

Testing nonstandard neutrino interaction parameters
with IceCube-DeepCore

Elisa Lohfink

Johannes Gutenberg-Universität



Fachbereich 08

Testing nonstandard neutrino interaction parameters with IceCube-DeepCore

Dissertation zum Erlangen des Grades
„Doktor der Naturwissenschaften“

Fachbereich für Physik, Mathematik und Informatik
der Johannes Gutenberg-Universität Mainz

von
Elisa Lohfink
geboren am 31.01.1992 in Mainz

Mainz, den 27. Januar 2023

Elisa Lohfink

Testing nonstandard neutrino interaction parameters with IceCube-DeepCore

Dissertation zum Erlangen des Grades

„Doktor der Naturwissenschaften“

Fachbereich für Physik, Mathematik und Informatik

der Johannes Gutenberg-Universität Mainz

von

Elisa Lohfink

geboren am 31.01.1992 in Mainz

Mainz, den 27. Januar 2023

Johannes Gutenberg-Universität

Fachbereich 08

Staudingerweg 9

55128 Mainz

I hereby declare that I have written the present thesis independently and without use of other than the indicated means. I also declare that to the best of my knowledge all passages taken from published and unpublished sources have been referenced.

Mainz, _____
Date Signature

Elisa Lohfink
ETAP
Institut für Physik
Staudingerweg 7
Johannes Gutenberg-Universität D-55099 Mainz
ellohfin@uni-mainz.de

Abstract

Currently, the clearest evidence for physics beyond the Standard Model is provided by observations that indicate non-zero neutrino masses. Numerous theories on how these masses are generated give rise to additional non-standard interactions (NSI) of neutrinos with quarks and charged leptons. Atmospheric neutrinos provide a sensitive probe for the neutrino flavor transitions resulting from the type of NSI investigated in this work. These comprise neutral current forward scattering of neutrinos of all flavors on first generation charged fermions in Earth matter. In order to maximize model independence, the NSI are parametrized using five effective coupling strengths.

In the IceCube-DeepCore detector, atmospheric neutrinos are detected indirectly via Cherenkov photons produced within the Antarctic glacier. The range of angles under which neutrinos enter the detector translates into propagation baselines of $\mathcal{O}(1 - 10^4)$ km. The data sample used in the presented analysis includes 9.3 years of DeepCore data, covering a neutrino energy range of 5 to 100 GeV.

Accuracy and performance of event property reconstruction from observed photons are a crucial factor for the analysis outcome. Therefore, this work includes a thorough study of the potential and shortcomings of likelihood based DeepCore event reconstruction algorithms.

The presented analysis relies on comparing binned observed event counts to simulation that is generated at different hypotheses. These include the individually considered NSI parameters as well as 17 nuisance parameters. In order to find the hypothesis that best describes the observation, the optimum of a test statistic is determined using a customized minimization strategy. The final analysis setup yields sensitivities to four effective NSI couplings that are competitive compared with existing results.

Acknowledgements

Contributions

This section serves to clarify the individual contributions by me and others to different parts of this work. Being part of the IceCube collaboration and working with the most recent DeepCore data sample allowed me to build upon a group effort, the aim of which is to achieve a whole new generation of analyses related to PMNS parameters and BSM physics to be performed on an updated set of DeepCore data. Along with the preceding DeepCore NSI analyses, this creates the basis for my work.

Development of the reconstructions discussed and applied for this work partly dates back many years, which is why only the most recent contributors are listed in the following. For the SANTA and LEERA reconstruction, these are [REDACTED] and [REDACTED]. The pegleg reconstruction was last extended by [REDACTED] and [REDACTED]. The RETRO reconstruction is a relatively recent development which was mainly conducted by [REDACTED] and [REDACTED].

The presented reconstruction comparison and the means to evaluating potential improvements were developed as part of this work, with [REDACTED] contributing significantly to the latter.

Building upon previous generations of samples, development of the general event selection was mainly implemented by [REDACTED], [REDACTED] and [REDACTED]. These people were also involved in the corresponding simulation, which was furthermore developed and operated by [REDACTED] and [REDACTED]. The FLERCNN reconstruction and related event selection steps were developed by [REDACTED] and [REDACTED]. The optimal data binning was evaluated for the purpose of this analysis.

The PISA software framework that is used for this analysis is being developed in a group-wide effort of the IceCube oscillations working group. I contributed to this in various aspects, for example by porting the NSI implementation originally realized by [REDACTED] to the current PISA version. His analysis of the preceding DeepCore sample also provided the fundamental analysis procedure and choice of NSI parametrization used in this work.

Code directly concerning this work is self-written or based on and adapted from existing scripts available through PISA or from present and past members of the

IceCube oscillations working group. The minimization approach that was developed as a part of this work includes and adapts techniques that were found to work in analyses by [REDACTED] and [REDACTED].

The utilized treatment of detector nuisance parameters was developed by [REDACTED] and [REDACTED] and was tested and adapted for the analysis and data sample used in this work.

The approach to computationally feasible Feldman and Cousins based confidence limits that is presented was developed in cooperation with and conceptually constructed by [REDACTED].

Applying Markov-Chain Monte Carlo techniques for DeepCore analyses was previously tested inconclusively by [REDACTED] and [REDACTED]. In this work, thorough optimization of the set-up of this approach for the first time resulted in its successful application to a DeepCore data sample.

Contents

1	Introduction	1
2	Neutrinos within the Standard Model	3
2.1	The Standard Model of Particle Physics	3
2.2	Neutrino interactions	6
3	Neutrinos beyond the Standard Model	11
3.1	Neutrino masses and mixing	11
3.1.1	Experimental evidence for neutrino mixing and oscillations	11
3.1.2	Neutrino mixing	12
3.1.3	Neutrino masses and mass ordering	13
3.1.4	Massive neutrino extensions to the standard model of particle physics (SM)	14
3.2	Neutrino oscillations and matter effects	15
3.2.1	Neutrino oscillations in vacuum	15
3.2.2	Matter effects	17
3.2.3	Earth effects on atmospheric neutrinos	19
3.2.4	Open questions on neutrino physics	20
3.3	Formalism of non-standard neutrino interactions	21
3.4	The impact of NSI	24
3.4.1	nonstandard interactions (NSI) phenomenology: Global behavior and remarks	24
3.4.2	Phenomenology of $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$	27
3.4.3	Phenomenology of $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$	29
3.4.4	Phenomenology of $\epsilon_{\mu\tau}^{\oplus}$	30
3.4.5	Phenomenology of $\epsilon_{e\mu}^{\oplus}$ and $\epsilon_{e\tau}^{\oplus}$	31
3.5	NSI measurements in IceCube and beyond	32
4	Neutrino detection	35
4.1	Neutrino sources	35
4.2	Atmospheric neutrino production	38
4.3	Detection principles	42

4.4	Neutrino detection in IceCube and DeepCore	45
4.4.1	The Digital Optical Module	46
4.4.2	The detector geometry	47
4.4.3	Ice as a detection medium for Cherenkov light	49
4.4.4	Event characteristics	52
4.4.5	Background and noise	54
4.4.6	Data acquisition and on-site signal processing	55
5	Simulation and data sample	59
5.1	Simulation	59
5.2	Event selection and processing	62
5.2.1	Run selection	62
5.2.2	Levels of event selection	63
5.2.3	Final level cuts	68
5.2.4	Composition of final sample	70
6	Event reconstruction	73
6.1	Reconstruction challenges in DeepCore	73
6.2	Likelihood based reconstruction	75
6.3	The pegleg reconstruction	76
6.3.1	Event representation	76
6.3.2	Tabulated hit probabilities	78
6.3.3	Hit representation	79
6.3.4	Likelihood formulation and optimization	80
6.4	Existing alternatives	82
6.4.1	single string ANTARES-inspired analysis (SANTA) and Low-Energy Energy Reconstruction Algorithm (LEERA)	83
6.4.2	Reverse Table Reconstruction (RETRO)	83
6.4.3	Fast Low Energy Reconstruction using Convolutional Neural Networks (FLERCNN)	84
6.5	Performance of pegleg with respect to other reconstructions	85
6.5.1	Run-times	86
6.5.2	Energy resolution	86
6.5.3	Angular resolution	87
6.5.4	Comparing the FLERCNN reconstruction	88
6.6	Potential and limits	90
6.6.1	Likelihood modality and reconstruction performance	90
6.6.2	Potential modeling deficits and their impact	95
6.6.3	Detector geometry induced bias	102

6.6.4	Minimization effects	103
7	Search for NSI	111
7.1	Neutrino signal in DeepCore	112
7.1.1	Standard oscillations footprint	112
7.1.2	NSI signal expectation	113
7.1.3	Binning in reconstructed observables	113
7.2	Statistical approach	116
7.2.1	Test statistic	118
7.2.2	Staged MC re-weighting in PISA	119
7.3	Uncertainties in describing particles	120
7.3.1	Flux modeling	121
7.3.2	Oscillation and matter effect implementation	123
7.3.3	Cross section uncertainties	124
7.3.4	Overall signal normalization	125
7.4	Parametrizing the detector	125
7.4.1	Module and hole ice characteristics	126
7.4.2	Bulk ice	127
7.4.3	Nuisance parameter interpolation	128
7.4.4	Evaluating the detector uncertainty treatment	131
7.5	Impact of individual nuisance parameters	136
7.6	Finding the hypothesis that fits the observation best	139
7.6.1	Navigating the parameter space	139
7.6.2	Recovering injected NSI hypotheses	142
7.6.3	Recovering fluctuated nuisance parameters	144
7.6.4	Handling pseudo-data	147
7.7	Sampling	149
7.7.1	The mechanics of MCMC	149
7.7.2	Finding the best settings	151
7.7.3	Challenging parameter space features	153
7.7.4	Performance	154
7.8	Expected sensitivities for the frequentist approach	157
7.9	Setting limits	159
7.9.1	Computing Feldman Cousins limits	161
7.9.2	Making Feldman Cousins limits computationally feasible	162
8	Conclusion	167
	Bibliography	169

A	Supplementary information	193
A.1	Massive neutrino extensions to the SM	193
A.2	Perturbation theory approach to the impact of individual NSI parameters	195
A.3	Coherent Elastic Neutrino-Nucleus Scattering ($\text{CE}\nu\text{NS}$)	195
A.4	Reconstruction within the IceCube Upgrade	195
A.5	Offset in nuisance parameter gradients	196
B	Supplementary figures and tables	199
B.1	NSI phenomenology	199
B.2	Reconstruction studies	204
B.3	NSI signal in DeepCore	206
B.4	Simulated sets with off-nominal nuisance parameter values	207
B.5	Detector nuisance parameter treatment	210
B.6	Nuisance parameter impact	214
B.7	Inject recovery test results	216
B.8	Fluctuating nuisance parameter values	217
B.9	Pseudo-data ensemble	220
B.10	Sensitivities	221
B.11	Additional figures concerning MCMC sampling	223
B.12	Limits on pseudo-data	224

Introduction

The way that neutrinos are included in the Standard Model of particle physics is not compatible with their measured properties. Specifically, observations of oscillation between neutrino flavor eigenstates indicate that, other than the Standard Model description, neutrinos possess non-zero masses [1]. This observation makes it opportune to search for NSI, which are neutrino interactions that are not included in the Standard Model and potentially arise from mechanisms that give neutrinos mass [2].

The variety of models that potentially give rise to NSI motivates this search to be preferably model-independent, which is achieved by using a parametrization of NSI through five effective coupling strengths [3]. These represent an effective matter potential that originates from interaction of neutrinos of all flavors with all first generation charged fermions.

Atmospheric neutrinos constitute a suitable probe for NSI, due to their sensitivity to neutrino flavor transition probabilities. These depend on the neutrino energy, flavor and the matter density profile they propagate through. The DeepCore low energy extension of the IceCube detector, a Cherenkov neutrino detector built into the Antarctic ice sheet, provides atmospheric neutrino events at an energy range of 5 GeV to 100 GeV. The data include neutrino trajectories from different directions, which implies oscillation baselines from a few kilometers up to the Earth diameter. The distinct signature that neutrino oscillations and NSI effects leave in DeepCore data allow for measuring even small deviations from the no-NSI null hypothesis. This analysis relies on the most recent DeepCore data set, which contains approximately 150000 preselected events that have been recorded within 9.3 years of detector operation. As a first step towards the presented analysis, the performance and potential of likelihood-based reconstructions of DeepCore events is explored.

The presented search for NSI relies on comparing binned data to simulated event rate maps for different physics hypotheses, which is done in a frequentist fitting approach. Bayesian sampling is explored as a supplementary or alternative technique. The key elements to the analysis are the minimization performed when fitting as well as the treatment of nuisance parameters. The latter describe neutrino flux and interaction properties as well as detector characteristics. A test statistic is minimized

with respect to nuisance and NSI parameters in order to find the hypothesis that best describes the observed data. In addition to this best fit value, limits are set for each individually investigated NSI parameter.

Neutrinos within the Standard Model

The postulation of neutrinos in 1930 by W. Pauli [4] was originally motivated by observations of the energy spectrum of electrons produced in β^- -decay. Instead of the discrete energy expected for two-body decays these showed a continuous energy distribution. This pointed towards an undetected third component, a light and neutral new particle. Although originally deemed undetectable by Pauli, the first detection of neutrinos succeeded in 1956 [5]. It was achieved in the Cowan–Reines neutrino experiment, a measurement of reactor neutrinos undergoing inverse beta decay within a water target [6].

The factor that impeded their detection for so long was not their scarceness - neutrinos originate from various sources and occur on earth in abundance, as will be discussed in section 4.1. It was rather how rarely they interact in a detectable manner, which has its origin in the fundamental neutrino properties. Their mathematical formulation, as included in the SM, is summarized in the following. This is followed by a discussion of the resulting detection-relevant neutrino interactions in section 2.2.

2.1 The Standard Model of Particle Physics

Mathematically, the SM is a quantum field theory, hence describing particles as excitations of their respective fields and particle interactions via the field's Lagrangian density [1]. It is regarded as an effective field theory, as there are known effects it fails to describe. While managing to include strong, electromagnetic and weak interactions between elementary particles, it prominently does not describe gravity. All elementary particles within the SM are either bosons, carrying integer spins, or fermions with half-numbered spins to which Fermi-Dirac statistics apply. The elementary fermions known today are shown in figure 2.1, along with their mass¹ and electric charge. Their anti-particles are not shown separately as they differ only in charge. The possible fundamental interaction vertices are shown in figures 2.2

¹Throughout this thesis, natural units will be used, e.g. when indicating particle masses.

and 2.3.

As a gauge theory, the SM includes invariance under certain symmetry transformations, leading to conservation of corresponding quantities, according to Noether's first theorem [7]. It can in total be described via the unitary product group $SU(3) \times SU(2) \times U(1)$ [1], which is subdivided into two parts: One is treating the strong nuclear force through the $SU(3)$ component, formulated as quantum chromodynamics (QCD), where the so-called color charge is conserved.

u 2.16 MeV 2/3	c 1.27 GeV 2/3	t 172.76 GeV 2/3	e 0.511 MeV -1	μ 105.658 MeV -1	τ 1776.86 MeV -1
d 4.67 MeV -1/3	s 93 MeV -1/3	b 4.18 GeV -1/3	ν_e 0 MeV 0	ν_μ 0 MeV 0	ν_τ 0 MeV 0

Fig. 2.1.: Elementary fermions as included in the SM, together with their respective mass (top right) and electric charge (bottom right). Three generations of quarks and leptons are shown in the left and right half, respectively. All values are taken from [8]. Note that the masses of the e and τ leptons are given with less digits than available therein.

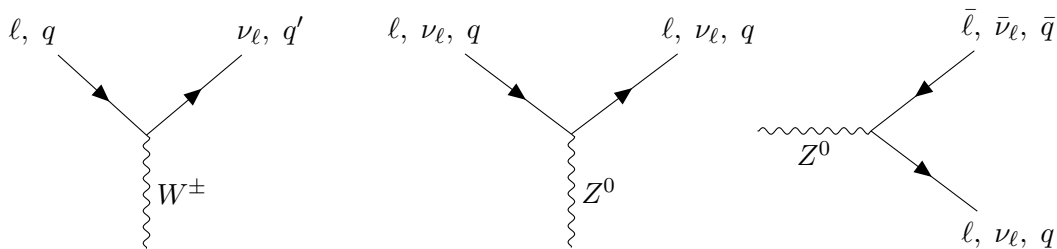


Fig. 2.2.: Fundamental weak vertices including a single boson: charged current (CC) (left) and neutral current (NC) (center) scattering and pair production (right). The time axis is set from left to right. Charged and neutral leptons are indicated as l and ν_l , respectively, and quarks as q , with q' indicating a different quark with respect to q . The same processes with reversed time axis or with particles interchanged with their antiparticles are allowed as well. Quantum number conservation laws are implicitly assumed and not explicitly visualized.

As the second part, electromagnetic (EM) and weak interactions are unified to the electroweak force, described in a $SU(2) \times U(1)$ group. This conserves electric charge, weak isospin and weak hyper-charge. The invariances of the electroweak force yield four gauge bosons, mediating the weak and electromagnetic force: The electromagnetically interacting massless photon γ , two W^\pm bosons that couple to weak and electromagnetic CC interaction and the purely weakly interacting, NC Z^0 boson² [1].

²This includes electroweak symmetry breaking and the Higgs mechanism [9].

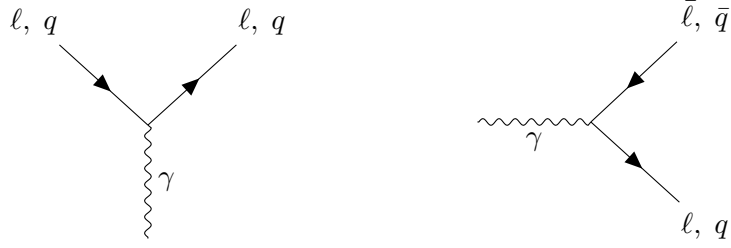


Fig. 2.3.: Fundamental EM vertices including a single boson: Scattering (left) and pair production (right). Annotations from figure 2.2 apply.

Within the SM, fermions are subdivided into two categories with three so-called generations each. The mathematical structure of the SM does not require exactly three generations, making their number purely measurement-based [10]. The first category of SM fermions contains six quarks. These are massive particles with electromagnetic charge of $2/3$ or $-1/3$ which interact via all forces. The second category consists of six leptons, of which three (electron, muon and tau) take part in EM, weak and gravitational interaction. Each of these shares its generation with a corresponding neutrino, which is a neutral massless lepton that only interacts weakly. Consequently, the Lagrangian densities that describe NC and CC neutrino interactions are [8]

$$\mathcal{L}_{CC} = \frac{-g_W}{\sqrt{2}} \sum_{\ell} \bar{\nu}_{\ell L} \gamma^{\mu} \ell_L^{-} W_{\mu}^{+} + h.c. \quad (2.1)$$

$$\mathcal{L}_{NC} = \frac{-g_W}{\sqrt{2} \cos(\theta_W)} \sum_{\ell} \bar{\nu}_{\ell L} \gamma^{\mu} \nu_{\ell L} Z_{\mu}^0, \quad (2.2)$$

where g_W is the weak coupling constant. The weak mixing angle or Weinberg angle, θ_W , describes how the boson fields after electroweak symmetry breaking are a linear combination of those before. The gamma matrices γ_{μ} correspond to fermions satisfying the Dirac equation [1]. In equation 2.1, h.c. indicates the Hermitian conjugate.

Note that for charged leptons and the respective neutrinos, indicated as l^{\pm} and ν_l , only components with negative (left) chirality are included. For massless particles, chirality is equal to helicity, which is positive if a particle's momentum and spin are aligned and negative otherwise. Antiparticles with negative chirality and particles with positive chirality do not interact. For non-zero particle masses, it is possible to apply a Lorentz-boost into a reference frame with higher velocity than the particle, which causes it to switch chirality. Therefore, for massive particles, chirality rather suppresses processes than forbidding them.

2.2 Neutrino interactions

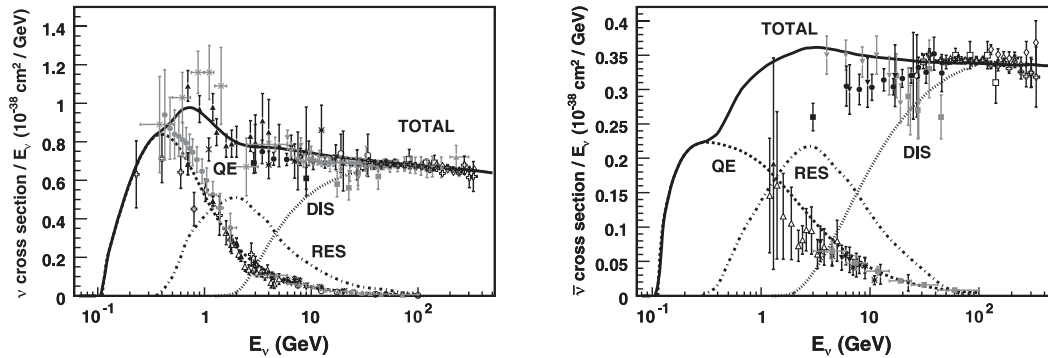


Fig. 2.4.: Cross sections of neutrino (left) and antineutrino (right) CC interaction per nucleon over energy. The total cross section (solid line) is shown together with contributing processes and the respective measurements. The three processes shown here are quasi-elastic scattering (QES) (dashed line, also with measurements), resonant scattering (RES) (dotted-dashed line) and deep inelastic scattering (DIS) (dotted line), which become dominant at different energies. Adapted from [11].

Today's SM knows neutrinos of three flavors, ν_e , ν_μ and ν_τ , which interact weakly as shown in figure 2.2. Their relevance is reflected in the Nobel Prizes awarded for the detection of electron neutrinos [6], muon neutrinos [12] and cosmic neutrinos [13]. The limitation to weak interaction results in overall small cross sections of neutrino interactions, as the Fermi coupling constant (G_F) of the weak force is more than three orders of magnitude lower than that of the EM interaction and enters cross sections quadratically [11] (see equation 2.2).

This section focuses on the neutrino interactions relevant for their detection in DeepCore, which are between nucleons and neutrinos with energies at the GeV scale. The higher the energy, the more details of the nucleus are resolved. Interactions therefore range from scattering on the entire nucleon below 1 GeV, to exciting the nucleon in the few GeV range, and even scattering on quarks at energies above that. Cross sections of these three processes are given in figure 2.4.

As expected from the quark-parton model [14], the total cross section shows an approximately linear energy dependence above ~ 10 GeV. The difference between neutrinos and antineutrinos stems from the relative amount of quarks and anti-quarks in matter, together with their chirality. cross sections also differ depending on the neutrino flavor, mostly due to the tau mass being large with respect to other leptons, which impedes tau neutrino interactions at low energies, as shown in figure 2.5.

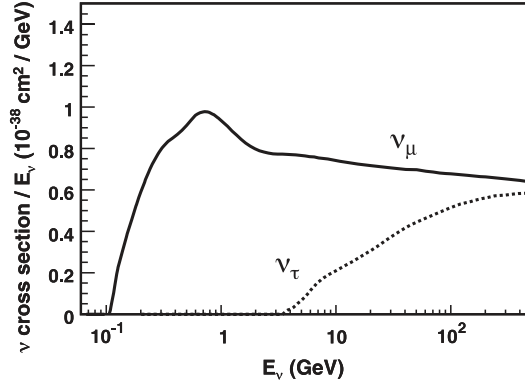
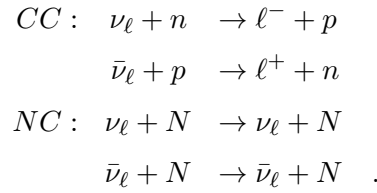


Fig. 2.5.: Total cross section of muon neutrino (solid) and tau neutrino (dashed) CC interactions, divided by neutrino energy. Taken from [11].

Quasi-elastic scattering (QES)

Of the three interactions mentioned above, elastic and quasi-elastic scattering are the NC and CC interactions relevant at the lowest energies. They constitute neutrinos scattering off an entire nucleon, undergoing



Here, n describes a neutron, p a proton and N any nucleon. This process is dominant below approximately 1 GeV, as shown in figure 2.4. The differential cross section of CC QES,

$$\frac{d\sigma_{QES}}{dQ^2} = \frac{G_F^2 M^2}{8\pi E_\nu^2} \left[A \pm \frac{4ME_\nu - Q^2 - m^2}{M^2} B + \frac{(4ME_\nu - Q^2 - m^2)^2}{M^4} C \right] \quad , \quad (2.3)$$

depends on the neutrino energy E_ν , Fermi coupling constant G_F , masses m and M of the lighter and heavier scattering partners, as well as the squared transferred four-momentum Q^2 [11]. Neutrinos and antineutrinos differ in the sign of the second summand, which is indicated as \pm . Nucleon form factors are present in the cross section through A , B and C . They depend on the effective axial mass M_A^{QE} , an empirical parameter that enters this analysis as a nuisance parameter, as discussed in detail in section 7.3.3.

Resonant scattering (RES)

Resonant pion production occurs dominantly at energies of a few GeV, producing an excited baryonic state N^* from the partaking nucleon. Otherwise the interaction mechanisms are similar to QES, as can be seen in the example

$$CC : \nu_\ell + N \rightarrow \ell^- + N^* \\ \rightarrow \ell^- + \pi^{\pm/0} + N' ,$$

where the baryonic state produces a pion. This can, in case of a π^0 , further decay into leptons, or, for π^\pm , become the origin of a hadronic cascade.

The dependencies of the RES cross section resemble the QES ones, with an effective axial mass M_A^{RES} .

Deep inelastic scattering (DIS)

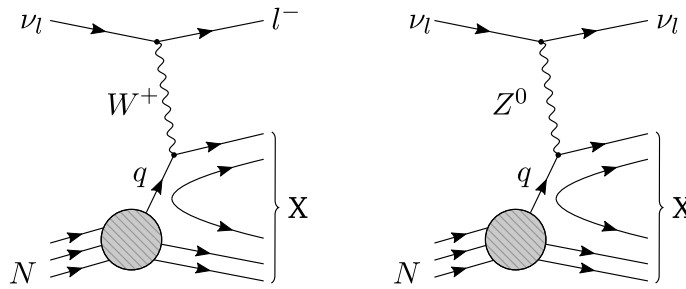


Fig. 2.6.: Diagram of DIS via CC (left) and NC (right) interaction, with N indicating a nucleon and X indicating an outgoing hadronic shower. The corresponding antineutrino processes yield respective anti-leptons. Based on [15].

The energy range above approximately 5 GeV is dominated by DIS, a process where sufficient momentum is transferred to the nucleus to be able to break it apart, resulting in a hadronic cascade X , as in the example

$$CC : \nu_\ell + N \rightarrow \ell^- + X , \quad (2.4)$$

which is shown as a diagram in figure 2.6.

Due to angular momentum conservation between initial and final state, together with the exclusion of positive-chiral neutrinos and negative-chiral anti-neutrinos, the cross section of antineutrino-quark interactions depends on the scattering angle. An additional factor is the difference in quark- versus antiquark-content in matter. The DIS cross section also depends on the inelasticity or Bjorken- y of the process,

which is the fraction of neutrino energy that goes into the hadronic cascade. These factors are relevant for the nuisance parameter treatment in section 7.3.3.

Hadronic and electromagnetic cascades

Hadronic or EM cascades have been mentioned above to be produced in higher energy interactions. They emerge from a single interaction which causes an avalanche-like particle production chain.

In EM cascades, photons decay into pairs of positrons and electrons. These individually produce photons in bremsstrahlung, which subsequently decay into electron-positron pairs, repeating the process. The cross section of this chain of processes is energy-dependent, with pair production ceasing below 1 MeV, giving showers a finite length based on the energy of the initial particle [15].

Hadronic cascades comprise a larger number of different processes, in which e.g. secondary hadrons such as pions are produced. As some of these, such as the decay of a π^0 , involve electromagnetic particles, a hadronic cascade always has an EM component.

Neutrinos beyond the Standard Model

Neutrinos are the particles providing the clearest evidence for physics beyond the standard model (BSM-physics), as will be discussed in the following. Most prominently, the observations connected to neutrino mixing and neutrino oscillations, which are introduced in sections 3.1 and 3.2, indicate that neutrinos have mass¹, allowing them to transition between their eigenstates. Different possible formalisms of how this mass is obtained entail additional BSM-physics effects, such as NSI. A model-independent NSI formalism is introduced in section 3.3 and the phenomenology of the parameters it comprises are discussed in section 3.4, followed by an overview of recent measurements in section 3.5.

3.1 Neutrino masses and mixing

The observation of neutrino oscillations and matter effects indicate non-zero differences between the neutrino mass eigenstates and hence massive neutrinos. The evidence and consequences of the existence of massive neutrinos as well as how they can be introduced to the SM will be discussed in the following.

3.1.1 Experimental evidence for neutrino mixing and oscillations

An early hint towards BSM-physics was the so-called solar neutrino deficit: In the 1990's, the solar electron neutrino flux observed by the Homestake experiment was found to be significantly below the solar standard model (SSM) prediction [18]. To disagree with the SSM implied either to contradict the understanding of the nuclear processes that occur inside the Sun and the resulting luminosity [19] or that electron neutrinos vanish in some way. Measurements at the Sudbury Neutrino Observatory (SNO) resolve both the electron neutrino and overall neutrino flux and supported

¹Neutrino oscillations are theoretically possible for massless neutrinos, for instance through NSI [16]. Measurements available today however exclude the respective models as the mechanism behind oscillations [17].

the hypothesis of flavor conversion rather than vanishing of electron neutrinos, as no deficit in the total flux was found [20].

These flux observations can be explained through adiabatic flavor conversion inside solar matter, which will be discussed in section 3.2.2, and, secondarily, neutrino oscillations, which are presented in section 3.2. Both of these effects are connected to mixing of neutrino states, a concept that was already postulated around 1960 by Pontecorvo [21, 22] as well as Maki, Nakagawa and Sakata [23]. Mixing requires non-zero neutrino masses, as will be detailed in the following, and therefore comprises BSM-physics.

3.1.2 Neutrino mixing

Neutrino properties are associated to two different bases. One is the flavor base of ν_e , ν_μ and ν_τ , summarized as ν_α , which was introduced in section 2. These are the neutrino states that take part in interactions. Secondly, the base of mass eigenvalues, ν_1 , ν_2 and ν_3 , denoted ν_i . The formalism for weak interaction allows for states, or rather the respective fields, to take part in interactions while being linear combinations of other fields that themselves do not take part. The neutrino flavor and mass base are connected as such a linear combination via

$$|\nu_\alpha\rangle = \sum_i U_{\alpha i} |\nu_i\rangle \quad . \quad (3.1)$$

The matrix connecting them is the Pontecorvo Maki Nakagawa Sakata mixing matrix (PMNS-matrix) U [21–23], which enters equation 3.1 complex conjugated for antineutrinos. It can be parameterized through three mixing angles θ_{ij} and one complex phase δ_{CP} as

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{CP}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.2)$$

$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{CP}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{CP}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{CP}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{CP}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{CP}} & c_{23}c_{13} \end{pmatrix} \quad . \quad (3.3)$$

This notation uses the abbreviations $c_{ij} = \cos(\theta_{ij})$ and $s_{ij} = \sin(\theta_{ij})$. Non-zero values of the δ_{CP} phase imply violation of charge and parity conjugation (CP) symmetry².

²A CP transformation corresponds to replacing a particle with its antiparticle and inverting its spatial coordinates. CP violation was observed in kaons [8] and is possible for neutrinos.

The mixing of neutrinos between their flavor and mass states links different observable behaviors, such as neutrino oscillations and matter effects, to neutrinos being massive.

3.1.3 Neutrino masses and mass ordering

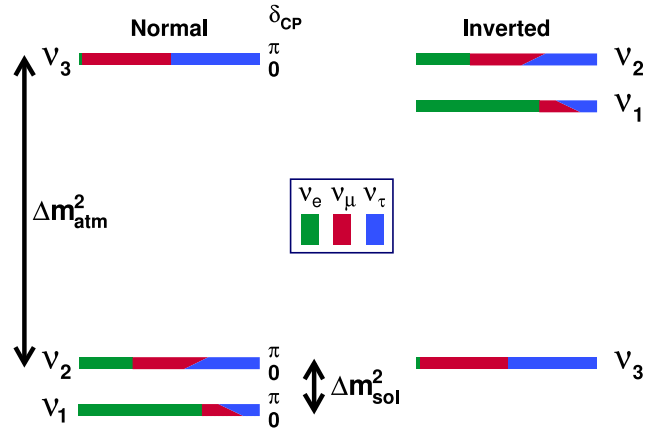


Fig. 3.1.: Schematic overview of the mass-squared differences in normal (left) and inverted (right) ordering, adapted from [24]. Heavier states are further at the top of the drawing. For normal ordering, the (larger) atmospheric mass splitting and (smaller) solar mass splitting are indicated. The neutrino flavor state content of the individual mass eigenstates is indicated in color. The dependence on δ_{CP} is represented for values of 0 and π at the lower and upper border, respectively, of the individual bars, yielding tilted edges.

For the absolute masses of neutrinos, only upper limits have been set so far. The currently lowest one amounts to $m_\nu < 0.9 \text{ eV}$ at 90% confidence limit (CL) and was set in Tritium decay measurements at the Karlsruhe Tritium Neutrino (KATRIN) experiment [25]. The necessity of non-zero mass-squared differences, however, is based on the observations around neutrino oscillations [17, 26], which are discussed in section 3.2.

The mass-squared differences Δm_{31}^2 and Δm_{21}^2 ³ are known to be at different scales, as is apparent from table 7.3. The smaller mass-squared difference, Δm_{21}^2 , is also known as the solar mass splitting, based on the origin of neutrino data that is most sensitive to this parameter [27].

Other than the sign of Δm_{21}^2 , which is set by convention to be positive [27], the sign of Δm_{31}^2 is unknown. As a result, there are two options on which neutrino mass eigenstate is the heaviest, as shown in figure 3.1. These two so-called neutrino mass orderings (NMOs) are normal ordering (NO), which describes the case where ν_3

³There are different conventions on which mass-squared differences to discuss. This work uses the one that is most common in IceCube analyses, by referring to Δm_{31}^2 and Δm_{21}^2 .

is the heaviest mass eigenstate, and inverted ordering (IO) if it is the lightest. The respective larger mass-squared difference is called the atmospheric mass splitting. Neutrino oscillations are not sensitive to the sign of the squared-mass differences. However, the NMO impacts them, as it changes the absolute squared-mass differences between the mass eigenstates ν_3 and $\nu_{1,2}$.

3.1.4 Massive neutrino extensions to the SM

While neutrinos are known to have non-zero mass differences, the origin and theoretical basis of their mass is still unknown. It might be a consequence of BSM-physics at some higher energy scale Λ that manifests at low energies. In order to include this, the SM can be seen as an effective model and extended in ways that are not disruptive with respect to SM effects [1].

The possibly most intuitive extension to the SM would be if neutrinos gained mass through the same mechanism as other elementary particles. The corresponding additional Dirac mass term is therefore seen as the minimal extension of the SM [28]. An alternative would be a Majorana mass term, in which neutrinos are their own anti-particles. Such a term would for example result from a generic extension in which the SM is considered an effective field theory with its components, the fields, acting as operators. Including BSM-physics implies adding operators of increasing exponent or dimension $n \geq 5$ which are suppressed by Λ^{n-4} . More detail on this can be found in [29]. At dimension six, this extension results in NSI [28, 30].

Unitarity of the neutrino mixing matrix U is given within the SM but not necessarily in its extensions. There, unitarity depends on the mechanism through which mass is provided. While being unitary for Dirac masses, U becomes non-unitary for Majorana masses. Also, extending the SM with a Majorana mass term would introduce two additional phases in U . However, these are not relevant to neutrino oscillations and therefore left out in this description [8].

A more general extension of the SM includes a Majorana and Dirac mass term. In such a case, the ratio between Majorana and Dirac mass eigenvalues provides the approximate neutrino mass scales through the so-called see-saw mechanism [27]. This yields an explanation for neutrino masses being at a significantly different scale, multiple orders of magnitude below that of charged fermions of the same particle generation [8].

The type-II see-saw mechanism similarly applies to Majorana masses above the vacuum expectation value of the Higgs field [1, 28]. This, like multiple other dimension-six SM extensions [29, 31], directly leads to the existence of NSI.

After having established that neutrinos have mass and how this might be included in the SM, the oscillation effects that mixing of massive neutrinos entails are described in the following. More details on neutrino mass SM extensions can be found in Appendix A.1.

3.2 Neutrino oscillations and matter effects

Neutrino oscillations are interference effects between neutrino mass eigenstates, leading to flavor transitions. They occur when neutrinos propagate in their mass eigenstates and the relative phases of the individual states change. The resulting flavor transitions require non-zero differences between neutrino masses, hence constituting BSM-physics. The first detection of neutrino oscillations was accomplished in the Super-Kamiokande experiment (Super-K)⁴ [26] and awarded the 2015 Nobel Prize [32].

The theoretical formulation of neutrino oscillations and flavor conversion effects in vacuum and in matter will be discussed in the following. A focus is set on oscillations and matter effects on atmospheric neutrinos in Earth.

3.2.1 Neutrino oscillations in vacuum

Ultrarelativistic neutrinos in vacuum can be described as plane waves, propagating in their mass eigenstates. Their relative phases change in propagation, which leads to interference effects and ultimately oscillations of the overall composition.

As neutrinos interact in their flavor states, this behavior can only be observed indirectly. The neutrino composition at production is given through equation 3.1 while the composition in which a neutrino interacts after propagating changes with time, or corresponding distance. In this section, a short introduction to neutrino oscillations in vacuum will be given, with non-vacuum effects covered in the following section. Assumptions and approximations that are made here for simplicity are legitimate for the neutrinos investigated in this thesis [8]. A more detailed derivation can be found in [1].

The time dependent behavior of a neutrino flavor eigenstate,

$$|\nu_\alpha(t)\rangle = \sum_i U_{\alpha i}^* |\nu_i(t)\rangle, \quad (3.4)$$

⁴See section 4.3 for a short discussion of the detection method of Super-K and others.

is given by the mass eigenstate with neutrino energy E_i , propagating as a plane wave

$$|\nu_i(t)\rangle = e^{-iE_i t} |\nu_i(t=0)\rangle . \quad (3.5)$$

In this notation, the time t is chosen as the free variable, implicitly including the dependency on travel distance L . The plane wave function 3.5 constitutes a solution to the Schrödinger equation

$$E_i |\nu_i(t)\rangle = i \frac{d}{dt} |\nu_i(t)\rangle \quad (3.6)$$

$$= \mathcal{H}_{vac} |\nu_i(t)\rangle \quad (3.7)$$

$$= \frac{1}{2E} U \text{diag} (m_1^2, m_2^2, m_3^2) U^\dagger |\nu_i(t)\rangle , \quad (3.8)$$

in natural units, with the vacuum Hamiltonian \mathcal{H}_{vac} and the masses m_i of the mass eigenstate ν_i [33]. The energy E_i depends on the momentum p_i through

$$E_i = \sqrt{p_i^2 + m_i^2} \quad (3.9)$$

$$\simeq p + \frac{m_i^2}{2E} , \quad (3.10)$$

$$\rightarrow E_i - E_j = \frac{\Delta m_{ij}^2}{2E} . \quad (3.11)$$

This uses the assumption that all neutrinos are ultra-relativistic, allowing for the approximation that different mass eigenstates have the same momentum $p_i \simeq p_j = p \simeq E$ with the neutrino energy E . As the mass basis is orthogonal, $\langle \nu_i | \nu_j \rangle = \delta_{ij}$, with δ_{ij} the Kronecker delta. The probability $P_{\alpha\beta}$ of a neutrino produced as ν_α to interact as the potentially different flavor ν_β then amounts to

$$P_{\nu_\alpha \rightarrow \nu_\beta}(t) = |\langle \nu_\beta | \nu_\alpha(t) \rangle|^2 = P_{\nu_\alpha \rightarrow \nu_\beta}(L, E) \quad (3.12)$$

$$= \sum_{i,j} U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^* \exp\left(-i \frac{\Delta m_{ij}^2 L}{2E}\right) \quad (3.13)$$

$$= \delta_{\alpha\beta} - 4 \sum_{i>j} \text{Re}\left(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^*\right) \sin^2\left(\frac{\Delta m_{ij}^2 L}{4E}\right) \quad (3.14)$$

$$+ 2 \sum_{i>j} \text{Im}\left(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^*\right) \sin\left(\frac{\Delta m_{ij}^2 L}{2E}\right) . \quad (3.15)$$

This assumes unitarity of U . For antineutrinos, this derivation would yield an expression with the transition matrix U replaced with U^* [27].

From the transition probability, it becomes clear that neutrino oscillations do not depend on the absolute masses of the mass eigenstates but on the mass-squared differences $\Delta m_{ij}^2 = m_i^2 - m_j^2$.

The transition between $P_{\nu_\alpha \rightarrow \nu_\beta}(t)$ and $P_{\nu_\alpha \rightarrow \nu_\beta}(L, E)$ assumes ultrarelativistic propagation velocity [1] and is motivated by the fact that in experiments, the distance L traveled by the neutrino as well as its energy E are measured rather than its propagation time t . Neutrino oscillations are hence commonly characterized through their dependence on L , E or L/E .

For the interpretation of neutrino data, being aware of the oscillation length,

$$L_{ij}^{osc} = \frac{4\pi E}{\Delta m_{ij}^2}, \quad (3.16)$$

and its relation to the oscillation baseline L can be of interest. The oscillation length is defined as the distance at which the oscillation phase,

$$\phi_{ij} = -\frac{\Delta m_{ij}^2 L}{2E}, \quad (3.17)$$

is equal to 2π .

This behavior motivates different kinds of neutrino oscillation experiments at different baselines and neutrino energies, optimized to resolve survival-, appearance- or disappearance-probabilities as a means to measuring oscillation parameters. The current experimental measurements of the values of the PMNS-matrix parameters θ_{ij} , δ_{CP} as well as mass-squared differences Δm_{ij} can be found in table 7.3. Neutrinos oscillations based on SM interactions will in the following be called standard oscillations.

3.2.2 Matter effects

Neutrino oscillations are not only dependent on baseline, neutrino energy and mass-squared differences, but also impacted by forward scattering when traversing matter. This can be understood as an effective potential, yielding different kinds of effects depending on the underlying density profile [1].

Matter effects are caused by coherent forward scattering of neutrinos - specifically by how this differs between neutrino flavors. Electron neutrinos show a different behavior in matter, as they are the only neutrino flavor that undergoes CC scattering on surrounding electrons, which is shown in figure 3.2.

In cases where neutrino oscillations occur, this difference between flavor components of a neutrino wave changes the oscillation behavior. The effect can be considered interference between forward-scattered and unscattered waves, which results in a change of phase velocity. Conceptually it is similar to vacuum oscillations, where relative phases of neutrino states change purely due to their different masses. The

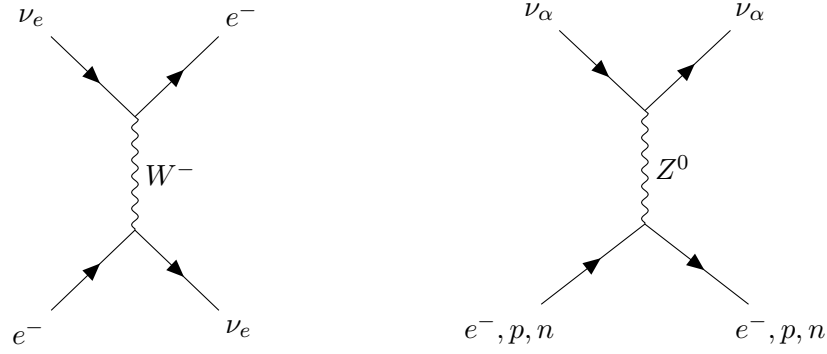


Fig. 3.2.: Forward scattering of neutrinos on regular matter as included in the SM. While NC scattering (right diagram) is possible for all neutrino flavors, electron neutrinos can also undergo CC scattering (left diagram) on surrounding electrons.

impact of surrounding matter is similar to the presence of an external potential [27]. The impact of varying density profiles can be described as density dependent effective neutrino mixing angles. Observable effects can be categorized approximately by the profile characteristics. Density variations that are monotonous and smooth can result in an adiabatic flavor transitions, independently of oscillation effects. Mixing, and therefore the composition of individual states, changes with density. A neutrino can therefore change its flavor purely by change of surrounding density, even if no interference effects in propagation occur.

For solar matter, this resolved the solar neutrino deficit [18, 27] mentioned in section 3.1.1. A first description of this by Wolfenstein [16] was extended by Smirnov and Mikheev [34], motivating the name Mikheyev-Smirnov-Wolfenstein (MSW) effect. An effective angle that results in maximal disappearance of a neutrino flavor is described as MSW resonance. The occurrence of this depends on the relative phase between the propagating neutrino eigenstates as well as the effective external potential.

The layer structure of Earth matter provides neighboring areas of almost constant density [35], something that potentially leads to resonant parametric enhancement of flavor transitions [36, 37]. In [33], a detailed review of the different matter effects can be found. A general description of matter effects can be provided by extending the Hamiltonian given in equation 3.8 to

$$\mathcal{H}_\nu = \mathcal{H}_{\text{vac}} + \mathcal{H}_{\text{mat}}(x) \quad (3.18)$$

$$\mathcal{H}_{\bar{\nu}} = [\mathcal{H}_{\text{vac}} - \mathcal{H}_{\text{mat}}(x)]^* \quad (3.19)$$

for neutrinos and antineutrinos, respectively [38]. The matter Hamiltonian $\mathcal{H}_{\text{mat}}(x)$ consists of a position dependent potential $V(x)$ which describes coherent forward scattering. In case of SM interactions, this is a diagonal matrix of the potentials

$V_\alpha(x)$, which each solely impact neutrinos of the α flavor. For matter that consists of protons, electrons and neutrons, only $V_e(x)$ differs from the other potentials. As standard oscillation effects only depend on the difference between potentials, $V_\mu(x)$ and $V_\tau(x)$ can be set to zero without observable impact. This leaves $V(x)$ with the CC component only,

$$V(x) = \text{diag}(V_e(x), V_\mu(x), V_\tau(x)) \quad (3.20)$$

$$= \text{diag}(V_{CC,e}(x), 0, 0) \quad (3.21)$$

$$= \sqrt{2}G_F N_e \text{diag}(1, 0, 0) \quad , \quad (3.22)$$

with Fermi's coupling constant G_F and the local electron number density N_e [33]. Matter effects in the Sun or Earth are what drives the sensitivity to oscillations effects in many cases, and also to NSI in most of the existing measurements [39].

3.2.3 Earth effects on atmospheric neutrinos

This work focuses on NSI detection through Earth matter effects on atmospheric neutrinos. The modeling of the radially symmetric Earth density profile $\rho(r)$ thereby affects neutrino flavor transitions.

Earth can be approximated as consisting of layers of different, constant densities, the values of which decrease with the radius. The density profile used in this work is adopted from the Preliminary Reference Earth Model (PREM) [35].

In atmospheric neutrino data, the zenith angle θ under which neutrinos reach a detector corresponds to the baseline at which they traverse the individual layers [1], as shown schematically in figure 3.4. The maximum total oscillation baseline is $\sim 1.3 \times 10^4$ km for Earth-crossing neutrinos.

The large and abrupt density changes between the Earth core and mantle form a so-called castle wall profile. The resulting resonant amplification of neutrino flavor transitions through parametric enhancement [36, 37] is depicted in figure 3.3 for a two-flavor neutrino model. An oscillogram showcasing the energy and zenith dependence of standard oscillations for DeepCore data can be found in section 7.1. In addition to solid Earth matter, the atmosphere has to be taken into account. In this and earlier IceCube NSI analyses [40], a neutrino production height of $h = 20$ km is assumed [41].

The factor ultimately affecting oscillations is the electron number density (see equation 3.22) which can be derived from the matter density and material composition. Details on how the Earth and especially its electron number density is modeled in this analysis are given in section 7.3.2.

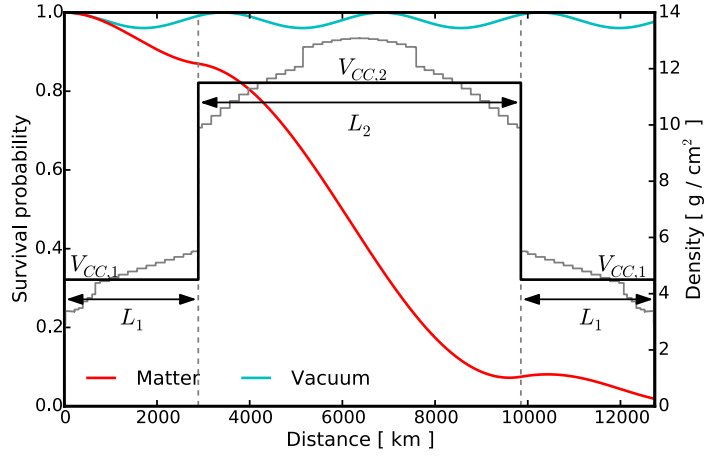


Fig. 3.3.: Survival probability in vacuum (cyan) and matter (red) for a two-flavor neutrino model. A detailed PREM Earth matter density profile is shown in gray and the simplified castle wall profile in black. The layer thickness and effective potential resulting from the castle wall profile are indicated as $V_{CC,i}$ and L_i , with the index $i = 1, 2$ denoting the mantle and core, respectively. The survival probability in this matter profile is calculated for neutrino energies of 1.37 GeV, at the two-neutrino oscillation parameter values $\sin^2 \theta = 0.01$ and $\Delta m^2 = 10^{-3} \text{ eV}^2$. The figure was taken from [15].

3.2.4 Open questions on neutrino physics

As mentioned in the above sections, there is a multitude of open questions revolving around neutrino properties, especially with regard to oscillations. These include the NMO, the existence of CP violation through $\delta_{CP} \neq 0$, the octant of the atmospheric mixing angle θ_{23} , and the sign of the atmospheric mass-squared difference. Upcoming experiments, such as the Deep Underground Neutrino Experiment (DUNE) [42], hope to resolve these questions. This, however, relies on the current assumptions on neutrino physics to be correct. A different open question is how neutrinos obtain mass, which entails unknown BSM-physics.

The existence of NSI could be consequence of a neutrino mass extension to the SM and might impact measurements of the standard oscillation parameters by introducing degeneracies in δ_{CP} , NMO and mixing angles. Single NSI scenarios⁵ have been shown to reduce or fully erase the sensitivity of existing and future oscillations experiments to e.g. δ_{CP} and the θ_{23} octant [3]. A complete measurement of NSI and standard oscillation parameters has also been shown perturbatively to be impossible in counting analyses like the one in this work [43]. Consequently, detection or exclusion of NSI is relevant for the interpretation of standard oscillation measurements.

⁵Specifically, small non-zero values of the effective NSI couplings $\epsilon_{e\mu}^\oplus$ and $\epsilon_{e\tau}^\oplus$ were tested.

3.3 Formalism of non-standard neutrino interactions

The concept of NSI is a very general one, including any neutrino interactions that are not included in the SM. It was first introduced together with that of matter effects by Wolfenstein [16]. As was discussed in section 3.1.4, NSI naturally result from several possible extensions of the SM that are designed to accommodate non-zero neutrino masses. The current understanding is that NSI affect flavor changes on a scale sub-leading with respect to standard oscillations.

The focus of this analysis is on NC NSI, which are observable mostly through matter effects. In contrast, CC NSI primarily impact neutrino detection and production [3]. In order to be independent of the specific underlying model, NC NSI can be parameterized as an effective matter potential with nine degrees of freedom, to eight of which oscillation effects are sensitive. In the following, this parametrization will be discussed as well as the phenomenology, signature and existing measurements of the individual NSI parameters.

As an effective theory, the SM would be the manifestation of a more complex underlying theory in the low-energy range, where BSM-physics effects of that theory would be suppressed. The latter would be expected to gain relevance at some higher energy scale Λ .

One commonly used way to model NSI is to assume a new mediator particle Z' with mass $m_{Z'}$. This is for example applied for searches that use accelerator neutrinos, which are directly sensitive to the mediator mass as they observe decays with Z' in the final state. Such an approach induces a strong dependence on the underlying model as well as the specific $m_{Z'}$, as the respective final state needs to be achievable within the collider energy range [44]. For oscillations experiments, this dependence is not necessary, and a less model oriented approach can be taken.

In order to yield the most general possible constraints in this analysis, an NSI parametrization is used that is based on effective coupling strengths ϵ . This approach is largely independent of the underlying model and energy scale Λ of the unknown interactions [2].

In this analysis, NSI are assumed to comprise coherent forward scattering, assuming negligible contributions from incoherent interactions. These can be subdivided into CC and NC scattering. At energies below the mass scale of weak bosons, neutrino interactions are described using dimension six operators [45]. The four-fermion Lagrangian densities describing NSI effects can be formulated largely model-independently [2] as

$$\mathcal{L}_{\text{NSI}}^{\text{NC}} = -2\sqrt{2}G_F\epsilon_{\alpha\beta}^{fC} (\bar{\nu}_\alpha\gamma^\mu P_L\nu_\beta) (\bar{f}\gamma_\mu P_C f) \quad , \quad (3.23)$$

$$\mathcal{L}_{\text{NSI}}^{\text{CC}} = -2\sqrt{2}G_F\epsilon_{\alpha\beta}^{ff'C} (\bar{\nu}_\alpha\gamma_\mu P_L\ell_\beta) (\bar{f}'\gamma^\mu P_C f) \quad . \quad (3.24)$$

These equations imply summation over the type of charged fermion $f \neq f'$, chirality $C = R, L$ and lepton flavor $\alpha, \beta = e, \mu, \tau$. Chiral projection is indicated by the operator $P_{R,L} = (1 \pm \gamma_5)/2$. The fermions that neutrinos interact with in Earth matter are first generation charged ones, namely e^- , u , and d , as depicted in figure 3.4.

Similarly to the standard model case, CC effects caused by NSI mainly impact production and detection of neutrinos. This makes neutrino sources with well known spectra, such as reactors, especially sensitive to them. This analysis is based on atmospheric neutrinos, which are more sensitive to effects during neutrino propagation and hence to NSI forward scattering from NC interactions [46].

Within equation 3.23, NSI are parameterized through the effective NC and CC coupling strengths $\epsilon_{\alpha\beta}^{fC}$ and $\epsilon_{\alpha\beta}^{ff'C}$. For propagation through matter only the vector part of NSI couplings is relevant [47], which can be summarized to

$$\epsilon_{\alpha\beta}^f \equiv \epsilon_{\alpha\beta}^{fL} + \epsilon_{\alpha\beta}^{fR} \quad . \quad (3.25)$$

The individual fermion components enter the overall local effective couplings $\epsilon_{\alpha\beta}$ scaled with the position dependent number density $N(x)$ of the respective fermion as

$$\epsilon_{\alpha\beta}(x) = \epsilon_{\alpha\beta}^f \sum_f \frac{N_f(x)}{N_e(x)} \quad . \quad (3.26)$$

The SM case is recovered for $\epsilon_{\alpha\beta} = 0$. Coupling strengths of $\epsilon_{\alpha\beta} = 1$ imply NSI effects at a scale comparable to that of weak interactions [2].

In the parameter description in equation 3.26, the ratio between number densities is motivated through $\epsilon_{\alpha\beta}(x)$ corresponding to the ratio of the respective effective NSI matter potential $V_{\text{NSI},\alpha\beta}(x)$ and the SM CC matter potential $V_{\text{CC}}(x)$ as

$$\epsilon_{\alpha\beta}(x) = \frac{V_{\text{NSI},\alpha\beta}(x)}{V_{\text{CC}}(x)} \quad . \quad (3.27)$$

While the SM matter potential purely depends on electron number density N_e , the effective NSI matter potential,

$$V_{\text{NSI},\alpha\beta}(x) = \sqrt{2}G_F \sum_f \epsilon_{\alpha\beta}^{fV} N_f(x) \quad , \quad (3.28)$$

includes the impact of coherent forward scattering off all three first generation charged fermions in a linear combination.

The overall effective coupling strengths $\epsilon_{\alpha\beta}(x)$ can be equivalently represented through the couplings to neutrons and protons instead of their partons⁶ as

$$\epsilon_{\alpha\beta}(x) = \sum_{N=e,n,p} \epsilon_{\alpha\beta}^N \frac{N_N(x)}{N_e(x)} \quad (3.29)$$

$$= \epsilon_{\alpha\beta}^e + \epsilon_{\alpha\beta}^p + \epsilon_{\alpha\beta}^n \frac{N_n(x)}{N_e(x)} \quad . \quad (3.30)$$

Equation 3.30 assumes neutral matter, yielding $N_e = N_p$. The ratio Y_n^\oplus between neutron and electron number densities in Earth is nearly constant [39, 40] at

$$Y_n^\oplus \equiv \langle N_n(x)/N_e(x) \rangle \approx 1.051 \quad . \quad (3.31)$$

As a result, an approximate independence of x can be assumed for the NSI couplings, yielding

$$\epsilon_{\alpha\beta} = \epsilon_{\alpha\beta}^e + \epsilon_{\alpha\beta}^p + Y_n^\oplus \epsilon_{\alpha\beta}^n \quad . \quad (3.32)$$

The generalized matter Hamiltonian, which includes both SM and NSI effects, can be described based on the underlying potentials analogously to that of SM matter effects discussed in section 3.2.2 as

$$H_{\text{mat}}(x) = V_{\text{CC}}(x) + V_{\text{NSI}}(x) \quad . \quad (3.33)$$

Assuming Hermiticity and equations 3.28 and 3.32, it can be expressed as

$$H_{\text{mat}}(x) \approx V_{\text{CC}}(x) \begin{pmatrix} 1 + \epsilon_{ee} & \epsilon_{e\mu} & \epsilon_{e\tau} \\ \epsilon_{e\mu}^* & \epsilon_{\mu\mu} & \epsilon_{\mu\tau} \\ \epsilon_{e\tau}^* & \epsilon_{\mu\tau}^* & \epsilon_{\tau\tau} \end{pmatrix} \quad . \quad (3.34)$$

The top left element contains a 1, which describes SM matter effects through CC electron neutrino scattering on electrons as shown in figure 3.2. Cases where other components are non-zero purely represent NSI effects. Scenarios where either single or multiple NSI parameters have non-zero values are theoretically possible.

Off-diagonal elements of the effective Hamiltonian in equation 3.34 are complex valued with a CP violating phase $\delta_{\alpha\beta}$, following

$$\epsilon_{\alpha\beta} = |\epsilon_{\alpha\beta}| e^{i\delta_{\alpha\beta}} \quad |\alpha \neq \beta \quad . \quad (3.35)$$

A non-zero magnitude of such a parameter would imply direct conversion between neutrino flavors at interaction, motivating the term flavor changing (FC) parameters [3].

⁶Quarks constitute nucleons in the combinations u, u, d for protons and u, d, d for neutrons.

Diagonal elements are real-valued and contain interactions where the in-going neutrino flavor equals the out-going one. They provide the option to break flavor universality and will in the following be called flavor diagonal (FD) parameters. As atmospheric neutrino data are only sensitive to this potential flavor non-universality [3], the effective parametrization can without observable consequences be changed by subtracting $\mathbb{1} \cdot \epsilon_{\mu\mu}$, yielding

$$H_{\text{mat}}(x) \approx V_{\text{CC}}(x) \begin{pmatrix} 1 + \epsilon_{ee} - \epsilon_{\mu\mu} & \epsilon_{e\mu} & \epsilon_{e\tau} \\ \epsilon_{e\mu}^* & 0 & \epsilon_{\mu\tau} \\ \epsilon_{e\tau}^* & \epsilon_{\mu\tau}^* & \epsilon_{\tau\tau} - \epsilon_{\mu\mu} \end{pmatrix}. \quad (3.36)$$

This leaves this analysis with a total of three complex and two real NSI parameters to investigate. These represent a variety of possible BSM-physics effects that cause NC forward scattering in matter.

3.4 The impact of NSI

The signal characteristics of NSI are diverse and have distinct features that will be discussed in the following, along with the current status of NSI measurements within and beyond IceCube and DeepCore.

3.4.1 NSI phenomenology: Global behavior and remarks

The impact of NC NSI investigated in this work becomes observable through their change of the effective matter potential seen by passing neutrinos. The resulting signal depends on neutrino energy and baseline and its phenomenology differs between individual NSI parameters and parameter values. In the following, the key features at probability level that are present in all five NSI parameters are discussed for the DeepCore energy range, largely based on [40, 48, 49]. Properties at data level are discussed in section 7.1.2 specifically for DeepCore.

The probabilities shown in the following are $P_{\alpha\beta}$ with $\alpha \in e, \mu$, as the flux of tau neutrinos in the atmosphere is negligible at the GeV range [41]. The selected scenarios are realizations of single non-zero NSI parameters, varied within the respective range that they can be constrained to based on existing measurements (see section 3.5). A discussion of current limits can be found in section 3.5.

In all shown scenarios $\delta_{CP} = 0$ is assumed and $\theta_{23} = 47.2$ deg is set close to maximal mixing. If not indicated otherwise, probabilities are given at a fixed zenith angle

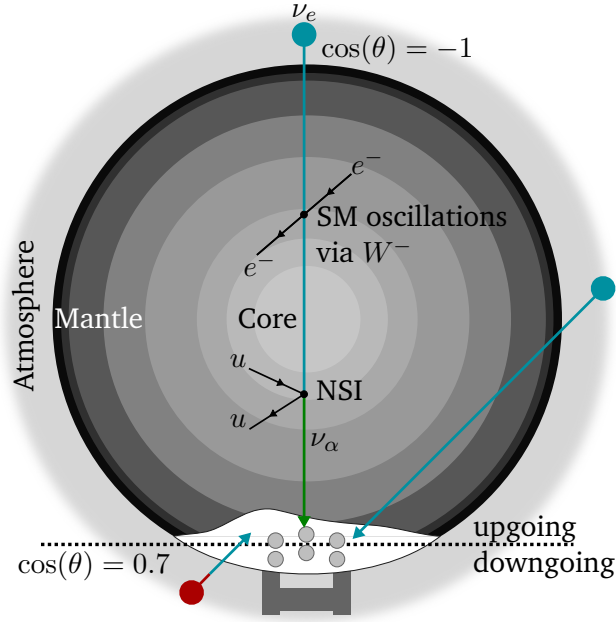


Fig. 3.4.: Schematic of SM and NSI based matter oscillations effects on atmospheric neutrinos. After being produced inside the atmosphere, neutrinos traverse Earth matter at baselines that depend on their zenith angle θ before potentially being detected inside IceCube. The baselines within layers of different density impact the oscillations behavior through the local electron and nucleon density. Within the SM, matter effects on oscillations are induced through CC scattering of electron neutrinos on electrons (see section 3.2.3). The NSI considered in this work comprise NC forward scattering of all neutrino flavors on first generation charged fermions.

that corresponds to $\cos(\theta) = -0.75$, which is diagonally up-going in the convention introduced in figure 3.4. This fixes the total baseline to $L \sim 9.6 \times 10^3$ km. Under this angle, neutrinos traverse the Earth mantle but not the core. This results in strong but not maximal NSI effects on oscillations, constituting a case similar to what most atmospheric neutrinos in a DeepCore data sample will undergo. Also, in this case, almost full electron neutrino disappearance through MSW resonance is observable in the case of no NSI at resonance energy $E_{res} \sim 6$ GeV [40, 50].

Assuming no CP violation ($\delta_{CP} = 0$) results in a spacial symmetry of the Earth matter, namely $V(x) = V(L - x)$ with L the total baseline, leading to $P_{\alpha\beta} \approx P_{\beta\alpha}$ [40]. With the assumed angle θ_{23} close to maximal mixing, electron neutrinos are converted to tau- and muon neutrinos at similar probabilities, $P_{e\mu} \approx P_{e\tau}$.

At energies above $E_\nu \simeq 20$ GeV, the evolution of electron neutrinos decouples from the rest of the system. The decoupling occurs due to the large mass differences between neutrino mass eigenstates [51] and yields $P_{ee} \sim 1$. This only applies for NSI in the electron sector significantly below the electroweak scale, with $\epsilon_{ee}^\oplus - \epsilon_{\mu\mu}^\oplus$,

$\epsilon_{e\mu}^\oplus, \epsilon_{e\tau}^\oplus \ll 1$, as is the case in most⁷ scenarios shown in the following discussion. As a result of this decoupling, the evolution of muon and tau neutrinos can be approximated as a two neutrino case, where standard oscillations are driven purely by $\Delta m_{31}^2, \theta_{23}$ and

$$V_{\text{NSI}} \propto \begin{pmatrix} 0 & \epsilon_{\mu\tau} \\ \epsilon_{\mu\tau}^* & \epsilon_{\tau\tau} - \epsilon_{\mu\mu} \end{pmatrix}. \quad (3.37)$$

At energies where $\Delta m_{31}^2/E_\nu \ll V_{\text{NSI}}$ ⁸, flavor changing NSI effects start to dominate over standard oscillations and flavor diagonal NSI [48].

The oscillation behavior of neutrinos differs from that of antineutrinos, as exchanging them corresponds to a change of sign of the effective potential and a complex conjugation of the PMNS-matrix. Since neutrinos and antineutrinos cannot be distinguished in IceCube the sensitivity to NSI at detector level stems from the combined signal of both. The larger cross section and flux [11, 41] of neutrinos, however, entails that NSI effects that are inverted for neutrinos and antineutrinos do not fully cancel out.

The MSW effect, described in section 3.2.2, can create resonances in flavor transitions that are pointed out in the following for individual NSI parameters. At the corresponding energies E_{res} , mixing between the ν_3 and ν_1 mass eigenstate becomes maximal due to the value of the effective θ_{13} mixing angle. This corresponds to full disappearance of electron neutrinos into muon and tau neutrinos. Depending on the specific NSI parameter value, the condition for MSW resonance can be fulfilled at different E_{res} .

In a perturbation theoretical approach [43], which will be referred to for individual phenomenological characteristics, all NSI parameters can be assumed to be at the same magnitude of $\epsilon \sim 10^{-2}$. Their impact can be estimated based on the order of ϵ at which they enter oscillation probabilities in a perturbative model. This is reflected in different orders of magnitude at which the individual NSI parameters can be constrained, and also, different ranges at which they are varied for this study. The full table of orders of ϵ for the individual parameters can be found in appendix section A.2. The transition probabilities given throughout this chapter show different realizations of individual NSI parameters. They assume the same Earth model as the overall analysis and realistic standard oscillations parameters, as detailed in the caption of figure 3.5.

⁷This is not given for some realizations of $\epsilon_{ee}^\oplus - \epsilon_{\mu\mu}^\oplus$ shown in the respective discussion in section 3.4.3.

⁸Under the assumptions made here this is the case at $E_\nu \sim 100$ GeV.

3.4.2 Phenomenology of $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$

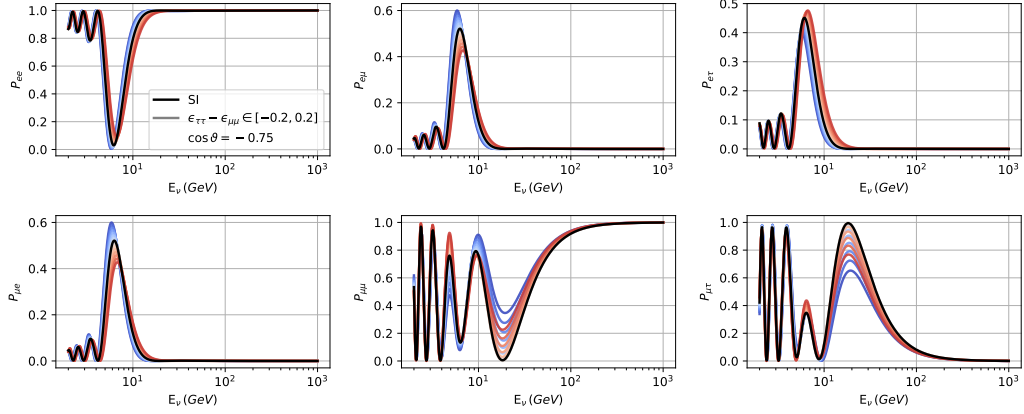


Fig. 3.5.: Probabilities for neutrino transitions between neutrino flavors $P_{\alpha\beta}$ at different values of the NSI parameter $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$. A zenith angle of $\cos(\theta) = -0.75$ is assumed. The individual probabilities, given on the vertical axes of the plots, depend on the neutrino energy, given on the horizontal axis. Values of $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ are between -0.2 and 0.2 , with negative values shown in blue and positive values in red. The standard oscillations case at $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus} = 0$ is highlighted in black. Note that all other NSI parameters are set to zero. The Earth is modeled as described in section 7.3.2. Standard oscillation parameters are set to $\theta_{12} = 33.62^\circ$, $\theta_{13} = 8.54^\circ$, $\theta_{23} = 47.2^\circ$, $\delta_{CP} = 0^\circ$, $\Delta m_{21}^2 = 7.40 \times 10^{-5} \text{ eV}^2$, $\Delta m_{21}^2 = 2.494 \times 10^{-3} \text{ eV}^2$ and NO is assumed.

Non-zero values of only the flavor diagonal NSI parameter $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ conserve flavors and introduce flavor non-universality. Figure 3.5 shows the oscillation probabilities $P_{\alpha\beta}$ at different values of $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$. As no flavor changes are introduced through this parameter, its effect is the same as altering the standard oscillations matter potential for 2-3 mixing, yielding the effective potential $V_{NSI} = (\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}) V_{CC}$.

The impact of some of the remarks in the previous section are visible here, such as the decoupling of the electron flavor from the μ - τ -sector. This becomes evident from the top three and bottom left panels in figure 3.5, where oscillation probabilities involving the electron flavor are constant above $E_\nu \sim 20 \text{ GeV}$. As a result, $P_{\mu\mu} \approx 1 - P_{\mu\tau}$.

In this sector, the main impact of varying $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ becomes apparent, which is damping of oscillations in the μ - τ -sector above the MSW resonance energy $E_R \sim 6 \text{ GeV}$. Above $E_\nu \sim 200 \text{ GeV}$, the impact of $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ is visibly negligible.

In figure 3.6, the effect on antineutrinos is shown. While the overall oscillation behavior differs substantially, the same arguments as in the neutrino case apply. Further on, this discussion of NSI phenomenology will focus on neutrinos. The respective figures for antineutrino behavior can be found in appendix B.1.

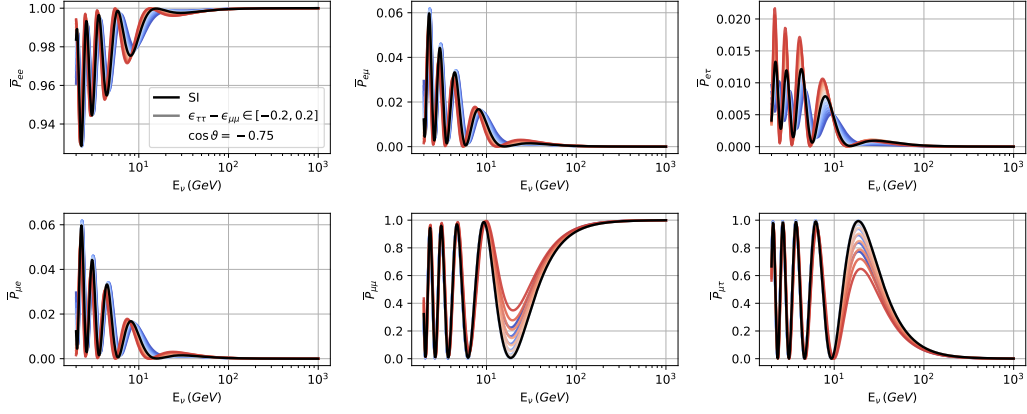


Fig. 3.6.: Same as figure 3.5, but for antineutrinos.

The observable signal for non-zero $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ from atmospheric neutrinos and antineutrinos combined is impacted strongest in the muon neutrino survival probability, even more so than for neutrinos and antineutrinos individually, as their impact on the flux in the electron neutrino sector partially cancels out.

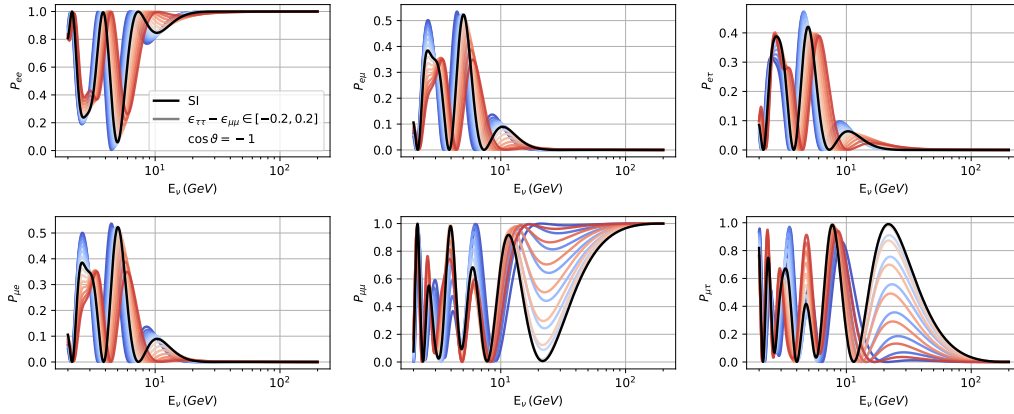


Fig. 3.7.: Same as figure 3.5, but for $\cos(\theta) = -1$. Figures showing other zenith angles can be found in appendix B.1.

The baselines at which neutrinos cross different layers within the Earth are fixed through their zenith angle. In addition to the geometrical increase of the overall baseline for zenith angles closer to $\cos(\theta) = -1$, which is shown in figure 3.7, layer densities increase towards the Earth core. Matter effects are consequently more pronounced, which can especially be seen at the first electron neutrino transition maximum around 3 GeV.

3.4.3 Phenomenology of $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$

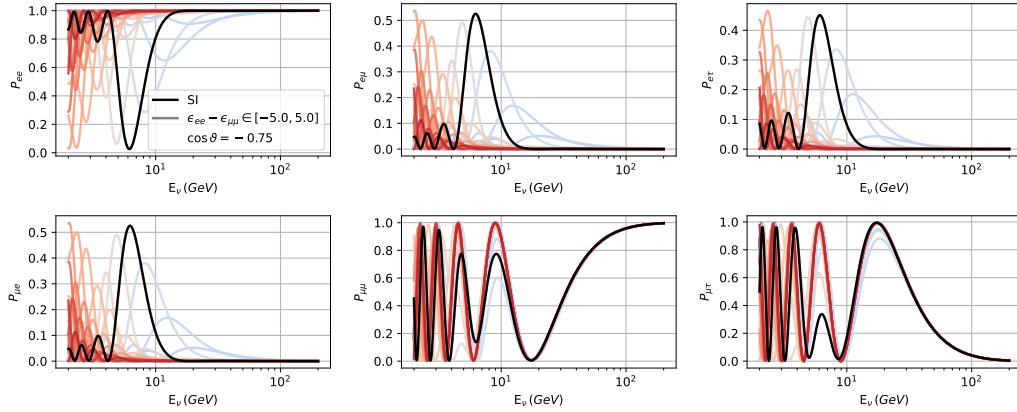


Fig. 3.8.: Same as figure 3.5, but for $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ values between -5 and 5 . Note that within the μ - τ -sector, most of the realizations of negative parameter values (shown in blue) are at the same position as the most extreme positive ones (in dark red).

Figure 3.8 shows different scenarios where $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ is the only source of NSI, with all assumptions kept the same as in the previous section on $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$. As a flavor diagonal parameter in the electron sector, $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ can be seen as a re-scaling of standard oscillations matter effects, since both are in the same matrix element, $1 - (\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus})$. As a result, $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus} = -1$ mimics vacuum oscillations, yielding the most significant difference in probabilities from the standard oscillations case. In vacuum, all transitions that involve electron neutrinos are suppressed at the GeV range. This originates from the different scales of the mass-squared differences leading to decoupling of the individual mass eigenstates in propagation [1].

Most negative $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ parameter values shown in figure 3.8 result in electron neutrino survival probabilities close to 1, while any transition probabilities $P_{e\alpha}$ ($\alpha \in [\mu, \tau]$) approximate to zero. This is, on the one hand, a consequence of negative $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ values yielding effects close to vacuum oscillations. On the other hand, negative $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ values in general lead to a suppression of oscillations effects that include electron neutrinos, while positive values lead to an enhancement.

The energy at which the MSW resonance criterion is satisfied can be shown to directly depend on $(\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus})^{-1}$ [48]. The result is a suppression or enhancement as well as a shift of the electron neutrino transition maximum towards lower energies for higher values of $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ [40, 49].

Probabilities that do not involve electrons experience a suppression of transitions for the most extreme parameter values that are shown. As mentioned before, they propagate mostly as in vacuum for $E_{\nu} \geq 20$ GeV. However, a full decoupling of the electron sector is not maintained for the larger values of $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ shown here.

Overall, the resulting signal makes atmospheric neutrinos mostly sensitive to $\epsilon_{ee}^\oplus - \epsilon_{\mu\mu}^\oplus$ outside the muon disappearance channel. This implies reduced sensitivity in IceCube.

3.4.4 Phenomenology of $\epsilon_{\mu\tau}^\oplus$

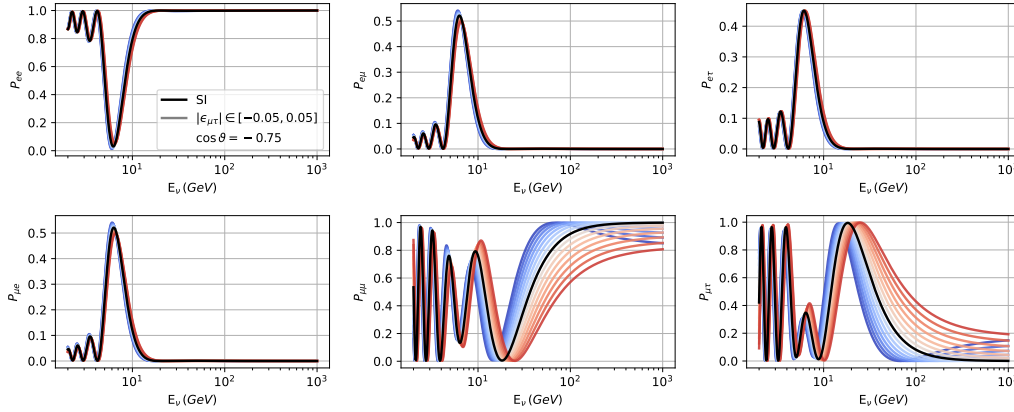


Fig. 3.9.: Same as figure 3.5, but for $|\epsilon_{\mu\tau}^\oplus| \in [-0.05, 0.05]$ at $\delta_{\mu\tau} = 0$.

If $\epsilon_{\mu\tau}^\oplus$ is the only non-zero coupling strength for NSI, flavor universality is conserved while flavor changing interactions are introduced in the μ - τ -sector. The impact of different $\epsilon_{\mu\tau}^\oplus$ realizations, shown in figure 3.9, concentrates almost exclusively on this sector, from which electron neutrino evolution mostly decouples. As mentioned above, flavor changing NSI becomes the dominant source of NSI effects at energies above $E_\nu \sim 100$ GeV [48]. Sensitivity to $\epsilon_{\mu\tau}^\oplus$, especially at energies above this, consequently lies in muon disappearance. The effect on antineutrinos is inverted with respect to that on neutrinos, reducing the detectable signal in detectors like IceCube, where no differentiation between neutrinos and antineutrinos is possible. The energy range shown for $\epsilon_{\mu\tau}^\oplus$ was chosen differently from that in other probability plots, reflecting how the signal continues at higher energies up to the TeV scale. While not relevant for DeepCore analyses, this energy range is available in IceCube data and used to determine $\epsilon_{\mu\tau}^\oplus$ in other analyses. This will be discussed in more detail in section 3.5.

An example for a non-zero complex phase value of a flavor changing NSI parameter is shown in figure 3.10. Only positive magnitude values are shown, as the phase space would otherwise be covered twice, with the complex phase covering $[0, 2\pi]$ and shifting it by π corresponding to changing the sign of the magnitude value. Phase changes, as a result, smoothly vary oscillation probabilities between these edge

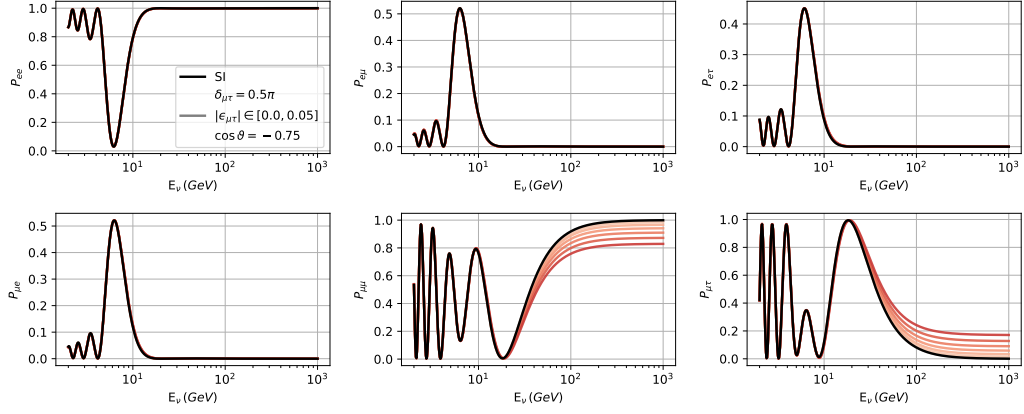


Fig. 3.10.: Same as figure 3.5, but for different magnitudes of $|\epsilon_{\mu\tau}^\oplus|$ at $\delta_{\mu\tau} = \pi/2$. Note that for non-zero complex phases of flavor changing NSI parameters, only positive magnitudes are shown. Parameter magnitude values hence are the positive-valued subset of what is shown in figure 3.9.

conditions. At $\delta_{\mu\tau} = \pi/2$, the coupling is purely imaginary. This makes MSW like enhancement impossible, as it does not allow for maximizing the effective mixing angle obtained from diagonalizing the Hamiltonian [33, 49, 50]. The transition probabilities for purely imaginary $\epsilon_{\mu\tau}^\oplus$ coupling hence only differ from standard oscillations in the μ - τ -sector.

3.4.5 Phenomenology of $\epsilon_{e\mu}^\oplus$ and $\epsilon_{e\tau}^\oplus$

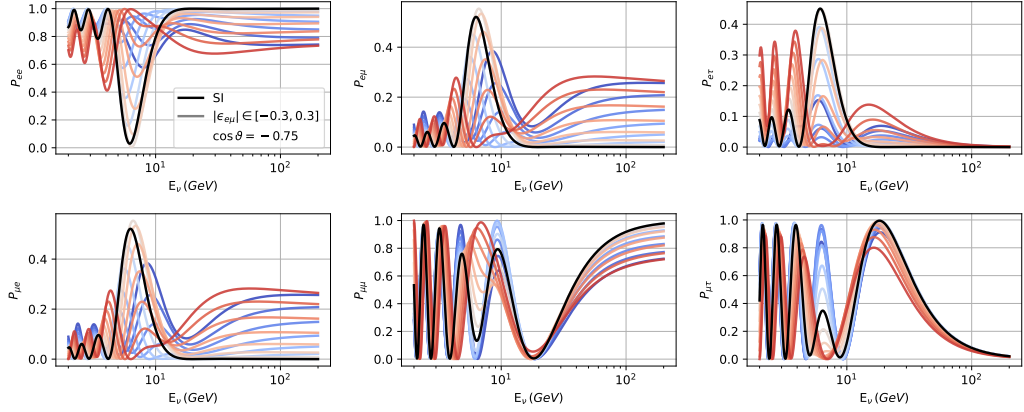


Fig. 3.11.: Same as figure 3.5, but for different magnitudes of $|\epsilon_{e\mu}^\oplus|$ at $\delta_{e\mu} = 0$.

In figures 3.11 and 3.12, different realizations of the flavor violating NSI parameters $\epsilon_{e\mu}^\oplus$ and $\epsilon_{e\tau}^\oplus$ are shown, respectively, at complex phases of 0° . The transition behavior induced by $\epsilon_{e\mu}^\oplus$ compared to $\epsilon_{e\tau}^\oplus$ is very similar when replacing τ with μ in the shown

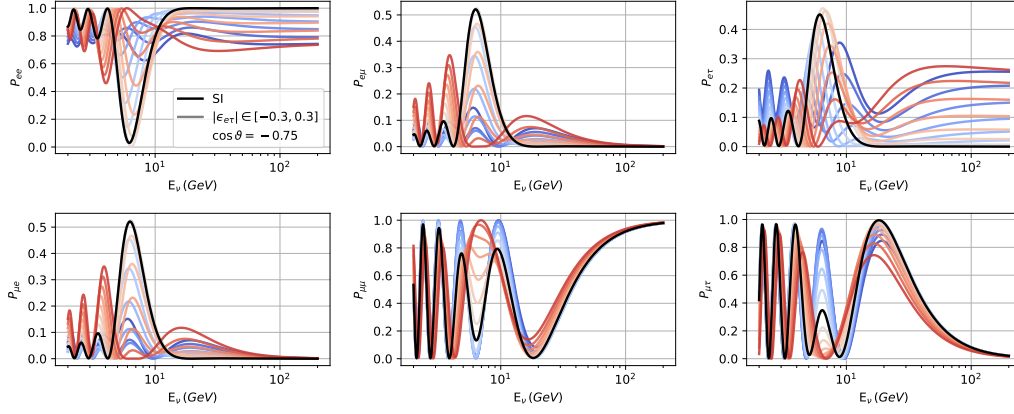


Fig. 3.12.: Same as figure 3.5, but for different magnitudes of $|\epsilon_{e\tau}^{\oplus}|$ at $\delta_{e\tau} = 0$.

probabilities. Figures showing other complex phase values and antineutrinos can be found in appendix B.1.

The subdominant impact of these NSI parameters can be motivated based on the perturbation theoretical approach mentioned previously. Therein, $\epsilon_{e\mu}^{\oplus}$ and $\epsilon_{e\tau}^{\oplus}$ only take part at ϵ^2 in any disappearance channel, while e.g. the impact of $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ on muon neutrino disappearance happens at first order (see appendix A.2). As a result, these parameters are not as much the focus of atmospheric neutrino experiments [43].

3.5 NSI measurements in IceCube and beyond

The different energy ranges of events detectable in IceCube and DeepCore allow for two types of NSI analyses. These differ in their data processing, nuisance parameter treatment and their simulation.

IceCube event sample energies are between 500 GeV and 10 TeV. Therefore, IceCube data can only constrain $\epsilon_{\mu\tau}^{\oplus}$, as can be seen in the phenomenological features discussed above. The most recent such IceCube analysis [52] currently provides the strongest limits on this parameter, $-0.0041 \leq |\epsilon_{\mu\tau}^{\oplus}| \leq 0.0031$.

The DeepCore analysis preceding this work [40] was the first to provide limits on all NSI parameters, which becomes possible due to the energy range of 5.6-100 GeV. It includes three years of DeepCore data and is the analysis that is most directly comparable to this work. Differences mostly stem from improvements in simulation, nuisance parameter treatment and event selection.

Tab. 3.1.: Current status of NSI measurements, restricted to experiments that are directly comparable to this work, obtained in DeepCore [40], IceCube [52] and ANTARES [53]. Note that ANTARES results use a different parametrization and are therefore re-scaled by a factor of three with respect to the published values.

	90 % CL	Experiment
FD NSI		
$\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$	$[-2.26, -1.27] \cup [-0.74, 0.32]$	DeepCore
$\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$	$[-0.041, 0.042]$	DeepCore
$\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$	$[-0.183, -0.042] \cup [0.042, 0.243]$	ANTARES
FC NSI		
$ \epsilon_{e\mu}^{\oplus} $	≤ 0.146	DeepCore
$ \epsilon_{e\tau}^{\oplus} $	≤ 0.173	DeepCore
$ \epsilon_{\mu\tau}^{\oplus} $	≤ 0.0232	IceCube
$\epsilon_{\mu\tau}^{\oplus}$ for $\delta_{\mu\tau} = 0^{\circ}, 180^{\circ}$	$[-0.0141, 0.087]$	ANTARES

Beyond IceCube and DeepCore, a multitude of different approaches to measuring NSI exists, as introduced in sections 2.2 and further on in section 4. In table 3.1, the current 90 % ranges of all NSI parameter magnitudes are given for results that are directly comparable with this work. No conclusive limits on NSI complex phases have been set to this point.

The details of parameter definition may change between the compared results. Therefore, values given here may differ from those in the respective reference. Mostly, this concerns a factor of ~ 3 which is introduced through defining NSI parameters per quark instead of per electron.

The two disjoint allowed ranges for $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ stem from $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus} = -1$ mimicking vacuum oscillations. While this hypothesis can be excluded exceedingly well, there is less sensitivity to parameter values above or below it.

The overall tightest bounds are on $\epsilon_{\mu\tau}^{\oplus}$, due to the large energy range at which this parameter impacts neutrinos. Similarly, as for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ the constraints profit from the strong signal provided in the muon neutrino disappearance channel.

A comparison to a larger number of results is given in section 7.8, including less directly comparable global analyses and the COHERENT experiment⁹ as well as the sensitivities obtained in this work. Note that experiments that can not reasonably be compared to this analysis are not shown therein either. This excludes e.g. collider limits, which are only valid for specific models or mediator mass ranges.

Summarizing, measurements of neutrino oscillations result in the necessity of BSM-physics, as they point towards massive neutrinos. Many of the possible respective extensions of the SM naturally entail NSI. Their existence could explain contradictory observations of the Tokai to Kamioka experiment (T2K) and NuMI Off-Axis ν_e

⁹Some detail on the COHERENT experiment can be found in appendix section A.3.

Appearance experiment (NO ν A), which yield conflicting best fit hypotheses of NMO, the octant of θ_{23} and δ_{CP} [54, 55].

This work takes a mostly¹⁰ model independent approach to searching for NSI by applying a parametrization that is based on effective matter potentials. This only includes NSI effects through NC interactions, which mainly impact neutrino propagation through coherent forward scattering.

Based on this, different approaches to measuring NSI will be discussed in the following.

¹⁰In general, model independence is retained except for the assumptions mentioned above. Investigating NSI parameters singly however reduces the generality of the approach.

Neutrino detection

The analysis approach taken in this work is based on comparing an expected and observed neutrino flux. The neutrino sources therefore need to be understood sufficiently well and their flux is required to be known with a high accuracy in order to be sensitive to NSI effects. Such data can be obtained from different neutrino sources which are discussed in sections 4.1 and 4.2. Corresponding to these sources, various approaches to neutrino detection are presented in section 4.3. This analysis uses data taken by IceCube DeepCore, an ice Cherenkov detector introduced in section 4.4, where atmospheric neutrino interactions are detected via the Cherenkov light produced by charged secondary particles.

4.1 Neutrino sources

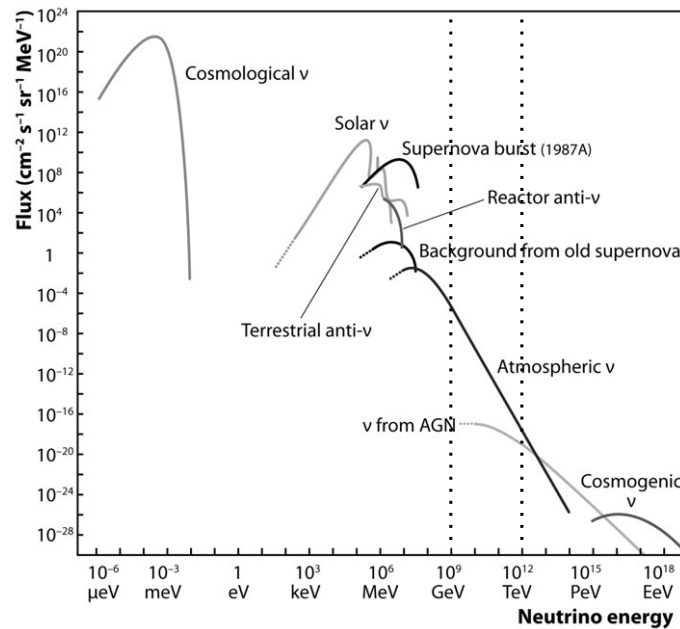


Fig. 4.1.: Flux from dominant neutrino sources over neutrino energy, adapted from [56]. The GeV range, which is most important in this work, is indicated with dotted lines.

Different production mechanisms and signal characteristics of sources are relevant with respect to NSI detection. Due to the L/E dependence of oscillation effects, NSI signal characteristics can be resolved at different energy scales, with oscillation baselines at vastly different orders of magnitude depending on the source.

The flux for a number of relevant neutrino sources is shown in figure 4.1 which showcases the differences in energy spectra. While the flux from these sources decreases, the interaction cross section increases with energy, as shown in figure 2.4. This, as well as the specific flavor compositions of the respective neutrino signal, results in diverse detection mechanisms, depending on the source.

Solar neutrinos

As shown in figure 4.2 solar neutrinos occur at energies of approximately 20 MeV and below, with detection usually being optimized for the higher energy range, due to the larger cross section. They propagate through solar matter and, at night, through Earth matter, at distances up to the full Earth diameter, yielding a large L/E range.

There are two main fusion processes in which solar neutrinos are produced, namely the proton-proton chain (pp chain) and the Carbon-Nitrogen-Oxygen cycle (CNO cycle) [57]. On a basic level, the processes involve proton fusion according to



The resulting signal consists purely of electron neutrinos. However, when reaching the Earth, depending on the observed energy, a majority of these electron neutrinos have oscillated into other flavors. This creates the electron neutrino flux deficit that was a first evidence of neutrino oscillation.

Reactor neutrinos

Nuclear power plants provide high fluxes of MeV-range electron antineutrinos. Their spectrum is a superposition of β -spectra of all present isotopes, with the basic decay mechanism



The resulting energy spectrum is shown in figure 4.3 to lie below ~ 10 MeV. Detection is oftentimes done via inverse β -decay (IBD)



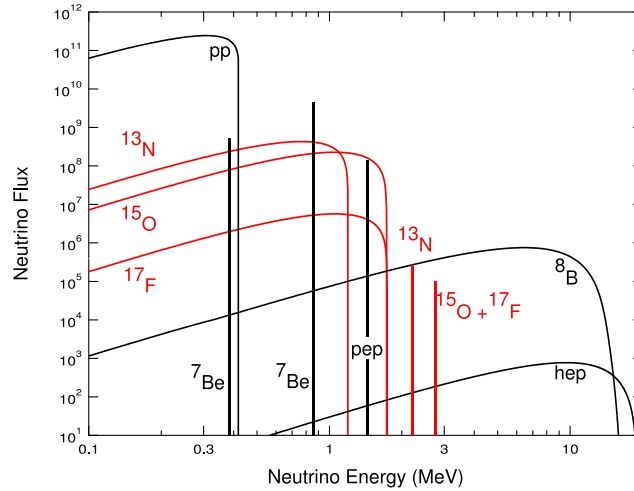


Fig. 4.2.: Predicted fluxes for different production processes of solar neutrinos. Neutrinos produced in the CNO cycle are shown in red, pp chain neutrinos in black. Monoenergetic fluxes are given in $\text{s}^{-1} \text{cm}^{-2}$ and continuous spectra in $\text{MeV}^{-1} \text{s}^{-1} \text{cm}^{-2}$. This figure is taken from [57].

which has a lower detection threshold of 1.8 MeV [1].

The overall flux and its time variations are known well from reactor monitoring. Uncertainties, for example on the energy spectrum, are addressed oftentimes by building dedicated detectors at small distances to the nuclear facilities.

Accelerator neutrinos

Accelerators can produce neutrinos as a secondary beam: Accelerated particles, typically protons at the GeV range, hit a target to produce charged mesons, which in turn decay into neutrinos. The charged mesons can potentially be redirected and selected by particle type, as a means to manipulate the beam of neutrinos that is produced in their decay. Such a selection would for example yield a mostly pure muon (anti-) neutrino beam, through a decay of

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \quad . \quad (4.4)$$

Similarly to reactor neutrinos, the expected neutrino flux and spectrum from accelerators is well known. The energy spectrum of the outgoing neutrinos depends on their emission angle, with the highest energies and energy variations expected on-axis (that is, parallel to the preceding meson beam). This motivates off-axis positioning of neutrino experiments, utilizing the resulting approximately monoenergetic signal [1].

A commonly used categorization of reactor and accelerator neutrino detectors is

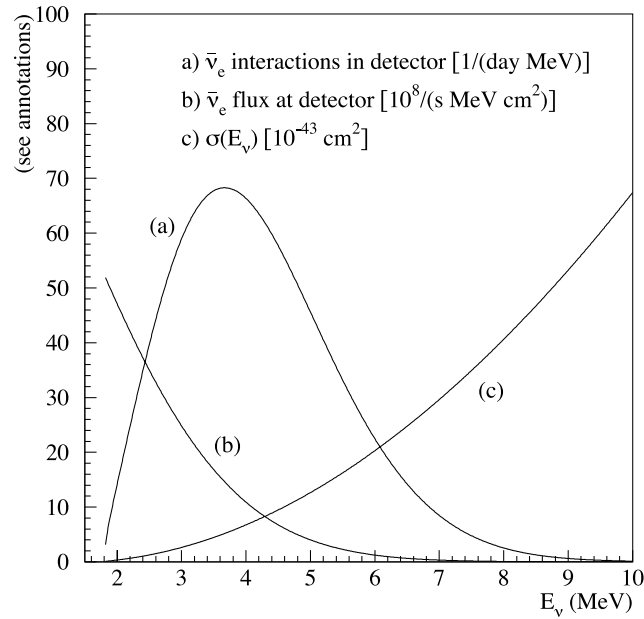


Fig. 4.3.: Interactions of electron antineutrinos in a hypothetical detector (a) based on IBD, with the corresponding cross section (c), for a reactor neutrino flux (b). The different units are given in the annotation. The hypothetical detector has 12 t fiducial mass and is positioned at 0.8 km distance from a reactor with 12 GW thermal power. Figure taken from [1].

into long baseline (LBL) and short baseline (SBL) experiments, with the associated distances differing between sources. For reactor neutrinos, a SBL is at $\mathcal{O}(10\text{ m})$ and a LBL at $\mathcal{O}(10\text{ km})$. The characteristics of accelerator neutrinos motivate to set SBL at 10 m to 1 km and LBL around 1000 km [58].

The NSI limits obtained from this analysis are not directly comparable to most accelerator experiments, as their sensitivity to the mass of a possible new mediator particle makes the results depend strongly on the underlying model [44].

4.2 Atmospheric neutrino production

Atmospheric neutrinos are produced in air showers that cosmic rays induce. In the following, the relevant characteristics of Cosmic Rays are discussed, followed by an overview of the produced neutrinos and background particles.

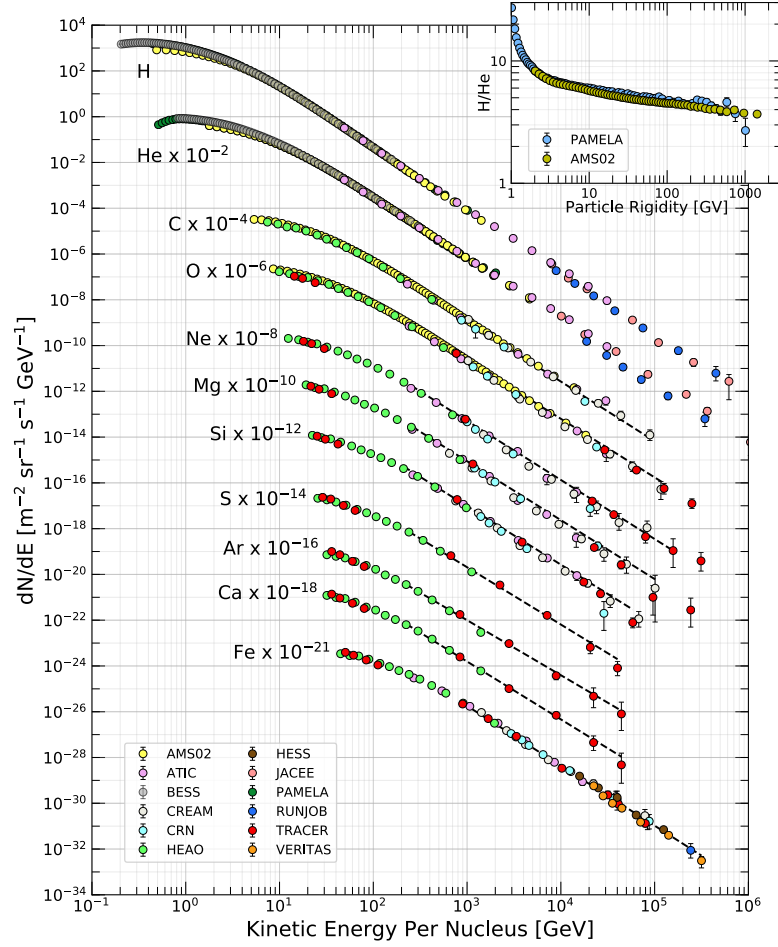


Fig. 4.4.: Fluxes of different cosmic ray (CR) nuclei in number of particles per energy-per-nucleus, plotted over energy-per-nucleus. Figure taken from [8].

Cosmic rays

Ionizing particles hitting the atmosphere are denoted CRs. Their overall energy spectrum can be described through the derivative of the flux ϕ by the energy E as a power law

$$\frac{d\phi}{dE} \propto E^{-\gamma} \quad . \quad (4.5)$$

The spectral index γ amounts to ~ 2.7 [8] within the GeV range. The uncertainty on this parameter will be introduced as a nuisance parameter in section 7.3.1.

Primary CRs, such as protons, are accelerated in astrophysical sources outside the solar system. The details of their production are subject of ongoing research [8]. When these primary CRs interact with interstellar gas, secondary CRs are produced. In figure 4.4, the fluxes of different CR nuclei are shown to decrease over energy. The vast majority are protons or Hydrogen, making up $\sim 90\%$ at 10 GeV.

In interactions of CRs within the atmosphere, cascades of secondary particles are produced including atmospheric neutrinos and muons which can be detected in IceCube [41].

Atmospheric neutrinos

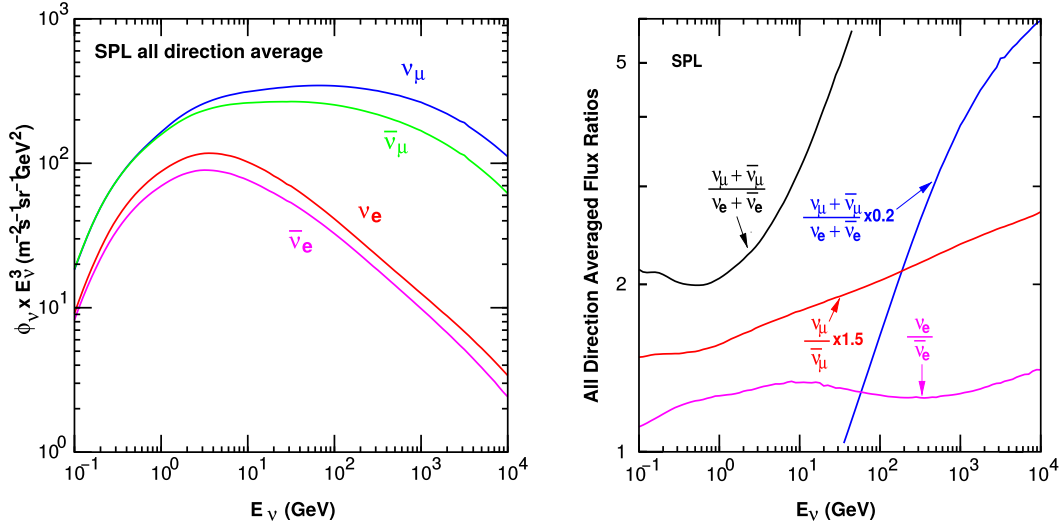


Fig. 4.5.: Fluxes (left) and flux ratios (right) of electron and muon (anti-) neutrinos over neutrino energy as predicted for the South Pole (SPL). Figures taken from [41].

In the range of 1 GeV to 100 GeV, CRs predominantly produce pions, with contributions of kaons. Charged pions primarily decay into neutrinos [8] in processes such as

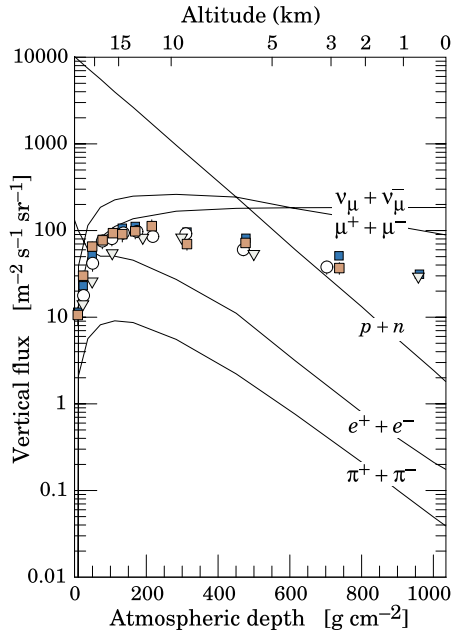
$$p + N \rightarrow \pi^+ + X \quad (4.6)$$

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \quad (4.7)$$

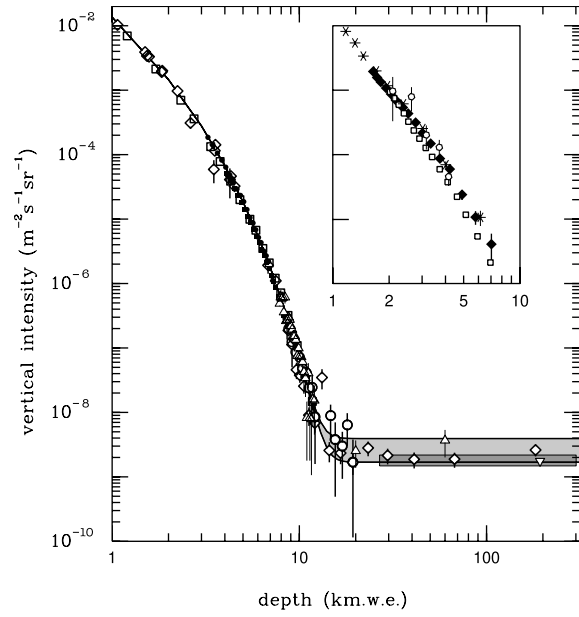
$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu \quad , \quad (4.8)$$

and equivalent for π^- . This implies a ratio between muon (anti-)neutrinos and electron (anti-)neutrinos of 2 : 1.

The details of atmospheric particle production and resulting flux are more complex than what can be covered in this section. This becomes apparent from the vertical neutrino flux, which is shown in figure 4.6a, showcasing the dependence on different CR primary and secondary particles and, as a result, on the atmospheric altitude. Figure 4.5 demonstrates the depth dependence of the muon flux. A detailed description of the resulting features in neutrino production rates can be found in [41].



(a) Vertical fluxes of different particles constituting or produced by CRs at >1 GeV. Markers show μ^- measurements. Primary CR particles are protons and electrons near the top of the atmosphere. Electrons further down are produced in π^0 decays. Neutrinos and muons are produced in charged lepton decays such as the one shown in equation 4.6. Figure taken from [59].



(b) Vertical muon intensity below ground as a function of depth in km w.e.. This includes atmospheric muons as well as those produced below ground level, which predominantly originate from neutrino interactions. Figure taken from [8].

Fig. 4.6.: Characteristics of CR particles and their secondaries with respect to their vertical position within the atmosphere or Earth.

Atmospheric muons

The atmospheric muons produced in reactions like 4.6 are a main background for underground neutrino detectors. At sufficiently high energies (see figure 4.15), their decay is suppressed such that they can penetrate through kilometers water equivalent (w.e.) of material. The resulting decrease in flux underground is shown in figure 4.6b. This illustrates that muons are not expected to enter underground experiments at directions substantially different from vertically down-going.

4.3 Detection principles

Depending on the neutrino source, there is a variety of detection techniques, all of which measure neutrinos indirectly through their secondaries. While this work focuses on IceCube, a basic understanding of other approaches to detecting neutrinos that are sensitive to NSI is necessary for comparing results between experiments in section 7.8. Although approaches differ fundamentally there are similarities between all neutrino detection methods.

Direct detection is not feasible, as neutrinos are only known to interact weakly and marginally gravitationally (see section 2.2). Instead, secondary particles produced in neutrino interactions are observed. Cherenkov and scintillator based experiments measure the light that charged secondary particles deposit within the detection medium while traversing it. Other than that, there is a variety of different strategies, depending on the neutrino flavor and energy, and hence their interaction cross section, spanning from calorimetric to radio detection and time projection chamber (TPC) approaches.

One more shared trait is that all neutrino detectors mentioned here require shielding or veto as signatures of neutrino interaction products can be mimicked by a variety of background sources, such as cosmic muons. Detector sites are hence commonly chosen to be underground, below kilometers w.e. of natural shielding material. Active veto is commonly provided using instrumented detector regions or specifically designed veto detectors around the designated fiducial detector volume.

Approaches relevant for NSI detection can be subdivided into three main groups [19]:

Radiochemical

In the radiochemical approach neutrinos interact with atoms of the detection medium, producing instable nuclei which are separated and counted when decaying. This only serves as a counting experiment, as the detection principle doesn't conserve properties of the neutrino such as energy or incidence angle. There are two elements that are commonly used, the first being gallium, which forms germanium through



with a threshold neutrino energy of 0.233 MeV. Examples of this are the Gallium Experiment/Gallium Neutrino Observatory (GALLEX/GNO) and the Soviet–American Gallium Experiment (SAGE) [19]. The second commonly used element is ${}^{37}\text{Cl}$, similarly forming ${}^{37}\text{Ar}$ at neutrino energies above 0.814 MeV. This was employed

in the Homestake chlorine experiment, providing the first hints towards the solar neutrino problem and hence towards matter effects [18].

Scintillator

Neutrino detection using scintillators is based on charged secondary particles produced in neutrino interactions. In the scintillating medium, the ionization energy transferred from the moving charged particle gets isotropically emitted as photons, which can be detected e.g. using photomultiplier tubes (PMTs). This has the potential to yield information on neutrino energy and direction - to the extent to which these are preserved in its interaction. The charged particle's light emission spectrum and threshold energy depend on the specific scintillator, with the latter commonly being dictated by background event rates.

Experiment designs vary strongly, depending on the usage of liquid or plastic scintillators and whether these are used as detection material only or double as neutrino target [19]. Examples for the latter are the Boron Experiment (Borexino), detecting solar neutrinos, and the reactor neutrino experiments Kamioka Liquid Scintillator Antineutrino Detector (KamLAND), Double-Chooz, Daya Bay and the Reactor Experiment for Neutrino Oscillation (RENO). All of these constitute realizations of a large liquid scintillator volume, surrounded by PMTs. The $\text{NO}\nu\text{A}$ experiment detects accelerator neutrinos using a segmented active volume, where light from single cells filled with liquid scintillator is detected using avalanche photodiodes.

An example for a fundamentally different scintillator-based neutrino detector is the Main Injector Neutrino Oscillation Search (MINOS), a sampling calorimeter built from alternating layers of steel and plastic scintillator, detecting accelerator and atmospheric muon neutrinos.

Cherenkov detectors

Similarly to scintillation based detectors, Cherenkov detectors observe charged secondaries from neutrino interactions with a target material. Relevant interactions are described in more detail in section 2.2.

The Cherenkov effect describes light emission caused by such a charged particle when traveling through a dielectric medium at a velocity v larger than the phase velocity of light within the medium $c' = \frac{c}{n}$, where c is the speed of light in vacuum and n the refractive index of the medium. As the charged particle propagates, it momentarily polarizes the medium, which subsequently emits light as the molecules return to a randomized polarization. For particle velocities sufficing the condition

$v > c'$ the produced light interferes constructively along a cone-shaped front which is aligned with the particle direction [60]. A schematic overview of the mechanism is shown in figure 4.7, with the opening angle θ_C following

$$\cos(\theta_C) = \frac{c'}{v} \approx \frac{1}{n} \quad , \quad (4.10)$$

with the approximation valid for $v \approx c$.

This introduces a dependence of the lower light production threshold E_{min} of the charged particle on its rest mass m_0 and the detection medium refractive index:

$$E_{min} \approx m_0 \frac{1}{\sqrt{1 - \left(\frac{c'}{c}\right)^2}} = \frac{m_0}{\sqrt{1 - \frac{1}{n^2}}} \quad . \quad (4.11)$$

In general, the choice of medium for a neutrino Cherenkov detector is strongly based on its availability as well as its optical properties, such as the absorption coefficient and scattering length, which predefine the potential efficiency of a light detection based experiment. Wavelength dependence of these properties needs to be considered with respect to the Cherenkov spectrum, the intensity of which increases towards smaller wavelengths [60]. The Cherenkov detectors most relevant for this work rely on water or ice. Heavy water, which is used for instance in SNO, is a less common alternative, with the advantage that the additional neutron available therein results in a larger number of neutrino interactions with nuclei of the target medium.

The lower particle energy threshold for emitting Cherenkov light in water or ice is in the MeV range for electrons and muons. Higher energies imply more light deposition within the target material. This allows for detection of neutrinos from all kinds of sources at different energy scales, mostly depending on the detector's size and, in order to resolve lower energy events, its instrumentation density.

At the MeV scale, the distance between optical units needs to be at the order of 1 m or below. Considering cost per module as well as mean free photon paths in the available detection media leads to experiment designs where the walls of a compact volume are instrumented with inwards facing modules. Such a setup, found for example in the Super-K experiment, enables detailed resolution of the Cherenkov cone, which drives the interaction reconstruction [61]. Super-K is consequently able to reconstruct event properties at neutrino energies of few MeV, including solar, atmospheric and accelerator neutrinos.

The size and shape of the photon pattern as well as the number of photons produced in an interaction depends on the energy it deposits. Hence a detector designed for larger event energies can be instrumented more sparsely.

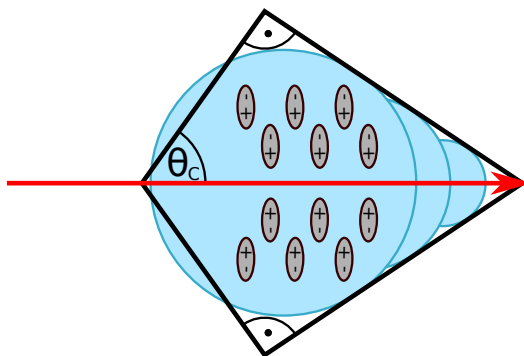


Fig. 4.7.: Schematic of the Cherenkov effect: The red arrow denotes a negatively charged particle traveling at a velocity larger than c' which polarizes the medium surrounding it, inducing light to be emitted and constructively interfere along a cone with the Cherenkov angle θ_C .

Secondary muons produced in neutrino interactions have the capacity to penetrate through matter at large lengths. Below ~ 100 GeV, the Cherenkov light output along the muon track is mostly constant and propagation lengths range up to $\mathcal{O}(100$ m). Estimating the energy deposited in such an event therefore requires kilometer-sized detectors that can resolve the full track length. Above the GeV range, significant fractions of the total energy are deposited within stochastic losses along the muon track, as can be seen in figure 4.15 and is explained further in section 4.4.4. The dependence of the loss intensity and number on the muon energy provides an additional measure of the total energy.

These event signature characteristics inspire a different experiment design for high energy neutrino detectors, using a target material that is available in high quantities and at low cost. Examples for this are ice and water, which are used in IceCube [62] and the Astronomy with a Neutrino Telescope and Abyss environmental Research project (ANTARES) [63], respectively. For these, large volumes are instrumented in a three-dimensional grid with inter-module distances at the order of 10 m.

This enables neutrino detection far beyond the TeV range [64], up to which oscillation phenomena can be observed. At these energies, the Earth becomes increasingly opaque to neutrinos, impacting the flux of neutrinos traversing it significantly [65].

4.4 Neutrino detection in IceCube and DeepCore

IceCube [62], together with DeepCore [66], is an example for such a Cherenkov detector, utilizing a cubic kilometer of ice at the South Pole as detection medium. Neutrino interactions within the ice produce secondary charged particles which in turn produce Cherenkov light. The optical properties of the instrumented ice,

which are discussed in 4.4.3, allow for detection of this light with little loss due to absorption and at correspondingly large distances. Ice with these properties is uniquely available in Antarctic glaciers [67], motivating the location of the experiment at the South Pole. As IceCube consists of a three-dimensional array of 5160 individual light detection modules, neutrino interactions (called events) are observed as extended three-dimensional patterns of photon detections. These need extensive triggering and processing routines in order to extract the relevant event data.

Detector geometry and individual modules are discussed in sections 4.4.2 and 4.4.1, respectively. An overview of the signal processing is given in section 4.4.6, with details on event and noise characteristics presented in sections 4.4.4 and 4.4.5.

4.4.1 The Digital Optical Module

The basic detection unit of the IceCube grid is the digital optical module (DOM), a pressure-resistant spherical detector designed around a single PMT. As shown in figure 4.8, this PMT faces downwards, in order to increase the sensitivity to upwards traveling photons, which are less likely to be caused by background [62].

The working principle of the PMT is shown in figure 4.9: Incident photons cause the emission of electrons from the cathode through the photoelectric effect. These electrons are amplified iteratively via multiple dynodes and ultimately read-out at the anode. The outgoing signal consequently has a time-dependent shape conserving the photon arrival time and an amplitude that depends approximately linearly¹ on the number of coincident photons [69].

Standard IceCube modules contain 252 mm R7081-02 Hamamatsu Photonics PMTs, which are responsive to light at 300 nm to 650 nm with a quantum efficiency (QE) of 25 % at 390 nm². The wavelength-dependent photon acceptance of a full IceCube DOM, as is shown in figure 4.10, is significantly below these 25 %, due to the module size as well as the optical properties of the pressure vessel and optical gel. Note that in order to detect Cherenkov radiation, the acceptance in the lower wavelengths that the DOM covers is of the highest importance.

Inside a DOM, as shown in figure 4.8, the PMT is shielded from magnetic field of the Earth by a Mu-metal grid. Dedicated boards provide communication, signal processing and transmission structures. Optical calibration across modules is done using LEDs installed on the so-called flasher board. In dedicated calibration detector runs,

¹The assumption of linearity can be made for the event energy range considered in this work [68].

²The QE of a PMT characterizes the number of electrons released from the cathode per incident photon.

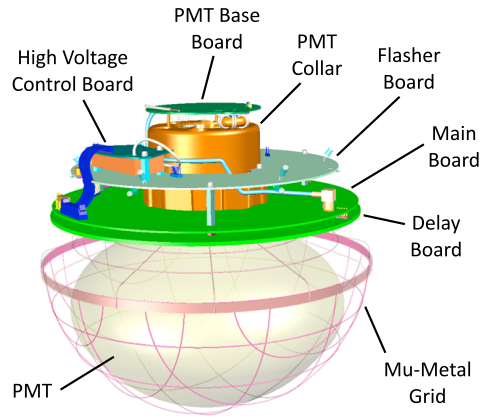


Fig. 4.8.: Mechanical layout of an IceCube DOM, showing the PMT as its core component as well as electronics for calibration, signal processing, communication and high voltage (HV). Figure taken from [62].

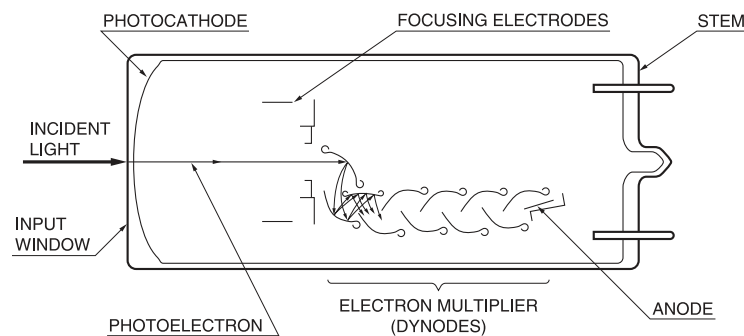


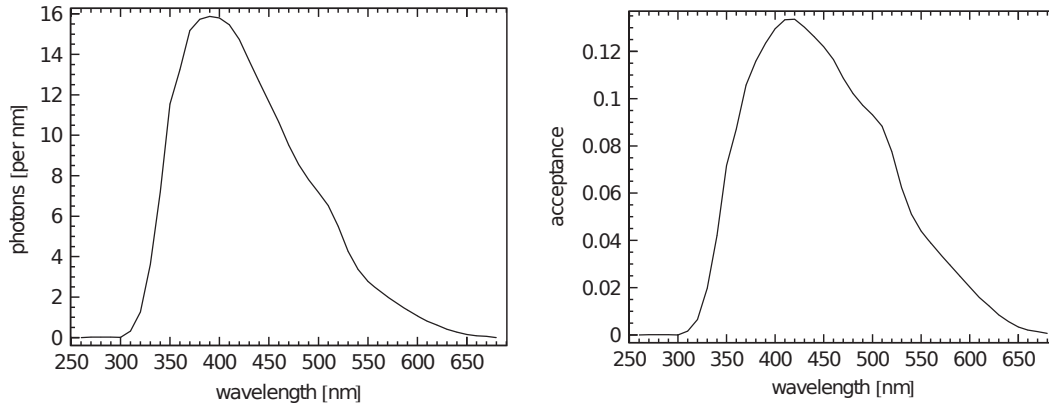
Fig. 4.9.: Schematic of the functioning principle of a PMT, taken from [69].

these LEDs are pulsed and the resulting light curve is measured by all surrounding DOMs.

Each DOM is connected to the surface for power and data transmission and is encapsulated in a spherical glass housing with a diameter of 330 mm and thickness of 13 mm to withstand the pressure of the ice around it. The overall detector capacity to observe neutrino interactions in specific energy ranges, however, does depend as much on the individual modules as on their relative distances and orientation, which is described in the following section.

4.4.2 The detector geometry

The 5160 DOMs that constitute IceCube and DeepCore are deployed within a cubic kilometer of antarctic ice as shown in figure 4.11. They are arranged along 86 cables, the so-called strings, at depths of approximately 1450 m to 2450 m, with the ice above providing shielding from particles other than neutrinos. At the surface all



- (a) Number of simulated Cherenkov photons emitted along one meter of muon track, not including any secondary cascades, for an IceCube DOM with the acceptance given in 4.10b.
- (b) Wavelength-dependent photon acceptance of an IceCube DOM with respect to the overall module size, including PMT properties as well as the impact of the DOM glass and optical gel.

Fig. 4.10.: Wavelength-dependent photon detection characteristics of an IceCube DOM. In comparison to the respective characteristics of the PMT only, the largest difference is a geometrical factor, as the shown DOM properties include all photons arriving at the DOM surface. Both figures are taken from [67].

strings terminate inside the IceCube laboratory (ICL), a facility where all necessary infrastructure, such as power supplies and data storage, but also the hardware used for online processing (see section 4.4.6), is located.

The original IceCube detector layout is a hexagonal grid with horizontal distances of ~ 125 m between strings and vertical distances of ~ 17 m between modules. Its overall geometry is designed for imaging light patterns from neutrino events with energies of $E_\nu \geq 100$ GeV.

IceTop is a surface extension of IceCube that was designed as an array of 81 stations which each hold four DOMs within two water tanks [62]. While being optimized for detection of air showers, IceTop signals can also be used as a veto for IceCube.

DeepCore is the more densely instrumented region in the bottom center of IceCube and was completed in 2011 [66]. Its horizontal and vertical instrumentation density exceeds that of the rest of IceCube, due to eight infill strings with reduced vertical inter-DOM distances of ~ 7 m. A subset of 50 DOMs deployed along these strings are high quantum efficiency (HQE) modules. These differ from standard IceCube DOMs only in the type of PMT, R7081-MOD, which provides approximately 40 % higher QE at 390 nm [66].

Deploying DeepCore at the center of IceCube allows for using the surrounding strings as veto, while the choice of vertical position at depths of 2100 m to 2450 m is motivated from the optical properties of the ice, which are discussed in section 4.4.3. This situates DeepCore below the so-called dust layer where impurities from the

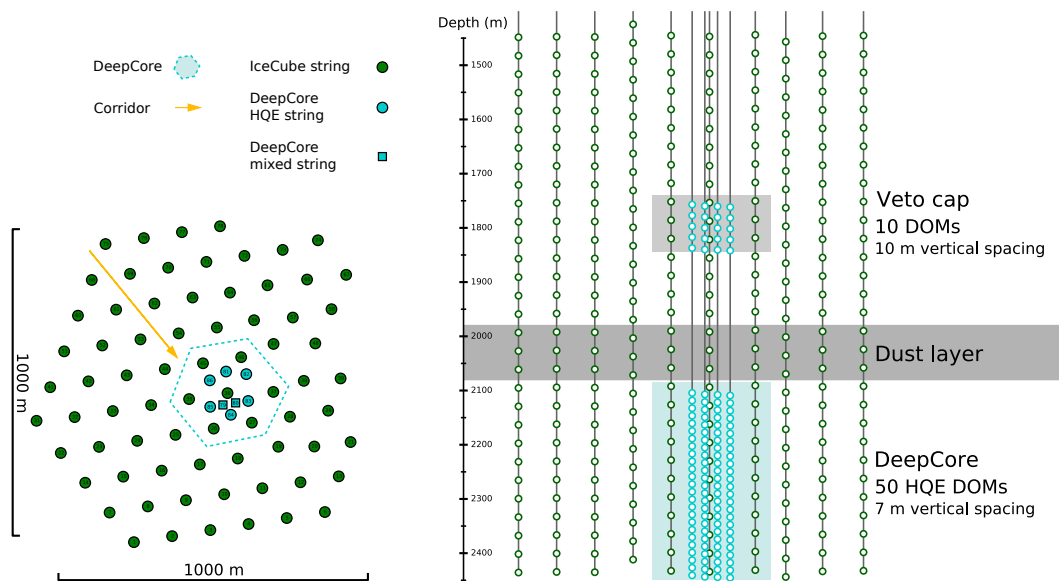


Fig. 4.11.: Schematic of the IceCube neutrino observatory, including the DeepCore sub-array. The DeepCore volume and strings are shown in blue, with the DeepCore veto layer indicated in light gray. The orange arrow shows a possibility for muon tracks to enter the detector without passing close to a string through a so-called corridor. The dust layer, indicated in dark gray, is shown in figure 4.12 to impact the optical properties of the ice. This figure is adapted from [70].

atmosphere are enclosed within the ice and impact its optical properties throughout an extended vertical range [66]. Above the dust layer, additional IceCube DOMs are deployed along the DeepCore strings to act as a veto cap for the DeepCore volume below.

In the coming years, seven additional strings with multi-PMT-modules will be deployed within the DeepCore region, constituting the IceCube-Upgrade. The more dense instrumentation as well as photon coincidences between PMTs of the same module will allow for individually³ detecting interactions of neutrinos with energies of 1 GeV and increase the detection rate throughout the GeV range. This is expected to significantly improve the sensitivity to oscillation effects [72].

4.4.3 Ice as a detection medium for Cherenkov light

The capacity of in-ice detectors to detect Cherenkov light from the mentioned signatures strongly depends not only on the deployed detector hardware detailed above, but also on the properties of the ice. These are unique in the 2.8 km thick Antarctic

³Lower energy events can not be distinguished from random background. Large numbers of coincident low energy events, which are for example produced by Supernovae, can be detected as an excess over the expected background [71].

glacier, which provides a large volume of potential detection material in a setting of extremely low background [67].

Measuring and modeling these optical properties in detail is a crucial and ongoing endeavor. The relevant quantities are the effective scattering and absorption coefficient (or respective length), which are shown in figure 4.12 as they are implemented in the model used for this work.

Pristine glacial ice

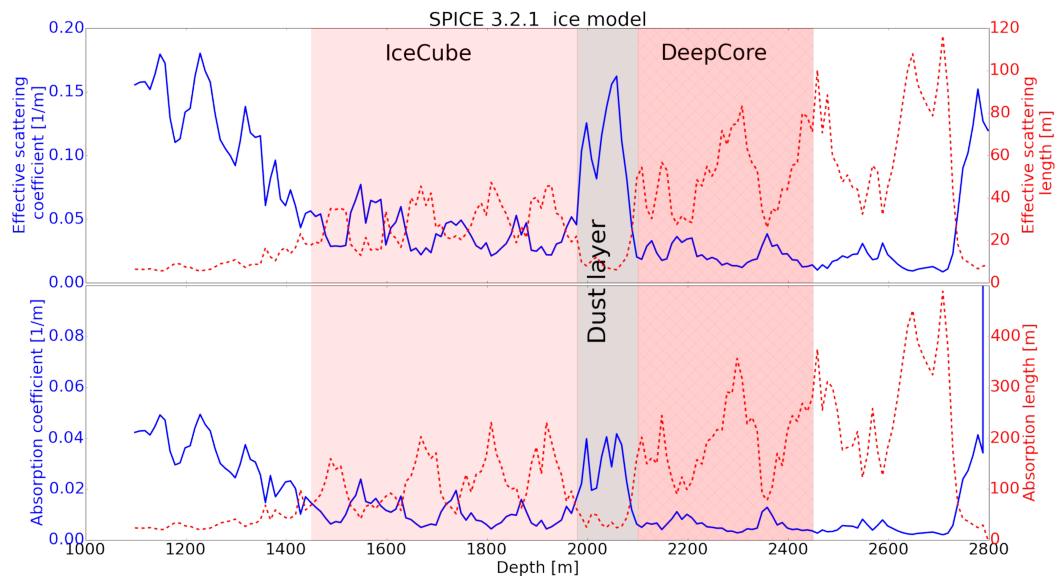
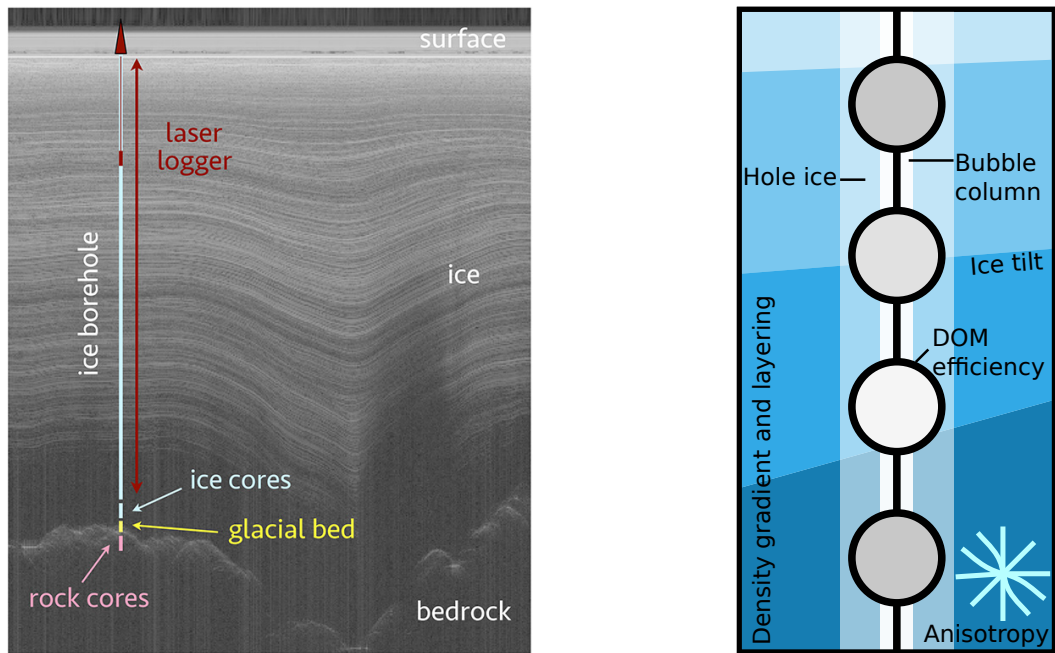


Fig. 4.12.: Ice scattering and absorption coefficients (left axis, in blue) and lengths (right axis, in red) at 400 nm for the entire vertical detector range, as implemented in current ice models [73]. Detector sections are indicated as background color. Taken from [74].

At depths where IceCube is deployed, the ice is approximately 40000 – 100000 years old [74]. Airborne dust particles that were present in the atmosphere at the time of deposition of the ice result in distinct layers. Consequently, the optical ice properties primarily vary over depth. The closer these layers are to the bedrock underneath, the more they adapt to its shape, resulting in vertical variations and an overall slope called the ice tilt [67], as can be seen in figure 4.13a.

A second reason for the mainly vertical variations of the ice optical properties is their dependence on pressure. The ice forms over time from snow accumulating in layers, with initially extremely short scattering lengths due to the contained air. At depths where IceCube is deployed, the pressure and temperature suffices for enclosed air bubbles to disintegrate: Their molecules get incorporated into the ice, forming so-called craigite, the refractive index of which differs negligibly from that



(a) Radar measurement of the Antarctic ice sheet, showing how the deeper ice layers adapt more to the contour of the bedrock [75].

(b) Schematic overview of ice properties around IceCube DOMs.

Fig. 4.13.: The Antarctic glacier ice has complex properties relevant for the performance of IceCube.

of ice [74]. As the difference in refraction index between air and ice otherwise is a main factor in scattering, effective scattering lengths reach up to 100 m once this effect takes place in the depth range starting few hundreds of meters above IceCube, as shown in figure 4.12.

Further complication in the optical properties of the ice is caused by the drift of the Antarctic glacier towards the coast at a velocity of approximately 10 m per year. This flow impacts the crystalline structure and its birefringent optical properties, which ultimately causes anisotropic light attenuation [76]. A rigorous study of ice properties and the respective detector calibration can be found in [74].

Ice properties after detector deployment

Unlike the pristine so-called bulk ice, the ice within drill holes has been re-frozen recently and abruptly, forming what is called hole ice. This process does not preserve the ice layering and changes the characteristics of air inclusions, which do not collapse under the ambient pressure. Re-freezing of the hole ice after string deployment starts at the boundary to the still frozen surrounding bulk ice. It is assumed that consequentially, air bubbles are pushed towards its center. They form the

so-called bubble column [67], the shape of which induces an angular dependence of the scattering probability. Modeling of the re-frozen ice as a major measurement uncertainty will be discussed in section 7.4.

Combining ice and hardware

To characterize detection capacities of a Cherenkov detector within the Antarctic glacier, properties of ice and detection hardware are treated jointly. As a result, angular module acceptance includes not only the DOM hardware and setup but also the local ice properties, which will be discussed in more detail in section 7.4.

Other parameters need to be attuned between detector and ice. The wavelength-dependence of the photon detection probability is impacted by the ice absorption as well as the PMT acceptance shown in figure 4.10, which are both optimal around 400 nm [67], a range that also contributes to the Cherenkov spectrum.

4.4.4 Event characteristics

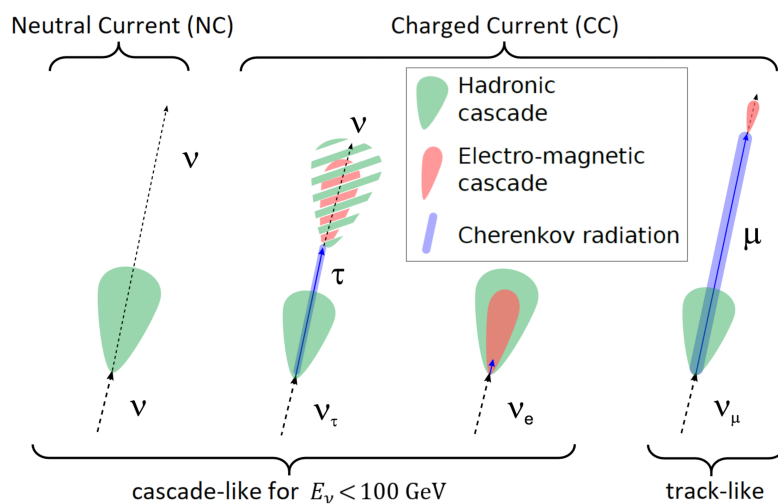


Fig. 4.14.: Photon signatures of neutrino events, with the interaction type indicated above and the topology as seen by IceCube indicated below. Modified from [77].

Neutrino events leave different photon patterns within IceCube, depending on the neutrino flavor and interaction type. The potential of a detector of distinguishing between these decreases with event energy, which implies less overall light output as well as smaller propagation distances of charged particles produced in the interaction. In figure 4.14, event signatures are shown for interactions relevant at the GeV range that were discussed in section 2.2.

In NC interactions of all neutrino flavors, a hadronic cascade is produced, resulting in an approximately spherical light deposition that is denoted as cascade-like. In contrast, signatures of CC interactions differ between flavors. The basic principle for all neutrino CC interactions is the same, being the production of an initial hadronic cascade together with a lepton of the same flavor as the neutrino. The lepton induces track-like light emission before being captured or decaying and producing a second cascade. The decay lengths of the leptons, however, differ significantly depending on the flavor and event energy, which results in more or less extended signatures. In tau neutrino CC interactions, the secondary tau that is produced along with the initial hadronic cascade decays after $\tau_{\text{life}}^{\tau} = 0.29$ ps [8]. At the GeV range considered in this work, this implies that the tau travels a mean distance of $\mathcal{O}(\text{cm})$ at most before decaying, which can not be resolved in DeepCore. Consequently, the overall interaction signature is cascade-like, consisting of the initial cascade, the light emitted along the tau track and the terminal hadronic or EM cascade, depending on the decay channel of the specific tau. As these tau decay channels potentially include secondary neutrinos, not all of the initial tau neutrino energy is necessarily deposited within the detector. The decay channels also include the possibility of producing charged leptons. As these carry only a fraction of the initial tau neutrino energy, they are not expected to induce observable tracks for event energies in the GeV range [78].

In electron neutrino CC interactions, the secondary electron is captured at a short distance [79], leading to overlap of the initial and terminal cascade and an overall cascade-like topology.

In contrast, the secondary muon produced in a muon neutrino CC interaction deposits a large fraction of its energy not in its terminal EM cascade but along an elongated path through the ice, constituting a so-called track-like signature. When assuming the energy loss of minimum ionizing muons, which is 0.22 GeV/m, as shown in figure 4.15, track lengths are approximately 4.5 m per GeV. This assumption is also made for the number of photons detected in an IceCube DOM per meter of muon track that is shown in figure 4.10a. A more realistic scenario, especially for the higher energies considered in this work, would also need to include secondary cascades that occur along the track, which can locally increase the energy loss.

As a result, in the GeV range, DeepCore can differentiate between muon neutrino CC interactions and all other types of interactions and specifically not between neutrinos or anti-neutrinos of any flavor [66]. Especially at low energies, the distinction between track-like and cascade-like events becomes error-prone due to the small track lengths as well as the low number of detected photons (see section

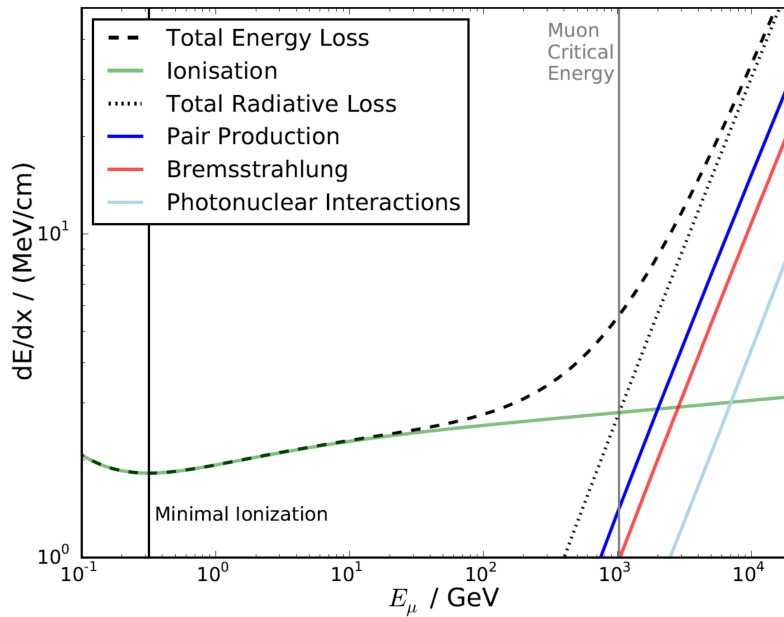


Fig. 4.15.: Average energy losses per distance as a function of the kinetic energy of muons in ice. Below 100 GeV, the behavior can be described well as that of a minimum ionizing particle. Figure taken from [80], based on [81].

5). Track-like events can therefore be selected more accurately than cascade-like events, which makes DeepCore data sets specifically sensitive to deviations from the expected muon neutrino CC event rate.

4.4.5 Background and noise

There are different sources of unwanted signals, originating from two basic categories: Firstly, background events, which are actual interactions producing Cherenkov light, but caused by something else than neutrinos. Secondly, charge or light produced randomly. Characteristics of these differ depending on their source, requiring different rejection methods.

Background events

Background events commonly originate from charged particles that have not been produced in neutrino interactions. While the approximately 1.5 km of ice above the detector effectively shield it from most such particles produced on or above the surface, e.g. from cosmic or atmospheric sources, muons with sufficiently high energies are able to enter the detector.

Using the fact that signals of neutrino-induced events start within the ice, this kind of background can be rejected by vetoing any events that start outside or close to the border of the detector. This needs to take into account the detector geometry, as it provides un-instrumented corridors where charged particles have a relevant chance to pass through without the light they emit being detected. Such a corridor is indicated in figure 4.11.

Random noise

A main random component in IceCube is dark noise, which describes a PMT observing a signal that was not caused by a photon but for example from thermal emission of electrons from the cathode. The dark noise rate depends on the cathode area and quantum efficiency of the PMT as well as the module temperature and the time since a module was deployed.

A second component are random bursts of correlated photons which originate from radioactivity of the glass pressure sphere [82]. Overall random noise rates in IceCube and DeepCore are on average 510 Hz and 590 Hz, respectively [62].

This kind of noise has a rate-dependent probability of passing the coincidence-based trigger criteria described in section 4.4.6, but can be reduced with more informed cuts that are given in section 5.2.

4.4.6 Data acquisition and on-site signal processing

Photons detected by a PMT are initially available as a voltage waveforms. Digitizing them, extracting information on the photon hit and combining single hits based on their origin - or rejecting hits that are not caused by neutrinos - is done through extensive data triggering and processing.

In-ice triggering

Once photons hit the cathode of a PMT, the produced photoelectrons are amplified by the dynodes shown in figure 4.9. The output signal is a time-dependent voltage curve which is captured if the observed voltage exceeds a threshold value. The total charge produced within the PMT upon photon arrival can be determined from integrating over the signal. It is interpreted as the number of initially released photoelectrons and therefore commonly given in units of photoelectrons (PE)⁴ [62].

⁴PE are calibrated per DOM via the PMT operating voltage.

The above mentioned threshold is typically set to a value corresponding to 0.25 PE. For each threshold-exceeding waveform, a total of $6.4\ \mu\text{s}$ are digitized coarsely binned in a Fast Analog-to-Digital Converter (FADC). In parallel, the first 427 ns of a threshold-exceeding waveform are digitized using Analog Transient Waveform Digitizers (ATWDs) [62]. The ATWD information of a waveform is sent to the ICL only if coincidence criteria are met. These are based on the so-called hard local coincidence (HLC) criterion which requires that two or more DOMs that are direct or next-nearest neighbors on the same string register hits within a time window of $\pm 1\ \mu\text{s}$. Any other signal passing the module triggering threshold is classified as soft local coincidence (SLC) hit.

A simple multiplicity trigger (SMT)- N requires N HLC hits within a certain time window. In case of DeepCore, that is an SMT-3 trigger applied to DOMs within the DeepCore fiducial volume⁵, with a time window of $2.5\ \mu\text{s}$. If trigger conditions are met, all hit data is combined into an event within a time window of $-4\ \mu\text{s}$ to $6\ \mu\text{s}$ ⁶ with respect to the triggering HLC hit [66, 77]. The trigger frequency for DeepCore is approximately 250 Hz and dominated by multiple magnitudes by muon background (see section 5).

Processing and filtering within the ICL

Within the ICL, meta-information such as the calibration and status of all individual DOMs for the respective run⁷ is used to extract information on single photons from the waveforms, such as photon arrival time and charge. The latter acts as a means to determine whether coincident photons were observed in the waveform. These are stored per event in a so-called pulse map.

The pulse maps are run through so-called hit-cleaning to reduce unphysical SLC hits. This is based on Radius-Time-cleaning (RT-cleaning), which is commonly seeded with HLC hits. Hits that are outside a 125 m radius and 500 ns time range of other hits are in this way rejected in an iterative approach [77].

Based on this cleaned pulse map, some event properties like vertex position and direction are obtained from a basic and fast reconstruction, since they are needed for event filtering. The DeepCore filter is designed to reject triggered dark noise and muon background and uses DOMs outside the DeepCore fiducial volume as veto. The filter rejects hits in veto DOMs if their vertex and time correspond to a positive velocity close to c with respect to the center of all observed hits. The reasoning behind this velocity cut is that such a behavior would be consistent with observing

⁵The volume indicated in figure 4.11, below 2100 m.

⁶Time windows can be extended if additional hits are triggered.

⁷A run commonly has a duration of 8 h.

atmospheric muons.

The rate of events passing the DeepCore filter amounts to approximately 15 Hz [83], which is low enough for data to be transferred north for further so-called offline filtering. For this, the pulse maps are merged into one file per run, together with the available meta-information and run data, which then is compressed. Based on these files, further data selection is done as described in the following section.

Simulation and data sample

In order to be sensitive to NSI effects on DeepCore data, a preferably large neutrino event sample with low background contamination is needed. This is achieved through nine stages of processing and filtering of pulses and events.

This analysis is performed by comparing data to expected event counts for different physics hypotheses. Consequently, large Monte Carlo (MC) simulated sets that capture all relevant data features are needed. In the following, the multiple steps of the simulation process will be introduced. Subsequently, the individual processing levels that are applied equally to data and MC sets are described.

There currently are three sub-samples based on all DeepCore data available to date. The sets differ as they either focus on adopting well understood techniques or on maximizing the event statistics. The latter is done in two sets that use different reconstruction techniques. For this analysis, a high statistics sample that uses convolutional neural network (CNN) based reconstruction is used.

5.1 Simulation

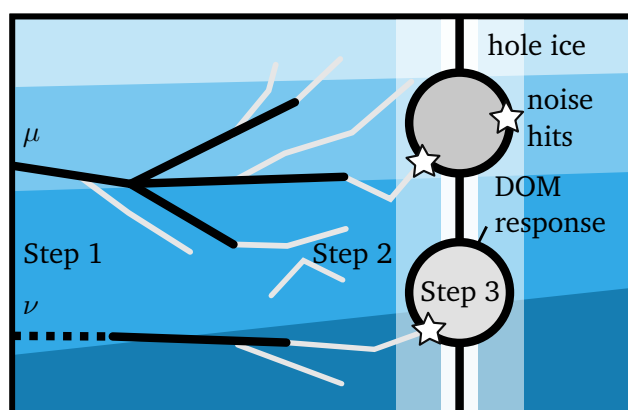


Fig. 5.1.: Schematic overview of the three steps of DeepCore event simulation. In step 1, particles are generated and propagated. Signal photons and their path through the ice are simulated in step 2. The detector response in step 3 applies to the mentioned photons as well as random noise within the modules.

The simulated pseudo-data sets used for DeepCore oscillations analyses are required to include signal as well as background and noise at sufficient statistics. Conformity with data in all relevant characteristics is needed as well, as the simulated sets are processed using the same triggering and filtering steps that are used for data. This agreement is mostly achieved directly in simulation, with the exception of some event categories that are poorly modeled there. These are rejected in the course of the subsequent event selection.

The simulation process described here is the same for the baseline set, which includes all nuisance parameters at their nominal value, and the off-nominal sets with varied nuisance parameters. More detail on their usage and parameter settings can be found in section 7.4.3.

Simulation of DeepCore data is split into the three steps shown in figure 5.1. Firstly particle production and propagation, secondly photon production and propagation and thirdly detector response. The first steps depend on whether neutrinos, atmospheric muons or random noise are simulated.

The respective generated live-times differ as well, due to different requirements and computational cost. For neutrinos, ~ 70 years of detection time are simulated in the baseline set, corresponding to approximately seven times that of the actual data sample used for this analysis.

Due to the high frequency at which atmospheric muons are detected, for these, simulating similarly long detection time would not be computationally feasible and therefore, only the equivalent to the live-time of the data sample is simulated. As atmospheric muons only contribute to the final sample at below 1%, statistical uncertainties originating from this reduction has negligible effect on analyses.

The homogeneous characteristics of random noise allow for a sample size with a shorter live-time of approximately one month.

Step 1: Particle generation, propagation and interactions

In a first step, atmospheric muons and neutrinos are generated and propagated, including production of secondary particles and energy depositions. While this is necessary for signal and background events, random noise hits are produced directly at detector response level in step 3.

Neutrinos are created isotropically on a cylinder surrounding the center of DeepCore and set to interact within the DeepCore instrumented volume. The cylinder size varies with neutrino energy such that any interactions that have a non-negligible probability of producing a signal passing the DeepCore trigger are included.

Neutrino cross sections are pretabulated and splined based calculations done with

the GENIE MC generator [84] (version 2.12.8). The ensuing detector response is simulated throughout IceCube. Neutrino energies are generated up to 10 TeV, following a $E^{-2.0}$ power law spectrum. They are afterwards weighted to match the expected atmospheric flux energy spectrum [41].

In order to keep computation time low, atmospheric muons are simulated at different degrees of detail, depending on whether they reach a DeepCore-centered cylinder with 500 m radius and 150 m length. This uses, directly or through pretabulated values, a detailed simulation of cosmic rays and associated particle showers using CORSIKA [85] and the SIBYLL2 [86, 87] model for atmospheric interactions. The tool providing tabulation and interfaces is the IceCube-internal MuonGun framework. Propagation of muons and taus, including stochastic energy losses, is done using the PROPOSAL [88] tool, which includes the local ice properties.

Step 2: Light production

Proceeding from the particles produced in step 1, Cherenkov emission is calculated using GEANT4 [89]. Due to the higher variance in particle showers at lower energies, full GEANT4 is evaluated below 30 GeV and 100 MeV for hadronic and electromagnetic showers, respectively. Above these energies, GEANT4 based shower parametrizations are used [90].

Subsequently, the produced photons are propagated through the ice until they are either absorbed or detected. This is done using the IceCube-specific OpenCL based CLSim [91] software.

Step 3: Detection

For any photons that are observed by a DOM through photoelectrons, the module response needs to be simulated. Properties of the DOM that enter this step are the PMT response and module electronics, readout and triggering. In addition, properties of the pressure vessel and the ice column directly surrounding the strings are included, a description of which can be found in section 4.4.3. Ice characteristics are measured through flasher calibration data and are treated together with module properties such as the intrinsic angular acceptance [76].

In addition to photons produced in step 2, module response to random noise needs to be simulated. This kind of noise is generated as Monte Carlo photoelectrons (MCPEs) using the VUVUZELA tool [82]. The photoelectrons are added to existing events or simulated separately for pure noise events.

Dark noise, such as thermal emission at the PMT, occurs at random times and

uniformly distributed throughout the event time window. An contribution of low energy muons that are not individually resolved is included in this as well. Contrarily, correlated random noise is caused by photon bursts from radioactive decays in the glass pressure sphere. This is simulated as a Poisson-drawn number of MCPE with hit times following a log-Gaussian distribution around a random emission.

All registered pulses simulated in the aforementioned steps are subsequently processed using the same steps that are used for data. This includes online triggers and filters, which are discussed in section 4.4.6, as well as the offline event selection that will be discussed in the following.

5.2 Event selection and processing

The DeepCore data sample used for this analysis consists of a selection of IceCube runs performed between 2012 and 2021 that are assumed to be free of defects. In general, the event selection performed on this data is similar to that of previous data samples, the main difference being its independence of the measured charge values¹ which were found to be instable [83].

On order to achieve sufficient neutrino purity of the sample, multiple levels of increasing complexity are performed on the successively shrinking number of events. The first levels utilize detector observables or computationally inexpensive reconstructions, while later ones include high level reconstruction of the contained events. In the following, after a short discussion of the run selection, levels 3 to 7 are introduced. This is largely based on the IceCube-internal documentation in [92] as well as on [83], where new developments with respect to previous procedures [70] are given. More detail on the reconstruction in level 6 can be found in section 6.

5.2.1 Run selection

The DeepCore data sample includes 9.27 years of DeepCore data, as compared to ~ 3 years in the sample used for the previous DeepCore NSI analysis [40]. Runs are selected based on multiple criteria, yielding approximately 95 % of effective total detector up-time. They are required to be marked as so-called good runs in detector monitoring, which implies that they are neither calibration runs nor flawed in some fundamental way. Additional criteria are a duration of at least 2 h with all strings

¹Instead, the number of DOMs that detect e.g. > 0.1 PE is commonly used.

and > 5035 DOMs active. A small number of runs is rejected manually for reasons such as missing documentation or unstable behavior.

5.2.2 Levels of event selection

The sample is optimized for an analysis of PMNS-matrix parameters through ν_μ disappearance. It therefore focuses on the ranges in energy and cosine of the zenith angle ($\cos(\theta)$) where respective oscillation effects in Earth occur. This largely coincides with regions where NSI effects are present, as described in section 7.1. Additional criteria are neutrino purity of the sample and agreement between data and simulation. While this agreement is tested for in numerous variables, it is most important in the three observables: Reconstructed energy and cosine of the zenith angle, which are obtained from event reconstruction, and particle identification (PID), which is derived separately.

Event selection in DeepCore data samples is structured in multiple levels of increasing complexity and computational cost per event, corresponding to the decreasing number of events entering a level.

Incoming signals are triggered and filtered on-site as described in section 4.4.6, constituting the first two levels of processing. The resulting data set at trigger level is dominated by three orders of magnitude by atmospheric muon background as well as random noise, as shown in table 5.1.

Levels 3 to 5 are equal for all three DeepCore samples and result in a sample where event rates from neutrinos and atmospheric muons are at approximately 1 mHz and random noise rates are more than one order of magnitude lower. In table 5.1, the development of neutrino and background event rates over the course of this event selection is shown.

The event selections diverge at level 6, which involves the final event reconstruction. For the sample used in this analysis, no additional cuts are applied levels 6 to 8. Instead, on these levels, additional information is derived for each event and variables are calculated on which the final event selection can be performed. The latter is described in section 5.2.3, with the final sample composition discussed in section 5.2.4.

Data and MC simulated events are processed equally through the entire selection. Good agreement between them is only expected in the final sample, as initially, event types that are not modeled well within the simulation are still present in the sample.

Tab. 5.1.: DeepCore data and simulated event rates at different levels of filtering and selection. Numbers are taken from [92].

	Rates [mHz]			
	L2	L3	L4	L5
Signal				
ν_e CC	1.61	0.95	0.84	0.48
ν_μ CC	6.16	3.77	3.11	1.39
ν_τ CC	0.193	0.129	0.119	0.071
ν NC (all)	0.86	0.53	0.46	0.23
Background				
MuonGun, atm. μ	7273	505	28.1	0.93
Vuvuzela, pure noise	6621	36.6	0.28	0.07
Total:				
MC	13903	547	32.9	3.16
Data	16168	582	31.6	3.15

Level 3

The first offline event selection level includes cuts on available variables that aim at reduction of muon background and random noise, especially in regions where the agreement between data and MC sets is known to be deficient. The selection techniques at this level are required to be computationally inexpensive, as they are applied to a basically unselected and therefore large event sample.

The origin of regions of the event parameter space with poor agreement between data and MC samples are effects that are modeled insufficiently in simulation, such as coincident independent muons or muon bundles. Such events have characteristic signatures, which allow for rejecting them effectively. Establishing good agreement is necessary prior to application of machine learning techniques in subsequent levels. To this end, a total of 12 variables are evaluated and cut upon. These revolve around the number, detection time and position of signal hits. Cuts that aim at noise reduction are based on temporal and spacial correlation between hits. As noise hits often result in low intensity and sparse signal, a 300 ns sliding time window is applied to observed pulses and is required to contain more than 2 hits at its maximum.

A fiducial DeepCore DOM range is defined, which excludes the three layers of strings closest to the detector edge. Vertically, it includes the bottom 50 and 21 DeepCore and IceCube DOMs, respectively. For the so-called cleaned pulse selection, minimum numbers of 6 and 3 hit DOMs are required within the overall and fiducial volume.

Atmospheric muon cuts focus mostly on the vertical position of hits in order to remove down-going tracks, since muons travel through the upper regions of the detector before reaching the fiducial volume. Events are rejected if they contain ≥ 10 pulses observed at $z > -200$ m at times smaller than the DeepCore trigger time. In addition, the vertical position of the first pulse in the cleaned pulse series must be below -120 m.

A second approach to rejecting atmospheric muons uses the number of hits in the veto region. The ratio between vetoed and fiducial hits is required to be below 1.5. An upper limit of 7 is set for veto hits that are causal with respect to the event vertex obtained from a low level reconstruction.

A temporal criterion that aims at the overall signal evolution is that the fraction of DOMs hit during the first 600 ns of the time window needs to be larger than 0.37, while ignoring the overall first two hit DOMs.

Cuts that specifically target coincident events use the time difference between the first and last pulse. Signals from multiple interactions that occur at slightly different times which are sufficiently close to be summarized as one event can be removed in this way. This is done for the cleaned and uncleaned pulse series, with upper thresholds of 5000 ns and 13 000 ns, respectively.

The number of hits vetoed by RT-cleaning is also cut upon in order to reduce events with multiple hit clusters, with thresholds depending on the number of fiducial hit DOMs.

Overall, level 3 cuts remove 95 % of atmospheric muons and >99 % of pure random noise events with respect to level 2 while keeping >60 % of neutrino events. The resulting overall event rate is at approximately 0.5 Hz and is dominated by atmospheric muons by a factor of ~ 10 with respect to noise and ~ 100 with respect to neutrino events.

Level 4

Based on the preceding rate reduction and with improved agreement between data and MC sets, applying machine learning (ML) techniques becomes feasible. Level 4 consists of two Boosted Decision Tree (BDT) classifiers [93] which are trained to reject random noise and atmospheric muons. Their input parameters are selected from more than 40 variables under the objective to maximal classification power. Training is performed on a subset of the MC simulated events and tested for over-training on the remaining MC set.

After applying the first BDT, which focuses on random noise reduction, noise-only

event rates are reduced by a factor of ~ 100 , to below 0.3 mHz while keeping 96 % of neutrino events.

The second classifier includes unfiltered measured data in its training, as this consists of >99 % atmospheric muons. It removes 94 % of atmospheric muons while keeping 87 % of neutrinos. The total fractions of kept events differ depending on the interaction type. As a secondary muon produced in muon neutrino CC interactions can have a signature similar to that of an atmospheric muon, for this kind of neutrino interaction, the smallest fraction of 82.5 % is kept. The resulting sample is still strongly dominated by atmospheric muons, as shown in table 5.1.

Level 5

Cuts at level 5 mainly focus on the remaining atmospheric muons in order to obtain a neutrino dominated sample. Muons that were not rejected at earlier levels are those of which no or little light is detected in the veto regions. Starting containment cuts target event signatures that start at the edge of the fiducial volume.

Corridor cuts apply to a scenario where muons pass largely undetected through uninstrumented parts of the detector which exist due to its almost exactly hexagonal geometry. An example of this is shown in figure 4.11. Such events are targeted through the scarce hits they produce in the vicinity of possible corridors. These are identified beforehand and compared to the approximate particle direction, which is derived in a simplified reconstruction that assumes an infinite track.

This reduces atmospheric muons by 96 % while keeping 48 % of neutrinos, yielding rates of approximately 1 mHz and 2 mHz, respectively, and hence achieving a sample that dominantly contains neutrino events.

Level 6

This is the first stage to differ between the three DeepCore event samples. The sample used in this analysis was optimized to maximize its statistics. Additional criteria are good agreement between data and simulation as well as good performance of the applied reconstruction algorithm, FLERCNN, throughout the included range of the observables. On this and subsequent levels, no selection criteria are applied. Instead, high level variables are determined which allow for an analysis specific event selection that is described in section 5.2.3.

Level 6 is restricted to calculation of computationally relatively inexpensive variables that are later on used in cuts or to determine the event PID. Some of these,

like the number of un-scattered detected photons, are associated to the SANTA reconstruction, which however is not executed for this sample.

Level 7

The final event reconstruction is one of the last stages of the event selection, as it is computationally expensive and scales with the number of events. This stage is only followed by steps that require reconstructed quantities as their input.

The most important reconstructed variables are the analysis observables, namely energy, cosine of the zenith angle and PID. While the latter is estimated separately, energy and cosine of the zenith angle are obtained by running the FLERCNN reconstruction [94] which is described in more detail in section 6.4. This reconstruction is applicable to all events that are still present at this level that meet the SMT-3 requirement discussed in section 4.4.6.

In addition to the event reconstruction, two classifications are performed in FLERCNN. One determines to what extent an event is track-like, corresponding to its PID. The performance of this classifier is discussed along with the final analysis binning in section 7.1.3. The second BDT classifies how muon-like an event is, a quantity that is later on used to reduce muon background.

Level 8

In order to further reduce muon background events, a BDT classifier is run at this level. The input variables of this are the reconstructed vertical position and radial distance to string 36, at the center of DeepCore, as well as four variables determined for muon classification in earlier stages. The latter originate from the FLERCNN muon classification at level 7, corridor identification, and the muon identifying BDT applied at level 4. The distinction power of the classification is shown in figure 5.2. Including the vertex position allows to target atmospheric muons that have not been captured in previous selection levels. Such muons typically enter the detector through corridors along which no significant light is detected until they reach the more densely instrumented DeepCore volume. As a result, their vertex position is reconstructed at the edge and towards the top of the DeepCore fiducial volume. In comparison, the vertex positions of neutrino events are distributed more evenly throughout the detector.

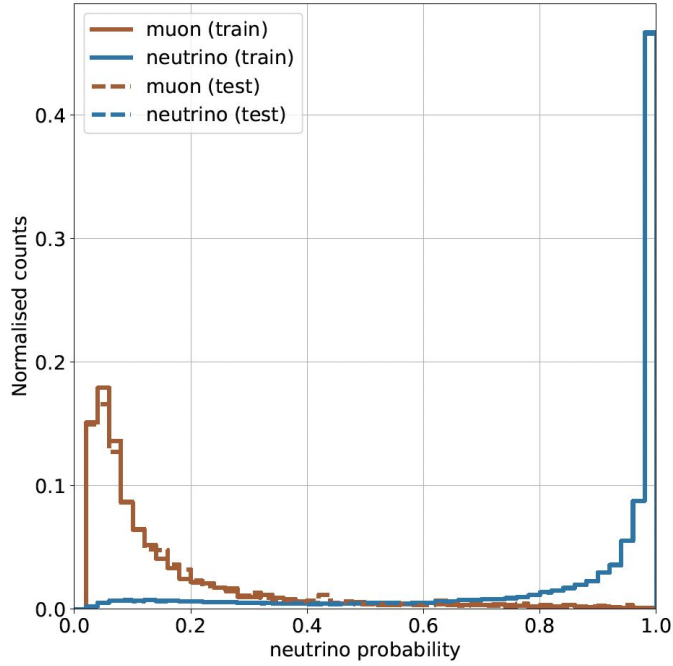


Fig. 5.2.: Rates of simulated neutrino and muon events over their muon BDT classification values. The training and test sample are shown in dashed and solid lines, respectively. At final level, events with a value below 0.8 are rejected. Plot taken from [95].

5.2.3 Final level cuts

The various analyses that can be performed on DeepCore data differ with respect to their signal regions and required data purity. In order to enable adopting analysis dependent selection criteria, the event sample processing up to level 8 does not apply all cuts that would be needed for one specific analysis. Instead, it only applies common cuts and provides all variables that might be used for further event selection.

The final level cuts applied on the data sample used in this analysis have been established for a muon neutrino disappearance analysis. They are specific to the FLERCNN reconstruction and follow two objectives. Firstly, to reject events for which the reconstruction has poor resolution. Secondly, to constrain the data sample to event parameter space regions with good agreement between data and simulation in reconstructed quantities. All involved variables and the applied selection criteria are listed in table 5.2 and will be described in the following.

Tab. 5.2.: Final level analysis sample selection criteria along with the individual removed fraction of the overall sample, based on [92]. Each of the fractions is calculated with respect to a sample where all preceding cuts are applied. In the final analysis binning, an upper limit on cosine of the zenith angle of 0.04 is implemented in order to improve agreement between data and simulation.

Cut name	Value	Removed fraction [%]
Resolution based		
Number of hit DOMs	≥ 7	1.05%
Radial distance	< 200 m	0.09%
Vertical position	-495 m $< z < -225$ m	5.48%
Energy	5 GeV $< E < 100$ GeV	20.70%
Agreement based		
cosine of the zenith angle	< 0.3	19.66%
Number of DOMs with direct pulses	> 2.5	10.50%
Number of hits in the top 15 layers	< 0.5	0.03%
Number of hits in the outer range	< 7.5	0.001%
Level 8 muon BDT	≥ 0.8	23.90%

Resolution motivated selection criteria

In order to utilize the improved reconstruction resolution for events that are fully contained in the densely instrumented DeepCore region, cuts are applied on the vertical position of the starting vertex as well as on the radial distance to string 36 at the center of DeepCore.

Training of the FLERCNN energy reconstruction network is less accurate for sparsely populated energy regions, which motivates including only events with reconstructed energies between 5 GeV and 100 GeV. While for this analysis, higher energies would be of interest, an improvement of the reconstruction accuracy in this range would imply partly redeveloping the reconstruction and exceeds the scope of this work. In addition, events within which less than seven DOMs registered hits are rejected, as this event property was found to be strongly correlated with energy resolution.

Selection criteria to optimize agreement between data and simulation

Within the data selection obtained after level 8, event parameter space regions with minor disagreement between data and simulation are apparent. Different event properties have been found to be correlated with this disagreement.

The mismatch for reconstructed incidence angles above the horizon hints towards remaining atmospheric muons that are poorly modeled within simulation. This finding was included by masking two of the originally 10 angular analysis bins described in section 7.1.3. The upper limit for including events correspondingly

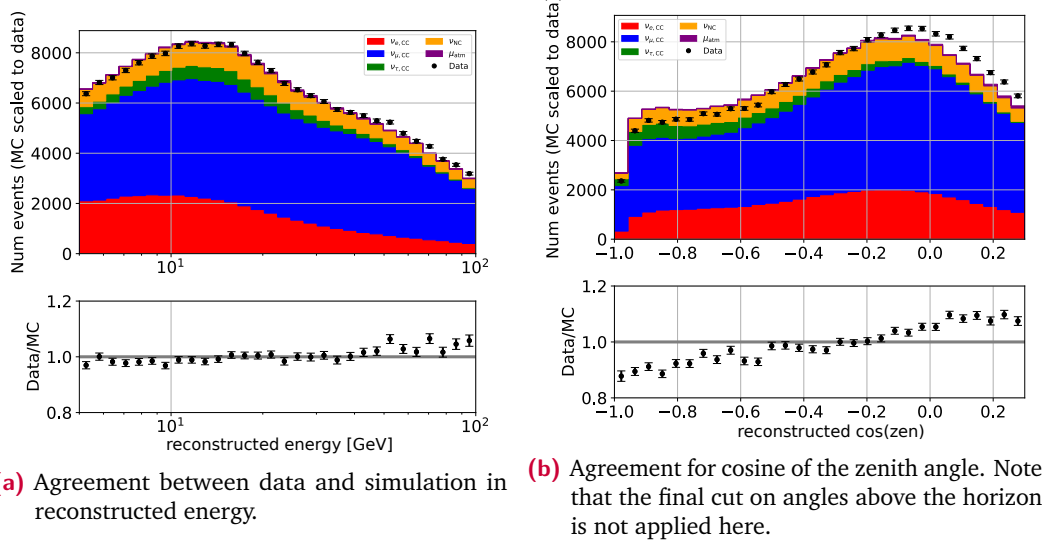


Fig. 5.3.: Comparisons of data and MC simulation in reconstructed observables, prior to fitting. The binning does not correspond to the analysis binning but is chosen significantly finer in order to resolve possible disagreement. Plots are taken from [92].

changes from 0.3 to 0.04 in the final binning.

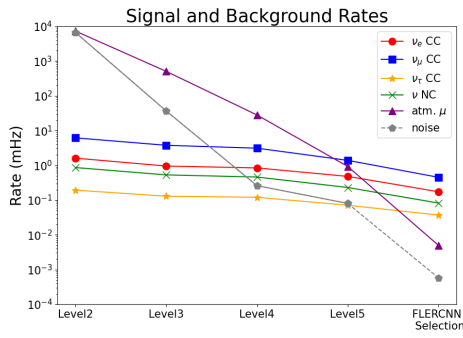
Similarly, coincident events are not included in the simulation. Their contribution in the data can be reduced by rejecting events with hits in the upper- and outermost IceCube DOM layers, as these are likely to be due to atmospheric muon tracks coinciding with the observed neutrino event.

For events that are composed of coincident random noise hits, individual hit times are not correlated. Such events can be rejected by requiring a minimum number of hits that are categorized as un-scattered based on their delay with respect to the reconstructed source position [83]. Selecting these so-called direct pulses was initially implemented for the SANTA reconstruction, which is described in section 6.4.

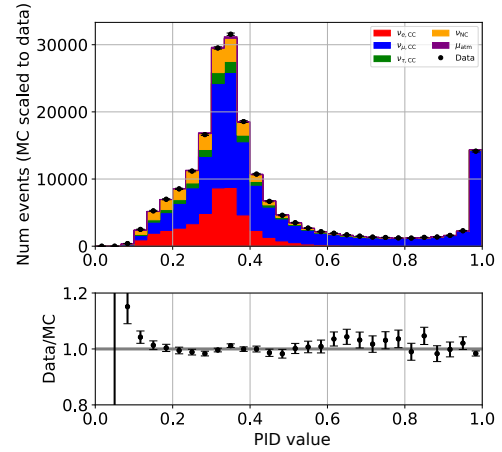
5.2.4 Composition of final sample

In figures 5.3 and 5.4b, the sample is shown in the three observables in which the analysis ultimately is performed, comparing the observed DeepCore data to the simulated baseline set². Since this includes simulated event rates prior to fitting, some degree of mismatch is expected. The disagreement for angles above the

²This assumes a simpler nuisance parameter interpolation than what is described in section 7.4.3, as this is used for the ν_μ disappearance analysis that this sample originally was optimized for.



(a) Development of the event rate over the course of the event selection, given separately for all neutrino interactions and background sources.



(b) Event counts over PID, compared between data and baseline simulation. The same remarks as on figure 5.3 apply.

Fig. 5.4.: Rates and event count agreement at final level. Plots are taken from [92].

horizon, at $\cos(\theta)$ above 0, motivates the respective cut described in the previous section.

The event rates at individual stages are shown in figure 5.4a, with the precise final sample event counts and rates given in table 5.3. The total signal event rate at final level amounts to approximately 0.6 mHz, with background contamination at $\sim 0.65\%$. In comparison to the previous NSI analysis sample [40], this represents an increase in signal event count by a factor of three to four.

With respect to the triggered events at level 2, which are shown in table 5.1, close to 7% of signal events are kept at final level. The largest reduction, by $\sim 96\%$, concerns muon neutrino CC interactions, which are rejected due to being falsely identified as atmospheric muons. For background events, depending on their source, a reduction by six to seven orders of magnitude is achieved.

Based on this approximately pure neutrino event sample that is obtained for both data and simulation, the analysis approach detailed in section 7 can be implemented. Firstly, however, more detail on the final event reconstruction is given.

Tab. 5.3.: Event rates and counts in the final sample, prior to fitting. The listed uncertainties on the expected event counts are based on the Poisson error of the number of simulated events prior to weighting. At this stage, disagreement between overall event rates for data and simulation is expected. Values are based on [92].

Selection	Rate [mHz]	Overall event count for 9.3 yr	Sample fraction [%]
Signal			
$\nu_{e,CC}$	0.1411	41299 ± 69	23.7
$\nu_{\mu,CC}$	0.3522	103063 ± 113	59.1
$\nu_{\tau,CC}$	0.0348	10187 ± 22	5.8
ν_{NC}	0.0667	19522 ± 47	11.2
Background			
μ_{atm}	0.0033	968 ± 57	0.55
noise	0.0006	178 ± 126	0.1
Total			
MC	0.5987	175217 ± 258	-
Data	0.5134	150257	-

Event reconstruction

In the context of this work, event reconstruction denotes the derivation of physics properties of an event from the associated spatiotemporal pattern of detected light that it includes. For DeepCore analyses, especially neutrino energy and incidence angle are relevant. Most reconstructions also derive the interaction time and vertex, yielding an eight-dimensional parameter space.

One part of this work is the development of a testing framework where different DeepCore event reconstructions can be compared with the purpose of finding the best reconstruction method for the current DeepCore data sample. Furthermore, the fundamental performance limits of likelihood based reconstruction methods are investigated and potential improvements are evaluated. A separate study on reconstruction performance within the IceCube Upgrade can be found in appendix A.4.

6.1 Reconstruction challenges in DeepCore

Reconstructing physics properties of DeepCore events can be approached in different ways, in all of which similar difficulties and trade-offs are present due to detector characteristics as well as computational limitations.

An event signature is defined through the interaction time t , the position in Cartesian coordinates x , y and z , the neutrino direction in zenith (θ) and azimuth (ϕ), and the energy. In the following, a set of values of these parameters, either in true or reconstructed parameter space, is denoted as event hypothesis. For neutrino oscillation signals, the main observables are zenith, energy and PID. The latter is not a dimension of the event parameter space but derived from reconstructed quantities. Event properties other than these three observables are used for event selection criteria, for example through containment within a region, temporal coincidence or direction-dependent instrumentation characteristics, as discussed in section 5.2.

The reconstructed energy can be represented in different ways, corresponding to the event topologies explained in section 4.4.4. Other than the total event energy, a subdivision into cascade- and track-energy is of interest, directly or through

reconstruction of the track length. Track energy represents the photons emitted by a potential secondary muon while cascade energy corresponds to the initial hadronic and potential electromagnetic cascade.

The approaches to reconstruct DeepCore events range from ones that very explicitly relate to physics processes by fitting the Cherenkov cone to more implicit solutions employing neural networks. While some reconstructions available for DeepCore events attempt to restore the full eight-dimensional event hypothesis, others concentrate on single event properties. The reconstruction techniques relevant for this work will be outlined in section 6.4.

A focus is set on the pегleg reconstruction, the function principle and shortcomings of which are discussed in the following. Based on this, studies on general limits and specific hindrances of reconstruction performance are presented.

As discussed in section 4.4.6, DeepCore data of a single event are available as a map of timed charge pulses at DOM positions. Reconstructing event properties from this hit pattern is increasingly challenging the less information on the event is available in form of hits. Most of the events that trigger DeepCore consist of $\mathcal{O}(10)$ hits and have energies of the order of single GeV.

In DeepCore analyses, event reconstruction is a limiting factor for the sensitivity to oscillations effects. This is true in more than one way: Apart from the challenge of reconstructing event properties accurately from only few hits, the size of the data set increases over time and with improved event selection techniques. Since the available computing resources do not scale at the same rate, the computational effort associated with reconstruction is becoming an increasingly relevant factor for analyses. As a result, a reconstruction needs to be fast, accurate and robust, which implies reliability of the outcome and applicability to all observed events. Robustness and speed are impacted by the high dimensional event parameter space. For most reconstructions, a metric is optimized in order to find the event hypothesis that best describes the observation. This proves difficult due to sharp variations in the applied metric throughout the parameter space, especially when aiming at reconstruction of the full event hypothesis. These variations are connected to detector features and can not be modeled analytically in full detail. They are associated to position dependent ice properties, the quantization of space and signal that the IceCube instrumentation entails and causality of hits with respect to tested event hypotheses.

Consequently, robustness and speed cannot simultaneously be at arbitrarily high levels, which motivates different strategies. There are robust and relatively slow reconstructions available as well as faster reconstructions that are only applicable for a fraction of the available events and concentrate on reconstructing single event

parameters.

For likelihood-based reconstruction approaches, which are discussed in the following section, the above mentioned trade-offs result in specific implementation choices. One concerns the choice of minimization technique, which will be discussed in more detail in section 6.3.4. Also, accuracy and speed contradict each other with respect to the amount of information included on the detector. Including more detailed information requires larger memory and more computational effort prior to the actual reconstruction. These basic considerations concerning robustness, accuracy and speed need to be considered in the choice of reconstruction approach and will come up in the following descriptions of existing approaches and possible improvements.

6.2 Likelihood based reconstruction

Most reconstruction methods mentioned in this work are likelihood based. A reason for this is the universality and robustness of this approach. It allows to reconstruct the full event hypothesis and, in an ideal case, has the potential of yielding the best possible resolution.

Likelihood based reconstruction is performed through optimization of a likelihood function with respect to the event hypothesis. Both the minimization process and the formulation of the likelihood are non-trivial tasks which will be described in detail further on.

The likelihood function $LH(\vartheta|\vec{x})$ maps the observed hit pattern \vec{x} and an event hypothesis ϑ to a likeliness of ϑ given \vec{x} [96]¹. Optimizing it with respect to ϑ yields the event hypothesis that best describes the observed data. For better numerical stability, the logarithm of $LH(\vartheta|x)$, the logarithmized likelihood (LLH), is commonly used. The information contained in the LLH function can be broken down to a set of PDFs $P(x_i|\vartheta)$ which connect any possible hit and event hypothesis. The LLH function for a given event consisting of n hits \vec{x} hence is

$$LLH(\vartheta|\vec{x}) = \log \left(\prod_{i=1}^n P(x_i|\vartheta) \right) = \sum_{i=1}^n \log (P(x_i|\vartheta)) \quad . \quad (6.1)$$

As $P(x_i|\vartheta)$ can not be described analytically for DeepCore data, no analytical formulation of the overall LLH function can be achieved. Instead, individual PDFs can be approximated numerically based on simulation. This simulation is either executed at

¹Note that $LH(\vartheta|\vec{x})$ is not normalized and therefore does not return probabilities or denote as probability density function (PDF).

run-time, which implies large computational cost, or based on previously simulated and tabulated events, which provide less accuracy [97].

6.3 The pegleg reconstruction

The reconstruction used for the DeepCore NSI analysis prior to this work [40] is called pegleg [82]². Part of this work was its maintenance as well as the evaluation of its current and potential performance.

Pegleg reconstructs the full event hypothesis in a likelihood based approach that is described in section 6.2. It utilizes pretabulated simulated information on hit probabilities for different event types and characteristics (see section 6.3.2).

Based on this, an eight-dimensional parameter space is defined. In section 6.3.1 the details are given of how these dimensions represent the physical properties of any event type. The data considered per event by pegleg are the associated hits, specifically their hit times and the positions of the DOMs observing them. This is discussed in section 6.3.3.

The likelihood formulation as well as different ways to optimize the parameter space are investigated in section 6.3.4. Since commonly the negative LLH is minimized, this optimization will further on be denoted as minimization.

6.3.1 Event representation

Neutrino interactions in IceCube and DeepCore produce different photon distributions, which are shown in figure 4.14 and discussed in section 4.4.4. Internal event representation in preferably few dimensions is needed for the likelihood formulation and minimization. The pegleg internal approach to this is shown in figure 6.1.

In order to represent all of the possible topologies, this event representation needs to include cascade- and track-like events. The latter consist of an initial hadronic cascade and outgoing muon track, the length of which depends on the muon energy. In pegleg, this is approximated by allowing for an initial cascade-like light deposition with variable energy and an outgoing track of variable length which is aligned with the initial neutrino direction.

The track length is subdivided into 5 m long parts to simplify LLH calculation. Each of these segments is assumed to deposit the same energy $E_{\text{min ion}} = 1.11$ GeV within

²The name initially implied this reconstruction to be an intermediate solution, which, however, was used and built upon persistently through several years.

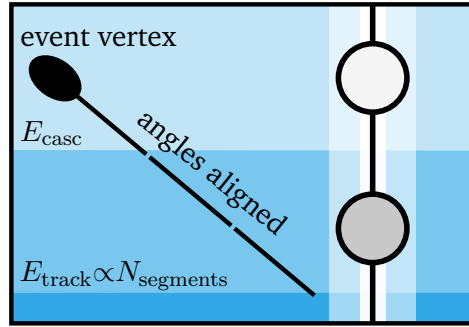


Fig. 6.1.: Schematic of the pegleg internal representation of a neutrino interaction with an initial cascade at the interaction vertex and an outgoing track which is subdivided into 5 m long segments. Cascade and track are assumed to be aligned. Only vertical variations in ice properties are included, as these can be described in tables.

the ice in the form of Cherenkov light. The underlying assumption is that the secondary muon traversing the ice acts as a minimum ionizing particle, as described in section 4.4.4, with no intermediate stochastic losses. An additional approximation is made for the photon production times. These are assumed to be the same for all photons produced in the same track segment or cascade.

As a result, a pegleg event hypothesis has eight dimensions:

- Two parameters, θ and ϕ , for the angle of the incoming neutrino. These are assumed to be the same for the resulting cascade and track.
- Three parameters, x , y and z , for the position of the interaction vertex.
- One parameter, t , for the interaction time.
- One parameter, E_{casc} , for the cascade energy.
- One parameter, N_{track} , for the number of track segments, from which the track energy E_{track} and length l_{track} can be derived as $E_{\text{track}} = l_{\text{track}} \frac{0.222 \text{ GeV}}{m} = N_{\text{track}} \cdot 1.11 \text{ GeV}$.

While this event representation is beneficial for a sufficiently simple parameter space, it has multiple shortcomings. For both the initial cascade and track, the directional alignment with the incident neutrino is only approximately correct at the energies considered here [56].

The most obvious shortcoming of the track representation is the quantization into 5 m segments. Furthermore, the assumption of constant muon energy loss per distance neglects the possibility of stochastic losses during propagation. Also, as shown in figure 4.15, the muon energy loss changes with its kinetic energy. The

assumed mean value of 0.22 GeV/m significantly underestimates losses within the first segments if initial muon energies approach 100 GeV .

6.3.2 Tabulated hit probabilities

The individual hit PDF values, $P(x_i|\vartheta)$, are taken from simulated events that are tabulated beforehand. The tables contain hit probabilities at a generic IceCube DOM for specific standard cascades and muon track segments. During reconstruction, any event hypothesis can be approximated by linking, transposing and rotating these as well as scaling the number of emitted photons.

In order to estimate the PDF of hits at each module, hypotheses are each simulated $100 - 1000$ times [98]. Shape and amplitude of the resulting hit time distribution are stored in separate files and spline-interpolated subsequently. Table dimensions are the source vertical position and zenith angle, observer position relative to the source and observation time Δt relative to the emission [98, 99].

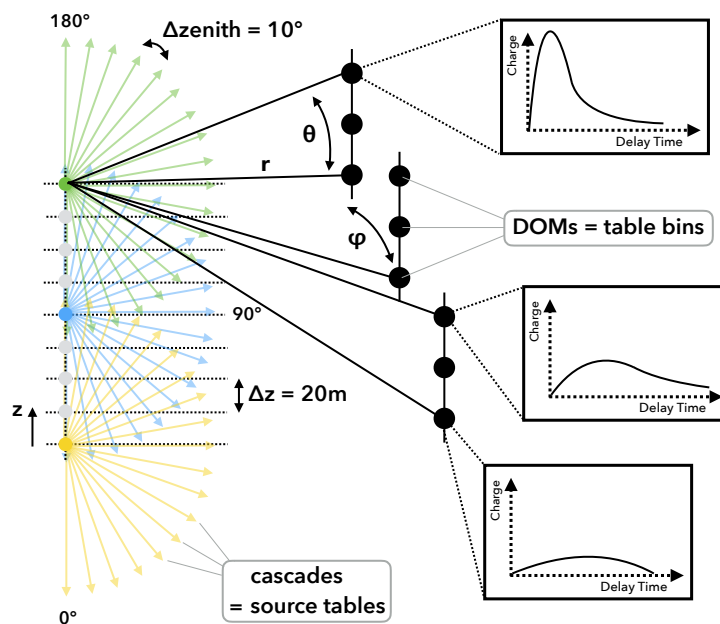


Fig. 6.2.: Schematic overview of how photospline tables for cascades are generated. Typically, 1 GeV cascade are simulated on a grid of zenith angles and vertical (z) positions. Photons from each cascade are simulated on a half-sphere, as symmetry in azimuth is assumed. Information is stored per zenith angle of the source, its z position and the DOM position, with the amplitudes and time profile of the resulting charge expectation values stored separately. Figure taken from [98].

In figure 6.2, an overview of the tabulation process for cascades is shown. Hit patterns at all DOMs for cascades (or muon track segments) at different angles

and positions are simulated multiple times to obtain the expected hit distributions. Simplifying assumptions are made in order to keep the resulting table file size and simulation time at a manageable scale. Simulated cascades typically all have an energy of 1 GeV, making re-scaling of the overall number of hits necessary during reconstruction. For muons, track segments of few meters length are simulated and tabulated in a similar manner, assuming minimum ionizing behavior.

In order to reduce dimensionality, the ice characteristics are assumed to be independent of the photon direction or horizontal position. Hence, simulating cascades at different vertical positions and zenith angles suffices to describe the entire detector. The assumed azimuthal symmetry allows for simulating only half of the solid angle for each cascade. Photons that are generated within this half-sphere are propagated through the ice using ray tracing, starting from the photon source [91, 100].

Instead of tabulating the photon flux throughout the ice, which was done in earlier approaches, tables used in `pegleg` only include the module response to any photons that reach them [99, 101]. A generic module is assumed, including information on module quantum efficiency and angular acceptance as well as local ice properties such as hole ice, which were discussed in sections 4.4.1 and 4.4.3.

The simplifications mentioned above induce several shortcomings of the information stored in the tables. Firstly, interpolating between a grid of event hypotheses may reduce precision. Also, as only the vertical variations of ice properties are included, the current understanding of the ice can not be fully represented in the tables, while being included in simulation and analyses otherwise. Simulating only cascades of a single energy and minimum ionizing muon track segments also induces imprecision. However, more accurate tabulation would result in larger table files. In the current state, table look-ups take up $\sim 70\%$ of the total `pegleg` runtime and a significantly increased reconstruction time is not computationally feasible.

6.3.3 Hit representation

The data from which `pegleg` derives an event hypothesis is a pattern of timed charge pulses and the associated DOM positions.

In `pegleg`, charge information is not used, as its representation in the simulation is not sufficiently correct to maintain good agreement with data. At the energy range considered here most events consist of $\mathcal{O}(10)$ hits, which are predominantly caused by single photons [77]. The mentioned mismatch between data and Monte Carlo is more important for such low-hit events, as its stochastic nature averages out its impact for large sets of hits.

In order to formulate a computationally feasible LLH function (see section 6.3.4),

individual charge pulses are binned in time. In other reconstructions, the first bin edge is set once a certain charge has accumulated [77]. With pegleg not including charge information, instead, the first time bin starts at the first hit time and continues for 5 ns. Since the behavior of a DOM observing additional signals within the following 45 ns³ is not modeled well, this period is set as a dead-time after any observed signal. Further hits are treated in the same way as the first one and each hit bin is attributed a charge of 1PE.

A shortcoming of this hit representation is that it does not make full use of the known charge pulse features and the detected information. It also differs from the spline representation used in the likelihood tabulation, which is discussed in section 6.3.2. However, the simplification it provides allows to formulate a likelihood function the computation of which is of manageable complexity, as is discussed in the following section.

6.3.4 Likelihood formulation and optimization

In order to find the best fitting event hypothesis, pegleg optimizes a logarithmized Poisson likelihood function, $\text{LLH}(\vartheta|\vec{x})$, which compares the measured charges \vec{x} to the expectation $\vec{\mu}$ per event. The latter is derived for an internal event hypothesis ϑ , with the expected signal read from the tables described in section 6.3.2.

LLH formulation

The logarithmized likelihood is given by

$$\text{LLH}(\vartheta|\vec{x}) = \sum_d^{N_{\text{DOMs}}} \sum_t^{T_d} q_{d,t} \ln(\mu_{d,t}(\vartheta)) - \mu_{d,t}(\vartheta) + \ln(\Gamma(q_{d,t} + 1)) \quad . \quad (6.2)$$

The first sum covers all DOMs that register hits in the observed event. The time bins of the signal at a respective DOM are iterated over in the second sum. In combination, this corresponds to the hit number n in equation 6.1. Hence q describes the observed charge obtained from the measurement \vec{x} , which in pegleg is set to 1PE. The introduction of the Γ -function would however allow for non-integer charge values [77].

Information on the detector and its response to light production in the ice enters

³These values have been determined from simulation [77].

the LLH function through the hit expectation PDFs. These are read from the pre-simulated tables described above and included in the response matrix R which enters the expectation values as

$$\mu_{d,t} = \sum_s^S R_{d,t,s}^{\text{type}}(\vec{x}_d, \vec{x}_s, \theta_s, \phi_s, \Delta t_{d,t,s}) \cdot E_s + n_{d,t} \quad . \quad (6.3)$$

The total of S sources s can be cascades or track-segments, which is indicated in the response matrix as type. This makes R a two-dimensional matrix with $S \times (N_{\text{DOM}s} \cdot T_d)$ elements, with T_d the number of time bins in DOM d . The content of the matrix depends on the position of DOM d , which is stored in data \vec{x}_d , the positions \vec{y}_s of the light sources that the event hypothesis is composed of, their azimuth and zenith angles ϕ_s and θ_s , and the differences between photon emission and detection time $\Delta t_{d,t}$. Energy E_s is deposited in form of photons by source s . The expected dark noise rate $n_{d,t}$ per DOM and time interval is drawn from calibration information.

LLH minimization

Within pegleg, the optimum of the negative LLH described above is found using a seeded minimization. The aim of this is to reconstruct the event hypothesis ϑ best fitting the observed data \vec{x} . Minimization therefore applies to the underlying set of light sources and their respective properties defined through ϑ within the boundaries given based on the event representation.

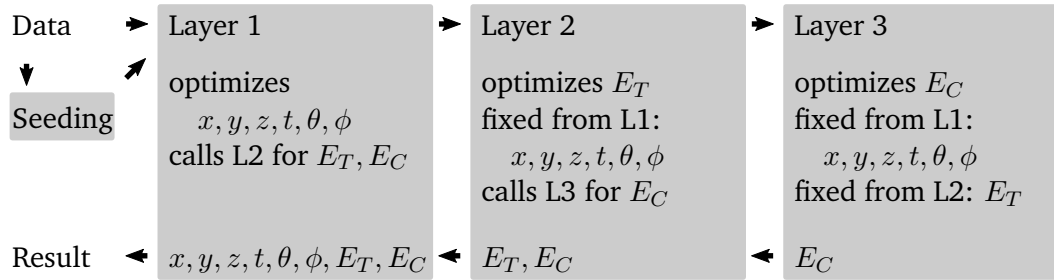


Fig. 6.3.: Schematic overview of the three layers of minimization within the pegleg event reconstruction. In a nested approach inner minimization layers are target functions of outer ones.

For technical reasons, minimization is subdivided into three layers, in which the energies of individual sources (layer 3), number of sources (layer 2), and other event parameters (layer 1) are optimized in a nested implementation [102, 103]. Parametrization and seeding algorithms are provided outside the three layer optimization [104]. Different fast seeding algorithms are used for the minimization in

layer 1, some of which are fast simplified reconstructions.

The three minimization layers are schematically described in figure 6.3. Each layer calls the subsequent one as the target function of the minimization it performs. Parameter values that constitute the hypothesis tested in layer i are handed over to layer $i + 1$ and fixed therein.

The likelihood space in layer 1 includes the three spacial dimensions, x, y, z , interaction time t and angles θ and ϕ . In these dimensions, the likelihood space has features that are hard to minimize (see section 6.6). This motivates the application of computationally expensive global algorithms such as Multimodal Nested sampling (MultiNest) [105]. Other minimization techniques have been investigated for example in [102] and [106]. More detail on the individual layer 1 minimization techniques can be found in section 6.6.4.

In layer 2, the track energy E_T is optimized through the number of minimum ionizing 5 m segments that represent an outgoing muon track. Given the relatively small number of possible track lengths this is done as a scan, starting at a pure cascade hypothesis and adding segments successively. The number of segments directly corresponds to the number of columns of the response matrix R , which has to be adapted accordingly in each step.

Layer 3 finds the optimal cascade energy given the parameter values from layers 1 and 2 by differentiating the LLH and applying the Newton method to find its minimum.

This way of minimizing the LLH yields reliable and reproducible results while being computationally expensive due to the multiple minimization layers and the globally operating algorithm in layer 1. In section 6.5, the performance of pegleg will be compared to other reconstruction algorithms that are introduced in the following section.

6.4 Existing alternatives

Beyond pegleg, there are multiple approaches to reconstructing DeepCore events that were developed later or in parallel. These can be subdivided into reconstructions that deduce the full event hypothesis in contrast to those which concentrate on a subset of the parameters. All implementations approach the trade-off mentioned in section 6.1, between robustness, accuracy and speed, differently. Likelihood based reconstructions mainly differ in how the hit PDFs are obtained and stored.

A study performed as a part of this work was originally aimed at finding the best performing reconstruction for oscillations analyses on the current DeepCore data

sample. To this end, pegleg was compared to the RETRO and the combined SANTA and LEERA reconstructions [97].

Since then, the FLERCNN [94] reconstruction was developed, which is used in the data sample that this NSI analysis is based on.

6.4.1 SANTA and LEERA

The main observables for oscillations analyses are the neutrino zenith angle and total energy. These are reconstructed by the SANTA and LEERA algorithms, respectively. The underlying approach focuses on speed, yielding a simplified algorithm that is only applicable to a fraction of the events in a DeepCore sample.

The principle used in SANTA is a χ^2 fit to the arrival times of photons that are assumed to lie on the Cherenkov cone, as shown in figure 6.4. Such photons are required to be only minimally scattered, which is why only a subset of the detected hits can be used for this reconstruction. The respective hit selection is based on the delay that scattering induces with respect to direct hits on neighboring DOMs [97]. In cases where less than five filtered hits remain in an event SANTA is not applicable. While SANTA is implemented with a focus on reconstructing the direction, it also returns an estimate of the track start- and end point.

The subsequently performed LEERA reconstruction of the event energy is table based. It optimizes a simplified likelihood that solely regards which DOMs observed any or no hits. Its event hypothesis constitutes an initial cascade and aligned outgoing track, with their angle and end point fixed to the values obtained beforehand [83, 92].

The fraction of events reconstructible in SANTA in comparison to pegleg strongly depends on the event type and data sample and is between 20 % and 40 % [83]. This reduced applicability makes SANTA and LEERA not universally comparable to reconstructions of the full event hypothesis, such as pegleg or RETRO, which is described in the following.

6.4.2 RETRO

In contrast to SANTA and LEERA, reconstructions like RETRO and pegleg are able to utilize scattered light. The relevance of this becomes apparent from the scattering lengths at typical Cherenkov wavelengths, which have been shown in section 4.4.3 to be below the distances between DOMs. The reconstruction approach of RETRO is similar to that of pegleg, both being likelihood-based and using presimulated tables.

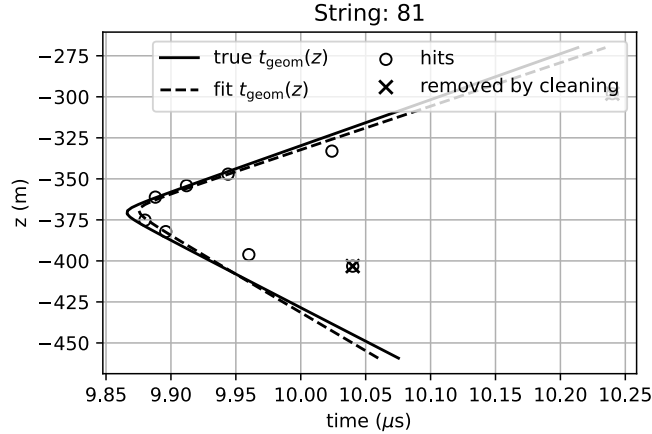


Fig. 6.4.: SANTA reconstructed muon neutrino CC event, where solely hits on string 81 are considered. Hit times are those of the first pulse in the respective DOM. Figure taken from [97].

It aims at being applicable to the same events as pegleg while being faster and more accurate. This similarity allows for a direct comparison in section 6.5.

Details that differ between the two reconstructions are the minimization algorithms, the fact that seeding is included in the RETRO reconstruction process, and the tabulation generation and compression.

The tabulation process of RETRO includes similar assumptions as that of pegleg, for example that ice properties purely change vertically. It differs mostly in being light-source-based: Photons are traced from DOMs through the ice to their source. In this way, only detectable photons are simulated [97].

6.4.3 FLERCNN

The FLERCNN reconstruction takes a fundamentally different approach from those described above. Instead of tabulated likelihood values, a neural network provides the connection between the hit pattern and the full event hypothesis [94]. A CNN is trained to identify event patterns, similarly to how this kind of network architecture is commonly applied for image classification in translation invariant settings [94, 107, 108]. The required network complexity is thereby correlated with the image complexity.

Apart from its specialization on event energies below 100 GeV, FLERCNN is fundamentally similar to preexisting IceCube event reconstruction CNNs [109]. It consists of multiple individual networks for different tasks, namely the regression for event

energy, zenith angle and vertex as well as classifications of how cascade-like and muon-like an event is [110].

6.5 Performance of pegasus with respect to other reconstructions

Since multiple reconstruction approaches are available for current DeepCore data samples, one part of this work was to set up an environment and decide on metrics to optimally compare different reconstructions. As the individual reconstructions have been developed separately, they do not all use one common interface, framework, programming language or data format. However, an accurate comparison requires that all considered reconstructions run on the same set of simulated events⁴, controlled through the same software and on the same hardware, since computational performance is evaluated as well.

To this end, the source code of all compared reconstructions was included in the OscNext metaproject, stored in [111] within the IceCube Apache Subversion (SVN) repository. This was set up on the IceCube high throughput (HT) Condor [112, 113] computing cluster with modular wrapper scripts. These use a common interface and are designed to make the available set of reconstructions easy to run, extend and evaluate [92, 114].

The comparison metrics need to express the requirements of oscillation analyses, which are strongly impacted by the resolution of reconstructed observables. Also, computational capacities need to be taken into account, since current DeepCore data sets that are optimized for high statistics contain approximately three times more events than previous samples [40, 92].

As a measure of the reconstructed parameter resolutions the range between the 25 % and 75 % quantile in each bin is used. This is called the 50 % inter-quantile range (IQR). A common binning scheme is applied for all reconstructions, using ten logarithmically spaced energy bins between 1 and 100 GeV.

In addition to the resolution, the median is investigated to account for systematic deviations. Both are regarded depending on the true neutrino energy as well as the number of hit DOMs, called channels, since resolutions are known to depend on these. Run-times are monitored externally for evaluation of computational performance.

⁴Here, a $\sim 3\%$ subset of the muon neutrino CC events simulated as described in section 5.1 is used to run all reconstructions on.

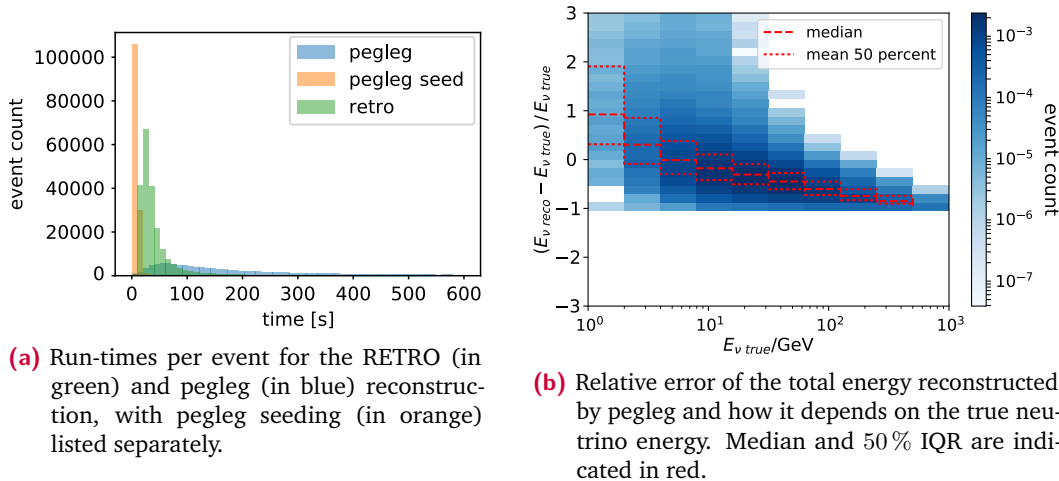


Fig. 6.5.: Evaluation of the reconstruction performance is based on computational cost (left) and resolution in the reconstructed parameters (right).

6.5.1 Run-times

Reconstruction time becomes more relevant with growing data samples. While SANTA and LEERA only need seconds to reconstruct single events, they are not applicable to most DeepCore events. More complex reconstructions, such as pegleg, take minutes per event. For the $\mathcal{O}(10^6)$ data- and Monte Carlo events within current DeepCore samples, one minute of reconstruction time per event implies at the order of 10^5 CPU hours of sample processing time. This motivates finding the fastest reconstruction for the full event hypothesis.

Figure 6.5a shows reconstruction durations per event for RETRO compared to pegleg and the run-time of its default seeding algorithm. The peak positions of the distributions differ by approximately a factor of two between pegleg and RETRO. In addition to that, the outliers towards longer reconstruction times in pegleg reaches up to 10 min and beyond, which is not the case in RETRO. Seeding takes $\mathcal{O}(s)$ and is therefore negligible.

Summarizing, reconstructions that concentrate on single event parameters are a factor of 10 – 100 faster those that reconstruct the full event hypothesis, among which RETRO is to be preferred.

6.5.2 Energy resolution

To compare resolutions of reconstructed parameters, the inter-quantile range is the preferred metric. Resolutions are known to depend on true neutrino energy. This is

apparent from figure 6.5b, where the relative error⁵ in reconstructed energy is given together with its 25 %, 50 % and 75 % quantiles.

For neutrino events with energies of 10 GeV and less to enter a data sample, over-fluctuation is likely, which means that an event contains more hits than expected on average. This increases its probability to be reconstructed with an energy larger than the truth. In order to represent such biases, comparisons between reconstructions include both the median and inter-quantile range, as shown in figure 6.6a for RETRO and pegleg.

The pegleg internal assumption that muon tracks behave as minimum ionizing particles causes an energy-dependent under-estimation of the track energy and corresponding over-estimation of the track length. The higher the event energy and consequentially the muon energy, the less exact is the description as a minimum ionizing particle. For the higher DeepCore event energies, the actual photon output is significantly larger than what a minimum ionizing muon track of the true length would produce. Reconstructing the correct track energy would imply a too large track length, which contradicts the observed hit positions.

While this applies to both compared reconstructions, RETRO implements appropriate corrections. As a result, larger biases are seen for pegleg. The inter-quantile range as a bias-independent measure for the resolution is similar for both, becoming smaller with increasing event energy as the increasing number of hits provides more information.

An energy resolution comparison including SANTA and LEERA shows similar resolutions and can be found in appendix section B.2.

6.5.3 Angular resolution

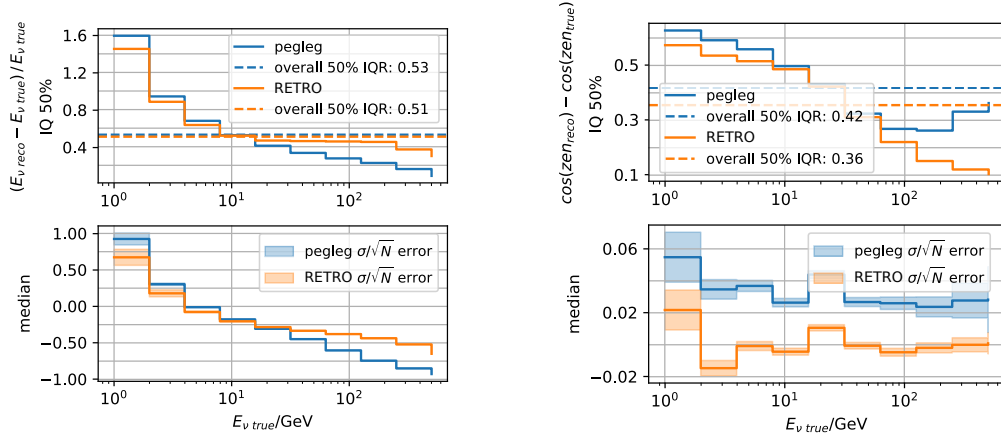
The second observable of DeepCore oscillations analyses is the zenith angle. In this parameter, less energy dependent biases are expected and apparent from figure 6.6b⁶. The energy dependence of resolutions persists.

Resolutions of SANTA are less energy dependent, making them better at lower energies and worse at higher ones. The respective plot can be found in appendix B.2.

Based on run-times and the two evaluated observables discussed above, RETRO was used to reconstruct one of the current high statistics DeepCore samples. The third observable of oscillations analyses, PID, is not included in this comparison,

⁵Using the relative rather than the absolute error serves to visualize solely the energy resolution in a way that is more easily comparable to that of other parameters, which do not directly depend on the true energy depicted on the horizontal axis.

⁶For reconstructed parameters other than energy, the absolute instead of the relative error is evaluated.



(a) Resolution comparison of reconstructed energy. (b) Resolution comparison of reconstructed cosine of the zenith angle.

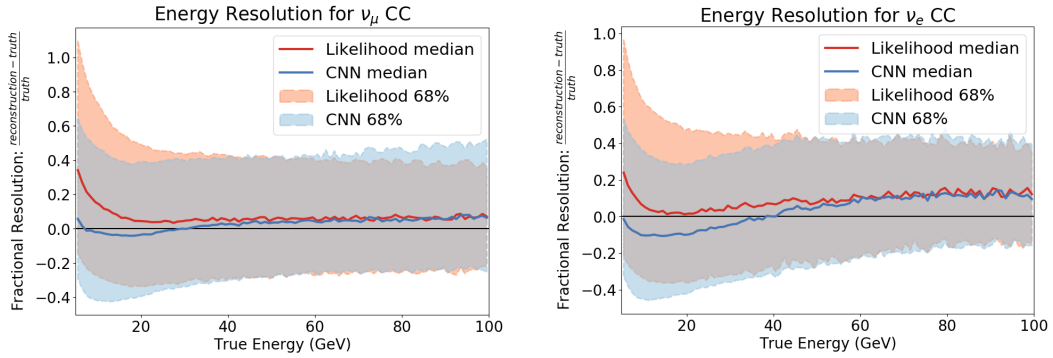
Fig. 6.6.: Comparison of the resolutions of DeepCore reconstructions of the full event hypothesis, namely pegleg and RETRO. The quantities being compared are the 50 % IQR and median of the reconstructed values per true energy bin. The overall 50 % IQRs, including all energy bins, are shown as dashed lines. Errors on the median are given as shaded regions, based on the standard deviation over the square root of the number N of entries.

since it is commonly determined based on reconstructed variables that are available independently of the specific reconstruction that is applied. Other parameters, such as interaction vertex position or time, do not have as much impact on analyses and are therefore not considered for this study.

The outcomes for RETRO and pegleg motivate a more fundamental study of reconstruction performance: These two implementations rely on very similar assumptions and approaches but perform very differently. The question arises what the theoretical potential of their basic approach is and whether it is already exhausted. The study performed in section 6.6 attempts to answer this based on the pegleg reconstruction.

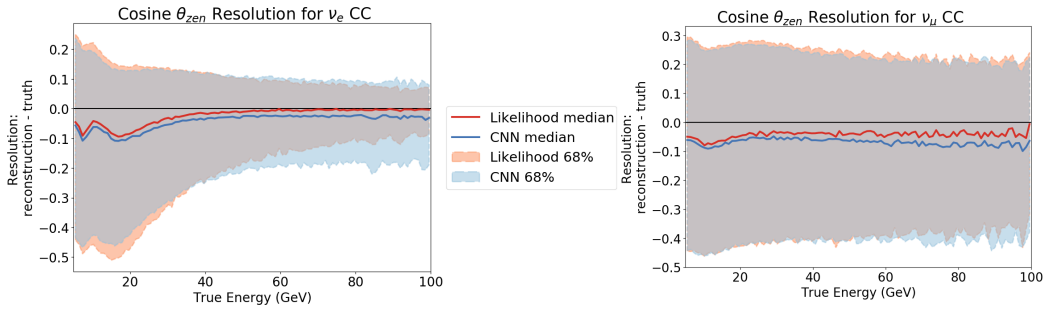
6.5.4 Comparing the FLERCNN reconstruction

From the comparison of table-based reconstructions in section 6.5, the RETRO reconstruction is clearly preferable with respect to its resolution and computational cost. As the FLERCNN reconstruction was under development when the respective study was conducted, it is not included in this original comparison. When compared to RETRO, however, FLERCNN performs similarly or better in the listed aspects. Therefore, FLERCNN is used in the event sample that this NSI search is based on. A



(a) Comparison of muon neutrino CC interactions, producing track-like events. (b) Comparison of electron neutrino CC interactions, producing cascade-like events.

Fig. 6.7.: Resolution comparison in reconstructed energy between the FLERCNN (CNN based) and RETRO (likelihood based) reconstruction. Event selection is applied as described in section 5.2.3, except for rejection of specific reconstructed energy ranges. Different interaction types are shown. These figures are taken from [115].



(a) Comparison for cascade-like events. (b) Comparison for track-like events.

Fig. 6.8.: Resolution comparison in reconstructed cosine of the zenith angle between the FLERCNN (CNN) and RETRO (likelihood) reconstruction. Event selection is applied as described in section 5.2.3, except for rejection of specific reconstructed energy ranges. Curving of the median is associated to oscillation effects. These figures are adapted from [115].

comparison between the energy and cosine of the zenith angle resolutions of the two reconstructions is shown in figures 6.7 and 6.8. Within the energy range of 5-100 GeV that is included in this analysis, only small differences in resolution are apparent. Depending on what subset of the sample is being considered, the relative performance of the two reconstructions varies.

A consistent and significant improvement of FLERCNN with respect to RETRO is the speed at which events can be reconstructed, depending on the employed hardware and specific case. A central processing unit (CPU) based test shows improvement by a factor of 400 [94]. This is a conservative case, since the parallelization obtainable on a graphics processing unit (GPU) can be exploited better for the CNN architecture.

6.6 Potential and limits

In order to evaluate the potential of likelihood-based reconstructions, the details of the pegleg reconstruction and its shortcomings are investigated in this study. In order to locate potential improvements it is necessary to understand the impact of individual assumptions as well as the resulting features of reconstructed values.

In a first step, the shape of the likelihood space close to the best fit value, which is the hypothesis returned by the reconstruction, is tested as a potential benchmark for the reconstruction accuracy. The observed likelihood shapes and behavior, however, motivate a more detailed study of the likelihood space, as they do not conform to Wilks' theorem [116]. This is likely due to the different assumptions entering the reconstruction and tabulation which were introduced in section 6.3.

Along with other effects, these assumptions result in biases and reduced resolution in individual reconstructed dimensions. An additional factor impacting these is how well the global likelihood minimum is located. This is approached by evaluating different alternatives and improvements to the pegleg minimization process.

6.6.1 Likelihood modality and reconstruction performance

There are multiple applications to correlating the likelihood space shape around the reconstructed parameter values with how accurately an event can be reconstructed. Firstly, such a measure for the accuracy could be used to prioritize or even reject events for an analysis. Secondly, this is a fundamental cross-check for the behavior of the likelihood space and the included simplifications.

When assuming that Wilks' theorem holds for the pegleg likelihood, the expected

Gaussian behavior of the likelihood minimum translates into a parabolic shape of the LLH around the global minimum. Stochastic uncertainties in the minimization would result in a corresponding Gaussian distribution of reconstructed parameter values around the respective truth. As a result, the width of the LLH minimum is expected to be correlated with the residual $|\theta_{i,\text{reco}} - \theta_{i,\text{true}}|$ of a reconstructed parameter θ_i .

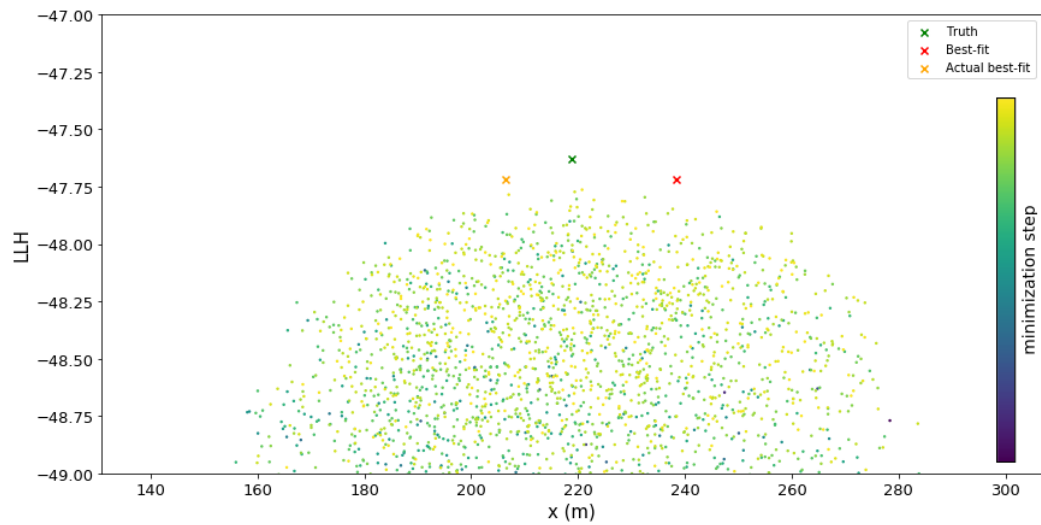
Figure 6.9 shows hypotheses that are evaluated during the pегleg reconstruction of two exemplary events, projected onto the x coordinate of the vertex position. In one of the cases shown there, the envelope of the hypotheses resembles the expected parabolic shape. Note that regarding the envelope instead of the entire ensemble of hypotheses is necessary due to the projection that is performed for this plot: Individual points may differ in more dimensions than are shown.

While in this case, the assumption of a parabolic likelihood optimum seems to hold, this could not be confirmed in other events, as shown in the second example. In general, no correlation to the parameter value residuals could be established, due to the non-parabolic shape found in most cases.

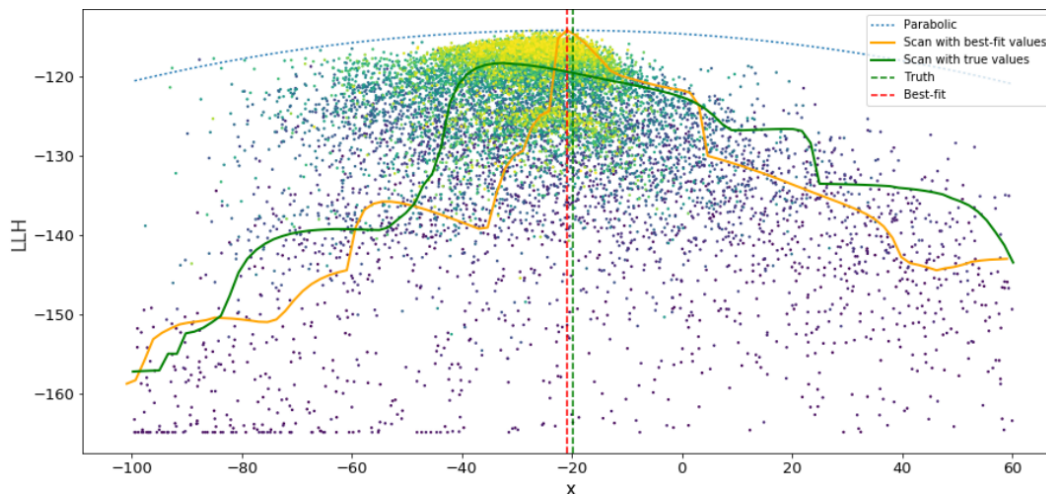
This behavior does not allow for correlating the width of a hypothetical parabolic LLH minimum with the reconstruction error, as was the initial aim. Other tested likelihood-based estimators, such as the one described in [117], also fail to provide good correlation. A simplified way to correlate the likelihood space shape around the best fit point with the reconstruction error without assuming a parabolic shape is pursued in the recently developed FreeDOM reconstruction [118, 119]. This reconstruction does not suffer from the modeling deficits in tabulation that are described above, as it is based on a neural network instead of tabulated values. The measure used therein is the standard deviation of the evaluated hypotheses within a fixed LLH range with respect to the best fit value. This is used to re-weight events at analysis level.

This observation of non-parabolic LLH minima hints at incorrectness of the underlying assumptions, specifically the applicability of Wilks' theorem. Hence, the reconstruction resolution is assumed to include effects that are non-Gaussian.

The applicability of Wilks' theorem to the pегleg LLH can furthermore be visualized through the ΔLLH distribution. These differences between the LLH values of the true and reconstructed parameter hypothesis would be expected to follow a χ^2 distribution with eight DOF, corresponding to the eight reconstructed parameters. Figure 6.10 illustrates that this is not the case for pегleg-reconstructed DeepCore events, which can not be matched with a χ^2 distribution of any DOF.



(a) For this set of hypotheses, only the LLH range close to the optimum is shown. The envelope of the projected hypotheses is approximately parabola-shaped, with its maximum at the true parameter value (green cross), as would be expected if Wilks' theorem applies. The indicated best-fit values show a small LLH difference of ~ 0.1 with respect to the true hypothesis. The best evaluated hypothesis (orange cross) differs marginally from the best-fit value that pegleg returns (red cross). Differences at this scale were observed repeatedly and are assumed to be due to pegleg-internal numerical inaccuracies.



(b) In this case, the envelope of the evaluated hypotheses is not parabolic. For a better understanding, one-dimensional scans through x are performed at different values of the parameters that are not shown here. These are fixed to their best fit values and true values, respectively, for the orange and green curve, and show some of the variation observed in the hypothesis distribution. A parabolic fit to the hypotheses is shown as a blue dashed line. The true and best-fit values are indicated as vertical lines.

Fig. 6.9.: Sets of all hypotheses that are evaluated to reconstruct two cascade events close to the center of IceCube. Shown is the projection onto the x dimension and LLH value. The color indicates the minimization step in which a hypothesis was evaluated.

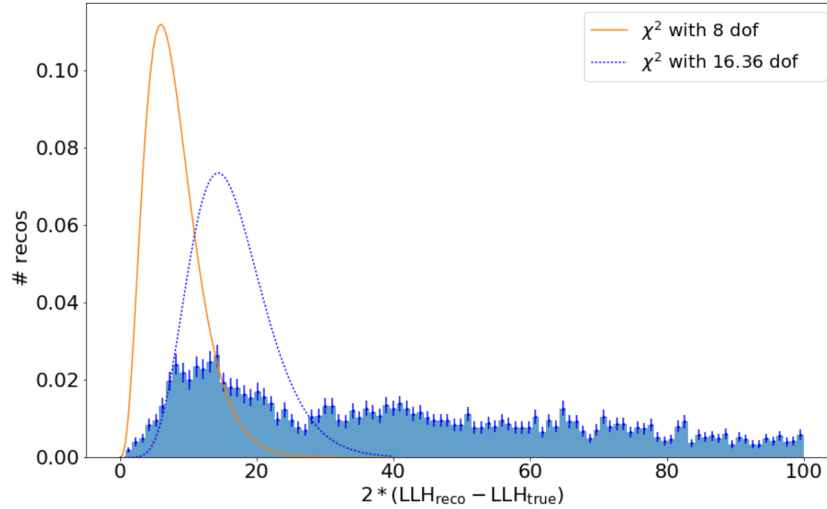


Fig. 6.10.: Histogram of the differences between LLH values of the true and reconstructed hypothesis. The reconstructed data set was simulated as described in section 5.1, with cuts applied up to level 2. It contains approximately 6000 events that are approximately 50% ν_e and ν_μ , mostly interacting via CC. For simplification, this assumes depth-independently homogeneous ice and no background other than random noise. Note that the event distribution is only shown up to ΔLLH values of 100. Other than expected if Wilks' theorem would apply, the distribution is not χ^2 shaped. This is shown from the best fitting χ^2 distribution (blue dashed line), which does not describe the histogram well. A distribution with 8 degrees of freedom (DOF) is shown in orange, for reference.

The origin of this behavior might hint towards deficits in modeling $P(x_i|\vartheta)$, resolving which could potentially improve the reconstruction. In investigating this, a first step is to discern the impact of the difference in assumptions used in tabulation and reconstruction. This is done by generating simplified events based on DeepCore simulation, as the true PDFs are not known.

In the following, events such as the ones shown in figure 6.10, which are produced using the full MC simulation procedure, are called simulated DeepCore events. For these, the observed detector signal is available as time-binned charge pulses, as was introduced in section 4.4.6. The simulated events include all current understanding of the involved physics processes, such as detailed ice modeling. In reconstruction, these events are compared based on the simplified event representation used in the tabulated PDF, which introduces a mismatch.

In order to exclude this difference in representation, further sets of events are drawn based on the tabulated expected charge distributions of a physics hypothesis. For these, event representations in simulation and tabulation are the same. In this way, for a simplified study, different realizations of hit patterns of single events are generated, based on the finely binned tabulated charge expectations of the associated physics hypothesis. In figure 6.11, multiple ways to draw and bin such events are

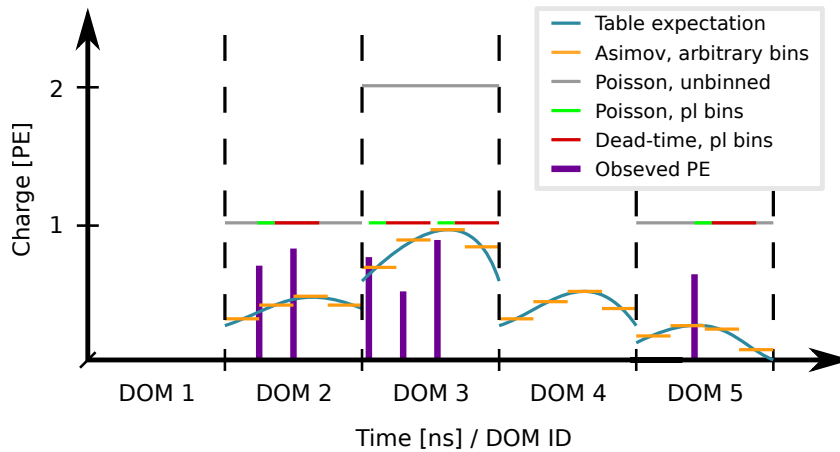


Fig. 6.11.: Schematic overview of the applied approaches to draw hit-patterns for single event hypotheses. The charge associated to a generated hit is shown on the vertical axis. The horizontal axis shows the individual DOMs, subdivided based on the applied time binning. An example for observed PE is shown as purple vertical lines at the respective hit time. The expected charge distributions at individual DOMs, as tabulated for the respective event hypothesis, are shown in blue. For this study, multiple hit pattern instances are drawn based on the expectation values of the same event hypothesis. For Asimov events, each hit contains the unaltered expected charge. Pseudo-events consist of hits that are drawn from Poisson-distributions around the respective charge expectations. The shown options for time binning are the pegleg (pl) binning described in section 6.3.1 or an arbitrary number of uniform bins throughout the triggered event time range.

shown schematically.

The first kind of such events are in the following called Asimov-events and are maximally idealized. By using a reconstructed hypothesis, these events adopt the pegleg-internal event representation. For this idealized hypothesis, the observed charge in all DOMs and time bins is set to the tabulated expectation value. This neglects the statistical nature of hits and that photoelectrons are an integer quantity. These properties are present in pseudo-events, which are an additional kind of simplified events. Other than Asimov-events, these consist of integer signal values that are drawn based on the charge expectation values using Poisson statistics. In the comparisons shown in the following, for each of the fully simulated DeepCore events, one pseudo- or Asimov-event is included, respectively.

For the artificially generated Asimov- and pseudo-event sets, the ΔLLH distributions of the reconstructed values are approximately χ^2 shaped. Asimov-events match a distribution with ~ 8 DOF, as would be expected based on Wilks' theorem for the eight-dimensional event parameter space. For the pseudo-events, 9 or 10 DOF fit best.

This supports the assumption of deficits in modeling the data, specifically such

deficits that are associated to the tables. An alternative cause could be the finite resolution of the MultiNest minimizer, which is caused by statistical fluctuations of the fit result due to the random components in minimization. This has however been ruled out as the sole origin.

The observed behavior has also been found to not be limited to the pegleg reconstruction. Other likelihood-based reconstructions, namely RETRO, also do not result in a χ^2 distribution when applied to simulated DeepCore events.

Overall, none of the two criteria that would be expected based on Wilks' theorem are observed, namely the parabolic shape of the likelihood minimum and the χ^2 distribution of the likelihood differences. The observed behavior leads to the supposition that the factors contradicting Wilks' theorem are due to the simplifications applied in tabulation, which are necessary for likelihood-based reconstructions. A more accurately modeled PDF would therefore potentially result in an improved reconstruction performance. The differences in the assumptions made for reconstruction and tabulation, such as the simplified ice model or hit representation, are investigated individually in the following study, in order to identify areas in which improvement of the modeling has significant impact on the reconstruction outcome.

6.6.2 Potential modeling deficits and their impact

At different points throughout the reconstruction, simplifications are applied in order to facilitate computation. The resulting ways that hits and events are represented within tables and reconstruction potentially impact the accuracy of reconstructed quantities, which will be assessed individually in the following.

Noise treatment

The treatment of random noise hits is one of the domains where the assumptions used in table generation differ from what is assumed in reconstruction. For the tabulated expectation values, noise is assumed to result in a fixed charge expectation throughout all DOMs. This has the additional effect that the LLH function 6.2 is defined for any number of hits taking part in an event, including 0. This treatment is motivated from the PMTs being the origin of dark noise, as discussed in section 4.4.5. Contrarily, data processing includes cuts on hits that occur outside a fixed radius r_B from the assumed event vertex, as described in section 4.4.6, in order to

reduce noise hits.

This mismatch between expectation and observation beyond r_B ideally only results in a constant offset of LLH values. Specifically, it is not a factor in the Δ LLH distribution, and thus not the cause of the likelihood differences not following Wilks' theorem. However, noise hits worsen reconstruction uncertainty and result in causality borders that cause sharp edges in the LLH space which can inhibit minimization.

When evaluating this in an idealized setting with no random noise, resolutions in the reconstructed quantities improve only marginally. This is expected, since the majority of events in low energy data samples contain no noise hits. Consequently, while noise is not modeled ideally, this is assumed to have no large impact.

Event representation in reconstruction

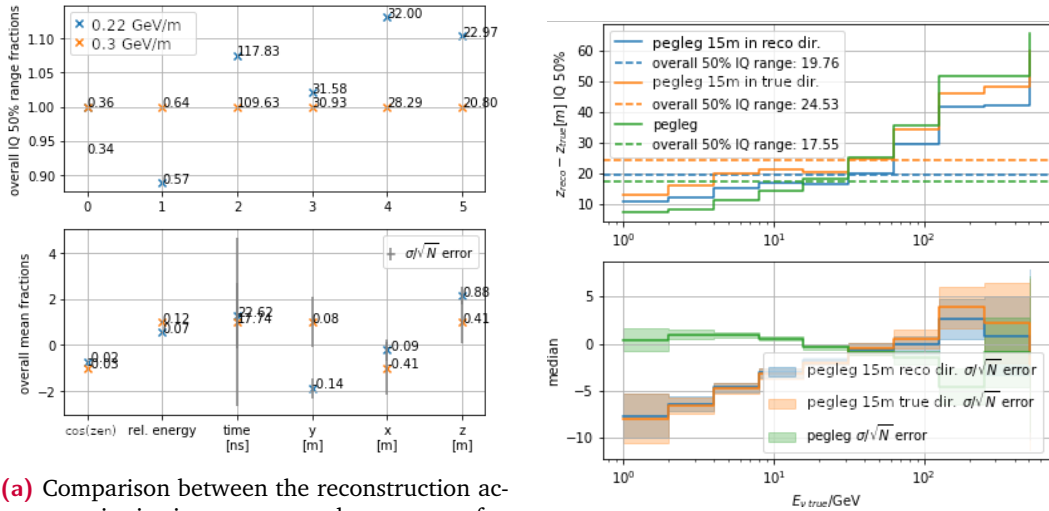
The pegleg internal event representation described in section 6.3.1 contains multiple mechanisms that induce biases in reconstructed variables.

Firstly, it includes a split of the overall energy into the initial cascade, E_{casc} , and outgoing track, E_{track} . However, this does not include any knowledge on the relation between both. Together with the limited spacial resolution of the shape of low energy DeepCore events, this causes a preference for reconstructing 0 GeV for either E_{casc} or E_{track} , attributing all observed light to just one of both. One way to resolve this would be to introduce a Bayesian Bjorken- y prior. It however seems preferable to find a less flawed event representation.

Secondly, the track is described as minimum ionizing. Especially for the higher event energies considered in this work, this does not describe the muon behavior well. It models neither stochastic losses along the muon track, nor the fact that the starting energy of these muons is significantly larger than what would be minimum ionizing and decreases subsequently. As a result, the relation between energy and track length is biased, as was discussed in section 6.5.2.

This bias manifests as a mismatch between the charge expected for the actual event and its pegleg representation. Since pegleg only considers integer PE, a matching description does not directly improve the resolution. It however can be used as a measure for the agreement between reconstruction and tabulation.

Two approaches were tested in order to estimate the potential of correcting this, targeting the single components of the mismatch between track length and energy. Firstly, increasing the muon energy from 0.22 GeV/m to 0.3 GeV/m. On average, this matches the actual energy deposition better, as the assumption of minimum ionizing muons with 0.22 GeV/m constitutes the smallest possible energy deposition. This increase impacts the resolution of each reconstructed parameter differently and



(a) Comparison between the reconstruction accuracies in six reconstructed parameters for assumed track energy losses of 0.22 GeV/m and 0.3 GeV/m. The impact on resolutions is characterized by the IQR relative to the 0.3 GeV/m case (upper plot). The median (lower plot) illustrates systematic biases.

(b) Energy dependent change in IQR and median when shifting reconstructed tracks by 15 m along the true and reconstructed direction. The vertical position is shown since the effect is most prominent there.

Fig. 6.12.: Impact of different attempts to reduce the incorrect relation between track energy and length that is induced by the pegleg internal track representation. The events that this is based on are a representative subset of current DeepCore MC.

no overall improvement can be concluded. The resulting summarized characteristics of the resolutions in all reconstructed parameters are shown in figure 6.12a. More detailed investigation shows that the resolution improves for high energy events while declining for lower energies, as expected from the underlying issue. Implementing segment energies that increase with overall track length, however, is not compatible with the existing optimization approach, as shown in figure 6.3.

In a second study, correction of the track length is imitated by shifting the position of the track segments along either the true or the reconstructed direction. This results in slightly larger improvement to both of the charge agreement and the spatial resolutions. The latter only concerns the median of the reconstructed distributions and not their widths, which are evaluated through the 50% IQR. Resolution characteristics of the reconstructed vertical positions can be found in figure 6.12b. Similar behavior was found when combining both approaches of positional shift and increased energy.

The ad-hoc corrections applied for this study show that the reconstruction internal representation has non-negligible impact on the resolution and median. While a fundamental improvement in the representation is preferable with respect to such corrections, this would imply fundamental changes to the reconstruction ap-

proach and is therefore only relevant for the development of future reconstruction algorithms.

Ice model

As discussed in section 6.3.2, the tabulated likelihood values that pegleg relies on assume an ice model that purely depends on the vertical position and zenith angle. However, current models, introduced in section 4.4.3, include significant deviation from this assumption.

The impact of this was evaluated by reconstructing the same events with different ice models, namely the South Pole ice (SPice 3.2.1) model and deep-homogeneous ice. While the first includes all known ice effects except for birefringence, the latter assumes homogeneous ice with the properties measured inside DeepCore.

The effect of this on the energy and angular resolution can be seen to be negligible, e.g. in figure B.17 in appendix B.2. This, however, is based on a reconstruction setting where the detailed pulse signals are ignored and all charge pulses are assumed to be 1 PE. In general, matching the ice models in reconstruction and tabulation yields a better agreement of the charge descriptions. An example for the disagreement is given in figure 6.13, where expected and observed charges differ due to a spacial shift that correlates with DOM numbers. While not constituting a large improvement in the charge-agnostic pegleg implementation, achieving better agreement in this is expected to be a significant factor in a reconstruction that includes the measured charge values.

Resolving this behavior would require solving further underlying reasons for the observed charge mismatch between data and models, which are detailed in the following section. Also, tabulating and reconstructing based on the same ice model implies that, for the current tabulation process, the common ice model can only include vertical dependence of ice properties, as described in section 6.3.2. Adapting the tabulation or raising the resources for significantly larger table files are both problems with no straight-forward solution [102].

Including charge information

Mismatches of the charge values are present in simulation with respect to data as well as in the expectation values for different event representations within reconstruction. This was discussed in the preceding sections as well as in the introduction of tabulation in section 6.3.2 and motivates the pegleg internal policy of assuming any pulses to have exactly 1 PE.

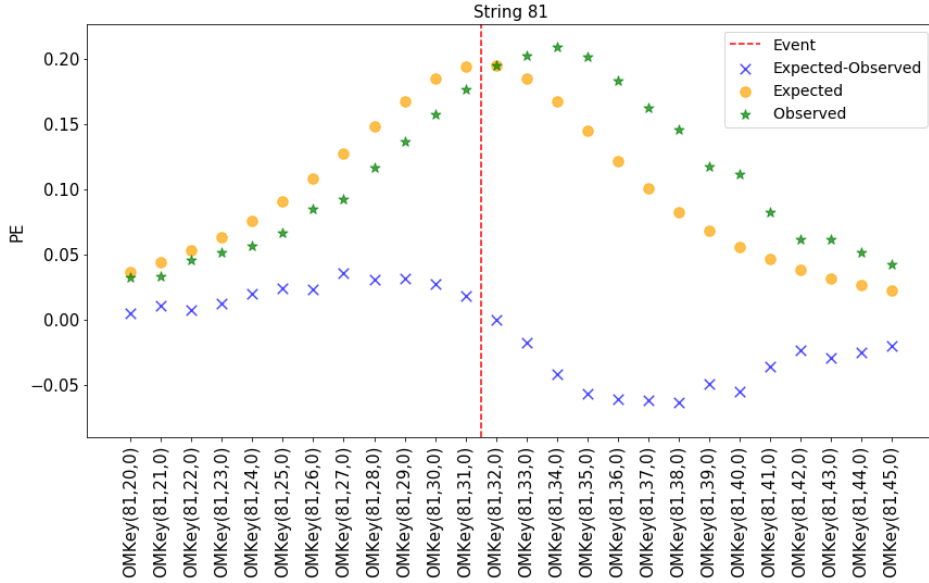
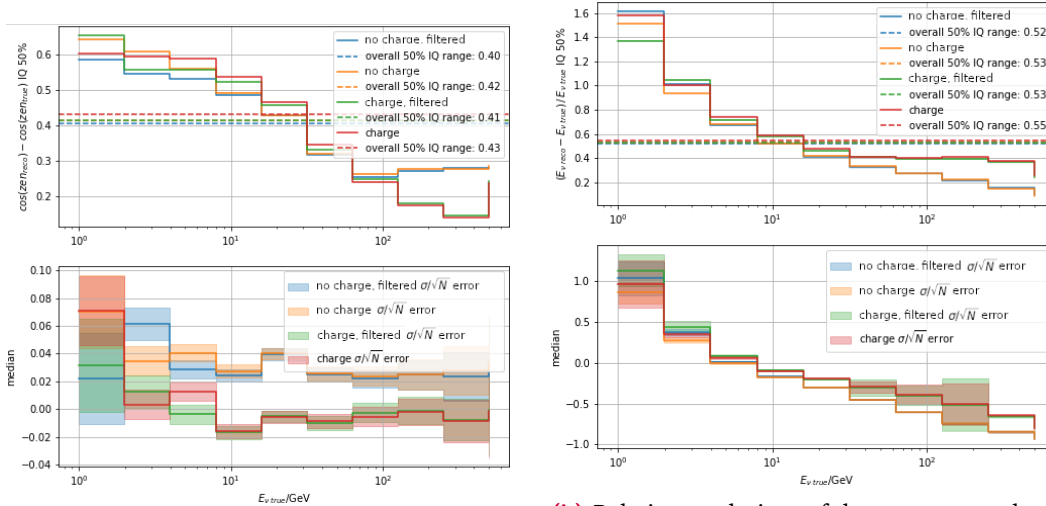


Fig. 6.13.: The expected and observed charge as well as the difference between both, shown per DOM for a single electron produced inside the DeepCore fiducial volume. This exemplifies how observed and expected charge disagree due to a spacial shift, originating from the differences between pegleg reconstruction and tabulation.

Including charge information in the current pegleg setup is the most simple test of the potential improvement. To this end, the likelihood function needs to be adapted to accommodate non-integer charge values, as discussed in [77]. The charge pulses used in this study are filtered in order to reduce noise as part of the event selection. Only pulses with more than 0.3 PE and events with at least 8 pulses are included. Utilizing unfiltered pulses has been found to lead to no significant change of the reconstructed distributions.

The effect observed for including detailed charge information and filtering is shown in figure 6.14 for reconstructed energy and zenith. In both, including charge leads to marginal improvement of the median. The resolutions are largely unchanged, except for the highest energy bins, which become slightly better in energy and slightly worse in zenith. The improved energy resolution and median can intuitively be attributed to the more accurate description of observed light. Figure 6.15 shows summarized IQR changes for the reconstructed time and vertex. Their resolutions improve significantly when including charge information.

As this study includes the existing charge mismatch, an idealized case where perfect charge information is available is investigated as well. A comparison between reconstructed DeepCore simulated events and pseudo-events, drawn with the tabulated charge expectation values, shows the impact of reconstructing events with perfect charge information. The relation of the residuals in these cases are mostly



(a) Resolutions of the reconstructed zenith angle.

(b) Relative resolutions of the reconstructed energy.

Fig. 6.14.: Impact of including detailed charge information on the reconstruction performance in cosine of the zenith angle and energy. Both cases are in addition shown with filters that were found in the RETRO reconstruction to improve the agreement between charge description and observation. The event sample shown here is a representative DeepCore simulated set.

energy independent at 80 % for energy and 50 % for the cosine of the zenith angle. This significant improvement of the reconstruction resolution implies that work on improving the charge description would be worthwhile, assuming that sufficient agreement with data can be obtained to allow for including charge information in the reconstruction.

During the detailed review of the charge pulses, an error in simulation was identified. This was observed as a 25 ns time shift between timestamps of SLC and HLC hits and could be traced back to an off-by-one error in the simulation of the FADC readout, which is the dominant readout for SLC hits. The impact of this was found to be significant for the current DeepCore sample, leading to a re-simulation of all corresponding Monte Carlo sets.

Signal time binning

The time binning of charge pulses used in pегleg constitutes a compromise between a desired high level of detail on one hand and, on the other hand, the correctness of modeling of the detector response. The impact of more accurate pulse description can be tested by introducing different time binning schemes.

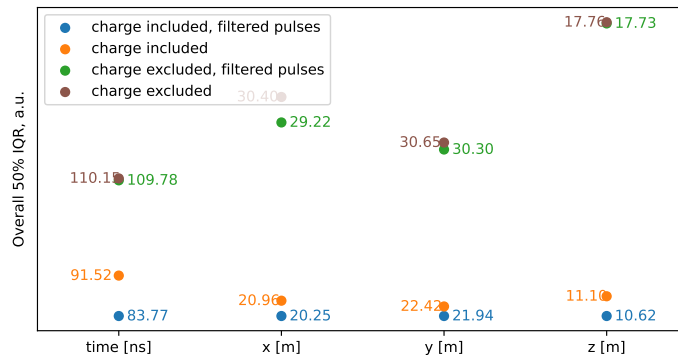


Fig. 6.15.: Reconstruction resolution in interaction position and time when including or excluding charge information, for filtered and unfiltered pulses. The overall 50 % IQR is given separately for each dimension.

As described in section 6.3.3, pegleg places a first 5 ns bin at the first hit time, followed by a 45 ns dead time bin, in which detector effects are known not to be modeled well in simulation. While the overall low number of bins reduces how much detail of the charge pulses this binning is able to describe, the small size of the hit bin causes single hits to have large impact on the overall likelihood value with respect to larger hit- or dead-time-bins, as shown in figure 6.16. This reduction of observed pulses to short isolated 1 PE peaks causes steps in the likelihood landscape that hinder minimization. In addition, a more detailed representation of charge behavior would make the reconstruction less dependent on single hit times.

In an ideal case, the entire pulse duration would be subdivided into preferably many short time bins with no dead-time. This is simulated by drawing pseudo-events as described in section 6.6.1 with different numbers of constant length time bins. All such events use the same pegleg internal event representation of an initial cascade and segmented outgoing track.

A resolution comparison between pseudo-events with 5 and 20 bins shows refinement of the 50 % IQR of typically $\sim 10\%$, depending on the specific reconstructed events as well as the considered parameter. With respect to the pegleg binning, these cases show 30-50 % worse resolutions of time and position and no significant change in other reconstructed parameters. This is expected as the signal accuracy in these dimensions strongly depends on the bin width. This motivates more, finer time bins, which, however, the current state of detector effect modeling does not allow.

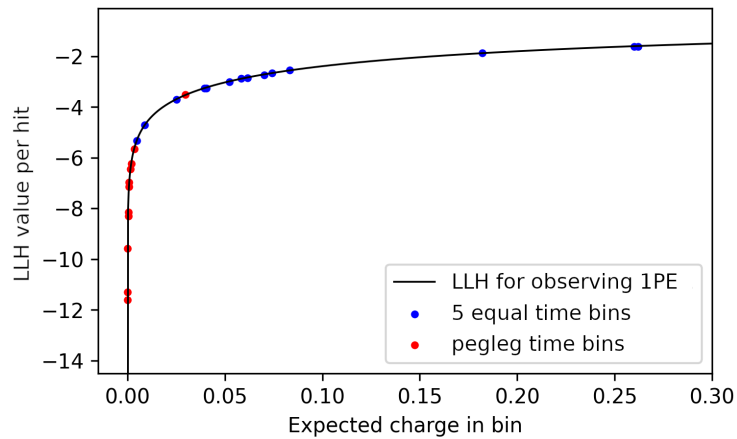


Fig. 6.16.: Dependence of the LLH for observing a charge of 1 PE on the expected charge in a single bin. In the pegleg time binning, a 5 ns bin starts at the first hit time. When using five bins of equal width throughout the event duration the individual bins have significantly longer durations. The expected charge throughout such a bin therefore integrates to larger values, corresponding to less extreme likelihood values in cases where 1 PE is observed.

6.6.3 Detector geometry induced bias

The detector geometry described in section 4.4.2, especially its discrete nature, is origin of various biases and inaccuracies within the pegleg reconstruction. The most prominent one is a spacial shift towards DOMs that observe hits. This is visualized in figure 6.17 for the distance between reconstructed positions and the closest DOM. Partially, the bias is induced by the event selection, as the events at lowest available energies are more likely to pass the cuts if their total number of hits is an over-fluctuation with respect to the overall expectation. This behavior is more likely for interactions at small distances to DOMs.

A second part of the spacial bias has been shown in an idealized simulation to occur for any IceCube-like detector with a spacial quantization of the instrumentation. With expectation values of charge per DOM below 1 PE for most hypotheses, a single hit comprises an over-fluctuation for the DOM observing it. The LLH function, as a result, strongly prefers positions close to the DOM. In the same way, non-observation in surrounding DOMs constitutes an under-fluctuation which disfavors reconstructing close-by event positions.

This spacial bias impacts other reconstructed dimensions, namely interaction time and angles. The angular bias is similar to the positional one, as it prefers hypotheses where the event is directed towards hit positions. For positions close to DOMs, small time differences between hits and interaction are deemed most likely, inducing a bias towards too large reconstructed interaction times.

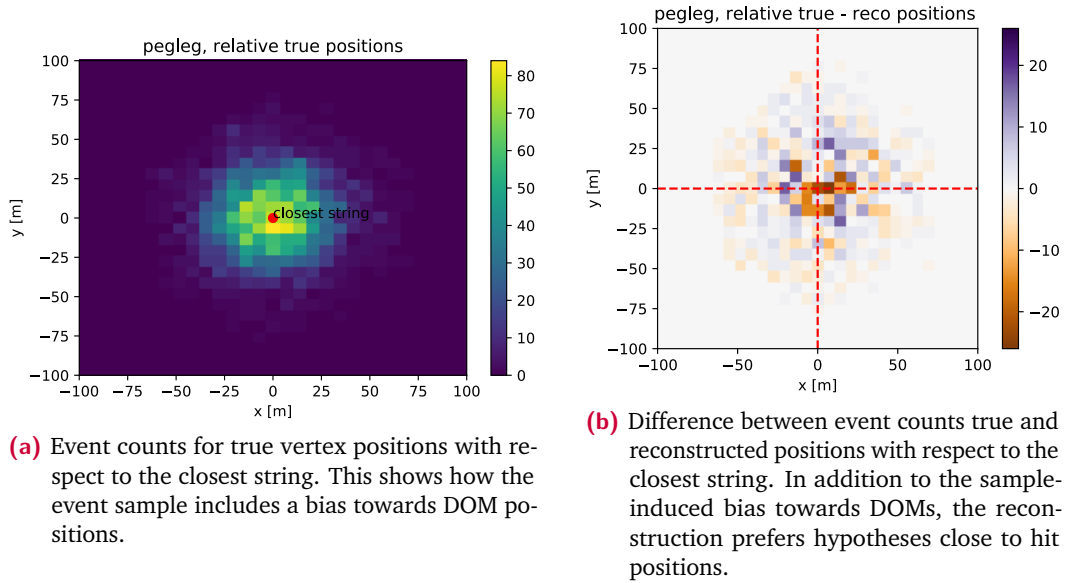


Fig. 6.17.: Positions of DeepCore events in the horizontal plane with respect to the closest string in red.

A possible improvement to the reconstruction would be to introduce a Bayesian prior [93] on the reconstructed position. In a related approach, requiring a minimum number of hit DOMs resulted in minor improvement. This leads to the assumption that the introducing a prior would be of small impact compared to other shortcomings of the pegleg reconstruction.

6.6.4 Minimization effects

The core of pegleg is its minimization process, where the optimum of the eight-dimensional likelihood space is found in a nested three-layer approach as described in section 6.3.4. Some of the likelihood space features discussed above complicate this minimization, as they introduce local extrema and discontinuities. An example for this is shown in figure 6.18, where causality induced behavior of the likelihood function is shown.

This can be addressed by using global or specialized local minimization algorithms. These as well as the performance and potential of the globally operating MultiNest algorithm are discussed in the following.

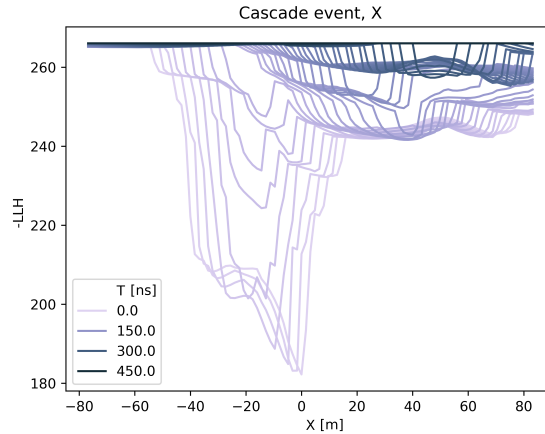


Fig. 6.18.: Scan of the x dimension of the vertex position at different times for a cascade-like event. As the times are larger than the true event time, causality induced edges occur. These as well as the overall non-smooth behavior of the likelihood space can impede minimization. This figure is taken from [102].

Basic minimization concepts

The multiple minima of the likelihood space encourage the usage of a so-called global minimization algorithm such as MultiNest, which is able to process and compare multiple minima simultaneously. This is achieved by evaluating a large number of hypotheses, the so-called live-points, in parallel, and clustering them in a Bayesian nested sampling approach.

In an initial step, MultiNest live-points are drawn either randomly within a range or with a prior around a seed hypothesis. For subsequent iterations, live-points are clustered in ellipsoids and therein newly drawn based on the previously evaluated hypotheses. This reduces the possible parameter space while robustly describing any shape of the considered areas [105].

This approach reliably finds the globally best fitting hypothesis, with a low statistical spread. It however evaluates a significantly larger number of hypotheses than algorithms that act locally.

While there are numerous ways to implement a local minimization, the fundamental difference to a global one is that a single hypothesis and area of the parameter space is evaluated at a time.

Examples that were evaluated in previous studies on improving the pegleg reconstruction [102] are gradient-based minimization techniques where each step is determined from the LLH gradient in the preceding hypothesis. In the underlying case, gradients can be obtained directly from the spline tables described in section 6.3.2. In order to overcome intermediate minima, a random component can be used.

This, however, does not accomplish a reliability with respect to its outcome that is similar to that of MultiNest.

Performance of the current implementation

The resolutions obtained when reconstructing an event sample have multiple random components. Firstly, the distribution of events within the parameter space, whose random character can be eliminated by simulating the same physics hypothesis multiple times. Secondly, the variation of hit patterns for a given event. In order to investigate the third factor, which originates from the minimization itself, the same hit pattern of the same physics hypothesis can be reconstructed multiple times. The outcome of this can, however, only be evaluated qualitatively, as it is strongly impacted by the choice of physics hypothesis and hit pattern and the computational capacities to reconstruct large numbers of events multiple times each are not available.

In figure 6.19, 195 reconstructed parameter values of the same physics hypothesis with different hit patterns are shown together with the same number of reconstruction results of the same hit pattern. For the latter, in an ideal minimization, a δ -distribution would be expected, as it purely includes the part of the resolution that is induced by imperfect minimization. In contrast to this, the reconstruction of different hit patterns also includes the random component within hit generation. The fraction of these components associated with the reconstruction varies between parameters and is smallest for the angles, which is expected to be different for track events. The change in width of the distribution of LLH values at the best fit point is expected and shows the stability of the minimization.

Resolutions vary depending on the minimization settings, as shown in figure 6.20, where different numbers of MultiNest live-points are tested. Note that the resolutions are not directly comparable to what is shown in figure 6.19, as they are obtained for a single hit pattern of a different event, namely a 25 GeV cascade at the same position. In this case, the change in the LLH distribution represents the improved capacity of the minimization to find the globally best fitting hypothesis. Resolutions in the individual parameters change by approximately 20% to 30%. The approximate factor ≤ 5 in run-time that the increased number of live-points entails however would make reconstruction of current sets computationally infeasible.

Overcoming LLH features

Some features of the LLH space that render minimization difficult have already been mentioned: Simplifications made in tabulation and signal representation, both in tables and reconstruction, are discussed in sections 6.6.2 and 6.1. Characteristics of the detector geometry, which are discussed in section 4.4.2, add to that. These result in secondary minima and non-continuous changes of the LLH space. These features occur within a small area of the parameter space, beyond which the LLH is largely constant, inhibiting any gradient-based optimization.

One main aspect of this are causality-induced effects. Due to the finite speed of particles and light, the time and location where a hit is observed sets one-sided boundaries on the temporal and spacial position of the interaction causing it. This is shown in figure 6.18. The one-sided-ness of this effect accounts for the capacity of photons to scatter and hence prolong their traveling time.

As a result, hypotheses that are temporally and spatially similar result in strongly differing likelihood values if one of them breaches the causality criterion for an observed hit. For hits that are not noise, this can only occur for reconstructed times beyond the true time value.

In figure B.15 in appendix B.2, the likelihood contributions of individual DOMs are evaluated in a scan through a spacial dimension. This visualizes how the overall likelihood instantly decreases when a hit is excluded through causality criteria. This behavior mostly poses a challenge to local minimization techniques.

A previous study [102] evaluated gradient-based minimization starting at causal hypotheses before the assumed true event time. This prevents crossing causality edges before reaching the global optimum and hence avoid the unreliable and extreme gradients there.

A different way to include the knowledge on this behavior in a reconstruction is implemented in the FreeDOM reconstruction [118, 119], where seeding is performed under mostly causality-preserving criteria.

A different aspect that challenges all minimization approaches is associated with the pegleg internal event representation. Treating the track as multiple segments of minimum ionizing muons results in an artificial quantization, which equally segments the likelihood space and hinders the minimization.

As described in section 6.3.4, the track length optimization proceeds at layer 2 of the pegleg minimization. It is implemented as a simple incrementation of the number of track segments and subsequent call to layer 3, which returns the respective best fitting cascade energy.

As a result, a hypothesis evaluated by the level 1 minimization may show a significant

likelihood difference with respect to the preceding hypothesis if their best fitting track lengths differ. This complicates the application of local minimization techniques. In addition, in a fraction of reconstructions, the optimization in level 2 stops before reaching the best fitting track length, due to the nested optimization approach. A part of this study was to find an efficient way to correct this by enforcing the evaluation of larger track lengths. However, this leads to minor improvement of the energy resolution, while increasing the reconstruction run-time per event, due to the larger number of evaluated track length hypotheses.

There are numerous leverage points to improve the pегleg reconstruction and likelihood based reconstructions in general. These are mostly associated to the simplifications made in tabulation and reconstruction that are necessary to provide sufficient computational performance. Mismatches between the data and their description in simulation, tables and reconstruction are an additional factor that leads to biases and inaccuracies in the reconstructed quantities.

The studies conducted in this work show in depth that the estimated potential of each individual component is either marginal or associated with significant changes to the framework, considerable resource requirements or uncertain success.

Some of these highlight the conceptual downsides of table based reconstructions, such as their incapacity to represent newer ice model developments or multi-PMT optical modules. This motivates the development of table-independent reconstructions that are for example based on neural networks [97].

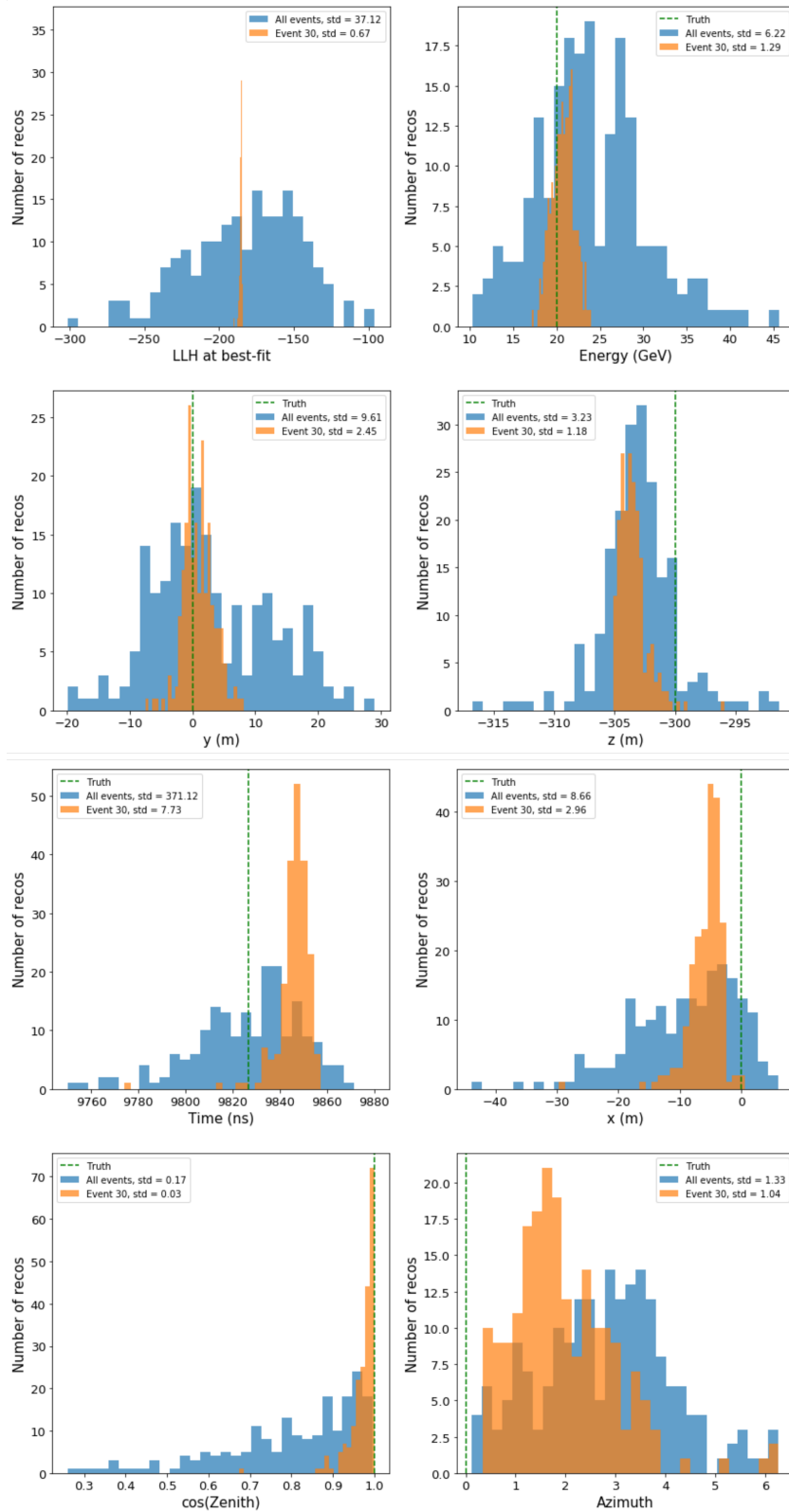


Fig. 6.19.: Multiple reconstructions of the same 20 GeV cascade event, located at the center of DeepCore. This is simulated as a pseudo-event, adopting the pegleg-internal event representation introduced in section 6.6.1. The widths of the distributions differ by between an approximate factor of 4 in energy to a factor of ~ 1.3 for the azimuth angle. Angular resolution is expected to be limited due to the cascade character of the event.

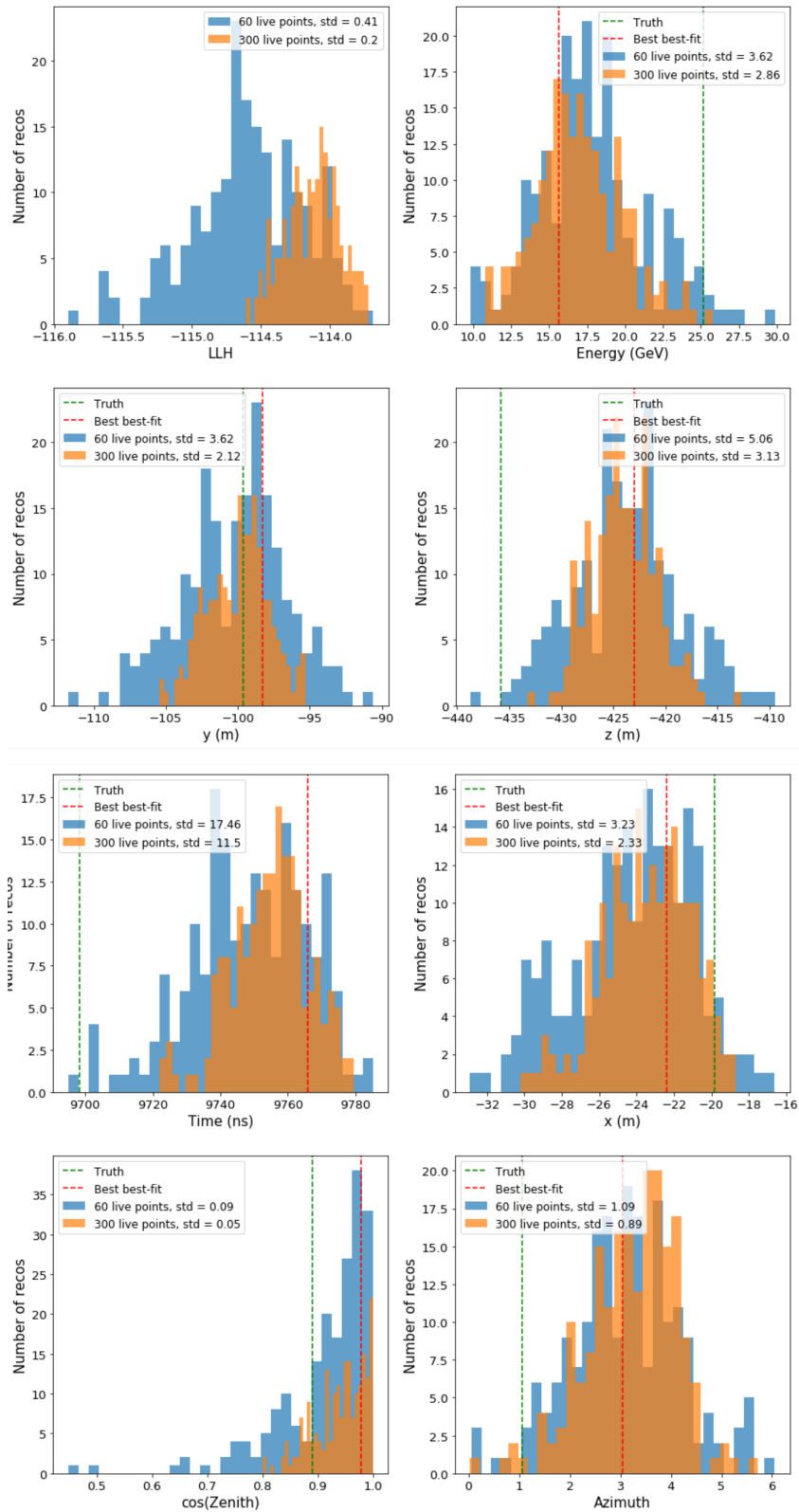


Fig. 6.20.: Multiple reconstructions of the same event and hit pattern. The resolution improves when changing the number of live-points that are evaluated in parallel in MultiNest from 60 (blue) to 300 (orange).

Search for NSI

The aim of this analysis is to find the NSI parameter values for which the expected signal best matches the DeepCore data set described in section 5. It is implemented as a comparison between binned event rates where appearance or disappearance in the individual neutrino channels is observed. This requires four basic elements: An understanding and representation of the signal, a statistical approach to measuring an expectation against data, an uncertainty treatment, and finally an optimization algorithm that finds the best fitting hypothesis within the parameter space. In addition, the statistical rejection significance of hypotheses throughout the NSI parameter ranges is determined in order to set confidence limits on the individual NSI parameters.

The signal is available as event counts per bin, with the binning defined in the observables that best capture Earth matter effects. These are reconstructed energy, zenith angle and interaction type. The binning as well as the expected signal features are discussed in section 7.1.

Comparison between simulated expected bin counts and observed data is done through calculating a test statistic between both binned signal maps. The PINGU Simulation and Analysis (PISA)¹ [120] framework implements this as re-weighting of existing MC sets in order to yield the specific expected bin-wise event counts for any possible tested hypothesis, as will be described in section 7.2.

Each hypothesis consists of NSI and nuisance parameter values. The latter mostly describe systematic uncertainties on the understanding of the detector and physics processes that can impact the expected event rates per bin. Their impact is modeled as a continuous function either of known physics parameters or based on interpolated simulated sets or external data, e.g. from calibration measurements. This is described in sections 7.3 and 7.4.

Nuisance parameters that are found to have significant impact on the analysis outcome are set free in the fitting process that determines the hypothesis for which the expectation best matches the observed data.

In a frequentist approach, each complex or real-valued NSI parameter is considered separately, while keeping all other NSI parameters fixed to 0. This is motivated from computational feasibility as well as comparability to other experiments. In theory,

¹This references PINGU, the Precision IceCube Next Generation Upgrade, a potential low-energy IceCube extension.

multiple non-zero NSI parameter values would be allowed.

With 17 nuisance parameters included and either one or two dimensions of the respective NSI parameter, finding the global minimum of the parameter space is a challenging task. In section 7.6, the complexity of this problem as well as the solution implemented in this analysis are discussed. In addition to the frequentist approach, Bayesian sampling of the overall parameter space is explored as a second fundamentally different technique in section 7.7.

With this overall setup, the NSI parameter values best describing the data, further on called best fit values, can be determined. A comparison between the sensitivity of this analysis and existing results can be found in section 7.8. Furthermore, NSI parameter regions can be excluded, by computing confidence limits within which the true parameter values are expected to be at different confidence levels, which is described in section 7.9. The details of how the latter can be done computationally effectively are given in section 7.9.1.

7.1 Neutrino signal in DeepCore

The signal investigated for this analysis is that of atmospheric neutrinos of all flavors which traverse the Earth before being detected in DeepCore. The rates at which individual flavors are detected change depending on neutrino energies and baselines due to oscillation and matter effects, making this approach sensitive to both PMNS-matrix based oscillations and NSI effects. The cosine of the zenith angle is used as an observable that is associated to the baseline as shown in figure 3.4, with up-going events that cross the entire Earth corresponding to $\cos(\theta) = -1$.

7.1.1 Standard oscillations footprint

Oscillograms showing $P_{\nu_\alpha \rightarrow \nu_\mu}$ for the standard oscillations case can be found in figure 7.1. This shows the expected signal as seen in true event properties, instead of reconstructed analysis observables. The cancellation of effects due to the combined signal of neutrinos and antineutrinos within DeepCore is not taken into account here, as only detections of ν_α are shown and not total of $\nu_\mu + \bar{\nu}_\mu$.

The overall signal structure corresponds to the $\frac{L}{E}$ dependence of oscillations effects. For baselines of the Earth diameter and below, the strongest oscillation signal can be observed at few GeV, close to the lower edge of the DeepCore energy range. At the boundary between outer Earth core and mantle, a sharp increase in electron density

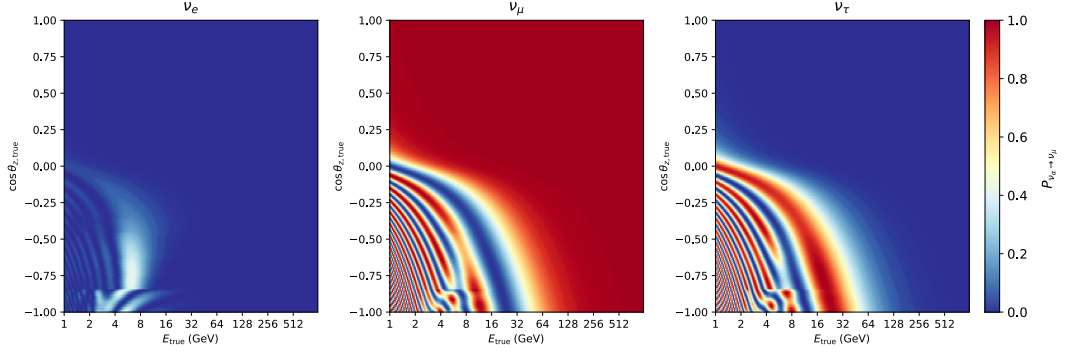


Fig. 7.1.: Probabilities of atmospheric neutrinos produced as the neutrino flavor indicated in the panel titles to be detected in DeepCore as muon neutrinos. This assumes the standard oscillations case with PMNS-matrix parameter values as defined in table 7.3.

occurs. Based on the matter effects discussed in section 3.2.3, this results in the sudden change in features visible at $\cos(\theta) \sim -0.85$.

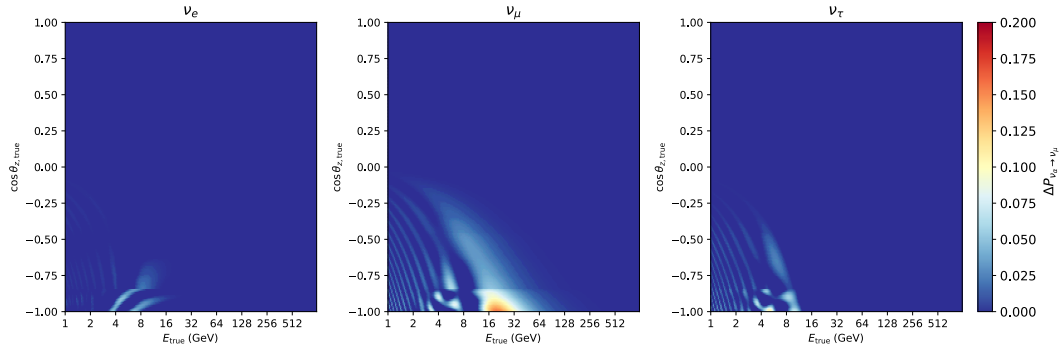
7.1.2 NSI signal expectation

As discussed in section 3.4, the individual variables of the NSI parametrization investigated for this analysis vary strongly with respect to their signal. In figure 7.2, two NSI scenarios are shown with their difference to the null hypothesis which is the standard oscillations case that only includes PMNS-matrix parameter effects. The NSI hypotheses are selected such that they are close to recent DeepCore limits [40]. Oscillograms of $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$, $\epsilon_{e\tau}^{\oplus}$ and $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ hypotheses can be found in the appendix in section B.3.

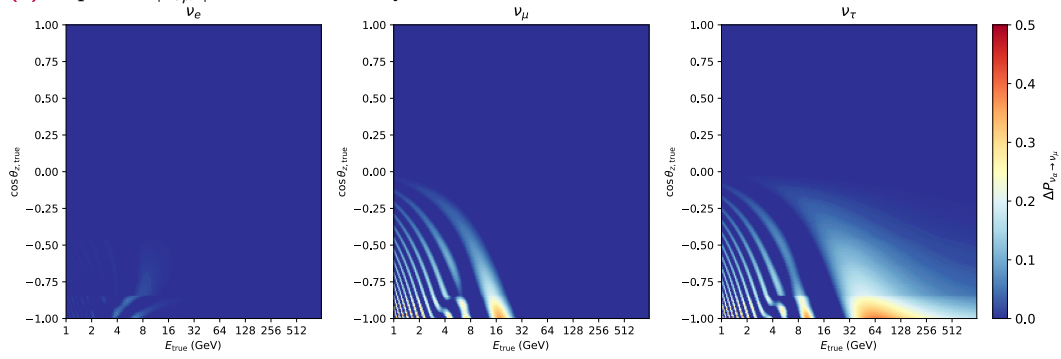
For most of the NSI parameters, the signal region is similar to that of the PMNS-matrix parameters. In contrast, a non-zero magnitude of the $\epsilon_{\mu\tau}^{\oplus}$ parameter yields a signal that exceeds this range in energy. This behavior is utilized for analyses of IceCube data that extends to the TeV range [52], as mentioned in section 3.5.

7.1.3 Binning in reconstructed observables

The level of detail shown in the oscillograms discussed above is not available in reconstructed parameters, due to the finite reconstruction resolution. Therefore, a relatively coarse analysis binning is applied. The observables that the signal is binned in are reconstructed energy, cosine of the zenith angle and PID, which represents the neutrino flavor and interaction type. These are determined per event as described



(a) Impact of $|\epsilon_{e\mu}^{\oplus}| = 0.15$ as the only source of NSI.



(b) Impact of $\epsilon_{\mu\tau}^{\oplus} = 0.02$ as the only source of NSI.

Fig. 7.2.: Difference in oscillation probabilities $\nu_{\alpha} \rightarrow \nu_{\mu}$, induced by NSI effects with respect to the standard oscillations case shown in figure 7.1. The shown NSI cases each include only one non-zero NSI parameter. Values are set approximately to the 90% confidence limit set in a previous DeepCore NSI analysis [40], while all other NSI parameters are set to 0. Note that the color scale does not cover the full range between 0 and 1 and differs between figures.

Tab. 7.1.: Analysis binning of the observables.

	Start	End	Number	Spacing / Edges
E_{reco} [GeV]	5	100	12	logarithmic
$\cos(\theta_{reco})$	-1	0.04	8	linear
PID	0	1	3	0.0, 0.25, 0.55, 1.0

in section 5.2. Note that, other than in the signal shown in the previous section, neutrinos can not be differentiated from anti-neutrinos in reconstructed variables.

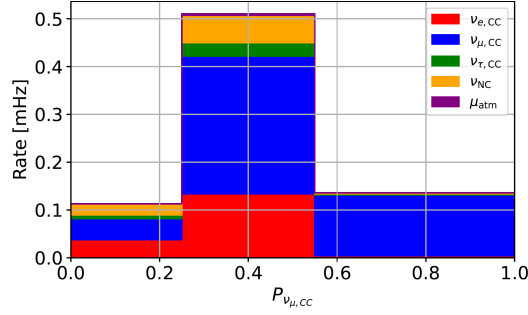


Fig. 7.3.: Separation of neutrino flavors and interactions based on the FLERCNN PID classification and implemented analysis binning. Charged current interactions are subdivided into flavors while neutral current ones are shown in total. Atmospheric muons constitute the only relevant background source in this sample. This plot is adapted from [92].

Reconstructed energies are subdivided into twelve logarithmically spaced bins between 5 GeV and 100 GeV, as specified in table 7.1. While the DeepCore data sample provides data up to 300 GeV, the FLERCNN reconstruction is known to perform poorly above 100 GeV, as discussed in section 5.2.2. As only the $\epsilon_{\mu\tau}^{\oplus}$ signal is present at energies above this, only the sensitivity to this parameter is impacted.

In reconstructed cosine of the zenith angle, eight linearly spaced bins between -1 and 0.04 are defined. This covers all directions from up-going to horizontal. Not including most down-going events is motivated by poor agreement between data and MC for muon-dominated regions. As apparent from the signal characteristics presented in the previous section, this has little influence on the sensitivity to any of the NSI parameters.

As a clean differentiation between track-like and cascade-like events is not feasible throughout the DeepCore energy range, the binning in PID is mostly optimized with respect to analysis performance. The three bins still represent how track-like events are, ranging from 0 for strongly cascade-like events to 1 for strongly track-like ones. In figure 7.3, the three bins are shown with their expected content based on simulation.

Individual bins with low expected event counts are masked out as invalid in order to prevent the resulting large relative uncertainties from impacting the minimizer stability, as can be seen in figure 7.4a. The four bins at lowest reconstructed energy and cosine of the zenith angle are masked due to imprecision of the nuisance parameter treatment in this region, which is discussed in section 7.4.4.

Statistics of atmospheric muon events in the finalized sample are low, with their overall expectation decreased to ~ 1000 . This potentially causes local fluctuations and abrupt changes in bin counts and errors of neighboring analysis bins. To resolve this and smooth the muon distribution, a kernel density estimation (KDE) [121, 122] is applied on the finalized sample. This step is presented in detail in [92] and [83], along with the tests performed to ensure that analyses are not impacted negatively by this.

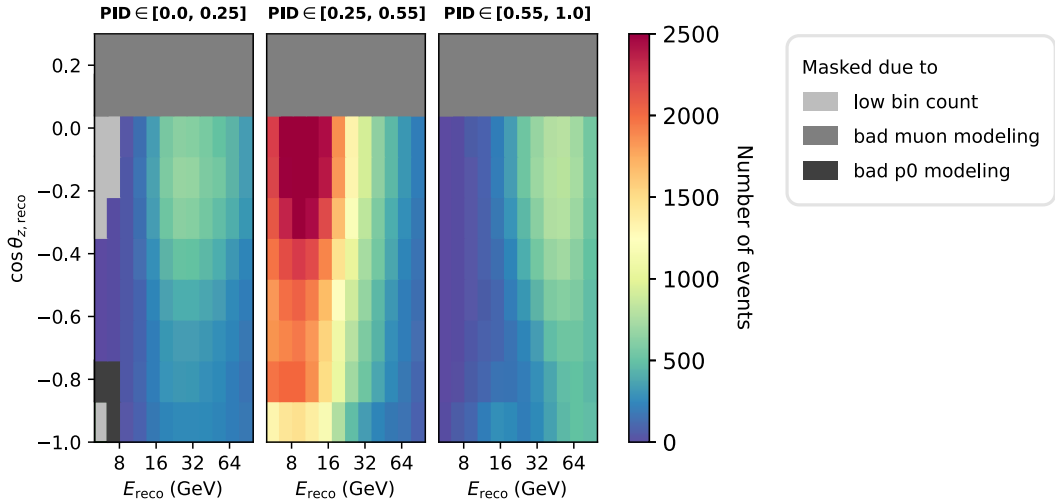
An example for the change in expected event counts due to NSI is shown in figure 7.4b, with the absolute differences amounting to $\mathcal{O}(10)$ at most. A comparison to figure 7.2b, which applies the same hypothesis of $|\epsilon_{\mu\tau}^{\oplus}| = 0.02$, showcases the impact of the limited resolution in reconstructed energy and cosine of the zenith angle.

A detailed study of the optimal bin numbers and edge positions was performed in all three dimensions for the NSI analysis as part of this work. Criteria to assess different binning strategies include not only NSI sensitivity but also computational cost, which scales with the number of individual bins. The bin numbers differ strongly between dimensions due to the different resolution therein, which is lowest in PID. Ultimately, the binning used for analyses of PMNS-matrix oscillation parameters on the same sample was found to perform well, while larger numbers of bins or variations in bin edge position only have small impact on the NSI analysis outcome.

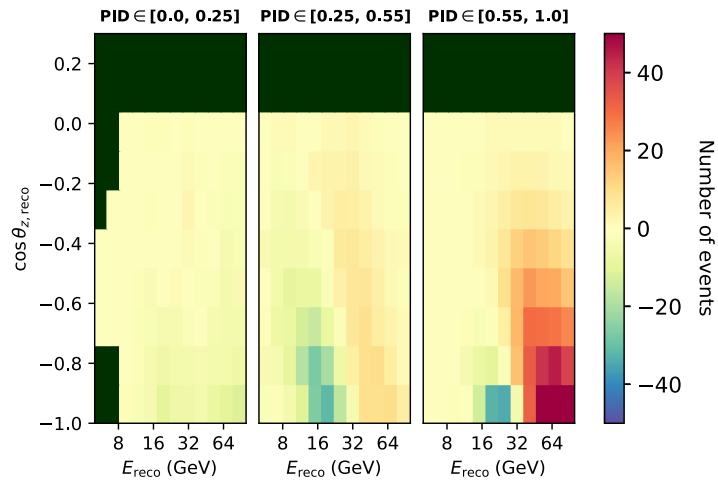
The final analysis binning differs only slightly from the one used in the previous DeepCore NSI analysis [40], specifically in the spacing of the energy bins and the range considered in cosine of the zenith angle. The PID metric differs fundamentally between the two analyses and can not be compared easily. However, using three instead of two PID bins provides additional signal resolution.

7.2 Statistical approach

This NSI search is implemented as a comparison between binned event counts. A histogram of observed data or pseudo-data is compared to one of expected rates that are calculated on what will be called a template for different hypotheses. The best fitting hypothesis is determined by minimizing a test statistic described



(a) Event counts in the standard oscillations case. Masked bins are shown in different shades of gray for different exclusion reasons. The final cut on cosine of the zenith angle is implemented as masking as well.



(b) Difference in expected event counts between the null hypothesis and a case where $|\epsilon_{\mu\tau}^{\oplus}| = 0.02$. Masked bins are shown in dark green.

Fig. 7.4.: Binned event counts expected for the sample used for this analysis, with bins in the reconstructed variables as shown in table 7.1.

in section 7.2.1. Generating expectation histograms is a staged process where a previously generated MC simulated template is re-weighted to represent the respective tested hypothesis, as discussed in section 7.2.2.

A hypothesis consists of physics and nuisance parameters. Details on the latter are given separately, in sections 7.3 and 7.4. The minimization process is discussed in section 7.6.

This general analysis approach is used for current and previous generations of DeepCore NSI [40] and oscillations [83] analyses and can be described as a forward-folded parameter estimation [120].

7.2.1 Test statistic

In order to compare expected and observed event count histograms, their modified Pearson's chi-squared (χ_{mod}^2) [40, 70] test statistic is calculated according to

$$\chi_{mod}^2 = \sum_{i \in \text{bins}} \frac{(n_i^{exp} - n_i^{obs})^2}{n_i^{exp} + (\sigma_i^{exp})^2} + \sum_{j \in \text{prior}} \frac{(\Delta s_j)^2}{\sigma_{s_j}^2} . \quad (7.1)$$

Here, n_i^{obs} is the observed event count in bin i and n_i^{exp} the respective expected event count, including neutrinos and antineutrinos of all flavors as well as background events. The uncertainty on the observation is determined from the error of the expectation as $\sqrt{(n_i^{exp})^2}$.

The expected bin count depends on the tested hypothesis and is calculated from the summed weights of all events contained within the respective bin. In addition to these terms, which are also present in unmodified Pearson's χ^2 [96], the variance σ_i^{exp} of the simulated bin count enters the denominator. It originates from finite statistics in the Monte Carlo event set and is typically smaller than n_i^{exp} . Signal and background terms are included such that

$$(\sigma_i^{exp})^2 = \sum_{i, \text{events}} w_i^2 + \sigma_{i, bg}^2 = \sigma_{i, \nu}^2 + \sigma_{i, bg}^2 . \quad (7.2)$$

The neutrino component of this, $\sigma_{i, \nu}$, originates from the weights w_i of all events contained in bin i . For atmospheric muons, which are the only background source relevant to this work, the variance $\sigma_{i, bg}$ is based on the applied KDE.

The second modification to Pearson's χ^2 is the second sum in equation 7.1. It constitutes a penalty term for those nuisance parameters where prior knowledge is included in form of Gaussian constraints. Deviations Δs_j from the nominal value are penalized based on the respective prior's standard deviation $\sigma_{s_j}^2$.

While for other test statistics, such as χ^2 or likelihood, the evaluation is slightly faster, these lack the error treatment in χ_{mod}^2 and therefore do not account for the non-negligible errors on bin counts which occur for the treatment of nuisance parameters applied in this analysis (see section 7.4.3).

7.2.2 Staged MC re-weighting in PISA

Like other recent DeepCore analyses, this work relies on PISA [120, 123] as the underlying software framework. The fundamental idea of this is to re-weight MC simulated event sets such that they correspond to arbitrary physics and nuisance parameter hypotheses. This effectively provides large MC statistics at each tested hypothesis without the associated computational cost.

The simulated event counts that are re-weighted using PISA are generated as described in section 5.1. The so-called baseline sets for atmospheric neutrinos and muons assume nominal values for all nuisance parameters, which corresponds to the current best knowledge of the involved processes. These event sets are re-weighted in consecutive stages corresponding to the involved physics processes, as shown schematically in figure 7.5.

A subset of nuisance parameters, which is discussed in section 7.5, has significant impact on the analysis outcome. These parameters are included floating freely within the final optimization, which means that the respective parameter value in a corresponding PISA stage can be varied.

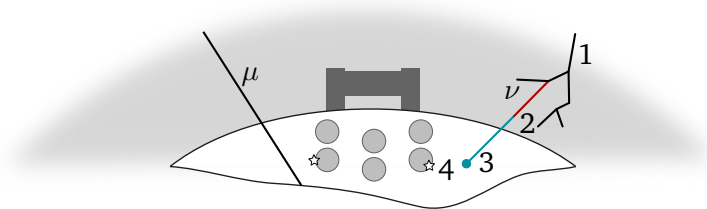


Fig. 7.5.: Schematic of the individual PISA stages. The numbers mark individual stages of re-weighting event counts based on different effects: (1) The atmospheric flux, (2) oscillations and matter effects during neutrino propagation, (3) neutrino interaction characteristics and (4) impact of the ice and detector response on the signal. Neutrinos and atmospheric muons are treated separately, with only a subset of said stages being relevant for muons.

For both, the neutrino signal and muon background, the baseline MC event set is re-weighted separately to fit the assumed flux in stage 1. For neutrino events, the resulting flux is then processed according to the assumed matter- and oscillations effects in stage 2. The subsequent neutrino interactions inside the detector

are weighted with respect to the adopted cross section model in stage 3. Then, in stage 4, changes in detection probability of the produced light due to ice and detector properties are calculated. An overall signal normalization is applied to atmospheric neutrinos and muons likewise, followed by allocation of the individual events to the histogram binning presented in section 7.1.3.

For muon events, due to the involved physics processes as well as their small contribution to the sample, the complexity of this queue is reduced to muon flux calculation and overall normalization.

Most stages operate on individual events while the more computationally expensive neutrino flux and oscillations calculations use internal binning to speed up computation. This reduces the number of weights that need to be calculated by a factor of $\mathcal{O}(10^4)$ from being equal to the number of processed events to the stage-specific number of bins. Optimizing the trade-off between computational cost and accuracy of these calculations motivates finer stage-internal binning than that of the final reconstructed parameters. The details of this internal binning have been evaluated and optimized for this analysis in a respective study. In comparison to event-based weight calculation, no significant negative impact on the analysis could be found for the internal binnings established in other analyses of the same data sample as the numbers of bins in all dimensions were found to be sufficient in all stages where internal binning is applied.

All parameter settings and information on the stages are stored in configuration files which can be found in a github repository created for this analysis [124]. The details on neutrino event modeling and assumptions concerning individual stages and associated sources of systematic uncertainty are given in the following.

7.3 Uncertainties in describing particles

In order to re-weight Monte Carlo sets based on a specific hypothesis, continuous functions in all varied dimensions are required. These are of different origin, depending on whether there are models describing a respective dimension, which is typically the case for particle interaction related dimensions. The models used in this analysis are discussed in the following, separated into neutrino flux modeling in section 7.3.1, oscillation effects in section 7.3.2, cross-sections in section 7.3.3 and overall signal and background normalization parameters in section 7.3.4.

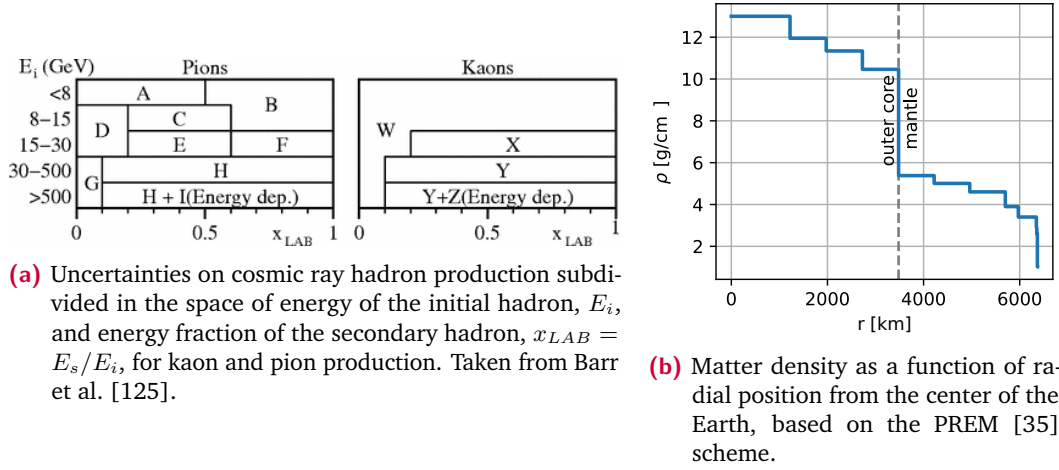


Fig. 7.6.: Depictions of assumptions going into the flux and oscillation modeling.

7.3.1 Flux modeling

In the first step of re-weighting, the atmospheric lepton flux is calculated using the model described by Honda et al. [41] as the baseline. Uncertainties arise based on the primary cosmic ray spectrum as well as the showers of secondary particles that cosmic rays produce. The latter concern properties of the atmosphere and of hadronic interactions. Multiple models are taken into account to ensure a preferably generally valid evaluation of the associated uncertainties, by calculating their outcome using the Matrix Cascade Equations (MCEq) [126] tool.

The main uncertainty of the primary cosmic ray spectrum [125] is modeled as a power-law correction $\Delta\gamma$ that directly impacts the flux of atmospheric leptons as

$$\Phi_{\text{mod}_{\nu/\mu}} = \Phi_{\nu/\mu} \left(\frac{E}{E_{\text{pivot}}} \right)^{\Delta\gamma_{\nu/\mu}}. \quad (7.3)$$

The pivot energy is set to 24 GeV in order to reduce the impact of the correction of the flux scale. This will compensate any normalization offset and is left as free parameter without a prior. Note that, while the structure of the correction is the same, neutrinos and muons are considered separately, with independent respective nuisance parameters $\Delta\gamma_{\nu}$ and $\Delta\gamma_{\mu}$.

Additional uncertainties on the neutrino flux, originating from hadron production, are described using phenomenological models [87, 125, 127]. The scheme shown in figure 7.6a shows the approach of Barr et al. [125] to segment the space defined by the initial hadron energy and relative secondary hadron energy into independent nuisance parameters. These differ for pions, kaons and their respective anti-particles. In case of pions and anti-pions, this difference amounts to a scaling factor that

Tab. 7.2.: Flux related nuisance parameter nominal values and, if applicable, Gaussian prior widths, based on [92]. The last column gives the status of the variable in the final minimization configuration. Note that, while most values given for Barr parameters use the prior ranges specified in the original paper, [125], the prior given for `barr_z_K` and `barr_z_antiK` is increased by a factor of 5 with respect to this, as it was found in IceCube internal simulation to be underestimating the actual variance.

Parameter	Nominal value \pm prior	Boundaries	Status
$\Delta\gamma_\nu$	0.0 ± 0.1	$\pm 3\sigma$	Free
$\Delta\gamma_\mu$	0.0 ± 1.0	$\pm 3\sigma$	Fixed
E_{pivot} [GeV]	24.09	-	Fixed
π^+/π^-	0.0 ± 0.05	$\pm 5\sigma$	Fixed
<code>barr_a_Pi</code>	0.0 ± 0.1	$\pm 3\sigma$	Fixed
<code>barr_b_Pi</code>	0.0 ± 0.3	$\pm 3\sigma$	Fixed
<code>barr_c_Pi</code>	0.0 ± 0.1	$\pm 3\sigma$	Fixed
<code>barr_d_Pi</code>	0.0 ± 0.3	$\pm 3\sigma$	Fixed
<code>barr_e_Pi</code>	0.0 ± 0.05	$\pm 3\sigma$	Fixed
<code>barr_f_Pi</code>	0.0 ± 0.1	$\pm 3\sigma$	Fixed
<code>barr_g_Pi</code>	0.0 ± 0.3	$\pm 3\sigma$	Fixed
<code>barr_h_Pi</code>	0.0 ± 0.15	$\pm 3\sigma$	Free
<code>barr_i_Pi</code>	0.0 ± 0.122	$\pm 3\sigma$	Free
<code>barr_w_K</code>	0.0 ± 0.4	$\pm 3\sigma$	Fixed
<code>barr_x_K</code>	0.0 ± 0.1	$\pm 3\sigma$	Fixed
<code>barr_y_K</code>	0.0 ± 0.3	$\pm 3\sigma$	Free
<code>barr_z_K</code>	0.0 ± 0.61	$\pm 3\sigma$	Free
<code>barr_w_antiK</code>	0.0 ± 0.4	$\pm 3\sigma$	Fixed
<code>barr_x_antiK</code>	0.0 ± 0.1	$\pm 3\sigma$	Fixed
<code>barr_y_antiK</code>	0.0 ± 0.3	$\pm 3\sigma$	Fixed
<code>barr_z_antiK</code>	0.0 ± 0.61	$\pm 3\sigma$	Fixed

corresponds to the pion ratio π^+/π^- , which allows for a reduction of the overall number of nuisance parameters. The resulting considered flux based nuisance parameters and their status in the final minimization are shown in table 7.2, along with their nominal values and the widths of the associated Gaussian priors.

The boundaries for all parameters modeling pion and kaon flux are set to $\pm 3\sigma$, which constitutes a difference with respect to other analyses using the same data sample. This is motivated from the fitting performance, which increases for a smaller parameter space. Such a change is assumed not to impact the analysis outcome, since fitting nuisance parameter values outside of this range would in any case be considered an indication of poor agreement between model and data and lead to re-evaluation of the analysis.

Tab. 7.3.: Nominal values and allowed ranges of PMNS-matrix parameters in both mass orderings, as used for this analysis. Oscillation parameter values are based on the NuFIT 4.0 results [130, 131] including Super-K atmospheric data.

Parameter	Nominal value \pm prior	Boundaries	Status
Y_e^I	0.4656	-	Fixed
Y_e^O	0.4656	-	Fixed
Y_e^M	0.4957	-	Fixed
Detector depth [km]	2.0	-	Fixed
Production height [km]	20.0	-	Fixed
θ_{23} [°]	45.3637	[20, 70]	Free
θ_{12} [°]	33.82	-	Fixed
δ_{CP} [°]	0	-	Fixed
Δm_{21} [eV ²]	0.0000739	-	Fixed
NO:			
θ_{13} [°]	8.61 ± 0.13	-	Fixed
Δm_{31} [eV ²]	0.00247996	[0.0015, 0.003]	Free
IO:			
$\sin^2(\theta_{13})$ [°]	8.65 ± 0.13	-	Fixed
Δm_{31} [eV ²]	-0.00247996	[-0.003, -0.0015]	Free

7.3.2 Oscillation and matter effect implementation

Once the atmospheric neutrino flux is known, matter and oscillations effects on it are calculated using an adaption of the prob3 [128] framework, which implements the approach taken in [129].

Effects of PMNS-matrix parameters are computed using the three flavor model introduced in section 3.2. For this analysis, they are included as nuisance parameters with no Gaussian priors. The nominal value of each as well as the range that it is allowed to vary in are shown in table 7.3 for both mass orderings. As the choice of mass ordering was found to have only a small impact on the NSI analysis outcome, all subsequent studies are shown for normal ordering.

In addition to the null hypothesis case, which includes Earth matter effects due to electron neutrino CC interactions, NSI effects are implemented as an extension to prob3. Their analytical formulation follows the parametrization discussed in section 3.3, which yields two real-valued FD and three complex-valued FC parameters. The PMNS-matrix oscillations hypothesis is hence nested within the NSI parameter space, corresponding to the case where all NSI parameter values are 0. A first implementation of this was performed for the previous DeepCore NSI analysis [40] and, as a part of this work, ported to the current PISA code base [132].

The electron and hence matter density profile of the Earth and atmosphere is an

input parameter for any matter effects. Atmospheric neutrinos are assumed to be produced at a height of 20 km [41] above ground and reach the detector at a depth of 2 km in good approximation. The matter density profile of the Earth is modeled as twelve concentric spherical shells with constant density each. The respective values, adapted from the PREM scheme introduced in section 3.2.3, can be found in figure 7.6b. Note that the largest offset in density between two subsequent layers is at the boundary of outer core and mantle, inducing the neutrino oscillation behavior discussed in section 7.1.

The radially dependent electron number density,

$$N_e(r) \propto Y_e(r)\rho(r) \quad , \quad (7.4)$$

depends on the matter density $\rho(r)$ as well as the matter composition through the relative number density of electrons per nucleon $Y_e(r)$. This work, like previous analyses [40], assumes an electron density that corresponds to iron $Y_e^I = Y_e^O = 0.4656$ for the inner and outer core and $Y_e^M = 0.4957$ for the Earth mantle.

For NSI, not only the electron number density but also that of up- and down-quarks is relevant. As the NSI parametrization applied in this analysis does not differentiate between coupling to individual fermions, all fermion number densities can be represented through that of electrons as shown in equation 3.30.

7.3.3 Cross section uncertainties

Neutrino interactions in the instrumented ice are subject to uncertainties concerning the parton distribution functions within the target nucleons which are given in table 7.4. These differ for the individual processes, namely deep inelastic (DIS), resonant (RES) and quasi-elastic scattering (QES), which are detailed in section 2.

Uncertainties in the cross sections of CC quasi-elastic scattering and resonant neutrino production are introduced in the GENIE [84] software through re-weighting with the axial mass form-factors M_A^{CCQE} and M_A^{CCRES} . The parton distribution function that GENIE is based on [133] is sufficiently accurate at the energies where RES and QES are dominant.

Deviations between the parton distribution functions used in GENIE and more recent measurements, however, need to be taken into account at energies where deep inelastic scattering becomes relevant. Of the implementations of parton distribution functions that were tested, the CSMS [134] model yields the largest differences with respect to GENIE. In current analyses such as this work, a correction is applied to interpolate between the GENIE and CSMS cross sections using the DIS CSMS

Tab. 7.4.: Nuisance parameters considered with respect to neutrino cross sections, along with their priors and allowed ranges as well as their status during minimization.

Parameter	Nominal Value	Boundaries	Status
M_A^{CCQE} (in σ)	0.0 ± 1.0	$\pm 2\sigma$	Free
M_A^{CCRES} (in σ)	0.0 ± 1.0	$\pm 2\sigma$	Free
DIS CSMS	0.0 ± 1.0	$\pm 3\sigma$	Free

Tab. 7.5.: Nuisance parameters that are constant scaling factors, along with their priors and allowed ranges as well as their status during minimization.

Parameter	Nominal Value	Boundaries	Status
N_ν	1	[0.5, 2]	Free
N_μ	1 ± 0.4	[-2.25σ , $+5\sigma$]	Free
$N_{\nu\tau}$	1	[0.0, 3.0]	Fixed
$N_{\nu,NC}$	1 ± 0.2	[0.5, 1.5]	Fixed

nuisance parameter [83, 115]. In recent studies, an error in the GENIE calculation of the charm quark fraction was found to be the origin of this behavior. The DIS CSMS nuisance parameter will, as a result, likely not be necessary in future DeepCore analyses.

7.3.4 Overall signal normalization

In order to account for any effects that impact all bin counts for a specific event kind equally, constant normalization factors are introduced as nuisance parameters. These apply to the overall neutrino event rate N_ν , that of NC interactions $N_{\nu,NC}$ and atmospheric tau neutrinos $N_{\nu\tau}$. The cosmic muon normalization, N_μ , is found to have negligible impact due to the muon contribution to the overall sample being below 1 %.

The fact that the overall signal normalization is allowed to vary in the analysis makes it sensitive to relative signal changes only. In other analyses using the same data sample, the lower boundary of the respective nuisance parameter is at $N_\nu = 0.1$. This was changed to 0.5 for this analysis, as allowing such an extreme parameter range is not expected to be representative of the true uncertainty on this parameter.

7.4 Parametrizing the detector

For all DeepCore analyses, the most relevant group of systematic uncertainties and corresponding nuisance parameters concerns the understanding of the detector itself.

This concerns characteristics of light propagation in the different kinds of ice, such as the pristine glacier ice, which is described in section 7.4.2. The re-frozen hole ice is treated in a joint description with the detector instrumentation, as discussed in section 7.4.1. The nominal values and ranges of all respective nuisance parameters are given in table 7.6.

In contrast to modeling particle interactions, for these detector parameters, evaluating the impact in a closed functional form is computationally intractable. Instead, Monte Carlo simulated sets are generated at hypotheses spanning the parameter space and interpolated between. This approach is described in section 7.4.3 and evaluated in section 7.4.4.

For the neutrino signal, the impact of the respective nuisance parameters has large effects at the analysis level. In contrast, the contribution of atmospheric muon events in the sample is sufficiently small to make detector related uncertainties on it negligible.

7.4.1 Module and hole ice characteristics

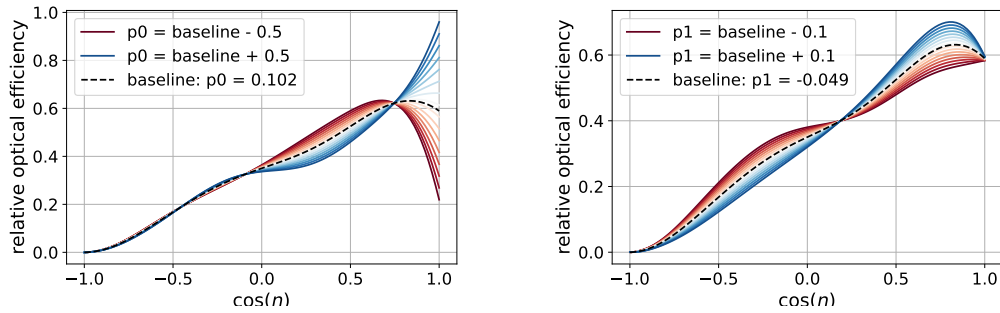
Detector and ice properties are measured mostly in flasher calibration runs. The calibration based understanding of the detector characteristics is included in the data sets for each detector run.

The overall scale of the DOM efficiency after calibration, ε_{dom} , accounts for any uncertainties that affect all modules alike. It is studied using minimum ionizing muons [74] and modeled in DeepCore analyses as a nuisance parameter.

The DOMs angular acceptance is considered equal for all modules and the applicable uncertainties are modeled as part of the hole ice parametrization. As described in section 4.4.3, hole ice is the melted and re-frozen ice column in which DOMs have been deployed. Its optical properties differ strongly from those of the surrounding pristine glacial ice, especially at its center. There, the bubble column shades off parts of the photocathode region. A hole ice model can therefore be considered a description of the relative uncertainty on DOM angular acceptance.

The current hole ice model [83, 135] unifies multiple data sets and previous descriptions by spline interpolating them and applying principal component analysis (PCA) [136]. The features and range of all ingoing curves are found to be describable at high accuracy with two parameters, p_0 and p_1 , the variations of which are shown in figure 7.7. Since p_0 and p_1 are eigenvectors of the covariance basis of the underlying models they are uncorrelated, which makes them especially suitable as nuisance parameters.

The behavior of p_0 , which mostly impacts the efficiency at incidence angles where



(a) Scan of the hole ice parameter p_0 within ± 0.5 of its baseline value, with p_1 fixed to its baseline. (b) Scan of the hole ice parameter p_1 within ± 0.1 of its baseline value, with p_0 fixed to its baseline.

Fig. 7.7.: Variations of the two free parameters of the PCA based hole ice model. For all realizations, the resulting relative optical DOM efficiency is shown over the cosine of the incidence angle η . The plots are adapted from [135].

$\cos(\eta) \sim 1$, suggests interpreting it as DOM acceptance in the forward direction with respect to the cathode orientation. Detection of light under these incidence angles is strongly impacted by hole ice properties, since the central part of it is assumed to induce the most scattering.

7.4.2 Bulk ice

The photon absorption and scattering properties of the pristine glacial ice outside the IceCube boreholes are modeled based on calibration data as shown in figure 4.12. The SPice 3.2.1 model which is used for this analysis is based on flasher data and includes Mie scattering [67]. A constant optical anisotropy of 6.9% at all depths is assumed [73], the axis of which coincides with the ice flow direction [74]. The main systematic uncertainties on this approach are associated to the flasher LED light emission and DOM characteristics. To account for them, two global scaling factors on scattering and absorption are introduced as nuisance parameters with Gaussian priors. While the nominal ice model is produced with both scaling factors at 1.0, further calibration studies suggest a scattering factor of 1.05 [83], which is further on set as the nominal value.

In the SPice 3.2.1 model, anisotropy is empirically modeled purely as a direction dependent modification. A better understanding of its origin from birefringence effects was only achieved recently [74]. Subsequently, the birefringence ice model (BFRv2) was constructed by including these effects ab initio with parameters derived in fits to flasher data [76].

While it is desirable to include this new understanding of light propagation in this

Tab. 7.6.: Detector and ice related nuisance parameter values, along with their priors and allowed ranges where applicable.

Parameter	Nominal Value	Boundaries	Status
ϵ_{DOM}	1.0 ± 0.1	[0.8, 1.2]	Free
hole ice p_0	0.101569	[-0.6, 0.5]	Free
hole ice p_1	-0.049344	[-0.15, 0.04]	Free
bulk ice absorption	1.0 ± 0.05	[0.9, 1.1]	Free
bulk ice scattering	1.05 ± 0.1	[0.85, 1.25]	Free
N_{bfr}	0.0	[0.0, 1.0]	Free

analysis, using the BFRv2 ice model as the baseline assumption in place of the SPice 3.2.1 model has multiple prerequisites that could not be met within the timeline of this thesis. These include a detailed study of the uncertainties on parameters in the BFRv2 ice model. Based on the outcome of this, Monte Carlo event sets would need to be simulated.

The uncertainty within the underlying ice model is therefore included in form of an effective nuisance parameter, N_{bfr} , that linearly interpolates between the bin counts nominally expected when assuming SPice 3.2.1 and BFRv2.

The common underlying flasher data measurements used for modeling birefringence and hole ice result in an overlap between the observed effects. This is reflected in correlation between the respective ice nuisance parameters. The nominal ice nuisance parameter values, which assume the SPice 3.2.1 model, can therefore be presumed to absorb some of the birefringence induced behavior. It hence seems plausible that the expected variation of N_{bfr} lies within the range defined through the SPice 3.2.1 and BFRv2 ice models.

7.4.3 Nuisance parameter interpolation

For detector related nuisance parameters that are continuous, ab initio modeling of the impact is computationally not feasible. In order to include them, expected event rates are calculated based on MC sets that are generated at discrete hypotheses throughout the detector nuisance parameter space. The parameter values at which each set is generated can be found in the appendix in section B.4. There are multiple ways to use these sets to obtain expectation values for the event counts at a given set of detector nuisance parameter values, which then are applied in the detector modeling stage of PISA.

For standard oscillations analyses, the expectation values from these simulated sets are linearly interpolated in each nuisance parameter in every analysis bin. The bin-wise revision is necessary as the nuisance parameter impact is known not to be

linear and therefore needs to be evaluated locally. The result are functions of the detector parameters for each bin in the analysis histogram. These functions return scaling factors by which the event count has to be multiplied in order to approximate the effect of the variations in the detector properties. In this way, event counts are adjusted to include the understanding of the impact of the respective detector nuisance parameter values. The functions are also referred to as hyper-surfaces [83]. This method operates on the bin counts expected at the null hypothesis and has been found to artificially introduce correlations between detector and physics effects. This makes it necessary to interpolate the hyper-surfaces in the relevant oscillation-related parameters. For BSM-physics analyses with additional parameters, this approach is computationally intractable, since it results in large hyper-surface file sizes and accordingly large memory requirements in computation.

The alternative strategy applied in this analysis uses likelihood-free inference and is briefly described in the following. More detail can be found in [137].

The aim of this treatment of nuisance parameters is to provide adapted weights for the individual events. These weights describe how likely each event is for any possible realization of detector-related nuisance parameters. This implies to find weights $w_{i,h}$ for each event i in the baseline MC set such that the overall even distribution matches a hypothetical set simulated at hypothesis h that includes off-nominal nuisance parameter values. This needs to be done in a way that is independent of assumptions that are made when generating the underlying MC simulated sets, such as the physics hypothesis or values of other nuisance parameters. In a case where the impact of all nuisance parameters was known analytically, weights could be calculated from the probability ratio $\frac{P(i|h)}{P(i|h_0)}$ with respect to the null hypothesis h_0 . While $P(i|h)$ is not known, the posterior probability $P(h|i)$ can be obtained from a classifier that derives the probability of an event to occur in a set of events simulated at nuisance parameter hypothesis h . Through this method of likelihood-free inference [96], the weights can be inferred as

$$w_{i,h} = \frac{P(i|h)}{P(i|h_0)} \frac{P(h)}{P(h_0)} \quad (7.5)$$

$$= \frac{P(h|i)}{P(h_0|i)} \quad (7.6)$$

with the second term in equation 7.5 serving normalization. The derivation of equation 7.6 uses Bayes' theorem,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (7.7)$$

The classification method used to calculate the posterior probabilities in equation 7.6 is the k nearest neighbors method. These derived posterior distributions allow for the events in the baseline Monte Carlo set to be re-weighted in such a way that they approximate the distribution of events in any of the off-nominal Monte Carlo sets [137]. In order to be useful for the analysis, these weights need to be interpolated between the discrete points. This interpolation is done by calculating gradients in logarithmic space for every event with respect to the detector parameters. These gradients are polynomials with two degrees of freedom for all nuisance parameters except N_{bfr} . For this, one degree is used, as Monte Carlo sets are only available at two points in this dimension, namely the two ice models. The final weights are calculated as a sum of exponential functions. They are based on gradients g at the nuisance parameter value difference $\Delta\Theta$ with respect to the nominal value as

$$w_i(\Delta\Theta) = \exp\left(\sum_j \Delta\Theta_j g_{i,j}\right), \quad (7.8)$$

for event i , using the total set of events j .

The Monte Carlo sets considered for this analysis are given in table B.5 in the appendix. Not all off-nominal MC simulated sets were included in generating the gradients used for this analysis. Nuisance parameter hypotheses that result in expected event rates which differ too strongly from those at the baseline hypothesis were found to introduce inconsistencies in the resulting gradients. This can be explained through fluctuations encountered when applying large weights. As fitting such extreme nuisance parameter values in the final analysis would in any case be considered a sign of bad description of the data and lead to further investigation, disregarding the respective MC sets is not concerning.

Of the sets that use the BFRv2 ice model, only the baseline set, in which all other nuisance parameters are set to their nominal values, is included. The impact of these choices is assessed in the following.

With this nuisance parameter treatment, the errors on expected bin counts are found to be sufficiently large to introduce a significant difference between evaluation based on different test statistics, namely modified and unmodified χ^2 , which were introduced in section 7.2.1. This is shown in figure 7.8 for a comparison between the baseline hypothesis and different DOM efficiency values. Due to this significant impact of bin count errors, χ_{mod}^2 is used as the test statistic in this work.

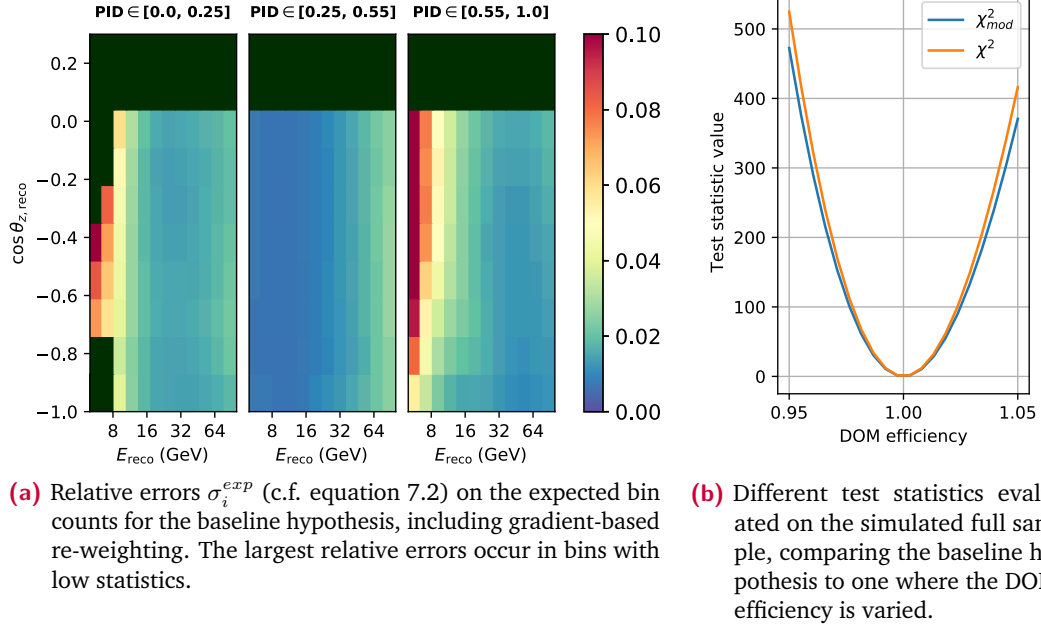


Fig. 7.8.: Impact of the errors on bin counts on the test statistic, given the applied nuisance parameter treatment.

7.4.4 Evaluating the detector uncertainty treatment

In order to verify that this treatment performs as intended, multiple metrics are assessed. Firstly, gradient results have been compared to the hyper-surface approach as spot-checks, showing that similar bin counts are obtained. In addition, the statistical pulls per analysis bin are evaluated as a measure of the difference between, firstly, MC simulated off-nominal sets, and secondly, the baseline set, which is re-weighted for the same nuisance parameter hypothesis. The pulls p are calculated from the weighted bin counts $n_{\text{off-nominal}}$ and n_{baseline} that are expected based on the off-nominal Monte Carlo set and the gradient-based re-weighted baseline set, respectively. The respective Monte Carlo statistics per bin, $s_{\text{off-nominal}}$ and s_{baseline} , are included as well, resulting in

$$p = \frac{n_{\text{off-nominal}} - n_{\text{baseline}}}{\sqrt{s_{\text{off-nominal}} + s_{\text{baseline}}}} \quad (7.9)$$

The respective distributions for the total neutrino flux in six of the tested simulated sets are shown in figure 7.9, including the detailed maps as well as histograms of the pulls. The nuisance parameter values at which the shown Monte Carlo sets are generated are listed in table 7.7. Summary plots of the distributions for all sets that are included in interpolating the detector nuisance parameters as well as further tested sets can be found in appendix section B.5.

For all tested MC sets, the pull distribution standard deviations are close to 1. The overall average is slightly negative, which is understood to originate from bins with low statistics and does not affect the analysis, as the overall normalization is kept free. This shows that the nuisance parameter impact is well modeled, even for hypotheses that are not directly included in generating the interpolation.

As expected for weights that well reproduce the expectation, the pull distributions mostly don't show prominent clustering. One feature is present in sets such as the shown sets 0105 and 1126, which include values below nominal of the p_0 hole ice parameter. A cluster of bins with larger event counts than expected is observed at the lowest cosine of the zenith angle, energy and PID values. This region is where p_0 has the largest impact on the bin count, as this parameter acts as a modulation to the DOM forwards acceptance. Further sets showing this behavior can be found in appendix section B.5.

The observed underestimation of the p_0 impact is likely an edge effect in the interpolation. This interpretation is supported by figure 7.10, in which the interpolated curves in two example bins are shown for all detector nuisance parameters. The first bin shown therein, in the panels on the left side, is located at the overall lowest energy, cosine of the zenith angle and PID, as indicated in figure 7.9. The incorrectly modeled impact of the lowest p_0 values is visible clearly therein from the discrepancy between Monte Carlo sets and interpolation. The bin shown in the panels on the right-hand side is located approximately at the center in all dimensions and does not show similar effects.

In order to circumvent that this behavior affects the analysis outcome, the four clearly compromised bins are excluded, as shown in figure 7.4a. As the expected bin counts in these bins are extremely small (below 100) the impact of this on the overall sensitivity of the analysis is assumed to be small.

The gradients shown in figure 7.10 partly show extreme behavior towards the edges of parameter ranges. This is only the case in regions where no off-nominal Monte Carlo sets are available. Due to the calculation of the weights from a sum of exponential functions given in equation 7.8, extreme behavior of the weights is possible in unrestricted parameter ranges. For this reason, the parameter ranges allowed in minimization within or negligibly beyond the values applied in the available off-nominal Monte Carlo sets. An example for this is the bulk ice scattering parameter, for which an upper limit of 1.15 is set in this analysis.

An additional feature that is apparent from figure 7.9 is that the mean values of the pull distributions of some sets are shifted towards positive or negative values. The extent of these shifts is significantly larger than the stochastic error of the mean. The implications of this can be assessed based on the shift that this induces in the

Tab. 7.7.: Nuisance parameter values at which the Monte Carlo sets shown in figure 7.9 are generated. Dashes indicate that the nominal value of a parameter is used. The respective tables for further Monte Carlo sets can be found in section B.4 of the appendix.

Set number	Ice model	ϵ_{DOM}	p_0	p_1	Absorption	Scattering
0105	-	0.93	-0.3729	0.0349	-	-
0107	-	1.03	0.1244	-0.1132	-	-
0500	-	-	-	-	1.05	1.1 -
1122	spice_bfr-v2	-	-	-	-	-
1126	spice_bfr-v2	-	-0.2	-	-	-

nuisance parameter values. For a set that is produced at a hypothesis where the off-nominal nuisance parameters n are at value η , the value at which the interpolation is evaluated is varied to η' such that the interpolation prediction matches the tested set. A schematic and realization of this is presented in section A.5 in the appendix, for set 0500, where ice scattering and absorption are varied. Since $\eta' - \eta$ is significantly below the scale of the prior of n and the fitted nuisance parameter values are no observables in this analysis, the shift of the mean can be considered not relevant. Apart from the above described behavior, no inconsistencies between the gradient-based expectation and the off-nominal Monte Carlo sets are present in the tested sets. The parameter space that the interpolated sets span is therefore considered to be modeled well.

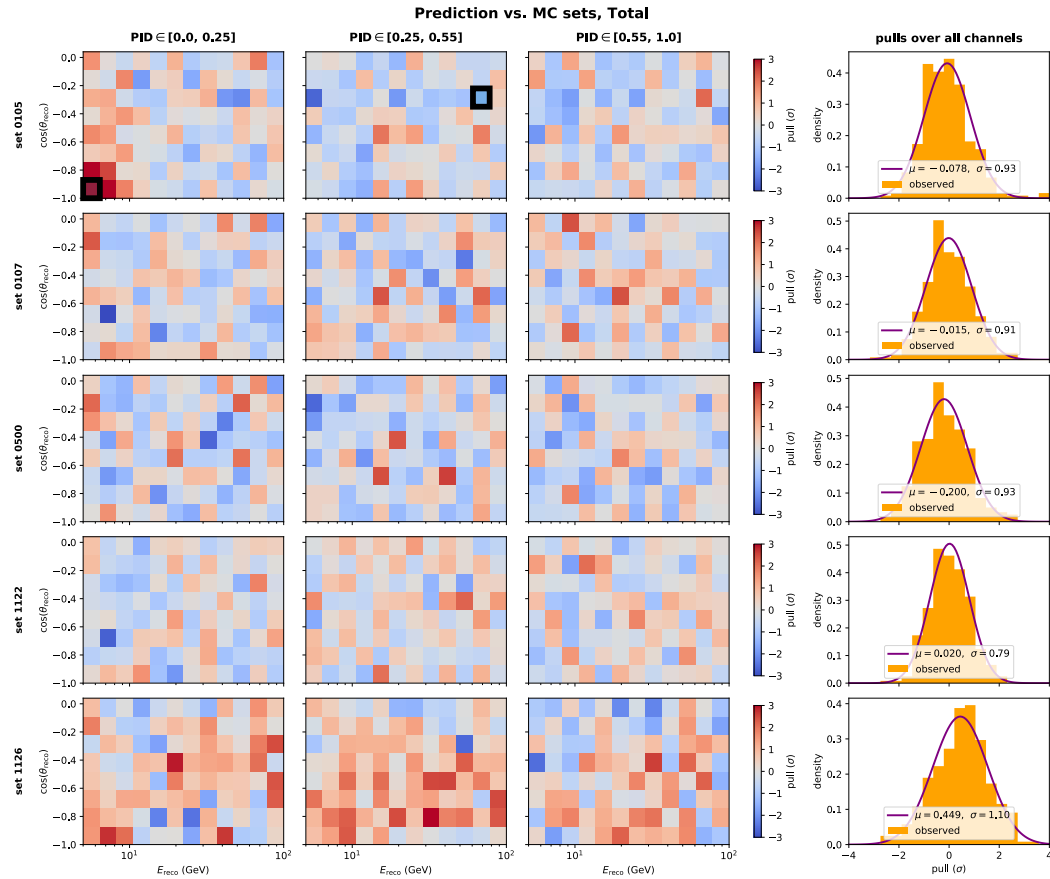


Fig. 7.9.: Maps and histograms of bin-wise pulls in the total neutrino event counts for a selection of simulated sets. All distributions are centered approximately around 0 and show a standard deviation close to 1. The clustering of positive pulls at low PID, energy and cosine of the zenith angle values is related to inaccuracies in the hole ice modeling and motivates excluding the respective bins. The bins for which the gradients are shown in figure 7.10 are indicated as black rectangles in the uppermost shown sample. All sets except 1126 were included in constructing the interpolation. This selection of sets covers variations in all interpolated dimensions. The nuisance parameter settings for each shown set can be found in table 7.7. Maps for further sets can be found in section B.5 of the appendix.

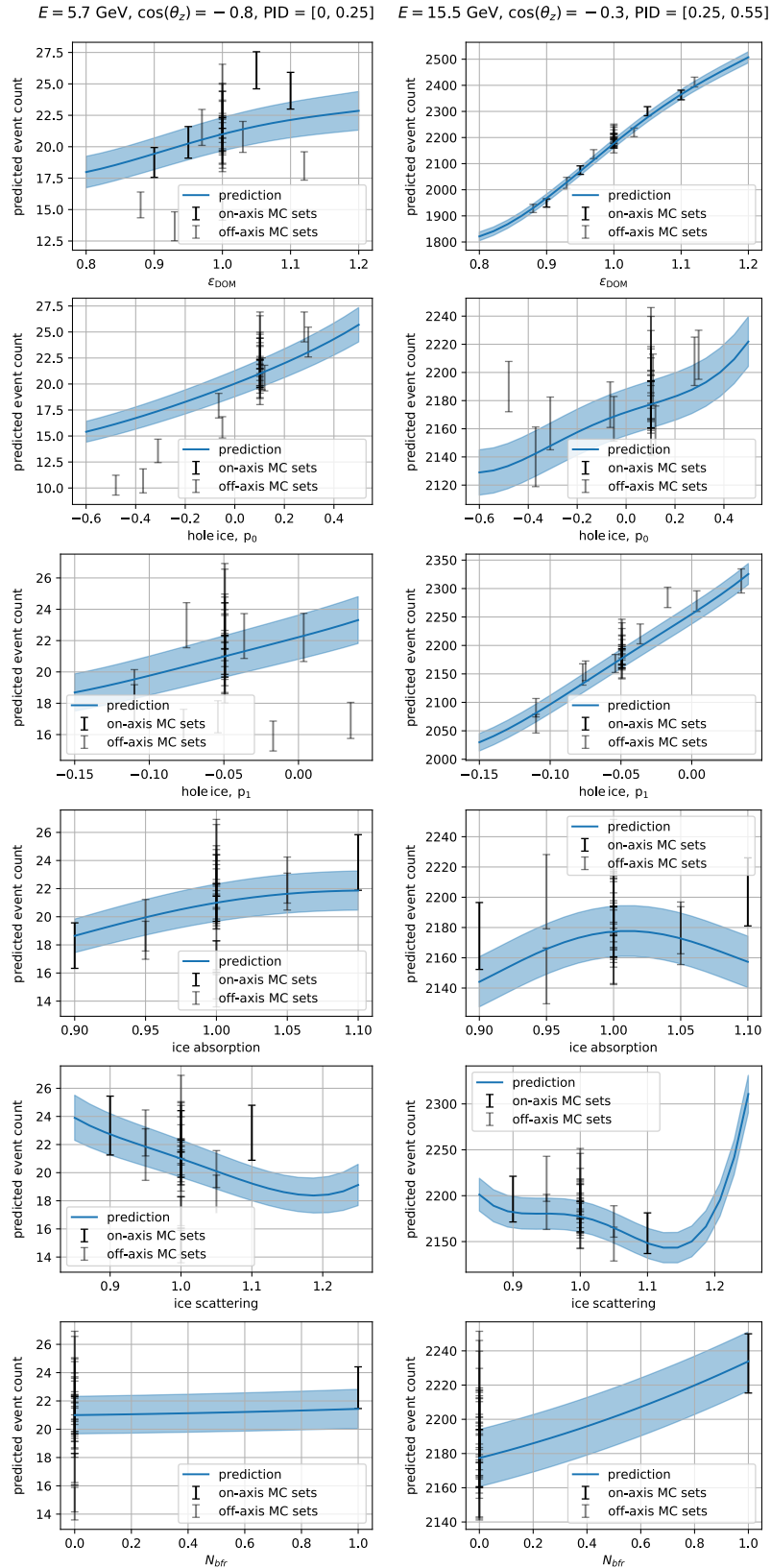


Fig. 7.10.: Detector nuisance parameter interpolation in two exemplary bins. The bin positions are given in the titles and indicated in figure 7.9. The curves and their standard deviation are shown in blue and light blue, respectively. Simulated sets included in the generation are shown as vertical error bars. On-axis sets, where only one nuisance parameter is varied, are shown in black and off-axis ones, with two or more varied parameters, in gray.

7.5 Impact of individual nuisance parameters

Not all of the nuisance parameters that are modeled are found to have an impact on the analysis outcome that is significant enough to motivate leaving them free floating in the minimization. Since each added dimension complicates the minimization process further², only including relevant parameters is essential.

The parameter impact is evaluated by exploring a hypothetical scenario where one nuisance parameter is fixed to its nominal value during minimization while its assumed true value is significantly off-nominal. The resulting impact on the fitted NSI parameter value illustrates an extreme case of how the analysis outcome is affected if the respective nuisance parameter is disregarded.

In order to test for this consistently for different nuisance parameters n , the off-nominal value η to which n is fixed during the fit needs to represent the expected variation of n . If a Gaussian prior is given, η corresponds to positive 1σ with respect to the nominal value. In cases where only a range is set for n , η is centered between the nominal value and the upper border of the range. If these values do not represent the expected variation of a nuisance parameter well, η is set to what is specified in appendix section B.6.

A set of nuisance parameters that were found in previous studies to likely be relevant are kept free throughout the fits. Their possible correlations with n are therefore included in the test.

At the fixed value of n , for each NSI parameter and at multiple true values thereof, two fits are performed. For one, the so-called free fit, the NSI parameter is kept free, while being fixed to its true value in the second, so-called null fit. The difference between the test statistic values obtained in the two fits is the so-called mis-modeling m , such that

$$m = \Delta\chi_{mod}^2 = \chi_{mod, \text{null fit}}^2 - \chi_{mod, \text{free fit}}^2 \quad . \quad (7.10)$$

In this test, the mis-modeling value is used to quantify the impact of fixing the nuisance parameter n to its nominal value, as shown schematically in figure 7.11. When assuming Wilks' theorem, mis-modeling values can be interpreted directly as significances. Nuisance parameters for which this results in a mis-modeling value that corresponds to $\leq 0.1\sigma$ for any NSI parameter and any tested value thereof are considered necessarily free in the final fit settings. As a result, the nuisance parameters that are ultimately set free are a super-set of those found to be relevant for each NSI parameter.

²In a naïve estimate, the minimization complexity scales as a power law with the number of parameter space dimensions.

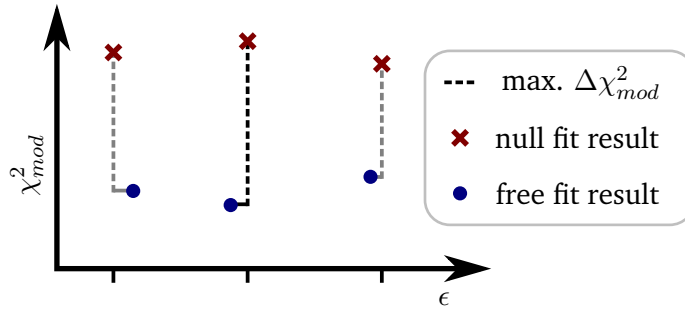
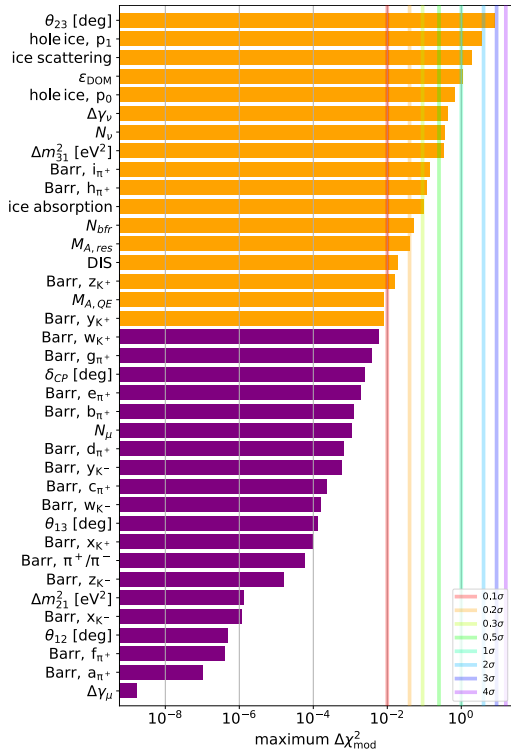


Fig. 7.11.: Schematic of how the impact of an individual nuisance parameter n is tested. Fits are performed at off-nominal values of n for multiple values of the tested NSI parameter ϵ , which is shown on the horizontal axis. In the null fit, ϵ is fixed to its true value, while floating freely in the free fit. The largest obtained mis-modeling, indicated as $\Delta\chi_{mod}^2$ between the two fit results, describes what impact fixing n to its nominal value has.

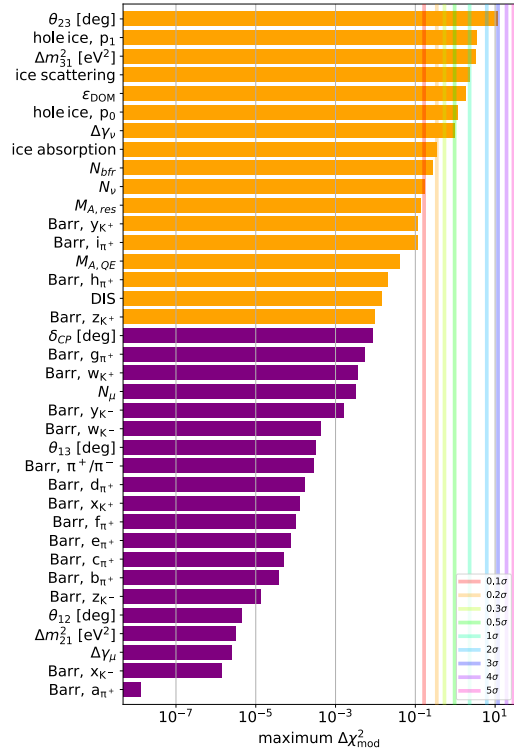
Figure 7.12 shows the mis-modeling values obtained in such tests for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ and $\epsilon_{e\tau}^{\oplus}$. The values for individual nuisance parameters differ little between NSI parameters and jointly confirm the set of nuisance parameters set free initially for the test.

Parameters with consistently largest impact are associated to uncertainties in the understanding of the detector modules and ice properties. A second large contribution lies within oscillation related parameters to which atmospheric neutrinos are sensitive, namely θ_{32} and Δm_{31} . Including these is crucial, as they potentially introduce effects similar to those of NSI. The overall normalization, which is a fundamental variable in the understanding of the signal, impacts the analysis outcome at similar significance. Only a subset of the parameters associated to neutrino flux and cross section result in significances of $\sim 0.1\sigma$, which motivates including them in the fit. As the contamination of atmospheric muons in the sample is reduced to 0.55% in the event selection, the respective nuisance parameters have in general small impact and are fixed.

From the super-set of all NSI parameters, a total of 17 nuisance parameters is determined that are set free-floating in all subsequent steps. These are shown in orange in figure 7.12 and stated as free in the nuisance parameter tables.



(a) Results for the real-valued $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ parameter.



(b) Results for the complex-valued $\epsilon_{e\tau}^{\oplus}$ parameter

Fig. 7.12.: Impact of fixing individual nuisance-parameters to their nominal value when fitting to an off-nominal hypothesis. The figures for $\epsilon_{e\mu}^{\oplus}$ and $\epsilon_{\mu\tau}^{\oplus}$ can be found in appendix section B.6.

7.6 Finding the hypothesis that fits the observation best

As a consequence of the aforementioned impact of individual nuisance parameters, the analysis parameter space is spanned by 17 nuisance parameters and the evaluated dimensions in NSI, which are either one or two if parameters are tested individually. The hypothesis that best fits the observed data needs to be located within this high-dimensional space. In addition, confidence limits on NSI parameters need to be calculated. The mechanics of this differ between the two approaches presented in this work, namely the frequentist approach of minimizing a test statistic function within the parameter space, which is discussed in this section, and secondly the Bayesian approach of sampling the posterior distribution, which is presented in section 7.7. The latter is not the focus of this analysis, but a secondary approach that was explored in this work. In the following, challenges in evaluating the parameter space will therefore be described with respect to minimization, although they mostly apply similarly to both approaches.

7.6.1 Navigating the parameter space

The main challenge in minimization arises from the high dimensional parameter space and the behavior of the test statistic function. There are known approximate test statistic mirror symmetries around $\theta_{23} = 45^\circ$ [27] and in the NSI parameters. The latter are present at $\epsilon_{\tau\tau}^\oplus - \epsilon_{\mu\mu}^\oplus = 0$, $\epsilon_{ee}^\oplus - \epsilon_{\mu\mu}^\oplus = -1$ and, for the complex phases of FC parameters, at $\delta_{\alpha\beta} = 180^\circ$. In addition to these symmetries, the test statistic does not progress smoothly along nuisance parameters that describe the detector. These characteristics increase the computational resources required for finding the test statistic minimum within the parameter space and a single iteration of the minimization process can run for up to tens of hours. Different tests that are run to ensure the reliability and accuracy of the analysis require $\mathcal{O}(10^4)$ repetitions in total. Fast minimization, which implies preferably few test statistic evaluations, is therefore necessary to stay within the resources available on the employed computing clusters [112, 138, 139].

The approach chosen for this analysis is to consider NSI parameters individually in a frequentist fitting method based on minimizing the test statistic described in section 7.2.1.

Treating NSI parameters individually reduces model independence by rejecting scenarios where multiple NSI parameters are non-zero at once. It however also reduces the dimensionality and symmetries within the parameter space, which

makes fitting computationally feasible. The reduced ranges of individual parameters discussed in section 7.3 furthermore reduces the size of the parameter space that a minimization needs to cover.

An additional factor is the treatment of errors on bin counts. These enter the applied test statistic and depend on the expected event count which varies depending on the evaluated hypothesis. This additional factor in the test statistic values was found to potentially complicate minimization.

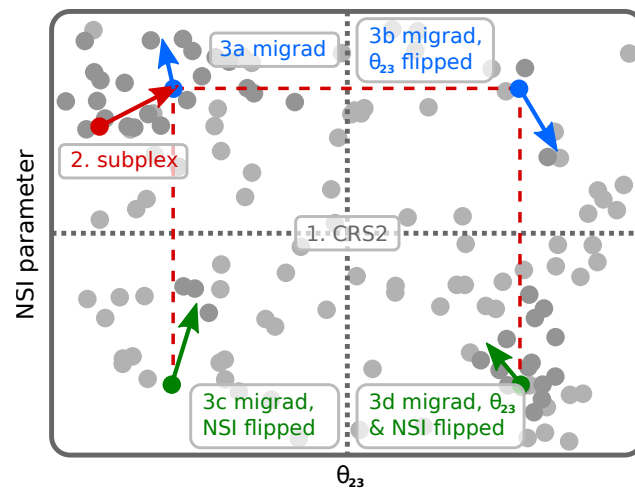


Fig. 7.13.: Schematic overview of the minimization process. The initial population of the first minimization step, a controlled random search (CRS2), is shown as light gray randomly distributed dots, while later iterations of the algorithm are shown in dark gray. Colored lines denote subsequent minimizations, namely migrad and subplex, with their start end points illustrated as dots and arrows, respectively. Symmetry axes in the two shown dimensions are indicated as black dashed lines, with inflections along them depicted as dashed lines in the color of the preceding minimization.

A multitude of different minimization approaches were tested for this work. The capacity of PISA to define recursive minimization strategies allows for combining multiple methods with different boundary conditions [123]. The minimization strategy ultimately used in this analysis is discussed in the following. It is motivated from the outcome of different Monte Carlo based tests that are discussed in subsequent sections.

The overall minimization process consists of two fundamental parts: First finding a good starting point and then optimizing this point further. This separation is motivated from the observation that no tested algorithm accomplishes both with sufficiently good accuracy and run-time. Algorithms that are able to find an approximate solution without any previously tuned starting point require prohibitively large computation resources to arrive at a hypothesis that fits the underlying truth within small tolerance. Algorithms that achieve the latter do so only if they start at a

hypothesis that is reasonably close to the global optimum. As a result, both kinds of algorithms are combined in the approach used in this work. A schematic overview of the minimizing procedure is shown in figure 7.13.

Finding a good starting point is addressed by using four instances of the CRS2 algorithm [140–142] that are confined to different quadrants of the parameter space. This algorithm operates with a large number of randomly initialized points that are evaluated in parallel in order to cover multi-modal parameter spaces. In this work, this so-called population size is set to 250.

The ranges are defined based on the approximate symmetries found at $\theta_{23} = 45^\circ$ as well as the NSI parameters. The four initial CRS2 instances only cover one half of the FC parameter complex phase range, dividing it into two ranges of $\delta_{\alpha\beta} \in [0^\circ, 90^\circ]$ and $[90^\circ, 180^\circ]$, in order to reduce the size of the parameter space.

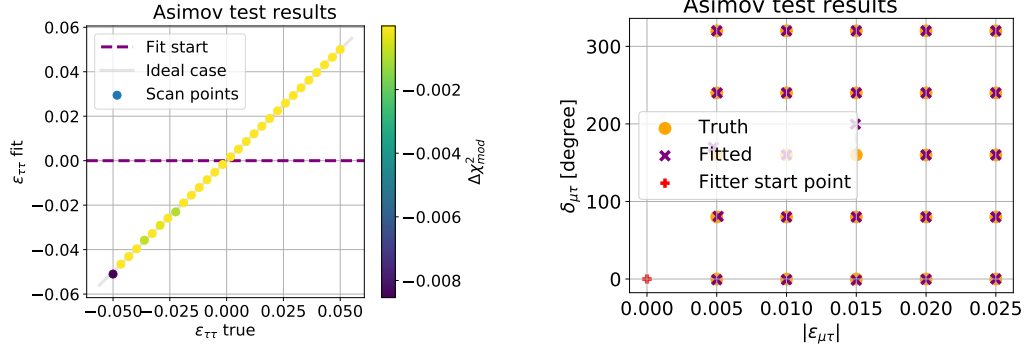
A CRS2 instance terminates once the relative or absolute tolerance of 0.002 is reached or, in most cases, when the maximum run-time of 7000 s (~ 2 h) is exceeded. The hypothesis with the best test statistic value is returned, with the overall best result of all individually optimized ranges being set as starting point for the subsequent minimization.

This is followed by an unrestricted iteration of the subplex [143] algorithm which is based on the Nelder-Mead simplex minimization [144]. It typically terminates after ~ 1000 evaluations, which corresponds to approximately 1 h of run-time, after achieving relative or absolute tolerance of 0.0001. In this intermediate step, the approximate symmetries within the parameter space are not treated separately as this is computationally costly and therefore done only when it is necessary. This necessity is not given in this step, which only refines the hypothesis up to a relatively coarse level.

Subsequently, four instance of the migrad [145, 146] algorithm are executed based on the θ_{23} and NSI parameter symmetries. These are not confined to specific ranges but start at the hypothesis returned by subplex and the corresponding hypotheses that are inflected at the respective symmetry axes. After typically 1000 – 3000 evaluations in each instance, the accuracy requirements are reached. These are a tolerance of 10^{-5} , precision of 10^{-14} and an error³ of 10^{-5} . The best of the four results is returned as the overall minimization result.

This approach yields sufficiently accurate results in most cases. There are, however, settings where the CRS2 instances are not reliably able to find a good starting point for the subsequent steps. This is discussed in the following, based on tests that were conducted to evaluate the minimization performance.

³A definition of these quantities can be found in [145].



- (a) Results for the real-valued $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ parameter. All scan points show good recovery, which is indicated with the gray diagonal line where the true value equals the fitted one. The color corresponds to the mis-modeling.
- (b) Results for the complex-valued $\epsilon_{\mu\tau}^{\oplus}$ parameter. The injected truth-values are shown in orange and are connected with gray dashed lines to the respective fitted value which is marked with a purple cross.

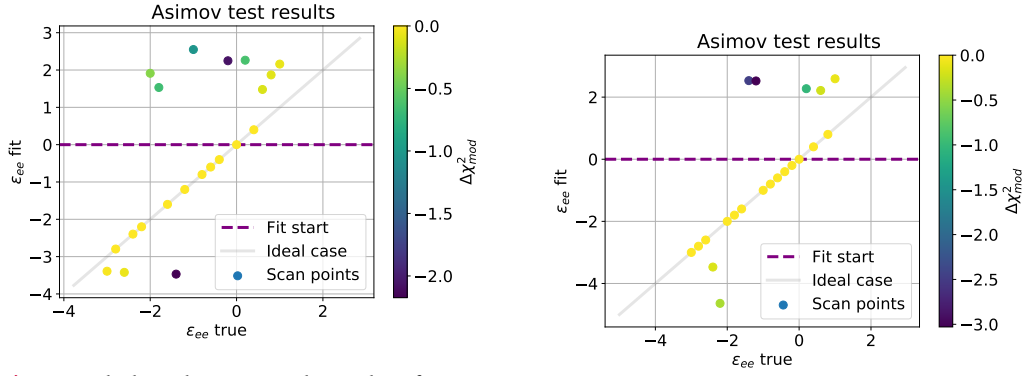
Fig. 7.14.: Test of the fit performance where pseudo-data are generated at different NSI parameter values and attempted to recover in an Asimov approach. Due to the random character of CRS2 as the initial fitting step, the start point of the fit is not expected to have any impact on the fit outcome. It is indicated as a purple dashed line for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ and as a red dot for $\epsilon_{\mu\tau}^{\oplus}$.

7.6.2 Recovering injected NSI hypotheses

As a basic test of the minimizer performance, pseudo-data are generated at different injected NSI hypotheses, which are attempted to recover through a fit. This is performed with an Asimov approach, where neither nuisance parameter values nor bin counts are fluctuated. The global optimum of the parameter space is therefore exactly at the injected hypothesis.

The drawn hypotheses cover the entire range that the NSI parameters are allowed to vary in while fitting, which significantly exceeds the confidence intervals given in previous studies [40]. Differences between injected and recovered parameter values are compared both in absolute scale and through their mis-modeling. Being the test statistic difference between the fit result where the tested NSI parameter is free and one where it is fixed to the injected value, the mis-modeling is required to be below 0.2σ . It measures what statistical impact the inaccuracy in recovery has, which is of more relevance to the interpretation of analysis results than the absolute error values.

Figure 7.14 shows the injected and fitted values exemplarily for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ and $\epsilon_{\mu\tau}^{\oplus}$. Respective figures for other NSI parameters as well as the corresponding mis-modeling values for $\epsilon_{\mu\tau}^{\oplus}$ can be found in section B.7 of the appendix. For all NSI parameters except for $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$, the evaluation criteria are met, as all truth values are recovered well and the mis-modeling is below 0.2σ for all tested hypotheses.



(a) Fits including the commonly used 17 free nuisance parameters. Most scan points diverge significantly from the optimal case that is indicated as the gray diagonal line, where the true value equals the fitted one.

(b) Results when including only the 8 nuisance parameters that yield the most significant impact on the analysis.

Fig. 7.15.: Test results for the real-valued $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ parameter in two settings. In both, the maximum run-time of the CRS2 instances is increased by a factor of 2 with respect to the otherwise used setting. The points with large errors result in large negative mis-modeling values, indicating failure of the minimization.

Small negative mis-modeling values, such as those found at some tested values of $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$, are not considered a problem.

For $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$, an improvement in performance in this test is achieved when doubling the run-time of the CRS2 instances and decreasing the number of free nuisance parameters to 8, as shown in figure 7.15. The reduced set of nuisance parameters includes those that result in $\geq 0.5\sigma$ mis-modeling in a test set-up as described in section 7.5. These are θ_{23} , Δm_{31} , N_{ν} , $p0$, $p1$, $\Delta\gamma_{\nu}$, DOM efficiency, ice scattering, and ice absorption.

The parameter value recovery for $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ is in this way shown to slightly improve in a case where the dimensionality is reduced and the minimization operates with significantly more computational resources. This indicates that $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ can likely be investigated with the minimization approach applied in this analysis, if vastly more computation time is available. Since this is currently not given, $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ will be disregarded in the following.

Reviewing the ice model

As the ice model nuisance parameter $N_{b,fr}$ is newly introduced in this generation of analyses, its capacity to model the detector under different assumptions is evaluated. To this end, the above described test is run with BFRv2 based pseudo-data and a SPice 3.2.1 based template. The implication of such a test is that SPice 3.2.1 can be kept as the model that this analysis is based on if that does not impede recovering

any NSI hypotheses, even in a scenario where the detector ice is perfectly described by BFRv2. For such a case, where the pseudo-data is generated at $N_{bfr} = 1$, fitted assuming $N_{bfr} = 0$ and N_{bfr} is kept free in the fit, the truth value recovery and the obtained mis-modeling values meet the given criteria of being positive or small as well as corresponding to $\leq 0.02\sigma$. If in such a setting, N_{bfr} is fixed for fitting, mis-modeling values up to $\mathcal{O}(1)$ are found, which furthermore motivates including this as a free-floating nuisance parameter.

7.6.3 Recovering fluctuated nuisance parameters

While the nominal values of nuisance parameters correspond to the current best knowledge of the underlying effect, variations in them are expected within their respective range. The capability of the minimization to recover off-nominal hypotheses is tested by fitting to pseudo-data that are generated with fluctuated nuisance parameter values and non-fluctuated bin counts. In figure 7.16, the fluctuated and fitted values for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ are shown for 100 trials. Respective test results for all further investigated NSI parameters can be found in the appendix, in section B.8. For all tested NSI parameter magnitudes, good recovery of the assumed truth values is observed.

In this test, all free parameters are fluctuated simultaneously based on their allowed ranges or Gaussian priors. In these cases, the global optimum is not at the pseudo-data truth hypothesis in all dimensions, as the Gaussian priors that are applied to a subset of the nuisance parameters result in a shift of the optimum towards the nominal value. This can become dominant for nuisance parameters to which there is low sensitivity. In addition, potential correlations between parameters can result in fitted values differing from to the underlying truth.

The evaluation criterion in this test is therefore less quantitative. Requirements are a clear and consistent correlation between injected and recovered values and few mis-modeling values above 0.1.

The negative mis-modeling values shown as red crosses in figure 7.16 indicate that some of the fits fail to find the global optimum: Negative values imply that the fit that includes the tested NSI parameter yields a worse test statistic value than the fit where this parameter is fixed. As in general, a larger number of free dimensions allows for a better fit to the pseudo-data, the fit with more degrees of freedom must have failed to find the global optimum. Such behavior was found to arise in cases where the CRS2 instances return a starting point for the subsequent minimization that is in a secondary, local optimum.

Tab. 7.8.: Fractions of negative mis-modeling occurrences and number of trials in different studies for all NSI parameters. A discussion of fluctuating binned event counts is given in section 7.6.4.

Ensemble	Fraction / Ensemble size			
	$\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$	$\epsilon_{e\mu}^{\oplus}$	$\epsilon_{e\tau}^{\oplus}$	$\epsilon_{\mu\tau}^{\oplus}$
Fluctuated parameter values	2 % / 100	1 % / 100	12 % / 100	32 % / 100
Fluctuated bin counts	2.6 % / 500	0.6 % / 500	0.2 % / 497	0.6 % / 500

Negative mis-modeling

The fractions of negative mis-modeling differ between the individual NSI parameters, as shown in table 7.8. Since the ensemble of pseudo-data trials drawn for this test does not represent the expected behavior of DeepCore data⁴, the fraction of fits yielding negative mis-modeling solely serves as an indication for parameter space regions where the minimization does not perform well. In the test results shown here, such cases are randomly distributed. Therefore, this behavior can likely not be addressed by further changing the minimization strategy. This issue is also present in other studies and makes partial repetitions necessary. For example, in order to meaningfully evaluate the impact of individual nuisance parameters in section 7.5, single tested hypotheses needed to be repeated. The rate at which cases of large negative mis-modeling are expected to occur in data is discussed in the following section.

⁴An ensemble representing this is discussed in section 7.6.4.

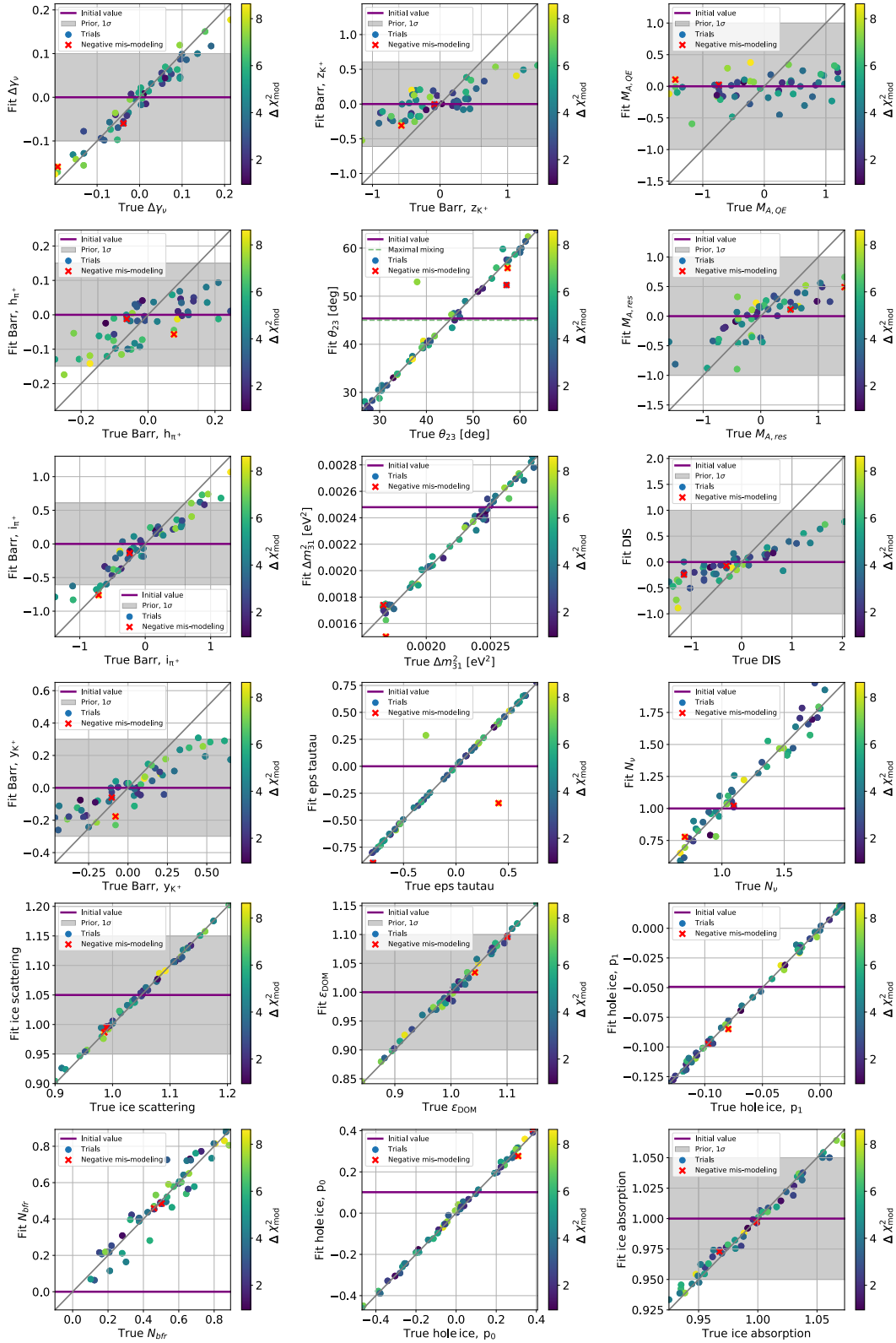


Fig. 7.16.: Fluctuated and fitted parameter values including $\epsilon_{\tau\tau}^\oplus - \epsilon_{\mu\mu}^\oplus$. The absolute mis-modeling obtained in each trial is shown in color. Red crosses mark cases of negative mis-modeling. The nominal value of the fluctuated parameter is shown in a purple line. For parameters with Gaussian priors, which are indicated as gray ranges, the global best fit point for off-nominal truth hypotheses is shifted towards the nominal value.

7.6.4 Handling pseudo-data

Pseudo-measurements can be obtained by Poisson-fluctuating each individual bin count based on its expectation value. This is done in the following under the assumption of the baseline model, in which all nuisance parameters are at their nominal values and all NSI parameters are zero. This non-Asimov approach tests the minimization performance on trials that simulate the fluctuation found in actual data.

For $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$, the test statistic distribution resulting from fits to the Poisson-trials is shown in figure 7.17a. Mean values and widths are similar for the NSI parameters not shown here, the test statistic distributions of which can be found in the appendix in section B.9. All mean values correspond approximately to the effective number of degrees of freedom, which is composed of the number of bins minus the number of free parameters in the fit.

In addition, if simulation and nuisance parameters model the reality correctly, the distribution of possible test statistic values obtained in fits to data would coincide with that of the pseudo-data ensemble. The test statistic values obtained for the ensemble of Poisson-trials then form a distribution within which the test statistic value obtained in a fit on data is expected to be situated. A low p-value [96] of the data fit test statistic is called a poor goodness of fit and implies potential incorrectness of the model. The convention of the IceCube oscillations working group sets a threshold of $p > 0.05$ for an analysis to proceed towards publication. The p-values resulting from fits to the DeepCore data sample⁵ that are performed for all NSI parameters are at $\sim 5\%$ ⁶. The threshold value of $p = 5\%$ is disputable, given that this is a comparison of an ensemble that is drawn at the null hypothesis and a fit result that includes off-nominal parameter values. It however hints towards possible incorrect modeling within the simulation.

This interpretation is supported by the relatively low p-value of 20% [147] found in an unpublished muon neutrino disappearance analysis that uses the exact same data sample as this work while including a different set of nuisance parameters. An equivalent analysis of a DeepCore data sample that differs from what is used in this work mostly in that it applies the RETRO reconstruction, yields a p-value of 0% at above 3σ [147].

This tendency towards low p-values motivates further detailed investigation of the data sample and physics modeling. The conventions of the IceCube oscillation working group with respect to blindness of the data dictate that this analysis may

⁵In order to assure that the minimization succeeded, these fits were performed three times for each NSI parameter.

⁶The exact values are $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$: 4.8%, $\epsilon_{e\mu}^{\oplus}$: 4.4%, $\epsilon_{e\tau}^{\oplus}$: 4.6%, $\epsilon_{\mu\tau}^{\oplus}$: 5.4%

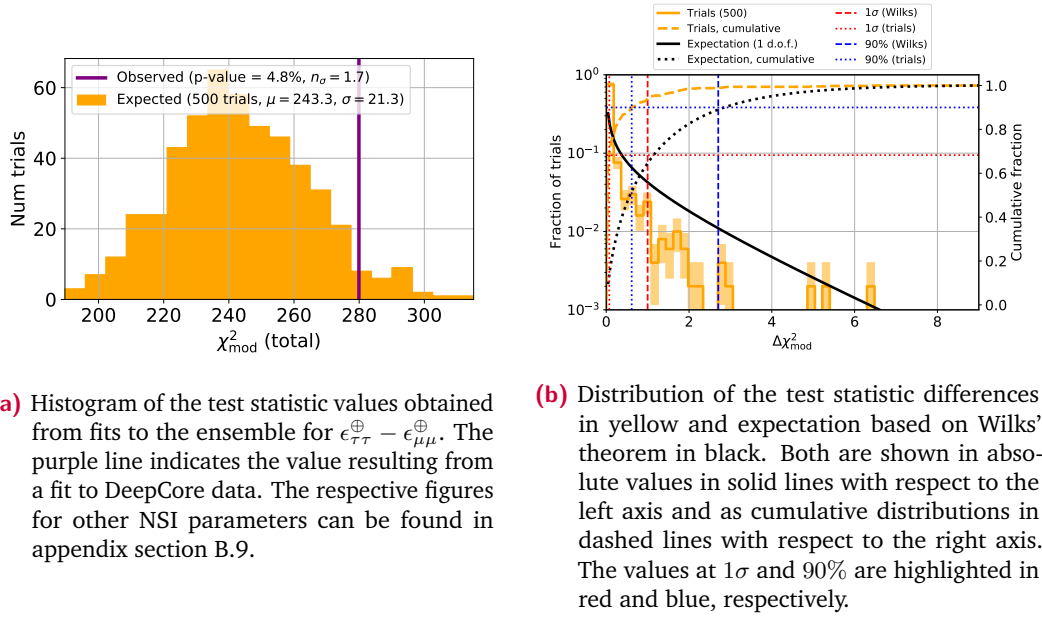


Fig. 7.17.: Test statistic distributions for an ensemble of pseudo-data trials where bin counts are Poisson-fluctuated.

not proceed on experimental data until these investigations are concluded. This is beyond the scope of the thesis presented here.

As a result, the aforementioned fits to data for all tested NSI parameters will not be evaluated in more detail and all further tests will be performed purely based on Monte Carlo simulation.

As shown in table 7.8, a small fraction of trials results in negative mis-modeling values. Since this ensemble is assumed to represent the distribution of possible data realizations, a non-zero fraction of negative mis-modeling cases implies a respective probability of failure when fitting DeepCore data. This motivates performing multiple redundant fits to the data.

In addition to the absolute test statistic values, null fits and corresponding mis-modeling values are available for this ensemble. From figure 7.17b, it is apparent that the mis-modeling values are not χ^2 distributed, which would be expected if Wilks' theorem applies to the overall analysis.

The non-applicability of Wilks' theorem could be reproduced in a toy Monte Carlo simulation, where different signal behaviors were tested. The existence of mirror symmetries in NSI parameters has the consequence that, broadly speaking, any non-zero NSI magnitudes result in bin count shifts to the same direction. For fluctuations in the overall bin count that are opposed to this shift, this results in a pile-up of fit results at the respective symmetry axis, leading to a shift of test statistic differences towards small values.

For further steps of this analysis, this implies that limits that are set assuming Wilks' theorem are conservative, and other methods are required to set representative confidence limits. This will be discussed in more detail in section 7.9.

7.7 Sampling

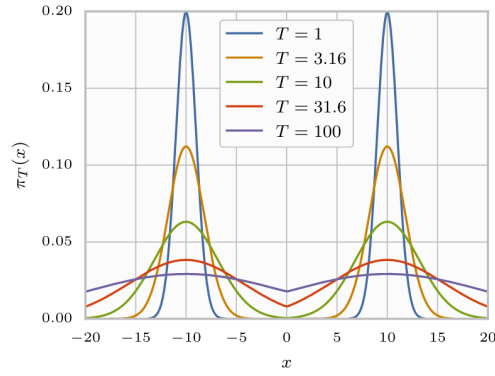
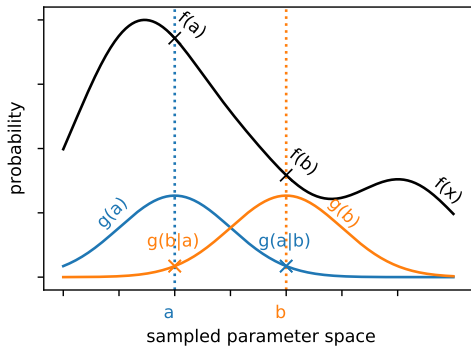
In the frequentist approach that is pursued in this analysis, setting confidence limits on NSI parameters is a separate task in which a large number of fits need to be performed, as described in section 7.9. This motivates exploring the fundamentally different method of sampling the parameter space, which provides both its optimum and confidence regions in a single instance.

In the following, the Bayesian approach of Markov Chain Monte Carlo (MCMC) sampling is evaluated, a technique that is known to perform well with large numbers of dimensions [148]. In addition to the option of this to replace or complement fitting, its potential with respect to evaluating multiple NSI parameters simultaneously is tested, which would decrease model dependence of this analysis.

7.7.1 The mechanics of MCMC

Markov-chains are sequences drawn from a set of possible elements within which the probability of each element solely depends on what element preceded it. In MCMC, this implies that from each point in an initial set of hypotheses, a sequence of steps emerges [148]. All of these so-called walkers proceed within the parameter space in simultaneous steps that are based on the sampled function $f(x)$. In this analysis, $f(x)$ is the test statistic function. Depending on the sampling mode either each walker is a fully self-contained Markov-chain or dependencies between walkers are introduced.

Sampling is performed such that ultimately, the set of evaluated hypotheses is distributed according to the posterior probability density of $f(x)$. Regions of the parameter space that are closer to a minimum are therefore more densely populated and confidence limits can be inferred directly from the final distribution of samples. The aim of MCMC, sampling from $f(x)$ at a frequency that is proportional to the function values throughout the parameter space, is achieved through the equilibrium function $g(x) \propto f(x)$. Proportionality is reached after sufficient sampling steps and persists for further steps as a result of the metric based on which sampling steps are accepted.



(a) Schematic of the Metropolis-Hastings algorithm. (b) Example of sampling a multi-modal function using ptemcee, taken from [149].

Fig. 7.18.: Schematic illustrations of two sampling implementations.

In a simple case, this is done using the Metropolis-Hastings algorithm that is schematically shown in figure 7.18a. The equilibrium function $g(x)$ is a sum of Gaussian distributions around each active hypothesis, which is the current final point of a single walker. Correspondingly, the equilibrium function depends on the distribution of sampled hypotheses and equals the posterior probability density of $f(x)$ when hypotheses are distributed accordingly.

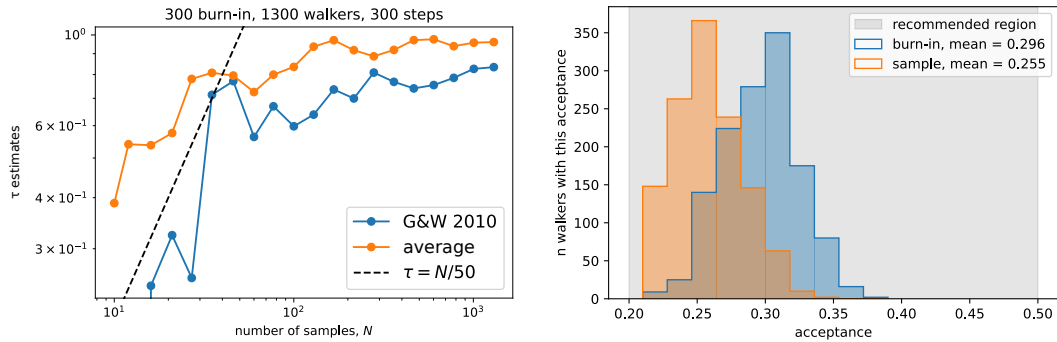
To obtain this distribution, a suggested step of a walker from current position a to position b depends solely on a , as b is sampled from the Gaussian distribution centered around a . The step is accepted based on the values of the sampled function and equilibrium function at the previous and suggested position following

$$\text{accept} = \min \left(1, \frac{f(b) g(a|b)}{f(a) g(b|a)} \right) . \quad (7.11)$$

In this way, starting from a random initial distribution, the sampling density successively adapts to $g(x) \propto f(x)$. Iterations performed until this proportionality is established are called the burn-in phase of sampling. These are not included in subsequent interpretation of the ensemble of samples, as their prevalence does not describe $f(x)$.

Of the so-called moves, which are the principles based on which steps are taken, different variants are available and can be adapted freely. The most commonly used one, which is also implemented in the following, is the so-called stretch move. This accelerates convergence by combining information from multiple walkers. The current position of a randomly selected second walker is included in drawing the suggested position b .

The adaptive parallel tempering MCMC implementation (ptemcee) implementa-



- (a) Two measures for the autocorrelation time τ dependence of the sampling steps, as described in [150, 151]. After following the dashed diagonal line, τ stabilizes at a constant value, which is the expected behavior.
- (b) The distributions of acceptance fractions per walker look similar for the burn-in and sampling phase. Both lie within the recommended region between 20% and 50%, shown in gray.

Fig. 7.19.: Two measures for the efficiency and accuracy of sampling. This uses slightly different settings from what is described in section 7.7.4, namely 1300 walkers and 300 burn-in and sampling steps.

tion [149] alters $f(x)$ in order to handle multi-modal distributions. In this approach, the temperature $T \geq 1$ is added as an additional dimension such that

$$f(x, T) = f(x)^{1/T} \quad . \quad (7.12)$$

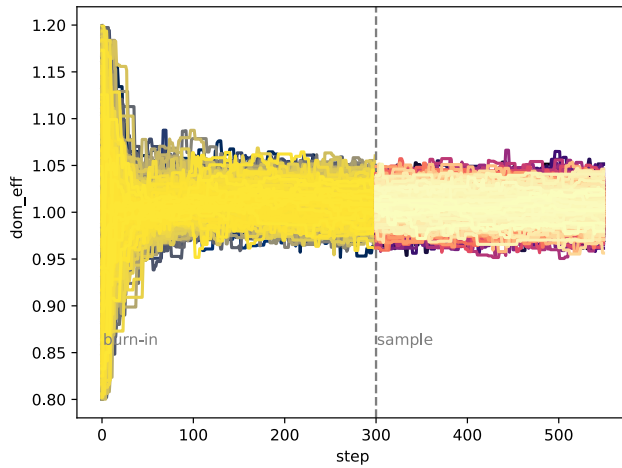
The ability to sample $f(x, T)$ implies sufficient knowledge of $f(x)$. The added temperature however has the advantage that, for $T > 1$, modes in $f(x)$ are less pronounced and therefore easier to sample, as shown in figure 7.18b.

In addition to finding the best implementation and move for the specific requirements of this analysis, multiple settings of the sampling can be adapted. The individual settings and corresponding metrics are discussed in the following.

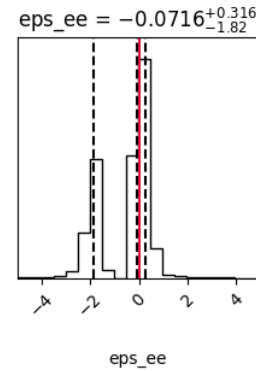
7.7.2 Finding the best settings

Multiple criteria can be used to determine whether MCMC is functioning correctly and uses sufficiently well optimized settings. In addition to the above-mentioned evaluation characteristics, namely the implemented move and whether $f(x)$ is altered to facilitate sampling, three quantities need to be adapted for the specific use case. These are the number of walkers and the number of steps per walker, the latter being set separately for the burn-in and the sampling phase.

Convergence of a walker towards the equilibrium function is related to its integrated autocorrelation time. This quantifies the number of steps after which effective



(a) Behavior of all individual walkers (colored lines) over the burn-in and sampling step. The sampling run includes the same parameters as shown in figure 7.21. Here, the DOM efficiency is shown. The randomly initialized walkers converge towards an equilibrium distribution within the first ~ 100 steps in this dimension.



(b) Sampling run with $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ as the only free parameter. The two minima have different heights as they are sampled with different numbers of walkers. The obtained optimum and 1σ range are given at the top of the plot.

Fig. 7.20.: Two visualizations for outcomes of different sampling runs.

independence of a subsequent step from the start point of the walker is reached. Two different measures for autocorrelation time are shown in figure 7.19a. The first directly implements the approach in [150]. The second is calculated per walker and then averaged [151]. Qualitatively, this quantity is expected to increase during the burn-in phase and stay constant for the subsequent sampling. A step independent autocorrelation time implies that the algorithm has converged and reached the equilibrium distribution. A small absolute autocorrelation time suggests quicker convergence and hence more efficient sampling [148].

An associated quantity is a comparison between halves of the sampled ensemble. If convergence is reached during the burn-in phase and kept throughout sampling, the distributions of sampled hypotheses in the first and second half of the overall sampled ensemble coincide within their statistical uncertainty.

A measure for the sampling efficiency that was also considered throughout this study is the fraction of accepted steps. This should lie between 20% and 50% [151] to ensure both variability of the walkers and efficiency of the overall sampling. An example for this in a mostly finalized setting is shown in figure 7.19b.

An example for a well performing sampling run is shown in figure 7.20a. The distribution of walker positions over steps shows first a convergence towards the optima of the sampled space. Once it stabilizes, the burn-in phase can be terminated and the subsequent sampling steps can be assumed to represent the posterior distribution. This convergence criterion as a result fixes the number of burn-in steps.

Destabilization in later steps would hint towards too few walkers. Their number, as well as the number of sampling steps, needs to be appropriate to resolve features of the parameter space.

A rule of thumb is to use as many walkers as computationally feasible, and necessarily more than the number of sampled dimensions. The number of sampling steps usually is similar to or larger than that of burn-in steps and is set such that the resulting ensemble of sampled hypotheses resolves the detail of $f(x)$ sufficiently well to yield a smooth distribution.

The hypotheses where walkers are initialized are by default distributed randomly throughout the parameter space. As this initial distribution is an optional input parameter to the sampling algorithm, it can however also be optimized with respect to sampling performance. A higher density at parameter space regions that are known to be of interest can improve autocorrelation time. In this analysis, drawing initial hypotheses according to the full or half widths of nuisance parameter priors or their allowed regions has been shown not to improve the overall sampling runtime or exactness significantly.

7.7.3 Challenging parameter space features

Similarly to the challenges observed with respect to minimization, intermediate minima within the test statistic can impede sampling. In addition, for a large number of simultaneously evaluated dimensions, sensitivity to individual parameters becomes small compared to biases induced by the distribution of initial walker positions.

Features of the parameter space that are relevant with respect to sampling are both multi-modalities and instabilities. Secondary minima of the parameter space can „trap“ individual walkers. This is countered by re-initializing walkers after the burn-in phase at the two thirds of previous positions with the best test statistic⁷ values, as shown in figure 7.22b.

When sampling all NSI parameters simultaneously and including a preliminary set of 18 nuisance parameters, the sensitivity to the complex phases of FC NSI parameters becomes negligible. As additional dimensions make more sampling points necessary, only the five NSI magnitudes are ultimately included in this approach.

Bimodalities, such as the symmetry in $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ around -1 , are hard to sample in the standard MCMC implementation as individual walkers are unlikely to migrate between minima of similar depth. The resulting fractions of walkers in each minimum strongly depend on their initial hypotheses and therefore does not properly

⁷This study uses LLH, as it is based on a preliminary version of the event sample and nuisance parameter treatment.

represent the sampled function. This is shown in figure 7.20b for a case where $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ is the only sampled parameter and the relative depths of the minima are not represented accurately.

Applying ptemcee reduces such features at a degree that can be adjusted via the temperature setting, as described above. Due to the low sensitivity to $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$, however, too lightly pronounced features were found to result in insensitivity to the parameter, while too little reduction of the features imply no improvement over the original MCMC approach. The method applied further on is to subdivide the parameter space along the $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ symmetry. A description of the necessary partitioning and merging can be found in [152].

Even for a subdivided parameter space, sampling in $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ is affected by a bias that originates from the mechanics of the stretch move, which introduces a dependency between multiple walkers. As each step of a walker is not only based on its previous steps but also on the position of a random second walker, a tendency towards the overall mean emerges. This effect can exceed the test statistic variation in a dimension of the sampled space if sensitivity in that dimension is low. The equilibrium distribution as a result does not reproduce the parameter space and instead corresponds strongly to the input distribution.

An indication to such behavior is that the sampled hypotheses don't reach small test statistic values in an Asimov case. This can be counteracted by increasing the number of burn-in steps and restricting the initial distribution to parameter ranges to which there is sensitivity. In case of $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$, the issue could be partly resolved by implementing a custom move that differs from the stretch move in that it includes no dependence between walkers in the $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ dimension.

Depending on the final set of free parameters and resulting sensitivity to $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$, in addition, separate burn-in of $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ can be necessary. In this approach, the final walker positions of a burn-in phase with $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ as the only free parameter are used to initialize the subsequent burn-in of all dimensions.

Due to the even smaller sensitivity to complex NSI parameter phases, these are not included as sampled dimensions. The magnitudes are however sampled symmetrically around 0, which corresponds to phases of 0° and 180° . This is done to avoid an additional difficulty for the sampling algorithm, which occurs if relevant parameter regions are close to the edge of their sampled range.

7.7.4 Performance

The finalized set-up implements MCMC with a custom move that largely coincides with the stretch move. After an initial burn-in phase with 300 steps, the 1500

walkers are re-initialized at the two thirds of current positions with the best test statistic values. Only the subsequent 250 sampling steps enter the determination of the parameter space optima and confidence intervals.

All five NSI magnitudes are sampled alongside 18 nuisance parameters, resulting in a total of 23 dimensions. Sampling of $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ is subdivided into two ranges. As this study was performed based on an earlier version of the data set, nuisance parameters that are considered herein differ from what is found to be relevant in section 7.5.

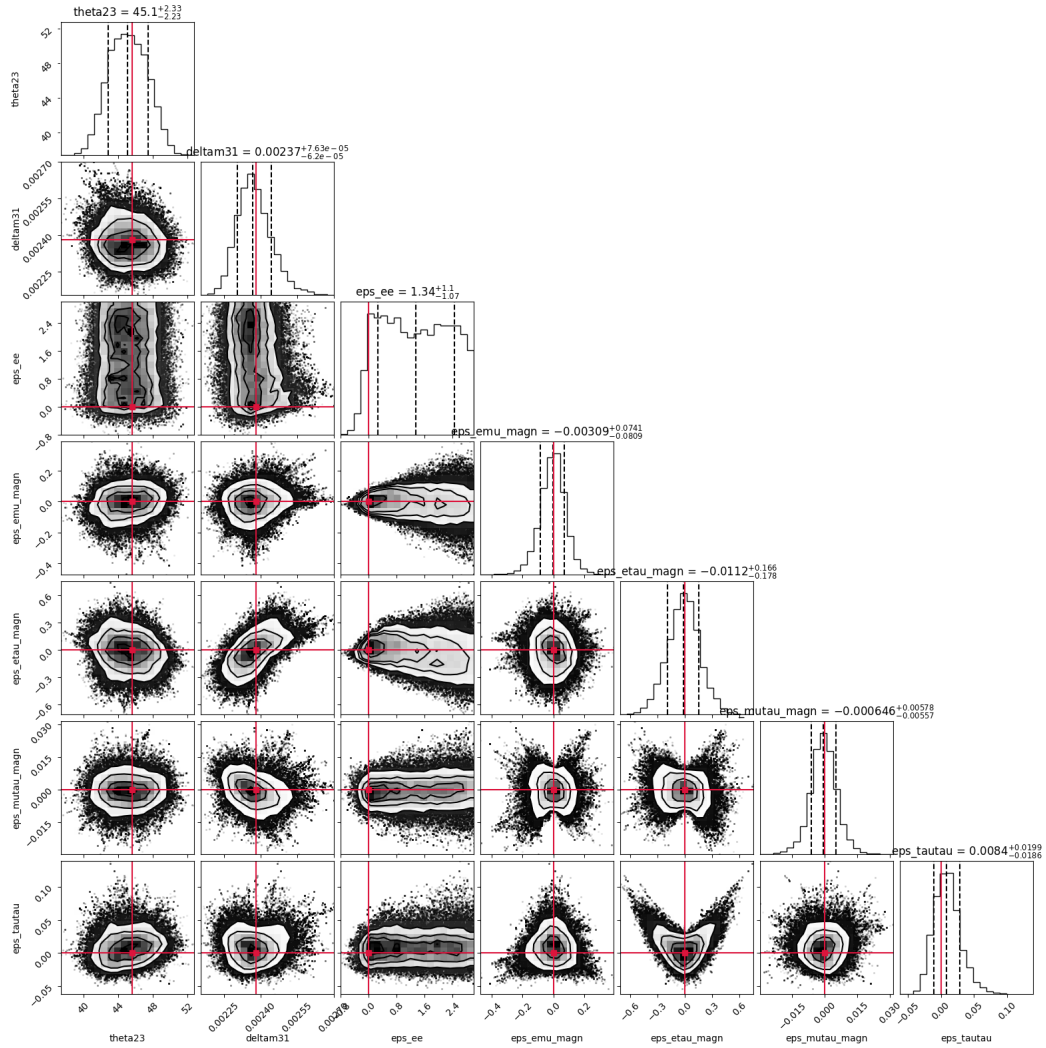
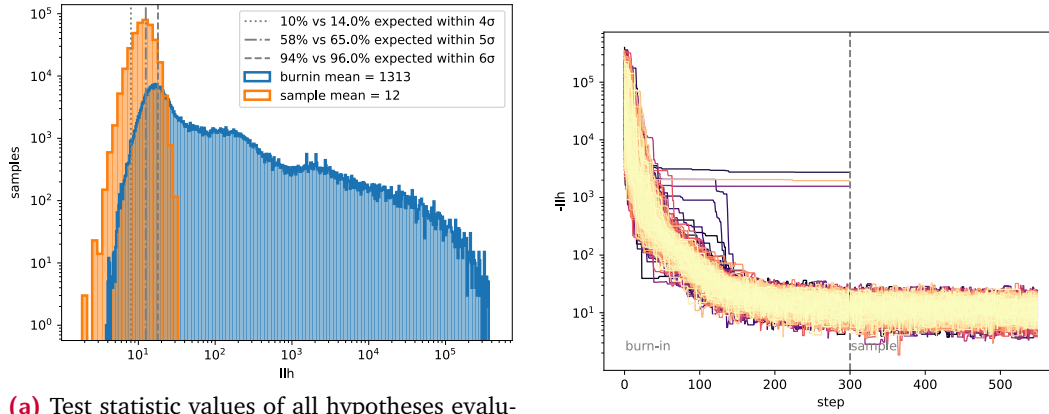


Fig. 7.21.: Sampling result for the MCMC settings described in the text including the $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ range above -1 . Above each one-dimensional histogram, the mean value and standard deviation are given. Only NSI parameters and oscillations-related nuisance parameters are shown. The sampling result for the 16 additional included nuisance parameters as well as the results from the corresponding burn-in phase are shown in appendix section B.11.



- (a) Test statistic values of all hypotheses evaluated during the burn-in (blue) and sampling (orange) phase. Lower values occur towards the later steps and are therefore more prevalent in the sampling phase. Assuming Wilks' theorem, good representation of the parameter space can be verified by comparing different σ ranges to the percentages within different LLH ranges.
- (b) Test statistic value over the step number throughout the entire burn-in and sampling process. The vertical line indicates the transition from burn-in to sampling phase, where walkers are re-initialized. The latter is showcased on the rejection of single walkers at large test statistic values.

Fig. 7.22.: Characteristics of the test statistic values throughout the sampling process that is shown in figure 7.21. The same settings as described in the text apply.

In figure 7.21, a full sampling run with the described settings is shown. The resulting confidence limits indicated as lines are relatively smooth, which hints towards sufficient sampling statistics. Their values coincide well with the sensitivities found in the frequentist approach for the four NSI magnitudes that are investigated there, which are given in section 7.8. The low sensitivity to $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ is visible from the limits as well as the reduced prominence of the peak with respect to the remaining parameter range. Correlations between parameters are expected, as discussed in section 7.5. They are visible as tilts or distortions in the sampled hypotheses. The associated test statistic values throughout the burn-in and sampling phase are shown in figure 7.22a. The fractions of samples within different σ ranges coincide well with the expectation assuming Wilks' theorem. The corresponding step dependent development of the test statistic values of the ensemble of walkers is shown in figure 7.22b. Convergence during the burn-in phase as well as re-initialization of walkers thereafter is visible therein.

The computational cost of a single sampling run of the full parameter space with the above described settings amounts to ~ 800 CPU hours, as compared to ~ 10 hours for a single frequentist fit. This is subdivided into a large number of parallel threads through openMPI [151, 153]. Parallelization however is limited, as applying the stretch move requires communication between walkers. Other moves are not computable in a parallelized setting for similar reasons.

Due to the computational cost of a single sampling run, thorough testing of the analysis, such as it is described for the frequentist approach, can not be conducted for MCMC. As a result, this can only serve as an additional Bayesian component to a well understood frequentist analysis.

7.8 Expected sensitivities for the frequentist approach

For the frequentist approach, the sensitivity at the null hypothesis allows to assess the potential of the analysis and data sample without fitting directly to the data. Different aspects of the analysis that are described above impact what sensitivity can be expected for the DeepCore data sample. These are mainly the resolution of the minimization and the uncertainties modeled as nuisance parameters. An additional factor are assumptions on neutrino characteristics, such as NMO. As presently, there is no clear preference for a mass ordering, this analysis will ultimately be performed separately for both.

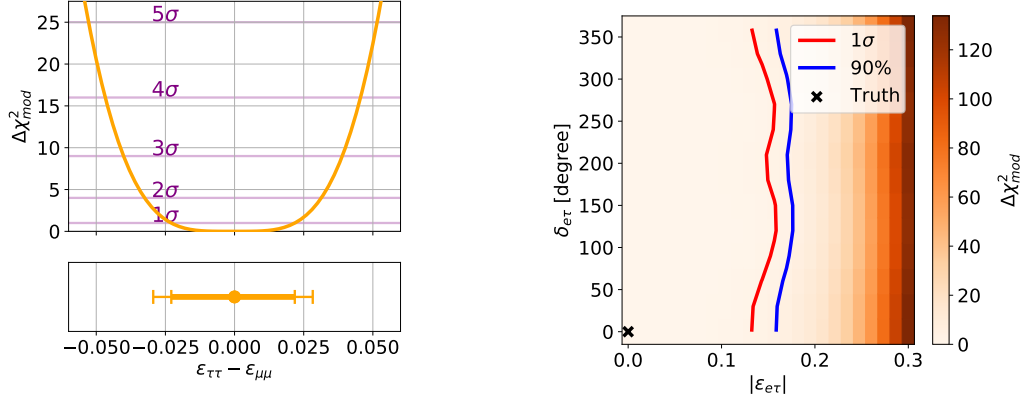
The sensitivity to individual NSI parameters can be assessed in an Asimov approach, including no statistical fluctuations. A template in which the respective parameter is fixed is fitted to a map of the expected bin counts assuming the null hypothesis. Performing this at different values of the tested parameter yields an understanding of what parameter range could be excluded at what significance in the null hypothesis case. This interpretation assumes Wilks' theorem and is therefore conservative, as will be discussed in section 7.9.

The obtained sensitivity mainly applies to the magnitudes of the tested NSI parameters. While for FC parameters, dependencies on the complex phases are present, it is not expected that the phases can be constrained in this analysis.

In figure 7.23, the sensitivities obtained for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ and $\epsilon_{e\tau}^{\oplus}$ are shown. For the former, the obtained confidence intervals are given at the bottom. The sensitivities for NO differ negligibly from those obtained assuming IO. The IO sensitivity scans can be found in appendix section B.10, preceded by NO sensitivities for NSI parameters

Tab. 7.9.: Sensitivity based 90% and 1σ confidence intervals of all tested NSI parameters. For complex parameter magnitudes $|\epsilon_{\alpha\beta}|$, bounds are given at $\delta_{\alpha\beta} = 0^\circ, 180^\circ$.

	NO 1σ	NO 90%	IO 90%
$\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$	[-0.023, 0.022]	[-0.029, 0.028]	[-0.029, 0.028]
$\epsilon_{e\mu}^{\oplus}$	[-0.087, 0.075]	[-0.106, 0.093]	[-0.107, 0.093]
$\epsilon_{e\tau}^{\oplus}$	[-0.148, 0.132]	[-0.176, 0.158]	[-0.176, 0.158]
$\epsilon_{\mu\tau}^{\oplus}$	[-0.0066, 0.0083]	[-0.0088, 0.0110]	[-0.0088, 0.0110]



- (a) Sensitivity for the real-valued $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ parameter. Confidence limits for 1σ and 90% are shown at the bottom as thick and thin orange lines, respectively.
- (b) Sensitivity for the complex-valued $\epsilon_{e\tau}^{\oplus}$ parameter. The 1σ and 90% confidence limits are shown as red and blue lines, respectively.

Fig. 7.23.: Sensitivities for tested NSI parameters assuming the Asimov null-hypothesis case. The NSI parameters that are not shown here can be found in appendix section B.10.

that are not shown in this section. This does not include $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$, which was not evaluated further for this work due to limited computational resources, as discussed in section 7.6.2.

Figure 7.24 shows a comparison between the resulting 90% confidence limits, which are listed in table 7.9, and results obtained in different experiments. For FC parameters, the two-dimensional confidence intervals are evaluated at phase values of $\delta_{\alpha\beta} = 0^\circ, 180^\circ$.

The comparison includes recent DeepCore and IceCube analyses, which are based on the energy ranges 5.6 – 100 GeV [40] and 500 GeV – 10 TeV [52], respectively. In addition, COHERENT [154] and ANTARES [53] results are shown for such NSI parameters where the respective confidence limits are competitive. The comparison of COHERENT results suffers from an implicit model dependence thereof: Due to coherent elastic neutrino-nucleus scattering (CE ν NS) being the underlying process, only NSI models with hypothetical mediator masses at or above the MeV scale are covered.

The additionally compared global analysis [39] includes data from solar and atmospheric neutrinos as well as long-baseline and medium-baseline reactor data. The concepts of all involved experiments are discussed in section 4.

In all tested NSI parameters, the presented sensitivity is comparable with or constitutes an improvement upon the limits set in other individual or global analyses. Final results obtained with this analysis can therefore be expected to be at least competitive in individual parameters.

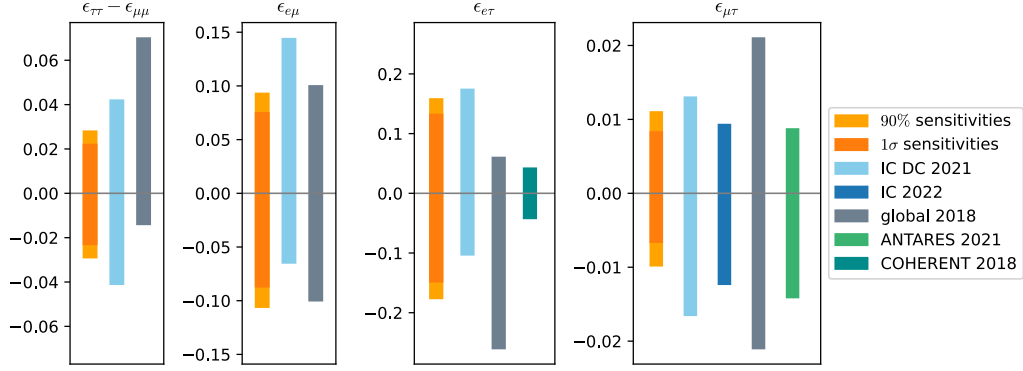


Fig. 7.24.: Comparison of the sensitivities obtained in this analysis, shown in orange, to existing 90% confidence intervals. These are set in previous DeepCore [40] and IceCube [52] measurements, results from ANTARES [53], COHERENT [154], and a global analysis [39]. Limits set in these analyses which are not competitive are not included in this figure. These are the ANTARES $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ limits as well as the COHERENT $\epsilon_{\mu\tau}^{\oplus}$ and $\epsilon_{e\mu}^{\oplus}$ limits.

7.9 Setting limits

The sensitivities and performed tests that were discussed in previous sections demonstrate that this analysis is in principle ready to run on DeepCore data. However, the following steps are only shown exemplarily on pseudo-data, as completing the analysis on data exceeds the scope of this thesis. Full review and approval of the analysis is a collaboration-wide process that, partly due to the involved review periods and partly due to the required computational resources to perform the appropriate tests, is expected to take several months. The pseudo-data that are therefore used in place of the DeepCore data set are a Monte Carlo simulated realization of the null hypothesis case where all bin counts are Poisson-fluctuated.

Since reasonable goodness of fit was obtained in a PMNS-matrix parameter analysis performed on the same data as this work [147], the sample is not expected to involve significant tension with respect to the null hypothesis. Therefore, no significant rejection of the standard-oscillations case is expected and the most relevant results of this analysis are the confidence limits that can be obtained for each tested NSI parameter.

Commonly, limits are set by fitting to the data at multiple NSI hypotheses that sample the full investigated NSI parameter space on a dense grid. From the resulting test statistic differences to the global best fit points, confidence limits are deduced based on Wilks' theorem. The resulting limits that this analysis can place on pseudo-data are given in table 7.10. For the realization of pseudo-data that is used here in place of data, all confidence intervals are slightly tighter than expected based on

Tab. 7.10.: Pseudo-data based best fit values and 90% confidence intervals of all tested NSI parameters for both neutrino mass orderings, assuming NO and using Wilks' theorem. Bounds for complex phases are given at $\delta_{\alpha\beta} = 0^\circ, 180^\circ$.

	confidence interval	best fit
$\epsilon_{\tau\tau}^\oplus - \epsilon_{\mu\mu}^\oplus$	$[-0.0164, 0.0167]$	-0.002
$\epsilon_{e\mu}^\oplus$	$[-0.087, 0.051]$	-0.023
$\epsilon_{e\tau}^\oplus$	$[-0.159, 0.096]$	-0.091
$\epsilon_{\mu\tau}^\oplus$	$[-0.0047, 0.0091]$	0.0022

sensitivities.

The dependency of the limits on FC NSI parameter magnitude on their complex phases is more pronounced than in the Asimov case, which results in stronger asymmetries of the magnitude limits at phases of 180° and 0° . Such variation in the phase dependent features are expected as these are strongly impacted by fluctuations in event counts as well as PMNS-matrix parameter values.

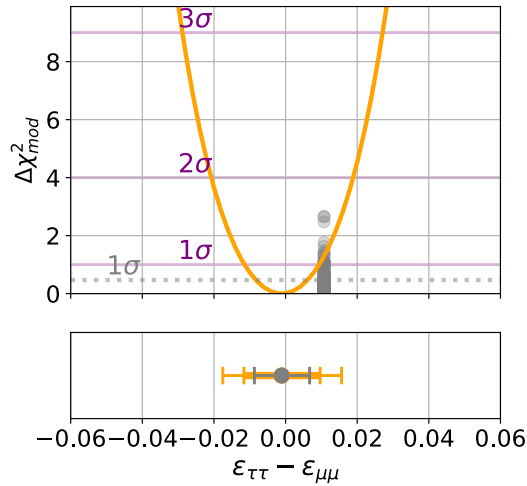


Fig. 7.25.: Test statistic deltas obtained for $\epsilon_{\tau\tau}^\oplus - \epsilon_{\mu\mu}^\oplus$ in orange and critical $n\sigma$ values assuming Wilks' theorem in purple. At the bottom, the width of the resulting 90% and 1 σ confidence intervals are given as thin and thick orange lines. At the upper 1 σ border, an ensemble of 200 pseudo-trials is drawn, from which a corrected critical 1 σ test statistic value is derived. Both are shown in gray, as dots and dotted line, respectively. At the bottom, the gray bar indicates the resulting confidence limit when assuming that the corrected critical value is constant throughout the parameter range. See appendix section B.12 for the limits obtained for FC parameters.

As discussed in section 7.6.4, the assumption of Wilks' theorem yields highly conservative estimates for this analysis. The alternative strategy described by Feldman and Cousins [155] involves performing ensembles of ~ 1000 trials at each tested NSI parameter value. For a $q\%$ confidence level, the test statistic value below which a

fraction of $q\%$ of these trials are located replaces the respective critical test statistic value determined through Wilks' theorem.

In a simplified approach to evaluating the difference with respect to Wilks' based limits, a single ensemble is drawn at the upper 1σ border for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$. The test statistic delta below which 68% of the trials are located serves as a corrected critical 1σ test statistic value that does not assume Wilks' theorem. The 1σ or 68% fraction of the drawn trials as well as the resulting $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ confidence interval estimated from this is shown in figure 7.25.

The full approach by Feldman and Cousins requires such ensembles at densely spaced hypotheses throughout the NSI parameter space, as the derived critical value can vary depending on the parameter value. Parameter ranges that are calculated based on a single critical test statistic value can therefore only be used as an approximation. In comparison to the Wilks' based 1σ limits of $[-0.0105, 0.0107]$, however, the ensemble-based limits of $[-0.0076, 0.0078]$ confirm the previous observation that Wilks' based limits on this analysis are highly conservative.

7.9.1 Computing Feldman Cousins limits

For cases like this analysis, where Wilks' theorem can not be applied without gaining significant bias, Feldman and Cousins [155] suggested a method based on sorted test statistic differences. This does not rely on the test statistic differences being distributed in a specific way, as is done for Wilks' theorem. Instead, the critical test statistic value corresponding to the desired quantile q is determined from pseudo-experiments at each parameter value μ_i that is tested with regard to being within or outside the respective CL. The approach as described in [155] is outlined in figure 7.26.

For a best fit result, the data are fitted with free nuisance and NSI parameters, yielding the overall hypothesis best describing the observation as well as the corresponding χ^2 value. In the following, NSI parameters are denoted as μ and nuisance parameters as α , with subscripts indicating their value to be fixed, in contrast to the parameter being free during a fit.

In order to calculate limits in the approach by Feldman and Cousins, for each tested NSI parameter value, a local fit to the data is performed, with free nuisance parameters while the NSI parameter is fixed to μ_i . The difference between the resulting local χ_{mod}^2 value and the overall best fit χ_{mod}^2 yields a $\Delta\chi_{mod}^2$ value that can later-on be compared to the $\Delta\chi_{mod}^2$ values obtained from pseudo-experiments.

An ensemble of n pseudo-experiments is drawn Poisson-distributed based on the

expected bin counts for the best fit hypothesis ($\mu_i, \alpha_{b.f. data}$). For each pseudo-experiment j , a $\Delta\chi_{mod i,j}^2$ is determined from two fits: One with free nuisance parameters and one with, in addition, free NSI parameters.

The requested quantile q of the $\Delta\chi_{mod i,*}^2$ for all pseudo-experiments within the ensemble is determined, and with that a $\Delta\chi_{mod i,q}^2$ limit value to which the $\Delta\chi_{mod}^2$ obtained from data can be compared. The NSI parameter value is within the CL corresponding to the requested quantile if $\Delta\chi_{mod i,data}^2$ is within $\Delta\chi_{mod i,q}^2$.

7.9.2 Making Feldman Cousins limits computationally feasible

The process described above is computationally expensive to a degree where it is not feasible, requiring $\mathcal{O}(10^8)$ CPU-hours, given the number of hypotheses to test for all five NSI parameters of $\mathcal{O}(10^5)$ as well as the approximately 10^3 trials per hypothesis. Different approaches to reducing this are described in the following.

One alternative to fitting the same hypotheses μ_i for data and pseudo-experiments is to calculate $\Delta\chi_{mod}^2$ values for the data on a grid, allowing for interpolation between the single points, and choosing the hypotheses where ensembles are drawn in a more informed way in order to reduce the overall computational cost. This is assumed to reduce the computation time by approximately a factor of five, depending on the implementation.

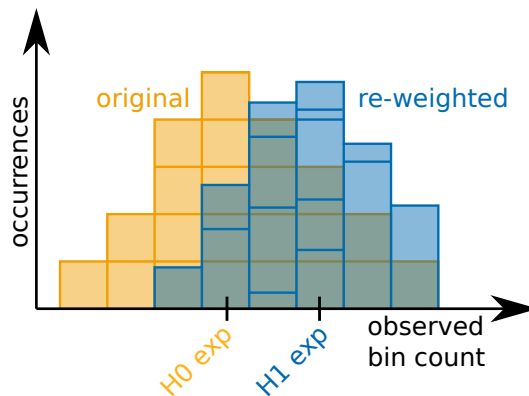


Fig. 7.26.: Schematic of pseudo-experiments drawn at hypothesis H_0 being re-weighted for a different hypothesis H_1 . For simplicity, the count in a single bin is shown, resulting in the pseudo-experiments being Poisson-distributed around the respective expectation value. In the re-weighted distribution, this is accomplished by multiplying each histogram entry with a weight based on the pseudo-experiment probabilities to be drawn for the two hypotheses.

For a given number of pseudo-experiments, the number of individual fits could be reduced by a factor of close to two by re-weighting pseudo-experiments for different

hypotheses μ_i . The basic concept is visualized in figure 7.26: Any pseudo-experiment pe_j drawn assuming a hypothesis $H0$ with a probability $P(pe|H0)$ has a probability $P(pe_j|H1)$ to be drawn based on a different hypothesis, $H1$. When re-weighting an ensemble of pseudo-experiments based on this, with weights $w_{pe_j} = \frac{P(pe_j|H1)}{P(pe_j|H0)}$, the resulting distribution for $H1$ closely matches that of an ensemble directly drawn at $H1$ in case that $H0$ and $H1$ differ little in their bin count expectations. For cases where $P(pe_j|H0) \ll P(pe_j|H1)$, extreme weights with large errors occur, making this approach unstable in a way that can not be handled easily, which was one reason to discard it for this analysis. A second reason was the incompatibility with the quantile regression method detailed below.

Quantile regression

The approaches to determining limits discussed above all rely on interpolating between quantiles that are calculated from ensembles of pseudo-experiments drawn at single hypotheses. Quantile regression, contrarily, assumes continuous limits which are fitted based on pseudo-experiments drawn at hypotheses distributed throughout the range of the tested NSI parameter, as depicted in figure 7.27. The CL is obtained from the intercept between an interpolated grid of fits to the data and the quantile.

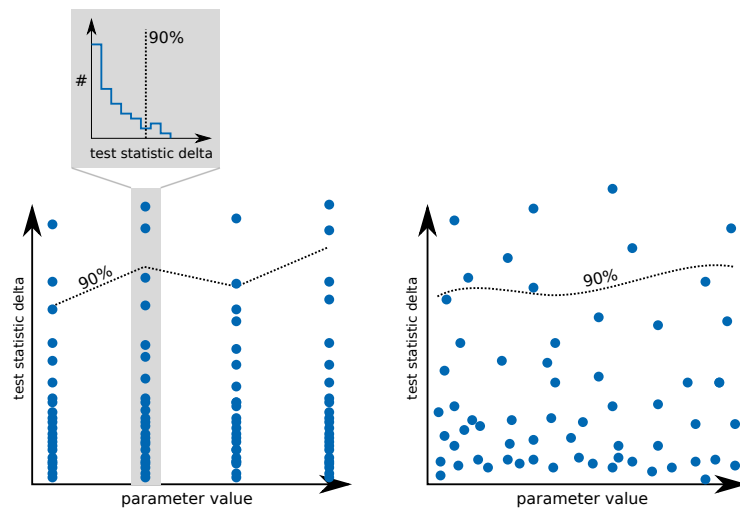


Fig. 7.27.: Visual comparison between calculating quantiles for individual hypotheses (left) and using quantile regression (right).

This is accomplished by shaping the problem at hand into a linear regression where the optimum solution is reached once the requested quantile is described cor-

rectly [93, 156]. A model $F(x)$, which describes the input values (x, y) , is optimized such that the metric PB of the absolute errors,

$$\sum_i PB_q(y_i - F(x_i)) \quad , \quad (7.13)$$

given the quantile q , becomes minimal. The Pinball loss, PB_q , is defined as

$$PB_q(t) = \begin{cases} qt, & t > 0 \\ 0 & t = 0 \\ (1 - q)t & t < 0 \end{cases} . \quad (7.14)$$

In this way, the term specified in equation 7.13 becomes minimal when $y < F(x)$ for a fraction of q of the input values.

The model $F(x)$ is a linear combination of the dimensions (or features) of x . A simple example for this describes a linear function,

$$F(x) = w_0 + w_1x \quad , \quad (7.15)$$

where the weights w_i are optimized. Different kinds of models become possible by choosing the features respectively: The space within which $F(x)$ is defined is linear, but can have e.g. x^2 as one of its dimensions.

In order to represent how a certain quantile varies depending on the NSI parameter value, the model chosen here uses splines. These are functions that are partially defined as $n_k + 1$ polynomials of a fixed degree n_{deg} , which are joined at n_k knots such that the overall spline is $n_{deg} - 1$ times differentiable [157].

For stability and error assessment, the bootstrap method [93] is used: Instead of only fitting splines to all input values and determining the CL from it, multiple randomized subsets are fitted and their mean and standard deviation are determined. Using the mean to obtain the CL grants less dependency of individual pseudo-experiments and the standard deviation indicates the CL accuracy.

The number of pseudo-experiments needed can be reduced by sampling them close to the edge of the confidence region. In an iterative approach, the intercepts between the data grid and the individual bootstraps of iteration n are sampled and Gauss-smearred to obtain the hypotheses for which the pseudo-experiments for iteration $n + 1$ are drawn.

In order to represent these more densely populated regions in more detail in the splines, their knots are positioned at quantiles of the hypotheses. As a result, the uncertainty on the CL decreases quicker with the number of pseudo-experiments than in a case where the hypotheses are drawn with less information.

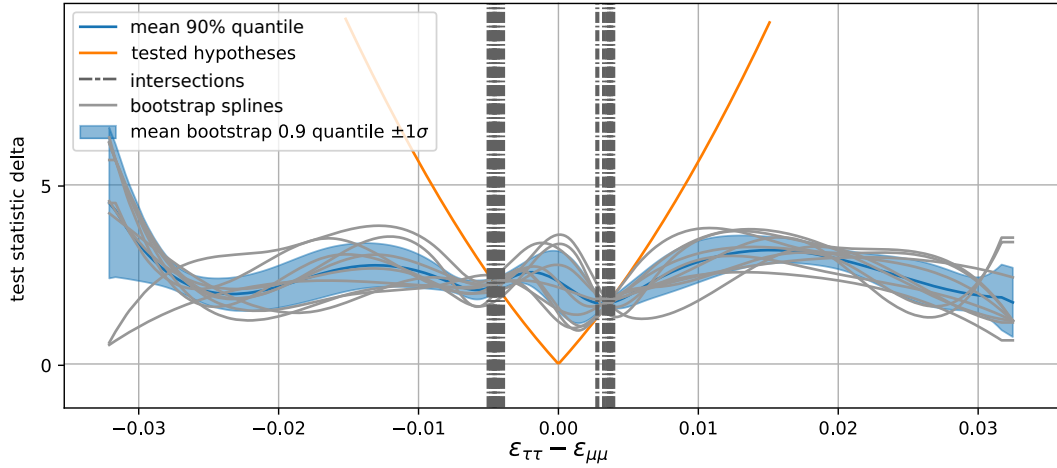


Fig. 7.28.: Second iteration of a quantile regression using B-splines. A 90% quantile is determined as the mean (blue line) of ten bootstrap splines (gray solid lines). Their standard deviation is shown in light blue as a measure for the error of the CL. The latter is obtained from the intersection of the mean and the $\Delta\chi^2$ values obtained from fitting the data at different parameter values, shown in orange. In the next iteration, new pseudo-experiments will be drawn around the intersections with the individual bootstrap splines (gray dotted-dashed line).

In figures 7.29 and 7.28, the outcome of a simulation with a simplified $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ signal is shown. In figure 7.28, the second iteration is shown as well as the individual bootstrapped splines and their intersections with the data scan. In iteration three, which can be seen in figure 7.29, pseudo-experiments are drawn around these intercepts. However, the pseudo-experiments from all previous iterations enter the fitting of the splines as well. The large density of pseudo-experiments close to intersections, which in addition causes a larger density of knots, results in increasingly small errors of the quantile. The lower density of knots throughout the rest of the parameter space becomes apparent as well, without however resulting in relevant bias, as can be seen from comparison to a BDT-based quantile that is shown purely for reference. The CL and its error, which result from this approach, are consequently assumed to be accurate, while needing significantly less pseudo-experiments than other methods for obtaining Feldman Cousins limits.

For complex-valued NSI parameters, one-dimensional limits of two simultaneously free dimensions can be determined in a way that is schematically shown in figure 7.30. All pseudo-experiments are drawn from hypotheses distributed throughout the two dimensions of the NSI parameter. The $\Delta\chi_{mod}^2$ values originate from fits performed with both parameters free versus both of them fixed. Quantiles are then fitted separately to the projected pseudo-experiments in each dimension.

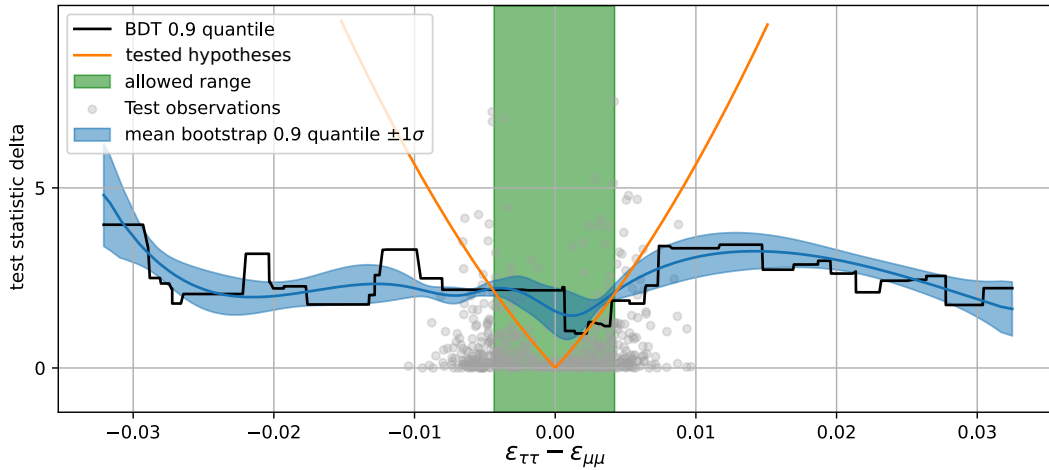


Fig. 7.29.: Third iteration of the same quantile regression shown in figure 7.28, with the blue line and range representing the same as therein. The pseudo-experiments newly drawn in this iteration (gray) cluster around the intersection between the splines and the $\Delta\chi_{mod}^2$ values from the data (orange). In black, a BDT-based quantile determined from the pseudo-experiments in the first iteration is given for reference.

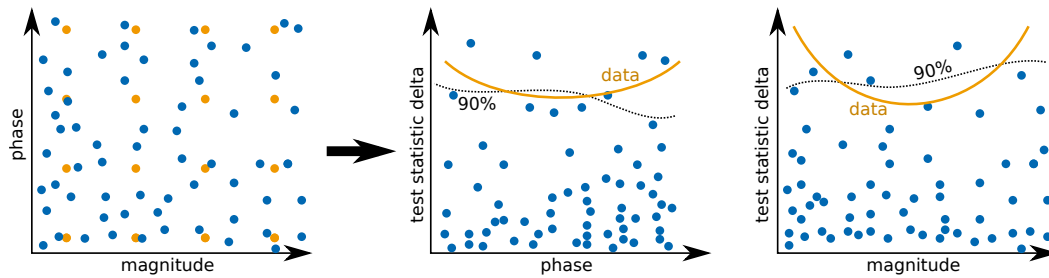


Fig. 7.30.: Illustration of how FC limits are computed for the simultaneously free phase and magnitude of flavor-changing NSI parameters. The resulting CL is the intersection of the $\Delta\chi_{mod}^2$ values determined from the data (orange) with the 90% quantile (black dotted line) which is fitted to the pseudo-experiments (blue) separately for phase and magnitude.

Following this or a similarly optimized implementation of Feldman Cousins limits is a necessary next step for this analysis, in order to compute NSI parameter limits that are not extremely conservative.

Conclusion

This work presents a search for neutrino interactions beyond the standard model of particle physics, or NSI. It is based on atmospheric neutrino data that is detected in the IceCube DeepCore detector [66], within which Earth matter effects on the neutrino flavor composition can be resolved. The observed variations depend on the neutrino energy, flavor and angle. The latter determines the Earth matter thickness and density profile through which the neutrino traverses [27].

The sample used in this work includes reconstructed neutrino energies of 5 GeV to 100 GeV and zenith angles from up-going to horizontal trajectories. For the 9.3 y of included DeepCore data, this amounts to a total of ~ 150000 events, exceeding previous samples [40] by approximately a factor of three.

The accuracy and computational requirements of DeepCore event reconstruction are of critical importance in this sample. In a first part of this work, the existing shortcomings and hypothetical potential of likelihood-based reconstruction algorithms for DeepCore events were explored in depth.

The main focus of this work, however, is a search for NSI originating from NC forward scattering of all neutrino flavors on first generation fermions in Earth matter. The utilized parametrization [2] includes two real-valued and three complex-valued NSI parameters which differ in their signal regions and features.

For the analysis, observed event counts binned in the observables are compared to corresponding template maps that are Monte Carlo generated based on different hypotheses. A test statistic serves as metric for the agreement between data and template [120]. This test statistic is optimized to find the hypothesis that best describes the data by evaluating differences in the respective expected event counts with respect to the observation. Each NSI parameter is treated individually in this analysis, with the magnitude and phase of complex-valued NSI parameters considered simultaneously. A minimization technique is presented that reliably masters the 18 to 19 dimensional parameter space. It is also shown that a newly developed nuisance parameter treatment that is based on likelihood-free inference can be employed on the present sample.

In evaluating the minimization performance, it is found that for fits to $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$, the available computational resources are not sufficient to consistently find the parameter space optimum. The resulting decision was to not further investigate this

NSI parameter in this work. The presented analysis principle is however assumed to be functional for $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$.

For all investigated NSI parameters, the Monte Carlo based sensitivities are found to be competitive with or exceed earlier results in IceCube and other experiments as well as global results. A goodness of fit of $p \sim 5\%$ was found in fits to the data sample for all NSI parameters. This may hint towards shortcomings in modeling the data and prompts for further investigation. Subsequent tests of the analysis are therefore performed on pseudo-data, including the calculation of limits based on Wilks' theorem [96]. This approach yields highly conservative limits, as is exemplarily shown for pseudo-data $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ limits in a simplified Feldman and Cousins [155] approach.

In a next step, possible reasons for the minor lack in goodness of fit need to be evaluated further. Once either these are known or it seems certain that the data sample is understood sufficiently well, the analysis can undergo the collaboration internal review process before being fully performed on data, potentially including $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$. Ultimately, full Feldman and Cousins based limits need to be set, in a resource-optimized manner that is outlined in this work. These steps require large computational resources and therefore could not be completed within the scope of this work. The final results for this analysis, namely limits on all NSI parameters, are expected to be competitive compared to other experiments.

In addition to the frequentist approach, within this work, Bayesian MCMC [148] sampling is explored and shown to be a viable secondary option, which allows to sample all NSI magnitudes simultaneously. Once a frequentist result is available, this can be included as a verification for the frequentist analysis and introduction of a conceptually different analysis approach.

Beyond this analysis, DeepCore data sets spanning even larger lifetimes will be available in the coming years. Furthermore, the IceCube upgrade [72], which is going to be build during the next years, will increase event detection rates, especially at the lowest detectable neutrino energies. This is expected to result in significant improvement of sensitivity to PMNS-matrix and NSI parameters.

In this way, NSI measurements might become a key element in understanding deficiencies of the current Standard Model of particle physics. Evidence for non-zero NSI could potentially resolve tensions between existing PMNS-matrix parameter measurements [55] and facilitate understanding the mechanisms behind non-zero neutrino masses [1].

Bibliography

- [1] Carlo Giunti and Chung W. Kim. *Fundamentals of Neutrino Physics and Astrophysics*. Oxford University Press, 2007.
- [2] Yasaman Farzan and Mariam Tórtola. “Neutrino Oscillations and Non-standard Interactions”. In: *Frontiers in Physics* 6 (2018). DOI: 10.3389/fphy.2018.00010.
- [3] P. S. Bhupal Dev, K. S. Babu, Peter B. Denton, et al. “Neutrino Non-Standard Interactions: A Status Report”. In: *SciPost Phys. Proc.* (2 2019), p. 1. DOI: 10.21468/SciPostPhysProc.2.001.
- [4] W. Pauli. *Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a. Band II: 1930–1939*. Springer-Verlag Berlin Heidelberg, 1985.
- [5] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire. “Detection of the Free Neutrino: a Confirmation”. In: *Science* 124.3212 (1956), pp. 103–104. DOI: 10.1126/science.124.3212.103.
- [6] Nobel Prize Press release. *Additional background material on the Nobel Prize in Physics*. 1995. URL: <https://www.nobelprize.org/prizes/physics/1995/advanced-information/> (visited on Jan. 19, 2023).
- [7] Emmy Noether. “Invariant variation problems”. In: *Transport Theory and Statistical Physics* 1.3 (1971), pp. 186–207. DOI: 10.1080/00411457108231446.
- [8] P.A. Zyla et al. “Review of Particle Physics”. In: *PTEP* 2020.8 (2020). and 2021 update, p. 083C01. DOI: 10.1093/ptep/ptaa104.
- [9] Mark Thomson. *Modern particle physics*. Cambridge, United Kingdom ; New York: Cambridge University Press, 2013. 554 pp.
- [10] S. Schael et al. “Precision electroweak measurements on the Z resonance”. In: *Phys. Rept.* 427 (2006), pp. 257–454. DOI: 10.1016/j.physrep.2005.12.006. arXiv: hep-ex/0509008.
- [11] J. A. Formaggio and G. P. Zeller. “From eV to EeV: Neutrino cross sections across energy scales”. In: *Rev. Mod. Phys.* 84 (3 Sept. 2012), pp. 1307–1341. DOI: 10.1103/RevModPhys.84.1307.
- [12] Nobel Prize Press release. *Press Release on the Nobel Prize in Physics*. 1988. URL: <https://www.nobelprize.org/prizes/physics/1988/press-release/> (visited on Jan. 19, 2023).
- [13] Nobel Prize Press release. *Advanced information on the Nobel Prize in Physics*. 2002. URL: <https://www.nobelprize.org/uploads/2018/06/advanced-physicsprize2002.pdf> (visited on Jan. 19, 2023).

- [14]Richard P. Feynman. “Very High-Energy Collisions of Hadrons”. In: *Phys. Rev. Lett.* 23 (24 Dec. 1969), pp. 1415–1417. DOI: 10.1103/PhysRevLett.23.1415.
- [15]Andrii Terliuk. “Measurement of atmospheric neutrino oscillations and search for sterile neutrino mixing with IceCube DeepCore”. PhD thesis. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2018. DOI: <http://dx.doi.org/10.18452/19304>.
- [16]L. Wolfenstein. “Neutrino oscillations in matter”. In: *Phys. Rev. D* 17 (1978), p. 2369.
- [17]A. Yu Smirnov. “Solar neutrinos: Oscillations or No-oscillations?” In: *arXiv:1609.02386 [astro-ph, physics:hep-ex, physics:hep-ph]* (Sept. 19, 2017). arXiv: 1609.02386.
- [18]Bruce T. Cleveland, Timothy Daily, Jr. Raymond Davis, et al. “Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector”. In: *The Astrophysical Journal* 496.1 (Mar. 1998), pp. 505–526. DOI: 10.1086/305343.
- [19]Lothar Oberauer and Judith Oberauer. *Neutrino physics: Grundlagen, Experimente und aktuelle Forschung*. de. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019. DOI: 10.1007/978-3-662-59335-6.
- [20]Q. R. Ahmad, R. C. Allen, T. C. Andersen, 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”. In: *Phys. Rev. Lett.* 87 (7 July 2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301.
- [21]B. Pontecorvo. “Inverse Beta Processes and Nonconservation of Lepton Charge”. In: *Sov. Phys. JETP* 7 (1958), p. 172.
- [22]B. Pontecorvo. “Mesonium and anti-mesonium”. In: *Sov. Phys. JETP* 6 (1957). [*Zh. Eksp. Teor. Fiz.* 33, 549 (1957)], p. 429.
- [23]Ziro Maki, Masami Nakagawa, and Shoichi Sakata. “Remarks on the Unified Model of Elementary Particles”. In: *Progress of Theoretical Physics* 28.5 (Nov. 1962), pp. 870–880. DOI: 10.1143/PTP.28.870. eprint: <https://academic.oup.com/ptp/article-pdf/28/5/870/5258750/28-5-870.pdf>.
- [24]X. Qian and P. Vogel. “Neutrino mass hierarchy”. In: *Progress in Particle and Nuclear Physics* 83 (July 2015), pp. 1–30. DOI: 10.1016/j.pnpnp.2015.05.002.
- [25]M. Aker et al. “Direct neutrino-mass measurement with sub-electronvolt sensitivity”. In: *Nature Phys.* 18.2 (2022), pp. 160–166. DOI: 10.1038/s41567-021-01463-1. arXiv: 2105.08533 [hep-ex].
- [26]Y. Fukuda, T. Hayakawa, E. Ichihara, et al. “Evidence for Oscillation of Atmospheric Neutrinos”. In: *Phys. Rev. Lett.* 81 (8 Aug. 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562.
- [27]P.A. Zyla et al. “Review of Particle Physics, 14. Neutrino Masses, Mixing, and Oscillations”. In: *PTEP* 2020.8 (2020). and 2021 update, p. 083C01. DOI: 10.1093/ptep/ptaa104.
- [28]André de Gouvêa. “Neutrino mass models”. In: *Ann. Rev. Nucl. Part. Sci.* 66 (2016), pp. 197–217. DOI: 10.1146/annurev-nucl-102115-044600.

- [29]Ingolf Bischer and Werner Rodejohann. “General neutrino interactions from an effective field theory perspective”. In: *Nuclear Physics B* 947 (2019), p. 114746. DOI: <https://doi.org/10.1016/j.nuclphysb.2019.114746>.
- [30]M. B. Gavela, D. Hernandez, T. Ota, and W. Winter. “Large gauge invariant nonstandard neutrino interactions”. In: *Physical Review D* 79.1 (Jan. 2009). DOI: 10.1103/physrevd.79.013007.
- [31]Tommy Ohlsson. “Status of non-standard neutrino interactions”. en. In: *Reports on Progress in Physics* 76.4 (Apr. 2013), p. 044201. DOI: 10.1088/0034-4885/76/4/044201.
- [32]S. Bilenky. “Neutrino oscillations: From a historical perspective to the present status”. In: *Nuclear Physics B* 908 (2016), pp. 2–13. DOI: <https://doi.org/10.1016/j.nuclphysb.2016.01.025>.
- [33]M. Blennow and A. Yu. Smirnov. “Neutrino propagation in matter”. In: *Adv. High Energy Phys.* 2013 (2013), p. 972485. DOI: 10.1155/2013/972485.
- [34]S. P. Mikheyev and A. Yu. Smirnov. “Resonant amplification of ν oscillations in matter and solar-neutrino spectroscopy”. In: *Il Nuovo Cimento C* 9 (1986), pp. 17–26. DOI: 10.1007/BF02508049.
- [35]A. M. Dziewonski and D. L. Anderson. “Preliminary reference Earth model”. In: *Phys. Earth Planet. Interiors* 25 (1981), pp. 297–356. DOI: 10.1016/0031-9201(81)90046-7.
- [36]E. Kh. Akhmedov, M. Maltoni, and A. Yu. Smirnov. “Neutrino oscillograms of the Earth: effects of 1-2 mixing and CP-violation”. In: *J. High Energy Phys.* 2008.6, 072 (June 2008), p. 072. DOI: 10.1088/1126-6708/2008/06/072.
- [37]E. K. Akhmedov, M. Maltoni, and A. Yu. Smirnov. “1-3 leptonic mixing and the neutrino oscillograms of the Earth”. In: *J. High Energy Phys.* 2007.5, 077 (May 2007), p. 077. DOI: 10.1088/1126-6708/2007/05/077.
- [38]M. C. Gonzalez-Garcia, M. Maltoni, and J. Salvado. “Testing matter effects in propagation of atmospheric and long-baseline neutrinos”. In: *J. High Energy Phys.* 2011.5, 75 (May 2011), p. 75. DOI: 10.1007/JHEP05(2011)075.
- [39]Ivan Esteban, M. C. Gonzalez-Garcia, Michele Maltoni, Ivan Martinez-Soler, and Jordi Salvado. “Updated constraints on non-standard interactions from global analysis of oscillation data”. In: *Journal of High Energy Physics* 2018.8 (2018), p. 180. DOI: 10.1007/JHEP08(2018)180.
- [40]R. Abbasi, M. Ackermann, J. Adams, et al. “All-flavor constraints on nonstandard neutrino interactions and generalized matter potential with three years of IceCube DeepCore data”. In: *Phys. Rev. D* 104 (7 Oct. 2021), p. 072006. DOI: 10.1103/PhysRevD.104.072006.
- [41]M. Honda, M. Sajjad Athar, T. Kajita, K. Kasahara, and S. Midorikawa. “Atmospheric neutrino flux calculation using the NRLMSISE-00 atmospheric model”. In: *Phys. Rev. D* 92 (2 July 2015), p. 023004. DOI: 10.1103/PhysRevD.92.023004.

- [42]DUNE Collaboration, R. Acciarri, M. A. Acero, et al. *Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) Conceptual Design Report Volume 2: The Physics Program for DUNE at LBNF*. 2015. DOI: 10.48550/ARXIV.1512.06148.
- [43]Takashi Kikuchi, Hisakazu Minakata, and Shoichi Uchinami. “Perturbation theory of neutrino oscillation with nonstandard neutrino interactions”. In: *Journal of High Energy Physics* 2009.3 (Mar. 23, 2009), pp. 114–114. DOI: 10.1088/1126-6708/2009/03/114.
- [44]K. S. Babu, Dorival Gonçalves, Sudip Jana, and Pedro A. N. Machado. “Neutrino non-standard interactions: Complementarity between LHC and oscillation experiments”. In: *Physics Letters B* 815 (2021), p. 136131. DOI: <https://doi.org/10.1016/j.physletb.2021.136131>.
- [45]Joachim Kopp. “New phenomena in neutrino physics”. PhD thesis. Combined Faculties of the Natural Sciences and Mathematics of the Ruperto-Carola-University of Heidelberg, 2009.
- [46]Tommy Ohlsson, He Zhang, and Shun Zhou. “Effects of nonstandard neutrino interactions at PINGU”. In: *Physical Review D* 88.1 (July 3, 2013), p. 013001. DOI: 10.1103/PhysRevD.88.013001. arXiv: 1303.6130.
- [47]O. G. Miranda and H. Nunokawa. “Non standard neutrino interactions: current status and future prospects”. en. In: *New Journal of Physics* 17.9 (Sept. 2015). Publisher: IOP Publishing, p. 095002. DOI: 10.1088/1367-2630/17/9/095002.
- [48]Arman Esmaili and Alexei Yu Smirnov. “Probing Non-Standard Interaction of Neutrinos with IceCube and DeepCore”. en. In: *Journal of High Energy Physics* 2013.6 (June 2013). arXiv: 1304.1042, p. 26. DOI: 10.1007/JHEP06(2013)026.
- [49]Thomas Ehrhardt. *internal communication*.
- [50]Esteban Roulet. “Mikheyev-Smirnov-Wolfenstein effect with flavor-changing neutrino interactions”. In: *Phys. Rev. D* 44 (4 Aug. 1991), R935–R938. DOI: 10.1103/PhysRevD.44.R935.
- [51]Mattias Blennow. “Theoretical and Phenomenological Studies of Neutrino Physics”. In: 2007.
- [52]R. Abbasi, M. Ackermann, J. Adams, et al. “Strong Constraints on Neutrino Nonstandard Interactions from TeV-Scale ν_μ Disappearance at IceCube”. In: *Phys. Rev. Lett.* 129 (1 June 2022), p. 011804. DOI: 10.1103/PhysRevLett.129.011804.
- [53]A. Albert et al. “Search for non-standard neutrino interactions with 10 years of ANTARES data”. In: *JHEP* 07 (2022), p. 048. DOI: 10.1007/JHEP07(2022)048. arXiv: 2112.14517 [hep-ex].
- [54]Ushak Rahaman, Soebur Razzaque, and Sankagiri Uma Sankar. “A Review of the Tension between the T2K and NOvA Appearance Data and Hints to New Physics”. In: *Universe* 8.2 (2022). DOI: 10.3390/universe8020109.

- [55] Ivan Esteban, M. C. Gonzalez-Garcia, and Michele Maltoni. “On the effect of NSI in the present determination of the mass ordering”. en. In: *arXiv:2004.04745 [hep-ex, physics.hep-ph]* (Apr. 2020). arXiv: 2004.04745.
- [56] Christian Spiering. “Towards high-energy neutrino astronomy”. In: *The European Physical Journal H* 37 (2012), pp. 515–565. DOI: <https://doi.org/10.1140/epjh/e2012-30014-2>.
- [57] E. G. Adelberger, A. Garcia, R. G. Hamish Robertson, et al. “Solar fusion cross sections. II. The pp chain and CNO cycles”. In: *Reviews of Modern Physics* 83.1 (Apr. 2011), pp. 195–245. DOI: 10.1103/revmodphys.83.195.
- [58] Carlos A. Argüelles, Mauricio Bustamante, Ali Kheirandish, et al. *Fundamental physics with high-energy cosmic neutrinos today and in the future*. 2019. DOI: 10.48550/ARXIV.1907.08690.
- [59] Lorenzo Bonechi, Raffaello D’Alessandro, and Andrea Giammanco. “Atmospheric muons as an imaging tool”. In: *Reviews in Physics* 5 (2020), p. 100038. DOI: <https://doi.org/10.1016/j.revip.2020.100038>.
- [60] I. M. Frank and Ig. Tamm. “Coherent visible radiation of fast electrons passing through matter”. In: *Physics-Uspekhi* 93 (1937), pp. 388–393.
- [61] M. Shiozawa. “Reconstruction algorithms in the Super-Kamiokande large water Cherenkov detector”. In: *Nucl. Instrum. Meth. A* 433 (1999). Ed. by A. Breskin, R. Chechik, and T. Ypsilantis, pp. 240–246. DOI: 10.1016/S0168-9002(99)00359-9.
- [62] IceCube Collaboration, M. G. Aartsen, M. Ackermann, et al. “The IceCube Neutrino Observatory: Instrumentation and Online Systems”. In: *Journal of Instrumentation* 12.3 (Mar. 14, 2017), P03012–P03012. DOI: 10.1088/1748-0221/12/03/P03012. arXiv: 1612.05093.
- [63] M. Ageron et al. “ANTARES: the first undersea neutrino telescope”. In: *Nucl. Instrum. Meth. A* 656 (2011), pp. 11–38. DOI: 10.1016/j.nima.2011.06.103. arXiv: 1104.1607 [astro-ph.IM].
- [64] M. G. Aartsen et al. “Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector”. In: *Science* 342 (2013), p. 1242856. DOI: 10.1126/science.1242856. arXiv: 1311.5238 [astro-ph.HE].
- [65] IceCube Collaboration, M. Aartsen, M. Ackermann, et al. “Measurement of the multi-TeV neutrino interaction cross-section with IceCube using Earth absorption”. In: *Nature* 551.7682 (Nov. 2017), pp. 596–600. DOI: 10.1038/nature24459.
- [66] R. Abbasi et al. “The design and performance of DeepCore”. In: *Astropart. Phys.* 35 (2012), pp. 615–624. DOI: 10.1016/j.astropartphys.2012.01.004.
- [67] M. G. Aartsen, R. Abbasi, Y. Abdou, et al. “Measurement of South Pole ice transparency with the IceCube LED calibration system”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 711 (2013), pp. 73–89. DOI: <https://doi.org/10.1016/j.nima.2013.01.054>.

- [68]R. Abbasi, Y. Abdou, T. Abu-Zayyad, et al. “Calibration and characterization of the IceCube photomultiplier tube”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 618.1 (2010), pp. 139–152. DOI: <https://doi.org/10.1016/j.nima.2010.03.102>.
- [69]Hamamatsu Photonics K.K. *Photomultiplier tubes and assemblies for scintillation counting & high energy physics*. Revised Apr. 2017.
- [70]M. G. Aartsen, M. Ackermann, J. Adams, et al. “Measurement of atmospheric tau neutrino appearance with IceCube DeepCore”. In: *Physical Review D* 99.3 (Feb. 15, 2019), p. 032007. DOI: [10.1103/PhysRevD.99.032007](https://doi.org/10.1103/PhysRevD.99.032007).
- [71]Lutz Köpke for the IceCube Collaboration. “Improved Detection of Supernovae with the IceCube Observatory”. In: *Journal of Physics: Conference Series* 1029 (May 2018), p. 012001. DOI: [10.1088/1742-6596/1029/1/012001](https://doi.org/10.1088/1742-6596/1029/1/012001).
- [72]Aya Ishihara and on behalf of the IceCube Collaboration. “The IceCube Upgrade - Design and Science Goals”. In: *Proceedings of 36th International Cosmic Ray Conference — PoS(ICRC2019)*. 36th International Cosmic Ray Conference. Vol. 358. SISSA Medialab, July 2, 2021, p. 1031. DOI: [10.22323/1.358.1031](https://doi.org/10.22323/1.358.1031).
- [73]IceCube collaboration. *SPICE 3.2.1 ice model*. URL: https://code.icecube.wisc.edu/projects/icecube/browser/IceCube/projects/ice-models/trunk/resources/models/spice_3.2.1 (visited on May 15, 2022).
- [74]Martin Rongen. “Calibration of the IceCube Neutrino Observatory”. PhD thesis. Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University, 2019.
- [75]RAID collaboration. *Ground penetrating radar*. 2018. URL: <http://www.rapidaccessicedrill.org/wp-content/uploads/2015/11/RAID-borehole-targets.jpg>.
- [76]R. Abbasi, M. Ackermann, J. Adams, et al. “In-situ estimation of ice crystal properties at the South Pole using LED calibration data from the IceCube Neutrino Observatory”. In: *The Cryosphere Discussions* 2022 (2022), pp. 1–48. DOI: [10.5194/tc-2022-174](https://doi.org/10.5194/tc-2022-174).
- [77]Martin Leuermann. “Testing the Neutrino Mass Ordering with IceCube DeepCore”. PhD thesis. Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University, 2018.
- [78]M. Sajjad Athar and S. K. Singh. *The Physics of Neutrino Interactions*. 1st ed. Cambridge University Press, June 30, 2020. DOI: [10.1017/9781108489065](https://doi.org/10.1017/9781108489065).
- [79]M.A. Zucker M.J. Berger J.S. Coursey and J. Chang. *Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions*. Tech. rep. NIST Standard Reference Database 124, NISTIR 4999. National Institute of Standards and Technology (NIST), Physics Laboratory (now Physical Measurement Laboratory (PML)), 7/2017. DOI: [10.18434/T4NC7P](https://doi.org/10.18434/T4NC7P).
- [80]Markus Vehring. “Measurement of Neutrino Oscillations with IceCube-DeepCore”. PhD thesis. Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University, 2017.

- [81]K.A. Olive. “Review of Particle Physics”. In: *Chinese Physics C* 40.10 (Oct. 2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.
- [82]Michael James Larson. “Simulation and Identification of Non-Poissonian Noise Triggers in the IceCube Neutrino Detector”. Masters thesis. University of Alabama, 2013.
- [83]The IceCube collaboration. “Measurement of Atmospheric Neutrino Mixing with Improved IceCube DeepCore Calibration and Data Processing”. [Unpublished]. 2023.
- [84]C. Andreopoulos, A. Bell, D. Bhattacharya, et al. “The GENIE neutrino Monte Carlo generator”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 614.1 (2010), pp. 87–104. DOI: <https://doi.org/10.1016/j.nima.2009.12.009>.
- [85]D. Heck, Johannes Knapp, J-N. Capdevielle, George C. Schatz, and T. J. Thouw. “CORSIKA: A Monte Carlo code to simulate extensive air showers”. In: 1998.
- [86]R. S. Fletcher, T. K. Gaisser, Paolo Lipari, and Todor Stanev. “sibyll: An event generator for simulation of high energy cosmic ray cascades”. In: *Phys. Rev. D* 50 (9 Nov. 1994), pp. 5710–5731. DOI: 10.1103/PhysRevD.50.5710.
- [87]Felix Riehn, Ralph Engel, Anatoli Fedynitch, Thomas K. Gaisser, and Todor Stanev. “Hadronic interaction model sibyll 2.3d and extensive air showers”. In: *Phys. Rev. D* 102 (6 Sept. 2020), p. 063002. DOI: 10.1103/PhysRevD.102.063002.
- [88]J-M Alameddine, J Soedingrekso, A Sandrock, M Sackel, and W Rhode. “PROPOSAL: A library to propagate leptons and high energy photons”. In: *Journal of Physics: Conference Series* 1690.1 (Dec. 2020), p. 012021. DOI: 10.1088/1742-6596/1690/1/012021.
- [89]S. Agostinelli, J. Allison, K. Amako, et al. “Geant4—a simulation toolkit”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. DOI: [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [90]Leif Rädcl and Christopher Wiebusch. “Calculation of the Cherenkov light yield from low energetic secondary particles accompanying high-energy muons in ice and water with Geant4 simulations”. In: *Astroparticle Physics* 38 (2012), pp. 53–67. DOI: <https://doi.org/10.1016/j.astropartphys.2012.09.008>.
- [91]Claudio Kopper. *clsim*. URL: github.com/claudiok/clsim (visited on Aug. 21, 2022).
- [92]IceCube internal documentation. *OscNext*. and technical notes linked therein. URL: <https://wiki.icecube.wisc.edu/index.php/OscNext> (visited on Aug. 28, 2022).
- [93]Ilya Narsky and Frank C. Porter. *Statistical Analysis Techniques in Particle Physics: Fits, Density Estimation and Supervised Learning*. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, Nov. 13, 2013. DOI: 10.1002/9783527677320.
- [94]Shiqi Yu and on behalf of the IceCube collaboration. “Direction reconstruction using a CNN for GeV-scale neutrinos in IceCube”. In: *Journal of Instrumentation* 16.11 (Nov. 2021), p. C11001. DOI: 10.1088/1748-0221/16/11/C11001.

- [95]Shiqi Yu and Jessie Micallef. *Internal presentation on Muon Classifiers for FLERCNN Sample*. URL: https://wiki.icecube.wisc.edu/index.php/Low_Energy_and_Oscillations_Phone_Meetings#Mar_14 (visited on Nov. 21, 2022).
- [96]G. Cowan. *Statistical Data Analysis*. Oxford science publications. Clarendon Press, 1998.
- [97]R. Abbasi, M. Ackermann, J. Adams, et al. 2022. DOI: 10.48550/ARXIV.2203.02303.
- [98]IceCube internal documentation by Marcel Usner. *SpiceLea Anisotropy Photospline Tables*. URL: <https://icecube.wisc.edu/~musner/docs/anisotropy/> (visited on Aug. 21, 2022).
- [99]IceCube internal documentation. *Photospline*. URL: wiki.icecube.wisc.edu/index.php/Photospline (visited on Aug. 22, 2022).
- [100]IceCube internal documentation. *clsim*. URL: docs.icecube.aq/icetray/main/projects/clsim/index.html (visited on Aug. 21, 2022).
- [101]J. Lundberg, P. Miočinić, K. Woschnagg, et al. “Light tracking through ice and water—Scattering and absorption in heterogeneous media with Photonics”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 581.3 (Nov. 2007), pp. 619–631. DOI: 10.1016/j.nima.2007.07.143.
- [102]Elisa Lohfink. “Low energy neutrino interaction reconstruction at the IceCube experiment”. Masters thesis. Johannes Gutenberg-Universität Mainz, 2019.
- [103]M G Aartsen, R Abbasi, M Ackermann, et al. “Energy reconstruction methods in the IceCube neutrino telescope”. In: *Journal of Instrumentation* 9.03 (Mar. 2014), P03009. DOI: 10.1088/1748-0221/9/03/P03009.
- [104]IceCube internal documentation and repositories linked therein. *Gulliver*. URL: wiki.icecube.wisc.edu/index.php/Gulliver (visited on Nov. 26, 2022).
- [105]F. Feroz, M. P. Hobson, and M. Bridges. “MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics”. In: *Monthly Notices of the Royal Astronomical Society* 398.4 (Sept. 2009), pp. 1601–1614. DOI: 10.1111/j.1365-2966.2009.14548.x. eprint: <https://academic.oup.com/mnras/article-pdf/398/4/1601/3039078/mnras0398-1601.pdf>.
- [106]Maicon Hieronymus. “Reconstruction of Low Energy Neutrino Events with GPUs at IceCube”. Masters thesis. Johannes Gutenberg-Universität Mainz, 2019.
- [107]Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton. “ImageNet Classification with Deep Convolutional Neural Networks”. In: *NeurIPS* 25 (Jan. 2012). DOI: 10.1145/3065386.
- [108]Eric Kauderer-Abrams. *Quantifying Translation-Invariance in Convolutional Neural Networks*. 2018.
- [109]“Deep Learning in Physics exemplified by the Reconstruction of Muon-Neutrino Events in IceCube”. In: *PoS ICRC2017* (2017), p. 1057. DOI: 10.22323/1.301.1057.

- [110] Jessie Micallef. *FLERCNN Low energy neural network*. URL: github.com/jessimic/LowEnergyNeuralNetwork (visited on Nov. 23, 2022).
- [111] Tom Stuttard et al. *OscNext software Metaproject*. URL: http://code.icecube.wisc.edu/svn/sandbox/stuttard/oscNext_meta/ (visited on Aug. 28, 2022).
- [112] The IceCube Collaboration UW Madison. *The IceCube NPX HTCondor computing cluster*.
- [113] Douglas Thain, Todd Tannenbaum, and Miron Livny. “Distributed computing in practice: the Condor experience.” In: *Concurrency - Practice and Experience* 17.2-4 (2005), pp. 323–356.
- [114] Elisa Lohfink. *OscNext reconstruction software Metaproject*. URL: <http://code.icecube.wisc.edu/svn/sandbox/elohfink/oscNextReco/> (visited on Aug. 29, 2022).
- [115] IceCube internal documentation and technical note linked therein. *FLERCNN*. URL: wiki.icecube.wisc.edu/index.php/FLERCNN (visited on Nov. 22, 2022).
- [116] S. S. Wilks. “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”. In: *The Annals of Mathematical Statistics* 9.1 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360.
- [117] Thomas Ehrhardt. “Studies of the Reconstruction of Cascade-Like Events in PINGU”. Masters thesis. Universität Bonn, 2014.
- [118] Jan Weldert. “Likelihood-free inference for IceCube low energy reconstruction”. PhD thesis. Johannes Gutenberg-Universität Mainz, Fachbereich Physik, Mathematik und Informatik, 2022.
- [119] Philipp Eller, Aaron Fienberg, Jan Weldert, et al. *A flexible event reconstruction based on machine learning and likelihood principles*. 2022. DOI: 10.48550/ARXIV.2208.10166.
- [120] M.G. Aartsen, M. Ackermann, J. Adams, et al. “Computational techniques for the analysis of small signals in high-statistics neutrino oscillation experiments”. en. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 977 (Oct. 2020), p. 164332. DOI: 10.1016/j.nima.2020.164332.
- [121] Emanuel Parzen. “On Estimation of a Probability Density Function and Mode”. In: *The Annals of Mathematical Statistics* 33.3 (1962), pp. 1065–1076. DOI: 10.1214/aoms/1177704472.
- [122] Murray Rosenblatt. “Remarks on Some Nonparametric Estimates of a Density Function”. In: *The Annals of Mathematical Statistics* 27.3 (1956), pp. 832–837. DOI: 10.1214/aoms/1177728190.
- [123] The IceCube collaboration. *PISA*. URL: <https://github.com/icecube/pisa> (visited on Nov. 11, 2022).
- [124] Elisa Lohfink. *OscNext NSI analysis repository*. URL: https://github.com/icecube/wg-oscillations-fridge/tree/master/analysis/oscNext_NSI (visited on Nov. 27, 2022).

- [125]G. D. Barr, S. Robbins, T. K. Gaisser, and T. Stanev. “Uncertainties in atmospheric neutrino fluxes”. In: *Phys. Rev. D* 74 (9 Nov. 2006), p. 094009. DOI: 10.1103/PhysRevD.74.094009.
- [126]Tania Kozynets Anatoli Fedynitch et al. *MCEq, version 1.1.3*. URL: <https://github.com/afedynitch/MCEq> (visited on Nov. 12, 2022).
- [127]Hans Dembinski, Ralph Engel, Anatoli Fedynitch, et al. “Data-driven model of the cosmic-ray flux and mass composition from 10 GeV to 10^{11} GeV”. In: (2017). DOI: 10.48550/ARXIV.1711.11432.
- [128]The Super-Kamiokande collaboration. *Prob3++*. URL: <https://webhome.phy.duke.edu/~raw22/public/Prob3++/> (visited on Nov. 13, 2022).
- [129]V. Barger, K. Whisnant, S. Pakvasa, and R. J. N. Phillips. “Matter effects on three-neutrino oscillations”. In: *Phys. Rev. D* 22 (11 Dec. 1980), pp. 2718–2726. DOI: 10.1103/PhysRevD.22.2718.
- [130]Ivan Esteban, M. C. Gonzalez-Garcia, Alvaro Hernandez-Cabezudo, Michele Maltoni, and Thomas Schwetz. “Global analysis of three-flavour neutrino oscillations: synergies and tensions in the determination of θ_{23} , δ_{CP} , and the mass ordering”. In: *Journal of High Energy Physics* 2019.1 (Jan. 2019). DOI: 10.1007/jhep01(2019)106.
- [131]NuFIT 4.0. www.nu-fit.org. 2018.
- [132]Elisa Lohfink Thomas Ehrhardt. *NSI implementation pull request to the PISA repository*. URL: <https://github.com/IceCubeOpenSource/pisa/pull/567> (visited on Nov. 13, 2022).
- [133]M. Glück, E. Reya, and A. Vogt. “Dynamical parton distributions revisited”. In: *The European Physical Journal C* 5.3 (Sept. 1998), pp. 461–470. DOI: 10.1007/s100529800978.
- [134]Amanda Cooper-Sarkar, Philipp Mertsch, and Subir Sarkar. “The high energy neutrino cross-section in the Standard Model and its uncertainty”. In: *Journal of High Energy Physics* 2011.8 (Aug. 2011). DOI: 10.1007/jhep08(2011)042.
- [135]Philipp Eller. *Hole ice model for effective angular DOM acceptance*. URL: github.com/icecube/angular_acceptance (visited on Nov. 16, 2022).
- [136]Karl Pearson F.R.S. “LIII. On lines and planes of closest fit to systems of points in space”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2.11 (1901), pp. 559–572. DOI: 10.1080/14786440109462720. eprint: <https://doi.org/10.1080/14786440109462720>.
- [137]Alexander Trettin. “Search for eV-scale sterile neutrinos with IceCube DeepCore”. [Unpublished]. 2023.
- [138]*The mogon computing cluster*. Johannes Gutenberg University Mainz (hpc.uni-mainz.de), AHRP (Alliance for High Performance Computing in Rhineland Palatinate, www.ahrp.info) and the Gauss Alliance e.V.
- [139]Institute for Cyber-Enabled Research at Michigan State University. *The ICER HPC computing cluster*.

- [140] Steven G. Johnson. *The NLOpt nonlinear-optimization package*. URL: <http://github.com/stevengj/nlopt> (visited on Dec. 11, 2022).
- [141] WL Price. “A controlled random search procedure for global optimization”. In: *The Computer Journal* 20 (4 1977), pp. 367–370. DOI: 10.1093/comjnl/20.4.367.
- [142] P. Kaelo and Montaz Ali. “Some Variants of the Controlled Random Search Algorithm for Global Optimization”. In: *Journal of Optimization Theory and Applications* 130 (Jan. 2006), pp. 253–264. DOI: 10.1007/s10957-006-9101-0.
- [143] T.H. Rowan. “Functional Stability Analysis of Numerical Algorithms”. PhD thesis. Department of Computer Sciences, University of Texas at Austin, 1990.
- [144] J. A. Nelder and R. Mead. “A Simplex Method for Function Minimization”. In: *The Computer Journal* 7.4 (Jan. 1965), pp. 308–313. DOI: 10.1093/comjnl/7.4.308. eprint: <https://academic.oup.com/comjnl/article-pdf/7/4/308/1013182/7-4-308.pdf>.
- [145] Hans Dembinski and Piti Ongmongkolkul et al. “scikit-hep/iminuit”. In: (Dec. 2020). DOI: 10.5281/zenodo.3949207.
- [146] F. James and M. Roos. “Minuit - a system for function minimization and analysis of the parameter errors and correlations”. In: *Computer Physics Communications* 10.6 (1975), pp. 343–367. DOI: [https://doi.org/10.1016/0010-4655\(75\)90039-9](https://doi.org/10.1016/0010-4655(75)90039-9).
- [147] The IceCube oscillations working group. *Internal communication*.
- [148] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. “emcee: The MCMC Hammer”. In: *Publications of the Astronomical Society of the Pacific* 125.925 (Feb. 2013), p. 306. DOI: 10.1086/670067.
- [149] W. D. Vousden, W. M. Farr, and I. Mandel. “Dynamic temperature selection for parallel tempering in Markov chain Monte Carlo simulations”. In: *Monthly Notices of the Royal Astronomical Society* 455.2 (Nov. 2015), pp. 1919–1937. DOI: 10.1093/mnras/stv2422.
- [150] Jonathan Goodman and Jonathan Weare. “Ensemble samplers with affine invariance”. In: *Communications in Applied Mathematics and Computational Science* 5.1 (2010), pp. 65–80. DOI: 10.2140/camcos.2010.5.65.
- [151] Dan Foreman-Mackey and contributors. *Documentation of the emcee python package*. URL: emcee.readthedocs.io (visited on Dec. 4, 2022).
- [152] Vasyl Hafych, Philipp Eller, Oliver Schulz, and Allen Caldwell. *Parallelizing MCMC Sampling via Space Partitioning*. 2020. DOI: 10.48550/ARXIV.2008.03098.
- [153] The Open MPI Development Team. *Open MPI: Open Source High Performance Computing*. URL: www.open-mpi.org (visited on Dec. 4, 2022).
- [154] Peter B. Denton, Yasaman Farzan, and Ian M. Shoemaker. “Testing large non-standard neutrino interactions with arbitrary mediator mass after COHERENT data”. In: *JHEP* 07 (2018), p. 037. DOI: 10.1007/JHEP07(2018)037. arXiv: 1804.03660 [hep-ph].

- [155] Gary J. Feldman and Robert D. Cousins. “Unified approach to the classical statistical analysis of small signals”. In: *Physical Review D* 57.7 (Apr. 1998), pp. 3873–3889. DOI: 10.1103/physrevd.57.3873.
- [156] scikit learn. *documentation on Quantile Regression*. URL: https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.QuantileRegressor.html (visited on Jan. 17, 2023).
- [157] J. H. Ahlberg, E. N. Nilson, and Joseph L. Walsh. “The theory of splines and their applications”. In: *Mathematics of Computation* 23 (1969), p. 443.
- [158] COHERENT Collaboration, D. Akimov, P. An, et al. “The COHERENT Experiment at the Spallation Neutron Source”. In: *arXiv:1509.08702 [hep-ex, physics:nucl-ex, physics:physics]* (Apr. 3, 2016). arXiv: 1509.08702.
- [159] D. Akimov, J. B. Albert, P. An, et al. “Observation of coherent elastic neutrino-nucleus scattering”. In: *Science* 357.6356 (Sept. 15, 2017), pp. 1123–1126. DOI: 10.1126/science.aao0990.
- [160] Tao Han, Jiajun Liao, Hongkai Liu, and Danny Marfatia. “Nonstandard neutrino interactions at COHERENT, DUNE, T2HK and LHC”. In: *Journal of High Energy Physics* 2019.11 (Nov. 6, 2019), p. 28. DOI: 10.1007/JHEP11(2019)028.
- [161] Peter B Denton, Yasaman Farzan, and Ian M Shoemaker. “Testing large non-standard neutrino interactions with arbitrary mediator mass after COHERENT data”. In: (2018), p. 21.

Acronyms

G_F Fermi coupling constant.

χ_{mod}^2 modified Pearson's chi-squared.

ANTARES the Astronomy with a Neutrino Telescope and Abyss environmental Research project.

ATWD Analog Transient Waveform Digitizer.

BDT Boosted Decision Tree.

BFRv2 birefringence ice model.

Borexino the Boron Experiment.

BSM-physics physics beyond the standard model.

CC charged current.

CE ν NS coherent elastic neutrino-nucleus scattering.

CL confidence limit.

CNN convolutional neural network.

CNO cycle Carbon-Nitrogen-Oxygen cycle.

CP charge and parity conjugation.

CPU central processing unit.

CR cosmic ray.

CRS2 controlled random search.

DIS deep inelastic scattering.

DOF degrees of freedom.

DOM digital optical module.

DUNE Deep Underground Neutrino Experiment.

EM electromagnetic.

FADC Fast Analog-to-Digital Converter.

FC flavor changing.

FD flavor diagonal.

FLERCNN Fast Low Energy Reconstruction using Convolutional Neural Networks.

GALLEX/GNO the Gallium Experiment/Gallium Neutrino Observatory.

GPU graphics processing unit.

HLC hard local coincidence.

HQE high quantum efficiency.

HT high throughput.

HV high voltage.

IBD inverse β -decay.

ICL IceCube laboratory.

IO inverted ordering.

IQR inter-quantile range.

KamLAND Kamioka Liquid Scintillator Antineutrino Detector.

KATRIN Karlsruhe Tritium Neutrino.

KDE kernel density estimation.

LBL long baseline.

LEERA Low-Energy Energy Reconstruction Algorithm.

LLH logarithmized likelihood.

MC Monte Carlo.

MCEq Matrix Cascade Equations.

MCMC Markov Chain Monte Carlo.

MCPE Monte Carlo photoelectron.

MINOS the Main Injector Neutrino Oscillation Search.

ML machine learning.

MSW Mikheyev-Smirnov-Wolfenstein.

MultiNest Multimodal Nested sampling.

NC neutral current.

NMO neutrino mass ordering.

NO normal ordering.

NO ν A NuMI Off-Axis ν_e Appearance experiment.

NSI nonstandard interactions.

PCA principal component analysis.

PDF probability density function.

PE photoelectrons.

PID particle identification.

PISA PINGU Simulation and Analysis.

PMNS-matrix Pontecorvo Maki Nakagawa Sakata mixing matrix.

PMT photomultiplier tube.

pp chain proton-proton chain.

PREM Preliminary Reference Earth Model.

ptemcee adaptive parallel tempering MCMC implementation.

QCD quantum chromodynamics.

QE quantum efficiency.

QES quasi-elastic scattering.

RAM random-access memory.

RDE relative DOM efficiencies.

RENO the Reactor Experiment for Neutrino Oscillation.

RES resonant scattering.

RETRO Reverse Table Reconstruction.

RT-cleaning Radius-Time-cleaning.

SAGE the Soviet–American Gallium Experiment.

SANTA single string ANTARES-inspired analysis.

SBL short baseline.

SLC soft local coincidence.

SM standard model of particle physics.

SMT simple multiplicity trigger.

SNO Sudbury Neutrino Observatory.

SPice 3.2.1 South Pole ice.

SSM solar standard model.

Super-K the Super-Kamiokande experiment.

SVN Apache Subversion.

T2K Tokai to Kamioka experiment.

TPC time projection chamber.

w.e. water equivalent.

List of Figures

2.1	The standard model of particle physics.	4
2.2	Fundamental weak vertices.	4
2.3	Fundamental electromagnetic vertices.	5
2.4	Cross sections of different neutrino interactions.	6
2.5	Total neutrino cross sections.	7
2.6	Deep inelastic scattering processes.	8
3.1	Neutrino mass ordering.	13
3.2	Neutrino forward scattering.	18
3.3	Neutrino survival probability within the Earth represented in the PREM scheme.	20
3.4	Neutrino interactions inside the Earth for SM and NSI.	25
3.5	Neutrino transition probabilities for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ at $\cos(\theta) = -0.75$	27
3.6	Antineutrino transition probabilities for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ at $\cos(\theta) = -0.75$	28
3.7	Neutrino transition probabilities for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ at $\cos(\theta) = -1$	28
3.8	Neutrino transition probabilities for $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$	29
3.9	Neutrino transition probabilities for $\epsilon_{\mu\tau}^{\oplus}$	30
3.10	Neutrino transition probabilities for $\epsilon_{\mu\tau}^{\oplus}$ at $\delta_{\mu\tau} = \pi/2$	31
3.11	Neutrino transition probabilities for $\epsilon_{e\mu}^{\oplus}$	31
3.12	Neutrino transition probabilities for $\epsilon_{e\tau}^{\oplus}$	32
4.1	Neutrino flux from different sources over energy.	35
4.2	Solar neutrino energy spectrum.	37
4.3	Neutrino flux and interaction in a hypothetical reactor experiment.	38
4.4	Cosmic ray flux.	39
4.5	Atmospheric neutrino flux at the South Pole.	40
4.6	Characteristics of cosmic rays.	41
4.7	Schematic of the Cherenkov effect.	45
4.8	Layout of a DOM.	47
4.9	Functioning principle of a photomultiplier tube.	47
4.10	Wavelength dependence of IceCube modules.	48
4.11	The IceCube detector.	49

4.12	Measured South Pole ice properties.	50
4.13	Effects within the South Pole ice.	51
4.14	Photon signatures of neutrino events.	52
4.15	Muon energy losses per distance over muon energy.	54
5.1	Event simulation steps.	59
5.2	Muon reduction BDT performance.	68
5.3	Agreement between data and Monte Carlo in energy and cosine of the zenith angle.	70
5.4	Event rates throughout selection levels and agreement between data and Monte Carlo in PID.	71
6.1	Reconstruction internal event representation.	77
6.2	Table generation for reconstruction.	78
6.3	Minimization in the pegleg reconstruction.	81
6.4	Event reconstruction with SANTA.	84
6.5	Event reconstruction run-time comparison.	86
6.6	Resolution comparison between table-based reconstructions.	88
6.7	Energy resolution comparison between the FLERCNN and RETRO re- construction.	89
6.8	Angular resolution comparison between the FLERCNN and RETRO reconstruction.	89
6.9	Projections of hypotheses evaluated during reconstruction of two events.	92
6.10	Distribution of likelihood differences.	93
6.11	Overview of different ways to generate simplified events.	94
6.12	Reconstruction performance for an improved relation between track length and energy.	97
6.13	Disagreement in charge expectation values.	99
6.14	Reconstruction performance in energy and angle for improved charge modeling.	100
6.15	Reconstruction performance for interaction position and time for im- proved charge modeling.	101
6.16	Impact of the pulse time bin size on LLH.	102
6.17	Reconstruction bias towards DOM positions.	103
6.18	Likelihood space behavior for non-causal interaction times.	104
6.19	Reconstructed parameter values for simplified events and multiple reconstructions of the same events.	108
6.20	Results from multiple reconstructions of the same event and hit pattern.	109
7.1	Neutrino oscillation probabilities with no NSI.	113

7.2	Neutrino oscillation probabilities with NSI.	114
7.3	Classification of PID within FLERCNN.	115
7.4	Expected event counts for the null hypothesis and $ \epsilon_{\mu\tau}^{\oplus} = 0.02$	117
7.5	Overview of PISA stages.	119
7.6	Depictions of assumptions going into the flux and oscillation modeling.	121
7.7	Impact of hole ice parameters.	127
7.8	Impact of the errors on bin counts on the test statistic, given the applied nuisance parameter treatment.	131
7.9	Maps and histograms of bin-wise pulls in the total neutrino flux for a selection of simulated sets.	134
7.10	Detector nuisance parameter interpolation in two exemplary bins.	135
7.11	Schematic of how the impact of an individual nuisance parameter is tested.	137
7.12	Impact of fixing individual nuisance-parameters to their nominal value when fitting to an off-nominal hypothesis for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ and $\epsilon_{e\tau}^{\oplus}$	138
7.13	The analysis minimization process.	140
7.14	Results of fit performance tests for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ and $\epsilon_{\mu\tau}^{\oplus}$	142
7.15	Results of fit performance tests for $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$	143
7.16	Fluctuated and fitted parameter values including $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$	146
7.17	Test statistic distributions for an ensemble of pseudo-data trials with $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ free.	148
7.18	Two MCMC sampling implementations.	150
7.19	Measures for the efficiency and accuracy of sampling.	151
7.20	Outcome of sampling runs.	152
7.21	Result of a MCMC sampling run.	155
7.22	Characteristics of the test statistic values throughout the sampling process.	156
7.23	Asimov sensitivities for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ and $\epsilon_{e\tau}^{\oplus}$	158
7.24	Comparison of sensitivities to existing results.	159
7.25	Pseudo-data based confidence limits on $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$	160
7.26	Re-weighting pseudo-experiments.	162
7.27	Different approaches to calculating quantiles.	163
7.28	Second iteration of a quantile regression using B-splines.	165
7.29	Third iteration of a quantile regression using B-splines.	166
7.30	Quantile regression for complex-valued NSI parameters.	166
A.1	Reconstruction improvement in multi-PMT modules.	196
A.2	Test of the gradient-based nuisance parameter interpolation.	197
A.3	Offset correction for nuisance parameter modeling.	198
B.1	Antineutrino transition probabilities for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ at $\cos(\theta) = -1$	199

B.2	Neutrino transition probabilities for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ at $\cos(\theta) = -0.5$	199
B.3	Antineutrino transition probabilities for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ at $\cos(\theta) = -0.5$	200
B.4	Neutrino transition probabilities for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ at $\cos(\theta) = -0.25$	200
B.5	Antineutrino transition probabilities for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ at $\cos(\theta) = -0.25$	200
B.6	Antineutrino transition probabilities for $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$	201
B.7	Antineutrino transition probabilities for $\epsilon_{\mu\tau}^{\oplus}$	201
B.8	Antineutrino transition probabilities for $\epsilon_{\mu\tau}^{\oplus}$ at $\delta_{\mu\tau} = \pi/2$	201
B.9	Neutrino transition probabilities for $\epsilon_{e\mu}^{\oplus}$ at $\delta_{\mu\tau} = \pi/2$	202
B.10	Antineutrino transition probabilities for $\epsilon_{e\mu}^{\oplus}$	202
B.11	Antineutrino transition probabilities for $\epsilon_{e\mu}^{\oplus}$ at $\delta_{e\mu} = \pi/2$	202
B.12	Neutrino transition probabilities for $\epsilon_{e\tau}^{\oplus}$ at $\delta_{e\mu} = \pi/2$	203
B.13	Antineutrino transition probabilities for $\epsilon_{e\tau}^{\oplus}$	203
B.14	Antineutrino transition probabilities for $\epsilon_{e\tau}^{\oplus}$ at $\delta_{e\tau} = \pi/2$	203
B.15	Single DOM contributions to the LLH.	204
B.16	Energy and cosine of the zenith angle reconstruction performance for four algorithms.	205
B.17	Performance comparison for different ice models used in the pegen reconstruction.	205
B.18	Oscillation probability difference for $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus} = -2.2$ as the only source of NSI.	206
B.19	Oscillation probability difference for $\epsilon_{e\tau}^{\oplus} = 0.2$ as the only source of NSI.	206
B.20	Oscillation probability difference for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus} = 0.04$ as the only source of NSI.	207
B.21	Nuisance parameter modeling performance for sets that include BFRv2.	210
B.22	Nuisance parameter modeling performance for sets where the ice scattering and absorption parameters are varied simultaneously.	211
B.23	Nuisance parameter modeling performance for sets where only DOM efficiency is varied.	211
B.24	Nuisance parameter modeling performance for sets where DOM efficiency and hole ice parameters are varied simultaneously, except for sets where the value of p0 is below its nominal value.	212
B.25	Nuisance parameter modeling performance for sets where the value of p0 is below its nominal value.	212
B.26	Detailed nuisance parameter modeling performance for sets where the value of p0 is below its nominal value.	213
B.27	Nuisance parameter impact for $\epsilon_{e\mu}^{\oplus}$	214
B.28	Nuisance parameter impact for $\epsilon_{\mu\tau}^{\oplus}$ and $\epsilon_{e\tau}^{\oplus}$	215
B.29	Fit performance test results for $\epsilon_{e\mu}^{\oplus}$	216
B.30	Fit performance test results for $\epsilon_{e\tau}^{\oplus}$	216

B.31	Mis-modeling values for fitting performance tests of $\epsilon_{\mu\tau}^{\oplus}$	217
B.32	Fluctuated and fitted parameter values including $\epsilon_{e\mu}^{\oplus}$	217
B.33	Fluctuated and fitted parameter values including $\epsilon_{e\tau}^{\oplus}$	218
B.34	Fluctuated and fitted parameter values including $\epsilon_{\mu\tau}^{\oplus}$	219
B.35	Test statistic behavior for fits to pseudo-experiments including $\epsilon_{e\tau}^{\oplus}$	220
B.36	Test statistic behavior for fits to pseudo-experiments including $\epsilon_{e\mu}^{\oplus}$	220
B.37	Test statistic behavior for fits to pseudo-experiments including $\epsilon_{\mu\tau}^{\oplus}$	221
B.38	Sensitivity for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ at IO.	221
B.39	Sensitivity for $\epsilon_{e\tau}^{\oplus}$ at IO.	222
B.40	Sensitivities of $\epsilon_{\mu\tau}^{\oplus}$ at different mass orderings.	222
B.41	Sensitivities of $\epsilon_{e\mu}^{\oplus}$ at different mass orderings.	222
B.42	Burn-in phase of a MCMC sampling run.	223
B.43	Sampling phase of a MCMC sampling run.	224
B.44	Pseudo-data based limits on $\epsilon_{\mu\tau}^{\oplus}$	225
B.45	Pseudo-data based limits on $\epsilon_{e\mu}^{\oplus}$ and $\epsilon_{e\tau}^{\oplus}$	225

List of Tables

3.1	Current status of NSI measurements.	33
5.1	DeepCore data and simulated event rates at different levels of filtering and selection.	64
5.2	Event rejection at final selection level.	69
5.3	Event rates and counts in the final sample.	72
7.1	Analysis binning of the observables.	115
7.2	Flux related nuisance parameters.	122
7.3	Neutrino mass and oscillation related nuisance parameters.	123
7.4	Cross section related nuisance parameters.	125
7.5	Normalization related nuisance parameters.	125
7.6	Detector and ice related nuisance parameters.	128
7.7	Nuisance parameter values of the Monte Carlo sets shown in figure 7.9.	133
7.8	Fractions of negative mis-modeling.	145
7.9	Sensitivity based confidence intervals.	157
7.10	Pseudo-data based best fit values and confidence intervals.	160
A.1	Perturbation theoretical approach to NSI parameter impact.	195
B.1	Monte Carlo sets with off-nominal DOM efficiency.	208
B.2	Monte Carlo sets with off-nominal DOM efficiency and hole ice parameters.	208
B.3	Monte Carlo sets with off-nominal bulk ice properties.	209
B.4	Monte Carlo sets using the BFRv2 ice model.	209
B.5	Off-nominal Monte Carlo sets included in nuisance parameter treatment	213
B.6	Tested off-nominal nuisance parameter values.	215

Supplementary information

A.1 Massive neutrino extensions to the SM

In this section, some of the numerous possible neutrino mass extensions to the standard model of particle physics will be introduced briefly, pointing out such mechanisms that naturally produce NSI.

In general, neutrino mass extensions to the SM that rely on dimension six operators potentially naturally include NSI. More detail on this can be found e.g. in [31].

Dirac masses

A Dirac mass term constitutes the minimal extension of the SM. While established for charged leptons, for which both chiralities take part in interactions, a Dirac neutrino mass term would imply a symmetry between neutrinos of opposite chirality. To accommodate this, the existence of sterile neutrinos, in addition to active ones, becomes necessary [1].

The mass term,

$$-\mathcal{L}_{M_\nu, Dirac} = M_{Dij} \bar{\nu}_{Ri} \nu_{Lj} + h.c. \quad , \quad (A.1)$$

generates the complex Dirac mass matrix M_D after spontaneous symmetry breaking through Yukawa coupling of the scalar Higgs doublet ϕ with the left-handed lepton doublet $L_{Li} = (\nu_i, l)_L^T$ [27] via

$$M_{Dij} \bar{\nu}_{Ri} \nu_{Lj} = Y_{ij}^\nu \bar{\nu}_{Ri} \tilde{\phi}^\dagger L_{Lj} \quad (A.2)$$

$$= Y_{ij}^\nu \frac{v}{\sqrt{2}} \bar{\nu}_{Ri} \nu_{Lj} \quad . \quad (A.3)$$

Here, $\nu^c = C\bar{\nu}^T$ is the charge conjugated neutrino field with C the charge conjugation matrix, Y denotes the hyper-charge and v is the Higgs field vacuum expectation value [27].

Detecting the sterile neutrinos required in this extension would provide a strong hint towards a Dirac mass term, while not excluding other, additional mass terms.

Majorana masses

A Majorana mass term could be added to the SM without requiring additional particles such as sterile neutrinos and would only break accidental SM symmetries leading to e.g. lepton number conservation. It arises for example from a generic extension of the SM with operators of increasing order. Such an extension assumes a new physics energy scale Λ , with BSM-physics effects suppressed by Λ^{-4} . New physics would accordingly start to manifest in dimension five operators above the energy scale Λ [1].

After spontaneous symmetry breaking, such a dimension five operator constitutes a Majorana mass term.

$$-\mathcal{L}_{M_\nu, Majorana} = \frac{1}{2} M_{Mij} \bar{\nu}_{Ri} \nu_{Rj}^c + h.c. \quad (\text{A.4})$$

It maintains gauge invariance as it regards left- and right-handed neutrinos as singlets under gauge transformations. As the mass term includes direct coupling of neutrinos to anti-neutrinos, it requires neutrinos being their own anti-particles.

The see-saw mechanism

One very general way to extend the SM Lagrangian is adding both a Dirac mass term and a Majorana mass term with the full Lagrangian

$$-\mathcal{L}_{M_\nu} = -\mathcal{L}_{M_\nu, Dirac} - \mathcal{L}_{M_\nu, Majorana} \quad (\text{A.5})$$

$$= \frac{1}{2} (\bar{\nu}_L^C, \bar{\nu}_R) \underbrace{\begin{pmatrix} 0 & M_D^T \\ M_D & M_M \end{pmatrix}}_{\equiv M_\nu} \begin{pmatrix} \vec{\nu}_L \\ \vec{\nu}_L^C \end{pmatrix} + h.c. \quad (\text{A.6})$$

The mass eigenvalues m_i of the three active neutrinos $\vec{\nu}_L$ and s sterile neutrinos $\vec{\nu}_R$ can be derived from diagonalizing M_ν .

If the Majorana mass eigenvalues of M_M are significantly larger than the vacuum expectation value of the Higgs field v , this diagonalization divides neutrinos into two weight groups. It yields three mostly left-handed lighter neutrinos with weights proportional to M_M^{-1} as well as s heavier, mostly right-handed neutrinos with their mass proportional to M_M [27]. This so-called type-I see-saw mechanism is one way to explain the different mass-scale of neutrinos with respect to other elementary particles.

The type-II see-saw mechanism [31] explains the neutrino mass scale similarly to type-I see-saw but for small non-zero Majorana masses and directly entails NSI.

A.2 Perturbation theory approach to the impact of individual NSI parameters

The impact of individual NSI parameters can be approximated using a perturbation theoretical approach where all NSI parameters $\epsilon_{\alpha\beta}$ are assumed to be $\epsilon \sim 10^{-2}$. This was done in [43]. The order of ϵ at which the respective parameter enters a signal channel is shown in table A.1.

Tab. A.1.: Orders of ϵ at which individual NSI parameters enter oscillation probabilities in ϵ perturbation theory, taken from and explained in detail in [43]. This assumes maximal 2-3 mixing.

Channel	$\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$	$\epsilon_{e\mu}^{\oplus}$	$\epsilon_{e\tau}^{\oplus}$	$\epsilon_{\mu\tau}^{\oplus}$	$\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$
$P_{e\alpha}$ ($\alpha = e, \mu, \tau$)	ϵ^3	ϵ^2	ϵ^2	ϵ^3	ϵ^3
$P_{\alpha\beta}$ ($\alpha, \beta = \mu, \tau$)	ϵ^3	ϵ^2	ϵ^2	$\epsilon^1(\epsilon^2)$	$\epsilon^1(\epsilon^2)$

A.3 Coherent Elastic Neutrino-Nucleus Scattering (CE ν NS)

The process of CE ν NS is dominant typically below 50 MeV [158]. Its detection relies on NC neutrino scattering off nuclei, measured through the recoil energy of the nuclei, which amounts to tens of keV at most and can be achieved with multiple complementary methods. The first detection was accomplished in 2017 in the COHERENT experiment [159], which measures MeV neutrinos produced in a spallation neutron source via pion and muon decays.

The sensitivity of CE ν NS to NSI [160, 161] covers NSI models with hypothetical mediator masses at the MeV scale and above [160].

A.4 Reconstruction within the IceCube Upgrade

As was introduced in section 4.4.2, the additional strings that will be deployed for the IceCube-Upgrade contain multi-PMT modules [72]. In comparison to hypothet-

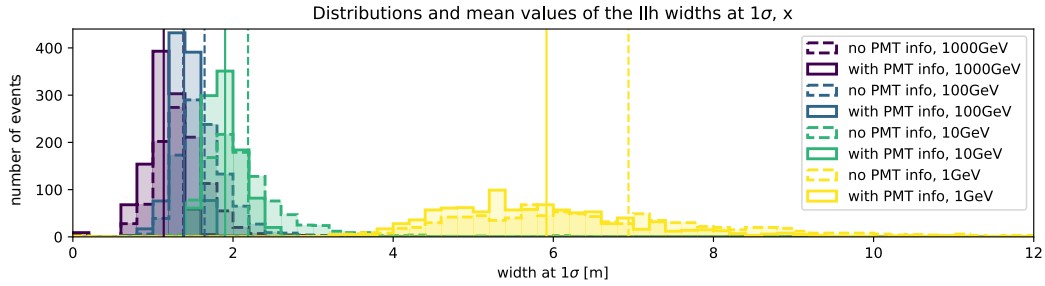


Fig. A.1.: Widths of the LLH minima in reconstructed x position for event sets simulated at different energies. In dashed lines, the behavior observed for unavailable PMT information is shown and can be compared to the cases where this is available, shown in solid lines. Vertical lines show the mean of the respective distributions.

ical IceCube DOMs with the same characteristics, such as their effective area, the advantage of multiple PMTs covering different directions is mainly in single-module coincidences. This is expected to facilitate triggering especially for extremely low event energies at the order of single GeV.

The impact of this was approximately quantified in a simplified simulation and reconstruction evaluation, based on the width of the likelihood minimum. Figure A.1 shows a comparison between including and excluding information of what specific PMT detected a hit, clearly preferring the former.

A.5 Offset in nuisance parameter gradients

In order to test the calculated gradients, simulated sets are compared to gradient-based predictions. They are evaluated based on the distributions of statistical pulls between both in individual analysis bins. The distribution of these pulls ideally fits a normal distribution around $\mu = 0$ with a standard deviation of $\sigma = 1$.

The observed outcome includes sets with overall shifts, resulting in $\mu \neq 0$. Such an overall bias between a single set and the gradient-based weights could potentially impact the analysis outcome, depending on whether it is equivalent to a significant shift in the respective off-nominal nuisance-parameters. In order to be significant, such a shift would have to be at the scale of the prior or the range of the respective nuisance parameter. The reason for that is that the exact values of the fitted nuisance parameters only become relevant if they exceed multiple sigma. In such cases, the fit result is considered not trustworthy.

In order to demonstrate that no significant impact is expected from the observed shifts, set 0500, which includes off-nominal ice scattering and absorption both at

1.05, is exemplarily modified as shown in figure A.3. The approach of this test is shown in figure A.2. When setting the nuisance parameter values assumed for this set to ice scattering of 1.045 and absorption of 1.039, the originally observed mean of -0.2 shifts to -0.01 . The mean values and standard deviations of all individual neutrino flavors and interactions appear reasonable at these settings. Based on the Gaussian priors of the nuisance parameters, their modification is at 0.05σ for scattering and 0.22σ for absorption. Since this is significantly below the multiple sigmas that would potentially render a fit implausible, this is considered a demonstration of how the observed shifts of the pull distributions are not relevant to the analysis outcome.

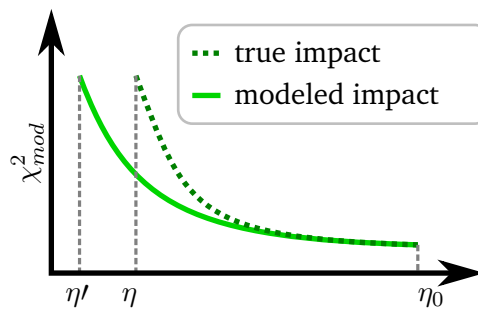
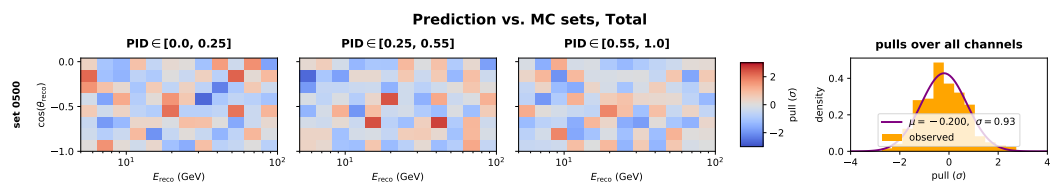
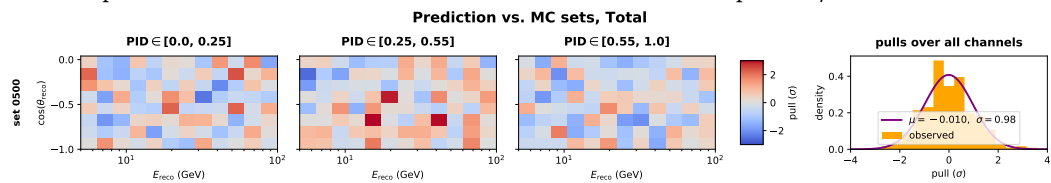


Fig. A.2.: Schematic of the test of the gradient-based nuisance parameter interpolation shown in figure A.3. The overall offset in gradient-based event count prediction for a off-nominal Monte Carlo set at nuisance parameter value η can be compensated for by shifting the parameter to η' . Since $\eta' - \eta$ is small compared to the prior of the nuisance parameter, the overall event count offset is considered not problematic.



(a) Maps for the correct nuisance parameter settings in the 0500 set, namely ice scattering and absorption of 1.05. This results in an overall shift of the statistical pulls of $\mu = -0.2$.



(b) Maps for the adapted nuisance parameter values, namely ice scattering of 1.045 and absorption of 1.039.

Fig. A.3.: Maps and histogram of statistical pulls for the 0500 data set. The pulls are between the simulated set and the nominal set which is re-weighted to fit the hypothesis assumed in the 0500 set. The weights are based on the gradients calculated for detector nuisance parameters.

Supplementary figures and tables

B.1 NSI phenomenology

Transition probabilities at different NSI hypotheses, supplementary to what is discussed in section 3.4. The standard oscillations case is shown in black, negative parameter values are shown in blue and positive ones in red. For non-zero phases of FC parameters, only positive magnitudes are shown.

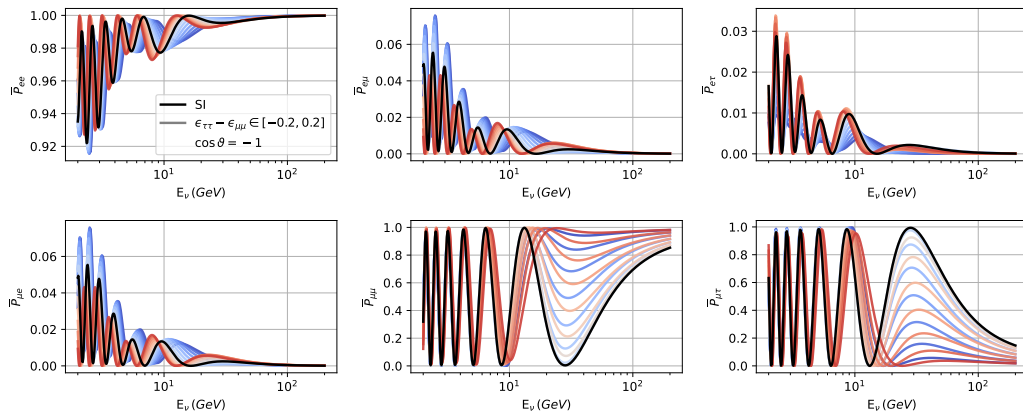


Fig. B.1.: Same as figure 3.5, but for antineutrinos and $\cos(\theta) = -1$.

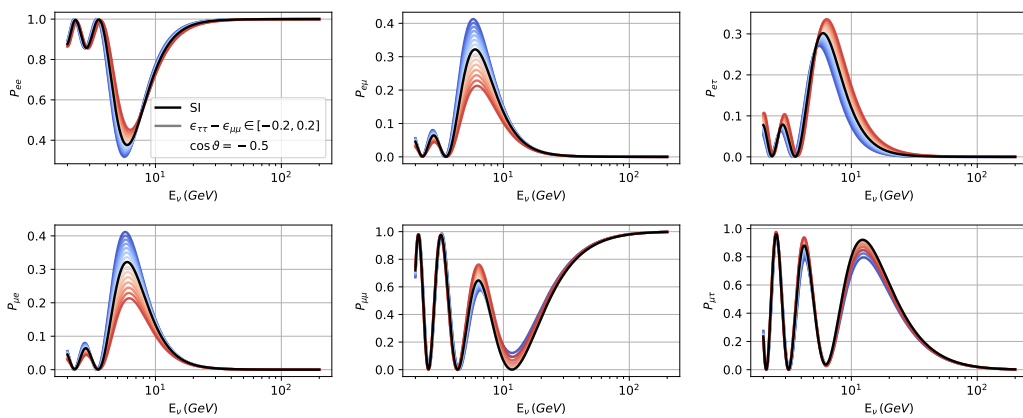


Fig. B.2.: Same as figure 3.5, but for $\cos(\theta) = -0.5$.

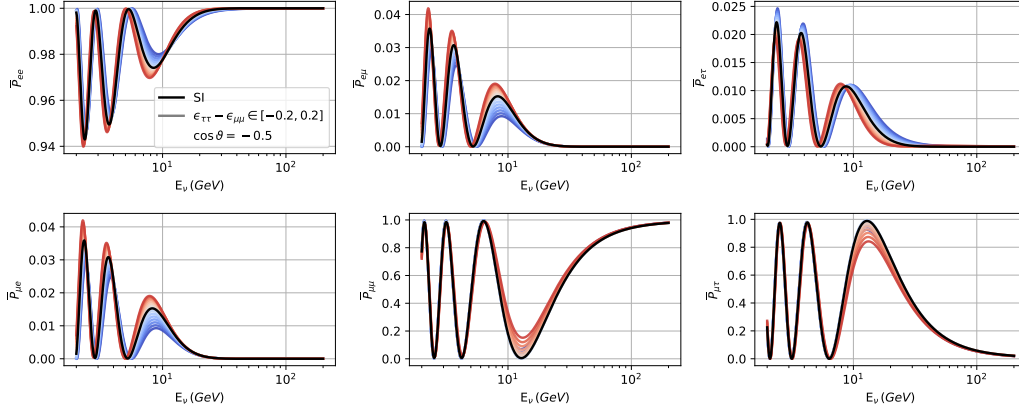


Fig. B.3.: Same as figure 3.5, but for antineutrinos and $\cos(\theta) = -0.5$.

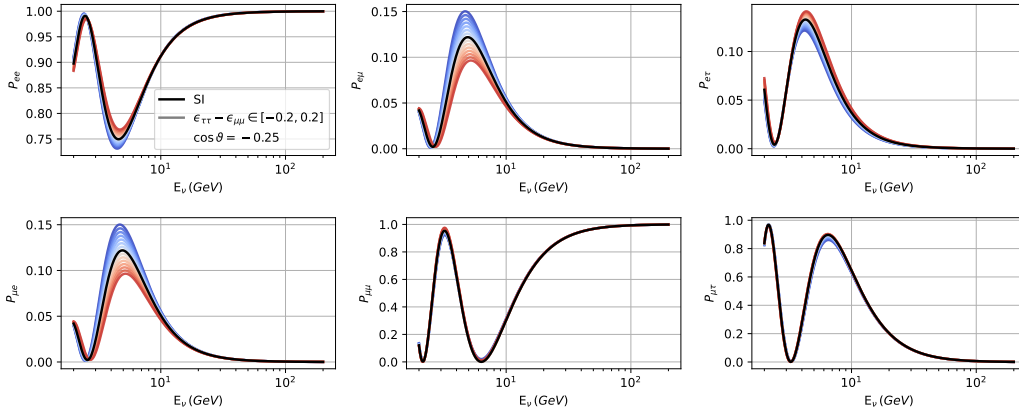


Fig. B.4.: Same as figure 3.5, but for $\cos(\theta) = -0.25$.

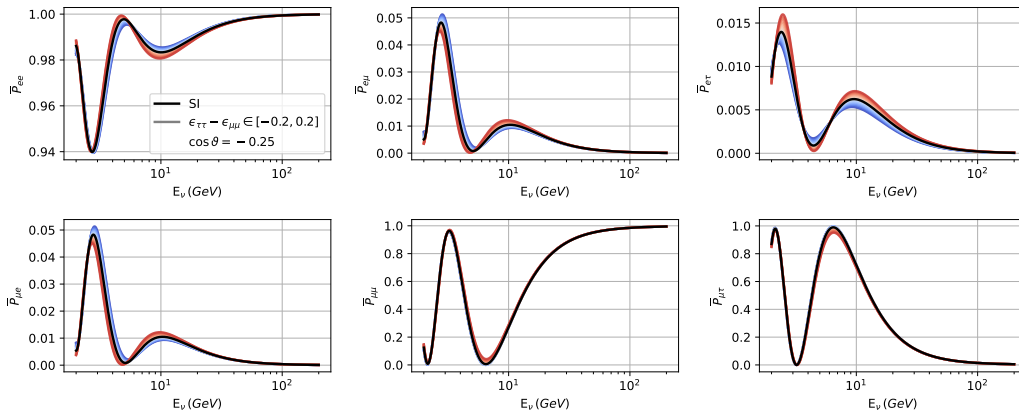


Fig. B.5.: Same as figure 3.5, but for antineutrinos and $\cos(\theta) = -0.25$.

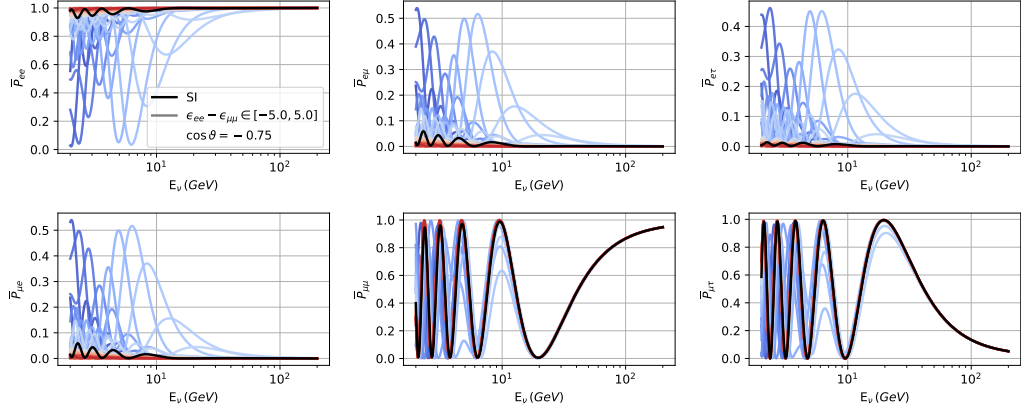


Fig. B.6.: Same as figure 3.5, but for $\epsilon_{ee}^{\oplus} - \epsilon_{\mu\mu}^{\oplus}$ and antineutrinos.

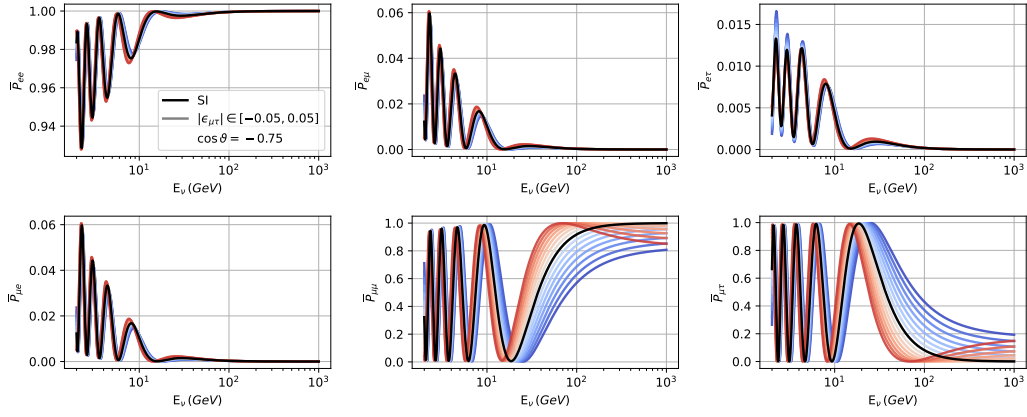


Fig. B.7.: Same as figure 3.5, but for $\epsilon_{\mu\tau}^{\oplus}$ with $\delta_{\mu\tau} = 0$, for antineutrinos.

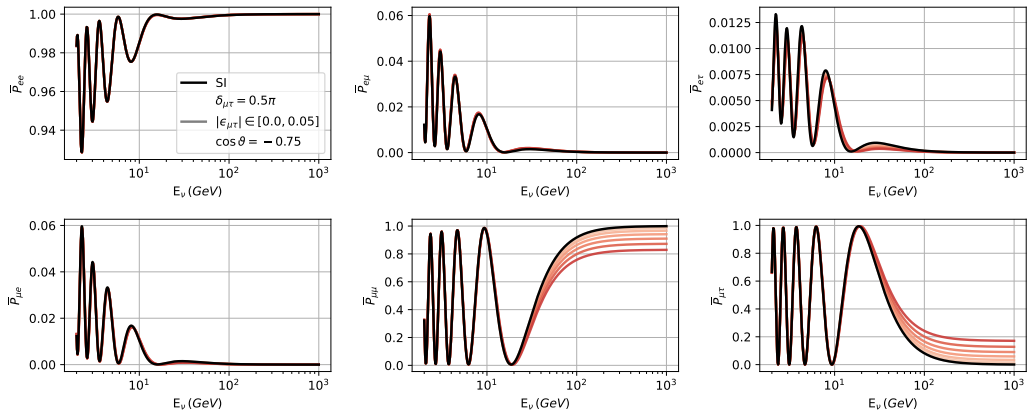


Fig. B.8.: Same as figure 3.5, but for $\epsilon_{\mu\tau}^{\oplus}$ with $\delta_{\mu\tau} = \pi/2$, for antineutrinos.

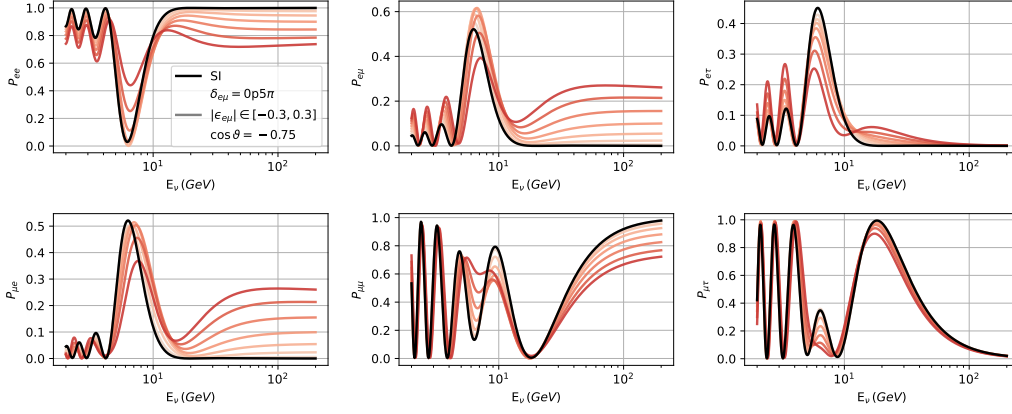


Fig. B.9.: Same as figure 3.5, but for $\epsilon_{e\mu}^{\oplus}$ with $\delta_{e\mu} = \pi/2$.

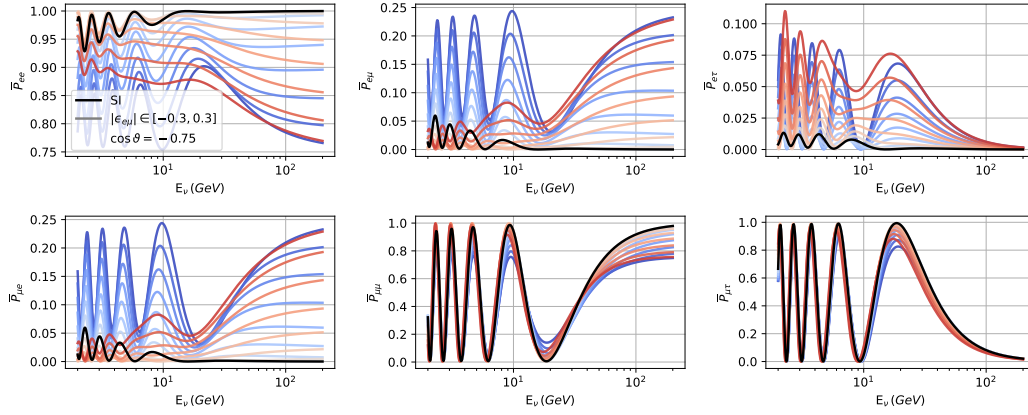


Fig. B.10.: Same as figure 3.5, but for $\epsilon_{e\mu}^{\oplus}$ with $\delta_{e\mu} = 0$, for antineutrinos.

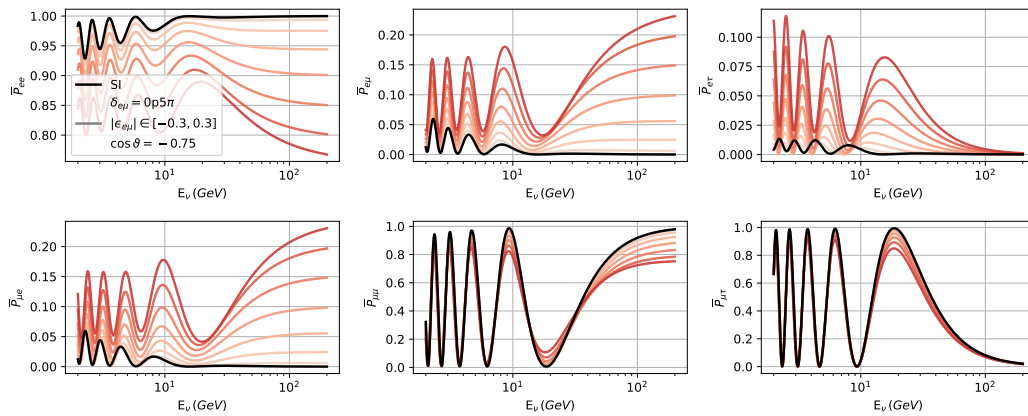


Fig. B.11.: Same as figure 3.5, but for $\epsilon_{e\mu}^{\oplus}$ with $\delta_{e\mu} = \pi/2$, for antineutrinos.

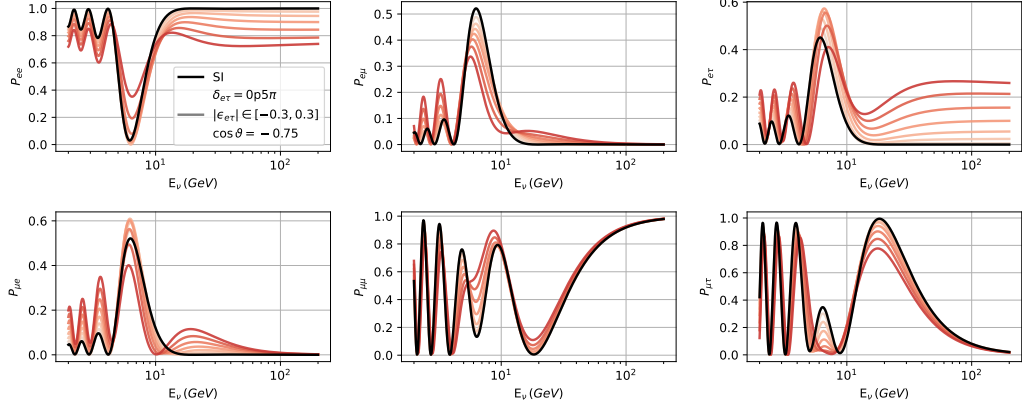


Fig. B.12.: Same as figure 3.5, but for $\epsilon_{e\tau}^{\oplus}$ with $\delta_{e\tau} = \pi/2$.

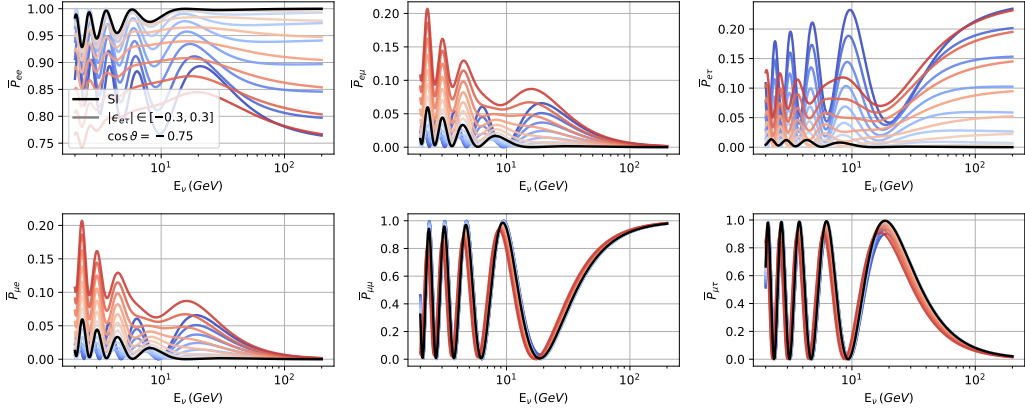


Fig. B.13.: Same as figure 3.5, but for $\epsilon_{e\tau}^{\oplus}$ with $\delta_{e\tau} = 0$, for antineutrinos.

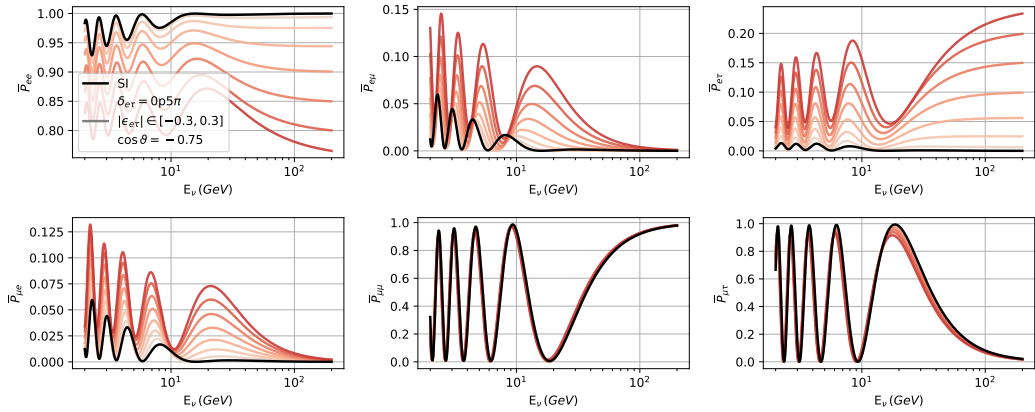


Fig. B.14.: Same as figure 3.5, but for $\epsilon_{e\tau}^{\oplus}$ with $\delta_{e\tau} = \pi/2$, for antineutrinos.

B.2 Reconstruction studies

Supplementary plots concerning the comparison between different reconstruction algorithms as well as the study on potential improvement in likelihood-based reconstructions performed in section 6.

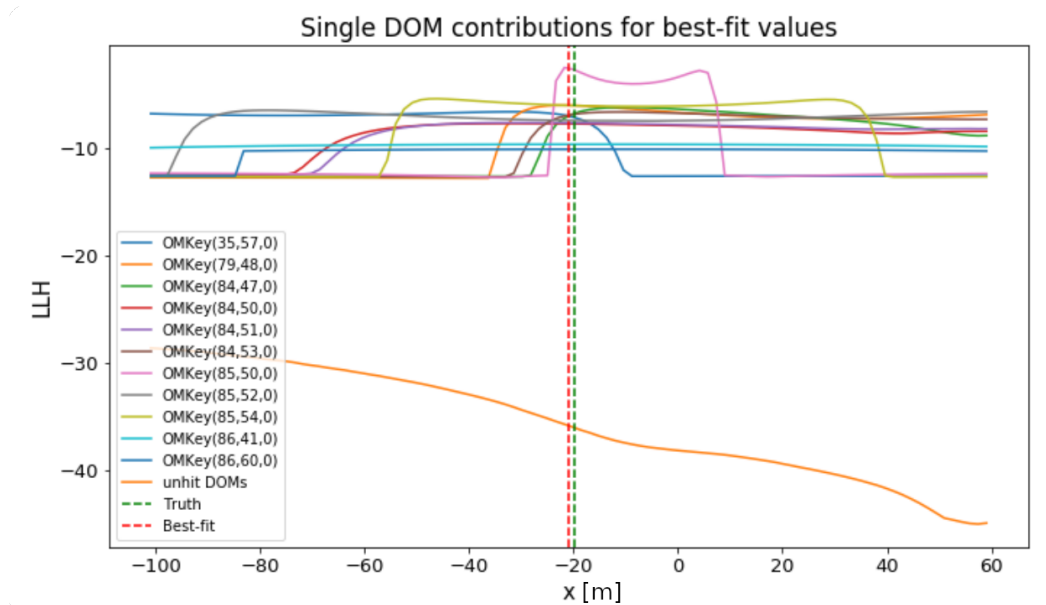
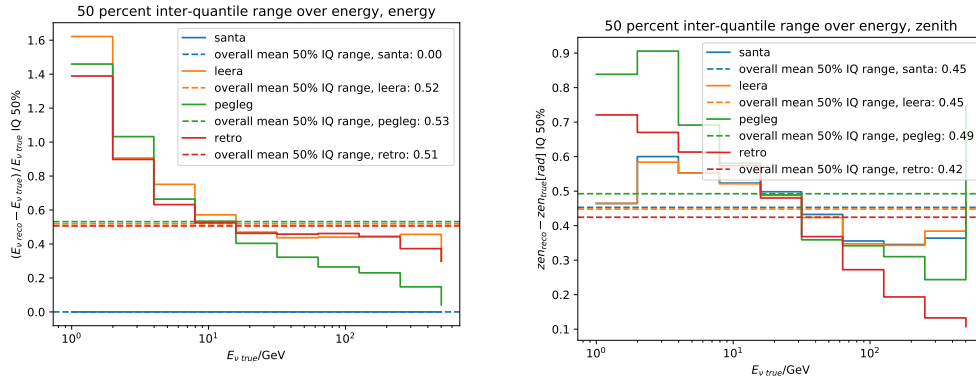


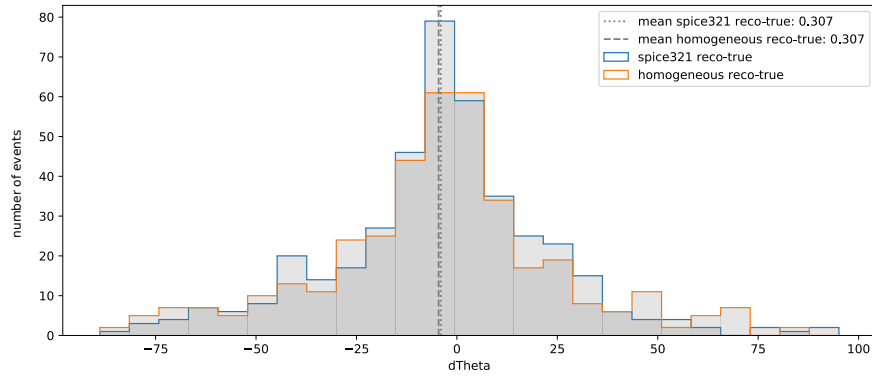
Fig. B.15.: Contributions to the overall LLH value by the most relevant DOMs, in a scan through the x parameter. DOMs identities are given through their key, which consists of the number of the string, module and PMT within the module. The same event as in figure 6.9b is shown.



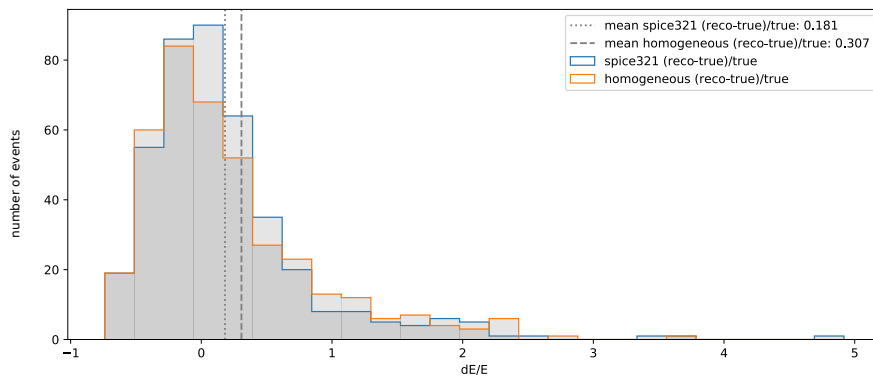
(a) Energy reconstruction comparison between RETRO, pegleg and LEERA. The SANTA reconstruction does not return reconstruct energies.

(b) Comparison between cosine of the zenith angle reconstruction in RETRO, pegleg and LEERA.

Fig. B.16.: Comparison between four reconstructions concerning their performance in the most relevant observables.



(a) Zenith angle reconstruction performance.



(b) Energy reconstruction performance.

Fig. B.17.: Comparison between reconstruction performances for different ice models.

B.3 NSI signal in DeepCore

The signal of NSI in DeepCore is similar to that of oscillation effects. It depends on which NSI parameter is varied and what value it is set to. Multiple examples of signal maps are shown in the following for true event properties. The binning shown here is significantly finer than the analysis binning, which is optimized for reconstructed properties. All figures show the difference in oscillation probabilities $\nu_\alpha \rightarrow \nu_\mu$, induced by the respective NSI, with respect to the standard oscillations case.

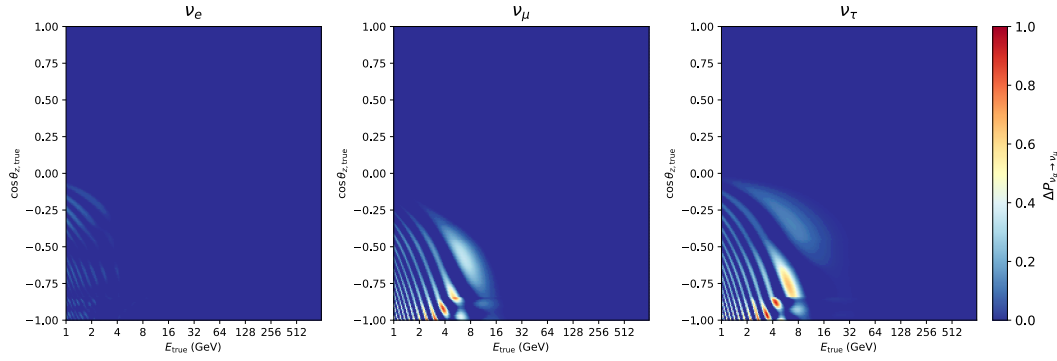


Fig. B.18.: Oscillation probability difference for $\epsilon_{ee}^\oplus - \epsilon_{\mu\mu}^\oplus = -2.2$ as the only source of NSI.

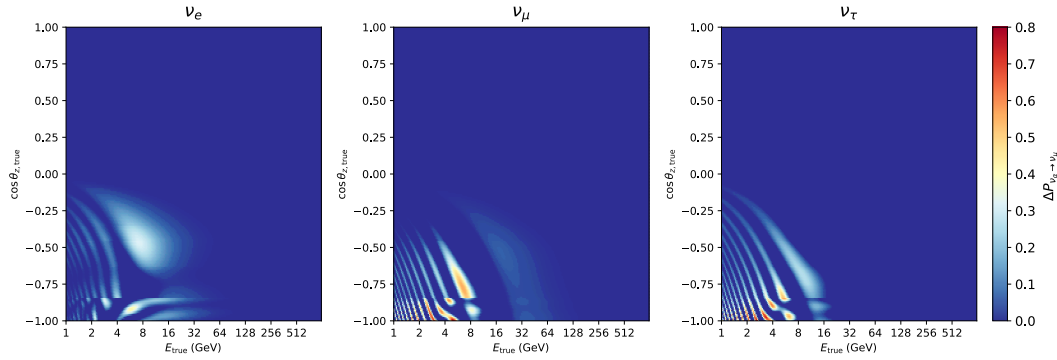


Fig. B.19.: Oscillation probability difference for $\epsilon_{e\tau}^\oplus = 0.2$ as the only source of NSI.

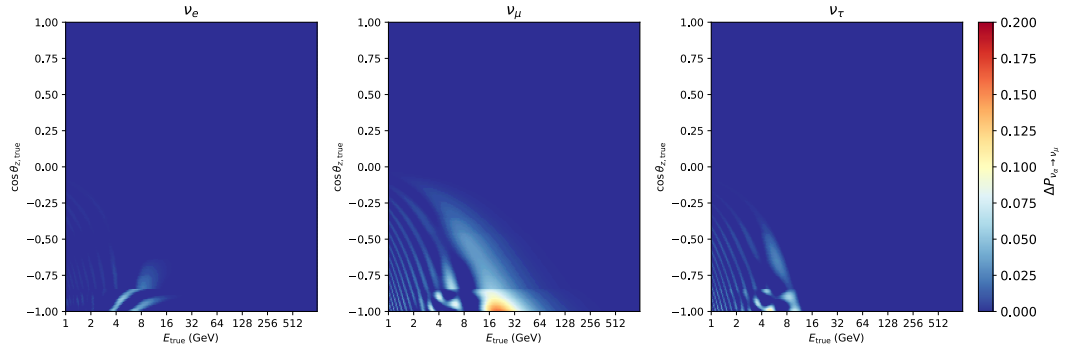


Fig. B.20.: Oscillation probability difference for $\epsilon_{\tau\tau}^{\oplus} - \epsilon_{\mu\mu}^{\oplus} = 0.04$ as the only source of NSI.

B.4 Simulated sets with off-nominal nuisance parameter values

In addition to the baseline Monte Carlo set, which is simulated at a hypothesis where all nuisance parameters are at their nominal values, off-nominal Monte Carlo sets are produced for detector related nuisance parameters. The set numbers and varied nuisance parameter values are shown in the following, subdivided based on which parameters are varied.

Tab. B.1.: List of set numbers and the corresponding DOM efficiency for sets where this is the only varied nuisance parameter.

Set number	DOM efficiency
0000 (baseline)	1.00
0001	0.90
0002	0.95
0003	1.05
0004	1.10
0005	0.50
0006	1.20
0007	0.70
0008	0.80

Tab. B.2.: List of set numbers and the corresponding settings of DOM efficiency and the hole ice parameters p0 and p1. The 0150, 0151 and 0152 sets are at extreme hypotheses and probe allowed but unlikely regions of the parameter space.

Set number	DOM efficiency	p0	p1
0000 (baseline)	1.00	+0.101569	-0.049344
0100	1.00	-0.0648	-0.1088
0101	1.00	-0.4839	-0.0171
0102	1.00	+0.2803	-0.0754
0103	1.00	+0.1122	+0.0035
0104	0.88	-0.0498	-0.0543
0105	0.93	-0.3729	+0.0349
0106	0.97	+0.2965	-0.0363
0107	1.03	+0.1244	-0.1132
0108	1.07	-0.0355	-0.0185
0109	1.12	-0.3142	-0.0773
0150	1.00	-1.3	0.15
0151	1.00	-1	-0.1
0152	1.00	0.5	0.15
0300	1.00	-1	-0.049344
0301	1.00	-0.5	-0.049344
0302	1.00	+0.1	-0.049344
0303	1.00	+0.3	-0.049344
0305	1.00	+0.101569	-0.15
0306	1.00	+0.101569	-0.1
0307	1.00	+0.101569	+0
0308	1.00	+0.101569	+0.05
0309	1.00	+0.101569	+0.1
0310	1.00	+0.101569	+0.15
0311	1.00	-0.2	-0.049344

Tab. B.3.: List of set numbers and the corresponding ice scattering and absorption values, which correspond to bulk ice model used for production. For better visibility, only the percentages by which the parameters are changed are given in this case.

Set number	Scattering	Absorption	Ice model
0500	+5%	+5%	spice_3.2.1/err_s+.05_a+.05
0501	+5%	-5%	spice_3.2.1/err_s+.05_a-.05
0502	-5%	+5%	spice_3.2.1/err_s-.05_a+.05
0503	-5%	-5%	spice_3.2.1/err_s-.05_a-.05
0504	-	+10%	spice_3.2.1/err_s+.00_a+.10
0505	-	-10%	spice_3.2.1/err_s+.00_a-.10
0506	+10%	-	spice_3.2.1/err_s+.10_a+.00
0507	-10%	-	spice_3.2.1/err_s-.10_a+.00
0508	-	+30%	spice_3.2.1/err_s+.00_a+.30
0509	-	-30%	spice_3.2.1/err_s+.00_a-.30
0510	+30%	-	spice_3.2.1/err_s+.30_a+.00
0511	-30%	-	spice_3.2.1/err_s-.30_a+.00
0512	-	+20%	spice_3.2.1/err_s+.00_a+.20
0513	-	-20%	spice_3.2.1/err_s+.00_a-.20
0514	+20%	-	spice_3.2.1/err_s+.20_a+.00
0515	-20%	-	spice_3.2.1/err_s-.20_a+.00
0516	-	+5%	spice_3.2.1/err_s+.00_a+.05
0517	-	-5%	spice_3.2.1/err_s+.00_a-.05
0518	+5%	-	spice_3.2.1/err_s+.05_a+.00
0519	-5%	-	spice_3.2.1/err_s-.05_a+.00

Tab. B.4.: List of set numbers that use the BFRv2 ice model, along with the applied DOM efficiency, hole ice parameters and ice absorption and scattering values. For better visibility, only the changes to the parameter values are given.

Set number	Ice model	ϵ_{DOM}	p0	p1	Absorption	Scattering
1122 (BFR baseline)	spice_bfr-v2	-	-	-	-	-
1124	spice_bfr-v2	+10%	-	-	-	-
1125	spice_bfr-v2	-10%	-	-	-	-
1126	spice_bfr-v2	-	-0.2	-	-	-
1127	spice_bfr-v2	-	+0.3	-	-	-
1128	spice_bfr-v2	-	-	-0.1	-	-
1129	spice_bfr-v2	-	-	0	-	-
1130	spice_bfr-v2	-	-	-	+5%	-
1131	spice_bfr-v2	-	-	-	-5%	-
1132	spice_bfr-v2	-	-	-	-	+5%
1133	spice_bfr-v2	-	-	-	-	-5%

B.5 Detector nuisance parameter treatment

Multiple sets have been tested in regard to whether the simulated bin counts match what the gradient-based interpolation generates for the same hypothesis. In the following, summary plots of these tests are shown. The detailed maps are only given for sets where they are assumed to be relevant, since the shown sets include off-nominal p_0 values that result in clustered pulls within these maps.

In the summary plots, for each set the mean value is given as yellow vertical line, with the 68% and 90% intervals indicated as a box and line, respectively. Outliers are shown as circles. The total of all sets included in each figure is shown at the top.

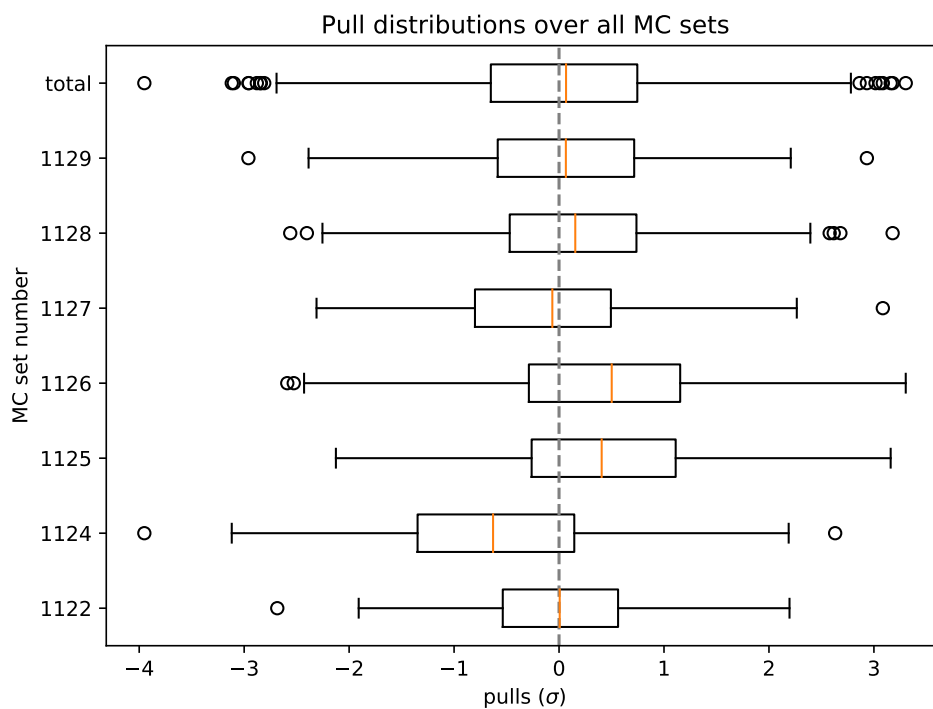


Fig. B.21.: Summary plot for all tested sets that include BFRv2. Note that only the 1122 set was included in calculating the gradients.

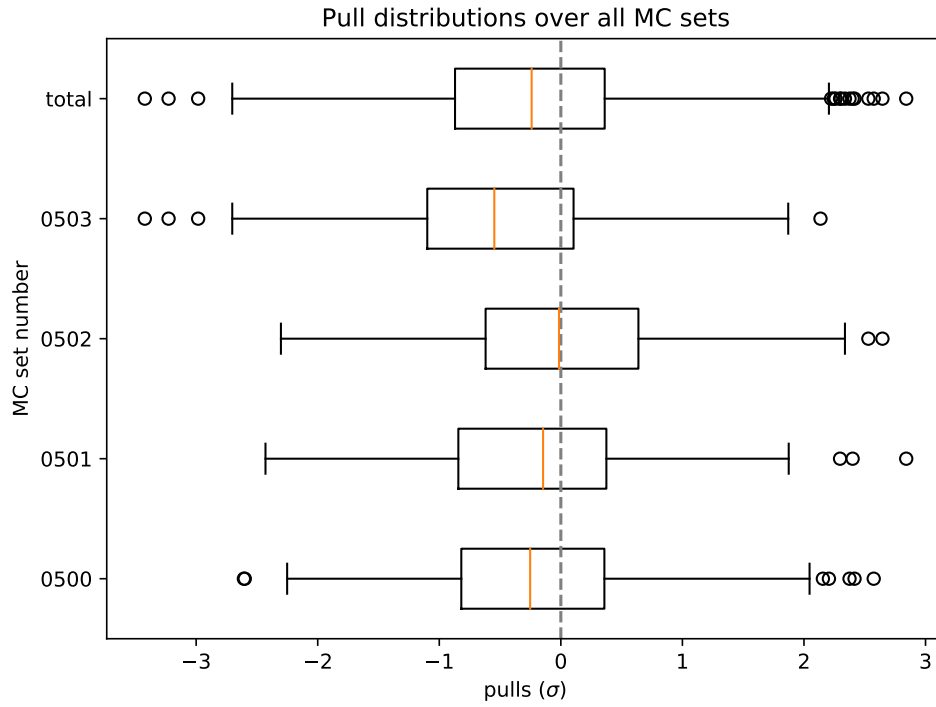


Fig. B.22.: Summary plot for all sets where the ice scattering and absorption parameters are varied simultaneously.

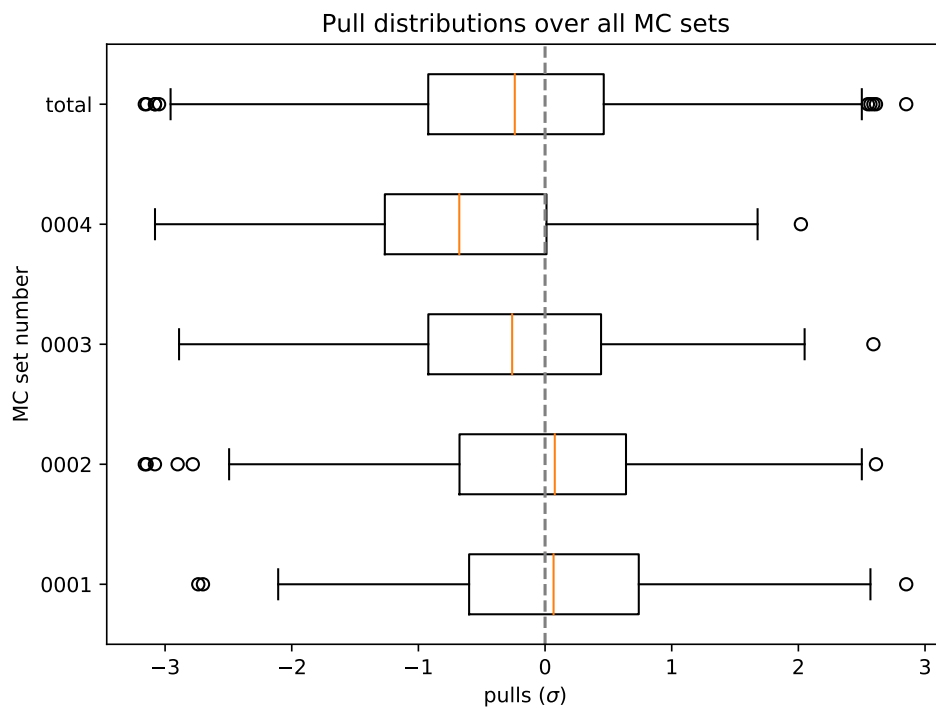


Fig. B.23.: Summary plot for all sets where only DOM efficiency is varied.

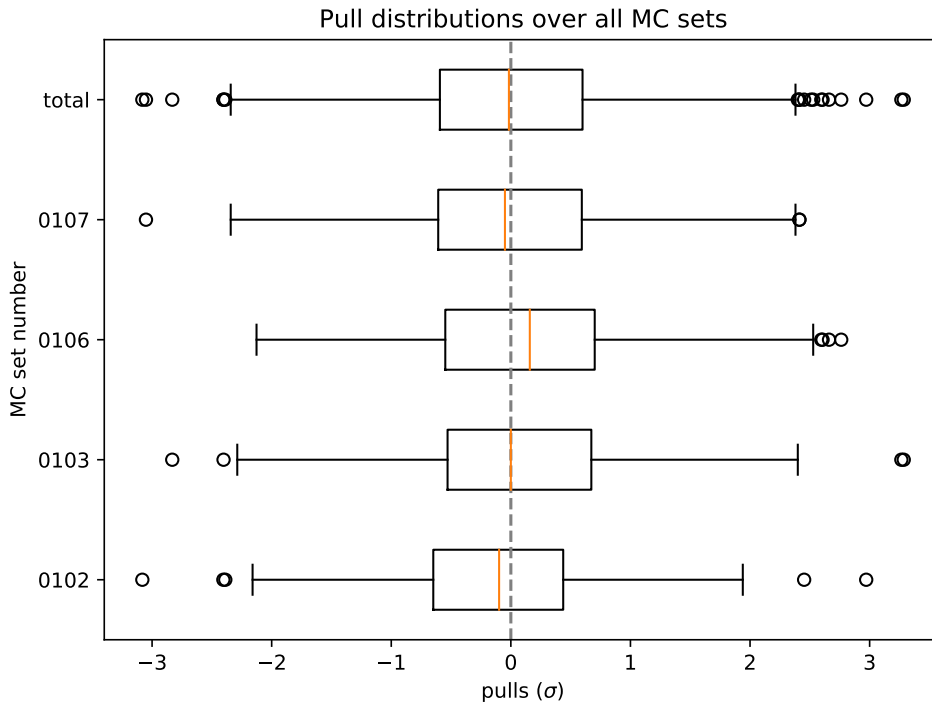


Fig. B.24.: Summary plot for all sets where DOM efficiency and hole ice parameters are varied simultaneously, except for sets where the value of p_0 is below its nominal value.

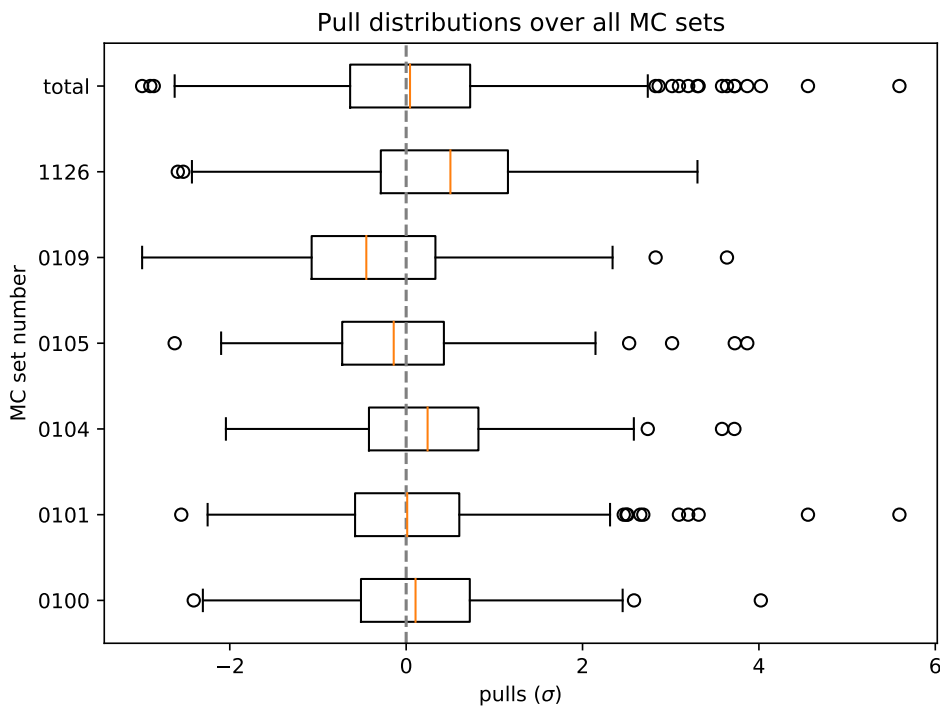


Fig. B.25.: Summary plot for all sets where the value of p_0 is below its nominal value.

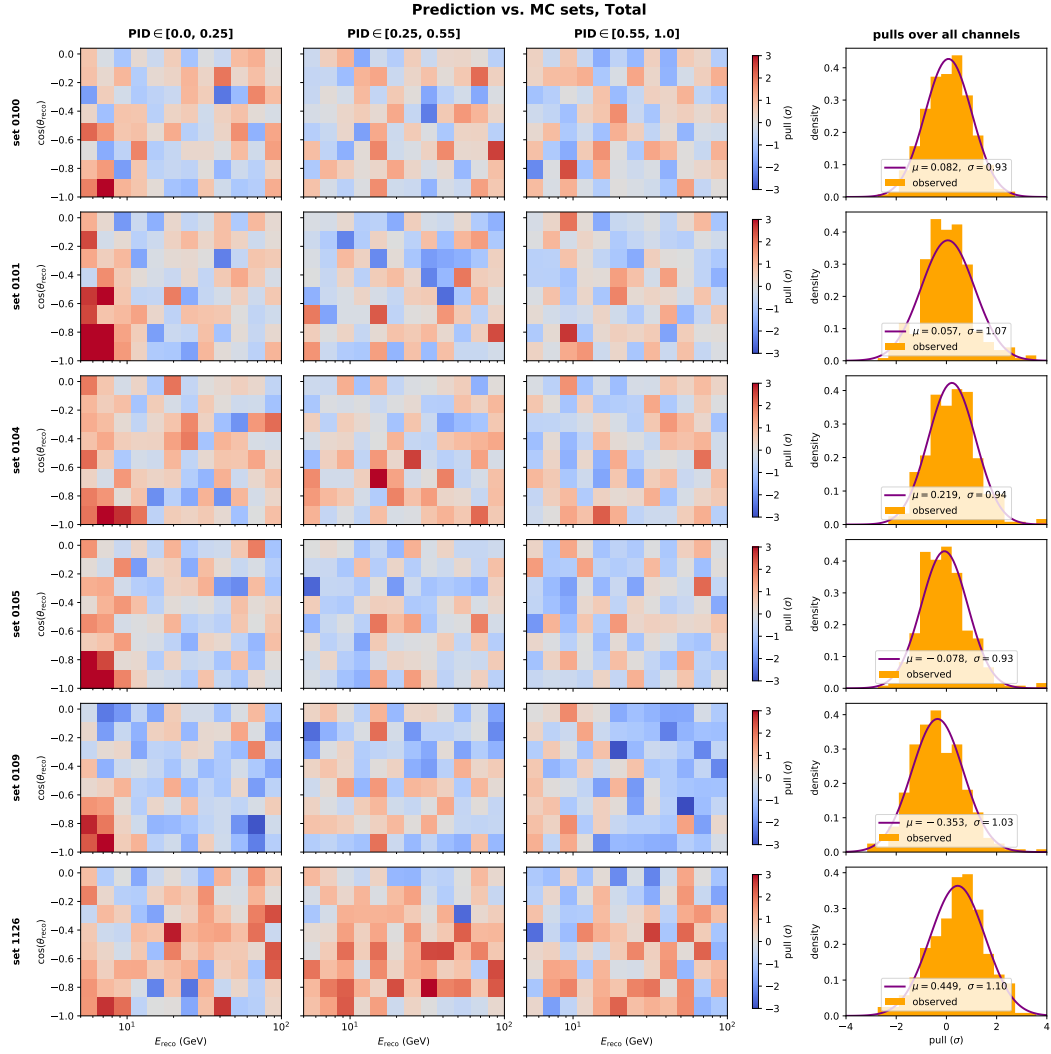


Fig. B.26.: Detailed maps of all sets where the value of p_0 is below its nominal value, showcasing the clustering of positive pulls in the lowest bins in all dimensions.

Tab. B.5.: Categories and numbers of MC simulated sets included in the generation of gradients used in the detector nuisance parameter treatment for this analysis. Detailed information on the parameter settings for all individual sets can be found in the appendix in section B.4.

Varied parameters	Set numbers
baseline SPice 3.2.1	0000
baseline BFRv2	1122
DOM efficiency	0001 0002 0003 0004
DOM efficiency, p_0 , p_1	0100 0101 0102 0103 0104 0105 0106 0107 0109
ice scattering, absorption	0500 0501 0502 0503 0504 0505 0506 0507

B.6 Nuisance parameter impact

In the following, results of the nuisance parameter impact test described in section 7.5 are given for NSI parameters not shown therein. In addition, the tested parameter values that do not conform to the general rule given there are listed.

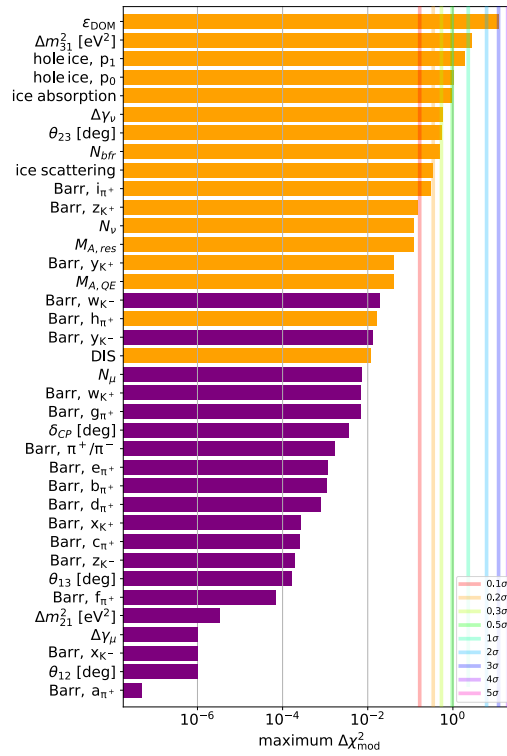


Fig. B.27.: Test result for the $\epsilon_{e\mu}^\oplus$ parameter.

Tab. B.6.: Tested parameter offsets for cases that deviate from the general rule described in the text. These are given as differences with respect to the nominal value.

Parameter	Value
$N_{\nu\tau}$	0.3
θ_{12} [°]	0.77
θ_{13} [°]	0.13
θ_{23} [°]	5.0
Δm_{21}^2 [eV ²]	$0.21 \cdot 10^{-5}$
Δm_{31}^2 [eV ²]	0.0001
δ_{CP} [°]	90
N_ν	0.4
p0	0.3
p1	0.05
N_{bfr}	0.5

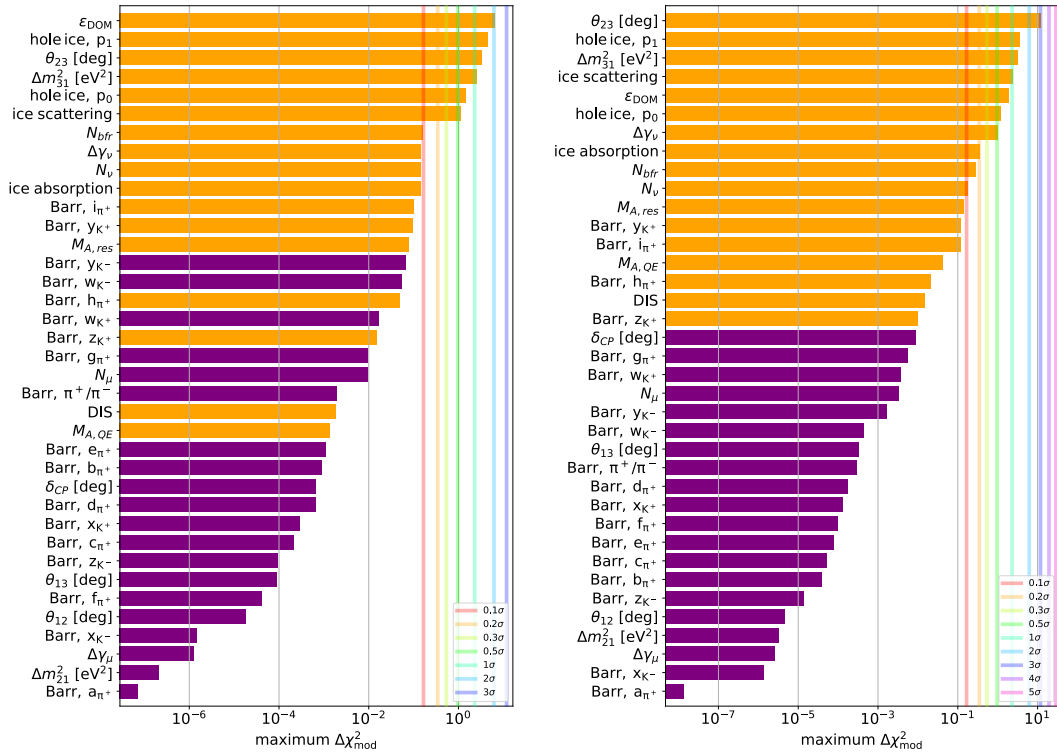
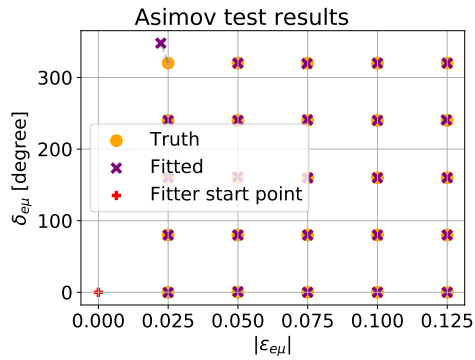


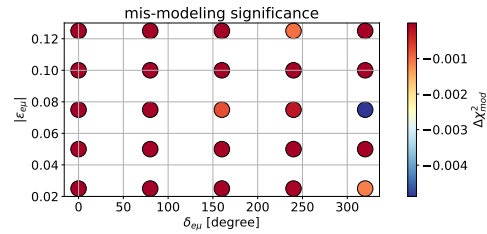
Fig. B.28.: Nuisance parameter impact for different NSI parameters.

B.7 Inject recovery test results

Results of fit performance tests for parameters that are not shown in section 7.6.2.

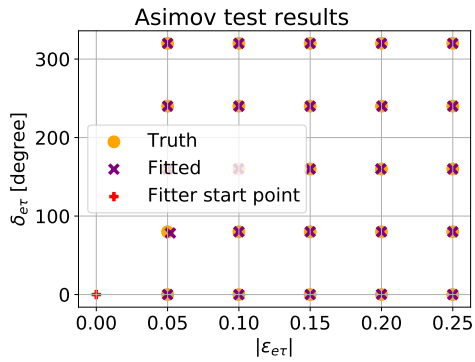


(a) True and fitted parameter values.

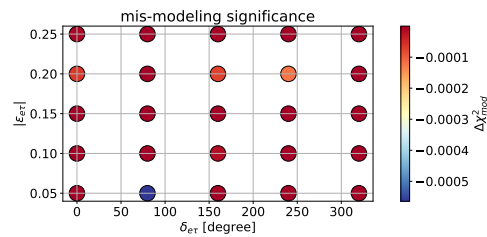


(b) Mis-modeling values.

Fig. B.29.: Fit performance test results for $\epsilon_{e\mu}^{\oplus}$.



(a) True and fitted parameter values.



(b) Mis-modeling values.

Fig. B.30.: Fit performance test results for $\epsilon_{e\tau}^{\oplus}$.

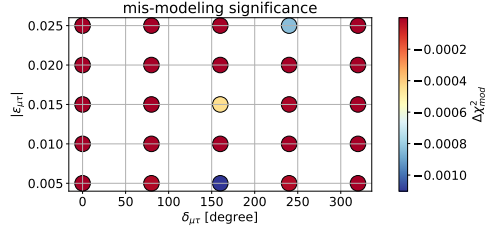


Fig. B.31.: Mis-modeling values for $\epsilon_{\mu\tau}^{\oplus}$ fitting performance tests.

B.8 Fluctuating nuisance parameter values

Results of fitting fluctuated nuisance and physics parameters that are now shown in section 7.6.3.

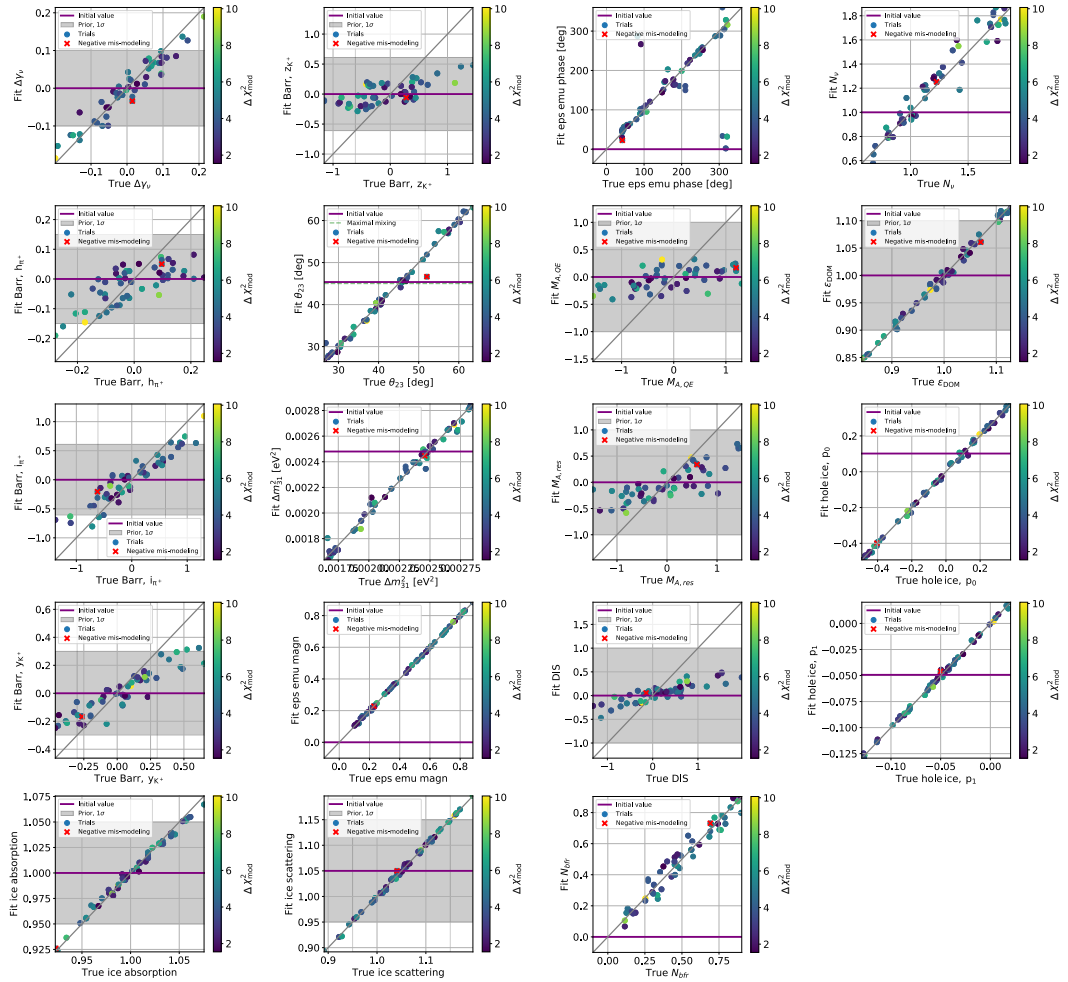


Fig. B.32.: Fluctuated and fitted parameter values including $\epsilon_{e\mu}$.

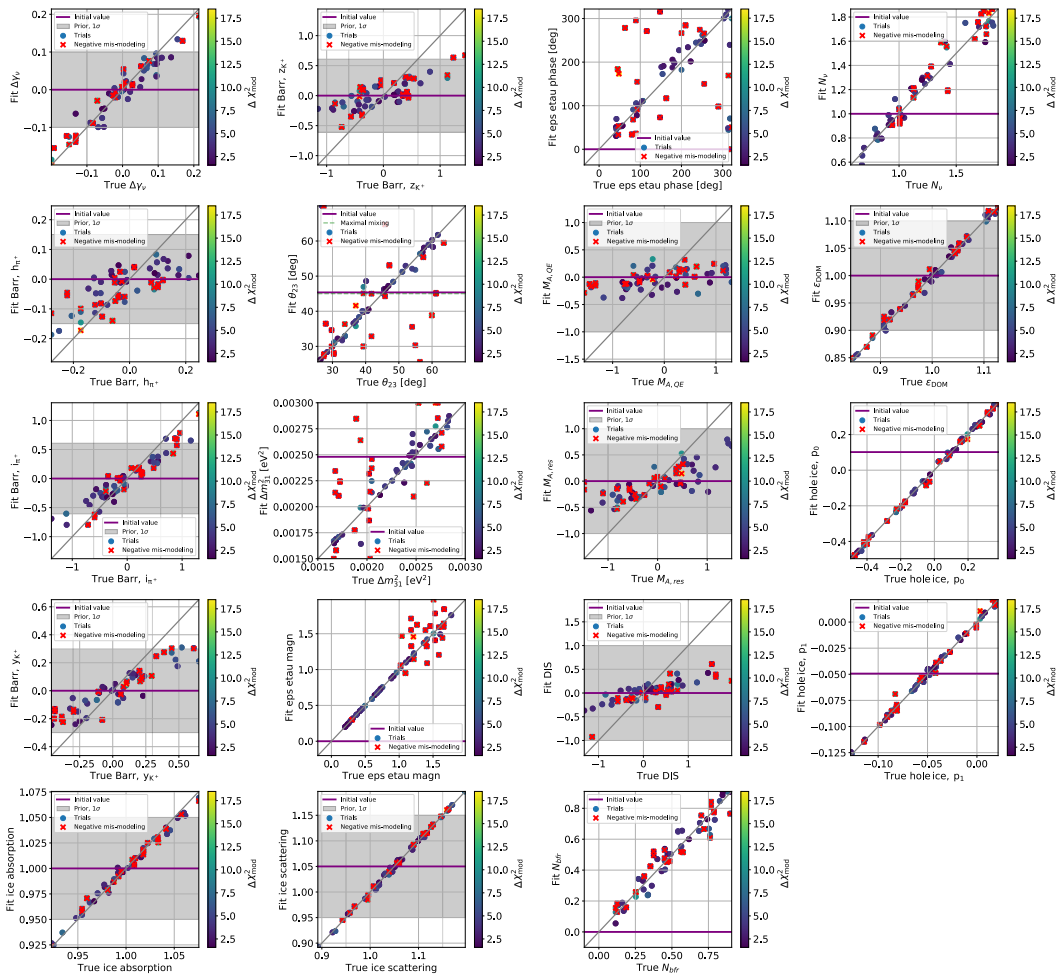


Fig. B.33.: Fluctuated and fitted parameter values including ϵ_{τ}^{\oplus} .

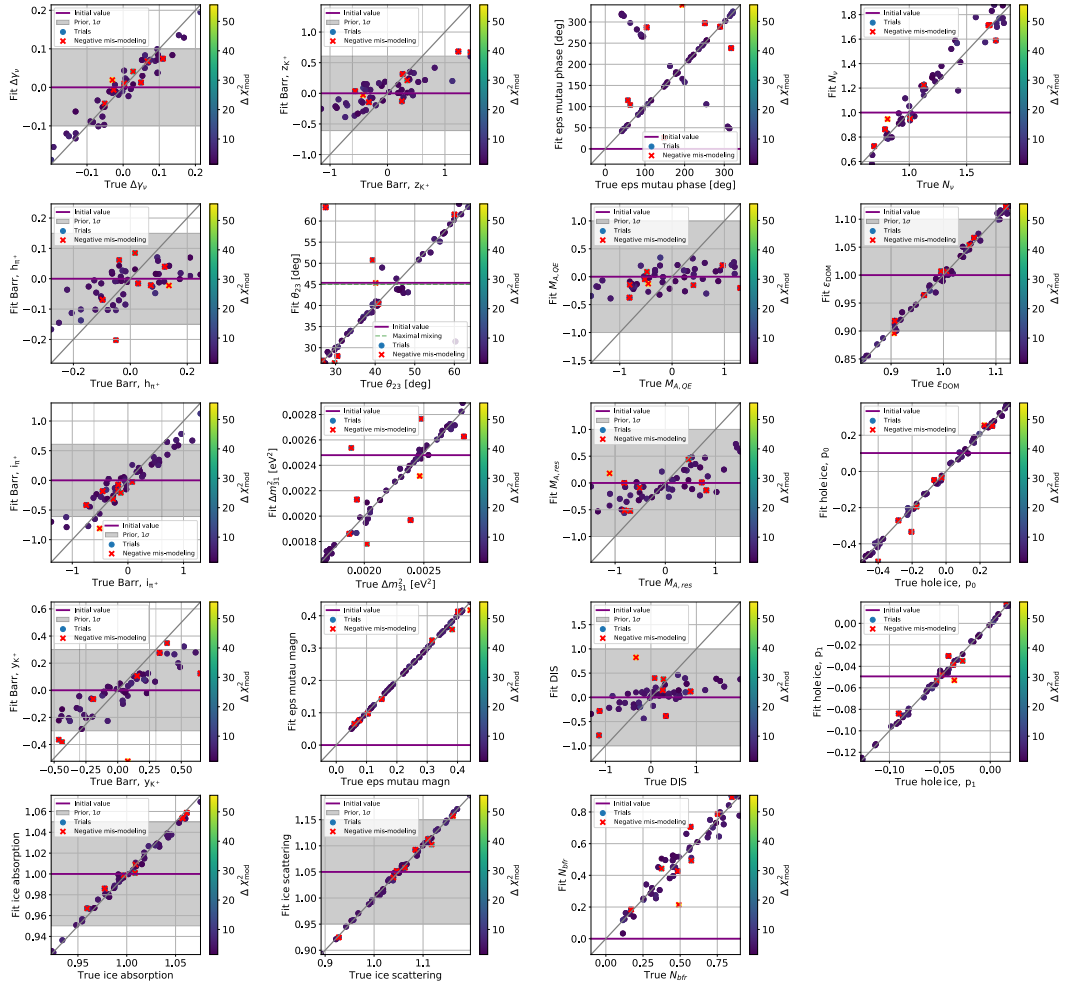
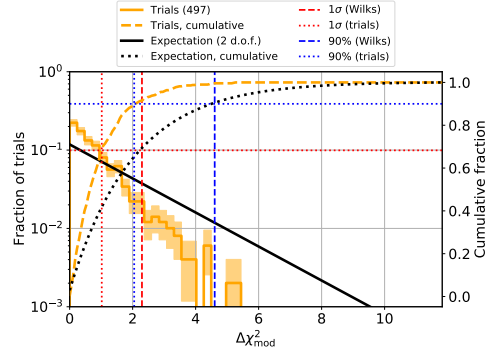


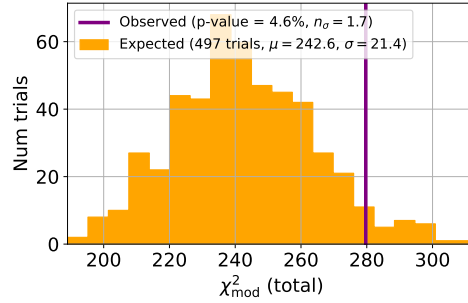
Fig. B.34.: Fluctuated and fitted parameter values including $\epsilon_{\mu\tau}^{\oplus}$.

B.9 Pseudo-data ensemble

Results of fitting Poisson-fluctuated pseudo-experiments for NSI parameters that are now shown in section 7.6.3.

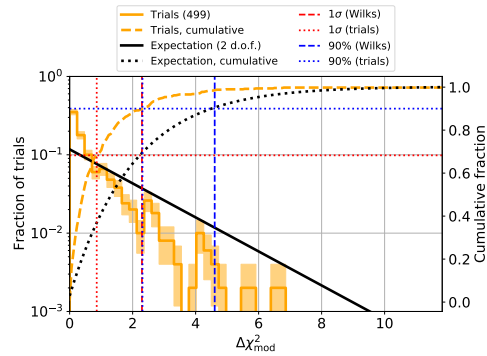


(a) Test statistic delta coverage with respect to the expectation based on Wilks' theorem.

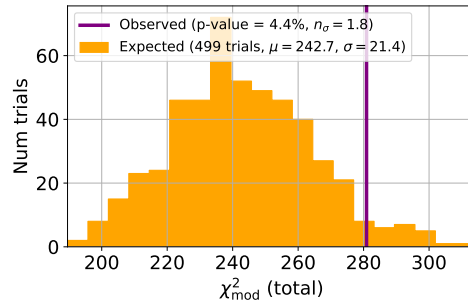


(b) Test statistic values from fits to pseudo-experiments and data.

Fig. B.35.: Test statistic behavior for fits to pseudo-experiments including $\epsilon_{e\tau}^\oplus$.

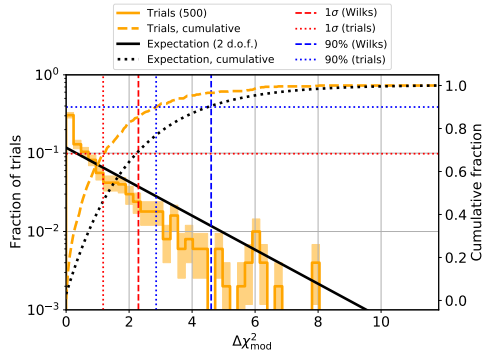


(a) Test statistic delta coverage with respect to the expectation based on Wilks' theorem.

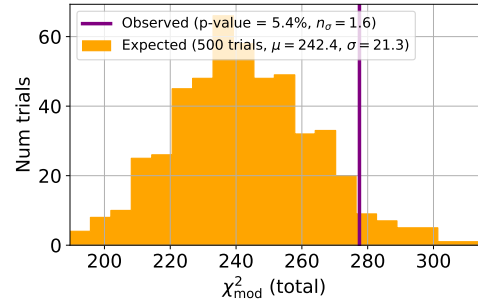


(b) Test statistic values from fits to pseudo-experiments and data.

Fig. B.36.: Test statistic behavior for fits to pseudo-experiments including $\epsilon_{e\mu}^\oplus$.



(a) Test statistic delta coverage with respect to the expectation based on Wilks' theorem.



(b) Test statistic values from fits to pseudo-experiments and data.

Fig. B.37.: Test statistic behavior for fits to pseudo-experiments including $\epsilon_{\mu\tau}^\oplus$.

B.10 Sensitivities

Asimov-based sensitivities to NSI parameters not shown in section 7.8.

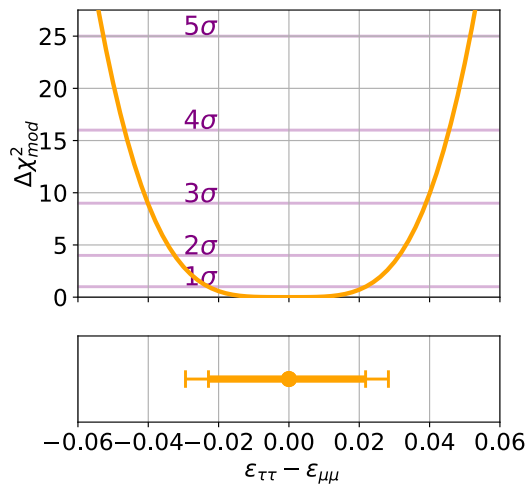


Fig. B.38.: Sensitivity for $\epsilon_{\tau\tau}^\oplus - \epsilon_{\mu\mu}^\oplus$ at IO.

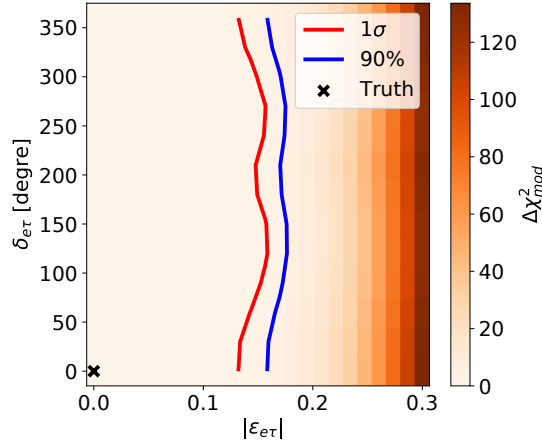
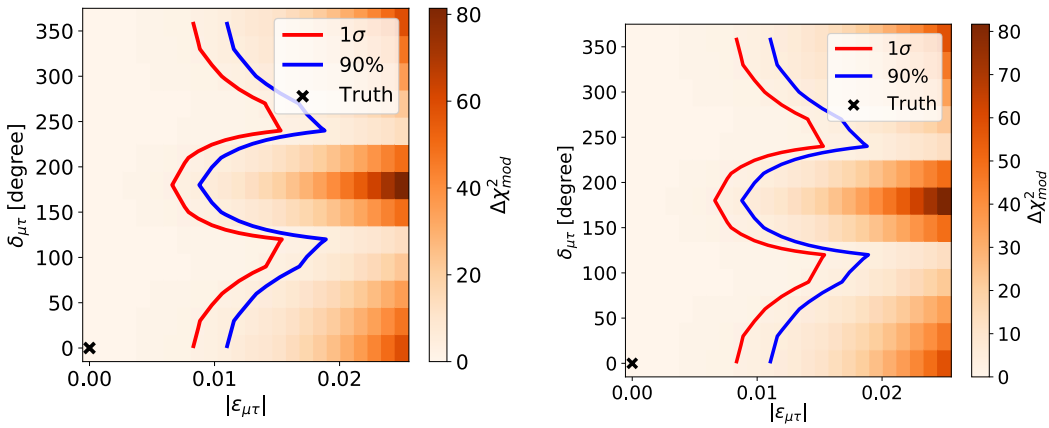


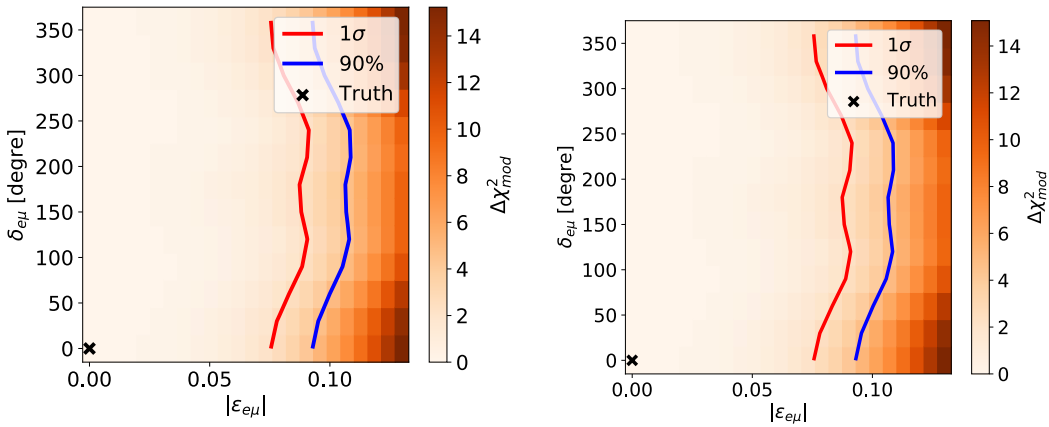
Fig. B.39.: Sensitivity for $\epsilon_{e\tau}^\oplus$ at IO.



(a) Sensitivity at NO.

(b) Sensitivity at IO.

Fig. B.40.: Sensitivities of $\epsilon_{\mu\tau}^\oplus$ at different mass orderings.



(a) Sensitivity at NO.

(b) Sensitivity at IO.

Fig. B.41.: Sensitivities of $\epsilon_{e\mu}^\oplus$ at different mass orderings.

B.11 Additional figures concerning MCMC sampling

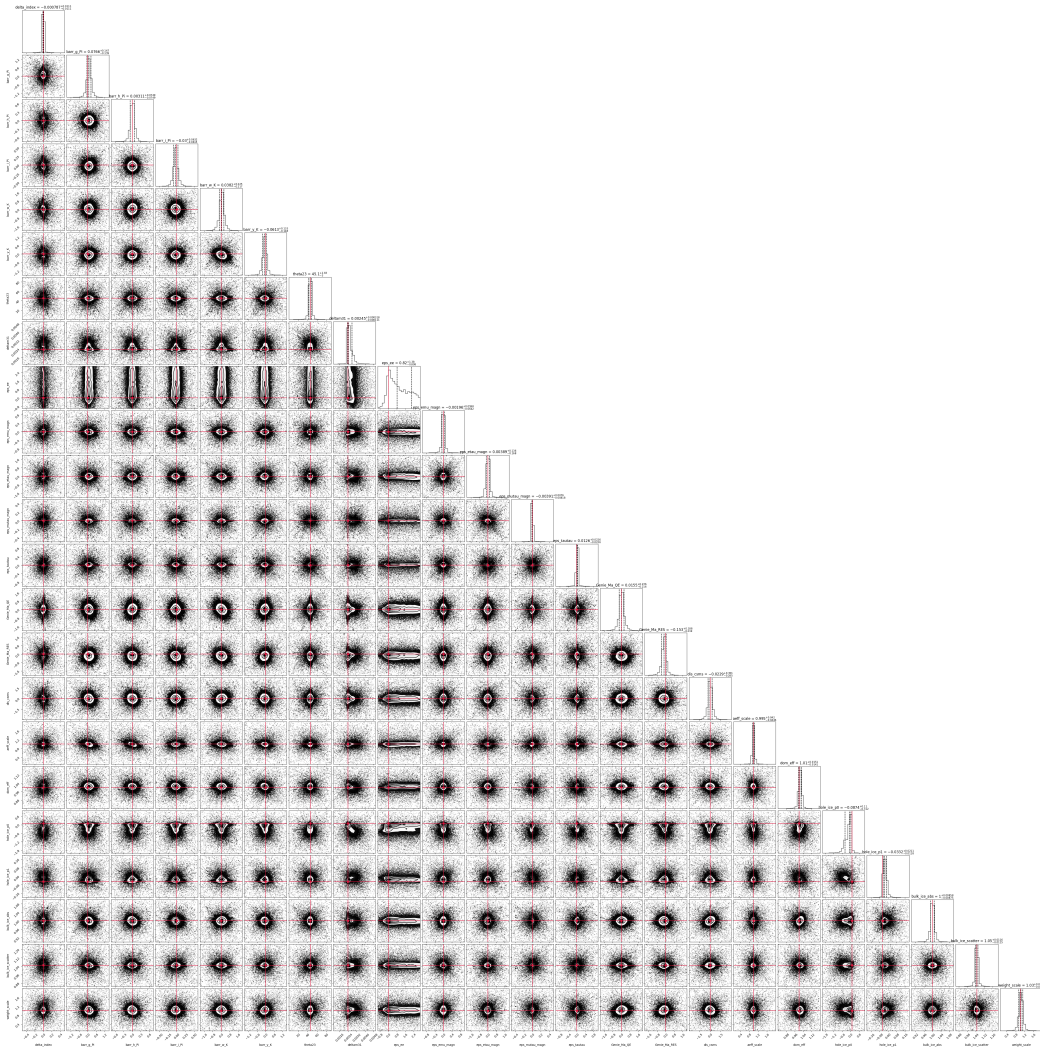


Fig. B.42.: Burn-in phase of the sampling described in section 7.7.4. The corresponding sampling phase is shown in figure B.43.

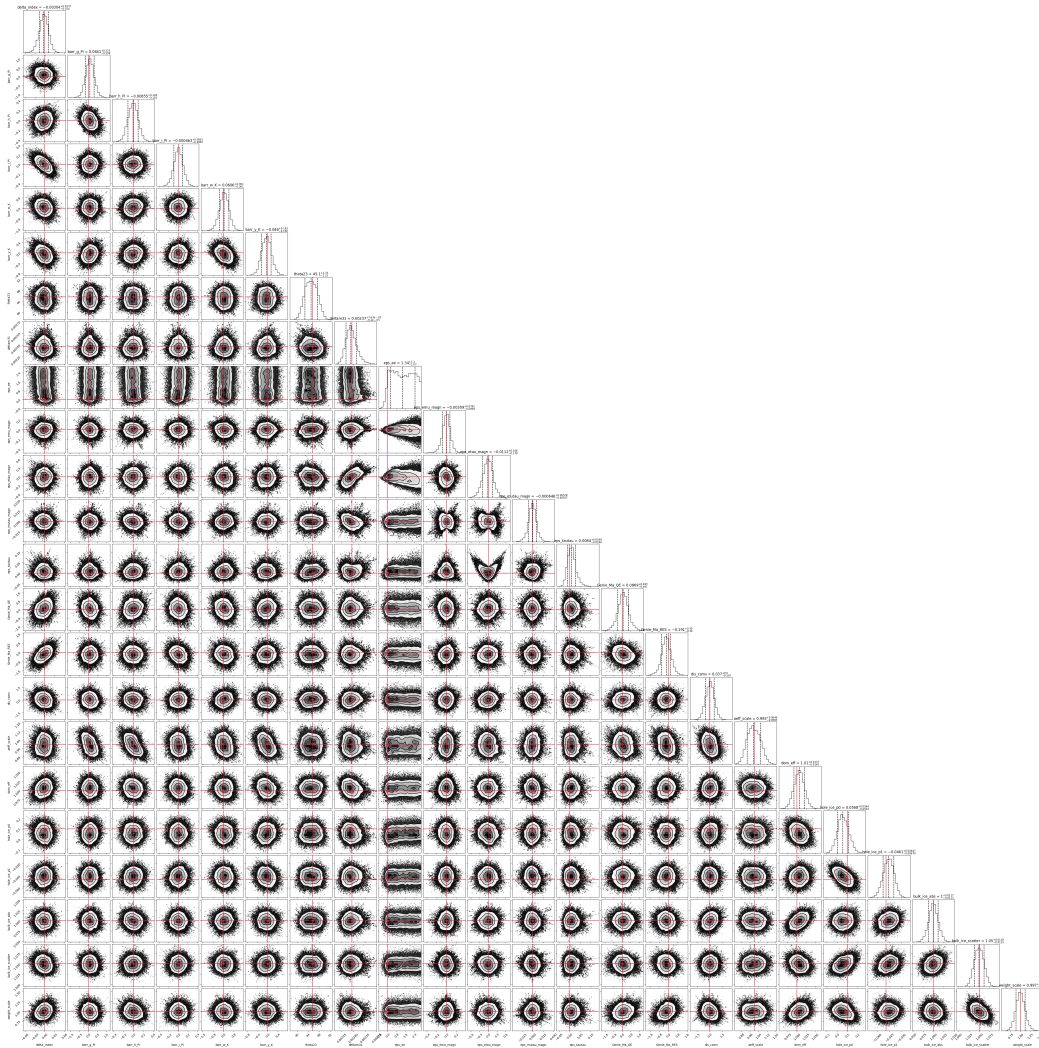


Fig. B.43.: Sampling phase of the sampling described in section 7.7.4.

B.12 Limits on pseudo-data

Pseudo-data based limits on NSI parameters that are not shown in section 7.9.

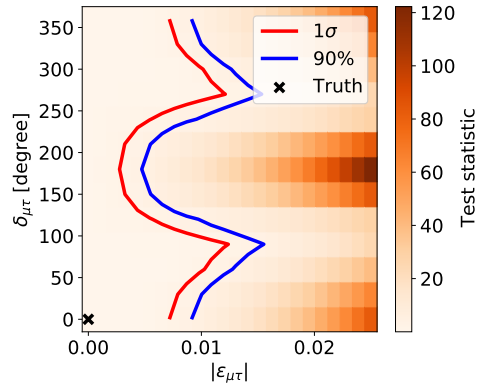
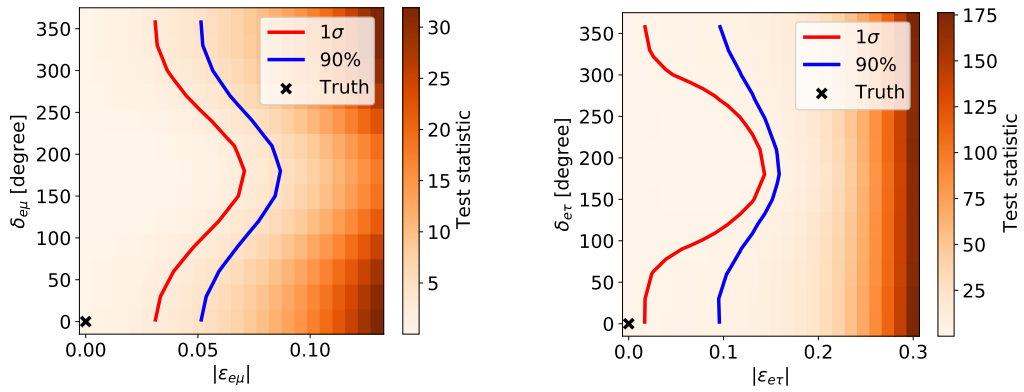


Fig. B.44.: Limits on $\epsilon_{\mu\tau}^{\oplus}$.



(a) Limits on $\epsilon_{e\mu}^{\oplus}$.

(b) Limits on $\epsilon_{e\tau}^{\oplus}$.

Fig. B.45.: Limits on different FC NSI parameters.

Colophon

This thesis was typeset with $\text{\LaTeX}2_{\epsilon}$. It uses the *Clean Thesis* style developed by Ricardo Langner. The design of the *Clean Thesis* style is inspired by user guide documents from Apple Inc.

Download the *Clean Thesis* style at <http://cleanthesis.der-ric.de/>.

