

First Search for Heavy Neutral Leptons with IceCube DeepCore

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

im Fach: Physik
Spezialisierung: Experimentalphysik

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

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

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

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

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 ...

Foreword

Some general remarks:

- ▶ Acronyms are introduced in *italic font* the first time they are mentioned and will be used in normal font from then on.
- ▶ Software packages will be introduced in **SMALL CAP FONT** and later used in normal font.
- ▶ On top of the references being listed in its full extent in the bibliography, a selection (but not necessarily all) of them will be highlighted in the margin next to where they appear to smoothen the flow of reading.

Thesis related:

- ▶ mention which chapters contain my own original work and which are collaborative efforts or contain work from others

Todo list

JVS: I would describe simulation sets you produced in past rather than present tense (ORANGE)	1
Make my own DC string positions/distances plot version, viable for the margin? (YELLOW)	1
Re-make plot with all energies (cascades and total, both samples (they are the same)) (RED)	2
Re-make plot with all decay lengths (both samples) (RED)	2
describe why these are shown to highlight some key aspect (uniformity in both energies to sample the whole space, decay length as a result of the z sampling etc..), also add this shortly to the caption (RED)	2
again, describe what is shown and why this is interesting (e.g. energy distribution between the cascades, realistic exponential decay length distribution..) also add this to the captions shortly (RED)	3
JVS: consider remaking figures at the native size so type size is consistent. Also, consider making the legend box semi-transparent so the lines do not obscure the text and markers. (ORANGE)	5
Maybe decide if I want to handle git repo urls/references in a different way.. currently (without date+title) they look ugly in the margin, so I just state them as regular cites, but Cris gave me some other ideas on how to handle them.. (YELLOW)	5
Re-make plot with 3 target masses and better labels/legends etc. (RED)	6
SB: emphasize the cut-off/suppression (ORANGE)	7
Add comparisons of SM cross-sections between NuXSSplMkr and genie? (YELLOW)	7
Calculate max BRs or remove.. (RED)	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 (RED)	8
What is the message of these plots? Make it clear in the text and also in the captions. Potentially cut down to one or two and leave the rest or move them to the appendix. (RED)	8
Combine low/high plots and remove all traces of the separation in the thesis (tables/text/etc.) (ORANGE)	8
JVS: The build-up of the weight expression is hard to follow without knowing where it's going. It may be better to start with the fact that the importance sampling weight is the ratio of PDFs, then write down each pdf, then drill down into each of the terms (basically, the standard "tell me what you're going to tell me, then tell me, then tell me what you told me" scheme). (RED)	8
add 3D data histograms (RED)	12
Add fractions of the different particle types in the bins for benchmark mass/mixing (another table?) (ORANGE)	12
link this to the BG 3-d histogram if I add it? (ORANGE)	12
fix caption (RED)	12
remake without the top row of masked bins? add label to cbar and remove from title (RED)	12
add bin-wise pulls and pull distribution for selection of sets and rest to backup (RED)	14

SB: generally you are missing a lot of references, and I don't think its very nice to read all these ranges and so on in the text. Surely you can just have a table with the systematics you include, their allowed range and prior if applicable? You can choose whether or not you want to discuss some parameters in more detail here, but it should be a discussion of the physics and impact on the analysis. e.g. you can include some plots showing how a change in some of the nuisance parameters changes the event counts in the analysis binning (RED)	14
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 (RED)	15
Need cite here! (RED)	15
I should add some final level effects of some systematics on the 3D binning and maybe discuss how they are different from the signal shape, or so? (ORANGE)	16
FInd first occurrence of "Asimov" and add reference and explain it there (RED)	17
Add 3D BFP-data pull distribution for one mass (they look the same, no?) (RED)	17
fix dimension to fit them in one row! (RED)	18
specify which they are, once I have them (RED)	18
add 1-d data/mc agreement for example mass sample (0.6?) and all 3 analysis variables (RED)	18
add table with reduced chi2 for all 1-d distributions (RED)	18
Cite (again)! (RED)	18
Compare here the best fit oscillation parameters to the FLERCNN results and try to quantify it, stating the pitfalls of the comparitions (statistically fully dependent) (RED)	18
Show best fit hole ice angular acceptance compared to nominal and flasher/in-situ fits, maybe? (YELLOW)	18
Discuss what it means that the parameters are at these values? Here, or somewhere else? (RED)	18
make summary plot (masses and mixing limits on one) and then discuss wrt to other experiments? (RED)	19

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
2 Search for Tau Neutrino Induced Heavy Neutral Lepton Events	11
2.1 Final Level Sample	11
2.1.1 Expected Rates/Events	11
2.1.2 Analysis Binning	12
2.2 Statistical Analysis	13
2.2.1 Test Statistic	13
2.2.2 Physics Parameters	14
2.2.3 Nuisance Parameters	14
2.2.4 Low Energy Analysis Framework	16
2.3 Analysis Checks	16
2.3.1 Minimization Robustness	16
2.3.2 Goodness of Fit	17
2.3.3 Data/MC Agreement	18
2.4 Results	18
2.4.1 Best Fit Nuisance Parameters	18
2.4.2 Best Fit Parameters and Limits	18
Bibliography	21

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	6
1.6	Model dependent simulation generation level distributions	10
2.2	Nuisance parameter mis-modeling impact ranking	15
2.3	Asimov inject/recover test	17
2.4	Pseudo-data trials fit metric distribution (0.6 GeV mass)	18
2.5	Best fit nuisance parameter distances to nominal	19
2.6	Best fit point TS profiles	20

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	9
2.1	Final level background event/rate expectation	12
2.2	Final level signal event/rate expectation	12
2.3	Analysis binning	13
2.4	Nuisance parameter nominal values and fit ranges	16
2.5	Staged minimization routine settings	17
2.6	Best fit mixing values and confidence limits	19

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 estimating 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 ?? and everything but the generation is applied identically to both neutrinos and HNLs.

1.1 Model Independent Simulation

1.1	Model Independent Simulation	1
1.2	Model Dependent Simulation	4

JVS: I would describe simulation sets you produced in past rather than present tense (ORANGE)

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.

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

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 is fixed to the center

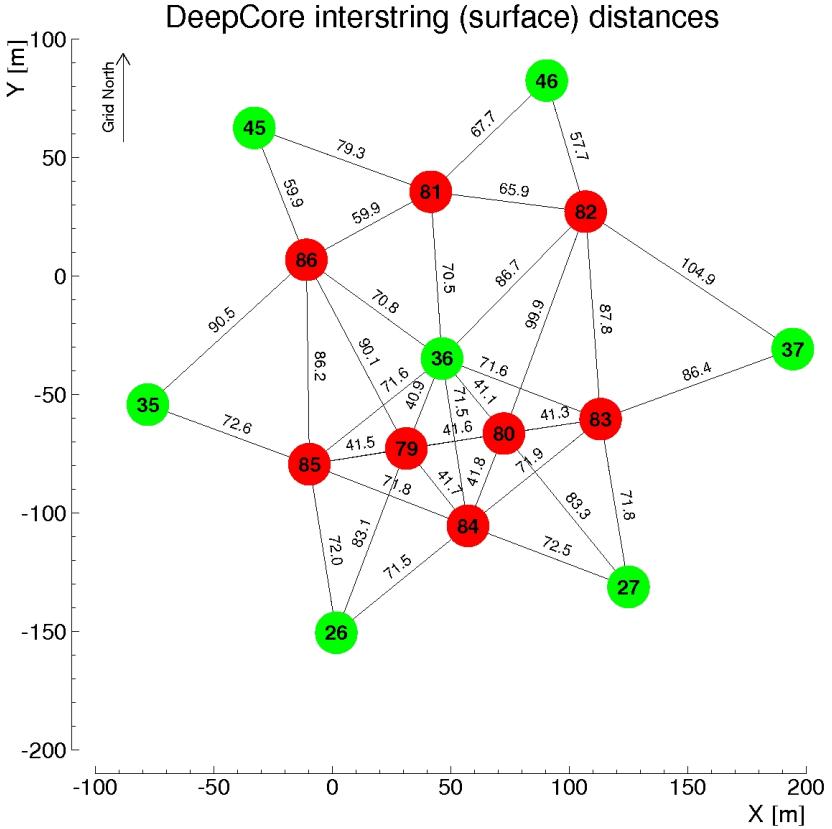


Figure 1.1: Horizontal positions and distances between DeepCore strings. Red strings are instrumented more densely (vertically) and partially have higher quantum efficiency (HQE) DOMs.

of string 81 while the z positions are each sampled uniformly along the axis of the string. Note here that this will therefore not produce a uniform length distribution between the cascades. The positions are defined in the IceCube coordinate system that was already introduced in Section ???. The energies are sampled uniformly between 0.0 GeV and 60.0 GeV. The specific sampling distributions/values for the cascades are listed in Table 1.1. The order of the cascades is chosen such that the lower one is first ($t_0 = 0.0$ ns) and the upper one is second ($t_1 = L/c$), assuming the speed of light c as speed of the heavy mass state, traveling between the two cascades.

Re-make plot with all energies (cascades and total, both samples (they are the same)) (RED)

Re-make plot with all decay lengths (both samples) (RED)

describe why these are shown to highlight some key aspect (uniformity in both energies to sample the whole space, decay length as a result of the z sampling etc..), also add this shortly to the caption (RED)

The second sample is used to investigate the reconstruction performance for horizontal events, where the spacing between DOMs is much larger. The cascades are placed uniformly on a circle centered in DeepCore. The direction is always horizontal and azimuth is defined by the connecting vector of both cascade positions. The energies are again sampled uniformly between 0.0 GeV and 60.0 GeV and the detailed sampling distributions/values are also listed in Table 1.1. Some examples of the generation level distributions of the simplified samples are shown in Figure 1.2, while further distributions can be found in Figure ??.

1.1.2 Realistic Sample

To thoroughly investigate the potential of IceCube DeepCore to detect double cascade events, a more realistic simulation sample is produced that aims to be as close as possible to the expected signal simulation explained in Section 1.2, while still allowing additional freedom to control the double

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.

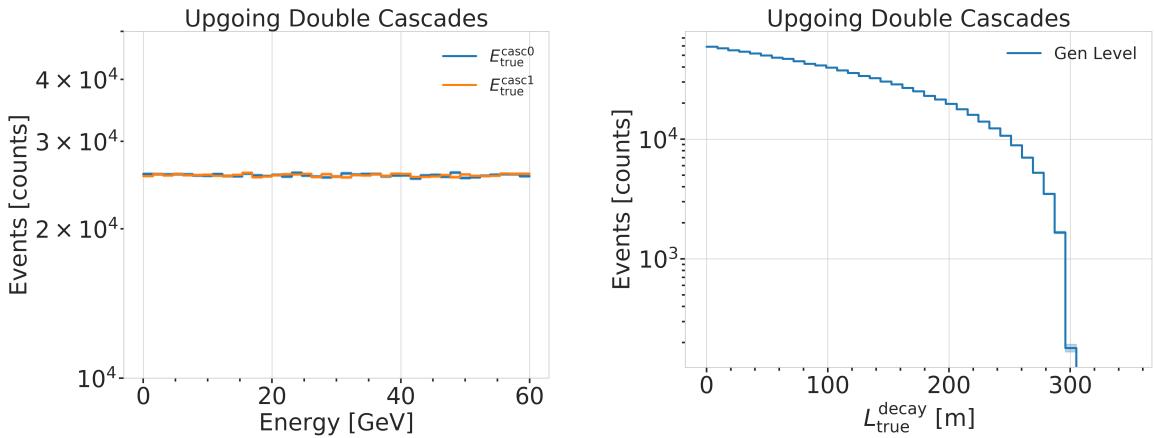


Figure 1.2: Generation level distributions of the simplistic simulation samples. Cascade and total energies (left) and decay lengths (right) of both samples are shown.

cascade kinematics. For this purpose the total energy is sampled from an E^{-2} power law, mimicking the energy spectrum of the primary neutrinos as stated in Section ???. The total energy is divided into two parts, by assigning a fraction between 0 % and 100 % to one cascade and the remaining part to the other cascade. This is a generic approximation of the realistic process described in Section 1.2, and chosen such that the whole sample covers various cases of energy distributions between the two cascades. To efficiently generate events in a way that produces distributions similar to what would be observed with DeepCore, one of the cascade positions is sampled inside the DeepCore volume by choosing its coordinates uniformly on a cylinder that is centered in DeepCore. This is similar to a trigger condition of one cascade always being inside the DeepCore fiducial volume. 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 choosing whether the cascade position that was sampled is the first cascade or the second with a 50 % chance. 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 1.2. Example distributions of the generation level variables are shown in Figure 1.3, while further distributions can be found in Figure ??.

again, describe what is shown and why this is interesting (e.g. energy distribution between the cascades, realistic exponential decay length distribution..) also add this to the captions shortly (RED)

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

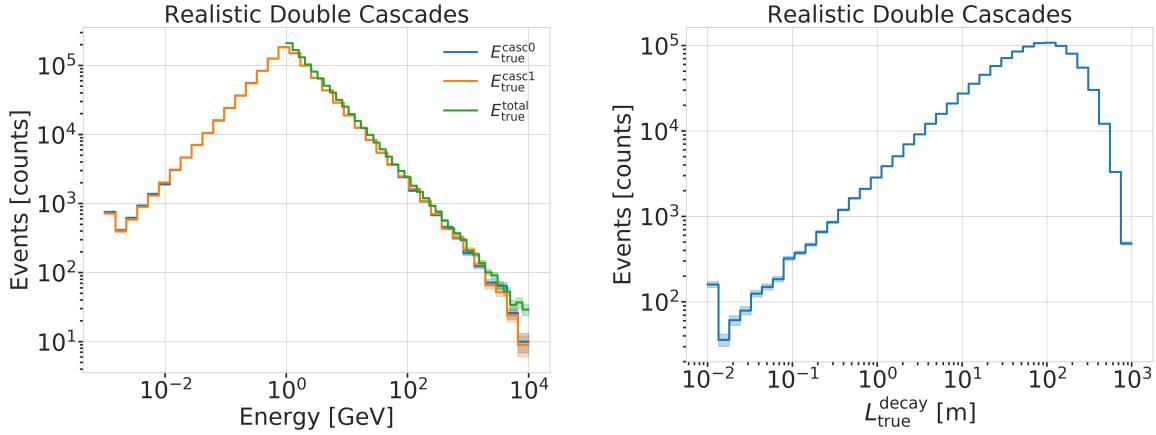


Figure 1.3: Generation level distributions of the simplistic realistic sample. Shown are the cascade and total energies (left) and decay lengths (right).

1.2 Model Dependent Simulation

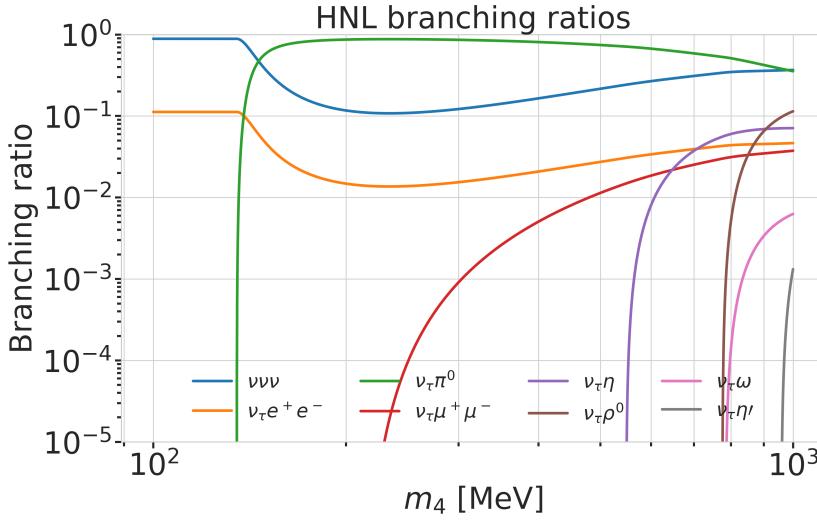
To estimate the HNL event expectation in IceCube DeepCore, depending on the specific model parameters, a generator was developed that is based on the HNL theory introduced in Section ???. For this work, only the interaction with the τ -sector was taken into account ($|U_{\alpha 4}^2| = 0$, $\alpha = e, \mu$), 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 ???. 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.



JVS: consider remaking figures at the native size so type size is consistent. Also, consider making the legend box semi-transparent so the lines do not obscure the text and markers. (ORANGE)

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 distance¹ to produce secondary SM particles, where the decay mode is chosen with a probability given by the mass dependent branching ratios from the kinematically accessible decay modes shown in Figure 1.4. 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.

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 fit to an N-dimensional tensor-product B-spline surface [6] to produce the splines that can be read and used with LI/LW. The tool was modified to produce the custom HNL cross-sections, where the main modification to calculate the cross-sections for the ν_τ -NC interaction into the new heavy mass state, is the addition of a kinematic condition to ensure

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

[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”

Maybe decide if I want to handle git repo urls/references in a different way.. currently (without date+title) they look ugly in the margin, so I just state them as regular cites, but Cris gave me some other ideas on how to handle them.. (YELLOW)

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

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

that there is sufficient energy to produce the heavy mass state. It is the same condition fulfilled for the CC case, where the outgoing charged lepton mass is non-zero. Following [7] (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. The custom version was made part of the open source NuXSSplMkr software and can thus be found in [5]. The result of this kinematic condition is that events cannot be produced for energy, x , y combinations that don’t have sufficient energy going into the outgoing, massive lepton. This results in a reduction of the cross-section towards lower energies, which scales with the assumed mass of the HNL. This effect can clearly be seen in Figure 1.5.

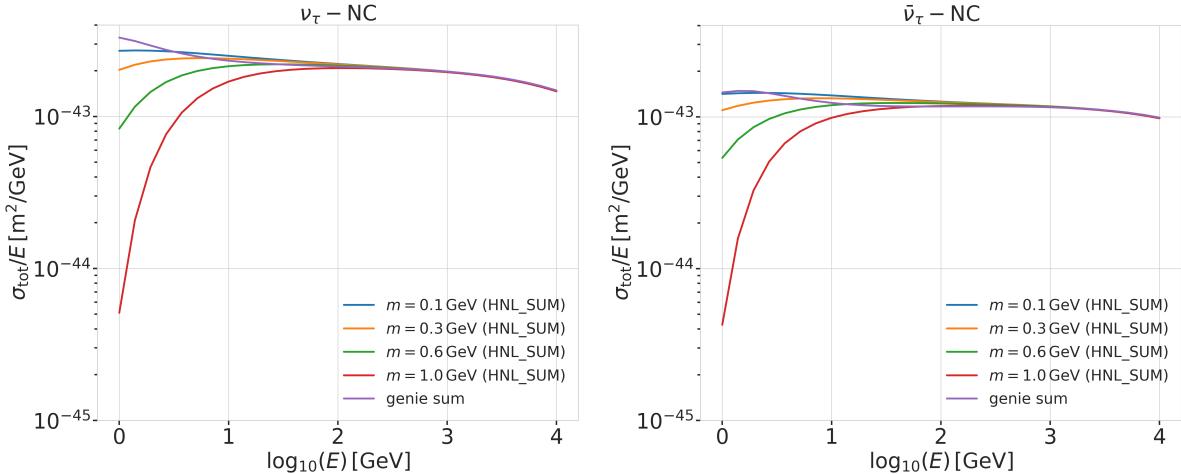


Figure 1.5: Custom HNL total cross-sections for the four target masses compared to the total ($\nu_\tau/\bar{\nu}_\tau$ NC) cross-section used for SM neutrino simulation production with GENIE.

Re-make plot with 3 star
get masses and better
labels/legends etc. (RED)

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 ???. The double-differential ($d\sigma/dxdy$) and total (σ) cross-sections were produced for the chosen target HNL masses and then splined in energy, x , and y and just in the energy, respectively. 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. They agree above a certain energy (~ 200 GeV), where the modification should not have any effect on the cross-sections, which is the desired result of using the identical input PDFs and confirms that the unmodified cross-sections produced with NuXSSplMkr agree with the GENIE cross-sections, which

were used to generate the SM background MC.

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 $\sin^2 \theta_w = 0.2229$ [8], where θ_w is the Weinberg angle.

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

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

SB: emphasize the cut-off/suppression (ORANGE)

Add comparisons of SM cross-sections between NuXSSplMkr and genie? (YELLOW)

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

Channel	Opens	$\hat{B}\bar{R}$
$\nu_4 \rightarrow \nu_\tau \nu_\alpha \bar{\nu}_\alpha$	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 or remove.. (RED)

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

with the constants defined as

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

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

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

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

3-Body Decay Kinematics with MadGraph

The specific MadGraph version used to produce the 3-body decay kinematics is **MADGRAPH4 v3.4.0** [9], which uses the decay diagrams calculated with **FEYNRULES 2.0** [10] and the Lagrangians derived in [3] as input. The *Universal FeynRules Output (UFO)* from **EFFECTIVE_HEAVYN_MAJORANA_v103** were used for our calculation. For each mass and corresponding decay channels, we produce 1×10^6 decay kinematic variations in the rest frame and store those in a text file. During event generation, we uniformly select an event from that list, to simulate the decay kinematics of a 3-body decay.

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 leads 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 . Figure 1.6 shows some selected generation level distributions. Additional distributions can be found in Figure ??.

1.2.3 Weighting Scheme

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

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

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

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

What is the message of these plots? Make it clear in the text and also in the captions. Potentially cut down to one or two and leave the rest or move them to the appendix. (RED)

Combine low/high plots and remove all traces of the separation in the thesis (tables/text/etc.) (ORANGE)

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

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

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

?? is also used. 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 ?? 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

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

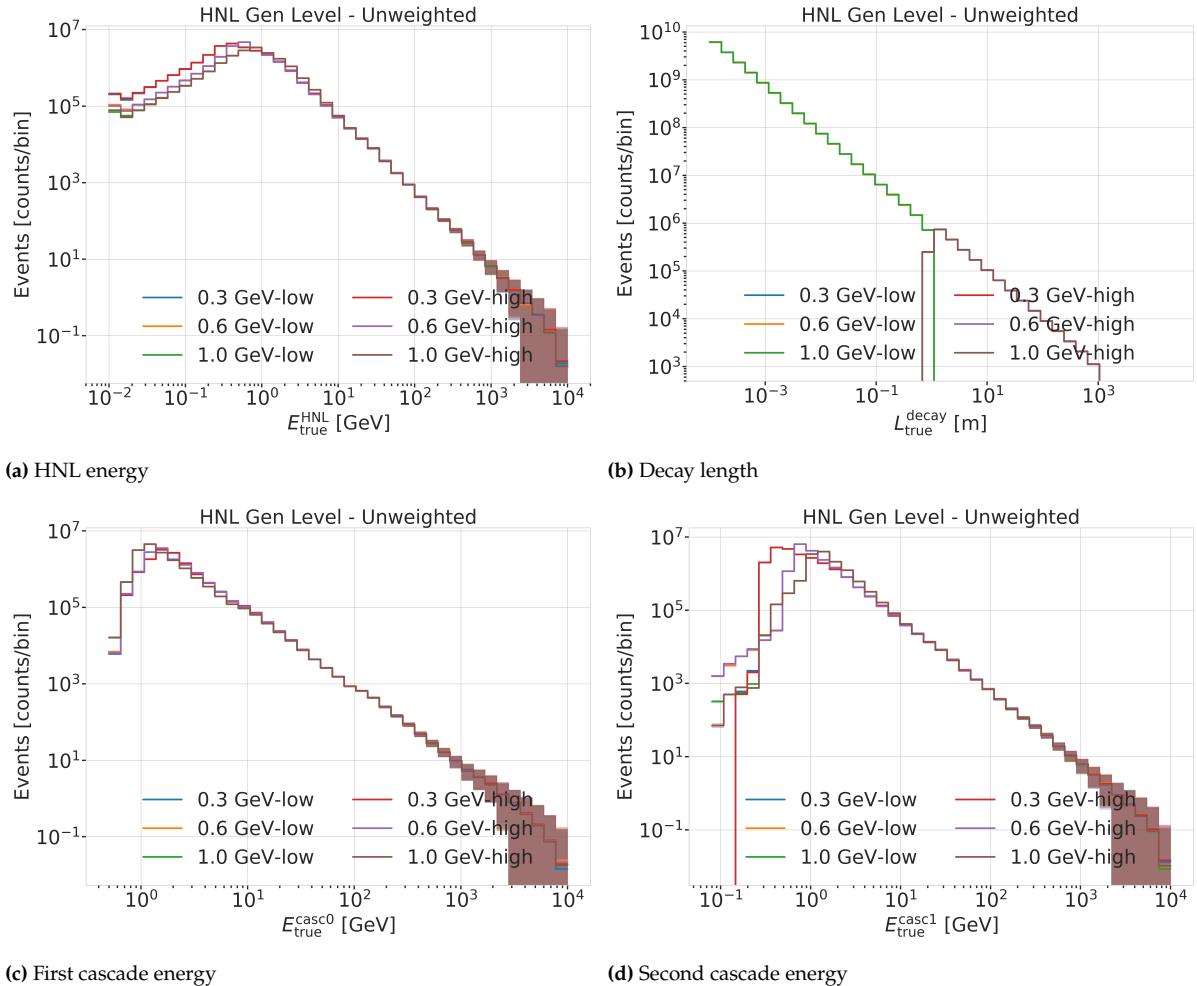


Figure 1.6: Generation level distributions of the model dependent simulation.

If this additional weighting factor is multiplied to a generation weight with units m^2 (like in Equation ??), 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).

Search for Tau Neutrino Induced Heavy Neutral Lepton Events

2

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

2.1 Final Level Sample

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

To get the neutrino expectation, the MC events are weighted according to their generation weight introduced in Section ??, multiplied by the total lifetime, and the expected neutrino flux. For the correct expectation at the detector, the events have to be weighted by the oscillation probability, depending on their energy and their distance traveled from the atmosphere to the detector. The oscillation probabilities are calculated using a PYTHON implementation of the calculations from [11], which use the matter profile of the Earth following the *Preliminary Reference Earth Model (PREM)* [12] as input. Apart from the energy and the distance, the two relevant parameters defining the oscillation probabilities are the atmospheric neutrino oscillation parameters θ_{23} and Δm_{31}^2 . Since the HNL events originate from the tau neutrinos that were produced as muon neutrinos in the atmosphere and then oscillated into ν_τ , this weighting is also applied in addition to the specific weighting scheme for the HNL events described in Section 1.2.3, which itself is defined by the mixing $|U_{\tau 4}^2|$ and the mass m_4 .

2.1	Final Level Sample	11
2.2	Statistical Analysis	13
2.3	Analysis Checks	16
2.4	Results	18

[11]: Barger et al. (1980), “Matter effects on three-neutrino oscillations”

[12]: Dziewonski et al. (1981), “Preliminary reference Earth model”

2.1.1 Expected Rates/Events

The rates and the expected number of events for the SM background are shown in Table 2.1 with around 175000 total events expected in the 10 years. Only data marked as good is used for the analysis, where *good* refers to

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

Table 2.1: Final level rates and event expectation of the SM background particle types.

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

Table 2.2 shows the rates and expected number of events for the HNL signal simulation. The expectation depends on the mass and the mixing and shown here are two example mixings for all the three masses that are being tested in this work. A mixing of 0.0 would result in no HNL events at all. It can already be seen that for the smaller mixing of $|U_{\tau 4}|^2 = 10^{-3}$ the expected number of events is very low, while at the larger mixing of $|U_{\tau 4}|^2 = 10^{-1}$ the number is comparable to the amount of atmospheric muons in the background sample.

Table 2.2: 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	3.3	975 ± 2
0.6 GeV	3.1	895 ± 2
1.0 GeV	2.5	731 ± 2
$ U_{\tau 4} ^2 = 10^{-3}$		
0.3 GeV	0.006	1.67 ± 0.01
0.6 GeV	0.022	6.44 ± 0.01
1.0 GeV	0.025	7.27 ± 0.01

add 3D data histograms (RED)

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

[13]: Yu et al. (2023), “Recent neutrino oscillation result with the IceCube experiment”

link this to the BG 3-d histogram if I add it? (ORANGE)

fix caption (RED)

remake without the top row of masked bins? add label to cbar and remove from title (RED)

2.1.2 Analysis Binning

An identical binning to the analysis performed in [13] is used. In total, there are three bins in PID (cascade like, mixed, and track like), 12 bins in reconstructed energy, and 8 bins in cosine of the reconstructed zenith angle as specified in Table 2.3. Extending the binning towards lower energies or increasing the number of bins in energy or cosine of the zenith angle did not improve the HNL sensitivities significantly, because the dominant signal region is already covered with a sufficiently fine binning to observe the shape and magnitude of the HNL events. This can be seen in the middle panel of Figure 2.1.

Some low energy bins in the cascade like region have very low MC expectations (<1 event) and are therefore not taken into account in the analysis, to prevent unwanted behavior in the fit. Those are shown in dark green in the three dimensional histogram.

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

Table 2.3: Three dimensional binning used in the analysis. All variables are from the FLERCNN reconstruction explained in Section ??.

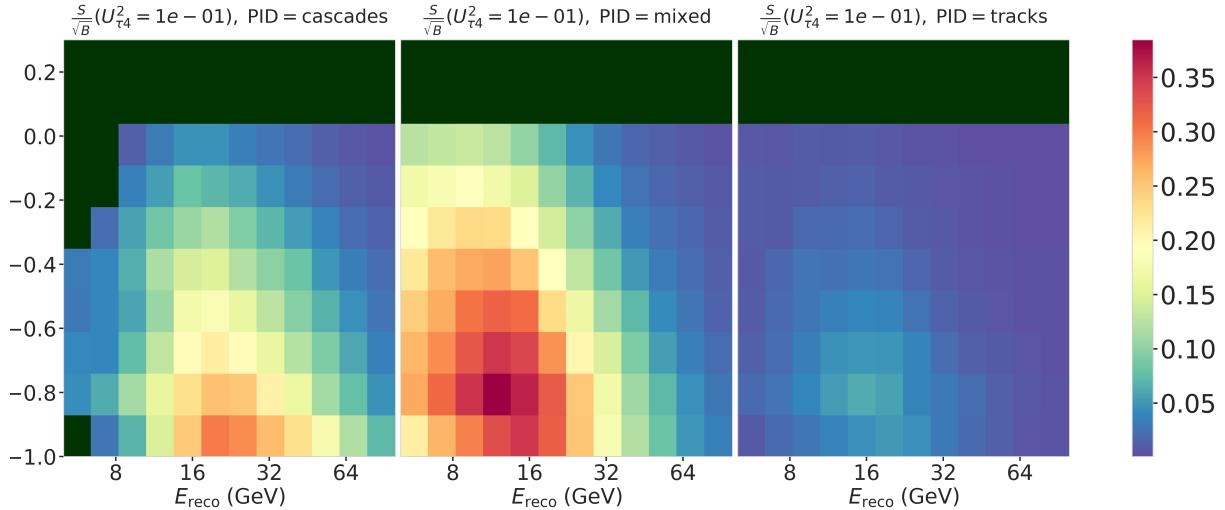


Figure 2.1

2.2 Statistical Analysis

2.2.1 Test Statistic

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

$$\chi^2_{\text{mod}} = \sum_{i \in \text{bins}} \frac{(N_i^{\text{exp}} - N_i^{\text{obs}})^2}{N_i^{\text{exp}} + (\sigma_i^{\nu})^2 + (\sigma_i^{\mu})^2 + (\sigma_i^{\text{HNL}})^2} + \sum_{j \in \text{syst}} \frac{(s_j - \hat{s}_j)^2}{\sigma_{s_j}^2}, \quad (2.1)$$

as the fit metric. It is designed such that taking the difference between a free fit and a fit with fixed parameters based on a chosen hypothesis, $\Delta\chi^2_{\text{mod}}$, can directly be used as a *test statistic* (TS) for hypothesis testing, due to its asymptotic behavior. The total even expectation is $N_i^{\text{exp}} = N_i^{\nu} + N_i^{\mu} + N_i^{\text{HNL}}$, where N_i^{ν} , N_i^{μ} , and N_i^{HNL} are the expected number of events in bin i from neutrinos, atmospheric muons, and HNLs, while N_i^{obs} is the observed number of events in the bin. The expected number of events from each particle type is calculated by summing the weights of all events in the bin $N_i^{\text{type}} = \sum_i^{\text{type}} \omega_i$, with the statistical uncertainty being $(\sigma_i^{\text{type}})^2 = \sum_i^{\text{type}} \omega_i^2$. The additional term in Equation 2.1 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.

2.2.2 Physics Parameters

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

2.2.3 Nuisance Parameters

interlude ...

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

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

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

add bin-wise pulls and pull distribution for selection of sets and rest to backup (RED)

SB: generally you are missing a lot of references, and I don't think its very nice to read all these ranges and so on in the text. Surely you can just have a table with the systematics you include, their allowed range and prior if applicable? You can choose whether or not you want to discuss some parameters in more detail here, but it should be a discussion of the physics and impact on the analysis. e.g. you can include some plots showing how a change in some of the nuisance parameters changes the event counts in the analysis binning (RED)

Treatment of Detector Systematic Uncertainties

Since the variations related to the detector calibration uncertainties introduced in Section ?? are estimated by simulating MC at discrete values of the systematic parameters, a method to derive continuous variations is needed to perform the fit. The method applied here was initially introduced in [14] and first used in the low energy sterile neutrino search in [15] (section 7.4.3). Using a *likelihood-free inference* technique, re-weighting factors are found for every event in the nominal MC sample, given a specific choice of detector systematic parameters. These factors quantify how much more or less likely the event would be for the corresponding change in detector response from the nominal parameters. Without going into the details of the method, which were already exhaustively discussed in [14] and [15], the performance is assessed here for the HNL signal simulation. In order to do so, the weights are applied to the nominal MC samples, choosing the detector systematic values used to produce the discrete samples and the resulting event expectations are compared to the expectations from the individual, discrete MC samples. The bin counts are compared by calculating the pull defined as

$$p = \frac{N_{\text{reweighted}} - N_{\text{sys}}}{\sqrt{\sigma_{\text{reweighted}}^2 + \sigma_{\text{sys}}^2}}, \quad (2.2)$$

where N are the bin-wise event expectations and σ are their MC uncertainty. For the SM BG simulation, the performance was already investigated in [16] (section 7.4.4, appendix B5) and the re-weighted nominal MC was shown to be in agreement with the discrete systematic sets at a sufficient level. Figure ?? shows the bin-wise pulls for the 1.0 GeV HNL mass sample at a mixing of 0.1 for a selection of the discrete systematic samples, where the DOM efficiency and the bulk ice absorption was varied by $\pm 10\%$. The distributions should follow a standard normal distribution, and no strong clustering or systematic deviations. It can be seen that spread of the distribution is slightly larger than 1.0 and the center is close to 0.0 as expected. A similar performance is found for the additional systematic sets and can be found in Section ??.

Free Parameters

To decide which systematic uncertainties should be included in the fit, we test the potential impact they have on the TS if they are neglected. The test is performed by creating Asimov data using the BG simulation and the HNL simulation of the 1.0 GeV mass sample at a mixing value of 0.1, which is chosen as a benchmark physics parameter, but the explicit choice does not have a significant impact on the test. The systematic parameter of interest is set to a value above its nominal expectation, either pulled up by $+1\sigma$ or by an educated estimate for parameters without a well-defined uncertainty. A fit is performed fixing the systematic parameter of interest and leaving all additional parameters free. The resulting TS is the mis-modeling significance between this fit and 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 2.2 shows the resulting significances of one of these tests. The parameters tested are the systematic parameters introduced in Section ?? and the atmospheric oscillation parameters mentioned in Section 2.1. In the final selection of free parameters the Barr h_{π^+} parameter was also left free and the ice absorption is still kept free, despite showing a small significance. This is done because the bulk ice parameters are not well constrained and are known to have a large impact, which might be concealed in this idealized test, due to correlations with the other parameters. In this test, the effect of correlations is challenging to consider, because only the impact of one parameter is tested at a time, using the overall mis-modeling significance as a measure. The mis-modeling could be reduced by a correlated parameter capturing the effect of the parameter of interest. For this reason a very conservative threshold of 0.1σ is chosen and some parameters below the threshold are still left free in the fit.

All nuisance parameters that are left free in the fit are summarized in Table 2.4, showing their nominal values, the allowed fit ranges, and their Gaussian prior, if applicable. The scaling parameter N_ν is included to account for the overall normalization of the neutrino rate, and it has the identical effect on the SM neutrino events and the BSM HNL events, because they both originate from the same neutrino flux. Despite being known to $\sim 5\%$ in this energy range, there is no prior applied to this parameter, because the fit itself is able to constrain it well, which can be seen by the large impact it shows in Figure 2.2. Concerning the atmospheric neutrino flux, the CR power law flux correction factor $\Delta\gamma_\nu$ introduced in Section ?? is included with nominal value of 0.0 which corresponds to the baseline flux model by Honda *et al* [17]. A slightly conservative prior of 0.1 is applied to the parameter, while latest measurements show an uncertainty of 0.05 [18].

All the detector systematic uncertainties discussed in Section ?? are included in the fit. The DOM efficiency ϵ_{DOM} is constrained by a Gaussian prior with a width of 0.1, which is a conservative estimate based on the studies of the optical efficiency using minimum ionizing muons from [19, 20].

The two atmospheric neutrino oscillation parameters θ_{23} and Δm_{31}^2 are also included in the fit with nominal values of 47.5° and $2.48 \times 10^{-3} \text{ eV}^2$ [13], respectively. Since they govern the shape and the strength of the tau neutrino

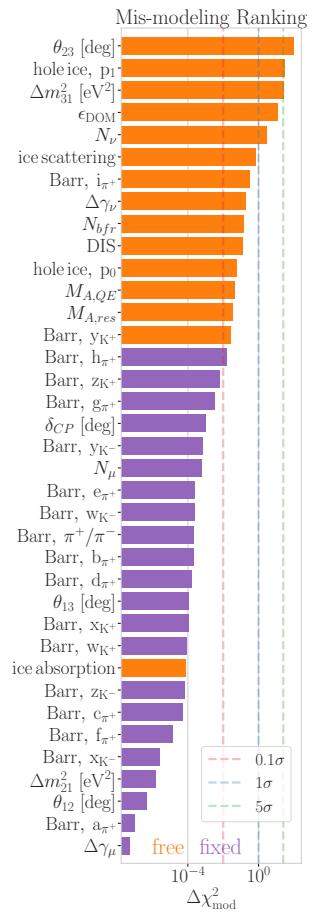


Figure 2.2: Mis-modelling impact ranking of the systematic parameters. The mis-modelling is calculated as the fit metric difference between a fit with the parameter fixed at its nominal value and a fit with the parameter pulled up by $+1\sigma$. The test was performed using Asimov data of the 1.0 GeV mass sample at a reference mixing of 0.1.

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 (RED)

Need cite here! (RED)

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

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

Table 2.4: Systematic uncertainty parameters that are left free to float in the fit. Their allowed fit ranges are shown with the nominal value and the Gaussian prior width if applicable.

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

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

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

[13]: Yu et al. (2023), “Recent neutrino oscillation result with the IceCube experiment”

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

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

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

flux, by defining the oscillation from ν_μ to ν_τ , they are also relevant for the HNL signal shape.

2.2.4 Low Energy Analysis Framework

The analysis is performed using the PISA [21] [22] 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 samples, which can then be compared to the observed data. The expectation for each MC sample is calculated by applying physics and nuisance parameter effects in a stage-wise manner, before combining them to the final expectation.

2.3 Analysis Checks

Fitting to data is performed in a *blind* manner, where the analyzer does not immediately see the fitted physics and nuisance parameter values, but first checks that a set of pre-defined *goodness of fit* (GOF) criteria are fulfilled. This is done to circumvent the so-called *confirmation bias* [23], where the analyzer might be tempted to construct the analysis in a way that confirms their expectation. After the GOF criteria are met to satisfaction, the fit results are unblinded and the full result can be revealed. Before these blind fits to data are performed, the robustness of the analysis method is tested using pseudo-data that is generated from the MC.

2.3.1 Minimization Robustness

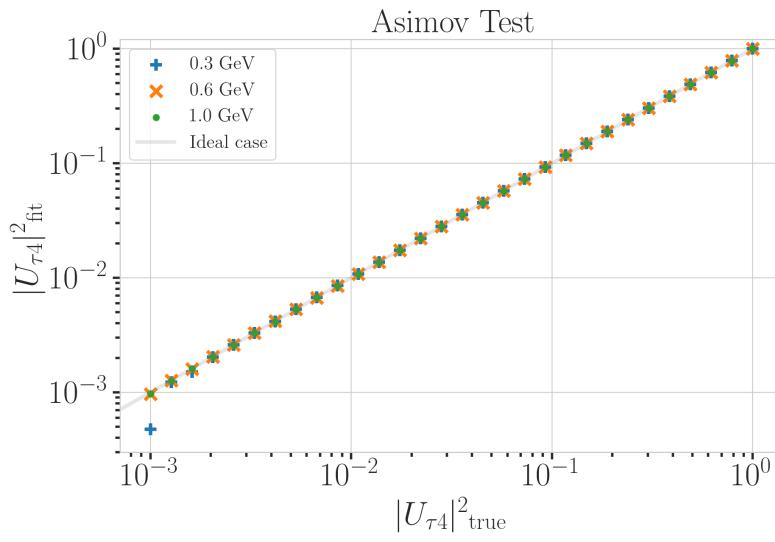
1: There is a degeneracy between the lower octant ($\theta_{23} < 45^\circ$) and the upper octant ($\theta_{23} > 45^\circ$), which can lead to fit metric minima (local and global) at two positions that are mirrored around 45° in θ_{23} .

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

To find the set of parameters that best describes the data, a staged minimization routine is used. In the first stage, a fit with coarse minimizer settings is performed to find a rough estimate of the *best fit point* (BFP). In the second stage, the fit is performed again in both octants¹ of θ_{23} , starting from the BFP of the coarse fit. For each individual fit the *MIGRAD* routine of *iminuit* [24]

is used to minimize the χ^2_{mod} fit metric defined in Equation 2.1. iminuit is a fast, python compatible minimizer based on the MINUIT2 C++ library [25]. The individual minimizer settings for both stages are shown in Table 2.5.

To test the minimization routine and to make sure it consistently recovers any physics parameters, pseudo-data sets are produced from the MC by choosing the nominal nuisance parameters and specific physics parameters, without adding any statistical or systematic fluctuations to it. These so-called *Asimov*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. 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 2.3 for all three mass samples. As desired, the fit is always able to recover the injected physics parameter and the nuisance parameters within the statistical uncertainty or at an insignificant fit metric difference.



[25]: 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 2.5: 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.

Find first occurrence of "Asimov" and add reference and explain it there (RED)

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

Figure 2.3: Asimov inject/recover test results for all three mass samples. Mixing values between 10^{-3} and 10^0 are injected and fit back with the full analysis chain. The injected parameter is always recovered within the statistical uncertainty or at an insignificant fit metric difference.

2.3.2 Goodness of Fit

To estimate the GOF, pseudo-data is generated from the MC by injecting the BFP parameters as true parameters and then fluctuating the expected bin counts to account for MC uncertainty and Poisson fluctuations in data. First, the expectation value of each bin is drawn from a Gaussian distribution centered at the nominal expectation value with a standard deviation corresponding to the MC uncertainty of the bin. Based on this sampled expectation value, each bin count is drawn from a Poisson distribution, independently, to get the final pseudo-data set. These pseudo-data sets are analyzed with the same analysis chain as the real data, resulting in a final fit metric value for each pseudo-data set. By comparing the distribution of fit metric values from this *ensemble* of pseudo-data trials to the fit metric of the fit to real data, a p-value can be calculated. The p-value is the probability of finding a value of the fit metric at least as large as the one from the data fit. Figure 2.4 shows the distribution from the ensemble tests for the 0.6 GeV mass sample and the observed value from the fit, resulting in a p-value

Add 3D BFP-data pull distribution for one mass (they look the same, no?) (RED)

of 28.5 %. The p-values for the 0.3 GeV and 1.0 GeV are 28.3 % and 26.0 %, respectively, and the corresponding plots are shown in Section ???. Based on this test, it is concluded that the fit result is compatible with the expectation from the ensemble of pseudo-data trials.

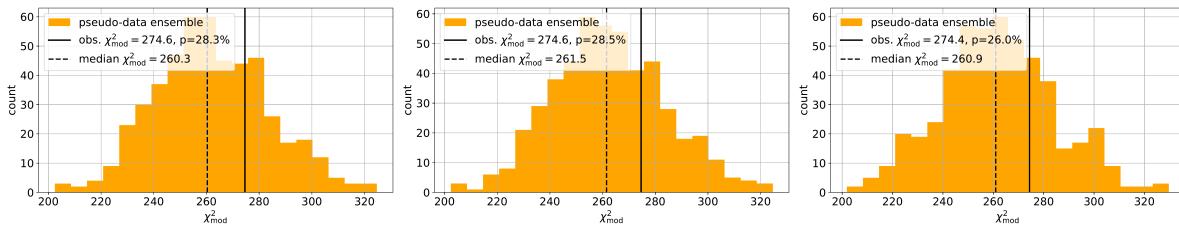


Figure 2.4: Observed fit metric and fit metric distribution from pseudo-data trials for the 0.6 GeV mass sample.
fix dimension to fit them
in one row! (RED)

2.3.3 Data/MC Agreement

At the BFP, the agreement between the data and simulation is probed by comparing the one dimensional analysis distributions for PID, energy, and cosine of the zenith angle. As an example, two distributions for the 0.6 GeV mass sample are shown in Figure ???. The data is compared to the total MC expectation, which is also split up into the individual signal and background components for illustration. Good agreement can be observed in the pull distributions, and is quantified by a reduced χ^2 , which is close to 1.0 for all distributions. The reduced χ^2 for all investigated distributions is listed in Table ???, while the distributions themselves can be found in Section ???.

specify which they are,
once I have them (RED)

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

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

Cite (again)! (RED)

Compare here the best
fit oscillation parameters
to the FLERCNN results
and try to quantify it,
stating the pitfalls of the
comparisons (statistically
fully dependent) (RED)

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

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

2.4 Results

2.4.1 Best Fit Nuisance Parameters

The resulting nuisance parameter values from the fits are illustrated in Figure 2.5, where the differences to the nominal values are shown, normalized by the distance to the closest boundary. The results from all three fits are shown in the same plot and the fits prefer values of the same size for all three mass samples. For parameters that have a Gaussian prior, the 1σ range is also displayed. As was already confirmed during the blind fit procedure, all fitted parameters are within this range. The effective ice model parameter, N_{bfr} , prefers a value of ~ 0.74 , indicating that the data fits better to an ice model that includes real birefringence effects. For completeness, the explicit results are listed in Table ???. There, the nominal values and the absolute differences to the best fit value are also presented.

2.4.2 Best Fit Parameters and Limits

The fitted mixing values are

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

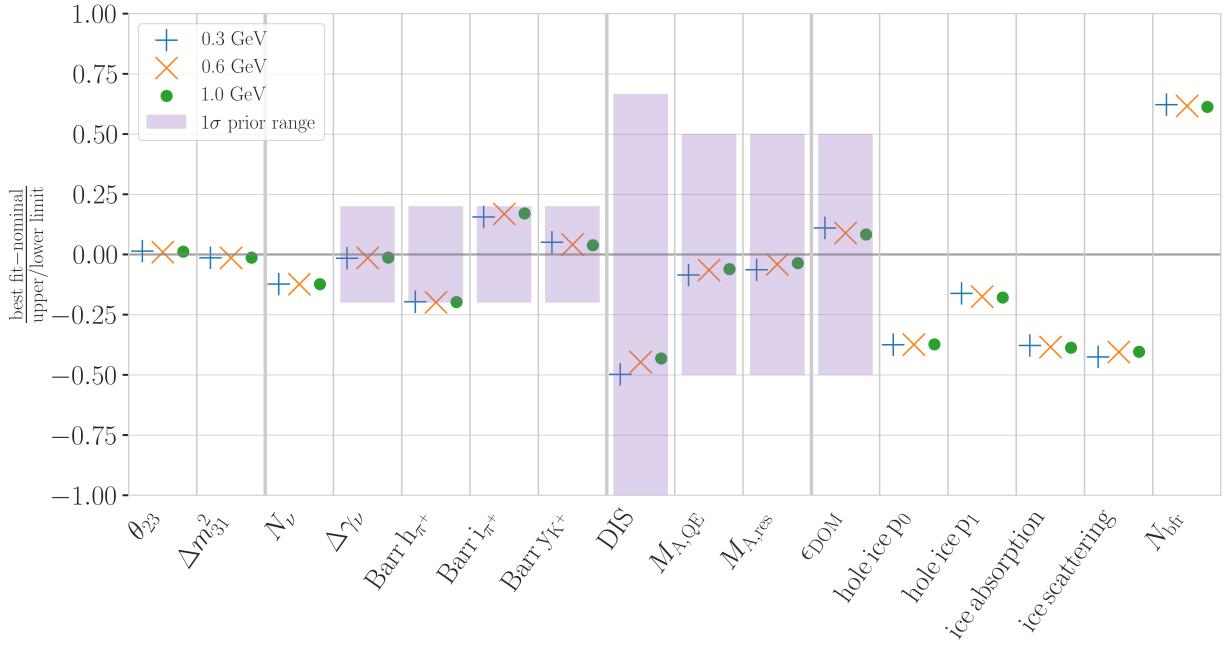


Figure 2.5: Best fit nuisance parameter distances to the nominal values, normalized by the distance to the closest boundary. For parameters with a Gaussian prior, the $+1\sigma$ range is also shown.

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 of 0.08 and 0.106, respectively. The best fit mixing values and the corresponding upper limits at 68 % and 90 % *confidence level* (CL) are listed in Table 2.6, also showing the *p*-value to reject the null hypothesis. The CLs and *p*-value are estimated by assuming that *Wilks' theorem* [26] holds, meaning that the TS follows a χ^2 distribution with one degree of freedom.

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

Table 2.6: Best fit mixing values and the corresponding upper limits at 68 % and 90 % confidence level, as well as the *p*-value to reject the null hypothesis, estimated by assuming that Wilks' theorem holds.

Figure 2.6 shows the observed TS profiles as a function of $|U_{\tau 4}|^2$ for all three fits. The TS profile is the difference in χ^2_{mod} between the free fit and a fit where the mixing is fixed to a specific value. Also shown is the expected TS profile, based on 100 pseudo-data trials, produced at the BFP and then fluctuated using both Poisson and Gaussian fluctuations, to include the data and the MC uncertainty as was explained in Section 2.3.2. The Asimov expectation and the 68 % and 90 % bands are shown and the observed TS profiles lie within the 68 % band for all three, confirming that they are compatible with statistical fluctuations of the observed data. For the 0.3 GeV fit, the observed contour is slightly tighter than the Asimov expectation, meaning that the observed upper limits in $|U_{\tau 4}|^2$ are slightly stronger than expected. For the 0.6 GeV the opposite is the case and the observed upper limit is therefore slightly weaker than expected. For the 1.0 GeV fit, the observed upper limit is very close to the Asimov expectation in the region where the 68 % and 90 % CLs thresholds are crossed. The observed upper limits are also shown in Table 2.6.

make summary plot
(masses and mixing limits on one) and then discuss wrt to other experiments? (RED)

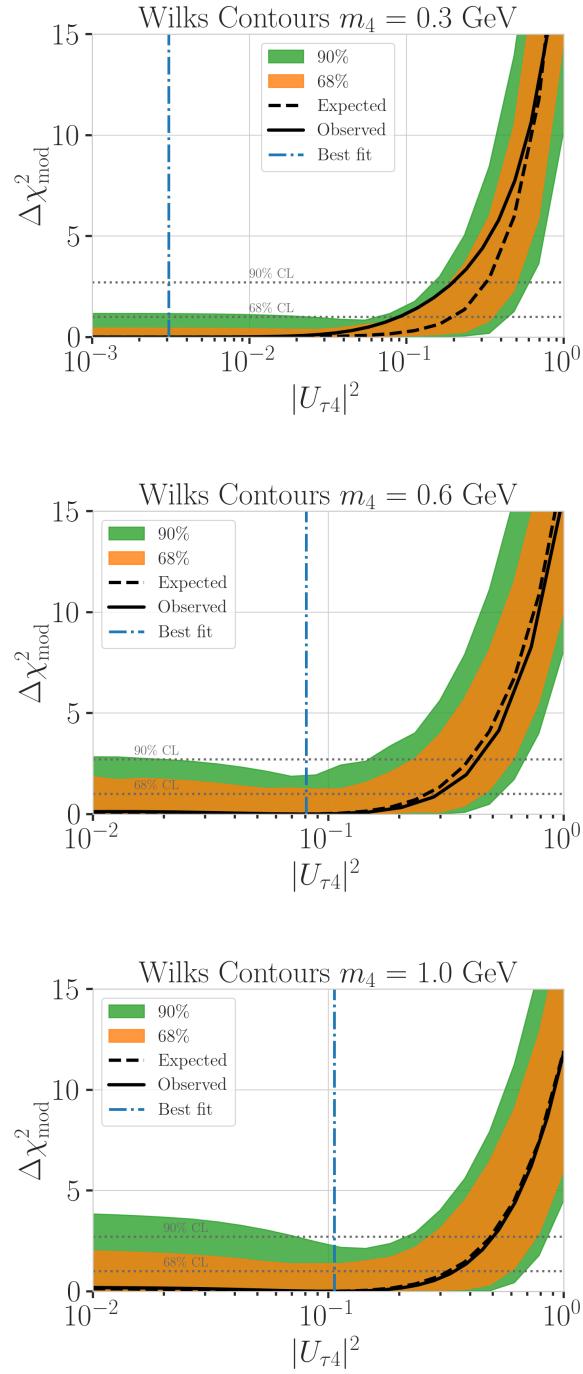


Figure 2.6: Best fit point TS profiles as a function of $|U_{\tau 4}|^2$ for the 0.3 GeV, 0.6 GeV, and 1.0 GeV mass samples. Shown are the observed profiles, the Asimov expectation at the best fit point, and the 68 % and 90 % bands, based on 100 pseudo-data trials. Also indicated are the 68 % and 90 % CL levels assuming Wilks' theorem.

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 pages 5, 6).
- [6] N. Whitehorn, J. van Santen, and S. Lafebre. "Penalized splines for smooth representation of high-dimensional Monte Carlo datasets". In: *Computer Physics Communications* 184.9 (2013), pp. 2214–2220. doi: <https://doi.org/10.1016/j.cpc.2013.04.008> (cited on page 5).
- [7] 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).
- [8] 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).
- [9] <https://launchpad.net/mg5amcnlo/3.0/3.3.x> (cited on page 8).
- [10] A. Alloul et al. "FeynRules 2.0 - A complete toolbox for tree-level phenomenology". In: *Comput. Phys. Commun.* 185 (2014), pp. 2250–2300. doi: [10.1016/j.cpc.2014.04.012](https://doi.org/10.1016/j.cpc.2014.04.012) (cited on page 8).
- [11] V. Barger et al. "Matter effects on three-neutrino oscillations". In: *Phys. Rev. D* 22 (11 Dec. 1980), pp. 2718–2726. doi: [10.1103/PhysRevD.22.2718](https://doi.org/10.1103/PhysRevD.22.2718) (cited on page 11).
- [12] A. M. Dziewonski and D. L. Anderson. "Preliminary reference Earth model". In: *Physics of the Earth and Planetary Interiors* 25.4 (1981), pp. 297–356. doi: [https://doi.org/10.1016/0031-9201\(81\)90046-7](https://doi.org/10.1016/0031-9201(81)90046-7) (cited on page 11).
- [13] S. Yu and J. Micallef. "Recent neutrino oscillation result with the IceCube experiment". In: *38th International Cosmic Ray Conference*. July 2023 (cited on pages 12, 15, 16).
- [14] 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](https://doi.org/10.1088/1748-0221/18/10/P10019) (cited on page 14).
- [15] 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 page 14).
- [16] E. Lohfink. "Testing nonstandard neutrino interaction parameters with IceCube-DeepCore". PhD thesis. Mainz, Germany: Johannes Gutenberg-Universität Mainz, Fachbereich für Physik, Mathematik und Informatik, 2023. doi: <http://doi.org/10.25358/openscience-9288> (cited on page 14).
- [17] 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 15).
- [18] 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](https://doi.org/10.1103/PhysRevD.95.023012) (cited on page 15).

- [19] J. Feintzeig. "Searches for Point-like Sources of Astrophysical Neutrinos with the IceCube Neutrino Observatory". PhD thesis. University of Wisconsin, Madison, Jan. 2014 (cited on pages 15, 16).
- [20] N. Kulacz. "In Situ Measurement of the IceCube DOM Efficiency Factor Using Atmospheric Minimum Ionizing Muons". MA thesis. University of Alberta, 2019 (cited on pages 15, 16).
- [21] 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](https://doi.org/10.1016/j.nima.2020.164332) (cited on page 16).
- [22] <https://github.com/icecube/pisa> (cited on page 16).
- [23] R. S. Nickerson. "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises". In: *Review of General Psychology* 2 (1998), pp. 175–220 (cited on page 16).
- [24] H. Dembinski et al. *scikit-hep/minuit: v2.17.0*. Version v2.17.0. Sept. 2022. doi: [10.5281/zenodo.7115916](https://doi.org/10.5281/zenodo.7115916) (cited on page 16).
- [25] 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](https://doi.org/10.1016/0010-4655(75)90039-9) (cited on page 17).
- [26] 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](https://doi.org/10.1214/aoms/1177732360) (cited on page 19).