

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

Copyright Notice

This book is released into the public domain using the CC-BY-4.0 code.

To view a copy of the CC-BY-4.0 code, visit:

<https://creativecommons.org/licenses/by/4.0/>

Colophon

This document was typeset with the help of KOMA-Script and L^AT_EX using the open-source kaobook template class.

The source code of this thesis is available at:

https://github.com/LeanderFischer/phd_thesis

Zusammenfassung

Zusammenfassung ...

Abstract

Abstract ...

Todo list

Make my own DC string positions/distances plot version, viable for the margin?	1
Maybe just reference the one from the detector description and drop it here?	1
Re-make plot with all energies (cascades and total, both samples (they are the same))	2
Re-make plot with all decay lengths (both samples)	2
In which variables are they splined?	5
Re-make plot with 3 target masses and better labels/legends etc.	6
SB: emphasize the cut-off/suppression	6
Add comparisons of SM cross-sections between NuXSSplMkr and genie?	6
add varied total cross-section for a few background HNL events (for QE/RES variations?!?)	6
Calculate max BRs	7
JVS: consider also writing down the (trivial) 2-body decay kinematics for completeness and consistency. This transition is a bit jarring as it is	8
add table with number of gen level files? mention the event number is smaller because of kinematic condition?	9
Combine low/high plots and remove all traces of the separation in the thesis (tables/text/etc.)	10
Adapt chapter to reflect switched chapter order	11
put a number on this significant increase?	15
put a number on the tilt angle?	16
Add SPE distribution plot	17
AT: Das klingt so, als würde die MC Simulation ein analoges Signal erzeugen, was auch digitalisiert wird. Vllt kann man das nochmal nachforschen, aber zumindest in meiner Arbeit habe ich geschrieben, dass die MCPEs direkt in ATWD und fADC Readouts umgewandelt werden.	17
add example plots (?) for L3 cut variables and applied cuts	19
add some figure showing the corridors?	20
add table with rates per level (split in flavor) - maybe better in analysis chapter to also show signal? .	20
add image with selected strings used for flercnn IC and DC	21
add some performance plots of the FLERCNN reconstruction	22
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?	22
add reference for flercnn analysis internal note	22
add final sample composition, but maybe also in analysis chapter to show signal and background? .	23
at some place I will want a selection efficiency plot for SM BG and HNL signal, but I'm not sure where to put it yet	23
add figure with Barr blocks?	24
which experiments measure the axial mass?	24

cite this?	27
add information about the matter profile used	29
add information about the oscillation probability calculation and the software used for it	29
Should I adapt the total numbers to match the sum of the rounded individual parts?	30
add 3D expectation and/or S/sqrt(B) plots	30
Add fractions of the different particle types in the bins for benchmark mass/mixing (another table?) .	30
I feel like I have to be a bit more precise on what is the fit metric (e.g. the mod chi2) and what is the TS, as in the mod chi2 difference, which is the actual TS, right?	31
Do I want/need to include the description of the KDE muon estimation?	31
Blow up labels/legend/title and make it more readable in the margin	31
I don't like this formulation, but don't know better right now..	31
elaborate why this is also done to cover the whole energy range for the pion production, referencing the Barr Block plot that I haven't included yet :D	31
I truly dislike this sentence, too, better ideas?	32
I'm just writing out the data from the table, but I need to mention/motivate the included priors here and maybe just point to the table for the ranges/nominal values? (Not quite sure about this) . . .	32
cite?!	32
I think I will need to mention here that I did no inclcude MA-RES and MA-QE for the HNL simulation..	32
add final level effects of varying the axial mass parameters (or example of one)	32
add final level effects of varying the DIS parameter (or example of one)	32
fix caption!	32
Do I want additional plots for this (fit diff, LLH distr, minim. stats, param. fits)?	33
Add bin-wise TS distribution? Add 3D TS maps?	34
fix caption for this figure	35
sort the variables also by type, same as in the table "best_fit_parameters"?	35
Show best fit hole ice angular acceptance compared to nominal and flasher/in-siuite fits, maybe? . . .	35
Discuss what it means that the parameters are at these values? Here, or somewhere else?	35
fix table caption	35
fix once I have them produced	36
fix caption for this figure	36
make plot with BFPs and limit, comparing to upper limits from other experiments	36
specify which they are, once I have them	36
add 1-d data/mc agreement for example mass set (0.6?) and all 3 analysis variables	36
add table with reduced chi2 for all 1-d distributions	36
Re-make plot with x,y for horizontal set one plot!	39
Re-make plot with x, y, z for both cascades in one.	39
Re-arrange plots in a more sensible way.	39

sort these by type of nuisance parameter? 43

Contents

Abstract	iii
Contents	ix
1 Heavy Neutral Lepton Signal Simulation	1
1.1 Model Independent Simulation	1
1.1.1 Simplistic Samples	1
1.1.2 Realistic Sample	2
1.2 Model Dependent Simulation	4
1.2.1 Custom LeptonInjector	4
1.2.2 Sampling Distributions	8
1.2.3 Weighting Scheme	8
1.2.4 Generation Level Distributions	10
2 Standard Model Background Simulation and Data Processing	13
2.1 Event Generation	13
2.1.1 Neutrinos	14
2.1.2 Muons	15
2.2 Detector Simulation	15
2.2.1 Photon Propagation	15
2.2.2 Detector Responses	16
2.3 Processing	17
2.3.1 Trigger and Filter	17
2.3.2 Event Selection	18
2.4 Reconstruction	20
2.4.1 Fast Low Energy Reconstruction using Convolutional Neural Networks	21
2.4.2 Analysis Selection	22
2.5 Systematic Uncertainties	23
2.5.1 Atmospheric Flux Uncertainties	23
2.5.2 Cross-Section Uncertainties	24
2.5.3 Detector Calibration Uncertainties	25
2.5.4 Muon Uncertainties	26
2.5.5 Treatment of Detector Response Uncertainties via a Likelihood-Free Inference Method	27
3 Search for an Excess of Heavy Neutral Lepton Events	29
3.1 Final Level Sample	29
3.1.1 Expected Rates/Events	29
3.1.2 Analysis Binning	30
3.2 Statistical Analysis	30
3.2.1 Low Energy Analysis Framework	30
3.2.2 Test Statistic	31
3.2.3 Physics Parameters	31
3.2.4 Nuisance Parameters	31
3.3 Analysis Checks	33
3.3.1 Minimization Robustness	33
3.3.2 Ensemble Tests	33
3.4 Results	34
3.4.1 Best Fit Nuisance Parameters	34
3.4.2 Best Fit Parameters and Limits	35
3.4.3 Data/MC Agreement	36

3.4.4 Likelihood Coverage	36
APPENDIX	37
A Heavy Neutral Lepton Signal Simulation	39
A.1 Model Independent Simulation Distributions	39
A.2 Model Dependent Simulation Distributions	40
B Analysis Checks	41
B.1 Minimization Robustness	41
B.1.1 Ensemble Tests	41
C Analysis Results	43
C.0.1 Best Fit Nuisance Parameters	43
C.0.2 Best Fit Parameters and Limits	43
Bibliography	45

List of Figures

1.1 DeepCore string spacing	2
1.2 Simplified model independent simulation generation level distributions	3
1.3 Realistic model independent simulation generation level distributions	4
1.4 HNL branching ratios	5
1.5 Custom HNL total cross-sections for the four target masses compared to the total ($\nu_\tau/\bar{\nu}_\tau$ NC) cross-section used for SM neutrino simulation production with GENIE.	6
1.6 Model dependent simulation generation level distributions	10
2.1 Depth dependent scattering and absorption lengths	16
2.2 IceCube trigger efficiencies	17
2.3 Level 4 classifier outputs (muon and noise)	20
2.4 FLERCNN architecture	21
2.5 FLERCNN muon classifier probability distributions	22
2.6 Inclusive total neutrino-nucleon cross-sections	25
2.7 Hole ice angular acceptance modification	26
3.1 xx	31
3.2 Asimov inject/recover test (0.6 GeV)	34
3.3 Pseudo-data trials TS distribution (0.6 GeV)	34
3.4 xx	35
3.5 xx	36
A.1 Simplified model independent simulation generation level distributions	39
A.2 Realistic model independent simulation generation level distributions	40
A.3 Model dependent simulation generation level distributions	40
B.1 Asimov inject/recover test (0.3 GeV, 1.0 GeV)	41
B.2 Pseudo-data trials TS distribution (0.3 GeV, 1.0 GeV)	41
C.1 xx	43

List of Tables

1.1	Simplified model independent simulation sampling distributions	3
1.2	Realistic model independent simulation sampling distributions	4
1.3	HNL mass dependent decay channels	7
1.4	Model dependent simulation sampling distributions	8
2.1	GENIE generation cylinder volumes	14
2.2	Vuvuzela noise simulation parameters	17
2.3	Final analysis cuts	23
3.1	Final level background event/rate expectation	29
3.2	Final level signal event/rate expectation	30
3.3	Analysis binning	30
3.4	xx	32
3.5	Staged minimization routine settings	33
3.6	xx	35
C.1	xx	44

Heavy Neutral Lepton Signal Simulation

1

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 expected performance in terms of reconstruction and classification accuracy which governs the 90 % confidence level on estimateing the $|U_{\tau 4}|^2$ mixing matrix element. This is the first HNL simulation developed for IceCube DeepCore. Two avenues of simulation generation were pursued in parallel. The physically accurate, model dependent simulation is described in Section 1.2 and a collection of model independent simulation samples was realized and is explained in Section 1.1. The latter is used for performance benchmarking and as a cross-check for the model dependent simulation. The SM simulation generation and the default low energy event selection and processing chain are introduced in Chapter 2 and everything but the generation is applied identically to both neutrinos and HNLs.

1.1	Model Independent Simulation	1
1.2	Model Dependent Simulation	4

1.1 Model Independent Simulation

To investigate the potential of IceCube to detect HNLs by identifying the unique double cascade morphology explained in Section ??, it is very valuable to have a simulation chain where the double cascade kinematics can be controlled directly. In a realistic model the decay kinematics and the absolute event expectation all depend on the specific model parameters chosen (see Section ??). To decouple the simulation from a specific parameter choice, a model independent double cascade generator was developed. Using this generator several simulation samples were produced to investigate the performance of IceCube DeepCore to detect low energetic 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 choosable locations in the detector. The results of this study will be discussed in Chapter ??.

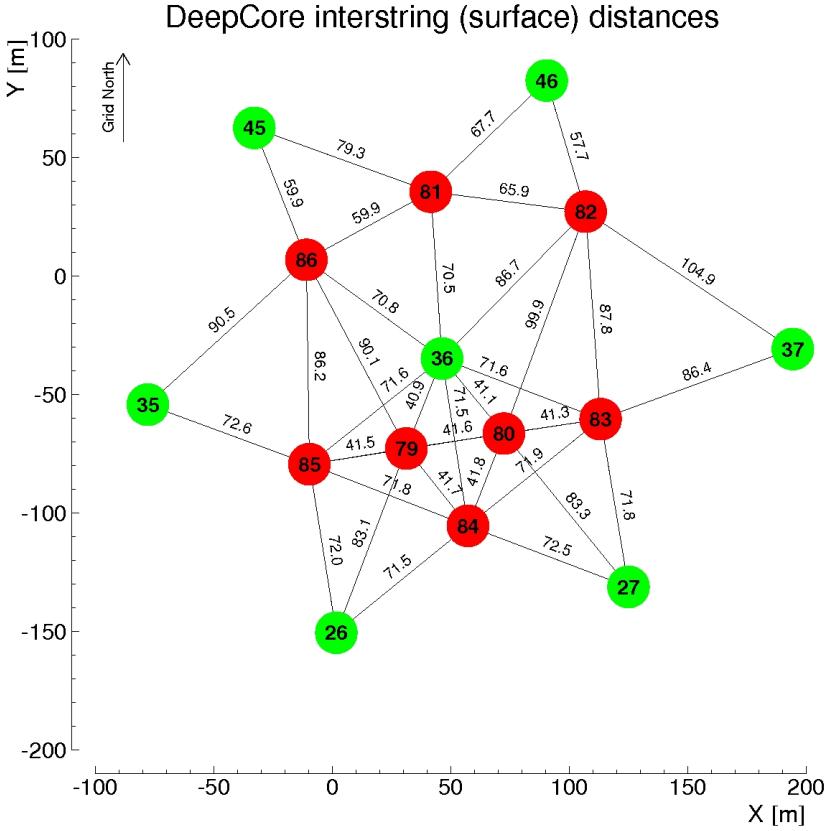
1.1.1 Simplistic Samples

To investigate some idealistic double cascade event scenarios, two samples are produced for straight up-going events that are centered on a string and horizontal events located inside DeepCore.

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 (namely string 81) and the directions are directly up-going. The horizontal positions and distances of all DeepCore fiducial volume strings are shown in Figure 1.1 and string 81 is at a medium distance of ~ 70 m to its neighboring strings. As already mentioned in Section ??, DeepCore strings have higher quantum efficiency DOMs and a denser vertical spacing, making them better to detect low energetic events that produce little light. To produce the events, the x, y position of the cascades

Make my own DC string positions/distances plot version, viable for the margin?

Maybe just reference the one from the detector description and drop it here?



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.

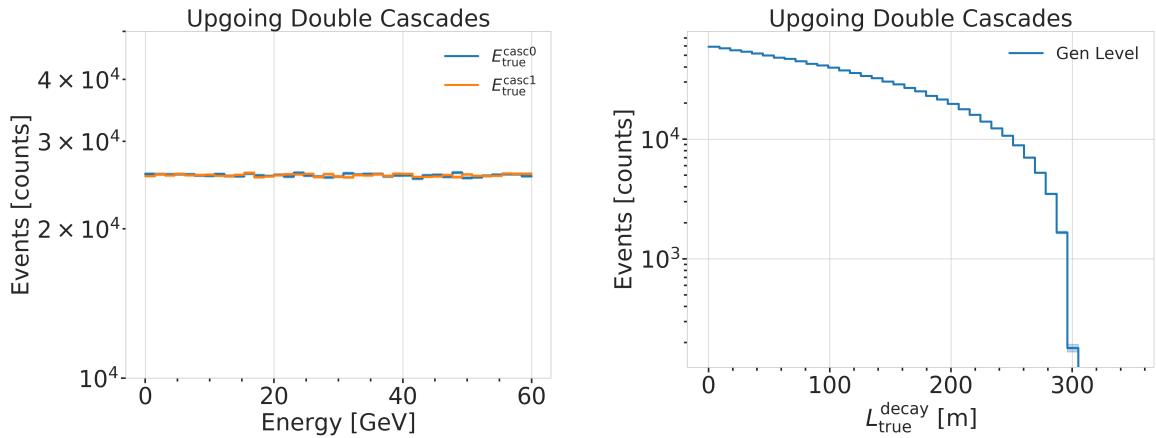


Figure 1.2: Generation level distributions of the simplistic simulation samples. Cascade and total energies (left) and decay lengths (right) of both samples are shown.

in Section 2.1.1. The total energy is divided into two parts, by randomly 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. 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 randomly 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. By 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 randomly choosing whether the cascade position that was sampled is the first cascade or the second. The decay length is sampled from an exponential distribution, as expected for a decaying heavy mass state. The sampling distributions/values are listed in Table 1.4. Example distributions of the generation level variables are shown in Figure 1.3, while further distributions can be found in Figure A.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

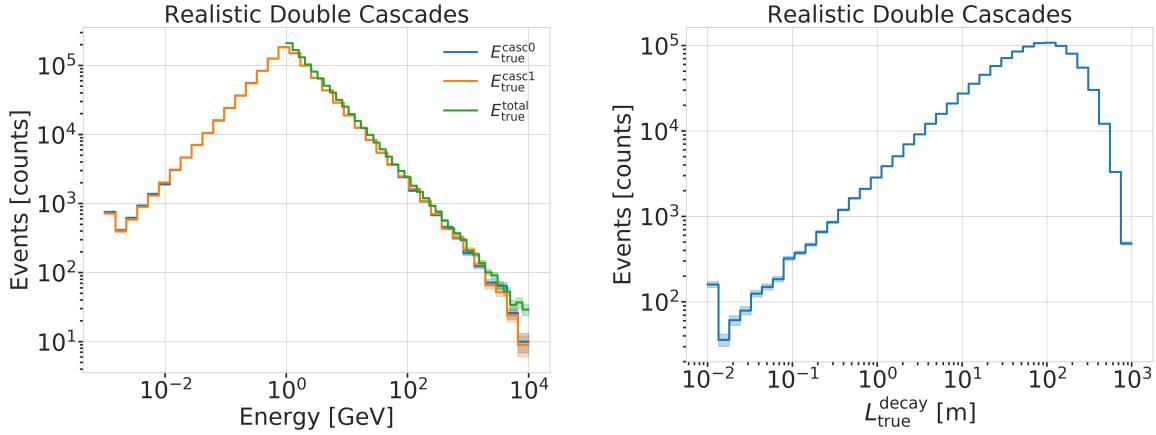


Figure 1.3: Generation level distributions of the simplistic realistic sample. Shown are the cascade and total energies (left) and decay lengths (right).

1.2 Model Dependent Simulation

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 simulation to the fourth heavy lepton mass, m_4 , and the mixing, $|U_{\tau 4}^2|$. 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 simulation 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 breaking nucleus. 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, identical to what will be described for SM neutrino simulation on Section 2.1.1. 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. 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.

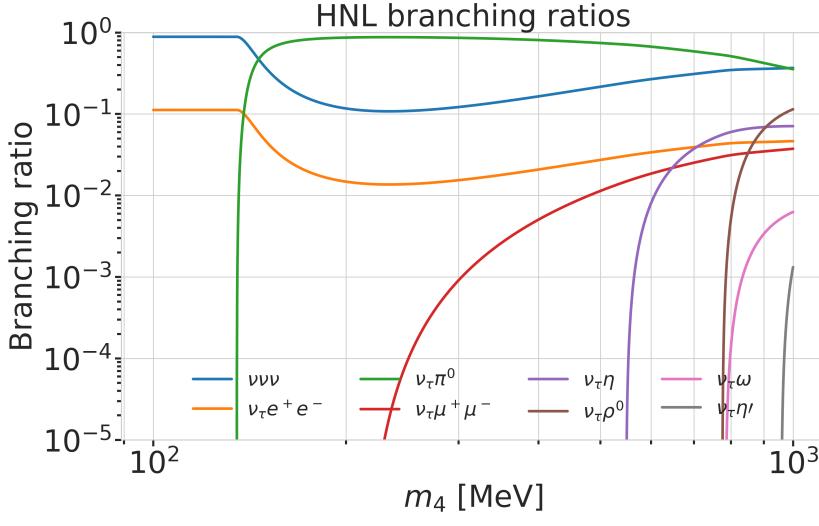


Figure 1.4: Branching ratios of the HNL within the mass range considered in this work, only considering $|U_{\tau 4}^2| \neq 0$, calculated based on the results from [3].

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 to produce secondary SM particles, where the decay mode is randomly chosen based on the mass dependent branching ratios from the kinematically accessible decay modes shown in Figure 1.4. The cross-section and decay width calculations were implemented for this purpose and will be explained in more detail in the following. Another needed addition to LI is that the decay products of the HNL are also added to the list of MC particles in each event. They are injected with the correctly displaced position and delayed time from the interaction vertex, given the HNL decay length. These HNL daughter particles form the second cascade, not as a single hadronic cascade object, but as the explicit particles forming the shower. The kinematics of the two-body decays are computed analytically, while the three-body decay kinematics are calculated with MADGRAPH [4], 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.

The injection is done using the LI *volume mode*, for the 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.

[4]: 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”

Cross-Sections

The cross-sections are calculated using the NUXSPLMKR [5] software, which is a tool to calculate neutrino cross-sections from *parton distribution functions* (*PDFs*) and then produce splines that can be read and used with LI/LW. The tool was customized 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

In which variables are they splined?

[6]: Levy (2009), “Cross-section and polarization of neutrino-produced tau’s made simple”

the same condition fulfilled for the CC case, where the outgoing charged lepton mass is non-zero. Following [6] (equation 7), the condition

$$(1 + x\delta_N)h^2 - (x + \delta_4)h + x\delta_4 \leq 0 \quad (1.1)$$

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

$$\delta_4 = \frac{m_4^2}{s - M^2}, \quad (1.2)$$

$$\delta_N = \frac{M^2}{s - M^2}, \text{ and} \quad (1.3)$$

$$h \stackrel{\text{def}}{=} xy + \delta_4, \quad (1.4)$$

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.

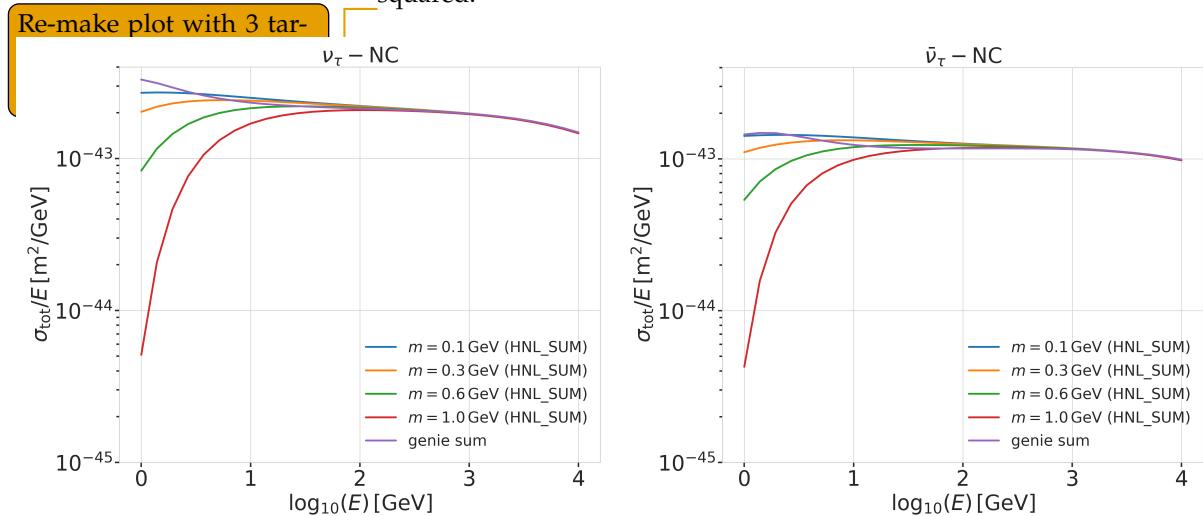


Figure 1.5: Custom HNL total cross-sections for the four target masses compared to the total ($\nu_\tau/\bar{\nu}_\tau$ NC) cross-section used for SM neutrino simulation production with GENIE.

The GRV98LO PDFs were added to the cross-section spline maker and used to create the HNL cross-sections for consistency with the SM neutrino simulation that will be explained in Section 2.1.1. The double differential ($d\sigma/dxdy$) and total (σ) cross-sections were produced for the chosen target HNL masses and then splined. The produced cross-section are added to the custom LI version and used for the simulation generation and weighting. Figure 1.5 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. Above $\sim 2 \times 10^2$ GeV they match, which is the desired result of using the identical input PDFs.

- SB: emphasize the cut-off/suppression
- Add comparisons of SM cross-sections between NuXSSplMkr and genie?
- add varied total cross-section for a few background HNL events (for QE/RES variations?!)

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 1.4. The individual branching ratio for a specific mass is calculated as $\text{BR}_i(m_4) = \Gamma_i(m_4)/\Gamma_{\text{total}}(m_4)$, where $\Gamma_{\text{total}}(m_4) = \sum \Gamma_i(m_4)$.

The individual decay widths Γ_i are computed using the state-of-the-art calculations from [3], which are described in the following.

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

$$\Gamma_{\nu_4 \rightarrow \nu_\tau P} = |U_{\tau 4}|^2 \frac{G_F^2 m_4^3}{32\pi} f_P^2 (1 - x_p^2)^2, \quad (1.5)$$

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

$$f_{\pi^0} = +0.1300 \text{ GeV}, \quad (1.6)$$

$$f_\eta = +0.0816 \text{ GeV}, \text{ and} \quad (1.7)$$

$$f_{\eta'} = -0.0946 \text{ GeV}, \quad (1.8)$$

while the decay to a neutral vector meson is given by

$$\Gamma_{\nu_4 \rightarrow \nu_\tau V} = |U_{\tau 4}|^2 \frac{G_F^2 m_4^3}{32\pi} \left(\frac{f_V}{m_V} \right)^2 g_V^2 (1 + 2x_V^2)(1 - x_V^2)^2, \quad (1.9)$$

with $x_V = m_V/m_4$,

$$f_{\rho^0} = 0.171 \text{ GeV}^2, \quad (1.10)$$

$$f_\omega = 0.155 \text{ GeV}^2, \quad (1.11)$$

and

$$g_{\rho^0} = 1 - 2 \sin^2 \theta_w, \quad (1.12)$$

$$g_\omega = \frac{-2 \sin^2 \theta_w}{3}, \quad (1.13)$$

and $\sin^2 \theta_w = 0.2229$ [7].

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

Channel	Opens	\hat{BR}
$\nu_4 \rightarrow \nu_\tau \nu_\alpha \bar{\nu}_\alpha$	0 MeV	1.0
$\nu_4 \rightarrow \nu_\tau e^+ e^-$	1 MeV	?
$\nu_4 \rightarrow \nu_\tau \pi^0$	135 MeV	?
$\nu_4 \rightarrow \nu_\tau \mu^+ \mu^-$	211 MeV	?
$\nu_4 \rightarrow \nu_\tau \eta$	548 MeV	?
$\nu_4 \rightarrow \nu_\tau \rho^0$	770 MeV	?
$\nu_4 \rightarrow \nu_\tau \omega$	783 MeV	?
$\nu_4 \rightarrow \nu_\tau \eta'$	958 MeV	?

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

Calculate max BRs

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

3-Body Decay Widths The (invisible) decay to three neutrinos is

$$\Gamma_{\nu_4 \rightarrow \nu_\tau \nu_\alpha \bar{\nu}_\alpha} = |U_{\tau 4}|^2 \frac{G_F^2 m_4^5}{192\pi^3}, \quad (1.14)$$

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

$$\Gamma_{\nu_4 \rightarrow \nu_\tau l_\alpha^+ l_\alpha^-} = |U_{\tau 4}|^2 \frac{G_F^2 m_4^5}{192\pi^3} [C_1 f_1(x_\alpha) + C_2 f_2(x_\alpha)], \quad (1.15)$$

with the constants defined as

$$C_1 = \frac{1}{4} (1 - 4s_w^2 + 8s_w^4), \quad (1.16)$$

$$C_2 = \frac{1}{2} (-s_w^2 + 2s_w^4), \quad (1.17)$$

the functions as

$$f_1(x_\alpha) = (1 - 14x_\alpha - 2x_\alpha^2 - 12x_\alpha^3)\sqrt{1 - 4x_\alpha} + 12x_\alpha^2(x_\alpha^2 - 1)L(x_\alpha), \quad (1.18)$$

$$f_2(x_\alpha) = 4[x_\alpha(2+10x_\alpha-12x_\alpha^2)\sqrt{1-4x_\alpha} + 6x_\alpha^2(1-2x_\alpha+2x_\alpha^2)L(x_\alpha)], \quad (1.19)$$

and

$$L(x) = \ln\left(\frac{1-3x_\alpha - (1-x_\alpha)\sqrt{1-4x_\alpha}}{x_\alpha(1+\sqrt{1-4x_\alpha})}\right). \quad (1.20)$$

JVS: consider also writing down the (trivial) 2-body decay kinematics for completeness and consistency.
This transition is a bit jarring as it is

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

3-Body Decay Kinematics with MadGraph

The specific MadGraph version used to produce the 3-body decay kinematics is MadGraph4 v3.4.0. As input, the decay diagrams calculated with FeynRules 2.0 using the Lagrangians derived in [3]. 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.

1.2.2 Sampling Distributions

In principle, the generation level sampling distributions should be chosen such that at final level of the 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 underestimate of the expected rates, because part 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. Because during development it became clear that the low lengths component is crucial to get a reasonable event estimate, each sample consists of a part that is generated for very short decay lengths and one for long decay lengths. 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 .

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

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	$r=600$ 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

1.2.3 Weighting Scheme

1: An analysis in PISA is a collection of functions that are written in so-called *stages*. The stages can be combined in a *pipeline* and are executed in a specific order. This particular stage is located in `pisa/stages/aeff/weight-hnl.py`.

To produce physically correct event distributions based on the simplified generation sampling distributions for the HNL simulation, the forward folding method that was already introduced for the SM simulation in Section 2.1 is also used. The weighting scheme that will be explained in the following is implemented in a custom stage¹ in the IceCube low energy analysis

framework PISA [8], which will be discussed in Section 3.2.1. 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

$$\gamma = \frac{\sqrt{E_{\text{kin}}^2 + m_4^2}}{m_4}, \quad (1.21)$$

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

$$v = c \cdot \sqrt{1 - \frac{1}{\gamma^2}}, \quad (1.22)$$

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

$$\tau_{\min/\max} = \frac{s_{\min/\max}}{v \cdot \gamma}. \quad (1.23)$$

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

$$\tau_{\text{proper}} = \frac{\hbar}{\Gamma_{\text{total}}(m_4) \cdot |U_{\tau 4}|^2}, \quad (1.24)$$

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 wanted exponential distribution

$$\text{PDF}_{\text{exp}} = \frac{1}{\tau_{\text{proper}}} \cdot e^{\frac{-\tau}{\tau_{\text{proper}}}}, \quad (1.25)$$

and the inverse distribution that was sampled from

$$\text{PDF}_{\text{inv}} = \frac{1}{\tau \cdot (\ln(\tau_{\max}) - \ln(\tau_{\min}))}. \quad (1.26)$$

This re-weighting factor is then calculated as

$$w_{\text{lifetime}} = \frac{\text{PDF}_{\text{exp}}}{\text{PDF}_{\text{inv}}} = \frac{\Gamma_{\text{total}}(m_4) \cdot |U_{\tau 4}|^2}{\hbar} \cdot \tau \cdot (\ln(\tau_{\max}) - \ln(\tau_{\min})) \cdot e^{\frac{-\tau}{\tau_{\text{proper}}}}. \quad (1.27)$$

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

$$w_{\text{total}} = |U_{\tau 4}|^2 \cdot w_{\text{lifetime}}. \quad (1.28)$$

If this additional weighting factor is multiplied to a generation weight with units m^2 (like in Equation 2.1), the livetime in s , and the oscillated primary neutrino flux in $\text{m}^{-2}\text{s}^{-1}$, it results in the number of expected events in the detector for this particular MC event for a chosen mixing (and mass).

add table with number of gen level files? mention the event number is smaller because of kinematic condition?

1.2.4 Generation Level Distributions

Figure 1.6 shows some selected generation level distributions. Additional distributions can be found in Figure A.3.

Combine low/high plots
and remove all traces

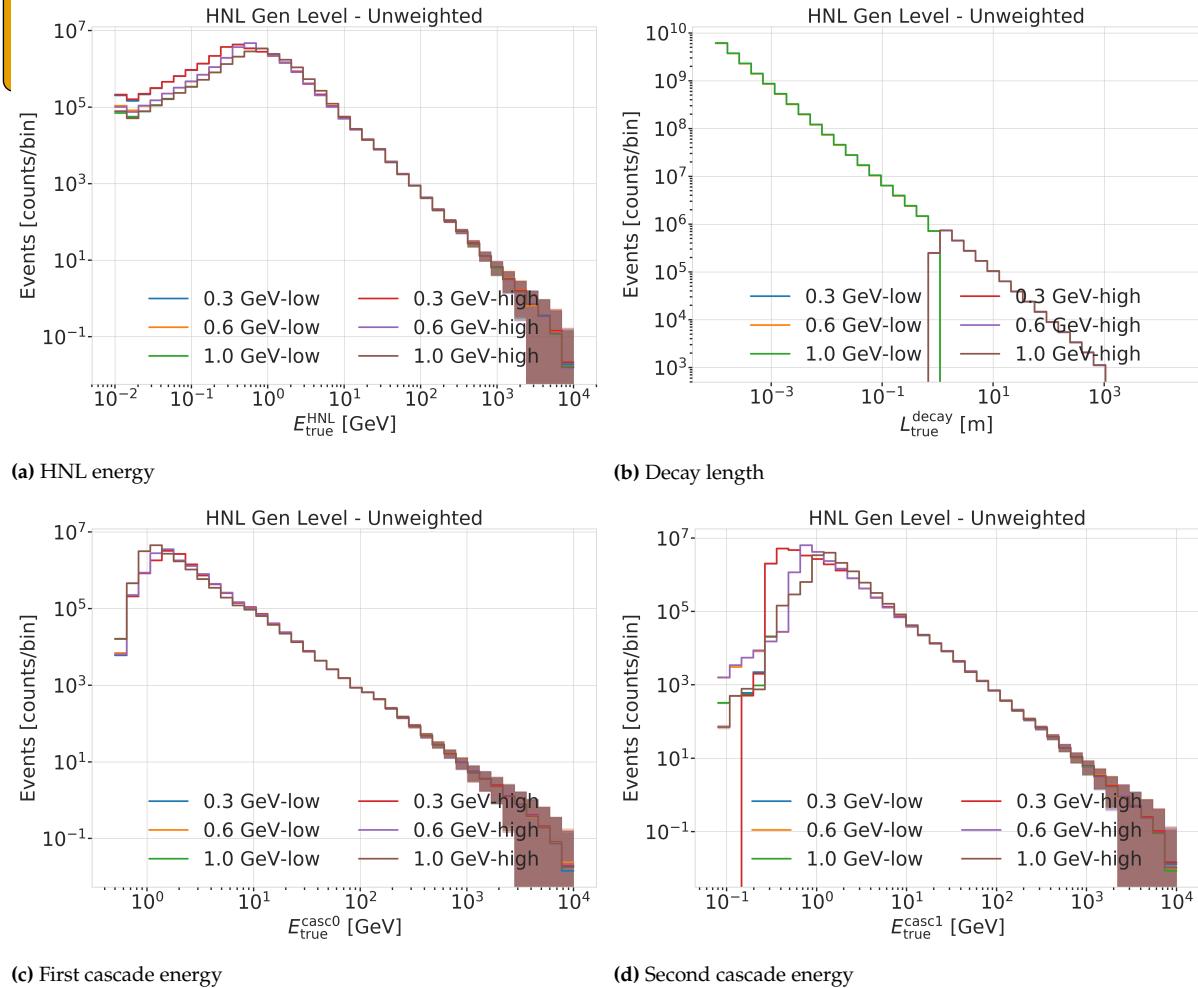


Figure 1.6: Generation level distributions of the model dependent simulation.

Adapt chapter to reflect
switched chapter order

Standard Model Background Simulation and Data Processing

2

The analysis presented in this thesis is highly dependent on an efficient 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 current simulation and event selection chain used for state-of-the-art IceCube neutrino oscillation measurements like [9]. The whole chain can be broadly split into 4 steps:

Step 1 Event Generation: 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 the particles produced in the interaction with the ice.

Step 2 Detector Simulation: The particles from the first step are propagated through the ice, producing Cherenkov photons, which are then propagated further until they reach a DOM or are absorbed. If they hit a DOM the detector response (acceptance and PMT) is simulated.

Step 3 Processing: 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.

Step 4 Reconstruction: Once the sample is small enough for more sophisticated reconstruction techniques to be feasible to run, the events are reconstructed using a CNN and some high level variables are computed. Based on these variables the final event selection is applied.

This chapter only describes the event generation for the SM background simulation (neutrinos and muons), while the signal simulation is described in Chapter 1. The detector simulation is identical for both signal and background events while processing and reconstruction are applied to all simulation and data in the same way. 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. Similar approach can be taken for varying detector response and through this a more efficient (reduced) use of computing resources can be achieved. The following sections describe the different steps in more detail and the last section, Section 2.5, describes the related systematic uncertainties considered for this work.

2.1 Event Generation

The MC is used in the analysis by applying a method called *forward folding*, where a very large number of events (signal and background) is produced using sampling distribution that are tuned to have a large selection efficiency. Those distributions don't have to be physically correct distributions, but they need to cover the full parameter space of interest for the analysis. To produce a physical distribution, the events are weighted given a specific

2.1	Event Generation	13
2.2	Detector Simulation	15
2.3	Processing	17
2.4	Reconstruction	20
2.5	Systematic Uncertainties	23

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

Table 2.1: 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	
	4-12				
	12-100	350	600	100000	650
	100-10000	550	1000	57500	
$\nu_\mu + \bar{\nu}_\mu$	1-5	250	500	408000	
	5-80	400	900	440000	
	80-1000	450		57500	1550
	1000-10000	550	1500	6700	
$\nu_\tau + \bar{\nu}_\tau$	1-4	250	500	1500000	
	4-10			300000	
	10-50	350	600	375000	350
	50-1000	450	800	200000	
	1000-10000	550	1500	26000	

choice of physics and nuisance parameters. The large number of raw MC events ensures a good estimation of the expected numbers and weighted distributions.

The analysis itself is then performed by comparing the weighted MC distributions to the observed data. This is done by binning them as described in Chapter 3 and calculating a loss function comparing the bin expectations to the data. The physics and nuisance parameters that best correspond to the observed data are estimated by minimizing this loss function. In order to achieve a reliable result with this method the MC needs to be precise and as close to the data as possible (at least at the final event selection).

2.1.1 Neutrinos

Due to the very low interaction rate of neutrinos, 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

$$w = \frac{1}{F_{\text{sim}}} \frac{1}{N_{\text{sim}}} , \quad (2.1)$$

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 baseline neutrino flux used in this thesis, computed for the South Pole, is taken from Honda *et al.* [10].

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

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 2.1. 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.

To simulate the neutrino interaction with the ice, the GENIE event generator [11] (version 2.12.8) is used, resulting in the secondary particles and the kinematic and cross-section parameters. As input, the outdated GRV98LO [12] *parton distribution functions (PDFs)* was used, because it was the only option that could incorporate extrapolations to lower Q^2 [13]. Muons produced in these interactions are propagated using PROPOSAL [14], also simulating their Cherenkov light output. The shower development of gamma rays, electrons, and positrons below 100 MeV and hadronic showers below 30 GeV is simulated using GEANT4 [15] while for higher energies an analytical approximation from [16] is used.

2.1.2 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 [17]. This work uses full CORSIKA [18] simulations of muons to tailor the parameterizations, starting from *cosmic ray (CR)* interactions with atmospheric nuclei using the CR flux model from [19] and producing the muons applying the *hadronic interaction (HI)* model SIBYLL 2.1 [20]. After the generation, they are propagated through the ice with PROPOSAL producing photons, treating them exactly like the muons produced in neutrino interactions.

Since the offline processing and selection steps described in Section 2.3.2 and Section 2.4 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 sets of muon simulation are produced. **A first set** including all events resulting from the above described generation to tune the lower level selection (up to L4) and **a second set** to estimate the muon contamination at higher levels (above L5), which only accepts muon events if they pass through a smaller cylinder centered in DeepCore (height of 400 m and radius of 180 m) and rejects events based on a KDE estimated muon density at L5 (in energy and zenith) increasing the simulation efficiency at L5 significantly .

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

[12]: Glück et al. (1998), “Dynamical parton distributions revisited”

[13]: Bodek et al. (2003), “Higher twist, $\chi_i(\omega)$ scaling, and effective LO PDFs for lepton scattering in the few GeV region”

[14]: Koehne et al. (2013), “PROPOSAL: A tool for propagation of charged leptons”

[15]: Agostinelli et al. (2003), “Geant4—a simulation toolkit”

[16]: 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”

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

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

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

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

put a number on this significant increase?

2.2 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).

2.2.1 Photon Propagation

Any photon that was produced in the event generation is individually traced through the ice, simulating scattering and absorption processes. The

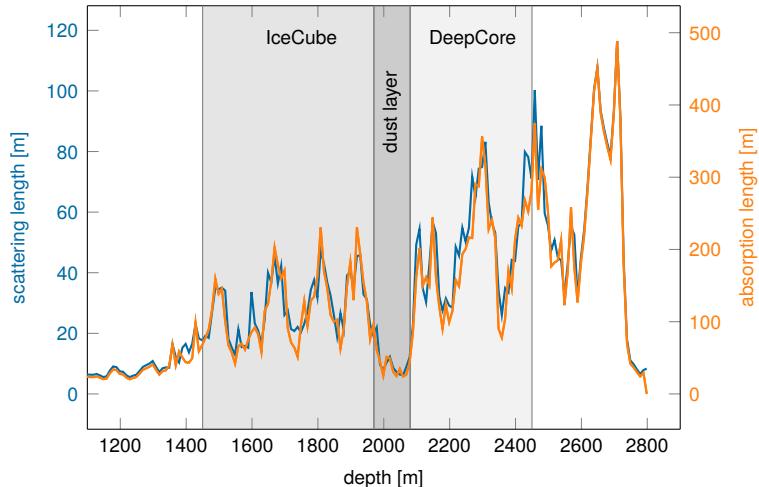


Figure 2.1: Scattering and absorption lengths in the SPICE model used for simulation production as a function of depth, modified from [24].

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

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

put a number on the tilt angle?

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

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

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

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

propagation is done using `clsim` [21] which is an implementation of the *Photon Propagation Code (PPC)* [22] in `OPENCL`. It is optimized to be run very efficiently on GPUs, which is what is done for IceCube simulation production. 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 [23] accounts for the layers being tilted by a small amount (δ) and the absorption and scattering lengths having a non-uniformity with respect to the azimuth direction. Figure 2.1 shows the values of this model for the different depths, indicating the location of IceCube, DeepCore, and the dust layer.

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 [25] and a Henyey-Greenstein distribution [26]. 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 reached a DOM or was absorbed¹. After all photons have been propagated in that manner, the final step is to output the photons that reached a DOM for further processing.

2.2.2 Detector Responses

The second part of simulating the IceCube detector is the DOM response. Whether a photon that reached a DOM produces a signal depends on the total efficiency and the angular acceptance curve of the specific DOM. The total efficiency includes effects of the DOM glass, PMT quantum and photo-electron collection efficiencies, and it is wavelength dependent. Additionally, there is another angle dependent effect called *hole ice* [27]. 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 a mixture of two exponential distributions and a normal distribution. This *single photo-electron (SPE)* distribution was tuned to match the observed distribution in each DOM in an *in-situ* calibration study [28]. Figure ?? shows the distribution compared to a lab measurement. 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*, which was already mentioned in Section ??.

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 [29, 30] is used to create additional MCPEs that are fed into the same simulation chain described above. This module takes into account thermal and non-thermal components and their times are sampled using parameterizations of the measured distributions, where the thermal noise component is uncorrelated photons and the non-thermal component is from burst of photons. 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 2.2 shows the average values for these parameters.

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

Add SPE distribuiton plot

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

Table 2.2: Typical parameter values used in the vuvuzela noise simulation. Averaged over all DOMs.

[29]: Larson (2013), "Simulation and Identification of Non-Poissonian Noise Triggers in the IceCube Neutrino Detector"

[30]: Larson (2018), "A Search for Tau Neutrino Appearance with IceCube-DeepCore"

[24]: Trettin (2023), "Search for eV-scale sterile neutrinos with IceCube DeepCore"

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

AT: Das klingt so, als würde die MC Simulation ein analoges Signal erzeugen, was auch digitalisiert wird. Vllt kann man das nochmal nachforschen, aber zumindest in meiner Arbeit habe ich geschrieben, dass die MCPEs direkt in ATWD und fADC Readouts umgewandelt werden.

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

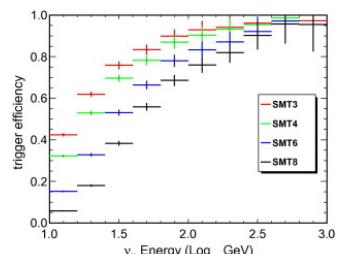


Figure 2.2: Efficiencies of different IceCube and DeepCore triggers, taken from [33].

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. Most parts of this processing are identical to the procedure already described in [24, 31]. 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.

2.3.1 Trigger and Filter

Before the data can be sent to the North, the initial signal coming from the PMT (for data) or from the detector response simulation (for MC) is a voltage waveform, which has to be digitized and information of photon hits has to be extracted. 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 [32].

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

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

[35]: Aartsen et al. (2017), “The IceCube Neutrino Observatory: instrumentation and online systems”

[33]: Abbasi et al. (2012), “The design and performance of IceCube DeepCore”

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

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 40 MHz. Additionally, the first 427 ns are digitized using an *Analog Transient Waveform Recorder* (ATWD) with a sampling rate of 300 MHz [34], 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 \sim 250 Hz [35], accepting \sim 70 % of ν_μ -CC events at 10 GeV and \sim 90 % at 100 GeV [33]. The trigger efficiencies for different SMT triggers, including the DeepCore SMT-3, are shown in Figure 2.2.

Online Filter

The digitized waveforms are sent to the ICL, where a further filter is applied *online*². 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.

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.

2.3.2 Event Selection

After the data was sent to the North, the *offline* filters and selection are applied to further 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 \sim 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 $\sim 0.5\text{ Hz}$.

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

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)* [36] 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 2.3. 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.

[36]: Friedman (2002), “Stochastic gradient boosting”

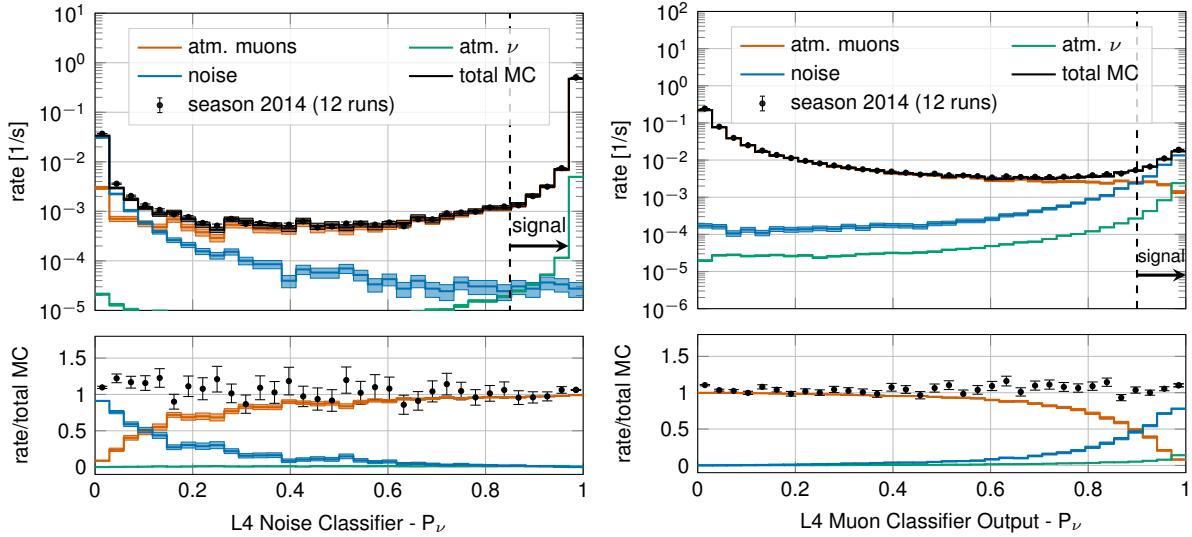


Figure 2.3: Distributions of Level 4 noise classifier output (left) and muon classifier output (right), where larger values indicate more neutrino-like and lower values more noise-like/muon-like. Taken from [9].

Level 5

add some figure showing the corridors?

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

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 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 question 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.

2.4 Reconstruction

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

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

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

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 [37] 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 [9]. 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 [38] is based on this reconstruction.

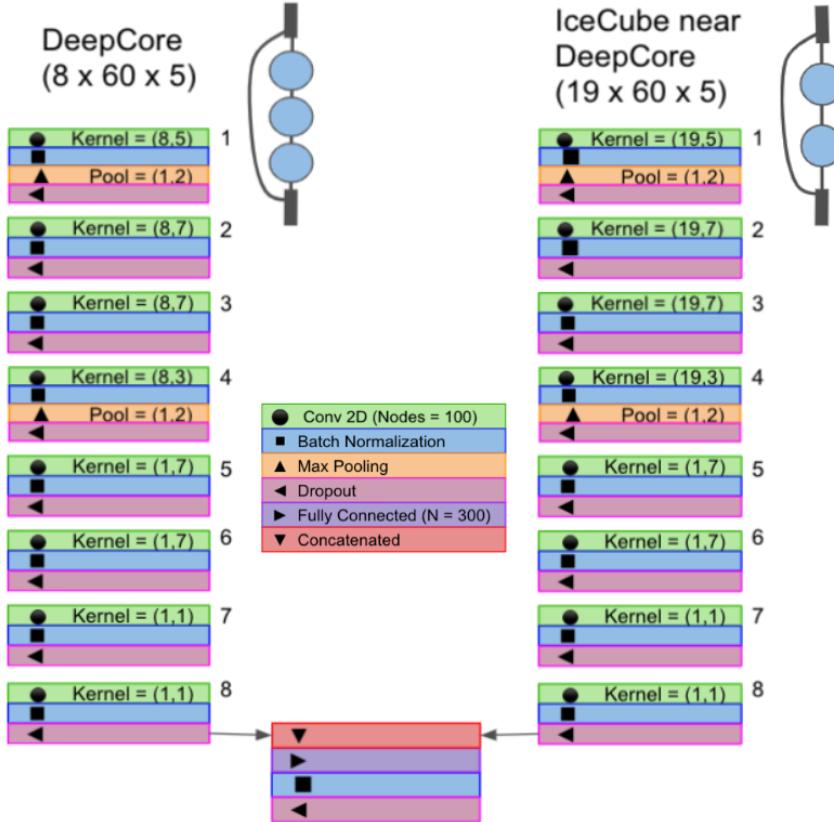


Figure 2.4: Architecture of the FLERCNN neural networks, taken from [39].

2.4.1 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 [39, 40] is a CNN optimized to reconstruct IceCube events at low energies ($<100\text{ GeV}$) in a fast and efficient manner, by leveraging the approximate translational invariance of event patterns within the detector. The architecture of the network is very similar to the preexisting IceCube CNN event reconstruction [41], 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 2.4. 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.

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

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

[40]: Micallef ()

[41]: Huenefeld (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

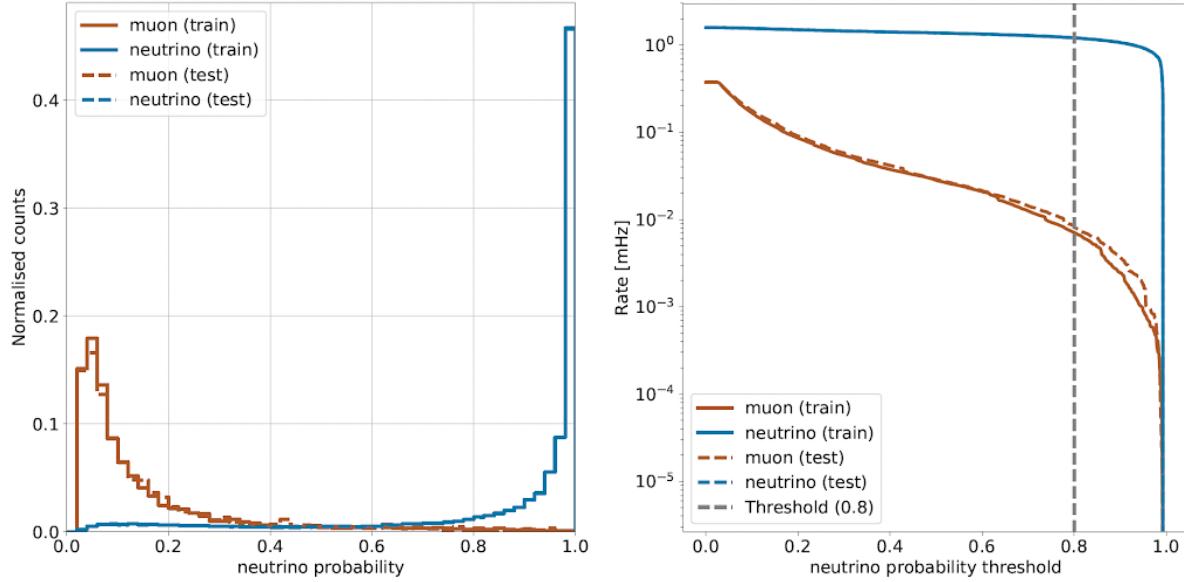


Figure 2.5: FLERCNN muon classifier output score (left) and rate of neutrinos and muons as function of muon classifier cut (right). Taken from [flercnn_analysis_internal_note]

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

add some performance plots of the FLERCNN reconstruction

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?

3: 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

2.4.2 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}^3 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 $\cos(\theta_{\text{zenith}}) \leq 0.3$. 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 2.5 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.

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 2.3. The FLERCNN reconstruction was optimized for atmospheric

Variable	Threshold	Removed
Number of hit DOMs	≥ 7	1.05 %
Radial distance	$< 200 \text{ m}$	0.09 %
Vertical position	$-495 \text{ m} < z < -225 \text{ m}$	5.48 %
Energy	$5 \text{ GeV} < E < 100 \text{ GeV}$	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 2.3: 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.

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 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 photons⁴. 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.

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

add final sample composition, but maybe also in analysis chapter to show signal and background?

at some place I will want a selection efficiency plot for SM BG and HNL signal, but I'm not sure where to put it yet

2.5 Systematic Uncertainties

There are multiple sources of systematic uncertainties related to the event generation and processing explained in this chapter. All uncertainties considered in this work are 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.

2.5.1 Atmospheric Flux Uncertainties

The flux of atmospheric neutrinos is influenced by multiple factors, the spectrum and composition of CRs, the assumed atmospheric conditions, and the HI 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 [9].

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

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

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

[43]: Barr et al. (2006), “Uncertainties in atmospheric neutrino fluxes”

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

[9]: Abbasi et al. (2023), “Measurement of atmospheric neutrino mixing with improved IceCube Deep-Core calibration and data processing”

[46]: Barr et al. (2006), “Uncertainties in Atmospheric Neutrino Fluxes”

add figure with Barr blocks?

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

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

5: The choice of flux and HI model have minor impact on the variations.

nuclei [42]. The uncertainty in this CR flux component can be described as a power law correction [43, 44]

$$\Phi'_\nu = \Phi_\nu \left(\frac{E}{E^*} \right)^{\Delta\gamma}, \quad (2.2)$$

where E^* 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^* was chosen to be 24 GeV as to minimize the dependence of the overall flux scale on $\Delta\gamma$ [9].

Hadronic interaction model: Neutrinos are produced in the decaying hadrons in CR air showers, 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 [45] 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 $d\Phi_l/dB$ for a variation dB of some parameter B . Scaling this variation by some value b , the modified total flux, s is then given by

$$\Phi'_l = \Phi_l + \left(b \cdot \frac{d\Phi_l}{dB} \right). \quad (2.3)$$

Matching the work in [46], 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 [47] HI model and the GSF CR flux [42] were used to calculate the related flux changes⁵ for the different regions in E_i and x_{lab} , resulting in 17 variables, encoding the possible changes.

2.5.2 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 DIS interactions, which are the dominant processes above 20 GeV, and *charged current resonance production* (CCRES) and *charged current quasi elastic scattering* (CCQE), which are relevant below 20 GeV where interactions with the nucleons as a whole are important. Three parameters are included to account for all relevant cross-sections uncertainties.

At low energies two parameters are included accounting 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

$$F(Q^2) \sim \frac{1}{(1 - (\frac{Q}{M_A})^2)^2}, \quad (2.4)$$

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 with GENIE, 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

which experiments measure the axial mass?

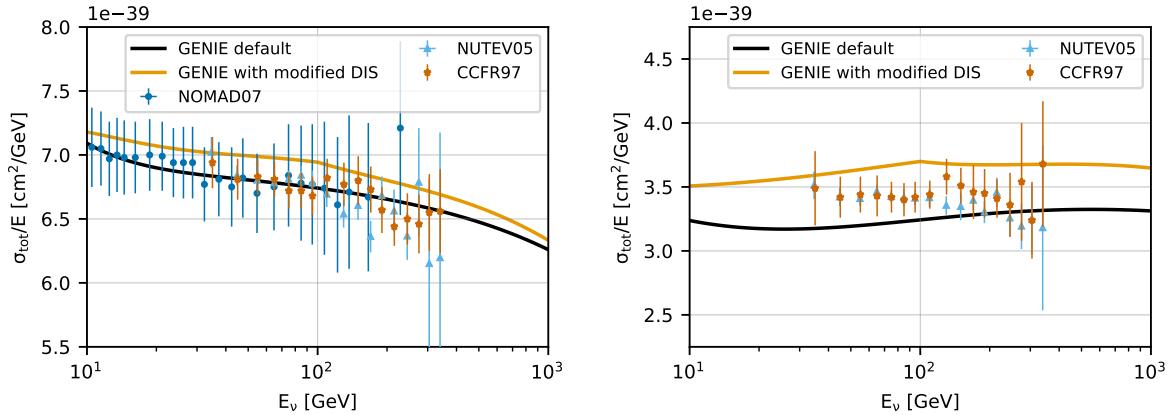


Figure 2.6: Inclusive total neutrino-nucleon cross-sections on an isoscalar target (black) for neutrinos (left) and antineutrinos (right) calculated with GENIE, comparing to measurements from NOMAD [49], NUTEV [50], and CCFR [51]. The scaled GENIE cross-section (orange) is also shown. Taken from [9].

$\pm 20\%$ for M_A^{CCRES} [11]. 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.

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 [48] 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. 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 2.6.

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

2.5.3 Detector Calibration Uncertainties

There are multiple sources of systematic uncertainties related to the detection process of neutrinos in IceCube. Dominant for this analysis are the effects of the properties of the ice itself and the optical efficiency of the DOMs. None of these uncertainties can be described by an analytic expression, so they have to be estimated using MC simulation. The method used to derive the continuous variations based on the MC simulation is described in Section 2.5.5. The five relevant uncertainty parameters are the absolute efficiency of the DOMs, a global scaling of bulk ice scattering and absorption lengths, and variations of the relative angular acceptance due to hole ice variations in two parameters.

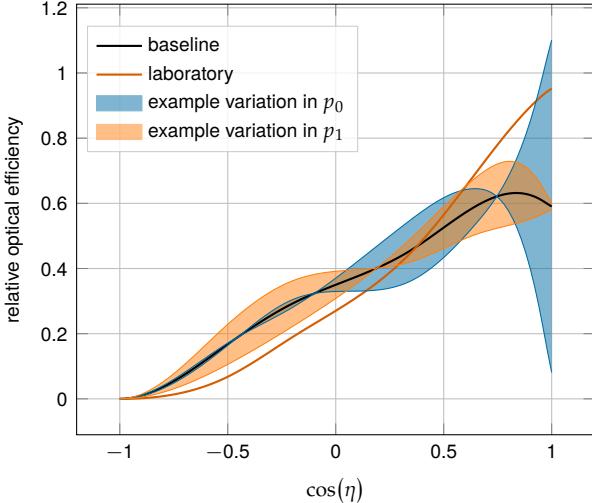


Figure 2.7: Relative angular acceptance modification due to hole ice. Shown is the current baseline model, the variations from changing p_0 and p_1 , and a laboratory measurement. Modified from [24].

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

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

[54]: Rongen, Martin (2016), “Measuring the optical properties of IceCube drill holes”

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

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 [52, 53] are at the order of 10 %, this systematic is highly relevant and has to be included in the analysis.

Bulk 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 2.2.1. To account for uncertainties on this model of the bulk ice coefficients, a global scaling for each of the two parameters (global absorption, global scattering) 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 [54]. This effectively decreases the chance of photons 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, normalized⁶ 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 2.7 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.

2.5.4 Muon Uncertainties

The muon fraction in the final level selection (see Section 2.4.2) 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 ??.

[cite this?](#)

2.5.5 Treatment of Detector Response Uncertainties via a Likelihood-Free Inference Method

[55]

Copy paste from OVS PRD about hypersurfaces (and interpolation of those):

To evaluate the expected impact of detection uncertainties, data sets are produced with different variations of detector response, processed to the final level of selection, and then they are parameterized following a model of the uncertainties to evaluate how the final sample would look like for any reasonable choice of parameters. The parametrizations are done at the analysis bin level, assuming that every effect considered is independent and that they can be approximated by a linear function. Under these assumptions we can compute a reweighting factor in every bin that depends on N parameters, which correspond to the number of systematic effects being considered, plus an offset c , as

$$f(p_1, \dots, p_N) = c + \sum_{n=1}^N m_n \Delta p_n. \quad (2.5)$$

Here m_n are the reweighting factors obtained from simulation sets with a systematic variation and Δp_n is the test value of a specific systematic variation.

The fit of the parameters m_n is done over all systematic MC sets, reducing the uncertainty on the MC prediction in each bin as a side effect since the error on the fitted function is smaller than the statistical error from the nominal MC set. The set of all fitted functions in all histogram bins are called "hypersurfaces". An example of such a fit from a single bin, projected onto one dimension, is shown in Fig. ??.

The event counts coming from different flavors and interactions have a different response to varying the same detector parameter. Therefore, the hypersurfaces in each bin are fit separately for three groups of events:

- ▶ $(\nu_{\text{all}} + \bar{\nu}_{\text{all}})$ NC + $(\nu_e + \bar{\nu}_e)$ CC: These events all produce cascade signatures in the detector.
- ▶ $(\nu_\tau + \bar{\nu}_\tau)$ CC: These interactions may differ from the previous group because they have a production threshold of $E_\nu \gtrsim 3.5$ GeV and also produce muons with a branching ratio of 17%.
- ▶ $(\nu_\mu + \bar{\nu}_\mu)$ CC: These interactions produce track-like signatures.

The distribution of $\chi^2/\text{d.o.f.}$ from the fits in all analysis bins is used as a diagnostic to ensure that the fitted, linear hypersurfaces provide a good estimate for the expected number of events for the full range of simulated detector configurations. We find that the means of these $\chi^2/\text{d.o.f.}$ distributions are all consistent with 1.0 as expected from good fits for each of the three categories described above (NC + ν_e CC, ν_τ CC and ν_μ CC). Attempts

[55]: Fischer et al. (2023), "Treating detector systematics via a likelihood free inference method"

to use higher order polynomial fits did not yield a significantly improved $\chi^2/\text{d.o.f.}$, and in fact often rendered the fits less stable.

To produce the histograms for fitting the hypersurfaces, a choice must be made for the values of flux, cross-section and oscillation parameters. We found that the hypersurface fits are sensitive to the choice of parameters that have correlations with the effect they encode. Most notably, this effect is observed between the mass splitting and DOM optical efficiency as demonstrated in Fig. ??, which shows the difference between fitted hypersurface gradients for the DOM efficiency dimension for two values of Δm_{32}^2 .

This problem arises because we are only fitting the hypersurfaces in reconstructed phase space, without accounting for the different true energy and zenith distributions of MC in each analysis bin, which change with each detector systematic variation. To mitigate this problem, we fit the hypersurfaces for 20 different values in mass splitting between $1.5 \times 10^{-3} \text{ eV}^2$ and $3.5 \times 10^{-3} \text{ eV}^2$, and then apply a piece-wise linear interpolation to all slopes, intercepts and covariance matrix elements. The oscillation parameter fit can then dynamically adapt the hypersurfaces for each value of Δm_{32}^2 that is tested using these interpolated functions. The effects of other parameter choices were evaluated as well, but none were found to introduce a significant bias.

Search for an Excess of Heavy Neutral Lepton Events

3

The measurement performed in this thesis is the search for an excess of HNL events in the 10 years of IceCube DeepCore data. In principle the two physics parameters to be probed are the mass of the HNL, m_4 , and the mixing between the fourth heavy mass state and the SM τ sector, $|U_{\tau 4}|^2$. Since the mass itself influences the production and decay kinematics of the event and the accessible decay modes, individual mass sets were produced as described in Section 1.2. The mass slightly influences the energy distribution, while the mixing both changes the overall scale of the HNL events and the shape in energy and PID. IceCube DeepCore is suited to measure the excess which appears around and below 20 GeV, due to its production from the atmospheric tau neutrinos, although a reduced lower energy threshold could improve the analysis. The measurement will be performed for the three mass sets individually, while the mixing is the parameter that can be varied continuously and will be measured in the fit.

3.1 Final Level Sample

The final level sample of this analysis always consists of the neutrino and muon MC introduced in Section 2.1 and one of the three HNL samples explained in Section 1.2. All of those simulation sets and the 10 years of IceCube DeepCore data are processed through the full processing and event selection chain described in Section 2.3 and Section 2.4 leading to the final level sample. Since applying the last cuts from Section 2.4.2 leaves an insignificant amount of pure noise events in the sample, the noise simulation is not included in the analysis and will not be listed here.

3.1.1 Expected Rates/Events

The rates and the expected number of events for the SM background are shown in Table 3.1. The explicit detector livetime in the 10 years data taking period is 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.

Type	Rate [mHz]	Events (in 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	0.0032	936 ± 15	0.5
total	0.5991	175336 ± 143	100.0

3.1 Final Level Sample	29
3.2 Statistical Analysis	30
3.3 Analysis Checks	33
3.4 Results	34

add information about the matter profile used

add information about the oscillation probability calculation and the software used for it

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

Should I adapt the total numbers to match the sum of the rounded individual parts?

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 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)
$ U_{\tau 4} ^2 = 10^{-1}$		
0.3 GeV	3.3298 ± 0.0053	974.5 ± 1.6
0.6 GeV	3.0583 ± 0.0058	895.0 ± 1.7
1.0 GeV	2.4988 ± 0.0059	731.3 ± 1.7
$ U_{\tau 4} ^2 = 10^{-3}$		
0.3 GeV	0.0057	1.67 ± 0.01
0.6 GeV	0.0220	6.44 ± 0.01
1.0 GeV	0.0248	7.27 ± 0.01

3.1.2 Analysis Binning

[38]: Yu et al. (2023), “Recent neutrino oscillation result with the Ice-Cube experiment”

The identical binning to the analysis performed in [38] is used. It was chosen such that the track-like bin has the largest ν_μ -CC fraction. Extend the binning towards lower energies or increasing the number of bins did not improve the HNL sensitivities significantly. It also has to be considered that sufficient data events need to end up in the individual bins to result in a good fit, which was already investigated in the previous analysis. To mitigate the low data statistics, a few bins were not taken into account in the analysis. 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. Originally, there were two more bins in $\cos(\theta)$, which were removed to reduce muons coming from the horizon and some low energy bins in the cascade-like bin are removed due to the low event statistics.

Table 3.3: Three dimensional binning used in the analysis. All variables are from the FLERCNN reconstruction explained in Section 2.4.

Variable	N _{bins}	Edges	Step
P_v	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

add 3D expectation and/or S/sqrt(B) plots

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

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

3.2 Statistical Analysis

3.2.1 Low Energy Analysis Framework

The analysis is performed using the PISA [56] [8] software framework, 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 sets, which can then be compared to the observed data. The expectation for each set is calculated in parallel, applying

physics and nuisance parameter effects in a stage-wise manner, before combining the final expectation from all the sets.

3.2.2 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. The comparison is done using a modified χ^2 defined as

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

as the *test statistic* (TS). 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 bin i . 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.

I feel like I have to be a bit more precise on what is the fit metric (e.g. the mod chi2) and what is the TS, as in the mod chi2 difference, which is the actual TS, right?

Do I want/need to include the description of the KDE muon estimation?

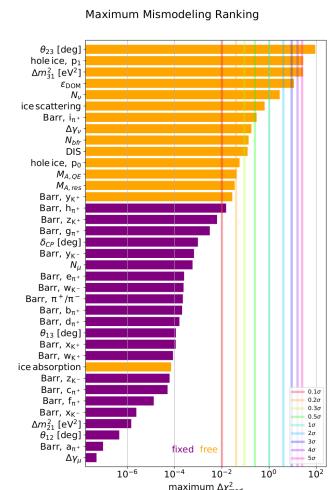


Figure 3.1: "calculated at a mixing of 0.1 and for the 1.0 GeV set"

3.2.3 Physics Parameters

The only variable physics parameter in this analysis is the mixing between the HNL and the SM τ sector, $|U_{\tau 4}|^2$. It can be changed 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 the nominal value of 0.0. The other physics parameter, the mass m_4 of the HNL, is 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.

3.2.4 Nuisance 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 1.0 GeV sample at a mixing value of 0.0, which is around the value where the analysis starts to become sensitive. 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 mis-modeling significance between a fit with all parameters free, which would result in a TS of 0.0 for this Asimov test. Parameters below a significance of 0.1σ are fixed and the test is performed in an iterative manner until the final set of free parameters is found. Figure 3.1 shows the resulting significances of one of these tests. In the final selection of free parameters the Barr h_{π^+} parameter was also left free and the ice absorption is kept free, despite showing a small significance. This is done

Blow up labels/legend/title and make it more readable in the margin

I don't like this formulation, but don't know better right now..

elaborate why this is also done to cover the whole energy range for the pion production, referencing the Barr Block plot that I haven't included yet :D

Table 3.4: xx

Parameter	Nominal	Range	Prior
$\Delta\gamma_\nu$	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
$\theta_{23} [\circ]$	47.5047	[0.0, 90.0]	-
$\Delta m_{31}^2 [\text{eV}^2]$	0.002475	[0.001, 0.004]	-
DIS	0.0	[-0.5, 1.5]	1.0
N_ν	1.0	[0.1, 2.0]	-
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]	-
bulk ice absorption	1.0	[0.85, 1.15]	-
bulk ice scattering	1.05	[0.9, 1.2]	-
N_{bfr}	0.0	[-0.2, 1.2]	-
$M_{A,\text{QE}}$	0.0	[-2.0, 2.0]	1.0
$M_{A,\text{res}}$	0.0	[-2.0, 2.0]	1.0

because the bulk ice parameters are not well constrained and are known to have large impact, which might be concealed in the test, due to correlations with the other parameters.

I truly dislike this sentence, too, better ideas?

I'm just writing out the data from the table, but I need to mention/motivate the included priors here and maybe just point to the table for the ranges/nominal values? (Not quite sure about this)

The scaling parameter N_ν is included to account for the unknown overall normalization of the neutrino rate. It has the identical effect on the SM neutrino events and the BSM HNL events and its nominal value is set to 1.0 with a wide range of [0.1, 2.0].

Concerning the atmospheric neutrino flux, the CR power law flux correction factor $\Delta\gamma_\nu$ is included with nominal value of 0.0 and a range of [-0.5, 0.5]. Additionally, the Barr h_{π^+} , Barr i_{π^+} , and Barr y_{K^+} parameters of the pion and kaon production uncertainties are included with nominal values of 0.0 and ranges of [-0.75, 0.75], [-3.05, 3.05], and [-1.5, 1.5], respectively.

From the cross-section uncertainties introduced in Section 2.5.2, all three parameters DIS, $M_{A,\text{QE}}$, and $M_{A,\text{res}}$ are included in the fit with nominal values of 0.0 for all of them and range [-0.5, 1.5] for DIS and [-2.0, 2.0] for the axial mass parameters $M_{A,\text{QE}}$, and $M_{A,\text{res}}$.

cite?!

I think I will need to mention here that I did no include MA-RES and MA-QE for the HNL simulation..

add final level effects of varying the axial mass parameters (or example of one)

add final level effects of varying the DIS parameter (or example of one)

fix caption!

All the detector systematic uncertainties are included in the fit. The DOM efficiency ϵ_{DOM} has a nominal value of 1.0 and a range of [0.8, 1.2]. It is constrained by a Gaussian prior with a width of 0.1. The hole ice model parameters p_0 and p_1 are included with nominal values of 0.101569 and -0.049344, respectively, and ranges of [-0.6, 0.5] and [-0.2, 0.2]. The bulk ice absorption and scattering parameters are included with nominal values of 1.0 and 1.05, respectively, and ranges of [0.85, 1.15] and [0.9, 1.2]. They are unconstrained in the fit and the ranges are set to be conservative determined from calibration data.

The two atmospheric neutrino oscillation parameters θ_{23} and Δm_{31}^2 are also included in the fit with nominal values of 47.5047° and $2.475 \times 10^{-3} \text{ eV}^2$, 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. Their ranges are set to [0.0°, 90.0°] and [0.001 eV², 0.004 eV²].

3.3 Analysis Checks

Fitting to data will be 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. At this point changes to the analysis can still be made, if the criteria are not met. This is done to circumvent the so-called *confirmation bias* [57], 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 run, the robustness of the analysis method is tested using pseudo-data that is generated using the MC sets.

3.3.1 Minimization Robustness

To find the set of parameters that describes the data best, 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 octants¹ of θ_{23} , starting from the BFP of the coarse fit. For each individual fit the *MIGRAD* routine of *IMINUIT* [58] is used to minimize the χ^2_{mod} TS defined in Equation 3.1. *Iminuit* is a fast, python compatible minimizer based on the *Mnuit2* C++ library [59]. The individual minimizer settings are shown in Table 3.5.

To test the minimization routine and to make sure it consistently recovers any injected 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*² 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. Even though this range is well within the excluded regions by other experiments, discussed in Section ??, this covers the current sensitive region of the analysis in IceCube DeepCore. 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 3.2 for the 0.6 GeV set. As expected, the fit is always able to recover the injected physcis parameter and the nuisance paramters. Additional plots for the other mass sets can be found in Section B.1.

3.3.2 Ensemble Tests

To estimate the goodness of fit, 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 then fit back with the analysis chain. By comparing the distribution of TS values from this *ensemble* of pseudo-data trials to the TS of the fit to real

[57]: Nickerson (1998), “Confirmation Bias: A Ubiquitous Phenomenon in Many Guises”

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

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

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

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.

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

Do I want additional plots for this (fit diff, LLH distr, minim. stats, param. fits)?

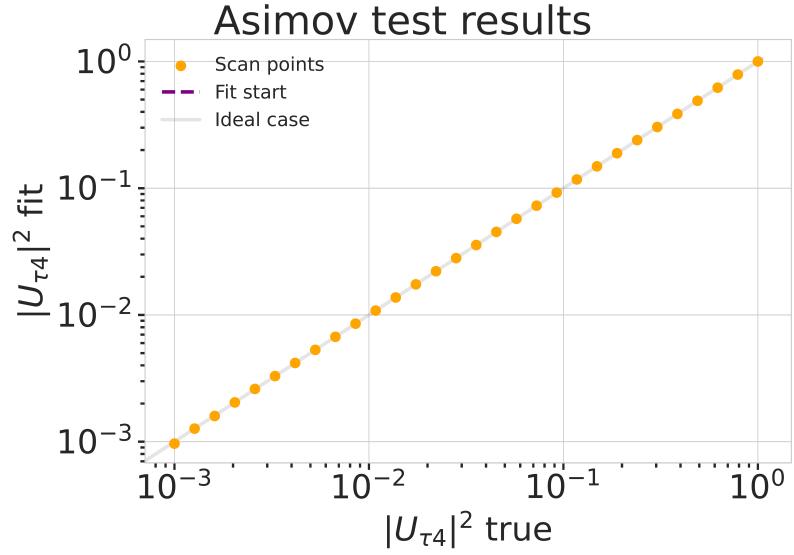


Figure 3.2: Asimov inject/recover test for the 0.6 GeV mass set. Mixing values between 10^{-3} and 10^0 are injected and fit back with the full analysis chain. The injected parameter is always recovered within the statistical uncertainty.

Add bin-wise TS distribution?
Add 3D TS maps?

data, a p-value can be calculated. The p-value is the probability of finding a TS value at least as large as the one from the data fit. Figure 3.3 shows the TS distribution from the ensemble tests for the 0.6 GeV mass set and the observed TS 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 B.1.1.

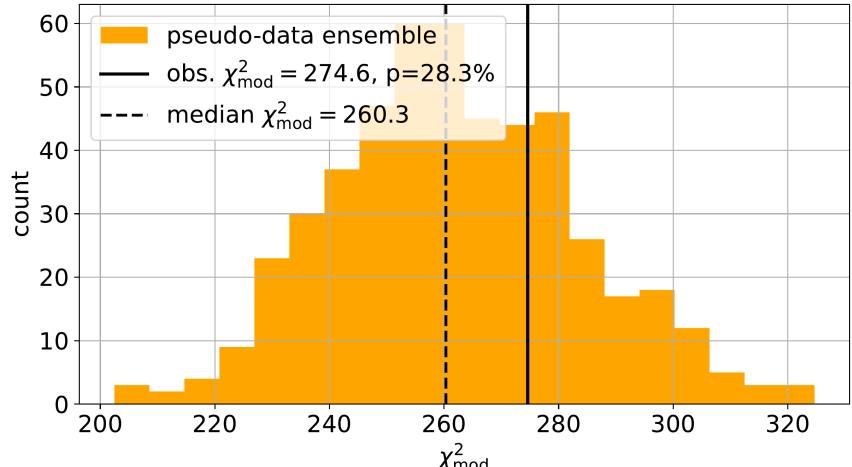


Figure 3.3: Observed fit TS and TS distribution from pseudo-data trials for the 0.6 GeV mass set.

3.4 Results

3.4.1 Best Fit Nuisance Parameters

The resulting nuisance parameter values from the fits are illustrated in Figure 3.4, 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 sets. For parameters that had a Gaussian prior, the 1σ range is also displayed. As was already confirmed during the blind fit procedure, all fitted parameters are within this range, but the Barr h_{π^+} parameter is smaller

and the Barr i_{π^+} is larger than expected, both being very close within the $+1\sigma$ and the -1σ range, respectively. The DIS parameter fits to a smaller value than the nominal and all ice parameters, both hole ice p_0 , and p_1 as well as bulk ice absorption, and scattering are found at values lower than the nominal. The effective ice model parameter, N_{bfr} , prefers a value of ~ 0.74 , indicating that the data is more *BFR*-like (value of 1.0) than *Spice 3.2.1*-like (value of 0.0). For completeness's sake, the explicit results are listed in Table C.1. There, the nominal values and the absolute differences to the best fit value are also presented.

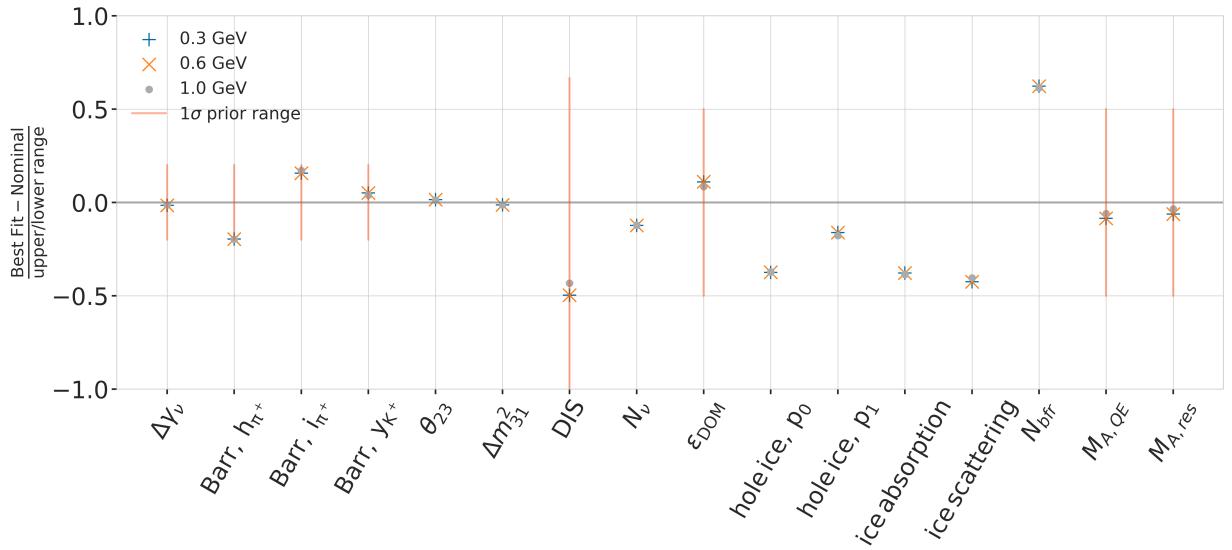


Figure 3.4: xx

3.4.2 Best Fit Parameters and Limits

The fitted mixing values are

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

with their $+1\sigma$ 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 around 0.09. The best fit mixing values and the corresponding upper limits at 68 % and 90 % confidence level (CL) are listed in Table 3.6, also showing the CL at the null hypothesis, which is the probability of excluding the null hypothesis with this fit. The CLs are estimated by assuming that Wilks' theorem [60] holds, meaning that the TS follows a χ^2 distribution with one degree of freedom.

HNL mass	$ U_{\tau 4} _{\text{BFP}}^2$	68 % CL	90 % CL	CL _{null hypo}
0.3 GeV	0.003	0.087	0.194	0.03 %
0.6 GeV	0.080	0.214	0.355	21.29 %
1.0 GeV	0.106	0.238	0.396	37.25 %

Figure 3.5 shows the observed likelihood profile for the 0.6 GeV, which is the difference in χ^2_{mod} between the best fit and each scan point in $|U_{\tau 4}|^2$. Also

fix caption for this figure

sort the variables also by type, same as in the table "best_fit_parameters"?

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

Discuss what it means that the parameters are at these values? Here, or somewhere else?

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

fix table caption

fix once I have them produced

shown is the expected likelihood profile, based on a scan over Asimov data, also produced at the BFP. The observed CLs are slightly larger/smaller than the expected CL. To ensure this is compatible with random fluctuations, the expected likelihood is also profiled for 100 pseudo-data sets, which are generated at the BFP and then fluctuated using both Poisson and Gaussian fluctuations, to include the data and the MC uncertainty, as was already done for the ensemble tests. The resulting CLs are shown as the colored areas and the observed contour is well within the 68 % band, confirming that it is compatible with data fluctuations. Figure C.1 shows the same likelihood profiles and bands for the other two mass sets. For both of them the observed CLs are also slightly larger/smaller than the expected, but still within the 68 % band of the pseudo-data trials, so they are also compatible with random fluctuations.

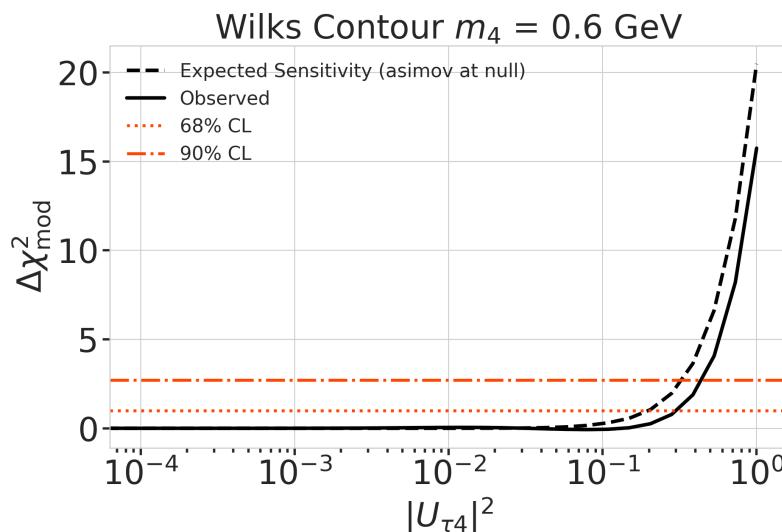


Figure 3.5: xx

fix caption for this figure

make plot with BFPs and limit, comparing to upper limits from other experiments

specify which they are, once I have them

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

add table with reduced chi2 for all 1-d distributions

3.4.3 Data/MC Agreement

At the BFP, the agreement between the data and simulation is probed by comparing the 1-dimensional analysis distributions for PID, energy, and cosine of the zenith angle. As an example, two distributions for the 0.6 GeV mass set are shown in Figure ???. The data is compared to the total MC expectation, which is also split up into its composing parts. 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 ???

3.4.4 Likelihood Coverage

To find the CLs, the profile likelihood was evaluated by applying Wilks' theorem. It isn't guaranteed however, that the theorem holds and it is therefore important to check the *coverage* of the likelihood.

APPENDIX

A

Heavy Neutral Lepton Signal Simulation

A.1 Model Independent Simulation Distributions

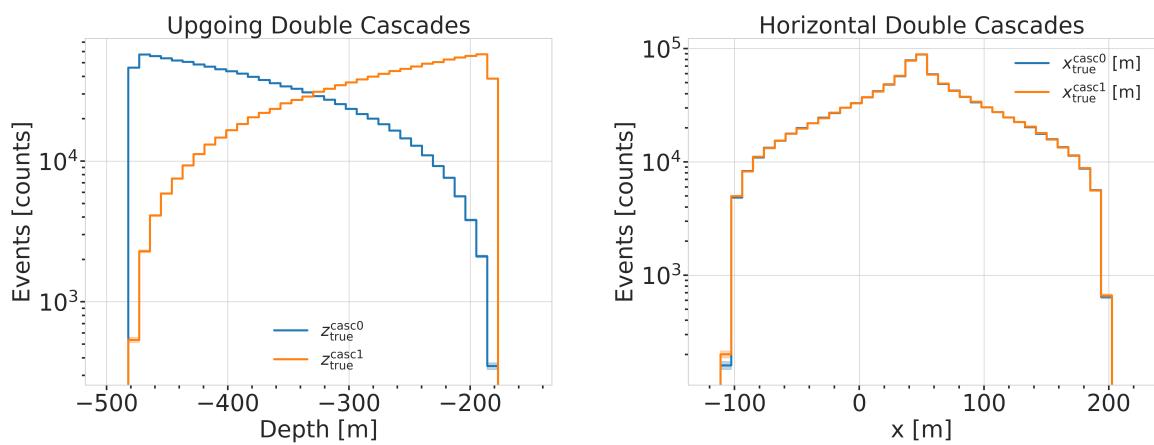


Figure A.1: Generation level distributions of the simplistic simulation sets. Vertical positions (left) and horizontal positions (right) of both sets are shown.

- Re-make plot with x, y for horizontal set one plot!
- Re-make plot with x, y, z for both cascades in one.
- Re-arrange plots in a more sensible way.

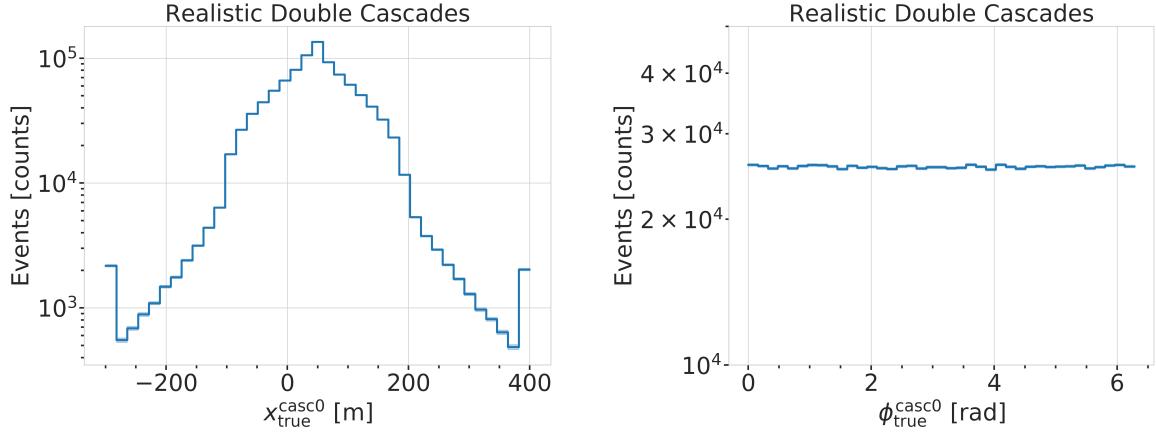


Figure A.2: Generation level distributions of the realistic simulation set. Shown are the cascade x, y, z positions (left) and direction angles (right).

A.2 Model Dependent Simulation Distributions

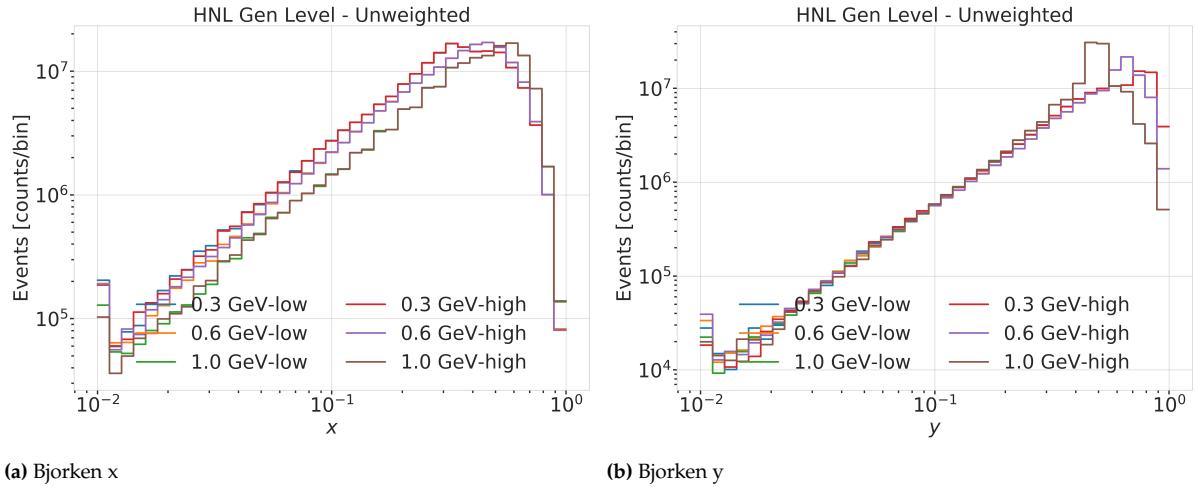


Figure A.3: Generation level distributions of the model dependent simulation.

B

Analysis Checks

B.1 Minimization Robustness

Figure B.1 shows additional Asimov inject/recover tests for the 0.3 GeV and the 1.0 GeV mass sets. The tests were described in Section 3.3.1.

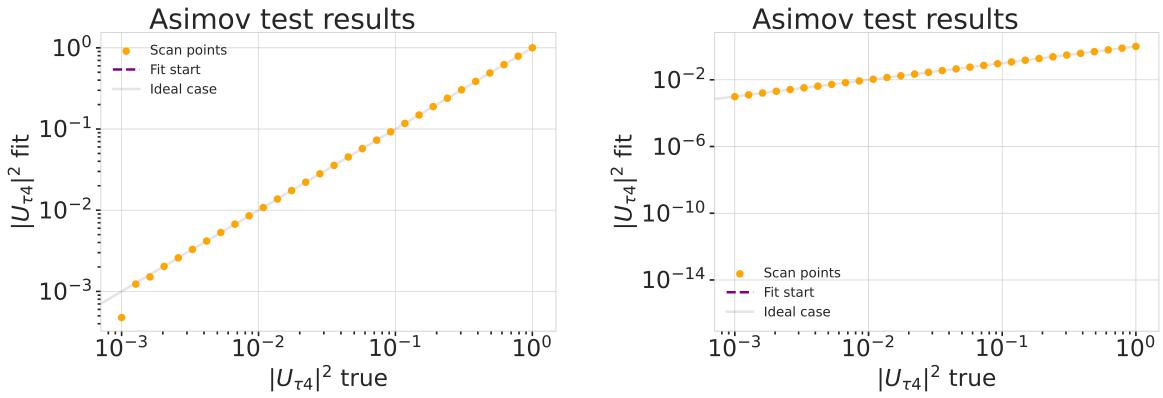


Figure B.1: Asimov inject/recover test for the 0.3 GeV (left) and the 1.0 GeV (right) mass sets. Mixing values between 10^{-3} and 10^0 are injected and fit back with the full analysis chain. The injected parameter is always recovered within the statistical uncertainty.

B.1.1 Ensemble Tests

Figure B.2 shows additional TS distributions from pseudo-data trials and the observed TS from the fit to the data for the ensemble for the 0.3 GeV and the 1.0 GeV mass sets. The tests were described in Section 3.3.2.

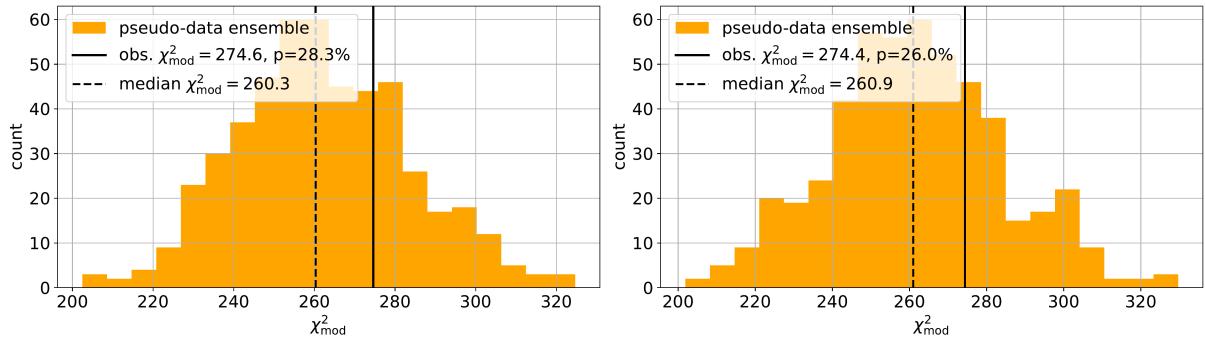


Figure B.2: Observed fit TS and TS distribution from pseudo-data trials for the 0.3 GeV (left) and the 1.0 GeV (right) mass set.

C

Analysis Results

C.0.1 Best Fit Nuisance Parameters

sort these by type of nuisance parameter?

C.0.2 Best Fit Parameters and Limits

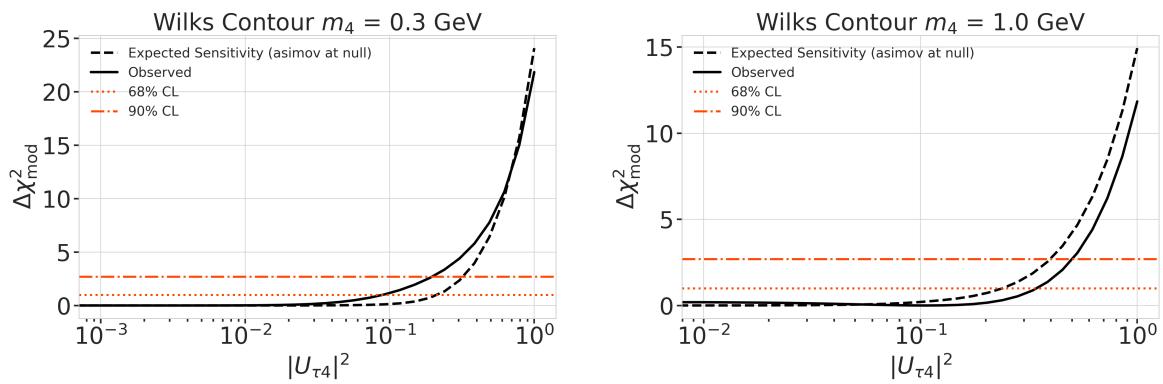


Figure C.1: xx

Table C.1: xx

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
$\Delta\gamma_\nu$	0.000000	-0.007927	-0.007957	-0.006622	0.007927	0.007957	0.006622
Barr h_{π^+}	0.000000	-0.147462	-0.147422	-0.148043	0.147462	0.147422	0.148043
Barr i_{π^+}	0.000000	0.475480	0.475323	0.520716	-0.475480	-0.475323	-0.520716
Barr y_{K^+}	0.000000	0.076161	0.076259	0.057927	-0.076161	-0.076259	-0.057927
$\theta_{23}[\circ]$	47.504700	48.117259	48.118713	48.013150	-0.612559	-0.614013	-0.508450
$\Delta m_{31}^2 [\text{eV}^2]$	0.002475	0.002454	0.002454	0.002455	0.000020	0.000020	0.000019
DIS	0.000000	-0.248768	-0.248845	-0.216247	0.248768	0.248845	0.216247
N_ν	1.000000	0.889145	0.889127	0.889543	0.110855	0.110873	0.110457
ϵ_{DOM}	1.000000	1.021987	1.022017	1.016791	-0.021987	-0.022017	-0.016791
hole ice p_0	0.101569	-0.161352	-0.161260	-0.160133	0.262921	0.262829	0.261702
hole ice p_1	-0.049344	-0.073700	-0.073682	-0.076212	0.024356	0.024338	0.026868
ice absorption	1.000000	0.943262	0.943271	0.942023	0.056738	0.056729	0.057977
ice scattering	1.050000	0.986150	0.986131	0.989374	0.063850	0.063869	0.060626
N_{bfr}	0.000000	0.746674	0.746852	0.736461	-0.746674	-0.746852	-0.736461
$M_{A,\text{QE}}$	0.000000	-0.170430	-0.170677	-0.121335	0.170430	0.170677	0.121335
$M_{A,\text{res}}$	0.000000	-0.125908	-0.126076	-0.071727	0.125908	0.126076	0.071727

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](https://doi.org/10.1016/j.cpc.2021.108018) (cited on page 4).
- [3] 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](https://doi.org/10.1140/epjc/s10052-021-08861-y) (cited on pages 5, 7, 8).
- [4] 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](https://doi.org/10.1007/JHEP07(2014)079) (cited on page 5).
- [5] C. Argüelles. <https://github.com/arguelles/NuXSSplMkr> (cited on page 5).
- [6] 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](https://doi.org/10.1088/0954-3899/36/5/055002) (cited on page 6).
- [7] 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](https://doi.org/10.1103/RevModPhys.93.025010) (cited on page 7).
- [8] <https://github.com/icecube/pisa> (cited on pages 9, 30).
- [9] 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](https://doi.org/10.1103/PhysRevD.108.012014) (cited on pages 13, 20, 23–25).
- [10] 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](https://doi.org/10.1103/PhysRevD.92.023004) (cited on page 14).
- [11] C. Andreopoulos et al. "The GENIE Neutrino Monte Carlo Generator: Physics and User Manual". In: (2015) (cited on pages 15, 25).
- [12] 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](https://doi.org/10.1007/s100529800978) (cited on page 15).
- [13] 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](https://doi.org/10.1088/0954-3899/29/8/369) (cited on page 15).
- [14] 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 15).
- [15] 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](https://doi.org/10.1016/s0168-9002(03)01368-8) (cited on page 15).
- [16] 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](https://doi.org/10.1016/j.astropartphys.2012.09.008) (cited on page 15).
- [17] 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 15).
- [18] D. Heck et al. "CORSIKA: A Monte Carlo code to simulate extensive air showers". In: (Feb. 1998) (cited on page 15).

- [19] 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](https://doi.org/10.1016/j.astropartphys.2012.02.010) (cited on page 15).
- [20] 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](https://doi.org/10.1051/epjconf/201614508001) (cited on page 15).
- [21] C. Kopper et al. <https://github.com/claudiok/clsim> (cited on page 16).
- [22] 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.2019.900050](https://doi.org/10.1109/eScience.2019.900050) (cited on page 16).
- [23] 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](https://doi.org/10.1016/j.nima.2013.01.054) (cited on page 16).
- [24] 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 16, 17, 26).
- [25] 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 16).
- [26] L. G. Henyey and J. L. Greenstein. "Diffuse radiation in the Galaxy." In: *apj* 93 (Jan. 1941), pp. 70–83. doi: [10.1086/144246](https://doi.org/10.1086/144246) (cited on page 16).
- [27] S. Fiedlschuster. "The Effect of Hole Ice on the Propagation and Detection of Light in IceCube". In: (Apr. 2019) (cited on page 16).
- [28] 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](https://doi.org/10.1088/1748-0221/15/06/P06032) (cited on page 17).
- [29] 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 17).
- [30] 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 17).
- [31] 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](https://doi.org/10.25358/openscience-9288) (cited on page 17).
- [32] 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](https://doi.org/10.1088/1748-0221/12/03/P03012) (cited on page 17).
- [33] 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](https://doi.org/10.1016/j.astropartphys.2012.01.004) (cited on pages 17, 18).
- [34] 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 18).
- [35] 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](https://doi.org/10.1088/1748-0221/12/03/P03012) (cited on page 18).
- [36] J. H. Friedman. "Stochastic gradient boosting". In: *Computational Statistics & Data Analysis* 38 (2002), pp. 367–378 (cited on page 19).
- [37] 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](https://doi.org/10.1140/epjc/s10052-022-10721-2) (cited on page 20).

- [38] S. Yu and J. Micallef. "Recent neutrino oscillation result with the IceCube experiment". In: *38th International Cosmic Ray Conference*. July 2023 (cited on pages 20, 30).
- [39] 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 (2021), p. C11001. doi: 10.1088/1748-0221/16/11/C11001 (cited on page 21).
- [40] J. Micallef. <https://github.com/jessimic/LowEnergyNeuralNetwork> (cited on page 21).
- [41] 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 21).
- [42] 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 24).
- [43] 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 24).
- [44] 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 page 24).
- [45] A. Fedynitch et al. <https://github.com/afedynitch/MCEq> (cited on page 24).
- [46] 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 page 24).
- [47] 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 24).
- [48] 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 25).
- [49] 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 25).
- [50] 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 25).
- [51] 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 25).
- [52] 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 page 26).
- [53] N. Kulacz. "In Situ Measurement of the IceCube DOM Efficiency Factor Using Atmospheric Minimum Ionizing Muons". MA thesis. University of Alberta, 2019 (cited on page 26).
- [54] 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 26).
- [55] 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 27).
- [56] 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 30).
- [57] R. S. Nickerson. "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises". In: *Review of General Psychology* 2 (1998), pp. 175–220 (cited on page 33).
- [58] H. Dembinski et al. *scikit-hep/iminuit: v2.17.0*. Version v2.17.0. Sept. 2022. doi: 10.5281/zenodo.7115916 (cited on page 33).
- [59] 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 33).

- [60] 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 35).