

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

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)	3
again, describe what is shown and why this is interesting (e.g. energy distribution between the cascades, realistic exponential decay length distribuion..) also add this to the captions shortly (RED)	4
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
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
put a number on this significant increase? (YELLOW)	12

Contents

Contents	v
1 Heavy Neutral Lepton Event Generation	1
1.1 Model-Independent Simulation	1
1.1.1 Simplistic Samples	1
1.1.2 Realistic Sample	3
1.2 Model-Dependent Simulation	4
1.2.1 Custom LeptonInjector	4
1.2.2 Sampling Distributions	8
1.2.3 Weighting Scheme	8
1.3 Background Simulation	10
1.3.1 Neutrinos	11
1.3.2 Muons	12
Bibliography	13

Heavy Neutral Lepton Event Generation

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. This chapter describes the first HNL event generation developed for IceCube DeepCore. Two avenues of generation were pursued in parallel. A collection of model-independent simulation samples is explained in Section 1.1. They were used for performance benchmarking and for cross-checks to validate the physically accurate, model-dependent simulation, which is described in Section 1.2. For completeness, the event generation for SM background events is briefly described in Section 1.3. The default low-energy event selection and processing chain, which is applied identically to both background and signal, is introduced in Chapter ??.

1.1	Model-Independent Simulation	1
1.2	Model-Dependent Simulation	4
1.3	Background Simulation	10

1.1 Model-Independent Simulation

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

1.1.1 Simplistic Samples

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

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

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.

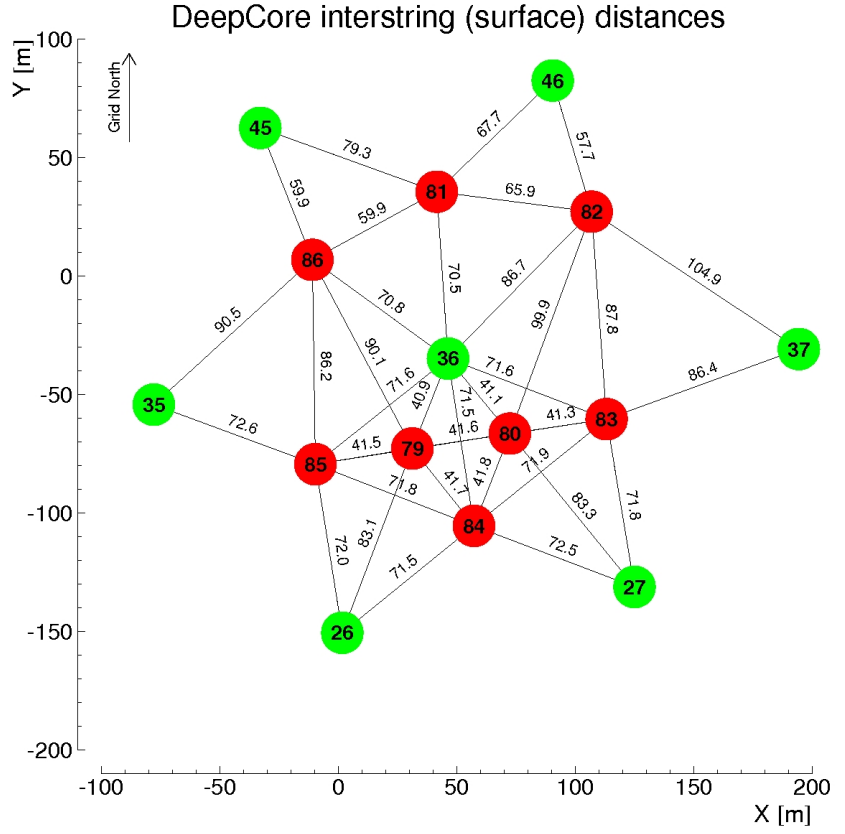


Figure 1.1: Horizontal positions and distances between DeepCore strings. Red strings are instrumented more densely (vertically) and partially have higher quantum efficiency (HQE) DOMs.

The first sample is used to investigate one of the most promising scenarios to detect a double cascade, where both cascade centers are located on a DeepCore string and the directions are directly up-going. From the DeepCore strings, string 81 was randomly chosen as the x - y coordinate for this sample. 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-energy events that produce little light. To produce the events, the x , y position of the cascades is fixed to the center of string 81 while the z positions are each sampled uniformly along the axis of the string. Note that this will therefore not produce a uniform length distribution between the cascades. The positions are defined in the IceCube coordinate system that was introduced in Section ??. The energies are sampled uniformly between 0.0 GeV and 60.0 GeV, to generously cover the region where $\nu_\mu \rightarrow \nu_\tau$ appearance is maximized. The specific sampling distributions/values for the cascades are listed in Table 1.1. The time of the lower cascade is set to $t_0 = 0.0$ ns and for the upper one to $t_1 = L/c$, assuming the HNL travels at the speed of light, c .

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)

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

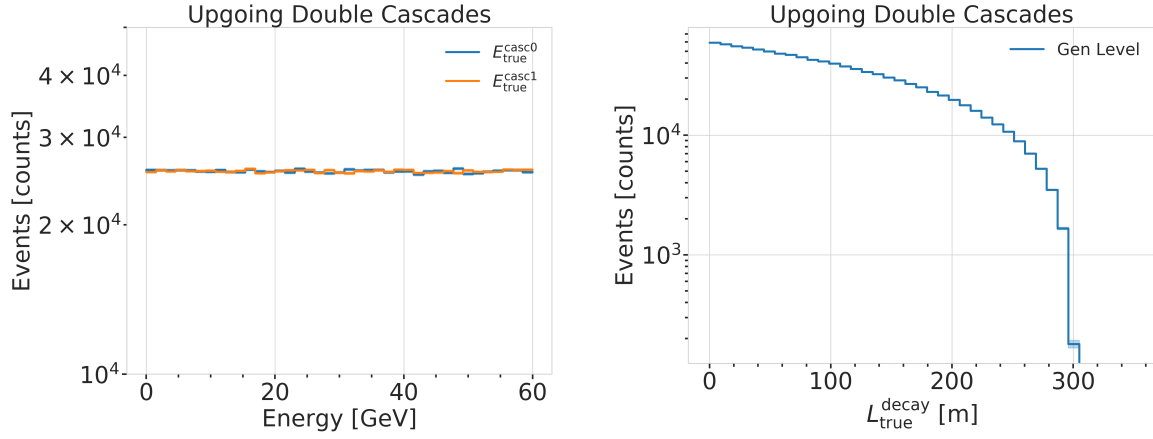


Figure 1.2: Generation level distributions of the simplistic simulation samples. Cascade and total energies (left) and decay lengths (right) of both samples are shown.

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 cascade kinematics. This sample is particularly useful for validating the model-dependent HNL simulation described in Section 1.2. For this purpose the total energy is sampled from an E^{-2} power law, mimicking the energy spectrum of the primary neutrinos as stated in Section 1.3.1. The total energy is divided into two parts, by assigning a fraction between 0% and 100% to one cascade and the remaining part to the other cascade. This is a generic approximation of the realistic process described in Section 1.2, and chosen such that the whole sample covers various cases of energy distributions between the two cascades. To efficiently generate events in a way that produces distributions similar to what would be observed with DeepCore, one of the cascade positions is sampled inside the DeepCore volume by choosing its coordinates uniformly on a cylinder that is centered in DeepCore. This is similar to a trigger condition of one cascade always being inside the DeepCore fiducial volume. Choosing the direction of the event by sampling zenith and azimuth uniformly between 70° and 180° and 0° and 360° , respectively, the position of the other cascade can be inferred for a given decay length, assuming a travel speed of c , and choosing whether the cascade position that was sampled is the first cascade or the second with a 50% chance. The zenith angle is chosen between straight up-going (zenith of 180°) and slightly down-going from above the horizon (70°) to mimic an event selection that reduces atmospheric muons by rejecting events coming from above the horizon, but still incorporates some down-going events. The decay length is sampled from an exponential distribution, as expected for a decaying heavy mass state. The sampling distributions/values are listed in

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)

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

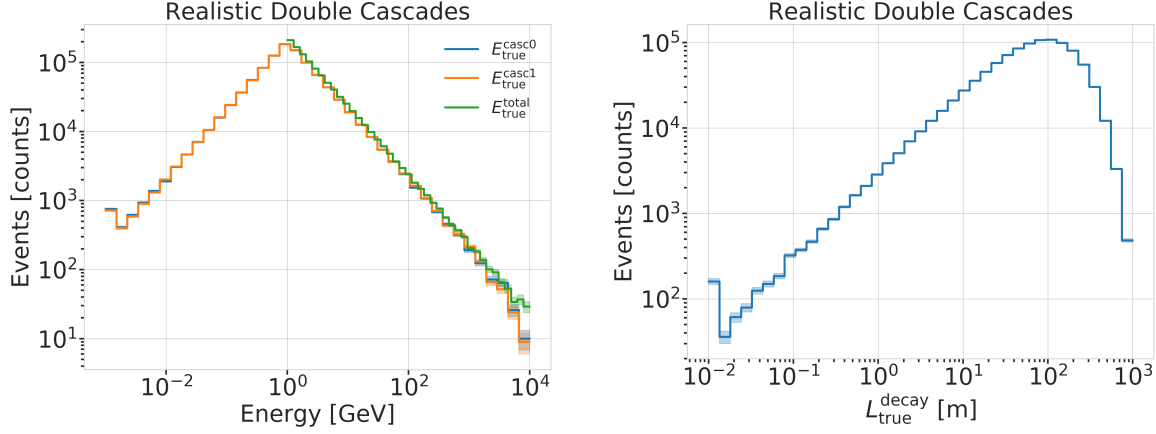


Figure 1.3: Generation level distributions of the simplistic realistic sample. Shown are the cascade and total energies (left) and decay lengths (right).

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)

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

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 generation.

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 is described for neutrinos

in Section 1.3.1. 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 as discussed in Section ?? . Both objects are injected with the same (x, y, z, t) coordinates and the kinematics are sampled from the differential and total cross-sections that are one of the inputs to LI.

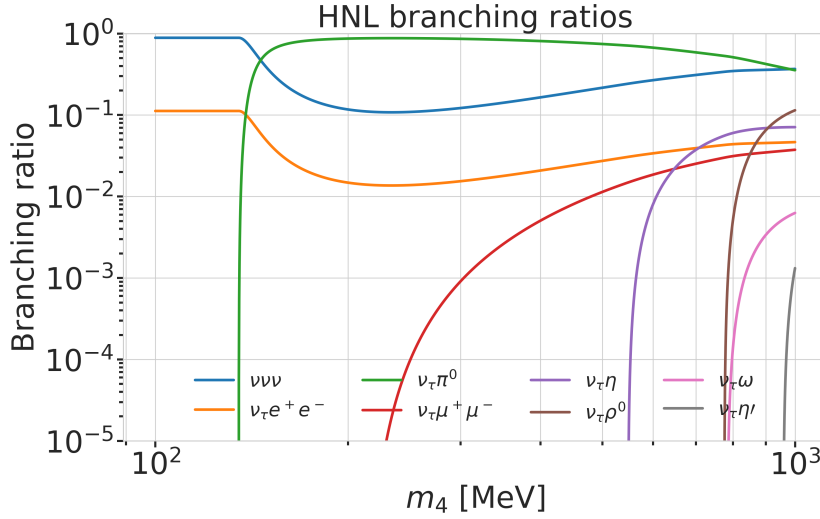


Figure 1.4: Branching ratios of the HNL within the mass range considered in this work, only considering $|U_{\tau 4}^2| \neq 0$, calculated based on the results from [3].

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 addition to LI is that the decay products of the HNL are also stored. These HNL daughter particles form the second cascade, not as a single hadronic cascade object, but as the explicit particles forming the shower. They are injected with the correctly displaced position and delayed time from the interaction vertex, given the HNL decay length. The kinematics of the two-body decays are computed analytically, while the 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 uniform injection of the primary particle on a cylindrical volume, adding 50 % of the events with ν_τ and the other half with $\bar{\nu}_\tau$ as primary particle types. The generator takes the custom double-differential/total cross-section splines described below and the parameters defining the sampling distributions as inputs.

Cross-Sections

The cross-sections are calculated using the NUXSSPLMKR [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

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”

[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”

[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 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 do not have sufficient energy to produce the outgoing, massive lepton. This results in a reduction of the cross-section towards lower energies, which scales with the assumed mass of the HNL. This effect can be seen in Figure 1.5.

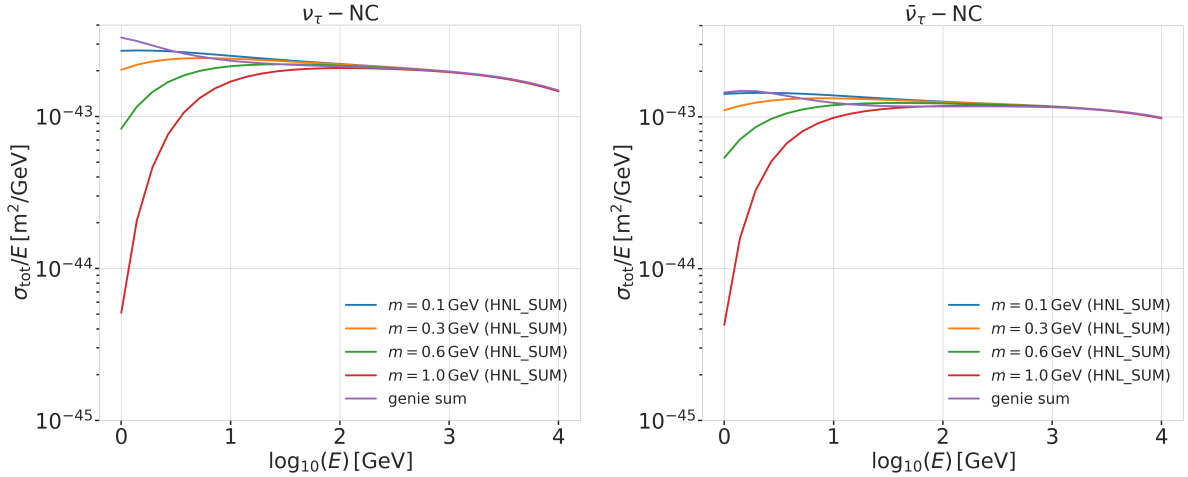


Figure 1.5: Custom HNL total cross-sections for the four target masses compared to the total ($\nu_\tau/\bar{\nu}_\tau$ NC) cross-section used for SM neutrino simulation production with GENIE.

Re-make plot with 5 target 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 neutrino simulation explained in Section 1.3.1. The double-differential ($d^2\sigma/dxdy$) and total (σ) cross-sections were produced for the chosen target HNL masses and then splined in energy, x , and y for $d^2\sigma/dxdy$ and σ in the energy. Figure 1.5 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 ~ 200 GeV, where the modification should not have any effect on the cross-sections. This is the desired result of using

the identical input PDFs, and confirms that the unmodified cross-sections produced with NuXSplMkr agree with the GENIE cross-sections.

Decay Channels

The accessible decay channels are dependent on the mass of the HNL and the allowed mixing. For this analysis, where only $|U_{\tau 4}|^2 \neq 0$, the decay channels considered are listed in Table 1.3 and the corresponding branching ratios are shown in Figure 1.4. 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.

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

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

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

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

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

Channel	Opens
$\nu_4 \rightarrow \nu_\tau \nu_\alpha \bar{\nu}_\alpha$	0 MeV
$\nu_4 \rightarrow \nu_\tau e^+ e^-$	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 $|U_{\tau 4}|^2 \neq 0$, and the mass at which each channel opens.

[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

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

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)

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_HEAVY_N_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 the final level of the event selection chain the phase space relevant for the analysis is covered with sufficient statistics to make a reasonable estimate of the event expectation. Initial distributions insufficiently covering the phase space leads to an underestimation of the expected rates, because some of the events that would pass the selection are not produced. This limits the expected analysis potential. Three discrete simulation samples were produced with HNL masses of 0.3 GeV, 0.6 GeV, and 1.0 GeV. During development of the analysis it became clear that short decay lengths were undersampled at the final selection level. Therefore, each discrete mass 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 at generation to result in sufficient MC statistics at final level. 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

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

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

folding method that was already introduced for the SM simulation in Section 1.3 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 $[s_{\text{min}}, s_{\text{max}}]$ can be converted into the rest frame lifetime range $[\tau_{\text{min}}, \tau_{\text{max}}]$ for each event

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

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

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

This re-weighting factor is then calculated as

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

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

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

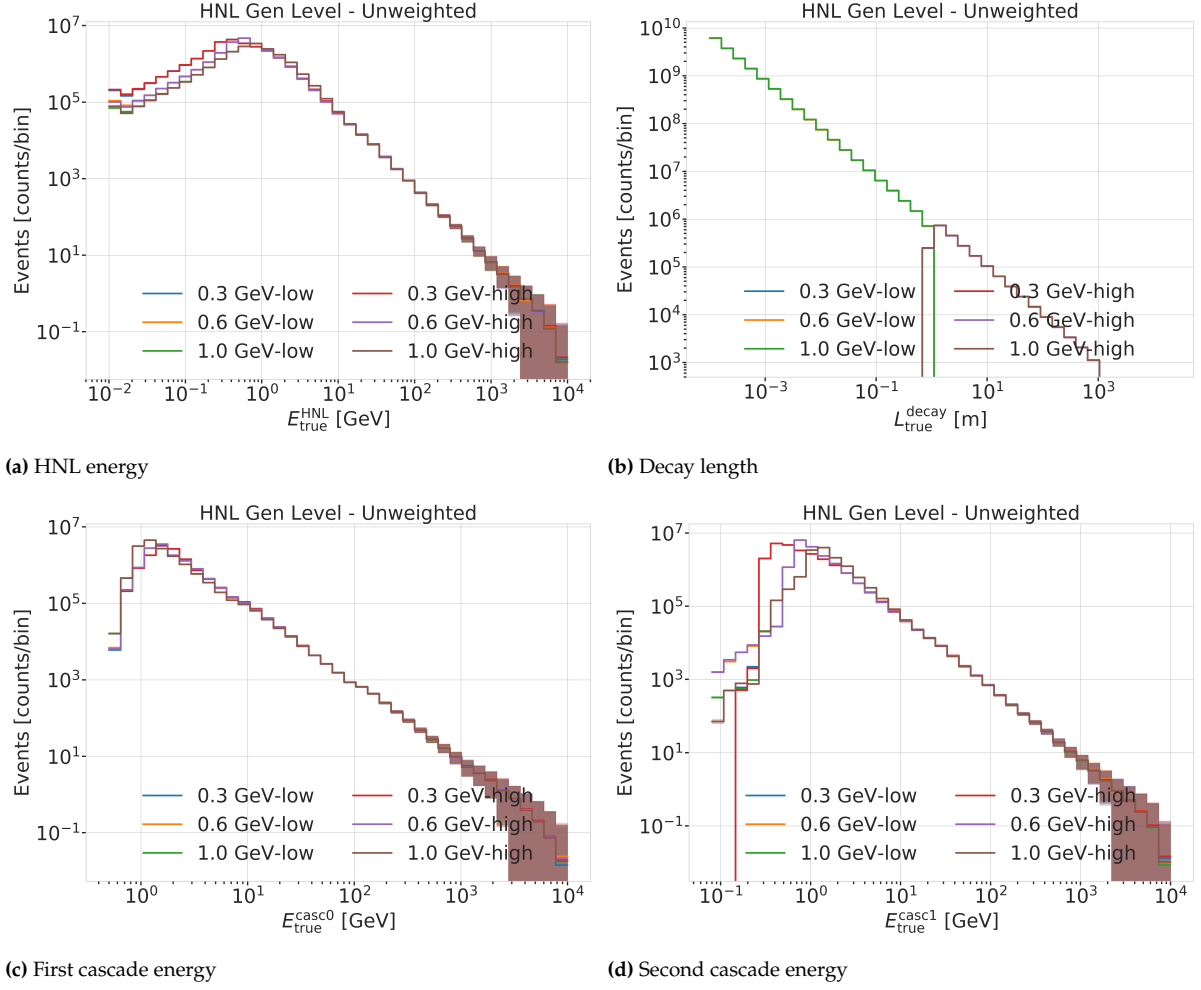


Figure 1.6: Generation level distributions of the model-dependent simulation.

If this additional weighting factor is multiplied to a generation weight with units m^2 (like in Equation 1.29), 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).

1.3 Background Simulation

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 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 ?? and calculating a loss function comparing the bin expectations to

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

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

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).

1.3.1 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

$$w_{\text{gen}} = \frac{1}{F_{\text{sim}}} \frac{1}{N_{\text{sim}}}, \quad (1.29)$$

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.* [11].

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

To simulate the neutrino interaction with the ice, the GENIE event generator [12] (version 2.12.8) is used, resulting in the secondary particles and the kinematic and cross-section parameters. As input, the outdated GRV98LO [13] *parton distribution functions* (PDFs) was used, because it was the only option that could incorporate extrapolations to lower Q^2 [14]. Muons produced in these interactions are propagated using PROPOSAL [15], also simulating their

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

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

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

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

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

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

[17]: Rädcl 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”

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

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

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

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

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 [16] while for higher energies an analytical approximation from [17] is used.

1.3.2 Muons

Atmospheric muons are generated on a cylinder surface enclosing the full IceCube detector array. The cylinder has a height of 1600 m and a radius of 800 m. The energy is sampled from an E^{-3} power law while the other sampling distributions (position, direction) are found from parameterizations based on [18]. This work uses full CORSIKA [19] simulations of muons to tailor the parameterizations, starting from *cosmic ray* (CR) interactions with atmospheric nuclei using the CR flux model from [20] and producing the muons applying the *hadronic interaction* (HI) model SIBYLL 2.1 [21]. 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 ?? and Section ?? 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 .

put a number on this significant increase? (YELLOW)

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 6).
- [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] 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 11).
- [12] C. Andreopoulos et al. “The GENIE Neutrino Monte Carlo Generator: Physics and User Manual”. In: (2015) (cited on page 11).
- [13] 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 11).
- [14] 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 11).
- [15] 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 11).
- [16] 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 12).
- [17] L. Rädcl 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 12).
- [18] Y. Becherini et al. “A parameterisation of single and multiple muons in the deep water or ice”. In: *Astroparticle Physics* 25.1 (2006), pp. 1–13. doi: <https://doi.org/10.1016/j.astropartphys.2005.10.005> (cited on page 12).

- [19] D. Heck et al. "CORSIKA: A Monte Carlo code to simulate extensive air showers". In: (Feb. 1998) (cited on page 12).
- [20] 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 12).
- [21] 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 12).