep inelastic scattering (DIS). Figure from [14].

When a neutrino does not have enough energy to break up the target nucleon, it can instead excite it to a *resonance state*. In such a resonance interaction, the resonance particle can decay into different final-states, though most often to a nucleon and a single pion. Some of the possible resonances and decay-channels are less well-studied than others and will be a possible source of systematic uncertainties for any neutrino oscillation study [14]. For even lower energies, neutrinos interact via *quasi-elastic* processes where the target nucleon is converted from a neutron to a proton for neutrinos and from a proton to a neutron for anti-neutrinos (both charged current processes). For this type of interaction, both the V-A structure of the weak interaction and the form factor of the nucleon are important for calculating the cross-section. In the "dipole approximation" the form factor will depend on the *axial mass* M_A, the value of which has been measured ranging from $0.65\,\mathrm{GeV}$ to $1.35\,\mathrm{GeV}$ with a tendency for newer experiments to measure higher values [14]. This tension also gives rise to systematic uncertainties that are important for this project.

2.2 Neutrino Oscillations

The neutrinos that are allowed to interact in the weak force are the eigenstates of the interaction Hamiltonian and corresponds by lepton number to the charged leptons e, μ, τ and are ν_e, ν_μ, ν_τ. These neutrinos were thought to be fundamental and massless and would thus fit into the Standard Model. It was however quickly

established from neutrino experiments that the standard model does not fully describe and explain the behavior of these neutral leptons [15].

The perhaps most well-known tension in the neutrino sector came from the *solar-neutrino problem*, where the expected flux of neutrinos produced in the beta-decay fueling the sun did not match the measured rate of neutrinos in the Homestake mine [16]. As the solar-neutrinos would have to have electron lepton flavor, the experiment was designed to measure the product of electron-neutrino - chlorine charged current interactions resulting in merely a third of the predicted rate. It wasn't until the Sudbury Neutrino Observatory (SNO) experiment [17] was able to measure not only the electron-neutrino rate, but also the total neutral current interaction rate from all neutrino flavors, that the results made sense. The total rate of neutrinos was compatible with the prediction of the electron-neutrino rate from solar nuclear reactions, but the neutrinos did not interact as electron type neutrinos when they reached Earth, and had instead *changed flavor* as was confirmed by also the Super-Kamiokande experiment [18]. This remarkable feature of neutrinos is known as neutrino oscillations and provides multiple ways to probe the basic properties of neutrinos, not yet fully described by the Standard Model.

With analogies to mixing in the quark sector and the neutral kaon system, mixing between the three neutrino mass eigenstates explained the oscillatory behavior of the weak eigenstates [5]. This mixing is described by the unitary PMNS matrix U, named after the pioneers of its development; Pontecorvo, Maki, Nagakawaa and Sakata [19, 20, 21]

$$|\nu_\alpha\rangle = U^*_{\alpha i} |\nu_i\rangle \,, \qquad U = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} , \tag{2.1}$$

where $|\nu_\alpha\rangle$ is a neutrino state of flavor $\alpha = e, \mu, \tau$ and $|\nu_i\rangle \,, i = 1, 2, 3$ is the i'th mass state.

2.2.1 Oscillations in Vacuum

The evolution of a flavor system is given by the collective evolution of the individual mass eigenstates and can be described by a matrix in flavor basis as [11]

$$i\frac{\mathrm{d}}{\mathrm{d}t} |\nu_\alpha, t\rangle = U^*_{\alpha i} i\frac{\mathrm{d}}{\mathrm{d}t} |\nu_i, t\rangle = U^*_{\alpha i} E_i |\nu_i, t\rangle = U^*_{\alpha i} E_i U_{\beta i} |\nu_\beta, t\rangle \,, \tag{2.2}$$

where repeated indices are summed over and E_i is the eigenvalue of the free particle Hamiltonian for the i'th mass state. Taking the eigenvalues in the ultra-relativistic limit and using $\hbar = c = 1$, we have

$$E_i = \sqrt{m_i^2 + p^2} \simeq p + \frac{m_i^2}{2E}, \tag{2.3}$$

where m_i is the mass of the $i'th$ state, p is the coherent momentum, and E is the energy of the neutrino to this order not including the mass. As any factors of the Hamiltonian proportional to the identity matrix only add an overall phase to the system we can express the vacuum Hamiltonian in the flavor basis by omitting the momentum term as

$$\mathcal{H}_0 = \frac{1}{2E} U^* M^2 U, \quad M^2 = \begin{pmatrix} m_1^2 & 0 & 0 \\ 0 & m_2^2 & 0 \\ 0 & 0 & m_3^2 \end{pmatrix}. \tag{2.4}$$

The probability of finding the system in a specific flavor state β after a time t is

$$P_{\alpha\to\beta}(t) = |\langle \nu_\beta, 0 | \nu_\alpha, t \rangle|^2 \tag{2.5}$$

$$= \delta_{\alpha\beta} - 4 \sum_{i<j} \mathrm{Re}(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}) \sin^2 \left(\frac{\Delta m^2_{ji} L}{4E} \right), \tag{2.6}$$

where $\Delta m_{ji} = m_j^2 - m_i^2$ and it has been used that for ultra-relativistic particles $t = L$ in natural units.

The PMNS matrix can be parameterized as a product of rotation matrices in the i, j-planes $U_{ij}, (\theta_{ij})$, and a complex phase matrix given by $I_\delta = \mathrm{diag}(1, 1, e^{i\delta_{CP}})$ as [22]

$$U = U_{23}(\theta_{23}) I_\delta U_{13}(\theta_{13}) I^*_\delta U_{12}(\theta_{12}). \tag{2.7}$$

Using the standard notation $c_{i,j} \equiv \cos\theta_{i,j}, s_{i,j} \equiv \sin\theta_{i,j}$ the full matrix is

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{CP}} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta_{CP}} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta_{CP}} & c_{13}c_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta_{CP}} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta_{CP}} & c_{13}c_{23} \end{pmatrix}. \tag{2.8}$$

The effective Hamiltonian for anti-neutrinos can be found by making the substitution $U^* \to U$. If all entries of U are real ($\delta_{CP} = 0$) the mixing of neutrinos and anti-neutrinos will be identical.

The parameterization is made possible by the difference in scale between the two mass splittings Δm_{21} and Δm_{31}, and the smallness of the mixing angle θ_{12}, the values

Parameter	Best-fit Value NO	$\pm 1\sigma$	Best-fit Value NO	$\pm 1\sigma$
θ_{12}	33.41°	$^{+0.012}_{-0.012}$	33.41°	$^{+0.012}_{-0.012}$
θ_{23}	42.2°	$^{+1.1}_{-0.9}$	49.0°	$^{+1.0}_{-1.2}$
θ_{13}	8.58°	$^{+0.11}_{-0.11}$	8.57°	$^{+0.11}_{-0.11}$
δ_{cp}	232°	$^{+36}_{-26}$	276°	$^{+22}_{-29}$
Δm^2_{21}	$7.41 \times 10^{-5}\,\mathrm{eV}^2$	$^{+0.21}_{-0.20}$	$7.41 \times 10^{-5}\,\mathrm{eV}^2$	$^{+0.21}_{-0.20}$
Δm_{31}	$+2.507 \times 10^{-3}\,\mathrm{eV}^2$	$^{+0.026}_{-0.027}$		
Δm_{32}			$-2.486 \times 10^{-3}\,\mathrm{eV}^2$	$^{+0.025}_{-0.028}$

TABLE 2.1.
NuFit [23, 24] best-fit values of oscillation parameters including data from Super-Kamiokande [25].

of which can be seen in Table 2.1. The scaling effect makes it a valid approximation in many cases, to only regard 2-flavor systems. IceCube is e.g. with baselines up to $1 \times 10^4\,\mathrm{km}$ and energies in the GeV scale, mainly sensitive to the "atmospheric" oscillation parameters θ_{23} and Δm_{32}, which describes the amplitude and phase of ν_μ, ν_τ oscillations respectively.

2.2.2 Oscillations in Matter

When neutrinos propagate through matter instead of vacuum, a few different mechanisms affect the oscillation pattern. One is related to the density of the traversed medium and the other is furthermore dependent on the specific profile of media consisting of multiple layers. This section will give an overview of the two effects in 2-flavor systems and discuss their relation to separating the two mass orderings in a full 3-flavor description.

Matter Potential

When neutrinos propagate through matter the Hamiltonian describing them can be expressed as the vacuum Hamiltonian plus a potential V, from coherent forward scattering of neutrinos in the medium [26, 27]. As all neutrino flavors interact similarly with matter in the neutral current sector, only the electron-neutrino charged current interaction contributes to the oscillation probability. In the flavor basis, we have for a medium with electron density N_e:

$$\mathcal{H}_f = \frac{1}{2E} U^* M^2_{\text{diag}} U + V, \quad V = \begin{pmatrix} \sqrt{2} G_F N_e & & \\ & 0 & \\ & & 0 \end{pmatrix}, \tag{2.9}$$

where G_F is the Fermi constant and the corresponding Hamiltonian for anti-neutrinos can be found by the substitutions $U \to U^*$ and $V \to -V$.

Constant Density Medium

Oscillations in matter are analogous to vacuum oscillations and can for a 2-flavor system be expressed using the effective mixing angle θ_m and the propagation (mass) states in matter ν_{1m} and ν_{2m}. The flavor states ν_α and ν_β can then be described as

$$\nu_\alpha = \cos\theta_m \nu_{1m} + \sin\theta_m \nu_{2m}, \qquad \nu_\beta = \cos\theta_m \nu_{2m} - \sin\theta_m \nu_{1m}. \tag{2.10}$$

The mixing angle θ_m can be found by diagonalizing the 2-flavor version of the flavor state Hamiltonian (2.9) and is [22]

$$\sin^2 2\theta_m = \frac{1}{R}\sin^2 2\theta_0, \qquad R \equiv \left(\cos 2\theta_0 - \frac{2VE}{\Delta m^2}\right)^2 + \sin^2 2\theta_0, \tag{2.11}$$

where θ_0 is the vacuum mixing angle. We see that for $V \to 0$ we re-find the vacuum oscillation angle as $R \to 1$. The dependence of $\sin^2 2\theta_m$ on E and V has a resonant character - the effective mixing term $\sin^2 2\theta_m$ can become 1 for vacuum mixing angles away from maximal mixing when

$$R = \sin^2 2\theta_0 \qquad \to \qquad \sin^2 2\theta_m = 1, \tag{2.12}$$

which is satisfied for

$$\frac{2VE}{\Delta m^2} = \cos 2\theta_0. \tag{2.13}$$

Using the expression for V_e in (2.9) we find the resonance density as a function of neutrino energy, mass splitting, and vacuum mixing to be

$$N_e^R = \frac{1}{2\sqrt{2}G_F}\frac{\Delta m^2 \cos 2\theta_0}{E}. \tag{2.14}$$

This resonant character of oscillations in matter is what is known as the MSW effect after Mikheyev, Smirnov [28] and Wolfenstein [26]. The effect is noticeable for different sets of neutrino-events in IceCube, dependent on the neutrino energy and which layers they traverse. Each layer with its specific density will give rise to oscillation amplifications for neutrinos with the corresponding resonance energy.

Discrete Density Layers

The constant density approximation is not entirely fitting for the neutrinos we are able to detect in IceCube, as Earth has multiple layers of varying density. Neutrino

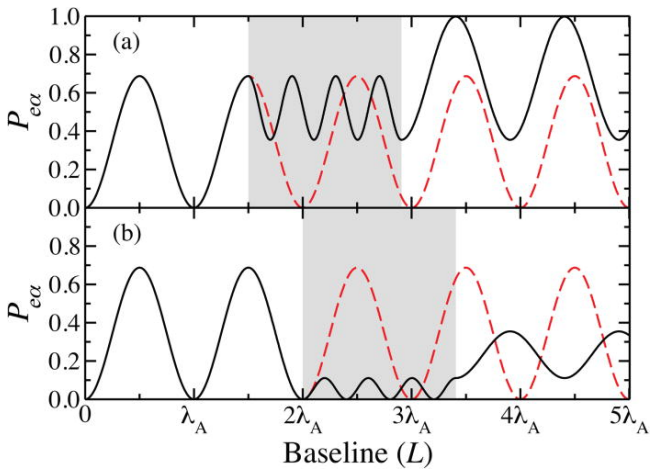

FIGURE 2.2.
Example of transition probability for neutrinos with fixed energy traveling through through two different matter profiles. Solid line shows the case of a dense middle medium (shaded gray) between two less but equally dense layers (shaded white). The dashed red lines indicate the transition probability without the dense middle section. Figure from [34].

oscillations in mediums consisting of a number of constant density layers will exhibit what is known as *parametric resonance*, where a phase is added at the borders between different layers [29, 30, 31, 32, 33]. Because the formalism of the effect is extremely dense and not necessary for the present analysis, only the general principles of the resonance will be discussed here.

Parametric resonance effects for Earth-crossing neutrinos are only well approximated by a two-flavor system for much smaller energies than can be detected in IceCube, even with the Upgrade, but we will use results from such a case to exemplify the effect [34]. For a 3-layer medium, consisting of layers of two distinct densities with the middle being denser than the first and third which are the same, neutrino oscillations are highly impacted by the placement and thickness of the dense middle layer. In Figure 2.2 the oscillation probability for a 2-flavor system is sketched against the density of the traversed medium. It can be seen that the placement of a higher density middle layer can either enhance or suppress oscillations depending on the corresponding neutrino energy and path length. At resonance the oscillation probability in such a two-flavor "castle-wall" case can be found to be [30]

$$P_{\alpha\beta} = 1 - \sin^2(4\theta_m - 2\theta_c). \tag{2.15}$$

Implicitly, in this equation is the expression for the matter mixing angles in the mantle and core respectively and from (2.11) we know that these are functions of the mass-squared difference.

2.3 The Neutrino Mass Ordering

The question of the neutrino mass ordering is the question of whether the third mass eigenstate is heavier or lighter than the two other mass eigenstates. As can be seen from (2.6) the mass-squared difference appears within a $\sin^2$-function in the vacuum

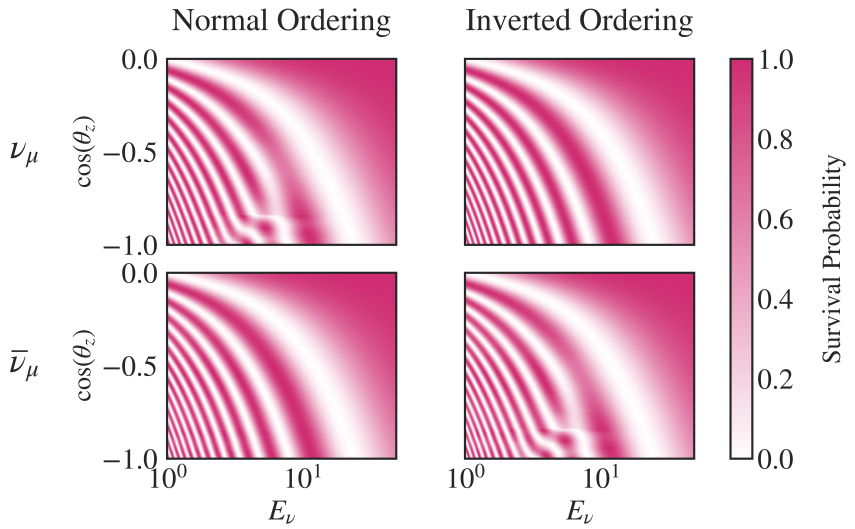

FIGURE 2.3.
Survival probability of muon-type neutrinos and anti-neutrinos as a function of zenith angle and energy under the two different mass orderings. The probabilities are calculated using the PISA software described in Section 3.4.1.

oscillation probabilities, removing any information about the *sign* of the difference. Luckily, matter effects are *also* a function of the sign of the mass-squared difference and in ways that makes it possible to discern the sign.

Matter effects are inherently tied to the electron-neutrino and therefore also to mixing of the first mass-state, which is a subdominant factor in atmospheric neutrino interactions in IceCube. Describing matter effects in IceCube therefore requires a 3 flavor system, but instead of writing out the full set of equations governing oscillations in constant density matter and then also applying that to any sort of realistic matter profile, we can simply plot the results of such calculations in what we call an *oscillogram* which can be seen in Figure 2.3. Recalling that matter effects enter the Hamiltonian with a different sign for neutrinos and anti-neutrinos, we find that this directly links with the sign of the mass-squared difference Δm_{23}. Depending on the true mass ordering, matter modulations will either affect neutrinos or anti-neutrinos traveling through Earth and interacting in IceCube. The question of the sign of the mass-squared difference thus becomes a question of separating the neutrino and anti-neutrino channel, as we will see in Chapter 3 is not as easy as one could hope.

3 Detecting Neutrinos With IceCube

At the South Pole one cubic-kilometer of ice constitutes the IceCube Neutrino Observatory. This detector makes it possible to detect neutrinos from all directions of the sky with energies from a few GeV to multiple TeV. Besides its use as a neutrino telescope to identify extraterrestrial neutrino sources [35], the densest part of the detector also enables us to study the fundamental behavior of these evasive, elementary particles. In this chapter, I will introduce the IceCube detector and the Upgrade, planned for deployment in the Antarctic summer 2025-26. I will briefly discuss the neutrino source used in this project and lastly examine how data is produced and measured in the IceCube Collaboration. This will include a brief overview of Monte-Carlo methods and parameterized analysis.

3.1 IceCube and the IceCube Upgrade

The IceCube Neutrino Observatory utilizes the dark and cold glacial ice to measure ionizing particles produced by atmospheric neutrinos interacting in its volume. When electrically charged particles travel trough a medium, faster than the speed of light in that medium they emit Cherenkov Radiation in the form of photons. By placing optical sensors in the ice, it is possible to measure the occurrence of such Cherenkov Radiation and map together photons detected in different places to belong to the same ionizing particle. Such digital optical modules (DOMs) consist of some form of photomultipler-tube (PMT) and a collection of electronics enabling the collection and transportation of data from a DOM deep in the ice to the IceCube Lab at the surface. The full IceCube Detector consists of different subsets of strings placed in the ice sheet, a top view of which can be seen in Figure 3.1. The primary in-ice array, labeled simply as IceCube in most publications, is the largest and least dense subset of the detector with 78 strings distributed with a horizontal spacing of $125\,\mathrm{m}$ [36]. 60 DOMs are placed on each string with a vertical spacing of $17\,\mathrm{m}$. 8 other strings constitute a specialized subset of the detector which was deployed alongside the primary array and optimized for oscillation measurements. These 8 *DeepCore* strings, each also with 60 DOMs with only $7\,\mathrm{m}$ to $10\,\mathrm{m}$ in between, are situated within the grid of the primary

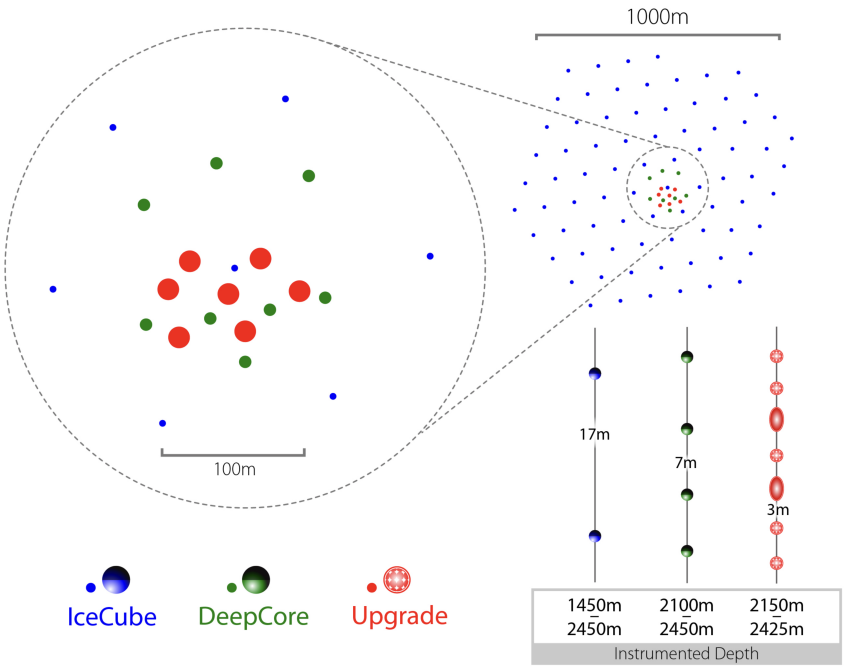

FIGURE 3.1.
Top view of the IceCube Detector. Each colored dot represents a string of Digital Optical Modules (DOMs) spaced according to their sub-detector specification. The Upgrade strings are more densely instrumented with not only downward-facing DOMs but also multi-directional mDOMs. The combination of IceCube, DeepCore and Upgrade strings in the middle of the instrumented volume represents the highest-resolution part of the detector. Figure from [7].

array resulting in horizontal distances of $42\,\mathrm{m}$ to $72\,\mathrm{m}$ between strings for the total of 15 (8 DeepCore + 7 normal) strings constituting the DeepCore volume [8].

While the primary array was designed for astrophysical neutrino measurements in the $\mathcal{O}(\mathrm{TeV}) - \mathcal{O}(\mathrm{PeV})$ energy-range, the DeepCore volume was optimized for $10\,\mathrm{GeV}$ to $100\,\mathrm{GeV}$. Though this enabled the measurement of atmospheric τ-appearance and μ-disappearance, the question of the Neutrino Mass Ordering is at even lower energies as can be seen in the oscillograms in Figure 2.3. The coming Upgrade subset will consist of 7 strings with a total of around 700 optical sensors for physics and calibration purposes. The optical modules on these strings will be improved from the 1st generation DOMs in IceCube and DeepCore. As is indicated on the lower right-hand side of Figure 3.1, the Upgrade strings shown in red will have both so called "D-eggs" consisting of 2 PMTs facing up and down respectively and the multi PMT "mDOMs" consisting of 24 smaller PMTs pointing in all directions as opposed to the original DOMs with a single PMT facing down towards the bedrock. These new sensors will improve both the angular resolution and photon detection efficiency, and with the new strings resulting in only $20\,\mathrm{m}$ horizontal spacing and $3\,\mathrm{m}$ vertical spacing, the energy threshold will be pushed towards $1\,\mathrm{GeV}$ and enable measurements of the matter effects described in Section 2.2.2.

3.2 Atmospheric Neutrinos

IceCube is designed to detect neutrinos produced from Cosmic Ray showers in the atmosphere. These extraterrestrial rays consist mainly of protons that enter and interact with the atmosphere producing *showers* of secondary mesons decaying to other lighter particles [37]. The main neutrino production arises from the decay of secondary pions.

As discussed in Section 2.1.1, pions decay mainly to muons and muon-neutrinos with some electrons and electron-neutrinos but no tau and tau-neutrinos. The muons produced in pion decays may either decay to electrons, electron-neutrinos and muon-neutrinos, or have enough energy to reach Earth before decaying. The neutrinos produced by these Cosmic Ray showers have a flavor composition of $2 : 1 : 0$ for $\nu_\mu : \nu_e : \nu_\tau$, and the muons that do not decay before reaching Earth will possibly decay inside the detector as *background* muon-events.

The rate of atmospheric neutrinos is highly dependent on energy and the neutrino flux has been found to follow a steep power law

$$\frac{\mathrm{d}\phi}{\mathrm{d}E} \propto E^\gamma, \quad \gamma = -3.7, \tag{3.1}$$

accounting both for the initial Cosmic Ray spectrum and for the energy-dependent lifetime of secondary mesons [38].

3.3 Neutrino Event-Types

Neutrino interactions have a variety of expressions and particle products depending on the type of the neutrino, the energy and the resolution with which it can be seen. Some of the features are possible to discern with the sensors of IceCube, others simply happen on too small scales for the detector to possibly be sensitive to them. Depending on the spatial and temporal signature in the detector, an event is classified as one of three event-types corresponding to the type of particle interacting in the event and therefore named particle identification (PID). The three types are cascade-like, track-like, and mixed, and will be explained in greater detail in this section.

For a neutral-current event a neutrino interacts with the ice and continues after the interaction as a neutrino, invisible to the detector. Only the hadronic cascade of the hit nucleon or the electromagnetic cascade of the excited electron will result in a signal in the detector. Hadronic and electromagnetic cascades create the same slightly oblong

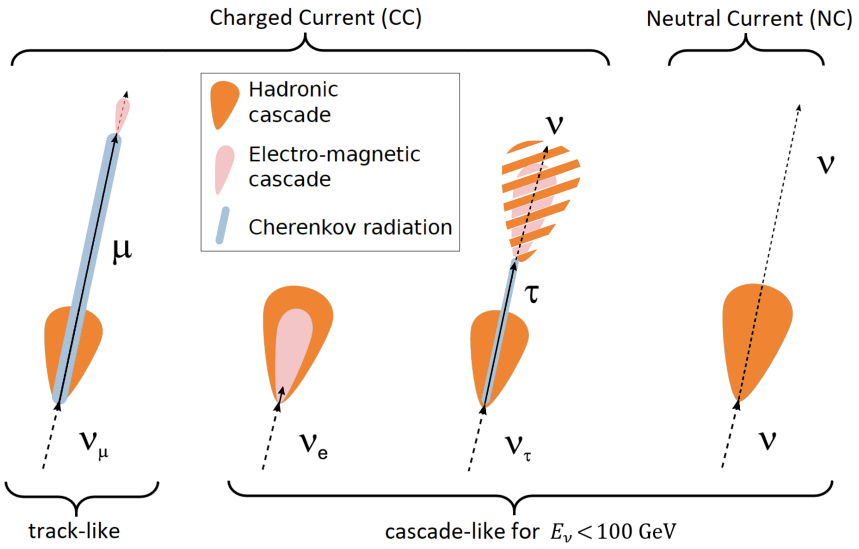

FIGURE 3.2.
Neutrino interaction sigatures in IceCube. Track-like events can be produced by charged current muon-neutrino interactions. Cascade-like events can be produced by all neutral current interactions and tau- and electron-neutrino charged current interactions. Modified from [38].

spherical signature and are therefore not distinguishable from each other, resulting in neutral current events of all flavors of neutrinos producing the same type of event in the detector namely cascade-like event-types [38].

Charged-current events differ by the flavor of the initial neutrino as this will determine the charged lepton produced in the ice. Electron-neutrinos produce electrons triggering an electromagnetic cascade making it look like any NC event. Muon-neutrinos produce a hadronic cascade and a charged muon. Because muons are close to minimally ionizing in the energies considered here, they are able to travel rather long distances in the detector setting of DOMs along their trajectories before decaying [39]. This will lead to a different looking event than the cascade-like events described in the previous paragraph. Most muon-neutrino events are therefore labeled as track-like.

Tau-neutrinos produce the heavy tau-lepton with mass $m_\tau = 1777\,\mathrm{MeV}$, which will only be able to travel few centimeters before decaying [6]. This means that the tau-lepton itself won't be visible in the detector, but the product of its decay will be. Tau-leptons can decay in modes resulting in electromagnetic and hadronic cascades so close to the initial interaction vertex, that the two merge together resulting in cascade-like events. It can also - though with a small branching ratio - decay to a muon and muon-neutrino pair from which a track-like signature may be observed from the muon. Such decay modes are however typically not visible at atmospheric neutrino energies.

The last event-type is *mixed*. In this PID bin we put the events that are not clearly either a cascade or a track. This is done to keep the track and cascade bins as clean as possible while allowing for uncertainties in the reconstruction of events. A schematic overview of the signatures of different neutrino-type events can be seen in Figure 3.2.

3.3.1 Reconstruction of Events

The individual detections of photons in the ice by different DOMs are collected and grouped together into events. Some photons will simply be noise e.g. from radioactive compounds on the DOMs, and an initial trigger-level is therefore set before accepting a signal. After further noise cleaning, events consisting of multiple DOM-hits at different times and different intensities undergo event reconstruction. The energy of an event is found by translating the amount of energy deposited in the detector in the form of photons to the energy of the incoming neutrino. Because of the sharp power-law of the atmospheric neutrino flux and the location of the NMO signal in low energies, events are binned by their reconstructed energy in 12 logarithmic bins from $3\,\mathrm{GeV}$ to $300\,\mathrm{GeV}$.

If enough DOMs are hit in an event, it is possible to reconstruct the direction of the initial neutrino to some precision, depending on the type of interaction and the energy deposited in the detector. The direction is translated into a zenith and an azimuth angle describing the angle with the horizon and the radial direction respectively. For oscillation measurements we are interested in the energy and baseline (distance traveled) of the neutrinos. The baseline can be found from the zenith angle alone and we therefore use the cosine of the zenith angle in our analysis binning. For this particular analysis mainly up-going neutrinos are considered as down-going events are both muon-contaminated and do not experience oscillations because of the short distances traveled. This results in a binning of 10 linear bins in $\cos\theta_z$ from -1 to 0.3. An example of the final analysis binning can be seen Figure 3.3.

3.4 Analysis Tools

In order to make any measurements from the obtained data, it is crucial to have some expectation to match the data with. For this analysis a software initially created for the Precision IceCube Next Generation Upgrade (PINGU) [40] extension of IceCube called PISA (PINGU Simulation and Analysis) has been used [41]. This software utilizes the fact that the final event rates in the detector can be described by independent stages

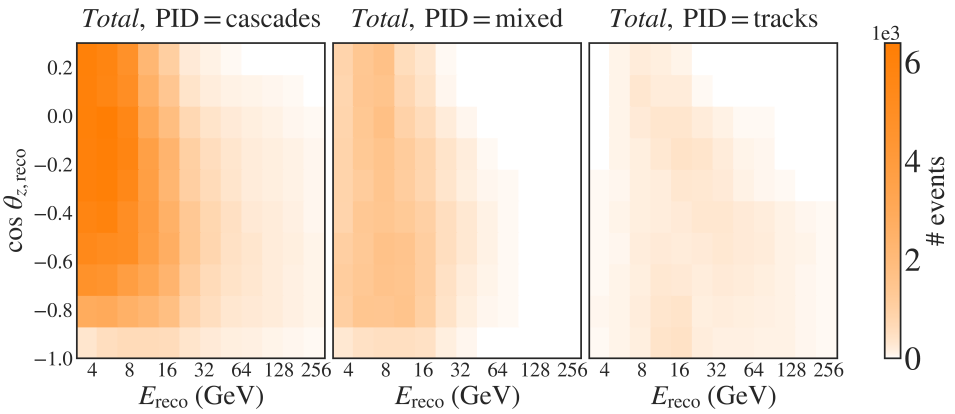

FIGURE 3.3.
Example of final stage event distributions simulated for the IceCube Upgrade. Produced with nominal values of Table 3.1, $\theta_{23} = 42.4°$, and for 3 year livetime. The events are binned by their reconstructed PID, energy, and coszen.

which multiplied together will provide a full picture without requiring simulation of all possible combinations of physics and nuisance parameters.

3.4.1 The PISA Software

The final number of events measured in the detector can be split into a product of different *stages* - flux, oscillation, detection and reconstruction. Each stage will provide a weight to a simulated event and can be handled separately from the simulation and therefore also independently from other stages only requiring parameters specific for that stage. These stages will be described here.

Flux

For the flux of atmospheric neutrinos we use Honda et. al as nominal values [42]. Such a flux model relies on a number of external parameters which can give rise to systematic uncertainties if not correct. These uncertainties are handled by calculating flux gradients using a software called MCEq (Matrix Cascade Equations) [43]. A list of the flux and other parameters used in the analysis can be found in Table 3.1. The flux parameters consist of parameters affecting the initial cosmic ray showers e.g. the spectral index, but also parameters associated with hadron production in these showers. Such uncertainties are parameterized based on Barr et. al [44] which separates the hadron production into pion and kaon production respectively, and parameterize uncertainties in these channels by energy. Pion uncertainties are calculated for pion and anti-pion combined and the ratio between the two production channels is included. For kaon production the kaon and anti-kaon uncertainties are varied completely independently of each other.

	Parameter	Nominal Value	Prior	Unit
Flux				
	$\Delta\gamma_\nu$	0.0	± 0.1	
	Barr, $d_{\pi+}$	0.0	± 0.3	
	Barr, $h_{\pi+}$	0.0	± 0.15	
	Barr, $i_{\pi+}$	0.0	± 0.122	
	Barr, y_{K+}	0.0	± 0.3	
	Barr, z_{K+}	0.0	± 0.122	
	Barr, y_{K-}	0.0	± 0.3	
Oscillation				
	θ_{12}	33.82		$^\circ$
	θ_{13}	8.65		$^\circ$
	θ_{23}	42.2(49.0)	$[0, 90]$	$^\circ$
	δ_{CP}	0.0		$^\circ$
	Δm_{21}	7.4×10^{-5}		eV^2
	Δm_{31}	$(-)2.5 \times 10^{-3}$	$[1, 4]$	eV^2
Detection				
	$M_{A,QE}$	0.0	± 0.1	
	$M_{A,res}$	0.0	± 0.1	
	DIS	0.0	± 0.1	
	N_ν	1.0	± 0.2	

TABLE 3.1.
Systematic parameters relevant for this analysis and their nominal values grouped by type. Priors are stated for parameters that are included as fit parameters. Symmetric Gaussian priors are indicated by $\pm\sigma$. Uniform priors are indicated by a range. For dimensionless parameters the unit is left blank. For parameters with different nominal values for normal and inverted ordering, the normal ordering value is stated and the inverted ordering value given in parenthesis. Only Barr parameters that are handled as free in the analysis are stated, see Section 5.1 for details on how they are chosen.

Oscillations

As described in Chapter 2, neutrino oscillations highly impact the flavor composition of the neutrinos detected in IceCube. This stage simply calculates the oscillation probability for a neutrino based on its true energy, true direction, and flavor using the Prob3++ package [45] originally written by the Super-Kamiokande collaboration based on the work of Barger et al. [27]. Besides the physics parameters related to oscillations, this stage also takes as input a model of the earth density profile. We use the Preliminary Earth Model (PREM) approximated by 12 constant density layers [46]. The oscillation probabilities are evaluated in bins of energy and zenith angle in a resolution much finer than the analysis binning but still less computationally demanding than evaluating the probability of each single neutrino.

Detection

The number of neutrinos that not only pass through the detector but also interact and are detected is estimated using large numbers of Monte-Carlo (MC) simulated neutrinos. Neutrinos are generated and their interactions propagated with the GENIE software [47]. These events are then passed through noise-cleaning and selection criteria optimized to get rid of noise events and background muons while retaining as many neutrino events as possible. The active livetime of the detector is applied and it is also at this stage that detector uncertainties should be applied. Because of the early stage of the Upgrade simulation effort, the necessary MC events generated under different detector realizations, are not at this time reliable enough to give an estimation of detector uncertainty effects on the final stage event distributions and such systematic uncertainties are therefore left out of this analysis.

Reconstruction

The reconstruction stage smears the events from the detection stage to the resolution of the final level. MC events used in the detection stage are used here to estimate the resolution functions in the three analysis observables i.e. reconstructed energy E_{reco}, reconstructed zenith angle $\theta_{z,reco}$, and PID (cascade, mixed, track). For this, different simple cuts are made while also machine learning algorithms are used to optimize the noise cleaning and event-reconstruction [48, 49].

3.5 Signature of the Neutrino Mass Ordering

If the IceCube upgrade is going to be sensitive to the mass ordering, the data we measure needs to be different depending on the true mass ordering. As discussed in the previous chapter, the mass ordering determines whether matter effects impact neutrinos or anti-neutrinos. Unfortunately IceCube can not tell the difference between neutrinos and anti-neutrinos. Luckily there are other differences between neutrinos and anti-neutrinos that make the distinction of the mass ordering possible, including the fact that the cross-sections for neutrinos and anti-neutrinos are different in the matter (and not anti-matter) dominated ice at the South Pole and that the incoming flux of neutrinos is larger than the corresponding flux of anti-neutrinos. These two effects cause the inverted ordering (affecting anti-neutrinos) to result in a *less pronounced* signature of matter effects while the normal ordering affecting neutrinos will result in a more pronounced signature of matter effects. Since matter effects are

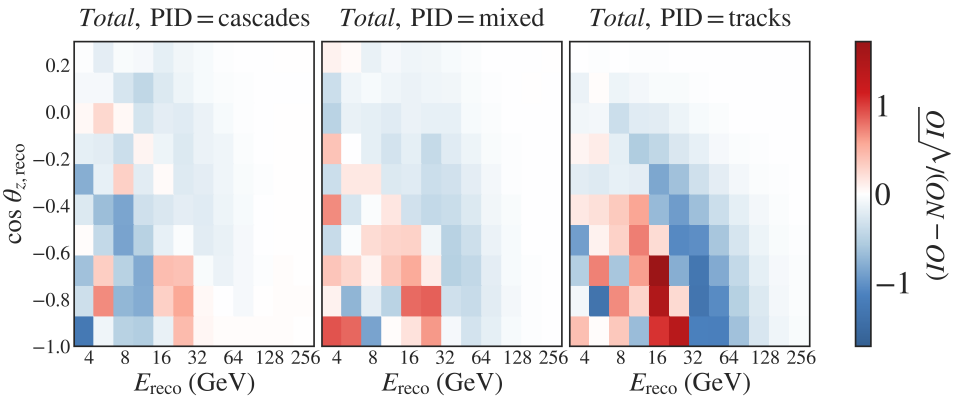

FIGURE 3.4.
Example of the bin-wise distinguishability of the two orderings. Produced with nominal values of Table 3.1, $\theta_{23} = 42.2°$, and for 3 year livetime.

essentially a distortion of the vacuum oscillation band, the measured neutrinos will either be more or less distorted based on the mass ordering.

To get a sense of the difference we can compare events expected to be measured in IceCube produced under the assumption of the two different orderings. In Figure 3.4 the pull in each bin for comparing Inverted Ordering to Normal Ordering is shown and as expected from the oscillograms in Figure 2.3 the signal is in the lower energy range and for more up-going events. The higher rate of neutrinos from Inverted Ordering in the band around $16\,\mathrm{GeV}$ is the result of enhanced matter oscillation in this region being less pronounced for the Inverted Ordering than for Normal Ordering.

4 Statistics

With data and theory comes statistics. It is the tool we use to quantify how well data fits theory, and how well we can trust that fit. For this project, we wish to quantify how well the coming Upgrade detector will be able to separate the two orderings i.e. whether the data we are collecting is a result of the Normal Ordering being true, or if it is a result of the Inverted Ordering being true. If the new detector cannot tell the two cases apart, we say that it is not sensitive to the NMO, but if it is able to separate them, we want to know *how well*, it can separate them. In this chapter I will present different interpretations of what such a *sensitivity* is and the steps needed to calculate it.

4.1 Basic Concepts

To be able to discuss the validity and interpretations of the statistics used to describe the Neutrino Mass Ordering, we need to discuss statistics. Such concepts, when merely stated and not explained can cause confusion and misinterpretation of results. Confidence levels for one, is not a well-defined unique concept[1].

4.1.1 Fitting Composite Hypotheses

The first thing we need to address is fitting. Assume we have a data-set $D = \{D_1, D_2, ..., D_N\}$ and a hypothesis H which is a function of several parameters $P = \{P_1, P_2, ..., P_k\}$. This hypothesis is then said to be *composite* if it can have different realizations for different values of the parameters P. A simple hypothesis could be "The next card drawn is a spades" where a composite hypothesis would be "The next card drawn is either a spades or a hearts depending on what the previously drawn card was". For the Neutrino Mass Ordering, the number of events we see in the detector is not only determined by whether the Normal Ordering hypothesis or the Inverted Ordering hypothesis is true, it is also a function of several physics and nuisance parameters such as the mixing angle θ_{23} and the cross-section of neutrinos. The two mass orderings are therefore composite hypotheses.

The purpose of the fitting stage is to find the parameters that make the output of a hypothesis look most like the observed data-set. We will call a realization of the hypothesis H for some set of parameters P a template $t(H, P)$. We quantify the difference between template and data-set with a metric, which returns more extreme values for more extreme differences. For a metric that returns larger values for larger differences, minimizing that metric over all parameters will therefore return the best-fit (bf) parameters P^{bf} that provides the template which most resembles the data-set. The choice of metric is not unique and it is not uncommon to see χ^2 metrics used in NMO studies. For this project we will use a likelihood based on the Poisson distribution,

$$p(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{\Gamma(k+1)}, \tag{4.1}$$

where $k!$ has been replaced by the gamma function $\Gamma(k+1)$ to properly work for non-integer bin-counts. The likelihood for a data-set of N bins is the product of the Poisson likelihoods of each bin with k_i, λ_i being the observed and expected values in the i'th bin respectively. The likelihood gives higher values for more likely outcomes, but for numerical reasons, instead of maximizing the likelihood, we minimize the negative log-likelihood given by

$$LLH = -\ln \mathcal{L} = -\sum_{i=1}^{N} k_i \ln \lambda_i - \lambda_i - (k_i \ln k_i - k_i), \tag{4.2}$$

where Stirling's approximation has been used. For a composite hypothesis the value of such a test-statistic will depend on the true parameters P_t that are in general not known. The test-statistic for a composite hypothesis must therefore be minimized over all parameters

$$LLH_{\min} = \min_{P} LLH(P). \tag{4.3}$$

There are different strategies for the type of minimization employed in such a fit of a multidimensional parameter space with possible local minima. For this project the MIGRAD method [50] implemented in the iminuit package [51] has been used. It is a variable metric method using inexact line search, see e.g. [52] for details.

4.1.2 Nested and Non-Nested Hypotheses

In many cases of physics when we consider hypotheses and whether or not a specific data-set can be said to match a hypothesis, we are dealing with nested hypotheses. If I measure the gravitational acceleration g, and want to say whether or not my measurements are in agreement with the standard value of g_{std} I am in reality comparing my data to two different hypotheses. Hypothesis (A) states that the value of g is g_{std}, while Hypothesis (B) states that g can be any positive number. We can easily get

TABLE 4.1.
Conversion table for a χ^2-distribution with one degree of freedom.

χ^2	Significance	Confidence Level
1	1σ	68.3 %
4	2σ	95.45%
9	3σ	99.37%

Hypothesis (B) from fixing a variable, g, in Hypothesis (A) and therefore the two are said to be nested.

In the case of nested hypotheses Wilk's Theorem [9] holds and the likelihood ratio (LHR) equivalent to the difference of the log-likelihoods

$$LHR = 2(LLH_B - LLH_A), \tag{4.4}$$

follows a χ^2-distribution with degrees-of-freedom related to the difference in fit parameters between the hypotheses. We can then decide to reject the hypothesis B at a confidence level $1-\alpha$, if we obtain a value of LHR within the α most extreme results. For the χ^2-distribution of 1 degree of freedom this is a straight forward conversion between measured test-statistic and resulting significance, shown for typical confidence levels in Table 4.1.

In the case of the neutrino mass ordering, the two hypotheses are not nested. The true ordering is *either* normal or inverted. This means that Wilk's Theorem does not hold and we need another method of determining significance. Exactly *which* other method should be used is and has been a subject of debate in the neutrino oscillation community [53, 10, 54, 55] and we will therefore discuss different options in the following sections.

Common for the methods is the use of a test-statistic based on the likelihood ratio of Wilk's theorem, using either χ^2 or LLH metrics for the fitting procedures. For the metric defined in (4.3) and for a given true ordering we define the test-statistic TS as

$$TS = \min_{P\in NO} LLH(P) - \min_{P\in IO} LLH(P) = LLH_{NO} - LLH_{IO}, \tag{4.5}$$

where the negative log-likelihood has been minimized over all fit parameters P assuming both Normal ($P \in NO$) or Inverted ($P \in IO$) Ordering. For a given data-set D the test-statistic becomes

$$TS(D) = LLH_{NO}(D) - LLH_{IO}(D). \tag{4.6}$$

If the data matches the two hypotheses equally well $TS(D) = 0$, if it matches Normal Ordering better $TS(D) < 0$, and oppositely if it better matches Inverted Ordering $TS(D) > 0$. The test-statistic can be converted to a χ^2-fitting procedure using

$TS = \Delta LLH_{NO-IO} = \frac{1}{2}\Delta\chi^2_{NO-IO} = \frac{1}{2}(\chi^2_{NO} - \chi^2_{IO})$, both will in general *not* be χ^2-distributed.

4.1.3 Bayesian and Frequentist Statistics

Interpretations of the sensitivity of an experiment to any type of measurement is affected by the interpretation of probability it is based on [2]. There are two major schools of interpretations in applied statistics: the frequentist and the Bayesian [1]. The frequentist approach considers probability as the limit of frequency, where an outcome is expected with some frequency for many repeated cases. Here the probability of an outcome is directly linked to the setting e.g the experimental setup and data-set considered.

A Bayesian understanding of probability can be thought of as a subjective probability. In this case a probability is not only affected by outcomes of some measurements or observables but also of the degree of belief one has in the theory. Bayesian statistics are therefore able to make statements about cases, to which the frequentist approach would yield nonsensical results. One of the problems with Bayesian Statistics however, is how to assign a prior knowledge to a theory. Assigning a uniform prior to describe "no knowledge" is non-trivial as a uniform prior in one parameter space will not look uniform in another parameter space.

As such there are different roads to follow in the case of defining sensitivity - a measure of how well we think we will be able to measure something. Bayesian statistics has been used for the NMO as it is possible to apply the knowledge of *one of the orderings must be true*, whereas such a statement holds no real value in the frequentist approach. Bayesian statistics applied to the neutrino mass ordering can be found in [55] and [53], a frequentist approach in [10] and a discussion of the two in [56].

For this project a strictly frequentist approach will be used, in line with previous work done for IceCube PINGU [38, 40, 57].

4.1.4 Confidence

Confidence levels and significance are terms used by many authors for many different goals. The trouble is that they are not uniquely defined. Confidence in a result means one thing in a frequentist's mind and another in a Bayesianist's [58]. For a Bayesianist a confidence may be the probability that a hypothesis is true given the measured data, or for a true value μ_t to lie in the interval $[\mu_1, \mu_2]$ with a probability

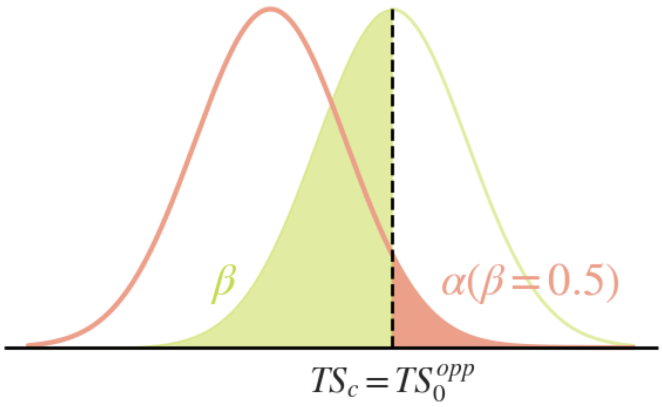

FIGURE 4.1.
A sketch of a confidence test for a true distribution (peach). The shaded area of the true distribution α is determined by the value of TS_c, here chosen to be the mean of a different distribution (light green) giving a power (shaded green area) of 0.5.

of α. For a frequentist however, a confidence interval contains the unknown μ_t in a fraction α of experiments. A confidence in the neutrino mass ordering in a frequentist understanding is a confidence with which we reject the wrong hypothesis. In figure 4.1 the case for a Gaussian distributed test-statistic is sketched. For some critical value of the test-statistic TS_c one would choose to reject the hypothesis which the distribution represents. This will be done at the confidence level $1 - \alpha$, where α is the fraction of times such a decision to reject the hypothesis would be a false negative i.e. an error of type I. At that value of the test-statistic some tests belonging to the wrong hypothesis may give the same, or a less extreme, test-statistic and therefore give rise to false positives i.e. errors of type II. The *power* of the test is one minus the fraction of these $(1 - \beta)$ and will be related to the significance by the choice of the critical value TS_c. Ideally one would like to construct a test that both has a high significance and a large power, but the two pull in different directions and deciding how to choose an appropriate power for the significance one wishes to state is non-trivial.

A common way to avoid this is to state the *median* sensitivity which is defined as the significance at a power of $1 - \beta = 0.5$. Such a test is sketched in Figure 4.1. Even though a power of 0.5 is quite low, the median sensitivity is used by other experiments for a basis of comparison and will therefore also be used in this project. In [10] a second sensitivity definition, the *crossing* sensitivity, is also laid out. As this is not easily translated to the case of a composite hypothesis, we have not included it here even though it is an interesting case as it describes the confidence when *exactly* one of the two-hypotheses can be excluded.

For composite hypotheses the choice of setting a critical value TS_c is non-trivial as discussed in [10]. If one wishes to report a confidence $1 - \alpha$, that confidence should be true for all possible parameters in the opposite hypothesis. One must therefore either state the confidence as a function of each parameter or minimize the confidence for these parameters and report the lowest. In this project and other IceCube NMO

studies [38, 40, 57, 59] we report the lowest confidence, the procedure for which will be explained in Section 4.2.

Converting to Number of σ

If we were dealing with a symmetric test-statistic with mean μ, a measured value of $\mu \pm x$ would be equally unlikely. We would therefore integrate *both* tails of the distribution from $\pm x$ to get the appropriate α. Our test-statistic is not symmetric and a measured value above or below the median of the distribution will not hold the same meaning. This is because the opposite distribution, which we are comparing to, will be located at either a lower or higher value of the test-statistic than the true distribution. For this reason we only integrate from the critical value of the test-statistic and *away* from the median of the true distribution being tested.

The value of α or equivalently $1 - \alpha$ can be converted to a number of standard deviations using the standard Gaussian distribution *even though the test-statistic is not Gaussian*. This is done to translate the obtained significance to a quantity commonly used in particle physics. The merit of this will be discussed in Chapter 6.

If we use the standard convention of setting a central limit, which follows the χ^2-distribution listed in Table 4.1, we would rather misleadingly convert a value of $\alpha = 0.5$, signifying that the medians of the two distributions have the same value, to a non-zero significance. It would therefore be more appropriate to translate the value of α to a one-sided Gaussian (upper or lower) limit were a value of $\alpha = 0.5$ will lead to a significance of 0. Using the Complimentary Error Function defined as

$$\text{erfc} = 1 - \text{erf}z = 1 - \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt, \tag{4.7}$$

where erf$z$ describes the probability for a normally distributed variable with mean $\mu = 0$ and standard deviation $\sigma = \frac{1}{\sqrt{2}}$ to fall in the range $[-z, z]$, we can define

$$n_{\sigma,one-sided} = \sqrt{2}\,\text{erfc}^{-1}(2\alpha). \tag{4.8}$$

For a central limit which we will call *two-sided* in this project the same definition can be used by making the substitution $2\alpha \to \alpha$.

4.2 Sensitivity to Neutrino Mass Ordering

At the time of this project the Upgrade detector has not been deployed and the goal of a sensitivity study is to map out the expected test-statistic distributions for the two orderings. This is done with several goals in mind: 1) to attain knowledge about the test-statistic landscape so that when data is taken, the steps to quoting a significance will be fewer, 2) to attain estimates of the significance that will be quoted once data is taken, and 3) to guide the choices being taken at this time in the project e.g. event selections and reconstruction resolutions, that will impact the achievable sensitivity. This section will described the specific procedure for such an analysis.

4.2.1 The Test-Statistic Distributions

Since we are dealing with composite non-nested hypotheses, we must take care that our handling of the sensitivity follows the prescriptions for a composite hypothesis. The critical value linked to a certain confidence level described in the previous section must give equal or better confidence for all free parameters in the hypothesis. The critical value of the test-statistic is found by setting the power to 0.5 for the median sensitivity. This means that we use the median of the "wrong" or *opposite* distribution to determine the critical TS-value. As discussed earlier this distribution will depend on the true parameters given that hypothesis, and we must therefore minimize the test-statistic over the free parameters. In practice, this is done by making a *fiducial-fit* defined as the best-fit of the *opposite* hypothesis to the true hypothesis being tested. The template from this fiducial fit will then be the opposite hypothesis looking most like the true hypothesis we are trying to reject,

$$TS_c^{true} = \min_{P \in opp} TS_c. \tag{4.9}$$

For a study like this project, which predates the data-taking process, we wish to quantify our believed sensitivity given either the Normal or Inverted Ordering. One of the orderings is manifested in nature and therefore the test-statistic we *will* measure once data is taken is believed to be a member of one of the two distributions $TS^{MO,true}$, $MO = NO, IO$. To quantify the separation between the two hypotheses we wish to populate these distributions for many possible realizations of the experiment. This is done by making data-sets from each of the orderings $D_0^{MO} = t(MO, P_0^{MO})$ representing the truth at some chosen set of parameters belonging to the respective true ordering P_0^{MO}. For each true data-set we make fits to both orderings and since this is MC generated pseudo-data, we know which hypothesis is true and which is wrong. The fit to the true ordering will give the true parameters the data was

made with because no fluctuations have been applied and the data can be completely matched by the hypothesis.

The fit to the opposite ordering will instead give a new set of parameters, P_f^{OP}, labeled with f for fiducial and OP for the opposite ordering to the true mass ordering MO. These parameters *will not be the same* as if we had produced a data-set assuming this opposite ordering using the same parameters as in producing the true ordering

$$P_0^{NO} \neq P_f^{NO}, P_0^{IO} \neq P_f^{IO}. \tag{4.10}$$

From the fiducial best fit parameters, we generate a fiducial data-set in the opposite ordering $D_{OP}^f = t(OP, P_f^{OP})$ and find the test-statistic for this data-set. It is important to stress that the fiducial data produced in an IO hypothesis from a true NO hypothesis will not, in general, be the same as would be produced from a true IO hypothesis using the same true parameters, $D_{NO}^0 \neq D_{NO}^f$. Following the frequentist method we wish to populate the distributions $TS^{MO,true}$ and $TS^{MO,opp}$ for many repeated experiments giving us the probability of obtaining a certain observed test-statistic TS_{obs} for each of the orderings. To do this we add Poisson fluctuations to the true and fiducial templates and calculate the test-statistic for that (random) realization of the experiment labeled with the subscript i,

$$TS_i^{MO,true} = TS\left(D_{MO,i}^0\right), \qquad TS_i^{MO,opp} = TS(D_{MO,i}^f). \tag{4.11}$$

Repeating this many, N, times essentially mimics carrying out the experiment that many times and is therefore referred to as *pseudo-experiments* or *pseudo-trials*. The full distributions become

$$TS^{MO,true} = \left\{TS_i^{MO,true}\right\}_{i=1}^{N}, \qquad TS^{MO,opp} = \left\{TS_i^{MO,opp}\right\}_{i=1}^{N}). \tag{4.12}$$

A schematic of this pseudo-trial process can be found in Figure 4.2. Making fiducial fits to properly account for the composite hypotheses encountered in oscillation experiments was introduced in IceCube in relation to PINGU [40] especially because of a degeneracy between the ordering and the octant of the atmospheric mixing angle θ_{23}. This will be discussed in detail in Section 5.2

It is important to note that other experiments do not make use of the same "double-fitting" method as IceCube and instead uses only the two distributions $TS^{NO,true}$ and $TS^{IO,true}$. This will also be discussed in the following chapters.

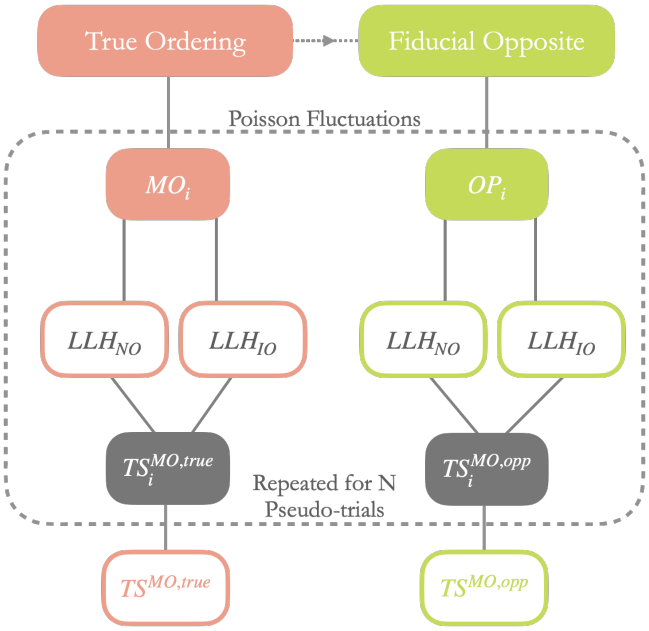

FIGURE 4.2.
A sketch of the analysis scheme used in this project. For a given true ordering, MO, a fiducial fit is made in the opposite ordering. Poisson fluctuations are added to both the true and fiducial templates giving new data-sets MO_i, OP_i. These are both fitted to the two orderings and the resulting test-statistic is saved. The steps inside the dashed line is repeated many times to populate the two distributions $TS^{MO,true}$ and $TS^{MO,opp}$ by re-applying Poisson fluctuations to the true and fiducial templates. The full process is done for both $MO = NO$ and $MO = IO$

4.2.2 The Asimov Approximation

In order to fully populate the test statistic distributions described in the previous section many pseudo-trials need to be run, requiring 4 fits for each pseudo-trial for each true ordering resulting in huge computation times. Furthermore are the tails of the distributions the important feature for our analysis while also being the hardest part of the distributions to populate. Given that a set of pseudo-trials is just for *one* set of true parameters P_0, it is very unattractive to repeat pseudo-trial experiments at many different event-selections, choices of systematic uncertainties and true oscillation parameters. To get such information we need to make some approximations about the shapes of the distributions.

Named after the author of the short story *Franchise* Isaac Asimov [60], the Asimov Approximation is the approximation that the median of a distribution is representative of the entire distribution [61]. Therefore we do not fluctuate the true and fiducial templates, but simply use the un-fluctuated templates and take the value of the test statistic obtained from these to be the median of the distributions. Labeling this with 0 in the subscript we define

$$TS_0^{MO,true} \equiv \text{median}\left(TS^{MO,true}\right) = TS\left(D^0_{MO,0}\right), \tag{4.13}$$

$$TS_0^{MO,opp} \equiv \text{median}\left(TS^{MO,opp}\right) = TS\left(D^f_{MO,0}\right). \tag{4.14}$$

When we have an un-fluctuated data-set created in some true ordering, we will be able to find the exact same template assuming that true ordering. This means that our test-statistic effectively reduces to a fit to the opposite ordering. For true NO the two test-statistics become $TS_0^{NO,true} = -LLH_{IO}(D^0_{NO,0})$ and $TS_0^{NO,opp} = LLH_{NO}(D^f_{NO,0})$,

where the fiducial data set is made under the assumption of inverted mass ordering using the best-fit values from the fiducial fit. For true inverted ordering the signs would change because we have defined our test-statistic as (4.6), and we will therefore use the notation of (4.13)-(4.14) for consistency.

Gaussianity

How to model the distributions in the Asimov approximation has been a matter of debate in the community with discussions on e.g. highly restrained parameter spaces [53] and the effects of nuisance parameters [55, 10]. In [10] it is shown that for a single free parameter, the test-statistic $T = \Delta\chi^2$ will be well-described by a gaussian with mean $\mu = T_0$ and standard deviation $\sigma = 2\sqrt{TS_0}$ if $T_0 \gg 1$. We do not satisfy $TS_0 \gg 1$ and have hypotheses that are functions of *multiple parameters*, but still choose to use the Gaussian approximation to follow previous work done in IceCube [38, 40, 57] and to compare to other experiments also applying this approximation. It would be preferred to make several pseudo-trial tests for different parameter-sets and compare the results of those tests with Gaussian approximations, but as time has been a limiting factor in the amount of feasible computations, this will only be done for a single realization of true parameters.

Median Sensitivity

Under the Gaussian approximation the test statistic distributions can be described by a normal distribution $\mathcal{N}(\mu, \sigma)$ with mean μ and standard deviation σ,

$$TS^{MO,hyp} = \mathcal{N}\left(2TS_0^{MO,hyp}, 2\sqrt{2TS_0^{MO,hyp}}\right), \tag{4.15}$$

where the superscript hyp can either be $true$ or opp and the true mass ordering MO either Normal or Inverted. This means that we can write up an analytical expression for α which we would otherwise have to simulate as described in the previous section. Using (4.15) we find

$$\alpha^{MO} = \frac{1}{2}\text{erfc}\left(\frac{\mu^{MO,true} - TS_c^{MO}}{\sqrt{2}\sigma^{MO,true}}\right) = \frac{1}{2}\text{erfc}\left(\frac{2TS_0^{MO,true} - TS_c^{MO}}{2\sqrt{4TS_0^{MO,true}}}\right). \tag{4.16}$$

To get the median sensitivity we choose $TS_c = \mu^{MO,opp} = 2TS_o^{MO,opp}$ and get

$$\alpha^{MO} = \frac{1}{2}\text{erfc}\left(\frac{2TS_0^{MO,true} - 2TS^{MO,opp}}{2\sqrt{4TS_0^{MO,true}}}\right), \tag{4.17}$$

$$= \frac{1}{2}\text{erfc}\left(\frac{TS_0^{MO,true} - TS^{MO,opp}}{2\sqrt{TS_0^{MO,true}}}\right) \tag{4.18}$$

Notice that this differs from eq. (3.4) in [10], where the denominator uses the mean of the opp and not $true$ distribution. After reaching out to the authors it has been confirmed that it is a typo and (4.17) is correct.

Taking the Asimov Approximation in a two-sided conversion gives you

$$\sigma_{2-sided} = \sqrt{2}\text{erfc}^{-1}(\alpha) = \sqrt{2}\text{erfc}^{-}1\left[\frac{1}{2}\text{erfc}\left(\frac{TS_0^{MO,true} - TS^{MO,opp}}{2\sqrt{TS_0^{MO,true}}}\right)\right]. \tag{4.19}$$

For a one-sided limited we get the more simple

$$\sigma_{1-sided} = \sqrt{2}\text{erfc}^{-1}(2\alpha) = \frac{TS_0^{MO,true} - TS^{MO,opp}}{\sqrt{2TS_0^{MO,true}}}. \tag{4.20}$$

In the limit where the two distributions for a given true mass ordering are symmetric around zero and therefore $TS_0^{MO,true} = -TS_0^{MO,opp} = TS_0^{MO}$ we find that the one-sided conversion gives the *standard* sensitivity one would expect from a nested hypothesis test, namely $\sigma_{simple} = \sqrt{2TS_0^{MO}}$.

The behavior of these different *interpretations* of the sensitivity will be discussed in Chapter 5.

5 Results

From the analysis described in the previous chapters a variety of different tests have been made to investigate both the nature of an NMO analysis and the specific case of the IceCube Upgrade's sensitivity to the Neutrino Mass Ordering. In this chapter, results on the impact of systematic parameters, the dependency of the sensitivity to θ_{23}, the validity of the Asimov approximation and projected sensitivities for the IceCube Upgrade will be presented.

5.1 Systematic Parameters

For any analysis, decisions have to be made about what effects to include and which to leave out and for this analysis it is no different. Because of the complexity of the full statistical analysis of the mass ordering, we have chosen to apply a simple *mis-modeling test* to the parameters described in Section 3.4.1 to determine which are important to keep as fit parameters and which can be fixed to their nominal value.

In this test each parameter is one by one pulled up 1σ from its nominal value for the production of pseudo-data and then fixed at its nominal value while a fit from a given true ordering to *the same ordering* is made. The value of $2LLH$ for this fit compared to when the parameter is not fixed is then compared to all other parameters giving a ranking of the maximal mis-modeling each parameter may cause. All parameters that cause a mis-modeling above 0.1σ are then kept free in the analysis. The result of this test for both orderings is shown in Figure 5.1, where gray bars are used for fixed parameters and for free parameters the mis-modelling for each tested ordering is shown with light-red for NO and light-blue for IO. We see that the non-atmospheric oscillation parameters (θ_{12}, θ_{13}, Δm_{21}) cause minimal mis-modeling and can be kept fixed at their nominal values shown in Table 3.1 for this analysis.

We also fix δ_{CP} to 0.0 as it has been found that it causes non-trivial local-minima in the full analysis while only affecting the obtained sensitivity by less than 0.1σ. The impact of δ_{CP} on the mass ordering is not unexpected as it describes whether or not neutrinos and anti-neutrinos mix in the same way. As the NMO signal in IceCube is either in the neutrino or anti-neutrino channel, including a parameter that can change

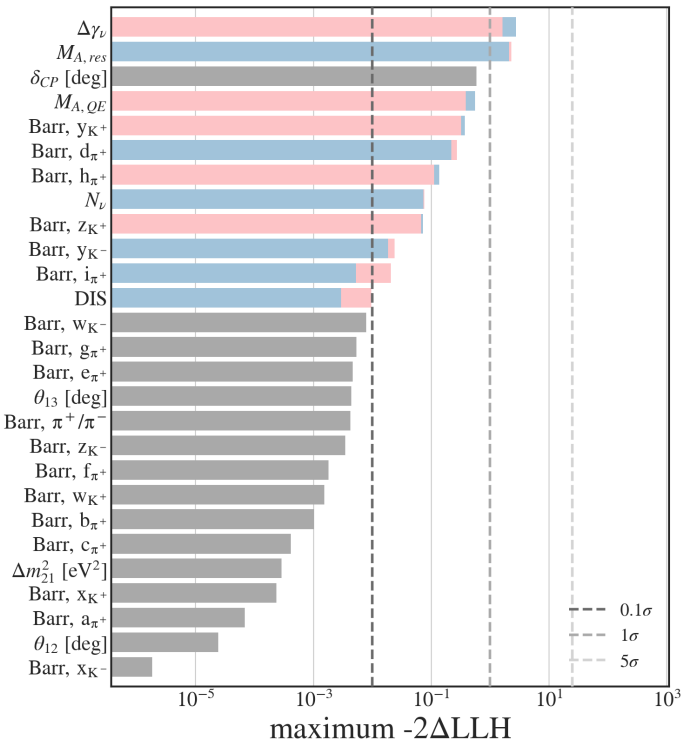

FIGURE 5.1.
Systematic impact test for both normal and inverted ordering. Each parameter is pulled 1σ from its nominal value for producing data and then fixed at its nominal value in a fit. Light-red bars indicate the resulting mis-modeling from a fit between true normal ordering and a hypothesis also of normal ordering. Light-blue bars show the same but for inverted ordering fitted to inverted ordering. Gray bars indicate parameters that are kept free in the rest of the analysis.

how these two contributions add up, but without being able to tell them apart the δ_{CP} and NMO signals intertwines.

5.2 Impact of θ_{23} on NMO Sensitivities

The number of events recorded in IceCube is not only a function of the mass ordering, but also several other both physics and *nuisance* parameters. The rather complicated method of mapping out the test-statistic distributions described in Chapter 4 is used to take into account the fact that these parameters may not have the same true values in the two different orderings. One parameter that causes some funny looking behavior is the atmospheric mixing angle θ_{23}.

θ_{23} can be understood as the oscillation *amplitude* effectively entering the oscillation probabilities as $\sin^2 2\theta_{23}$. The angle that would cause maximal mixing is $\theta_{23,max} = 45°$ but because $\sin^2$ is a symmetric function, values just above and just below $45°$ will lead to similar oscillation amplitudes which makes it difficult for experiments such as IceCube to determine the *octant* of θ_{23}.

This also forces us to be aware of how the minimization is done in practice as there will be local minima in *both* octants. This is handled by minimizing the log-likelihood separately for θ_{23} values in each octant (below and above $45°$), and then choosing the fit of those two with the lowest log-likelihood as the "true" best-fit. It has also been

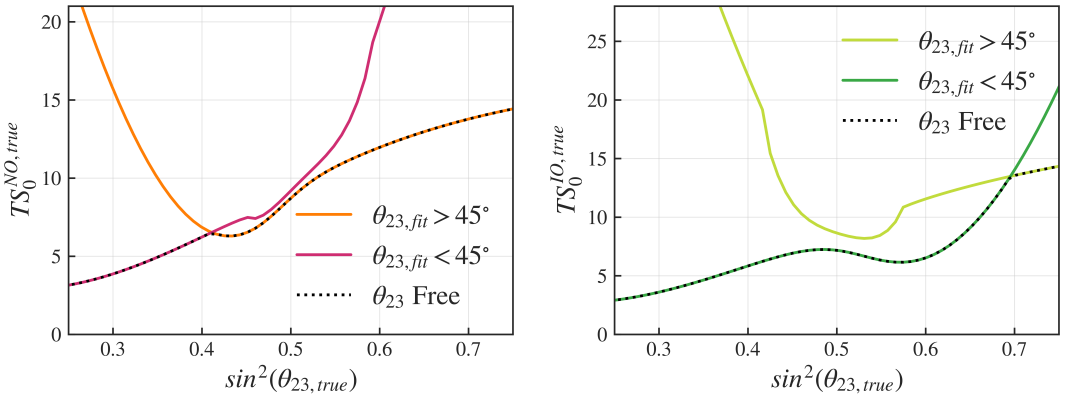

FIGURE 5.2.
The Asimov true test-statistic for different allowed regions of the fitted value of θ_{23} as a function of the true value. (Left) True Normal Ordering, (Right) True Inverted Ordering.

well established in the neutrino oscillation community that the octant degeneracy should be highlighted and not hidden in the formulation $\sin^2 2\theta_{23}$, and this term is therefore expressed in a combination of $\sin^2 \theta_{23}$ and $\cos^2 \theta_{23}$ and the value of $\sin^2 \theta_{23}$ fitted - without the factor of two inside.

5.2.1 Degeneracy with NMO

When this analysis tries to determine which ordering is the *true* ordering, we effectively test how much we could make the *wrong* ordering look like the true one by tuning the parameters our results are impacted by. As discussed in Section 3.5 the effective *signal* of the mass ordering is a scaling of the number of oscillated events in specific areas of the θ_{23} and energy range of our data. This means that comparing data from normal ordering to an inverted ordering template, the higher number of events in the oscillation bands in the low energy, could to some degree be "washed out" by turning up the oscillation amplitude θ_{23} in the inverted ordering template. The opposite effect is the case for fitting a normal ordering template to data from the inverted ordering, and this is why we have a *degeneracy* between the true value of θ_{23} and the true mass ordering.

5.2.2 Octant Flipping

This degeneracy can be seen in the value of the test-statistics obtained for the two orderings. In Figure 5.2 the Asimov test-statistic values for true mass ordering are plotted as a function of the true value of $\sin^2 \theta_{23}$. For true normal ordering we see that the point where the best-fit points in the two octants cross is lower than maximal mixing whereas for true inverted ordering the cross-over is located above maximal mixing. These crossings causes the lowest possible test-statistic to switch from one

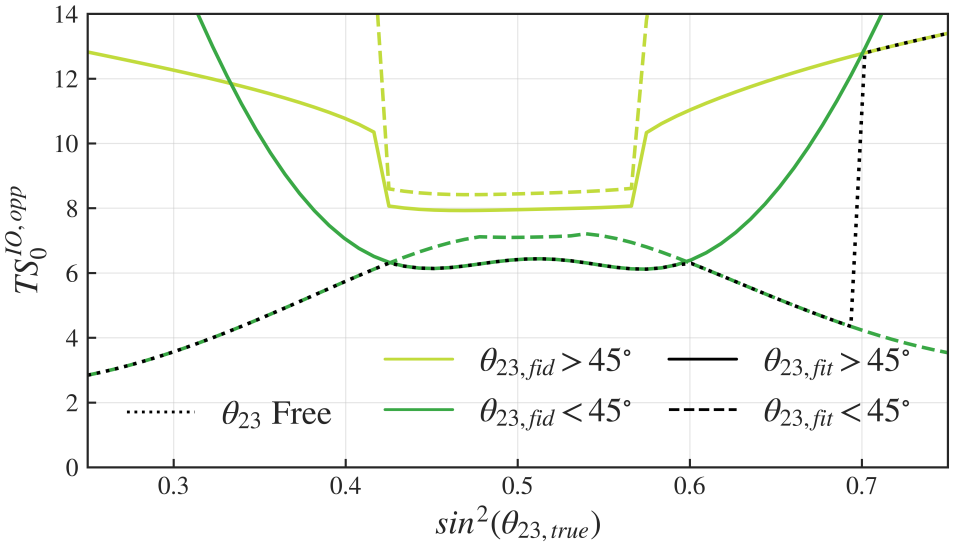

FIGURE 5.3.
The Asimov *opposite* test-statistic for true Inverted Ordering for different allowed regions of θ_{23} in both the fiducial fit and in the $TS_0^{IO,opp}$ fit.

octant to the other and introduces sharp bends in the *free* test-statistic that would otherwise not be expected.

For both orderings we also see that the test-statistic restricted to the upper octant for NO and lower octant for IO exhibit sharp bends. This can be shown to be because of the degeneracy discussed earlier, as both of these restrictions are in the octant *not* looking like the opposite ordering and therefore force the fitted value of θ_{23} to $45°$ as the true best-fit point is in the octant not allowed. This edge-effect is only seen for $\theta_{23,true}$ values around maximal mixing.

The effect of octant flipping in the *true* test-statistic is also visible in the *opposite* test-statistic. The fiducial template in the opposite ordering is created from the best-fit value to this ordering, which is exactly the fit that provides the lowest *true* test-statistic. This means that the fiducial fit from which $TS_0^{MO,opp}$ is found is made in either the lower or the upper octant depending on which gives the lowest value of $TS_0^{MO,true}$. This can be seen in 5.3 where the value of the asimov test-statisctic $TS_0^{MO,opp}$ is shown for different cases of allowed θ_{23} in the fiducial fit and the fit for $TS_0^{MO,opp}$. The striking jump in the upper octant of the case where θ_{23} is free in both fits is exactly because of the *choice of octant* implicitly embedded into the fiducial fit. Even though the lower octant in both fits would provide the lowest value of $TS_0^{MO,opp}$, the fiducial fit *forces* the template to be created in the upper octant. The same behavior can be found for true normal ordering but because the lower octant is preferred for fiducial fits in both octants, the flipping does not cause a jump as seen in the case for true IO, but only a sharp kink. The behavior is translated into the sensitivities and the sensitivity for true inverted ordering therefore also show this *jump* in the upper octant

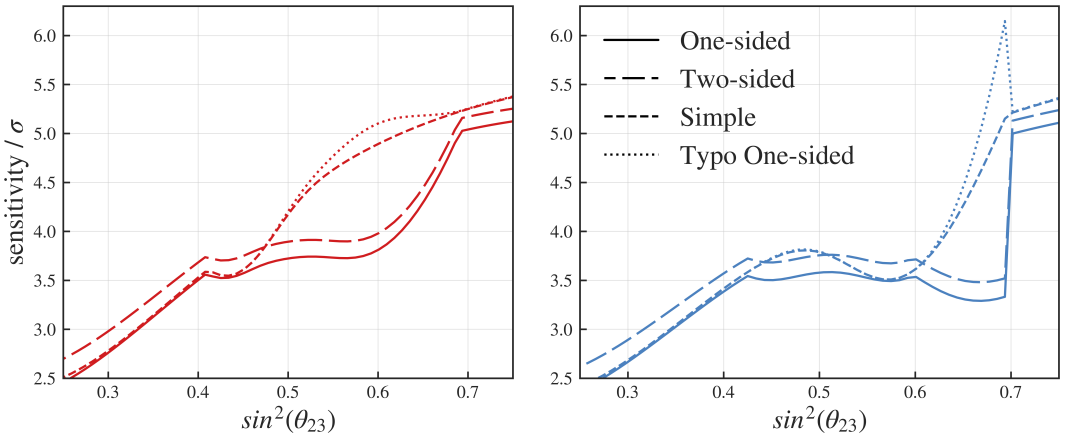

FIGURE 5.4.
Sensitivity to the Neutrino Mass Ordering using different definitions defined in Section 4.2.2. The linestyles corresponds to the same labels in both panels. Left panel is true Normal Ordering, right panel is True Inverted Ordering.

while for true normal ordering it only shows up as a kink in the lower octant. Exactly *how* the effect shows up in the sensitivity depends on how it is calculated, but is has been found that the effect is not restricted to Asimov analyses but is also present in the pseudo-trial method described in Figure 4.2.

5.3 Asimov Definitions

The sensitivity obtained from the test-statistic values is derived for different assumptions in Section 4.2.2. Applying these definitions to the asimov test-statistic values we obtain by letting θ_{23} vary freely in both fits provides the sensitivities shown in Figure 5.4 for true normal and inverted ordering. The one-sided and two-sided definitions give similar results but with slightly higher values for the two-sided conversion for all tested true θ_{23} values. This is consistent with the underlying differences between a one-sided and two-sided conversion as for a two-sided conversion the obtained α is divided between both sides of a normal Gaussian giving a higher confidence than if the obtained α is in effect put in both tails of the Gaussian.

We also see that the simple approximation of symmetry around $TS = 0$ causes higher sensitivities in the upper octant for NO and lower octant for IO. By only using the value of the *true* test-statistic this definitions misses out on the octant degeneracy experienced by the opposite ordering and therefore overestimates the separation between the two distributions.

A one-sided conversion using the typo described in Section 4.2.2 is also plotted for completeness. Here we see that properly including the width of the *true* distribution

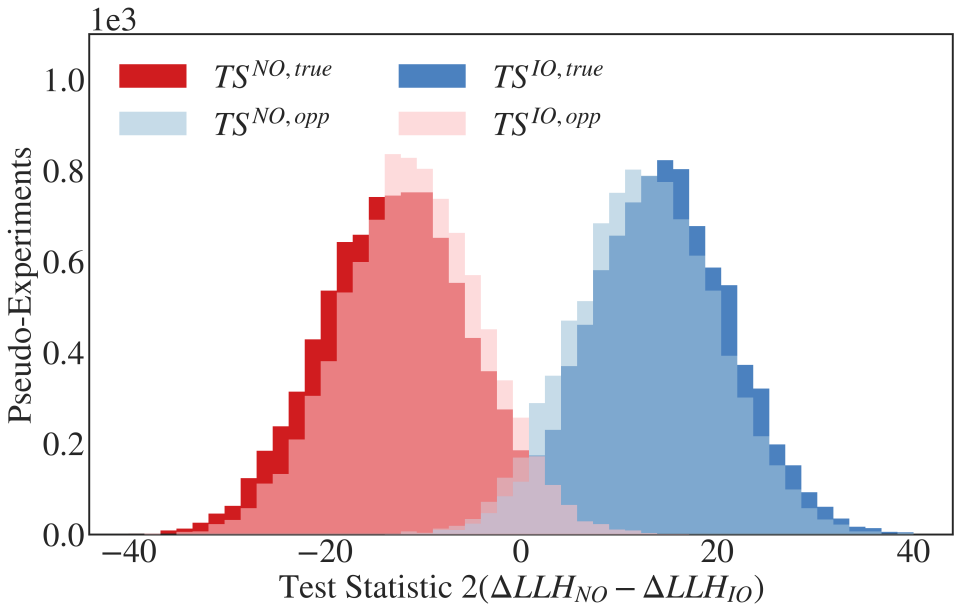

FIGURE 5.5.
Pseudo-trial distributions for the true (solid) and opposite (see-through) test statistics for both true Normal (red) and true Inverted Ordering (blue). The true templates are created at $\sin^2(\theta_{23}) = 0.45$ and the rest of the parameters at the nominal values listed in Table 3.1.

and not the *opposite* describes the octant degeneracy whereas the typo causes the sensitivity to seem closer to the simple definition. The rather unnatural looking "spike" for true inverted ordering is the effect of the octant flipping described in the previous section, also present for the correct sensitivities. The reason for the spike to be turned down in the correct sensitivities is that as the value of $TS_0^{NO,opp}$ jumps, as can be seen in Figure 5.3, the distance between the distributions grow and the sensitivity therefore increases.

5.4 The Test-Statistic Distributions for the Upgrade

The key to determining the sensitivity to the neutrino mass ordering is describing the test-statistic distributions we expect to measure. As described in Section 4.2 the distributions one wishes to compare is a matter of choice. Following the arguments for a composite hypothesis test, a certain true ordering $TS^{MO,true}$ should be compared not to the different true hypothesis, but to the opposite hypothesis that looks most like the true hypothesis. The four different test-statistic distributions $TS^{MO,true}$, $TS^{MO,opp}$, $MO \in NO, IO$ are shown in Figure 5.5 for a single set of parameters P. We see that the solid distributions describing the *true* distributions at this point in the parameter-space are in fact further away from $TS = 0$ than the fiducial *opposite* distributions $TS^{MO,opp}$. This difference can also be seen in the Asimov approximation

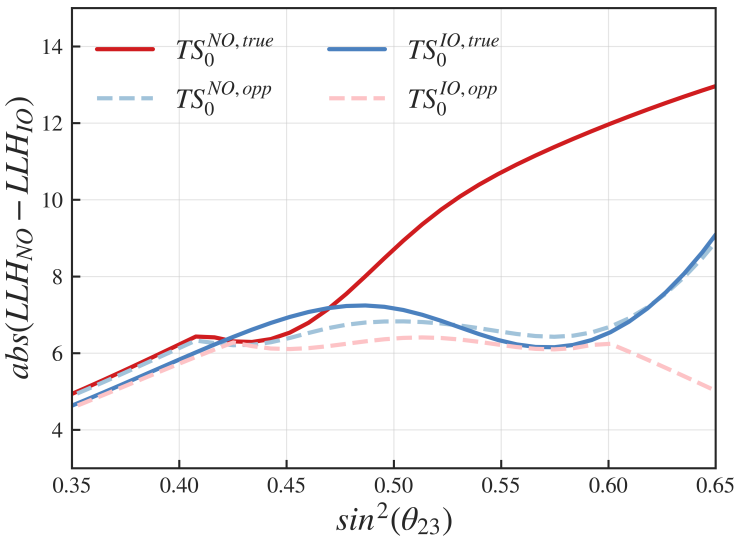

FIGURE 5.6.
Absolute values of the Asimov predictions of the test-statistic distributions in Figure 5.5 as a function of the true value of θ_{23}.

by plotting the Asimov median values of the four distributions from Figure 5.5 for multiple points of the true parameter space. A scan over true values of $\sin^2(\theta_{23})$ can be seen in Figure 5.6 where the absolute values of the test statistic is plotted so that lower values correspond to less separation between the orderings. It is found that at $\sin^2(\theta_{23}) = 0.45$ where the pseudo-trial in Figure 5.5 is run, also the Asimov approximation shows that the value of $TS_0^{NO,opp}$ is in fact closer to zero than simply using $TS_0^{IO,true}$. It is, however, also found that for a short range of values between $\sin^2(\theta_23) \approx 0.55 - 0.60$ the *true* opposite distribution is closer to zero. From running a limited number of pseudo-experiments at $\sin^2\theta_{23} = 0.57$, see Figure 5.7, these results have been confirmed showing a small difference between the means of $TS_0^{NO,wrong} - TS_0^{IO,true} = 0.3$.

The correction from using $TS_0^{MO,opp}$ is not very large for true Normal Ordering in the lower octant where the solid blue and dashed light-blue lines are close together, but a big difference can be seen in the upper octant for true Inverted Ordering. Here the solid red line describing $TS_0^{NO,true}$ is much larger than the corresponding dashed light-red line describing the case where a fiducial fit has been made and this is also confirmed by the pseudo-trial shown in Figure 5.7.

From the pseudo-trials in Figure 5.5 at $\sin^2\theta_{23} = 0.45$ we can calculate a sensitivity based on how many of our "pseudo-experiments" resulted in a test-statistic more extreme than the median of the opposite distribution. For the case of true Normal Ordering we find that this is the case for $1:9000$ experiments. Converting this value of α to a sensitivity using (4.8) we find

$$\sigma_{pseudo}^{NO} = 3.40^{+0.03}_{-0.03}, \qquad \sigma_{pseudo}^{IO} = 3.51^{+0.08}_{-0.06}, \tag{5.1}$$

using an uncertainty on the number of experiments n in α of $1/n$.

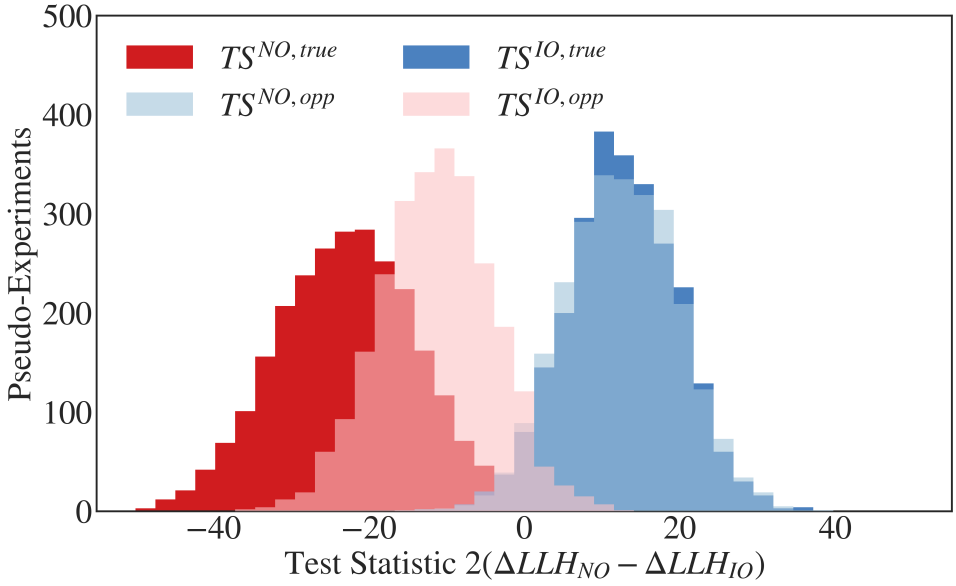

FIGURE 5.7.
Pseudo-trial distributions for the true (solid) and opposite (see-through) test statistics for both true Normal (red) and true Inverted Ordering (blue). The true templates are created at $\sin^2(\theta_{23}) = 0.57$ and the rest of the parameters at the nominal values listed in Table 3.1.

5.5 Asimov Sensitivity Scans

Using the Asimov approximation described in Section 4.2.2 we are able to quantify the sensitivity for multiple different possible scenarios. We will here present sensitivities as a function of both the true value of θ_{23} and the *livetime* of the detector in number of years.

5.5.1 Consistency With Pseudo-Trials

In Figure 5.8 the estimated sensitivity after 3 years of livetime is shown for a range of $\theta_{23,true}$ consistent with the expected range from current global best-fits [23]. These sensitivities are found using the free parameters from Figure 5.1 and the nominal values of all other parameters as stated in Table 3.1. No detector systematics are included which would be expected to lower the sensitivity. We see that with the correct one-sided conversion from significance to sensitivity the impact of the true value of θ_{23} is minimal in the expected range.

At the point of $\sin^2\theta_{23} = 0.45$ which was analyzed with rigorous pseudo-trials, the sensitivity from the Asimov approximation is slightly higher for Normal Ordering but consistent within the estimated errors of Inverted Ordering. From computational limitations the pseudo-trials at $\sin^2\theta_{23} = .57$ have not been converted to sensitivities.

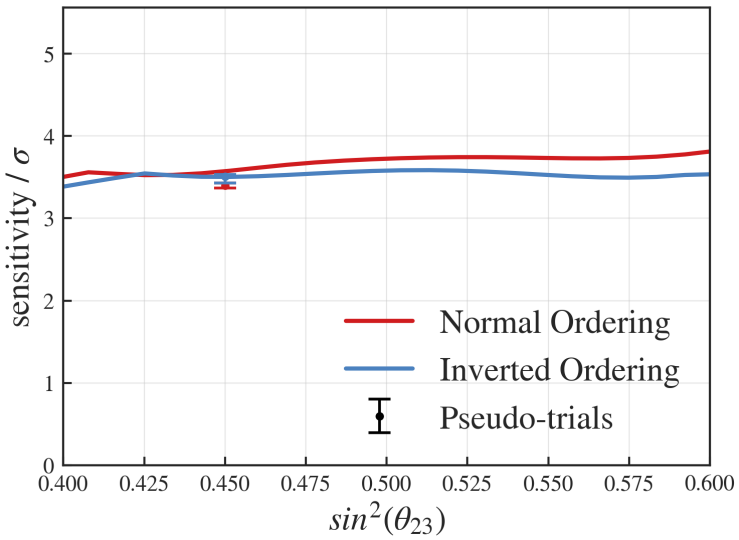

FIGURE 5.8.
Estimated sensitivity to the Neutrino Mass Ordering for 3 years data-taking of the IceCube Upgrade. A one-sided conversion is used to obtain the sensitivity as stated in eq. (4.20). Also shown is the sensitivities as found by the Pseudo-experiments of Figure 5.5.

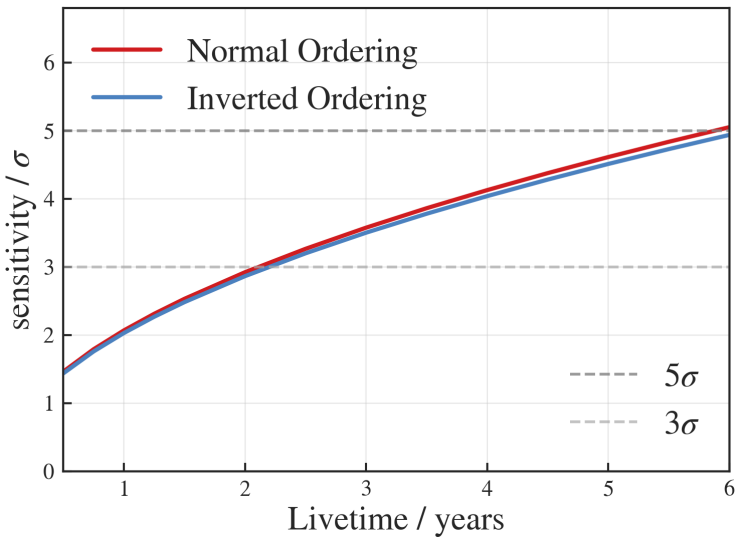

FIGURE 5.9.
Sensitivity to the Neutrino Mass Ordering as a function of livetime for the IceCube Upgrade. The sensitivities are calculated using the settings described in Section **??** and for a single value of the atmospheric mixing angle $\theta_{23} = 42.2°$ which is the NuFit best-fit value assuming normal ordering. Consistent results were obtained using the best-fit values assuming inverted ordering.

5.5.2 Livetime

The sensitivities in this analysis are estimated from Monte-Carlo data which is handled as described in Chapter 3. This can be done using different values of the active *livetime* of the detector and we can therefore estimate when the expected sensitivities will cross different milestones. Such an analysis is shown in Figure 5.9 for a single value of the true atmospheric mixing angle θ_{23}. The value has been chosen to be the NuFit [24] best-fit value under normal ordering and it has been checked that consistent results are achieved using the best-fit value under the inverted ordering as expected from the minimal θ_{23} dependence found in Figure 5.8. It is found that for this analysis, IceCube Upgrade will reach 3σ sensitivity to both orderings within 3 years and 5σ in about 6 years.

6 Discussion

The discussion of the results presented in Chapter 5 is split into three sections - the problems involved with making a statistical test for the Neutrino Mass Ordering, what such a statistical test can tell us about the IceCube Upgrade's ability to separate the two orderings, and lastly how that relates to current and future NMO sensitivities.

6.1 NMO Statistics in General

6.1.1 The Frequentist's Bane

The sensitivities presented in this project are limited by their origin in the frequentist methodology. As has been pointed out by multiple authors [53, 55] the frequentist approach has the undesirable downside that it cannot answer the question "Which Mass Ordering is the True Ordering?". And within this framework can we then even say *anything* about the mass ordering? The answer to that question is obviously yes - if not, frequentist statistics would not be as used as it is today. It does however set very high standards for us when we share our results with others since such statistical results can easily be misinterpreted to signify more than it actually does. These problematic sides of statistics and in particular the choice of being *bayesianist* or a *frequentist* is not new nor confined to the problem of the Neutrino Mass Ordering [1, 2, 62].

The subtle but impactful differences between the framework of bayesian and frequentist statistics may however have exceedingly good circumstances for unfolding in full bloom in an analysis like the NMO. Normally when a *standard* frequentist physicist sets up a statistical test for some hypotheses they are faced with the difficult problem of setting a critical value of their test-statistic, but they are not faced with having an alternative hypothesis *that must be true* to compare their experiment to. This makes it very difficult for any reasonable person to not conclude that from rejecting one hypothesis the other must be true. But this can simply *not* be inferred from a frequentist test. It would otherwise be very reasonable to state that an experiment's sensitivity to Normal Ordering is determined by how well it will be able to reject Inverted Ordering - if Normal Ordering is in fact true then we should be able to reject

Inverted Ordering. But since we only construct a test of whether or not we can reject that Inverted Ordering is true, we cannot from that test gain any knowledge about whether or not Normal Ordering is true. Therefore, the frequentist "sensitivities" in this project are labeled by the ordering for which we would expect to reject it with the stated significance.

From the considerations about a frequentists approach to the neutrino mass ordering a natural next step would be to turn to bayesian statistics. Though it has been outside the scope of this project the two frameworks are expected to give complementary information and it would therefore be interesting to compare the results of each. Such results should however clearly be separated and not confused with eachother as one will describe how well the respective measurement is consistent with the two mass ordering (frequentist) whereas the other will describe how this measurement will influence our belief in which ordering is the true ordering manifest in nature (bayesianist).

6.1.2 Discovery, Evidence or Nothing?

Working within the frequentist framework, what should we take away from the values of our sensitivities? In this and other NMO studies the obtained significance is converted to a number of standard deviations *as if* the test-statistic had been normally distributed. The argument for this is related to the convention of when a measured signal is assumed to be evidence or discovery of some underlying theory. This conversion forces a choice of setting either a central-limit or an upper/lower limit with no direct relation to the measured quantity which could cause more confusion as such a result may look like one would expect from a non-nested simple hypothesis. It could also wash away the possibility of physics beyond the standard model (BSM) as a measured test-statistic lying further away from zero than any of the two orderings would not be taken as sign of inconsistency with the preferred ordering. We would argue that when converting to a number of σ it should be made very clear why the analysis *itself* is not nested. We would furthermore argue that a one-sided limit should be stated as to not misleadingly present a non-zero significance for total confusion between the two orderings as described in Section 4.1.4.

The convention is to state a discovery when background can be rejected with more than 5σ confidence and evidence when this confidence is above 3σ but below 5σ. If we were to measure a value of the test-statistic which rejects Normal Ordering with 3σ confidence but also rejects Inverted Ordering with 2.5σ, what should the take-away be from such a measurement? As IceCube has not previously been expected to separate the orderings in such a significant way, the nuances of a possible measurement has only

recently begun to surface. With increasingly sensitive experiments the complications of the non-nested nature of the analysis becomes more crucial.

In Blennow et. al [10] it is proposed that a different sensitivity than the *median sensitivity* used in this project, could be made to include the sensitivity to *both* orderings by choosing the critical value of the test-statistic satisfying that only *one* of the orderings can be rejected at that value. This test however does not easily translate to a composite hypothesis, because as we have seen, using the two *true* test-statistic distributions to determine the separation actually overestimates the sensitivity. This means that setting a critical test-statistic value that rejects only one ordering is a complicated task and has been outside the scope of this project. Previous work in IceCube [38, 57, 40, 59] has used a CL_s method based on [63] to describe how much *more unlikely* an observed value is under the hypothesis least favored compared to the favored. Such an approach could be relevant for this analysis as well and should be considered in future work.

6.1.3 The Impact of the True Parameter Space

The composite nature of the hypotheses does not only complicate setting a critical value of a test-statistic but it also complicates many otherwise straightforward statistical tests. We have for example seen that the impact of allowing θ_{23} to have different true values in the two orderings causes un-intuitive behavior of the resulting sensitivity. It could therefore be discussed whether the approach of letting a parameter such as θ_{23} be free is the correct approach or if one should set up tests for each possible combination of the octant of θ_{23} and state those sensitivities as complementary to each other. The problem with that approach is how to decide between and interpret the possible ways of restricting θ_{23}. If we knew the true octant, such a choice would be simple as we could restrict the octant in both the fiducial and other fits as that octant would be true in both orderings. But as it is now, we do not know the octant and restricting it to either the lower or the upper octant in the full analysis would not be sufficient, as the true value could be different in the two orderings. The best choice therefore seems to state the lowest possible sensitivity accounting for the possible confusion between ordering and octant until such a time where the octant has been determined.

6.2 NMO with the IceCube Upgrade

6.2.1 The Trouble of the Octant of θ_{23}

As described in Section 4.2 and the previous section, this analysis uses a *fiducial* fit to account for the fact that the two mass orderings may have different true parameters associated with them. The impact of this has been found to be quite drastic in some regions of the expected value of $\sin^2\theta_{23}$. As was found from the pseudo-trials at $\sin^2\theta_{23} = 0.57$ shown in Figure 5.7 the fiducial fit from true Inverted Ordering to Normal Ordering *halves* the absolute value of the critical test-statistic used to calculate the sensitivity. Though the corrections due to a fiducial fit may be small in the lower octant this result shows that not accounting for the composite nature of the non-nested hypotheses overestimates the sensitivity in the upper octant and in the expected range true value of $\sin^2\theta_{23}$. The global NuFit 3σ range is around $0.408 - 0.603$ or $0.412 - 0.613$ depending on the assumed mass ordering corresponding roughly to the x-axis of Figure 5.8.

Due to computational limitations in the number of pseudo-trials conducted, it has only been possible to carry out a rigorous investigation of the test-statistic distribution at one set of true parameters at $\sin^2\theta_{23} = 0.45$. Even though the Asimov approximation at this point seems to deliver reasonable results, more pseudo-trials should be run at other points in the parameter space.

We have found that for the Gaussian approximation in the expected range of θ_{23} the sensitivity is minimally affected by the true value, but this is only because of the inclusion of the fiducial fit as explained above. This can be contrasted to the other Asimov sensitivity definitions explained in Section 4.2.2 and shown in Figure 5.4. We find that assuming the test-statistics are symmetric around the test-statistic $TS = 0$ (the Simple sensitivity) overestimates the sensitivity by up to 1σ in the upper octant for true Normal Ordering. This definition of the Asimov sensitivity does not utilize the fiducial fit as it assumes symmetry and only relies on the median value of the true test-statistic considered. Symmetry between the two *true* distributions has only been found to be valid for true values of $\sin^2\theta_{23} < 0.4$ from the Asimov approximation of the medians in Figure 5.6. If we consider the *opposite* distributions we find that symmetry between $TS_0^{NO,true}$ and $TS_0^{NO,opp}$ is valid for $\sin^2\theta_{23} < 0.45$ and symmetry between $TS_0^{IO,true}$ and $TS_0^{IO,opp}$ is valid for $0.55 < \sin^2\theta_{23} < 0.6$. It would therefore be wrong to say that the symmetric approximation is valid for IceCube Upgrade based on this analysis.

6.2.2 The Early Stages of the Analysis

The parameters that have been chosen to be handled as fit parameters in this analysis shown in Figure 5.1, are based on simple mismodeling tests within the two mass orderings and their influence on the final result of the analysis are therefore not known. As seen with δ_{CP} which had a high level of mismodeling within both orderings but a negligible impact on the final sensitivity, the simple test does not directly translate to the full analysis. This makes sense as the simple mismodeling test describes how much impact a parameter will have on finding the initial true parameters of the template. A parameter's impact on the NMO signal on the other hand should be described by how much it can make the two orderings look more or less like each other. The choice of fit parameters should therefore be reconsidered using a more sophisticated method of ranking their impact on the sensitivity.

This analysis has no detector uncertainty systematics included, as their implementation is also complicated by the non-nested nature of the analysis. Because detector uncertainties, like the efficiency of the DOMs and the refractive index of the ice, affect the detection of neutrinos in non-trivial ways, their impact has to be simulated and then applied to the events considered. This has to be done for both orderings and it will again lead to non-trivial ways in which a given uncertainty may pull one mass ordering closer or further away from the other. Implementation of detector systematics are expected to lower the sensitivities as they should make it easier to make the orderings look like each other, by accounting for some of the discrepancies that could otherwise only be attributed to the mass ordering.

6.3 Perspectives

6.3.1 The Impact of the Upgrade's Sensitivity

The sensitivities presented in this project are very promising, even though as discussed in the previous sections, inclusion of detector systematics are expected to result in lower sensitivities. Early studies indicate that such corrections will be less than 1σ and in that case these results suggest that the IceCube Upgrade will be able to reject the wrong mass ordering with 3σ confidence within four years. These results are somewhat consistent with the expected sensitivity of the proposed, but not funded, more ambitious IceCube upgrade PINGU [40]. This is somewhat surprising as PINGU was proposed to consist of 26 new strings and would be expected to perform better than the less dense Upgrade. One possible explanation for this can be linked to the

new optical modules designed for the Upgrade which were not part of the PINGU simulation and therefore not included in the sensitivities stated. The effect of not including detector systematics and background muons is also expected make the results of this analysis higher than analyses including such effects. It should also be noted that the sensitivities in [40] as well as other previous IceCube NMO studies [59] have inherited the typo discussed in Section 4.2.2 causing the θ_{23} dependence to be similar with the "Typo One-sided" sensitivity shown in Figure 5.4.

For the present analysis an Upgrade bespoke event-selection has been used that effectively removes both noise and atmospheric muon-background. Current efforts are furthermore put into the inclusion of *inelasticity* in the reconstruction of neutrino events. The inelasticity of an event describes the amount of energy transferred to the track- and cascade-portion of the interaction respectively. Since neutrinos and anti-neutrinos have slightly different inelasticity distributions including it in the reconstruction could possibly enable IceCube to somewhat distinguish between neutrino and anti-neutrinos and therefore enhance the NMO signal [64].

6.3.2 Combining Experiments

The high level of sensitivity that the IceCube Upgrade will be able to attain on its own is very promising. Besides adding to the already obtained significance of the IceCube DeepCore volume, combined significance of multiple experiments are also a possibility. Synergy effects from combining experiments are expected to significantly improve joint sensitivity. Combining sensitivities from IceCube and the Jiangmen Underground Neutrino Observatory (JUNO) have previously been explored [65] with synergy effects arising from the different ways the mass ordering manifests in the two experiments. JUNO is a medium-baseline liquid scintillator experiment that measures MeV electron-antineutrinos from nuclear reactors $\sim 53\,\mathrm{km}$ from the detector. At these energies and baselines the NMO signal is a short wavelength correction to the electron-antineutrino survival probability and its detection relies on precise energy spectrum measurements [66]. The synergy effect from combining the two experiments comes from differences in how the *wrong* ordering best-fit parameters will pull to make the hypothesis look most like the true ordering. In effect combining the two experiments confines some of the allowed parameter space in the wrong ordering making it harder to attain test-statistic values on similar levels with the true ordering. Even though previous examinations of combined sensitivities have used the symmetric assumption of the test-statistics proven to not be appropriate for the IceCube Upgrade, combining the two experiments are still expected to significantly boost the global sensitivity to the Neutrino Mass Ordering.

7 Conclusion

This project has compared different methods of quantifying the sensitivity of an experiment to the Neutrino Mass Ordering and applied them to simulated data of the IceCube Upgrade.

It has been found that a correction to the Gaussian approximation of the median sensitivity defined in [10] causes a lower sensitivity to be obtained in the upper octant for Normal Ordering and in the lower octant for Inverted Ordering as seen in Figure 5.4. This correction properly accounts for the degeneracy between the NMO and the octant of θ_{23} discussed in detail in Sections 5.2, 6.2.1. The impact of the true value of θ_{23} on the size of the sensitivity has been found to be minimal in the global fit expected range, but only if a fiducial fit is made between the true ordering being tested and the opposite ordering. In the upper octant for true Inverted Ordering the test-statistic distribution obtained for the normal ordering is significantly further away from the distribution for true inverted ordering if the pseudo-data in the normal ordering is produced using the same parameters as the true inverted ordering pseudo-data, see Figure 5.7. It can therefore be concluded that the method of applying a *fiducial* fit to the true ordering being tested gives a more well-founded sensitivity result as it uses the lowest possible value of the critical test-statistic.

Using the fiducial fit introduces shape to the Asimov approximation defined in Section 4.2.2 which would otherwise be expected to be smooth. It has been found that the shapes arise from the octant degeneracy and can be explained by considering in which octant the fiducial fit finds the best-fit value. For the simulated data used in this project the effects of the octant degeneracy on the Asimov sensitivity have been found to not over- or under-estimate the sensitivity.

From extensive pseudo-trials at $\sin^2\theta_{23} = 0.45$, the Asimov approximation has been found to provide reasonable results. This point in the parameter space was chosen as it is the global best-fit point assuming Normal Ordering from NuFit [23], other points should however be investigated using the same approach before a conclusion on the validity of the Asimov approximation for the IceCube Upgrade can be made.

Using the a fiducial fit method to estimate the Asimov sensitivity of the IceCube Upgrade, it has been found that a 3σ sensitivity will be reached within three years

if no detector uncertainties are applied. Even with the inclusion of such systematic uncertainties the Upgrade is expected to reach 3σ in a comparable time-frame to what was proposed with PINGU [40]. From initial analyses on the impact of δ_{CP} on the obtained sensitivity it is suggested that a more sophisticated method of ranking parameters' impact should be derived.

From considerations on what a sensitivity means and how underlying assumptions on what a probability is, it has been found that no satisfying answer to "How well will the IceCube Upgrade be able to determine the NMO?" can be given within the frequentist framework. The sensitivities found in this project are therefore indicative of, with which confidence should we be able to reject the given mass ordering. Rejecting the Normal Ordering does not mean the Inverted Ordering can be accepted and vice versa. It is therefore the recommendation that sensitivity analyses, like the one at hand, performed before an experiment has started collecting data, should be treated as indicative *only* of how well data may be used to say something about the Mass Ordering. Once data has been taken other statistical methods will need to be utilized to obtain a measure of preference for one ordering over the other.

Bibliography

8

[1] R. J. Barlow. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*. John Wiley & Sons, Dec. 1993.

[2] Robert D. Cousins. "Why Isn't Every Physicist a Bayesian?" In: *American Journal of Physics* 63.5 (May 1995), pp. 398–410.

[3] Pantone. *PANTONE® USA | Fashion Color Trend Report: New York Fashion Week Spring/Summer 2023*. https://www.pantone.com/articles/fashion-color-trend-report/new-york-fashion-week-spring-summer-2023.

[4] Guillermo Franco Abellán, Zackaria Chacko, Abhish Dev, Peizhi Du, Vivian Poulin, and Yuhsin Tsai. "Improved Cosmological Constraints on the Neutrino Mass and Lifetime". In: *Journal of High Energy Physics (Online)* 2022.FERMILAB-PUB-21-779-T; arXiv:2112.13862 (Aug. 2022).

[5] Mark Thomson. *Modern Particle Physics*. Illustrated edition. Cambridge, United Kingdom ; New York: Cambridge University Press, Oct. 2013.

[6] Particle Data Group, R L Workman, V D Burkert, *et al.* "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2022.8 (Aug. 2022), p. 083C01.

[7] Aya Ishihara. *The IceCube Upgrade – Design and Science Goals*. Aug. 2019. arXiv: 1908.09441 [astro-ph, physics:physics].

[8] The IceCube Collaboration. "The Design and Performance of IceCube DeepCore". In: *Astroparticle Physics* 35.10 (May 2012), pp. 615–624.

[9] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *The Annals of Mathematical Statistics* 9.1 (Mar. 1938), pp. 60–62.

[10] Mattias Blennow, Pilar Coloma, Patrick Huber, and Thomas Schwetz. "Quantifying the Sensitivity of Oscillation Experiments to the Neutrino Mass Ordering".

In: *Journal of High Energy Physics* 2014.3 (Mar. 2014), p. 28. arXiv: 1311.1822 [hep-ex, physics:hep-ph].

[11] Guido Fantini, Andrea Gallo Rosso, Francesco Vissani, and Vanessa Zema. "The Formalism of Neutrino Oscillations: An Introduction". In: *arXiv:1802.05781 [hep-ph]* (Feb. 2018). arXiv: 1802.05781 [hep-ph].

[12] Richard Slansky, Stuart Raby, Terry Goldman, and Gerry Garvey. "An Introduction to Neutrino Masses and Mixings". In: *Los Alamos Science* 25 (1997). Ed. by Necia Grant Cooper.

[13] B. R. Martin. "Electroweak Interactions". In: *Nuclear and Particle Physics*. John Wiley & Sons, Ltd, 2006, pp. 181–216.

[14] J. A. Formaggio and G. P. Zeller. "From eV to EeV: Neutrino Cross Sections across Energy Scales". In: *Reviews of Modern Physics* 84.3 (Sept. 2012), pp. 1307–1341.

[15] Alessandro De Angelis and Mário Pimenta. "The Properties of Neutrinos". In: *Introduction to Particle and Astroparticle Physics: Multimessenger Astronomy and Its Particle Physics Foundations*. Ed. by Alessandro De Angelis and Mário Pimenta. Cham: Springer International Publishing, 2018, pp. 543–574.

[16] Raymond Davis. "A Review of the Homestake Solar Neutrino Experiment". In: *Progress in Particle and Nuclear Physics* 32 (Jan. 1994), pp. 13–32.

[17] A. Bellerive, J. R. Klein, A. B. McDonald, A. J. Noble, and A. W. P. Poon. "The Sudbury Neutrino Observatory". In: *Nuclear Physics B* 908 (July 2016), pp. 30–51. arXiv: 1602.02469 [hep-ex, physics:nucl-ex, physics:physics].

[18] The Super-Kamiokande Collaboration and Y. Ashie. "Evidence for an Oscillatory Signature in Atmospheric Neutrino Oscillation". In: *Physical Review Letters* 93.10 (Sept. 2004), p. 101801. arXiv: hep-ex/0404034.

[19] B. Pontecorvo. "Mesonium and Anti-Mesonium". In: *Sov. Phys. JETP* 6 (1957), p. 429.

[20] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. "Remarks on the Unified Model of Elementary Particles". In: *Prog. Theor. Phys.* 28 (1962), pp. 870–880.

[21] B. Pontecorvo. "Neutrino Experiments and the Problem of Conservation of Leptonic Charge". In: *Zh. Eksp. Teor. Fiz.* 53 (1967), pp. 1717–1725.

[22] Mattias Blennow and Alexei Yu. Smirnov. “Neutrino Propagation in Matter”. In: *Advances in High Energy Physics* 2013 (2013), pp. 1–33.

[23] Ivan Esteban, M.C. Gonzalez-Garcia, Michele Maltoni, Thomas Schwetz, and Albert Zhou. “The Fate of Hints: Updated Global Analysis of Three-Flavor Neutrino Oscillations”. In: *Journal of High Energy Physics* 2020.9 (Sept. 2020), p. 178.

[24] *NuFIT | NuFIT 5.2*. http://www.nu-fit.org/. 2022.

[25] M Jiang, K Abe, C Bronner, *et al.* “Atmospheric Neutrino Oscillation Analysis with Improved Event Reconstruction in Super-Kamiokande IV”. In: *Progress of Theoretical and Experimental Physics* 2019.5 (May 2019), 053F01.

[26] L. Wolfenstein. “Neutrino Oscillations in Matter”. In: *Physical Review D* 17.9 (May 1978), pp. 2369–2374.

[27] Vernon D. Barger, K. Whisnant, S. Pakvasa, and R. J. N. Phillips. “Matter Effects on Three-Neutrino Oscillations”. In: *Phys. Rev. D* 22 (1980), p. 2718.

[28] S. P. Mikheyev and A. Yu. Smirnov. “Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos”. In: *Sov. J. Nucl. Phys.* 42 (1985), pp. 913–917.

[29] E. K. Akhmedov. “On Neutrino Oscillations in a Nonhomogeneous Medium”. In: *Sov. J. Nucl. Phys. (Engl. Transl.); (United States)* 47:2 (Feb. 1988).

[30] Q. Y. Liu, S. P. Mikheyev, and A. Yu Smirnov. “Parametric Resonance in Oscillations of Atmospheric Neutrinos?” In: *Physics Letters B* 440.3-4 (Nov. 1998), pp. 319–326. arXiv: `hep-ph/9803415`.

[31] E. Kh Akhmedov. “Parametric Resonance of Neutrino Oscillations and Passage of Solar and Atmospheric Neutrinos through the Earth”. In: *Nuclear Physics B* 538.1-2 (Jan. 1999), pp. 25–51. arXiv: `hep-ph/9805272`.

[32] M. Chizhov, M. Maris, and S. T. Petcov. *On the Oscillation Length Resonance in the Transitions of Solar and Atmospheric Neutrinos Crossing the Earth Core*. Oct. 1998. arXiv: `hep-ph/9810501`.

[33] M. V. Chizhov and S. T. Petcov. “Enhancing Mechanisms of Neutrino Transitions in a Medium of Nonperiodic Constant - Density Layers and in the Earth”. In: *Physical Review D* 63.7 (Mar. 2001), p. 073003. arXiv: `hep-ph/9903424`.

[34] Margaret A. Millhouse and David C. Latimer. “Neutrino Tomography”. In: *American Journal of Physics* 81.9 (Sept. 2013), pp. 646–654.

[35] ICECUBE COLLABORATION. “Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector”. In: *Science* 342.6161 (Nov. 2013), p. 1242856.

[36] IceCube Collaboration, M. G. Aartsen, M. Ackermann, *et al.* “The IceCube Neutrino Observatory: Instrumentation and Online Systems”. In: *Journal of Instrumentation* 12.03 (Mar. 2017), P03012–P03012.

[37] T. K. Gaisser and M. Honda. “Flux of Atmospheric Neutrinos”. In: *Annual Review of Nuclear and Particle Science* 52.1 (2002), pp. 153–199.

[38] Martin Leuermann. “Testing the Neutrino Mass Ordering with IceCube DeepCore”. PhD thesis. RWTH Aachen University, Nov. 2018.

[39] Donald E. Groom, Nikolai V. Mokhov, and Sergei I. Striganov. “Muon Stopping Power and Range Tables 10 MeV–100 TeV”. In: *Atomic Data and Nuclear Data Tables* 78.2 (July 2001), pp. 183–356.

[40] The IceCube-PINGU Collaboration. “Letter of Intent: The Precision IceCube Next Generation Upgrade (PINGU)”. In: *arXiv:1401.2046* (Sept. 2017). arXiv: 1401.2046.

[41] IceCube Collaboration, M. G. Aartsen, M. Ackermann, *et al. Computational Techniques for the Analysis of Small Signals in High-Statistics Neutrino Oscillation Experiments*. Dec. 2019. arXiv: 1803.05390 [astro-ph, physics:hep-ex, physics:physics].

[42] M. Honda, M. Sajjad Athar, T. Kajita, K. Kasahara, and S. Midorikawa. “Atmospheric Neutrino Flux Calculation Using the NRLMSISE00 Atmospheric Model”. In: *Physical Review D* 92.2 (July 2015), p. 023004. arXiv: 1502.03916 [astro-ph].

[43] Anatoli Fedynitch, Ralph Engel, Thomas K. Gaisser, Felix Riehn, and Todor Stanev. *Calculation of Conventional and Prompt Lepton Fluxes at Very High Energy*. Mar. 2015. arXiv: 1503.00544 [astro-ph, physics:hep-ph].

[44] G. D. Barr, S. Robbins, T. K. Gaisser, and T. Stanev. “Uncertainties in Atmospheric Neutrino Fluxes”. In: *Physical Review D* 74.9 (Nov. 2006), p. 094009.

[45] Roger Wendell and Luke Pickering. *Prob3++*. http://webhome.phy.duke.edu/∼raw22/public/P 2018.

[46] Adam M. Dziewonski and Don L. Anderson. “Preliminary Reference Earth Model”. In: *Physics of the Earth and Planetary Interiors* 25.4 (June 1981), pp. 297–356.

[47] C. Andreopoulos, A. Bell, D. Bhattacharya, *et al.* “The GENIE Neutrino Monte Carlo Generator”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 614.1 (Feb. 2010), pp. 87–104. arXiv: 0905.2517 [hep-ph].

[48] Rezvan Abbasi, M. Ackermann, Jim Adams, *et al.* “Graph Neural Networks for Low-Energy Event Classification & Reconstruction in IceCube”. In: *Journal of Instrumentation* 17 (Nov. 2022), P11003.

[49] Andreas Søgaard, Rasmus F. Ørsøe, Leon Bozianu, Morten Holm, Kaare Endrup Iversen, Tim Guggenmos, Martin Ha Minh, and Philipp Eller. *GraphNeT*. Zenodo. June 2022.

[50] F. James and M. Roos. “Minuit - a System for Function Minimization and Analysis of the Parameter Errors and Correlations”. In: *Computer Physics Communications* 10.6 (Dec. 1975), pp. 343–367.

[51] Hans Dembinski, Piti Ongmongkolkul, Christoph Deil, *et al. Scikit-Hep/Iminuit*. Zenodo. Apr. 2023.

[52] Ladislav Lukšan and Emilio Spedicato. “Variable Metric Methods for Unconstrained Optimization and Nonlinear Least Squares”. In: *Journal of Computational and Applied Mathematics* 124.1-2 (Dec. 2000), pp. 61–95.

[53] X. Qian, A. Tan, W. Wang, J. J. Ling, R. D. McKeown, and C. Zhang. “Statistical Evaluation of Experimental Determinations of Neutrino Mass Hierarchy”. In: *Physical Review D* 86.11 (Dec. 2012), p. 113011. arXiv: 1210.3651 [hep-ex, physics:hep-ph, physics:nucl-ex].

[54] Mattias Blennow. “On the Bayesian Approach to Neutrino Mass Ordering”. In: *Journal of High Energy Physics* 2014.1 (Jan. 2014), p. 139. arXiv: 1311.3183 [hep-ex, physics:hep-ph].

[55] Emilio Ciuffoli, Jarah Evslin, and Xinmin Zhang. “Sensitivity to the Neutrino Mass Hierarchy”. In: *Journal of High Energy Physics* 2014.1 (Jan. 2014), p. 95. arXiv: 1305.5150 [hep-ex, physics:hep-ph, stat].

[56] Emilio Ciuffoli. *Statistical Methods for the Neutrino Mass Hierarchy*. Apr. 2017. arXiv: 1704.08043 [hep-ph].

[57] Steven Wren. “Neutrino Mass Ordering Studies With IceCube-Deepcore”. PhD thesis. University of Manchester, 2018.

[58] Gary J. Feldman and Robert D. Cousins. “A Unified Approach to the Classical Statistical Analysis of Small Signals”. In: *Physical Review D* 57.7 (Apr. 1998), pp. 3873–3889. arXiv: physics/9711021.

[59] M. G. Aartsen, M. Ackermann, J. Adams, *et al.* “Development of an Analysis to Probe the Neutrino Mass Ordering with Atmospheric Neutrinos Using Three Years of IceCube DeepCore Data”. In: *The European Physical Journal C* 80.1 (Jan. 2020), p. 9. arXiv: 1902.07771.

[60] Isaac Asimov. “Franchise”. In: *If: Worlds of Science Fiction*. Vol. August 1955. Quinn Publishing, 1955.

[61] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. “Asymptotic Formulae for Likelihood-Based Tests of New Physics”. In: *The European Physical Journal C* 71.2 (Feb. 2011), p. 1554. arXiv: 1007.1727 [hep-ex, physics:physics].

[62] R. D. Cousins. “Comments on Methods for Setting Confidence Limits”. In: *Workshop on Confidence Limits*. Aug. 2000, pp. 49–61.

[63] X. Qian, A. Tan, J. J. Ling, Y. Nakajima, and C. Zhang. “The Gaussian CL$_s$ Method for Searches of New Physics”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 827 (Aug. 2016), pp. 63–78. arXiv: 1407.5052 [hep-ex, physics:physics].

[64] Mathieu Ribordy and Alexei Yu Smirnov. “Improving the Neutrino Mass Hierarchy Identification with Inelasticity Measurement in PINGU and ORCA”. In: *Physical Review D* 87.11 (June 2013), p. 113007. arXiv: 1303.0758 [hep-ex, physics:hep-ph].

[65] IceCube-Gen2 Collaboration, M. G. Aartsen, M. Ackermann, *et al.* “Combined Sensitivity to the Neutrino Mass Ordering with JUNO, the IceCube Upgrade, and PINGU”. In: *Physical Review D* 101.3 (Feb. 2020), p. 032006. arXiv: 1911.06745.

[66] David V. Forero, Stephen J. Parke, Christoph A. Ternes, and Renata Zukanovich Funchal. “JUNO’s Prospects for Determining the Neutrino Mass Ordering”. In: *Physical Review D* 104.11 (Dec. 2021), p. 113004.