First Search for Heavy Neutral Leptons with IceCube DeepCore

Dissertation
zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)

im Fach: Physik
Spezialisierung: Experimentalphysik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Leander Fischer M. Sc.
geboren am 24. Oktober 1992
in Heidelberg

Präsidentin der Humboldt-Universität zu Berlin
Prof. Dr. Julia von Blumenthal

Dekanin der Mathematisch-Naturwissenschaftlichen Fakultät
Prof. Dr. Caren Tischendorf

Colophon
This document was typeset with the help of KOMA-Script and LaTeX using the open-source kaobook template class.

The source code of this thesis is available at:
https://github.com/LeanderFischer/phd_thesis

Todo list

Contents

Monte Carlo Event Generation and Detector Simulation

1

Like many analyses in IceCube, this work is based on MC simulations. The initial step for all particle (non-noise) simulation is the generation of events from selected initial distributions and fluxes. Events are the primary particle and all particles produced in the interaction with the ice. The particles are then propagated through the ice, producing Cherenkov photons, which are propagated further until they reach a DOM or are absorbed in the ice. If they hit a DOM the detector response is simulated. Splitting the simulation steps has the advantage of reusing the outputs of for example the generation step to propagate the particles with different ice model, in order to estimate the systematic impacts of uncertainties of the ice properties. A similar approach can be taken for varying detector response, before starting the event selection. Through this a more efficient (reduced) use of computing resources can be achieved.

The central part of this thesis is the HNL signal simulation itself. Since this is the first search for HNLs with IceCube DeepCore, there was no prior knowledge of the number of events expected per year nor of the performance in terms of reconstruction and classification accuracy. This chapter describes the first HNL event generation developed for IceCube DeepCore. Two avenues of generation were pursued in parallel. A collection of model-independent samples is explained in Section 1.1. They were used for performance benchmarking and for cross-checks to validate the physically accurate, model-dependent event generation, which is described in Section 1.2. The event generation for SM background events is briefly described in Section 1.3, followed by the detector response simulation in Section 1.4. The detector response is identically applied to both signal and background events.

1.1 Model-Independent Heavy Neutral Lepton Event Generation

To investigate the potential of IceCube to detect HNLs by identifying the unique double cascade morphology explained in Section **??**, a model-independent double cascade generator was developed, where the kinematics of each cascade can be controlled directly. Using this generator, several simulation samples were produced to investigate the performance of IceCube DeepCore to detect low-energy double cascades, dependent on their properties. All samples are produced using a collection of custom generator functions [1] that place two EM cascade vertices with variable energy and direction at configurable locations in the detector.

1.1.1 Simplistic Samples

To investigate the best-case and the worst-case double cascade event scenarios, two samples are produced in the DeepCore volume: straight up-going events ($\cos(\theta) = -1$) that are centered on a string and horizontal events ($\cos(\theta) = 0$).

add/move event views to here? or draw a schematic of these events? (ORANGE)

The first sample is used to investigate one of the most promising scenarios to detect a double cascade, where both cascade centers are located on a DeepCore string and the directions are directly up-going. One of the DeepCore strings was randomly chosen as the x-y coordinate for this sample. As already mentioned in Section **??**, DeepCore strings have higher quantum efficiency DOMs and a denser vertical spacing, making them better to detect low-energy events that produce little light. To produce the events, the x, y position of the cascades is fixed to the center of the string while the z positions are each sampled uniformly along the axis of the string. Note that this will therefore not produce a uniform length distribution between the cascades. The positions are defined in the IceCube coordinate system that was introduced in Section **??**. The energies are sampled uniformly between 0.0 GeV and 60.0 GeV, to generously cover the region where $\nu_\mu \longrightarrow \nu_\tau$ appearance is maximized. The time of the lower cascade is set to $t_0 = 0.0\,\text{ns}$ and for the upper one to $t_1 = L/c$, assuming the HNL travels at the speed of light, c. Figure 1.1 shows the resulting energy distributions and the decay length distribution, where it can be seen how the uniform cascade energies sum into a non-uniform total energy, and the decay length distribution is also non-uniform due to the uniform z sampling of both cascades, which sets the distance between them.

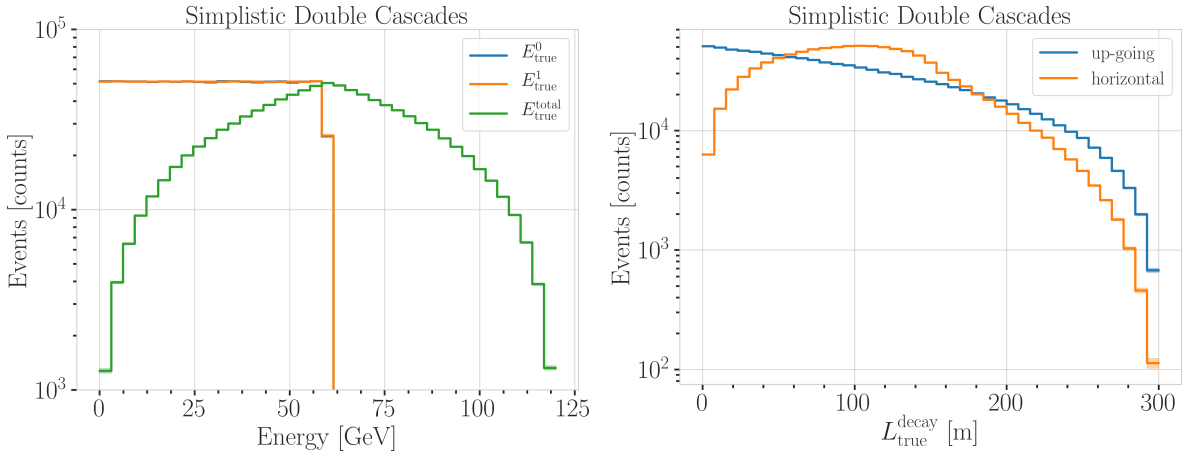


Figure 1.1: Generation level distributions of the simplistic simulation samples. Cascade and total energies (left) and decay lengths (right) of both samples are shown.

The second sample is used to investigate the reconstruction performance for horizontal events, where the spacing between DOMs is much larger. The cascades are placed uniformly on a circle with radius of $r = 150\,\text{m}$ centered in DeepCore at the depth of $z = -330\,\text{m}$. The direction is always horizontal and azimuth is defined by the connecting vector of both cascade positions. The energies are again sampled uniformly between 0.0 GeV and 60.0 GeV. The specific sampling distributions/values for the cascades are listed in Table 1.1, for both samples and for completeness, all distributions are shown in Figure A.2.

1.1.2 Realistic Sample

To thoroughly investigate the potential of IceCube DeepCore to detect double cascade events, a more realistic simulation sample is produced that aims to be as close as possible to the expected signal simulation explained in

Sample	Variable	Distribution	Range/Value
Up-going			
	energy	uniform	0.0 GeV to 60.0 GeV
	zenith	fixed	180.0°
	azimuth	fixed	0.0°
	x, y position	fixed	(41.6, 35.49) m
	z position	uniform	−480.0 m to −180.0 m
Horizontal			
	energy	uniform	0.0 GeV to 60.0 GeV
	zenith	fixed	90.0°
	azimuth	uniform	0.0° to 360.0°
	x, y position	uniform (circle)	c=(46.29, -34.88) m, r=150.0 m
	z position	fixed	−330.0 m

Table 1.1: Generation level sampling distributions and ranges/values of up-going and horizontal model-independent simulation.

Section 1.2, while still allowing additional freedom to control the double cascade kinematics. This sample is particularly useful for validating the model-dependent HNL simulation described in Section 1.2.

For this purpose the total energy is sampled from an E^{-2} power law, mimicking the energy spectrum of the primary neutrinos as stated in Section 1.3.1. The total energy is divided into two parts, by assigning a fraction between 0 % and 100 % to one cascade and the remaining part to the other cascade. This is a generic approximation of the realistic process described in Section 1.2, and chosen such that the whole sample covers various cases of energy distributions between the two cascades. The decay length is sampled from an exponential distribution, as expected for a decaying heavy mass state. The resulting energy distributions and the decay length distribution are shown in Figure 1.2, where it can be seen that the individual cascade energies can be very small, due to splitting the total energy, and the decay lengths spans across several orders of magnitude.

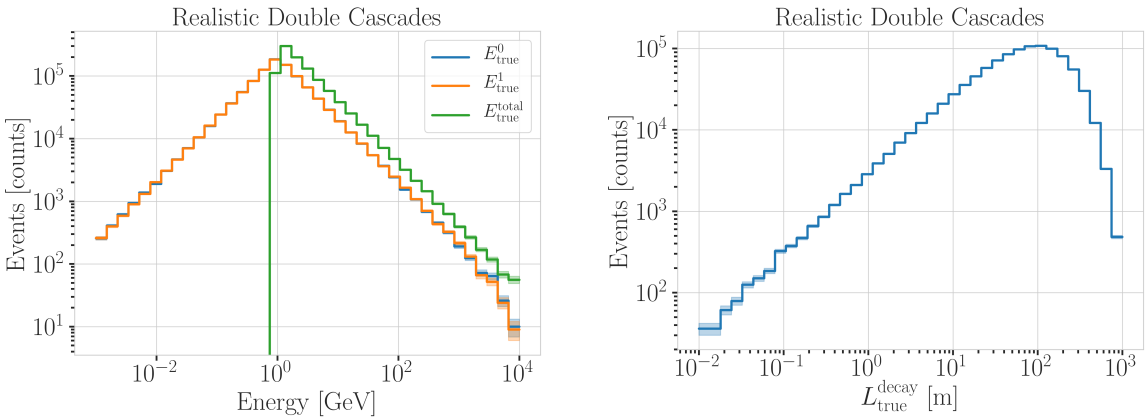


Figure 1.2: Generation level distributions of the realistic sample. Shown are the individual cascade energies and total energy (left) and decay lengths (right). It can be seen how the cascade energies can get very small, and the decay length follows a more realistic distribution spanning across several orders of magnitude.

To efficiently generate events in a way that produces distributions similar to what would be observed with DeepCore, one of the cascade positions is sampled inside the DeepCore volume by choosing its coordinates uniformly on a cylinder that is centered in DeepCore. This is similar to a trigger condition of one cascade always being inside the DeepCore fiducial volume. Choosing the direction of the event by sampling zenith and azimuth uniformly between 70° and 180° and 0° and 360°, respectively, the position of the other cascade

can be inferred for a given decay length, assuming a travel speed of c, and choosing whether the cascade position that was sampled is the first cascade or the second with a 50 % chance. The zenith angle is chosen between straight up-going (zenith of 180°) and slightly down-going from above the horizon (70°) to mimic an event selection that reduces atmospheric muons by rejecting events coming from above the horizon, but still incorporates some down-going events. All distributions are shown in Figure A.2, and the sampling distributions/values are listed in Table 1.2.

Table 1.2: Generation level sampling distributions and ranges/values of the realistic model-independent simulation.

Variable	Distribution	Range/Value
energy (total)	power law E^{-2}	1 GeV to 1000 GeV
decay length	exponential $e^{-0.01L}$	0 m to 1000 m
zenith	uniform	70° to 180°
azimuth	uniform	0° to 360°
x, y (one cascade)	uniform (circle)	c=(46.29, -34.88) m, r=150 m
z (one cascade)	uniform	−480.0 m to −180.0 m

1.2 Model-Dependent Heavy Neutral Lepton Event Generation

To estimate the HNL event expectation in IceCube DeepCore, depending on the specific model parameters, a generator was developed that is based on the HNL theory introduced in Section **??**. For this work, only the interaction with the τ-sector was taken into account ($|U_{\alpha 4}^2| = 0, \alpha = e, \mu$), which reduces the physics parameters of interest and relevant for the generation to the fourth heavy lepton mass, m_4, and the mixing, $|U_{\tau 4}^2|$.

Due to the very low interaction rate of neutrinos, which are the source of HNL production, the event generation is performed in a way that forces every event to interact in a chosen sampling volume. The weight of each event is then calculated as the inverse of the simulated neutrino fluence

$$w_{\text{gen}} = \frac{1}{F_{\text{sim}}} \frac{1}{N_{\text{sim}}} , \tag{1.1}$$

where F_{sim} is the number of neutrino events per energy, time, area, and solid angle, and N_{sim} is the number of simulated events. If this weight is multiplied by the livetime and the theoretically expected neutrino flux for a given physical model, it results in the number of expected events in the detector for this particular MC event.

The generator uses a customized *LeptonInjector (LI)* version to create the events and *LeptonWeighter (LW)* to weight them [2]. The modified LI and the essential components needed for the HNL simulation are described in the next sections, followed by the description of the weighting scheme and the sampling distributions chosen for the generation.

[2]: Abbasi et al. (2021), "LeptonInjector and LeptonWeighter: A neutrino event generator and weighter for neutrino observatories"

1.2.1 Custom LeptonInjector

In its standard version, the LI generator produces neutrino interactions by injecting a lepton and a hadronic cascade at the interaction vertex of the neutrino, where the lepton is the charged (neutral) particle produced in a CC (NC) interaction and the cascade is the hadronic cascade from

the nucleus that is breaking apart. The hadronic cascade is stored as a specific object of type *Hadrons*, which triggers the correct simulation of the shower development in the following simulation steps. Below 30 GeV the individual hadrons are simulated using GEANT4 [3] while for higher energies an analytical approximation from [4] is used. The main differences to an EM cascade is that part of the energy will not be observed, because it goes into neutral particles, and that the spatial development of the shower is different as discussed in Section **??**. Both objects are injected with the same (x, y, z, t) coordinates and the kinematics are sampled from the differential and total cross-sections that are one of the inputs to LI.

[3]: Agostinelli et al. (2003), "Geant4—a simulation toolkit"

[4]: Rädel et al. (2012), "Calculation of the Cherenkov light yield from low energetic secondary particles accompanying high-energy muons in ice and water with Geant4 simulations"

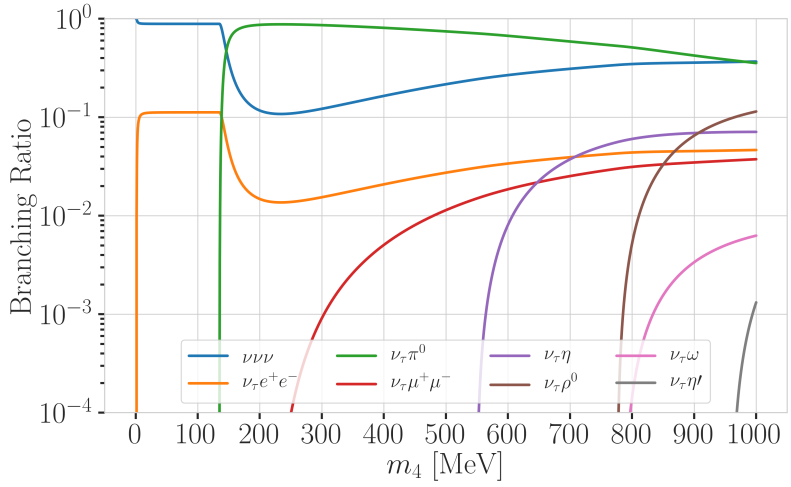


Figure 1.3: Branching ratios of the HNL within the mass range considered in this work, only considering $|U_{\tau 4}^2| \neq 0$, calculated based on the results from [5].

In the modified version, the SM lepton at the interaction vertex is replaced by the new HNL particle, where the interaction cross-sections are replaced by custom, mass dependent HNL cross-sections. The HNL is forced to decay after a chosen distance[1] to produce secondary SM particles, where the decay mode is chosen with a probability given by the mass dependent branching ratios from the kinematically accessible decay modes shown in Figure 1.3. The cross-section and decay width calculations were implemented for this purpose and will be explained in more detail in the following. Another addition to LI is that the decay products of the HNL are also stored. These HNL daughter particles form the second cascade, not as a single hadronic cascade object, but as the explicit particles forming the shower. They are injected with the correctly displaced position and delayed time from the interaction vertex, given the HNL decay length. The kinematics of the two-body decays are computed analytically, while the 3-body decay kinematics are calculated with MADGRAPH [6], which will also be explained further below. Independent of the number of particles in the final state of the HNL decay, the kinematics are calculated/simulated at rest and then boosted along the HNL momentum.

1: The explicit sampling distributions and ranges can be found in Section 1.2.2.

[6]: Alwall et al. (2014), "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations"

Muons produced in those decays are propagated using PROPOSAL [7], also simulating their Cherenkov light output. The shower development of gamma rays, electrons, and positrons below 100 MeV is also simulated using Geant4 and for higher energies the analytical approximation is used again [4].

[7]: Koehne et al. (2013), "PROPOSAL: A tool for propagation of charged leptons"

The injection is done using the LI *volume mode*, for the uniform injection of the primary particle on a cylindrical volume, adding 50 % of the events with ν_τ and the other half with $\bar{\nu}_\tau$ as primary particle types. The generator takes

the custom double-differential/total cross-section splines described below and the parameters defining the sampling distributions as inputs.

Cross-Sections

The cross-sections are calculated using the NuXSSplMkr [8] software, which is a tool to calculate neutrino cross-sections from *parton distribution functions (PDFs)* and then fit to an N-dimensional tensor-product B-spline surface [9] to produce the splines that can be read and used with LI/LW. The tool was modified to produce the custom HNL cross-sections, where the main modification to calculate the cross-sections for the ν_τ-NC interaction into the new heavy mass state, is the addition of a kinematic condition to ensure that there is sufficient energy to produce the heavy mass state. It is the same condition fulfilled for the CC case, where the outgoing charged lepton mass is non-zero. Following [10] (equation 7), the condition

[9]: Whitehorn et al. (2013), "Penalized splines for smooth representation of high-dimensional Monte Carlo datasets"

[10]: Levy (2009), "Cross-section and polarization of neutrino-produced tau's made simple"

$$(1 + x\delta_N)h^2 - (x + \delta_4)h + x\delta_4 \leq 0 \tag{1.2}$$

is implemented for the NC case in the NuXSSplMkr code. Here

$$\delta_4 = \frac{m_4^2}{s - M^2}\,, \tag{1.3}$$

$$\delta_N = \frac{M^2}{s - M^2}\,, \text{and} \tag{1.4}$$

$$h \overset{def}{=} xy + \delta_4\,, \tag{1.5}$$

with x and y being the Bjorken variables, m_4 and M the mass of the heavy state and the target nucleon, respectively, and s the center of mass energy squared. The custom version was made part of the open source NuXSSplMkr software and can thus be found in [8]. The result of this kinematic condition is that events cannot be produced for energy, x, y combinations that do not have sufficient energy to produce the outgoing, massive lepton. This results in a reduction of the cross-section towards lower energies, which scales with the assumed mass of the HNL. This effect can be seen in Figure 1.4.

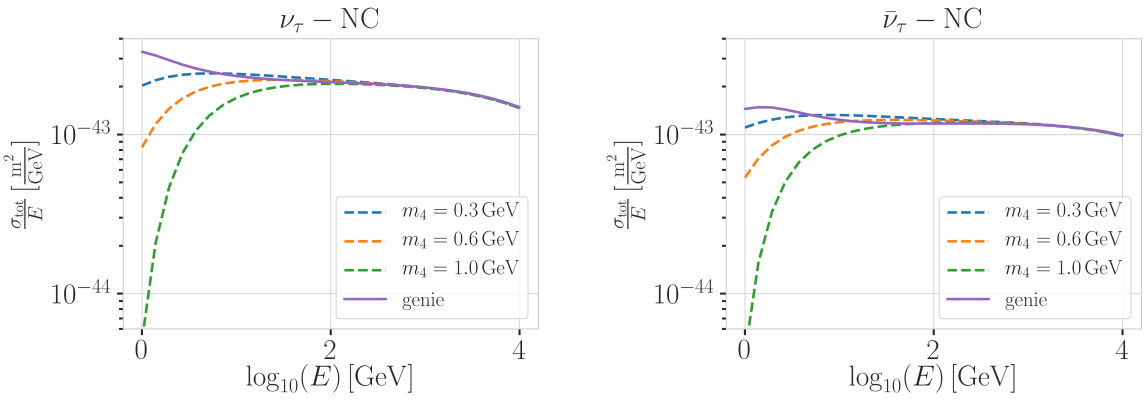


Figure 1.4: Custom HNL total cross-sections for the three target masses compared to the total ($\nu_\tau/\bar{\nu}_\tau$ NC) cross-sections used for SM neutrino simulation production with GENIE.

The GRV98LO PDFs were added to the cross-section spline maker and used to create the HNL cross-sections for consistency with the neutrino

simulation explained in Section 1.3.1. The double-differential ($d^2\sigma/dxdy$) and total (σ) cross-sections were produced for the chosen target HNL masses and then splined in energy, x, and y for $d^2\sigma/dxdy$ and σ in the energy. Figure 1.4 shows the total cross-sections that were produced compared to the cross-section used for the production of the SM $\nu_\tau/\bar{\nu}_\tau$ NC background simulation. They agree above $\sim$ 200 GeV, where the modification should not have any effect on the cross-sections. This is the desired result of using the identical input PDFs, and confirms that the unmodified cross-sections produced with NuXSSplMkr agree with the GENIE cross-sections.

Decay Channels

The accessible decay channels are dependent on the mass of the HNL and the allowed mixing. For this analysis, where only $|U_{\tau 4}|^2 \neq 0$, the decay channels considered are listed in Table 1.3 and the corresponding branching ratios are shown in Figure 1.3. The individual branching ratio for a specific mass is calculated as $\mathrm{BR}_i(m_4) = \Gamma_i(m_4)/\Gamma_{\mathrm{total}}(m_4)$, where $\Gamma_{\mathrm{total}}(m_4) = \sum \Gamma_i(m_4)$. The individual decay widths Γ_i are computed using the state-of-the-art calculations from [5], which are described in the following.

[5]: Coloma et al. (2021), "GeV-scale neutrinos: interactions with mesons and DUNE sensitivity"

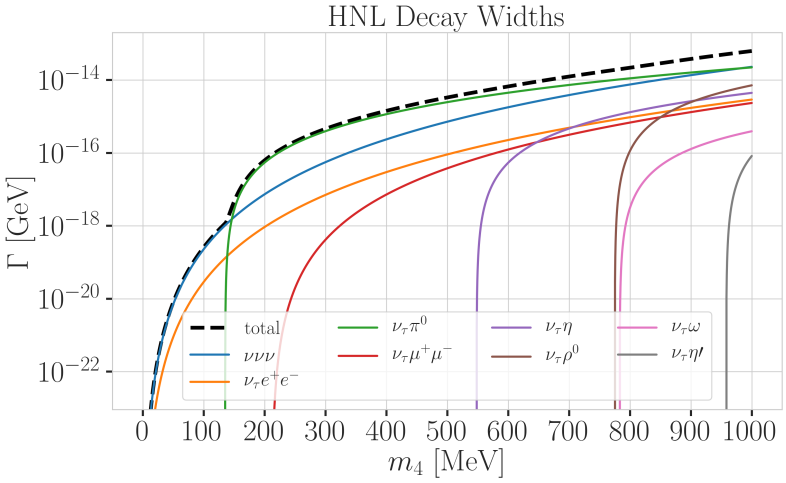


Figure 1.5: Decay widths of the HNL within the mass range considered, calculated based on the results from [5]. Given the existing constraints on $|U_{e4}|^2$ and $|U_{\mu 4}|^2$, we consider that the corresponding decay modes are negligible.

Channel	Opens
$\nu_4 \longrightarrow \nu_\tau \nu_\alpha \bar{\nu}_\alpha$	0 MeV
$\nu_4 \longrightarrow \nu_\tau e^+ e^-$	1 MeV
$\nu_4 \longrightarrow \nu_\tau \pi^0$	135 MeV
$\nu_4 \longrightarrow \nu_\tau \mu^+ \mu^-$	211 MeV
$\nu_4 \longrightarrow \nu_\tau \eta$	548 MeV
$\nu_4 \longrightarrow \nu_\tau \rho^0$	770 MeV
$\nu_4 \longrightarrow \nu_\tau \omega$	783 MeV
$\nu_4 \longrightarrow \nu_\tau \eta'$	958 MeV

Table 1.3: Possible decay channels of the HNL, considering only $|U_{\tau 4}|^2 \neq 0$, and the mass at which each channel opens.

says something about the decay widht plot, mention it in the text.. (RED)

2-Body Decay Widths The decay to a neutral pseudoscalar meson is

$$\Gamma_{\nu_4 \to \nu_\tau P} = |U_{\tau 4}|^2 \frac{G_F^2 m_4^3}{32\pi} f_P^2 (1 - x_p^2)^2 \,, \tag{1.6}$$

with $x_P = m_P/m_4$ and the *effective decay constants* f_P given by

$$f_{\pi^0} = +0.1300\,\mathrm{GeV}\,, \tag{1.7}$$
$$f_\eta = +0.0816\,\mathrm{GeV}\,, \text{and} \tag{1.8}$$
$$f_{\eta'} = -0.0946\,\mathrm{GeV}\,, \tag{1.9}$$

while the decay to a neutral vector meson is given by

$$\Gamma_{\nu_4 \to \nu_\tau V} = |U_{\tau 4}|^2 \frac{G_F^2 m_4^3}{32\pi} \left(\frac{f_V}{m_V}\right)^2 g_V^2 (1 + 2x_V^2)(1 - x_V^2)^2 \,, \tag{1.10}$$

with $x_V = m_V/m_4$,

$$f_{\rho^0} = 0.171\,\text{GeV}^2\,, \tag{1.11}$$

$$f_{\omega} = 0.155\,\text{GeV}^2\,, \tag{1.12}$$

and

$$g_{\rho^0} = 1 - 2\sin^2\theta_w\,, \tag{1.13}$$

$$g_{\omega} = \frac{-2\sin^2\theta_w}{3}\,, \tag{1.14}$$

and $\sin^2\theta_w = 0.2229$ [11], where θ_w is the Weinberg angle.

[11]: Tiesinga et al. (2021), "CODATA recommended values of the fundamental physical constants: 2018"

3-Body Decay Widths The (invisible) decay to three neutrinos, one of flavor τ and two of any flavor α, is

$$\Gamma_{\nu_4\to\nu_\tau\nu_\alpha\bar{\nu_\alpha}} = |U_{\tau 4}|^2\frac{G_F^2 m_4^5}{192\pi^3}\,, \tag{1.15}$$

while the decay to two charged leptons (using $x_\alpha = (m_\alpha/m_4)^2$) of the same flavor reads

$$\Gamma_{\nu_4\to\nu_\tau l_\alpha^+ l_\alpha^-} = |U_{\tau 4}|^2\frac{G_F^2 m_4^5}{192\pi^3}\left[C_1 f_1(x_\alpha) + C_2 f_2(x_\alpha)\right]\,, \tag{1.16}$$

with the constants defined as

$$C_1 = \frac{1}{4}(1 - 4\sin^2\theta_w + 8\sin^4\theta_w)\,, \tag{1.17}$$

$$C_2 = \frac{1}{2}(-\sin^2\theta_w + 2\sin^4\theta_w)\,, \tag{1.18}$$

the functions as

$$f_1(x_\alpha) = (1 - 14x_\alpha - 2x_\alpha^2 - 12x_\alpha^3)\sqrt{1-4x_\alpha} + 12x_\alpha^2(x_\alpha^2 - 1)L(x_\alpha)\,, \tag{1.19}$$

$$f_2(x_\alpha) = 4[x_\alpha(2 + 10x_\alpha - 12x_\alpha^2)\sqrt{1-4x_\alpha} + 6x_\alpha^2(1 - 2x_\alpha + 2x_\alpha^2)L(x_\alpha)]\,, \tag{1.20}$$

and

$$L(x) = \ln\left(\frac{1 - 3x_\alpha - (1-x_\alpha)\sqrt{1-4x_\alpha}}{x_\alpha(1+\sqrt{1-4x_\alpha})}\right)\,. \tag{1.21}$$

Analytical 2-Body Decay Kinematics

[12]: Workman et al. (2022), "Review of Particle Physics"

Following the review of [12], the 4-vector defining the kinematics of a particle is $p = (E, \vec{p})$, with its energy, E, and 3-momentum, $\vec{p}$. Squaring it gives the mass, $p^2 = E^2 - \vec{p}^2 = m^2$, while the velocity is $\vec{\beta} = \vec{p}/E$. If the HNL with mass m_4 decays into two particles with masses m_1 and m_2, their 3-momenta in the rest frame of the HNL are given by

$$|\vec{p}_1| = |\vec{p}_2| = \frac{\lambda^{1/2}(m_4^2, m_1^2, m_2^2)}{2m_4}\,, \tag{1.22}$$

where $\lambda(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2xz - 2yz$. The energy of the particles is then given by

$$E_1 = \frac{m_4^2 + m_1^2 - m_2^2}{2m_4} , \tag{1.23}$$

and equivalently for E_2. The 4-vectors of the particle are then boosted to the lab frame, where the HNL is moving with velocity $\vec{\beta}$.

Simulated 3-Body Decay Kinematics

The 3-body decay kinematics cannot be computed analytically, instead, we employ MadGraph4 (v3.4.0) [13] for this purpose. MadGraph is a tool to simulate particle collisions and decay processes, and is widely used in the high-energy physics community. The 3-body decay kinematics are calculated in the rest frame of the HNL, using decay diagrams calculated with FeynRules 2.0 [14] and the Lagrangians derived in [5] as input. The *Universal FeynRules Output (UFO)* from effective_HeavyN_Majorana_v103 were used for our calculation. For each mass and corresponding decay channels, we produce 1×10^6 decay kinematic variations in the rest frame and store those in a text file. During event generation, we uniformly select an event from that list, to simulate the decay kinematics of a 3-body decay.

[14]: Alloul et al. (2014), "FeynRules 2.0 - A complete toolbox for tree-level phenomenology"

[5]: Coloma et al. (2021), "GeV-scale neutrinos: interactions with mesons and DUNE sensitivity"

1.2.2 Sampling Distributions

Variable	Distribution	Range/Value
energy	E^{-2}	[2 GeV, 1×10^4 GeV]
zenith	uniform (in $\cos(\theta)$)	[80°, 180°]
azimuth	uniform	[0°, 360°]
vertex x, y	uniform (on circle)	r=600 m
vertex z	uniform	−600 m to 0 m
m_4	fixed	[0.3, 0.6, 1.0] GeV
L_{decay}	L^{-1}	[0.0004, 1000] m

Table 1.4: Generation level sampling distributions and ranges/values of the model-dependent simulation samples.

In principle, the generation level sampling distributions should be chosen such that at the final level of the event selection chain the phase space relevant for the analysis is covered with sufficient statistics to make a reasonable estimate of the event expectation. Initial distributions insufficiently covering the phase space lead to an underestimation of the expected rates, because some of the events that would pass the selection are not produced. This limits the expected analysis potential. Three discrete simulation samples were produced with HNL masses of 0.3 GeV, 0.6 GeV, and 1.0 GeV. The remaining sampling distributions are identical for all samples and are listed in Table 1.4. The target number of events for each sample was 2.5×10^9 at generation to result in sufficient MC statistics at final level. Figure 1.6 shows the true cascade energies, which result from the custom interaction cross-sections and the decays discussed above. Note here that these are the full true energies going into the cascades. For the first cascade some energy goes into invisible particles produced in the hadronic shower development and for the HNL decay, there always is at least one invisible neutrino. Additional sampling distributions can be found in Figure A.3.

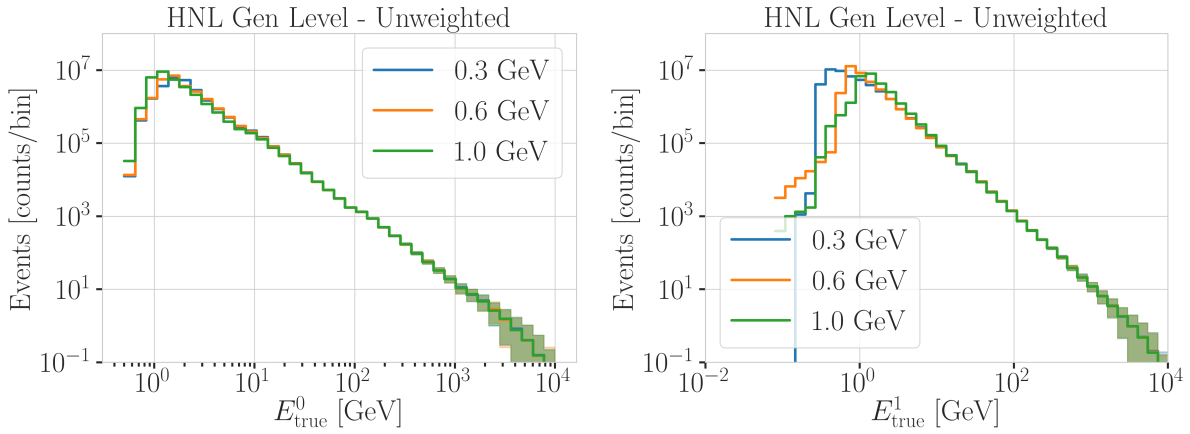


Figure 1.6: Generation level distributions of the model-dependent simulation. Shown are the true energies of both cascades including the energy that goes into invisible particles.

JVS: The build-up of the weight expression is hard to follow without knowing where it's going. It may be better to start with the fact that the importance sampling weight is the ratio of PDFs, then write down each pdf, then drill down into each of the terms (basically, the standard "tell me what you're going to tell me, then tell me, then tell me what you told me" scheme). (RED)

1.2.3 Weighting Scheme

To produce physically correct event distributions based on the simplified generation sampling distributions, each event has to be re-weighted. The only required input is the mixing strength $|U_{\tau 4}|^2$, which is the variable physics parameter in this analysis. For each event the gamma factor

$$\gamma = \frac{\sqrt{E_{\text{kin}}^2 + m_4^2}}{m_4}\,, \tag{1.24}$$

is calculated, with the HNL mass m_4, and its kinetic energy E_{kin}. The speed of the HNL is calculated as

$$v = c \cdot \sqrt{1 - \frac{1}{\gamma^2}}\,, \tag{1.25}$$

where c is the speed of light. With these, the lab frame decay length range $[s_{\text{min}}, s_{\text{max}}]$ can be converted into the rest frame lifetime range $[\tau_{\text{min}}, \tau_{\text{max}}]$ for each event

$$\tau_{\text{min/max}} = \frac{s_{\text{min/max}}}{v \cdot \gamma}. \tag{1.26}$$

The proper lifetime of each HNL event can be calculated using the total decay width Γ_{total} from Section 1.2.1 and the chosen mixing strength $|U_{\tau 4}|^2$ as

$$\tau_{\text{proper}} = \frac{\hbar}{\Gamma_{\text{total}}(m_4) \cdot |U_{\tau 4}|^2}\,, \tag{1.27}$$

where $\hbar$ is the reduced Planck constant. Since the decay lengths or lifetimes of the events are sampled from an inverse distribution instead of an exponential, as it would be expected from a particle decay, we have to re-weight accordingly to achieve the correct decay lengths or lifetimes distribution. This is done by using the desired exponential distribution

$$\text{PDF}_{\text{exp}} = \frac{1}{\tau_{\text{proper}}} \cdot e^{\frac{-\tau}{\tau_{\text{proper}}}}\,, \tag{1.28}$$

and the inverse distribution that was sampled from

$$\mathrm{PDF}_{\mathrm{inv}} = \frac{1}{\tau \cdot (\ln(\tau_{\mathrm{max}}) - \ln(\tau_{\mathrm{min}}))} \,. \tag{1.29}$$

This re-weighting factor is then calculated as

$$w_{\mathrm{lifetime}} = \frac{\mathrm{PDF}_{\mathrm{exp}}}{\mathrm{PDF}_{\mathrm{inv}}} = \frac{\Gamma_{\mathrm{total}}(m_4) \cdot |U_{\tau 4}|^2}{\hbar} \cdot \tau \cdot (\ln(\tau_{\mathrm{max}}) - \ln(\tau_{\mathrm{min}})) \cdot e^{\frac{-\tau}{\tau_{\mathrm{proper}}}} \,. \tag{1.30}$$

Adding another factor of $|U_{\tau 4}|^2$ to account for the mixing at the interaction vertex the total re-weighting factor becomes

$$w_{\mathrm{total}} = |U_{\tau 4}|^2 \cdot w_{\mathrm{lifetime}} \,. \tag{1.31}$$

If this additional weighting factor is multiplied to the generation weight from Equation 1.1 with units m^2, the livetime in s, and the oscillated primary neutrino flux in $\mathrm{m}^{-2}\mathrm{s}^{-1}$, it results in the number of expected events in the detector for this particular MC event for a chosen mixing (and mass).

1.3 Standard Model Event Generation

1.3.1 Neutrinos

The simulation volume is a cylinder centered in DeepCore with radius and height chosen such that all events possibly producing a signal are contained. The different sizes, chosen depending on energy and neutrino flavor, are shown in Table 1.5. The directions of the neutrinos are sampled isotropically and the energies are sampled from an E^{-2} power law. The number of simulated events is chosen such that the livetime is more than 70 years for each flavor. Neutrinos and antineutrinos are simulated with ratios of 70% and 30%, respectively, which is roughly the ratio expected from the atmospheric neutrino flux [15].

To simulate the neutrino interaction with the ice, the GENIE event generator [16] (version 2.12.8) is used, resulting in the secondary particles and the

[15]: Honda et al. (2015), "Atmospheric neutrino flux calculation using the NRLMSISE-00 atmospheric model"

[16]: Andreopoulos et al. (2015), "The GENIE Neutrino Monte Carlo Generator: Physics and User Manual"

Table 1.5: Cylinder volumes used for GENIE neutrino simulation generation. Cylinder is always centered in DeepCore at $(x, y, z) = (46.29, -34.88, -330.00)$ m.

Flavor	**Energy** [GeV]	**Radius** [m]	**Length** [m]	**Events/File**	**Files**
$\nu_e + \bar{\nu_e}$	1-4	250	500	450000	650
	4-12			100000	
	12-100	350	600		
	100-10000	550	1000	57500	
$\nu_\mu + \bar{\nu_\mu}$	1-5	250	500	408000	1550
	5-80	400	900	440000	
	80-1000	450	1500	57500	
	1000-10000	550		6700	
$\nu_\tau + \bar{\nu_\tau}$	1-4	250	500	1500000	350
	4-10			300000	
	10-50	350	600	375000	
	50-1000	450	800	200000	
	1000-10000	550	1500	26000	

kinematic and cross-section parameters. As input, the outdated GRV98LO [17] PDFs were used, because they were the only option that could incorporate extrapolations to lower Q^2 [18]. The propagation of the secondary particles and of the shower development is performed identical to the description in Section 1.2.1 and produces the energy losses and event morphologies introduced in Section **??**.

[17]: Glück et al. (1998), "Dynamical parton distributions revisited"

[18]: Bodek et al. (2003), "Higher twist, xi(omega) scaling, and effective LO PDFs for lepton scattering in the few GeV region"

1.3.2 Muons

Atmospheric muons are generated on a cylinder surface enclosing the full IceCube detector array. The cylinder has a height of 1600 m and a radius of 800 m. The energy is sampled from an E^{-3} power law while the other sampling distributions (position, direction) are found from parameterizations based on [19]. This work uses full CORSIKA [20] simulations of muons to tailor the parameterizations, starting from *cosmic ray (CR)* interactions with atmospheric nuclei using the CR flux model from [21] and producing the muons applying the hadronic interaction model SIBYLL 2.1 [22]. After the generation, they are propagated through the ice with PROPOSAL producing photons, treating them exactly like the muons produced in the HNL and neutrino event generation.

[19]: Becherini et al. (2006), "A parameterisation of single and multiple muons in the deep water or ice"

[20]: Heck et al. (1998), "CORSIKA: A Monte Carlo code to simulate extensive air showers"

[21]: Gaisser (2012), "Spectrum of cosmic-ray nucleons, kaon production, and the atmospheric muon charge ratio"

[22]: Engel et al. (2017), "The hadronic interaction model Sibyll – past, present and future"

Since the offline processing and selection steps described in Section 2.1.2 and Section 2.2 reduce the muon contamination to an almost negligible level, the statistical uncertainty on the number of expected muon events at the final selection level is large and therefore two separate samples of muon simulation are produced. **A first sample** is used to tune the lower level selection (up to Level 4), therefore including all events resulting from the above described generation. **A second sample** is then produced to estimate the muon contamination at higher levels (above Level 5). It only consists of muon events that pass through a smaller cylinder centered in DeepCore (height of 400 m and radius of 180 m), and additionally rejects events based on a KDE estimated muon density at Level 5 (in energy and zenith). This increased the simulation efficiency at Level 5 significantly, making it feasable to use this sample to estimate the muon contamination at higher levels.

1.4 Detector Simulation

The detector simulation is performed after the event generation, where the initial particles and the resulting photons and secondary particles from their propagation were produced. This part of the simulation chain is applied to all muon and neutrino simulation as well as the HNL signal simulation explained in detail in Chapter 1. The detector simulation can be split into two parts: the propagation of the photons and the simulation of the detector response (including internal noise).

1.4.1 Photon Propagation

Any photon that was produced in the event generation is individually traced through the ice, simulating scattering and absorption processes. The propagation is done using CLSIM [23] which is an implementation of the *Photon Propagation Code (PPC)* [24] in OPENCL. It is optimized to be

[24]: Chirkin et al. (2019), "Photon Propagation using GPUs by the IceCube Neutrino Observatory"

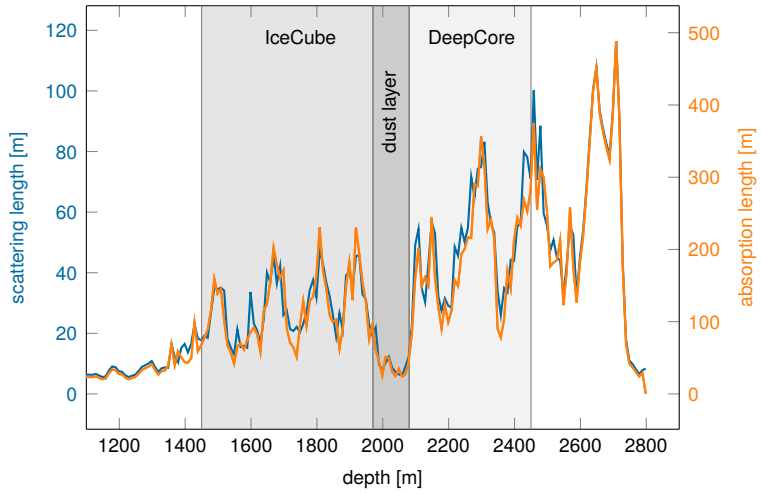


Figure 1.7: Scattering and absorption lengths as a function of depth in the SPICE model used for simulation. Modified from [26].

run efficiently on GPUs. The ice is modeled as a set of 10 m thick, almost horizontal layers with specific absorption and scattering lengths. The *South Pole ice (SPICE)* model [25] accounts for the layers being tilted by a small amount and the absorption and scattering lengths having a non-uniformity with respect to the azimuth direction. Figure 1.7 shows the values of this model for the different depths, indicating the location of IceCube, DeepCore, and the dust layer.

[25]: Aartsen et al. (2013), "Measurement of South Pole ice transparency with the IceCube LED calibration system"

add half sentence why this is (RED)

In an initial step, each photon's absorption length is sampled from an exponential distribution with the expectation value at the current layer's absorption length. The following propagation steps are performed in parallel for all photons. In each of those steps, corresponding to a single scattering event, the photon travels a length that is sampled from an exponential distribution with the expectation value at the scattering length of the current layer and the scattering angle chosen based on a combination of a simplified Mie scattering distribution [27] and a Henyey-Greenstein distribution [28] [29]. The parameters defining the shape of these distributions were calibrated using data from *in-situ* LED calibration runs. These steps are continuously repeated until each photon reaches a DOM or is absorbed[2]. After all photons have been propagated in that manner, the final step is to store the photons that reached a DOM for further processing.

[27]: Mie (1908), "Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen"

[28]: Henyey et al. (1941), "Diffuse radiation in the Galaxy."

[29]: Abbasi et al. (2022), "In-situ estimation of ice crystal properties at the South Pole using LED calibration data from the IceCube Neutrino Observatory"

2: A photon is absorbed, when it traveled its full absorption length, sampled in the initial step of the photon propagation.

1.4.2 Detector Responses

The second part of simulating the IceCube detector is the DOM response. For a photon that reaches a DOM, the probability to produce a signal depends on the total efficiency and the angular acceptance of the specific DOM. The total efficiency includes effects of the DOM glass, PMT quantum and photoelectron collection efficiencies, and it is wavelength dependent. Additionally, there is another angle dependent effect called *hole ice* [30]. This effect is due to varied ice properties resulting from the re-freezing process of the water column inside the borehole after deployment of the string. Accepted photons are converted into a so-called *Monte Carlo photo-electron (MCPE)*. The amount of charge measured for each MCPE is determined by sampling from the so-called *single photo-electron (SPE)* distribution, which was tuned to match the observed distribution in each DOM in an *in-situ* calibration study [31]. Figure 1.8 shows the spread of the distribution measured over

[30]: Fiedlschuster (2019), "The Effect of Hole Ice on the Propagation and Detection of Light in IceCube"

[31]: Aartsen et al. (2020), "In-situ calibration of the single-photoelectron charge response of the IceCube photomultiplier tubes"

all DOMs compared to lab measurements of a specific PMT type. Based on the sampled charges and times of MCPEs, the voltage waveforms for the (two) different readout channels are simulated and passed on to the trigger simulation starting with *WaveDeform* (see Section **??**).

Parameter	Value
Therm. rate λ_{th}	180 Hz
Decay rate λ_{dec}	80 Hz
Decay hits η	8.5
Decay μ	4.3 $\log_{10}$(ns)
Decay σ	1.8 $\log_{10}$(ns)

Table 1.6: Typical parameter values used in the vuvuzela noise simulation. Averaged over all DOMs.

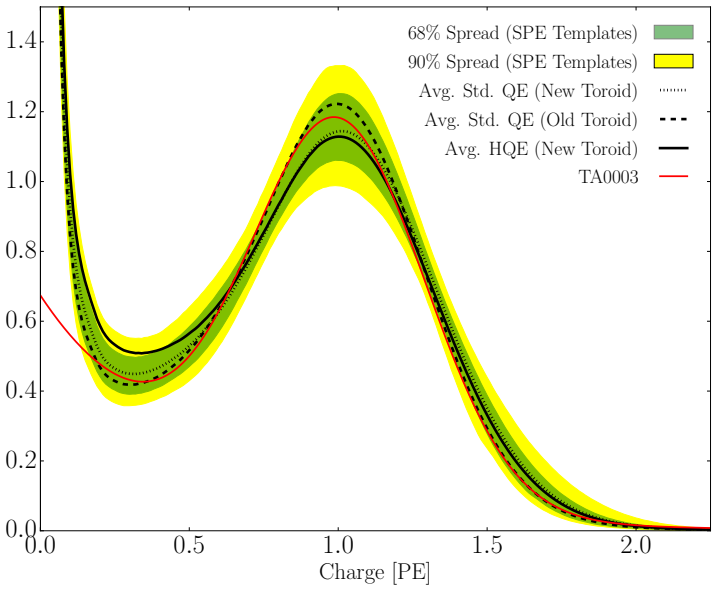


Figure 1.8: Single photo-electron charge distribution shown for a lab measurement in red (TA0003), various hardware configurations in black dashed, dotted, and solid lines, and the 68 % and 90 % spread of the measured charged templates for all DOMs. All curves are normalized to the same area. The figure is taken from [31].

Besides the Cherenkov photons, IceCube also observes photons that are produced in radioactive decays inside the DOMs, both in the glass housing sphere and the PMT glass itself. To simulate this internal noise, the *Vuvuzela* module [32, 33] is used to create additional MCPEs that are fed into the same simulation chain described above. The noise hits are simulated by drawing the times from a constant rate Poisson process and the number of photons from a Poisson distribution. Then the time differences between the individual photons per hit is found, based on a Log-Normal distribution. The simulation is defined by 5 parameters that are calibrated for each DOM individually. Table 1.6 shows the average values for these parameters.

[32]: Larson (2013), "Simulation and Identification of Non-Poissonian Noise Triggers in the IceCube Neutrino Detector"
[33]: Larson (2018), "A Search for Tau Neutrino Appearance with IceCube-DeepCore"

Event Processing and Reconstruction

2

The analysis presented in this thesis is highly dependent on an efficient filtering and event selection to reduce the raw IceCube trigger data to a usable atmospheric neutrino sample. Based on this selection, a precise estimation of both expected SM background and expected BSM signal events can be made using MC simulations. This chapter describes the event selection chain used for state-of-the-art IceCube neutrino oscillation measurements like [34]. Starting from the PMT output, both real data and simulation are processed through the in-ice trigger, the online filter and processing, and the low-energy event selection to produce a neutrino dominated sample. Once the sample is small enough for more sophisticated reconstruction techniques to be feasible to run, the events can be reconstructed with the existing IceCube reconstruction algorithms. Using the reconstruction outputs and some high level variables that are also computed, the final event selection is performed.

[34]: Abbasi et al. (2023), "Measurement of atmospheric neutrino mixing with improved IceCube DeepCore calibration and data processing"

re work with combined reco and so on? (RED)

2.1 Processing

After the detector simulation is performed, all MC and data are processed in exactly the same way. This section explains the trigger and event selection that is applied starting from the raw voltage measured by the PMTs. It is split in different steps run inside the ice, at the South Pole, and after the data was transferred to the North. The complexity and computational cost of the processing increases with each step, while the total number of events reduces, making it feasible and reducing the use of computational resources on events that are not of interest for the analysis.

2.1.1 Trigger and Filter

Before the data can be sent to the North, the initial signal coming from the PMT is a voltage waveform that has to be digitized (for data) and then information of photon hits has to be extracted (also for the MC coming from the detector response simulation). The trigger and filter explained here are tailored to select events that passed through the DeepCore volume, while rejecting background events (either from atmospheric muons or from random noise). There are other filters used in IceCube which will not be explained here, since they are not relevant for this work. A full description of the instrumentation and the online systems can be found in [35].

[35]: Aartsen et al. (2017), "The IceCube Neutrino Observatory: Instrumentation and Online Systems"

Include some low level plots like the trigger efficiency for the HNL simulation (ORANGE)

In-ice Trigger

The trigger is applied inside the DOM in the ice before sending the information to the ICL on the surface. The time dependent voltage curves are captured if a pre-defined threshold value is exceeded. Once the threshold set to the equivalent of 0.25 PE is crossed, 6.4 µs of the waveform are coarsely digitized by a *Fast Analog-to-Digital Converter (FADC)* with a sampling rate of

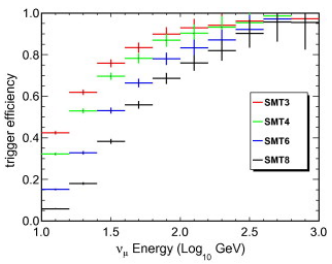


Figure 2.1: Efficiencies of different IceCube and DeepCore triggers, taken from [36].

40 MHz. Additionally, the first 427 ns are digitized using an *Analog Transient Waveform Recorder (ATWD)* with a sampling rate of 300 MHz [37], but only if some trigger condition is met, because this readout frequency is too high to be sampled directly and requires some buffering. For DeepCore, the HLC condition already mentioned in Section **??** has to be met for three DOMs inside the fiducial volume within a time window of 5 µs. If this is the case, all waveforms that crossed the threshold within a 20 µs time window around the trigger are digitized and sent to the ICL for further processing. This trigger is called *Simple Multiplicity Trigger 3 (SMT-3)*. The DOM hits that are read out in this process, but do not meet the HLC condition, are called *soft local coincidence (SLC)* hits. The rate of the DeepCore SMT-3 trigger is ~250 Hz [38], accepting ~70 % of ν_μ-CC events at 10 GeV and ~90 % at 100 GeV [36]. The trigger efficiencies for different SMT triggers, including the DeepCore SMT-3, are shown in Figure 2.1.

[37]: Abbasi et al. (2009), "The IceCube data acquisition system: Signal capture, digitization, and timestamping"

[38]: Aartsen et al. (2017), "The IceCube Neutrino Observatory: instrumentation and online systems"

[36]: Abbasi et al. (2012), "The design and performance of IceCube DeepCore"

Online Filter

The digitized waveforms are sent to the ICL, where a further filter is applied *online*[1]. First, the WaveDeform algorithm is run to extract photon arrival times and charge from the waveforms, then the DeepCore filter is applied, which is an iterative hit cleaning starting from HLC hits and removing any hits outside a 125 m radius and a 500 ns time window (called *radius-time cleaning (RT-cleaning)*) of the initial hit. This mainly rejects unphysical SLC hits, which are potentially caused by random noise. The following selection steps are done using the resulting cleaned pulses.

1: Where *online* means running on hardware at the South Pole.

Next, an additional cut is applied to reject events that are likely to be caused by atmospheric muons. This is done by splitting the hits depending on whether they were inside the DeepCore fiducial volume or outside and then calculating the speed of each hit outside the fiducial volume towards the *center of gravity (COG)* of the hits inside. If one of them has a speed close to the speed of light, the whole event is rejected, because this is a strong indication for a muon event.

As input for the further selection levels, a few event properties, like vertex position and direction, are determined using fast and simple event reconstructions. After the DeepCore online filter, the rate is about 15 Hz, which can be sent to the North via satellite for further processing.

2.1.2 Event Selection

After the data was sent to the North, the *offline* filters and selection are applied to further recude reduce the background of atmospheric muons and noise. The selection is split into three levels referred to as *Level 3-5 (L3-L5)*, which bring down the neutrino and muon rate to ~1 mHz, while the remaining fraction of random noise is below 1 %.

Level 3

At the first offline filtering level, Level 3, 1D cuts are used to reduce atmospheric muons, pure noise, and coincident muons. These cuts are targeting regions where the data/MC agreement is poor, so that more sophisticated

machine learning (ML) techniques can be applied at later levels. The cuts are made using 12 control variables, that are inexpensive to compute for the very large sample at this stage. The variables are related to position, time, and overall number of hits in the event.

Pure noise hits, that are temporally uncorrelated, are cleaned by applying a 300 ns sliding window, requiring the containment of more than 2 hits at its maximum. Additionally, an algorithm is run to check whether the hits show some directionality, accepting them only if they do.

To reduce the amount of muons a series of cuts is applied using spatial and temporal information. Events that have more than 9 hits observed above −200 m or the first HLC hit above −120 m are rejected as well as events where the fraction of hits in the first 600 ns of the event is above 0.37, ignoring the first two hit DOMs. Additionally, the ratio between hits in the veto region and the DeepCore fiducial volume is required to be below 1.5.

If a muon enters the detector after the data acquisition was already triggered, it causes events that span over a much larger time range. To reduce those coincident events, the time difference between first and last pulse cannot be above 5000 ns. This cut mainly affects a region of very poor data to MC agreement, because coincident events are not simulated at all.

The L3 cuts remove 95 % of the atmospheric muons and >99 % of pure noise hits, while keeping >60 % of the neutrino events. The sample now roughly contains muons/neutrinos/noise at a ratio of 100:10:1 with a total rate of ~0.5 Hz.

add example plots (2?) for L3 cut variables and applied cuts (YELLOW)

Level 4

After the total rate was reduced by the simple cuts of L3 and the overall agreement between data and MC is established, ML techniques can be applied to further reduce the background. For Level 4, two *Boosted Decision Trees (BDTs)* [39] classifier are trained to separate neutrino events from atmospheric muons and noise hits, separately. The output of each classifier, a probability score, can be seen in Figure 2.2. The noise filter is applied first and an event passes the score if it is larger than 0.7, reducing the noise hits by a factor of 100, while keeping 96 % of neutrinos. Then the second BDT classifier is applied to reject muons. It was trained partly on unfiltered data, which consists of >99 % atmospheric muons, to reject the data and keeping the neutrinos from the simulation. Rejecting events with a score smaller than 0.65 removes 94 % of atmospheric muons while keeping 87 % of neutrinos. This fraction varies depending on the flavor and interaction type, ν_μ-CC events for example, which have a muon in the final state, are therefore reduced to 82.5 %. After applying the L4 cuts based on the BDT classifier outputs, the sample is still dominated by atmospheric muons, while the noise rate dropped to below most neutrino types.

[39]: Friedman (2002), "Stochastic gradient boosting"

Level 5

Level 5 is the final selection level, before event reconstructions are applied. This level aims to reduce the remaining atmospheric muon rate below the rate of neutrinos. Muons not rejected by the earlier levels are those that produced little or no light in the veto regions. One possible reason is that

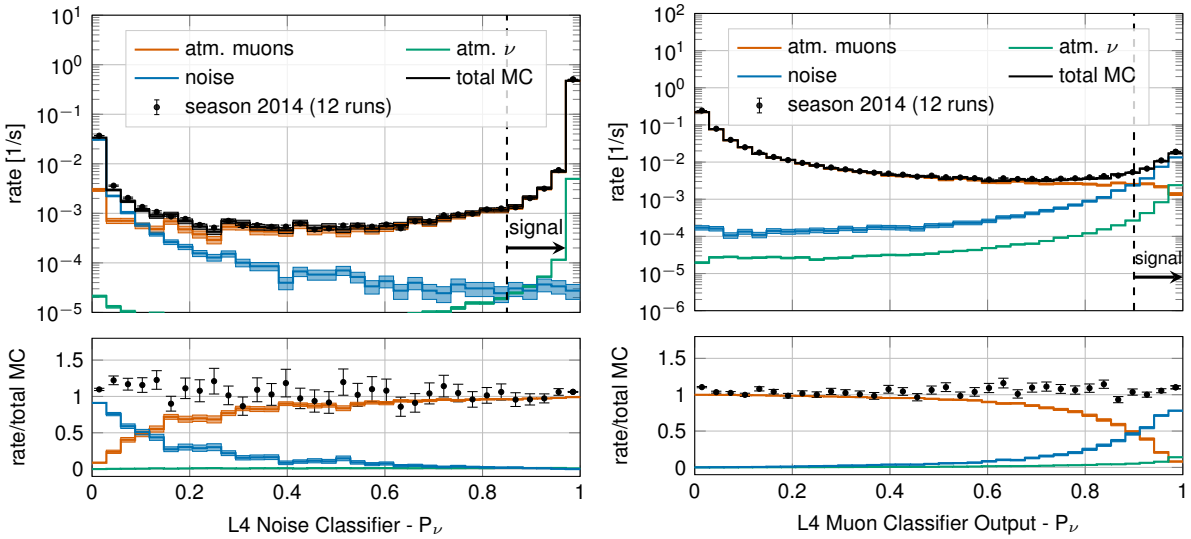


Figure 2.2: Distributions of Level 4 noise classifier output (left) and muon classifier output (right), where larger values indicate more neutrino-like and lower values more noise-like/muon-like. Taken from [34].

add some figure showing the corridors? (YELLOW)

they passed through one of the un-instrumented regions between the strings called *corridors*. To reject those, special corridor cuts, based on the number of hits they produced close to a potential corridor they passed through. The potential corridor in questions is identified based on a simple infinite track fit. In addition to the corridor cuts, starting containment cuts are applied to reject events that start at the edge of the fiducial volume. Events with more than seven hits in the outermost strings of the detector or those that have a down going direction in the uppermost region are rejected. This further reduces the fraction of muons by 96 % while keeping 48 % of neutrinos. The rates after this level are 1 mHz and 2 mHz for neutrinos and muons, respectively, making it a neutrino dominated sample.

add table with rates per level (split in flavor) - maybe better in analysis chapter to also show signal? (RED)

2.2 Reconstruction

In the energy range most relevant for this work, between 10 GeV and 100 GeV, the light deposition is very low and only a few DOMs detect light, making the reconstructions difficult. In [40] two classical methods are described, which have partly been applied in one recent IceCube atmospheric neutrino oscillation measurement using a sub-sample of the DeepCore sample [34]. The algorithm used in this work on the other hand, is a newer method that applies a *convolutional neural network (CNN)* to reconstruct the events and determine some discriminating quantities. The latest muon neutrino disappearance result from IceCube [41] is based on this reconstruction.

[40]: Abbasi et al. (2022), "Low energy event reconstruction in IceCube DeepCore"

[34]: Abbasi et al. (2023), "Measurement of atmospheric neutrino mixing with improved IceCube DeepCore calibration and data processing"

[41]: Yu et al. (2023), "Recent neutrino oscillation result with the IceCube experiment"

2.2.1 Fast Low-Energy Reconstruction using Convolutional Neural Networks

As the name *Fast Low-Energy Reconstruction using Convolutional Neural Networks (FLERCNN)* already indicates, the FLERCNN reconstruction [42, 43] is a CNN optimized to reconstruct IceCube events at low energies (<100 GeV) in a fast and efficient manner, by leveraging the approximate translational

[42]: Yu et al. (2021), "Direction reconstruction using a CNN for GeV-scale neutrinos in IceCube"

[43]: Micallef (),

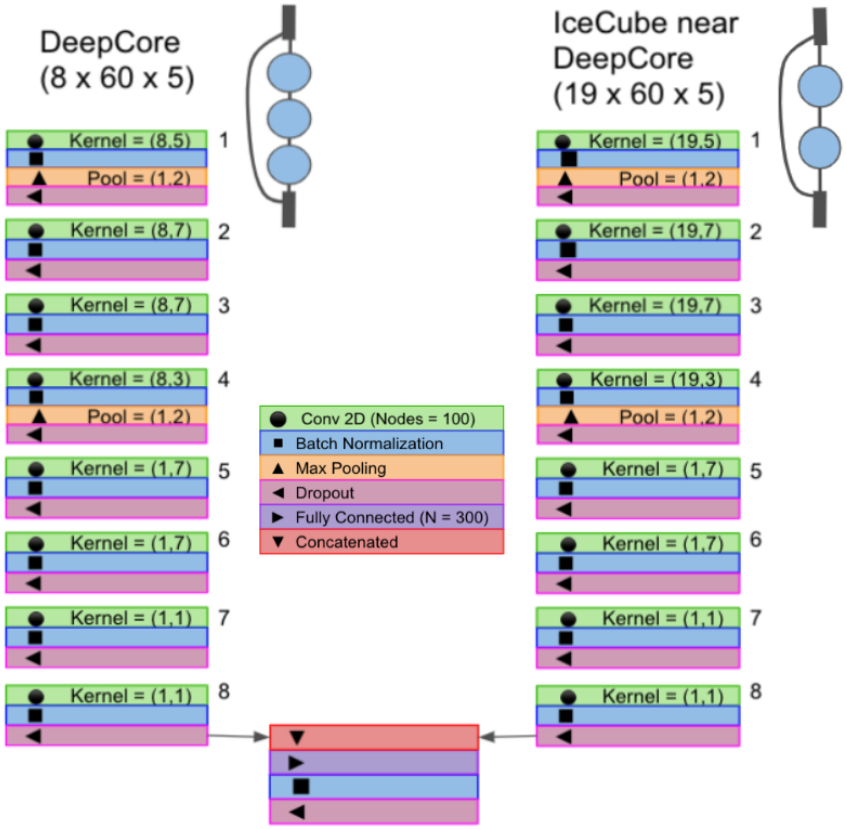


Figure 2.3: Architecture of the FLERCNN neural networks, taken from [42].

invariance of event patterns within the detector. The architecture of the network is very similar to the preexisting IceCube CNN event reconstruction [44], but optimized on low-energy events and specifically tailored to include the DeepCore sub-array. Only the eight DeepCore strings and the central 19 IceCube strings are used for the reconstruction (compare to Figure **??**). Because of the different z-positions of the DeepCore and IceCube DOMs, they are divided into two networks that are combined in the final layer of the network. The full architecture is shown in Figure 2.3. The first dimension of the network is the string index, while the second dimension is the order of the DOMs along the vertical axis. The horizontal position of the DOMs is not used, since the strings are arranged in an irregular pattern. The information from the DOM hits is summarized into five charge and time variables, which make up the last dimension of the input layer. The variables are the total summed charge, the time of the first hit, the charge weighted mean time of the hits, the time of the last hit, and the charge weighted standard deviation of the hit times.

[44]: Huennefeld (2017), "Deep Learning in Physics exemplified by the Reconstruction of Muon-Neutrino Events in IceCube"

add image with selected strings used for flercnn IC and DC (YELLOW)

Five different networks are trained using this architecture. Three networks do the regression of the events' energy, zenith angle, and the starting vertex (x, y, z position), while two of them are used for classification. One is trained to predict the probability of the event being a track (used as PID) and the other to predict the probability of the event being a muon. Each network is trained with an MC sample modified to have a flat distribution in the target variable, to be unbiased for that variable and ideally extending outside the target reconstruction region. For the classification tasks the loss function is the *binary cross entropy* and the activation function is a *sigmoid*. To perform the regression of zenith and vertex position, the loss function is the *mean squared error (MSE)*, while for the energy it is the *mean absolute percentage*

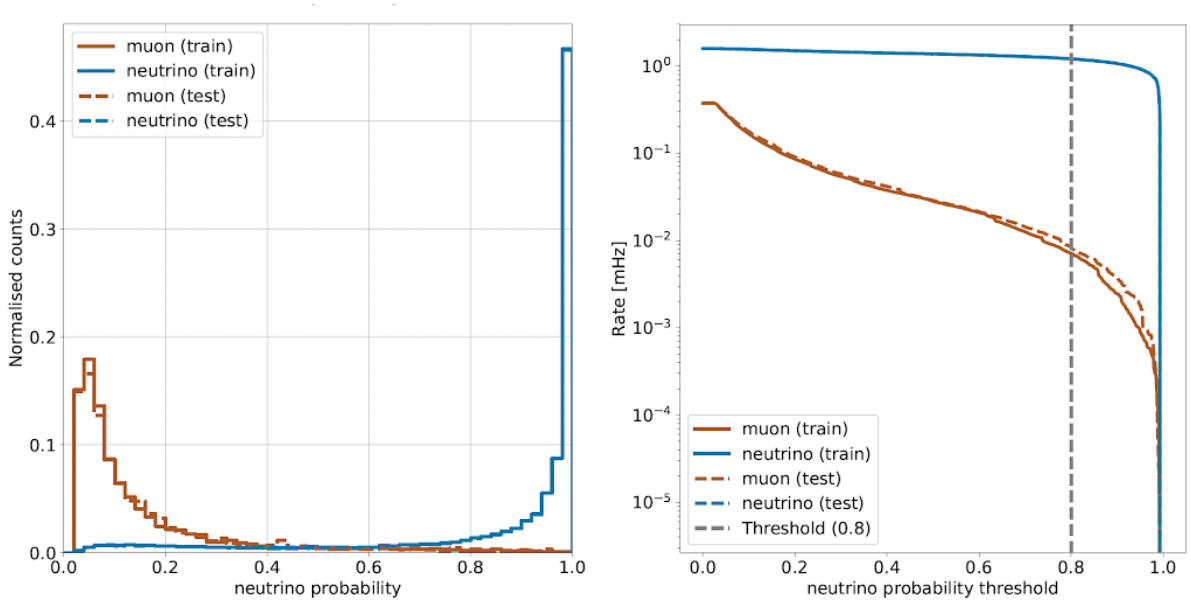


Figure 2.4: FLERCNN muon classifier output score (left) and rate of neutrinos and muons as function of muon classifier cut (right). Taken from **[flercnn_analysis_internal_note]**

add some performance plots of the FLERCNN reconstruction (ORANGE)

There is more information on pre-processing the samples and preparing the input features, and training each cnn, but I'm not sure if that might be too much detail? (YELLOW)

error. The activation for all regression tasks is *linear*.

2.2.2 Analysis Selection

Before the reconstruction is applied a few additional high level variables are computed, which are from fast and inexpensive algorithms. Then the reconstruction is performed by applying the trained FLERCNN networks to get the output quantities. After that, another BDT classifier is trained to further reduce the muon background for the final sample. The BDT is trained on five high level variables, where three are FLERCNN reconstruction variables (vertex z, ρ_{36}[2], and muon probability) and two are lower level variables (L4 muon classifier output and L5 corridor cut variable). To train the BDT, the FLERCNN nominal simulation set is used, only using events with $\cos(\theta_{zenith}) \leq 0.3$. The output of the BDT is the neutrino probability and a cut at 0.8 is applied to reject events with a high probability of being a muon. Figure 2.4 shows the output of the BDT classifier, where the neutrinos in both training and testing sets are gathered at 1 and muons are around 0, which shows great classification power.

2: A radial variable that is often used in IceCube, is the horizontal distance to string 36 called ρ_{36}, which is basically the distance to the center of IceCube.

add reference for flercnn analysis internal note (ORANGE)

To get the final, pure sample of well reconstructed neutrinos another set of cuts is applied. The first cuts are meant to reject events with poor reconstruction quality, by requiring the events to fall into the DeepCore volume, where the denser, better instrumented detector leads to enhanced resolution. The cuts are applied on the vertex z and ρ_{36} and are listed in Table 2.1. The FLERCNN reconstruction was optimized for atmospheric neutrino analyses which are mainly in the region below 100 GeV and there are very few events with energies below 5 GeV, so the reconstructed energy is required to be in that range. Additionally, rejecting events with fewer than seven hits in the selected DOMs used for FLERCNN showed to increase the resolution.

Another set of cuts is applied to make sure the agreement between data

Variable	Threshold	Removed
Number of hit DOMs	≥ 7	1.05 %
Radial distance	$< 200\,\mathrm{m}$	0.09 %
Vertical position	$-495\,\mathrm{m} < z < -225\,\mathrm{m}$	5.48 %
Energy	$5\,\mathrm{GeV} < E < 100\,\mathrm{GeV}$	20.70 %
Cosine of zenith angle	< 0.04	19.66 %
Number of direct hits	> 2.5	10.50 %
Number of hits in top layers	< 0.5	0.03 %
Number of hits in outer layer	< 7.5	0.001 %
Muon classifier score	≥ 0.8	23.90 %

Table 2.1: Cuts performed to select the final analysis sample. Parts of the cuts are meant to increase the data/MC agreement, while others are meant to reject events with poor reconstruction quality.

and MC is good. To remove coincident muon and neutrino events, cuts are applied to the number of hits in the top 15 layers of IceCube DOMs and the number of hits in the outermost IceCube strings. Coincident random noise events are removed by requiring more than three hit DOMs from direct photons[3]. Neither of the two coincident event types are simulated, which can be seen as bad agreement between data and MC. The last cut is on the reconstructed cosine zenith, which is required to be smaller than 0.04 to reject down-going muons.

3: *Direct photons* are photons that were not scattered on their way from the interaction vertex to the DOM.

Search for Tau Neutrino Induced Heavy Neutral Lepton Events

3

This chapter describes the search for HNL events using 10 years of IceCube DeepCore data. The expected number of HNL events in the data sample depends on the mass of the additional heavy state, m_4, and the mixing element $|U^2_{\alpha 4}|$, with $\alpha = e, \mu, \tau$, between the SM flavors and the new mass state. As discussed in Section **??**, this work focuses on the mixing to the tau sector, $|U^2_{\tau 4}|$, which has the weakest constraints to date. Since the mass itself influences the production and decay kinematics of the event and the accessible decay modes, individual mass samples were produced as described in Section 1.2. The mass influences the decay length and energy distributions, while the mixing both changes the overall expected rate of the HNL events and the shape in energy and length. We perform three independent searches for each mass sample, where the mixing is measured in each of the fits.

work in the parts about MC re-weighting (forward folding) etc. from below (RED)

3.1 Final Level Sample

The final level simulation sample of this analysis consists of the neutrino and muon MC introduced in Section 1.3 and one of the three HNL samples explained in Section 1.2, while the data are the events measured in 10 years of IceCube DeepCore data taking. All simulation and the data are processed through the full event selection chain described in Section 2.1 and Section 2.2 leading to the final level sample. As described in Section 2.2.2, event triggers consisting purely of random coincidences induced by noise in the DOMs have been reduced to a negligible rate, and will not be discussed further.

To get the neutrino expectation, the MC events are weighted according to their generation weight introduced in Section 1.3.1, multiplied by the total lifetime, and the expected neutrino flux. For the correct expectation at the detector, the events have to be weighted by the oscillation probability, depending on their energy and their distance traveled from the atmosphere to the detector. The oscillation probabilities are calculated using a PYTHON implementation of the calculations from [45], which use the matter profile of the Earth following the *Preliminary Reference Earth Model (PREM)* [46] as input. Apart from the energy and the distance, the two relevant parameters defining the oscillation probabilities are the atmospheric neutrino oscillation parameters θ_{23} and Δm^2_{31}. Since the HNL events originate from the tau neutrinos that were produced as muon neutrinos in the atmosphere and then oscillated into ν_τ, this weighting is also applied in addition to the specific weighting scheme for the HNL events described in Section 1.2.3, which itself is defined by the mixing $|U^2_{\tau 4}|$ and the mass m_4.

[45]: Barger et al. (1980), "Matter effects on three-neutrino oscillations"
[46]: Dziewonski et al. (1981), "Preliminary reference Earth model"

3.1.1 Expected Rates/Events

The rates and the expected number of events for the SM background are shown in Table 3.1 with around 175000 total events expected in the 10 years. Only data marked as good is used for the analysis, where *good* refers to

measurement time with the correct physics run configuration and without other known issues. The resulting good detector livetime in this data taking period was 9.28 years. The rates are calculated by summing the weights of all events in the final level sample, while the uncertainties are calculated by taking the square root of the sum of the weights squared. The expected number of events is calculated by multiplying the rate with the livetime. The individual fractions show that this sample is neutrino dominated where the majority of events are ν_μ-CC events.

Table 3.1: Final level rates and event expectation of the SM background particle types.

Type	**Rate** [mHz]	**Events** (9.28 years)	**Fraction** [%]
ν_μ^{CC}	0.3531	103321 ± 113	58.9
ν_e^{CC}	0.1418	41490 ± 69	23.7
ν^{NC}	0.0666	19491 ± 47	11.1
ν_τ^{CC}	0.0345	10094 ± 22	5.8
μ_{atm}	0.0032	936 ± 15	0.5
total	0.5992	175332 ± 143	100.0

Table 3.2 shows the rates and expected number of events for the HNL signal simulation. The expectation depends on the mass and the mixing and shown here are two example mixings for all the three masses that are being tested in this work. A mixing of 0.0 would result in no HNL events at all. It can already be seen that for the smaller mixing of $|U_{\tau 4}|^2 = 10^{-3}$ the expected number of events is very low, while at the larger mixing of $|U_{\tau 4}|^2 = 10^{-1}$ the number is comparable to the amount of atmospheric muons in the background sample.

Table 3.2: Final level rates and event expectations of the HNL signal for all three masses and two example mixing values.

HNL mass	**Rate** [μHz]	**Events (in** 9.28 years**)**
$\lvert U_{\tau 4}\rvert^2 = 10^{-1}$		
0.3 GeV	3.3	975 ± 2
0.6 GeV	3.1	895 ± 2
1.0 GeV	2.5	731 ± 2
$\lvert U_{\tau 4}\rvert^2 = 10^{-3}$		
0.3 GeV	0.006	1.67 ± 0.01
0.6 GeV	0.022	6.44 ± 0.01
1.0 GeV	0.025	7.27 ± 0.01

3.1.2 Analysis Binning

[41]: Yu et al. (2023), "Recent neutrino oscillation result with the IceCube experiment"

Add fractions of the different particle types in the bins for benchmark mass/mixing (another table?) (ORANGE)

An identical binning to the analysis performed in [41] is used. In total, there are three bins in PID (cascade like, mixed, and track like), 12 bins in reconstructed energy, and 8 bins in cosine of the reconstructed zenith angle as specified in Table 3.3. Extending the binning towards lower energies or

Table 3.3: Three-dimensional binning used in the analysis. All variables are from the PLEReCNN reconstruction explained in Section 2.2.

Variable	N_{bins}	**Edges**	**Spacing**
P_ν	3	[0.00, 0.25, 0.55, 1.00]	linear
E	12	[5.00, 100.00]	logarithmic
$\cos(\theta)$	8	[-1.00, 0.04]	linear

increasing the number of bins in energy or cosine of the zenith angle did not improve the HNL sensitivities significantly, because the dominant signal region is already covered with a sufficiently fine binning to observe the shape and magnitude of the HNL events on top of the SM background. This

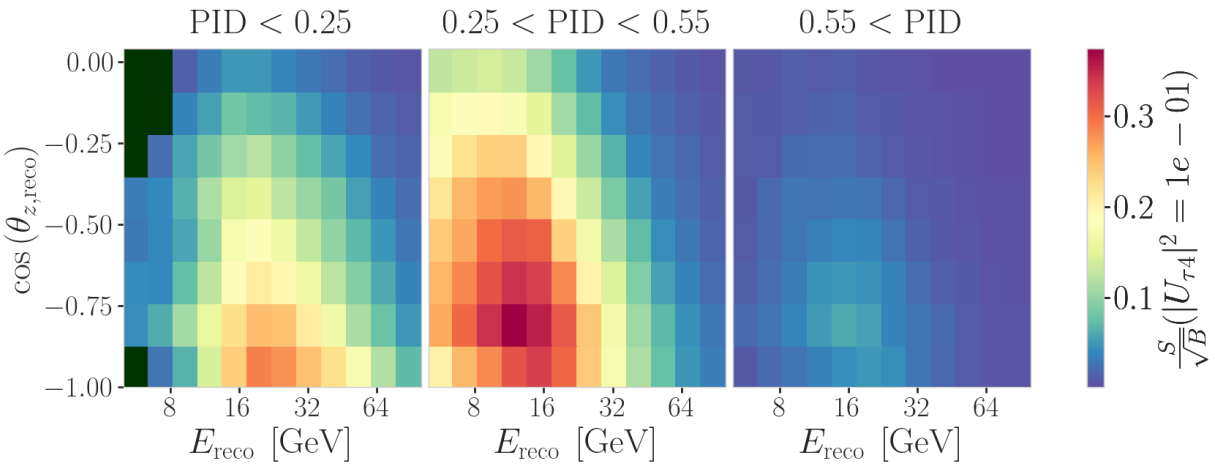


Figure 3.1: Signal over square root of background expectation in 9.28 years for the 1.0 GeV mass sample at a mixing of 0.1, while all other parameters are at their nominal values.

can be seen in the middle panel of Figure 3.1, which shows the expected signal events divided by the square root of the expected background events for every bin used in the analysis. The signal expectation is using the 1.0 GeV mass sample at a reference mixing of 0.1, with the corresponding three dimensional histogram shown in Figure B.1. Both the nominal background expectation used to calculate the signal to square root of background ratio and the detector data can be seen in Figure 3.2.

Some low energy bins in the cascade like region have very low MC expectations (<1 event) and are therefore not taken into account in the analysis, to prevent unwanted behavior in the fit. Those are shown in dark green in the three dimensional histograms, and both background and data histograms show a strong decrease of events towards low energies in the cascade like bin. This background expectation is not necessarily supposed to agree with the data, because this is the distributions assuming nominal parameter values, before performing the fit to find the parameters that describe the data best. All parameters used in the analysis are discussed in Section 3.2.2, and post-fit data to MC comparisons are shown in Section 3.3.3.

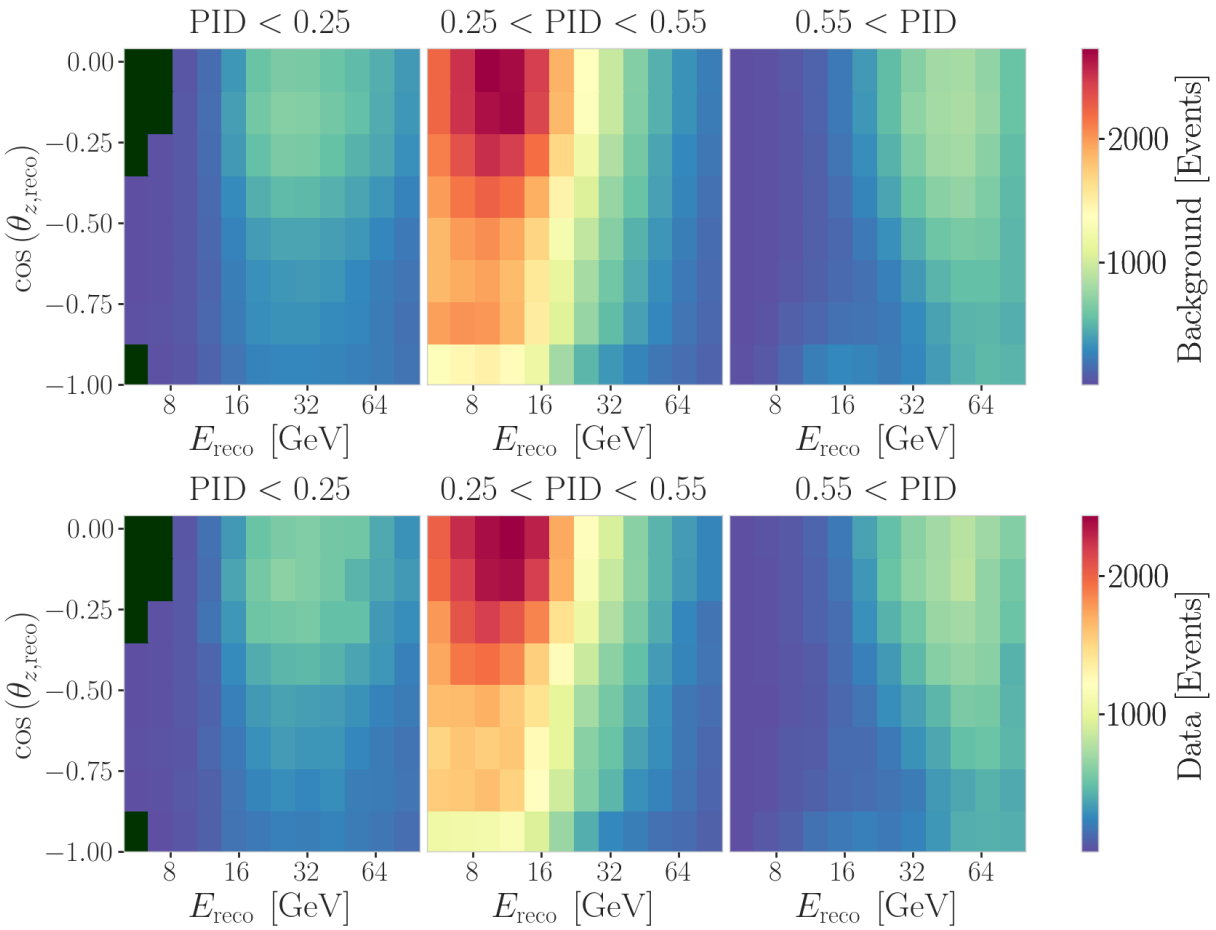


Figure 3.2: Background expectation in 9.28 years for all other parameters are at their nominal values (top) and observed data (bottom).

3.2 Statistical Analysis

3.2.1 Test Statistic

The measurements are performed by comparing the weighted MC to the data. Through variation of the nuisance and physics parameters that govern the weights, the best matching set of parameters can be found, by optimizing a fit metric. The comparison is done using a modified χ^2, defined as

$$\chi^2_{\text{mod}} = \sum_{i \in \text{bins}} \frac{(N_i^{\text{exp}} - N_i^{\text{obs}})^2}{N_i^{\text{exp}} + (\sigma_i^{\nu})^2 + (\sigma_i^{\mu})^2 + (\sigma_i^{\text{HNL}})^2} + \sum_{j \in \text{syst}} \frac{(s_j - \hat{s}_j)^2}{\sigma_{s_j}^2}, \tag{3.1}$$

as the fit metric. It is designed such that taking the difference between a free fit and a fit with fixed parameters based on a chosen hypothesis, $\Delta\chi^2_{\text{mod}}$, can directly be used as a *test statistic (TS)* for hypothesis testing, due to its asymptotic behavior. The total even expectation is $N_i^{\text{exp}} = N_i^{\nu} + N_i^{\mu} + N_i^{\text{HNL}}$, where N_i^{ν}, N_i^{μ}, and N_i^{HNL} are the expected number of events in bin i from neutrinos, atmospheric muons, and HNLs, while N_i^{obs} is the observed number of events in the bin. The expected number of events from each particle type is calculated by summing the weights of all events in the bin $N_i^{\text{type}} = \sum_i^{\text{type}} \omega_i$, with the statistical uncertainty being $(\sigma_i^{\text{type}})^2 = \sum_i^{\text{type}} \omega_i^2$. The additional term in Equation 3.1 is included to apply a penalty term for

prior knowledge of the systematic uncertainties of the parameters where they are known. s_j are the systematic parameters that are varied in the fit, while $\hat{s_j}$ are their nominal values and σ_{s_j} are the known uncertainties.

3.2.2 Physics Parameters

The variable physics parameter in this analysis is the mixing between the HNL and the SM τ sector, $|U_{\tau 4}|^2$. It is varied continuously in the range of [0.0 , 1.0] by applying the weighting scheme described in Section 1.2.3. The fit is initialized at an off-nominal value of 0.1. The other physics parameter, the mass m_4 of the HNL, is implicitly fixed to one of the three discrete masses to be tested, by using the corresponding sample of the HNL simulation described in Section 1.2.

3.2.3 Nuisance Parameters

There are multiple sources of systematic uncertainties related to the event generation and detector simulation explained in Chapter 1. All uncertainties considered in this work need to be implemented with parameters that can be varied continuously so that a simultaneous fit of the physics and systematic parameters can be performed. Where possible, a correct model of the effect is used, but in many cases the variations are captured by effective parameters. Uncertainties that solely scale the total event rate are not included individually, since the analysis only uses the relative distribution of events and a single scaling parameter N_ν is used to scale the total neutrino rate instead.

Atmospheric Flux Uncertainties

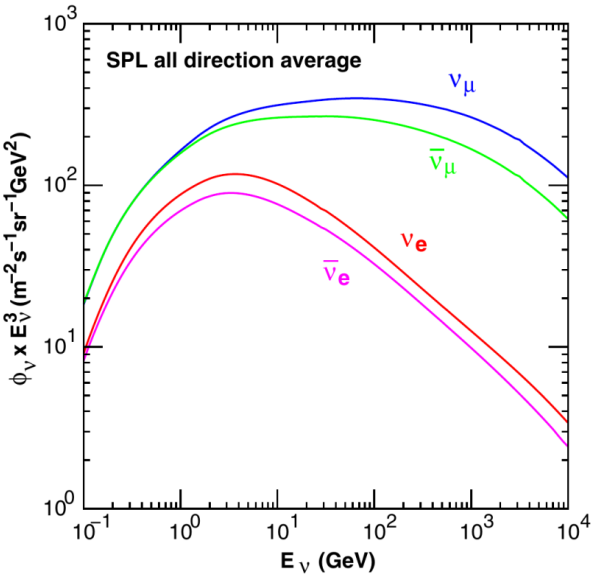


Figure 3.3: Atmospheric neutriono flux computed at the South Pole. Shown are the neutrino and antineutrino flux for ν_e and ν_μ. Taken from [15].

The flux of atmospheric neutrinos is influenced by multiple factors, the spectrum and composition of primary CRs, the atmospheric conditions, and the hadronic interaction model used to describe the air showers development. Uncertainties of the neutrino flux are therefore dictated by the uncertainties

on these components, where the variations in atmospheric conditions were found to have negligible effect [34]. The baseline neutrino flux used in this thesis is taken from [15]. Figure 3.3 shows the flux for neutrinos and antineutrinos, computed at the South Pole.

[34]: Abbasi et al. (2023), "Measurement of atmospheric neutrino mixing with improved IceCube DeepCore calibration and data processing"

[15]: Honda et al. (2015), "Atmospheric neutrino flux calculation using the NRLMSISE-00 atmospheric model"

Cosmic ray flux: The selected sample of atmospheric neutrinos lies around energies of up to 100 GeV. The initial primary particles in the CR flux can have 100 times larger energies and therefore the CR flux between 10 GeV and 10 TeV is important, which dominantly consists of hydrogen and helium nuclei [47]. The uncertainty in this CR flux component can be described as a power law correction [48, 49]

$$\Phi'_\nu = \Phi_\nu \left(\frac{E}{E^\star} \right)^{\Delta\gamma} , \tag{3.2}$$

where $E^\star$ is the pivot energy and $\Delta\gamma$ is the correction to the power law exponent. This modification propagates into the neutrino flux, which is therefore corrected in the same way. $E^\star$ was chosen to be 24 GeV as to minimize the dependence of the overall flux scale on $\Delta\gamma$ [34].

[47]: Dembinski et al. (2017), "Data-driven model of the cosmic-ray flux and mass composition from 10 GeV to 10^{11} GeV"

[48]: Barr et al. (2006), "Uncertainties in atmospheric neutrino fluxes"

[49]: Evans et al. (2017), "Uncertainties in atmospheric muon-neutrino fluxes arising from cosmic-ray primaries"

[34]: Abbasi et al. (2023), "Measurement of atmospheric neutrino mixing with improved IceCube DeepCore calibration and data processing"

Hadronic interaction model: Neutrinos are produced from the decaying hadrons (dominantly pions and kaons) in CR air shower, spanning a large parameter space that is sparsely evaluated by experimental data. To include uncertainties based on energy, direction, and neutrino flavor, the MCEq package [50] is used to compute the distribution of atmospheric leptons and to estimate the impact of varying their contributions. The calculations result in the change in flux $\mathrm{d}\Phi_\mathrm{l}/\mathrm{d}B$ for a variation $\mathrm{d}B$ of some parameter B. Scaling this variation by some value b, the modified total flux, s is then given by

$$\Phi'_\mathrm{l} = \Phi_\mathrm{l} + \left(b \cdot \frac{\mathrm{d}\Phi_\mathrm{l}}{\mathrm{d}B} \right) . \tag{3.3}$$

Matching the work in [51], the parameter space is divided in regions of the primary energy, E_i, and the energy fraction of the secondary meson, x_lab, with varying uncertainties, derived from fixed target experiment data. The Sibyll2.3c [52] hadronic interaction model and the GSF CR flux [47] were used to calculate the related flux changes[1] for the different regions in E_i and x_lab, resulting in 17 variables, encoding the possible changes. Figure 3.4 shows the selected regions of the parameter space and the names given to the uncertainties. At the energies relevant for this work, the flux is dominantly affected by the pion uncertainties. The variational term in Equation 3.3 is applied for each of these parameters and the total variation is the sum of all individual variations.

[51]: Barr et al. (2006), "Uncertainties in Atmospheric Neutrino Fluxes"

[52]: Riehn et al. (2020), "Hadronic interaction model sibyll 2.3d and extensive air showers"

[47]: Dembinski et al. (2017), "Data-driven model of the cosmic-ray flux and mass composition from 10 GeV to 10^{11} GeV"

1: The choice of flux and hadronic interaction model have negligible impact on the variations.

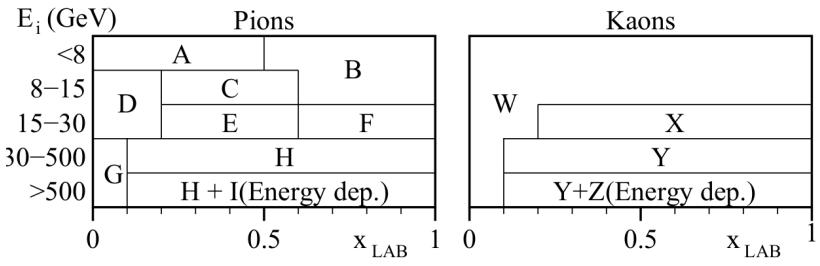


Figure 3.4: Flux uncertainty regions of the hadronic interaction model in the phase space of the primary energy E_i and the energy fraction of the secondary meson x_lab. Taken from [51].

Cross-Section Uncertainties

The uncertainties related to the cross-sections are split into low and high energy components, since there is no coherent model to explain both regimes. Below 20 GeV, *charged current resonance production (CCRES)* and *charged current quasi elastic scattering (CCQE)* interactions with the nucleons as a whole are important, while above 20 GeV DIS interactions are the dominant processes. Three parameters are included to account for all relevant cross-section uncertainties.

At low energies two parameters are included to account for uncertainties in form factors of CCQE and CCRES events. These uncertainties are due to uncertainties in the *axial mass* M_A, which enters the form factor as in

$$F(Q^2) \sim \frac{1}{(1-(\frac{Q}{M_A})^2)^2}\,, \tag{3.4}$$

where Q^2 is the momentum transfer squared. The axial mass can be determined experimentally and to include uncertainties on the values of M_A^{CCQE} and M_A^{CCRES}, the cross-sections are computed for each event, where the form factors are calculated varying the axial mass by $\pm 20\%(1\sigma)/\pm 40\%(1\sigma)$ around the nominal value. This is an approximation of the recommended uncertainties by the GENIE collaboration, which are -15%, $+25\%$ for M_A^{CCQE} and $\pm 20\%$ for M_A^{CCRES} [16]. To apply a continuous uncertainty variation of the axial mass in a fit, the total cross-section is fit with a quadratic function to interpolate between the cross-sections computed with the different axial masses.

which experiments measure the axial mass? (ORANGE)

Even though the DIS interactions can be calculated very precisely, there are still uncertainties in the input PDF, describing the probability of finding a specific parton (quark) with a specific momentum fraction x inside a nucleon. To account for differences between the used method and more sophisticated methods using newer PDFs seen at high energies, an uncertainty parameter is introduced. The parameter is based on the discrepancy between the cross-sections computed with GENIE and the ones computed with CSMS [53] above 100 GeV. The included parameter scales the cross-section from the GENIE values to the CSMS values, which are considered more accurate above 100 GeV. The scaling is done as a function of energy and inelasticity and to guarantee continuity, the scaling is extrapolated linearly below 100 GeV[2]. The parameter is designed such that a value of 0.0 corresponds to the GENIE cross-sections and a value of 1.0 gives an approximation of the CSMS cross-sections. A comparison of the total cross-sections GENIE (scaled/unscaled) with the data is shown in Figure 3.5.

[53]: Cooper-Sarkar et al. (2011), "The high energy neutrino cross-section in the Standard Model and its uncertainty"

2: Multiple functional extrapolations were tested, but the choice was found to have negligible impact.

Muon Uncertainties

The muon fraction in the final level selection (see Section 2.2.2) is below 1 %, therefore additional muon systematic uncertainties apart from the spectral index are not implemented, but rather a total muon scaling parameter is added. This total scale is somewhat degenerate with the DOM efficiency, since an increased DOM efficiency leads to better muon rejection. Both the total muon scaling and the muon spectral index have a very small impact on the analysis as will be shown in Section 3.

cite this? (YELLOW)

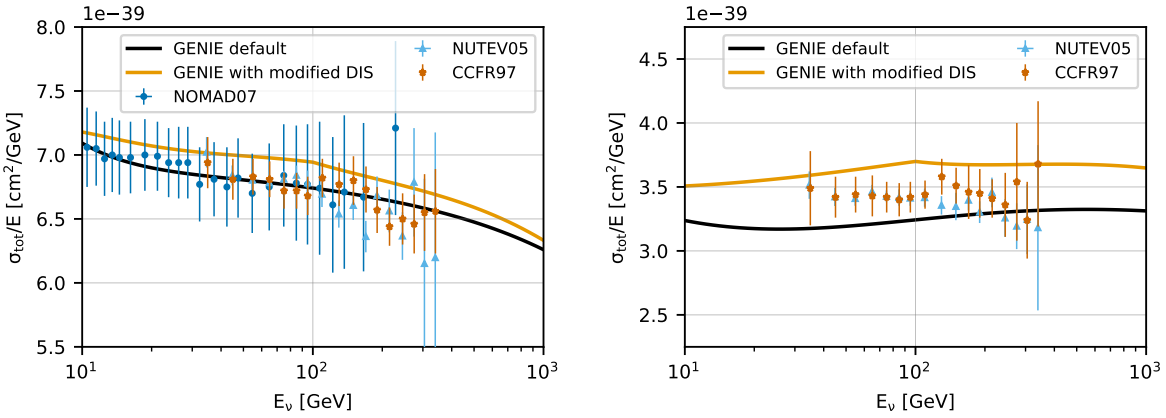


Figure 3.5: Inclusive total neutrino-nucleon cross-sections on an isoscalar target (black) for neutrinos (left) and antineutrinos (right) calculated with GENIE, comparing to measurements from NOMAD [54], NUTEV [55], and CCFR [56]. The scaled GENIE cross-section (orange) is also shown. Taken from [34].

Detector Calibration Uncertainties

The detection process of neutrinos in IceCube has several sources of uncertainties, where the effects of the properties of the ice itself and the optical efficiency of the DOMs are dominant for this analysis. None of these uncertainties can be described by an analytic expression, so instead their effects are estimated using MC simulation. This is done by producing additional simulation samples at discrete values of those parameters. The five relevant uncertainty parameters are the absolute efficiency of the DOMs, a global scaling of ice scattering and absorption lengths, and variations of the relative angular acceptance due to hole ice variations in two parameters. To perform the fit, continuous variations with respect to these parameters will be derived with a method explained in Section 3.

Relative DOM efficiency: As was already mentioned in Section **??**, the absolute efficiency of the DOMs, ϵ_{DOM} is calibrated using minimum ionizing muons from air showers, due to the lack of a calibrated light source in the detector. Using the muons as a steady, controlled source of light, the efficiency can be estimated by comparing simulated muon data sets with varied DOM response to the measured data. Since the uncertainties found in multiple iterations of this study [57, 58] are at the order of 10 %, this systematic is highly relevant and is included in the analysis.

[57]: Feintzeig (2014), "Searches for Point-like Sources of Astrophysical Neutrinos with the IceCube Neutrino Observatory"
[58]: Kulacz (2019), "In Situ Measurement of the IceCube DOM Efficiency Factor Using Atmospheric Minimum Ionizing Muons"

Ice scattering and absorption: Absorption and scattering length are the most important properties that govern the propagation of photons through the ice. The simulation principle and how the depth dependent absorption and scattering coefficients are used was already explained in Section 1.4.1. To account for uncertainties on this model of the ice coefficients, a global scaling for each of the two parameters (global absoprtion, global scatteriong) is applied.

Hole ice angular acceptance: Due to bubble formation in the re-freezing process of the boreholes, the hole ice seems to be less transparent in the center of the columns [59]. This effectively decreases the chance of photons

[59]: Rongen, Martin (2016), "Measuring the optical properties of IceCube drill holes"

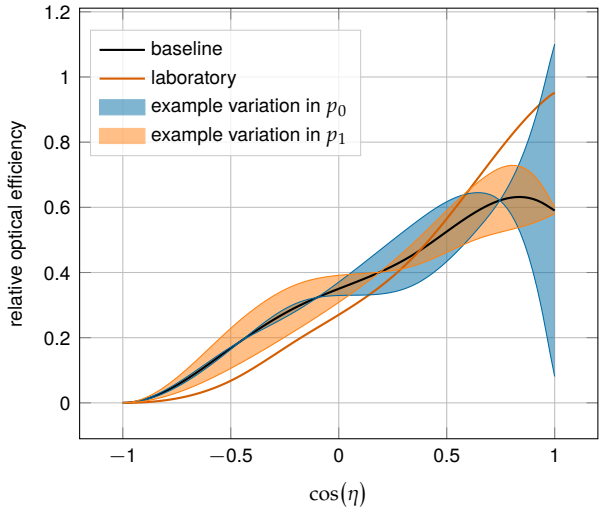


Figure 3.6: Relative angular acceptance modification due to hole ice. Shown is the current baseline model, the variations from changing p_0 and p_1, and a laboratory measurement. Modified from [26].

hitting the DOMs directly from below, which can be described as an additional angular modification of the DOM acceptance. The modification is parameterized by a two dimensional, normalized[3] function, where the two dominant of the parameters (p_0, p_1), dictating its form, are enough to describe all past and the current hole ice models from both *in-situ* and laboratory measurements. Figure 3.6 shows the acceptance modification as a function of the incident photon angle $\cos(\eta)$. The current baseline model, the variations achieved through modifying p_0 and p_1, and a laboratory measurement can be seen.

3: The hole ice angular acceptance modification is normalized so that it does not affect the total charge.

Ice Model: The ice model used in IceCube is continuously improved, and the recent models incorporate the birefringent polycriystalline microstructure [60] into the ice properties. To account for the uncertainty, due to this unmodeled effect in the ice model used for the simulation production, an additional simulation sample is produced using the newer version of the ice model, that incorporates the *birefringence (BFR)* effect.

[60]: Abbasi et al. (2024), "In situ estimation of ice crystal properties at the South Pole using LED calibration data from the IceCube Neutrino Observatory"

Treatment of Detector Systematic Uncertainties Since the variations related to the detector calibration uncertainties introduced in Section 2 are estimated by simulating MC at discrete values of the systematic parameters, a method to derive continuous variations is needed to perform the fit. The method applied here was initially introduced in [61] and first used in the low energy sterile neutrino search in [26] (section 7.4.3). Using a *likelihood-free inference* technique, re-weighting factors are found for every event in the nominal MC sample, given a specific choice of detector systematic parameters. These factors quantify how much more or less likely the event would be for the corresponding change in detector response from the nominal parameters. Without going into the details of the method, which were already exhaustively discussed in [61] and [26], the performance is assessed here for the HNL signal simulation. In order to do so, the weights are applied to the nominal MC samples, choosing the detector systematic values used to produce the discrete samples and the resulting event expectations are compared to the expectations from the individual, discrete MC samples. The

[61]: Fischer et al. (2023), "Treating detector systematics via a likelihood free inference method"
[26]: Trettin (2023), "Search for eV-scale sterile neutrinos with IceCube DeepCore"

bin counts are compared by calculating the pull defined as

$$p = \frac{N_{\text{reweighted}} - N_{sys}}{\sqrt{\sigma^2_{\text{reweighted}} + \sigma^2_{sys}}}\,, \tag{3.5}$$

where N are the bin-wise event expectations and σ are their MC uncertainty. For the SM BG simulation, the performance was already investigated in [62] (section 7.4.4, appendix B5) and the re-weighted nominal MC was shown to be in agreement with the discrete systematic sets at a sufficient level. Figure **??** shows the bin-wise pulls for the 1.0 GeV HNL mass sample at a mixing of 0.1 for a selection of the discrete systematic samples, where the DOM efficiency and the ice absorption was varied by ±10%. As expected, the pull distributions follow a standard normal distribution, without strong clustering or any systematic deviations. A similar performance is found for the additional systematic variations and the detailed figures can be found in Section B.2.

[62]: Lohfink (2023), "Testing non-standard neutrino interaction parameters with IceCube-DeepCore"

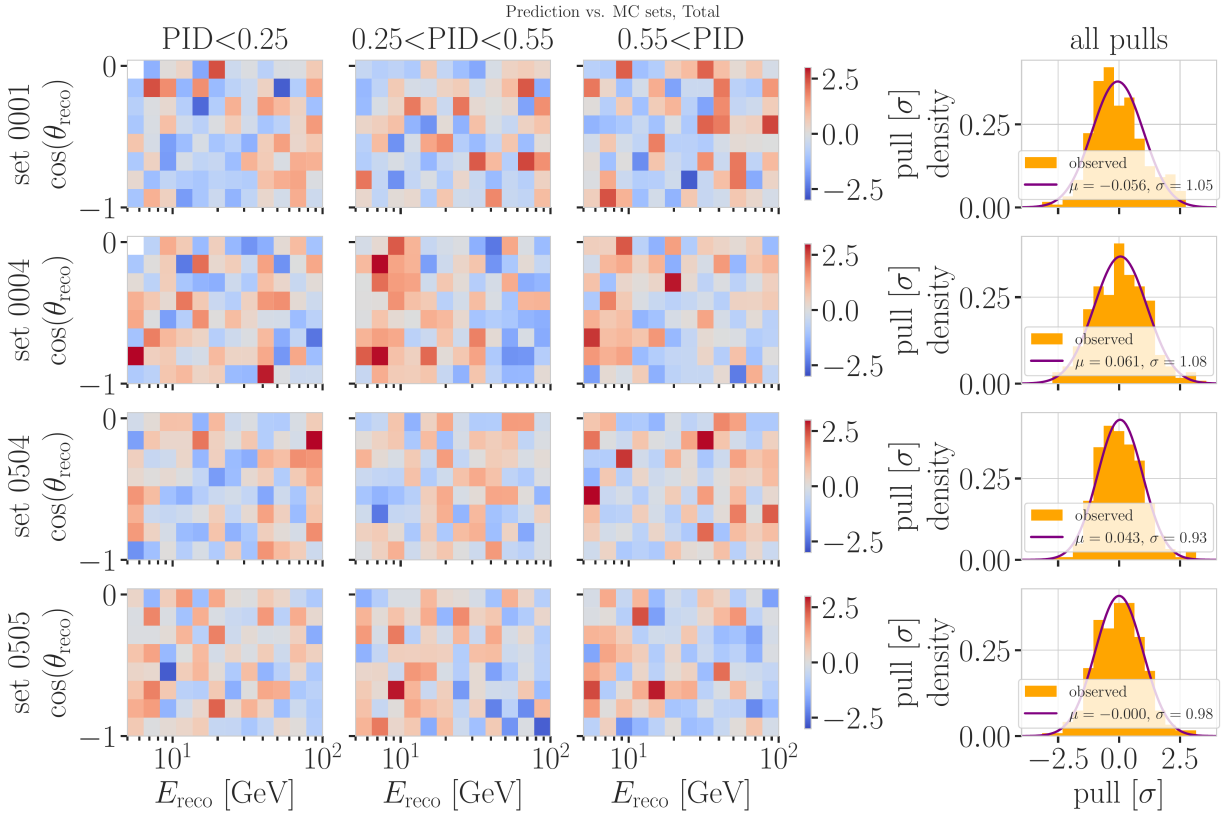
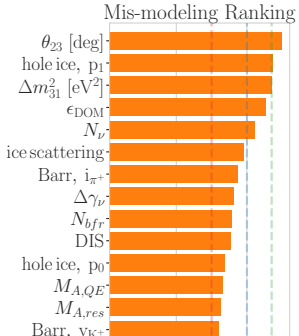


Figure 3.7: Three dimensional pulls and set-wise pull distributions between the nominal set and the specific systematic sets, after the nominal set was re-weighted to the correspoinding systematic parameter value.

fix caption (add description of what specific set number means (RED))

Free Parameters

To decide which systematic uncertainties should be included in the fit, we test the potential impact they have on the TS if they are neglected. The test is performed by creating Asimov data using the BG simulation and the HNL simulation of the 1.0 GeV mass sample at a mixing value of 0.1, which is chosen as a benchmark physics parameter, but the explicit choice does not have a significant impact on the test. The systematic parameter of interest



is set to a value above its nominal expectation, either pulled up by $+1\sigma$ or by an educated estimate for parameters without a well-defined uncertainty. A fit is performed fixing the systematic parameter of interest and leaving all additional parameters free. The resulting TS is the fit metric difference between between this fit and a fit with all parameters free, which would result in a fit metric of 0.0 for this Asimov test. This difference is called mis-modeling significance and parameters below a significance of 0.1 σ are fixed. The test is performed in an iterative manner until the final set of free parameters is found.

Figure 3.8 shows the resulting significances of one of these tests. The parameters tested are the systematic parameters introduced in Section 3.2.3 and the atmospheric oscillation parameters mentioned in Section 3.1. In the final selection of free parameters the Barr h_{π^+} parameter was also left free, to sufficiently cover the relevant energy production range of the Pions, as can be seen in Figure 3.4, where both for Kaons and Pions the uncertainties are included for primary energies above 30 GeV and $x_{\text{lab}} > 0.1$. Additionally, the ice absorption is still kept free, despite showing a small significance, which is done because the ice parameters are not well constrained and are known to have a large impact, which might be concealed in this idealized test, due to correlations with the other parameters. In this test, the effect of correlations is challenging to consider, because only the impact of one parameter is tested at a time, using the overall mis-modeling significance as a measure. The mis-modeling could be reduced by a correlated parameter capturing the effect of the parameter of interest. For this reason a very conservative threshold of 0.1 σ is chosen and some parameters below the threshold are still left free in the fit.

All nuisance parameters that are left free in the fit are summarized in Table 3.4, showing their nominal values, the allowed fit ranges, and their Gaussian prior, if applicable. The scaling parameter N_ν is included to account for the overall normalization of the neutrino rate, and it has the identical effect on the SM neutrino events and the BSM HNL events, because they both originate from the same neutrino flux. Despite being known to ~5 % in this energy range, there is no prior applied to this parameter, because the fit itself is able to constrain it well, which can be seen by the large impact it shows in Figure 3.8. Concerning the atmospheric neutrino flux, the CR power law flux correction factor $\Delta\gamma_\nu$ introduced in Section 3.2.3 is included with nominal value of 0.0 which corresponds to the baseline flux model. A slightly conservative prior of 0.1 is applied to the parameter, while latest measurements show an uncertainty of 0.05 [49]. The Barr parameters are contrained by a Gaussian prior, taken from [51]. All the detector systematic uncertainties discussed in Section 2 are included in the fit. The DOM efficiency ϵ_{DOM} is constrained by a Gaussian prior with a width of 0.1, which is a conservative estimate based on the studies of the optical efficiency using minimum ionizing muons from [57, 58]. The two atmospheric neutrino oscillation parameters θ_{23} and Δm^2_{31} are also included in the fit with nominal values of 47.5° and 2.48×10^{-3} eV^2 [41], respectively. Since they govern the shape and the strength of the tau neutrino flux, by defining the oscillation from ν_μ to ν_τ, they are also relevant for the HNL signal shape.

Need cite here! (RED)

I could add some final level effects of some systematics on the 3D binning and maybe discuss how they are different from the signal shape, or so? (ORANGE)

3.2.4 Low Energy Analysis Framework

The analysis is performed using the PISA [63] [64] software framework,

[49]: Evans et al. (2017), "Uncertainties in atmospheric muon-neutrino fluxes arising from cosmic-ray primaries"

[51]: Barr et al. (2006), "Uncertainties in Atmospheric Neutrino Fluxes"

[57]: Feintzeig (2014), "Searches for Point-like Sources of Astrophysical Neutrinos with the IceCube Neutrino Observatory"

[58]: Kulacz (2019), "In Situ Measurement of the IceCube DOM Efficiency Factor Using Atmospheric Minimum Ionizing Muons"

[41]: Yu et al. (2023), "Recent neutrino oscillation result with the IceCube experiment"

[63]: Aartsen et al. (2020), "Computational techniques for the analysis of small signals in high-statistics neutrino oscillation experiments"

Table 3.4: Systematic uncertainty parameters that are left free to float in the fit. Their allowed fit ranges are shown with the nominal value and the Gaussian prior width if applicable.

Parameter	Nominal	Range	Prior
$\theta_{23}[°]$	47.5047	[0.0, 90.0]	-
$\Delta m^2_{31}[\text{eV}^2]$	0.002475	[0.001, 0.004]	-
N_ν	1.0	[0.1, 2.0]	-
$\Delta\gamma_\nu$	0.0	[-0.5, 0.5]	0.1
Barr h_{π^+}	0.0	[-0.75, 0.75]	0.15
Barr i_{π^+}	0.0	[-3.05, 3.05]	0.61
Barr y_{K^+}	0.0	[-1.5, 1.5]	0.3
DIS	0.0	[-0.5, 1.5]	1.0
$M_{\text{A,QE}}$	0.0	[-2.0, 2.0]	1.0
$M_{\text{A,res}}$	0.0	[-2.0, 2.0]	1.0
ϵ_{DOM}	1.0	[0.8, 1.2]	0.1
hole ice p_0	0.101569	[-0.6, 0.5]	-
hole ice p_1	-0.049344	[-0.2, 0.2]	-
ice absorption	1.0	[0.85, 1.15]	-
ice scattering	1.05	[0.9, 1.2]	-
N_{bfr}	0.0	[-0.2, 1.2]	-

which was developed to perform analyses of small signals in high-statistics neutrino oscillation experiments. It is used to generate the expected event distributions from several MC samples, which can then be compared to the observed data. The expectation for each MC sample is calculated by applying physics and nuisance parameter effects in a stage-wise manner, before combining them to the final expectation.

3.3 Analysis Checks

Fitting to data is performed in a *blind* manner, where the analyzer does not immediately see the fitted physics and nuisance parameter values, but first checks that a set of pre-defined *goodness of fit (GOF)* criteria are fulfilled. This is done to circumvent the so-called *confirmation bias* [65], where the analyzer might be tempted to construct the analysis in a way that confirms their expectation. After the GOF criteria are met to satisfaction, the fit results are unblinded and the full result can be revealed. Before these blind fits to data are performed, the robustness of the analysis method is tested using pseudo-data that is generated from the MC.

3.3.1 Minimization Robustness

To find the set of parameters that best describes the data, a staged minimization routine is used. In the first stage, a fit with coarse minimizer settings is performed to find a rough estimate of the *best fit point (BFP)*. In the second stage, the fit is performed again in both octants[4] of θ_{23}, starting from the BFP of the coarse fit. For each individual fit the *MIGRAD* routine of IMINUIT [66] is used to minimize the χ^2_{mod} fit metric defined in Equation 3.1. Iminuit is a fast, python compatible minimizer based on the MINUIT2 C++ library [67]. The individual minimizer settings for both stages are shown in Table 3.5.

Fit	Err.	Prec.	Tol.
Coarse	1e-1	1e-8	1e-1
Fine	1e-5	1e-14	1e-5

Table 3.5: Migrad settings for the two stages in the minimization routine. *Err.* are the step size for the numerical gradient estimation, *Prec.* is the precision with which the LLH is calculated, and *Tol.* is the tolerance for the minimization.

To test the minimization routine and to make sure it consistently recovers any physics parameters, pseudo-data sets are produced from the MC by choosing the nominal nuisance parameters and specific physics parameters, without adding any statistical or systematic fluctuations to it. These so-called *Asimov*[5]

FInd first occurance of "Asimov" and add reference and explain it there (RED)

[65]: Nickerson (1998), "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises"

4: There is a degeneracy between the lower octant ($\theta_{23} < 45°$) and the upper octant ($\theta_{23} > 45°$), which can lead to fit metric minima (local and global) at two positions that are mirrored around 45° in θ_{23}.

[66]: Dembinski et al. (2022), *scikit-hep/iminuit: v2.17.0*

[67]: James et al. (1975), "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations"

5: A pseudo-data set without statistical fluctuations is called Asimov data set.

data sets are then fit back with the full analysis chain. This type of test is called *Asimov inject/recover test*. A set of mixing values between 10^{-3} and 10^{0} is injected and fit back. Without fluctuations the fit is expected to always recover the injected parameters (both physics and nuisance parameters). The fitted mixing values from the Asimov inject/recover tests are compared to the true injected values in Figure 3.9 for all three mass samples. As desired, the fit is always able to recover the injected physics parameter and the nuisance parameters within the statistical uncertainty or at an insignificant fit metric difference.

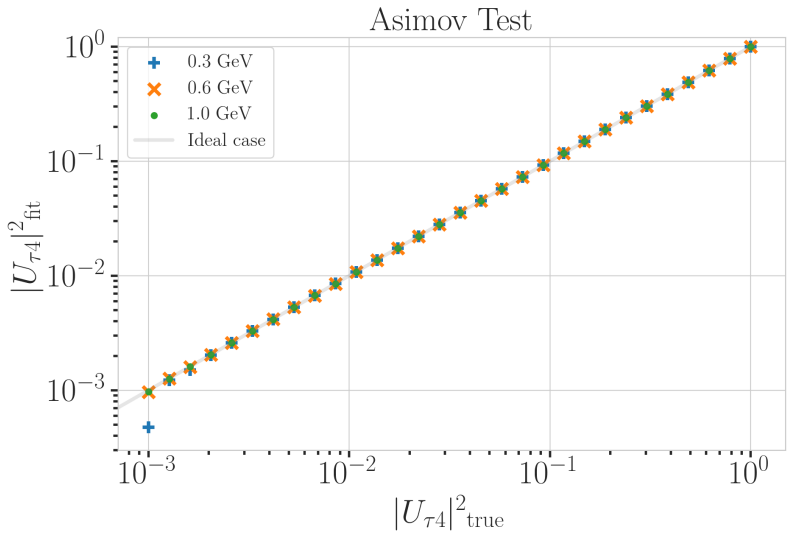


Figure 3.9: Asimov inject/recover test results for all three mass samples. Mixing values between 10^{-3} and 10^{0} are injected and fit back with the full analysis chain. The injected parameter is always recovered within the statistical uncertainty or at an insignificant fit metric difference.

3.3.2 Goodness of Fit

To estimate the GOF, pseudo-data is generated from the MC by injecting the BFP parameters as true parameters and then fluctuating the expected bin counts to account for MC uncertainty and Poisson fluctuations in data. First, the expectation value of each bin is drawn from a Gaussian distribution centered at the nominal expectation value with a standard deviation corresponding to the MC uncertainty of the bin. Based on this sampled expectation value, each bin count is drawn from a Poisson distribution, independently, to get the final pseudo-data set. These pseudo-data sets are analyzed with the same analysis chain as the real data, resulting in a final fit metric value for each pseudo-data set. By comparing the distribution of fit metric values from this *ensemble* of pseudo-data trials to the fit metric of the fit to real data, a p-value can be calculated. The p-value is the probability of finding a value of the fit metric at least as large as the one from the data fit. Figure 3.10 shows the distribution from the ensemble tests for the 0.6 GeV mass sample and the observed value from the fit, resulting in a p-value of 28.5 %. The p-values for the 0.3 GeV and 1.0 GeV are 28.3 % and 26.0 %, respectively, and the corresponding plots are shown in Section **??**. Based on this test, it is concluded that the fit result is compatible with the expectation from the ensemble of pseudo-data trials.

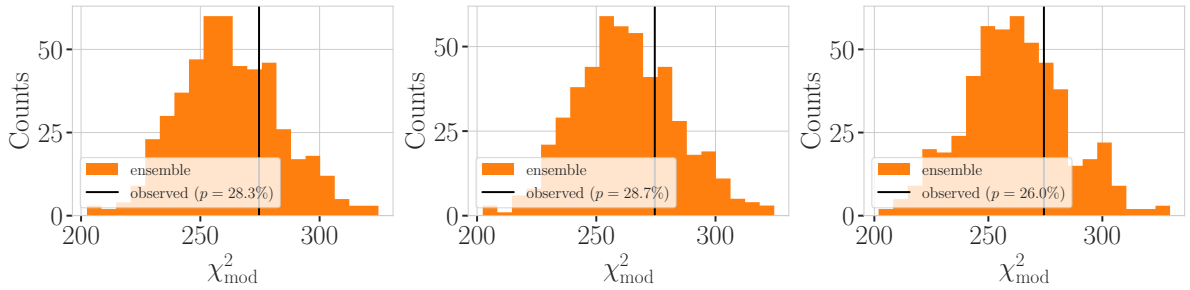


Figure 3.10: Observed fit metric (data fit) and fit metric distribution from pseudo-data ensemble generated around the best fit point. Shown are the results for all three mass samples, with the ensemble distribution on orange, the observed value in black, and the p-value in the legend.

3.3.3 Data/MC Agreement

At the BFP, the agreement between the data and simulation is probed by comparing both the one dimensional analysis distributions for PID, energy, and cosine of the zenith angle as well as the full three dimensional distributions. Figure 3.11 shows the three dimensional pull distribution between data and the total MC expectation for the 0.6 GeV mass sample at the BFP. The pulls are evenly spaced and show no strong clustering. In Figure **??**, two examples of the one dimensional distributions for the energy and cosine of the zenith angle are shown for the 0.6 GeV mass sample. The data is compared to the total MC expectation, which is also split up into the individual signal and background components for illustration. Good agreement can be observed in the pull distributions, and is quantified by a reduced χ^2, which is close to 1.0 for all distributions. The reduced χ^2 for all investigated distributions is listed in Table **??**, while the distributions themselves can be found in Section **??**.

add 1-d data/mc agreement for example mass sample (0.6?) and all 3 analysis variables (RED)

add table with reduced chi2 for all 1-d distributions (RED)

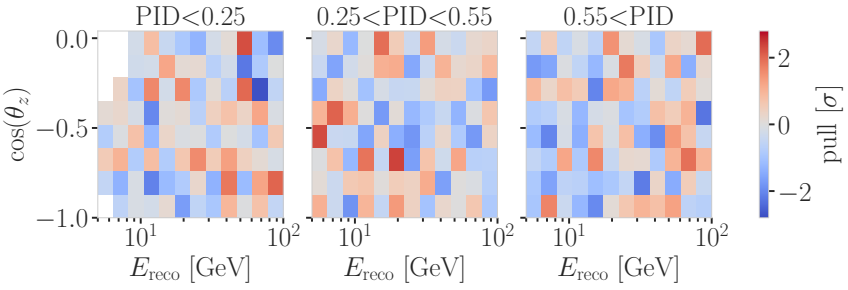


Figure 3.11

fix caption and describe plot in text (RED)

3.4 Results

3.4.1 Best Fit Nuisance Parameters

The resulting nuisance parameter values from the fits are illustrated in Figure 3.12, where the differences to the nominal values are shown, normalized by the distance to the closest boundary. The results from all three fits are shown in the same plot and the fits prefer values of the same size for all three mass samples. For parameters that have a Gaussian prior, the 1 σ range is also displayed. As was already confirmed during the blind fit procedure, all fitted parameters are within this range. The effective ice model parameter, N_{bfr}, prefers a value of ~0.74 , indicating that the data fits better to an ice model that includes real birefringence effects [60]. For completeness, the

[60]: Abbasi et al. (2024), "In situ estimation of ice crystal properties at the South Pole using LED calibration data from the IceCube Neutrino Observatory"

explicit results are listed in Table B.1. There, the nominal values and the absolute differences to the best fit value are also presented.

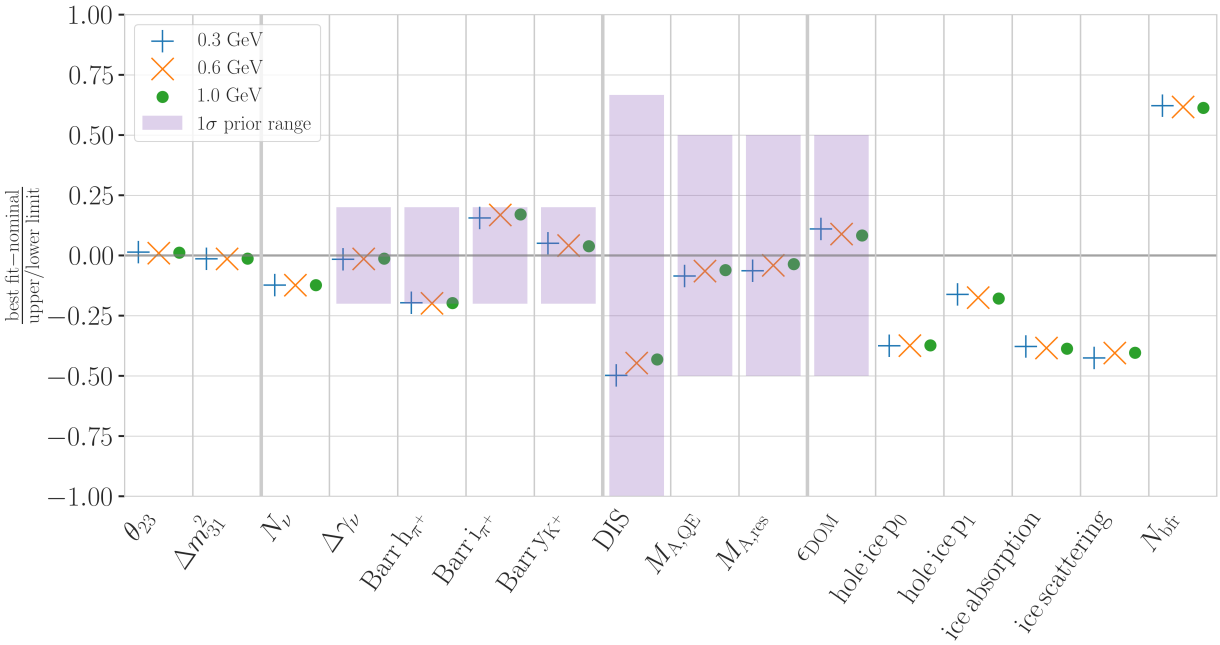


Figure 3.12: Best fit nuisance parameter distances to the nominal values, normalized by the distance to the closest boundary. For parameters with a Gaussian prior, the +1 σ range is also shown.

Show best fit hole ice angular acceptance compared to nominal and flasher/in-situ fits, maybe? (YELLOW)

3.4.2 Agreement with Standard Model Three-Flavor Oscillation Measurement

The recently performed atmospheric neutrino oscillation measurement by the IceCube collaboration resulted in a best fit point of $\sin^2\theta_{23} = 0.544^{+0.030}_{-0.096}$ and $\Delta m^2_{32} = 2.40^{+0.03}_{-0.06} \times 10^{-3}\text{eV}^2$ [41]. The result used the identical 10 years of data at the same final level selection, assuming normal mass ordering. The differences to this analysis are the choice of fit metric and the use of the previous treatment of detector systematic uncertainties. Both the choice of the fit metric and the detector systematic uncertainty treatment should not influence the best fit values and a naive comparison of the results to the results from this work is done, to validate this is the case. The best fit values are listed in Table 3.6 and are all compatible with the IceCube result within the uncertainties. Since they are statistically fully dependent, a more rigorous quantitative comparison would be more involved and is not performed here. The agreement is still interpreted as a first validation of the new detector systematics treatment.

m_4	$\sin^2\theta_{23}$	Δm^2_{32} [eV^2]
0.3 GeV	0.554	0.0238
0.6 GeV	0.551	0.0238
1.0 GeV	0.553	0.0238

Table 3.6: Best fit oscillation parameters from the three mass sample fits. The values are compatible with the IceCube result within the uncertainties.

[41]: Yu et al. (2023), "Recent neutrino oscillation result with the IceCube experiment"

3.4.3 Best Fit Parameters and Limits

The fitted mixing values are

$$|U_{\tau 4}|^2(0.3\,\text{GeV}) = 0.003^{+0.084}\,,$$
$$|U_{\tau 4}|^2(0.6\,\text{GeV}) = 0.080^{+0.134}\,,\text{and}$$
$$|U_{\tau 4}|^2(1.0\,\text{GeV}) = 0.106^{+0.132}\,,$$

with their +1 σ uncertainty. All of them are compatible with the null hypothesis of 0.0 mixing, although the 0.6 GeV and 1.0 GeV fits indicate a mixing value of 0.08 and 0.106 , respectively. The best fit mixing values and the corresponding upper limits at 68 % and 90 % *confidence level (CL)* are listed in Table 3.7, also showing the p-value to reject the null hypothesis. The CLs and p-value are estimated by assuming that *Wilks' theorem* [68] holds, meaning that the TS follows a χ^2 distribution with one degree of freedom.

[68]: Wilks (1938), "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses"

Table 3.7: Best fit mixing values and the corresponding upper limits at 68 % and 90 % confidence level, as well as the p-value to reject the null hypothesis, estimated by assuming that Wilks' theorem holds.

HNL mass	$\lvert U_{\tau 4}\rvert^2$	68 % CL	90 % CL	NH p-value
0.3 GeV	0.003	0.09	0.19	0.97
0.6 GeV	0.080	0.21	0.36	0.79
1.0 GeV	0.106	0.24	0.40	0.63

Figure 3.13 shows the observed TS profiles as a function of $|U_{\tau 4}|^2$ for all three fits. The TS profile is the difference in χ^2_{mod} between the free fit and a fit where the mixing is fixed to a specific value. Also shown is the expected TS profile, based on 100 pseudo-data trials, produced at the BFP and then fluctuated using both Poisson and Gaussian fluctuations, to include the data and the MC uncertainty as was explained in Section 3.3.2. The Asimov expectation and the 68 % and 90 % bands are shown and the observed TS profiles lie within the 68 % band for all three, confirming that they are compatible with statistical fluctuations of the observed data. For the 0.3 GeV fit, the observed contour is slightly tighter than the Asimov expectation, meaning that the observed upper limits in $|U_{\tau 4}|^2$ are slightly stronger than expected. For the 0.6 GeV the opposite is the case and the observed upper limit is therefore slightly weaker than expected. For the 1.0 GeV fit, the observed upper limit is very close to the Asimov expectation in the region where the 68 % and 90 % CLs thresholds are crossed. The observed upper limits are also shown in Table 3.7.

discuss how this compares to other experiments (it's far, far below the existing limits..) (RED)

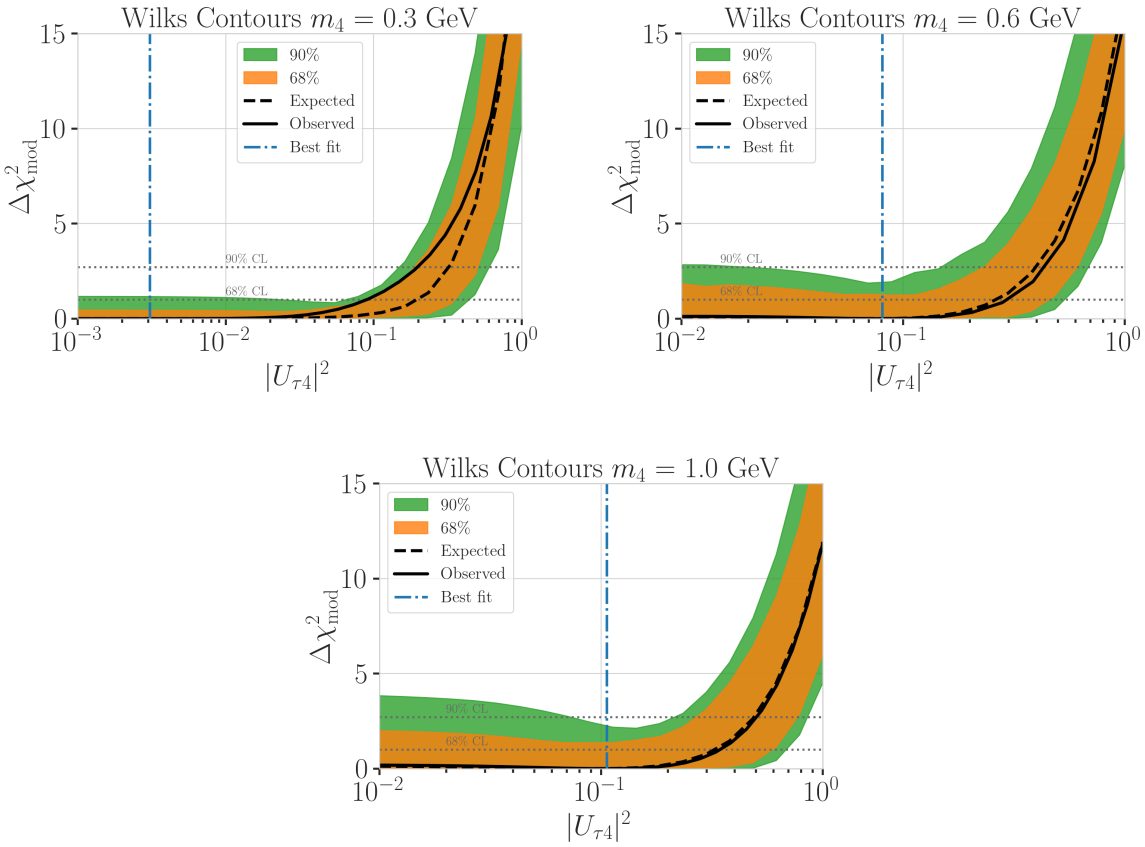


Figure 3.13: Best fit point TS profiles as a function of $|U_{\tau 4}|^2$ for the 0.3 GeV, 0.6 GeV, and 1.0 GeV mass samples. Shown are the observed profiles, the Asimov expectation at the best fit point, and the 68 % and 90 % bands, based on 100 pseudo-data trials. Also indicated are the 68 % and 90 % CL levels assuming Wilks' theorem.

3.5 Outlook

3.5.1 Analysis Improvements

- estimate full contribution from cascade only events (underestimated due to limited sampling distributions)
- include double cascade classifier into Binning
- further optimize binning

Test Coupling to Electron/Muon Flavor

Test Additional Coupling Processes

IceCube Upgrade

Appendix

A

Heavy Neutral Lepton Event Generation

A.1 Model-Independent Simulation Distributions

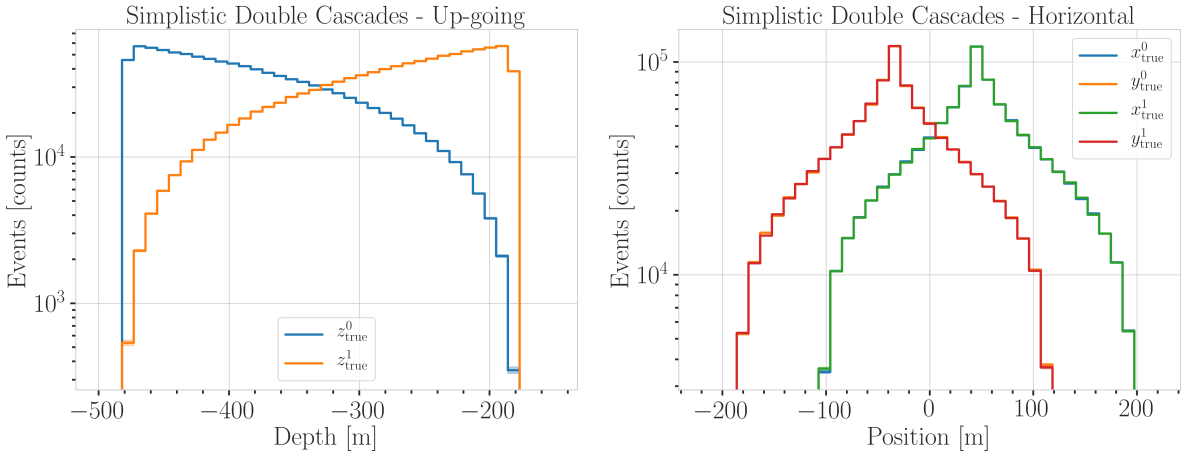


Figure A.1: Generation level distributions of the simplistic simulation sets. Vertical positions of the cascades in the up-going sample (left) and horizontal positions of the cascades in the horizontal sample (right).

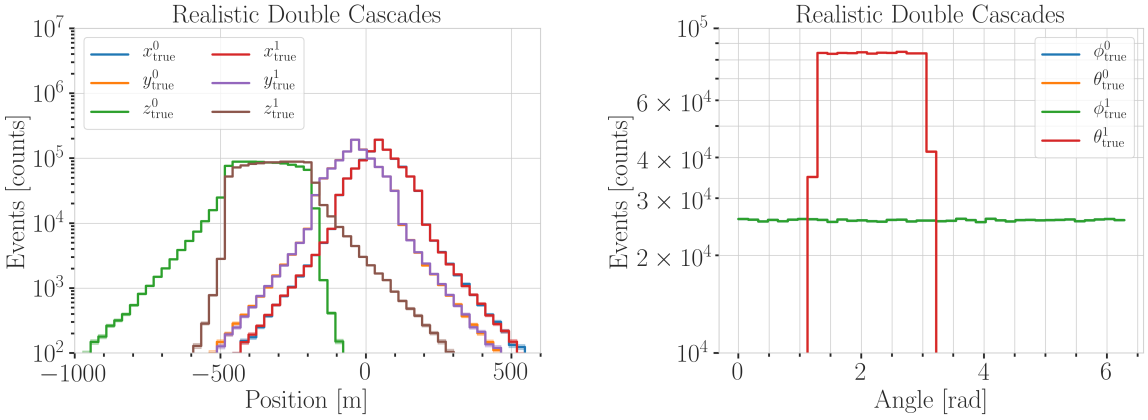
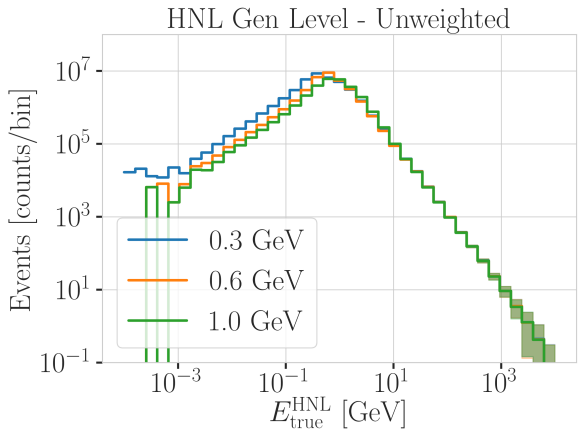
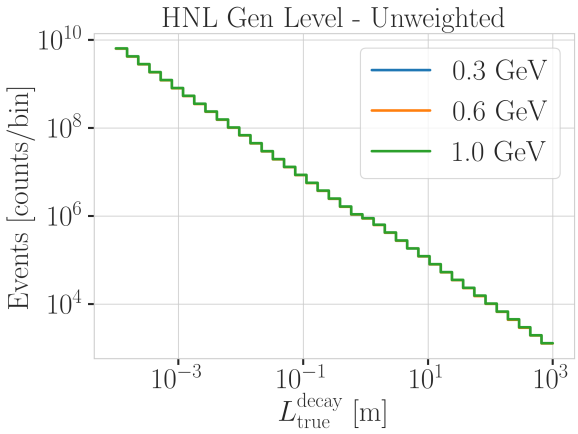


Figure A.2: Generation level distributions of the realistic simulation set. Shown are the cascade x, y, z positions (left) and the cascade direction angles (right).

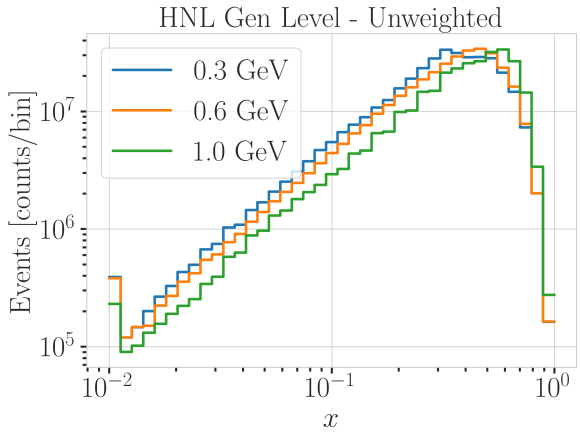
A.2 Model-Dependent Simulation Distributions



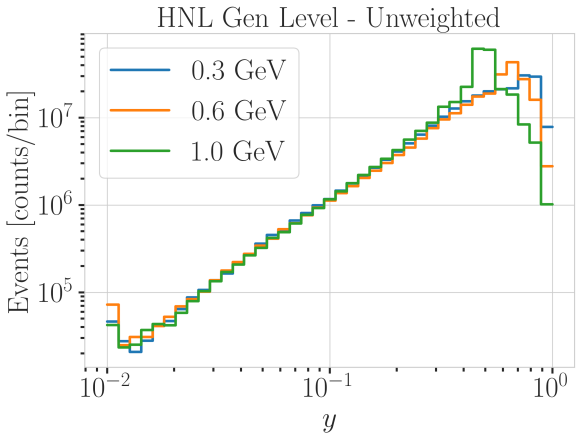
(a) HNL energy



(b) Decay length



(c) Bjorken x



(d) Bjorken y

Figure A.3: Generation level distributions of the model dependent simulation.

B

Analysis Results

B.1 Final Level Simulation Distributions

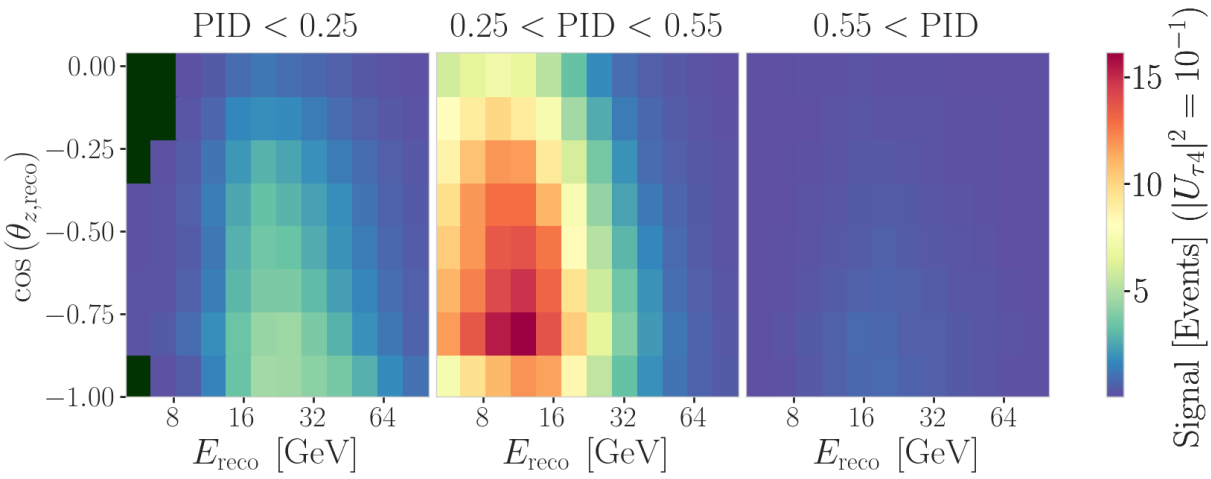


Figure B.1: Signal expectation in 9.28 years for the 1.0 GeV mass sample at a mixing of 0.1, while all other parameters are at their nominal values (top) and observed data (bottom).

B.2 Treatment of Detector Systematic Uncertainties

add description of what specific set number means (RED)

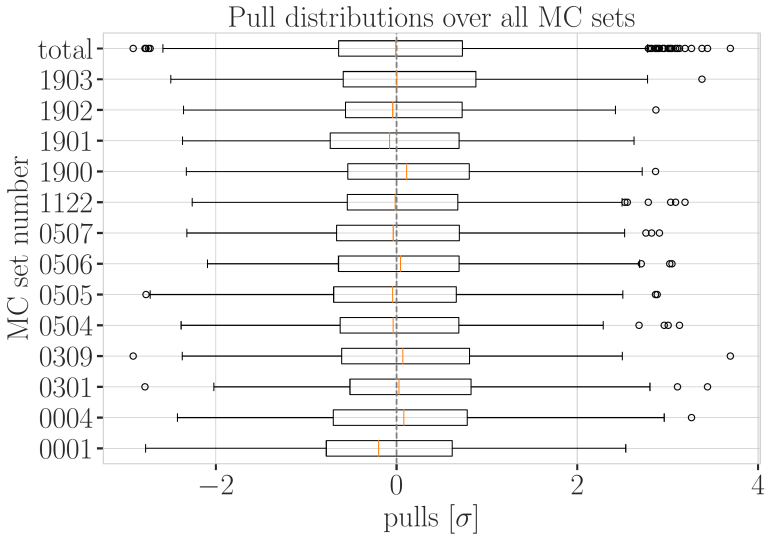


Figure B.2: Overall performance of the detector systematic uncertainty treatment. Shown are the pull distributions of the three dimensional pulls shown in Figure 3.7 and Figure B.3 between the nominal set and the specific systematic set, after the nominal set was re-weighted to the correspoinding systematic parameter value.

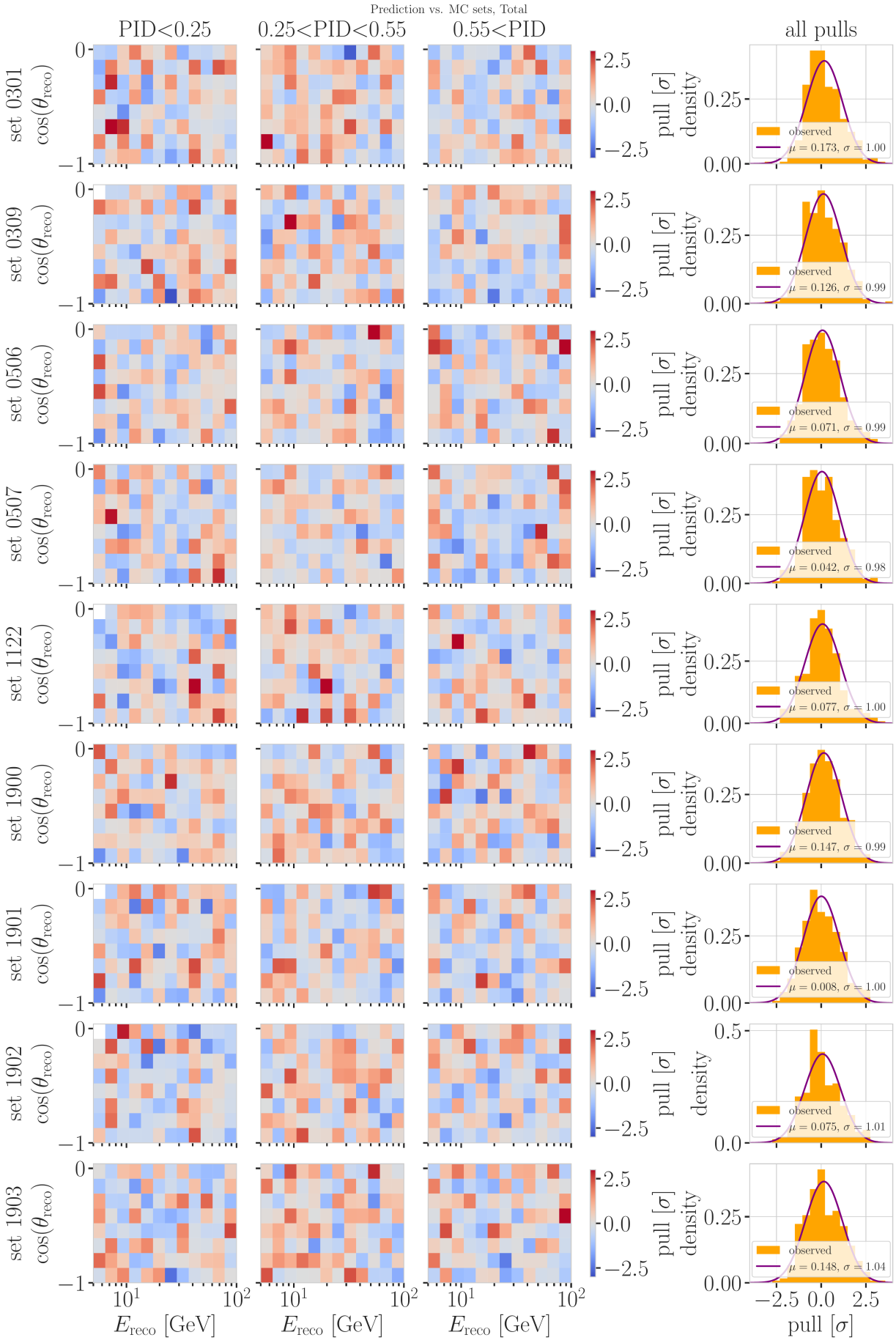


Figure B.3: Three dimensional pulls and set-wise pull distributions between the nominal set and the specific systematic sets, after the nominal set was re-weighted to the correspoinding systematic parameter value.

B.3 Best Fit Nuisance Parameters

Table B.1: Best fit nuisance parameters for the three mass samples. Also shown is the nominal value and the difference between the nominal and the best fit.

Parameter	Nominal	Best Fit			Nominal - Best Fit		
		0.3 GeV	0.6 GeV	1.0 GeV	0.3 GeV	0.6 GeV	1.0 GeV
$\|U_{\tau 4}\|^2$	-	0.003019	0.080494	0.106141	-	-	-
$\theta_{23}[^\circ]$	47.5047	48.117185	47.918758	48.010986	−0.612485	−0.414058	−0.506286
$\Delta m^2_{31}[\mathrm{eV}^2]$	0.002475	0.002454	0.002454	0.002455	0.000020	0.000021	0.000019
N_ν	1.0	0.889149	0.889055	0.889559	0.110851	0.110945	0.110441
$\Delta\gamma_\nu$	0.0	−0.007926	−0.006692	−0.006596	0.007926	0.006692	0.006596
Barr h_{π^+}	0.0	−0.147475	−0.148481	−0.148059	0.147475	0.148481	0.148059
Barr i_{π^+}	0.0	0.475448	0.513393	0.521626	−0.475448	−0.513393	−0.521626
Barr y_{K^+}	0.0	0.076176	0.062893	0.057548	−0.076176	−0.062893	−0.057548
DIS	0.0	−0.248709	−0.223302	−0.215666	0.248709	0.223302	0.215666
$M_{\mathrm{A,QE}}$	0.0	−0.170528	−0.128150	−0.120345	0.170528	0.128150	0.120345
$M_{\mathrm{A,res}}$	0.0	−0.125855	−0.080875	−0.070716	0.125855	0.080875	0.070716
ϵ_{DOM}	1.0	1.021984	1.017789	1.016689	−0.021984	−0.017789	−0.016689
hole ice p_0	0.101569	−0.161341	−0.161051	−0.160129	0.262910	0.262620	0.261698
hole ice p_1	−0.049344	−0.073701	−0.075596	−0.076261	0.024357	0.026252	0.026917
ice absorption	1.00	0.943261	0.942463	0.942000	0.056739	0.057537	0.058000
ice scattering	1.05	0.986152	0.989289	0.989438	0.063848	0.060711	0.060562
N_{bfr}	0.0	0.746684	0.740255	0.736215	−0.746684	−0.740255	−0.736215

fix design + significant digits to show (ORANGE)

maybe show range/prior and then deviation in sigma, or absolute for the ones without prior

Bibliography

Here are the references in citation order.

[1] L. Fischer. https://github.com/LeanderFischer/icetray_double_cascade_generator_functions (cited on page 1).

[2] R. Abbasi et al. "LeptonInjector and LeptonWeighter: A neutrino event generator and weighter for neutrino observatories". In: *Comput. Phys. Commun.* 266 (2021), p. 108018. DOI: 10.1016/j.cpc.2021.108018 (cited on page 4).

[3] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nucl. Instr. Meth. Phys. Res.* 506.3 (July 2003), pp. 250–303. DOI: 10.1016/s0168-9002(03)01368-8 (cited on page 5).

[4] L. Rädel and C. 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 (Oct. 2012), pp. 53–67. DOI: 10.1016/j.astropartphys.2012.09.008 (cited on page 5).

[5] P. Coloma et al. "GeV-scale neutrinos: interactions with mesons and DUNE sensitivity". In: *Eur. Phys. J. C* 81.1 (2021), p. 78. DOI: 10.1140/epjc/s10052-021-08861-y (cited on pages 5, 7, 9).

[6] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079 (cited on page 5).

[7] J.-H. Koehne et al. "PROPOSAL: A tool for propagation of charged leptons". In: *Computer Physics Communications* 184.9 (2013), pp. 2070–2090. DOI: https://doi.org/10.1016/j.cpc.2013.04.001 (cited on page 5).

[8] C. Argüelles. https://github.com/arguelles/NuXSSplMkr (cited on page 6).

[9] N. Whitehorn, J. van Santen, and S. Lafebre. "Penalized splines for smooth representation of high-dimensional Monte Carlo datasets". In: *Computer Physics Communications* 184.9 (2013), pp. 2214–2220. DOI: https://doi.org/10.1016/j.cpc.2013.04.008 (cited on page 6).

[10] J.-M. Levy. "Cross-section and polarization of neutrino-produced tau's made simple". In: *J. Phys. G* 36 (2009), p. 055002. DOI: 10.1088/0954-3899/36/5/055002 (cited on page 6).

[11] E. Tiesinga et al. "CODATA recommended values of the fundamental physical constants: 2018". In: *Rev. Mod. Phys.* 93 (2 July 2021), p. 025010. DOI: 10.1103/RevModPhys.93.025010 (cited on page 8).

[12] R. L. Workman et al. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2022.8 (Aug. 2022), p. 083C01. DOI: 10.1093/ptep/ptac097 (cited on page 8).

[13] https://launchpad.net/mg5amcnlo/3.0/3.3.x (cited on page 9).

[14] A. Alloul et al. "FeynRules 2.0 - A complete toolbox for tree-level phenomenology". In: *Comput. Phys. Commun.* 185 (2014), pp. 2250–2300. DOI: 10.1016/j.cpc.2014.04.012 (cited on page 9).

[15] M. Honda et al. "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 (cited on pages 11, 27, 28).

[16] C. Andreopoulos et al. "The GENIE Neutrino Monte Carlo Generator: Physics and User Manual". In: (2015) (cited on pages 11, 29).

[17] M. Glück, E. Reya, and A. Vogt. "Dynamical parton distributions revisited". In: *The European Physical Journal C* 5 (Sept. 1998), pp. 461–470. DOI: 10.1007/s100529800978 (cited on page 12).

[18] A. Bodek and U. K. Yang. "Higher twist, xi(omega) scaling, and effective LO PDFs for lepton scattering in the few GeV region". In: *Journal of Physics G: Nuclear and Particle Physics* 29.8 (2003), p. 1899. DOI: 10.1088/0954-3899/29/8/369 (cited on page 12).

[19] Y. Becherini et al. “A parameterisation of single and multiple muons in the deep water or ice”. In: *Astroparticle Physics* 25.1 (2006), pp. 1–13. DOI: https://doi.org/10.1016/j.astropartphys.2005.10.005 (cited on page 12).

[20] D. Heck et al. “CORSIKA: A Monte Carlo code to simulate extensive air showers”. In: (Feb. 1998) (cited on page 12).

[21] T. K. Gaisser. “Spectrum of cosmic-ray nucleons, kaon production, and the atmospheric muon charge ratio”. In: *Astropart. Phys.* 35 (2012), pp. 801–806. DOI: 10.1016/j.astropartphys.2012.02.010 (cited on page 12).

[22] R. Engel et al. “The hadronic interaction model Sibyll – past, present and future”. In: *EPJ Web Conf.* 145 (2017). Ed. by B. Pattison, p. 08001. DOI: 10.1051/epjconf/201614508001 (cited on page 12).

[23] C. Kopper et al. https://github.com/claudiok/clsim (cited on page 12).

[24] D. Chirkin et al. “Photon Propagation using GPUs by the IceCube Neutrino Observatory”. In: *2019 15th International Conference on eScience (eScience)*. 2019, pp. 388–393. DOI: 10.1109/eScience.2019.00050 (cited on page 12).

[25] M. G. Aartsen et al. “Measurement of South Pole ice transparency with the IceCube LED calibration system”. In: *Nucl. Instrum. Meth. A* 711 (2013), pp. 73–89. DOI: 10.1016/j.nima.2013.01.054 (cited on page 13).

[26] A. Trettin. “Search for eV-scale sterile neutrinos with IceCube DeepCore”. PhD thesis. Berlin, Germany: Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2023. DOI: https://github.com/atrettin/PhD-Thesis (cited on pages 13, 31).

[27] G. Mie. “Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen”. In: *Annalen der Physik* 330.3 (1908), pp. 377–445. DOI: https://doi.org/10.1002/andp.19083300302 (cited on page 13).

[28] L. G. Henyey and J. L. Greenstein. “Diffuse radiation in the Galaxy.” In: *apj* 93 (Jan. 1941), pp. 70–83. DOI: 10.1086/144246 (cited on page 13).

[29] R. Abbasi 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 (cited on page 13).

[30] S. Fiedlschuster. “The Effect of Hole Ice on the Propagation and Detection of Light in IceCube”. In: (Apr. 2019) (cited on page 13).

[31] M. G. Aartsen et al. “In-situ calibration of the single-photoelectron charge response of the IceCube photomultiplier tubes”. In: *Journal of Instrumentation* 15.6 (June 2020), P06032. DOI: 10.1088/1748-0221/15/06/P06032 (cited on pages 13, 14).

[32] M. Larson. “Simulation and Identification of Non-Poissonian Noise Triggers in the IceCube Neutrino Detector”. Available at https://ir.ua.edu/handle/123456789/1927. MA thesis. University of Alabama, Tuscaloosa, AL, USA, 2013 (cited on page 14).

[33] M. Larson. “A Search for Tau Neutrino Appearance with IceCube-DeepCore”. available at https://discoverycenter.nbi.ku.dk/teaching/thesis_page/mjlarson_thesis.pdf. PhD thesis. University of Copenhagen, Denmark, 2018 (cited on page 14).

[34] R. Abbasi et al. “Measurement of atmospheric neutrino mixing with improved IceCube DeepCore calibration and data processing”. In: *Phys. Rev. D* 108 (1 July 2023), p. 012014. DOI: 10.1103/PhysRevD.108.012014 (cited on pages 15, 18, 28, 30).

[35] M. G. Aartsen et al. “The IceCube Neutrino Observatory: Instrumentation and Online Systems”. In: *JINST* 12.03 (2017), P03012. DOI: 10.1088/1748-0221/12/03/P03012 (cited on page 15).

[36] R. Abbasi et al. “The design and performance of IceCube DeepCore”. In: *Astropart. Phys.* 35.10 (2012), pp. 615–624. DOI: 10.1016/j.astropartphys.2012.01.004 (cited on pages 15, 16).

[37] R. Abbasi et al. “The IceCube data acquisition system: Signal capture, digitization, and timestamping”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 601.3 (2009), pp. 294–316. DOI: https://doi.org/10.1016/j.nima.2009.01.001 (cited on page 16).

[38] M. G. Aartsen et al. "The IceCube Neutrino Observatory: instrumentation and online systems". In: *Journal of Instrumentation* 12.3 (Mar. 2017), P03012. DOI: 10.1088/1748-0221/12/03/P03012 (cited on page 16).

[39] J. H. Friedman. "Stochastic gradient boosting". In: *Computational Statistics & Data Analysis* 38 (2002), pp. 367–378 (cited on page 17).

[40] R. Abbasi et al. "Low energy event reconstruction in IceCube DeepCore". In: *Eur. Phys. J. C* 82.9 (2022), p. 807. DOI: 10.1140/epjc/s10052-022-10721-2 (cited on page 18).

[41] S. Yu and J. Micallef. "Recent neutrino oscillation result with the IceCube experiment". In: *38th International Cosmic Ray Conference*. July 2023 (cited on pages 18, 24, 33, 37).

[42] S. 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 (cited on pages 18, 19).

[43] J. Micallef. https://github.com/jessimic/LowEnergyNeuralNetwork (cited on page 18).

[44] M. Huennefeld. "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 (cited on page 19).

[45] V. Barger et al. "Matter effects on three-neutrino oscillations". In: *Phys. Rev. D* 22 (11 Dec. 1980), pp. 2718–2726. DOI: 10.1103/PhysRevD.22.2718 (cited on page 23).

[46] A. M. Dziewonski and D. L. Anderson. "Preliminary reference Earth model". In: *Physics of the Earth and Planetary Interiors* 25.4 (1981), pp. 297–356. DOI: https://doi.org/10.1016/0031-9201(81)90046-7 (cited on page 23).

[47] H. Dembinski et al. "Data-driven model of the cosmic-ray flux and mass composition from 10 GeV to 10^{11} GeV". In: *PoS* ICRC2017 (2017), p. 533. DOI: 10.22323/1.301.0533 (cited on page 28).

[48] G. D. Barr et al. "Uncertainties in atmospheric neutrino fluxes". In: *Phys. Rev. D* 74 (9 Nov. 2006), p. 094009. DOI: 10.1103/PhysRevD.74.094009 (cited on page 28).

[49] J. Evans et al. "Uncertainties in atmospheric muon-neutrino fluxes arising from cosmic-ray primaries". In: *Phys. Rev. D* 95 (2 Jan. 2017), p. 023012. DOI: 10.1103/PhysRevD.95.023012 (cited on pages 28, 33).

[50] A. Fedynitch et al. https://github.com/afedynitch/MCEq (cited on page 28).

[51] G. D. Barr et al. "Uncertainties in Atmospheric Neutrino Fluxes". In: *Phys. Rev. D* 74 (2006), p. 094009. DOI: 10.1103/PhysRevD.74.094009 (cited on pages 28, 33).

[52] F. Riehn et al. "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 (cited on page 28).

[53] A. Cooper-Sarkar, P. Mertsch, and S. Sarkar. "The high energy neutrino cross-section in the Standard Model and its uncertainty". In: *JHEP* 08 (2011), p. 042. DOI: 10.1007/JHEP08(2011)042 (cited on page 29).

[54] Q. Wu et al. "A precise measurement of the muon neutrino–nucleon inclusive charged current cross section off an isoscalar target in the energy range 2.5<E<40 GeV by NOMAD". In: *Physics Letters B* 660.1 (2008), pp. 19–25. DOI: https://doi.org/10.1016/j.physletb.2007.12.027 (cited on page 30).

[55] M. M. Tzanov. "Precise measurement of neutrino and anti-neutrino differential cross sections on iron". PhD thesis. University of Pittsburgh, Pennsylvania, Jan. 2005 (cited on page 30).

[56] W. G. Seligman. "A next-to-leading-order QCD analysis of neutrino-iron structure functions at the Tevatron". PhD thesis. Columbia University, New York, Aug. 1997 (cited on page 30).

[57] J. Feintzeig. "Searches for Point-like Sources of Astrophysical Neutrinos with the IceCube Neutrino Observatory". PhD thesis. University of Wisconsin, Madison, Jan. 2014 (cited on pages 30, 33).

[58] N. Kulacz. "In Situ Measurement of the IceCube DOM Efficiency Factor Using Atmospheric Minimum Ionizing Muons". MA thesis. University of Alberta, 2019 (cited on pages 30, 33).

[59] Rongen, Martin. "Measuring the optical properties of IceCube drill holes". In: *EPJ Web of Conferences* 116 (2016), p. 06011. DOI: 10.1051/epjconf/201611606011 (cited on page 30).

[60] R. Abbasi 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* 18.1 (2024), pp. 75–102. DOI: 10.5194/tc-18-75-2024 (cited on pages 31, 36).

[61] L. Fischer, R. Naab, and A. Trettin. “Treating detector systematics via a likelihood free inference method”. In: *Journal of Instrumentation* 18.10 (2023), P10019. DOI: 10.1088/1748-0221/18/10/P10019 (cited on page 31).

[62] E. Lohfink. “Testing nonstandard neutrino interaction parameters with IceCube-DeepCore”. PhD thesis. Mainz, Germany: Johannes Gutenberg-Universität Mainz, Fachbereich für Physik, Mathematik und Informatik, 2023. DOI: http://doi.org/10.25358/openscience-9288 (cited on page 32).

[63] M. G. Aartsen et al. “Computational techniques for the analysis of small signals in high-statistics neutrino oscillation experiments”. In: *Nucl. Instrum. Meth. A* 977 (2020), p. 164332. DOI: 10.1016/j.nima.2020.164332 (cited on page 33).

[64] https://github.com/icecube/pisa (cited on page 33).

[65] R. S. Nickerson. “Confirmation Bias: A Ubiquitous Phenomenon in Many Guises”. In: *Review of General Psychology* 2 (1998), pp. 175–220 (cited on page 34).

[66] H. Dembinski et al. *scikit-hep/iminuit: v2.17.0*. Version v2.17.0. Sept. 2022. DOI: 10.5281/zenodo.7115916 (cited on page 34).

[67] F. James and M. Roos. “Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations”. In: *Comput. Phys. Commun.* 10 (1975), pp. 343–367. DOI: 10.1016/0010-4655(75)90039-9 (cited on page 34).

[68] 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 (cited on page 38).