Testing the Neutrino Mass Ordering with IceCube DeepCore

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Martin Leuermann, M.Sc.

aus

Moers

Berichter Univ.-Prof. Dr. Christopher Wiebusch
Univ.-Prof. Dr. Michael Krämer

Tag der mündlichen Prüfung: 11.12.2018

Abstract

The *Neutrino Mass Ordering* (NMO) describes the energetic ordering of the three neutrino masses m_1, m_2 and m_3. Today, the ordering is constrained up to the question of m_3 being the heaviest or the lightest of all neutrino masses, which is commonly called *Normal* (NO) and *Inverted Ordering* (IO), respectively. One way to determine the NMO is to measure matter effects in the oscillation pattern of atmospheric neutrinos. While oscillations arise mainly at neutrino energies of $E_\nu \lesssim 30\,\mathrm{GeV}$, matter effects play an important role for $E_\nu \lesssim 15\,\mathrm{GeV}$. Such analysis of the NMO is proposed by several future neutrino experiments, that aim for a measurement of the NMO with $> 3\,\sigma$ within the next decade. One of these experiments is the *Precision IceCube Next Generation Upgrade* (PINGU), which is a low-energy extension of the existing *IceCube DeepCore* detector. However, already DeepCore is capable of probing the NMO with its energy threshold of $E_\nu \sim 5\,\mathrm{GeV}$.

In this work, an analysis of three years of DeepCore data is presented to test the Neutrino Mass Ordering. The data sample consists of 43 214 events, which comprises *Charged Current* and *Neutral Current* interactions of all flavors and backgrounds from atmospheric muons and triggered noise. It is tuned to provide high statistics of all flavors around $\sim 10\,\mathrm{GeV}$, while reducing the contamination from background. For this sample, a new event reconstruction was developed in this work, which provides excellent resolutions in neutrino energy and zenith-angle at the lowest energies. Moreover, it allows to distinguish different event signatures, which gives a handle for separating neutrino flavors.

The Neutrino Mass Ordering is fit in a *maximum-likelihood* method, which includes systematic uncertainties in the atmospheric fluxes, the oscillation parameters, the detector response and the neutrino-nucleon interactions as nuisance parameters in the likelihood fit. Moreover, the uncertainties from limited simulation statistics are included into the likelihood function. The resulting sensitivity to the NMO is found to be on the $\sim 0.4 - 0.7\,\sigma$-level, while the precise value depends strongly on the true ordering and the value of the atmospheric mixing angle θ_{23}. Moreover, the matter effects, which allow to probe the NMO, can be distinguished from vacuum oscillations on a $\sim 0.6 - 1.1\,\sigma$-level. All of the analysis chain was developed *blindly*, i.e. without applying the analysis to experimental data before finalizing it, and supervised by the *IceCube Collaboration*.

In experimental data, a preference for Normal over Inverted Ordering is found with p-values of $p_{\mathrm{NO}} = 71.1\%$ ($\mathrm{CL}_s = 83.0\%$) and $p_{\mathrm{IO}} = 15.2\%$ ($\mathrm{CL}_s = 53.3\%$) for the Normal and Inverted Ordering hypotheses. This is inline with recent observations of the *NOνA*, *T2K* and *Super-Kamiokande* experiments. Depending on the ordering, the fit prefers matter effects (MA) over vacuum oscillations (VA) with p-values of $p_{\mathrm{MA}} = 62.3\%$ and $p_{\mathrm{VA}} = 12.3\%$ in case of Normal Ordering and $p_{\mathrm{MA}} = 53.2\%$ and $p_{\mathrm{VA}} = 22.2\%$ in case of Inverted Ordering.

The impact of various systematic uncertainties on the result is studied in terms of a potential bias on the result and in terms of potential improvements of the sensitivity. For potential improvements, further constraints on systematic uncertainties are found to be subordinate at the current sensitivity level, while the sensitivity could be increased strongly by an improved event reconstruction.

Besides the experimental result, this work provides a *proof-of-concept* for an analysis of the Neutrino Mass Ordering with a future low-energy extension of IceCube, such as PINGU. It tests the full analysis-chain, including the statistical interpretation of the experimental result and the understanding of systematic uncertainties. Thus, it provides a benchmark analysis for these more sensitive future measurements.

Zusammenfassung

Die *Neutrino Massenordnung* (NMO) beschreibt die energetische Anordnung der drei Neutrinomassen m_1, m_2 und m_3. Nach heutigem Stand verbleiben zwei mögliche Anordnungen, wobei m_3 entweder die schwerste oder die leichteste der drei Neutrinomassen ist, was als *Normale* (NO) und *Invertierte Ordnung* bezeichnet wird. Eine Möglichkeit zur Bestimmung der NMO ist die Messung von Materieeffekten in den Oszillationen atmosphärischer Neutrinos. Während atmosphärische Oszillationen hauptsächlich bei Neutrinoenergien von $E_\nu \lesssim 30\,\text{GeV}$ auftreten, werden Materieeffekte bei $E_\nu \lesssim 15\,\text{GeV}$ erwartet. Eine solche Messung der NMO wird von mehreren, zukünftigen Neutrinoexperimenten mit $> 3\,\sigma$ Sensitivität beabsichtigt. Eines dieser Experiment ist das *Precision IceCube Next Generation Upgrade* (PINGU), eine Niederenergie-Erweiterung des existierenden *IceCube DeepCore* Detektors, der bereits in der Lage ist, erste Anzeichen für die NMO zu beobachten.

In dieser Arbeit wird eine Analyse zur Bestimmung der Neutrino Massenordnung mit drei Jahren Daten von IceCube DeepCore entwickelt. Die Daten umfassen 43 214 Ereignisse von *Charged Current* und *Neutral Current* Wechselwirkungen aller Neutrino-Flavor sowie einen Hintergrund von atmosphärischen Myonen und Detektorrauschen. Die Datenselektion wurde für hohe Statistik im Bereich um $E_\nu \sim 10\,\text{GeV}$ und auf die Unterdrückung des Hintergrunds optimiert. Für die Bestimmung der Energie und der Ankunftsrichtung der Neutrino-Ereignisse wurde eine neue Ereignis-Rekonstruktion entwickelt, welche für die niedrigsten Energien eine exzellente Auflösung erreicht. Darüber hinaus erlaubt sie die Separation von Ereignissignaturen zur indirekten Unterscheidung der Neutrino-Flavor.

Die Neutrino Massenordnung wird in einem *maximum-likelihood* Verfahren bestimmt, das systematische Unsicherheiten im atmosphärischen Fluss, in den Oszillationsparametern, in den Wechselwirkungen und in der Detektorantwort als Störparameter behandelt. Unsicherheiten durch die begrenzte Simulationsstatistik werden in der Likelihood-Funktion berücksichtigt. Die Sensitivität auf die NMO liegt bei $\sim 0.4 - 0.7\,\sigma$ mit einer hohen Abhängigkeit vom atmosphärischen Mischungswinkel θ_{23}. Darüber hinaus können Materieeffekte von Vakuum-Oszillationen mit $\sim 0.6 - 1.1\,\sigma$ unterschieden werden. Die gesamte Analyse wurde *blind* entwickelt, d.h. ohne sie vor ihrer Finalisierung auf experimentelle Daten anzuwenden, und von der *IceCube Collaboration* begutachtet.

Die experimentellen Daten präferieren die Normale über die Invertierte Ordnung mit p-Werten von $p_{\text{NO}} = 71.1\%$ ($\text{CL}_s = 83.0\%$) und $p_{\text{IO}} = 15.2\%$ ($\text{CL}_s = 53.3\%$). Dies deckt sich mit jüngsten Resultaten der *NOνA*, *T2K* und *Super-Kamiokande* Experimente. Darüber hinaus werden Materieeffekte (MA) gegenüber Vakuum-Oszillationen (VA) mit p-Werten von $p_{\text{MA}} = 62.3\%$ und $p_{\text{VA}} = 12.3\%$ für Normale und $p_{\text{MA}} = 53.2\%$ und $p_{\text{VA}} = 22.2\%$ für Invertierte Ordnung bevorzugt.

Der Einfluss systematischer Unsicherheiten auf das Ergebnis (als systematischer Fehler) und die Sensitivität werden untersucht, wobei sich bei dem derzeitigen Sensitivitätsniveau die Reduktion von systematischen Unsicherheiten als nachrangig gegenüber einer Verbesserung der Rekonstruktions-Auflösung für eine Steigerung der Sensitivität herausstellt.

Neben dem experimentellen Resultat stellt diese Arbeit eine wichtige Machbarkeitsstudie (*proof-of-concept*) für die Messung der Neutrino Massenordnung mit einer zukünftigen Niederenergie-Erweiterung von IceCube, wie z.B. PINGU, dar. Sie testet die vollständige Analyse-Kette, einschließlich der statistischen Interpretation der Resultate und der Behandlung systematischer Unsicherheiten, und ist damit eine wichtige *Benchmark*-Analyse für eine zukünftige, sensitivere Messung der NMO.

Alles hat seine Zeit
und jegliches Vornehmen
unter dem Himmel seine Stunde.

Bibel, Prediger 3, 1

Contents

1 Motivation for Measurements of the Neutrino Mass Ordering

Mass is one of the fundamental properties of elementary particles. At the same time it is encountered in all of our everyday life so frequently that a physicist might have difficulties in explaining to people, that there is a particle, which is studied in detail, but still its absolute mass is unknown. It sounds even more confusing that while the mass differences between various types of this particle are known rather precisely, their abolute masses can neither be measured nor excluded to be zero for each of them - even though they are large enough to affect the evolution of the universe. Such a particle is the neutrino.

During the last decade, only few observations were found that indicate that the Standard-model of Particle Physics is incomplete and requires a further generalization. The discovery of non-zero neutrino masses is one of them and was awarded with the Nobel Prize in physics in 2015 for Arthur McDonald and Takaaki Kajita, two of the main contributors to this discovery [1, 2].

The non-zero neutrino masses lead to the phenomenon of neutrino oscillation, which is the change of neutrino flavor over time and propagated distance. Today, the parameters describing the oscillations between the three neutrino states are known rather precisely due to large experimental efforts within the last decades. However, there are still missing pieces to the puzzle: The Neutrino Mass Ordering, the size of the CP-violation in the neutrino sector, the absolute value of the three neutrino masses m_1, m_2 and m_3 and the question of neutrinos being Majorana or Dirac particles.

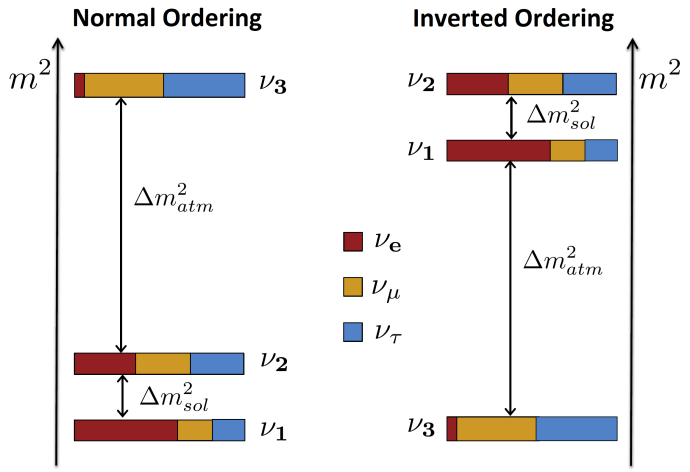


Figure 1.1: Ordering of the three neutrino masses in case of Normal (left) and Inverted (right) Ordering; the colors illustrate the mixing with the flavor states that are described in Section 2.1 (from [3]).

This work addresses the question of the *Neutrino Mass Ordering* (NMO), which describes

the energetic ordering of the three neutrino mass eigenstates. Today, there are only two options remaining, known as *Normal Ordering* (NO) and *Inverted Ordering* (IO), which refer to the large (atmospheric) squared-mass difference $\Delta m^2_{\text{atm}} = m^2_3 - m^2_{1,2}$ being positive or negative, at the top or at the bottom, while the small (solar) squared-mass difference $\Delta m^2_{\text{sol}} = m^2_2 - m^2_1$ is at the bottom or top, respectively. The corresponding ordering of the three mass eigenstates is illustrated in Figure 1.1 [4].

Note that the term *Neutrino Mass Hierarchy* (NMH) is often used synonymously for Neutrino Mass Ordering [4]. However, it is not used in this work to avoid confusion with statements on the absolute scale of the neutrino masses, which is not discussed here.

For many fields of physics, the Neutrino Mass Ordering is important to know: For neutrino oscillation measurements, it changes the oscillation pattern, while for the double-beta decay, it changes the strength of a potential experimental signature [5]. Moreover, since it also influences the minimum absolute mass of the neutrinos, it relates the neutrino number density and the neutrino mass density in our universe, which play an important role in cosmology [6]. As a result, many future experiments, like the proposed *Precision IceCube Next Generation Upgrade* (PINGU), are aiming for a $> 3\,\sigma$ measurement within the next decade [7]. However, already the predecessor of PINGU, the currently data-taking IceCube DeepCore detector, has some sensitivity to the NMO.

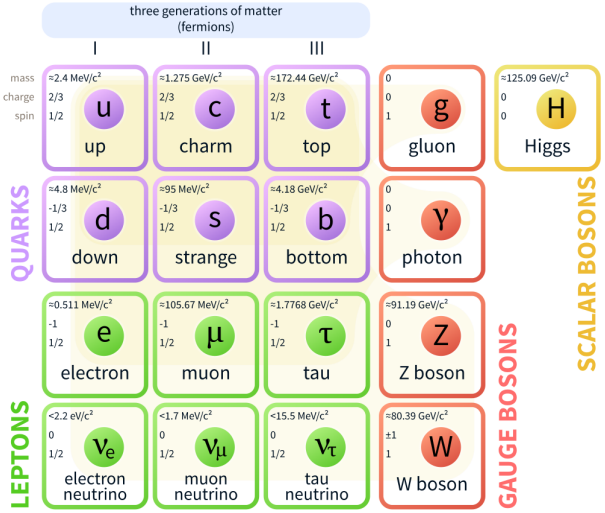
In this work, a measurement of the Neutrino Mass Ordering with three years of data from IceCube DeepCore is presented. It includes the development of a new, low-energy event reconstruction and uses a binned likelihood method to fit the NMO on an extreme low-energy data sample. It is similar to the benchmark analysis proposed for PINGU [7]. Thus, besides providing a first, low-significance measurement of the NMO with IceCube DeepCore, this work also explores this type of measurement for the future PINGU detector by testing the full analysis-chain and the understanding of systematic uncertainties, which validates the analysis method.

2 Introduction

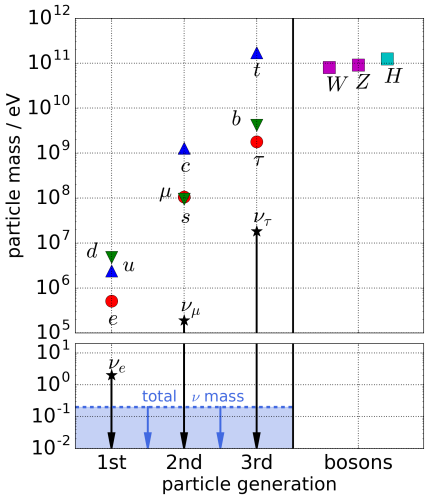
2.1 Neutrinos in the Standard Model of Particle Physics

The neutrino is one of the fundamental particles in the Standard Model of particle physics. It was postulated in 1930 by Wolfgang Pauli to explain the continuous energy spectrum of beta-decay electrons. Afterwards, it was experimentally discovered in 1956 [8] and is a fundamental particle of the Standard Model ever since.

In the Standard Model, the behavior of all particles is derived from the Lagrange formalism, in which the *Action* $\mathcal{S}$, given by the spacetime integrated *Lagrangian density* $\mathcal{L}$, is minimized with respect to the fields it contains. In this Lagrangian density, the neutrinos are represented by the three fundamental fields ν_1, ν_2 and ν_3 (labelled by latin index), which are often referred to as neutrino mass eigenstates. In the following, only the behavior of the neutrino fields is discussed, while the behavior of the anti-neutrino fields $\bar{\nu}_1$, $\bar{\nu}_2$ and $\bar{\nu}_3$ is analogous. The physics of these neutrino fields is described by terms in the Lagrangian density $\mathcal{L}$ for their interaction and free propagation [9]. The only interaction terms for the three neutrino



(a) Overview of particle properties (from [10])



(b) Overview of particle masses

Figure 2.1: Overview of all Standard Model particles and their properties; for (b), the particle masses are taken from [11] and [12]; note that the ν_e limit is based on $\bar{\nu}_e$ measurements, while direct ν_e measurements are less constraining by ~ 2 orders of magnitude [11].

fields are given by weak interactions, transmitted by charged $W^{\pm}$- or neutral Z^0-Bosons [13]. Thus, the total cross-section for neutrino interactions is very small, due to the small *Fermi*

coupling constant G_F, describing the strength of the weak force, which makes measurements of the neutrino properties challenging [11]. In these interaction terms, the fields ν_1, ν_2 and ν_3 show up as rotated states, i.e. linearly independent superpositions, which are described by the unitary, 3×3 *Pontecorvo-Maki-Nakagawa-Sakata* (mixing) matrix U, such that

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U \cdot \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}. \tag{2.1}$$

These superpositions are often considered as the natural choice for describing the three types of neutrinos, since they represent the states in which all experimental observations take place. These superpositions are called neutrino flavors. They are commonly labelled as electron, muon and tau neutrino, or ν_e, ν_μ, ν_τ (labelled by greek index), respectively, where the label corresponds to the charged lepton, each neutrino flavor couples to via Charged Current (CC) interactions [9].

In Figure 2.1(a) an overview of all Standard Model particles is shown, where the neutrinos are conventionally shown in their flavor states, next to the corresponding charged lepton.

Besides the interactions, the Lagrangian density $\mathcal{L}$ contains terms for the free propagation of the mass eigenstates. The structure of these terms depends on whether neutrinos are Dirac or Majorana particles, which is a question of ongoing research and directly related to the origin of their masses (cf. Section 2.3.1) [14].

In the Standard Model, neutrinos were first believed to be massless as this would avoid the question of right-handed neutrinos, described in Section 2.2. However, the existence of non-zero neutrino masses is known from the measurement of neutrino oscillations that are described in more detail in Section 2.3.2. The measured masses of all fundamental particles of the Standard Model are illustrated in Figure 2.1(b). Note that the black stars mark only the direct constraints on the neutrino masses m_{ν_e}, m_{ν_μ} and m_{ν_τ} of the flavor states, while more constraining measurements exist from β-decay, $\beta\beta$-decay and cosmological data for other superpositions of the mass eigenstates m_1, m_2 and m_3. For example, a limit on the total neutrino mass $m_1 + m_2 + m_3$ [12] is shown as blue box in Figure 2.1(b). The limits on the neurino masses are further discussed in Section 2.3 [12, 9].

Thus, the neutrino masses are substantially smaller than any other fundamental particle mass in the Standard Model. This might be an indication for a different mechanism giving rise to the neutrino masses than for the other particles, which is discussed in Section 2.3.1.

2.2 Neutrino Interactions

Neutrinos carry neither electrical charge nor color. Thus, they do not participate in strong or electromagnetic interactions. In the Standard Model Lagrangian, they only interact with other Standard Model particles via a weak interaction term [9]. This term is given by

$$\begin{aligned} \mathcal{L}_{\text{weak}} = &- \frac{g_W}{\sqrt{2}} \bar{\ell}_\alpha \gamma^\mu \frac{1}{2}(1-\gamma^5)\nu_\alpha W_\mu^- \\ &- \frac{g_W}{\sqrt{2}} \bar{\nu}_\alpha \gamma^\mu \frac{1}{2}(1-\gamma^5)\ell_\alpha W_\mu^+ \\ &- \frac{g_W}{2\cos(\theta_W)} \bar{\nu}_\alpha \gamma^\mu \frac{1}{2}(1-\gamma^5)\nu_\alpha Z_\mu^0, \end{aligned} \tag{2.2}$$

where $W^{\pm}$ and Z^0 represent the gauge bosons of the Charged Current (CC) and Neutral Current (NC) interactions, respectively, while ℓ_α with $\alpha \in \{e,\mu,\tau\}$ is a charged lepton. The coupling constants of CC and NC interactions are given by g_W and $g_Z = g_W/\cos(\theta_W)$, respectively, where θ_W is the weak mixing angle [9].

One should note that the fundamental fields ν_i do not show up individually in the Lagrangian density $\mathcal{L}_{\text{weak}}$ of the interaction. Instead, they show up as linear combinations $\nu_\alpha = \sum_{k=1}^{3} U_{\alpha k}\nu_k$ describing the flavor states. Therefore, the fundamental fields cannot be observed directly in particle experiments, but are accessible only via their contributions to the flavor states.

Due to the structure of the interaction terms, the Lagrangian density can be rewritten purely in terms of left-handed neutrinos $\nu_{\alpha L}$ and right-handed anti-neutrinos $\bar{\nu}_{\alpha R}$, using $\bar{\ell}_\alpha\gamma^\mu\frac{1}{2}(1-\gamma^5)\nu_\alpha = \bar{\ell}_{\alpha L}\gamma^\mu\nu_{\alpha L}$. Thus, only left handed neutrinos participate in weak interactions, causing the weak parity violation. Right-handed neutrinos, if they exist, do not participate in these interactions and are therefore not observable [9].

The fact, that only left-handed neutrinos are experimentally accessible, affects the question of neutrinos being Dirac or Majorana particles, as discussed in Section 2.3.1.

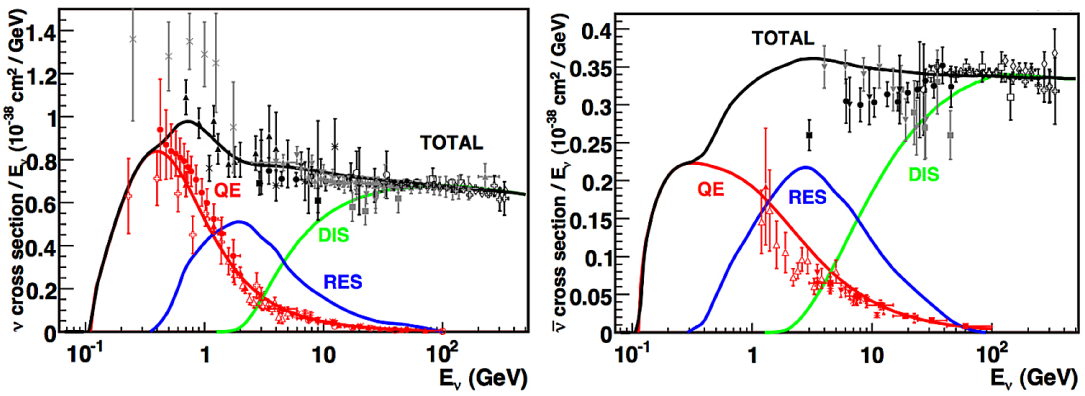


Figure 2.2: Energy-dependent cross-section for CC neutrino-nucleon interactions split into resonant (RES), quasi-elastic (QE) and deep inelastic scattering (DIS) for neutrinos (left) and anti-neutrinos (right) (from [15]).

In principle, neutrinos couple to all Standard Model fermions via Charged and Neutral Current interactions. However, the most relevant processes for the detection of neutrinos above $\sim 1\,\text{GeV}$ are neutrino-nucleon interactions. Therefore, these are discussed in more detail in the following [15].

Neutrino-nucleon interactions can be split into three different processes: At high energies of $E_\nu \gtrsim 100\,\text{GeV}$, the cross-section is almost exclusively given by *deep inelastic scattering* (DIS, cf. Section 2.2.1). At lower energies of $E_\nu \lesssim 10\,\text{GeV}$, *resonant* (RES), *elastic and quasi-elastic* (QE) interactions start dominating (cf. Section 2.2.2). The energy-dependent cross-section of CC interactions is shown in Figure 2.2, where different processes dominate at different energies. The overall cross-section is notably smaller for neutrinos than for antineutrinos, while the ratio between neutrino and anti-neutrino remains roughly constant at ~ 2. This is shown for the total cross-section in Figure 2.3 and becomes relevant for the atmospheric signature of the NMO, described in Section 2.4.2 [15].

Besides these three contributions, the neutrino interaction can generate a charged meson without exciting the nucleon state. This process is typically referred to as *coherent meson production*. Due to its small cross-section, it is not relevant for this work and therefore not discussed here [16].

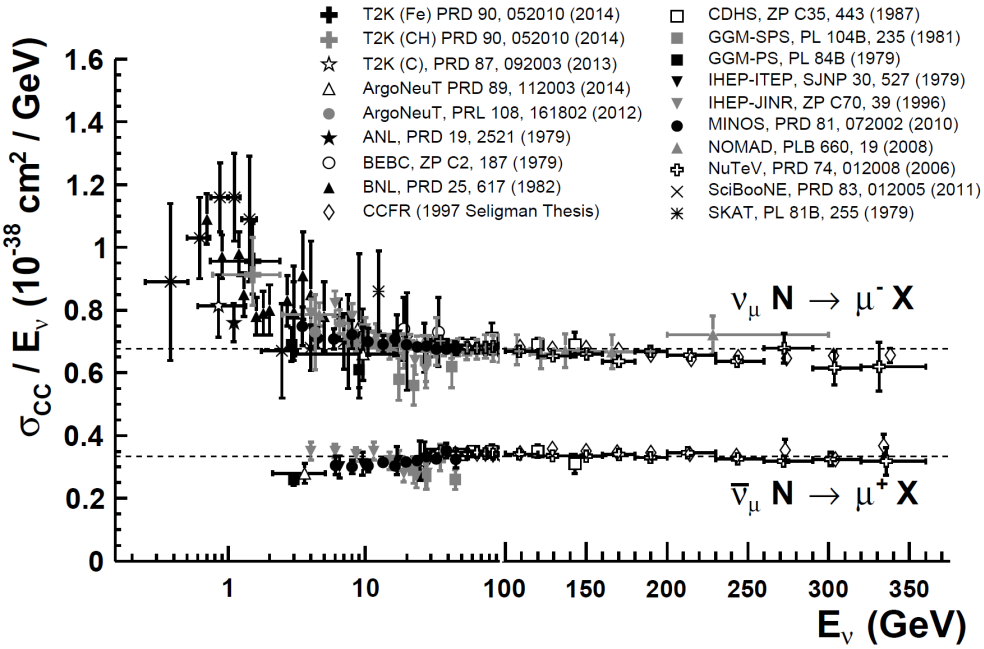


Figure 2.3: Total muon neutrino and anti-neutrino CC cross-section for neutrino-nucleon scattering; note the switch between log- and linear scale at $E_\nu = 100\,\text{GeV}$; overall, the neutrino cross-section is roughly twice as high as the one for anti-neutrinos (from [17], based on [11]).

In the following sections, DIS, QE and RES interactions are briefly discussed for the most interesting energy regime of this work, which is $E_\nu \sim 1\,\text{GeV} - 1\,\text{TeV}$. The impact of their uncertainty on the analysis is discussed in Section 6.3.4.

2.2.1 Deep Inelastic Scattering

Deep inelastic scattering is the dominant process for CC interactions above $E_\nu \sim 10\,\text{GeV}$. For DIS interactions, the CC transfers enough energy to an individual quark to break the bound state of the nucleon. This is the case, if the momentum/energy transfer is large compared to the nucleon mass of $m_N \sim 1\,\text{GeV}$. The resulting remnants of the nucleon N generate a *hadronic cascade* of secondary particles. The cascade is called *hadronic*, since the production of secondary particles is dominated by strong interactions, in contrast to *electromagnetic cascades*, discussed later [15, 18].

The primary interaction of DIS is given by Equation 2.3

$$\overset{(-)}{\nu}_\alpha + N \rightarrow \ell_\alpha^\pm + X, \tag{2.3}$$

where $\ell_\alpha^\pm$ represents a charged lepton and X represents the hadronic cascade generated by the nucleon remnants.

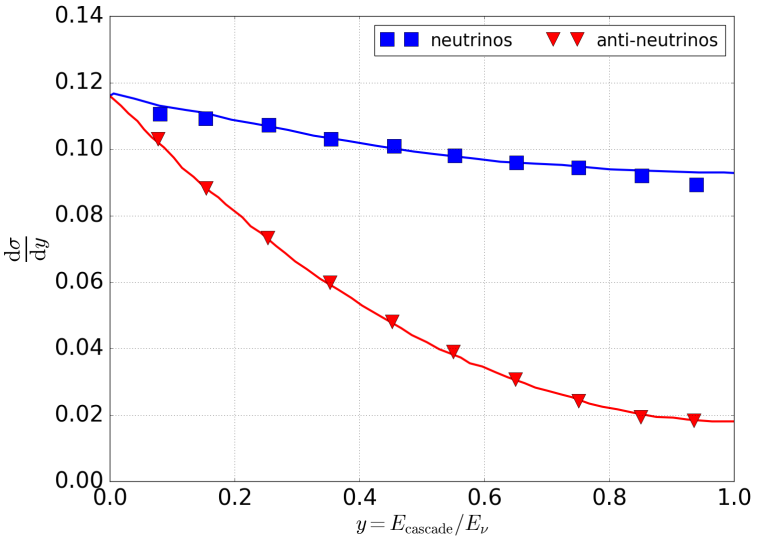


Figure 2.4: Inelasticity (Bjorken-y) distribution of CC neutrino-nucleon scattering in case of DIS (modified from [19]).

This equation is symmetric for neutrinos and anti-neutrinos. However, the density of quarks is increased compared to anti-quarks in the nucleons, due to the presence of the valence quarks. This affects the cross-section of the interaction. The cross-section depends on the relative energy transfer y to the nucleon, called *inelasticity* or *Bjorken-y*. It is given by Equation 2.4

$$y = \frac{E_\nu - E_{\ell_\alpha}}{E_{\nu_\alpha}} = \frac{E_{\text{cascade}}}{E_{\nu_\alpha}}. \tag{2.4}$$

The inelasticity describes the fraction of the neutrino energy E_ν, transferred towards the hadronic cascade. In the following, it is used without distinguishing between momentum and energy transfer, since $m_{\ell_\alpha} \ll E_\nu$ is assumed implicity [9, 18].

The y-dependent cross-sections for neutrinos and anti-neutrinos are given by Equation 2.5 and 2.6

$$\frac{\mathrm{d}\sigma_{\nu N}}{\mathrm{d}y} = \frac{G_F^2}{2\pi} s \left[f_q + (1-y)^2 f_{\bar{q}} \right], \tag{2.5}$$

$$\frac{\mathrm{d}\sigma_{\bar{\nu} N}}{\mathrm{d}y} = \frac{G_F^2}{2\pi} s \left[(1-y)^2 f_q + f_{\bar{q}} \right], \tag{2.6}$$

where f_q and $f_{\bar{q}}$ describe the total momentum fraction carried by quarks and anti-quarks, respectively, while $G_F = g_W^2/(4\sqrt{2}m_W^2)$ is the *Fermi coupling constant* and $\sqrt{s}$ the center-of-mass energy. Note that both cross-sections are composed of a term proportional to $(1-y)^2$ and a y-independent term. The first term dominates in both equations, since $f_q > f_{\bar{q}}$ holds due to the presence of the valence quarks [9].

As a result, the neutrino interactions are mostly flat in y with only a small $(1-y)^2$-contribution, while the anti-neutrino interactions are dominated by the $(1-y)^2$ term with only a small y-independent contribution. The cross-sections for neutrinos and anti-neutrinos are shown in Figure 2.4, where the different y-dependencies are clearly visible.

The different y-depencence of the neutrino and anti-neutrino cross-section has some impact on the later-on presented analysis. It leads to small differences in the experimental signature of neutrino and anti-neutrino events, which is discussed in more detail in Section 5.3.4.

2.2.2 Resonant and Quasi-Elastic Scattering

Below $E_\nu \sim 5\,\text{GeV}$, the neutrino-nucleon cross-section is dominated by resonant (RES) and quasi-elastic scattering (QE) (cf. Figure 2.2). The resonant interactions arise from generating an excited state of the nucleon $N \in \{p,n\}$ by the neutrino energy transfer. This is shown in Equation 2.7 and 2.8 for any resonant state N^* of the nucleon [18]

$$\textbf{CC:} \quad \overset{(-)}{\nu}_\alpha + N \to \alpha^\pm + N^* \to \alpha^\pm + \pi^{\mp/0} + N', \tag{2.7}$$

$$\textbf{NC:} \quad \overset{(-)}{\nu}_\alpha + N \to \overset{(-)}{\nu}_\alpha + N^* \to \overset{(-)}{\nu}_\alpha + \pi^{\pm/0} + N'. \tag{2.8}$$

The excited state N^* is not stable and decays into $N' \in \{p,n\}$ and a single pion π of any charge in most cases. Due to the pion, $N' = N$ and $N' \neq N$ are possible, independent of the interaction being CC or NC.

The cross-section is often evaluated for several resonances N^*, although it is dominated by $N^* = \Delta(1232)$ in the most interesting energy regime, where it dominates the total neutrino-nucleon cross-section. The pion π and nucleon N' can trigger a following hadronic cascade at the primary vertex, analogously to the DIS interaction discussed in Section 2.2.1 [18]. Furthermore, in case of charged pions, the pion decays into a muon and a (anti-)neutrino with a branching ratio of 99.98% [11].

The energy range in which the contribution from resonant scattering is notable is very limited (cf. Figure 2.2). It is driven by the excitation energy of the nucleon, which is typically $\mathcal{O}(1\,\text{GeV})$.

In contrast, elastic and quasi-elastic interactions do not degenerate or excite the nucleon and their momentum transfer is typically small. For NC interactions, they are called elastic, as no particle type is changed and thus no energy is needed for such transformations. In contrast, for CC interactions they are called quasi-elastic, since a charged lepton ℓ_α is generated and the nucleon charge is changed according to Equation 2.9 [18]

$$\textbf{CC:} \quad \overset{(-)}{\nu}_\alpha + p/n \to \ell_\alpha^\pm + n/p, \qquad \textbf{NC:} \quad \overset{(-)}{\nu}_\alpha + p/n \to \overset{(-)}{\nu}_\alpha + p/n. \tag{2.9}$$

As shown in Figure 2.2, QE interactions peak at even lower energies than RES ones.

2.3 Neutrino Masses and Neutrino Oscillations

As mentioned in Section 2.1, the neutrino masses are known to be non-zero for at least two of the three neutrino masses m_1, m_2 and m_3 from neutrino-oscillation measurements.

At the same time, the neutrino masses are known to be small from observations of the β-decay spectrum [20] and the non-observation of *neutrinoless double-beta decays* ($0\nu\beta\beta$-decays) [5]. Furthermore, the neutrino masses are constrained from cosmology [21].

In β-decay experiments, the energy spectrum of the escaping charged lepton is observed. A non-zero value of the electron neutrino mass m_{ν_e} limits the maximum energy that can be transfered to the charged lepton. Observing such a cut-off would provide a measurement of m_{ν_e} and thus an indirect measurement of the masses m_1, m_2 and m_3. In Figure 2.5, such cut-off is sketched for a β-decay spectrum. However, current measurements are compatible with $m_{\nu_e} = 0$, constraining the electon neutrino mass to $m_{\nu_e} \lesssim 2\,\text{eV}$ [20].

In case of Majorana neutrinos (cf. Section 2.3.1), $0\nu\beta\beta$-experiments can measure the effective neutrino mass $m_{\beta\beta} = \left|m_1|U_{e1}|^2 + m_2|U_{e2}|^2 e^{i\rho_1} + m_3|U_{e3}|^2 e^{i\rho_2}\right|$ from the observation of two charged leptons without energy being transfered to an escaping neutrino. The effective mass $m_{\beta\beta}$ is given by the the mixing matrix U and two majorana phases ρ_1 and ρ_2, that are described in more detail in Section 2.3.2. Since no $0\nu\beta\beta$-decays have been observed so far, limits are in the order of $m_{\beta\beta} \lesssim 0.5\,\text{eV}$ [5].

Finally, the structure formation of the universe is influenced by the gravitational impact of the *Cosmic Neutrino Background* (CNB). The mass of the CNB is linked to the cosmological neutrino mass $m_{\text{cosmo}} = \sum_{i=1}^{3} m_i$. Thus, the value of m_{cosmo} can be estimated by observing the structure formation in our universe and comparing it to corresponding simulations for different neutrino masses. Although the limits depend strongly on the choice of model and cosmological data, recent cosmological limits are in the order of $m_{\text{cosmo}} \lesssim 0.2\,\text{eV}$ [21].

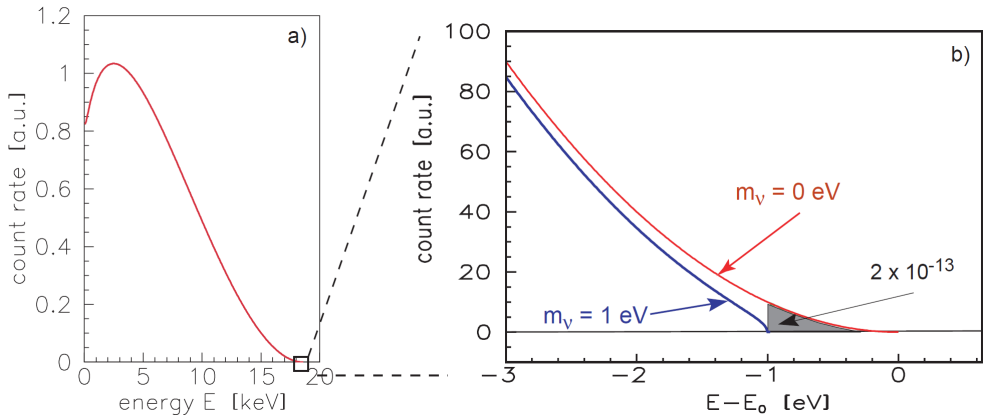


Figure 2.5: Energy spectrum of electrons released in Tritium beta-decays, as expected by the *KATRIN* experiment; a cut-off in the spectrum at the far right, which was not yet observed, could provide evidence for a non-zero electron neutrino mass (from [20]).

Note that these constraints on the absolute neutrino masses are closely connected to constraints on the NMO, which are discussed in more detail in Section 2.5.

From all these measurements, the neutrino mass is known to be small. This is remarkable, because the mass scale for all of the neutrino masses is no larger than $m_i \sim \mathcal{O}(1\,\text{eV})$ with $i = 1,2,3$, which is substantially lower than the mass scale $m_H \sim \mathcal{O}(1\,\text{GeV})$ observed for all other particles of the Standard Model (cf. Figure 2.1(b)), which arises from the coupling to the Higgs boson [15, 9].

Thus, the neutrino masses are of special interest for a potential, future generalization of the Standard Model. In the following, a possible mechanism is described, that explains the low value of the neutrino masses.

2.3.1 Origin of the Neutrino Masses

The most general mass term in the Lagrangian density $\mathcal{L}$, that is commonly assigned to the neutrino fields $\nu_{1,2,3}$, is given by Equation 2.10 [9]

$$\mathcal{L}_{DM} = \underbrace{-m_D\left(\bar{\nu}_R\nu_L + \bar{\nu}_L\nu_R\right)}_{=\mathcal{L}_D} + \underbrace{-\frac{1}{2}M\left(\bar{\nu}_R^c\nu_R + \bar{\nu}_R\nu_R^c\right)}_{=\mathcal{L}_M}, \tag{2.10}$$

with $\bar{\nu}_i = \nu^\dagger \gamma^0$ being the Dirac-adjoint of ν_i and $\nu_i^c = i\gamma^2\gamma^0\nu_i^*$ being the CP-conjugate of the field ν_i [9].

The Lagrangian is composed of a Majorana mass term $\mathcal{L}_M$ and a Dirac mass term $\mathcal{L}_D$. The Dirac mass term is generated analogously to all other massive fermions of the Standard Model. It drops out from spontaneous symmetry breaking after coupling the corresponding neutrino field ν_i to the Higgs field. As a consequence, right-handed neutrinos must exist, which is not supported by any experimental evidence, because they do not participate in any interaction of the Standard Model. However, without them the gauge-invariance of the Standard Model Lagrangian is not maintained [9].

In contrast, the Majorana mass term can be added directly to the Standard Model Lagrangian. Since ν_L and ν_R transform as singlets under all Standard Model gauge transformations, this maintains the gauge invariance of $\mathcal{L}_{DM}$. The corresponding mass term can be seen as a direct coupling of a neutrino field $\nu_{L/R}$ to an anti-neutrino field $\nu_{R/L}^c$. Thus, in the Majorana scenario neutrinos would be their own anti-particle [9].

For neutrinos being purely Majorana or Dirac particles, the corresponding other mass term vanishes. More generally, it is often assumed that both terms exist. In that case, an interesting mechanism takes places that gives an explanation for the neutrino masses being very different from all other masses in the Standard Model. It is often referred to as *seesaw mechanism* and briefly explained in the following.

The combined mass term in the Lagrangian density can be re-written as in Equation 2.11, using $\bar{\nu}_L\nu_R = \bar{\nu}_R^c\nu_L^c$ [9]

$$\mathcal{L}_{DM} = -\frac{1}{2}\begin{pmatrix}\bar{\nu}_L & \bar{\nu}_R^c\end{pmatrix}\underbrace{\begin{pmatrix}0 & m_D\\ m_D & M\end{pmatrix}}_{=S}\begin{pmatrix}\nu_L^c\\ \nu_R\end{pmatrix} + h.c., \tag{2.11}$$

where $h.c.$ stands for the hermitian conjugate of the previous expression. The physical states of the system can be derived by diagonalizing the mass matrix S. The resulting eigenvalues on the diagonal give the masses of the physical neutrino states. Solving $\det(S - \lambda\mathbb{1}) = 0$, one obtains the mass eigenstates in Equation 2.12

$$m_\pm = \lambda_\pm = \frac{1}{2}\left(M \pm \sqrt{M^2 + 4m_D^2}\right) \overset{m_D \ll M}{\approx} \frac{1}{2}M \pm \frac{1}{2}\left(M + \frac{2m_D^2}{M}\right), \tag{2.12}$$

where the last equality holds approximately for the Majorana mass being much larger than the mass obtained from the Higgs mechanism. In the seesaw mechanism, this is assumed to be the case to obtain two very different neutrino masses

$$|m_\nu| \approx \frac{m_D^2}{M} \quad \text{and} \quad m_N \approx M. \tag{2.13}$$

Thus, for a Dirac mass of the same order of magnitude as for all other fermions, i.e. $\mathcal{O}(1\,\text{GeV})$, and a Majorana mass $M \sim 10^{11}\,\text{GeV}$, one obtains a small mass $m_\nu \sim 0.01\,\text{eV}$, which is compatible with the expected mass of the neutrino states. In contrast, the larger mass m_N must be $m_N \sim M \sim 10^{11}\,\text{GeV}$ [9].

The resulting neutrino states are described by the mixing angle $\tan(\alpha) = m_D/M$ and are given by Equation 2.14 and 2.15 [9]

$$\nu = \cos(\alpha)\,(\nu_L + \nu_L^c) - \sin(\alpha)\,(\nu_R + \nu_R^c) \overset{m_D \ll M}{\approx} (\nu_L + \nu_L^c) - \frac{m_D}{M}\,(\nu_R + \nu_R^c), \tag{2.14}$$

$$N = \cos(\alpha)\,(\nu_R + \nu_R^c) - \sin(\alpha)\,(\nu_L + \nu_L^c) \overset{m_D \ll M}{\approx} (\nu_R + \nu_R^c) - \frac{m_D}{M}\,(\nu_L + \nu_L^c). \tag{2.15}$$

Thus, the heavy neutrino state N is almost exclusively right-handed and does not participate in any weak interactions. In contrast, the light neutrino state ν is almost exclusively left-handed, giving the observed couplings for neutrinos in the Standard Model [9].

Although the seesaw mechanism is only a hypothesis, it provides an interesting mechanism to explain the large gap between the neutrino masses and the masses of all other Standard Model particles, shown in Figure 2.1(b). Since neutrinos are the only Standard Model particles capable of having a Majorana term, this mechanism would only take place for neutrino masses and explain why neutrino masses are the only ones deviating from the typical mass scale $\mathcal{O}(1\,\text{GeV})$ observed for masses generated by the Higgs mechanism [9].

A possible way to determine the nature of neutrino masses would be the observation of $0\nu\beta\beta$-decay. Such decay would be possible, if neutrinos were their own anti-particles and thus, could confirm the Majorana nature of neutrinos. However, such a process has not been observed yet by any experiment [14, 5].

2.3.2 Neutrino Oscillations in Vacuum

Flavor transitions between different neutrino flavors during the propagation of the neutrino fields are called neutrino oscillations. They are possible, because the fundamental fields ν_1, ν_2 and ν_3 are not identical to the flavor states ν_e, ν_μ, ν_τ. As a result, a neutrino generated in a flavor state $\alpha \in \{e,\mu,\tau\}$ by a weak interaction is a linear combination of three fundamental fields, each following an independent time expansion. Thus, each component of the flavor state develops differently in time, such that the linear composition of ν_1, ν_2 and ν_3 changes with time and propagated distance. As a result, transitions are possible in the flavor basis [9].

These flavor transitions were first predicted in 1957 by Bruno Pontecorvo [22] and observed by the *Super-Kamiokande* [2] experiment in 1998 and the *Sudbury Neutrino Observatory* [1] in 2001, which proved the existence of neutrino masses.

Neutrino oscillations in vacuum are described by six parameters. For propagation through matter, the treatment is more complicated and discussed in Section 2.3.3. Three of these six parameters are called *mixing angles* θ_{ij} with $i \in \{1,2,3\}$ and $i \neq j$. They describe the elements of the mixing matrix U in Equation 2.1, that links the neutrino mass and flavor states. The matrix U can be written as a product of three unitary mixing matrices $\tilde{U}_{23}$, $\tilde{U}_{13}$ and $\tilde{U}_{12}$, as shown in Equation 2.16 [4]

$$U = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix}}_{\tilde{U}_{23}} \underbrace{\begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\text{CP}}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{\text{CP}}} & 0 & c_{13} \end{pmatrix}}_{\tilde{U}_{13}} \underbrace{\begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\tilde{U}_{12}} \underbrace{\begin{pmatrix} e^{-i\rho_1} & 0 & 0 \\ 0 & e^{-i\rho_2} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{R}, \tag{2.16}$$

with $s_{ij} = \sin(\theta_{ij})$ and $c_{ij} = \cos(\theta_{ij})$. In addition to the mixing angles θ_{ij}, Equation 2.16 contains a CP-violating phase δ_{CP} as free parameter. In principle, one could assign several phases to these matrices, which are only constrained by their unitarity. However, the choice of additional phases is irrelevant here, since the individual phases are not observable and can be represented by the single phase δ_{CP}. The matrix R inserts two additional phases ρ_1 and ρ_2. It exists only in case of Majorana neutrinos and is irrelevant here, as it drops out from the calculation of oscillation probabilities. Thus, these phases are assumed to be zero in the following [23, 4].

The remaining two parameters are given by the neutrino masses m_1, m_2 and m_3, corresponding to the neutrino mass eigenstates ν_1, ν_2 and ν_3. As shown later, neutrino oscillations depend only on the difference of the squared masses, reducing these three masses to only two parameters, while the third squared-mass difference can be derived by the other two.

The time expansion of each mass eigenstate i can be obtained by solving *Schrödinger's Equation*, as stated in Equation 2.17 (in natural units)

$$-i\frac{\mathrm{d}}{\mathrm{d}t}\left|\nu_i(t,\vec{x})\right\rangle = \mathcal{H}\left|\nu_i(t,\vec{x})\right\rangle = E_i\left|\nu_i(t,\vec{x})\right\rangle, \tag{2.17}$$

where $\mathcal{H}$ is the neutrino Hamiltonian and E_i the energy of the mass eigenstate i [23]. The solutions to this equation are plane waves, given by

$$\left|\nu_i(L(t))\right\rangle = e^{-iE_i t}\left|\nu_i(t=0)\right\rangle \approx e^{-i\left(p_\nu+\frac{m_i^2}{2p_\nu}\right)L(t)}\left|\nu_i(t=0)\right\rangle, \tag{2.18}$$

where $L(t) = |\vec{x}(t)|$ corresponds to the distance the neutrino has propagated after the time t, while $\vec{p}_\nu$ (with $p_\nu = |\vec{p}_\nu|$) and $\vec{x}$ are the momentum and location vector of the neutrino, respectively [23]. Note that for this derivation of vacuum oscillations, the momentum vector $\vec{p}_\nu$ is assumed to be the same for all neutrino states i. However, several other assumptions lead to the same result for the oscillation equations [9]. Moreover, more detailed calculations consider wave packages instead of plane waves, which are not discussed here for spatial reasons [24].

The approximation in the second step uses the relativistic identity $E_i^2 = \vec{p}_\nu^{\,2} + m_i^2$ and the fact that the neutrino masses m_i are substantially smaller than the total energy of the investigated neutrinos E_i. Thus, the energies E_i can be approximated by Equation 2.19 [9]

$$E_i \approx p_\nu + \frac{m_i^2}{2p_\nu}. \tag{2.19}$$

The resulting solutions for the mass eigenstates i can be converted to a solution for the flavor states α by calculating the corresponding linear combinations, as done in Equation 2.20 [9]

$$\left|\nu_\alpha(L)\right\rangle = \sum_{i=1}^{3} U_{\alpha,i}\left|\nu_i(t)\right\rangle \approx \sum_{i=1}^{3} U_{\alpha,i}e^{-ip_\nu L(t)}e^{-i\frac{m_i^2}{2p_\nu}L(t)}\left|\nu_i(t=0)\right\rangle, \tag{2.20}$$

where the phase $e^{-ip_\nu L(t)}$ is common to all neutrino states i and thus drops out from the following calculation of probabilities. The probability $P_{\nu_\alpha\to\nu_\beta}$ for a flavor transition between two flavors α and β for a neutrino with energy $E_\nu \approx p_\nu$ after a propagated distance L can be calculated according to Equation 2.21

$$P_{\nu_\alpha\to\nu_\beta}(E_\nu, L) = \left|\langle\nu_\beta|\nu_\alpha(L)\rangle\right|^2 = \sum_{i=1}^{3}\sum_{j=1}^{3} U_{\beta,i}^* U_{\alpha,j}^* U_{\alpha,i} U_{\beta,j} e^{-i\frac{\Delta m_{ij}^2 L}{2E_\nu}}, \tag{2.21}$$

where Δm_{ij}^2 is defined as $\Delta m_{ij}^2 = m_i^2 - m_j^2$. Thus, the oscillation probabilities depend only on the difference of the squared neutrino masses, not on the masses themselves. As a result, a measurement of the absolute neutrino mass is not possible in neutrino oscillation experiments [9].

Note that in literature, the convention, which mass difference is chosen to be expressed by the other two, varies. In particular, the choice of using Δm_{31}^2 or Δm_{32}^2 to describe atmospheric

oscillations often depends on the NMO, such that Δm^2_{31} (Δm^2_{32}) is conventionally chosen for NO (IO). This is discussed in more detail for the results in Chater 8. Table 2.1 summarizes the current state of the knowledge of these neutrino oscillation parameters as obtained from [25].

Table 2.1: Summary of the current knowledge of the neutrino oscillation parameters; best-fit values and uncertainties are taken from [26, 25]; the atmospheric mass difference is given in terms of $\Delta m^2_{3\ell} = \Delta m^2_{31} > 0$ for NO and $\Delta m^2_{3\ell} = \Delta m^2_{32} < 0$ for IO.

parameter	best fit value ($\pm 1\,\sigma$) NO	IO
$\sin^2(\theta_{12})$	$0.307^{+0.013}_{-0.012}$	$0.307^{+0.013}_{-0.012}$
$\sin^2(\theta_{23})$	$0.538^{+0.033}_{-0.069}$	$0.554^{+0.023}_{-0.033}$
$\sin^2(\theta_{13})$	$0.02206^{+0.00075}_{-0.00075}$	$0.02227^{+0.00074}_{-0.00074}$
$10^5 \frac{\Delta m^2_{21}}{\mathrm{eV}^2}$	$7.40^{+0.21}_{-0.20}$	$7.40^{+0.21}_{-0.20}$
$10^3 \frac{\Delta m^2_{3\ell}}{\mathrm{eV}^2}$	$+2.494^{+0.033}_{-0.031}$	$-2.465^{+0.032}_{-0.031}$
$\delta_{\mathrm{CP}}/^\circ$	234^{+43}_{-31}	278^{+26}_{-29}

Using the unitarity of U, the probability $P_{\nu_\alpha \to \nu_\beta}$ can be rewritten as in Equation 2.22

$$\begin{aligned} P_{\nu_\alpha \to \nu_\beta}(E_\nu, L) =& \delta_{\alpha,\beta} - 4 \sum_{i=1}^{2} \sum_{j=i+1}^{3} \mathrm{Re}\left\{U^*_{\beta,i} U^*_{\alpha,j} U_{\alpha,i} U_{\beta,j}\right\} \sin^2\left(\frac{\Delta m^2_{ij} L}{4E_\nu}\right) \\ & \pm 2 \sum_{i=1}^{2} \sum_{j=i+1}^{3} \mathrm{Im}\left\{U^*_{\beta,i} U^*_{\alpha,j} U_{\alpha,i} U_{\beta,j}\right\} \sin\left(\frac{\Delta m^2_{ij} L}{2E_\nu}\right), \end{aligned} \tag{2.22}$$

where the sign corresponds to neutrinos (+) and anti-neutrinos (-), respectively [27, 9].

One should note that Equation 2.22 is sensitive to the NMO in two ways. The first term of the equation includes Δm^2_{ij} only in sine-squared terms. Thus, each term is insensitive to the NMO. However, if all three mass differences Δm^2_{ij} are observed at the same time, the NMO could be determined by finding $\Delta m^2_{31} > \Delta m^2_{32}$ or vice versa. Thus, the observation of several oscillation channels at the same time allows to determine the NMO. Furthermore, the second term is sensitive to the NMO, since it includes sine-terms that are sensitive to the sign of Δm^2_{ij} in their argument. This requires that the mixing matrix U features elements with non-vanishing imaginary parts and thus $\delta_{\mathrm{CP}} \neq 0$. As a result, the sensitivity of many future experiments to the NMO depends strongly on the true value of δ_{CP} (cf. Section 2.5) [27, 28].

Apart from the sine-terms, Equation 2.22 depends only on constants given by the matrix elements of U. The sine-terms can be rewritten in SI-units to faciliate their understanding, as done in Equations 2.23 and 2.24 [27, 9]

$$\sin^2\left(\frac{\Delta m^2_{ij} L}{4E_\nu}\right) \overset{\text{SI-units}}{\approx} \sin^2\left(1.27 \cdot \frac{\Delta m^2_{ij}}{\mathrm{eV}^2} \cdot \frac{L/\mathrm{km}}{E_\nu/\mathrm{GeV}}\right), \tag{2.23}$$

$$\sin\left(\frac{\Delta m^2_{ij} L}{4E_\nu}\right) \overset{\text{SI-units}}{\approx} \sin\left(2.54 \cdot \frac{\Delta m^2_{ij}}{\mathrm{eV}^2} \cdot \frac{L/\mathrm{km}}{E_\nu/\mathrm{GeV}}\right). \tag{2.24}$$

Note that the oscillations depend only on the ratio of propagated distance over neutrino energy L/E_ν. For the values of Δm^2_{21} and Δm^2_{32} stated in Table 2.1, the oscillations become visible on two different L/E_ν-scales of $\sim 10^4\,\mathrm{km/GeV}$ and $\sim 10^3\,\mathrm{km/GeV}$ [27, 23].

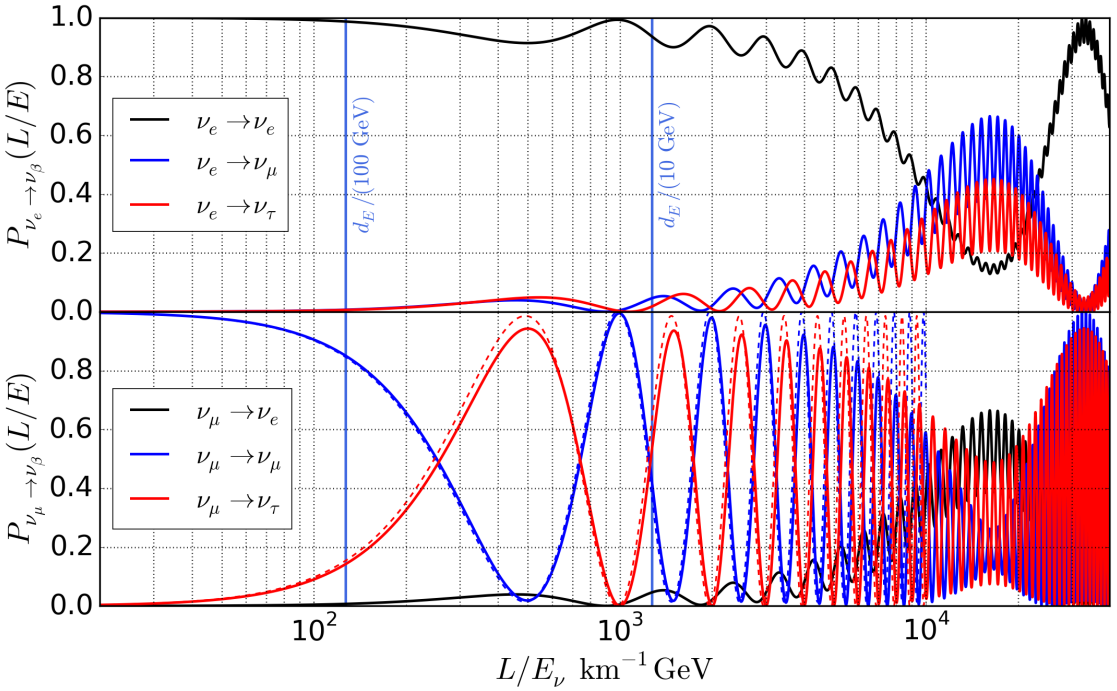


Figure 2.6: Vacuum oscillation probabilities for an electron (top) and muon neutrino (bottom) being injected; the solid lines show the full three-flavor oscillations, while the dashed lines represent the two-flavor approximation from Equation 2.26 up to $L/E_\nu < 10^4\,\mathrm{km\,GeV}^{-1}$; vertical solid lines indicate the ratio of the Earth's diameter d_E with certain neutrino energies.

In Figure 2.6, the oscillations are shown for various ratios of L/E_ν. The vertical, dashed lines illustrate the corresponding neutrino energies for using the diameter of the Earth $d_\mathrm{E} \approx 12\,742\,\mathrm{km}$ as baseline L [29].

As one can see, the oscillations arise first in the $\nu_\mu \to \nu_\tau$ channel for low values of L/E_ν, while the mixing of ν_μ or ν_τ with ν_e is much smaller for $L/E_\nu < 1000\,\mathrm{km\,GeV}^{-1}$. Thus, oscillations of atmospheric neutrinos, discussed in more detail in Section 2.4, with energies $E_\nu > 5\,\mathrm{GeV}$ and baselines of $L \sim d_E$ are mostly visible in the $\nu_\mu \to \nu_\tau$ channel. This decoupling of the oscillations channels is due to the small ratio of $\Delta m^2_{21}/\Delta m^2_{32} \sim 0.03$, causing oscillations between different flavors to dominate in different energy regimes.

Since the $\nu_\mu \to \nu_\tau$ channel is the dominant effect visible in atmospheric neutrino oscillations, the oscillation pattern can be approximated in a simple two-flavor model. It is briefly discussed in the following, as it is used for discussing the observed oscillation parameters in Section 8.1.2.

In the two-flavor model, the neutrino mixing matrix from Equation 2.1 simplifies to

$$U_{2\nu} = \begin{pmatrix} \cos(\theta_{2\nu}) & \sin(\theta_{2\nu}) \\ -\sin(\theta_{2\nu}) & \cos(\theta_{2\nu}) \end{pmatrix}, \tag{2.25}$$

with only one mixing angle $\theta_{2\nu}$. The matrix $U_{2\nu}$ can be plugged into Equation 2.22 to obtain the oscillation probability $P_{\nu_\mu \to \nu_\tau}(E_\nu, L)$ in Equation 2.26 [9]

$$P_{\nu_\mu \to \nu_\tau}(E_\nu, L) = \sin^2(2\theta_{2\nu}) \cdot \sin^2\left(1.27 \cdot \frac{\Delta m^2_{2\nu}}{\mathrm{eV}^2} \cdot \frac{L/\mathrm{km}}{E_\nu/\mathrm{GeV}}\right). \tag{2.26}$$

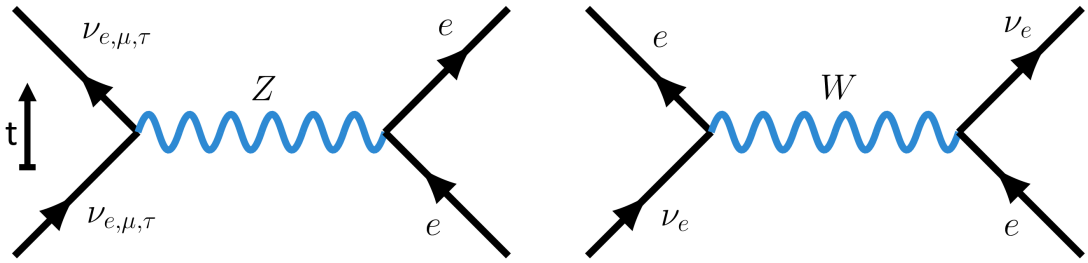
Note that the oscillation parameters $\theta_{2\nu}$ and $\Delta m^2_{2\nu}$ control the amplitude and the location of the flavor transition in L/E_ν, respectively. The two-flavor oscillations are shown in Figure 2.6 as dashed line up to $L/E_\nu < 10^4\,\mathrm{km\,GeV}^{-1}$. They resemble the full three-flavor oscillations for sufficiently small values of L/E_ν [9].

The oscillation amplitude in Equation 2.26 is proportional to $\sin^2(2\theta_{2\nu})$. Thus, the amplitude is maximal for $\theta_{2\nu} = 45°$, while positive and negative deviations from this value lead to a reduction of this amplitude. To first order, this is also the case for three-flavor oscillations. Thus, the value of $\theta_{23} = 45°$ or $\sin^2(\theta_{23}) = 0.5$ is often referred to as *maximum mixing*. As a result, the parameter space is split into $\sin^2(\theta_{23}) < 0.5$ and $\sin^2(\theta_{23}) > 0.5$, which is referred to as *first* (*left, lower*) and *second* (*right, upper*) *octant*, respectively.

The oscillation effects described above are valid for the propagation of neutrinos in vacuum. These effects can be modulated by matter effects, that arise for neutrinos propagating through matter. These modulations are essential for the signature of the NMO and discussed in the following section.

2.3.3 Matter Effects in Neutrino Oscillations

The vacuum oscillations discussed in the last section hold only for neutrinos propagating through vacuum or sufficiently thin matter. In reality, matter effects arise for typical oscillation baselines through Earth. The interactions of neutrinos with the electrons in the surrounding matter are the most relevant effect during the propagation, leading to a shift in the electron neutrino potential and thus, the neutrino oscillation pattern. The relevant Feynman diagrams for these interactions are those of *coherent forward scattering*, sketched in Figure 2.7. Note that the NC forward scattering is common to all neutrino flavors, while the CC one is only possible for electron neutrinos and anti-neutrinos. Thus, only the latter one changes the relative potential between the different neutrino flavors.



(a) Neutral Current (NC) (b) Charged Current (CC)

Figure 2.7: Leading order Feynman diagrams of neutrino forward scattering: the NC contribution (left) is independent of the neutrino flavor, while the CC contribution (right) is only possible for electron neutrinos.

These interactions modify the Hamiltonian from Equation 2.17 by a potential term V_e for electron neutrinos, as shown in Equation 2.27 [30]:

$$\mathcal{H} = \frac{1}{2E_\nu}\left[U\begin{pmatrix}0 & & \\ & m_2^2 - m_1^2 & \\ & & m_3^2 - m_1^2\end{pmatrix}U^\dagger + \begin{pmatrix}V_e & & \\ & 0 & \\ & & 0\end{pmatrix}\right]. \tag{2.27}$$

Note, that the Hamiltonian is now expressed in the flavor space $(\nu_e, \nu_\mu, \nu_\tau)^T$ by the transformation U to simplify the structure of the potential term. The electron neutrino potential V_e can be expressed by $V_e(\vec{x}) = \pm\sqrt{2}G_F N_e(\vec{x})$, where $N_e(\vec{x})$ is the local electron density in the medium and the sign corresponds to neutrinos and anti-neutrinos, respectively [30].

This gives rise to two different effects during the propagation: first, *resonant enhancement* and second, *parametric enhancement*. The resonant enhancement is often called *Mikhaev-Smirnov-Wolfenstein-effect* (MSW), labelled after the first people pointing out the impact of such matter interactions on the neutrino propagation in 1978 [31] and 1986 [32].

For three-flavor oscillations in an arbitrary electron density $N_e(\vec{x})$, the oscillation probabilities can only be derived numerically. However, the MSW effect can be expressed analytically in case of two-flavor oscillations in a constant neutrino profile $N_e(\vec{x}) = N_e$. As a result, the vacuum oscillation angle θ and squared-mass difference Δm^2 are modified according to Equations 2.28 and 2.29.

$$\Delta m^2 \rightarrow \Delta m^2_m = C \cdot \Delta m^2 \qquad \textbf{with } C = \sqrt{(\cos(2\theta) - A)^2 + \sin^2(2\theta)} \tag{2.28}$$

$$\sin(2\theta) \rightarrow \sin(2\theta_m) = \frac{\sin(2\theta)}{C} \qquad A = \pm\frac{2\sqrt{2}G_F N_e E_\nu}{\Delta m^2} \tag{2.29}$$

where the sign in the last line is for neutrinos (+) and anti-neutrinos (−). Therefore, the disappearance minima of neutrinos and anti-neutrinos are slightly shifted, where the size of the effect depends on the electron- and thus matter-density. The effect is maximal for a vanishing first term in Equation 2.28 for the definition of C, such that Equation 2.30 is satisfied

$$E_\nu = \pm\frac{\Delta m^2}{2\sqrt{2}G_F N_e}\cos(2\theta). \tag{2.30}$$

Again, the sign is for neutrinos (+) and anti-neutrinos (-), respectively. This resonance is often called *MSW-resonance*. It turns out that for neutrinos travelling through the Earth's core and mantle, the resonance is observed at energies of $E_\nu \sim 1-10\,\text{GeV}$, which is discussed in more detail in Section 2.4.2 [33].

In contrast to the MSW effect, parametric enhancement arises only in non-constant matter profiles $N_e(\vec{x})$, which vary periodically over the distance S, such that

$$V_e(x) = V_e(x + S), \tag{2.31}$$

for the electron neutrino potential, where for simplicity $x = |\vec{x}|$ is measured in direction of the neutrino propagation. The effect arises from a phase-shift $\Delta\phi$ that iteratively increases with each period of the matter profile, such that $\phi_n = n\Delta\phi$ is obtained after n iterations of $V_e(x)$. Although the Earth's matter profile is not really periodic, it features an increasing density towards the Earth's core followed by a decrease towards the surface. This mantle-core-mantle structure can be seen as a periodicity with $n = 1$, which is sufficient to give rise to parametric enhancement effects. However, these effects barely contribute above $\sim 1\,\text{GeV}$, such that the effect is of less importance for the analysis presented in the following chapters [34].

The matter effects in neutrino oscillations give a handle for separating the two Neutrino Mass Orderings. This is due to the fact that a non-zero electron density leads to an increase (decrease) in the energy of the electron neutrino (anti-neutrino) state. As a result, shifting the mass of the flavor state does not have the same effect for both of them:

For Normal Ordering, one has $\Delta m^2 = \Delta m^2_{31} > 0$. Thus, the sign in the definition of A in Equation 2.29 allows to satisfy the resonance condition in Equation 2.30 for neutrinos, while for anti-neutrinos no positive value leads to this resonance. In contrast, for Inverted Ordering one has $\Delta m^2 = \Delta m^2_{31} < 0$, which cannot give positive values of E_ν in Equation 2.30 for neutrinos, but only for anti-neutrinos [33].

Thus, matter effects arise mainly in the neutrino and anti-neutrino channel for Normal and Inverted Ordering, respectively. For atmospheric neutrinos, this is further discussed in Section 2.4.2.

2.4 Atmospheric Neutrinos

2.4.1 Cosmic Rays and Atmospheric Neutrino Generation

In 1912, Victor Hess discovered ionizing particles, hitting the Earth's atmosphere from outer space [35]. These particles were henceforth called *Cosmic Rays*. They give rise to atmospheric neutrinos, which are generated in the Earth's atmosphere by interactions of the primary Cosmic Rays with the molecules of the air.

The origin of Cosmic Rays is still unknown, but object of vast investigations. The most common theories predict Cosmic Rays to be accelerated in astrophysical objects like Supernovae or Blazars by shock-acceleration. At Earth, the Cosmic Ray spectrum above $\sim 10^{10}\,\text{eV}$ consists mainly of protons ($\approx 85\%$) that follow a broken power law behavior in the differential flux $\frac{d\Phi}{dE} \propto E^{-\gamma}$ over more than 10 orders of magnitude in Cosmic Ray energy. Beside protons, also heavier nuclei and electrons contribute to the total flux [36].

The broken power law changes its spectral index from $\gamma \approx 2.7$ to $\gamma \approx 3$ at $\sim 10^{15.5}\,\text{eV}$ and back to $\gamma \approx 2.7$ at $\sim 10^{18.5}\,\text{eV}$, which is often referred to as *knee* and *ankle*, respectively [11]. In addition, it features a cut-off at $\sim 5 \cdot 10^{19}\,\text{eV}$, which can be explained by the limited size of the accelerating astrophysical objects and/or the *Greisen-Zatsepin-Kuzmin* (GZK) cut-off [38, 39]. The change in the spectral index is often seen as an indication for different Cosmic Ray origins: while below the knee, the spectrum might be dominated by galactic sources, the galactic contribution dies out between knee and ankle, leaving a dominant extragalactic contribution above the ankle behind. In Figure 2.8, the broken power law is shown for several components of the flux below the knee, which is the relevant energy regime for the observation of atmospheric neutrino oscillations [36, 40].

In the Earth's atmosphere, these Cosmic Rays interact with the air's molecules via strong and electromagnetic interactions. Such process is sketched in Figure 2.9. These interactions generate charged mesons, which also interact with the molecules of the atmosphere or decay, leaving neutrinos and anti-neutrinos behind. At low neutrino energies of $E_\nu \lesssim 100\,\text{GeV}$, this process is dominated by pions, while at higher energies of $E_\nu \sim 100\,\text{GeV}$ kaons take over, as shown in Figure 2.10. At even higher neutrino energies of $E_\nu \gtrsim 100\,\text{TeV}$, a contribution from the production of heavier mesons is expected, carrying at least one charm- as a valence-quark. This contribution is commonly called *prompt flux*. However, there is no experimental evidence for such contribution so far [42].

The mesons decay into neutrinos and other light particles. For charged pions, the branching ratio of $\Gamma \approx 99.99\%$ guarantees an almost entire conversion into muons. The most important decay chains for a pion produced by a primary Cosmic Ray proton p and a nucleon $N \in \{p,n\}$

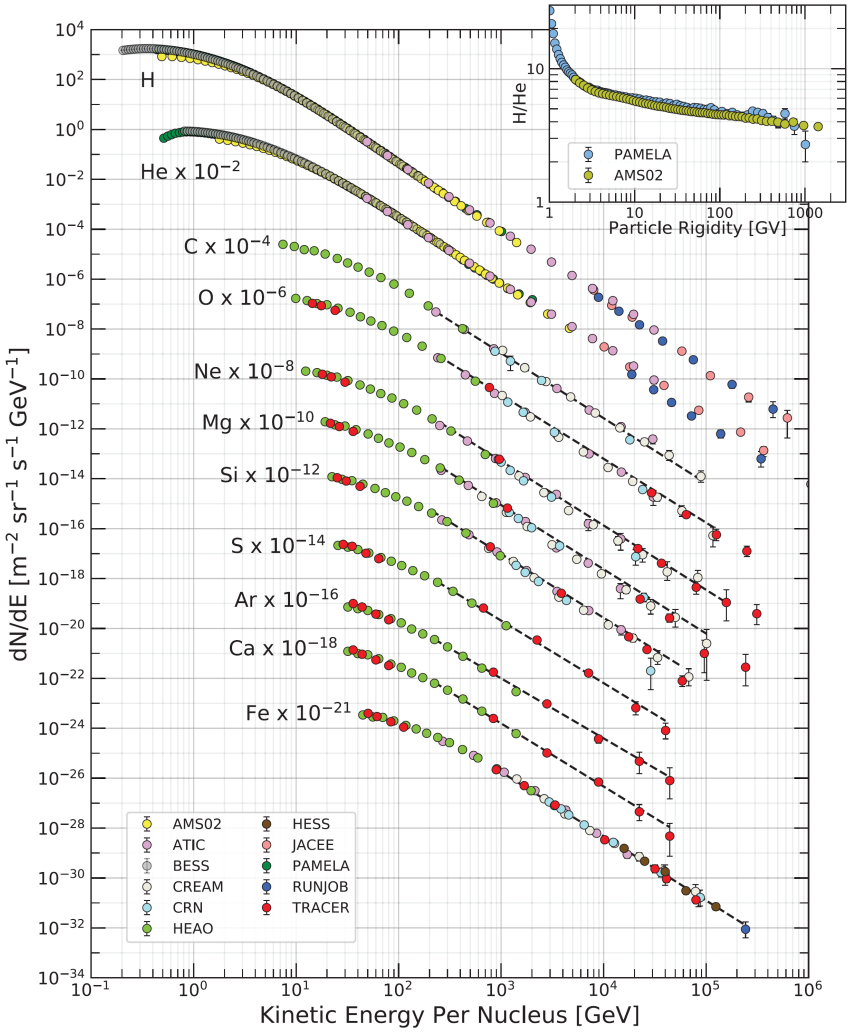


Figure 2.8: Differential Cosmic Ray spectrum for several nuclei as observed by various experiments below the knee at $\sim 10^{15.5}$ eV (from [37]).

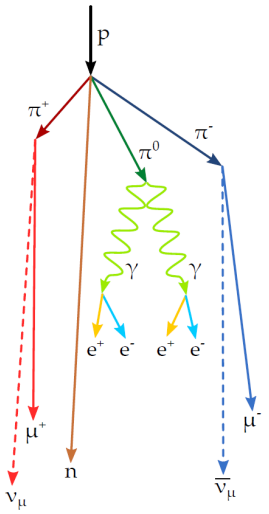


Figure 2.9: Sketch of an air shower generated in the Earth's atmosphere; the primary proton generates charged pions that decay, leaving neutrinos and anti-neutrinos as well as muons behind (from [41]).

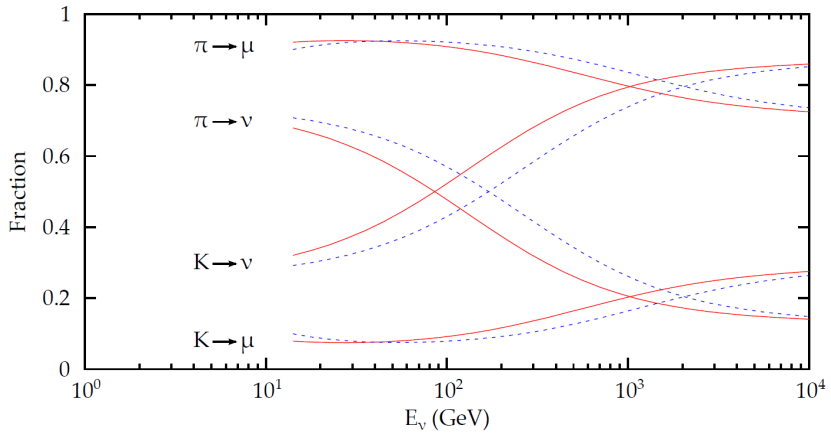


Figure 2.10: Energy-dependent production of atmospheric neutrinos and muons via pion and kaon decays for zenith-angles of 0° (solid) and 60° (dashed) at the surface (from [43]).

of the Earth's atmosphere are summarized in Equation 2.32 with the subsequent muon decays [43, 11]

$$\begin{aligned} p^+ + N \rightarrow \pi^\pm + X & \\ \pi^+ \rightarrow \mu^+ + \nu_\mu & \quad (\Gamma \approx 100\%) \\ \mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu & \\ \pi^- \rightarrow \mu^- + \bar{\nu}_\mu & \quad (\Gamma \approx 100\%) \\ \mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu . & \end{aligned} \tag{2.32}$$

One should note that these decay chains suggest a $(\nu_\mu + \bar{\nu}_\mu)/(\nu_e + \bar{\nu}_e)$-ratio of 2 : 1 in case all pions decay before losing a substantial amount of their energy by interactions with the Earth's atmosphere, which is a valid assumptions for low energies. Also $\nu_\mu/\bar{\nu}_\mu \sim 1$ and $\nu_e/\bar{\nu}_e \sim \mu^+/\mu^-$ can be estimated for the flavor-dependent neutrino to anti-neutrino ratio [43].

In contrast to pions, the kaon production and decay is not dominated by one process, but a combination of several non-negligible $K^\pm$, K^0_L and K^0_S processes. The most relevant kaon production and decay chains are summarized in Equations 2.33 and 2.34 [11]

$$\begin{aligned} p^+ + N \rightarrow K^\pm + X & \\ K^\pm \rightarrow \mu^\pm + \overset{(-)}{\nu}_\mu & \quad (\Gamma \approx 64\%) \rightarrow \text{cf. pion decay chain} \\ K^\pm \rightarrow \pi^\pm + \pi^0 & \quad (\Gamma \approx 21\%) \rightarrow \text{cf. pion decay chain} \\ K^\pm \rightarrow \pi^\pm + \pi^\pm + \pi^\mp & \quad (\Gamma \approx 6\%) \rightarrow \text{cf. pion decay chain} \\ K^\pm \rightarrow \pi^0 + e^\pm + \overset{(-)}{\nu}_e & \quad (\Gamma \approx 5\%) \rightarrow \text{cf. pion decay chain} \\ K^\pm \rightarrow \pi^0 + \mu^\pm + \overset{(-)}{\nu}_\mu & \quad (\Gamma \approx 4\%) \rightarrow \text{cf. pion decay chain} \end{aligned} \tag{2.33}$$

$$\begin{aligned} p^+ + N \rightarrow K^0 + X & \\ K^0_L \rightarrow \pi^\pm e^\mp + \overset{(-)}{\nu}_e & \quad (\Gamma_L \approx 41\%) \rightarrow \text{cf. pion decay chain} \\ K^0_L \rightarrow \pi^\pm \mu^\mp + \overset{(-)}{\nu}_\mu & \quad (\Gamma_L \approx 27\%) \rightarrow \text{cf. pion decay chain} \\ K^0_S \rightarrow \pi^\pm \pi^\mp & \quad (\Gamma_S \approx 69\%) \rightarrow \text{cf. pion decay chain,} \end{aligned} \tag{2.34}$$

where the subsequent decays of muons and pions are analogous to the pions in Equation 2.32.

The slope of the atmospheric neutrino spectrum is driven by two processes: first, the decay of the mesons and second their interaction between generation and decay with the Earth's atmosphere. With increasing energy, the lifetime of pions and kaons increases in the laboratory frame. As a result, they have more time to interact with the surrounding molecules before they decay. Thus, the mesons lose more energy, having less energy left to be passed to the neutrino and the muon after decay. The same is true for the decay of the muons, which lose a crucial amount of their energy, as soon as they hit the Earth's surface and then drop out from the production of high-energy neutrinos [40, 43].

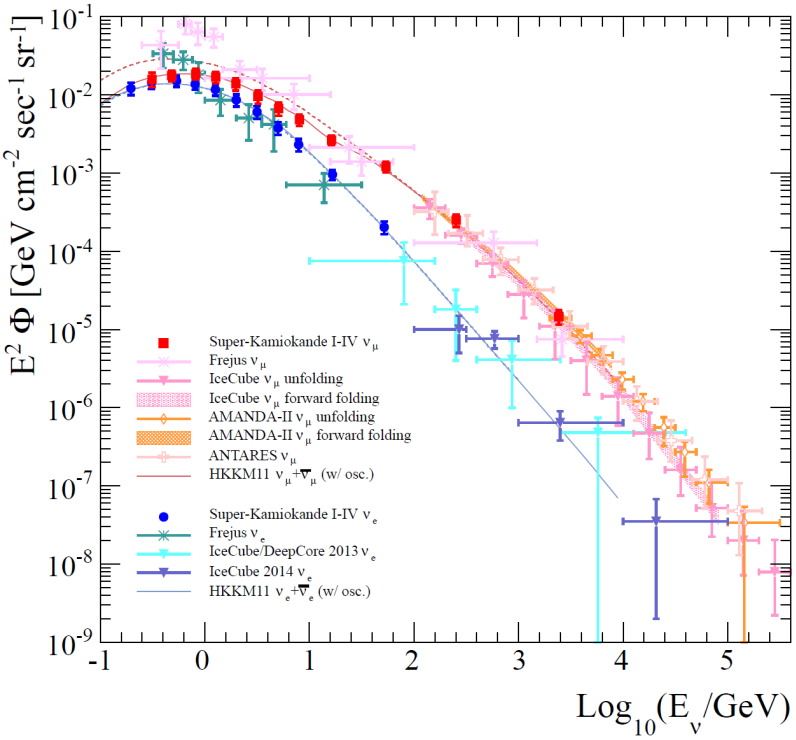


Figure 2.11: Energy spectrum of the atmospheric electron and muon neutrino flux as predicted by the *HKKM research group* [44] and seen by various experiments (from [45]).

Furthermore, the amount of matter passed by a meson/muon per unit time depends on the zenith-angle. Since the air is thinner at higher altitudes, a pion that was generated in the upper atmosphere loses less energy under a flat zenith-angle than it does when crossing the atmosphere vertically. As a result, the energy-dependence of atmospheric neutrinos features a zenith-dependent structure [40].

The total spectrum can be described approximately by a $\frac{\mathrm{d}\Phi_\nu}{\mathrm{d}E_\nu} \propto E_\nu^{-3.7}$ power law. The spectral index arises from the spectral index of the Cosmic Ray primary spectrum ($\gamma \approx 2.7$) and an additional factor of E_ν, describing the energy-dependent lifetime of the meson. For high energies, the elongated lifetime of the muon and the constant interaction probability lead to an increased energy-loss and thus decreased remaining energy of the decaying meson, which softens the observed atmospheric neutrino spectrum [40].

More advanced calculations of the atmospheric fluxes take the composition of the primary spectrum, energy-dependent cross-sections, the inclination angle, the point of observation and global, atmospheric conditions into account. As a result, these predictions come with non-negligible uncertainties from multiple sources, feature seasonal variations and have to be

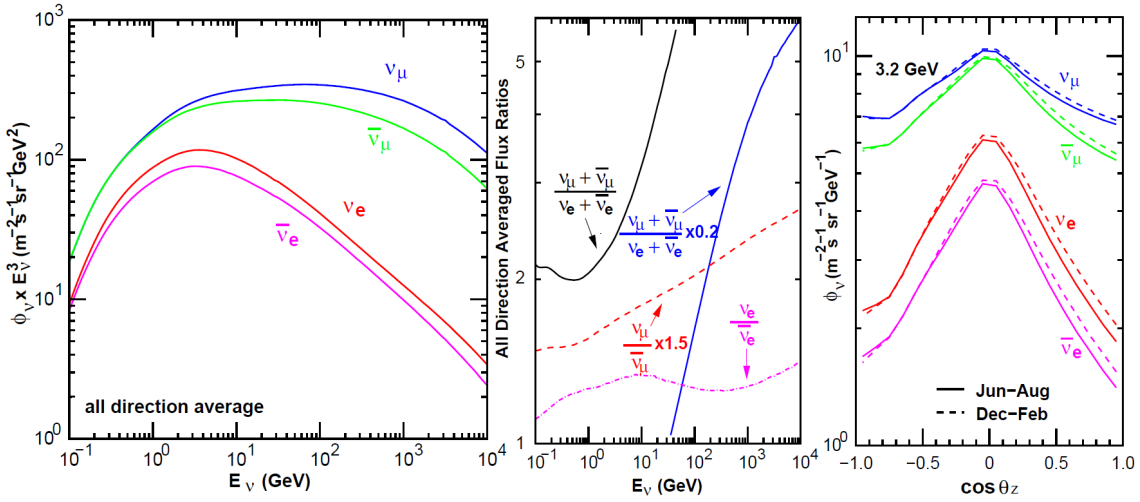


Figure 2.12: Left: energy spectrum of atmospheric neutrinos for $\nu_e, \nu_\mu, \bar{\nu}_e$ and $\bar{\nu}_\mu$; center: energy-dependent flavor- and neutrino to anti-neutrino-ratio; right: zenith-angle spectrum of atmospheric neutrinos for various flavors and two different seasons at $E_\nu = 3.2\,\text{GeV}$; all fluxes are calculated for the Geographic South Pole (modified from [46]).

benchmarked by comparisons to experimental data. Some measurements of the atmospheric neutrino energy spectrum are shown in Figure 2.11. It comprises measurements from various experiments stretching over ~ 6 orders of magnitude in neutrino energy [40, 46].

The atmospheric neutrino- and anti-neutrino fluxes used in this work are based on predictions from [46] for the South Pole. In Figure 2.12, the resulting energy spectrum (left), the flavor and neutrino to anti-neutrino ratio (center) and the neutrino zenith spectrum are shown for the predicted fluxes [46]. The uncertainties on these fluxes are parametrized based on [47], which estimates the uncertainties on the energy-, flavor- and zenith-spectrum for neutrinos to anti-neutrinos based on [48]. The resulting systematic uncertainties for this work are discussed in detail in Section 6.3.2.

Note that the atmospheric fluxes in Figure 2.12 are given only for electron and muon neutrinos, since tau neutrinos are not generated in the pion and kaon decays. For high energies, tau neutrinos can be generated in the prompt contribution by the decay of charmed mesons like D_s. However, even for these processes, the tau neutrino flux is an order of magnitude below the one obtained from muon neutrinos [49]. In contrast, the tau neutrinos observed in atmospheric neutrino measurements arise from neutrino oscillations, as described in Section 2.4.2.

The measurement of the appearance of tau neutrinos in the atmospheric neutrino flux is a key goal of many experiments and allows for probing the unitarity of the PMNS matrix U. First measurements of the tau-appearance by the *Super-Kamiokande* detector are consistent with the standard-oscillation prediction [50], while future experiments like the *IceCube-Gen2* detector are aiming for more precise measurements of the tau-neutrino normalization [51].

Moreover, Figure 2.12 illustrates the difference between the neutrino and anti-neutrino fluxes. At low energies, the $\nu_\mu/\bar{\nu}_\mu$-ratio is ~ 1, as the neutrino flux is predominantly produced by the decay of charged pions, with each pion generating one muon neutrino and anti-neutrino within its decay chain (cf. Equation 2.32). At higher energies, the ra-

tio changes, as muons are capable to reach the ground. In that case, muons lose most of their energy before they decay, such that the neutrino to anti-neutrino ratio depends on the π^+/π^--ratio and an increasing contribution from kaon decays. The resulting neutrino fluxes exceed the anti-neutrino fluxes, due to the dominance of positive charges in the Cosmic Ray primaries and in the nuclei of the Earth's atmosphere. As a result, the hadronic interactions of the air shower produce positive mesons more frequently than negative ones, leading to the observed preference for neutrinos over anti-neutrinos at high energies [52, 40]. The difference between the neutrino and anti-neutrino fluxes and cross-sections give rise to an important feature of the NMO signature presented in Section 2.4.2.

Besides atmospheric neutrinos, atmospheric muons are a relevant background for this work. Since these muons are capable of travelling several kilometers through solid matter, even underground neutrino detectors are affected by an atmospheric muon background. In the following analysis, these muons are not removed completely from the data sample and thus, need to be modelled accurately. This is described in Chapter 4.

Note that the fluxes of atmospheric neutrinos and muons feature seasonal variations: As the temperature in the atmosphere changes over the year, the neutrino and muon production are affected by the changes in the air density. However, these changes are small and the following analysis is based on a multi-year data-sample, such that seasonal changes are averaged out. Thus, they are not discussed further in the following, as their impact was found to be negligible for the presented analysis [53, 40].

2.4.2 Oscillations of Atmospheric Neutrinos

Once generated, atmospheric neutrinos propagate almost undisturbed through the atmosphere and the Earth's matter. Their low cross-section (cf. Section 2.2) avoids a notable decrease in the atmospheric neutrino flux, due to interactions with the surrounding matter. Thus, they can be measured anywhere on Earth coming from all zenith-angles. Nonetheless, underground observations are of course preferable, due to the otherwise large background from atmospheric muons.

During this propagation through Earth, neutrinos can undergo neutrino oscillations, as discussed in Section 2.3, before being measured by an underground detector.

The neutrino propagation from the point of their generation through the Earth is sketched in Figure 2.13. Since the Earth's geometry is known, the zenith-angle of an incoming neutrino can directly be linked to the distance it propagated. For a spherical Earth and neutrinos being generated at an average height of $L_1 \approx 19\,\text{km}$, the corresponding relation $L_\nu(\theta_\nu)$, can be written as

$$L_\nu(\theta_\nu) = \sqrt{(r_E - L_1)^2 + r_E^2 - \left[2r_E(r_E - L_1)\cos\left(\arcsin\left(\left(1 - \frac{L_1}{r_E}\right)\sin(\theta_\nu)\right)\right)\right]} \tag{2.35}$$

where θ_ν and L_ν are the zenith-angle and the propagated distance of the incoming neutrino, respectively, while $r_\text{E} \approx 6371\,\text{km}$ is the Earth radius (cf. Section 2.3.2) [29, 56].

As shown in Figure 2.13, the Earth features an inhomogeneous density profile, which is described by the *Preliminary Reference Earth Model* (PREM). It parametrizes the Earth's density by a core-mantle inhomogenity, where the core is additionally split into an inner and an outer part of higher and lower density, respectively. The resulting density profile is

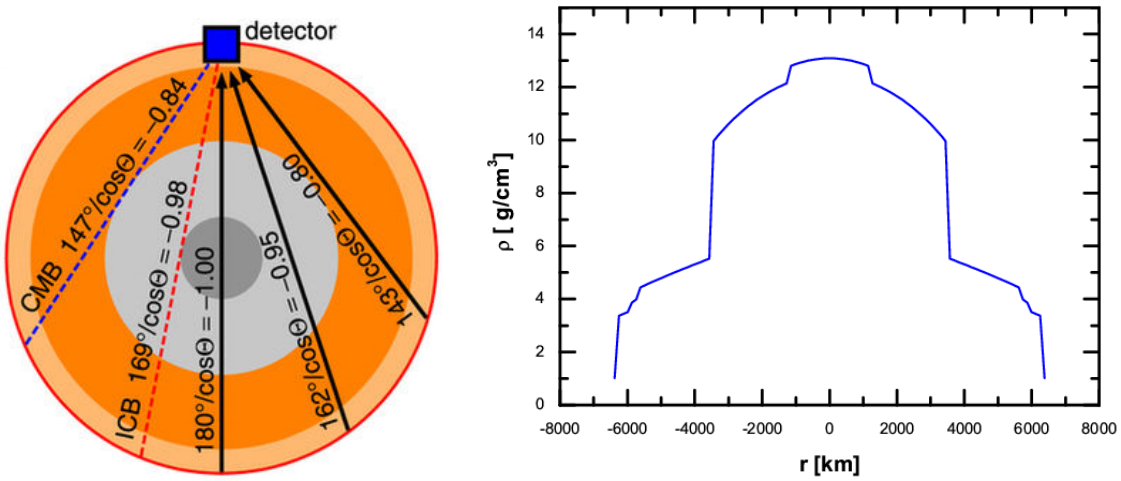


Figure 2.13: Left: visualization of neutrinos propagating through the Earth at Core-Mantle-Boundary (CMB) and Inner-Core-Boundary (ICB) in the PREM Earth model (modified from [54]); right: density profile of the PREM Earth model, consisting of a an inner core, an outer core, a mantle and some minor features (modified from [55]).

shown in Figure 2.13 (right). For the following studies, the *PREM12* model is used, which approximates the PREM model by 12 layers of constant density [55, 57].

Although the neutrino flux is not notably reduced during propagation, the matter profile has some impact on the neutrino oscillations, as discussed in Section 2.3.3. Due to the MSW-effect and parametric enhancement, the oscillation pattern is modified with respect to vacuum oscillations. Depending on the Neutrino Mass Ordering, these matter effects arise mainly in the neutrino or anti-neutrino sector. This is visualized in Table 2.2, where the observation channels undergoing matter effects are underlayed in blue, while observation channels featuring no or negligible matter effects are underlayed in red.

Table 2.2: Oscillation patterns relevant for atmospheric neutrinos; the hypotheses of NO, IO and vacuum oscillations are shown for neutrinos and anti-neutrinos; blue background color: matter effects in this channel; red background color: vacuum (or similar) oscillations in this channel

	NO	IO	Vacuum
neutrinos	$\nu_e \to \nu_\alpha$ $\nu_\mu \to \nu_\alpha$ (Fig. 2.14(a),(b))	$\nu_e \to \nu_\alpha$ $\nu_\mu \to \nu_\alpha$ (analogous)	$\nu_e \to \nu_\alpha$ $\nu_\mu \to \nu_\alpha$
anti-neutrinos	$\bar{\nu}_e \to \bar{\nu}_\alpha$ $\bar{\nu}_\mu \to \bar{\nu}_\alpha$ (Fig. 2.14(c),(d))	$\bar{\nu}_e \to \bar{\nu}_\alpha$ $\bar{\nu}_\mu \to \bar{\nu}_\alpha$ (analogous)	$\bar{\nu}_e \to \bar{\nu}_\alpha$ $\bar{\nu}_\mu \to \bar{\nu}_\alpha$
overall signature	high matter effects	low matter effects	no matter effects

For these processes, the transition probabilities for a generated electron or muon neutrino to show up as an electron, muon or tau neutrino after propagating through Earth under a certain zenith-angle θ_ν and a neutrino energy E_ν are shown in Figure 2.14. However, since

all oscillation channels of the same background color in Table 2.2 are almost identical, they are shown only for the case of Normal Ordering for $\nu_{e/\mu} \to \nu_\alpha$ (blue) and $\bar{\nu}_{e/\mu} \to \bar{\nu}_\alpha$ (red) with $\alpha \in \{e,\mu,\tau\}$. However, the remaining processes can be derived approximately using Table 2.2. In this work, the oscillation probabilities with matter effects are calculated with the *Prob3* software [58], which is based on [59]. Moreover, an implementation of the analytic solution from Equation 2.22 is used for vacuum oscillations.

For a generated muon neutrino, the mixing with tau-neutrinos is strong below $E_\nu \sim 30\,\mathrm{GeV}$, which leads to the effect of ν_μ-disappearance and ν_τ-appearance. Both effects can be observed with similar analyses as the one presented in this work. Since atmospheric muon neutrinos are abundant, the transition probability $p(\nu_\mu \to \nu_\tau)$ is high, and no ν_τ-neutrinos are produced in the atmosphere, both processes can be measured with high significances with recent atmospheric data [60, 61, 62].

The matter effects can easily be understood by comparing Figure 2.14(a) to 2.14(c) and Figure 2.14(b) to 2.14(d). In the electron neutrino sector, an increased oscillation probability arises in case of matter effects, while the electron anti-neutrinos follow almost vacuum oscillations, in which they barely mix with the other flavors. The sudden jump at $\cos(\theta_\nu) \approx -0.8$ arises at the core-mantle boundary, which changes the observed matter effects. The increased transition probabilities at $\log_{10}(E_\nu/\mathrm{GeV}) \approx 0.7$ are mainly due to the MSW-resonance. However, the mixing with ν_μ and ν_τ neutrinos is in general small for neutrinos and anti-neutrinos.

In the muon neutrino channel, matter effects are mainly visible for long oscillation baselines with $\cos(\theta_\nu) \lesssim -0.6$ and low energies of $E_\nu \lesssim 15\,\mathrm{GeV}$, i.e. below the first oscillation extremum. For anti-neutrinos, the matter effects are not visible in case of NO, such that only vacuum-like oscillations are seen in Figure 2.14(d), while for IO, this behavior flips for neutrinos and anti-neutrinos. As a result, a detector capable of distinguishing neutrinos and anti-neutrinos, only needs to show in which of these channels matter effects arise to determine the NMO.

Note that this might lead to the naive impression, that NO and IO are not distinguishable in a detector that does only observe $\nu + \bar{\nu}$ and is not capable of separating neutrinos from anti-neutrinos on an event-by-event basis. However, this is not true: due to the increased cross-section for neutrino-nucleon interactions compared to anti-neutrino-nucleon interactions (cf. Section 2.2), the strength of the signature from matter effects is different for NO and IO. This asymmetry is enhanced by the fact that the atmospheric neutrino flux is higher than the corresponding anti-neutrino flux, as discussed in Section 2.4. Thus, in an atmospheric neutrino detector more neutrino than anti-neutrino events are expected. This means that in case of NO strong matter effects are observed in the combined $\nu + \overline{\nu}$-channel, while for IO, weak matter effects are observed. This is summarized in the last row of Table 2.2.

Additionally, neutrinos and anti-neutrinos can be distinguished on a statistical basis, even if the charge of the secondary lepton cannot be observed. The difference in the signature is based on measuring the fraction of neutrino energy passed to the secondary lepton. This inelasticity, i.e. Bjorken-y, distribution is different for neutrinos and anti-neutrinos, as shown in Figure 2.4. This difference allows to estimate the neutrino-to-anti-neutrino ratio from the reconstructed Bjorken-y distribution. However, this effect is small and only observable for muon neutrinos in CC interactions, as discussed in Section 3.3.

Beside the measurement of the NMO, these matter effects can also be used for neutrino Earth tomography. Such measurements are investigated for many future experiments aiming for a

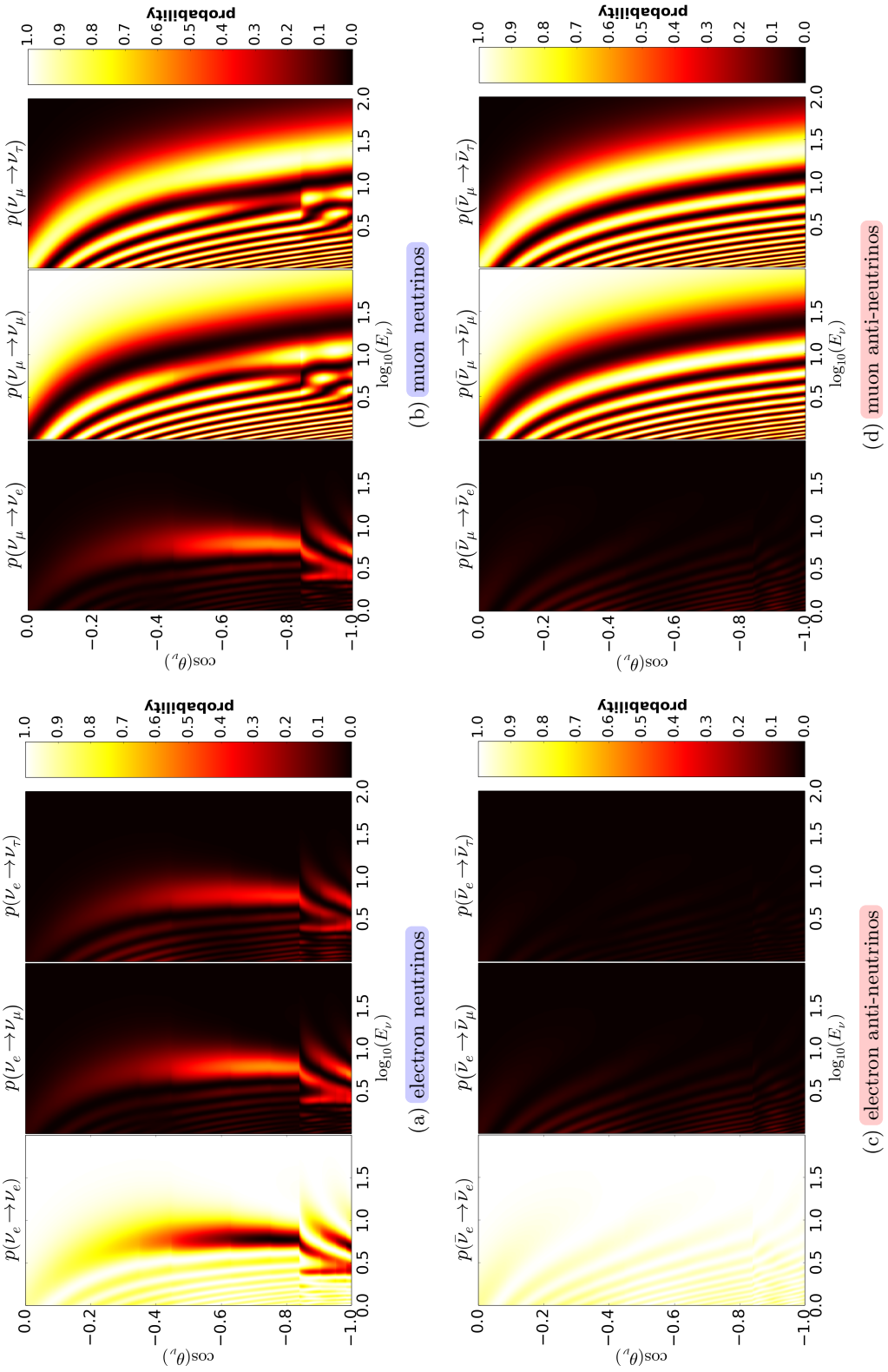


Figure 2.14: Atmospheric oscillation probabilities for $\overset{(-)}{\nu}_e \to \overset{(-)}{\nu}_\alpha$ and $\overset{(-)}{\nu}_\mu \to \overset{(-)}{\nu}_\alpha$ with $\alpha \in \{e,\mu,\tau\}$, depending on the zenith-angle θ_ν and the neutrino energy E_ν, for neutrinos (top) and anti-neutrinos (bottom); labels with blue (red) background color indicate matter effects (vacuum oscillations) in this channel; the oscillation parameters used are taken from Table 2.1.

measurement of the NMO with atmospheric neutrinos [63, 64]. However, this is beyond the scope of this work and requires a high sensitivity towards matter effects, as not only their existence needs to be observed, but also modulations on top of them. Such sensitivity is not reached for the analysis presented in the following.

2.5 Status of the Neutrino Mass Ordering Measurements

There are two common approaches to determine the NMO: First, using oscillation experiments, which are sensitive to the mass differences between the neutrino states, and second, using experiments that are directly sensitive to the neutrinos' absolute masses [4].

The first group comprises reactor, accelerator (long-baseline) and atmospheric neutrino experiments. All of them measure the NMO via modulations in the observed pattern of neutrino oscillations. However, while accelerator and atmospheric neutrino experiments use mainly matter effects to determine the NMO, reactor experiments use the small difference between Δm^2_{32} and Δm^2_{31} (i.e. Δm^2_{21}) by measuring multiple oscillation channels at once. Thus, they are capable to determine the NMO purely from vacuum oscillations [4].

The second group comprises β- and $\beta\beta$-experiments as well as cosmological data. All of them are capable to constrain an *effective neutrino mass* that is given by a superposition of the neutrino masses m_1, m_2 and m_3 (e.g. the effective mass $m_{\text{cosmo}} = m_1 + m_2 + m_3$ for cosmological data). Assuming the lightest of the three neutrino masses to be zero, the assumption of NO or IO implies a minimum value for such effective mass. Thus, these experiments can constrain the NMO by measuring or excluding the corresponding minimum mass predicted by one of the ordering hypotheses. An overview of these constraints from cosmology, β- and $\beta\beta$-decays is added to Appendix A [4, 6].

For accelerator experiments, recent indications on the NMO come from the *T2K* [65, 66], *MINOS/MINOS+* [67, 68], and *NOνA* [69, 70] Collaborations, with the most significant result being a slight preference of $1.8\,\sigma$ for NO by NOνA [70]. In addition, the *Super-Kamiokande* Collaboration provides some result on the NMO based on a measurement of atmospheric neutrinos, with the observed preference for NO being on a $\sim 2\,\sigma$-level [71]. In contrast to NOνA, the result is based on a full statistical analysis of their χ^2-fit, which is similar to the Pseudo-Experiments method described in Section 7.1. However, all of these measurements are at a $\lesssim 2\,\sigma$-level. Thus, from current oscillation experiments only weak evidence exists on the nature of the NMO.

To provide a more significant, global best-fit of the NMO, the oscillation results from several experiments are combined by *NuFit* [26, 25]. Besides the best-fit oscillation parameters in Table 2.1, they provide a global-fit for the NMO based on the ordering-dependent combination of the oscillation measurements. The resulting best-fit prefers NO over IO by $\sim 2\,\sigma$ for the oscillation parameters stated in Table 2.1 [26, 25]. Note that this global-fit does not yet include the most recent results from all above mentioned experiments. Thus, an increased significance for NO is expected from global best-fits in the near future.

For non-oscillation experiments, the strongest indication on the NMO comes from cosmological data, which prefers NO over IO at the level of $\Delta\chi^2_{\text{NO}-\text{IO}} \approx 0 - 1.6$ [12]. However, it strongly depends on the included data and cosmological model, causing the observed range in the NMO preference (cf. Appendix A) [6, 12].

Besides these currently data-taking experiments, several future experiments are proposed,

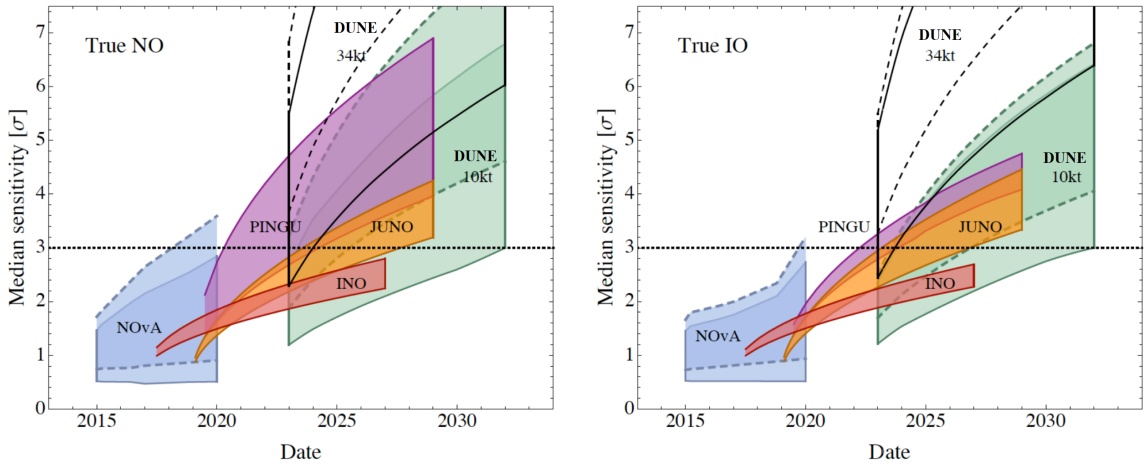


Figure 2.15: Projected median sensitivity to the NMO for several future experiments as by 2014; the shaded areas indicate the dependence on δ_{CP} for NOνA and DUNE, the dependence on θ_{23} for PINGU and INO and the energy resolution uncertainty for JUNO; for DUNE and NOνA, the solid (dashed) lines are obtained for $\theta_{23} = 40°$ (50°) (from [72]).

which aim for an NMO measurement within the next 5 to 15 years. Their time-dependent projected sensitivity to the NMO is summarized in Figure 2.15 for both orderings. Note that these projected sensitivities change constantly and depend on several assumptions. They are shown here by the state of 2014 just to illustrate the general state of the field and the dependence of the sensitivities on different uncertainties. However, the expected start of data-taking and sensitivities changed in the meantime for several experiments, e.g. for the PINGU projection [7].

For accelerator experiments, the most prominent proposal is the *Deep Underground Neutrino Experiment* (DUNE), which reaches $5 - 20\,\sigma$ projected sensitivity for 10 years of data, depending on the true value of δ_{CP} [73, 74]. For reactor experiments, one of the leading efforts is the *Jiangmen Underground Neutrino Observatory* (JUNO), which aims to be taking data by 2020. Its projected sensitivity reaches $\sim 3\,\sigma$ for the NMO with 6 years of data, which is mostly independent of the value of the other oscillation parameters, in particular δ_{CP} [75, 76].

Finally, for atmospheric neutrino experiments, several ice and water Cherenkov detectors are proposed like the *Precision IceCube Next Generation Upgrade* (PINGU) [7], the *Oscillation Research with Cosmics in the Abyss* (ORCA) [3] experiment, the *India-based Neutrino Observatory* (INO) or the *Hyper-Kamiokande* [77] experiment.

They all feature a similar measurement principle. However, only the PINGU detector and its projected sensitivity are described in more detail in Section 3.6, after the introduction of the IceCube and DeepCore detector. Its design is similar to and motivated by the design of the currently running IceCube DeepCore detector. Besides the experimental result, the following work is largely dedicated to providing a full-chain analysis of the Neutrino Mass Ordering on DeepCore data as a *proof-of-concept* for a future PINGU detector.

3 The IceCube DeepCore Detector

3.1 The IceCube Detector Set-Up

The *IceCube Neutrino Observatory* is a $\sim 1\,\mathrm{km}^3$ size neutrino detector at the Geographic South Pole. Completed in December 2010, it is a multi-purpose Cherenkov detector with the primary goal to measure the astrophysical neutrino flux and its sources [78].

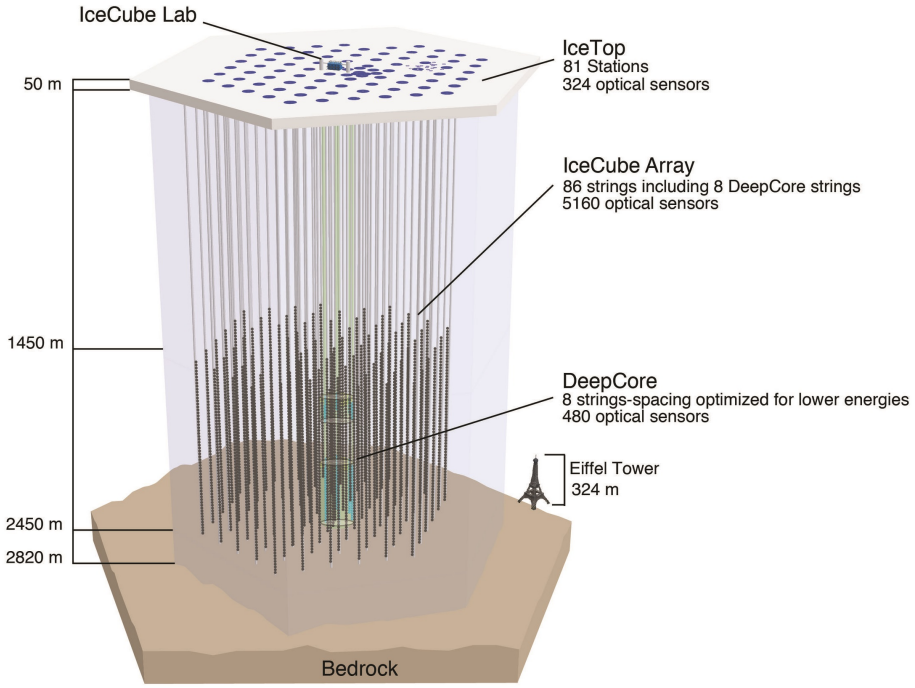


Figure 3.1: The IceCube detector set-up, showing the IceCube, DeepCore and IceTop arrays (from [79]).

The reason for its location at the South Pole is the need for a large mass of a transparent, heavy medium. Since the neutrino cross-section and the astrophysical fluxes are low, neutrino measurements require a large active mass to obtain reasonable neutrino rates. In addition, the detector medium needs to be transparent, since neutrino interactions are detected via the *Cherenkov light* emission of their secondary particles [80]. The transparent medium allows the photons to propagate towards a photo sensor. At the South Pole, the $\sim 2.8\,\mathrm{km}$ thick ice shield provides such a large volume of clear, transparent medium [78].

The IceCube detector consists of 86 strings running through the clear Antarctic ice vertically

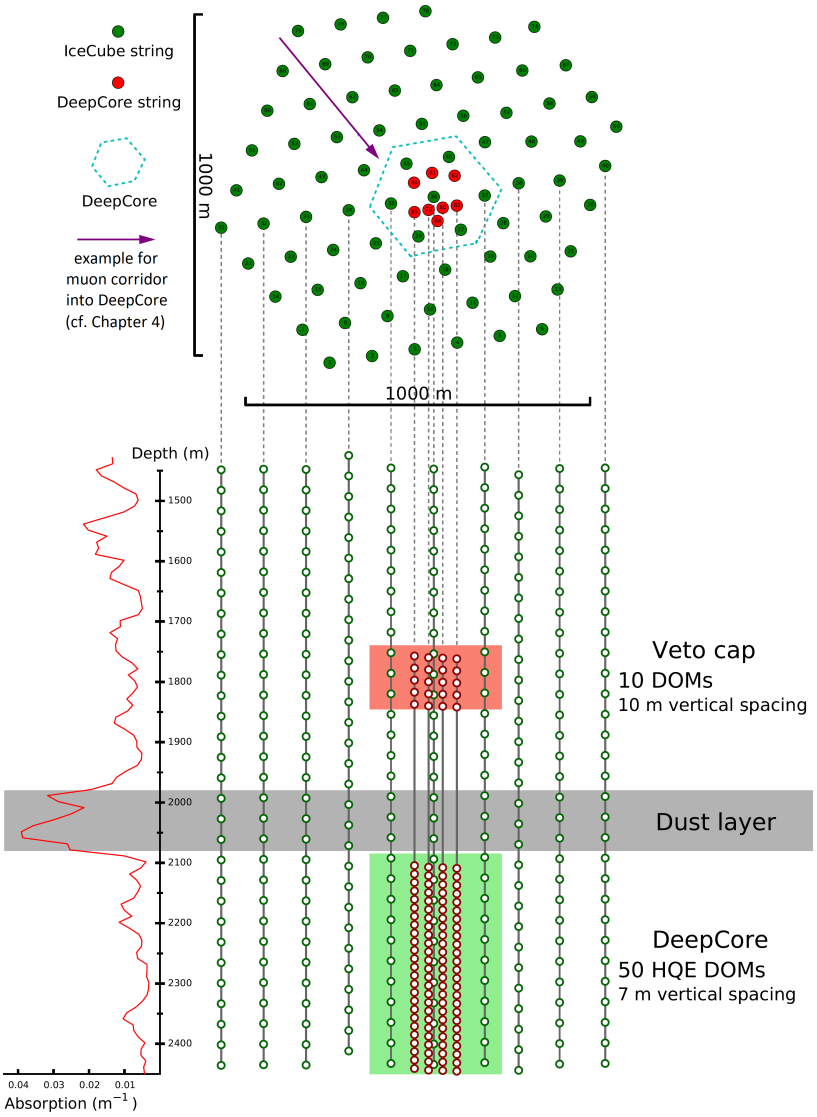


Figure 3.2: Projection of the IceCube detector set-up, sketching the string and DOM densities of the IceCube and DeepCore array (modified from [81]).

from the ice surface to almost the bedrock, carrying 5160 *Digital Optical Modules* (DOMs) in depth between 1450 m and 2450 m. These DOMs are equipped with *photomultiplier-tubes* (PMTs) [82] to measure the emitted Cherenkov light from neutrino interactions, as discussed in Section 3.2. The detection hardware of the DOMs is described in more detail in Section 3.5.1 [78, 82].

The strings are arranged in a hexagonal structure, leaving ~ 125 m space between strings horizontally, while the vertical distance between DOMs is only ~ 17 m. The set-up of the IceCube array is shown in Figure 3.1 and 3.2. With this set-up, the IceCube array is capable to detect neutrinos above $E_\nu \sim 100$ GeV.

Additionally, a more densly instrumented volume exists within the IceCube array, forming the *DeepCore* sub-detector, which is located roughly in the center of the IceCube array. It is vertically split into a shallow veto-cap for atmospheric muons at a depth of ~ 1800 m and a deeper, fiducial volume in a depth of $\sim 2100 - 2450$ m, where the ice is very clear. It con-

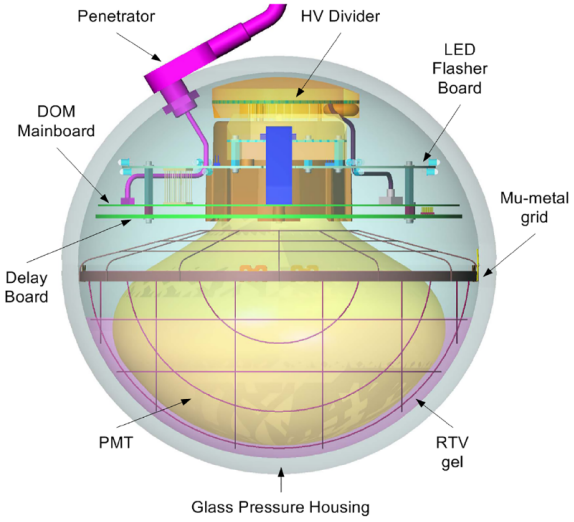


Figure 3.3: Structure of the Digital Optical Modules (DOMs), showing the PMT and the subsequent digitizing electronics (from [79]).

sists of eight DeepCore-strings surrounding one IceCube string. On six of these strings, the vertical distance between DOMs is decreased to only $\sim 7\,\mathrm{m}$, while the horizontal spacing between strings in DeepCore is roughly $\sim 75\,\mathrm{m}$. In addition to the increased density of DOMs, the DeepCore DOMs feature an enhanced quantum-efficiency, as described in Section 3.5.1. The remaining two strings of DeepCore consist of a mix of IceCube and DeepCore DOMs with a vertical spacing of $\sim 10\,\mathrm{m}$. Due to the more densly instrumented volume, the high transparency of the ice and the increased quantum-efficiency, DeepCore lowers the energy threshold of IceCube to $\sim 5\,\mathrm{GeV}$. The arrangement of DOMs on the different IceCube and DeepCore strings is illustrated in Figure 3.2 [83, 78].

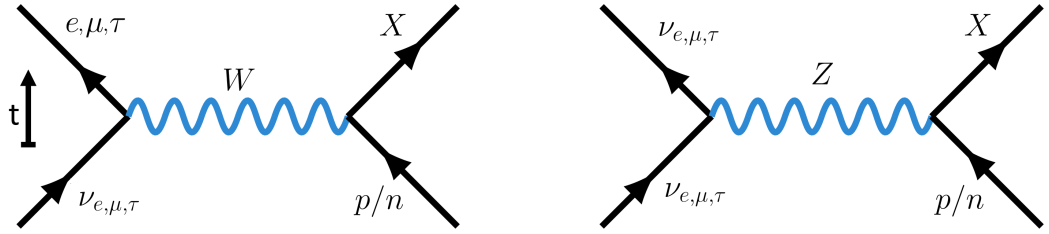
All IceCube and DeepCore strings are connected to the *IceCube Laboratory* (ICL) at the surface. In the ICL, the data from all DOMs is gathered and the trigger conditions are tested. It serves as an online system running at the detector location and pre-processing the data before sending it to the Northern Hemisphere via satellite [78].

Beside these sub-surface arrays, there is a surface detector on top of the ice. It is called *IceTop* and consists of 81 stations of pairwise deployed frozen water-tanks. Each of these water tanks carries two DOMs, shielded from outer light, leading to a total number of 324 DOMs for IceTop. It is capable of measuring Air Showers generated by Cosmic Rays between the knee and the ankle. Additionally, it serves as an atmospheric muon veto for the underlying IceCube detector and allows for calibration of the IceCube detector using coincident muon events in both detectors. However, for low-energy analyses it is of no further relevance and therefore not discussed in this work [84].

3.2 Detection Principle

Within IceCube, neutrinos and anti-neutrinos are detected via their CC and NC interactions with the nucleons of the surrounding ice. These interactions can be deep inelastic (DIS), resonant (RES) or quasi-elastic (QE) scattering, as discussed in Section 2.2. The two corresponding Feynman diagrams are shown in Figure 3.4 for CC and NC.

These interactions generate charged secondary particles. The secondary particles emit



(a) CC (b) NC

Figure 3.4: Feynman diagrams of CC and NC neutrino interactions as seen by IceCube.

Cherenkov radiation, if they move faster than the speed of light in the medium $c_n = \frac{c}{n}$, where $n \approx 1.32$ is the refraction index of the Antarctic ice [85, 80].

The light emission arises from the polarization of ice molecules. While propagating, a charged particle arranges the molecules in its vicinity by its electromagnetic field. As the particle passes, the molecules return to randomized orientations. This randomization process moves with the speed of light through the medium and leads to photon emission by the corresponding molecules. These photons can interfere constructively, if the charged particle moves with a velocity v larger than the speed of light c_n in the medium. In that case, the photons generate a shock-cone coaxial to the movement of the particle and similar to a jet plane passing the sonic barrier. It was first discovered in 1934 by Pavel Alekseyevich Cherenkov, giving name to the resulting radiation. The corresponding process is sketched in Figure 3.5 for a particle with $v > c_n$ [80, 86].

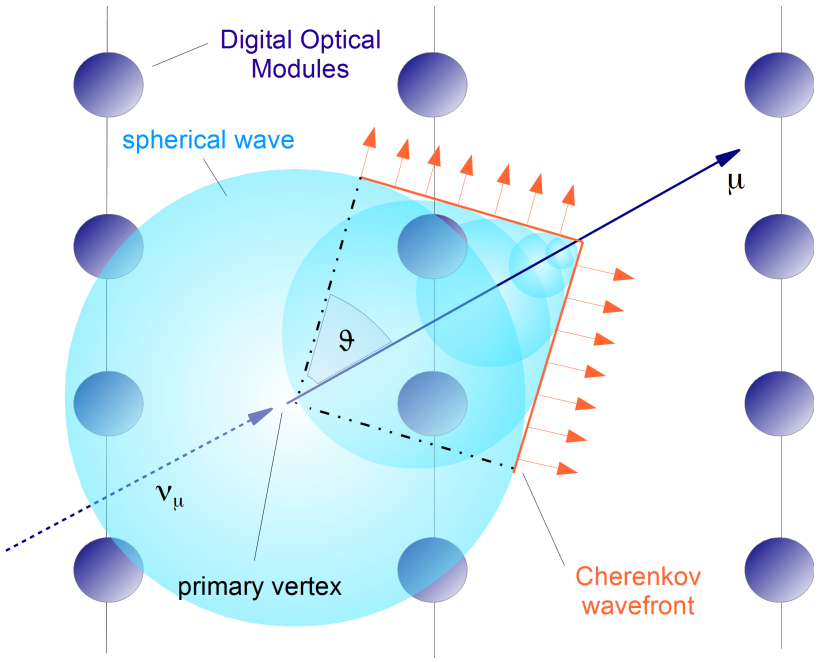


Figure 3.5: Sketch of a muon generated by a CC muon neutrino interaction, emitting Cherenkov radiation; the spherical blue waves centered on the muon track are following the *Huygens-Fresnel Principle* leading to constructive interference along the Cherenkov cone, which is aligned with the muon track (pre-published in [87]).

The opening of the Cherenkov cone is described by the Cherenkov angle ϑ in Equation 3.1. It is derived by the ratio of the velocity of the shock-front $v_s = c \cdot n$ and the velocity of the particle $v_p \approx c$

$$\cos(\vartheta) = \frac{v_s}{v_p} = \frac{1}{\beta n} \approx \frac{1}{n}, \tag{3.1}$$

where the last equality assumes a highly relativistic particle. Since the resulting expression depends only on c and n, it is roughly constant within IceCube at $\cos(\vartheta) \approx 41°$. After propagating through the ice, the resulting Cherenkov photons are eventually measured by the PMTs, carried by the DOMs within IceCube. The number of observed photons and their arrival time can then be used to reconstruct the direction, energy and flavor of the preceding neutrino event, as described in Chapter 5 [80, 86].

Note that a single Cherenkov cone is only obtained for a single particle, moving on a straight line. In case of multiple, charged particles, the resulting light emission is not a pure Cherenkov cone, but rather a diffuse emission, due to the superposition of several Cherenkov cones. It is described in more detail in the following section about the event signatures [86].

The amount of energy emitted by Cherenkov radiation per unit length x travelled by the particle and per unit of frequency ω is given by the so-called *Frank-Tamm formula* [88]. Although most of the light observed by the IceCube detector is generated via Cherenkov radiation, the particle's energy loss due to Cherenkov radiation is small. Instead, charged particles lose most of their energy via ionization, Bremsstrahlung, photo-nuclear effects and pair-production, which are described in more detail in Section 3.3 [88, 11].

3.3 Event Signatures for Different Neutrino Flavors

3.3.1 Track-Like Events

In case of CC muon neutrino interactions, the muon generated at the primary vertex can travel long distances through the detector, generating a distinct Cherenkov cone. Due to the resulting, elongated shape of these events, they are called *track-like* in the following [78].

During the propagation, the muon loses energy by ionization, Bremsstrahlung, photo-nuclear effects and pair-production. The total, energy-dependent, average loss is shown in Figure 3.6. In the vicinity of the minimum at $E_\mu \approx 0.3\,\text{GeV}$, the energy loss is dominated by ionization processes, which are described by the *Bethe-Bloch formula* for heavy, charged particles [11]. Therefore, muons at this energy are called *minimum-ionizing particles* (MIP). Moreover, the energy losses are of similar size between $E_\mu \sim 0.1\,\text{GeV}$ and $\sim 100\,\text{GeV}$ [11].

The muon energy $E_\mu = (1 - y)E_\nu$ can be obtained from the neutrino energy E_ν and the inelasticity y, which follows the distribution in Figure 2.4 in case of DIS interactions. As a result, the muon energy is typically of the same order of magnitude as the neutrino energy. Thus, for CC muon neutrino interactions with $E_\nu \sim 0.1 - 100\,\text{GeV}$ the muon generated at the primary vertex is well-approximated as MIP, which is used for the event reconstruction in Chapter 5.

In general, the energy loss in ice for muons above $\sim 1\,\text{GeV}$ per distance x is described by

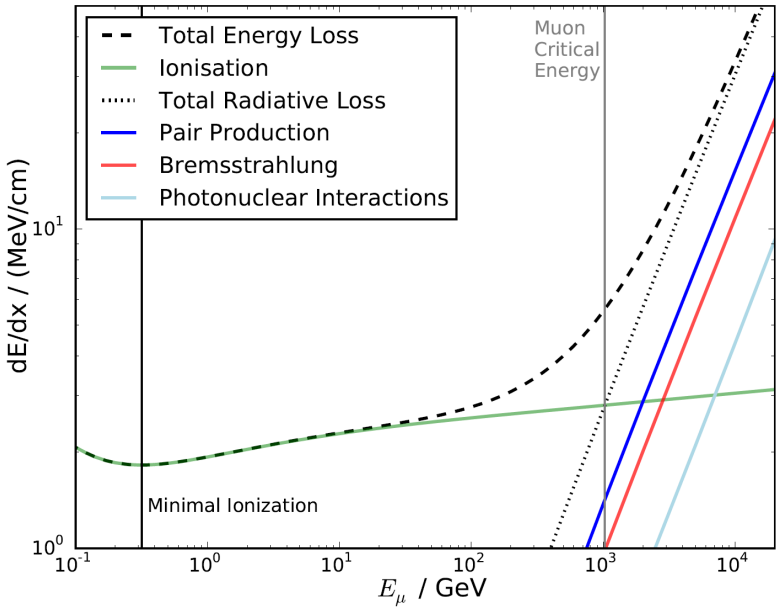


Figure 3.6: Average energy loss per propagated distance x for a muon in ice; close to the minimum, at $E_\mu \sim 0.1 - 100\,\mathrm{GeV}$, the muon is well-approximated as *minimum-ionizing particle* (MIP) (from [89], based on [11]).

Equation 3.2 [90]

$$-\left\langle \frac{\mathrm{d}E_\mu}{\mathrm{d}x} \right\rangle_{\mathrm{ice}} = a + b \cdot E_\mu \approx 0.238\,\frac{\mathrm{GeV}}{\mathrm{m}} + 3.34 \cdot 10^{-4}\,\mathrm{m}^{-1} \cdot E_\mu, \tag{3.2}$$

assuming an ice density of $\rho_{\mathrm{ice}} \approx 0.92\,\frac{\mathrm{g}}{\mathrm{cm}^3}$, which is roughly true for all of the IceCube detector. The energy loss is composed of an almost energy-independent term a from ionization processes and an energy-dependent term $b \cdot E_\mu$ from Bremsstrahlung, photo-nuclear effects and pair production (*radiative losses*), where the coefficients a and b are only weakly energy-dependent [90]. The energy $E_\mu^{\mathrm{crit}} \approx a/b$, at which the radiative loses start dominating the ionization ones, is called *critical energy*. While the term $b \cdot E_\mu$ is small for minimum-ionizing muons, this is not the case for electrons, which lose energy dominantly by Bremsstrahlung and pair production and thus can not be treated as MIPs analogously [90, 11].

Apart from the average loss, the processes feature different stochastic behaviors: For ionization processes, dominating below E_μ^{crit}, the energy loss is almost continuous, while for the radiative processes, dominating above E_μ^{crit}, the energy losses feature strong stochastical fluctuations [11].

In addition to the outgoing muon, CC muon neutrino interactions generate a *hadronic cascade* at the primary vertex, caused by the energy transfer to the nucleon. By the following strong and electromagnetic interactions, multiple charged secondary particles like pions and electrons are produced, which also generate Cherenkov emission but are incapable of travelling large distances like the muon. Due to the large number of particles, the resulting photons do not form a single Cherenkov cone, but cause a rather diffuse, total emission. The resulting, nearly spherical light emission is similar to the one described in Section 3.3.2 and therefore called *starting cascade* [86].

Besides the starting cascade, the muon decay causes an additional electromagnetic cascade at the end of the muon track. However, due to its long lifetime of $\tau_{\mathrm{life}}^{\mu} = 2.2\,\mu\mathrm{s}$, the muon does not decay before losing most of its initial energy, such that the electromagnetic cascade is irrelevant for the presented signature and not discussed here [11].

Thus, in this work *track-like* events comprise the signature of a track and a starting cascade,

if both are contained inside the detector volume [91, 11]. The signature of track-like events in sketched in Figure 3.7.

Note that for $E_\nu < 100\,\text{GeV}$, the length of the muon track is $L_\mu \lesssim 400\,\text{m}$ with most of the tracks being substantially shorter. As a result, upgoing neutrino events are typically contained within the IceCube detector volume, if the primary vertex is inside DeepCore, which is the case for most events used in this work. Note that in case of *contained events* no energy is deposited outside the detector, which allows for a good energy resolution compared to uncontained events, as discussed in Chapter 5. In contrast, *uncontained events* can arise e.g. from CC muon neutrino interactions below DeepCore with the generated muon entering the detector.

3.3.2 Cascade-Like Events

Besides CC muon neutrino events, also CC electron and tau neutrino events are measured in IceCube. For electron neutrinos, the electron generated at the primary vertex is not capable of travelling large distances through the detector like the muon. Instead, it triggers a cascade of secondary particles, which is dominated by electromagnetic interactions, and therefore called *electromagnetic cascade*.

Like hadronic cascades, the electromagnetic cascade features a rather local energy deposition and thus, a light emission of nearly spherical symmetry (cf. Section 3.3.1). However, the electromagnetic cascade is on average brighter, since hadronic cascades deposit more energy in nuclear bindings and neutral particles, like neutrons, that provide no output in Cherenkov light. Moreover, hadronic cascades typically feature electromagnetic sub-showers, that are induced by the generation of π^0 and other neutral mesons. However, the differences between electromagnetic and hadronic cascades are not resolvable on an event-by-event level in DeepCore. Thus, hadronic and electromagnetic cascades are not distinguished in this work [92, 86].

For tau neutrinos, the CC primary interaction generates a tau lepton. For energies $E_\nu \lesssim 100\,\text{GeV}$ the resulting cross-section is reduced compared to the other flavors by kinematic effects, which arise from the large mass of the tau lepton $m_\tau = 1\,777\,\text{MeV}$ [11, 93]. Due to the low energies of $E_\tau \leqslant E_\nu \lesssim 1\,\text{TeV}$ considered in this work, the generated tau lepton decays close to the primary vertex, as the mean distance between generation and decay is only $\ell_\tau \approx (E_\tau/(m_\tau c))\tau^\tau_{\text{life}} = 4.9\,\text{cm}$ for a tau energy of $E_\tau = 1\,\text{TeV}$, using the tau mass m_τ and its lifetime $\tau^\tau_{\text{life}} = 0.29\,\text{ps}$ [11]. This tiny distance is not resolvable with a spacing of tens to hundreds of meters between PMTs. Instead, the tau decays close to its generation vertex, following one of the decay channels in Equations 3.3 to 3.5 [11, 94]:

$$\tau^\pm \to \overset{(-)}{\nu}_\tau + n\,h^{\pm/0} \qquad (\Gamma \approx 64.7\%), \tag{3.3}$$

$$\tau^\pm \to \overset{(-)}{\nu}_\tau + \overset{(-)}{\nu}_\mu + \mu^\pm \quad (\Gamma \approx 17.4\%), \tag{3.4}$$

$$\tau^\pm \to \overset{(-)}{\nu}_\tau + \overset{(-)}{\nu}_e + e^\pm \quad (\Gamma \approx 17.8\%), \tag{3.5}$$

where $n\,h^{\pm/0}$ represents all possible sets of $n \in \mathbb{N}$ mesons of any charge [11]. The low-energy leptons and mesons, generated in the tau decay, can trigger an electromagnetic or hadronic cascade. However, the additional neutrinos lead to invisible energy, escaping from the process. Thus, in contrast to electron and muon neutrino interactions, tau neutrino interactions do not deposit all of the original neutrino energy within the detector, leading

to a systematic shift in the reconstructed energy for tau neutrinos, which is discussed in Section 5.3.4.

Note that the tau decay may also generate a muon, that can potentially cause a track-like signature in DeepCore. However, due to the reduced muon energy and the small branching ratio, these muons are typically not visible at energies $E_\nu < 100\,\text{GeV}$ relevant for tau neutrinos arising from atmospheric neutrino oscillations.

For high-energy analyses, going up to energies of $E_\tau \sim E_\nu \sim 1\,\text{PeV}$, the mean distance between the generation and decay of the tau lepton grows to $\ell \approx 50\,\text{m}$, which can be resolved with current reconstruction methods. This so called *double-bang* signature is targeted by several analyses searching for astrophysical tau neutrino events [94].

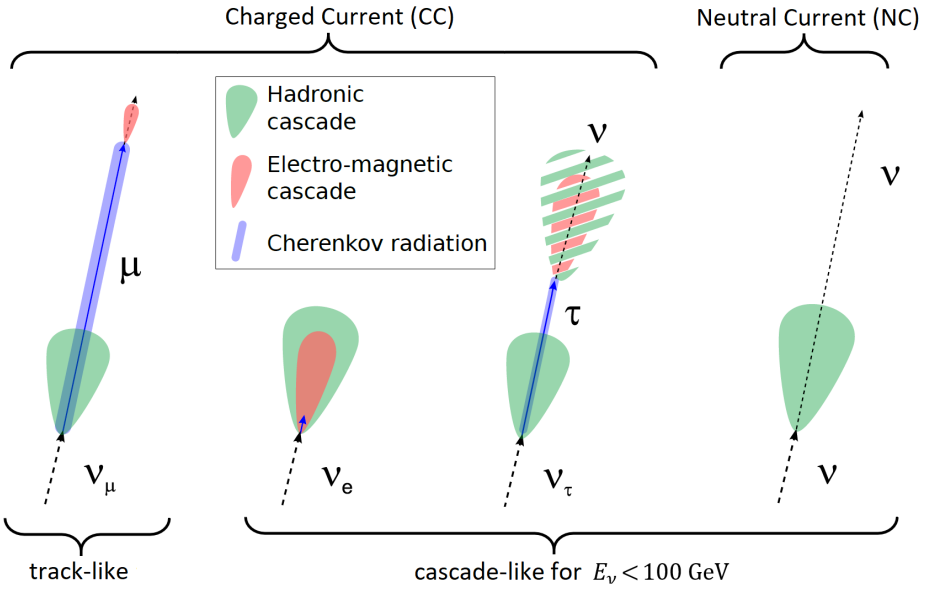


Figure 3.7: Event signatures in IceCube: the first three (from left to right) arise from CC and the last one from NC interactions; track-like events feature a starting cascade, but are distinguishable from the others due to the outgoing muon track; in contrast, the other types are not distinguishable for energies $E_\nu < 100$ GeV and called cascade-like (modified from [95] and [96]).

In addition to the electromagnetic cascade, triggered by the electron and tau lepton, CC electron and tau neutrino interactions generate a hadronic cascade at the primary vertex analogous to the CC muon neutrino interactions. Since these cascades share the same primary vertex and the same, nearly spherical symmetry, they can not be distinguished in IceCube, but appear as a single cascade of increased energy. Therefore, all CC electron and tau neutrino events are considered *cascade-like* in the following [91].

For NC interactions, the energy, transferred to the nucleon by the Z-Boson, leads to a hadronic cascade similar to the ones described above. However, the remaining energy is carried away by the neutrino, that escapes from the detector. Thus, analogous to CC tau neutrino interactions, a large fraction of the neutrino energy stays invisible, leading to a systematic shift in the reconstructed energy, as seen in Section 5.3.4. Since NC interactions cause a cascade without outgoing track, they are also called cascade-like in the following [91].

All event signatures, discussed above, are summarized in Figure 3.7. Additionally, a track-

like and a cascade-like event are shown in Appendix B, using an *event viewer* of the IceCube detector, i.e. visualization software, for illustration purposes.

One should note that the event signature has some notable impact on the energy and direction resolution. Since tracks provide a long lever arm for the reconstruction, their angular resolution is typically better than the one for cascade-like events. Moreover, above $\sim$ 1TeV the energy resolution for cascades is notably better than for tracks, where the muon can leave the detector, causing energy being deposited outside the instrumented volume. However, below $\sim$ 1TeV the energy resolution of tracks and cascades is similar (cf. Section 5.3.4).

3.4 The Deep Ice within IceCube

The Antarcic ice is the active medium used for the neutrino detection in IceCube. Moreover, its properties affect the propagation of Cherenkov photons by absorption and scattering. Thus, precise measurements of atmospheric neutrinos require a profound understanding of the ice.

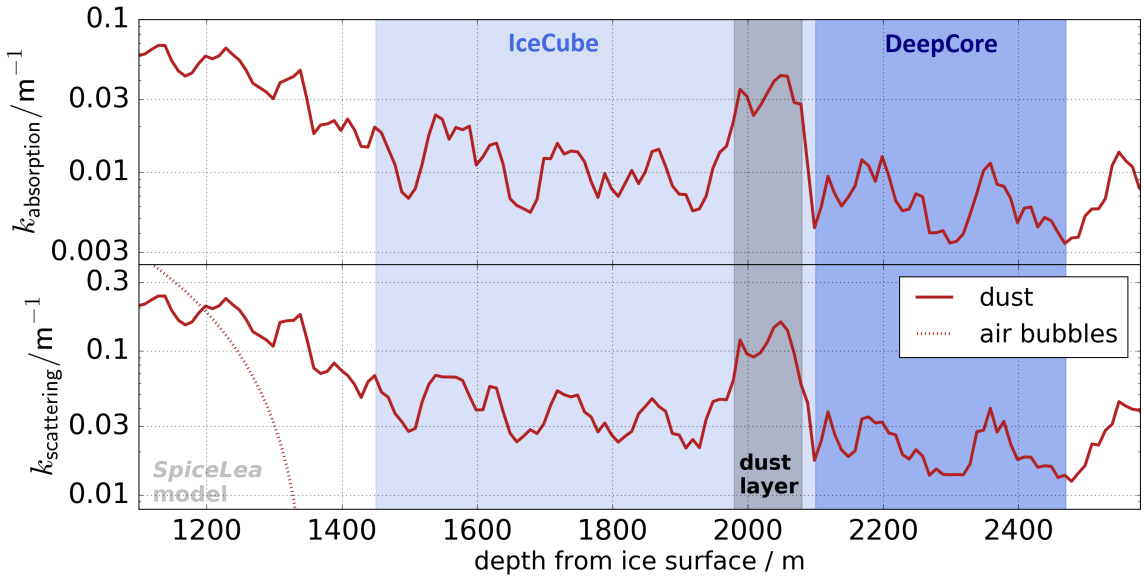


Figure 3.8: Optical properties of the *bulk-ice* in terms of the depth-dependent, effective scattering and absorption coefficients at a wavelength of 400 nm; DeepCore and IceCube detector are indicated by blue shaded areas; at shallow depth, a notable contribution from *air bubbles* exists, which vanishes at larger depth, while *dust* comprises all other impurities.

The Antarctic ice at the South Pole has grown over tens of thousands of years by compressing fallen snow into dense ice through gravitational pressure. During this process, different amounts of impurities, called *dust*, were frozen into these ice-layers, where the amount and properties of the frozen dust depend on the global weather conditions in the Earth's atmosphere for each epoch. As a result, the properties of the ice vary mostly depending on its depth. The depth-dependent, effective scattering and absorption coefficients of the ice are shown in Figure 3.8 for a typical Cherenkov wavelength of 400 nm [97, 78].

Note that above $\sim$ 1.3 km, the ice becomes substantially less transparent than deep in the detector. One reason for this change are air bubbles frozen into the Antarctic ice, which are

a major source of photon scattering. However, below a depth of $\sim 1.3\,\text{km}$ the high pressure causes the air contamination to be incorporated into the ice crystal structure over time, featuring the same refractive index as ice [98]. As a result, the scattering contribution from air bubbles vanishes, while scattering and absorption remain present due to other impurities of the ice, as shown in Figure 3.8. Note that the DeepCore detector was deployed in a region of very clear ice, while a pronounced *dust-layer* is located right above its fiducial volume. Thus, the absorption and scattering coefficients are small within the fiducial volume, while they are large in the dust-layer above [97].

The ice in the detector is limited from below by the bedrock, starting $\sim 300\,\text{m}$ below DeepCore. It is opaque for photons, but features ~ 3 times higher neutrino-interaction rates, due to the increased density of rock compared to ice. Since muons, produced by CC muon neutrino interactions with $E_\nu \gtrsim 100\,\text{GeV}$ in the bedrock, are capable to reach and trigger the DeepCore detector, the bedrock needs to be simulated accurately for neutrino analyses, which is discussed in more detail in Section 4.2.1 [97].

Vertically, the ice properties change rapidly, due to the layered structure of the Antarctic ice. However, the depth at which certain ice features become visible varies slightly over the horizontal extent of the detector, leading to the conclusion that a small *tilt* in the ice layers is observed, caused by the landscape of the underlying bedrock. This landscape is only washed-out slowly, when going to more shallow layers, but is still very visible within the IceCube detector. Thus, the same ice properties are found at slightly different depths for different strings [97].

Horizontally, the ice properties are affected by glacial movements. Since the ice at the South Pole is a moving glaciar, the horizontal isotropy is broken by the direction of movement of the glaciar. This movement induces an anisoptropy in the absorption and scattering properties of the Antarctic ice. The underlying processes, that cause these changes, are still not understood, but object of major, ongoing investigations [97, 99].

This Antartic ice, that has grown naturally without any human interference, is called *bulk-ice* in the following.

While the bulk-ice makes up the majority of the ice in the detector, the DOMs are located in a different type of ice, called *hole-ice*. The hole-ice is the refrozen ice within the vertical holes, that were drilled while deploying the strings. Although the melted water was re-used to fill the drill holes, the resulting ice behaves differently than the bulk-ice. This change in the ice properties is not yet understood completely, but is likely related to the relatively quick re-freeze of the ice, which was not generated over thousands of years under gravitational pressure. As a result, the air contained in the hole water could not escape during the freezing process. Instead, it was pushed towards the center of the hole-ice, forming the so-called *bubble-column*, which is confirmed by in-ice camera observations [100, 101]. Additionally, the vertical layers found in the bulk-ice were washed-out by the melting and re-freezing process. As a result, the hole-ice differs in scattering and absorption compared to the bulk-ice, shown in Figure 3.8 [97].

The hole-ice properties are a major systematic uncertainty for the DeepCore detector. Its models differ mainly by their assumption on the radial-dependence of the optical hole-ice properties: For example, a non-uniform radial-dependence of the optical properties is motivated by the observed bubble-column, which features a much smaller radius than the drill holes. Thus, the outer parts of the drill hole might have similar optical properties to the bulk-ice (or better), while deviations from the bulk-ice are mostly contained in the

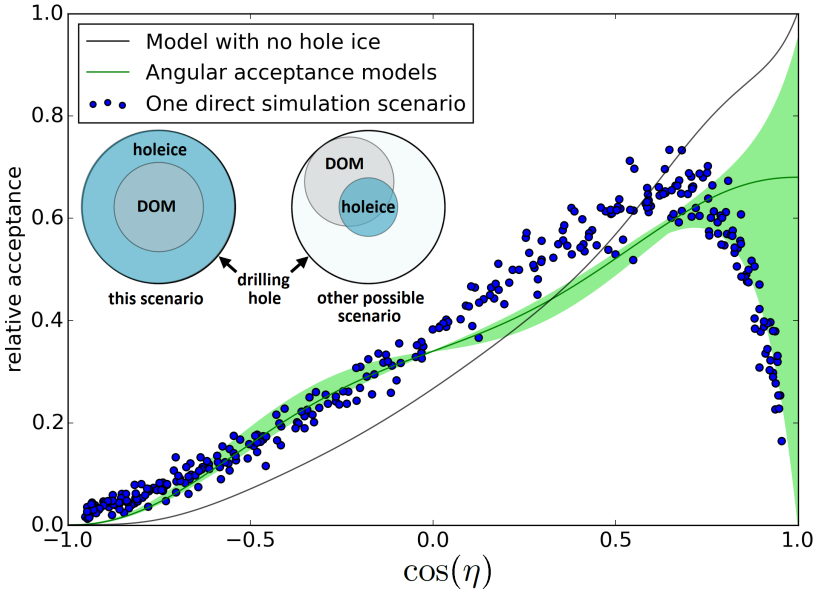


Figure 3.9: Parametrization of hole-ice uncertainties by an effective photon detection efficiency (*relative acceptance*), depending on the inclination angle $\cos(\eta)$ of the incoming photon: the blue dots show the simulation of a homogeneous hole-ice with all DOMs centered in their drill hole; the green shaded area shows the range covered by the systematic parametrization (modified from [81] and [101]).

bubble-column in the center. In contrast, the hole-ice properties might also be homogeneous throughout the entire hole, such that the DOMs are completely contained within a homogeneous ice column. Moreover, the horizontal location of each DOM within its drill hole might vary, as well as its orientation with respect to the cable, that shadows photons arriving from a certain azimuth angle. In general, all these effects are highly correlated. The resulting models are mostly constrained by *in-situ* calibration measurements with *LED flashers* attached to the DOMs, which are described in Section 3.5.1 [100, 81].

For physics analyses, the holeice uncertainties are parametrized by an effective photon detection efficiency, that depends on the inclination-angle η of the incoming photon. It is derived from propagating photons through pure bulk-ice and comparing the resulting number of detected photons for all DOMs to LED flasher measurements. The η-dependent acceptance is then added to the bulk-ice simulation to optimize the agreement with flasher data and parametrized by two systematic parameters (cf. Section 6.3.5). The resulting photon acceptance is shown in Figure 3.9, where the parametrization covers the green shaded area. Note that the parametrization is least constraining at $\cos(\eta) \approx 1$, which is for upwards moving photons, where the LED measurements provide almost no constraints. Therefore, the parametrization is free to vary the acceptance over a large range for these photons.

In addition, the hole-ice can be simulated directly by using different optical properties within the drill hole for the photon simulation. For comparison, one of these simulations is shown as blue dots in Figure 3.9 for the case of homogeneous hole-ice properties and all DOMs centered within the drill holes [97, 81].

The nuisance parameters used in this work for the hole-ice uncertainties, are described in more detail in Section 6.3.5.

3.5 The IceCube Data Acquisition and Processing

3.5.1 The IceCube DOM

The photon detection in IceCube is performed by the *Digital Optical Module* (DOM), which can be seen as the fundamental detection unit of the detector. As sketched in Figure 3.3, it consists of a 10" photomultiplier tube (PMT) and digitizing electronics, placed in a glass pressure sphere. Within the sphere, the PMT is arranged looking downwards, increasing the detection efficiency for photons moving upwards [78].

The PMT is sensitive to photons between 300 nm and 600 nm wavelength with a peak detection efficiency of 25 % for IceCube and 33 % for DeepCore DOMs at 390 nm. The PMT amplification is $\sim 10^7$, generating signals from ~ 8 mV to the linearity limit at ~ 2 V. However, the low-energy events discussed in this work do typically not contain signals outside the linearity regime [78, 82].

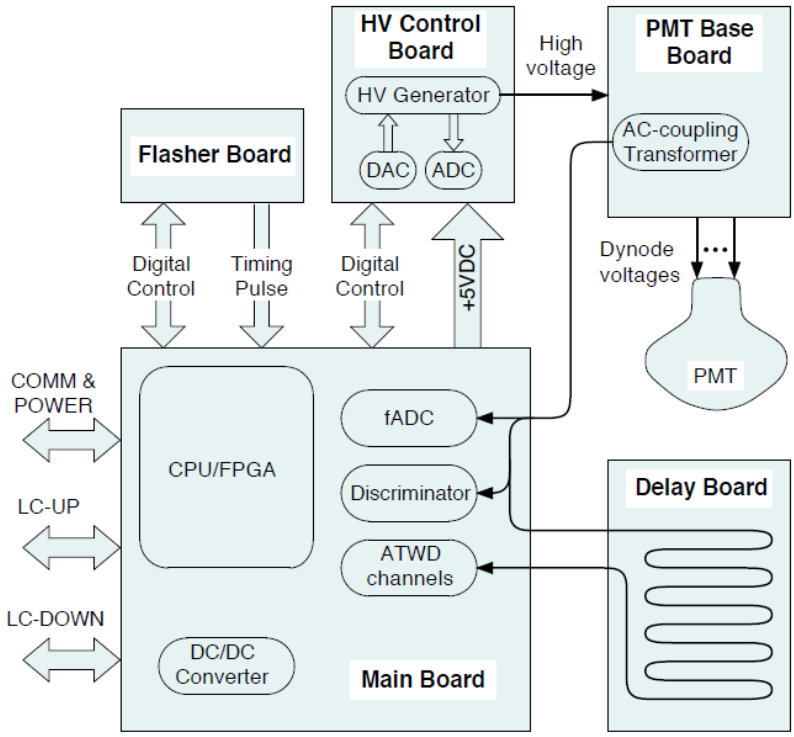


Figure 3.10: Raw-data processing within a DOM: the PMT response is passed to an FADC and two ATWD digitizers and tested for local coincidences (LC) up- and downstream the string; an additional high-voltage supply and flasher LEDs are deployed in each DOM, where the latter is used for calibration purposes (from [78]).

The PMT is attached to the DOM mainboard, as sketched in Figure 3.10. The voltage curve, measured by the PMT, when a photon hits the photocathode, is called *waveform*. On the mainboard, these waveforms are digitized by two kinds of digitizers in parallel: first, two *Analog Transient Waveform Digitizer* (ATWD) on the first 422 ns of each waveform and second, a *Fast Analog-To-Digitcal Converter* (FADC) for any longer signal. Two ATWD chips are implemented per mainboard to allow a waveform to be digitized by one ATWD, while the other one is being read out, which reduces the dead-time of the digitization. The

typical width of a digitized waveform ranges from 12 ns to $\sim$ 1.4 μs [78]. The system is capable of resolving up to 300 photoelectrons in 25 ns, being limited only by the linearity regime of the PMT [97].

The integral over these waveforms is interpreted as observed charge of the PMT in units of photoelectrons (PE), giving an estimate for the number of observed photons on the photocathode. Instead of calculating the integral, it is approximated by the amplitude of the waveform, where $\sim$ 8 mV amplitude correspond to 1 PE. The waveform is read out, if the estimated charge exceeds 0.25 PE, in which case the corresponding DOM is labelled as *hit*. Furthermore, the time of each waveform is stored and used as proxy of the time of arrival for the corresponding Cherenkov photons. These quantities are later used to reconstruct the neutrino events in Chapter 5 [78].

In addition to the digitizing electronics, each DOM is equipped with 12 LED flashers on a flasher board, which are pointing into different directions into the surrounding ice. These LEDs are capable of producing pulsed light of up to 70 ns pulse-length and up to 10^{10} photons for calibration purposes. Here, calibration includes measurements of the absorption and scattering properties of the ice (cf. Section 3.4), the DOM positions and the linearity of the DOM response in photon intensity. The wavelength $\lambda_{LED} = 405$ nm emitted by the LED flashers was chosen to be close to the typical wavelength of Cherenkov photons, while 16 DOMs were equipped with flashers of different wavelength. The knowledge and uncertainties of the ice properties, included as systematic parameters in Section 6.3.5, are largely based on such LED measurements (cf. Section 3.4) [97].

The noise rates for IceCube and DeepCore DOMs are $\sim$ 500 Hz and $\sim$ 650 Hz, respectively, where the precise value depends on the individual DOM. To reduce this random noise, each hit is tested for local coincidences (LC) with its nearest or next-to-nearest neighbor on the same string within 1 μs. In case such coincidence is found, ATWD and FADC read-outs are sent to the surface, which is called *Hard Local Coincidence* (HLC). Otherwise, only the FADC read-out is sent to the surface, which is called *Soft Local Coincidence* (SLC) [78].

At the surface, this data is used in the trigger criteria of the detector, which are described in more detail in Section 3.5.2.

3.5.2 Online Processing at the Pole

3.5.2.1 Triggering

If a *Hard Local Coincidence* (HLC) is detected on a string, the data is sent to the *IceCube Laboratory* (ICL) at the surface. At the ICL, the HLC data from all strings is tested to match one of the trigger conditions. Most triggers used in IceCube are *Simple Multiplicity Triggers* (SMT). They require a certain number of HLCs in a given time-window to trigger the full read-out of the detector. In case of the IceCube array, the most relevant trigger is the SMT-8 trigger, requiring 8 HLCs being seen within 5 μs in IceCube DOMs. It triggers at a frequency of 2.5 kHz. Data used in IceCube analyses is mostly based on this trigger, while other triggers exist for specific physics cases [78].

In contrast, the DeepCore sub-detector used in this work is based on an SMT-3 trigger. It requires 3 HLCs being seen within 2.5 μs among the lower 50 DeepCore DOMs on each string (without veto cap) and the lower 22 DOMs of the 7 adjacent IceCube strings. This leads to a trigger frequency of 250 Hz. In contrast to the IceCube array, it is the only trigger

run on this sub-detector. For the DeepCore trigger, the corresponding read-out window is $[-4\,\mu s, +6\,\mu s]$, which is centered at the HLC hit triggering the detector read-out [83, 102].

It is common, that a single event matches multiple triggers. To handle this, the ICL checks for all individual triggers first, before running a full detector read-out. Different triggers being launched within $1\,\mu s$ are combined in one global trigger. There, the read-out window is adjusted such that the read-out windows of all individual triggers are contained. The resulting event is then flagged with all triggers that were launched [78].

At trigger level, the data is clearly dominated by atmospheric muons, which feature $\sim 10^6$ times higher rates than the atmospheric neutrino contribution. As a result, subsequent filtering and processing is needed to purify neutrino samples [78, 83].

3.5.2.2 Processing and Compressing Data

In the ICL, *Processing and Filtering* (PnF) scripts are run on the triggered data. This processing includes [78]:

1. extraction of data for detector and quality monitoring,
2. generation of realtime alerts for events of astrophysical interest,
3. calibration of the raw DOM response data,
4. generation of data files and meta-information for the data interpretation.

Within this process, the waveform data is compressed into the *Super Data Storage and Tranfer* format (SuperDST) to be sent to the Northern Hemisphere. This format is known to lead to small information losses due to discretization, while the resulting loss is small compared to the calibration uncertainties. However, the format leads to a reduction of the data size by a factor of ~ 10. In addition to the DOM reponse data, the ICL adds information on the detector *Geometry, Calibration and Detector Status* (GCD) to each run of the detector, which is usually every 8 h. This information is later used to identify problematic DOMs in reconstruction and data selection and to interpret and re-calibrate the DOM response.

During the calibration (point 3), the best knowledge of the DOM response is used to deconvolve the digitized waveform into single photon responses. In this process, the photon arrival time information is extracted and stored for each DOM in each event as so-called *pulses* or *hits*. In addition to the photon arrival times, the amplitude/charge information is stored, since the photon extraction from the continuous waveform is not perfect and the amplitude/charge information allows to estimate the number of coincident photons that might have ended up in the same extracted pulse. Thus, the integrated charge over all pulses in a single DOM within a certain time frame is a more robust estimator for the observed number of photons than the number of pulses itself.

The entirety of all pulses in one event is called *pulse-map*.

3.5.2.3 Hit Cleaning

For each event, the pulse-map is run through a cleaning algorithm, called *hit cleaning*, which is supposed to remove unphysical hits from the corresponding pulse-map. Since DeepCore events are triggered by HLC hits, but also include a large number of SLC hits in the read-out

data, they include a large number of random noise hits not being related to any physical light observed by the corresponding DOM [78].

These hits are removed by the so-called *SeededRT cleaning*. The SeededRT cleaning starts with a sub-set of pulses in the pulse-map. These are typically chosen to be the HLC hits, as they are assumed to be of physical origin. Then, each pulse in the sub-set is tested for other pulses in its spatial and time vicinity, in the same or other DOMs. If such pulses are found, they are added to the sub-set. After all original pulses were tested, the process is repeated iteratively with the new, extended set. The algorithm converges, if no more pulses are added to the sub-set during one interation. Thus, the sub-set contains only pulses, that are close to each other in space and time.

To define the vicinity of a pulse, a radius of 250 m and a time difference of 1 μs is used for pulses in the IceCube array. For DeepCore, these values are adjusted to 125 m and 500 ns. Due to the fact that pulses can not be removed from the seeded pulse-map, the final pulse-map is a super-set of all HLC hits.

The cleaned map is afterwards used for the DeepCore filter, described in the following.

3.5.2.4 Filtering

In the ICL, every event is run through a set of filter criteria using observables from simple low-level reconstructions like the vertex position, direction, energy or topology. Depending on these filter criteria, the event is classified as matching or not matching, while events can match several filters at the same time. These filter decisions split the events into different data streams, targeting different physics analyses [78].

The data sample used in this work is based on the DeepCore filter. The DeepCore filter reduces the amount of data by a factor of $\sim$ 100 by removing obvious background events from triggered dark noise and atmospheric muons [83].

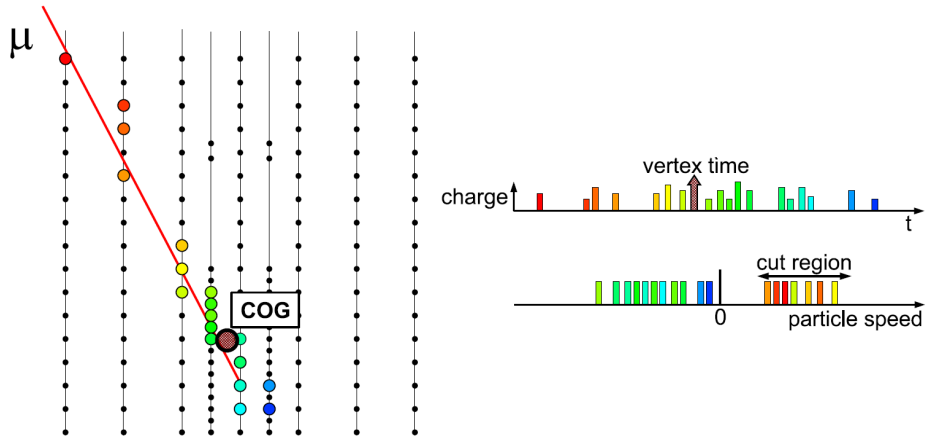


Figure 3.11: Schematic view of the DeepCore filter: pulses passing the hit cleaning are split into fiducial and veto hits; while the fiducial hits are used to calculate a time and space COG within DeepCore, the veto hits are used to calculate the hypothetical velocity of a particle to this COG coordinates; in case one of these velocities is consistent with an atmospheric muon hypothesis, the event is rejected (modified from [83]).

To do this, all hits that remain in the SeededRT cleaned pulse-map (s. Section 3.5.2.3)

are split into fiducial hits and veto hits. The fiducial hits are all hits that are within the DeepCore fiducial volume, excluding the DeepCore veto cap, while all remaining hits are defined as veto region.

Afterwards, the following algorithm is applied, as visualized in Figure 3.11 [83]:

1. The *center of gravity* (COG) for all pulses i within the fiducial volume is calculated. Since these are not weighted by their observed charge, this is equivalent to calculating the mean value $\langle \vec{x} \rangle$ of the position vector $\vec{x}_j$ for the corresponding DOMs $j(i)$ and the mean $\langle t \rangle$ of their corresponding hit times t_i.

2. The mean value $\langle \vec{x} \rangle$ is refined by taking only hits that are within one standard deviation of the mean time $\langle t \rangle$.

3. The mean value $\langle t \rangle$ is refined by taking the average of the causally corrected hit times for all hits in the fiducial volume. This causal correction is defined by subtracting the time that unscattered light would need to travel from the refined COG position $\langle \vec{x} \rangle$ to the corresponding DOM from each hit time.

4. The refined COG position and time, calculated in step 2 and 3 from the fiducial hits, are then used to calculate a velocity for each pulse in the veto region. Therefore, a hypothetical particle is assumed to travel from the corresponding hit to the refined COG. For a pulse i in the DOM $j(i)$ of the veto region, the resulting velocity is

$$v_i = \frac{\left| \vec{x}_{j(i)} - \langle \vec{x} \rangle \right|}{\langle t \rangle - t_i}, \tag{3.6}$$

   where t_i is the time of the pulse i and $\vec{x}_{j(i)}$ is the position of the DOM j observing pulse i. The velocity is positive in case the hit happens before the COG average time and negative otherwise.

5. In case of atmospheric muons causing the veto hits and the hits in the fiducial volume, the resulting velocities v_i are expected to be close to the speed of light and positive. Thus, a cut is applied, such that an event is rejected, if the velocity v_i for one pulse i in the veto region satisfies:

$$\frac{5}{6}c \leqslant v_i \leqslant \frac{8}{6}c, \tag{3.7}$$

   as it is likely caused by an atmospheric muon.

The DeepCore filter is a simple topological criterion to reduce the amount of data sent via satellite. Although simple, it allows to remove the vast majority of triggered atmospheric muons, while keeping almost all atmospheric neutrino events.

The remaining DeepCore data is sufficiently small to be sent via satellite, although the vast majority of the remaining data is still induced by atmospheric muons. These remaining muons are later removed by a carefully chosen set of selection criteria, that define the final data sample. Since this is an analysis-specific selection and not done at the pole, the further processing is described separately in Chapter 4.

3.6 The Precision IceCube Next Generation Upgrade (PINGU)

The *Precision IceCube Next Generation Upgrade* is a proposed, low-energy extension of the currently data-taking DeepCore detector. It aims for several physics goals, while many of them extend already existing DeepCore analyses with increased sensitivity [7].

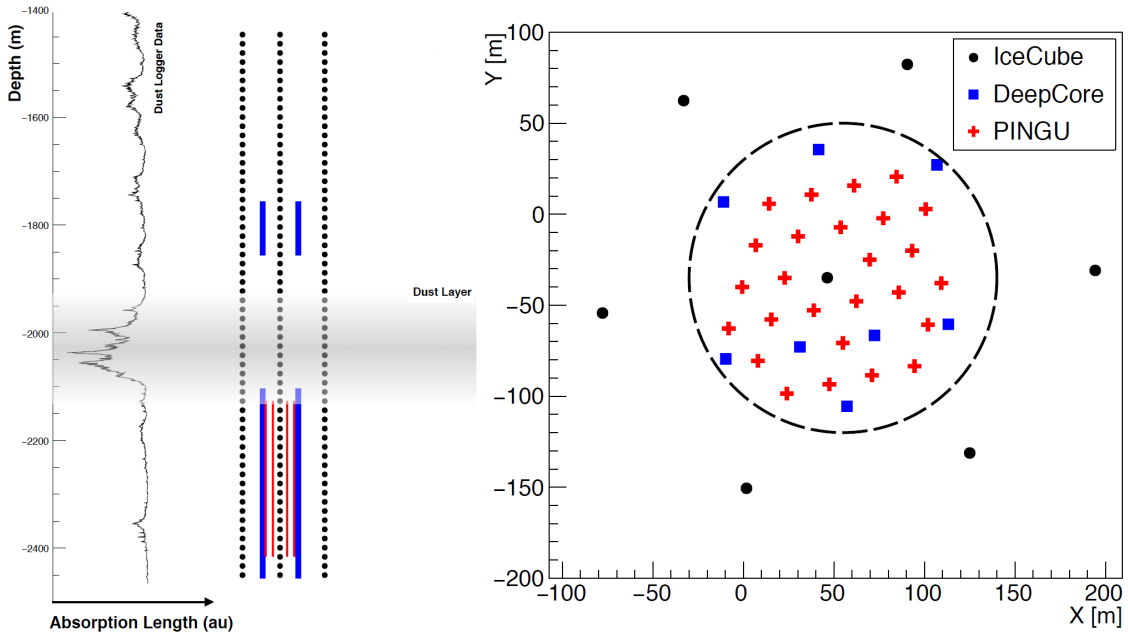


Figure 3.12: Proposed baseline detector set-up for the PINGU detector, shown as side view (left) and on-top view (right) (modified from [7]).

Placed inside the DeepCore volume, the currently proposed baseline scenario for the PINGU detector consists of 26 additional strings within the DeepCore volume. Each of these strings carries 192 optical modules, leading to a horizontal, inter-string spacing of 24 m and a vertical, inter-module spacing of 1.5 m, which is a much denser instrumentation than used for the DeepCore volume. For the optical modules, several options are currently being discussed: They range from an updated version of the currently used, high-quantum-efficiency DeepCore DOMs up to more advanced, possibly multi-PMT modules. In addition to improving the light detection, several calibration devices are planned to be used during the deployment of the PINGU strings (while the drill holes are open) and during the data-taking of the PINGU detector. This will lead to an improved understanding of the ice properties and allow to reduce systematic errors on the measurements. The PINGU detector set-up from the most recent *Letter of Interest* of the PINGU Collaboration is shown in Figure 3.12 [7].

Due to the higher instrumentation density and at least small modifications to the optical modules, the PINGU detector will on average detect more Cherenkov light than the currently running DeepCore detector for low-energy events. The additional Cherenkov light will improve the noise suppression and the reconstruction of events and thus reduce the uncertainties on the extracted physics parameters.

By lowering the energy threshold to $E_\nu \sim 1\,\text{GeV}$ and improving the resolution of the event reconstruction, PINGU will be capable to produce competitive measurements of the oscillation parameters $\sin^2(\theta_{23})$ and the atmospheric squared-mass difference Δm^2_{31}. It will also be capable to detect tau neutrino appearance (cf. Section 2.3.2) with high significance and measure the corresponding tau neutrino normalization with an excellent accuracy. Furthermore, it aims for a $> 3\,\sigma$ measurement of the NMO with less than fours years of data-taking. The corresponding $\sin^2(\theta_{23})$-dependent sensitivity is shown in Figure 3.13 for both ordering hypotheses and four years of data [63, 7].

Although measuring the NMO is a highly competitive field, the PINGU detector is a rela-

tively cost-efficient ($\sim$ 50 M\$) detector that could come online on a short timescale. Thus, it might be taking data before other, highly sensitive experiments like DUNE or Hyper-Kamiokande (cf. Section 2.5). However, the time-window for such early measurement is closing [7].

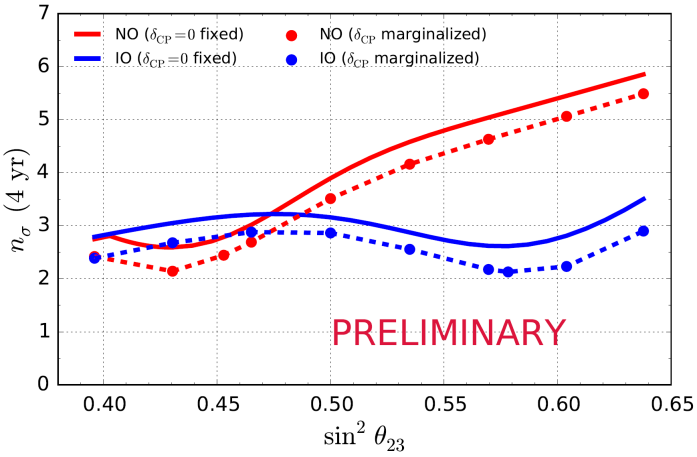


Figure 3.13: Projected sensitivity of the PINGU detector to the NMO for four years of data, depending on the atmospheric oscillation parameter $\sin^2(\theta_{23})$ (from [7]).

While PINGU is still in its planning stage, its predecessor DeepCore has been taking data for several years now. Thus, the already available DeepCore data can be used to prototype such NMO analysis for the future PINGU detector.

This is done in this work by identifying potential future challenges and testing the full analysis chain for such measurement. Since the DeepCore and PINGU detector set-ups are similar, except for their spacing of optical modules, they suffer from similar systematic uncertainties:

1. uncertainties in the (atmospheric) neutrino oscillation parameters (cf Section 2.3),
2. uncertainties in the atmospheric neutrino flux (cf Section 2.4),
3. uncertainties in the neutrino-nucleon interactions (cf Section 2.2),
4. uncertainties in the detector response of the optical modules (cf Section 3.5),
5. uncertainties due to unknown, local, optical properties of the ice (cf Section 3.4).

In Section 6.3, these uncertainties are discussed in detail for this work and the DeepCore detector [7]. However, the impact of some of these uncertainties varies between DeepCore and PINGU due to the differences between the detectors, e.g. the energy range that is used for the measurement.

Beside neutrino oscillations, PINGU is also aiming for other fields of physics, like dark-matter detection and Earth tomography, that are not described here [7, 63].

4 The GRECO Data Sample

4.1 General Purpose and Strategy of the Sample

The *GeV Reconstructed Events with Containment for Oscillations* (GRECO) sample is an extreme low-energy data sample based on the DeepCore trigger and filter. It aims for high statistics at low energies close to the DeepCore threshold at $E_\nu \sim 5\,\text{GeV}$.

At these energies, one of the challenges is the weak experimental signature from neutrino events. The majority of DeepCore events features pulses in only $\mathcal{O}(10)$ DOMs, which are mostly single photons. Due to this limited signature, reconstructions at these energies provide only poor resolutions for the zenith-angle, neutrino energy and track against cascade separation. Still, DeepCore is capable to measure atmospheric oscillations competitively, due to the high statistics of $\mathcal{O}(10^4)$ events per year on analysis level.

Table 4.1: Monte Carlo (MC) generators used for low energies (LE) and high energies (HE) in GRECO, optimized for the most accurate description in their energy regime; the last column gives an estimate for the fraction of the final sample, while the precise values depend on several nuisance parameters (cf. Section 6.3 and Table 4.4).

type	contribution	LE generator	HE generator	oscillations	fraction
neutrinos (signal)	CC ν_μ	GENIE	NuGen	yes	$\sim 60\%$
	CC ν_e	GENIE	NuGen	yes	$\sim 25\%$
	CC ν_τ	GENIE	-	yes	$\sim 5\%$
	NC $\nu_{e,\mu,\tau}$	GENIE	NuGen	no	$\sim 5\%$
background	atmospheric μ	MuonGun+CORSIKA		no	$\sim 5\%$
	noise	Vuvuzela		no	$\sim 0.1\%$

Another challenge at these energies is the separation and modelling of different components of the incoming flux. While high-energy samples in IceCube reach purities of 99.7% in CC muon neutrino events [103], low-energy samples typically comprise non-negligible contributions from:

1. CC muon neutrino events
2. CC electron neutrino events
3. CC tau neutrino events
4. NC neutrino events of all flavors
5. atmospheric muon events
6. triggered dark noise.

For measuring the NMO, all these contributions need to be modelled accurately in *Monte Carlo* simulations (MC). The MC generators used for this are summarized in Table 4.1, while the MC simulations are described in more detail in the following section before introducing the data selection in Section 4.4.

In the data selection, experimental data and Monte Carlo are run through the same processing, such that no difference is made between them. The *final sample*, i.e. the sample after the data selection, is then used for a likelihood analysis of the NMO in Chapter 6.

4.2 Monte Carlo Generators

The Monte Carlo generators are summarized in Table 4.1. They are split into high- and low-energy generators to use the most reliable generator at every energy. In the following, all neutrino events are classified as *signal*, while *background* comprises atmospheric muons and triggered dark noise. Moreover, NC events are classified as non-oscillating, since the NC interaction is independent of the neutrino flavor.

4.2.1 Neutrino Simulations

For all neutrino flavors, *GENIE 2.8.6* [104] is used to simulate events between 1 GeV and 1 TeV. It includes simulations of QE, RES and DIS interaction. Within the simulation, neutrinos are forced to interact within the detector volume to avoid a time-intense simulation of neutrino events that do not interact, due to the small neutrino-nucleon cross-section. At the primary vertex, the simulation of secondary particles is done by the *PYTHIA* framework [105], which is closely linked to the GENIE framework.

To further reduce the computation time, neutrino events are generated according to power-laws with the spectral index depending on the energy range. On analysis level, the events are then re-weighted according to the expected flux from atmospheric neutrinos and their interaction probability. The energy ranges and spectral indices used for the simulation are summarized in Table 4.2. Overall, the simulated statistics correspond to an effective livetime of ~ 30 years, which is roughly an order of magnitude above the experimental sample size.

Each data sample in Table 4.2 features a unique *sample ID* to distinguish these *baseline samples* from the *systematic samples*, introduced in Section 6.3.5. While the baseline samples in Table 4.2 were generated with the default settings of all detector uncertainties, the systematic samples were generated with varying settings for these systematic parameters to estimate their impact. The systematic samples are described in more detail in Section 6.3.5.

The sample ID consists of five digits, where the first two digits describe the simulated neutrino flavor and the remaining three digits identify the detector simulation (640 corresponds to the baseline settings). A detailed list of all baseline and systematic samples is added to Appendix H.1.

For the GENIE simulation, the generation volume, i.e. the volume in which the neutrino must interact, is shrunk to a cylinder of length ℓ and radius r, which is aligned with the neutrino direction. The value of ℓ and r depend on the type of simulated event and the corresponding energy scale, as summarized in Table 4.2. This is done, since low-energy events within DeepCore feature a very local signature. Thus, neutrino interactions, that cause light deposition more that ~ 250 m away, are rare. While cascades are incapable to

Table 4.2: Energy ranges, spectral indices and detector volume used in the neutrino simulation with GENIE; the generation volume is given by a cylinder of radius r and length ℓ; the last column gives the simulated effective livetime of the data sample.

sample ID	type	E_ν/GeV	spec. index	r/m	ℓ/m	effective livetime / y
12640	$\nu_e + \bar{\nu}_e$	1-4	1.9	250	500	~ 30
	$\nu_e + \bar{\nu}_e$	4-12	2.0	250	500	~ 30
	$\nu_e + \bar{\nu}_e$	12-1000	2.5	250	500	~ 30
14640	$\nu_\mu + \bar{\nu}_\mu$	1-5	2.0	250	500	~ 30
	$\nu_\mu + \bar{\nu}_\mu$	5-80	2.0	330	900	~ 30
	$\nu_\mu + \bar{\nu}_\mu$	80-1000	2.0	330	1500	~ 30
16640	$\nu_\tau + \bar{\nu}_\tau$	4-10	1.5	250	500	~ 30
	$\nu_\tau + \bar{\nu}_\tau$	10-30	2.0	250	500	~ 30
	$\nu_\tau + \bar{\nu}_\tau$	30-1000	3.5	250	1000	~ 30

generate such light depositions, CC muon neutrino and atmospheric muon events can cause such signatures, which is the reason for the elongated volume used for the generation of CC muon neutrinos.

In addition, these far travelling muons cause issues for upgoing events, since the bedrock is not implemented in GENIE. Instead, the ice extends down into the bedrock region. As a result, the cross-section of the neutrinos and the energy loss of the secondary particles is not modelled correctly for these events. However, as muons from that depth with $E_\mu \lesssim 100\,\text{GeV}$ do typically not lead to notable light depositions in DeepCore, this effect is only relevant for CC muon neutrinos of high energy.

To account for these inadequacies of GENIE, a second generator, called *NuGen*, is patched in for these critical parts of the parameter spaces (cf. Table 4.1). NuGen is based on the *ANIS* simulation framework [106] and is the standard neutrino generator above $E_\nu \sim 100\,\text{GeV}$ within the IceCube Collaboration. It has been found to be reliable over several orders of magnitude in neutrino energy up to $E_\nu \sim 1\,\text{PeV}$. Moreover, it is not limited by its generation volume and includes an accurate description of the bedrock below the detector. However, it includes only DIS interactions, leading to a growing deficit in the neutrino rates below $\sim 100\,\text{GeV}$, due to the lack of QE and RES events (cf. Figure 2.2).

NuGen simulations are used to extend the GENIE sets in the following analysis. To do this, NuGen sets are generated for the same baseline settings as for GENIE. Then, all GENIE events are replaced by NuGen events if they feature one of the following conditions:

$$E_\nu \geqslant 500\,\text{GeV} \tag{4.1}$$

$$z_\text{vertex} \leqslant -650\,\text{m} \tag{4.2}$$

$$\rho_\text{vertex} \geqslant 300\,\text{m}, \tag{4.3}$$

where the last two lines describe a cylindrical volume around the DeepCore detector. The vertical z-axis is given in *IceCube detector coordinates* with the origin at a depth of $\sim 1800\,\text{m}$, while ρ_vertex is the radial distance from string 36 (cf. Figure 3.2), which is roughly in the center of the DeepCore detector. When z and ρ are used in the following, they refer to this cylindrical coordinate system centered at DeepCore.

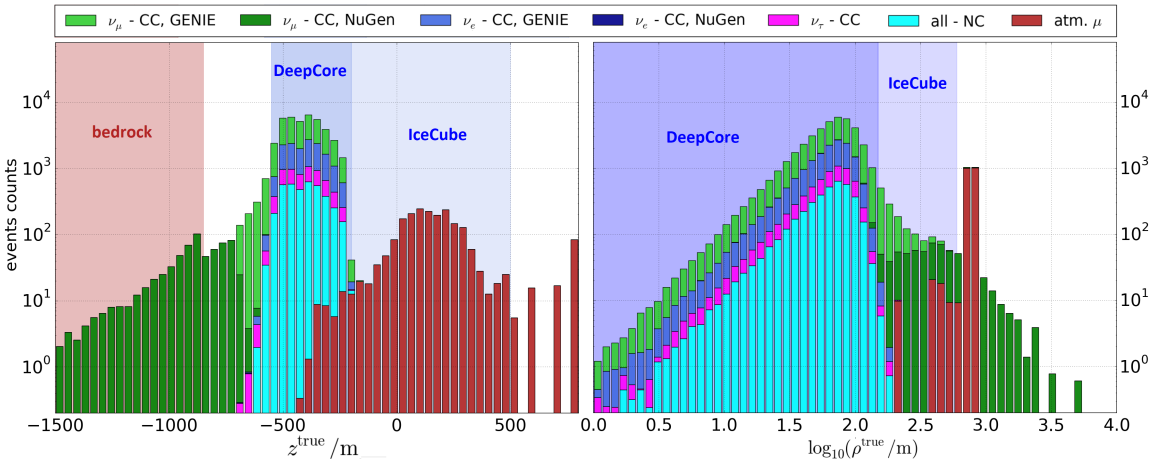


Figure 4.1: Distribution of depth z (left) and radial distance ρ (right) of the primary vertex to the center of DeepCore; the bedrock, DeepCore and the IceCube array are shown as shaded areas; note that noise events have no interaction vertex and are thus not shown, while atmospheric muons enter the detector from far outside, but are plotted at their position on the generation cylinder, discussed in Section 4.2.2.

Thus, far travelling and high-energy events in the GENIE samples are replaced by NuGen events. The resulting z and ρ distribution of the primary vertex are shown separately for NuGen and GENIE events in Figure 4.1 for the final data sample. As one would expect, events in the tails of the two distributions are mostly CC muon neutrino events, as other neutrinos do not produce far-travelling muons. However, the total contribution from NuGen is only on the order of $\sim 1\%$ of the total sample. In the following, NuGen and GENIE events are not distinguished, but treated as a single simulation sample.

For both generators, the simulated neutrino interactions generate secondary particles, which emit Cherenkov photons (cf. Section 3.2). The Cherenkov photons are then propagated through the ice, taking the local ice properties for absorption and scattering effects into account [107].

In case a photon is not absorbed but hits the surface of a PMT, the corresponding photon is stored in the so-called *photon-level*. After the propagation of all photons, the DOM hardware simulation is launched for each DOM. The DOM simulation drops some of the photons from the photon-level according to the photon-detection efficiency of the DOM. Then, it adds a waveform template to the voltage curve for each of the detected photons. This template is obtained from lab measurements by averaging over the response of all DOMs [82].

The resulting voltage curves are then run through the same online processing as actual data, described in Section 3.5.2. Therefore, the treatment of data and MC is identical after the measurement of the waveforms and the following data selection in Section 4.4 is described without distinguishing between data and MC.

4.2.2 Atmospheric Muon Simulations

The rate of atmospheric muons firing the DeepCore trigger is more than $\sim 10^6$ times higher than the corresponding rate from atmospheric neutrinos. However, the atmospheric muon

background is suppressed to the percent level by the subsequent filtering and selection chain. This means that a high number of atmospheric muons needs to be run through the simulation to obtain a single muon event in the final level sample. Thus, obtaining a sufficiently large MC sample for atmospheric muons is computationally challenging [83].

This challenge is faced by two generators: *CORSIKA* and *MuonGun*. The *COsmic Ray SImulation for KAscade* framework (CORSIKA) [108] runs detailed simulations of extensive air-showers, tracking secondary particles through different layers of the Earth's atmosphere. It was developed for the *KASCADE* experiment and became a *quasi-standard* for the simulation of atmospheric air-showers used by several air-shower experiments [109, 110]. It requires an assumption on the primary Cosmic Ray spectrum, which is set to the *Polygonato-Hoerandel* model described in [111]. Although more recent parametrizations exist, this model is sufficiently accurate for the description of the small atmospheric muon contamination in DeepCore. Here, only the muons generated within these air-showers are relevant, since the atmosphere and the thick ice above the detector shield it from any other background contamination. Since CORSIKA stops the simulation of atmospheric muons at the ice surface, the propagation from the surface to the IceCube and DeepCore detectors is done separately, including an accurate description of the muon energy loss and the corresponding photon production.

However, these CORSIKA simulations are computationally intense, since they require the simulation of the entire air-shower. Therefore, MuonGun was developed within the IceCube Collaboration to allow for a fast simulation of a large number of muons without the simulation of time-intense air-showers. It *bypasses* the CORSIKA simulation by directly generating muons on a vertically oriented generation cylinder centered at DeepCore with a radius of 800 m and a length of 1600 m. The generated muons are forced to aim towards the DeepCore fiducial volume, parametrized by a smaller, aligned target cylinder with a radius of 150 m and a length of 500m.

To link the generated muons to a realistic atmospheric muon flux, the templates used to generate muons randomly on the generation cylinder are obtained from CORSIKA. To do this, large CORSIKA simulations are run and the generated muons are propagated towards the surface of the generation cylinder. From these CORSIKA muons, the probability density function (pdf) of atmospheric muons in energy, direction and position on the cylinder is derived and used for the generation in MuonGun.

However, MuonGun does not include the simulation of multiple, coincident muons (*muon-bundles*) or muons, that do not pass the target cylinder, but still deposit charge in the DeepCore detector. To account for this, several CORSIKA sets were merged and checked for events that would not be generated in MuonGun. These events make up a fraction of $\sim 5\%$ of all generated muons. Since the muon contamination itself is $\sim 5\%$ of the total sample (cf. Table 4.1), this corresponds to a fraction of $\sim 0.25\%$ of the final sample. For these CORSIKA events, the final-level observables, later used in this work, were compared to the distributions obtained from MuonGun. Since no significant deviations from the MuonGun template were found, the CORSIKA events are not used for the final muon template.

Although the MuonGun simulation is substantially faster than CORSIKA, atmospheric muon simulations are still the bottle-neck in reducing the statistical errors on the final templates, since the simulated MC statistics correspond to only ~ 1 year of detector livetime. To reduce these uncertainties on the background, *Kernel Density Estimation* is used within the likelihood analysis, which is described in more detail in Section 6.2.1.

4.2.3 Triggered Detector Noise Simulations

The term *dark noise* refers to noise in the DOMs that is not caused by the observation of a Cherenkov photon. Such noise is mainly produced by thermal electronic noise and the decay of radioactive elements within the glass pressure housing of the DOM. It is added to the simulation of neutrino and muon events. In addition, such noise is capable to fire the DeepCore trigger itself in rare cases, called *triggered noise*. This is possible, since the SMT3 trigger (cf. Section 3.5.2) is tuned to far lower energies than the IceCube triggers and thus is far more sensitive to dark noise.

The simulation of such triggered noise is time-intense, since it requires a continuous simulation of all of the detector without any physics event. It is done using the so-called *Vuvuzela* tool, which includes an accurate estimate for the noise rate of every DOM based on in-situ measurements. As soon as random coincidendes fire the DeepCore trigger, the event is run through the same processing chain as neutrinos.

The total contamination of triggered noise is on the level of $\sim 0.1\%$ of the total sample. Contaminations on that level have typically been dropped in former DeepCore analyses, since they were negligible for the physics results. For the GRECO sample, this was not done for historical reasons, as the contribution from triggered noise was previously found to be higher, but substantially reduced during the last steps of the data selection.

The simulated MC statistics correspond to only ~ 0.3 years of detector livetime. Thus, it is the worst ratio of all simulated contributions. However, due to the tiny noise contribution, its impact on the total MC error is much smaller than the one from atmospheric muons.

In the same way as for the atmospheric muon background, the statistical uncertainties on the noise MC is reduced using Kernel Density Estimation, described in Section 6.2.1.

4.3 Experimental Data Sample

The experimental data used in this work was taken between April 2012 and May 2015 with IceCube in its 86 string configuration. It is split into three subsamples corresponding to the three seasons, called *years*, of data taking. Furthermore, each year is split into runs of typically 8 hours length. For each run, calibration data is taken and stored in a GCD file as described in Section 3.5.2. This data is used in combination with monitoring data to decide whether the run is flagged as *good* or *bad*, where all bad runs are removed from the following analysis.

Table 4.3: Summary of the good livetime for the three years of data used in the GRECO sample.

season	livetime/s	livetime/y
2012	28 257 176	0.895
2013	28 640 727	0.908
2014	29 987 404	0.950
total	86 885 307	2.753

A run is marked as bad in case of problematic behavior of the data acquisition, planned downtime of all or parts of the detector or unexpected events influencing the reliability of the

data acquisition. Since oscillation measurements are very sensitive to systematic influences, only very clean uptime is used in the following.

The resulting good livetime for each of the years is summarized in Table 4.3.

In addition to flagging whole runs as bad, the *Geometry, Calibration and Detector Status* (GCD) information (cf. Section 3.5.2) allows to identify single, problematic DOMs. This way, DOMs can be flagged that are broken, saturated, have an unreliable calibration or suspicious behavior in some of the monitoring data. Such DOMs can be excluded for all of the year, several months or just a single run. In the reconstruction presented in Chapter 5, these DOMs are removed from the likelihood to obtain an unbiased result in the reconstruction.

In addition, this work lead to the discovery of two so-called *Flaring DOMs* within DeepCore. These DOMs show a very abnormal behavior in the observed charges and their time distribution. Their discovery and their behavior is discussed in more detail in Section 4.5.1.

Besides the discovery of the Flaring DOMs, this work also revealed a significant mismatch between data and Monte Carlo in the observed charge per DOM, arising from an inaccurate *Single-Photo-Electron* (SPE) template used to assign charges to simulated photons, hitting the DOMs' photocathode. This is discussed in more detail in Section 4.5.2.

4.4 The GRECO Data Selection

The data processing is split into discrete steps, called *Levels*. The processing up to Level 2 (L2) is given by the online processing, while the subsequent levels are run *offline*, i.e. on the Northern Hemisphere. Level 3 is a common level shared by several analyses aiming on similar physics cases. Thus, the GRECO sample starts at the Level 3 dedicated to oscillation analyses. In total, it consists of 7 Levels, that are described in the following. For this work, it was adapted at Level 5 as developed by Michael Larson, while the subsequent Level 6 and Level 7 were developed in collaboration with Michael Larson [112] and Elim Thompson [113].

4.4.1 DeepCore Level 3

After the online processing, the data is run through the physics specific selection on Level 3, so-called *DeepCore Level 3*, which is used by all oscillation analyses [83]. It consists of several direct cuts on simple observables to reduce the amount of data sufficiently to run more advanced and computationally intense methods on higher levels.

It runs several cleaning methods on the raw pulse-map obtained from Level 2. These methods are similar to the SeededRT-Cleaning described in Section 3.5.2.3 and remove pulses from the original pulse-maps based on different causality criteria, i.e. requiring other hits to be within certain times and distances.

These pulse-maps are used for the following cuts on Level 3:

1. $N_{\text{MicroCount}} > 2$: This quantity counts the number of hit DOMs within DeepCore within an aggressively cleaned pulse-map. It is supposed to assure that there is a significant cluster of pulses within the DeepCore fiducial volume.
2. $Q_{\text{MicroCount}} > 2\,\text{PE}$: This quantity measures the corresponding charge analogously to

the previous cut. Since a detected photon causes $\sim 1\,\mathrm{PE}$ charge, these two cuts are highly correlated.

3. $Q_{z>-200} < 12\,\mathrm{PE}$: This quantity measures the deposited charge above $z = -200\,\mathrm{m}$ in a loosely cleaned pulse-map within $2\,\mu\mathrm{s}$ before the DeepCore trigger. Thus, it removes events, where a large amount of charge has been deposited in the upper IceCube detector, which is an indication for a downgoing atmospheric muon.

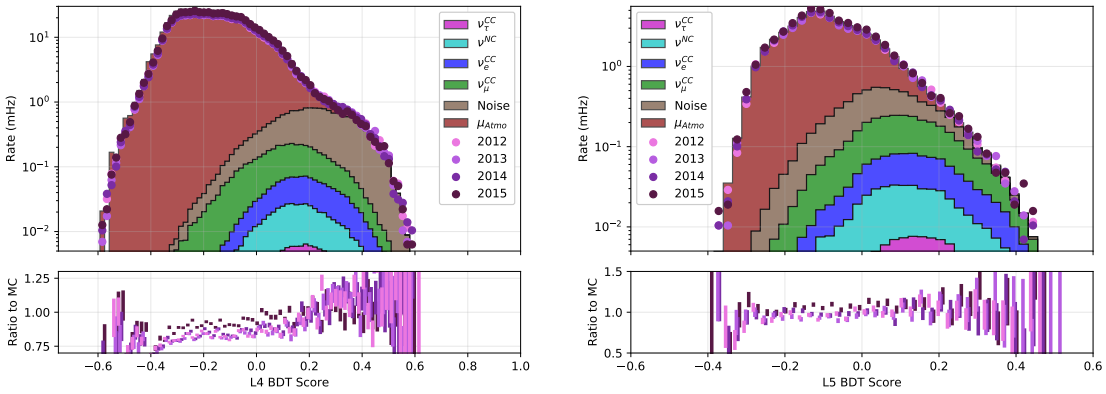
4. $z_{\mathrm{VertexGuess}} < -120$: This quantity gives the depth of the first hit in the SeededRT cleaned pulse-map (cf. Section 3.5.2.3). It is introduced, since hits in the upper detector are an indication for downgoing muons.

5. $Q_{\mathrm{uncleaned,veto}} < 7\,\mathrm{PE}$: This quantity assures that there are no more than $7\,\mathrm{PE}$ charge deposited in the veto region, i.e. in the IceCube array and the DeepCore veto cap, that can be causally connected to the DeepCore hits by the hypothesis of a downgoing atmospheric muon. One should note that this cut is analogous to the DeepCore filter. However, this cut uses a raw pulse-map without any filtering.

These simple cuts on Level 3 reduce the amount of atmospheric muons by almost two orders of magnitude, such that the following levels can apply more time-intense methods.

4.4.2 Processing to Level 6

On Level 4 and Level 5, the GRECO selection applies *Boosted Decision Trees* (BDTs), to further remove the background from atmospheric muons and triggered noise. Both BDTs are trained with a small number of variables, that are well-understood and -established within the IceCube Collaboration. The split into Level 4 and Level 5 is due to historical reasons, while in principle a single BDT could be trained to combine both levels. The BDTs are trained with Monte Carlo independent of the one used for this analysis and tested for only decent overtraining, to maintain a high efficiency. The development of these BDTs is discussed in more detail in the work by Michael Larson [112].



(a) BDT score on Level 4

(b) BDT score on Level 5

Figure 4.2: BDT scores on Level 4 and 5 of the GRECO data selection. The scores are used to remove atmospheric muons and triggered noise. The chosen cut is at a score of 0.04 for both levels, while events with higher values are kept for Level 6 (modified from [112]).

Before training the BDT, the GRECO selection requires more than 3 PE charge to be left in the SeededRT cleaned pulse-map and more than 6 PE in a more loosely cleaned pulse-map. This is necessary to calculate all quantities used in the following selection.

Level 4 BDT: The BDT trained on Level 4 is based on six variables:

1. $Q_{z>-200}$: Defined in Section 4.4.1.
2. $z_{\text{VertexGuess}}$: Defined in Section 4.4.1.
3. *QR6*: This quantity is the fraction of charge seen within the first 600 ns compared to the total charge of an event. It is motivated by the fact that the light emission in cascade-like events happens almost instantaneously. Thus, the corresponding, deposited charge is expected to be more peaked in time compared to the charge depositions seen from muons, which deposit energy slowly during their propagation.
4. *C2QR6*: This quantity is calculated analogously to QR6, but removing the first two observed hits from the calculation. This gives a handle on a possible noise contamination in the early hits.
5. β_{LineFit}: This quantity gives the reconstructed velocity of a hypothetical particle travelling through the detector and causing the observed charge depositions. It is based on a simple, low-level reconstruction. Values close to c indicate a real particle causing the charge depositions, which could only be a muon.
6. $I_{\text{max}}/(I_x+I_y+I_z)$: This quantity uses the tensor of inertia I calculated by weighting the individual DOM positions by their observed charge, analogously to the mass-weighted tensor of inertia used in mechanics. The resulting three eigenvalues of the tensor are called I_x, I_y and I_z. Due to their elongated shape, muons are expected to have one of these three Eigenvalues I_{max} dominating the others. Thus, it gives a handle on the atmospheric muon background and flavor separation.

One might note that the above variables are similar to those used on Level 3. This is due to the fact that the GRECO Level 4 was adapted as Level 3 for all DeepCore analyses. However, the BDT was replaced by a series of straight cuts on Level 3 to simplify the understanding and allow for easy checks and reproducibility. Thus, the GRECO Level 4 can be seen as a GRECO-specific refinement of the Level 3 selection.

The final BDT score S_{BDT4} is shown in Figure 4.2(a). The applied cut on Level 4 is $S_{\text{BDT4}} > 0.04$. As one can see, this removes most of the atmospheric muon background, while it was tuned to keep a high fraction of tau neutrinos, appearing from atmospheric oscillations. This is achieved with nearly 100% efficiency.

Level 5 BDT: The BDT trained on Level 5 is based on six variables:

1. d_{Q1-Q4}: This quantitiy gives the distance between the center of gravity calculated for the first 25% and the last 25% of the observed charge. An elongated shape of the charge deposition would give large values for this distance, while a low value indicates a spherically symmetric deposition.
2. Δz_{Q1-Q4}. This quantity describes the vertical distance between the center of gravity for the first 25% and the last 25% of deposited charge. Since atmospheric muons are downgoing events, this gives a handle for the reduction of atmospheric muons.

3. $t_{75\%Q}$: This quantity gives the time at which 75% of the total charge of the event is deposited within the detector. It is used, since neutrinos induce peaked charge depositions, while triggered noise and atmospheric muons give a rather uniform distribution in time.

4. Q_{VICH}: This quantity gives the total charge from *Veto Identified Causal Hits* in a loosely cleaned pulse-map. These are hits before the DeepCore trigger, that are causally connected to the first DOM contributing to the trigger condition. The idea is analogous to $Q_{\mathrm{uncleaned,veto}}$ discussed for Level 3.

5. ρ_{FirstHit}: This quantity gives the radial distance of the first hit DeepCore DOM from the center of DeepCore. It uses a loosely cleaned pulse-map to allow for more hits to veto the event than using the strict SeededRT cleaning.

6. $\cos(\theta_{\mathrm{SPE11}})$: This quantity gives the cosine of the reconstructed zenith-angle θ_{SPE11} from an algorithm using the times of the first pulses in each DOM in a likelihood method, where the observed times are compared to the *Single-Photo-Electron* (SPE) expectation for an infinite muon track. The 11-label indicates that 11 seeds are used, while the best one with respect to the likelihood value is chosen to assure that the global minimum of the likelihood space is found.

The final BDT score S_{BDT5} is shown in Figure 4.2(b). The cut is applied at $S_{\mathrm{BDT4}} > 0.04$, which is the same value as for Level 4 (only by coincidence). The cut reduces both, the amospheric muon and the triggered dark noise contamination.

Level 6 Straight Cuts: Level 6 refines the selection from Level 5 by several straight cuts, that are described in the following:

1. $8 \leqslant N_{\mathrm{Channel}} < 100$: This is a simple cut on the total number of hit DOMs in the final SeededRT cleaned pulse-map. The lower bound strongly reduces the noise contamination. The upper bound at 100 is introduced for computational reasons only. Events with more than 100 active channels are never reconstructed into the energy range $< 90\,\mathrm{GeV}$ used in this work, but take an exceptionally long time in the reconstruction on Level 7. Thus, they are removed on Level 6 before the reconstruction.

2. $\mathrm{FR}(r) < 0.05$: The *Fill-Ratio* calculates the ratio of the number of hit DOMs over the number of all DOMs within a sphere with radius r around a hypothetical vertex. This vertex is estimated by the position of the first hit DOM in DeepCore using a decently cleaned pulse-map. The radius r is chosen to be 1.6 times the average distance of all hit DOMs within this pulse-map towards this vertex. For atmospheric mouns and triggered noise, the ratio is small compared to the very local signature from neutrino events. Thus, a minimum Fill-Ratio reduces the noise and muon backgrounds.

3. $z_{\mathrm{FR}} - \rho_{\mathrm{FR}}$: This cut is a two-dimensional cut on the position of the reconstructed vertex, given by clindrical coordinates (z, ρ) centered at DeepCore. The used vertex reconstruction is called *FiniteReco* [41]. It uses the *SPE11*-fit of an infinite-muon hypothesis, described in Level 5. FiniteReco extends this fit by an estimate for the start and end point of the previously infinite track. To do this, it projects the position of all hit DOMs in a cleaned pulse-map on the track. For this, the direction of projection is required to form the Cherenkov angle θ_C with the direction of the track hypothesis. The first and the last projections on the track are then taken as the track's start and end point. In a refinement step, the start and end point are then shifted to

minimize a likelihood, that takes the hit and no-hit probabilities for each DOM into account. The two-dimensional cut requires z and ρ to be below a given value. Thus, vertices far outside or high in the detector are rejected, as they are likely atmospheric muons surviving up to Level 6. An analogous cut is repeated after the reconstruction on Level 7 (cf. Figure 4.4).

4. $N_{Corridors} = 0$: This cut requires no hit to be observed along so-called *corridors*. These corridors run radially from the center of DeepCore to outside of IceCube through uninstrumented Detector volume and are due to the geometry used for the deployment of strings (cf. Figure 3.2). Through these corridors, atmospheric muons can reach DeepCore without causing veto hits in the surrounding IceCube array. These corridors are known to be the origin of a high fraction of the remaining atmospheric muon contamination. Thus, events with any hit DOM on a string of the IceCube array, which is at the edge of such corridor, are removed from the sample.

5. $\mathrm{Asym}_Q > 0.85$: This cut was introduced after finding the *Flaring DOMs*, described in Section 4.5.1. These DOMs were found to occasionally observe large charge depositions, that are not causally connected to any neutrino interaction. Events including such *flares* are driven by these large charge depositions and supposed to be removed from the analysis. The *charge asymmetry* is defined by Equation 4.4

$$\mathrm{Asym}_Q = \frac{\sqrt{\sum_{i=1}^{N_{\mathrm{DOMs}}} q_i^2}}{\sum_{i=1}^{N_{\mathrm{DOMs}}} q_i} = \sqrt{\frac{1}{N_{\mathrm{DOMs}}}\left(\frac{\mathrm{Var(q_i)}}{\langle q_i \rangle^2} + 1\right)}, \tag{4.4}$$

where q_i is the total charge observed in DOM $i = 1, ..., N_{\mathrm{DOMs}}$. From the first equation, it is easy to see that $\mathrm{Asym}_Q \in [0..1]$, where values close to 0 correspond to very homogeneous charge distribution throughout all DOMs, while values close to 1 correspond to few DOMs dominating the total charge. The distribution is shown in Figure 4.3. Since the flares are characterized by large charge depositions, uncorrelated with all other DOMs, such flare events show up close to 1. The decent cut at 0.85 was found to remove all of the events that were obviously caused by such flares, while keeping almost all of the remaining data. More details on the Flaring DOMs can be found in Section 4.5.1. The charge asymmetry distribution on final level is added to Appendix C.

The cuts on Level 6 were largely developed by Michael Larson, while the discovery of the Flaring DOMs and the charge asymmetry cut are part of this work.

After Level 6, the event-selection is finalized except for cuts depending on the final event reconstruction.

4.4.3 Processing to Final Level 7

After Level 6, the GRECO sample is run through the *Pegleg* reconstruction, described in Chapter 5. It returns an estimate for the vertex position and time (x,y,z,t), the neutrino direction (θ_ν, ϕ_ν) and the neutrino energy E_ν, which are used for the post-reconstruction cuts listed below. To present all of the data selection at once, these quantities are assumed to be available here, while the reconstruction is presented afterwards in the following chapter.

The cuts on Level 7 are not applied on any earlier level, since the time spent on the Pegleg reconstruction is the bottle-neck in the total computation time of the GRECO processing.

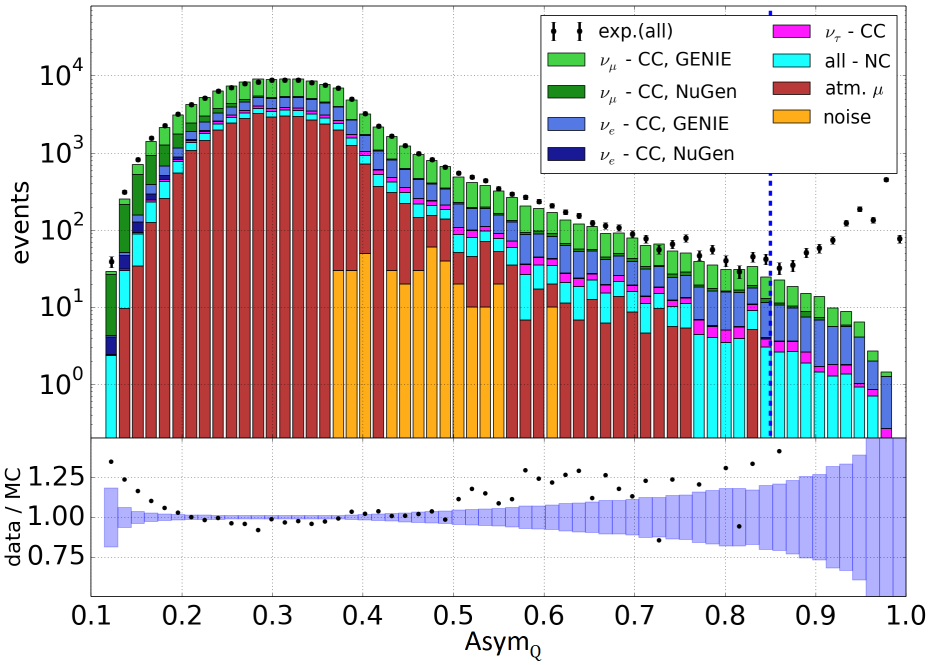


Figure 4.3: Distribution of the charge asymmetry calculated from Equation 4.4; it is used to remove events with a very asymmetric charge distribution (Asym$_Q \sim 1$); a cut is applied, removing events with Asym$_Q > 0.85$, which removes the large peak at ~ 1 caused by the two Flaring DOMs discussed in Section 4.5.1.

To reconstruct all MC samples used in this work, the reconstruction took $\sim$ 6 months using several computer clusters in parallel. Thus, any cut reducing the amount of data before reconstruction is useful. However, once the reconstruction is done, it can be used to substantially reduce the background compared to Level 6.

While Level 4, 5 and most of Level 6 were developed by Michael Larson, this work contributed strongly on the design of Level 7.

The post-reconstruction cuts applied on Level 7 are listed below:

1. $z_{\mathrm{PL}} - \rho_{\mathrm{PL}}$: This cut is analogous to the $z - \rho$-cut applied on Level 6 using FiniteReco. However, since the Pegleg reconstruction is substantially more accurate in estimating the true event vertex, it allows for an additional background reduction. The cut removes large values of ρ and z in a two-dimensional cut shown in Figure 4.4. One should note that it does not remove events entering the detector from below. Since atmospheric muon contamination is exclusively downgoing, the resulting contamination from atmospheric muons is small. Additionally, any lower containment cut would strongly reduce the amount of good data, since DeepCore is not shielded by the IceCube array from below. Thus, for a lower veto the DeepCore fiducial volume would have to be reduced.

2. $\sqrt{\mathrm{Var}(t_i)} - E_\nu^{\mathrm{PL}}/N_{\mathrm{Channel}}$: This cut is a two-dimensional cut reducing the noise and atmospheric muon contamination, as shown in Figure 4.5. The first observable is the standard deviation of the first hit times of all DOMs within the SeededRT cleaned pulse-map. The idea is analogous to the $t_{75\%Q}$-cut on Level 5. For noise events, a hit is equally likely at any time within the readout-window except for selection effects.

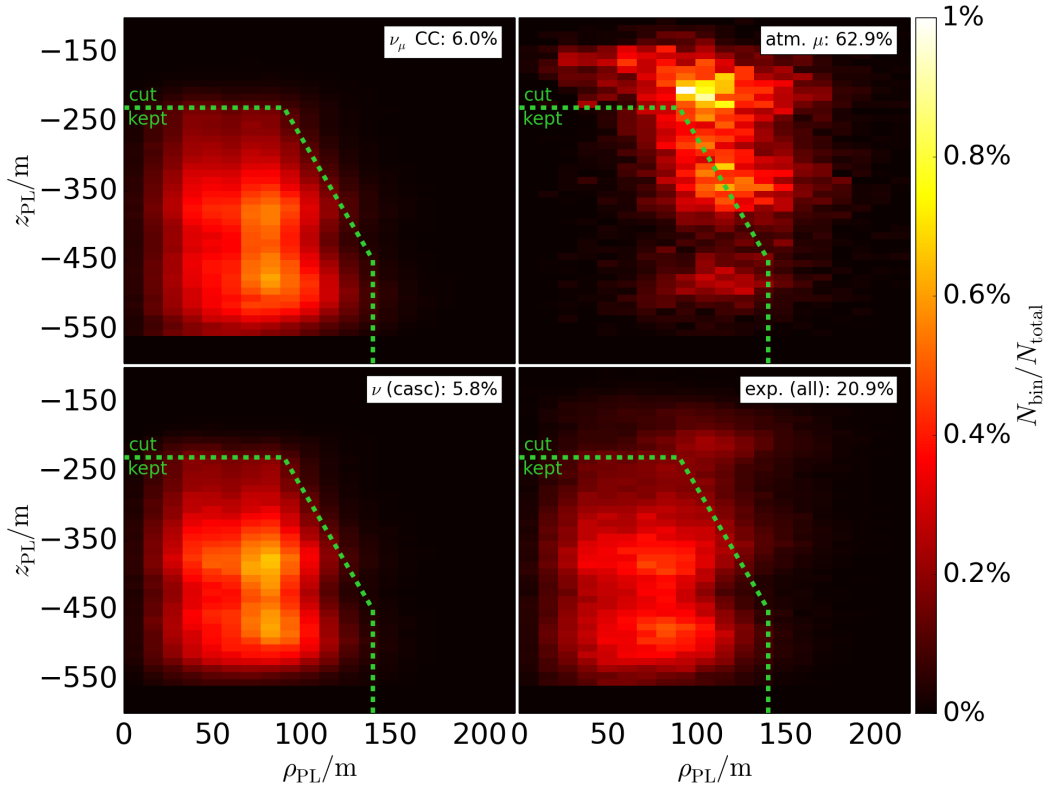


Figure 4.4: Visualization of the Level 7 cut on the reconstructed vertex position obtained from Pegleg in cylindrical coordinates z and ρ; the plot is shown for CC muon neutrino events, i.e. tracks (top left), atm. muons (top right), all other neutrino events, i.e. cascades (bottom left), and experimental data (bottom right); each plot is labelled by the fraction of events that are removed by the two-dimensional cut given by the green dashed line.

In contrast, the light emission caused by neutrinos is very peaked in time. Thus, a large standard deviation in the first hit times is an indication for a noise event. A similar argument can be made for the ratio of reconstructed neutrino energy per hit DOM, i.e. $E_\nu^{\mathrm{PL}}/N_{\mathrm{Channel}}$: The number of hit DOMs, N_{Channel}, is strongly correlated with the neutrino energy. Thus, the above ratio is a ratio of two energy estimators, where the nominator includes the knowledge of topology, ice properties and detector response, while the denominator does not. As a result, large ratios mean that that the Pegleg reconstruction favors unusually large energy depositions per channel. This happens, if the event is reconstructed into obscure, dark regions of the detector, i.e. into corridors or regions of bad ice-quality. These events are typically atmospheric muons sneaking in through these dark regions or triggered noise does not follow the expected hit pattern for a neutrino event. Since noise was not found to cause events with $N_{\mathrm{Channel}} > 11$, this cut is applied only for $N_{\mathrm{Channel}} \leqslant 12$.

3. $E_\nu^{\mathrm{PL}} \in [3.5, 90]$ GeV: This cut is applied on the reconstructed neutrino energy obtained from the Pegleg reconstruction (cf. Chapter 5). It is not a general GRECO selection cut, but the range of reconstructed energies used in this analysis. It can be seen as final level cut, since only these events enter the likelihood analysis, described in Chapter 6.

4. $\cos\left(\theta_\nu^{\mathrm{PL}}\right) \leqslant 0$: This cut is applied on the reconstructed neutrino zenith-angle obtained from the Pegleg reconstruction. Like for the previous cut, this is not a general GRECO

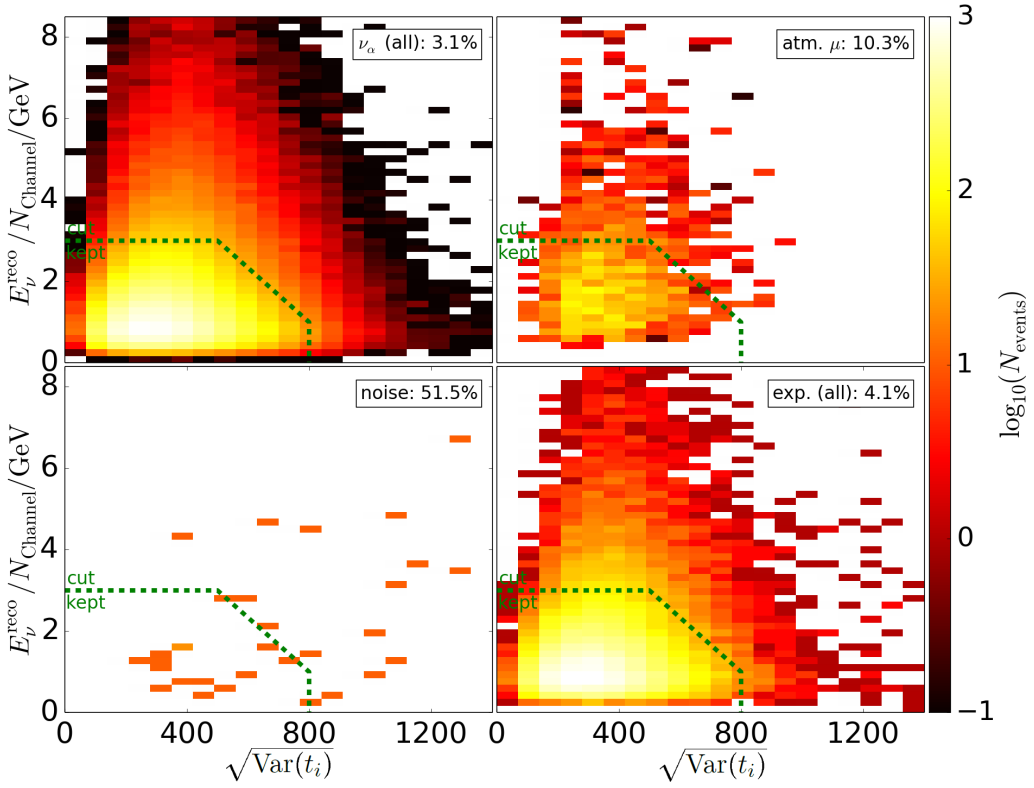


Figure 4.5: Visualization of the Level 7 cut on the variance of the first hit times and reconstructed energy per channel for events with $N_{\mathrm{Channel}} \leqslant 12$; the plot is shown for all neutrino events (top left), atm. muons (top right), triggered noise (bottom left) and data (bottom right); each plot is labeled by the fraction of events that are removed by the two-dimensional cut given by the green dashed line; note the logarithmic color-scale.

cut, but the range of reconstructed zenith-angles used in the following analysis.

After these cuts, the data selection is finalized. The resulting rates of the final level sample are summarized in Table 4.4.

Table 4.4: Final level rates of the GRECO selection; the given values correspond to the best-fit values of the systematic parameters obtained from the fit to experimental data in Section 8.1.

expectation	ν_e CC	ν_μ CC	ν_τ CC	all NC	atm. μ	noise	Σ
rate of events / mHz	0.108	0.291	0.031	0.044	0.024	0.0006	0.497
tot. number of events	9 390	25 255	2 674	3 794	2 048	52	43 214
fraction of sample	21.7%	58.4%	6.2%	8.8%	4.7%	0.1%	100%

From this sample, three observables are used in the following likelihood analysis: the reconstructed zenith-angle, neutrino energy and a flavor separating variable, called *Particle Identification* or *PID* (cf. Section 5.5). These three observables are shown in Figure 4.6 for data and Monte Carlo. The corresponding true energy and zenith-angle distributions are

shown for comparison.

The values in Table 4.4 and the distributions in Figure 4.6 depend on several systematic effects, fitted as nuisance parameters in the likelihood analysis. The shown values are obtained from the best-fit value of the nuisance parameters to experimental data. The fit is described in more detail in Section 8.1.

In addition to the observables shown in Figure 4.6, several control distributions were produced. A selection of these variables can be found in Appendix C.

However, before using this sample in the likelihood analysis in Chapter 6, two more things are presented: First, in Section 4.5 two hardware issues are discussed, which were discovered in experimental data and previously lead to a major data-MC disagreement. Second, in Chapter 5 the reconstruction algorithm is presented that was used to obtain the observables in Figure 4.6.

4.5 Observation and Fix of Data-Monte-Carlo Disagreement

During the development of this analysis, two issues with the experimental data used in DeepCore analyses were found. They are presented in the following along with the corresponding fix, used for this work.

4.5.1 Flaring DOMs

The term *Flaring DOMs* refers to two Digital Optical Modules (DOMs) on string 83 in depth of 2270 and 2440 m, i.e. deep within the DeepCore fiducial volume. They did not show up as being suspicious in any standard monitoring, but are now suspected to occassionally emit light in irregular time steps, caused by an unknown hardware effect. At the same time, this light is observed by the PMT in the same DOM, causing unexpectedly large charge depositions without any correlation to a neutrino event.

These DOMs first showed up in a comparison of data and MC distributions in this work. In particular, they became visible in a two-dimensional distribution of reconstructed event vertices, as shown in Figure 4.7 for data and Monte Carlo. While Monte Carlo does not show any suspicious peak arising from single misbehaving DOMs, the two Flaring DOMs are clearly visible (center plot), because the large charge depositon in these DOMs influences the reconstruction to push the event vertex into the corresponding DOM. In the same way, the one-dimensional, event-wise charge-per-channel distribution features an excess at high values, which is almost exclusively due to these two DOMs. The charge-per-channel distribution is shown in Figure 4.8.

In this context, the charge asymmetry cut from Level 6 was developed (cf. Section 4.4.2). This cut was found to remove both, the excess in Figure 4.8 and the two peaks in Figure 4.7, at the cost of losing only a negligible fraction of neutrino events. In Figure 4.7 and 4.8 the distributions after the charge asymmetry cut are shown for comparison.

Moreover, during a general re-processing of the sample, the two Flaring DOMs were removed completely from the pulse-maps of Monte Carlo and data, that enter the reconstruction from Chapter 5. Still, the charge asymmetry cut was kept for safety reasons. In addition, these DOMs were intensely studied by the Calibration Group of the IceCube Collaboration, testing

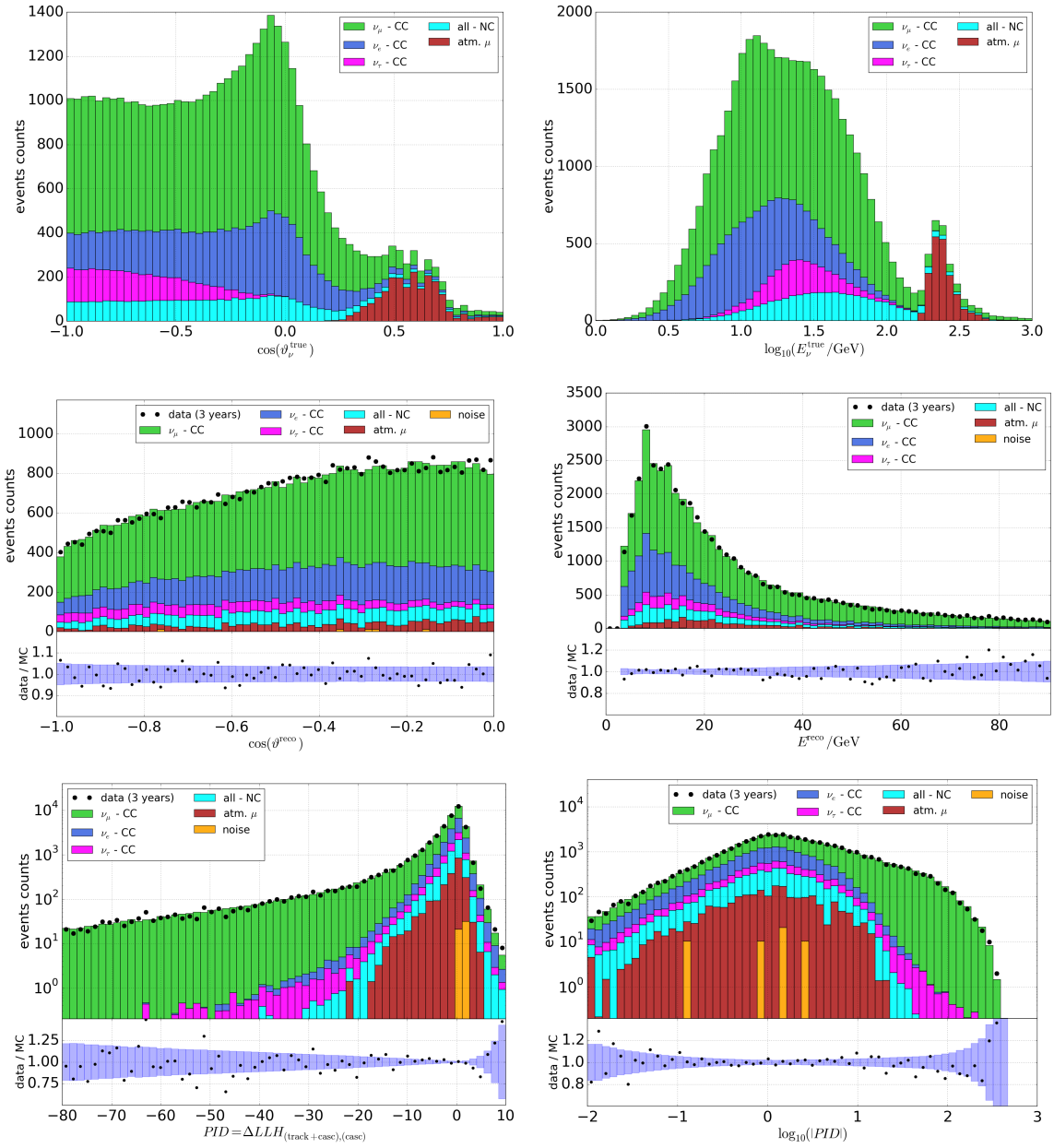


Figure 4.6: One-dimensional distributions of Level 7 observables of the GRECO sample; the first two plots show the true distribution of energy and zenith-angle of Monte Carlo events according to the best-fit values of all nuisance parameters discussed in Section 6.3; the remaining plots show the three observables (reconstructed energy, zenith-angle and PID) later used in the likelihood analysis in Chapter 6; note that the PID (cf. Section 5.5) is shown on a linear-scale (bottom, left) and on a log-scale for its absolute value (bottom, right) for illustration purposes.

all DOMs for such flares and re-analysing their properties. However, a similar behavior was not found for any other DOMs.

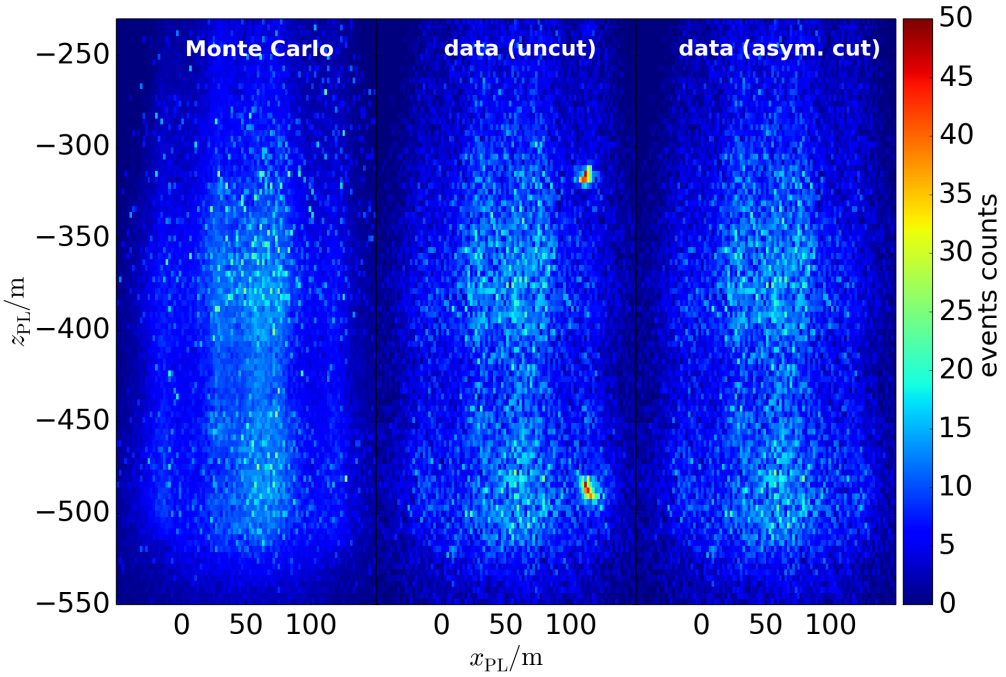


Figure 4.7: Two-dimensional distribution of reconstructed event vertices in one horizontal direction x and depth z for one of three slices in y; the comparison of MC (left) and data without the charge asymmetry cut (center) clearly shows two clusters of events, drawn to the position of the Flaring DOMs by the large, unphysical charge excesses; after the charge asymmetry cut (right) these events are removed from data.

4.5.2 Single-Photo-Electron (SPE) Template Mismatch

Besides the Flaring DOMs, another mismatch was found in the charge description between data and Monte Carlo, leading to a tilt in the charge-per-channel distribution from Figure 4.8. The corresponding charge-per-event distribution is attached in Appendix C.

Since the charge-per-channel distribution is the charge-per-DOM distribution averaged over all hit DOMs within one event, this disagreement indicates a mismatch in the underlying charge-per-DOM distribution, which is shown in Figure 4.9. Note that in the charge-per-DOM distribution each entry corresponds to a single hit DOM such that each event enters the histogram several times, depending on the number of hit DOMs.

The charge-per-DOM distribution is very stable with respect to the systematic uncertainties, discussed in Section 6.3. As a result, the disagreement in Figure 4.9 cannot be compensated by any reasonable value of the nuisance parameters from Section 6.3.

After the charge-per-DOM disagreement was discovered, the issue was brought to the *Calibration Group* of the IceCube Collaboration, triggering follow-up studies by Spencer Axani [114] and Martin Rongen [115]. They found that the *Single-Photo-Electron* (SPE) distribution, used to assign a random charge to MC pulses, does not describe the real DOM response. As the charge per pulse is not accurately simulated, this leads to the observed disagreement in the charge-per-DOM, charge-per-channel and charge distributions [115].

Note that in Figure 4.9, the charge-per-DOM distribution is also shown for applying the GRECO selection to 2015 data: Due to a re-calibration of the observed charges in early 2015

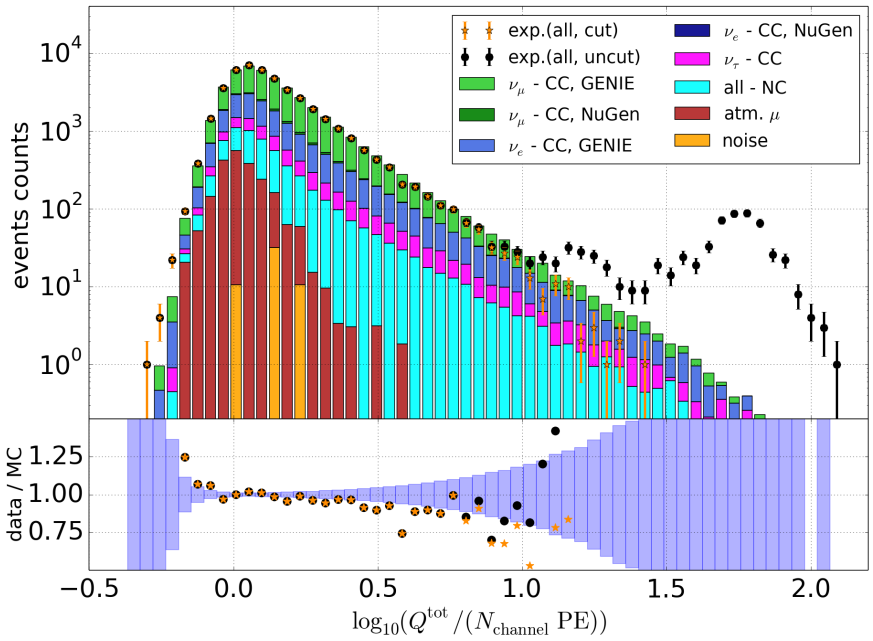


Figure 4.8: Charge-per-channel distribution for MC and data before (*uncut*) and after (*cut*) the charge asmmetry cut; the blue shaded areas in the lower subplot indicate the statistical uncertainty; the disagreement at $\log_{10}(Q^{\mathrm{tot}}/(N_{\mathrm{channel}}/\mathrm{PE})) > 1$ is caused by the *Flaring DOMs* and removed by the cut; note that the MC distributions are shown before the charge asymmetry cut, but change only slightly after the cut.

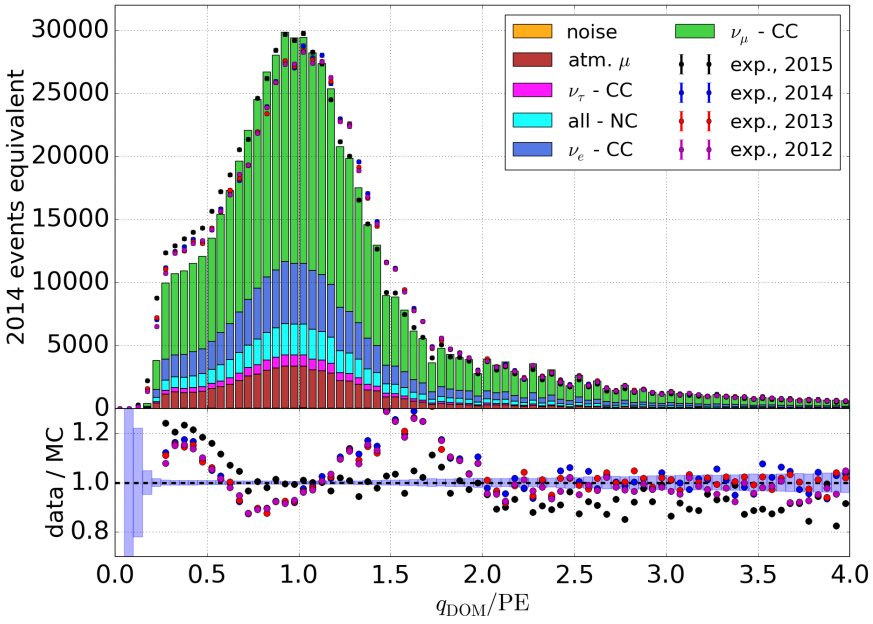


Figure 4.9: Charge per DOM for data and MC: note that entries do not correspond to events but hit DOMs, such that events enter the distribution several times; all years of data and MC are scaled to the observed integral value of 2014 data; the lower subplot shows the data over MC ratio with the statistical uncertainties indicated as blue shaded areas; the observed mismatch between data and MC is caused by the mismodelling of the SPE template; in 2015, a re-calibration was applied, causing the observed shift between 2012-2014 and 2015.

(unrelated to the above findings), the charge per DOM for 2015 deviates from the curves observed for 2012 – 2014. Although the resulting data fits more accurately to MC around $q_{\mathrm{DOM}} \sim 1\,\mathrm{PE}$, the agreement between data and MC on analysis level is worse for 2015. This is due to the fact that the mean value of the charge per DOM was approximately correctly described by Monte Carlo for 2012 – 2014 data, which is not the case for 2015. However, the agreement of the mean charge per DOM is crucial for an accurate description of data, since most selection relevant quantities are calculated event-wise by integrating charges or charge-weighted quantities over several DOMs.

Producing a more consistent description of the charge per DOM in MC and data for all years is a major goal of ongoing collaboration efforts. However, Monte Carlo samples using corrected SPE template are not expected to be available within the scope of this work.

Instead, the impact of the charge on final level observables was removed by re-phrasing the likelihood used in the final-level reconstruction from Chapter 5. By re-phrasing its likelihood, the reconstruction was made independent of the charge observed in the individual DOMs. Using this charge-independent reconstruction, data and Monte Carlo were found to match, while the re-phrased likelihood obtains the same resolution in energy and zenith as obtained by the charge-dependent one. The changes to the likelihood are discussed in more detail in the following chapter.

5 Development of a Low-Energy Event Reconstruction

The term *reconstruction* refers to a method, extracting physical properties of an event from the observed pattern of charge. In IceCube, these properties are typically the direction of the incoming neutrino and its energy. In addition, the neutrino flavor or interaction type (CC or NC) is estimated by a parameter called *Particle Identification* (PID), which separates experimental signatures as tracks or cascades.

Within IceCube, several reconstructions are available for energy, zenith and PID. While some of them reconstruct all of these quantities simultaneously, others focus only on one of them. Since IceCube covers 6 orders of magnitude in neutrino energy, their performance depends strongly on the considered energy regime.

In this chapter, the reconstructions commonly used below $E_\nu \sim 100\,\text{GeV}$ are briefly presented, which is a subset of all available reconstructions. Then, a new reconstruction is presented that was developed for the simultaneous reconstruction of energy, zenith and PID of events at extreme low energies, i.e. down to $E_\nu \sim 5\,\text{GeV}$. Although developed for this analysis, it has recently been adopted by almost all DeepCore analyses and is also used as standard reconstruction by the PINGU Collaboration. It provides high performance in the reconstruction of all three, above mentioned variables, but comes at the price of high computation times. Parts of the reconstruction were published in [60], [116] and [81].

Note that for oscillation measurements, angular reconstructions are typically stated only for the zenith-angle, since the atmospheric oscillation pattern is independent of the neutrino's azimuth-angle (cf. Section 2.4.2). Thus, the reconstructed azimuth-angle (if any is returned by the reconstruction algorithm) is ignored in the following.

5.1 Pre-existing Reconstructions

5.1.1 SANTA Zenith-Angle Reconstruction

The *Single String ANTARES-Inspired Analysis* (SANTA) is a reconstruction algorithm adapted from the *Astronomy with a Neutrino Telescope and Abyss Environmental Research* (ANTARES) experiment in [117] and modified for IceCube/DeepCore purposes by Juan Pablo Yáñez [118].

It uses so-called *direct photons* to reconstruct a zenith-angle for a neutrino event. Direct photons are photons that are assumed to be unscattered, i.e. that traveled on a straight line between their generation and their detection at a DOM's photo-cathode. Thus, they are

on the Cherenkov cone discussed in Section 3.2. To identify such photons, each string with more than three hit DOMs is tested individually, whether the arrival times of the first hits in these $\geqslant 3$ DOMs agree with the hypothesis of a Cherenkov wavefront passing the string. This is visualized in Figure 5.1. Depending on the zenith-angle of the charged particle, the corresponding Cherenkov cone is rotated with respect to the vertical string. Thus, from the depth at which the charged particle is closest to the string, the photon arrival times at the DOMs up-string and down-string differ and encode the information of the particle's zenith-angle. Thus, knowing at what time and depth unscattered photons were seen allows for a reconstruction of the particle's zenith-angle [118].

Note that for a single string, the system is symetric with respect to rotations around the string axis. Thus, no azimuth-angle is reconstructed and the reconstruction depends only on 4 parameters: the zenith-angle θ_ν, the closest distance to the considered string r_c, the time of closest approach t_c and the depth at the point of closest approach z_c. By assuming an infinitely extended track of such charged particle, these four parameters are sufficient to derive the estimated time of arrival $t_{\text{exp,i}}$ in case of direct photons for each DOM i on the string [118].

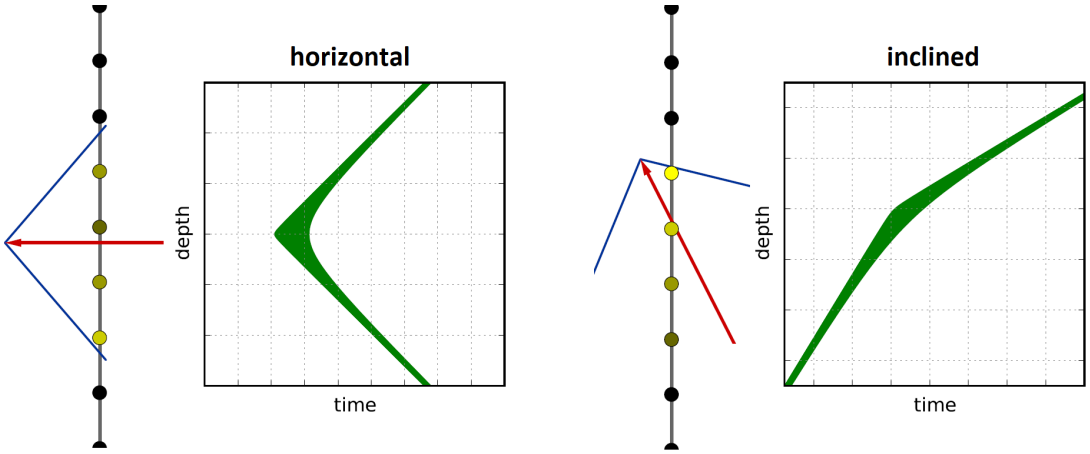


Figure 5.1: Illustration of the SANTA reconstruction for a horizontal (left) and an inclined, upgoing muon (right), emitting a Cherenkov cone observed by several DOMs on the same string; the resulting light observation of unscattered photons is given by the green shaded area in the two-dimensions of depth and time (from [118]).

The observed arrival times $t_{\text{obs,i}}$ and the expected arrival time $t_{\text{exp,i}}$ are then compared for all hit DOMs in a χ^2-method. The uncertainty used for the χ^2 calculation is typically chosen to be $\sigma_t = 20\,\text{ns}$. An additional term is included to penalize large charge deposition being far from the point of closest approach, which is of minor relevance and not discussed here. The resulting χ^2-function is minimized with respect to the four parameters described above, resulting in a reconstructed value for the zenith-angle of the event [118].

In some cases, several strings provide $\geqslant 3$ DOMs with direct photons being observed. In that case, SANTA provides a multi-string algorithm that breaks the rotation symmetry and adds another parameter to determine a supporting point for the track hypothesis. In this case, the χ^2-method is extended to include the arrival times from both strings [118].

The SANTA algorithm was found to provide good resolutions at $E_\nu \sim 15 - 100\,\text{GeV}$, including the first oscillation minimum (cf. Figure 2.14). However, it suffers from the necessity of

a large number of direct photons. In case no string provides $\geqslant 3$ DOMs consistent with a direct photon hypothesis, the reconstruction fails. For the GRECO sample, this is the case for $\sim 95\%$ of all final-level events with $E_\nu < 15\,\text{GeV}$. Since these events carry most information on the NMO, it is unpractical to use the SANTA reconstruction for the following NMO analysis.

5.1.2 LEERA Energy Reconstruction

The *Low-Energy Energy-Reconstruction Algorithm* (LEERA) is an energy reconstruction used below $\sim 100\,\text{GeV}$, which was developed by Andrii Terliuk [119]. It requires a seeded direction, a start point (vertex) and an end point of a track hypothesis and uses only DOMs that are within a cylinder of infinite length and 200 m radius, aligned with the seeded track hypothesis. It is based on a so-called $p(\text{hit}) - p(\text{noHit})$ likelihood, that optimizes the binary probabilities of DOMs being either hit or not hit [119].

The fitted hypothesis consists of a finite track and a starting cascade at the primary vertex. The seeded track direction is typically obtained from SANTA. The algorithm then modifies the energy of the starting cascade and the start and end point of the track along the track direction. Since the energy of the track is directly obtained from its length, assuming a minimum-ionizing track (cf. Equation 3.2), and the cascade direction is assumed to be coaxial with the track, the resulting optimization problem consists of only three free parameters: the start and end point of the track and the cascade energy [119].

These parameters are modified to optimize the $p(\text{hit}) - p(\text{noHit})$ likelihood. The best fit hypothesis is then converted into an energy estimate by converting the final track length into a track energy and adding it to the cascade energy. The combined energy is taken to be the reconstructed neutrino energy from LEERA [119].

Similar to SANTA, LEERA has proven to give good resolutions for events below 100 GeV. However, its performance depends strongly on the direction and vertex of the tested hypothesis. Although it gives similar resolutions as the reconstruction developed in this work, it requires an accurate direction and vertex reconstruction to do so, which was previously not available below $\sim 15\,\text{GeV}$ [119].

5.2 The Millipede Framework

Millipede is a software framework used for high-level reconstructions in IceCube. It consists of the Millipede likelihood and several auxiliary tools. Since the reconstruction developed in this chapter is based on it, the framework is briefly described in the following.

5.2.1 The Millipede Likelihood Space

The Millipede likelihood is a Poissonian likelihood, comparing the observed number of photons $N_{i\tau}$ in each DOM $i = 1,...,M$ in several time bins $\tau = 1,...,T_i$ to the expected number of photons $\mu_{i\tau}$ for a given hypothesis. The resulting negative log-likelihood LLH_mp is shown in Equation 5.1

$$\text{LLH}_\text{mp} = -\sum_{i}^{M}\sum_{\tau}^{T_i}\left[N_{i\tau}\cdot\log(\mu_{i\tau}) - \mu_{i\tau} + \log\left(\Gamma(N_{i\tau}+1)\right)\right]. \tag{5.1}$$

The Poissonian statistics arise from the discreteness of the photon counts. However, the continuous charge is used as a proxy for the observed and predicted number of photons as previously discussed (cf. Section 3.5.2). Thus, the factorial term of the Poissonian is replaced by a Γ-function to account for the non-integer values of $N_{i\tau}$.

In Millipede, the expected number of photons $\mu_{i\tau}$ is expressed in terms of discrete energy losses E_k with $k = 1,...K$ in the detector volume, as summarized in Equation 5.2

$$\mu_{i\tau} = \sum_{k}^{K} \Lambda^{\mu/c}_{i\tau k}(\vec{x}_i, \vec{y}_k, \theta_k, \phi_k, \Delta t_\tau) \cdot (E_k/\mathrm{GeV}) + n_i, \tag{5.2}$$

where the expected number of photons $\mu_{i\tau}$ is the sum of the photons caused by the individual energy depositions E_k and a time-independent noise term n_i, that is obtained from multiplying the known noise rate for each DOM with the length of the read-out window. The link between energy losses and expected photons is illustrated in Figure 5.2.

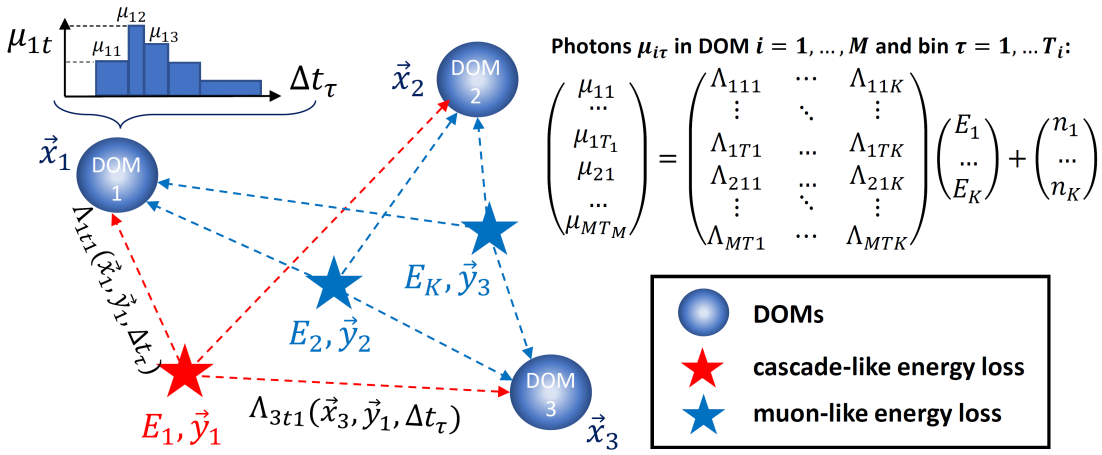


Figure 5.2: Sketch of how Millipede estimates the expected number of photons (i.e. charge) in each DOM i and time-bin τ for a set of energy depositions E_k at positions $\vec{y}_k$: the energy losses are multiplied with the response matrix $\Lambda^{\mu/c}_{i\tau k}$, which describes the seen photons in DOM i and time-bin τ caused by a loss of $E_k = 1\,\mathrm{GeV}$ at $\vec{y}$.

The proportionality constant $\Lambda^{\mu/c}_{itk}(\vec{x}_i, \vec{y}_k, \theta_k, \phi_k, \Delta t_\tau)$ between the E_k and $\mu_{i\tau}$ gives the expected charge per deposited energy in PE/GeV and depends on several parameters:

1. the position of the DOM $\vec{x}_i$,
2. the position of the energy deposition $\vec{y}_k$,
3. the time difference between the energy deposition and the photon detection,
4. the energy deposition being cascade- or muon-like (indicated by the indices μ and c),
5. the *direction* of the energy deposition given by the spherical coordinates θ_k and ϕ_k, which are described in the following.

The values of $\Lambda^{\mu/c}_{itk}(\vec{x}_i, \vec{y}_k, \theta_k, \phi_k, \Delta t_\tau)$ are read from large photon tables, that were splined to obtain a continuous value for $\Lambda^{\mu/c}_{itk}(\vec{x}_i, \vec{y}_k, \theta_k, \phi_k, \Delta t)$ in all of the parameters. To obtain these tables, cascades and short segments of minimum-ionizing muons are generated on a grid in all of the above mentioned parameters and the resulting photons are propagated towards

the IceCube and DeepCore DOMs via photon simulations. Thus, the expected number of photons includes a simple model of the local ice properties. The parameters θ_k and ϕ_k describe the orientation of these cascades and muon segments in spherical coordinates. The probability density functions (PDFs) used for the photon arrival times and the total number of photons are described in more detail in Appendix D.

Note that the matrix Λ, expressed by a combined row-index $\{i,\tau\}$ and a column-index k, is not quadratic and typically features substantially more rows ($\sim M$) than columns ($= K$). As a result, varying the energy losses E_k does not allow to produce every set of predicted photons $\mu_{i\tau}$. More specifically, the matrix Λ can not be inverted to obtain energy losses that produce precisely the observed number of photons $\mu_{i\tau} = N_{i\tau}$.

Moreover, the time difference Δt_τ used in the likelihood is binned, where the number of bins T_i and the bin edges are DOM-dependent. The binning is based on the number of observed photons in each DOM. To do this, for every DOM the algorithm starts at the beginning of the read-out window with $\tau = 1$. Then, the time is increased until the observed charge of all pulses included in the resulting bin exceeds a predefined value $n_{\text{PhotonsPerBin}}$. At this point, the bin is terminated and the next bin $\tau + 1$ is started. This way, the binning is determined purely from the observed number of photons, while the expected number of photons is calculated for this binning.

Thus, the Millipede likelihood maps a set of energy losses given by $\{E_k, \vec{y}_k, t_k, \theta_k, \phi_k\}$ to a single likelihood value:

$$\text{LLH}_{\text{mp}} : \{E_k, \vec{y}_k, t_k, \theta_k, \phi_k\} \mapsto \text{LLH}. \tag{5.3}$$

Such set of energy losses $\{E_k, \vec{y}_k, t_k, \theta_k, \phi_k\}$ is called *internal hypothesis/parameters* in the following.

5.2.2 Energy Optimization within Millipede

One of the auxiliary features of the Millipede likelihood is the internal optimization of the energy losses E_k for a given internal hypothesis $\{E_k, \vec{y}_k, t_k, \theta_k, \phi_k\}$. Thus, the energies are removed from the minimization problem, simplifying the mapping from Equation 5.3 to

$$\text{LLH}^E_{\text{mp}} : \{\vec{y}_k, t_k, \theta_k, \phi_k\} \mapsto \text{LLH}, E_k. \tag{5.4}$$

Here, the energy losses E_k are not passed to the likelihood function, but internally optimized and returned as K additional, positive values from the Millipede framework.

Note that this internal minimization introduces a two-layer optimization process: While some parameters are minimized by an external minimizer, the energy losses E_k are optimized internally for each call of the likelihood. In most cases, such layered optimization slows down an optimization process. However, the Millipede framework exploits the fact that most of the time for calling LLH_{mp} is spend on evaluating the matrix elements $\Lambda_{itk}(\vec{x}_i, \vec{y}_k, \Delta t_\tau)$ from the photon tables. Since these are independent of the energy losses E_k (per construction, as they are derived as 'observed photons **per** unit energy'), several combinations of energies E_k can be tested quickly, once the matrix Λ is generated for a given hypothesis $\{\vec{y}_k, t_k, \theta_k, \phi_k\}$.

Additionally, the internal minimization of the energy depositions is a well-known problem from medical imaging. As such, specialized optimization methods exist for such problems, like *Preconditioned Conjugate Gradient* methods (PCG), which was implemented into Millipede to accelerate the internal optimization [120]. PCG is a powerful tool allowing to minimize up to ~ 100 energy losses for each call of LLH^E_{mp} in an acceptable time.

However, for a likelihood call of $\mathrm{LLH}^{E}_{\mathrm{mp}}$, the internal energy optimization is the computational bottleneck, requiring substantially more time than any other part of the likelihood.

5.2.3 Implemented Hypotheses

By converting a physics hypothesis $\mathcal{H}$ (called *external hypothesis*) into a set of energy depositions $\{E_k, \vec{y}_k, t_k, \theta_k, \phi_k\}$ (internal hypothesis), it can be mapped to a Millipede likelihood value. If the hypothesis $\mathcal{H}$ is given by the set of parameters $\{p_s\}$ with $s = 1, ..., S$, the mapping $P : \{p_s\} \mapsto \{E_k, \vec{y}_k, t_k, \theta_k, \phi_k\}$, is called *parametrization*.

If such a parametrization exist, a hypothesis can be reconstructed by minimizing the Millipede likelihood with respect to the physics parameters $\{p_s\}$ of the hypothesis.

At the beginning of this work, three parametrizations $P : \{p_s\} \mapsto \{E_k, \vec{y}_k, t_k, \theta_k, \phi_k\}$ were already implemented in Millipede. Each parametrization describes a physical hypothesis and one of the experimental signatures shown in Figure 3.7:

- **Monopod**: The Monopod parametrization is the most simple parametrization possible. It consists of only one cascade-like energy deposition. Thus, the external parameters p_s are identical to the internal parameters, describing the energy E_k, the vertex $\{\vec{y}_k, t_k\}$ and the direction $\{\theta_k, \phi_k\}$ of the cascade. If the energy is internally optimized, the external hypothesis has six free parameters. Monopod is typically used for fitting cascade-like events in IceCube (electromagnetic and hadronic).
- **Taupede**: The Taupede parametrization fits two cascades that are separated by the distance L and aligned in the direction $\{\theta_k, \phi_k\}$. Thus, it features seven parameters describing the vertex of the first cascade $\{\vec{y}, t\}$, the distance between the cascades L, and the direction of both cascades $\{\theta, \phi\}$. The first and second cascade are inserted at the times $t_1 = t$ and $t_2 = t + L/c$, respectively.
- **MuMillipede**: The MuMillipede parametrization describes a muon moving on an infinite track through the detector. It digests the corresponding muon hypothesis into short muon segments of typically $\sim 15\,\mathrm{m}$ length, which are all aligned with the muon direction. Thus, the external hypothesis consists of the muon direction $\{\theta, \phi\}$ and an arbitrary supporting point of the track $\{\vec{y}, t\}$, which are six free parameters. However, the energy depositions E_k are optimized internally to describe the stochastic losses on the track, which is a time-intense minimization of up to $K \sim 100$ internal parameters.

5.3 The Pegleg Reconstruction

The Pegleg reconstruction was developed in this work to reconstruct events within $E_\nu \sim 5 - 100\,\mathrm{GeV}$. However, the lower bound is only given by the energy threshold of DeepCore, while above $\sim 100\,\mathrm{GeV}$, its performance suffers from the incorrect assumption of a minimum-ionizing muon. The reconstructed hypothesis, the implementation and the performance are discussed in the following.

Note that Pegleg was one of two independent efforts reconstructing the same hypothesis on a similar likelihood. The second, independently developed reconstruction, called *Hybrid Reco*, was earlier used for the reconstruction of low-energy events, but was finally replaced by the reconstruction presented in this work. The relation between Pegleg and HybridReco

is discussed in Appendix E.1.

5.3.1 Reconstructed Hypothesis

Pegleg reconstructs the hypothesis of a finite, minimum-ionizing track with a starting cascade at the primary vertex, that is aligned with the muon direction. Thus, the hypothesis is given by the primary vertex position and time $\{\vec{y}, t\}$, the direction of cascade and muon $\{\theta, \phi\}$, the starting cascade energy E_c and the length of the minimum-ionizing muon track L. As for minimum-ionizing muons in ice, the energy loss is set to $0.22\,\mathrm{GeV/m} \cdot L$. Moreover, the track is split into muon segments of 5 m length, while the total number of segments depends on the track length L.

Although similar to the hypotheses in Section 5.2.3, this hypothesis could not be implemented into Millipede directly for two reasons:

1. The energy losses of the muon track E_k^μ are not optimized. Thus, the hypothesis is a mixture of a free energy parameter E_c and fixed energy losses E_k^μ, which could not be handled in Millipede.
2. The number of energy losses E_k and thus internal parameters changes depending on the length L, which could not be handled by Millipede previously.

These challenges lead to the three-layer optimization process described in Section 5.3.2.

Note that the Monopod and Pegleg hypotheses are identical, if the track length for Pegleg is forced to $L = 0$ and both are run on the modified, charge-independent Millipede likelihood from Section 5.3.3. Thus, the Monopod hypothesis is nested in the Pegleg hypothesis.

The log-likelihood difference between the Pegleg and Monopod fit is called *Particle Identification* (PID):

$$\mathrm{PID} \equiv \mathrm{LLH(Pegleg)} - \mathrm{LLH(Monopod)}. \tag{5.5}$$

It is used to distinguish track- from cascade-like events in the following analysis. For the final level sample, the PID distribution is shown in Figure 4.6 (bottom). More details on the reconstruction method, the seeds and the minimization are added to Appendix E.2.

5.3.2 Three-Layer Optimization in Pegleg

In contrast to the two optimization layers in Millipede, the optimization in Pegleg is done in three layers, as discussed below. It is sketched in Figure 5.3.

First Layer - Outer Minimizer: The outer minimizer uses the Pegleg likelihood function f_{L1}^{PL}as target function of the minimization. It is given by

$$f_{L1}^{\mathrm{PL}} : \{x,y,z,t,\theta,\phi\} \mapsto \mathrm{LLH}, E_c, L. \tag{5.6}$$

Thus, it only has six external parameters, while the optimum cascade energy E_c and the optimum track length L are obtained internally and returned with the corresponding log-likelihood value LLH.

For the outer minimization, the *Multinest* minimizer was used, which is an advanced sampling algorithm, that was originally developed for applications in astrophysics [121]. Multinest was found to efficiently solve the minimization problem, while local search algorithms

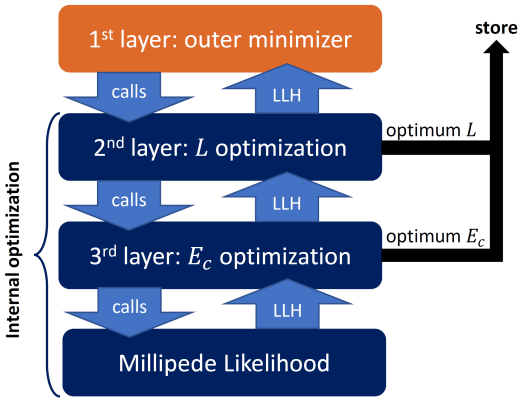


Figure 5.3: Sketch of optimization in three layers, where each layer optimizes some of the parameters of the hypothesis, such that the above layer obtains an optimum value for all parameters optimized below and the corresponding likelihood value.

(like *Migrad* [122], *Simplex* [123] or *L-BFGS-B* [124]) typically failed, due to the large number of side minima in the likelihood. The minimizer settings used for this outer minimizations are listed in Appendix E.2.

Second Layer - Length Optimization: On this layer, the parameters $\{x,y,z,t,\theta,\phi\}$ are fixed by the likelihood call of the outer minimizer. The target function of the minimization on this layer is given by Equation 5.7

$$f_{L2}^{\mathrm{PL}} : L \mapsto \mathrm{LLH}, E_c \Big|_{x,y,z,t,\theta,\phi \text{ fixed}} . \tag{5.7}$$

Thus, the length L is changed, while for each change the cascade energy E_c is internally optimized. Here, changing the length L means adding/removing muon segments to/from the internal hypothesis. Thus, it corresponds to adding/removing columns to/from the detector response matrix Λ, as illustrated in Figure 5.4, while for each change of the matrix Λ, the new likelihood LLH and optimum cascade energy E_c are calculated.

The optimization in L is done as a scan: The algorithm starts with only one column in Λ for the starting cascade. Then, muon segments are successively added as long as this improves the value of LLH. If the likelihood did not improve during the last two added segments, the procedure is stopped and the optimum length L is returned together with the corresponding value for LLH and E_c.

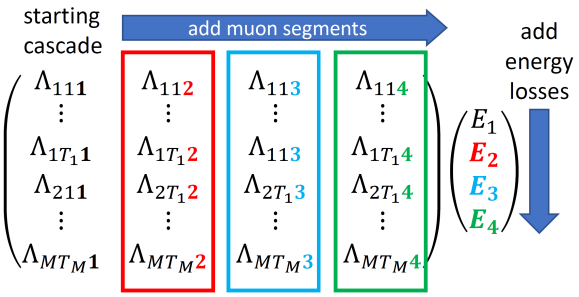


Figure 5.4: Sketch of successively extending the detector response matrix Λ and the energy losses $\vec{E}$ by adding additional muon segments during the length optimization.

This scan in L is computationally efficient, since the table read-outs for Λ dominate the

time needed for a likelihood call. Thus, starting with only a cascade and then subsequently adding up to K muon segments is similarly fast as testing only the hypothesis of K segments directly. Therefore, as soon as the marix Λ for a length L, corresponding to K energy losses, is calculated, estimating all shorter tracks than L comes at no computational cost.

Third Layer - Cascade Energy Optimization: On this layer, the parameters $\{x,y,z,t,\theta,\phi, L\}$ are fixed from the two layers above. Thus, the target function of the minimization is given by Equation 5.8

$$f_{L3}^{\mathrm{PL}} : E_c \mapsto \mathrm{LLH}\Big|_{x,y,z,t,\theta,\phi,L \text{ fixed}} . \tag{5.8}$$

Since the matrix elements $\Lambda_{i\tau k}^{\mu/c}$ are independent of the values of E_k, the energy can be optimized internally. To do this, the muon elements with $k = 2,...K$ are fixed to minimum-ionizing. Then, the expected number of photons caused by muon segments is derived and added to the noise term $\vec{n}$ as shown in Equation 5.9

$$\vec{m} \equiv \vec{n} + \Lambda \cdot (0,E_2, E_3, ..., E_K)^T . \tag{5.9}$$

Thus, the effective noise term $\vec{m}$ includes the photon expectation due to the muon track. Afterwards, the Λ matrix is cut to only its first column, representing the starting cascade. This reduces the optimization problem to just one energy deposition E_c, which is analogous to Monopod with an enhanced noise vector $\vec{m}$ (cf. Section 5.2.3).

The energy E_c can then be obtained internally using PCG (cf. Section 5.2.2). However, for one parameter, running PCG is inefficient, which was designed to optimize a large number of energy depositions at the same time. Instead, the likelihood from Equation 5.1 is simply differentiated with respect to E_c, as shown in Equation 5.10

$$\frac{\mathrm{d}f_{L3}^{\mathrm{PL}}}{\mathrm{d}E_c} = \frac{\mathrm{dLLH}}{\mathrm{d}E_c} = \left(\sum_i^M \sum_\tau^{T_i} \underbrace{\frac{N_i \Lambda_{i\tau 1}}{\Lambda_{i\tau 1} E_c + m_i}}_{\propto (E_c + \text{const.})^{-1}} \right) - \underbrace{\left(\sum_i^M \sum_\tau^{T_i} \Lambda_{i\tau 1} \right)}_{\text{const.}} \stackrel{!}{=} 0, \tag{5.10}$$

where the last equality is required for the optimization, but cannot be solved analytically for E_c. However, note that the first term in Equation 5.10 is strictly decreasing in E_c, while the second term is constant. Thus, there is (at most) one positive value of E_c that satisfies the equation.

To find this value, a simple *Newton method* [125] is applied that uses the second derivative $\frac{\mathrm{d}^2 f_{L3}^{\mathrm{PL}}}{\mathrm{d}^2 E_c}$ to quickly find the optimum value of E_c. It is guaranteed to find the optimum value, if a positive value exists (otherwise chooses $E_c = 0$), and chosen to stop the optimization at a tolerance of $\delta E_c = 10^{-6}\,\mathrm{GeV}$, which is far more accurate than needed in the following. This accuracy is typically reached with $\mathcal{O}(10)$ steps of the Newton method. By seeding the Newton Method with the optimum value of E_c found at the previous call, the number of steps is further decreased to $\mathcal{O}(3)$, such that the time, spent on this optimization, is negligible.

5.3.3 Redefinition of the Millipede Likelihood

For the Pegleg reconstruction, an additional feature was implemented that modifies the Millipede likelihood. It was added due to the findings in Section 4.5.2 that the charge per

DOM is not modelled correctly in the Monte Carlo. Although the mismatch is hardly visible when averaging over all DOMs within an event, it causes some crucial disagreement in the reconstructed quantities, as the Millipede likelihood is strongly charge-dependent.

Note that this disagreement is less relevant when having a large number of hit DOMs and a large charge per DOM, as the average charge per DOM is approximately correct. Thus, the disagreement arises mainly from the stochastical nature of the charges in case of only few photons being observed.

To overcome this charge-dependence, for this work the Millipede likelihood was re-phrased in a charge-independent way:

Instead of determining the time-binning by aggregating pulses up to a threshold value $n_{\text{PhotonsPerBin}}$ (cf. Section 5.2.1), the first observed pulse terminates the time-bin. The charge of the pulse is ignored and a value of $N_{i\tau} = 1$ is used as the observed number of photons in the bin τ. Afterwards, the DOM is treated as *dead* for the following $\delta t_{\text{dead}} = 45\,\text{ns}$. This means, the time-window is completely removed from the likelihood, since the subsequent DOM behavior after the photon detection is not expected to be well-described by Monte Carlo. After δt_{Dead}, the next bin is started allowing for a observation of late photons. In case another photon is observed after δt_{Dead}, the procedure above is repeated.

This way, any unexpected behavior of the DOM within δt_{Dead} after the photon detection is completely ignored. Moreover, the charges assigned to each pulse in data and MC are removed from the likelihood, allowing only for integer values in the number of observed photons. The value of $\delta t_{\text{dead}} = 45\,\text{ns}$ is obtained from data-MC comparisons and described in more detail in Appendix E.3.

Although the change to the likelihood seems large, the corresponding resolutions stay almost unchanged. This is due to the fact that most of the information for the angular reconstruction is contained in the arrival times of the first observed photons, which are kept for the charge-independent likelihood. The energy resolution even improves slightly, which is due to the removed randomness induced by the assignment of charges to the photons using the SPE template, which is more relevant than the loss of information due to removed pulses.

After these modifications, the reconstruction was found to show no disagreement between data and MC and the charge-independent likelihood was adapted by all analyses based on the GRECO sample [61, 116].

5.3.4 Reconstruction Performance

The final level distributions of the reconstructed energy, zenith-angle and PID are shown in Figure 4.6. In addition, the zenith-angle and neutrino energy resolution of Pegleg are discussed in the following. For producing the following resolution plots, the final zenith and energy cuts on the GRECO sample were removed to not bias the resolutions by a cut on the reconstructed quantities.

In Figure 5.5 (top), the zenith-angle resolution is shown for muon neutrinos and anti-muon neutrinos separately for different true neutrino energies E_ν. As one would expect, the zenith-angle resolution increases with energy, especially between 10 and 40 GeV, due to the increased amount of observed light. Moreover, the zenith-angle resolution for muon anti-neutrinos is slightly better than for muon neutrinos. This is due to the inelasticity distribution in Figure 2.4: The good angular resolution is primarily driven by the outgoing

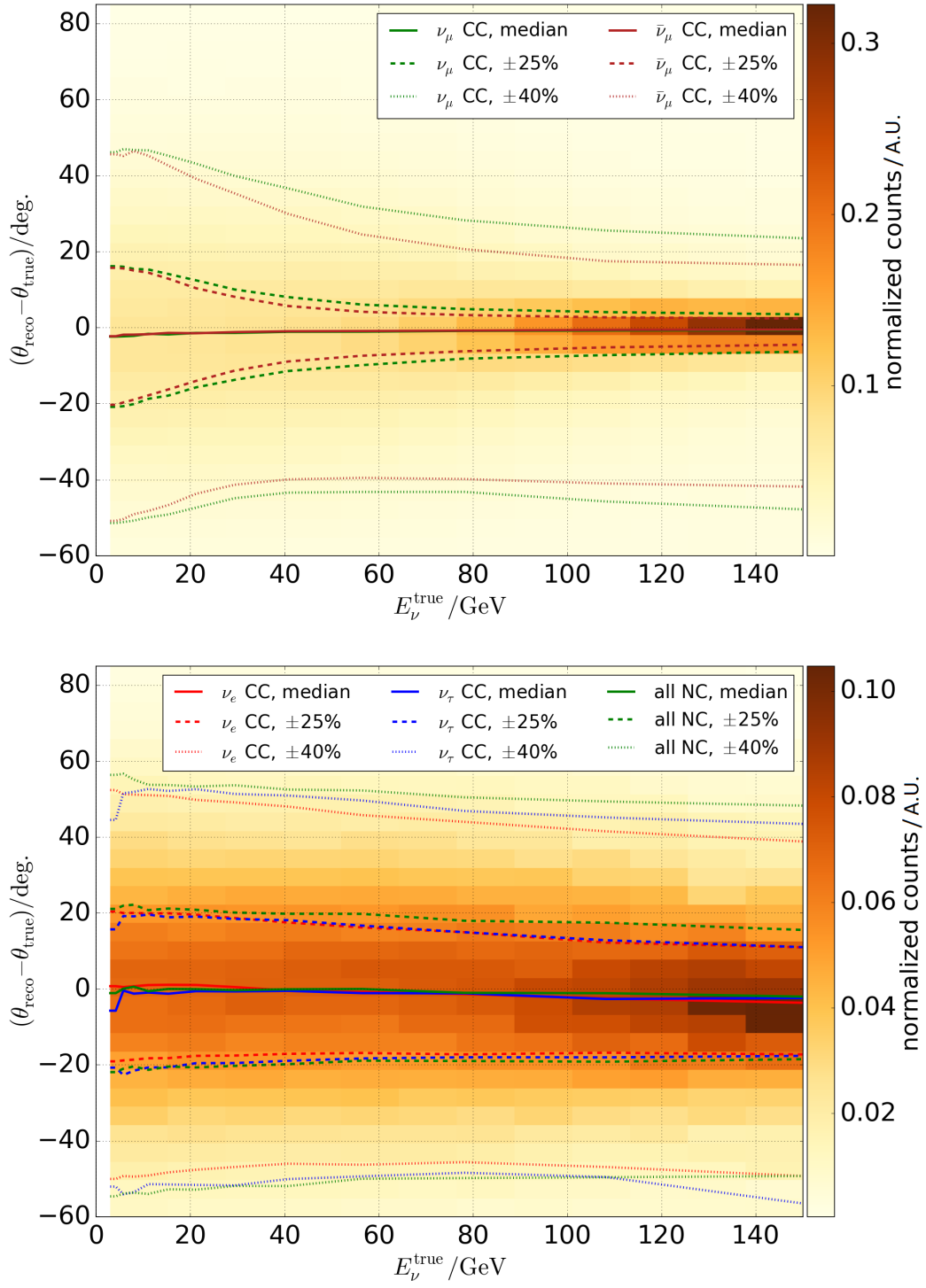


Figure 5.5: Energy-dependent zenith-angle resolution shown for CC muon neutrinos and anti-neutrinos (top) and for the other flavors, combining neutrinos and anti-neutrinos (bottom); solid lines give the median angular error in each vertical slice of true energy, while the dashed and dotted lines enclose the central 50% and 80% quantiles; the background colormap shows the underlying, colwise-normalized two-dimensional probability distribution for CC muon neutrinos (top) and CC electron neutrinos (bottom).

muon, causing a long lever arm for the angular reconstruction, while the angular resolution for cascades is poor, as discussed in Section 3.3. Since anti-neutrinos on average pass more energy to the muon (small value of inelasticity y), they are on average more track-like than neutrinos at the same energy and thus feature a better angular resolution.

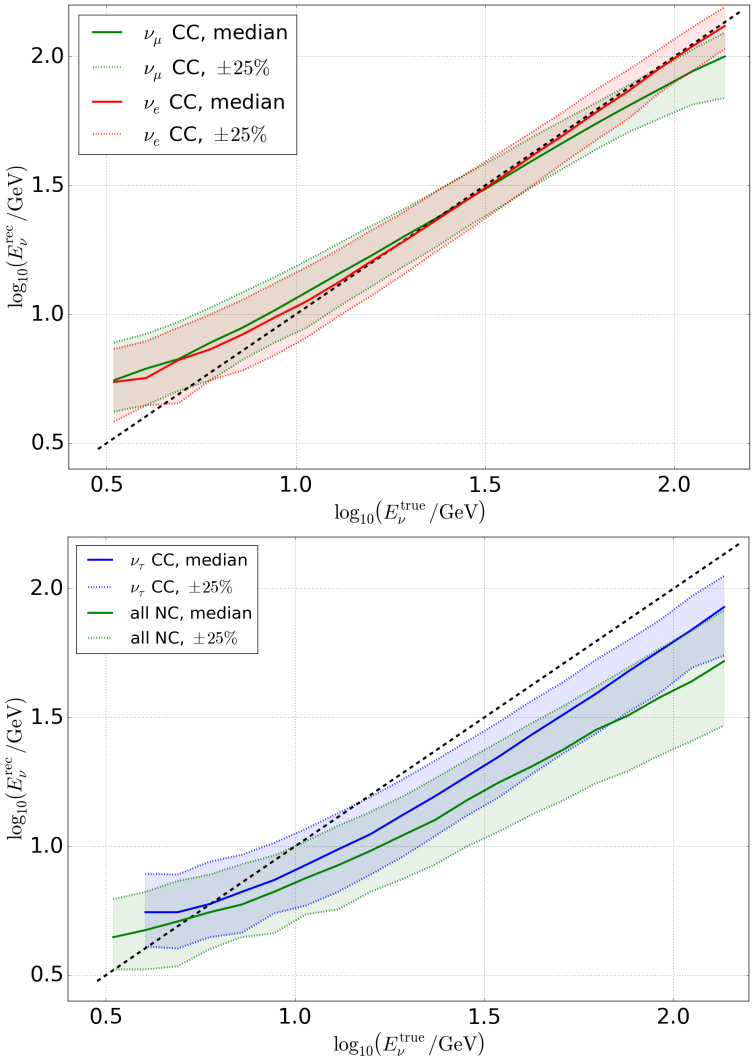


Figure 5.6: Neutrino energy resolution with Pegleg for CC electron and muon neutrinos (top) and CC tau and NC events (bottom); the solid lines indicates the vertical median for each slice in true energy, while the dashed line and the shaded areas indicate the central 50% quantiles.

In Figure 5.5 (bottom), the resolutions are shown for CC electron neutrino, CC tau neutrino and NC interactions, which are the same for neutrinos and anti-neutrinos and therefore not shown separately. As one can see, the resolution is clearly worse than for the muon neutrinos, since no lever arm is obtained from an outgoing muon. Additionally, the resolution for CC electron neutrinos is slightly better than for CC tau neutrinos.

For NC events, the resolution in Figure 5.5 is worse than for CC electron and tau neutrinos. This can be explained by the invisible energy of the escaping neutrino in NC interactions (cf. Section 3.3.2), and the cascade being of hadronic origin: Since the Cherenkov light stored in the reconstruction tables is obtained from the simulation of electromagnetic cascades, the hypothesis is slightly incorrect in case of hadronic cascades.

Note that all flavors show a slight trend towards negative values, especially for the central 80% quantiles. This is mostly due to the zenith-dependent selection efficiency of the data sample. To illustrate this, the error is shown depending on the true zenith-angle in

Appendix E.4.

In Figure 5.6, the neutrino energy resolution is shown for CC muon and electron neutrino events (top) and CC tau and NC neutrino events (bottom). As one would expect, the energy resolution is the best for CC muon and electron neutrino events, where no invisible energy arises from escaping neutrinos. The drop in reconstructed energy for high-energy muon neutrinos is due to the resulting muon leaving the IceCube detector volume. Moreover, the assumption of a minimum-ionizing track is inaccurate leading to an underestimation of the lost energy per unit of track length.

For CC tau and NC events, the deficit from escaping neutrinos is clearly visible. This deficit is typically larger in NC events, leading not only to a reduced median in reconstructed energy, but also to a larger variance.

In Figure 5.7 (top), the zenith-angle resolution for CC muon neutrinos is compared to the resolution for SANTA. Note that also the Pegleg resolution improves compared to previous figures, because only events that are reconstructable with SANTA are used here, which is a strong cut on the event quality. The SANTA resolution is clearly worse than the one obtained for Pegleg on the same set of events. However, it is comparable to the Pegleg resolution from Figure 5.5. The fraction of events that can be reconstructed with SANTA is shown as solid, black line in Figure 5.7 (top) to be read from the right vertical axis. It decreases to only a few percent when going below 15 GeV, which makes it unusable for the GRECO sample.

In Figure 5.7 (bottom), the Pegleg energy resolution is compared to LEERA. Pegleg provides a similar resolution at the lowest energies and an improved resolution at higher energies. Note that the constant bias towards higher energies observed in LEERA does not reduce the reconstruction performance, since it could simply be removed by multiplying a constant correction factor to all of the reconstructed neutrino energies.

For CC electron neutrino events, the comparison to SANTA/LEERA is added to Appendix E.4, leading to similar conclusions as for CC muon neutrinos.

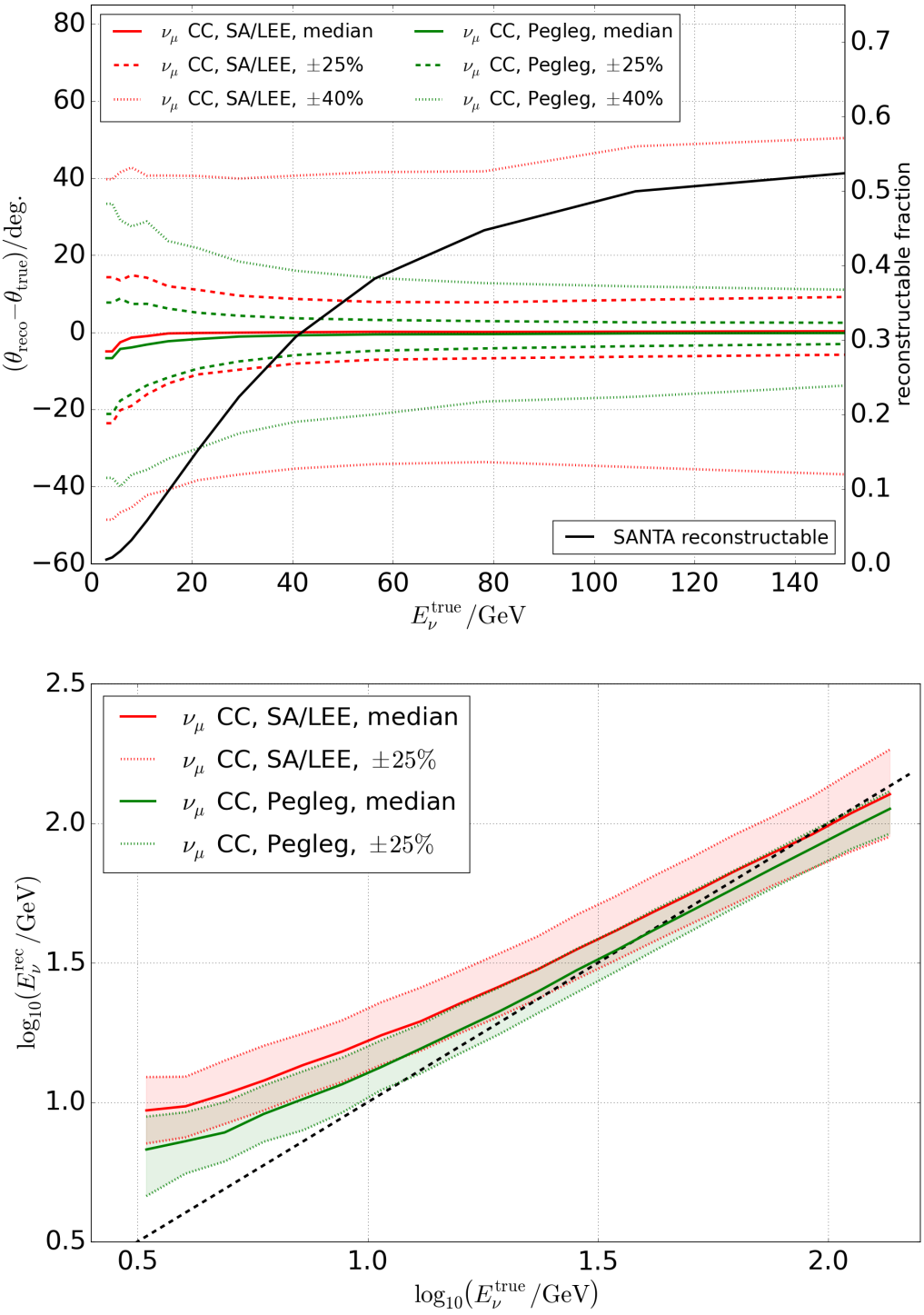


Figure 5.7: Comparison of zenith-angle (top) and neutrino energy resolution (bottom) for Pegleg and SANTA/LEERA (*SA/LEE*), discussed in Section 5.1; in the top plot, the black, solid line, to be read from the right vertical axis, indicates the energy-dependent fraction of the GRECO CC muon neutrino sample that is reconstructable with SANTA.

6 Likelihood Analysis of the Neutrino Mass Ordering

In this work, the Neutrino Mass Ordering (NMO) is fitted in a binned, *maximum-likelihood* analysis on the final level GRECO sample. All of the analysis, presented in the following chapter, was developed *blindly*, i.e. without running the analysis or relevant parts of it on experimental data before finalizing the analysis procedure. The full analysis chain and the subsequent *unblinding procedure* were supervised and reviewed by the IceCube Collaboration. The analysis presented in the following and the results from Chapter 8 were published in [126], [127] and [116], in combination with a second analysis of the Neutrino Mass Ordering, that was developed independently and in parallel by Steven Wren [128].

The idea of the following analysis is to measure matter effects in the oscillation pattern from Figure 2.14. Since the true neutrino energy, zenith-angle and flavor are not accessible, the analysis uses the corresponding reconstructed observables. To do this, the reconstructed energy, zenith-angle and PID from the Pegleg reconstruction presented in Chapter 5 are used. Note that here the PID is no direct proxy for the neutrino flavor, but instead distinguishes between track-like and cascade-like events (cf. Section 5.3.1).

As mentioned before, IceCube is not capable to separate neutrinos and anti-neutrinos on an event-wise level. Therefore, no attempt is made to distinguish these within the analysis binning. Instead, the observed signature is a combination of the neutrino and anti-neutrino channels, shown in Figure 2.14.

6.1 Analysis Principle

In the following analysis, only upgoing events with $\cos(\theta_\nu^{\text{reco}}) \in [-1,0]$ are used, because downgoing events do not oscillate in this energy range (cf. Figure 2.14). In principle, downgoing events could be added to constrain nuisance parameters in such *off-signal* region. However, the gain in sensitivity was found to be small, while including the downgoing region strongly increases the atmospheric muon contamination. Instead, this analysis extends to higher energies than previous DeepCore analyses, which also allows to constrain nuisance parameters outside the oscillation regime [60].

The three-dimensional distribution of energy, zenith-angle and PID, that is used in the likelihood analysis, is shown in Figure 6.1, where all contributions from neutrinos, atmospheric muons and triggered noise are combined. The PID bins are labelled as *track-bin*, *transition-bin* and *cascade-bin*. The track-bin is strongly dominated by CC muon neutrinos, while the other neutrino flavors dominate in the cascade-bin. The transition-bin contains events, that are neither clear tracks nor clear cascades.

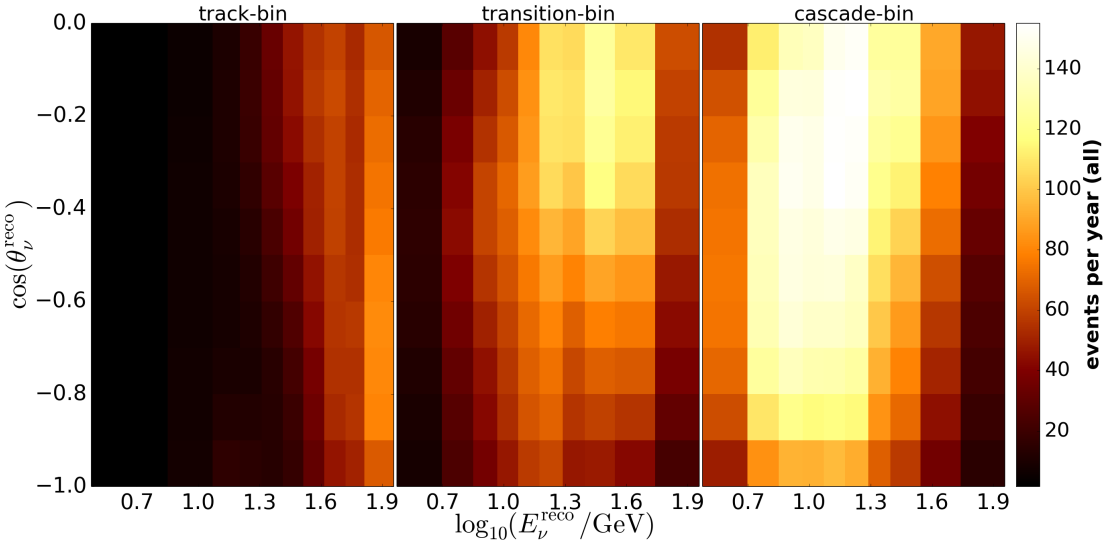


Figure 6.1: Three-dimensional *analysis distribution* of reconstructed energy, zenith-angle and PID for Monte Carlo simulations; this distributions is fitted in the likelihood analysis; the shown distribution is the sum of CC and NC neutrino interactions of all flavors, atmospheric muons and triggered noise, while the individual distributions are shown in Figure 6.3 and 6.4 (bottom) (pre-published in [127]).

The bin-edges, chosen for this three-dimensional *analysis distribution*, are given by:

- PID: $\{-\infty, 1.0, 8.0, \infty\}$ (defined by Equation 5.5)
- $\cos(\theta_\nu^{\mathrm{reco}})$: 10 equidistant bins between -1 and 0
- $E_\nu^{\mathrm{reco}}/\mathrm{GeV}$: 10 non-equidistant bins between 3 GeV and 90.5 GeV, where the bin-edges are chosen differently for the track-bin and the other two PID bins, which share the same binning.

Note that the reconstructed energy and zenith ranges correspond to the cuts on energy and zenith-angle introduced in Section 4.4.3.

The energy bin-edges in the track-bin are different from the ones used in the other two PID bins. The binning for the cascade- and transition-bin were obtained by trying to obtain sufficient statistics at all energies, while at the same time having a small binning at $\sim 10\,\mathrm{GeV}$, which is the most interesting regime for an NMO measurement. For the track-bin, the bin-edges were manually adjusted, as the PID is highly correlated with the neutrino energy. As a result, the track-bin is dominated by high-energy events, because events at low energies are unlikely to give a strong separation between the track- and cascade-hypothesis. To account for this, the edges are shifted to follow the statistics with larger bins at low energies and smaller bins at high energies (cf. Figure 6.1).

The resulting signature expected from the Neutrino Mass Ordering (NMO) is shown in Figure 6.2. Here, the expected pulls are shown for testing Normal Ordering (NO) in case the true ordering is inverted (IO), which is defined as $(n_{ijk}(\mathcal{H}_{\mathrm{IO}}) - n_{ijk}(\mathcal{H}_{\mathrm{NO}}))/\sqrt{n_{ijk}(\mathcal{H}_{\mathrm{NO}})}$ for the bin-content n_{ijk}. As one can see, the signature is primarily obtained in the track-like PID bin for vertically upgoing events. Moreover, the signature is reversed for cascade-like events, as expected from Figure 2.14. In contrast, horizontal events carry almost no information about the NMO, as they are nearly unaffected by neutrino oscillations.

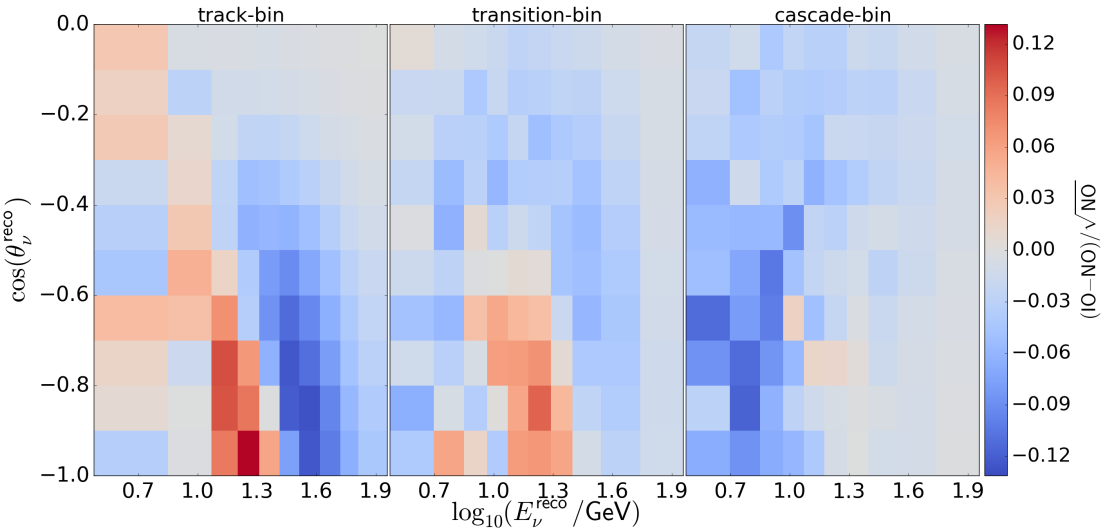


Figure 6.2: Expected signature from the Neutrino Mass Ordering in the analysis distribution from Figure 6.1: the signature is shown in terms of the *expected pulls* defined by $(n_{ijk}(\mathcal{H}_{\text{IO}}) - n_{ijk}(\mathcal{H}_{\text{NO}}))/\sqrt{n_{ijk}(\mathcal{H}_{\text{NO}})}$ with n_{ijk} being the bin-content of the bin i,j,k, which is a commonly used measure of the NMO signature [7] (pre-published in [127]).

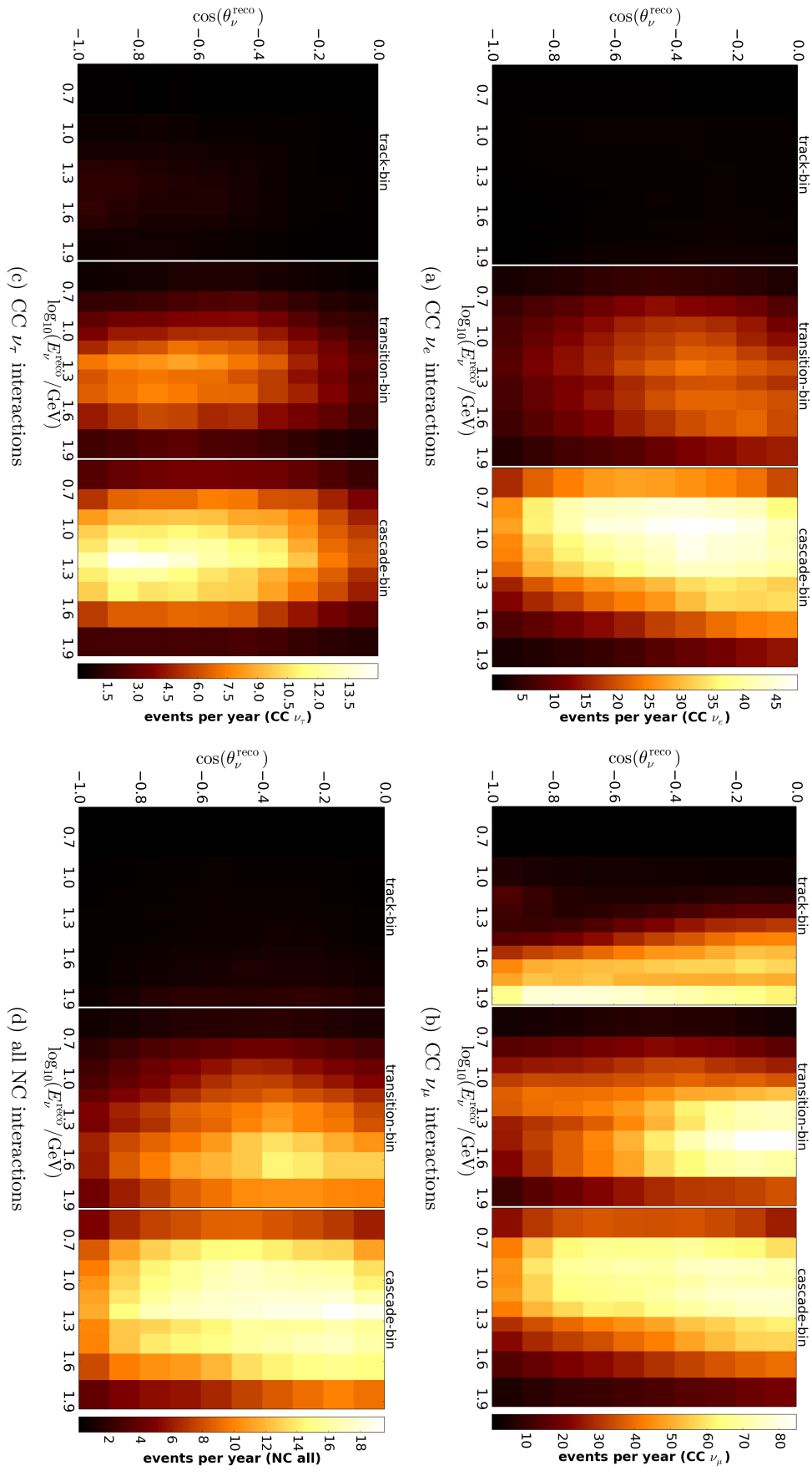
Besides the combined analysis distribution, the individual distributions for CC muon neutrino, CC electron neutrinos, CC tau neutrinos and NC neutrinos are shown in Figure 6.3. The background distributions for atmospheric muons and triggered noise are shown in Figure 6.4 (bottom). Note that these distributions were generated using *Kernel Density Estimation* (KDE), which is described in more detail in the following Section 6.2.

In the following likelihood analysis, these distributions are used for a *template-fit* of the corresponding data distribution. In this template-fit, the combined distribution in Figure 6.1 is compared to the data histogram in a maximum-likelihood method. This likelihood includes the expected Poissonian fluctuations on the bin-content in data and the uncertainties on the MC template due to the limited amount of Monte Carlo events. Due to the *Kernel Density Estimation* (KDE) method, used to derive the MC distributions, the MC uncertainties are strongly reduced compared to histograms, as described in Section 6.2. Thus, the likelihood is clearly dominated by the Poissonian fluctuations, while the MC uncertainties are only a minor correction. The likelihood is described in more detail in Section 6.4.

In the template fit, systematic uncertainties on the neutrino flux, the neutrino cross-sections, the detector-response and the oscillation parameters are included as nuisance parameters. To do this, their impact on the analysis distribution in Figure 6.1 is parametrized and for each systematic, a nuisance parameters is inserted into the template-fit.

This way, the unknown value of the systematic uncertainties is fitted simultaneously with the NMO, which reduces the sensitivity to the NMO, but includes the uncertainty on these parameters into the fit result. A list of all systematic parameters is presented in Section 6.3, while the optimization of the likelihood with respect to these parameters is described in Section 6.5.

The likelihood is optimized with respect to all nuisance parameters for NO and IO, separately. The difference of the two negative log-likelihood values $\Delta\text{LLH}_{\text{NO}-\text{IO}}$ is then used as a test statistic to obtain the experimental result in Chapter 8. Moreover, it is used in two independent methods in Section 7.4 to derive the corresponding sensitivities. Finally, it is



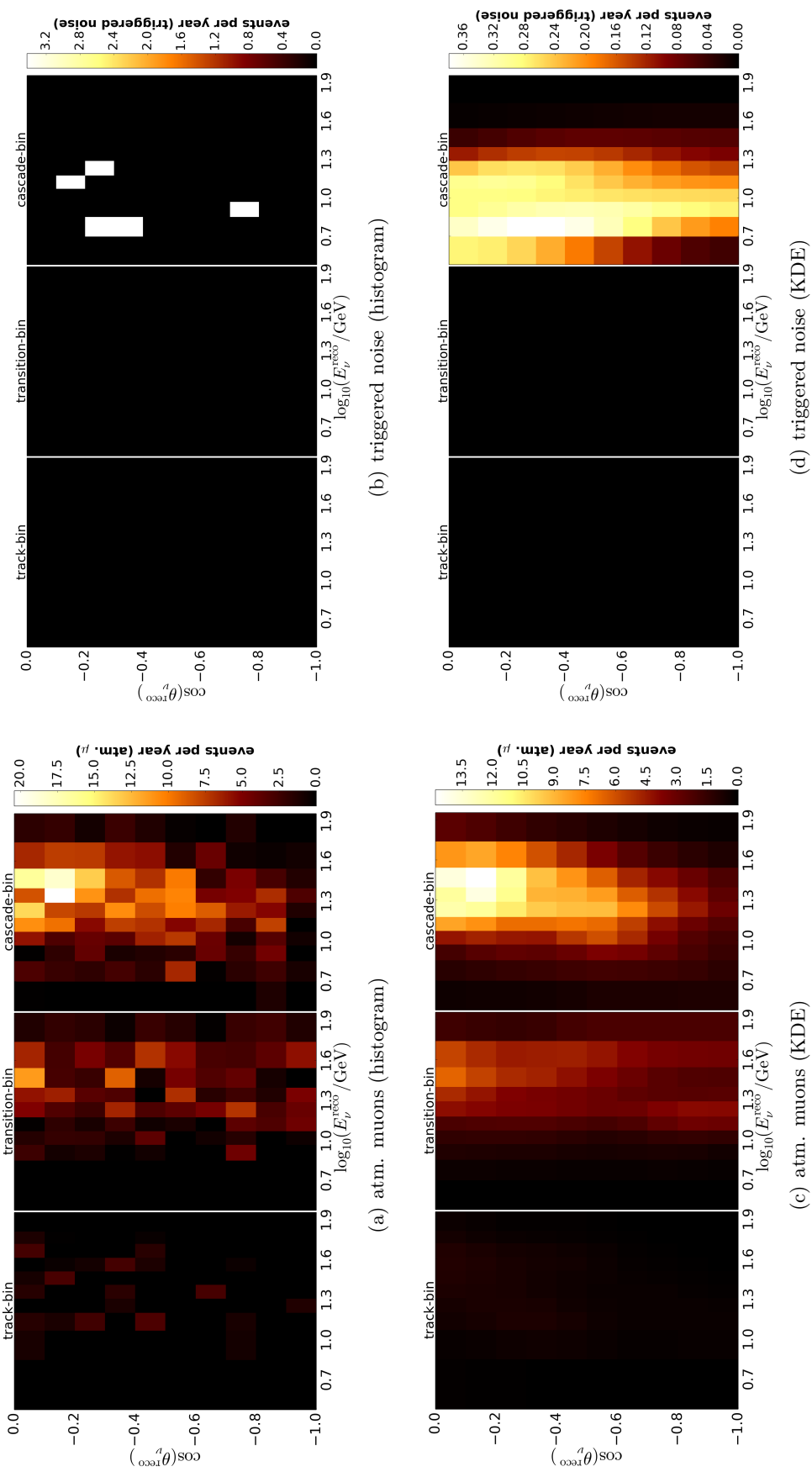
(a) CC ν_e interactions

(b) CC ν_μ interactions

(c) CC ν_τ interactions

(d) all NC interactions

Figure 6.3: Three-dimensional *analysis distribution* for CC ν_μ, CC ν_e, CC ν_τ and NC neutrino events; note that the track-like PID bin is populated mostly by CC ν_μ events, while all remaining neutrino events populate mostly the cascade-like PID bin.



(a) atm. muons (histogram)

(b) triggered noise (histogram)

(c) atm. muons (KDE)

(d) triggered noise (KDE)

Figure 6.4: Comparison of the three-dimensional *analysis distribution* for atmospheric muons (left) and triggered noise (right) obtained by histograms (top) and Kernel Density Estimation (bottom); note the clearly visible Monte Carlo fluctuations in case of histograms that are smoothed by the KDE method.

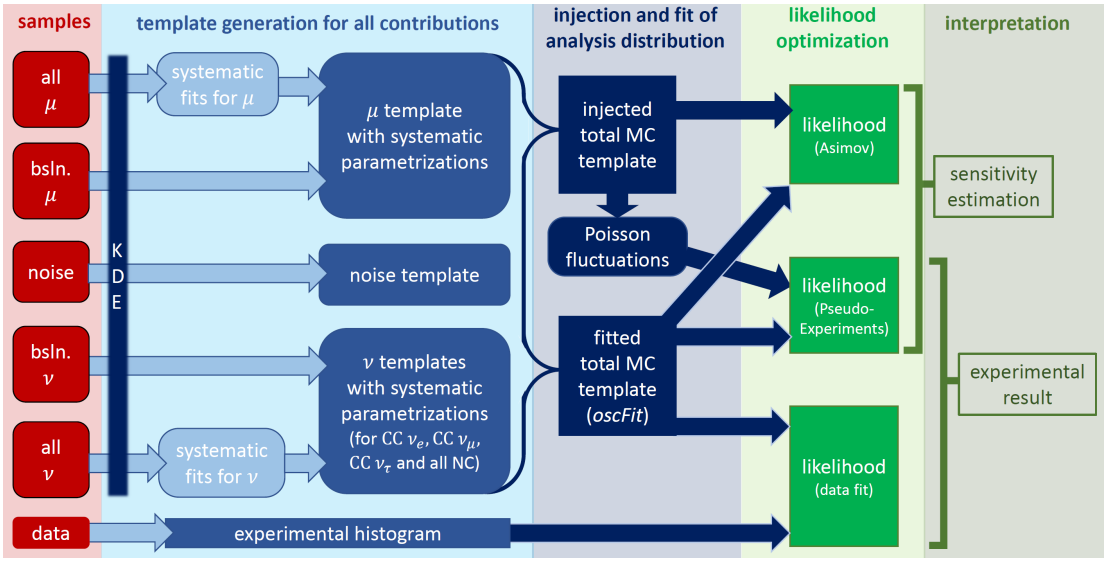


Figure 6.5: Illustration of the analysis procedure: first, KDEs are used for all MC samples to derive the three-dimensional *analysis distribution* for neutrinos (*bsln* ν), atmospheric muons (*bsln* μ) and *noise* – these distributions are used as templates in the following fit; second, the parametrizations of the detector systematics are derived from the baseline and all systematic samples (*all* ν/μ), as described in Section 6.3.5; third, these templates and parametrizations are then used to fit an injected distribution either from MC (*sensitivity estimation*) or data (*experimental result*) in a maximum-likelihood method.

used in Section 7.6, to estimate the impact of each systematic on the NMO result.

The analysis procedure is summarized in Figure 6.5, while the individual steps of this diagram are described in more detail in the following sections.

6.2 Template Generation Using Kernel Density Estimation

From a discrete set of MC events, the templates in Figure 6.4 and 6.3 can be obtained using various methods:

A prominent method to generate templates is the *histogram*. The histogram is a *cut-and-count* method, that sums the weights of all events within one patch of the parameter space, i.e. bin, and takes the resulting sum as predicted bin-content y_i. In case of small bins, this leads to large relative errors on the predicted bin-content, as the prediction is limited by the statistical power of the events contained within the bin.

A more advanced method for the generation of templates is *Kernel Density Estimation* (KDE). In contrast to histograms, it is an unbinned method, that calculates a continuous probability density function (PDF) instead of an aggregated probability. The PDF can then be integrated over any chosen interval, i.e. bin, to obtain a probability distribution. This allows for an arbitrarily small binning without loss in predictive power [129]. For low statistics and most practical applications, KDEs are known to outperform histograms clearly in an accurate estimation of PDFs [130].

In this work, a *weighted, adaptive KDE* is used to reduce the statistical error of the Monte Carlo templates. The method is described in more detail in the following section, where the

terms *weighted* and *adaptive* are also defined.

Since the neutrino MC provides large statistics for estimating the templates in Figure 6.3, the differences between KDEs and histograms are small. In contrast, for atmospheric muons and triggered noise the MC statistics is small, leading to large differences between histogram and KDE templates. In Figure 6.4, the resulting templates for histograms (top) and KDEs (bottom) are compared, where the statistical fluctuations for histograms are clearly visible.

6.2.1 Multidimensional Adaptive Kernel Density Estimation

The weighted, adaptive KDE method, that is used for the template generation, is based on [131]. The method can be applied to an arbitrary number of dimensions. However in the following, the distributions are assumed to be two-dimensional ($d = 2$) in reconstructed energy and zenith-angle, which are calculated separately for each PID bin (cf. Figure 6.4).

At any point $\vec{y} = (y_1, y_2)^T$, the value of the KDE, generated from a set of N Monte Carlo events at $\vec{x}_i = (x_{i,1}, x_{i,2})^T$, is defined by Equation 6.1 [131]

$$\mathrm{KDE}(\vec{y}) = \sum_{i=1}^{N} w_i D_i \exp\left(-\frac{1}{2}\left(\vec{x}_i - \vec{y}\right)^T C^{-1} \left(\vec{x}_i - \vec{y}\right) \frac{1}{(\lambda_i h)^d}\right), \tag{6.1}$$

where w_i is the weight of the MC event at $\vec{x}_i$ and C^{-1} is the inverted, weighted covariance matrix of the two-dimensional MC data $\{\vec{x}_i\}$. The term *weighted KDE* refers to this weighting of the input data.

As shown in Equation 6.1, the KDE is calculated by convolving each data point $\vec{x}_i$ with a two-dimensional Gaussian, so-called *kernel-function* or *kernel*. The integral of the kernel is given by the event weight w_i. Moreover, the width of the Gaussian, so-called *bandwidth*, is given by the covariance matrix C and two scaling factors h and λ_i.

The quantities λ_i, h and D_i are defined by [131]:

$$D_i = \sqrt{\frac{\det(C^{-1})}{2\pi h \lambda_i}} \qquad h = \left(\frac{N(d+2)}{4}\right)^{\frac{-1}{d+4}} \quad (\textit{Silverman's-Rule}) \tag{6.2}$$

$$\lambda_i = \left(\frac{\mathrm{KDE}(\vec{x}_i)|_{\lambda_j=1,\forall j}}{g}\right)^{-\alpha} \qquad g = \exp\left(\frac{1}{N}\sum_{i=1}^{N} \log\left(\mathrm{KDE}(\vec{x}_i)|_{\lambda_j=1,\forall j}\right)\right) \tag{6.3}$$

Thus, the bandwidth is given by the global scaling factor h and an event-wise correction λ_i. The global factor h is estimated using *Silverman's-Rule* from [132], which is a common, heuristic choice for the global bandwidth in case of a Gaussian kernel-function.

The idea of an event-wise correction λ_i is motivated by distributions with pronounced high- and low-statistics regions:

In high-statistics regions, small features in the distribution may become very significant. There, a small bandwidth is necessary to reproduce these feature in the corresponding KDE. In low-statistics regions, small features may not be resolvable with the available statistics. Instead, the bandwidth should be chosen to be large to remove statistical features from individual MC events. Thus, an ideal bandwidth depends on the statistics in the vicinity of the event. The term *adaptive KDE* refers to this adaption of the bandwidth for every MC event [129].

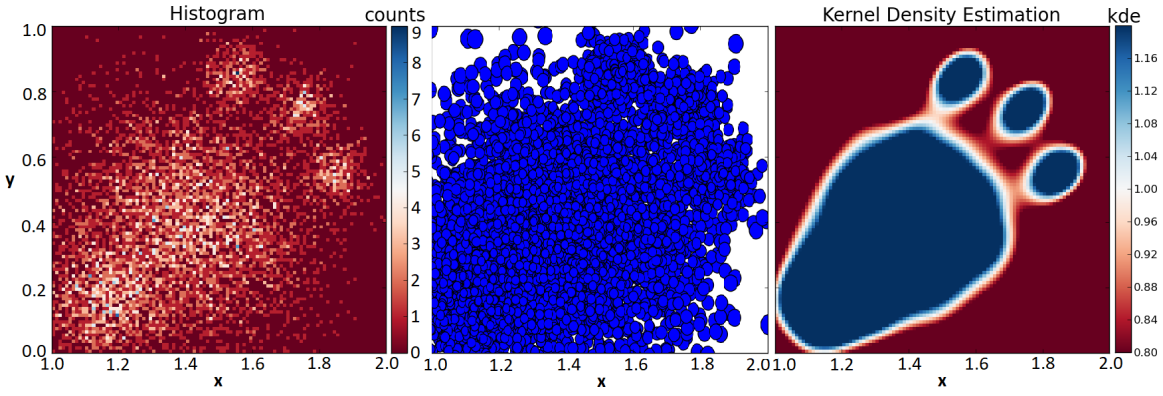


Figure 6.6: Sketch of the procedure of Kernel Density Estimation (KDE); left: histogram of two-dimensional data pulled from an arbitrarily chosen, two-dimensional PDF; center: width of the individual kernels for each data point, obtained from evaluating a non-adaptive KDE; the size of each circle is proportional to the bandwidth factor λ_i of the data point $\vec{x}_i$; right: resulting KDE, given by the weighted sum of all kernels.

To do this, the KDE is evaluated in two iterations: In the first iteration, the KDE is generated with $\lambda_j = 1, \forall j$, such that the bandwidth is h for all events (non-adaptive). Then the KDE is evaluated at each data point $\vec{y} = \vec{x}_i$ and the resulting density $\mathrm{KDE}(\vec{x}_i)|_{\lambda_j=1,\forall j}$ is used to estimate the statistics in the vicinity of the event $\vec{x}_i$. In the second iteration, this estimate is used to calculate the adaptive KDE with the bandwidth factor λ_i for every MC event i.

Finally, the parameter α in Equation 6.3 determines the extent to which the kernel width is following the statistical power of the data points $\vec{x}_i$. The value chosen in this work is $\alpha = 0.3$, motivated by [129]. The general idea of KDEs is sketched in Figure 6.6 for an arbitrarily chosen PDF.

Before this work, this KDE method was implemented in a *Python* framework by Sebastian Schoenen [133]. For this work it was re-implemented in *C++* and *CUDA* for several reasons, discussed in Appendix F.1. All three implementations (Python, C++, CUDA) were tested to give the same results for one and two-dimensional KDEs, while the C++ implementation is used for all calculations in this work.

6.2.2 Parametrization Boundaries and Renormalization

To use KDEs for the template generation, two additional steps are required:

First, a boundary treatment must be applied, since KDEs tend do underestimate the PDF at its parametric boundary (e.g. $\cos(\theta_\nu^{\mathrm{reco}}) = -1$). To account for this, the KDE is reflected at all boundaries, such that no density is *flowing out* of the parameter space. This reflection method is commonly used for parametric boundaries [134] and was investigated in detail by Eric per Vogel [135].

Second, the KDE is integrated over the extent of every bin from the binning in Section 6.1. The resulting distribution obtained from integrating the two-dimensional KDE is then normalized to the sum of the weights of all MC events. The boundary treatment and the renormalization, are described in more detail in Appendix F.2.

Note that both, histograms and KDEs, converge to the true probability distribution for $N \to \infty$. In that case, the Gaussian kernels become infinitely small, approximating a delta-distribution with normalization w_i. Thus, solving the integral is equivalent to summing the weights of all events within that bin, which is identical to a histogram.

A comparison of templates, generated with histograms and KDEs, is shown in Figure 6.4 for atmospheric muons and triggered noise. For histograms, the statistical fluctuations from limited MC are clearly visible, while the KDE returns a smooth template estimate. Additionally, a comparison of the performance for histograms and KDEs is added to Appendix F.3 for PDFs similar to those used in this work.

The uncertainties on the templates, generated with KDEs, are estimated using *bootstrapping* [136]. In the following, whenever a KDE uncertainty is used, it refers to such errors obtained from bootstrapping the Monte Carlo sample and recalculating the KDE several times from the bootstrapped samples [129].

6.3 Implementation of Systematic Parameters

The measurement of the NMO is affected by several systematic uncertainties. In this work, systematic uncertainties are treated by inserting nuisance parameters into the template fit. To do this, the impact of each systematic on the analysis distribution is parametrized and the resulting parameters are used as nuisance parameters. These parameters comprise:

1. uncertainties in the normalization of the contributions in Figure 6.3 and 6.4,
2. uncertainties in the atmospheric neutrino and muon fluxes, discussed in Section 2.4,
3. uncertainties in the oscillation parameters, discussed in Section 2.3,
4. uncertainties in the neutrino-nucleon cross-sections, discussed in Section 2.2,
5. uncertainties in the detector response, discussed in Chapter 3.

Thus, these uncertainties extend from the neutrino generation (1-2), propagation (3) and interaction (4) to their detection in DeepCore (5).

The systematic parameters are parametrized either

1. *event-wise*, where the effect on the analysis distribution is obtained from re-weighting Monte Carlo events or
2. *bin-wise*, where for each bin an analytical function is used to describe the relative change of the bin-content depending on the nuisance parameter.

For the event-wise parametrization, only one MC sample is needed for each contribution. It is called *baseline sample*, as it was generated with all systematic uncertainties being at their baseline value.

For the bin-wise parametrization, the bin-wise function is obtained from so-called *systematic samples*, i.e. Monte Carlo samples including the investigated systematic effect. To do this, the analysis distribution is generated for each systematic sample and the observed change in the bin-content is fitted against the value of the systematic parameter for each bin.

In the following, the bin-wise parametrization is used only for the detector systematics, while normalization, cross-section, flux and oscillation systematics are parametrized event-

wise. All systematic parameters are summarized in Table 6.1. The first half of the table

Table 6.1: Systematics treated as nuisance parameters in the likelihood analysis, including normalization (N), detector response (D), oscillation (O), atmospheric flux (F) and neutrino-nucleon interaction (I) uncertainties, where the last column gives the baseline value and the estimated uncertainty: values in brackets are used only for systematic studies in Section 7.6, while uncertainties without brackets are also used as Gaussian prior in the fitted likelihood; if possible, a reference for the assumed baseline value and uncertainty is given.

label	type	short description of parameter	baseline ± uncert.
		fitted parameters	
N_ν	N	normalization of total neutrino template [46][c]	$1.00(\pm 0.25)$[a]
N_{ν_e}	N	normalization of ν_e flux (before propagation) [46][c][137][d]	1.00 ± 0.05[a]
N_{NC}	N	normalization of the NC template	1.0 ± 0.2[a]
N_μ	N	normalization of atmospheric muon template	$1.0(\pm 0.5)$[a]
ϵ_{opt}	D	photon detection (optical) efficiency [78][c,d]	1.0 ± 0.1[a]
$L^{\mathrm{hi}}_{\mathrm{scat}}$	D	photon scattering length in hole-ice [78, 60][c,d]	25 ± 10
$k^{\mathrm{hi}}_{\mathrm{fwd}}$	D	head-on optical efficiency [60][c,d]	$0.0(\pm 1.0)$[b]
Δm^2_{31}	O	atmospheric squared-mass difference	$2.5 \cdot 10^{-3}\,\mathrm{eV}^2$
θ_{23}	O	atmospheric mixing angle [25, 26][d]	$0.740^{(+0.036)}_{(-0.195)}$ rad
$\Delta\gamma_\nu$	F	neutrino spectral index uncertainty [46][c][137][d]	0.0 ± 0.1
γ_μ	F	atmospheric muon template uncertainty [138][d]	0.0 ± 1.0[b]
R_θ	F	atm. neutrino flux uncertainty [46][c][47][d]	0.0 ± 1.0[b]
$R_{\nu/\bar{\nu}}$	F	atm. neutrino flux uncertainty [46][c][47][d]	0.0 ± 1.0[b]
M^{res}_A	I	axial mass uncertainty of RES events [139][c,d]	$1.12 \pm 0.22\,\mathrm{GeV}$
M^{qe}_A	I	axial mass uncertainty of QIS events [139][c,d]	$0.99 \pm 0.25\,\mathrm{GeV}$
		additional, tested parameters (not fitted)	
N_{noise}	N	normalization of triggered noise template	1.0 ± 1.0[a]
N_{coin}	N	normalization of events containing a coincident atmospheric muon	1.0 ± 1.0[a]
$\sin^2(\theta_{13})$	O	neutrino oscillation parameter [25, 26][c,d]	$0.02206^{+0.00225}_{-0.00225}$
$\sin^2(\theta_{12})$	O	neutrino oscillation parameter [25, 26][c,d]	$0.307^{+0.039}_{-0.035}$
Δm^2_{21}	O	neutrino oscillation parameter [25, 26][c,d]	$7.40^{+0.62}_{-0.60} \cdot 10^{-5}\,\mathrm{eV}^2$
δ_{CP}	O	neutrino oscillation parameter [25, 26][c,d]	$180° \rightarrow 270°$
$\sigma^{\nu}_{\mathrm{DIS}}$	I	x_{Bjorken}-dependent DIS cross-section for neutrinos [140][c,d]	0 ± 0.0757[b]
$\sigma^{\overline{\nu}}_{\mathrm{DIS}}$	I	x_{Bjorken}-dependent DIS cross-section for anti-neutrinos [140][c,d]	0 ± 0.1008[b]

[a] relative to the nominal value of this parameter
[b] *ad-hoc* parametrization in arbitrary units with baseline value at 0
[c] nominal value motivated by the provided reference
[d] uncertainty motivated by the provided reference

consists of all parameters fitted in the NMO analysis, which are described in this section. The second half of the table comprises additional parameters that were tested, but removed from the likelihood fit, since their impact was found to be negligible. They are discussed in

more detail in Section 7.6.

For parameters used with a Gaussian prior in the template fit, the width of the Gaussian prior is stated in the last column, while the prior is centered at the systematic's baseline value. These priors are the same for all ongoing DeepCore analyses, fitting these parameters. For priors based on external information, a reference is provided to motivate the choice of the prior. All priors are chosen conservatively in the sense that their width is equal to or larger than the expected uncertainty of this parameter. This is done to support the convergence of the fit to the true value, but avoid the fit being driven primarily by the value of the prior, but the observed data.

The parameters from Table 6.1 are described in more detail in the following.

6.3.1 Normalization Uncertainties

The analysis includes CC electron, muon and tau neutrinos as well as NC events, atmospheric muons and triggered noise (cf. Figure 6.3 and 6.4). To avoid fitting a systematically wrong hypothesis, by under- or overestimating the normalization of one of these contributions, the normalizations $N_\nu, N_{\nu_e}, N_{\mathrm{NC}}$ and N_μ are used as nuisance parameters in the likelihood fit.

The parameter N_ν describes the total normalization of the neutrino template, i.e. it scales the combined templates from CC electron, CC muon, CC tau and NC events. The parameter N_{ν_e} describes the normalization of electron neutrinos, generated in the atmosphere, i.e. before applying oscillation effects. This allows to account for uncertainties in the ν_μ/ν_e-ratio generated in the Earth's atmosphere by meson decays (cf. Section 2.4.1). The parameter N_{NC} describes the total normalization of NC events, including all flavors. Finally, the parameter N_μ scales the normalization of the atmospheric muon template.

The CC tau normalization N_{ν_τ} and noise normalization N_{noise} are not used as nuisance parameters: Since tau neutrinos are not generated in the atmosphere, N_{ν_τ} depends only on the electron and muon neutrino fluxes and the oscillation parameters. Deviations from this expectation could only arise from non-standard oscillations (e.g. non-unitarity of the mixing matrix U) or unknown, systematic selection effects, which should be small, since the cascade-like signatures from CC tau neutrinos, CC electron neutrinos and NC interactions are indistinguishable with IceCube at these energies. As the measurement of tau-appearance is a strong physics case by itself, this analysis focuses only on the NMO, while keeping the tau normalization to the expectation $N_{\nu_\tau} = 1$ for standard oscillations. However, a similar analysis is conducted on this sample to measure the tau normalization by Michael Larson [112].

The noise normalization is not fitted, since within any reasonable range of N_{noise}, this analysis is not sensitive to it. It can only affect the analysis distribution notably, if it is free to choose completely unrealistic noise contributions like $N_{\mathrm{noise}} \sim 10 - 100$. Therefore, the normalization is fixed in the following.

Moreover, note that the atmospheric muon normalization N_μ is used without prior. This is done to give enough freedom to the likelihood fit to describe the amospheric muon contamination, due to the limited MC description and the muons' strong sensitivity to the detector response.

6.3.2 Atmospheric Flux Uncertainties

The atmospheric electron and muon (anti-)neutrino fluxes used in Monte Carlo are taken from the predictions for the South Pole in [46] (cf. Section 2.4.1). The uncertainties on these fluxes are mostly taken from [47], which discusses the energy-, zenith- and flavor-dependent uncertainties relative to the flux predicted in [48]. Since the used model of atmospheric fluxes is an updated version of the model, that was assumed for the calculation of the uncertainties in [47], the actual errors on the predicted Monte Carlo fluxes are expected to be smaller than the used one.

The dominant systematic effects on the atmospheric neutrino production arise from uncertainties in the kaon and pion production process in the atmosphere (cf. Section 2.4.1) and the unknown properties of the primary Cosmic Ray flux [47]. In the following, these effects are parametrized by the parameters $\Delta\gamma_\nu$, $R_{\nu/\bar{\nu}}$ and R_θ.

The parameter $\Delta\gamma_\nu$ modifies the weight w_i of each MC event i according to its energy $E_{\nu,i}$ by $w_i \rightarrow w_i(E_{\nu,i}/\mathrm{GeV})^{\Delta\gamma_\nu}$. Thus, it modifies the spectral index of the atmospheric neutrino flux for all flavors and neutrinos and anti-neutrinos in the same way. The value of the prior is set to $\sigma_{\Delta\gamma_\nu} = 0.1$ motivated by [137].

The parameters $R_{\nu/\bar{\nu}}$ and R_θ describe the energy- and zenith-dependent uncertainties in the ratios of neutrino flavors and neutrinos to anti-neutrinos. They are based on an *ad-hoc* parametrization of the relevant uncertainties discussed in [47]. The uncertainties on these ratios from [47] are shown in Figure 6.7.

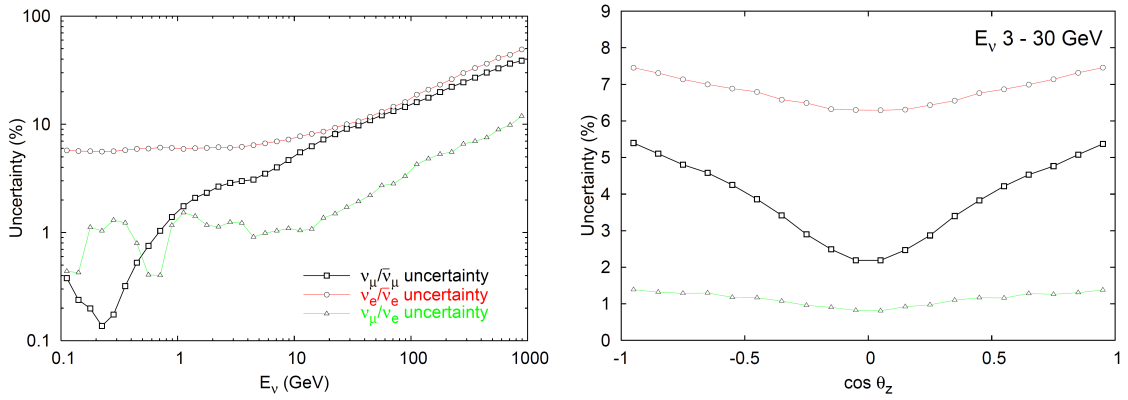


Figure 6.7: Energy-dependent (left) and zenith-dependent (right) uncertainties on ratio of different neutrino types; left: for all zenith-angles; right: at energies of $E_\nu \in [3{,}30]\,\mathrm{GeV}$ (from [47]).

Note that changes in the normalization of the individual contributions are already captured by the parameters from Section 6.3.1, while a global, energy-dependent tilt in all neutrino templates is already captured by $\Delta\gamma_\nu$. Moreover, uncertainties at energies $E_\nu \lesssim 1\,\mathrm{GeV}$ are irrelevant for this work.

The ad-hoc parametrization reduces the 16 independent uncertainties from [47] to the two parameters $R_{\nu/\bar{\nu}}$ and R_θ by focussing only on relevant effects for DeepCore. The most relevant parameter $R_{\nu/\bar{\nu}}$ describes the energy-dependent ratio of neutrinos and anti-neutrinos, which changes the zenith-dependent ratio at the same time. The second parameter R_θ is a minor correction to the zenith-dependent change, which is used to guarantee that the

changes cover the uncertainties given by [47]. This two-dimensional parametrization is the common treatment of atmospheric flux uncertainties in DeepCore analyses [60, 61].

For the ad-hoc parametrization, $R_{\nu/\bar{\nu}}$ and R_θ were chosen to be 0.0 at the baseline point and ± 1 at $\pm 1\,\sigma$ uncertainty. Thus, the parameter value gives the pull with respect to the Gaussian prior. The impact of a $+1\,\sigma$-shift in both parameters on the final analysis distribution is shown in Figure 6.8. As illustrated, the impact of $R_{\nu/\bar{\nu}}$ is clearly larger than the one for R_θ. In fact, R_θ has only a small effect on the template fit, while $R_{\nu/\bar{\nu}}$ is an important parameter, as discussed in Section 7.6.

A detailed comparison between the uncertainty provided by [47] and the ad-hoc parametrization was conducted by Juan Pablo Yáñez [118]. Another study of all parameters was done by the PINGU collaboration, which found no relevant effects beyond the ones provided by the above mentioned ad-hoc parametrization. Thus, the ad-hoc parametrization is also used for sensitivity studies of the future PINGU detector [7].

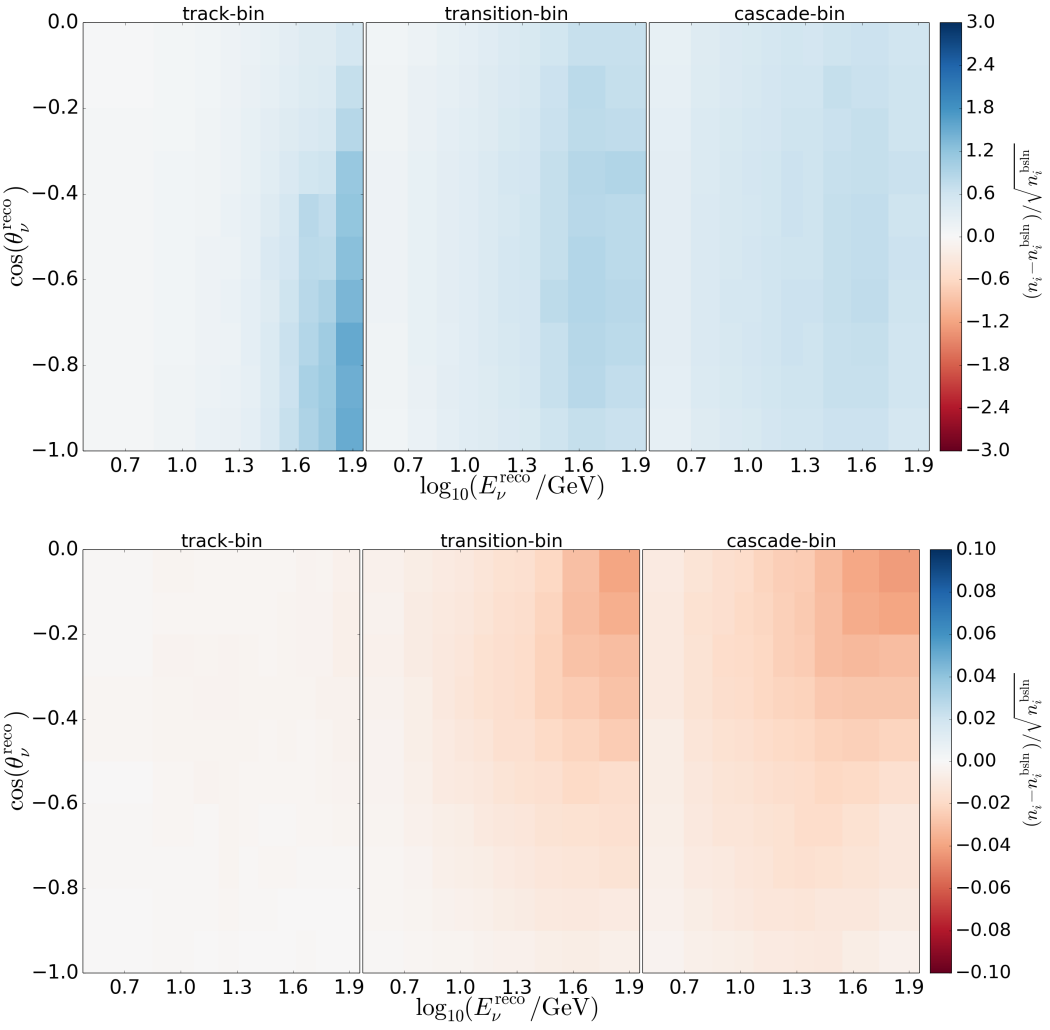


Figure 6.8: Expected pulls from changing $R_{\nu/\bar{\nu}}$ (top) and R_θ (bottom) by $+1\,\sigma$ for one year of data; note the different scale of the colorbars, indicating the larger impact of $R_{\nu/\bar{\nu}}$ compared to R_θ; the pulls from changing the parameters by $-1\,\sigma$, are roughly the shown ones with an inverted sign, since their impact is nearly linear within $[-1\sigma, +1\sigma]$.

Finally, a nuisance parameter is used for the atmospheric muon template, called γ_μ. The

parameter γ_μ propagates the uncertainties on the primary Cosmic Ray flux towards the muon template according to [138]. This comprises uncertainties in the composition of Cosmic Rays as well as on their energy spectrum.

To obtain the effect on the final muon template, CORSIKA simulations were run for varying assumptions on the Cosmic Ray primary spectrum. The atmospheric muons were then propagated to the MuonGun generation cylinder. The covered change in the density of CORSIKA muons was then splined and parametrized by the parameter γ_μ. Finally, this parametrized density is used to re-weight the MuonGun events on the generation cylinder, changing the atmospheric muon template. The parameter is chosen to be $\gamma_\mu = 0$ at the baseline point and ± 1 at $\pm 1\,\sigma$ [138].

It turns out, that the parameter has only negligible impact on the template fit and the fitted values are always close to the baseline value for Pseudo-Experiments (cf. Appendix H.2), which is due to the assumed prior. However, the parameter is kept to check for unexpected, extreme deviations in the atmospheric muon template.

6.3.3 Uncertainties of Oscillation Parameters

For this work, the oscillation parameters from Table 2.1 are treated as nuisance parameters. However, most of them have no impact on the oscillation measurement with DeepCore (cf. Section 7.6). Thus, the parameters Δm^2_{12}, θ_{13}, θ_{12} and δ_{CP} are fixed to the best-fit values.

The remaining, free parameters are the so-called atmospheric oscillation parameters Δm^2_{31} and θ_{23}. In contrast to the other oscillation parameters, their baseline values are not taken from the most recent NuFit result [25, 26]. Instead, they were kept unchanged during the development of this work, but are still compatible with the most recent global fits. These baseline values are only used for sensitivity and systematic studies and do not affect the experimental result. In the following, θ_{23} is mostly replaced by $\sin^2(\theta_{23})$, which is conventionally chosen for the illustration of the mixing angle. Note that Δm^2_{32} is no additional degree of freedom, as it is given by $\Delta m^2_{31} = \Delta m^2_{32} + \Delta m^2_{21}$ (cf. Section 2.3.2).

As shown in Table 6.1, the atmospheric oscillation parameters enter the likelihood without prior. In principle, such prior could be assigned by taking recent, global best-fit contours from other experiments, although they would be highly non-Gaussian in θ_{23}. However, this is not done, since IceCube's sensitivity towards these parameters is competitive to those measurements and thus the oscillation parameters are of interest themselves. Moreover, this allows to produce an independent confidence interval on these parameters, so-called *contour*, as done in Section 8.1.2. Such contour can still be combined with results from other experiments, as done in Section 8.2, but also allows to study the result individually.

Although these parameters are nuisance parameters for the NMO measurement, the experimental fit of the atmospheric oscillation parameters is discussed in Section 8.1.2.

6.3.4 Uncertainties of Interactions

Neutrino-nucleon interactions can be split into deep inelastic (DIS), quasi-elastic (QE) and resonant (RES) interactions, as discussed in Section 2.2. The GRECO sample on final level is clearly dominated by DIS interactions, contributing $\sim 81.4\%$ of all CC neutrino events and $> 99.9\%$ of all NC neutrino events. However, a $\sim 11.4\%$ resonant and $\sim 7.2\%$ quasi-elastic contribution for CC events remains - especially at low energies (cf. Figure 2.2).

The theoretical value of the QE and RES cross-sections depends strongly on the nucleon form factors. The axial mass M_A arises as a phenomenological parameter, when writting these form factors in a standard dipole parametrization. Thus, the axial mass rescales the cross-section of RES and QE neutrino-nucleon interactions, depending on the value of the transferred, squared four-momentum Q^2 [139].

In GENIE, the QE and RES interactions are simulated with

$$\begin{aligned} M_A^{\text{res}} &= 1.12\,\text{GeV}\ (1.12 \pm 0.03\,\text{GeV}) \\ M_A^{\text{qe}} &= 0.99\,\text{GeV}\ (0.96 \pm 0.03\,\text{GeV}), \end{aligned} \tag{6.4}$$

where the values in brackets correspond to the best-fit values reported by [139]. Moreover, GENIE provides the coefficients of a second order polynomial for every generated MC event. With this polynomial, the event can be re-weighted for any change of the axial mass parameters M_A^{res} and M_A^{qe}.

Besides RES and QE uncertainties, uncertainties in the DIS cross-section are not included in the likelihood fit. For DIS interactions, the uncertainties from [141] were parametrized by Shivesh Mandalia and Teppei Katori [140]. However, the resulting parameters were found to have no relevant impact on the likelihood and are therefore not used as nuisance parameters (cf. Section 7.6).

6.3.5 Detector Uncertainties

The detector systematics describe uncertainties in the detector response. This comprises the propagation of Cherenkov photons through the ice as well as the DOM response to a photon hit. In this work, they are treated bin-wise by producing systematic samples for every uncertainty. These systematic samples include the effect from varying the systematic within its uncertainties and propagate the effect to the final analysis histogram. A full list of all systematic samples, used in this work, is attached in Appendix H.1.

In this work, three detector uncertainties are used:

1. *optical efficiency* ϵ_{opt}: The optical efficiency describes the efficiency of the photo detection. This includes the efficiency of the PMT and other DOM hardware as well as optical properties of the surrounding ice (cf. Section 3.5). Its uncertainty is assumed to be $\pm 10\%$ of the baseline value. To estimate the impact of the optical efficiency, the MC samples in Appendix H.1 are generated with high photon efficiency and propagated through the selection chain in Chapter 4. Then, the optical efficiency is reduced by removing some of the observed photons randomly, where the fraction of removed photons corresponds to the reduction in optical efficiency. After removing the given fraction of photons, the sample is processed analogously.

2. *hole-ice* $L_{\text{scatter}}^{\text{hi}}$: The hole-ice parameter is the first of two parameters, describing the optical properties of the hole-ice. Since the DOMs are located within the drill holes, its impact is parametrized by changing the photon detection efficiency, depending on the inclination-angle η of the incoming photons, as described in Section 3.4. The baseline value and uncertainty of the resulting systematic parameter are $L_{\text{scatter}}^{\text{hi}} = 25 \pm 10$. For illustration purposes, this value is sometimes interpreted as scattering length in centimeters, but is in fact not directly linked to the scattering length. However, it was defined in a way that resembles the scattering length parameter of a previously used hole-ice model to simplify comparisons. The impact of $L_{\text{scatter}}^{\text{hi}}$ on the η-dependent

detection efficiency is shown in Figure 6.9. Together with the high-forward parameter, it covers the parametrized uncertainty from Figure 3.9 (green shaded band). The hole-ice systematic samples are produced analogously to the optical efficiency parameter by removing photons according to their inclination-angle η and chosen values of $L^{\text{hi}}_{\text{scatter}}$.

3. *high-forward* $k^{\text{hi}}_{\text{scatter}}$: The high-forward is the second of the two parameters, describing the optical properties of the hole-ice. It changes the η-dependent photon detection efficiency for vertically upwards moving photons, as shown in Figure 6.9. For these photons, the acceptance is not well-constrained by string-to-string LED measurements and therefore largely unknown (cf. Section 3.4). The high-forward parameter is an *ad-hoc* parametrization with one dimensionless parameter. Its baseline value $k^{\text{hi}}_{\text{scatter}} = 0$ is used without any prior from external knowledge. For values $k^{\text{hi}}_{\text{scatter}} > 0$ ($k^{\text{hi}}_{\text{scatter}} < 0$), it enhances (reduces) the acceptance for vertically upwards moving photons. Together with the hole-ice parameter, it covers the uncertainty given by the green shaded band in Figure 3.9. The systematic samples are produced analogously to the cases for $L^{\text{hi}}_{\text{scatter}}$ and ϵ_{opt}.

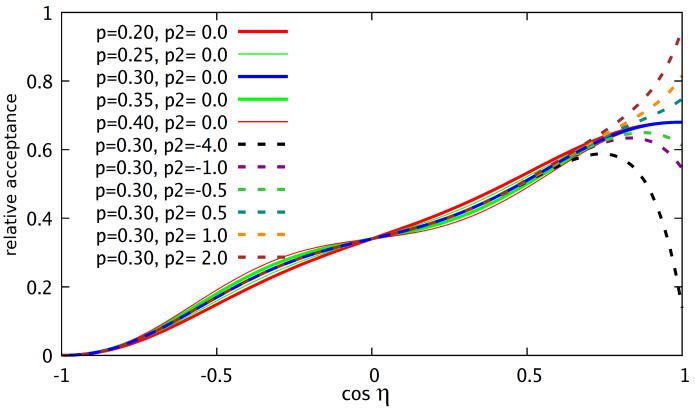


Figure 6.9: η-dependent photon acceptance for various values of p = $L^{\text{hi}}_{\text{scat}}/100$ and p2 = $k^{\text{hi}}_{\text{fwd}}$, where $\cos(\eta) = 1$ (-1) corresponds to photons travelling vertically upwards (downwards); the two parameters cover the η-dependent uncertainty from Figure 3.9 (from [142]).

For each systematic sample, the analysis distributions from Figure 6.3 and 6.4 are generated. For every bin, the bin-content of all systematic samples is then used to parametrize the systematic effect continuously. The parametrization is done separately for each contribution $c \in \{\nu_e, \nu_\mu, \nu_\tau, \nu_{\text{NC}}, \mu\}$, while the triggered noise is not affected by these detector systematics.

For each contribution, systematic and bin, a first order polynomial is fitted to the values of the bin-content, obtained from the systematic samples, as a function of the systematic parameter, except for the parametrization of the optical efficiency for atmospheric muons. There, an exponential function is fitted for the following reasons:

In contrast to neutrinos, where an increased optical efficiency leads to an increased number of events, the atmospheric muon contamination decreases in optical efficiency. This is due to the fact that the muon contamination is largely suppressed by the data selection (cf. Chapter 4). However, the suppression depends strongly on the efficiency of the photon detection. If the optical efficiency is increased, this also increases the veto-efficiency and thus reduces the muon contamination. For decreased optical efficiency, the opposite is the case.

In Figure 6.10, the parametrization of the bin-content y, compared to the baseline sample y_0, is shown for one analysis bin against the optical efficiency ϵ_{opt}. Each marker corresponds to the bin-content of one systematic sample, while the errorbars are obtained from boot-

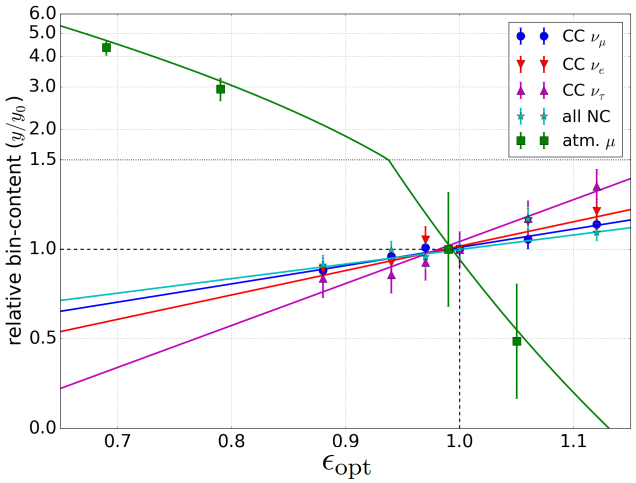


Figure 6.10: Example for the parametrization of the bin-content y relative to the baseline value y_0 for one bin, depending on ϵ_{opt}: each marker represents a systematic sample with an errorbar obtained from bootstrapping, while the solid, colored lines show the resulting parametrization of the effect for every contribution; note the transition from linear to logarithmic scale at $y/y_0 = 1.5$.

strapping the corresponding KDE (cf. Section 6.2.1). The solid lines passing through the markers of the same color are the parametrizations, obtained for each contribution.

For each contribution, the baseline sample is located at $y/y_0 = 1$, as the bin-content of the baseline sample $y = y_0$ is used for the normalization. However, y_0 might be under- or overfluctuating due to small MC fluctuations. To account for this, the parametrization p is free to choose $p(x_0) \neq y_0$ and thus to correct for fluctuations in the baseline sample.

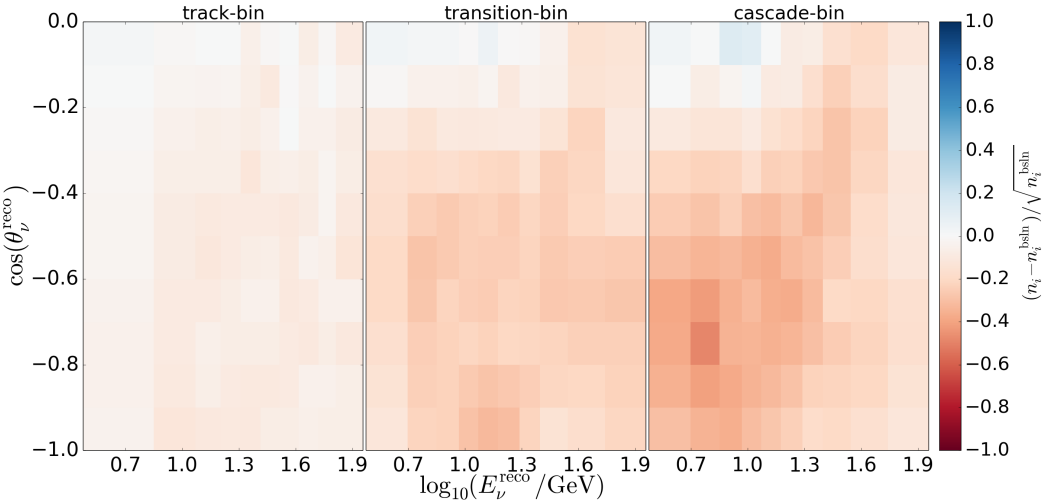


Figure 6.11: Pulls on analysis distribution from changing $k_{\text{scatter}}^{\text{hi}}$ by $+1\,\sigma$; the behavior of the parameters is nearly linear within $[-1\sigma, +1\sigma]$ such that the pulls from changing $k_{\text{scatter}}^{\text{hi}}$ by $-1\,\sigma$ are roughly the shown ones with an inverted sign.

By applying such fit to every bin, contribution and systematic, the impact of each detector systematic on the analysis distribution in Figure 6.1 is parametrized. In Figure 6.11, the pulls on the analysis distribution are examplarily shown for shifting $k_{\text{scatter}}^{\text{hi}}$ by $+1\sigma$. Note that this is the sum of the individual effects from all contributions $c \in \{\nu_\mu, \nu_e, \nu_\tau, \text{NC}, \mu\}$.

The detector systematics are assumed to be independent close to the baseline values. Thus, they factorize such that the bin-content y_{ci} for a component c and bin i is given by the

product

$$y_{ci}(\epsilon_{\text{opt}}, L^{\text{hi}}_{\text{scatter}}, k^{\text{hi}}_{\text{scatter}}) = b_{ci} \cdot p^{\text{DE}}_{ci}(\epsilon_{\text{opt}}) \cdot p^{\text{HI}}_{ci}(L^{\text{hi}}_{\text{scatter}}) \cdot p^{\text{HF}}_{ci}(k^{\text{hi}}_{\text{scatter}}) \cdot y^{ci}_0, \tag{6.5}$$

where p^{DE}_{ci}, p^{HI}_{ci} and p^{HF}_{ci} are the parametrizations of the optical efficiency, hole-ice and high-forward, respectively, while y^{ci}_0 is the *raw* bin-content of the baseline sample.

A so-called *baseline-correction* b_{ic} is added to align the parametrizations for all three detector systematics at the baseline sample, since each parametrization passes its baseline value x_0 at a different value y. The correction b_{ic} is in general small and described in more detail in Appendix G.

6.4 Modified Poissonian Likelihood

For comparing the total MC template from Figure 6.1 to the data histogram, two statistical effects are taken into account: The first one are the statistical fluctuations on the bin-content for data. The underlying statistic is Poissonian, such that the number of observed events N_i might deviate from the expectation value $\hat{\mu}_i$. The true expectation value $\hat{\mu}_i$ is not known, but estimated from MC simulations as μ_i with an uncertainty σ_{μ_i}, due to the finite number of MC events. The error σ_{μ_i}, obtained from bootstrapping KDEs, is shown in Figure 6.12 (left) for the total MC template.

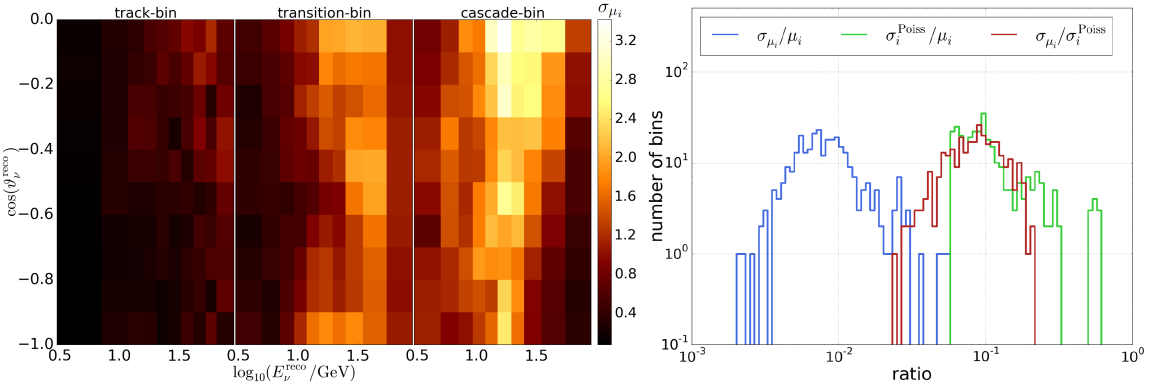


Figure 6.12: Left: template error σ_{μ_i} (in expected event counts) from limited MC statistics, obtained from the combination of the individual MC errors for each contribution; Right: template error σ_{μ_i} (blue) and Poissonian error $\sigma_i^{\text{Poiss}} = \sqrt{\mu_i}$ (green) relative to total number of predicted events μ_i and ratio of both (red), where the MC error is roughly an order of magnitude smaller than the Poissonian error on the bin-content.

In this work, the uncertainty from limited MC statistics is much smaller than the Poissonian fluctuations $\sigma_i^{\text{Poiss}} = \sqrt{\mu_i}$ on the bin-content. Thus, $\sigma_{\mu_i} \ll \sigma_i^{\text{Poiss}}$ holds for every bin of the MC template. The ratio of $\sigma_{\mu_i}/\sigma_i^{\text{Poiss}}$ is shown in Figure 6.12 (right).

Although the errors are small, they are conventionally included in DeepCore analyses. This is done by combining the uncertainty σ_{μ_i} with the Poissonian error. To do this, the Poissonian distribution is convolved with a log-normal distribution according to Equation 6.6

$$p^{\text{tot}}_i(N_i, \mu_i, \sigma_{\mu_i}) = \int_0^\infty \mathrm{d}\tilde{\mu} \left(p^{\text{Poiss}}_i(N_i, \tilde{\mu}) \cdot \mathcal{LN}(\tilde{\mu}, \mu_i, \sigma_{\mu_i}) \right), \tag{6.6}$$

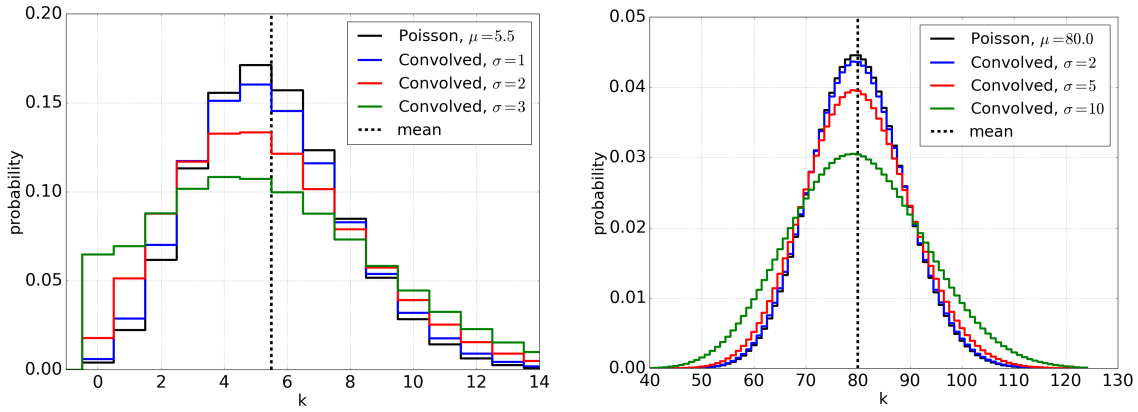


Figure 6.13: Example of Poissonian probabilities of mean $\mu = 5.5$ (left) and $\mu = 80$ (right) convolved with a log-normal distributions of different width $\sigma = \sigma_{\mu_i}$; the black distributions show the underlying Poissonian distributions that are obtained in the limit $\sigma \to 0$.

where $p_i^{\text{Poiss}}(N_i, \tilde{\mu})$ is the Poissonian probability to observe N_i events, in case $\tilde{\mu}$ are predicted, and $\mathcal{LN}(\tilde{\mu}, \mu_i, \sigma_{\mu_i})$ is a (narrow) log-normal distribution in $\tilde{\mu}$ with an expectation value μ_i and a standard deviation σ_{μ_i}. For illustration, the resulting convolution of the Poissonian and a narrow log-normal distribution are shown in Figure 6.13.

The probabilities p_i^{tot} are used to define the negative log-likelihood in Equation 6.7

$$\text{LLH} = \left[- \sum_{i \in \{\text{bins}\}} \log \left(\frac{p_i^{\text{tot}}(N_i, \mu_i, \sigma_{\mu_i})}{p_i^{\text{tot}}(N_i, N_i, \sigma_{\mu_i})} \right) \right] + \frac{1}{2} \left[\sum_{s \in \{\text{sys}\}} \left(\frac{s - s_0}{\sigma_s} \right)^2 \right]. \tag{6.7}$$

The first sum runs over all $3 \times 10 \times 10 = 300$ bins in PID, reconstructed energy and reconstructed zenith-angle. It compares the observed bin-content N_i to the prediction μ_i, taking the uncertainty of the prediction σ_{μ_i} into account and normalizing the probability by $p_i^{\text{tot}}(N_i, N_i, \sigma_{\mu_i})$. This way, a *normalized likelihood* [143] is obtained, which allows for a comparison of log-likelihood values for different realisations of N_i. The normalized log-likelihood values are used in Chapter 8 as a proxy for the agreement of data N_i with the Monte Carlo prediction μ_i. Such proxies are known as *goodness-of-fit* estimators [144].

The second sum in Equation 6.7 adds the LLH-contribution from the Gaussian priors on the systematic parameters in Table 6.1. Each systematic s is compared to its baseline value s_0 in terms of its uncertainty σ_s and the resulting disagreement is taken into account in the log-likelihood.

In Chapter 8, the fit is repeated with the Poissonian likelihood from Equation 6.8

$$\text{LLH} = \left[- \sum_{i \in \{\text{bins}\}} \log \left(\frac{p_i^{\text{Poiss}}(N_i, \mu_i)}{p_i^{\text{Poiss}}(N_i, N_i)} \right) \right] + \frac{1}{2} \left[\sum_{s \in \{\text{sys}\}} \left(\frac{s - s_0}{\sigma_s} \right)^2 \right]. \tag{6.8}$$

It differs from the convolved likelihood in Equation 6.7 by using only Poissonian probabilities, without taking the uncertainties σ_{μ_i} on the MC template into account. In the following, only the convolved likelihood is used to obtain the NMO result, while the Poissonian likelihood is used to validate the result in Chapter 8.

6.5 Likelihood Optimization within *oscFit*

The NMO is fitted simultaneously with the nuisance parameters described in Section 6.3. It consists of four separate minimizations of the negative log-likelihood. This is done, since the likelihood space features four separate minima: one in each octant (i.e. for $\sin^2(\theta_{23}) \leqslant 0.5$ and $\sin^2(\theta_{23}) > 0.5$) and ordering ($\Delta m^2_{31} > 0$ and $\Delta m^2_{31} < 0$).

Note that in case of maximum mixing ($\sin^2(\theta_{23}) = 0.5$), the two octant-dependent minima converge to just one minimum at the boundary of both octants. For the two NMO-dependent minima this is only the case for $\Delta m^2_{31} = 0$, which is clearly excluded by DeepCore and several other experiments (cf. Section 2.3.2).

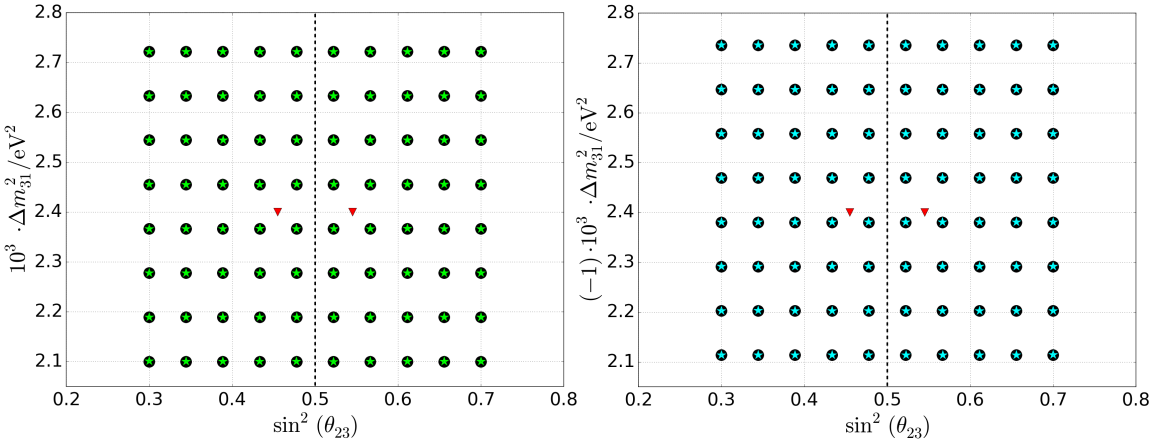


Figure 6.14: Validation of the minimization process; each black dot represents a set of oscillation parameters injected either in the Normal (left) or Inverted (right) Ordering hypothesis; the stars on top show the resulting best-fit values for NO (green star) and IO (cyan star) being injected; as only green (cyan) stars show up in the NO (IO) plane, they recover the true ordering; the seed value in both octants and orderings is indicated by a red triangle.

To account for this non-convexity of the likelihood space, the fit is repeated for every combination of the Neutrino Mass Ordering NMO $\in \{\mathrm{NO}, \mathrm{IO}\}$ and octant $O \in \{\mathrm{L}, \mathrm{R}\}$, while for each NMO+$O$ hypothesis, all remaining systematic parameters are optimized. The fit giving the optimum log-likelihood value is taken as result of the fit, while the likelihood differences between the four fits is used to determine the sensitivity in Chapter 7 and the result of the NMO measurement in Chapter 8.

The fit in each octant and ordering is done using *oscFit*, which is a standard fitting tool within the IceCube Collaboration. For the likelihood optimization, *L-BFGF-B* is chosen as a minimization algorithm (*minimizer*), which is based on a *Quasi-Newton-Method* [124].

To validate the minimization performance, the oscillation parameters are injected on a grid, as shown in Figure 6.14 as black dots for NO (left) and IO (right). The minimizer is seeded once in each octant and ordering at the values indicated by the red triangles. For all other systematic parameters, the seed is perturbed randomly within the uncertainty of the corresponding parameter. Then, the minimizer is tested to recover the correct values of the oscillation parameters, all other systematic parameters and the correct octant and ordering hypothesis.

This is visualized in Figure 6.14, where each star corresponds to the optimum of four minimizations (in both orderings and octants). The best-fit value is indicated by the position of the star, while the color corresponds to NO (green) and IO (cyan) being injected. The stars line-up with the corresponding black dots, such that every injected value is recovered by the minimizer. Moreover, all green stars are found in the NO plane (left), while all cyan stars end up in the IO plane (right). Thus, the minimizer always chooses the right octant and ordering, while also recovering the correct oscillation parameters. Note that it was also validated that all other systematic parameters were recovered at their injected values.

From the four fits, the best NO and IO fits are taken to derive the log-likelihood difference from Equation 6.9

$$\Delta\mathrm{LLH}_{\mathrm{NO-IO}} \equiv \mathrm{LLH}_{\mathrm{NO}} - \mathrm{LLH}_{\mathrm{IO}} = -\log\left(\frac{\mathcal{L}_{\mathrm{NO}}}{\mathcal{L}_{\mathrm{IO}}}\right) = \log(\mathcal{L}_{\mathrm{IO}}) - \log(\mathcal{L}_{\mathrm{NO}}), \tag{6.9}$$

where $\mathcal{L}_{\mathrm{TO}}$ is the likelihood obtained for the $\mathrm{TO} \in \{\mathrm{NO}, \mathrm{IO}\}$ hypothesis, optimizing over both octants and all systematic parameters. This log-likelihood difference is used in Chapter 8 to derive the p-values for the NO and IO hypotheses in the experimental result and in Chapter 7 to calculate sensitivities.

Moreover, by considering all four, above mentioned likelihood values separately, any combinations of ordering and octant can be tested, which is done in Section 7.4.

7 Sensitivity for Measuring the Neutrino Mass Ordering

The sensitivities, presented in the following, are calculated with two different methods: first, using *Pseudo-Experiments*, described in Section 7.1, and second, using an *Asimov method*, described in Section 7.3.

The Pseudo-Experiments method is very time consuming, but close to what is done for experimental data. It is the most accurate estimate of the sensitivity used in the following. However, since its calculation is time intense, it is used only to validate the result from the Asimov method for some points of the parameter space.

The Asimov method is used for quick estimations of the sensitivity. It allows to test the sensitivity for various parameters in a suitable amount of time.

Both methods are described in the following Sections 7.1 to 7.3, before the resulting sensitivities are presented in Section 7.4 and 7.5. Afterwards the impact of systematics and potential improvements on the sensitivity are discussed in Section 7.6 and 7.7.

The shown sensitivities are published in [126], [127] and [116].

7.1 Experimental Result and Pseudo-Experiments

For the experimental result, the negative log-likelihood difference from Equation 6.7 is used as a test statistic, while the sign determines the preference for Normal Ordering ($\Delta\text{LLH}_{\text{NO}-\text{IO}} < 0$) or Inverted Ordering ($\Delta\text{LLH}_{\text{NO}-\text{IO}} > 0$).

From the observed value of $\Delta\text{LLH}_{\text{NO}-\text{IO}}$, a p-value for both hypotheses is derived. The p-value gives the probability to observe an excess in positive or negative direction of the test statistic that is at least as strong as the observed one for a given hypothesis.

To derive the p-value, the best-fit values of all systematic parameters for Normal (NO) and Inverted (IO) Ordering are used to produce *Pseudo-Experiments*. Pseudo-Experiments are produced by generating a Monte Carlo (MC) template for the given set of systematic parameters, scaling it to the observed livetime and adding Poissonian fluctuation on the content of each bin. These Pseudo-Experiments are supposed to resemble the actual experimental histogram, such that repeating the experiment N times would lead to a similar distribution of fitted parameters as obtained from fitting the N Pseudo-Experiments, in case all systematic uncertainties are described correctly.

From the Pseudo-Experiments, two distributions of $\Delta\text{LLH}_{\text{NO}-\text{IO}}$ are obtained for NO and IO being the injected ordering hypothesis. From these two distributions and the experimentally

observed value $\Delta\text{LLH}^{\text{exp}}_{\text{NO-IO}}$, a p-value for the NO (IO) hypothesis is derived as the fraction of Pseudo-Experiments that are right (left) of $\Delta\text{LLH}^{\text{exp}}_{\text{NO-IO}}$. The procedure is sketched in Figure 7.1 for the case of NO being observed ($\Delta\text{LLH}^{\text{Exp}}_{\text{NO-IO}} < 0$). It is often referred to as *Feldman-Cousin* method, based on [145].

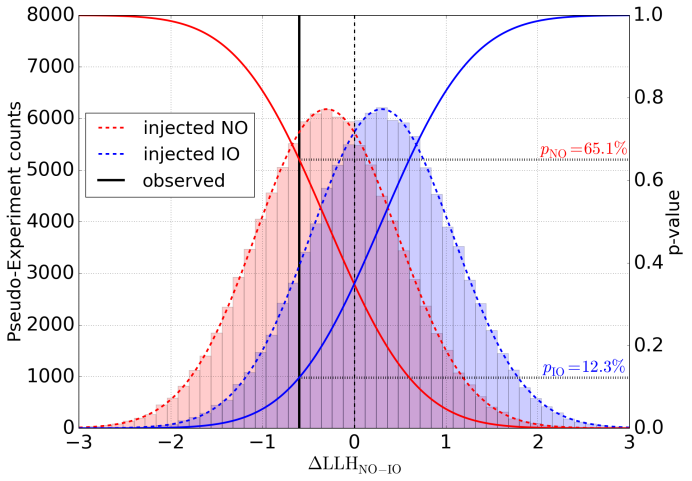


Figure 7.1: Sketch of procedure to determine the p-value for NO and IO, using $\Delta\text{LLH}_{\text{NO-IO}}$ as test statistic: the black, solid line marks an observed value, while the red and blue Gaussian distributions are obtained from Pseudo-Experiments; the red and blue, solid lines, to be read from the right, vertical axis, determine the p-value from the cummulative distributions.

Thus, an observed log-likelihood difference $\Delta\text{LLH}^{\text{exp}}_{\text{NO-IO}}$ can be converted into a p-value for each hypothesis, using the simulation of Pseudo-Experiments with the best-fit systematic parameters from data.

Additionally, the expected p-values, called *sensitivity*, are calculated to estimate the analysis performance before fitting experimental data. The sensitivity to NO (IO) is defined by the median p-value obtained for the exclusion of the opposite IO (NO) hypothesis, in case NO (IO) is true. It is calculated from the median likelihood difference $\overline{\Delta\text{LLH}}_{\text{NO-IO}}$ for a given hypothesis. The median p-value is estimated by the fraction of the IO (NO) distribution that is left (right) of the median value $\overline{\Delta\text{LLH}}_{\text{NO-IO}}$, obtained for the NO (IO) hypothesis.

Note that another common choice for the definition of sensitivities is replacing the median by the mean of the $\Delta\text{LLH}_{\text{NO-IO}}$-distributions from Figure 7.1. However, the distribution of $\Delta\text{LLH}_{\text{NO-IO}}$ is assumed to be symmetric in the following, such that both definitions are identical. This assumption is discussed in more detail in Section 7.3, while for Pseudo-Experiments the median value is used in the following.

The p-value is converted into one-sided Gaussian sigmas according to Table 7.1. The sensitivities stated in the following correspond to these p-values.

Table 7.1: Conversion of p-values p into one- and two-sided Gaussian sigmas.

$1-p$	0%	30%	50%	68.27%	90%	95%	99%
one-sided	$-\infty$	$-0.524\,\sigma$	$0\,\sigma$	$0.476\,\sigma$	$1.282\,\sigma$	$1.645\,\sigma$	$2.326\,\sigma$
two-sided	$0\,\sigma$	$0.385\,\sigma$	$0.674\,\sigma$	$1\,\sigma$	$1.645\,\sigma$	$1.960\,\sigma$	$2.576\,\sigma$

Note that the choice of a one-sided instead of a two-sided Gaussian is natural for this measurement: In case the analysis is not sensitive to the NMO at all, the distributions in

Figure 7.1 overlap at 0.0. In that case, an expected p-value of 50% is obtained for NO and IO, which corresponds to $0\,\sigma$ for a one-sided Gaussian, but $0.476\,\sigma$ in case of a two-sided Gaussian. In other words, even a random choice of NO or IO would get the true NMO right in 50% of the cases, which should not be interpreted as sensitivity to the NMO.

To estimate a sensitivity for NO (IO) at a specific value of the oscillation parameters, Pseudo-Experiments are generated at these parameters with all other systematic parameters being set to the baseline value. Then, the opposite hypothesis IO (NO) is generated at a set of oscillation parameters that resembles the original NO (IO) hypothesis most closely. To obtain this set of parameters, the original NO (IO) template is fitted with the opposite IO (NO) hypothesis excluding statistical fluctuations. The resulting values of all parameters are then used to generate the Pseudo-Experiments for the opposite IO (NO) hypothesis.

For all Pseudo-Experiments (injecting NO and IO), both hypotheses are fitted in both octants and the value of $\Delta\mathrm{LLH}_{\mathrm{NO-IO}}$ is calculated (cf. Equation 6.9).

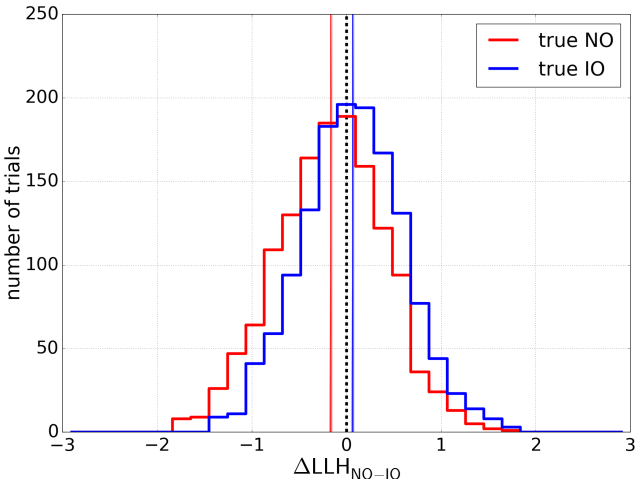


Table 7.2: Example of the distribution of $\Delta\mathrm{LLH}_{\mathrm{NO-IO}}$ obtained from Pseudo-Experiments, generated at the baseline value of all systematic parameters for NO and the best-matching parameters for IO.

The resulting distributions for NO and IO are shown in Figure 7.2 for the case of the baseline parameters from Table 6.1 being chosen for NO. The vertical solid lines correspond to the median value of $\Delta\mathrm{LLH}_{\mathrm{NO-IO}}$. Thus for NO being true, the expected p-values for testing the NO and IO hypotheses are

$$p_{\mathrm{NO}} = 65.1\% \qquad\qquad p_{\mathrm{IO}} = 32.3\%, \tag{7.1}$$

where p_{IO} gives the sensitivity for NO, i.e. the expected p-value for the exclusion of the IO hypothesis.

Besides the log-likelihood difference, the distribution of the oscillation parameters Δm^2_{31} and $\sin^2(\theta_{23})$ are shown in Figure 7.2. Here, red stars correspond to NO being injected, while green stars correspond to IO being injected. Due to the sensitivity to the NMO, the number of red stars clustering at $\Delta m^2_{31} > 0$ is larger than at $\Delta m^2_{31} < 0$, while for green stars, the behavior is inverted.

Note that the absolute value for positive Δm^2_{31} is typically larger than for negative Δm^2_{31}, which is mostly due to the mass-difference Δm^2_{21}.

Moreover, one can see a clustering of events at $\sin^2(\theta_{23}) = 0.5$, i.e. at maximum mixing. This is due to the fact, that muon neutrino disappearance is the dominant oscillation effect in DeepCore. The effect can be understood from the two-flavor approximation in Equation 2.26. There, the disappearance amplitude is proportional to $\sin^2(2\theta_\nu)$. Thus, it is

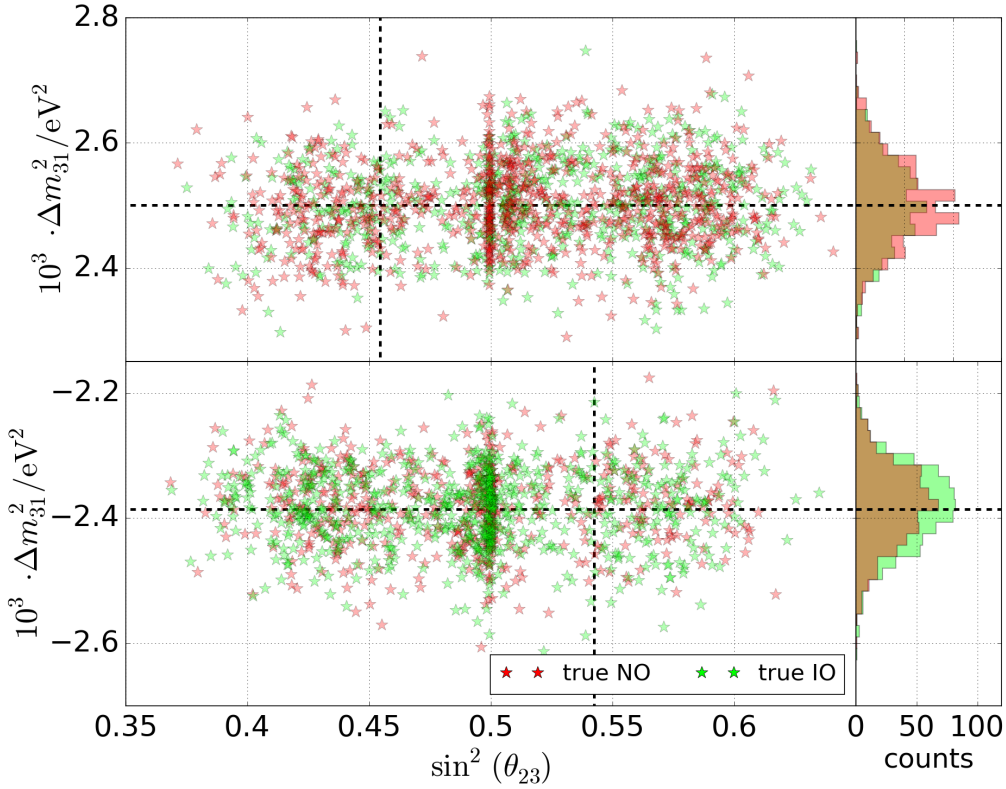


Figure 7.2: Distribution of $\sin^2(\theta_{23})$ and Δm^2_{31} for Pseudo-Experiments generated at the baseline values for NO (red stars) and the best-matching values for IO (green stars); Pseudo-Experiments preferring NO (IO) show up in the top (bottom) figure; the injected value for NO and IO are indicated by the dashed black lines; the clustering at maximum mixing is expected due to the described parametrization effect.

maximal for $\theta_{2\nu} \mathrel{\widehat{=}} \theta_{23} = \pi/4$, which leads to $\sin^2(\theta_{23}) = 0.5$. As a result, the oscillation effect reduces when moving away from $\sin^2(\theta_{23}) = 0.5$ in positive or negative direction, making $\sin^2(\theta_{23}) = 0.5$ a boundary of the oscillation effect. For Pseudo-Experiments simulated close to maximum mixing, events can prefer more or less muon neutrino disappearance due to statistical fluctuations. However, more disappearance than at $\sin^2(\theta_{23}) = 0.5$ is not parametrized and thus, these Pseudo-Experiments cluster at $\sin^2(\theta_{23}) = 0.5$.

The distributions of the remaining systematic parameters are not discussed here. They are added to Appendix H.2 as well as a correlation matrix for all systematic parameters for the baseline scenario discussed above.

Although Pseudo-Experiments are very close to what is done with experimental data, they are not used for all sensitivity estimations in this work. Since they are very time consuming, the simulation of Pseudo-Experiments is used only for experimental data and to validate the results of the Asimov sensitivities, presented in Section 7.3.

In contrast, the *Asimov method* uses $\Delta\mathrm{LLH}_{\mathrm{NO-IO}}$ not only as a test statistic. Instead, it exploits some of its properties that arise from its definition as a negative log-likelihood-ratio. Such methods are commonly used in particle physics and mostly based on *Wilks' Theorem*, described in the following Section. They allow for a much faster sensitivity estimation and are used here to test the sensitivity to the NMO for different values of the oscillation parameters in a suitable amount of time.

Although Wilks' Theorem is not valid for the measurement of the Neutrino Mass Ordering, it is used in Section 8.1.2 to interpret the result on the oscillation parameters. It is described in Section 7.2, while the Asimov method, used for the NMO, is described in Section 7.3.

7.2 Wilks' Theorem and Asimov Datasets

Hypotheses tests in particle physics are commonly done using *Wilks' Theorem*, based on the work by S.S. Wilks [146] and A. Wald [147]. In Wilks' Theorem, two hypotheses $\mathcal{H}_G$ (G for *general*) and $\mathcal{H}_S$ (S for *specific*) are tested in a log-likelihood method, where $\mathcal{H}_S$ is nested in $\mathcal{H}_G$ such that they differ by n_{dof} degrees of freedom. The additional conditions on the likelihood, required by Wilks' Theorem, are not discussed here, but assumed to be satisfied for the given likelihood function [147].

When fitted to data D, one inevitabably finds $\text{LLH}(D|\mathcal{H}_G) \leqslant \text{LLH}(D|\mathcal{H}_S)$ for the negative log-likelihood values, as $\mathcal{H}_S$ is nested in $\mathcal{H}_G$ and thus the more general hypothesis must outperform the specific one. The degree to which the general hypothesis outperforms the specific one is given by $\Delta\text{LLH}(D) = \text{LLH}(D|\mathcal{H}_S) - \text{LLH}(D|\mathcal{H}_G) \geqslant 0$. Note that besides the n_{dof} parameters for the $\mathcal{H}_G$ hypothesis, both hypotheses might share a common set of nuisance parameters that are optimized to obtain the above likelihood values.

With statistical fluctuations, Wilks' Theorem predicts that the values of $2\Delta\text{LLH}(D)$, obtained for repeating the experiment $N \to \infty$ times follows a χ^2-distribution with n_{dof} degrees of freedom. As a result, the p-value to obtain $\Delta\text{LLH} > x$ in case $\mathcal{H}_S$ is true is given by

$$p(\Delta\text{LLH} > x|\mathcal{H}_S) = 1 - \text{cdf}_{\chi^2}(2\Delta\text{LLH}, n_{\text{dof}}), \tag{7.2}$$

where $\text{cdf}_{\chi^2}(2\Delta\text{LLH}, n_{\text{dof}})$ is the cummulative density function of a χ^2-distribution with n_{dof} degrees of freedom, evaluated at $2\Delta\text{LLH}$.

Note that in literature, Wilks' Theorem is often stated in terms of $\Delta\chi^2 = 2\Delta\text{LLH}$ in case of Gaussian statistics. In case $n_{\text{dof}} = 1$, this leads to the more widely known formula for the number of Gaussian standard deviations $n_\sigma = \sqrt{2\Delta\text{LLH}} = \sqrt{\Delta\chi^2}$.

Table 7.3: Conversion of ΔLLH values into p-values according to Wilks' Theorem for different numbers of degrees of freedom n_{dof} [146].

n_{dof}	30%	68.3%	90%	95%	99%
1	0.074	0.500	1.353	1.921	3.317
2	0.357	1.148	2.303	2.996	4.605
3	0.712	1.763	3.126	3.907	5.672

An overview over the p-values for different values of ΔLLH is given in Table 7.3.

For Monte Carlo simulations, the templates can be fitted without statistical fluctuations. In that case, a deterministic likelihood value $\overline{\text{LLH}}$ is obtained, which is 0.0, in case the injected hypothesis is the fitted one (or nested in the fitted hypothesis) and > 0.0 otherwise, due to the normalization of the likelihood (cf. Equation 6.6). Such fits of the templates without injecting statistical fluctuations are called *Asimov fits*, while the template is sometimes called *Asimov dataset* [1]. A proper motivation for using the Asimov datasets can be found in [149].

[1] referring to the non-scientific short-story *Franchise* by Isaac Asimov [148], where an election is held by the vote of only one (the most-representative) voter

For the Asimov method, it is assumed that the value $\overline{\text{LLH}}$, obtained without statistical fluctuations, is a valid estimator for the median (or mean) value of LLH, obtained with statistical fluctuations. Thus, using LLH = $\overline{\text{LLH}}$ in Wilks' Method, one obtains an estimate for the median p-value from Equation 7.2, where small p-values indicate a preference of $\mathcal{H}_G$ over $\mathcal{H}_S$. For analyses, that test a specific hypothesis nested in a more general one, the sensitivity is typically defined by such median (or mean) p-value.

However, Wilks' Theorem cannot be used to estimate the sensitivity to the NMO for two reasons: First, the NO and IO hypotheses are not nested within each other. Second, they are not connected by a continuous parameter, such that $n_{\text{dof}} = 0$.

Nonetheless, Wilks' Theorem and the values from Table 7.3 are used in Section 8.1.2 to derive *confidence-level* (CL) contours depending on the oscillation parameters.

7.3 Asimov Sensitivities for Discrete, Non-Nested Hypotheses

In this work, an alternative Asimov method is used to determine the sensitivity to the Neutrino Mass Ordering, which is based on [150].

Since the NO and the IO hypotheses are not nested, $\Delta\text{LLH} = \text{LLH}_{\text{NO}} - \text{LLH}_{\text{IO}}$ is neither χ^2-distributed nor strictly positive. Instead, the sign of ΔLLH determines the preference for NO ($\Delta\text{LLH} < 0$) or IO ($\Delta\text{LLH} > 0$). In [150], it is shown that the mean $\overline{\Delta\text{LLH}}$ and the standard deviation $\sigma_{\Delta\text{LLH}}$ of the ΔLLH-distribution are related by Equation 7.3, in case the statistics in the underlying bins are Gaussian:

$$\sigma_{\Delta\text{LLH}} = \sqrt{2\overline{\Delta\text{LLH}}}. \tag{7.3}$$

It is tempting to identify this with the corresponding relation of χ^2-distributions. However, $2\Delta\text{LLH}$ is clearly not χ^2-distributed. Instead, it is found in [150] to be well-approximated by a Gaussian distribution. This motivates the assumption from Section 7.1 about the mean and median of the ΔLLH-distribution to be approximately identical.

Due to its Gaussian behavior, any observed value ΔLLH can be converted into a number of one-sided Gaussian sigmas n_σ by Equation 7.4

$$n_\sigma = \frac{\Delta\text{LLH} - \overline{\Delta\text{LLH}}(\mathcal{H}_{\text{TO}})}{\sqrt{2\overline{\Delta\text{LLH}}(\mathcal{H}_{\text{TO}})}}, \tag{7.4}$$

where $\overline{\Delta\text{LLH}}(\mathcal{H}_{\text{TO}})$ is the expected log-likelihood difference for injecting $\text{TO} \in \{\text{NO}, \text{IO}\}$ as the true ordering.

For calculating sensitivities, ΔLLH is replaced by its expectation value $\overline{\Delta\text{LLH}}(\mathcal{H}'_{\text{CO}})$ for injecting the complementary ordering $\text{CO} \in \{\text{IO}, \text{NO}\}$ to $\text{TO} \in \{\text{NO}, \text{IO}\}$. Additionally, the dash in $\mathcal{H}'$ indicates that the two hypotheses $\mathcal{H}_{\text{TO}}$ and $\mathcal{H}'_{\text{CO}}$ differ in the set of injected systematic parameters.

The values $\overline{\Delta\text{LLH}}(\mathcal{H}_{\text{TO}})$ and $\overline{\Delta\text{LLH}}(\mathcal{H}'_{\text{CO}})$ are obtained by the following steps, which are sketched in Figure 7.3:

1. generate an Asimov template for the NO (IO) hypothesis with a set of parameters P,
2. fit this template in both octants of the complementary ordering IO (NO) (2 fits),

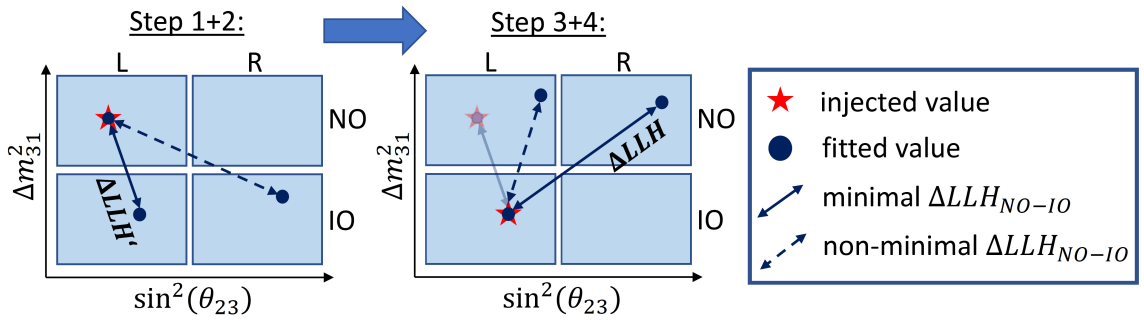


Figure 7.3: Illustration of the four steps applied to calculate the Asimov sensitivity n_σ^{NO}, where L and R represent the left and right octant, respectively

3. generate an Asimov template for the resulting, best fitting systematic parameters P' for the complementary ordering IO (NO),
4. fit this template of the complementary ordering in both octants of the original ordering NO (IO) (2 fits).

From the best-fit values in Step 2, the value of $\Delta\mathrm{LLH}(\mathcal{H}_{\mathrm{NO/IO}})$ is calculated, while the value of $\Delta\mathrm{LLH}(\mathcal{H}'_{\mathrm{IO/NO}})$ is obtained from fitting the best-fit complementary ordering from step 2 in the original ordering NO (IO) in step 4.

Thus, the sensitivity from Equation 7.4 for NO and IO is given by:

$$n_\sigma^{\mathrm{NO}} = \frac{\overline{\Delta\mathrm{LLH}}(\mathcal{H}'_{\mathrm{NO}}) - \overline{\Delta\mathrm{LLH}}(\mathcal{H}_{\mathrm{IO}})}{\sqrt{2\overline{\Delta\mathrm{LLH}}(\mathcal{H}_{\mathrm{IO}})}} \qquad n_\sigma^{\mathrm{IO}} = \frac{\overline{\Delta\mathrm{LLH}}(\mathcal{H}'_{\mathrm{IO}}) - \overline{\Delta\mathrm{LLH}}(\mathcal{H}_{\mathrm{NO}})}{\sqrt{2\overline{\Delta\mathrm{LLH}}(\mathcal{H}_{\mathrm{NO}})}}, \tag{7.5}$$

where $n_\sigma^{\mathrm{NO}} \geqslant 0$ and $n_\sigma^{\mathrm{IO}} \leqslant 0$ due to the definition of $\Delta\mathrm{LLH} = \mathrm{LLH_{NO}} - \mathrm{LLH_{IO}}$. Note that the sign is kept here to simplify the discussion of systematic influences in Section 7.6.

A simple way to understand the Equations 7.5 is the following:

The two distributions in Figure 7.1 are Gaussian. Their mean value is at $\overline{\Delta\mathrm{LLH}}$, while their width $\sqrt{2\overline{\Delta\mathrm{LLH}}}$ is directly linked to their mean value. Thus, the average NO experiment, which gives a value of $\overline{\Delta\mathrm{LLH}}(\mathcal{H}_{\mathrm{NO}}) < 0$, is $\overline{\Delta\mathrm{LLH}}(\mathcal{H}_{\mathrm{NO}}) - \overline{\Delta\mathrm{LLH}}(\mathcal{H}_{\mathrm{IO}})$ off the expected value for the average IO experiment. Expressing this distance in terms of the standard deviations of the IO distribution, one obtains Equation 7.5 for n_σ^{NO} up to one hypothesis $\mathcal{H}$ being dashed, which indicates the different sets of injected systematic parameters, used to compare the best-matching NO and IO hypotheses.

Although the Asimov method allows for a quick estimation of the sensitivity for several values of parameters, one should note that it is not expected to give perfect agreement with the Pseudo-Experiments method from Section 7.1. For example, the Gaussian approximation of the ΔLLH-distribution in Figure 7.3 does not take the degeneracy of the $\sin^2(\theta_{23})$-parameter into account. As a result, the binary choice of the octant is not modelled accurately, which leads to a $\sin^2(\theta_{23})$-dependent deviation from the Gaussian assumption and thus, a $\sin^2(\theta_{23})$-dependent deviation from the expected sensitivity. However, the differences between the Asimov sensitivity and the result from Pseudo-Experiments are found to be small in Section 7.4. Therefore, the Asimov sensitivity is kept as a valid sensitivity proxy, that can be derived in a suitable amount of time.

Finally, matter effects can be tested against vacuum oscillations. Again, this is a test of two non-nested, discrete hypotheses. It is done with the same Asimov method by replacing the log-likelihood-ratio ΔLLH between the NO and the IO hypothesis by the log-likelihood ratio between matter effects (MA) and vacuum oscillations (VA), as done in Equation 7.6

$$\Delta\text{LLH} = \text{LLH}_{\text{MA}} - \text{LLH}_{\text{VA}} = -\log\left(\frac{\mathcal{L}_{\text{MA}}}{\mathcal{L}_{\text{VA}}}\right) = \log(\mathcal{L}_{\text{VA}}) - \log(\mathcal{L}_{\text{MA}}). \tag{7.6}$$

Like the binary hypothesis test between NO and IO is done by assuming matter effects, the binary hypothesis test between matter effects and vacuum oscillations is done by assuming a Neutrino Mass Ordering. Thus, the likelihood values $\mathcal{L}_{\text{MA}}$ and $\mathcal{L}_{\text{VA}}$ are calculated under the assumption of NMO $\in \{\text{NO}, \text{IO}\}$. The resulting sensitivity for matter effects is estimated by Equations 7.7

$$\begin{aligned} n_\sigma^{\text{MA}}(\mathcal{H}_{\text{NMO}}) &= \frac{\overline{\Delta\text{LLH}}'(\mathcal{H}_{\text{MA}}|\mathcal{H}_{\text{NMO}}) - \overline{\Delta\text{LLH}}(\mathcal{H}_{\text{VA}}|\mathcal{H}_{\text{NMO}})}{\sqrt{2\overline{\Delta\text{LLH}}(\mathcal{H}_{\text{VA}}|\mathcal{H}_{\text{NMO}})}}, \\ n_\sigma^{\text{VA}}(\mathcal{H}_{\text{NMO}}) &= \frac{\overline{\Delta\text{LLH}}'(\mathcal{H}_{\text{VA}}|\mathcal{H}_{\text{NMO}}) - \overline{\Delta\text{LLH}}(\mathcal{H}_{\text{MA}}|\mathcal{H}_{\text{NMO}})}{\sqrt{2\overline{\Delta\text{LLH}}(\mathcal{H}_{\text{MA}}|\mathcal{H}_{\text{NMO}})}}. \end{aligned} \tag{7.7}$$

In the following, when the Neutrino Mass Ordering is tested, matter effects are always assumed implicitly, while for the test of matter effects against vacuum oscillations, a specific NMO is assumed explicitly. Moreover, for the fit of vacuum oscillations, the ordering is set to Normal Ordering for simplicity reasons: This is done, since DeepCore is not sensitive to the NMO without matter effects, such that $\text{LLH}_{\text{VA}}(\mathcal{H}_{\text{VA}}|\mathcal{H}_{\text{NO}}) \approx \text{LLH}_{\text{VA}}(\mathcal{H}_{\text{VA}}|\mathcal{H}_{\text{IO}})$.

Note that the sensitivity to matter effects is larger for NO than for IO, as the observable signature from matter effects is increased for NO compared to IO.

7.4 Sensitivity to the Neutrino Mass Ordering

In the following, three types of sensitivities are calculated with the Asimov method:

1. testing NO against IO as described in the previous section,
2. testing NO against IO for a specific combination of ordering and octant,
3. testing matter effects (MA) against vacuum oscilations (VA) for a specific choice of NMO $\in \{\text{NO}, \text{IO}\}$ (MA is replaced by NO and IO in the following to indicate the assumed ordering).

For the first scenario, the sensitivities n_σ^{NO} and n_σ^{IO} are calculated according to Equation 7.5. The resulting $\sin^2(\theta_{23})$-dependent sensitivities are shown in Figure 7.4 as red, dot and blue, square markers for NO and IO being the true ordering, respectively.

Moreover, the sensitivity is verified at some values of $\sin^2(\theta_{23})$ using the Pseudo-Experiments method from Section 7.1. The obtained sensitivities are shown as star-like markers, while the vertical error indicates the statistical uncertainty due to the limited number of Pseudo-Experiments. Both methods are roughly consistent in their results and feature a similar $\sin^2(\theta_{23})$-dependency, while some expected deviations from the Asimov sensitivity are observed (cf. Section 7.3).

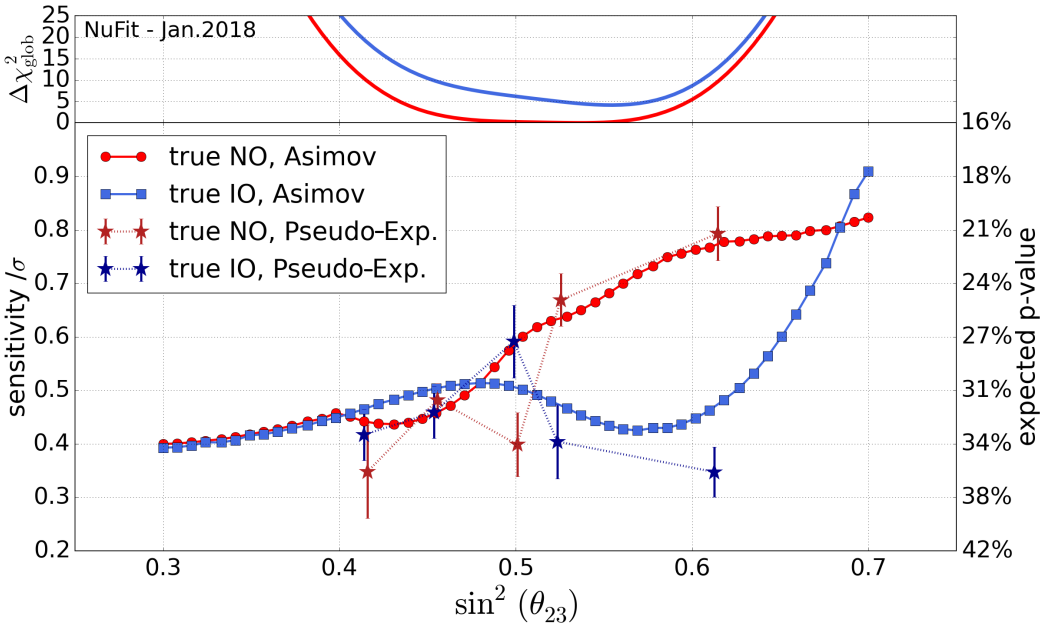


Figure 7.4: $\sin^2(\theta_{23})$-dependent sensitivity to NO and IO obtained from the Asimov method in terms of one-sided Gaussian sigmas (left, vertical axis) and the expected p-value (right, vertical axis); the star markers indicate the sensitivity as obtained from Pseudo-Experiments to validate the Asimov method at selected values of $\sin^2(\theta_{23})$; in the subplot above, the $\Delta\chi^2$-scan from global best-fits are shown, as provided by NuFit [26] (pre-published in [127, 116]).

Recall that the sensitivity n_σ^{NO} means that the IO hypothesis is on average rejected at a confidence level corresponding to a one-sided p-value of n_σ^{NO} Gaussian standard deviations. For the most interesting region around maximum mixing, the obtained sensitivities in Figure 7.4 are $\sim 0.4 - 0.7\,\sigma$ for both orderings.

The characteristic shape in $\sin^2(\theta_{23})$ is due to the θ_{23}-dependence of the matter effects (cf. Section 2.3.3). It resembles the shape of the sensitivity curve, stated by the PINGU Collaboration in Figure 3.13. However, the DeepCore and PINGU sensitivities are not expected to match perfectly, since PINGU will measure oscillation effects down to $\sim 1\,\text{GeV}$ with a different data selection and different reconstruction resolutions for PID, energy and zenith-angle. Thus, the NMO signature is observed by PINGU over a broader enery range and with higher resolution.

The shape of the curves in Figure 7.4 can be understood from considering the left and right octants separately. This is shown in Figure 7.5 (left) in terms of $\overline{\Delta\text{LLH}}_{\text{NO-IO}}$, separated for both octants. By forcing a comparison with the same or the opposite octant for the two orderings, the resulting curves of $\overline{\Delta\text{LLH}}_{\text{NO-IO}}$ feature a nearly linear or nearly parabolic behavior, respectively.

For IO being the true ordering, the best fitting NO hypothesis is always in the left octant, since the blue dots outperform the blue stars in the left octant, while the stars outperform the dots in the right octant. For NO being true, this is not the case: In the left octant, the best fitting IO hypothesis is either in the left or right octant, while in the right octant, the best fitting IO hypothesis is always in the right octant as well.

The shape of the sensitivity curves in Figure 7.4 is a result of the intersections of these parabolic and linear behaviors for both octants and orderings.

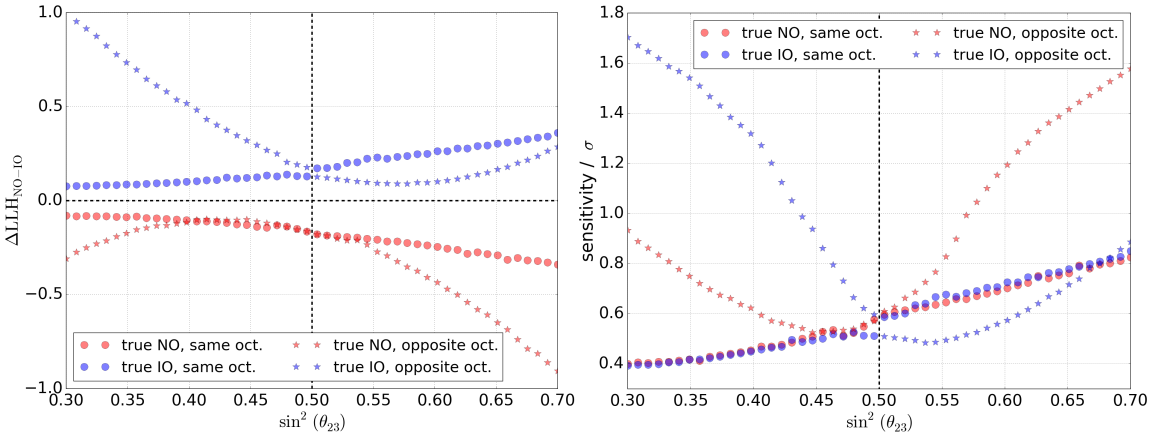


Figure 7.5: Left: octant-dependent $\Delta\mathrm{LLH}_{\mathrm{NO-IO}}$ for the true ordering TO $\in$ {NO,IO}: dotted markers are obtained by taking the same octant for the complementary ordering CO $\in$ {IO,NO}, while stars are obtained by taking the opposite octant; right: the resulting Asimov sensitivity, which is the expected p-value for excluding the same (dots) and the opposite octant (stars) of the CO hypothesis, in case the true ordering is TO.

Furthermore, Figure 7.5 (left) allows to calculate an octant-dependent sensitivity to the NMO. This is done according to Equation 7.5 by enforcing the comparison with a specific octant of the opposite ordering, i.e. by avoiding the optimization with respect to the octant. The resulting sensitivities are shown in Figure 7.5 (right). Here, the n_σ^{NO} sensitivity for the same (opposite) octant gives the expected p-value for the exclusion of the same (opposite) octant in the opposite ordering.

As one can see, the octant-dependent NMO sensitivity is typically larger than the one from Figure 7.4, since the optimization has less freedom to account for differences between the orderings by choosing the octant. As a result, the sensitivity increases, especially for large differences between both octants, i.e. far-off maximum mixing.

7.5 Sensitivity to Matter Effects

Besides testing the NMO, matter effects can be tested against vacuum oscillations according to Equation 7.7. In Figure 7.6, the resulting sensitivity is shown for MA and VA being true and for both ordering hypotheses NO and IO.

As expected, the separation between matter effects and vacuum oscillations is larger for NO than for IO being the true ordering. Moreover, a peak is observed at maximum mixing, in case matter effects exist. The peak is due to the fact that fitting vacuum oscillations on templates, generated with matter effects, leads to a systematic shift of the mixing angle towards maximum mixing. Thus, if the true mixing angle is already close to maximum mixing, the vacuum fit is incapable of compensating for the difference with respect to matter effects by pushing $\sin^2(\theta_{23})$ closer to maximum mixing. As a result, the log-likelihood difference between MA and VA is increased, leading to an increased sensitivity for excluding VA, close to $\sin^2(\theta_{23}) = 0.5$.

One might be tempted to think that the sensitivity to matter effects should exceed the sensitivity to the NMO, as the NMO is seen as a result of matter effects. However, this is

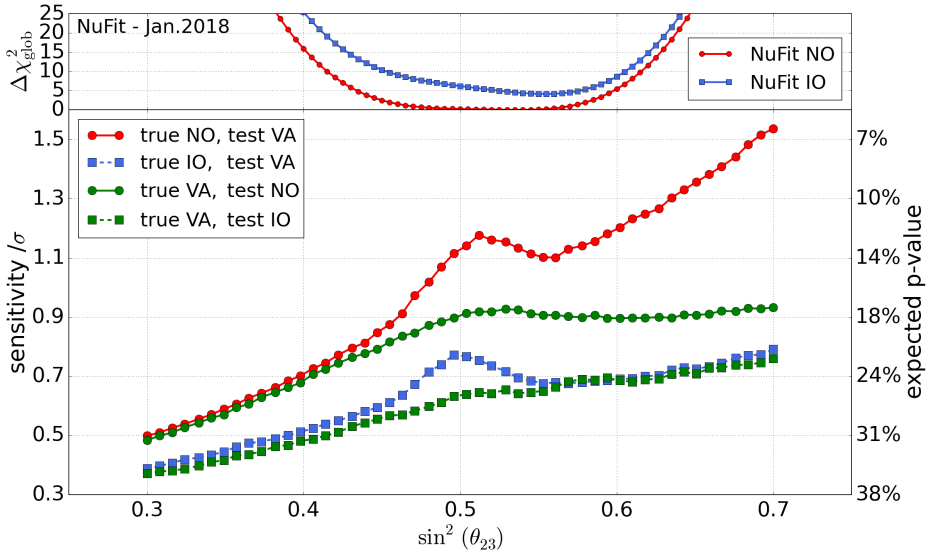


Figure 7.6: $\sin^2(\theta_{23})$-dependent sensitivity to matter effects in case of Normal (NO) and Inverted Ordering (IO) and nearly ordering-independent vacuum oscillations (VA), as obtained from the Asimov method (cf. Equation 7.7) for matter effects in NO or IO or VA being true.

a-priori not necessarily the case.

As shown in Table 2.2, matter effects arise in the neutrino channel in case of NO and in the anti-neutrino channel in case of IO. Moreover, the GRECO sample consists of 65.0% neutrino and only 30.1% anti-neutrino events. The flavor-dependent composition of neutrinos and anti-neutrinos is shown in Figure C.4 of Appendix C. As a result, the signature of matter effects is large for NO and small for IO (cf. Table 2.2). However, the gap between 65.0% and 30.1% of events featuring matter effects is comparable to the difference between 30.1% and 0% (vacuum oscillations) of events featuring matter effects.

This leads to the peculiar feature that it is possible to e.g. prefer Inverted Ordering over Normal Ordering, assuming matter effects, at a higher significance than matter effects over vacuum oscillations.

7.6 Impact of Systematic Uncertainties on the Sensitivity

In addition to estimating the sensitivity, the Asimov method is used to quantify the impact of the systematic parameters. To do this, three tests are performed for each parameter, regarding the following questions:

1. How relevant is the inclusion of the parameter into the fit to obtain an unbiased result?
2. How much is the sensitivity affected by the true value of this systematic parameter?
3. How much sensitivity could be gained by knowing the parameter perfectly?

These three questions are answered by so-called $(N-1)$-*tests*. The $(N-1)$ refers to the fact that for each test, a single systematic parameter is fixed to and/or shifted from its baseline value and the effect on the Neutrino Mass Ordering (NMO) result is studied. The following

$(N-1)$-*tests* are applied to answer the questions from above:

1. *Fitting Relevance Test* (FRT): In the FRT, a parameter is injected at $\pm 1\,\sigma$ from its baseline value. In addition, the parameter is fixed to the baseline value in the fitter. Note that this leads to a systematically wrong hypothesis being fitted. The resulting shift in $\Delta\mathrm{LLH}_{\mathrm{NO-IO}}$ is the bias that would be obtained by fixing this parameter in the fitter, while a $\pm 1\,\sigma$ value is realized in nature. Thus, it answers the first of the above questions.
2. *Sensitivity Stability Test* (SST): In the SST, a parameter is injected at $\pm 1\,\sigma$ from its baseline value. The fitter corrects for this shift by moving the corresponding systematic parameter. The resulting change in the sensitivity is used to answer the second of the above questions. Note that for $\sin^2(\theta_{23})$, this is also done in the sensitivity curve in Figure 7.4.
3. *Hidden Potential test* (HPT): In the HPT, a parameter is fixed to its baseline value, which is also the injected value. Thus, the parameter cannot be used in the fit to compensate for differences between the orderings. However, the true value of the parameter is used for all fits. As a result, the sensitivity to the NMO increases, which answers the third of the above questions.

The uncertainties used for all parameters are listed in Table 6.1. For parameters used with a Gaussian prior, the assumed uncertainty is identical to the value of the prior. For the remaining parameters, the assumed uncertainty is estimated from external knowledge or internal parameters studies within the IceCube Collaboration. Where possible, the table provides a reference for the assumed baseline value and uncertainty.

For the following FRT, more systematics were tested than finally added to the NMO fit. The remaining parameters were excluded from the likelihood fit, because they were found to have no impact in the FRT, as shown in the following. They are listed in the lower half of Table 6.1. This includes the oscillation parameters $\theta_{13}, \theta_{12}, \Delta m^2_{21}$ and δ_{CP}, the uncertainties on the DIS neutrino and anti-neutrino cross-sections $\sigma^{\nu}_{\mathrm{DIS}}$ and $\sigma^{\bar{\nu}}_{\mathrm{DIS}}$ and two normalization parameters, N_{noise} and N_{coin}. The normalization parameters describe the noise contribution, relative to the MC prediction, and the contribution from *coincident events*, which are events featuring an atmospheric muon, coincident with a neutrino event. For N_{noise} and N_{coin}, the assumed uncertainty is 100% of the value estimated from Monte Carlo. For most of the oscillation parameters, the $3\,\sigma$-range from [26] is adapted for the following tests, while δ_{CP} is tested at $\delta_{\mathrm{CP}} = 180°$ and $\delta_{\mathrm{CP}} = 270°$.

Moreover, two parameters, that are included in the NMO fit, are not tested in the following: Δm^2_{31} and N_{ν}. For Δm^2_{31}, the $(N-1)$-tests are not clearly defined, since the NMO is given by the sign of Δm^2_{31}. Thus, fixing Δm^2_{31} as systematic parameter also fixes the NMO. Moreover, the absolute value of Δm^2_{31} differs notably for the most similar hypotheses of NO and IO, such that the absolute value and the sign of Δm^2_{31} are closely entangled. Therefore, Δm^2_{31} is not tested in the following. However, the gain from external knowledge on this parameter is discussed for the experimental result in Section 8.2.

The N_{ν} parameter is not tested, as it normalizes the number of observed neutrino events. Since the total normalization is driven by many effects and not supposed to be fully constrained by any external knowledge, but only by the observed data, it is also excluded from the following tests.

Fitting Relevance Test (FRT): The results of the FRT are shown in Figure 7.7, where for several systematic parameters (horizontal axis), the *hypothesis pull* $n_\sigma^{\mathrm{NO/IO}}$ (vertical axis) is shown according to Equations 7.5. For the hypothesis pull, positive (negative) values correspond to preferring NO (IO) by the corresponding value of Gaussian sigmas, while the dashed, horizontal, red and blue lines correspond to the baseline sensitivity for NO and IO, respectively. The blue and red markers, enframing the blue and red shaded areas, are the values obtained in the FRT by varying the injected parameter by $\pm 1\,\sigma$, but fixing it to the baseline value in the fitter. Thus, the vertical size of the shaded areas estimates the bias in an NMO measurement, where the parameter is not fitted, but off by $\pm 1\,\sigma$ from the true value. The black, solid line separates the fitted and not fitted parameters analogously to Table 6.1.

As a result, parameters with small shaded areas can be fixed in the fitter without biasing the NMO result. This is the case for all parameters left of the solid, black line.

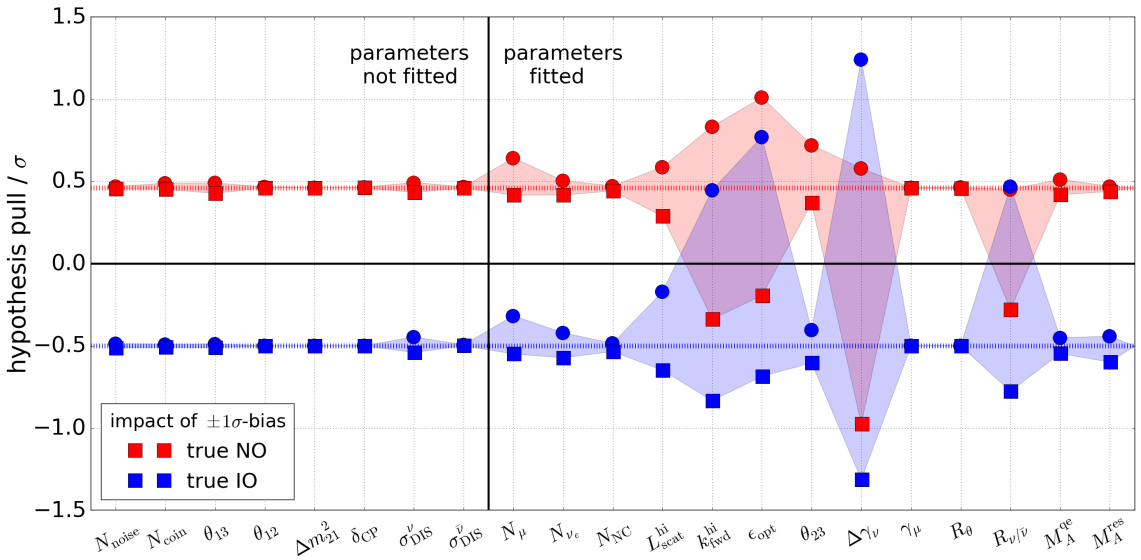


Figure 7.7: *Fitting Relevance Test* (FRT), showing the hypothesis pull (i.e. $n_\sigma^{\mathrm{NO/IO}}$ from Equation 7.5) for testing various systematic parameters (horizontal axis): the shaded red and blue areas indicate the range of hypothesis pulls obtained by fitting a hypothesis that is off the injected one by $\pm 1\,\sigma$ in this parameter; parameters with large shaded areas need to be fitted to obtain an unbiased result, while small shaded areas indicate that an unbiased result is also obtained for fixing the parameter to its baseline value (pre-published in [126]).

For the fitted parameters, the most relevant ones are the detector parameters $L_{\mathrm{scat}}^{\mathrm{hi}}$, ϵ_{opt} and $k_{\mathrm{fwd}}^{\mathrm{hi}}$ as well as the atmospheric flux parameters γ and $R_{\nu/\bar{\nu}}$. Thus, fitting these parameters is crucial to obtain an unbiased result for the NMO.

The parameters R_θ and γ_μ are found to have almost no impact on the analysis result. Still, they are included as free parameters for the following reasons:

The atmospheric flux uncertainties from [47] were previously parametrized by two other parameters, which were non-negligible in the fit of the NMO. However, a collaboration-wide correction of this parametrization lead to $R_{\nu/\bar{\nu}}$ carrying most of the relevant uncertainties, while R_θ only caused a marginal correction to the resulting zenith-spectrum. For consistency reasons, both parameters were kept, although the latter was found to have no relevance for any ongoing oscillation analysis.

The γ_μ parameter was kept, since it is the only parameter changing the shape of the atmospheric muon template. Since the prior of ± 1 is strong with respect to the sensitivity to this parameter, a strong pull is only expected in case of a large disagreement between data and Monte Carlo. Otherwise, a value close to $\gamma_\mu = 0$ is expected (cf. Appendix H.2), while large deviations could indicate an inaccurate modelling of atmospheric muons.

One should note that the relevance of fitting a parameter depends on the assumed uncertainty. For example, if the optical efficiency ϵ_{opt} was assumed to be known by $\pm 2\%$, the resulting shaded area would shrink substantially. Since the uncertainties are estimated according to or larger than the expected true uncertainty (cf. Section 6.3), their impact on the NMO is estimated conservatively.

Sensitivity Stability Test (SST): The Sensitivity Stability Test ist shown in Figure 7.8. The plot is read analogously to Figure 7.7. It shows the change in sensitivity depending on the true value of each of the systematic parameters. To do this, the true value is injected at $\pm 1\,\sigma$ (analogous to the FRT) but left as floating parameter in the subsequent likelihood fit. Thus, the fit corrects for the change in the systematic parameter. Naturally, the resulting shaded area for NO and IO is smaller than for the FRT.

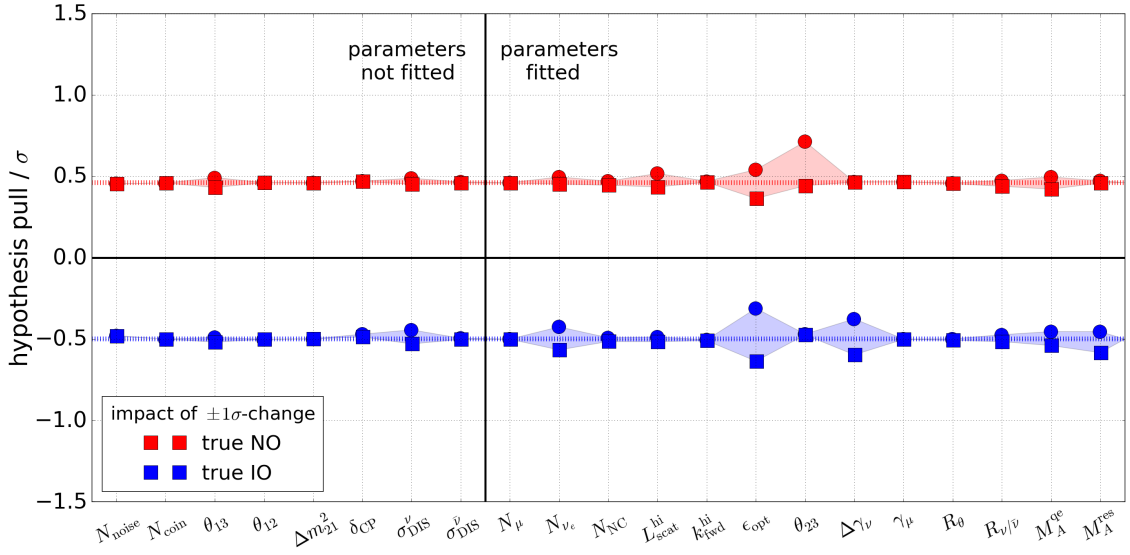


Figure 7.8: *Sensitivity Stability Test* (SST), showing the change in sensitivity from varying the value of a single systematic parameter: the vertical axis is analogous to Figure 7.7 for NO (red) and IO (blue) being injected.

Note that the injected value is the fitted one, only if it is used without prior in the fitter. The priors insert a pull on each parameter towards the baseline value, which is not the true value in the SST. Thus, the fitted parameter deviates slightly from the injected one (for example, for injecting $\epsilon_{\text{opt}} = 1.1$, the fitted value is $\epsilon_{\text{opt}} = 1.092$). However, the impact of these pulls on the NMO result is small.

As shown in Figure 7.8, the NMO sensitivity depends on the true value of only a few of the fitted parameters. One of these parameters is θ_{23}, where an effect of the true value is only visible for one of the two markers in NO. This is due to illustration reasons: The two red and two blue markers for θ_{23} in Figure 7.8 correspond to two points on the NO and IO sensitivity curve in Figure 7.4. The uncertainty on θ_{23} in Table 6.1 causes a similar value

of n_σ^{IO} for the baseline value and for $\pm 1\sigma$, which is only a coincidence and does not mean that θ_{23} has in general no impact in case of IO.

Besides θ_{23}, the optical efficiency ϵ_{opt} and the spectrum parameter $\Delta\gamma_\nu$ have the largest impact on the NMO sensitivity. For both parameters, this can be explained by their large impact on the neutrino energy spectrum, which is crucial for an NMO measurement - especially at the lowest energies. Moreover, the electron neutrino normalization N_{ν_e} has a non-vanishing impact on the sensitivity, which can be explained by the fact that the matter effects arise from interactions in the electron neutrino state. Thus, a large flux of electron neutrinos before the propagation through Earth increases the observable matter effects after propagation and vice versa.

Overall, the sensitivity to the NMO does not depend strongly on the value of any of the systematic parameters. Instead, the sensitivity is roughly stable with respect to the true value of all systematics.

Hidden Potential Test (HPT): Finally, the Hidden Potential Test estimates the potential improvement in the NMO measurement, that is obtained from knowing one of the N systematic parameters perfectly. To do this, the parameter is fixed in the fitter and the resulting gain in sensitivity is observed. This way, the fitter is free to use only $N-1$ parameters to cover the difference between NO and IO, while all N parameters are available in the standard fit.

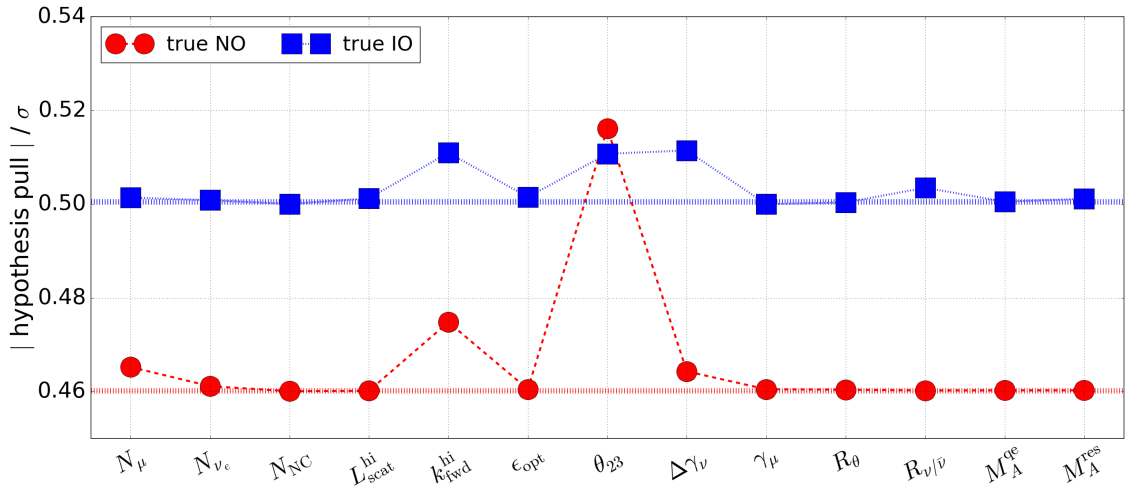


Figure 7.9: *Hidden Potential Test* (HPT), showing the increase in sensitivity from fixing a single systematic parameter: the vertical axis is similar to Figure 7.7, but shows the absolute value of the hypothesis pull from Equation 7.5 for illustration purposes.

The resulting gain in sensitivity is shown in Figure 7.9. In contrast to the FRT and SST, the vertical axis shows the absolute value of the hypothesis pull for illustration purposes. The largest gains are obtained from the mixing angle θ_{23} and the high-forward parameter $k^{\mathrm{hi}}_{\mathrm{fwd}}$, which give an increase of up to $\sim 10\%$ and $\sim 3\%$, respectively. However, the absolute gain from knowing one parameter is in general small.

The result of the HPT is sobbering: Since the knowledge of no systematic parameter increases the sensitivity substantially, external knowledge of any single systematic parameter is unlikely to increase the sensitivity crucially. However, a more severe gain can be obtained

from an improved understanding of several systematics at the same time (cf. Section 7.7).

7.7 Potential Improvements of the Sensitivity

Since the presented Neutrino Mass Ordering (NMO) sensitivity of DeepCore is small (cf. Section 7.4) and no single systematic parameter was found to affect the sensitivity strongly (cf. HPT in Section 7.6), further potential improvements of the sensitivity are investigated. This comprises:

1. testing the combined impact of several systematic uncertainties,
2. testing the impact of the atmospheric muon and noise contamination,
3. testing the impact of the resolution of the event reconstruction.

The results of all tests are summarized in Table 7.4.

First, the Hidden Potential Test (HPT) is repeated by fixing groups of systematics instead of single systematic parameters. To do this, the systematic parameters are grouped according to Table 6.1 into normalization parameters (N), atmospheric flux parameters (F), detector response parameters (D), oscillation parameters (O) and interaction parameters (I).

Each group of systematic parameters N, F, I and D is then fixed in the fitter and the resulting gain in sensitivity is calculated. Note that the oscillation parameters O are always fitted, since $\sin^2(\theta_{23})$ and Δm^2_{31} are closely connected to the NMO sensitivity, as already discussed in the previous section. Moreover, the total neutrino normalization N_ν is fitted for all of the shown cases, such that the normalization parameters determine only the composition of the different contributions, but not the overall number of events. This is done, since a pure rate difference between data and Monte Carlo should not lead to any NMO sensitivity.

The sensitivity is shown in Table 7.4 as tests 2–5 for three combinations (columns) of the oscillation parameters: in the left (first) octant, in the right (second) octant and at maximum mixing. The values in brackets show the corresponding, relative gain compared to the baseline case (test 1). For all groups of systematics, the found gain in sensitivity is small, with the largest gains being in the order of $\lesssim 6\%$ for the detector systematics. Thus, an improved knowledge on one group of systematic uncertainties is insufficient to reach a substantial gain in the NMO sensitivity.

Furthermore, the inverse test is applied by fixing all groups of systematics N, F, I and D except one, where the oscillation parameters O and the total normalization are again fitted for all cases. The resulting gain is shown as tests 6–9 in Table 7.4 and in Figure 7.10 for the full range in $\sin^2(\theta_{23})$.

As expected, the resulting gain in sensitivity is larger than for fixing only one group of systematic parameters. Moreover, the largest gain in Figure 7.10 is obtained for regions of $\sin^2(\theta_{23})$ that are already the most sensitive in the baseline scenario. As a result, the $\sin^2(\theta_{23})$-dependent features of the sensitivity curve become more pronounced.

This can be understood in the following way: If the signature of the NMO, i.e. the difference between the total NO and IO templates, is small, there is no pull on the systematic parameters due to the NMO. Instead, the fit of NO and IO return nearly the same value for all parameters, since the small difference between the orderings is insufficient to pull any of the parameters notably. Now, if both fits return the same value, the parameters are

Table 7.4: Summary of potential improvements in NMO sensitivity: test 2-5: by fixing one group of systematics; test 6-9: by fitting only one group of systematics and the oscillation parameters; test 10: by fitting only the oscillation parameters; test 11: by removing the background (atm. muons and noise); test 12-16: by improving the resolution in reconstructed neutrino energy ($\delta_{\log E}$) and zenith-angle ($\delta_{\log E}$); for each test, the sensitivity is given for both orderings in the left and right octant and at maximum mixing; values in brackets gives the relative gain w.r.t to the presented analysis (test 1).

#	applied test fitted parameters					sensitivity (NO|IO) left octant[a]	max. mixing[b]	right octant[c]
1[d]	N	F	I	D	O	$0.4371\sigma \| 0.4830\sigma$	$0.5747\sigma \| 0.5089\sigma$	$0.7160\sigma \| 0.4287\sigma$
2		F	I	D	O	$0.4486\sigma \| 0.4833\sigma$ $(+2.6\% \| +0.6\%)$	$0.5833\sigma \| 0.5100\sigma$ $(+1.5\% \| +0.2\%)$	$0.7253\sigma \| 0.4273\sigma$ $(+1.3\% \| +0.3\%)$
3	N		I	D	O	$0.4394\sigma \| 0.4938\sigma$ $(+0.5\% \| +2.2\%)$	$0.5905\sigma \| 0.5238\sigma$ $(+2.7\% \| +2.9\%)$	$0.7482\sigma \| 0.4284\sigma$ $(+4.5\% \| +0.7\%)$
4	N	F		D	O	$0.4370\sigma \| 0.4838\sigma$ $(+0.0\% \| +0.1\%)$	$0.5771\sigma \| 0.5099\sigma$ $(+0.4\% \| +0.2\%)$	$0.7184\sigma \| 0.4251\sigma$ $(+0.3\% \| +0.8\%)$
5	N	F	I		O	$0.4594\sigma \| 0.4968\sigma$ $(+5.1\% \| +2.9\%)$	$0.5859\sigma \| 0.5304\sigma$ $(+1.9\% \| +4.2\%)$	$0.7482\sigma \| 0.4514\sigma$ $(+4.5\% \| +5.3\%)$
6	N				O	$0.4735\sigma \| 0.5678\sigma$ $(+8.3\% \| +17.6\%)$	$0.6414\sigma \| 0.6267\sigma$ $(+11.6\% \| +23.1\%)$	$0.8762\sigma \| 0.4737\sigma$ $(+22.4\% \| +10.5\%)$
7		F			O	$0.4631\sigma \| 0.5022\sigma$ $(+5.9\% \| +4.9\%)$	$0.6056\sigma \| 0.5379\sigma$ $(+5.4\% \| +5.7\%)$	$0.7561\sigma \| 0.4543\sigma$ $(+5.6\% \| +5.9\%)$
8			I		O	$0.4637\sigma \| 0.5277\sigma$ $(+6.1\% \| +9.3\%)$	$0.6227\sigma \| 0.5648\sigma$ $(+8.3\% \| +11.0\%)$	$0.8037\sigma \| 0.4574\sigma$ $(+12.2\% \| +6.7\%)$
9				D	O	$0.4838\sigma \| 0.5877\sigma$ $(+10.7\% \| +21.7\%)$	$0.6977\sigma \| 0.6772\sigma$ $(+21.4\% \| +33.1\%)$	$0.9827\sigma \| 0.4573\sigma$ $(+37.2\% \| +6.7\%)$
10					O	$0.4873\sigma \| 0.6266\sigma$ $(+11.5\% \| +29.7\%)$	$0.7145\sigma \| 0.6934\sigma$ $(+24.3\% \| +36.3\%)$	$0.9989\sigma \| 0.4940\sigma$ $(+39.5\% \| +15.2\%)$
11	N	F	I	D	O	$0.4576\sigma \| 0.4852\sigma$	$0.5876\sigma \| 0.5249\sigma$	$0.7527\sigma \| 0.4420\sigma$
	no background					$(+4.7\% \| +0.5\%)$	$(+2.2\% \| +3.1\%)$	$(+5.1\% \| +3.1\%)$
12	N	F	I	D	O	$0.5453\sigma \| 0.5706\sigma$	$0.6808\sigma \| 0.6068\sigma$	$0.8231 \| 0.5446$
	$\delta_{\log E} \downarrow 25\%$					$(+24.8\% \| +18.1\%)$	$(+18.5\% \| +19.2\%)$	$(+43.2\% \| +7.0\%)$
13	N	F	I	D	O	$0.7161\sigma \| 0.7055\sigma$	$0.8184\sigma \| 0.7737\sigma$	$0.9638\sigma \| 0.6766\sigma$
	$\delta_{\log E} \downarrow 50\%$					$(+63.8\% \| +46.1\%)$	$(+42.4\% \| +52.0\%)$	$(+67.7\% \| +33.0\%)$
14	N	F	I	D	O	$0.6267\sigma \| 0.5781\sigma$	$0.7421\sigma \| 0.6130\sigma$	$0.8689\sigma \| 0.5369\sigma$
	$\delta_\theta \downarrow 25\%$					$(+43.4\% \| +19.6\%)$	$(+29.1\% \| +20.5\%)$	$(+51.1\% \| +5.5\%)$
15	N	F	I	D	O	$0.8150\sigma \| 0.7592\sigma$	$0.8988\sigma \| 0.7923\sigma$	$1.0203\sigma \| 0.7028\sigma$
	$\delta_\theta \downarrow 50\%$					$(+86.5\% \| +57.1\%)$	$(+56.4\% \| +55.7\%)$	$(+77.5\% \| +38.1\%)$
16	N	F	I	D	O	$0.9532\sigma \| 0.9797\sigma$	$1.0739\sigma \| 1.0701\sigma$	$1.1552\sigma \| 0.9509\sigma$
	$\delta_{\log E}, \delta_\theta \downarrow 50\%$					$(+118\% \| +103\%)$	$(+86.9\% \| +110\%)$	$(+101\% \| +86.9\%)$

[a] at $\sin^2(\theta_{23}) = 0.43$, $\Delta m^2_{31} = 2.5 \cdot 10^{-3}\,\mathrm{eV}^2$ (NO), $\Delta m^2_{31} = -2.424 \cdot 10^{-3}\,\mathrm{eV}^2$ (IO)
[b] at $\sin^2(\theta_{23}) = 0.50$, $\Delta m^2_{31} = 2.5 \cdot 10^{-3}\,\mathrm{eV}^2$ (NO), $\Delta m^2_{31} = -2.424 \cdot 10^{-3}\,\mathrm{eV}^2$ (IO)
[c] at $\sin^2(\theta_{23}) = 0.57$, $\Delta m^2_{31} = 2.5 \cdot 10^{-3}\,\mathrm{eV}^2$ (NO), $\Delta m^2_{31} = -2.424 \cdot 10^{-3}\,\mathrm{eV}^2$ (IO)
[d] standard fit using all parameters

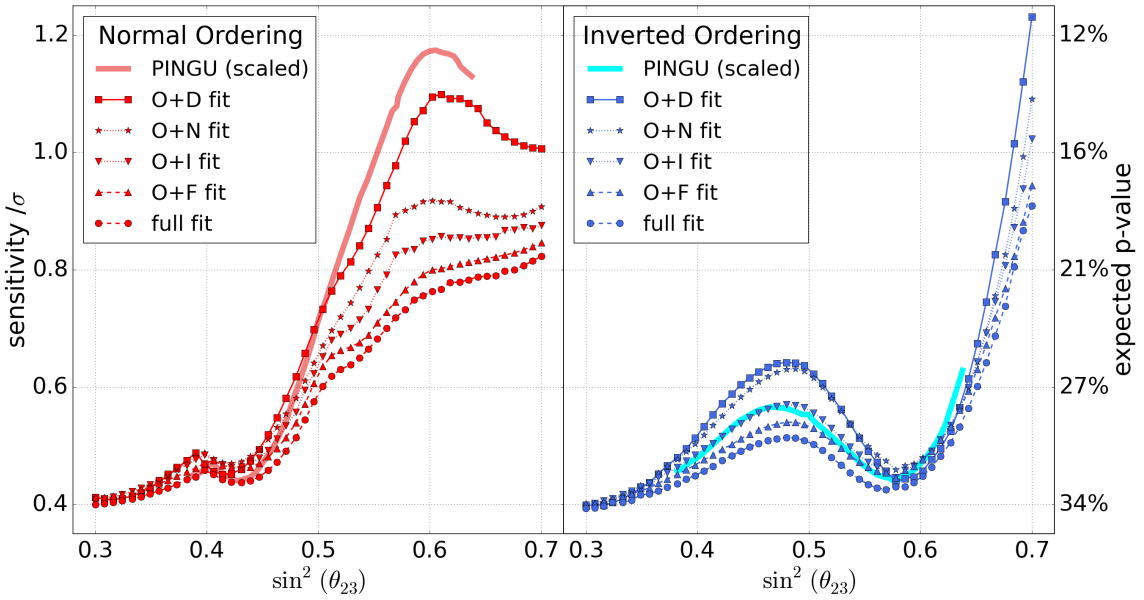


Figure 7.10: Sensitivities for fitting only the total normalization, the oscillation parameters (O) and one of the parameter groups from Table 6.1 (N,F,I,D); for comparison, the projected PINGU sensitivity [7] is shown, scaled by 0.17 to illustrate the shape differences between DeepCore and PINGU.

not used to compensate for the difference between the orderings and thus, do not affect the NMO sensitivity. In other words, the signature of the NMO is too weak to affect the fit of the systematic parameters and thus, fixing these parameters does not cause a major change in the fit.

Additionally, Figure 7.10 shows the projected sensitivity for PINGU, obtained from [7] and scaled by a factor of 0.17. This is done only to illustrate how shape differences can arise from variations in the fitted nuisance parameters, while the PINGU scaling is arbitrary. However, there are of course more differences between DeepCore and PINGU, like the observed energy range and the reconstruction resolution.

Finally, test 10 shows the sensitivity for the case of fitting only the oscillation parameters O and the total normalization. Naturally, the resulting sensitivity outperforms the sensitivities obtained by the tests 2–9. However, even with the observed gain of $\sim 12\% - 40\%$, the resulting sensitivity does not exceed the 1σ-threshold.

Overall, these tests illustrate that the absolute gain in sensitivity from fixing systematic parameters is small, although the relative gain might reach $\sim 40\%$ for the most progressive cases. Therefore, any improvement in the understanding of systematic uncertainties is unlikely to result in a substantial gain in the NMO sensitivity for DeepCore, unless it is combined with further improvements in the analysis.

Second, the impact of background on the NMO sensitivity is tested in test 11 of Table 7.4. To do this, the background (atmospheric muons and triggered noise) is removed from Monte Carlo and the corresponding systematic parameters (N_μ and γ_μ) are fixed in the fitter. This is supposed to test the impact of a perfect background veto, removing all events caused by atmospheric muons and triggered noise. As visible in Table 7.4, the resulting gain in sensitivity is on the $\lesssim 5\%$-level, such that the background contamination has only negligible effect on the NMO sensitivity.

Third, the potential gain from an improved event reconstruction is investigated. This is motivated by the weak resolutions in neutrino energy and zenith-angle at $E_\nu < 15\,\mathrm{GeV}$ (cf. Figure 5.5 and 5.6). To test such potential improvements, the reconstruction errors $\delta_{\log E}$ and δ_θ of the neutrino energy E_ν and zenith-angle θ_ν are calculated for Monte Carlo according to Equation 7.8. Then, new values for the reconstructed zenith angle $\tilde{\theta}_\nu^{\mathrm{reco}}$ and energy $\tilde{E}_\nu^{\mathrm{reco}}$ are calculated according to Equation 7.9:

$$\delta_{\log E} = \log(E_\nu^{\mathrm{reco}}/\mathrm{GeV}) - \log(E_\nu/\mathrm{GeV}) \qquad \delta_\theta = \theta_\nu^{\mathrm{reco}} - \theta_\nu \tag{7.8}$$

$$\tilde{E}_\nu^{\mathrm{reco}} = \exp\left(\log(E_\nu/\mathrm{GeV}) + \alpha_{\log E} \cdot \delta_{\log E}\right)\,\mathrm{GeV} \qquad \tilde{\theta}_\nu^{\mathrm{reco}} = \theta_\nu + \alpha_\theta \cdot \delta_\theta. \tag{7.9}$$

This way, the reconstructed observables are shifted towards their *true* value with a shifting parameter $\alpha \in \{\alpha_{\log E}, \alpha_\theta\}$. Thus, for $\alpha = 1$ ($\alpha = 0$) one obtains the previous reconstruction (the true value). One should note that this modifies the resolution of the energy and zenith-angle in the same way for all events. Although this is an unrealistic assumption for future developments, e.g. with the PINGU detector, it is used here as a benchmark for the impact of possible improvements in the reconstruction.

Moreover, the above procedure is applied only to neutrinos, while atmospheric muons and triggered noise are kept unchanged in their reconstructed observables. Naturally, an improved reconstruction, e.g. in the zenith-angle, would also allow for a better suppression of atmospheric muons. However, choosing $\alpha_\theta = 0.5$ would remove almost all atmospheric muons from the template, as they were shifted into the downgoing region, such that the remaining muons are insufficient to obtain a suitable Monte Carlo template. Moreover, the *true* energy is not well-defined for atmospheric muons, as they are produced far outside the DeepCore volume with initial energies of $E_\mu \gg 100\,\mathrm{GeV}$. Thus, the muon energy and zenith resolutions are kept unchanged for the following tests.

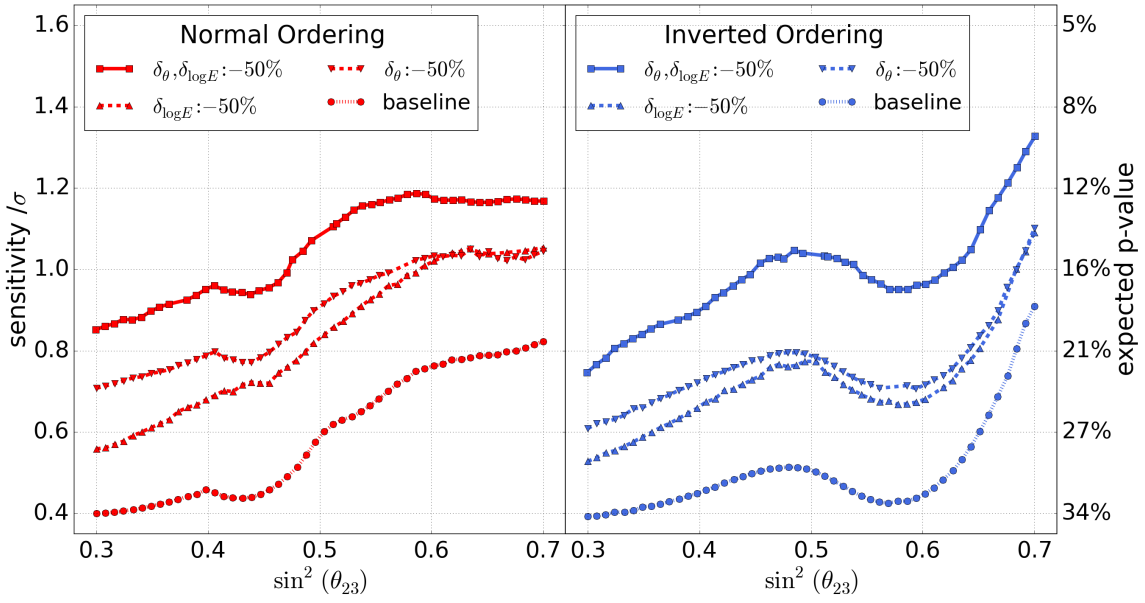


Figure 7.11: Sensitivities obtained by scaling the resolutions of the event reconstruction in $\log(E_\nu)$ and θ_ν by $\alpha = 0.5$ (δ_θ,$\delta_{\log E_\nu}$: -50%) for NO (left) and IO (right).

From the resulting set of reconstructed observables, the Asimov sensitivities are re-calculated according to Section 7.4. To do this, the detector systematics are re-parametrized by applying the above procedure to all systematic samples. This way, the whole analysis chain is repeated with the modified reconstructions $\tilde{E}_\nu^{\mathrm{reco}}$ and $\tilde{\theta}_\nu^{\mathrm{reco}}$.

The resulting gain in sensitivity is shown in Table 7.4 as tests 12–16 for reducing the reconstruction errors δ_E and δ_θ to $\alpha = 0.75$ ($\downarrow 25\%$) and $\alpha = 0.5$ ($\downarrow 50\%$). Moreover, they are shown for $\alpha = 0.5$ in Figure 7.11 for the full range of $\sin^2(\theta_{23})$.

The gain in sensitivity is typically large compared to the gain from fixing systematic parameters. Moreover, it is clearly visible over all the range of $\sin^2(\theta_{23})$, since an improved event reconstruction increases the sensitivity for all values of the oscillation parameters. In case of $\alpha_{\log E} = 0.5$ and $\alpha_\theta = 0.5$, the resulting NMO sensitivity is at the $1\,\sigma$-level for all values of $\sin^2(\theta_{23})$. Although this is still a small, absolute sensitivity, the relative gain with respect to the baseline sensitivity is $> 100\%$ for some cases.

In addition, the sensitivity can be increased by an improved particle separating variable (PID). However, defining an artificial PID improvement, analogous to Equation 7.9 for zenith-angle and energy, is complicated, because the PID encodes information about the neutrino flavor, energy, interaction type and its neutrino or anti-neutrino nature. Instead, a realistic, improved PID was developed by Marvin Beck [151] and Saskia Philippen [152], using machine-learning techniques. In these first, explorative studies of the PID, a potential gain of $\sim 20\%$ was achieved for the sensitivity of this NMO analysis. However, such improved PID parameter requires further investigations.

In summary, an improved understanding of systematic uncertainties is unlikely to increase the DeepCore sensitivity substantially, as long as the observed signature from the NMO in DeepCore is too weak to be affected by them. However, the signature depends on the resolution of the event reconstruction in energy and zenith-angle. In case these could be improved notably in the future, a substantial gain in the NMO sensitivity with DeepCore may be obtained. Moreover, the sensitivity can be increased by an improved PID, separating track- and cascade-like events.

For the proposed PINGU detector (cf. Section 3.6), several of the above mentioned improvements are planned [7]:

1. Additional calibration devices are supposed to be used during the detector deployment and afterwards to reduce the uncertainties on the detector response.
2. The surrounding DeepCore detector allows for an efficient veto of atmospheric muons, reducing the background for oscillation analyses. Additionally, due to the increased number of optical modules, that are required to trigger the detector, triggered noise is less likely to occur than for DeepCore.
3. Events at $E_\nu \sim 5 - 10\,\text{GeV}$ will on average cause more light being observed in PINGU than in DeepCore, which allows for an improved reconstruction of these events.
4. The more densly instrumented volume and the potentially improved optical modules will lower the energy threshold to $E_\nu \sim 1\,\text{GeV}$, such that neutrino oscillations and matter effects can be observed over an extended energy range [7].

Thus, although a highly significant measurement of the NMO with DeepCore is unlikely, PINGU will be capable to reach $\sim 3\,\sigma$ sensitivity within four years, which is a comparable time of data-taking [7].

For DeepCore, an additional gain might come from statistics: While the presented analysis includes three years of data, eight years of data will be available soon. However, the combination of all eight years of data in a single analysis is non-trivial for calibration reasons, which are not discussed in more detail here. Besides the extended livetime, recent studies

indicate that a statistical gain of $\sim 50 - 100\%$ may be reached by improvements in the data selection. Thus in an optimistic scenario, a ~ 4 times larger DeepCore sample may be available within the next years.

This increase in statistics would lead to a factor of ~ 2 gain in sensitivity, as the sensitivity scales $\sqrt{N}$ with the number of data events N. The relation holds, because all systematic uncertainties are fitted by nuisance parameters, such that any increase in statistics allows for a more accurate determination of the systematic parameters. As a result, in the limit of infinite data statistics all systematic parameters would be known precisely. Thus, a systematic limit to the sensitivity is not present, as long as the parametrization of systematics is reliable and complete. Otherwise, the uncertainty in the parametrization and the existence of unknown systematics limits the potential gain from increased statistics. For future analyses with increased statistics, such limitation must still be investigated.

8 Fit of Experimental Data

8.1 Independent Fit of DeepCore Data

The experimental data, described in Section 4.3, is fit with the likelihood method described in Section 7.1. The results are published in [127] and [116] with a more detailed discussion presented here.

The best-fit parameters are listed in Table 8.1 for the hypothesis of matter effects in Normal (NO) and Inverted (IO) Ordering and vacuum oscillations (VA) separately. Note that in the following, matter effects (MA) are assumed implicitly for the NO and IO fits. Moreover, for the fit of vacuum oscillations (VA), the Neutrino Mass Ordering (NMO) is assumed to be Normal ($\Delta m_{31} > 0$), since for DeepCore the two ordering hypotheses are degenerated without matter effects, such that $\mathcal{H}_{\mathrm{NO}} \approx \mathcal{H}_{\mathrm{IO}}$ (cf. Section 7.3).

Table 8.1: Overview of the best-fit values obtained from the fit of the Normal Ordering (NO), Inverted Ordering (IO) and vacuum oscillations (VA) hypotheses; these values are used for the generation of Pseudo-Experiments in Section 8.1.1.

		convolved LLH			Poissonian LLH		
parameter	prior	fit NO	fit IO	fit VA	fit NO	fit IO	fit VA
N_ν	-	0.832	0.836	0.819	0.830	0.835	0.820
N_{ν_e}	1.00 ± 0.05	1.004	1.004	1.005	1.004	1.004	1.006
N_{NC}	1.0 ± 0.2	0.741	0.749	0.734	0.740	0.747	0.735
N_μ	-	1.348	1.339	1.360	1.350	1.343	1.359
ϵ_{opt}	1.0 ± 0.1	1.003	1.002	1.008	1.003	1.003	1.007
$L^{\mathrm{hi}}_{\mathrm{scat}}$	25 ± 10	31.84	31.76	32.11	31.80	31.65	32.16
$k^{\mathrm{hi}}_{\mathrm{fwd}}$	-	-1.014	-1.014	-1.131	-0.973	-1.012	-1.135
$10^3 \Delta m^2_{31}/\mathrm{eV}^2$	-	2.626	−2.511	2.516	2.625	-0.2511	2.518
$10^3 \Delta m^2_{32}/\mathrm{eV}^2$	-	2.551	−2.585	2.442	2.550	-2.585	2.444
θ_{23}/rad	-	0.7611	0.7701	0.7892	0.7608	0.7701	0.7849
$\sin^2(\theta_{23})$	-	0.4757	0.4847	0.5038	0.4754	0.4847	0.4995
$\Delta\gamma_\nu$	0.0 ± 0.1	0.0727	0.0705	0.0735	0.0731	0.0711	0.0735
R_θ	0.0 ± 1.0	−0.120	−0.113	-0.124	-0.111	-0.111	-0.110
$R_{\nu/\bar{\nu}}$	0.0 ± 1.0	−1.034	−1.019	-1.029	-1.037	-1.024	-1.030
γ_μ	0.0 ± 1.0	0.0409	0.0406	0.0322	0.0460	0.0369	0.0265
$M^{\mathrm{res}}_A/\mathrm{GeV}$	1.12 ± 0.22	1.091	1.095	1.126	1.091	1.096	1.117
$M^{\mathrm{qe}}_A/\mathrm{GeV}$	0.99 ± 0.25	0.862	0.867	0.892	0.868	0.875	0.909
LLH	-	146.692	147.061	147.561	148.008	148.373	148.886

For confirmation, the VA fit was repeated with $\Delta m_{31} < 0$: The log-likelihood difference between the two orderings was found to be more than an order of magnitude below the already small log-likelihood difference in case of matter effects. Thus, it is neglected in the following.

Since the NO, IO and VA hypotheses differ only slightly, the resulting values for all systematic parameters are very similar. For comparison, the fit is repeated with the Poissonian likelihood (cf. Section 6.4), i.e. without including the uncertainties from limited Monte Carlo. This is done to quantify the impact of limited MC statistics on the final experimental result, which is found to be small: The resulting values for all systematic parameter and the value of LLH are almost identical with only small deviations, as shown in Table 8.1.

The pulls on all fitted parameters, except for Δm^2_{31} and $\sin^2(\theta_{23})$, are shown in Figure 8.1 with respect to the baseline values and uncertainties from Table 8.1, while the obtained oscillation parameters are discussed in Section 8.1.2. The parameters used with (without) a Gaussian prior in the likelihood are indicated by the green (grey) background color. Most parameters are fitted within their prior width, while only the NC normalization leads to an undershooting, notably below $-1\,\sigma$. However, N_{NC} has only small impact on the NMO measurement, as found in Section 7.6. Moreover, the NC events are only a small contribution to the total sample and degenerated with the electron neutrino normalizations, since both lead to normalization changes in the cascade channel. Thus, the result is unaffected by the observed pull in N_{NC}.

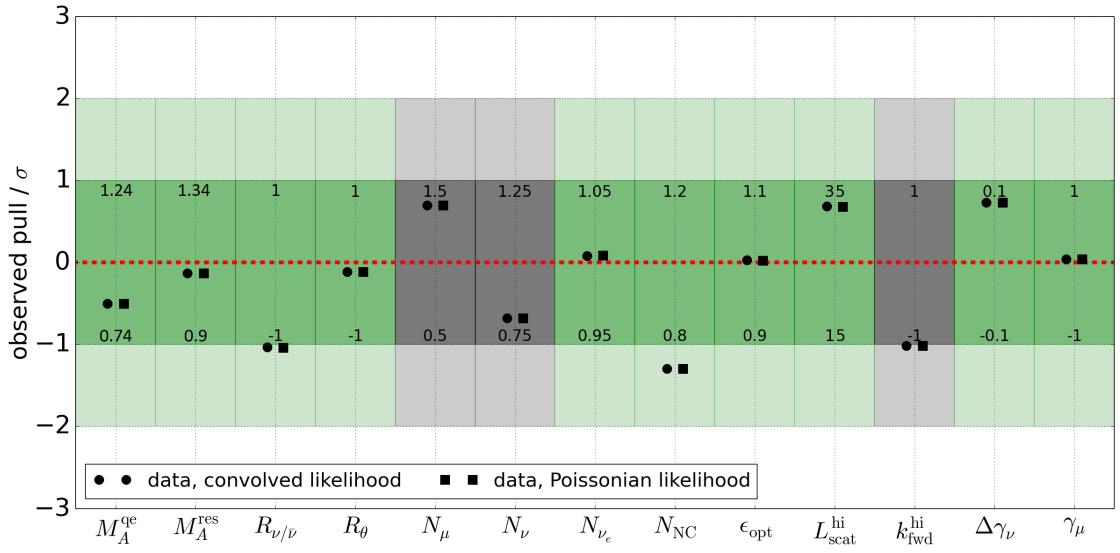


Figure 8.1: Best-fit values of the systematic parameters for fitting data with the convolved likelihood (cf. Equation 6.7) and the Poissonian likelihood (cf. Equation 6.8): each column corresponds to one systematic parameter from Table 8.1, while the vertical axis gives the pull with respect to the baseline value of each parameter; the green (grey) background color indicates parameters used with (without) Gaussian prior in the likelihood.

Besides the NC normalization, the $R_{\nu/\bar{\nu}}$ parameter shows a $\sim -1\,\sigma$ deviation from the baseline value. Although the pull is small, it might indicate that the properties of the atmospheric flux, especially the neutrino to anti-neutrino ratio, might deviate from the baseline value, but could be constraint by a measurement similar to this one.

Moreover, the atmospheric muon normalization N_μ shows some excess compared to the baseline value. However, the total muon contamination is low and thus the statistical

fluctuations on this parameters are large, as shown in Appendix H.2. The observed deviation is consistent with statistical fluctuations. Moreover, the simulation based on MuonGun and CORSIKA (cf. Section 4.2.2) features uncertainties in the total normalization of the simulated flux, which are non-negligible [112]. Thus, the observed normalization is consistent with the expectations from Monte Carlo simulations.

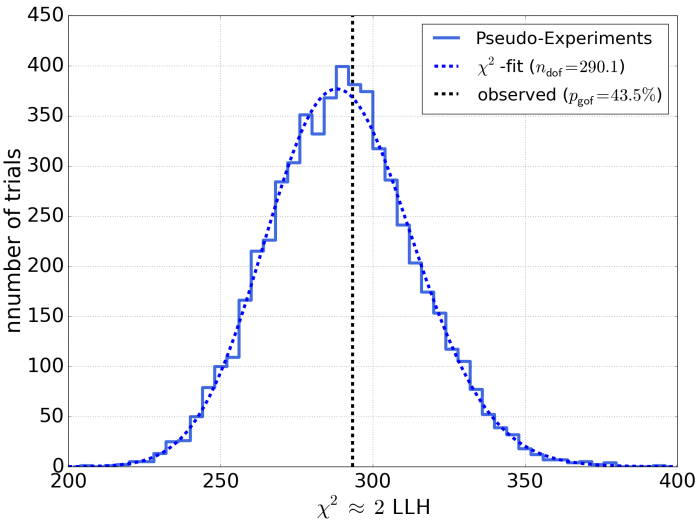


Figure 8.2: 2LLH-distribution used to estimate the *goodness-of-fit* (GOF): the blue histogram is obtained from Pseudo-Experiments, while the vertical, black line marks the experimental result with a p-value of 43.5%.

Finally, the agreement between data and the MC description is tested using 2LLH as a *goodness-of-fit* estimator [144], which is shown in Figure 8.2. Note that in case of Gaussian statistics for the bin-content of all bins, one would obtain $\chi^2 = 2\mathrm{LLH}$. Thus, the test is analogous to testing the goodness-of-fit in a χ^2-method.

The observed value for the experimental fit is shown as a vertical, black, dashed line. The blue background distribution is obtained by generating Pseudo-Experiments with the best-fit values of the NO hypothesis. For illustration, a χ^2-distribution (dashed, blue line) is fitted to the histogram, giving $n_{\mathrm{dof}} \approx 290$ for its scale parameter, also known as *number of degrees of freedom*. As expected, the 2LLH-histogram is well-described by the χ^2-distribution [125]. Moreover, $n_{\mathrm{dof}} \approx 290$ is compatible with the expectation for 300 bins in the analysis histogram and 15 parameters being optimized (cf. Table 8.1). Note that some parameters have only small impact on the analysis distribution, are mostly constrained by priors or are largely degenerated with other ones. As a result, the number of effective parameters $300 - n_{\mathrm{dof}} \approx 10$ is slightly smaller than the actual number of fitted parameters.

A disagreement between data and Monte Carlo would lead to a shift of the observed value to the right hand side. This is not observed, while the p-value of 43.5% is well-compatible with statistical fluctuations.

This χ^2-method is the conventional goodness-of-fit estimator used for DeepCore analysis, requiring a p-value of $> 5\%$ to accept the fit. More detailed studies on the observed pulls for the analysis distribution are added to Appendix I.1.

8.1.1 Results on Neutrino Mass Ordering and Matter Effects

The fit of experimental data prefers Normal over Inverted Ordering by $\Delta\mathrm{LLH}_{\mathrm{NO-IO}} = -0.369$. Moreover, it prefers the left (first) octant. The likelihood values for both orderings are given in Table 8.1.

To estimate the resulting p-values for both orderings, Pseudo-Experiments are generated with the best-fit values for both cases (cf. Table 8.1). As described in Section 7.1, the Pseudo-Experiments are then fitted in both orderings and octants and the resulting distribution of $\Delta\text{LLH}_{\text{NO}-\text{IO}}$ is used to derive the experimental p-values.

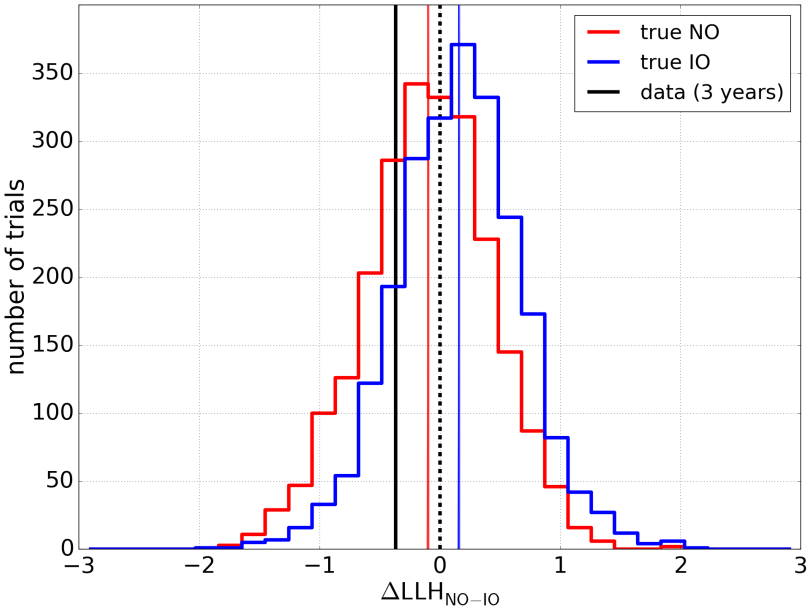


Figure 8.3: $\Delta\text{LLH}_{\text{NO}-\text{IO}}$-distribution used as test statistic to separate NO and IO (cf. Section 7.1): the red and blue distributions are obtained from Pseudo-Experiments generated with the values from Table 8.1, while the vertical, black, solid line marks the experimental result; the red and blue vertical, solid lines mark the mean value of the NO and IO distributions, respectively (pre-published in [127, 116]).

The distributions of $\Delta\text{LLH}_{\text{NO}-\text{IO}}$ obtained for NO and IO are shown in Figure 8.3 as red and blue histograms, respectively. The experimental value is indicated by a solid, black, vertical line. The resulting quantiles q and p-values p for NO and IO are (cf. Section 7.1):

$$\begin{aligned} q_{\text{NO}} &= 28.9\% & q_{\text{IO}} &= 15.3\%, \\ p_{\text{NO}} &= 1 - q_{\text{NO}} = 71.1\% & p_{\text{IO}} &= q_{\text{IO}} = 15.3\%, \end{aligned} \tag{8.1}$$

where q_{NO} and q_{IO} are the quantiles of the NO and IO distributions, that are left of the experimental value.

Note that the fit could return small quantiles q (or $1 - q$) for both orderings, if both hypotheses were unlikely to produce the value observed in data. For example, a value of $\Delta\text{LLH}_{\text{NO}-\text{IO}} > 2$ or $\Delta\text{LLH}_{\text{NO}-\text{IO}} < -2$ is neither compatible with the NO nor IO hypothesis. In that case, q_{NO} and q_{IO} may differ only slightly, while the exclusion of the disfavored hypothesis is very strong in terms of the p-value.

A common way to account for this is the calculation of CL_s-values, according to Equation 8.2 [153]

$$\text{CL}_s\left(\mathcal{H}_1\right) = \frac{q_1}{q_2} \qquad \text{CL}_s\left(\mathcal{H}_2\right) = \frac{1 - q_2}{1 - q_1}, \tag{8.2}$$

where q_i with $i = 1{,}2$ are the quantiles of the two tested hypotheses. The method is adapted from [153]. Note that the quantiles are defined such that $q_1 \leqslant q_2$, which can always be

obtained by switching the definition of $\mathcal{H}_1$ and $\mathcal{H}_2$ or by replacing $q_i \to 1 - q_i$. This leads to $\mathrm{CL}_s \approx 0$, if $\mathcal{H}_2$ is prefered over $\mathcal{H}_1$, while values $\mathrm{CL}_s \approx 1$ allow for no statement concerning the two hypotheses. In a more illustrative manner, the CL_s-value determines, *how much less likely* the observed excess is obtained under the assumption of the disfavored hypothesis, compared to the favored one.

The resulting CL_s-values for the p-values from Equation 8.1 are

$$\mathrm{CL}_s\left(\mathcal{H}_{\mathsf{NO}}\right) = \frac{1 - q_{\mathrm{IO}}}{1 - q_{\mathrm{NO}}} = 83.0\%, \tag{8.3}$$

$$\mathrm{CL}_s\left(\mathcal{H}_{\mathsf{IO}}\right) = \frac{q_{\mathrm{NO}}}{q_{\mathrm{IO}}} = 53.3\%. \tag{8.4}$$

Thus, the preference for NO over IO becomes visible in $\mathrm{CL}_s\left(\mathcal{H}_{\mathsf{IO}}\right) = 53.3\%$, stating that the observed value is roughly twice as likely to occur in case the true ordering is Normal than in case the true ordering is Inverted.

Although not significant itself, one should note that the observed preference for NO over IO is in-line with observations by MINOS/MINOS+ [68], T2K [66], Super-Kamikande [71] and NoνA [70], which all report a minor preference for NO over IO in their most recent results. However, even a potentially more significant, global combination of these results should not be taken as a final statement on the NMO, but rather as an indication in the absence of more sensitive, future experiments. Unfortunatly, the global best-fits used in Section 8.2 do not yet include all of the above results.

In addition to testing the two ordering hypotheses, matter effects (MA) are tested against vacuum oscillations (VA), as described in Section 7.3. The best-fit parameters for the VA hypothesis are listed in Table 8.1.

From the LLH values in Table 8.1, the negative log-likelihood differences $\mathrm{LLH}_{\mathrm{NO-VA}} = -0.869$ and $\mathrm{LLH}_{\mathrm{IO-VA}} = -0.500$ are calculated. Analogously to the previous test of the NMO, Pseudo-Experiments are generated for the best-fit VA hypothesis. For the MA hypothesis, the previously generated Pseudo-Experiments for the best-fit NO and IO parameters are fitted with vacuum oscillations to calculate the above log-likelihood differences.

The distribution of $\Delta\mathrm{LLH}_{\mathrm{NO-VA}}$ is shown in Figure 8.4 (left) for injecting matter effects in Normal Ordering and vacuum oscillations, while Figure 8.4 (right) shows the distribution of $\Delta\mathrm{LLH}_{\mathrm{IO-VA}}$ for injecting matter effects in Inverted Ordering and vacuum oscillations. As one can see, matter effects are preferred over vacuum oscillations for both orderings, while the preference is larger for Normal than for Inverted Ordering being assumed.

The resulting p- and CL_s-values for the hyothesis of matter effects, in Normal (NO) and Inverted Ordering (IO), and vacuum oscillations (VA) are given by Equations 8.5.

$$\begin{aligned}
p\left(\mathcal{H}_{\mathsf{MA}}|\mathcal{H}_{\mathsf{NO}}\right) &= 62.3\% & \qquad p\left(\mathcal{H}_{\mathsf{MA}}|\mathcal{H}_{\mathsf{IO}}\right) &= 53.2\% \\
p\left(\mathcal{H}_{\mathsf{VA}}|\mathcal{H}_{\mathsf{NO}}\right) &= 12.3\% & p(\mathcal{H}_{\mathsf{VA}}|\mathcal{H}_{\mathrm{IO}}) &= 22.2\% \\
\mathrm{CL}_s\left(\mathcal{H}_{\mathsf{MA}}|\mathcal{H}_{\mathrm{NO}}\right) &= 71.0\% & \mathrm{CL}_s\left(\mathcal{H}_{\mathsf{MA}}|\mathcal{H}_{\mathrm{IO}}\right) &= 68.4\% \\
\mathrm{CL}_s\left(\mathcal{H}_{\mathsf{VA}}|\mathcal{H}_{\mathrm{NO}}\right) &= 32.6\% & \mathrm{CL}_s\left(\mathcal{H}_{\mathsf{VA}}|\mathcal{H}_{\mathrm{IO}}\right) &= 47.4\%
\end{aligned} \tag{8.5}$$

The observed preference for matter effects over vacuum oscillations is natural in case NO is preferred over IO. As summarized in Table 2.2, NO and IO differ mainly by the amount of matter effects, while NO features more matter effects than IO. Thus, a preference of NO over IO already indicates a preference for strong matter effects. As a result, the VA hypothesis is disfavored compared to the MA hypothesis.

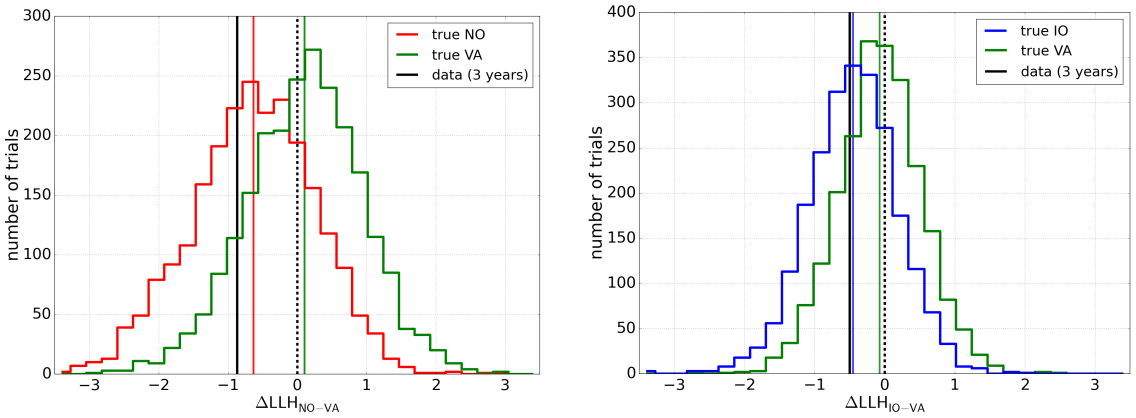


Figure 8.4: ΔLLH_{MA-VA}-distribution used as test statistic to separate matter effects and vacuum oscillations (VA) in case of NO (left) and IO (right); analogous to Figure 8.3, the red, green and blue distributions are obtained from Pseudo-Experiments generated with the best-values from Table 8.1, while the vertical, black, solid line marks the experimental fit.

Note that in Figure 8.4, the distributions seem to be slightly shifted towards negative values. The effect arises from the different mixing angles obtained for the MA and VA hypotheses in Table 8.1: If a template is generated with matter effects, but fitted with vacuum oscillations, the vacuum fit is biased towards maximum mixing. Thus, the VA hypothesis is fitted closer to $\sin^2(\theta_{23}) = 0.5$ than the NO and IO fits using matter effects. In the observed case, the NO and IO fits are already close to maximum mixing. As a result, the VA fit gets stuck at the parametric boundary at $\sin^2(\theta_{23}) = 0.5$ (cf. Section 2.3.2), losing its freedom to account for over- and under-fluctuations in the same way. Since the fit is incapable of increasing the oscillation amplitude beyond maximum mixing, the hypotheses using matter effects are at an advantage, as they can vary the mixing angle to compensate for over- or under-fluctuations in the oscillation region.

Note that this effect is to first order independent of the injected hypothesis, since for every injected hypothesis, both tested hypotheses (MA and VA) are fitted. As a result, all distributions of ΔLLH_{MA-VA} are shifted slightly to the left.

8.1.2 Results on Oscillation Parameters

The best-fit oscillation paramters Δm^2_{31} and $\sin^2(\theta_{23})$ are stated in Table 8.1 for NO and IO. To visualize the $\sin^2(\theta_{23})$-dependent preference for NO over IO, the likelihood space is scanned with respect to $\sin^2(\theta_{23})$, while for each value all other systematic parameters are optimized. The resulting values of LLH for NO and IO are shown in Figure 8.5, relative to the global minimum LLH_{min}.

The best-fit value of $\sin^2(\theta_{23})$ is compatible with maximum mixing, while featuring only a weak preference for the left octant. The preference for NO over IO is visible by the offset between the NO and IO curves at their individual minima. For comparison, the upper subplot shows the $\Delta\chi^2$-contours of the global fit of oscillation parameters provided by *NuFit* (cf. Section 2.3) [26, 25]. As one can see, the curves are compatible within their uncertainties. Moreover, they agree in their preference for NO, while global fits indicate a slight preference for the right (second) octant, which is not the case for this work.

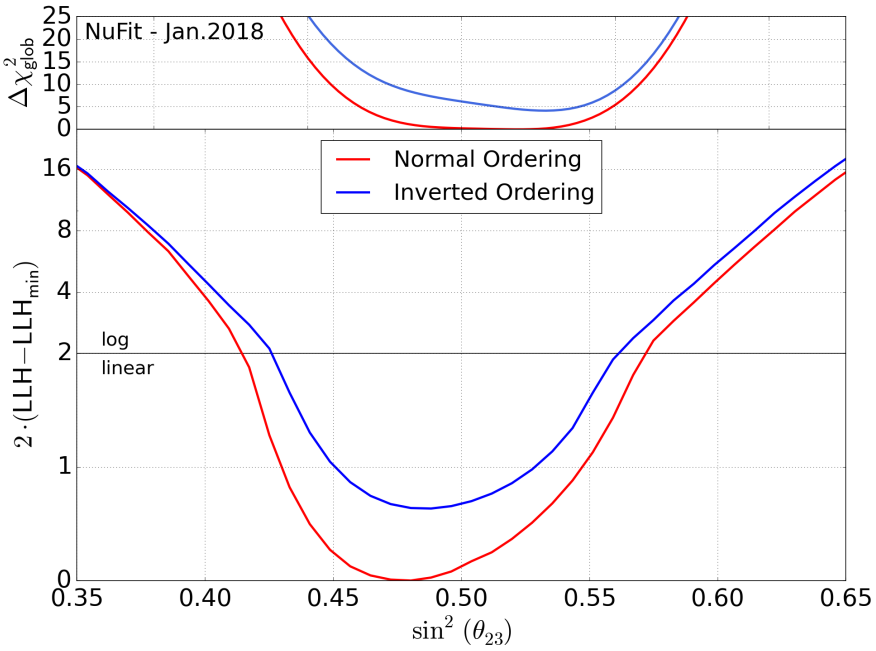


Figure 8.5: One-dimensional log-likelihood scan in $\sin^2(\theta_{23})$ for NO (red) and IO (blue), showing the $\sin^2(\theta_{23})$-dependent preference for the NMO; for each point on the curve, all other systematic parameters are optimized; in the upper subplot, the $\Delta\chi^2_{\text{glob}}$-scan in $\sin^2(\theta_{23})$ is shown for the global best-fit, provided by *NuFit* [26, 25] (pre-published in [127, 116]).

Besides the one-dimensional scan in the mixing angle, a full two-dimensional scan of the log-likelihood landscape is done in $\sin^2(\theta_{23})$ and Δm^2_{31}. It is shown in Figure 8.6 for NO (top) and IO (bottom). Again, for each point of the log-likelihood landscape, all other $N-2$ systematic parameters are optimized. Note that for NO (IO), the vertical axis is chosen to be Δm^2_{31} (Δm^2_{32}). This is a common choice for showing the ordering-dependent, atmospheric squared-mass difference, which is adapted here for consistency reasons [25]. It is motivated by the fact that for NO $|\Delta m^2_{31}| > |\Delta m^2_{32}|$, while for IO $|\Delta m^2_{31}| < |\Delta m^2_{32}|$, such that the two differences switch their relative size depending on the NMO. Thus, with the above choice the values of the fitted parameters for both hypotheses become more comparable.

The shown contours (magenta lines) are obtained by converting the ΔLLH values into confidence levels using Wilks' Theorem for two degrees of freedom (cf. Section 7.2). Naturally, the best-fit in the left octant for NO (star marker) agrees with the result of the one-dimensional scan. The 25%-contour is added to indicate the preference for NO by showing a confidence level, that is completely contained within the NO hypothesis.

An analogous, two-dimensional $\Delta\chi^2$-scan for the atmospheric oscillation parameters is provided by NuFit [25, 26]. Converting these into log-likelihood contours, using $2\Delta\text{LLH} = \Delta\chi^2$, *confidence-level* (CL) contours are derived for the global best-fits. These contours are shown in Figure 8.6 for the fits, published in January 2018 (*NuFit 3.2*), as green lines for different confidence intervals. Like this work, the NuFit result prefers NO over IO with a rather small region of parameters being allowed in the IO plane. The contours derived in this work agree reasonably well with the NuFit results. However, a slightly higher value of $\Delta m^2_{3\ell}$ in the left octant is preferred in contrast to the presented NuFit result.

Note that although the presented contour is competitive with global measurements of $\Delta m^2_{3\ell}$ and $\sin^2(\theta_{23})$, this analysis does not aim for a measurement of the oscillation parameters.

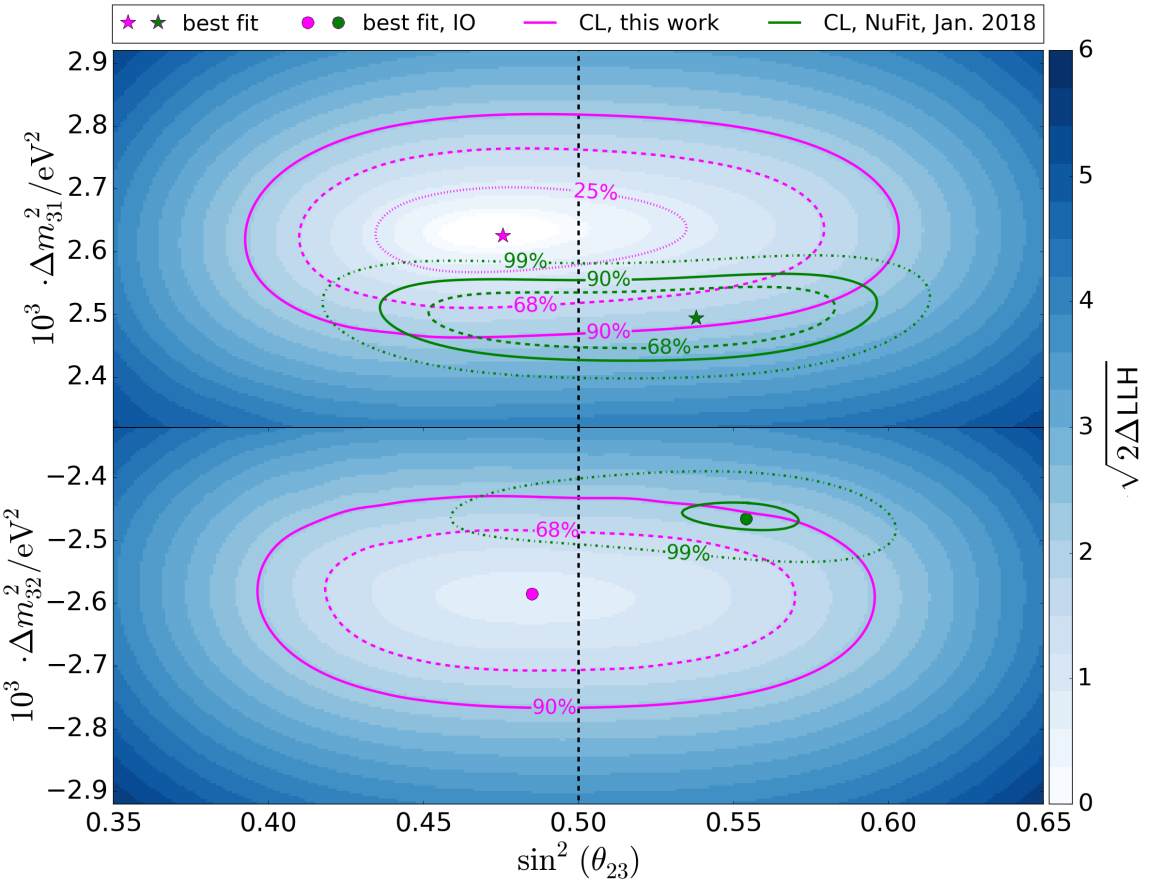


Figure 8.6: Two-dimensional log-likelihood scan in $\sin^2(\theta_{23})$ and $\Delta m^2_{3\ell}$ with $\ell = 1$ for NO (top) and $\ell = 2$ for IO (bottom): the various lines show 99%-, 90%-, 68%- and 25%-contours obtained by Wilks' Theorem (cf. Table 7.3) for this work (magenta) and the global best-fits by NuFit (green) [26, 25]; the choice of $\ell = 1$ (2) for NO (IO) is conventionally used and adapted here.

Instead, the shown contour provides a visualization of the $\sin^2(\theta_{23})$- and $\Delta m^2_{3\ell}$-dependent preference for the NMO. However, the oscillation parameters are determined in a similar analysis of the IceCube Collaboration, driven by Elim Thompson [113]. Still, the analysis features several differences to the NMO analysis presented in this work and is not discussed here in detail.

Finally, the effect from neutrino oscillations is illustrated in so-called *L-over-E-plots*. To do this, the ratio of the reconstructed baseline L_ν^{reco}, obtained from the reconstructed zenith-angle (cf. Equation 2.35), and the reconstructed neutrino energy E_ν^{reco} is calculated for every event. The resulting distribution of $L_\nu^{\mathrm{reco}}/E_\nu^{\mathrm{reco}}$ approximates the oscillograms in Figure 2.6 in reconstructed observables. It is shown in Figure 8.7 (left) for the events entering the likelihood analysis, where the black dots represent the observed data and the colored, stacked histograms shows the different Monte Carlo contributions.

For comparison, the dashed, black line shows the predicted total distribution from Monte Carlo, if neutrino oscillations are not present. Note that it only illustrates the total effect of disappearing events, in case neutrino oscillations are present. It does **not** allow to estimate the significance of the effect, since the dashed, black line is not optimized with respect to the nuisance parameters for the case of no oscillations. These would compensate for

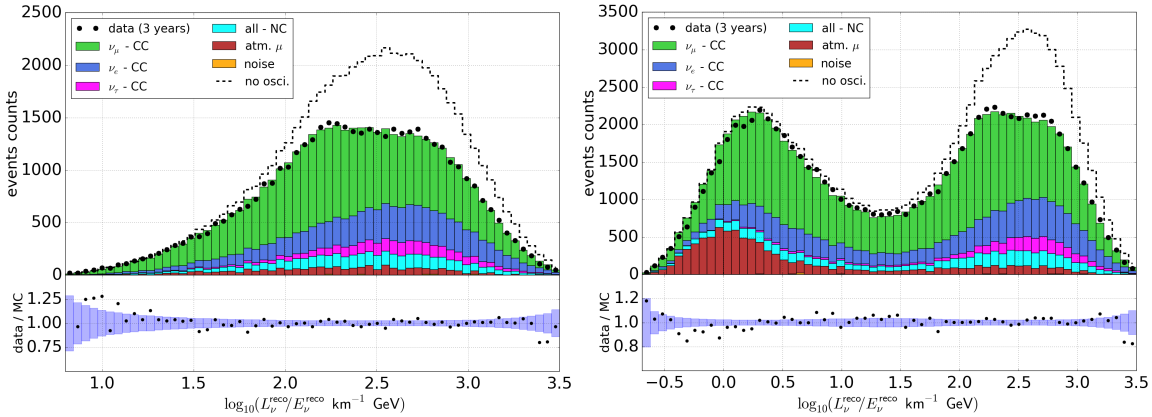


Figure 8.7: Visualization of the oscillation effect, showing the ratio of the propagated distance L_ν^{reco} as obtained from the reconstructed zenith-angle and the reconstructed neutrino energy E_ν^{reco} (cf. Figure 2.6); left: for the events entering the likelihood analysis (upgoing events); right: including up- and downgoing events.

most of the observed difference, but fit a systematically wrong hypothesis. Instead, the shown distributions illustrate the strength of the observed oscillation effects, given that the oscillation picture is correct.

Furthermore, the distribution is shown for combining both hemispheres in Figure 8.7 (right), i.e. for removing the cut on the reconstructed zenith-angle on Level 7 (cf. Section 4.4.3). Thus, the distribution includes up- and downgoing events, extending the range of $L_\nu^{\text{reco}}/E_\nu^{\text{reco}}$ to much smaller values. This is done to visualize the difference between the up- and downgoing events, where downgoing events (mostly at $\log_{10}(L_\nu^{\text{reco}}/E_\nu^{\text{reco}}\text{km}^{-1}\text{GeV}) < 1$) feature almost no neutrino oscillations and thus, show no major difference between the observed data and the dashed, black line.

Note that the pulls at $\log_{10}(L_\nu^{\text{reco}}/E_\nu^{\text{reco}}\text{km}^{-1}\text{GeV}) < 0.5$ in Figure 8.7 (right) feature expected deviations from the MC prediction, due to the extrapolation of systematics to the upper hemisphere: Since all systematic parameters are optimized on the upgoing sample, they are not expected to describe the downgoing region perfectly. Especially parameters which are mostly relevant for the downgoing region, like the atmospheric muon normalization, can lead to a notable disagreement there, while they have only little impact on the upgoing regime.

Besides the three-flavor fit of matter effects, in Normal and Inverted Ordering, and vacuum oscillations, the data is also fit in the approximation of two-flavor oscillations in vacuum (cf. Equation 2.26). In this case, $\sin(2\theta_{2\nu})$ is fitted, instead of the mixing angle $\theta_{2\nu}$, as $\sin^2(2\theta_{2\nu})$ directly scales the effect of muon neutrino disappearance in the two-flavor scenario. As a result, the fit is free to choose values $\sin(2\theta_{2\nu}) > 1$, which is an unphysical result that cannot be interpreted by a mixing angle $theta_{2\nu}$ in standard oscillations. However, this is done to identify issues in the background description that can lead to the observation of more muon neutrinos disappearing than expected for maximum mixing. The test is attached in Appendix I.2, where the observed p-value of $p = 19.1\%$ is well-compatible with statistical fluctuations and thus, no indication for a mismodelling of backgrounds is found.

Moreover, a short overview of additional tests, that were applied, is attached to Appendix I.3.

8.2 Combination with *NuFit* Results

The two-dimensional likelihood scan from Figure 8.6 can be combined with the $\Delta\chi^2$-scan provided by NuFit [26]. To do this, the NuFit 3.2 scan is converted into a log-likelihood landscape using

$$\Delta\chi^2 = \chi^2 - \chi^2_{\text{min}} = 2\Delta\text{LLH} = 2(\text{LLH} - \text{LLH}_{\text{min}}).$$

Then, the two independent scans are combined as

$$\text{LLH}_{\text{total}}(\theta_{23}, \Delta m^2_{3\ell}) = \text{LLH}_{\text{DeepCore}}(\theta_{23}, \Delta m^2_{3\ell}) + \text{LLH}_{\text{NuFit}}(\theta_{23}, \Delta m^2_{3\ell}), \tag{8.6}$$

where $\text{LLH}_{\text{DeepCore}}$ is the likelihood scan from Figure 8.6 and $\text{LLH}_{\text{NuFit}}$ is the converted $\Delta\chi^2$-landscape from NuFit.

Afterwards, the combined likelihood LLH_{total} is normalized by shifting the minimum towards zero. Then, the new log-likelihood landscape is used to derive confidence-level contours using Wilks' Theorem (cf. Table 7.3), which is analogous to the previous Section. The scan of the combined likelihood LLH_{total} and the resulting confidence-level contours are shown in Figure 8.8.

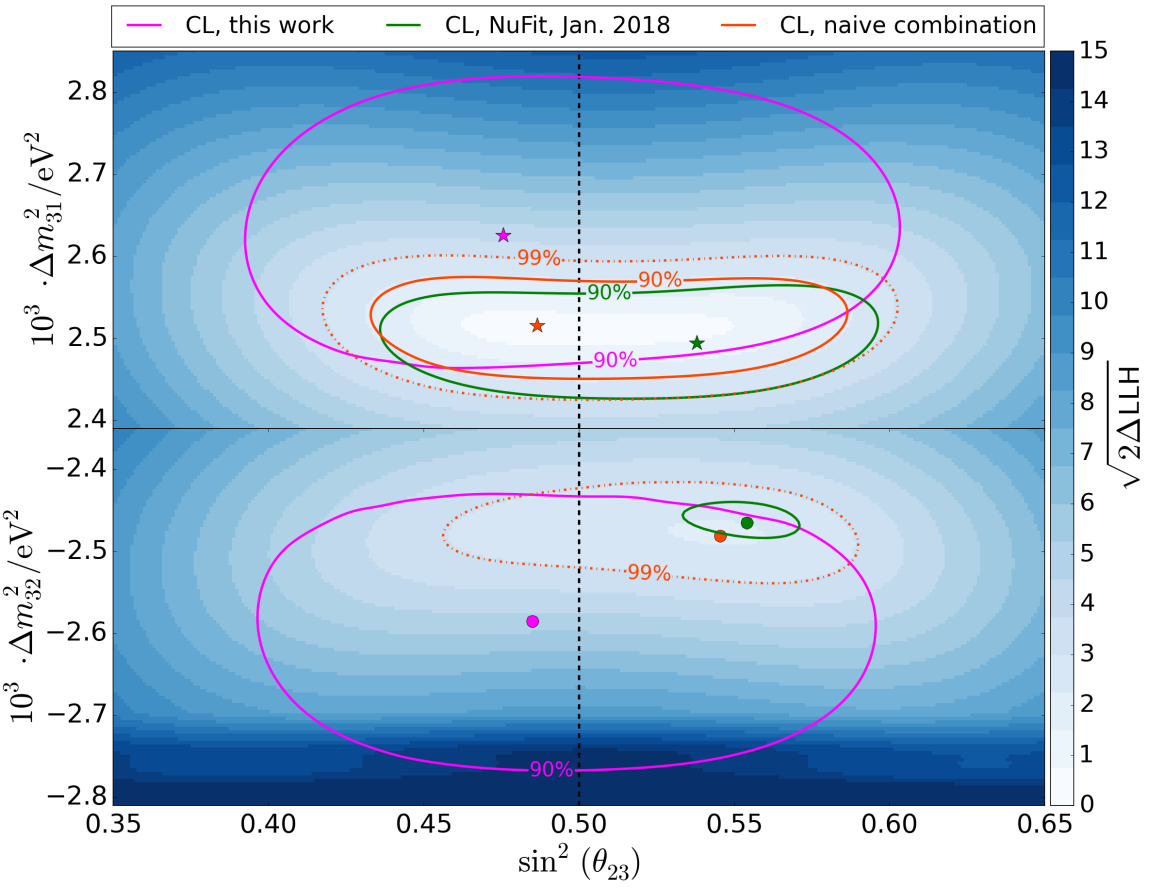


Figure 8.8: Combination of NuFit contours (green) from [25, 26] and the results of this work (magenta): the combination is shown as 90%- and 99%-CL contours (red), while the combined likelihood, $\text{LLH}_{\text{total}}(\theta_{23}, \Delta m^2_{3\ell})$ (cf. Equation 8.6), is shown in the background.

As expected, the red (combined) contours are very close to the contours provided by NuFit, as NuFit is more constraining on the oscillation parameters than the DeepCore result.

Moreover, the combined result prefers higher values of $|\Delta m^2_{3\ell}|$ and smaller values of $\sin^2(\theta_{23})$ than NuFit, which is expected, due to the preference of the DeepCore result. For the NO case, the combination flips the contour from a slight preference for the right (second) octant to the left (first) octant.

Finally, both results prefer Normal over Inverted Ordering. As a result, also the combined contour features a clear preference for Normal Ordering, such that the 90%-contour of the combined result (red, solid line) is completely contained within the NO hypothesis. In particular, the small *island* of parameters in the right octant of the IO hypothesis, that is not excluded by the 90%-contour from NuFit, is excluded at 90% confidence level in the combined result. Moreover, the minimum value of the normalized LLH for the IO hypothesis changes from $\min_{\Delta m^2_{31},\theta_{23}} \mathrm{LLH}_{\mathrm{NuFit}}(\mathcal{H}_{\mathrm{IO}}) = 2.07$ for NuFit to $\min_{\Delta m^2_{31},\theta_{23}} \mathrm{LLH}_{\mathrm{total}}(\mathcal{H}_{\mathrm{IO}}) = 2.38$ for the combined result.

Thus, the combination of DeepCore with the most recent NuFit results strengthens the globally seen preference for NO and supports the exclusion of the IO hypothesis.

Note that this *naive* combination of the results features some subtleties: According to NuFit, the $\Delta\chi^2$-scan includes the DeepCore result from [154], which was published in 2015. Thus, the above combination is not strictly correct, since the NuFit contour depends partially on DeepCore results, which are correlated with the results of this work. In terms of systematics, the missing alignment of nuisance parameters can be seen as conservative, as it flattens the likelihood. However, the DeepCore result from [154] uses a data sample that overlaps with 9.99% of the GRECO sample, leading to a statistical dependence between the original (green) NuFit contour and the (magenta) contour of this work.

Nevertheless, the resulting correlation between the NuFit contour and this work is estimated to be very small for several reasons: First, the DeepCore result from [154] does not contribute notably to the global understanding of oscillation parameters. Its 90%-contour in $\sin^2(\theta_{23})$ and $\Delta m^2_{3\ell}$ is substantially larger than the contours provided by this work or NuFit and does not contribute to other oscillation parameters of the global fit. Thus, the NuFit χ^2-scan in these parameters is only weakly affected by the previous DeepCore result. Second, even the small pull of the previous DeepCore result on the NuFit contour is mostly independent of this work, since the overlap between the DeepCore samples is only $\sim 10\%$ of the GRECO sample. Third, a different reconstruction algorithm was used for the previous result, which does not use the same information to reconstruct energy, zenith-angle and PID as the reconstruction presented in Chapter 5. Thus, even for the small number of events in the overlap, the reconstructed observables are not fully correlated.

As a result, the naively combined contours from Figure 8.8 can be seen as a valid approximation of a more accurate global fit.

9 Summary and Outlook

The existence of non-vanishing neutrino masses is a well-established fact in modern particle physics. Although the absolute neutrino masses m_1, m_2 and m_3 are unknown, the squared-mass differences between these states can be explored using neutrino oscillations. Moreover, neutrino oscillations allow to test for the Neutrino Mass Ordering (NMO) (cf. Chapter 1).

One way to test the NMO with oscillations is to measure the modulations from matter effects in the oscillation pattern of atmospheric neutrinos. These atmospheric neutrinos are generated in Cosmic Ray interactions with the Earth's atmosphere. During their propagation through Earth, their oscillation probabilities are affected by the Earth's matter profile, leading to a NMO-dependent signature from matter effects below $E_\nu \sim 15\,\text{GeV}$ (cf. Chapter 2). One detector capable of probing this effect is the IceCube Neutrino Observatory with its low-energy sub-detector DeepCore (cf. Chapter 3).

In this work, a low-energy analysis is presented to probe the NMO with three years of IceCube DeepCore data. To do this, a low-energy data sample, originally developed by Michael Larson [112], was adapted and refined in collaboration with Michael Larson and Elim Thompson [113]. It is characterized by high statistics and reaches down to neutrino energies of $E_\nu \sim 5\,\text{GeV}$. It comprises non-negligible contributions from CC muon, electron and tau neutrinos, NC events of all flavors and atmospheric muons. The contribution from triggered noise was modelled and found to be negligible. Moreover during this work, a previously seen disagreement between data and MC was found to be caused by a mismodelling of the observed charges in Monte Carlo and two light emitting (*Flaring*) DOMs (cf. Chapter 4).

For this data sample, a new event reconstruction was developed, which achieves an excellent performance in the reconstruction of the neutrino zenith-angle and the neutrino energy. Moreover, it allows to separate different event signatures, giving a handle on the neutrino flavors. As a result, the event reconstruction was adapted by several other DeepCore and PINGU analyses (cf. Chapter 5).

Finally, a binned, *maximum-likelihood* method was developed to probe the NMO, using the reconstructed neutrino energy, zenith-angle and a flavor-separating variable (PID). For this, uncertainties on the Monte Carlo templates from limited MC statistics were reduced using *Kernel Density Estimation*. The remaining uncertainties were included into the likelihood-function by convolving the Poissonian likelihood with the bin-wise Monte Carlo uncertainty (cf. Section 6.1, 6.2 and 6.4).

In the likelihood fit, several nuisance parameters were included to account for systematic uncertainties on the atmospheric neutrino and muon fluxes, the neutrino oscillation parameters, the neutrino-nucleon interactions and the detector response. The impact of each systematic uncertainty on the analysis distribution was parametrized and the resulting parameter was added to the likelihood optimization. The nuisance parameters were tested for

their relevance in the NMO measurement and all non-negligible parameters were added to the NMO fit (cf. Section 6.3 and 7.6).

The sensitivity to the NMO was estimated using two methods: first, a slow but accurate method using Pseudo-Experiments and second, a substantially faster Asimov method. Additionally, the impact of future improvements on the understanding of systematics and the event reconstruction were investigated. The resolution of the event reconstruction was found to be crucial for an increased NMO sensitivity. In contrast, an improved understanding of systematics gives only a small gain in sensitivity, as long as the total NMO signature is too weak to be affected by the systematic uncertainties (cf. Chapter 7).

All of the analysis was developed blindly with the unblinding-procedure supervised and reviewed by the IceCube Collaboration.

Finally, the analysis was applied to experimental data: The fit was found to prefer Normal (NO) over Inverted Ordering (IO) with a p-value of $p_{\mathrm{NO}} = 71.1\%$ and a CL_s-value of $\mathrm{CL}_s\left(\mathcal{H}_{\mathsf{NO}}\right) = 83.0\%$. In contrast, the IO hypothesis lead to a p-value of $p_{\mathrm{IO}} = 15.3\%$ and a CL_s-value of $\mathrm{CL}_s\left(\mathcal{H}_{\mathsf{IO}}\right) = 53.3\%$. Moreover, the hypothesis of vacuum oscillation was tested against the hypothesis of matter effects for both orderings. The resulting fit prefers matter effects in both cases with p-values of 12.3% (NO) and 22.2% (IO) for vacuum oscillations, compared to 62.3% (NO) and 53.2% (IO) for matter effects (cf. Section 8.1).

Finally, a likelihood-scan was conducted in $\sin^2(\theta_{23})$ and Δm^2_{31} to validate the preference for NO over IO, depending on the values of the atmospheric oscillation parameters (cf. Section 8.1.2). The best-fit was found to prefer $\sin^2(\theta_{23}) = 0.4757$ and $\Delta m_{31} = 2.626 \cdot 10^{-3}\,\mathrm{eV}$, which is compatible with the global-fit results provided by NuFit [26, 25]. The naive combination of both results excludes every combination of oscillation parameters in the IO plane at a 90% confidence level (cf. Section 8.2).

To validate the likelihood fit, the *goodness-of-fit* was estimated in several ways. To do this, the observed values of all nuisance parameters and the pulls on the analysis distribution were investigated. Overall, no indication was found for the NMO fit being affected by unmodelled systematic effects or an underestimation of backgrounds (cf. Section 8.1 and Appendix I).

For future measurements, the DeepCore detector is unlikely to provide a high-significance measurement of the Neutrino Mass Ordering. Due to its relatively high energy threshold, it is only capable of observing a small fraction of the total signature from matter effects with low reconstruction resolution. However, the DeepCore sensitivity might still increase over the next years: While this work is based on three years of data, eight years might be available soon. In combination with an improved data selection, this may increase the sensitivity by a factor of up to ~ 2. Moreover, an improved energy or zenith-angle reconstruction could lead to an additional improvement on the NMO sensitivity (cf. Section 7.7).

In parallel, the IceCube Neutrino Observatory will be extended to a next generation observatory, called *IceCube-Gen2*. Part of this instrument is the *Precision IceCube Next Generation Upgrade* (PINGU) [7]. This upgrade is supposed to improve the calibration of the detector response substantially, while lowering the energy threshold and improving the event reconstruction. A first step to such a detector is the extension of IceCube by eight additional strings with a dense, vertical spacing of DOMs, that will be deployed within the next years. It will lower the energy threshold of DeepCore substantially. Moreover, calibration measurements will allow to constrain the current uncertainties on the optical properties of the ice and to re-calibrate the already collected data (cf. Section 3.6).

Some of the key goals of this future extension are an improved measurement of the atmospheric oscillation parameters and a precise measurement of tau neutrino appearance, probing the unitarity of the PMNS mixing matrix U. Besides these goals, the extension will allow for an improved measurement of the Neutrino Mass Ordering, compared to DeepCore (cf. Section 3.6).

For these future upgrades, this work provides a *proof-of-concept* for a more significant NMO measurement. It tests the full analysis-chain, including the treatment of systematic uncertainties and the statistical evaluation of the experimental result, that could be adapted by such more significant, future measurement of an IceCube upgrade.

Appendix

A Constraints from Non-Oscillation Experiments

Besides the observation of neutrino oscillations, the NMO could be determined by constraints on the absolute neutrino masses. These constrains are briefly discussed in the following.

One of these constraints arises from cosmological observations: In the early universe, neutrinos decoupled from the other particles early due to their small cross-section, forming the *Cosmic Neutrino Background* (CNB), similar to the *Cosmic Microwave Background* (CMB) for photons. These CNB neutrinos contributed significantly to the total mass density of the universe during its evolution and thus influenced the subsequent structure formation. However, for a given number density n_ν, the minimum mass density ρ_ν depends on the NMO: In case of NO, the two lower neutrino states could have negligible mass compared to the third one. In case of IO, only one state might be negligible compared to the other two. Thus, a higher minimum mass density is expected for IO compared to NO [6, 12].

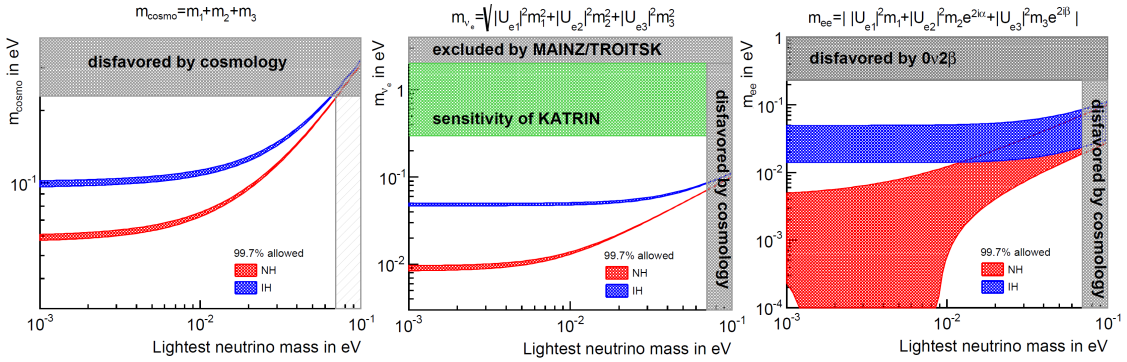


Figure A.1: Constraints from cosmology (left), β-decay experiments (center) and $\beta\beta$-decay experiments (right) on the lightest neutrino mass and an effective mass, being defined as a superposition of the three fundamental neutrino masses m_i (from [4]).

By observing the structure formation in the universe, a limit on the effective mass $m_{\text{cosmo}} = m_1 + m_2 + m_3$ for cosmology can be calculated. Since NO and IO correspond to some minimum, effective mass m_{cosmo}, IO could be excluded, if the observed m_{cosmo} is below the minimum value required for IO. A recent summary of the cosmological constraints on m_{cosmo} and the lightest neutrino mass is shown in Figure A.1 (left). The limit is insufficient to exclude IO, while recent, more progressive limits go down to even $m_{\text{cosmo}} \leqslant 0.17\,\text{eV}$ at 95% Confidence Level [6, 12].

However, one should note that these limits depend on cosmological assumptions, that might

be wrong. Depending on these assumptions and the data being used, the preference for NO over IO in recent cosmological findings is in the range of $\Delta\chi^2_{\mathrm{NO-IO}} \approx 0 - 1.6$ [6, 12].

In addition to cosmological constraints, β- and $\beta\beta$-decay experiments provide a handle on the NMO, using their sensitivity to the absolute masses (cf. Section 2.3). Similarly to the effective mass m_{cosmo}, they feature effective masses m_{ν_e} and m_{ee} that are linear combinations of the mass eigenstates m_1, m_2 and m_3. Some recent limits on these effective masses are shown in Figure A.1 for β- (center) and $\beta\beta$-experiments (right). As for cosmology, they could exclude IO, if their limit on the corresponding effective mass was small enough to exclude the IO phase-space (red). However, in general they are less compelling in determining the NMO than the most recent cosmological constraints [4, 12].

B Event View of Typical Low-Energy Events

In Figure B.2, a typical *track-like* and *cascade-like* event is shown from Monte Carlo, using the IceCube event visualization software *steamshovel*. For the cascade-like event, a CC electron neutrino interaction is used, while the track-like event represents a CC muon neutrino interaction (inelasticity $y \approx 0.3$).

The event visualization shows the DeepCore volume, where the individual strings are indicated by thin, grey, vertical lines. On these lines, the individual DOMs are indicated by spheres of different size and color: *Unhit* DOMs, that did not observe any pulse during the event, are represented by small, grey spheres, while *hit* DOMs are represented by colored spheres, where the size of the sphere corresponds to the amount of observed total charge and the color encodes the time of the first observed photon (time ordering: red, yellow, green, blue).

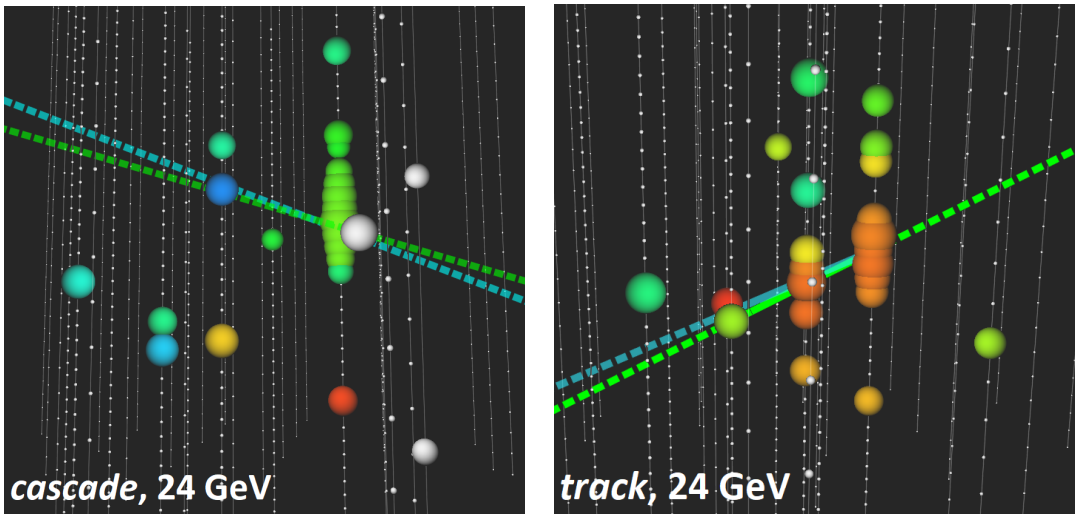


Figure B.2: Example of a *cascade-like* (left) and *track-like* (right) event at $E_\nu \approx 24\,\mathrm{GeV}$, where the cyan (green) lines respresent the reconstructed (true) direction of the neutrino event; for the track-like event, the solid fraction of this line represents the reconstructed (true) length of the outgoing muon track (modified from [142]).

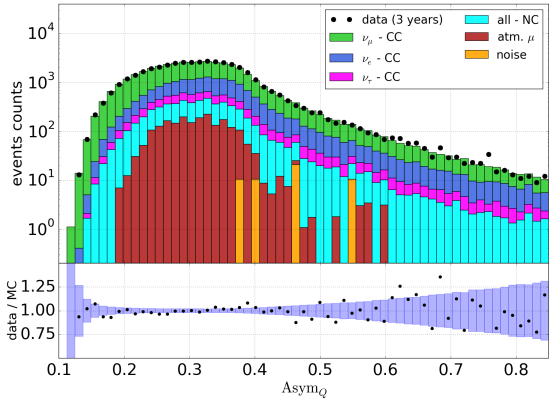
The dashed, cyan (green) lines represent the reconstructed (true) direction of the neutrino event. Moreover, for the track-like event, the solid fraction of the line represents the recon-

structed (true) length of the outgoing muon track.

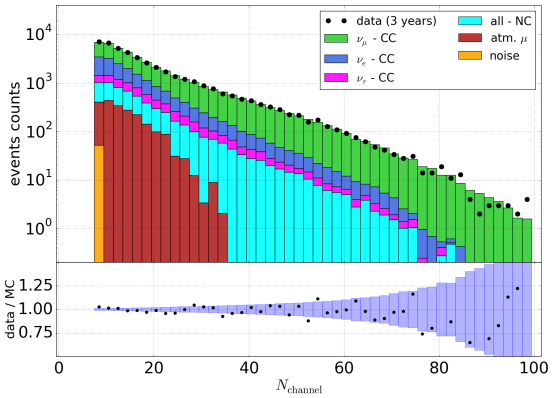
As visible in Figure B.2, the event signatures of tracks and cascades are very similar at low energies. However, minor differences can be spotted along the muon track. These differences are used to separate track- and cascade-like signatures in the likelihood reconstruction method from Section 5.3.1.

C Supplemental Material on GRECO Selection

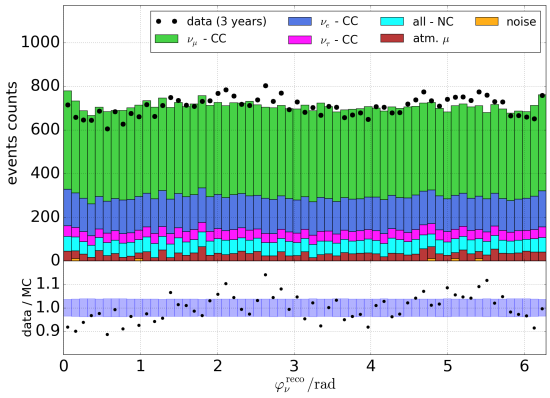
The control distributions in Figure C.3 are a selection of high-level variables, shown for the final level GRECO sample. They compare the observed data to the best-fit Monte Carlo (MC) predictions. All shown distributions are event-wise quantities, i.e. each entry in the histogram corresponds to an observed or predicted event.



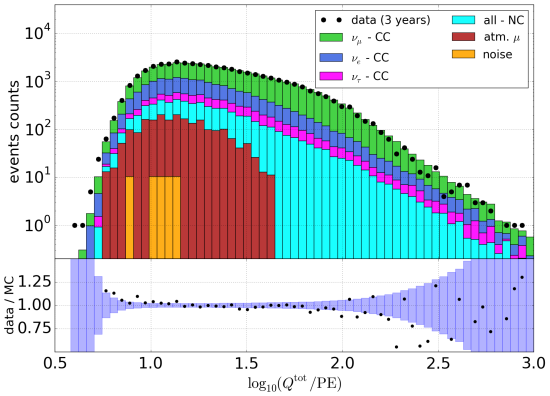
(a) charge asymmetry Asym$_Q$



(b) number of hit DOMs N_{Channel}



(c) reconstructed azimuth-angle ϕ_ν^{reco}



(d) integrated charge Q^{tot}/PE

Figure C.3: Overview of some *control distributions* at GRECO Level 7 that are used to identify unexpected data or MC behavior in the final level data sample.

The shown distributions (from top left to bottom right) are the charge asymmetry discussed in Section 4.4.2, the number of hit DOMs, also called number of channels N_{Channel}, the Pegleg reconstructed azimuth-angle of the events ϕ_ν^{reco} and the observed total charge per event Q^{tot}.

Note that not all of these distributions are expected to show perfect agreement between data and MC. For instance, the charge distribution is known to be slightly tilted between data and MC, due to the SPE template issue, discussed in Section 4.5.2. Moreover, the azimuth-angle is expected to show a slight preference direction in data, that is not present in MC, due to a weak anisotropy in the ice (cf. Section 3.4).

However, these distributions are conventionally tested to identify unexpected disagreement besides known deviations.

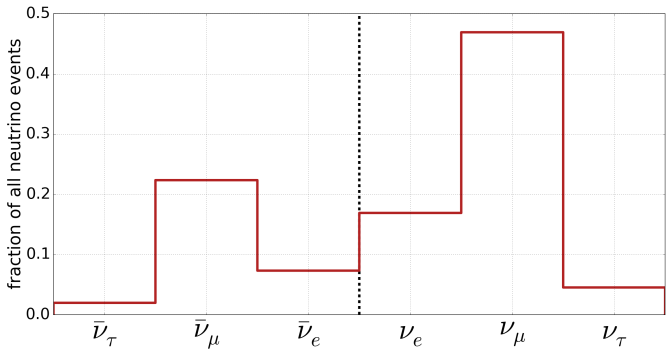


Figure C.4: Composition of the final level GRECO sample of neutrinos and anti-neutrinos of different flavors; the vertical axis gives the fraction of all observed neutrino events in the sample.

Finally, Figure C.4 shows the composition of the final level GRECO sample of neutrinos and anti-neutrinos for different flavors. As one can see, the neutrino contribution dominates, while the neutrino to anti-neutrino ratio is roughly $2:1$, as expected from Section 2.4.1.

D Photon Arrival Time Expectations

If the scattering and absorption properties of the ice were perfectly homogeneous and the photon acceptance for all DOMs was rotationally symmetric, the arrival time distribution of photons at a DOM would follow a Gamma distribution. This Gamma distribution is often refered to as *Pandel Function*. It is given by Equation D.1 [155].

$$p(\rho, \xi, t) = \frac{\rho^\xi t^{\xi-1}}{\Gamma(\xi)} e^{-\rho t} \tag{D.1}$$

where $t = t_{\text{hit}} - t_{\text{geom}}$ measures the time, relative to the geometric time t_{geom}, that an unscattered Cherenkov photon needs to reach the DOM. Thus before $t = 0$, no light can be observed from the interaction vertex. The parameter $\xi = d/(\lambda \sin(\theta_c))$ scales the distance d between generation and detection of the photons by the mean scattering length λ and the sine of the Cherenkov angle θ_c [155].

The optical properties of the surrounding ice are very approximately described by the parameters λ and ρ. For the IceCube-predecessor *AMANDA*, they were found to be $\lambda \approx 33.3\,\text{m}$ and $\rho \approx 0.004\,\text{ns}^{-1}$ [155].

Although the Pandel Functions allow for an estimate of the photon arrival times, the assumption of homogeneous and isotropic ice and rotationally symmetric DOMs is unphysical. Instead, current IceCube reconstructions rely on photon arrival time distributions that are obtained from the time-intense simulation of photon propagation.

For this, electromagnetic cascades and short muon tracks are simulated and the resulting Cherenkov photons are propagated through the detector. The cascades and muons vertices are generated on a three-dimensional grid, that spans through all of the detector volume and beyond with a 15 m spacing. Moreover, the orientation of the cascades and muons are varied, leading to two additional degrees of freedom. The subsequent photon-propagation includes local ice properties and an accurate simulation of the DOMs photon acceptance and hardware response. For each of these five-dimensional grid points and each DOM, the total number of observed photons and the observed arrival time distribution are then stored in large tables.

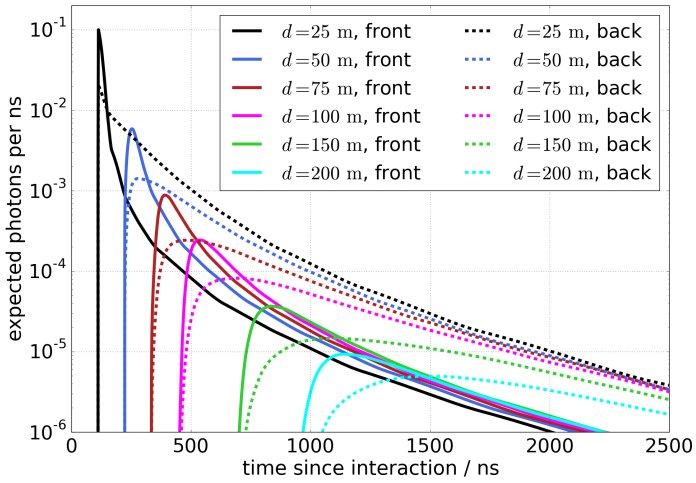


Figure D.5: Distribution of photon arrival times, relative to the geometric time t_{geom}, as obtained from the splined photon-tables, for the two cases of pointing towards (*front*) and away from (*back*) the DOM position.

For the *Millipede likelihood*, these tables are loaded and splined to obtain a continuous value in the total number of expected photons and the arrival time distribution. An example for such a continuous distribution is shown in Figure D.5 for an electromagnetic cascade of 1 GeV at different distances d to the observing DOM and for the two cases of pointing towards (*front*) and away from (*back*) the DOM position.

Integrating the arrival time distribution over a given binning, one obtains the expected number of photons $\Lambda^{\mu/c}_{i\tau k}(\vec{x}_i, \vec{y}_k, \theta_k, \phi_k, \Delta t_\tau)$ for each time bin Δt_τ, that is used in Equation 5.2.

E Pegleg Supplemental Material

The following material provides additional information on the Pegleg reconstruction introduced in Chapter 5.

E.1 Relation to HybridReco

The Pegleg reconstruction was one of two independent efforts to reconstruct a track+cascade hypothesis on the Millipede likelihood. The second, independently developed reconstruction, called *HybridReco*, was earlier used for the reconstruction of low-energy events, but was ultimately not used in any journal publication of DeepCore oscillation results.

Instead, the Pegleg reconstruction was gradually adapted by all DeepCore and PINGU analyses, previously using HybridReco, for the following reasons:

First, the Pegleg reconstruction provides the charge-independent Millipede likelihood from Section 5.3.3, while such likelihood does not exist in HybridReco. Thus, the disagreement between data and MC, described in Section 4.5.2, was never resolved using HybridReco.

Second, HybridReco uses the Multinest minimizer (and first introduced it to IceCube) to directly minimize all eight parameters of the track+cascade hypothesis in the outer minimizer, without any internal optimizations. This eight-dimensional optimization takes $\sim 2-3$ times longer than the three-layer optimization used in Pegleg (cf. Section 5.3.2). Since the reconstruction is typically the computational bottle-neck in such analyses, with reconstruction times of several month up to a year for all of the Monte Carlo samples, this gain was crucial for the adaption of Pegleg.

Third, the zenith-angle resolutions for Pegleg and HybridReco are almost identical, while the energy resolution of Pegleg features less outliers due to failed minimizations. This is due to the fact that the cascade energy is guaranteed to be at the optimum, for the given set of remaining parameters, the way it is derived within Pegleg, while this is not the case for the eight-dimensional optimization. However, the difference is very small, such that all resolutions of Pegleg and HybridReco are nearly identical.

Fourth, the Pegleg reconstruction was implemented within the IceCube software framework for event reconstructions. Thus, its implementation was substantially shorter and more consistent with other collaboration software than HybridReco, which was developed as stand-alone tool and called the Millipede likelihood from outside the IceCube framework.

E.2 Reconstruction Settings

The settings used in the Pegleg reconstructions are shown in Table E.1 and E.2, where the first focusses on the settings of the Multinest minimizer [121], while the latter gives the settings related to the likelihood space and the parametrization, discussed in Chapter 5.

E.3 The Charge-Independent Millipede Likelihood

The charge-independent Millipede likelihood, described in Section 5.3.3, is motivated by the finding from Section 4.5.2 that the observed charge is no reliable proxy for the observed number of photons in each DOM.

The charge assigned to a detected photon in Monte Carlo is obtained as a random variable following the *Single-Photo-Electron* (SPE) distribution. The SPE distribution is obtained from lab measurements. These templates were found to be insufficiently accurate to obtain reliable Monte Carlo. Thus, they lead to a disagreement in the final level data sample. While large efforts were started by Martin Rongen [115] and Spencer Axani [114] to replace the existing SPE templates by *in-situ* measurements, this work was aiming for a solution on a much shorter time scale.

To do this, the charge-dependent Millipede likelihood, used for the event reconstruction, was replaced by a charge-independent likelihood.

The charge-independent likelihood is obtained by replacing the charge of the first observed pulse in each DOM by 1 PE. After this pulse, a time-window of $\delta t_{\text{dead}} = 45\,\text{ns}$ is removed from the likelihood, such that any pulse, observed in this time-window, is ignored. This means the DOM is treated as *offline* or *dead*, which accounts for an unexpected *post-hit*

Table E.1: Settings of the Pegleg reconstruction that are related to the likelihood space or the parametrization; the parameters are labelled (italic font) as in the IceCube software to faciliate the reproduction of the final level reconstruction used for the GRECO sample.

Parameter	Setting / Value
PulseMap	SRT-, TimeWindow-cleaned OfflinePulses (DeepCore)
ExcludeDOMs	Bad, Saturated, Bright, Omitted, Flaring
TimeWindow	OfflinePulsesTimeRange
LikelihoodService	Pegleg
DomEff	0.99
PhotonsPerBin	1
PENormalization	0.85
PartialExclusion	False
UseUnhitDOMs	True
maxSegSplits	0
minTrackSegments	0
minTrackSegmentsUp	2
minTrackSegmentsDown	200
SegmentSpacing	5 m
PrereadTables	False
fitStochasticLosses	False
fitContinuousLength	False
usePCG	False

Table E.2: Settings of the Multinest minimizer, used for the Pegleg reconstruction; the parameters are labelled (italic font) as in the Multinest optimization software [121].

parameter	setting / value	parameter	setting / value
MaxModes	10	*NIS*	False
Tolerance	0.5	*Efficiency*	1.0
NLive	60	*Periodic*	Azimuth
MaxIterations	80 000	*ModeSeparated*	t, x, y, z

behavior of the DOM hardware, after the first photon detection.

This is motivated by Figure E.6, showing the time difference between the first observed pulse and all following pulses in the same DOM for the final level GRECO sample (Flaring DOMs removed). Note that events enter this histogram several times, if they feature multiple DOMs with $\geqslant 2$ pulses. In contrast, events with no more than one pulse for all DOMs do not enter the histogram at all.

As one can see, the distributions for data and MC agree roughly for $\Delta t > 45\,\text{ns}$. The disagreement for $\Delta t < 45\,\text{ns}$ indicates the incorrect description of the DOM behavior after the first pulse being observed. Therefore, the dead time-window was chosen to extend over $\delta t_{\text{Dead}} = 45\,\text{ns}$ to remove the region, where the predicted DOM behavior is unreliable.

Already before this work, it was known that the number of pulses per DOM was no reliable quantity, since the *pulse-splitting* was assumed to work differently on data and MC. This means, a single photon might be split into two pulses in data, but not in MC. In the same way, several photons might be merged in one pulse in data, but not in MC. However, the integrated charge per DOM in a given time-window was assumed to be a reliable estimator

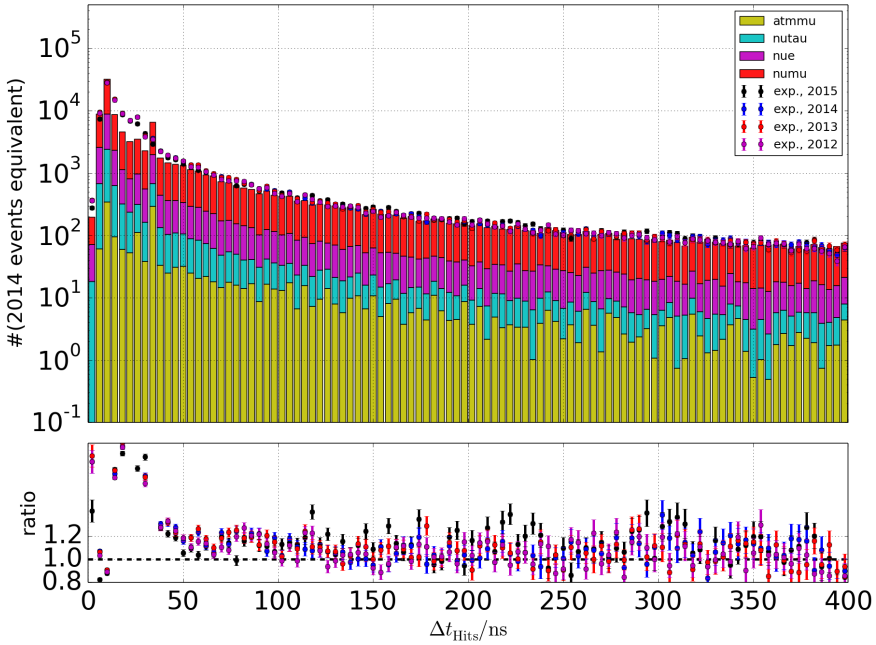


Figure E.6: Time difference between the first and all following pulses within a DOM in data and Monte Carlo; the distribution is well-described by MC for $\Delta t > 45\,\text{ns}$, while for $\Delta t < 45\,\text{ns}$ the DOMs observe far more pulses in data than simulated in MC; thus, this region is removed from the likelihood using a *dead* time-window of $\delta t_{\text{Dead}} = 45\,\text{ns}$.

for the observed number of photons. As discussed in Section 4.5.2, this is not the case for low energies, where the total number of observed photons is small.

Thus, neither the integrated charge per time-window nor the individual pulses are well-described in Monte Carlo.

However, the first pulse observed in a DOM is unaffected by any previous pulses and can be seen as a reliable proxy for the time of the first observed photon. Moreover, after δt_{Dead} the DOMs seem to return to their normal behavior. Therefore, the DOMs are turned on after the dead time-window to potentially observe another pulse, triggering a further dead time-window and so on. Note that this way, the observed charge is always an integer number in terms of photo-electrons (PE).

Moreover, one should note that the dead time-window δt_{Dead} is long, compared to the typical arrival time distributions for low-energy events. Thus, more than one pulse per DOM is rarely observed and the majority of events contains DOMs with no photons (0 PE) or only one photon (1 PE) being detected.

E.4 Reconstruction Performance

Figure E.7 shows the zenith-dependent zenith-resolution of the Pegleg reconstruction, developed in Chapter 5. It illustrates the zenith-dependent bias due to the boundary of the parameter space. While horizontal events with $\cos(\theta_{\text{true}}) = 0$ are reconstructed almost unbiased, the bias increases for more vertically up- or downgoing events. This is due to the parametric boundaries of $\cos(\theta_{\text{true}})$ at -1 and $+1$.

For example, a neutrino entering the detector from below might be close to $\cos(\theta_{\text{true}}) = -1$. Thus, it can hardly be reconstructed to more vertical values, while the phase-space for reconstructions towards more horizontal directions is much larger. This causes the shift of all quantiles in Figure E.7 towards negative (positive) values for $\cos(\theta_{\text{true}}) < 0$ ($\cos(\theta_{\text{true}}) > 0$). This shift in the zenith-angle reconstruction could be interpreted as reconstruction bias. However, it is not due to the algorithm itself, but the choice of events used to determine the resolution.

Therefore, in Section 5.3.4 the zenith and energy cuts from Level 7 are dropped to obtain a mostly unbiased result for the energy- and zenith resolutions.

For CC electron neutrino events, the energy-dependent zenith- and energy-resolution for the Pegleg and *SANTA/LEERA* reconstructions (cf. Section 5.1) are compared in Figure E.8. The comparison is analogous to the one for muon neutrinos in Figure 5.7.

As one can see, the Pegleg reconstruction provides an almost unbiased estimate of the true zenith-direction and the true neutrino energy. A small bias is expected and seen at low energies due to selection effects, i.e. events surviving the selection up to Level 7 tend to be exceptionally bright and are therefore reconstructed to higher energies, while events that are less bright do not survive the data selection.

In contrast, the zenith reconstruction by SANTA and the energy reconstruction by LEERA are biased compared to the true quantities. Moreover, LEERA provides an almost similar resolution to the Pegleg reconstruction, while SANTA leads to a clearly worse reconstruction for the zenith-angle, except at the lowest energies.

Finally, the fraction of reconstructable events with SANTA is shown in Figure E.8 (top) as black, solid line to be read from the right vertical axis. Similar as for muon neutrinos in Figure 5.7 (bottom), the fraction of reconstructable events drops quickly, such that only a small fraction of events is reconstructable at the lowest energies. As for muon neutrinos, the reconstruction is therefore badly suited for an analysis at $E_\nu < 15\,\text{GeV}$.

F Supplemental Material on Kernel Density Estimation

F.1 Implementation of Kernel Density Estimation

For this work, the KDE method from Section 6.2.1 was implemented in *C++* and *CUDA* [156]. It was used instead of the *Python* implementation by Sebastian Schoenen [133] for the following reasons:

1. The C++ implementation is much faster than the Python one ($< 50\%$ computation time).
2. The required memory for the Python implementation scales as $\propto N^2$, which lead to an unfeasible large memory consumption for $N \gtrsim 20\,000$.

Therefore, the C++ implementation was used for all purposes of this work.

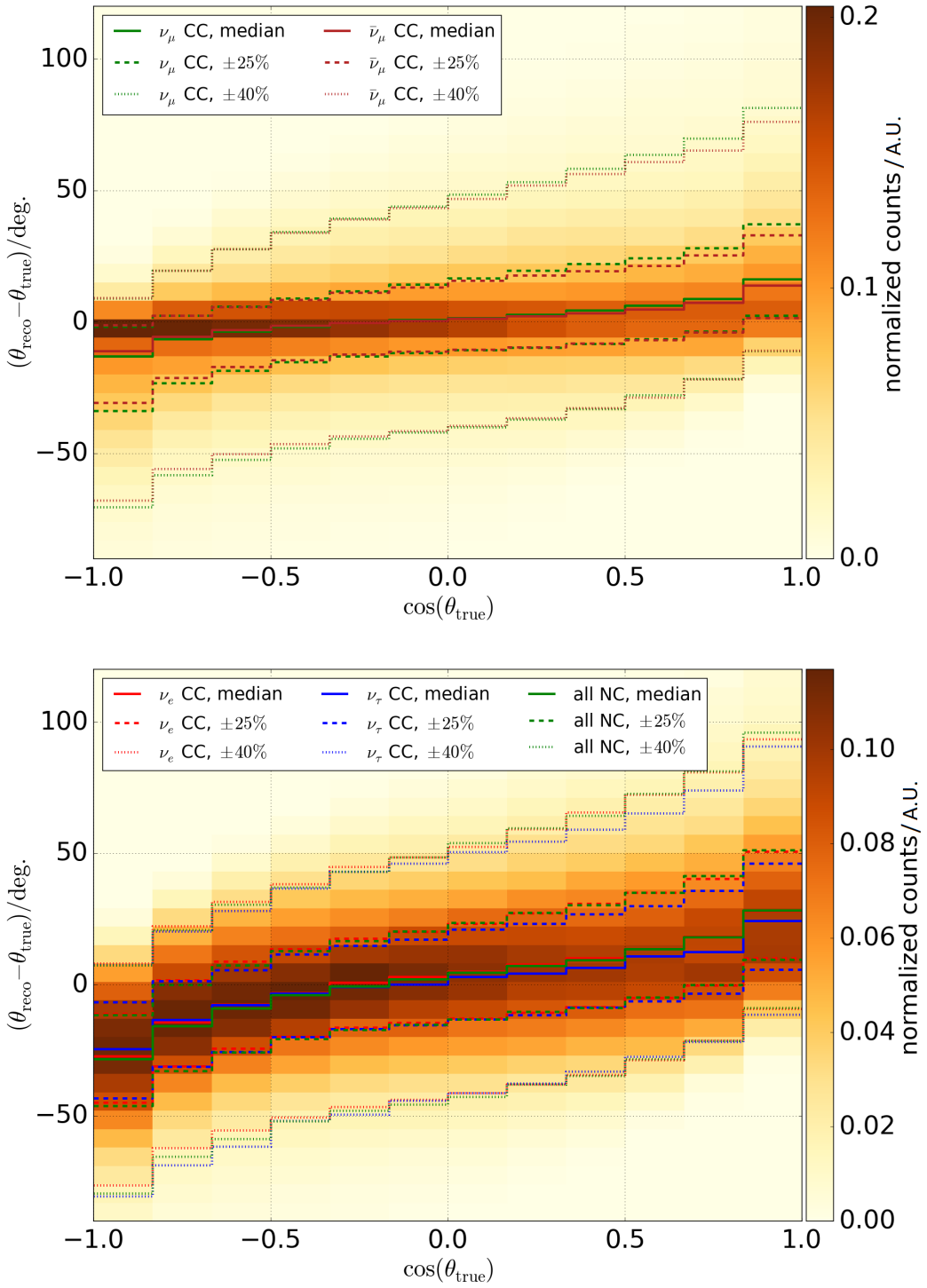


Figure E.7: Zenith-dependent zenith resolution of the Pegleg reconstruction for CC muon neutrinos and anti-neutrinos separately (top) and all other flavors, combining neutrinos and anti-neutrinos (bottom); the dashed and dotted lines give the vertical, central 50%- and 80%-quantile, respectively, while the solid line gives the corresponding median; at $\cos(\theta_{\mathrm{true}}) = 0$, the reconstruction is nearly unbiased, while moving towards the poles introduces a bias due to the asymmetry of the reconstructed phase-space.

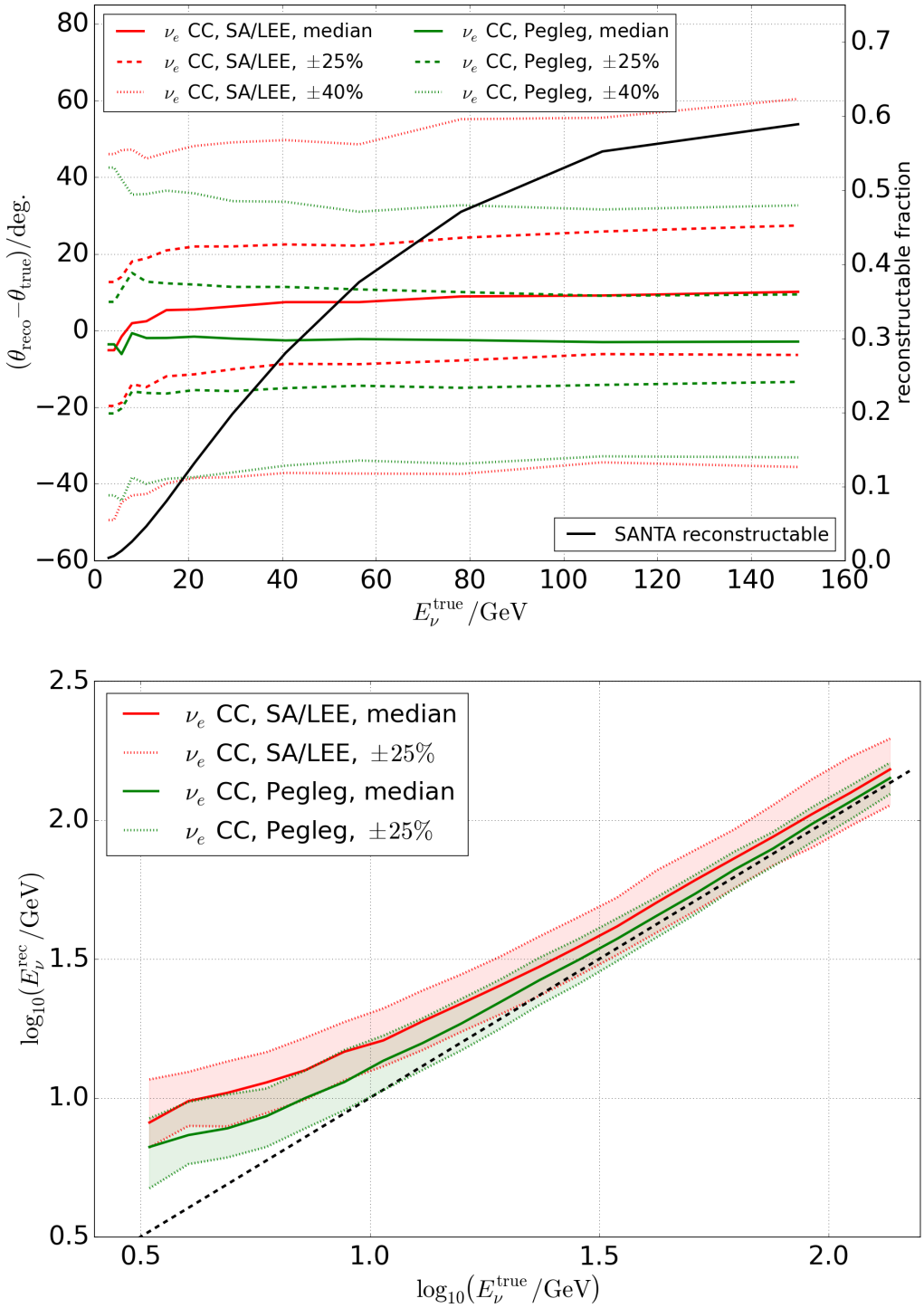


Figure E.8: Energy-dependent zenith (top) and energy (bottom) resolutions for electron neutrino events, comparing *SANTA/LEERA* to Pegleg (analogous to Figure 5.7 for muon neutrinos); in the top plot, the central, vertical 50%-quantile (80%) is enframed by the dashed (dotted) lines, while in the bottom plot, only the central 50%-quantile is shown; in both plots, the solid, colored lines indicate the vertical median, while the black, solid line (top plot), to be read from the right, vertical axis, gives the fraction of events reconstructable with SANTA.

F.2 Parametrization Boundaries and Renormalization

To generate a template of the analysis distribution (cf. Figure 6.1) from the KDEs, calculated in Section 6.2.1, two additional steps are required:

First, a boundary treatment must be applied. This is necessary, since KDEs tend do underestimate the PDF at a parametric boundary. Such boundary is for example the value of $\cos\left(\theta_\nu^{\mathrm{PL}}\right) = -1$. Below -1, no events enter the KDE calculation. Thus, the true PDF suddenly drops to zero at -1, while the value at $-1 + \epsilon$ with $\epsilon > 0$ and $\epsilon \to 0$ is substantially higher. The KDE has no information about this logical boundary and thus, tends to *smear-out* the edge at -1 according to the local bandwidth of the kernel. In other words, the KDE can not resolve jumps in the PDF and thus features probability density *flowing out* of the used parameter space, while no density is *flowing in* from outside events [129].

To account for this, the KDE is reflected at the boundary, such that no density is *flowing out* of the parameter space. Additionally, this maintains the integral of the KDE over the used parameter space to be 1.0. This KDE-reflection mechanism was investigated in details by Eric per Vogel [135] and is a commonly used method to account for boundaries of the underlying parameter space [134].

Second, the KDE needs to be integrated over the extent of every bin from the binning in Section 6.1. This is done by evaluating the KDE on a fine grid of 300×300 sampling points and generating a cubic, two-dimensional spline through these points. This two-dimensional spline can then easily be integrated numerically over the rectangular extent of each bin.

The resulting distribution obtained from integrating the two-dimensional KDE is then normalized to the sum of the weights of all MC events.

F.3 Performance Comparison of KDEs and Histograms

To benchmark the performance of the adaptive, weighted Kernel Density Estimation (KDE), presented in Section 6.2, a *Toy Monte Carlo* is generated.

To do this, a two-dimensional probability density function (PDF) is chosen, that resembles the distribution in energy and zenith used in the actual analysis. To simplify the analogy, the corresponding two parameters are labelled θ_{fake} and E_{fake} in the following.

The zenith-distribution is assumed to be uniformly distributed in $\cos(\theta_{\mathrm{fake}})$, while the energy distribution is generated as a modified Gamma distribution in $\log_{10}(E_{\mathrm{fake}}/\mathrm{GeV})$. The precise choice is of no relevance for the following Toy Monte Carlo. The resulting probability density function (PDF) is shown in Figure F.9.

From this two-dimensional PDF, Monte Carlo *events* are pulled randomly, while the number of events N is varied. Afterwards, the Monte Carlo sample is used to generate a probability distribution by using first a histogramm and second the adaptive, weighted KDE mentioned above.

For $N = 10^3$, the resulting two-dimensional probability distributions are shown in Figure F.10. The fluctuations in the histogram are clearly visible by eye, while the KDE returns a much smoother estimate of the true PDF.

Since the true PDF is known, the true probability distribution is obtained by integrating the PDF over the extent of each bin from Figure F.10. Then, the deviation of the histogram

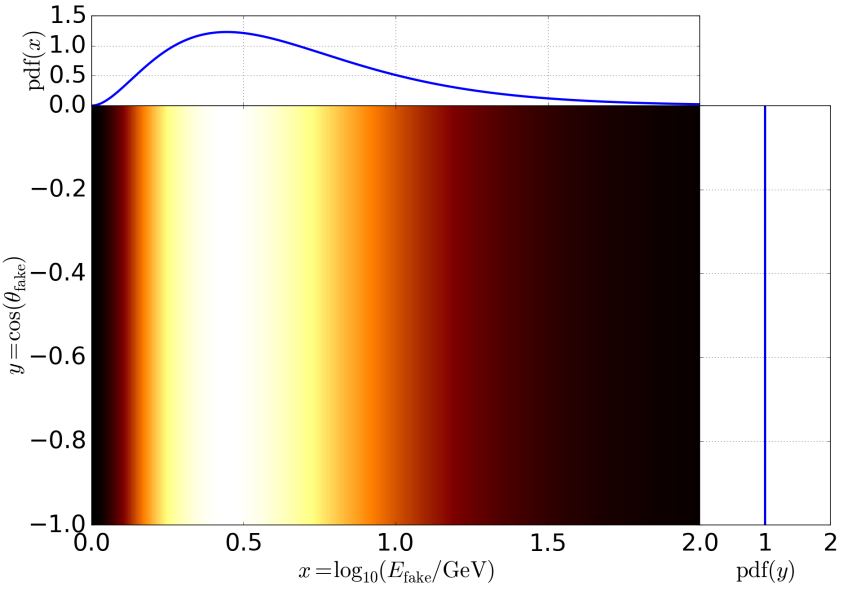


Figure F.9: Randomly chosen, two-dimensional PDF, where the two parameters $\log_{10}(E_{\text{fake}}/\text{GeV})$ and $\cos(\theta_{\text{fake}})$ resemble the reconstructed energy and zenith-angle distribution of oscillation analyses with DeepCore; the PDF is constructed from a uniform distribution in $\cos(\theta_{\text{fake}})$ and a Gamma-distribution in $\log_{10}(E_{\text{fake}}/\text{GeV})$.

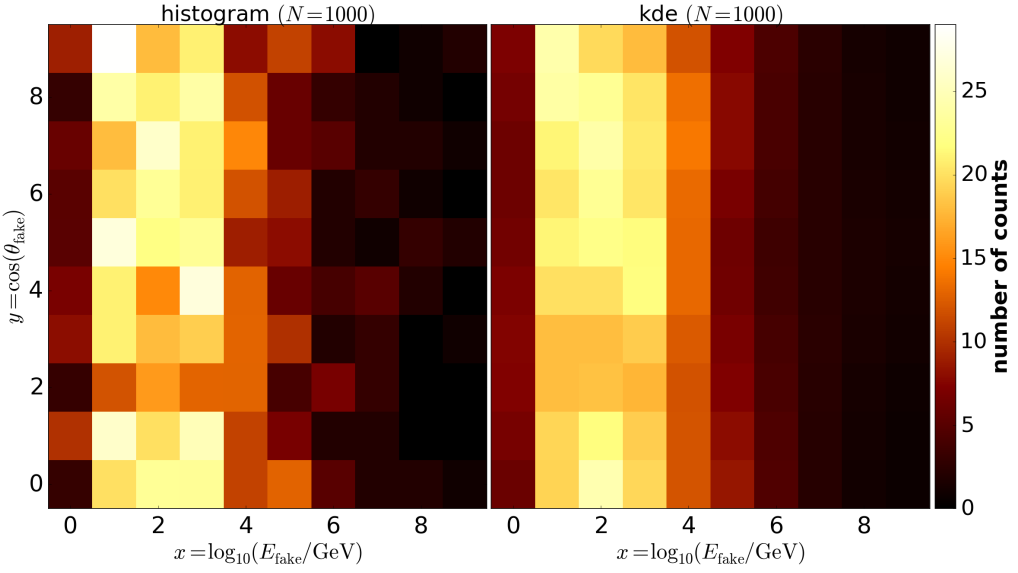


Figure F.10: Comparison between estimated probability distribution using a histogram (left) and Kernel Density Estimation (right): the Monte Carlo fluctuations for histograms are clearly visible, when compared to the underlying, true PDF in Figure F.9.

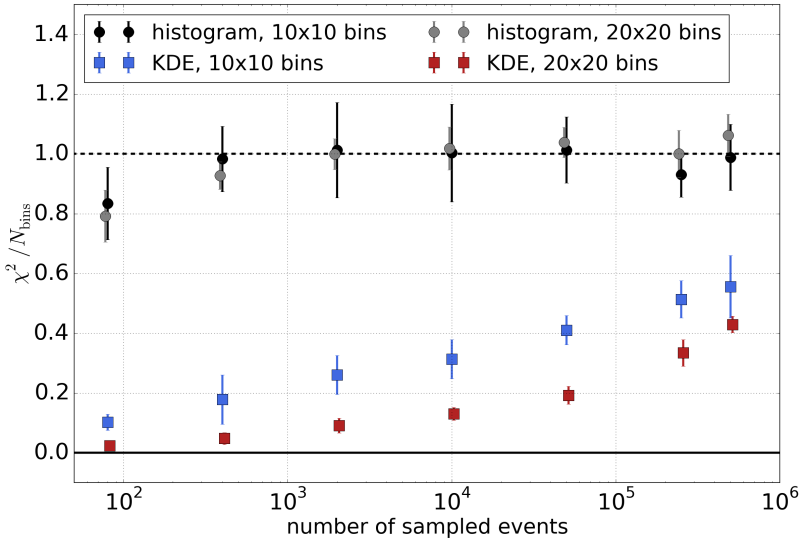


Figure F.11: Comparison of the performance of KDEs and histograms in estimating the true PDF of the Toy Monte Carlo from a given sample for 10×10 and 20×20 bins; the χ^2 compares the bin-content between the KDE or histogram and the true PDF from Figure F.9, assuming Poissonian statistics.

and the KDE from the true PDF are calculated in a χ^2-method. The result is shown in Figure F.11 as χ^2/N_{bins}, where N_{bins} is the number of used bins (e.g. for Figure F.10, this is $N_{\text{bins}} = 100$).

In Figure F.11, the deviations for histograms are clearly larger than for KDEs for any number of events N. However, the gain from KDEs decreases with an increasing number of events and with a decreasing number of bins. Both arises from the same effect: As the number of events increases, the kernel width decreases (cf. Equations 6.3). As soon as the width is small compared to the binning, the KDE converges to the sum of the weights of all events within that bin. Thus, the KDE approaches the known behavior from histograms, while the only events that contribute notably to several bins are those, that are close to one of the bin-edges.

The same effect occurs for increasing the number of bins: If the number of bins increases, the size of an individual bin shrinks, such that the typical kernel width becomes larger compared to the binning. Thus, the effect from in- and outflowing probability densities increases, while the fraction of events, that contribute almost exclusively to one bin, decreases.

Another way to see this is that the relative uncertainty of the bin-content increases for histograms with the number of bins. However, this is not the case for KDEs, which do not predict a probability distribution, but a PDF and derive the probability distribution simply by integrating the estimated PDF.

G Baseline-Correction for Detector Systematics

For each bin of the analysis distribution in Figure 6.1, three one-dimensional parametrizations p_s with s = 1,2,3 of the bin-content are derived to parametrize the effect of the optical

efficiency, the hole-ice and the high-forward parameters. At their baseline value x_0^s, the parametrizations are not required to pass the bin-content of the baseline sample y_0. The resulting offset $\delta y_s = y_s - y_0 = p_s(x_0^s) - y_0$ for the systematic s corrects for the statistical fluctuations of the baseline sample.

For example, in case the bin-content of the baseline sample is underfluctuating, the parametrizations p_s are expected to give $\delta y_s > 0$ for each of the systematics s. The size of the ratio $c_s = y_s/y_0 > 0$ gives an estimate for the correction c_s of the baseline bin-content y_0.

If all parametrizations were simply multiplied to the value of the baseline sample, the resulting bin-content would be given by Equation G.2

$$y(\{x_0^s\}) = \left(\prod_{s=1}^{3} p_s(x_0^s)\right) \cdot y_0 = \left(\prod_{s=1}^{3} (y^s/y_0)\right) \cdot y_0 \tag{G.2}$$

for all systematics s being at their baseline value x_0^s. Thus, the correction of the baseline value would be applied three times, once for each systematic, such that the baseline fluctuations are overcorrected. The resulting bias increases with the number of parametrized systematics.

To account for this, the correction terms c_s are not simply multiplied, but combined. To do this, they are averaged with respect to their uncertainty, as stated in Equation G.3

$$b = \frac{\sum_{s=1}^{3} c_s \sigma_s^{-2}}{\sum_{s=1}^{3} \sigma_s^{-2}}, \tag{G.3}$$

where b is the resulting *baseline-correction* and σ_s is the uncertainty of the correction c_s at the baseline point x_0^s. The uncertainties σ_s are estimated in Equation G.4 by the covariance matrix Cov^s returned from the individual fits of the parametrizations p_s:

$$\sigma_s = \frac{1}{y_0} \cdot \sqrt{\sum_{\{\tilde{q} \in p_s\}} \sum_{\{q \in p_s\}} \left(\frac{\mathrm{d}p_s}{\mathrm{d}\tilde{q}}(x_0^s)\right) \mathrm{Cov}^s_{\tilde{q},q} \left(\frac{\mathrm{d}p_s}{\mathrm{d}q}(x_0^s)\right)}, \tag{G.4}$$

where q and $\tilde{q}$ are the fitted parameters of the parametrization p_s.

For example, in case of a first-order polynomial $y_s = p_s(x_0^s) = a_1 \cdot x_0^s + a_0$ the two parameters are a_0 and a_1, such that Equation G.4 simplifies to:

$$\sigma_s = \frac{1}{y_0} \cdot \sqrt{\begin{pmatrix} x_0^s \\ 1 \end{pmatrix} \mathrm{Cov}^s \begin{pmatrix} x_0^s \\ 1 \end{pmatrix}}. \tag{G.5}$$

Thus, the baseline-correction b combines the corrections c_s, obtained for each systematic, and applies the correction once (instead of three times) to the baseline value y_0. This is done for every bin i and component $c \in \{\nu_e, \nu_\mu, \nu_\tau, \nu_{\mathrm{NC}}, \mu\}$, resulting in the factor b_{ci} in Equation 6.5. Moreover, the factor c_s is removed from each parametrization p_s, such that $p_s(x_0^s) = 1.0$ is obtained at the baseline value x_0^s.

Note that the resulting corrections b_{ci} are small and tested to have only minor impact on the result in Chapter 8. However, in previous studies with much smaller Monte Carlo samples (and thus larger corrections δy_s), combining the baseline correction factors c_s lead to a substantial improvement in describing statistically independent Monte Carlo compared to the three-times (over-correcting) method.

H Systematic Parameters

H.1 Systematic Data Samples

The following two tables summarize the Monte Carlo (MC) samples used to parametrize the detector systematics in Section 6.3.5. The neutrino samples are listed in Table H.3, while the atmospheric muon samples are listed in Table H.4.

The neutrino samples are labelled by the *sample ID*, introduced in Section 4.2.1. The letter $X \in \{2,4,6\}$ indicates the separate samples for electron, muon and tau neutrinos, respectively. Moreover, each set comprises $\sim 70\%$ neutrinos and $\sim 30\%$ anti-neutrinos.

Table H.3: Systematic samples used to parametrize the detector systematics for neutrinos: here X in the *sample ID* represents either 2, 4 or 6 for electron, muon and tau neutrinos, respectively; the individual columns state the parameters used for the generation of the samples, while the last column (*parametrized*) states whether the sample was used for the parametrization of detector systematics in Section 6.3.5.

sample ID	N_{coin}	ϵ_{opt}	$L^{\text{hi}}_{\text{scat}}$	$k^{\text{hi}}_{\text{fwd}}$	livetime	parametrized
1X640 (baseline)	0%	100%	25	0	30 years	yes
1X641	0%	88%	25	0	30 years	yes
1X643	0%	94%	25	0	30 years	yes
1X644	0%	97%	25	0	10 years	yes
1X646	0%	106%	25	0	10 years	yes
1X648	0%	112%	25	0	10 years	yes
1X660	0%	100%	15	0	10 years	yes
1X661	0%	100%	20	0	10 years	yes
1X662	0%	100%	30	0	10 years	yes
1X663	0%	100%	35	0	10 years	yes
1X670	0%	100%	25	2	10 years	yes
1X671	0%	100%	25	-5	10 years	yes
1X672	0%	100%	25	-3	10 years	yes
1X673	0%	100%	25	1	10 years	yes
1X674	0%	100%	25	-1	10 years	yes
1X640C	100%	100%	25	0	30 years	no
NG (NuGen)	0%	99%	alter.	model	10 years	yes

Besides the sample ID, Table H.3 shows the coincident fraction (N_{coin}), the optical efficiency (ϵ_{opt}), the hole-ice parameter ($L^{\text{hi}}_{\text{scat}}$), the high-forward parameter ($k^{\text{hi}}_{\text{fwd}}$) and the Monte Carlo livetime, used for the simulation of the sample. The final column (*parametrization*) states whether the sample was used for the parametrization of the detector systematics in Section 6.3.5, which is the case for all samples except 1X640C.

The coincident fraction N_{coin} is the fraction of neutrino events that contain a coincident atmospheric muon event. It is 0% for all samples, except 1X640C, as the standard MC samples are produced without taking coincident atmospheric muons into account. For 1X640C a pure sample of coincident events is generated to test their impact on the analysis distribu-

tion. From the rates of neutrinos and atmospheric muons, the fraction of coincident events is naively estimated to be $\sim 10\%$ of the data sample.

In Section 7.6, the 1X640C sample is used to derive the N_{coin} parameter for the template fit. There, the coincident fraction is found to have no impact on the analysis, since coincident muons are typically too faint to be seen or removed by one of the veto cuts. As a result, the coincident fraction is not used in the template fit (cf. Section 7.6).

The atmospheric muon samples are shown in Table H.4 with columns analogous to Table H.3.

Table H.4: Systematic samples used to parametrize the detector systematics for atmospheric muons with columns analogous to Table H.3.

sample ID	ϵ_{opt}	$L^{\text{hi}}_{\text{scat}}$	$k^{\text{hi}}_{\text{fwd}}$	livetime	parametrized
BS (baseline)	99%	25	0	1 year	yes
DE1	69.3%	30	0	1 year	yes
DE2	79.2%	30	0	1 year	yes
DE3	89.1%	30	0	1 year	yes
DE4	105%	25	0	3 years	yes
HI1	99%	15	0	1 year	yes
HI2	99%	20	0	1 year	yes
HI3	99%	30	0	1 year	yes
HF1	99%	30	-2	1 year	yes
HF2	99%	30	-4	1 year	yes

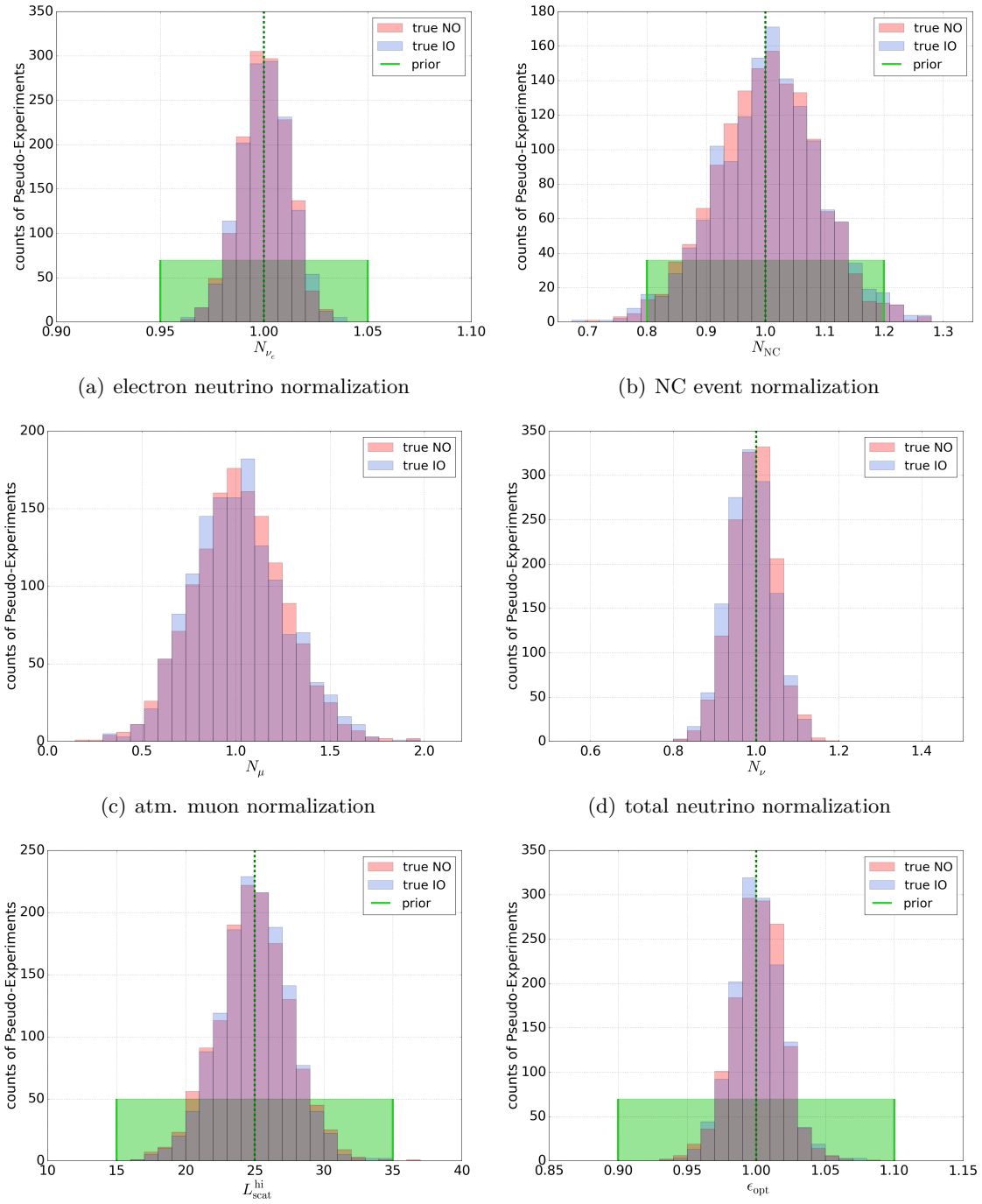
H.2 Distributions of Systematic Parameters for Pseudo-Experiments

The one-dimensional distributions of fitted systematic parameters for Pseudo-Experiments are shown in Figure H.12, H.13 and H.14. The width of each distribution describes the uncertainty expected from statistical fluctuations. It is driven by two effects: First, the sensitivity of the analysis towards the given parameter and second the Gaussian prior on the likelihood from Equation 6.7. If a parameter is used with a prior, the prior width is indicated by a green shaded band.

The interplay between the prior and the sensitivity of the analysis distribution leads to a rather non-intuitive effect: If the prior is much stronger than the sensitivity of the Pseudo-Data, the optimization of the parameter effectively minimizes the likelihood with respect to the prior. Thus, the parameter distribution becomes very peaked compared to the width and at the central value of the prior. However, if the prior has almost no impact and the Pseudo-Data is very sensitive to the parameter, the distribution also becomes very peaked compared to the prior width.

Thus, a very peaked distribution (compared to the width of the prior) can indicate either a very strong or a very weak prior with respect to the sensitivity of the Pseudo-Data. Note that this behavior changes, if the injected baseline value is different from the central value of the prior: In that case, the prior and the Pseudo-Data pull the parameter towards different values, leading to a more complicated picture.

In contrast, the width of the prior is comparably large, if both (the prior and the analysis



(a) electron neutrino normalization

(b) NC event normalization

(c) atm. muon normalization

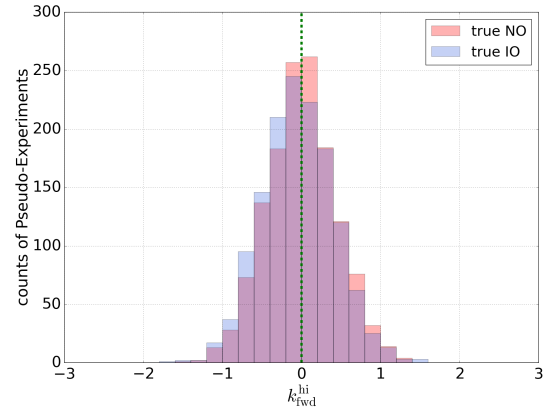
(d) total neutrino normalization

(e) hole-ice parameter

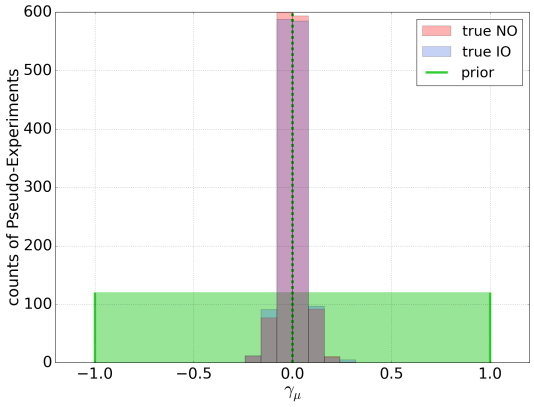
(f) optical efficiency

Figure H.12: One-dimensional distributions of systematic parameters (part I) for Pseudo-Experiments generated for NO (red) and IO (blue); the vertical dashed line indicates the injected value; for parameters used with a prior in the likelihood fit, the prior is indicated by a green shaded band.

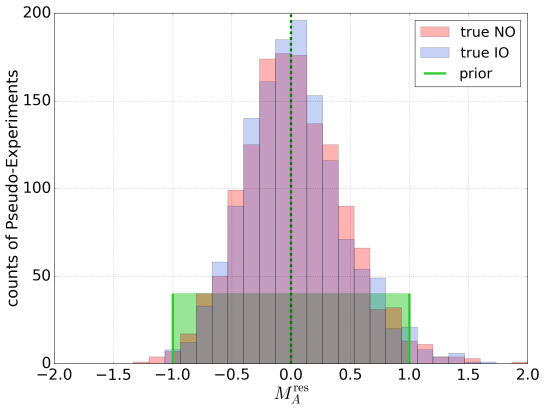
distribution) constrain the value of the parameter. However, it must obviously be smaller than the width of the prior for all cases.



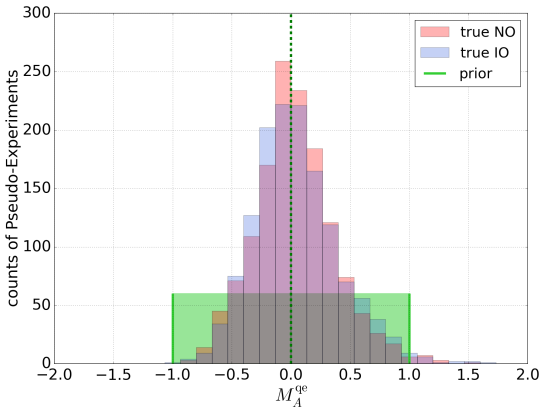
(a) high-forward parameter



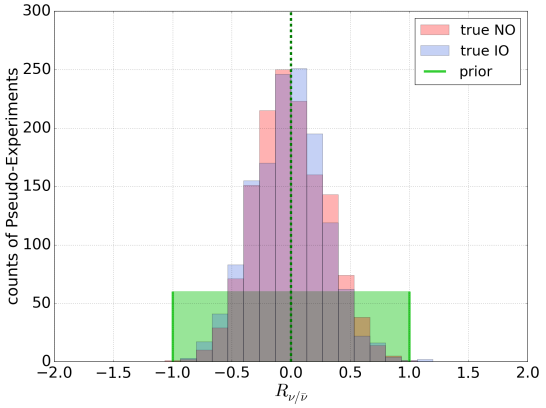
(b) atm. muon template uncertainty



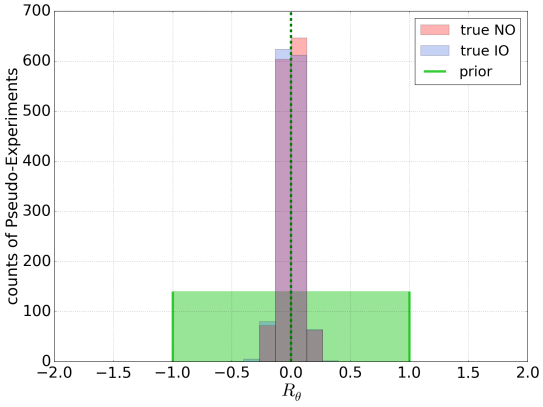
(c) axial mass for RES events



(d) axial mass for QE events



(e) energy-dependent $\nu/\bar{\nu}$ ratio



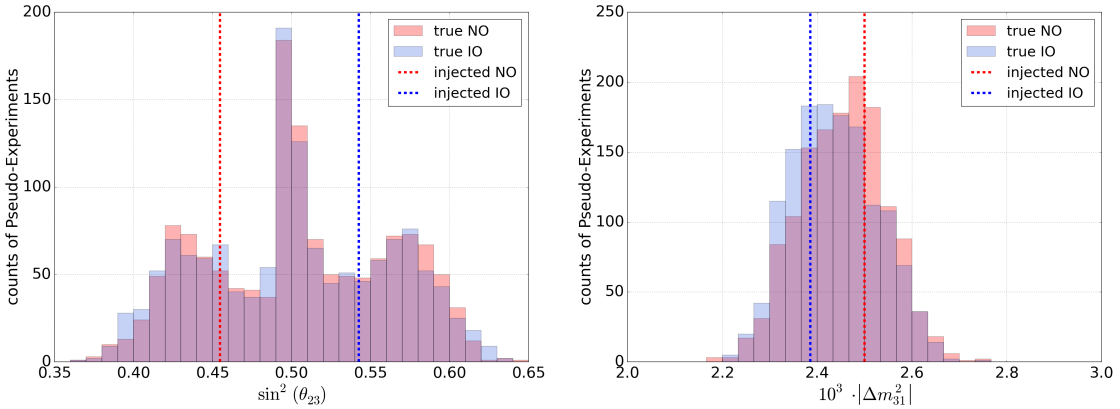
(f) zenith-dependent $\nu/\bar{\nu}$ ratio

Figure H.13: One-dimensional distributions of systematic parameters (part II) for Pseudo-Experiments generated for NO (red) and IO (blue); the vertical dashed line indicates the injected value; for parameters used with a prior in the likelihood fit, the prior is indicated by a green shaded band.

Note that in Figure H.14, the distributions of the fitted oscillation parameters $\sin^2(\theta_{23})$ and Δm^2_{31} do not peak at their injected values. This is due to the following reasons:

The colors of the blue and red distributions indicate the *injected* Neutrino Mass Ordering for the corresponding Pseudo-Experiments. Thus, each distribution in Δm^2_{31} (blue and red) contains a mixture of Pseudo-Experiments, prefering Normal and Inverted Ordering in the fit. Since the absolute value of Δm^2_{31} changes, when flipping the NMO, the blue distribution has a tail to larger values of $|\Delta m^2_{31}|$, caused by the Pseudo-Experiments prefering NO, although IO was injected. In the same way, the red distribution has a tail to smaller values of $|\Delta m^2_{31}|$, caused by the Pseudo-Experiments prefering IO, although NO was injected. This leads to the distributions observed for Δm^2_{31} in Figure H.14 (right).

For $\sin^2(\theta_{23})$, multiple effects are observed: First, a large fraction of Pseudo-Experiments gets stuck at the parametric boundary of $\sin^2(\theta_{23}) = 0.5$ from maximum mixing. Moreover, a maximum is observed in both octants, i.e. for $\sin^2(\theta_{23}) < 0.5$ and $\sin^2(\theta_{23}) > 0.5$, which is due to the non-convexity of the parameter space (cf. Section 6.5). Note that the maximum in both octants is not at the injected value. This is due to the binning of $\sin^2(\theta_{23})$, instead of $\sin(2\theta_{23})$ or θ_{23}: Since the injected value is close to maximum mixing, the peak and mean values of $\sin^2(\theta_{23})$, $\sin(2\theta_{23})$ and θ_{23} do not correspond to the same mixing angle. As a result, the peak in $\sin^2(\theta_{23})$ is no unbiased estimator for the injected value, as observed in Figure H.14 (left).



(a) atmospheric mixing angle

(b) atmospheric squared-mass difference

Figure H.14: One-dimensional distribution of oscillation parameters $\sin^2(\theta_{23})$ and Δm^2_{31} for Pseudo-Experiments, generated for NO (red) and IO (blue); the vertical dashed line indicates the injected value for NO and IO; the observed features are discussed in Appendix H.2.

In addition to these one-dimensional distributions, the correlation coefficients between all systematic parameters are shown in Figure H.15. These correlations are derived separately from Pseudo-Experiments for NO (upper triangle) and IO (lower triangle).

Note that all correlation coefficients are within $[-0.8, +0.8]$, while most coefficients are substantially smaller. In particular, the atmospheric oscillation parameters feature only very small correlations with other parameters. Thus, no pair of the parameters is completely degenerated. Moreover, the correlation coefficients feature the same structure for NO and IO, as one would expect due to the small difference between the orderings.

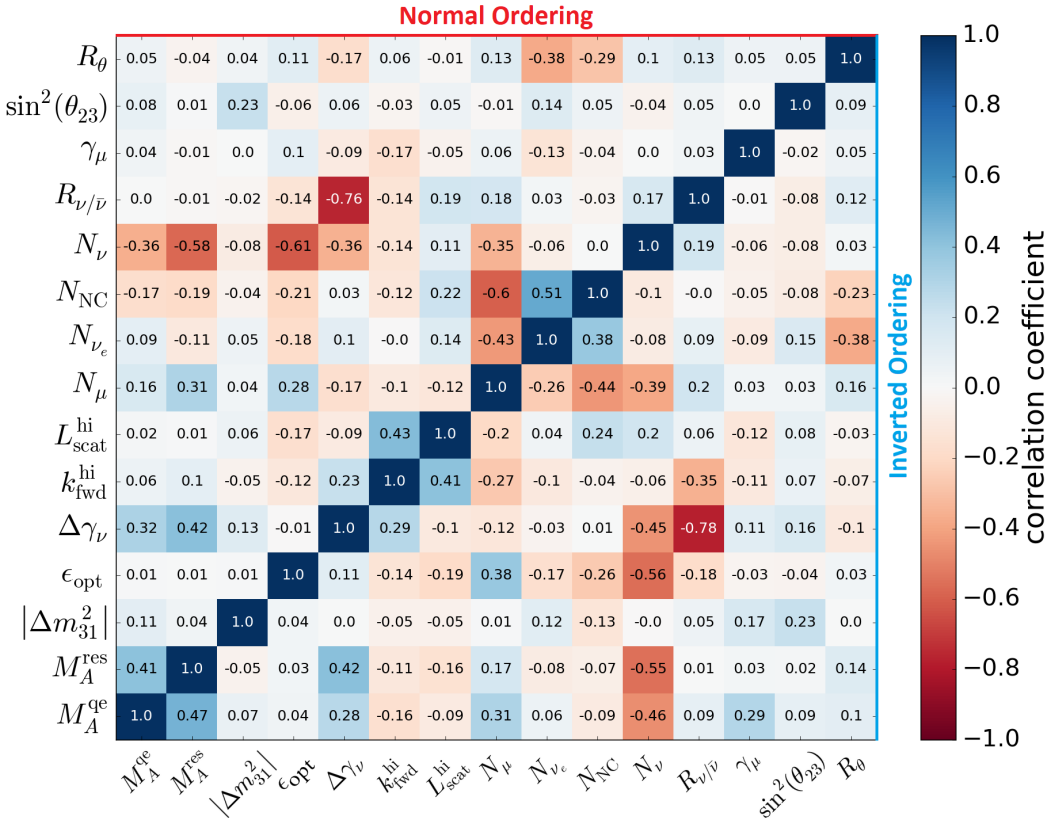


Figure H.15: Correlation matrix of systematic parameters obtained from the fits of Pseudo-Experiments for NO (upper triangle) and IO (lower triangle); the value in each cell gives the corresponding correlation coefficient.

I Validation of Experimental Fit

I.1 Pulls on Analysis Histogram

The pulls on the experimental analysis histogram are defined by the difference between the observed, experimental bin-content n^{exp}_{ijk} and the total MC prediction n^{MC}_{ijk} divided by the total uncertainty on the bin-content. The indices i,j,k describe the three-dimensions in reconstructed energy, zenith-angle and PID. The total uncertainty is given by the dominant Poissonian error $\sqrt{n^{\mathrm{MC}}_{ijk}}$ and the minor uncertainty on the MC template $\sigma^{\mathrm{MC}}_{ijk}$ (cf. Section 6.4), while both are combined quadratically in the denominator, such that

$$\mathrm{pull}_{ijk} = \frac{n^{\mathrm{exp}}_{ijk} - n^{\mathrm{MC}}_{ijk}}{\sqrt{n^{\mathrm{MC}}_{ijk} + (\sigma^{\mathrm{MC}}_{ijk})^2}}. \tag{I.6}$$

The sum over all pulls is conventionally taken as *goodness-of-fit* estimator for DeepCore analyses, as discussed in Section 8.1. Moreover, the one-dimensional distribution of pulls is shown in Figure I.16 (left) as black histogram. For comparison, a Gaussian distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$ is shown as a blue curve on top.

As expected, the pulls follow a nearly Gaussian behavior. Note that a small skewness, tilting the peak to the left-hand-side, is expected for these pulls, as the actual distribution of the bin-content is nearly Poissonian, instead of Gaussian. Moreover, the width of the pulls is

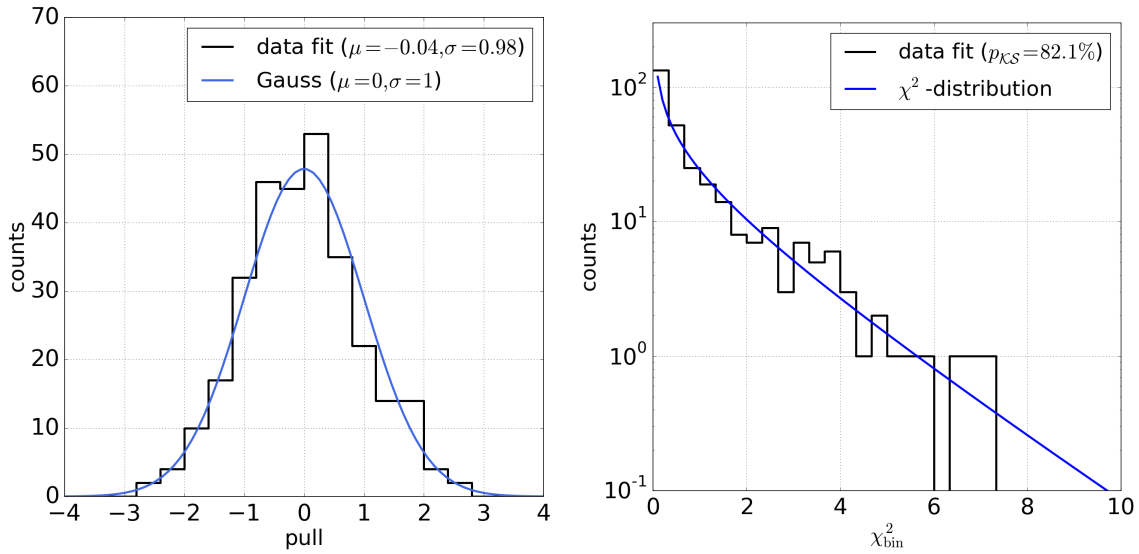


Figure I.16: Left: distribution of observed pulls in analysis distribution (black histogram), compared to a standard Gaussian distribution (blue curve) with mean value $\mu = 0$ and standard deviation $\sigma = 1$; right: corresponding distribution of squared pulls χ^2_{bin}, compared to a χ^2-distribution with $n_{\text{dof}} = 1$.

expected to be slightly below $\sigma = 1$, due to the fit of the systematic parameters.

In Figure I.16 (right), the corresponding distribution of squared pulls χ^2_{bin} is shown as black histogram and compared to a χ^2-distribution with a scale parameter of $n_{\text{dof}} = 1$, which is shown on top as a blue curve. The agreement of the pulls with the χ^2-distribution is tested in a *Komarov-Smirnov-Test* (KS-Test) [157], which tests for unexpected over- and undershooting of the observed values in any region of the distribution, independent of the bin-edges. The resulting p-value of $p_{\mathcal{KS}} = 82.1\%$ is well-compatible with the background expectation.

Overall, the observed fluctuations on the experimental distributions agree well with the total Monte Carlo template within the statistical uncertainties.

Finally, the full three-dimensional distribution of pulls is shown in Figure I.17. For all PID bins, the statistical fluctuations are clearly visible. To check whether any observed cluster of over- or underfluctuations is inconsistent with the statistical expectations, an additional test is applied.

To do this, a two-dimensional *filter* is defined by two integers $i,j = 1, ..., 10$. These integers describe a two-dimensional window that is moved over the analysis distribution, aggregating the bin-content of all bins within the filter. The filter width is given by the number of bins i and j in energy and zenith-angle, respectively. Obviously, for each PID bin with $N \times N$ bins in energy and zenith-angle, the number of possible filters is $(N + 1 - i) \times (N + 1 - j)$ with $N = 10$ for this case. For example, for $i = j = 1$, each bin is tested individually, such that the number of tested filters $(N + 1 - i) \times (N + 1 - j) = 100$ is the number of energy and zenith bins in each PID bin. In contrast, for $i = j = 10$ all 100 bins are aggregated, such that only $(N + 1 - i) \times (N + 1 - j) = 1$ single filter is tested. Note that the aggregated bin-content from adjacent filters is in general highly correlated.

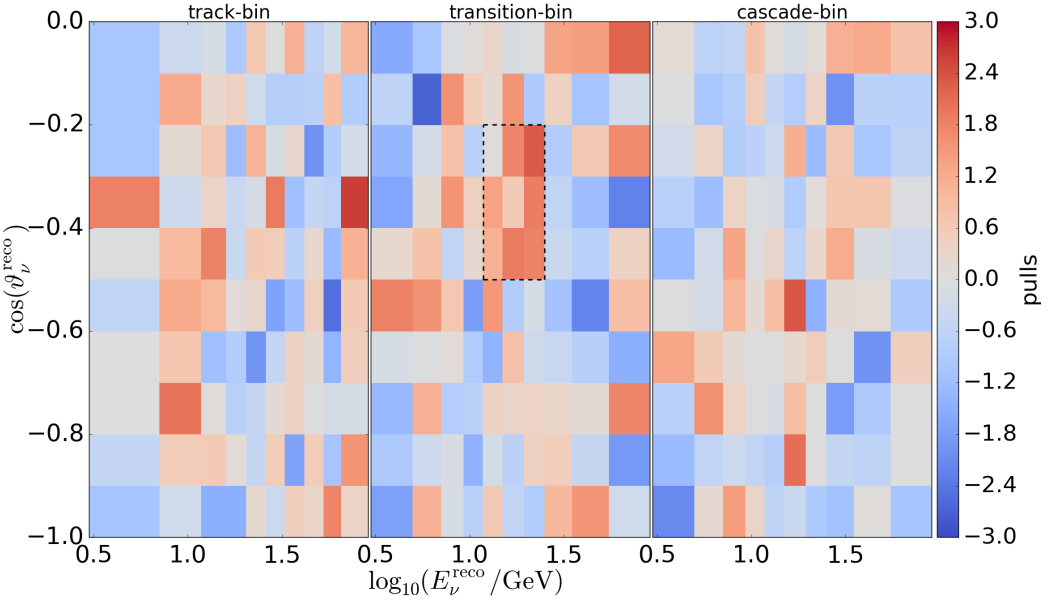


Figure I.17: Pulls on analysis distributions obtained from fitting three years of DeepCore data; these pulls are tested for unexpected clusters: the most significant deviation from MC is found in the transition-bin for $\cos(\theta_\nu^{\mathrm{reco}}) \in [-0.5, -0.2]$ and $\log_{10}(E_\nu^{\mathrm{reco}}/\mathrm{GeV}) \in [1.1,1.4]$ (marked by black, dashed line).

For every filter size $i \times j$, the maximum excess from all $(N+1-i) \times (N+1-j)$ filters is calculated for the experimental distribution in Figure I.17.

The same procedure is applied to Pseudo-Experiments. Thus, for each filter size $i \times j$ the observed values in data can be assigned a p-value derived from the distribution of excesses seen in Pseudo-Experiments. Finally, the most significant p-value of all filter sizes $i \times j$ is taken as the most significant deviation observed in the data distribution.

For the pulls shown in Figure I.17, the most significant excess is found for a 3×3 filter in the transition bin for $\cos(\theta_\nu^{\mathrm{reco}}) \in [-0.5, -0.2]$ and $\log_{10}(E_\nu^{\mathrm{reco}}/\mathrm{GeV}) \in [1.1,1.4]$ (filter marked by a black, dashed line). The aggregated excess in this filter corresponds to a $\sim 4.2\,\sigma$ deviation. The p-value of observing such excess in a 3×3 filter anywhere in the two-dimensional distribution is $\sim 0.1\%$. The p-value to observe such excess in any filter size $i \times j$ is 2.4%. The p-value of observing such excess in any filter of any size in any PID bin is calculated to 7.1%. Thus, although the p-value is still small compared to other *goodness-of-fit* estimators, it is still compatible with a rare statistical fluctuation.

In addition to the above test, the data events populating these 9 bins of the analysis distribution were tested for unexpected behavior in any of the control distributions. However, no suspicious behavior of these data events was found, that could indicate a systematic bias.

Finally, the excess is found in an off-signal region, i.e. in a region that does not show a strong signature for determining the NMO (cf. Figure 6.2). To assure that the NMO measurement is not affected by this excess, the corresponding 3×3 bins were masked in the likelihood calculation (i.e. removed from the sum over all bins in Equation 6.7) to derive a fit of the NMO without these events being included. The resulting best-fit remains in the left octant of the NO hypothesis, while the value of $\mathrm{LLH}_{\mathrm{NO-IO}}$ changed by -0.072 (-19%), i.e. increasing the preference for NO over IO slightly. Thus, even if the nature of the observed

excess was systematic and not statistical, it would be unlikely to affect the NMO result. Still, it seems reasonable to look for a similar behavior in future data.

I.2 Two-Neutrino Fit

The experimental result is tested for unphysical behavior in the mixing angle θ_{23}. This is done by fitting the two-flavor approximation from Equation 2.26 to experimental data. For this, $\sin(2\theta_{2\nu})$ is treated as free parameter without boundaries. More specifically, it is allowed to take values $\sin(2\theta_{2\nu}) > 1$, which is not compatible with any mixing angle $\theta_{2\nu} = \theta_{23}$. However, this is done for the following reason: The term $\sin^2(2\theta_{2\nu})$ in Equation 2.26 scales the effect of muon neutrino disappearance. Thus, $\sin^2(2\theta_{2\nu}) = 1$ is the maximum disapperance observable, that is compatible with standard oscillations. In case the background is overestimated in Monte Carlo, this can lead to more disappearance being observed than expected for $\sin(2\theta_{2\nu}) = 1$. Additionally, $\sin(2\theta_{2\nu}) > 1$ can be fitted due to statistical fluctuations without any bias in the fitted background templates. Thus, if $\sin(2\theta_{2\nu})$ is fitted significantly into the unphysical region of $\sin(2\theta_{2\nu}) > 1$, this is an indication for a mismodelling of the non-oscillating background.

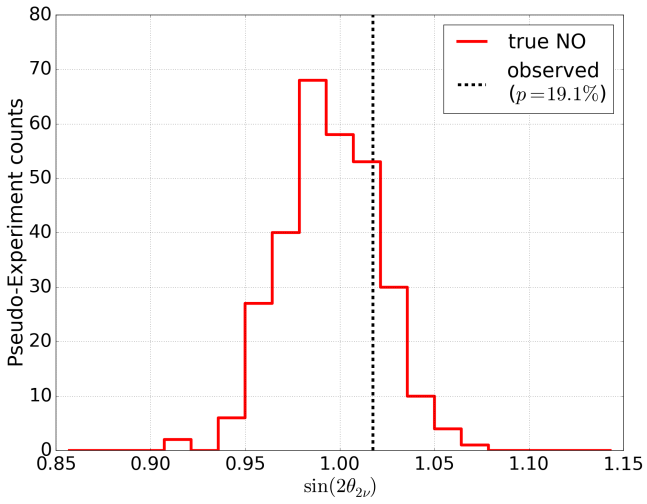


Figure I.18: Mixing angle $\theta_{2\nu}$ obtained from the fit of two-flavor vacuum oscillations (cf. Section 2.3.2), where the vertical, black line indicates the observed value and the background distributions is obtained from Pseudo-Experiments, generated with the best-fit parameters in the NO hypothesis, but fitted in the simplified, two-flavor model.

The resulting best-fit value for $\sin(2\theta_{2\nu})$ is shown in Figure I.18 as vertical, dashed, black line. To test whether the observed shift is compatible with standard oscillations, best-fit Pseudo-Experiments for NO are generated and fitted in the two-flavor approximation. The resulting distribution is shown in the background of Figure I.18. The resulting p-value for the observed data is 19.1%, which is well-compatible with standard oscillations. Thus, no indication for a mismodelling of the non-oscillating background is found by an unphysical fit of the two-flavor approximation.

I.3 Additional Tests of Unfitted Systematic Influences

Besides the systematic uncertainties that were discussed in Section 6.3 and 7.6, additional systematic influences were investigated that were not parametrized.

These additional tests are briefly summarized in the following and test the effects from:

- absorption and scattering in the bulk-ice,
- deviations in the optical properties of the hole-ice,
- variations in the properties of the *bubble-column* (cf. Section 3.4),
- fitting with and without KDE template generation and *baseline correction*,
- injection and fit of distributions from independent Monte Carlo.

The impact of the uncertainties on the optical bulk-ice properties were tested in two ways. First, neutrino simulations were produced with an updated bulk-ice model, called *Spice 3.2.1*. The corresponding GENIE samples were injected into the fitter and fitted with the baseline bulk-ice model. The differences in the depth-dependent scattering and absorption properties of the two models are on the $\lesssim 3\%$ level. For these fits, no notable pull on any systematic parameter was found, while the correct NMO and octant was recovered, independent of the injected NMO.

Second, systematic samples were generated with modified absorption and scattering properties of the ice. To do this, the scattering/absorption coefficients were varied by $+10\%/0\%$, $0\%/+10\%$ and $-7\%/-7\%$. Since these sets became available only after the *unblinding* of the presented analysis, they were not included in the list of fitted parameters. Instead, their impact on the analysis was tested after the presented analysis was finalized:

As the systematic samples for absorption and scattering were not produced *on-axis* (i.e. the absorption and scattering coefficients were varied at the same time), the parameters could not be parametrized independently, as done in Section 6.3. Instead, all detector systematics were parametrized in a five-dimensional *hyperplane*, where the bulk-ice properties were represented by two parameters $\lambda_{\text{scat}}^{\text{bulk}}$ and $\lambda_{\text{abs}}^{\text{bulk}}$ for the scattering and absorption properties of the bulk-ice, respectively. However, both parameters were found to deviate in the data fit by only $\sim 1\%$ from the baseline value, which is consistent with uncertainties stated by the Calibration Group of the IceCube Collaboration, but substantially smaller than the assumed (conservative) prior of 10%. As a result, no change in the observed ordering or octant was found and the original, unblinded result was kept unchanged.

Concerning the parametrization of the hole-ice properties, an alternative ice-model was previously used for the generation of GENIE Monte Carlo, which is commonly called *H2* [100]. However, this model was consistently found to give worse agreement between data and MC in several independent analysis. Thus, it was dropped at an early stage of this analysis.

Moreover, the impact of the existence of a bubble-column in the hole-ice was investigated at an early stage of this analysis and similar analyses. It was found that all models that could lead to a flip of the NMO could be excluded by bad data-Monte-Carlo agreement in the analysis or one of the control distributions.

Finally, it was tested that the preference for NO and IO is independent of the usage of KDEs (cf. Appendix 6.2.1) and the *baseline correction* (cf. Appendix G). Thus, even a fit using no baseline correction and no KDE to estimate the MC templates would recover the NO preference shown in this work. The same could be shown for splitting the MC sample into two samples of equal size and using one half for injecting an analysis template and one half for fitting it. There, the injected NMO could be recovered with independent MC for both ordering hypotheses.

Acknowledgements

It is an unexpected feeling to be done. All data is analysed, all corrections are made, and only the stack of papers, resting on my desk, awaits to be finally handed in.

Looking back on these years, it seems unreal that all of this journey is supposed to end now. I still remember my first collaboration meeting in Geneva. I was despairing of neither seeing a spot for me in the low-energy group nor a way forward for my analysis. Or summer 2017, when the data/MC agreement was so bad that Michael, Elim and I were almost sure, we would never graduate – or at least not with an analysis that was anyhow publishable.

These days seem ages ago now, and even the most meticulous reader won't find them, when looking through these pages. All frustration, all disappointment, all triumph and relief vanished in the sterile wording of scientific language or the irrelevance of *technical details*. From this, it is easy to believe that all of this was just *my* achievement and so I must spend a few lines on reminding myself that it wasn't.

First of all, I would like to thank Michael Larson and Elim Thompson (formerly Cheung). During the last years, barely a day went by without our slack-channel being flooded with endless discussions about our analyses, the data sample and the low-energy group. Looking back, I think it was a miracle, how the three of us came together in such a unique combination of personalities, capabilities and time. To the outside, this might have seemed ordinary, but I cannot imagine any other group of students to solve the problems we solved and work the way we worked.

Second, I would like to thank the Aachen group, especially Rene Reimann, who shared $\sim$ 5000 cups of coffee with me, during the last years. I will probably miss these coffees within the next months, wherever life may take me. The discussions we had in these breaks were a major source of encouragement – especially when things went bad. Moreover, I want to thank Christopher Wiebusch. Although discussions sometimes turn harsh and emotional in his group, he always supported all of his students to the end – even if this caused major conflicts with all the rest of IceCube. His strong belief in his group and his – sometimes irrational – optimism is probably one reason for his group to flourish the way it does.

However, beyond anything else I want to thank my parents. Their unlimited, unconditional, unexplainable love and support for me is one of the corner stones of all I am and did and there are no words to ever thank for this adequately. They made me the person I am today and without them, nothing I did would have been possible.

Finally, I would like to thank Evelyn Schneider for all her support and encouragement. Especially in 2017, she had to sacrifice a lot of weekends together, while I was sitting at my desk working, and she mercifully endured endless hours of me talking about neutrino physics, she probably found more boring than I'll ever know. In the belief, that the following months will give us more time together, I am looking forward to them gratefully.

Bibliography

[1] Q. R. Ahmad *et al.*, "Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory", *Phys. Rev. Lett.*, vol. 87, p. 071 301, 7 2001. DOI: 10.1103/PhysRevLett.87.071301.

[2] T. Kajita, "Atmospheric neutrino results from Super-Kamiokande and Kamiokande - evidence for ν_μ oscillations", *Nucl. Phys. B*, vol. 77, no. 1, pp. 123 –132, 1999. DOI: 10.1016/S0920-5632(99)00407-7.

[3] S. Adrian-Martinez *et al.*, "Letter of intent for KM3NeT 2.0", *J. Phys. G*, vol. 43, no. 8, p. 084 001, 2016. DOI: 10.1088/0954-3899/43/8/084001.

[4] X. Qian and P. Vogel, "Neutrino mass hierarchy", *Prog. Part. Nucl. Phys.*, vol. 83, no. Supplement C, pp. 1 –30, 2015. DOI: 10.1016/j.ppnp.2015.05.002.

[5] P. Guzowski *et al.*, "Combined limit on the neutrino mass from neutrinoless double-β decay and constraints on sterile Majorana neutrinos", *Phys. Rev. D*, vol. 92, no. 1, p. 012 002, 2015. DOI: 10.1103/PhysRevD.92.012002.

[6] F. Couchot *et al.*, "Cosmological constraints on the neutrino mass including systematic uncertainties", *A&A*, vol. 606 A, 2017. DOI: 10.1051/0004-6361/201730927.

[7] M. G. Aartsen *et al.*, "Letter of intent: the Precision IceCube Next Generation Upgrade (PINGU) (version 2)", 2017. arXiv: 1401.2046v2.

[8] C. L. Cowan Jr. *et al.*, "Detection of the free neutrino: a confirmation", *Science*, vol. 124, pp. 103–104, 1956. DOI: 10.1126/science.124.3212.103.

[9] M. Thomson, *Modern particle physics*. New York: Cambridge University Press, 2013, ISBN: 978 110 703 426 6.

[10] 'MissMJ', *Standard model of elementary particles*, accessed at Dec. 12th, 2017. [Online]. Available: https://upload.wikimedia.org/wikipedia/commons/0/00/Standard_Model_of_Elementary_Particles.svg.

[11] C. Patrignani *et al.*, "Review of particle physics", *Chin. Phys. C*, vol. 40, no. 10, p. 100 001, 2016. DOI: 10.1088/1674-1137/40/10/100001.

[12] F. Capozzi *et al.*, "Global constraints on absolute neutrino masses and their ordering", *Phys. Rev. D*, vol. 95, p. 096 014, 9 2017. DOI: 10.1103/PhysRevD.95.096014.

[13] G. Bhattacharyya, "Electroweak symmetry breaking and BSM physics (a review)", *Pramana*, vol. 72, pp. 37–54, 2009. DOI: 10.1007/s12043-009-0004-0.

[14] M. Czakon, M. Zralek, and J. Gluza, "Are neutrinos Dirac or Majorana particles?", *Acta Phys. Polon. B*, vol. 30, pp. 3121–3138, 1999. arXiv: hep-ph/9910357.

[15] J. Hewett *et al.*, "Neutrino scattering", in *2011 workshop on Fundamental Physics at the Intensity Frontier, Rockville, MD, USA*, 2012. DOI: 10.2172/1042577.

[16] E. A. Paschos, A. Kartavtsev, and G. J. Gounaris, "Coherent pion production by neutrino scattering off nuclei", *Phys. Rev. D*, vol. 74, p. 054 007, 2006. DOI: 10.1103/PhysRevD.74.054007.

[17] T. Nakaya *et al.*, *Neutrino interactions*, ICISE group, accessed at Apr. 25th, 2018. [Online]. Available: http://ifirse.icise.vn/nugroup/Neutrino-Interactions.html.

[18] J. A. Formaggio and G. P. Zeller, "From eV to EeV: neutrino cross sections across energy scales", *Rev. Mod. Phys.*, vol. 84, pp. 1307–1341, 3 2012. DOI: 10.1103/RevModPhys.84.1307.

[19] K. S. McFarland, "Neutrino interactions", in *61st Scottish Universities Summer School in Physics, , St. Andrews, UK*, 2008, pp. 65–90, ISBN: 978 142 008 239 5. [Online]. Available: http://www.crcpress.com/product/isbn/9781420082395.

[20] A. Osipowicz *et al.*, "KATRIN: a next generation tritium beta decay experiment with sub-eV sensitivity for the electron neutrino mass: letter of intent", 2001. arXiv: hep-ex/0109033.

[21] Couchot, F. *et al.*, "Cosmological constraints on the neutrino mass including systematic uncertainties", *A&A*, vol. 606, A 104, 2017. DOI: 10.1051/0004-6361/201730927.

[22] B. Pontecorvo, "Mesonium and antimesonium", *Sov. J. Exp. Theor. Phys.*, vol. 6, no. 2, p. 429, 1958.

[23] C. Giunti and C. W. Kim, *Fundamentals of neutrino physics and astrophysics*. 2007, ISBN: 978 019 850 871 7.

[24] Y.-L. Chan *et al.*, "Wave-packet treatment of reactor neutrino oscillation experiments and its implications on determining the neutrino mass hierarchy", *Eur. Phys. J. C*, vol. 76, no. 6, p. 310, 2016. DOI: 10.1140/epjc/s10052-016-4143-4.

[25] I. Esteban *et al.*, "Updated fit to three neutrino mixing: exploring the accelerator-reactor complementarity", *J. High Energ. Phys.*, vol. 2017, no. 1, p. 87, 2017. DOI: 10.1007/JHEP01(2017)087.

[26] I. Esteban *et al.*, *NuFIT 3.2*, best-fit, Jan. 2018. [Online]. Available: www.nu-fit.org.

[27] A. Upadhyay and M. Batra, "Phenomenology of neutrino mixing in vacuum and matter", *ISRN High Energy Phys.*, p. 206 516, 2013. DOI: 10.1155/2013/206516.

[28] H. Nunokawa, S. J. Parke, and J. W. F. Valle, "CP-violation and neutrino oscillations", *Prog. Part. Nucl. Phys.*, vol. 60, pp. 338–402, 2008. DOI: 10.1016/j.ppnp.2007.10.001.

[29] NASA, *Earth fact sheet*, accessed at Dec. 12th, 2017, 2016. [Online]. Available: https://nssdc.gsfc.nasa.gov/planetary/factsheet/earthfact.html.

[30] K. Hagiwara, N. Okamura, and K.-i. Senda, "The earth matter effects in neutrino oscillation experiments from Tokai to Kamioka and Korea", *J. High Energy Phys.*, vol. 2011, no. 9, p. 82, 2011. DOI: 10.1007/JHEP09(2011)082.

[31] L. Wolfenstein, "Neutrino oscillations in matter", *Phys. Rev. D*, vol. 17, pp. 2369–2374, 9 1978. DOI: 10.1103/PhysRevD.17.2369.

[32] S. P. Mikheyev and A. Y. Smirnov, "Resonant amplification of ν oscillations in matter and solar-neutrino spectroscopy", *Il Nuovo Cimento C*, vol. 9, no. 1, pp. 17–26, 1986. DOI: 10.1007/BF02508049.

[33] I. Mocioiu and R. Shrock, "Matter effects on neutrino oscillations in long baseline experiments", *Phys. Rev. D*, vol. 62, p. 053 017, 5 2000. DOI: 10.1103/PhysRevD.62.053017.

[34] Q. Liu, S. Mikheyev, and A. Smirnov, "Parametric resonance in oscillations of atmospheric neutrinos?", *Phys. Lett. B*, vol. 440, no. 3, pp. 319 –326, 1998. DOI: 10.1016/S0370-2693(98)01102-2.

[35] V. F. Hess, " Ueber Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten ", *Z. Phys.*, vol. 13, p. 1084, 1912. [Online]. Available: http://cds.cern.ch/record/262750.

[36] K. Olive and P. D. Group, "Review of particle physics", *Chin. Phys. C*, vol. 38, no. 9, p. 090 001, 2014.

[37] M. Tanabashi *et al.*, "Review of particle physics: cosmic rays", *Phys. Rev. D*, vol. 98, p. 030 001, 2018. [Online]. Available: pdglive.lbl.gov/.

[38] G. T. Zatsepin and V. A. Kuz'min, "Upper limit of the spectrum of cosmic rays", *Sov. J. Exp. Theor. Phys.*, vol. 4, p. 78, 1966.

[39] K. Greisen, "End to the cosmic-ray spectrum?", *Phys. Rev. Lett.*, vol. 16, pp. 748–750, 17 1966. DOI: 10.1103/PhysRevLett.16.748.

[40] T. Gaisser, R. Engel, and E. Resconi, *Cosmic rays and particle physics*. Cambridge University Press, 2016, ISBN: 978 131 659 843 6.

[41] S. Euler, "Observation of oscillations of atmospheric neutrinos with the IceCube Neutrino Observatory", PhD thesis, III. Phys. Inst. B, RWTH Aachen University, 2014.

[42] F. Halzen and L. Wille, "Charm contribution to the atmospheric neutrino flux", *Phys. Rev. D*, vol. 94, no. 1, p. 014 014, 2016. DOI: 10.1103/PhysRevD.94.014014.

[43] T. K. Gaisser and M. Honda, "Flux of atmospheric neutrinos", *Annu. Rev. Nucl. Part. Sci.*, vol. 52, no. 1, pp. 153–199, 2002. DOI: 10.1146/annurev.nucl.52.050102.090645.

[44] M. Honda *et al.*, "Improvement of low energy atmospheric neutrino flux calculation using the JAM nuclear interaction model", *Phys. Rev. D*, vol. 83, p. 123 001, 12 2011. DOI: 10.1103/PhysRevD.83.123001.

[45] E. Richard *et al.*, "Measurements of the atmospheric neutrino flux by Super-Kamiokande: energy spectra, geomagnetic effects, and solar modulation", *Phys. Rev. D*, vol. 94, p. 052 001, 5 2016. DOI: 10.1103/PhysRevD.94.052001.

[46] M. Honda *et al.*, "Atmospheric neutrino flux calculation using the NRLMSISE-00 atmospheric model", *Phys. Rev. D*, vol. 92, p. 023 004, 2 2015. DOI: 10.1103/PhysRevD.92.023004.

[47] G. D. Barr *et al.*, "Uncertainties in atmospheric neutrino fluxes", *Phys. Rev. D*, vol. 74, p. 094 009, 2006. DOI: 10.1103/PhysRevD.74.094009.

[48] M. Honda *et al.*, "New calculation of the atmospheric neutrino flux in a three-dimensional scheme", *Phys. Rev. D*, vol. 70, p. 043 008, 4 2004. DOI: 10.1103/PhysRevD.70.043008.

[49] A. Bulmahn and M. H. Reno, "Secondary atmospheric tau neutrino production", *Phys. Rev. D*, vol. 82, p. 057 302, 2010. DOI: 10.1103/PhysRevD.82.057302.

[50] Z. Li *et al.*, "A measurement of the tau neutrino cross section in atmospheric neutrino oscillations with Super-Kamiokande", 2017. arXiv: 1711.09436.

[51] J. de André, "Sensitivity of IceCube-Gen2 Phase I to tau neutrino appearance", in *APS Meeting Abstracts*, accessed at Jun. 23rd, 2018, 2017. [Online]. Available: http://adsabs.harvard.edu/abs/2017APS..APRM11002D.

[52] P. Adamson *et al.*, "Measurements of atmospheric neutrinos and antineutrinos in the MINOS far detector", *Phys. Rev. D*, vol. 86, p. 052 007, 2012. DOI: 10.1103/PhysRevD.86.052007.

[53] T. Gaisser, "Seasonal variation of atmospheric neutrinos in IceCube", in *33rd International Cosmic Ray Conference (ICRC2013): Rio de Janeiro, Brazil*, p. 0492. [Online]. Available: http://www.cbpf.br/%7Eicrc2013/papers/icrc2013-0492.pdf.

[54] C. Rott, A. Taketa, and D. Bose, "Spectrometry of the earth using neutrino oscillations", *Scientific Reports*, vol. 5, 2015. DOI: 10.1038/srep15225.

[55] Y. D. O. Arcos and S. Sahu, "Earth matter effect on GeV neutrino propagation", 2015. arXiv: 1510.07103.

[56] M. Wallraff and C. Wiebusch, "Calculation of oscillation probabilities of atmospheric neutrinos using nuCraft", *Comput. Phys. Commun.*, vol. 197, pp. 185–189, 2015. DOI: 10.1016/j.cpc.2015.07.010.

[57] A. Dziewonski and D. Anderson, "Preliminary reference earth model", *Phys. Earth Plan. Int.*, vol. D 25, pp. 297–356, 1981, doi:10.17611/DP/9991844.

[58] R. Wendell, *Prob3++ software for computing three flavor neutrino oscillation probabilities*, accessed at Apr. 1st, 2018, 2012. [Online]. Available: http://webhome.phy.duke.edu/~raw22/public/Prob3++.

[59] V. Barger *et al.*, "Matter effects on three-neutrino oscillations", *Phys. Rev. D*, vol. 22, pp. 2718–2726, 1980. DOI: 10.1103/PhysRevD.22.2718.

[60] M. G. Aartsen *et al.*, "Measurement of atmospheric neutrino oscillations at 6-56 GeV with IceCube DeepCore", *Phys. Rev. Lett.*, vol. 120, no. 7, p. 071 801, 2018. DOI: 10.1103/PhysRevLett.120.071801.

[61] F. Huang and IceCube Collaboration, "Atmospheric tau neutrino appearance analysis with IceCube/DeepCore", in *APS Meeting Abstracts*, 2017, J11.001. [Online]. Available: http://adsabs.harvard.edu/abs/2017APS..APRJ11001H.

[62] G. Giordano, O. Mena, and I. Mocioiu, "Atmospheric neutrino oscillations and tau neutrinos in ice", *Phys. Rev. D*, vol. 81, p. 113 008, 2010. DOI: 10.1103/PhysRevD.81.113008.

[63] M. G. Aartsen *et al.*, "PINGU: a vision for neutrino and particle physics at the South Pole", *J. Phys. G*, vol. 44, no. 5, p. 054 006, 2017. DOI: 10.1088/1361-6471/44/5/054006.

[64] S. Bourret *et al.*, "Neutrino oscillation tomography of the earth with KM3NeT-ORCA", *J. Phys. Conf. Ser.*, vol. 888, no. 1, p. 012 114, 2017. DOI: 10.1088/1742-6596/888/1/012114.

[65] K. Abe *et al.*, "Measurements of neutrino oscillation in appearance and disappearance channels by the T2K experiment with 6.6×10^{20} protons on target", *Phys. Rev. D*, vol. 91, p. 072 010, 7 2015. DOI: 10.1103/PhysRevD.91.072010.

[66] M. Wascko, "T2K status, results, and plans", in *Neutrino 2018, Heidelberg, Germany*, 2018. DOI: 10.5281/zenodo.1286752.

[67] A. Holin, "Results from the MINOS experiment and new MINOS+ data", *PoS*, vol. NUFACT2014, p. 028, 2014. arXiv: 1507.08564.

[68] A. Aurisano, "Recent results from MINOS and MINOS+", in *Neutrino 2018, Heidelberg, Germany*, 2018. DOI: 10.5281/zenodo.1286760.

[69] P. Adamson *et al.*, "Constraints on oscillation parameters from ν_e appearance and ν_μ disappearance in NOνA", *Phys. Rev. Lett.*, vol. 118, no. 23, p. 231 801, 2017. DOI: 10.1103/PhysRevLett.118.231801.

[70] M. Sanchez, "NOνA results and prospects", in *Neutrino 2018, Heidelberg, Germany*, 2018. DOI: 10.5281/zenodo.1286758.

[71] Y. Hayato, "Atmospheric neutrino results from Super-Kamiokande", in *Neutrino 2018, Heidelberg, Germany*, 2018. DOI: 10.5281/zenodo.1298567.

[72] M. Blennow *et al.*, "Quantifying the sensitivity of oscillation experiments to the neutrino mass ordering", *J. High Energy Phys.*, vol. 03, p. 028, 2014. DOI: 10.1007/JHEP03(2014)028. arXiv: 1311.1822.

[73] E. Kemp, "The Deep Underground Neutrino Experiment – DUNE: the precision era of neutrino physics", in *4th Caribbean Symposium on Cosmology, Gravitation, Nuclear and Astroparticle Physics (STARS2017) Havana, Cuba*, vol. 338, 2017, pp. 993–999. DOI: 10.1002/asna.201713417.

[74] J. Martin-Albo, "Sensitivity of DUNE to long-baseline neutrino oscillation physics", in *2017 European Physical Society Conference on High Energy Physics (EPS-HEP 2017) Venice, Italy*, vol. EPS-HEP2017, 2017, p. 122. DOI: 10.22323/1.314.0122. arXiv: 1710.08964.

[75] G. Yang, "Neutrino mass hierarchy determination at reactor antineutrino experiments", in *12th Conference on the Intersections of Particle and Nuclear Physics (CIPANP 2015): Vail, Colorado, USA*, 2015. arXiv: 1509.08747.

[76] F. An *et al.*, "Neutrino physics with JUNO", *J. Phys. G*, vol. 43, no. 3, p. 030 401, 2016. DOI: 10.1088/0954-3899/43/3/030401.

[77] K. Abe *et al.*, "A long baseline neutrino oscillation experiment using J-PARC neutrino beam and Hyper-Kamiokande", 2014. arXiv: 1412.4673.

[78] M. Aartsen *et al.*, "The IceCube Neutrino Observatory: instrumentation and online systems", *J. Instrum.*, vol. 12, no. 03, P03012, 2017. DOI: 10.1088/1748-0221/12/03/P03012.

[79] The IceCube Collaboration, *IceCube internal gallery*, accessed at Jan. 10th, 2018, 2012. [Online]. Available: http://gallery.icecube.wisc.edu.

[80] P. A. Cherenkov, "Visible emission of clean liquids by action of γ radiation", *Doklady Akademii Nauk SSSR*, vol. 2, p. 451, 1934.

[81] M. Aartsen *et al.*, "Measurement of atmospheric tau neutrino appearance with IceCube DeepCore", *Phys. Rev. D*, 2018, publication pending.

[82] R. Abbasi *et al.*, "Calibration and characterization of the IceCube photomultiplier tube", *Nucl. Instrum. Methods Phys. Res. A*, vol. 618, no. 1, pp. 139 –152, 2010. DOI: 10.1016/j.nima.2010.03.102.

[83] R. Abbasi *et al.*, "The design and performance of IceCube DeepCore", *Astropart. Phys.*, vol. 35, no. 10, pp. 615 –624, 2012. DOI: 10.1016/j.astropartphys.2012.01.004.

[84] R. Abbasi *et al.*, "IceTop: the surface component of IceCube", *Nucl. Instrum. Methods Phys. Res. A*, vol. 700, pp. 188–220, 2013. DOI: 10.1016/j.nima.2012.10.067.

[85] I. Bartos, Z. Marka, and S. Marka, "Infused ice can multiply IceCube's sensitivity", 2017. arXiv: 1706.06589.

[86] L. Rädel and C. Wiebusch, "Calculation of the Cherenkov light yield from electromagnetic cascades in ice with Geant4", *Astropart. Phys.*, vol. 44, pp. 102–113, 2013. DOI: 10.1016/j.astropartphys.2013.01.015.

[87] M. Leuermann, "Search for diffuse neutrino point sources using a multipole analysis in IceCube", Master's thesis, III. Phys. Inst. B, RWTH Aachen University, 2013.

[88] I. Frank and I. Tamm, "Coherent visible radiation of fast electrons passing through matter", in *Selected Papers*, Springer, 1991, pp. 29–35.

[89] M. Vehring, "Measurement of neutrino oscillations with IceCube-DeepCore", PhD thesis, III. Phys. Inst. B, RWTH Aachen University, 2017.

[90] D. Chirkin and W. Rhode, "Muon Monte Carlo: a high-precision tool for muon propagation through matter", 2004. arXiv: hep-ph/0407075.

[91] R. Abbasi *et al.*, "First search for atmospheric and extraterrestrial neutrino-induced cascades with the IceCube detector", *Phys. Rev. D*, vol. 84, no. 7, p. 072 001, 2011.

[92] A Grant, "A Monte Carlo calculation of high energy hadronic cascade in matter", *Nucl. Instrum. Meth.*, vol. 131, no. 1, pp. 167–172, 1975.

[93] S. Kretzer and M. H. Reno, "Tau neutrino deep inelastic charged current interactions", *Phys. Rev. D*, vol. 66, p. 113 007, 11 2002. DOI: 10.1103/PhysRevD.66.113007.

[94] M. G. Aartsen *et al.*, "Search for astrophysical tau neutrinos in three years of IceCube data", *Phys. Rev. D*, vol. 93, no. 2, p. 022 001, 2016. DOI: 10.1103/PhysRevD.93.022001.

[95] M. Wallraff, "Design, implementation and test of a new feature extractor for the IceCube Neutrino Observatory", Master's thesis, III. Phys. Inst. B, RWTH Aachen University, 2010.

[96] P. Hallen, "On the measurement of high-energy tau neutrinos with IceCube", Master's thesis, III. Phys. Inst. B, RWTH Aachen University, 2013.

[97] M. Aartsen *et al.*, "Measurement of South Pole ice transparency with the IceCube LED calibration system", *Nucl. Instrum. Methods Phys. Res. A*, vol. 711, pp. 73 –89, 2013. DOI: 10.1016/j.nima.2013.01.054.

[98] H Craig, H Shoji, and C. C. Langway, "Nonequilibrium air clathrate hydrates in Antarctic ice: a paleopiezomdter for polar ice caps", *Proc. Natl. Acad. Sci. USA*, vol. 90, no. 23, pp. 11 416–11 418, 1993. DOI: 10.1073/pnas.90.23.11416.

[99] M. G. Aartsen *et al.*, "The IceCube Neutrino Observatory part VI: ice properties, reconstruction and future developments", in *33rd International Cosmic Ray Conference (ICRC2013): Rio de Janeiro, Brazil*, 2013. arXiv: 1309.7010.

[100] M. Rongen, "Measuring the optical properties of IceCube drill holes", *EPJ Web Conf.*, vol. 116, p. 06 011, 2016. DOI: 10.1051/epjconf/201611606011.

[101] M. Jurkovic *et al.*, "A precision optical calibration module (POCAM) for IceCube-Gen2", *EPJ Web Conf.*, vol. 116, p. 06 001, 2016. DOI: 10.1051/epjconf/201611606001.

[102] M. G. Dunkman, "Measurement of atmospheric muon neutrino disappearance with IceCube-DeepCore", PhD thesis, Pennsylvania State University, 2015, ISBN: 978 136 - 963 602 4. [Online]. Available: https://etda.libraries.psu.edu/catalog/27226.

[103] M. G. Aartsen *et al.*, "Observation and characterization of a cosmic muon neutrino flux from the Northern Hemisphere using six years of IceCube data", *Astrophys. J.*, vol. 833, no. 1, p. 3, 2016. DOI: 10.3847/0004-637X/833/1/3.

[104] C. Andreopoulos *et al.*, "The GENIE neutrino Monte Carlo generator: Physics and user manual", 2015. arXiv: 1510.05494.

[105] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual", *J. High Energy. Phys*, vol. 2006, no. 05, p. 026, 2006. DOI: 10.1088/1126-6708/2006/05/026.

[106] A. Gazizov and M. Kowalski, "ANIS: high energy neutrino generator for neutrino telescopes", *Comput. Phys. Commun.*, vol. 172, no. 3, pp. 203 –213, 2005. DOI: 10.1016/j.cpc.2005.03.113.

[107] D. Chirkin, "Photon propagation with GPUs in IceCube", in *GPU Computing in High-Energy Physics (GPUHEP2014), Pisa (Italy)*, 2015, pp. 217–220. DOI: 10.3204/DESY-PROC-2014-05/40.

[108] D. Heck *et al.*, *CORSIKA: A Monte Carlo code to simulate extensive air showers*, FZKA 6019, 1998. [Online]. Available: https://www.ikp.kit.edu/corsika/70.php.

[109] The Telescope Array Collaboration, "CORSIKA simulation of the Telescope Array surface detector", 2014. arXiv: 1403.0644.

[110] J. Vicha and J. Chudoba, "Data processing at the Pierre Auger Observatory", *J. Phys. Conf. Ser.*, vol. 608, no. 1, p. 012 077, 2015. DOI: 10.1088/1742-6596/608/1/012077.

[111] J. R. Hörandel, "On the knee in the energy spectrum of cosmic rays", *Astropart. Phys.*, vol. 19, no. 2, pp. 193 –220, 2003. DOI: 10.1016/S0927-6505(02)00198-6.

[112] M. Larson, "A search for tau neutrino appearance with IceCube-DeepCore", PhD thesis, Niels Bohr Institute, University of Copenhagen, 2018.

[113] E. Thompson (formerly Cheung), "Measurements of atmospheric neutrino oscillation parameters using three years of IceCube data", Pending, PhD thesis, University of Maryland, 2018.

[114] S. Axani, *Private communication*, Massachusetts Institute of Technology, May 2017.

[115] M. Rongen, *Private communication*, III. Phys. Inst. B, RWTH Aachen University, May 2017.

[116] M. Aartsen *et al.*, "Testing the neutrino mass ordering with atmospheric neutrinos from three years of IceCube DeepCore data", *Eur. Phys. J. C*, 2018, publication pending.

[117] J. A. Aguilar *et al.*, "A fast algorithm for muon track reconstruction and its application to the ANTARES neutrino telescope", *Astropart. Phys.*, vol. 34, no. 9, pp. 652 –662, 2011. DOI: 10.1016/j.astropartphys.2011.01.003.

[118] J. P. Yáñez, "Measurement of neutrino oscillations in atmospheric neutrinos with the IceCube DeepCore detector", PhD thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät I, 2014. DOI: 10.18452/17016.

[119] R. Nahnhauer, A. Terliuk, and J. P. Yáñez, *Energy reconstruction of ν_μ events in DeepCore*, IceCube Collaboration - Internal Report: icecube/201304001-v2, 2013.

[120] G. Angelis *et al.*, "The performance of monotonic and new non-monotonic gradient ascent reconstruction algorithms for high-resolution neuroreceptor PET imaging", *Phys. Med. Biol.*, vol. 56, no. 13, p. 3895, 2011. DOI: 10.1088/0031-9155/56/13/010.

[121] F. Feroz, M. P. Hobson, and M. Bridges, "MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics", *Mon. Notices Royal Astron. Soc.*, vol. 398, no. 4, pp. 1601–1614, 2009. DOI: 10.1111/j.1365-2966.2009.14548.x.

[122] F. James and M. Roos, "MINUIT: a system for function minimization and analysis of the parameter errors and corrections", *Comput. Phys. Commun.*, vol. 10, no. CERN-DD-75-20, pp. 343–367, 1975. DOI: 10.1016/0010-4655(75)90039-9.

[123] J. A. Nelder and R. Mead, "A simplex method for function minimization", *Comput. J.*, vol. 7, no. 4, pp. 308–313, 1965. DOI: 10.1093/comjnl/7.4.308.

[124] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization", *Math. Program.*, vol. 45, no. 1-3, pp. 503–528, 1989. DOI: 10.1007/BF01589116.

[125] V. Blobel and E. Lohrmann, *Statistische und numerische Methoden der Datenanalyse*. Springer-Verlag, 2013, ISBN: 978 366 305 690 4.

[126] M. Leuermann, "Testing the neutrino mass ordering with multiple years of IceCube/DeepCore", in *Topics in Astroparticle and Underground Physics (TAUP), Sudbury, Canada*, publication pending, 2017.

[127] M. Leuermann, "Results from testing the neutrino mass ordering with three years of IceCube DeepCore data", in *Neutrino 2018, Heidelberg, Germany*, 2018. DOI: 10.5281/zenodo.1300550.

[128] S. Wren, "Neutrino mass ordering studies with IceCube-DeepCore", PhD thesis, School of Physics and Astronomy, University of Machester, Great Britain, 2018.

[129] Y.-C. Chen, "A tutorial on Kernel Density Estimation and recent advances", 2017. arXiv: 1704.03924.

[130] D. W. Scott, "On optimal and data-based histograms", *Biometrika*, vol. 66, no. 3, pp. 605–610, 1979. DOI: 10.2307/2335182.

[131] B. Wang and X. Wang, "Bandwidth Selection for Weighted Kernel Density Estimation", 2007. arXiv: 0709.1616.

[132] B. W. Silverman, *Density estimation for statistics and data analysis*. CRC press, 1986, vol. 26, ISBN: 978 041 224 620 3.

[133] S. Schoenen, "Discovery and characterization of a diffuse astrophysical muon neutrino flux with the IceCube Neutrino Observatory", PhD thesis, III. Phys. Inst. B, RWTH Aachen University, 2017.

[134] D. B. H. Cline and J. D. Hart, "Kernel estimation of densities with discontinuities or discontinuous derivatives", *Statistics*, vol. 22, no. 1, pp. 69–84, 1991. DOI: 10.1080/02331889108802286.

[135] E. P. Vogel, "Implementation and test of Kernel Density Estimation on Graphics Processing Units (GPUs) for the data analysis of the IceCube Neutrino Observatory", Bachelor's Thesis, III. Phys. Inst. B, RWTH Aachen University, 2014.

[136] B. Efron and R. Tibshirani, "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy", *Stat. Sci.*, vol. 1, pp. 54–75, 1986. DOI: doi:10.1214/ss/1177013815.

[137] M. Honda *et al.*, "Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data", *Phys. Rev. D*, vol. 75, p. 043 006, 2007. DOI: 10.1103/PhysRevD.75.043006.

[138] J. Evans *et al.*, "Uncertainties in atmospheric muon-neutrino fluxes arising from cosmic-ray primaries", *Phys. Rev. D*, vol. 95, no. 2, p. 023 012, 2017. DOI: 10.1103/PhysRevD.95.023012.

[139] K. S. Kuzmin, V. V. Lyubushkin, and V. A. Naumov, "Axial masses in quasielastic neutrino scattering and single-pion neutrinoproduction on nucleons and nuclei", in *20th Max Born Symposium, Wroclaw, Poland*, vol. 37, 2006, pp. 2337–2348. arXiv: hep-ph/0606184.

[140] T. Katori *et al.*, "First look at PYTHIA8 hadronization program for neutrino interaction generators", *JPS Conf. Proc.*, vol. 12, p. 010 033, 2016. DOI: 10.7566/JPSCP.12.010033.

[141] M. Tzanov *et al.*, "Precise measurement of neutrino and anti-neutrino differential cross sections", *Phys. Rev. D*, vol. 74, p. 012 008, 2006. DOI: 10.1103/PhysRevD.74.012008.

[142] The IceCube Collaboration - oscillation working group, *Internal material*, accessed at Mar. 15th, 2018. [Online]. Available: https://wiki.icecube.wisc.edu/.

[143] Y. Pawitan, *In all likelihood: statistical modelling and inference using likelihood*. Oxford University Press, 2001, ISBN: 978 019 967 122 9.

[144] N. Cressie and T. R. C. Read, "Multinomial goodness-of-fit tests", *J. Royal Stat. Soc. B*, vol. 46, no. 3, pp. 440–464, 1984.

[145] G. J. Feldman and R. D. Cousins, "Unified approach to the classical statistical analysis of small signals", *Phys. Rev. D*, vol. 57, no. 7, p. 3873, 1998. DOI: 10.1103/PhysRevD.57.3873.

[146] S. S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses", *Ann. Math. Statist.*, vol. 9, no. 1, pp. 60–62, 1938. DOI: 10.1214/aoms/1177732360.

[147] A. Wald, "Tests of statistical hypotheses concerning several parameters when the number of observations is large", *Trans. Amer. Math. Soc.*, vol. 54, no. 3, pp. 426–482, 1943. DOI: 10.2307/1990256.

[148] I. Asimov, *The Complete Stories*. Broadway Books, 1990, vol. 1, ISBN: 978 143 951 298 2.

[149] G. Cowan *et al.*, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C*, vol. 71, p. 1554, 2011, [Erratum: Eur. Phys. J.C73,2501(2013)]. DOI: 10.1140/epjc/s10052-011-1554-0,10.1140/epjc/s10052-013-2501-z.

[150] E. Ciuffoli, J. Evslin, and X. Zhang, "Confidence in a Neutrino Mass Hierarchy determination", *J. High Energy Phys.*, vol. 01, p. 095, 2014. DOI: 10.1007/JHEP01(2014)095.

[151] M. Beck, "Untersuchung zur Messung der Inelastizität bei Wechselwirkungen atmosphärischer Neutrinos in IceCube", Bachelor's Thesis, III. Phys. Inst. B, RWTH Aachen University, 2017.

[152] S. Philippen, "Anwendung von maschinellem Lernen für die Klassifikation niederenergetischer Ereignistopologien in IceCube/DeepCore", Bachelor's Thesis, III. Phys. Inst. B, RWTH Aachen University, 2017.

[153] X. Qian *et al.*, "The Gaussian CL_s-method for searches of new physics", *Nucl. Instrum. Methods Phys. Res. A*, vol. 827, pp. 63–78, 2016. DOI: 10.1016/j.nima.2016.04.089.

[154] M. G. Aartsen *et al.*, "Determining neutrino oscillation parameters from atmospheric muon neutrino disappearance with three years of IceCube DeepCore data", *Phys. Rev. D*, vol. 91, p. 072 004, 7 2015. DOI: 10.1103/PhysRevD.91.072004.

[155] N. van Eijndhoven, O. Fadiran, and G. Japaridze, "Implementation of a Gauss convoluted Pandel PDF for track reconstruction in neutrino telescopes", *Astropart. Phys.*, vol. 28, no. 4, pp. 456 –462, 2007. DOI: 10.1016/j.astropartphys.2007.09.001.

[156] NVIDIA Corporation, *CUDA programming guide*, accessed at Feb. 17th, 2018, 2010. [Online]. Available: https://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf.

[157] F. J. Massey Jr, "The Kolmogorov-Smirnov test for goodness of fit", *J. Am. Stat. Assoc.*, vol. 46, no. 253, pp. 68–78, 1951. DOI: 10.2307/2280095.

List of Figures

List of Tables