First Search for Heavy Neutral Leptons with IceCube DeepCore

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

im Fach: Physik
Spezialisierung: Experimentalphysik

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

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

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

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

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

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

Zusammenfassung

Zusammenfassung . . .

Abstract

Abstract . . .

Todo list

Contents

List of Figures

List of Tables

Detecting Low Energetic Double Cascades

1

1.1 Reconstruction

All existing reconstruction algorithms applied for low energetic atmospheric neutrino events mentioned in Section **??** are either assuming a single cascade hypothesis or a track and cascade hypothesis, which are the two SM morphologies observable at these energies, as was described in Section **??**. A HNL being produced and decaying inside the IceCube detector however, will produce two cascade like light depositions. The morphology and how the cascade properties and their spatial separation depend on the model parameters was introduced in Section **??**. To investigate the performance of the detector to observe these events, a low energetic double cascade reconstruction algorithm was developed, based on a pre-existing algorithm used to search for double cascades produced from high energetic astrophysical tau neutrinos [1] that was established in [2], but first mentioned in [3].

[1]: Abbasi et al. (2020), "Measurement of Astrophysical Tau Neutrinos in IceCube's High-Energy Starting Events"
[2]: Usner (2018), "Search for Astrophysical Tau-Neutrinos in Six Years of High-Energy Starting Events in the IceCube Detector"
[3]: Hallen (2013), "On the Measurement of High-Energy Tau Neutrinos with IceCube"

1.1.1 Table-Based Minimum Likelihood Algorithms

The reconstruction is relying on a maximum likelihood algorithm, which is the *classical* approach to IceCube event reconstructions, as opposed to ML based methods. A Poissonian likelihood is constructed, which compares the observed photon numbers, n, with their arrival times to the expected light depositions, μ, for a given even hypothesis as

maybe I want a figure for this, or not so important? (YELLOW)

$$\ln(L) = \sum_j \sum_t n_{j,t} \cdot \ln\big(\mu_{j,t}(\Theta) + \rho_{j,t}\big) - \big(\mu_{j,t}(\Theta) + \rho_{j,t}\big) - \ln\big(n_{j,t}!\big) \,, \quad (1.1)$$

where ρ are the number of expected photons from noise, Θ are the parameters governing the source hypothesis, and the likelihood is calculated summing over all DOMs j splitting observed photons into time bins t. The light expectations are calculated using look-up tables [4] that contain the results from MC simulations of reference cascade events or track segments. By varying the parameters defining the event hypothesis, the likelihood of describing the observed light pattern by the expected light depositions is maximized to find the reconstructed event. Algorithms of this kind used in IceCube are described in great detail in [5]. For the table production a specific choice of ice model has to be made, while the calibrated DOM information is taken from the measurement itself.

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

[5]: Aartsen et al. (2014), "Energy Reconstruction Methods in the IceCube Neutrino Telescope"

Based on the tabulated light expectations for cascades and track segments, various event hypothesis can be constructed, like the common cascade only or the track and cascade hypotheses. The hypothesis describing the double cascade signature of the HNL is using two reference cascades that are separated by a certain distance. The whole hypothesis is defined by 9 parameters and assumes that the two cascades are aligned with each other, which is a safe assumption for strongly forward boosted interactions. The parameters are the position of the first cascade x, y, z, the direction of both cascades ϕ, θ, and its time t as well as the decay length L between the

Elaborate whether this is the case (show it in a plot?). Discuss directionality of cascades in general. (ORANGE) two cascades. Assuming the speed of the HNL to be the speed of light, c, this already defines the full signature. The HNL particle does not produce any light while traveling, as it is electrically neutral. The full 9 parameters describing the event are $\Theta = (x, y, z, t, \theta, \phi, E_0, E_1, L)$. To compute the full likelihood the term in Equation 1.1 is summed over both cascade parts, i, as $\sum_i \ln(L_i)$.

1.1.2 Optimization for Low Energy Events

Optimizing the double cascade reconstruction for low energetic events was done in parallel to the development of the model dependent simulation generator introduced in Section **??**. A preliminary sample of HNL events was used, containing a continuum of masses between 0.1 GeV and 1.0 GeV and lab frame decay lengths sampled uniformly in the range from 5 m to 500 m. Even though this sample is not representative of a physically correct model and therefore not useful to predict the event expectation, it can still be used to optimize the reconstruction. The double cascade nature of the individual events and the evenly spaced decay length distribution are especially useful for this purpose.

The simulation is processed up to Level 5 of the selection chain described in Section **??** and one of the reconstructions from [6] is applied to the events, fitting a cascade and a track and cascade hypothesis. The results from this reconstruction are used as an input for the double cascade reconstruction, where the position of the vertex, the direction of the event, and its interaction time are used as the input quantities for the first cascade, and the length of the track reconstruction is used as a seed for the distance between the two cascades.

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

Decay Length Seeds

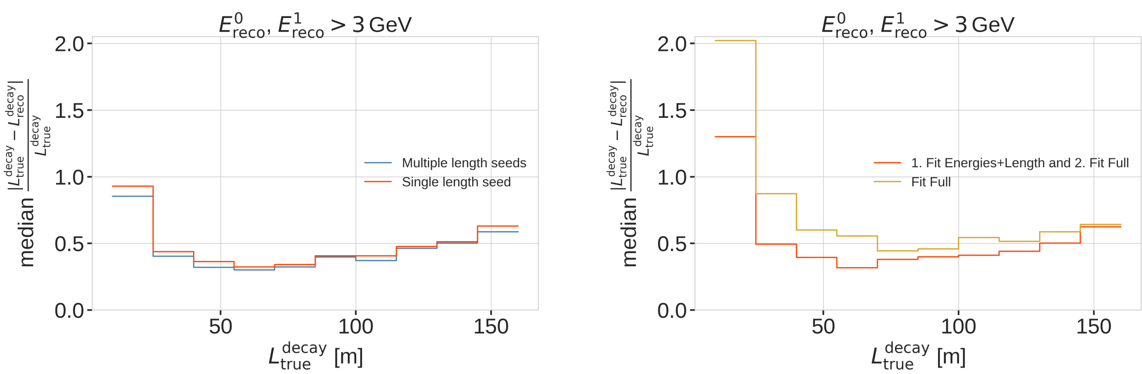


Figure 1.1

fix caption of this figure (RED)

The full 9 dimensional likelihood space is very complex and can have many local minima, depending on the specific event and its location in the detector. Especially the seed value of the length between the two cascades was found to have a very strong impact on whether the global minimum was found during the minimization. To mitigate this effect, multiple fits are performed, seeding with variations of the input length at different orders of magnitude.

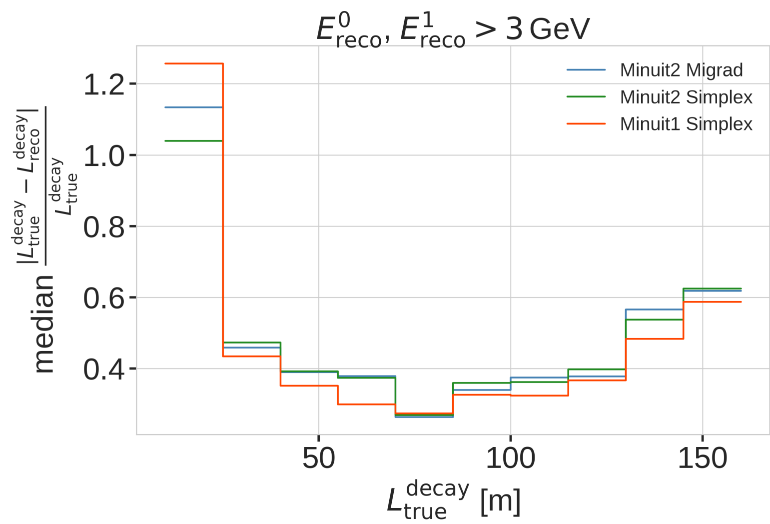
The best result is used, selected based on the total likelihood value of the best fit parameter set. A small improvement in the decay length resolution can be found by using this approach as compared to a single length seed. The effect can be seen in the left part of Figure 1.1, which shows the median, absolute, fractional decay length resolution.

Fit Routine

Because the length seed showed to have such a large impact on the reconstruction performance, a more sophisticated fit routine, than just fitting all 9 parameters at once, was tested. In a first fit iteration, some parameters are fixed and the resulting best fit point is used to fit all 9 parameters in a second iteration. In the right part of Figure 1.1 it can be seen how a fit split into two consecutive steps, where the first step fits only both cascade energies and the decay length and the second step fits the full 9 parameters, performs better as compared to a single, full 9 parameter fit. The initial seed for both routines is the same.

Minimizer Settings

To investigate the effect of the minimizer used to find the best fit parameters, the reconstruction was performed using three different minimizers, which were easily accessible within the reconstruction framework. The minimizers used were Minuit1 Simplex, Minuit2 Simplex, and Minuit2 Migrad. The results can be seen in Figure 1.2, where the Minuit1 Simplex minimizer performs best. The initial idea was to test a global minimizer, or a routine that can find the rough position of the global minimum first and then a local minimizer to find the exact minimum, but unfortunately this was not possible with the minimizers available in the framework. From the three tested minimizers, Minuit1 Simplex performed best and was chosen as the default for the reconstruction. The comparison of the decay length resolutions can be seen in Figure 1.2.



fix caption of this figure (RED)

Figure 1.2: title

1.1.3 Performance

The chosen reconstruction chain used to test the performance of the detector to observe low energetic double cascades is the following; Minuit1 Simplex is used as the minimizer, the decay length is seeded with 3 different values, 0.5x, 1.0x, and 1.5x the length of the track reconstruction, and the fit routine is split into two steps, where the first step fits the energies and the decay length and the second step fits the full 9 parameters. In the first step, the number of time bins in Equation 1.1 is set to 1, so just the number of photons and their spatial information is used. The second step is seeded with the best results from the first fit and here the number of time bins is chosen such that each photon falls into a separate time bin, which means all time information is used. The average runtime per event is ~16 s on a single CPU core, but is very dependent on the number of photons observed in the event, since the likelihood calculation in the second step scales with this number and a table lookup has to be performed for each photon.

To get a more realistic estimate of the reconstruction performance, it is run on a second preliminary sample of HNL events, containing masses between 0.1 GeV and 3.0 GeV and the lab frame decay length is sampled from an inverse distribution in the range from 1 m to 1000 m, which is a better approximation of the expected exponential decay distribution of the HNL. The performance is shown for events where the reconstruction chain was successfully run, the event selection criteria up to level 7 are fulfilled, and the reconstructed energy of both cascades is above 3.0 GeV. This is done to only investigate well reconstructed events with two significant light depositions at a usual final selection level of the oscillation analyses.

one half sentence on why this number was chosen? (ORANGE)

Energy Resolutions

The energy resolution is inspected by looking at the 2-dimensional distribution of reconstructed energy versus the true energy as shown in Figure 1.3. The bin entries are shown as well as the median and ±25 % calculated per vertical column, to get an idea of the distribution for a given energy slice. The color scale is showing the PDF along in each true energy slice, which is the full information combined into the median±25 % lines. The reconstructed energy is only the energy that is observable from photons, while the true energy is the total cascade energy, including the parts that go into EM neutral particles that do not produce light. It is therefore not expected that the median lines up with the axis diagonal, but rather the reconstructed energy is going to be lower.

The histogram for the first cascade energy is shown on the left and above an energy of ~10 GeV the reconstruction performs well, with the median being parallel to the diagonal and the spread in the ±25 % quantile being small. Below this energy the reconstruction is over-estimating the true energy, which is a known effect in IceCube, where the reconstruction is biased towards higher energies around the energy detection threshold, because events that enter the sample are events with an over fluctuation in their light deposition, which makes them pass into the selection and being reconstructible in the first place.

For the second cascade the overall behavior is similar, only that the energy where the reconstruction starts to perform good is higher around ~20 GeV.

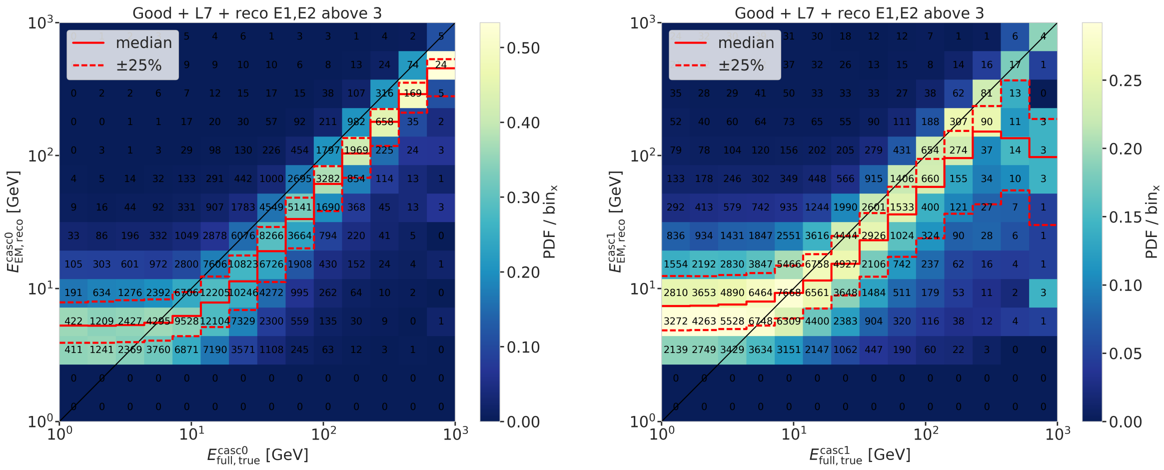


Figure 1.3: Reconstructed (EM) energy versus true energy (full) energy for the first cascade (left) and second cascade (right). The color scale is according to the PDF in each vertical true energy slice, with the solid and dashed lines showing the median±25 % quantiles. The bin entries are shown as numbers.

The spread around the median is also larger and starts to expand a lot above 200 GeV, where the statistics are lower as can be seen from the bin counts. It is also very apparent that the majority of energies of the second cascade are at lower true energy values between 1 GeV and 20 GeV.

For both cascade resolutions the effect of the reconstruction being biased towards lower values, due to the comparison of the full true energy to the reconstructed EM energy can be seen.

Length Resolutions

The decay length resolution is also investigated by looking at a similar style of 2-d histograms as for the energies, where the reconstructed decay length is plotted versus the true decay length. The left part of Figure 1.4 shows the distributions after the same selection criteria from Section 1.1.3 are applied. It can be observed that for short true lengths the reconstruction is over-estimating the length, while for long true lengths the reconstruction is strongly under-estimating the length. There is a region between true lengths of 20 m and 80 m where the median reconstruction is almost unbiased, but the 50 % interquartile range is large and increasing from ∼50 m to ∼70 m with true decay lengths.

The over-estimation at small true lengths can be explained for multiple reasons, one being that the shortest DOM spacing is ∼7 m, vertically for DeepCore strings, but mostly larger than that, so resolving lengths below this is very complicated, and the reconstruction tends to be biased towards estimating the length around where the light was observed. Another reason is a similar argument to why the energies are over-estimated at small true values, namely that events that passed the selection and were reconstructed in those cases, probably have an over fluctuation in light deposition, extending further out from the vertices, so the reconstructed length is larger. Additionally, approaching a length of 0.0, the reconstructed length will of course always be a one-sided distribution, because the lengths have to be positive.

The under-estimation at large true lengths is more puzzling, and it seems like the distribution becomes bimodal in the reconstructed lengths, with

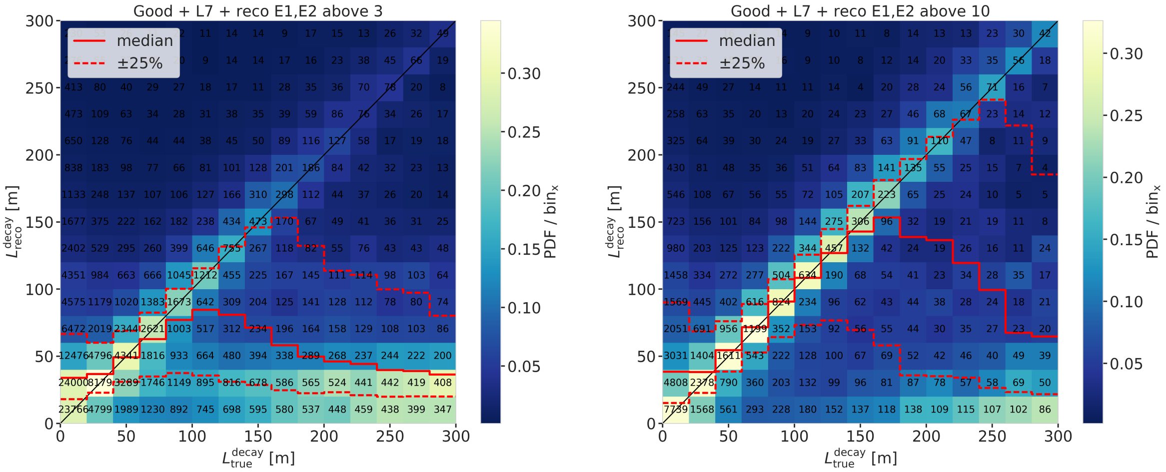


Figure 1.4: Reconstructed decay length versus true decay length for ∼3 GeV (left) and ∼10 GeV (right) minimum reconstructed cascade energies. The color scale is according to the PDF in each vertical true length slice, with the solid and dashed lines showing the median±25 % quantiles. The bin entries are shown as numbers.

one population around the diagonal, meaning that they are properly reconstructed, and one population at very short reconstructed lengths, which are badly reconstructed. Above 150 m the badly reconstructed population starts to dominate, and the median resolution drops off strongly. The assumption is that for these events, only one cascade was observed with enough light to be reconstructed, and the reconstruction describes the one observed cascade in two parts, separated by a short distance, driven by similar factors as mentioned before. A quick check to confirm whether this is the case, was to increase the selection criteria to minimum reconstructed cascade energies of 10 GeV, which is shown in the right part of Figure 1.4. It can be seen that the median resolution is already much better, aligning with the expectation between 40 m and 160 m. Judging from the median resolution and the +25 % quantile in this range, there is very few events with an over-expectation in the energy, since both of them are alidgning with the diagonal. The spread towards lower reconstructed lengths, on the other hand, is still very large and above 200 m the badly reconstructed population starts to dominate again.

Badly Reconstructed Cascade Population

make on figure, where the two selected regions are highlighted? (RED)

fix caption (RED)

To investigate the badly reconstructed population further, a rough separation was made to find out what the cause of the difference is. It was already established that a larger reconstructed energy in both cascades, which is related to a larger true energy in form of more deposited light, leads to a better reconstruction in more events. As shown in Figure 1.5, only events with true decay length larger than 80 m are used, where the populations are split by the reconstructed decay length being larger or smaller than 80 m. Based on a few potential sourced of bad reconstruction, several variables were checked to see if there is a difference between the two populations.

The left part of Figure 1.6 shows the true horizontal distance of the second cascade from string 36. The distance is denoted as ρ_{36} and is a very good proxy for the distance to the center of the detector, because string 36 is almost at the center. While the distributions looks very similar for the first cascade

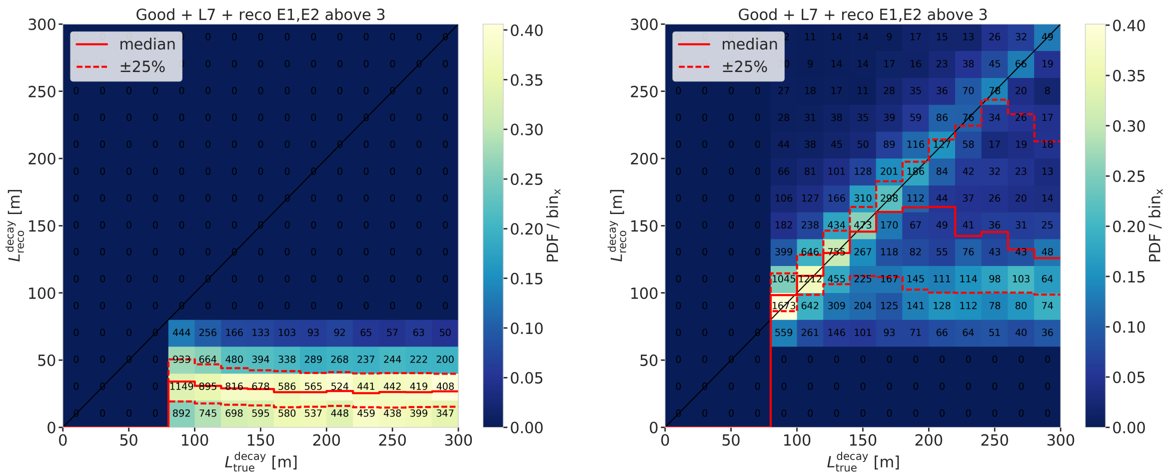


Figure 1.5

(not shown), for the second cascade the badly reconstructed population extends to larger values. Considering that the DeepCore strings are roughly inside a 70 m radius from the center, and the next layer of IceCube strings is at a radius of 125 m, this is a plausible explanation for a worse reconstruction, because for the badly reconstructed population the second cascades are more often in regions without DOMs, so less or no light is observed from them.

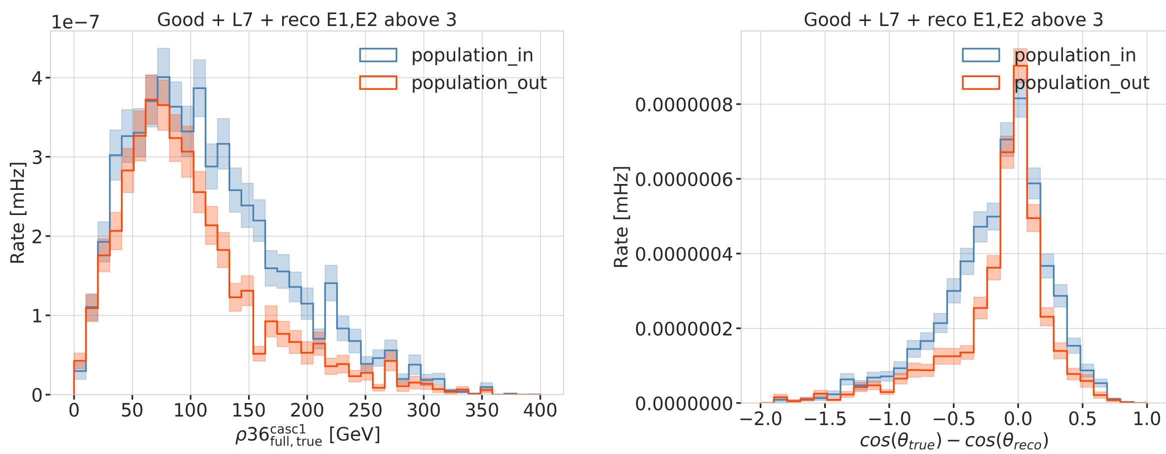


Figure 1.6

Another possible reason why the reconstruction underperforms could be that the initial seed direction was off and therefore one of the cascades cannot be found properly. Looking at the error of the cosine of the reconstructed zenith angle shown in the right of Figure 1.6, we see that the badly reconstructed population has a larger error, and is less peaked around 0.0. This could be a hint that the direction is worse for the badly reconstructed population, which could be due to a bad seed direction, or just the result of one cascade not being observed properly.

re-make normalized? (ORANGE)

fix caption (RED)

The true energies of both cascades are shown in Figure 1.5, where it can be observed that the first cascade energy is generally much larger than the second, peaking between 10 GeV and 20 GeV, while the second cascade peaks

re-make normalized? (ORANGE)

fix caption (RED)

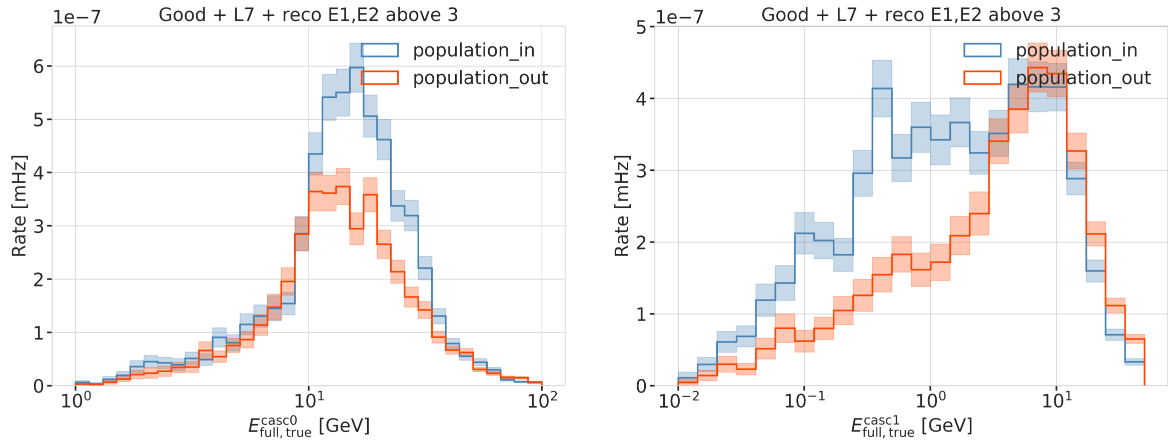


Figure 1.7

below 10 GeV. For the first cascade there is no significant difference beteween the two populations, but for the second cascade the badly reconstructed population has a larger fraction of events with lower energies and the distribution is almost uniform in the range of 2 GeV to 10 GeV, while the well reconstructed population has a peak around 10 GeV and falls off faster towards lower energies. This is a strong indication that the main reason for the bad reconstruction is the low energy of the second cascade.

Despite the fact that the split into the two populations was very rudimentary, it is clear that the main reason for the bad reconstruction is the low energy of the second cascade, while other factors, like the position of the second cascade, or the potentially bad input seed direction are also contributing to the bad reconstruction. For a more thorough investigation, a more sophisticated separation would be needed, but this is sufficient to conclude that the main reason for the bad reconstruction is the low energy of the second cascade.

circle back to the energy distributions I showed somehwere else, and state how the cascade energies are distributed in general (ORANGE)

1.2 Double Cascade Classification

Even though the performance results show that it is very complicated to reconstruct these low energetic double cascade events, the attempt to identify them over the background of SM neutrino events was made

tested training a classifier to distinguish between HNL events and SM neutrino events, using the same preliminary sample of HNL events as was used for the reconstruction performance

for the training a set of cuts was applied to make sure the classifier is trained on well reconstructed events, which are events with a minimum reconstructed energy of both cascades of 5 GeV, a minimum reconstructed decay length of 40 m

fix caption (RED)

additionally for the HNL events some cuts on the true energy and decay length were applied, which are a minimum true energy of both cascades of 5 GeV and a minimum true decay length of 40 m, these were chosen to make sure the HNL events were double cascade like

add table with cuts? (YELLOW)

the classifier used was a Boosted Decision Tree (BDT) from the SCIKIT-LEARN *(sklearn)* package and the input features are taken from the double cascade

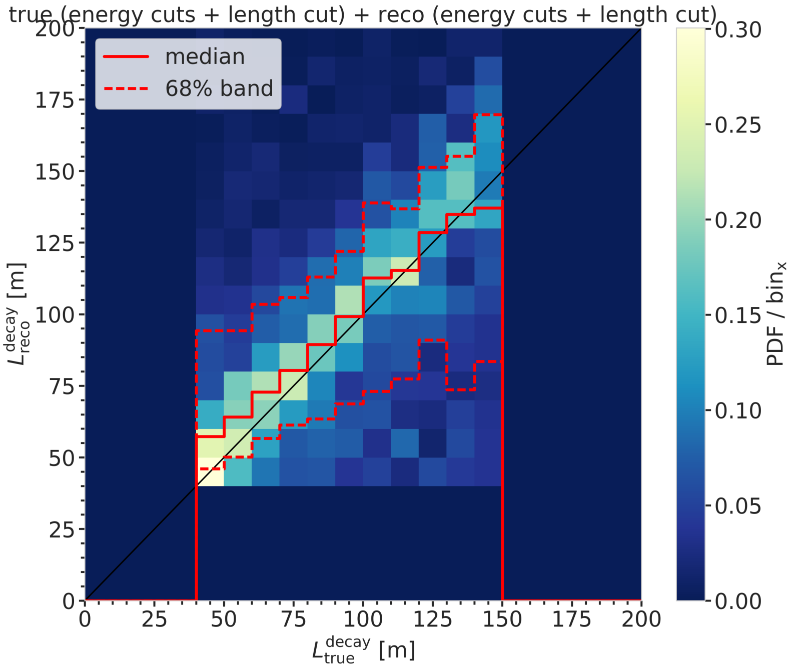


Figure 1.8

reconstruction explained in Section 1.1 as well as some additional variables from earlier levels of the processing explained in Section **??**

add example features? maybe like 2 or so? (RED)

both a single classifier trained to distinguish between HNL and all SM neutrino events was tested and two classifiers trained to distinguish HNLs from SM track-like and cascade-like events separately were tested

takeaway? (not sure how to deal with the caveat of the weight being kind of wrong for these..)

1.3 Generic Double Cascade Performance

Ideas:

- why do these checks again with the model independent samples? (controllable parameter space and distributions, especially energy and length)
- benchmark some edge cases (e.g. "best case" of up-going string centered, purely horizontal, and realistic case)
- investigate where the reconstructions breaks down (in terms of cascade energies, which are the main factor as shown before)

1.3.1 Idealistic Performance

- for the perfect edge case of the event being directly on top of a DeepCore string with an up-going direction the length can be very well reconstructed above a true length of around 20 m
- 2-d histogram shows that up to a true decay length of ~210 m, there is no under-estimation of the length

- this means that the under-estimation observed before is only possible if there is no DOMs in between the two cascades
- this makes a lot of sense, since DOMs being present, but not observing any light is affecting the light expectation that goes into the reconstruction likelihood and therefore makes these hypotheses less likely and therefore incompatible with the data
- do I want more detailed plots here? I feel like this is enough..

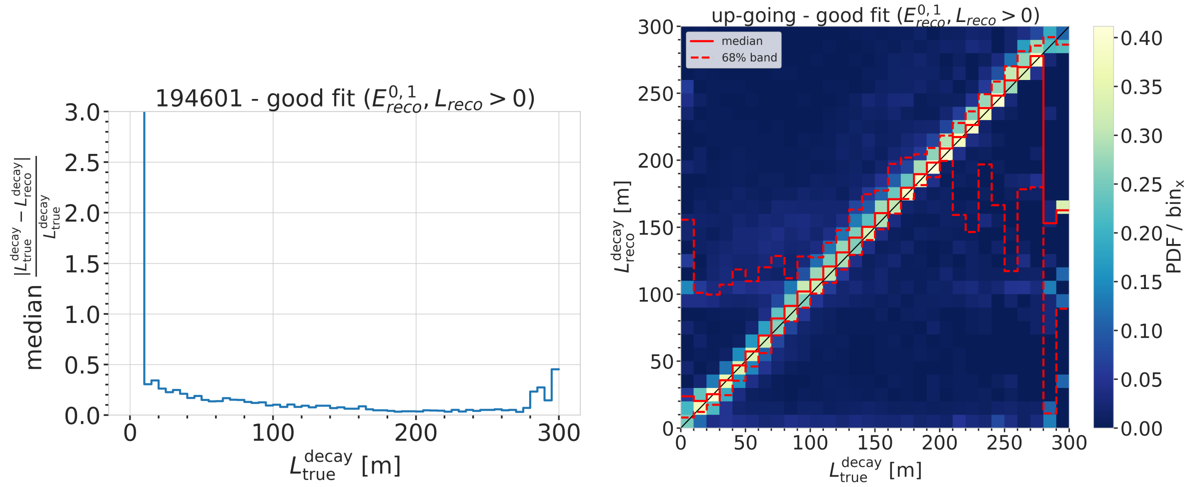


Figure 1.9

fix caption of this figure (RED)

1.3.2 Realistic Performance

Energy/Decay Length Resolution

describe the "good fit" selection when I first show plots that use it (ORANGE)

add some 1D resolution plots (ORANGE)

2-D Histograms

Things to mention about the 2d-hists:

- total energy resolution looks very good, above 10 GeV it's almost unbiased and the 1-sigma resolution band is below 20 %
- individual cascade resolutions mirror this behavior, but are starting to stabilize in energy at lower energies around 5 GeV to 6 GeV with a broader resolution band of 50 %, but reducing drastically with increasing energy (down to 20 % at 100 GeV)
- interestingly, the second cascade energy reconstruction performs slightly worse, although they have the same energy ranges. This could hint at an asymmetry in the reconstruction process (might relate to how the two cascades are parameterized) or be due to the different positions and the dominantly up-going direction used in the sampling combined with the DOMs looking down (relate this to the sampling distribuitons explained/shown in the previous chapter)
- the decay length resolution looks much worse. In the region between 20 m and 80 m it's roughly unbiased, but 1-sigma resolution band is quite wide with a lot of outliers towards short reconstructed lengths. Below 20 m the reconstructed lengths are always over-estimating the true and above 80 m a population of events start to dominate where the decay lengths isn't getting reconstructed at all, which might indicate

that one of the cascades wasn't observers. (Relate to the fact that this marginalizes over all energies, meaning also all events which have one cascade with very low energy are included here.)

- another interesting feature is the band of reconstructed lengths around 100 m, which is probabaly related to the spacing between most of the strings, which favors the reconstruction to be around this value, because that's the distance at which light can be observed, just from the fact that the DOMs are spaced at this distance (for low energetic cascades, this can dominate the reconstruction)

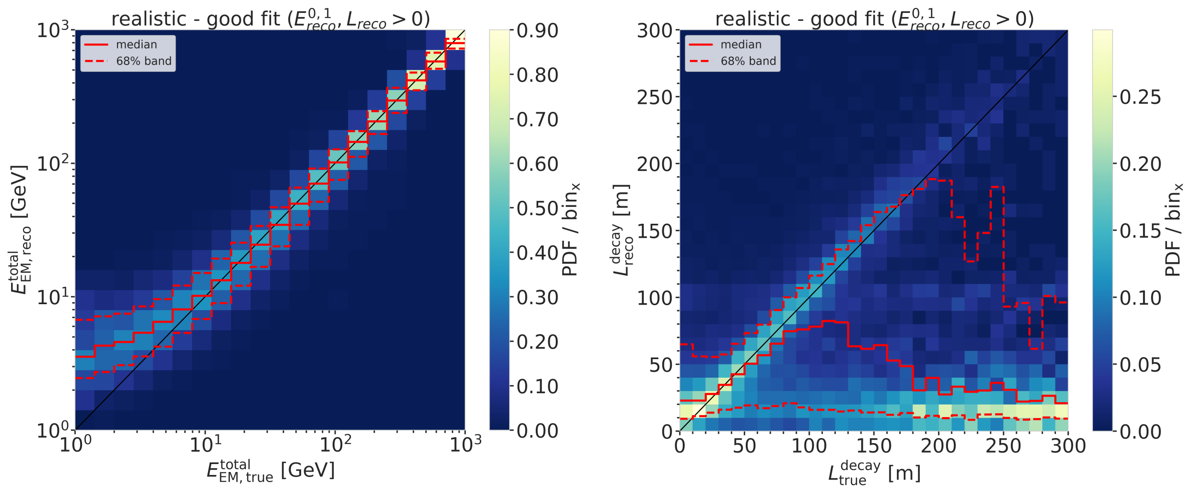


Figure 1.10

fix caption of this figure

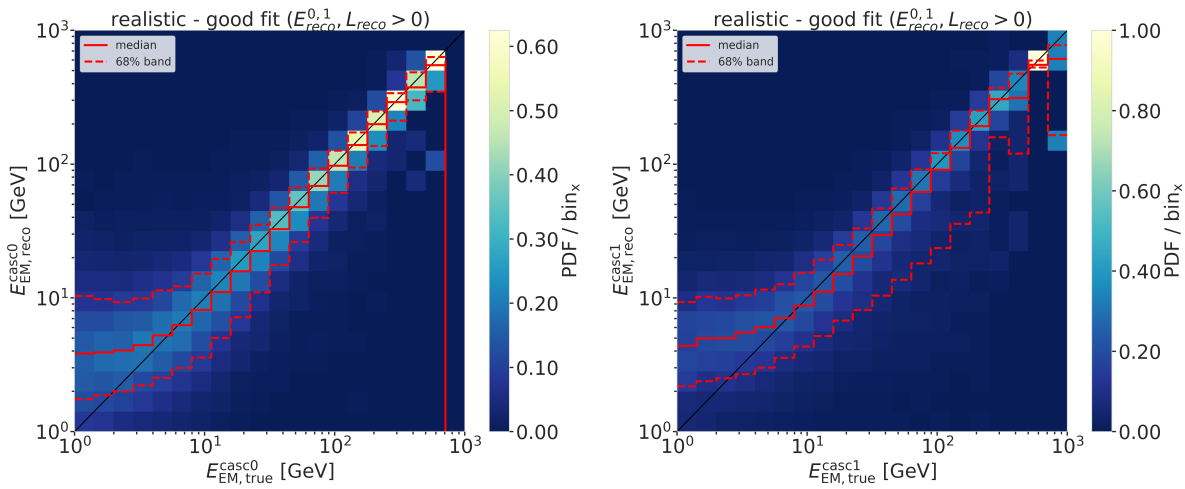


Figure 1.11

fix caption of this figure (RED)

Energy Resolution **Things to mention about the energy resolution:**

- Here it can be seen more clearly how the median total energy resolution starts to stabilize around 0.0 at 10 GeV, while for lower energies the reconstruction is over-estimating the true energy. This is a known behavior of energy reconstructions in IceCube, which is mainly due to a selection effect. Only events with a certain amount of light can be reconstructed, which means that the ones with true small energies that are still in the sample are events with over average light production due to fluctuations or other effects?

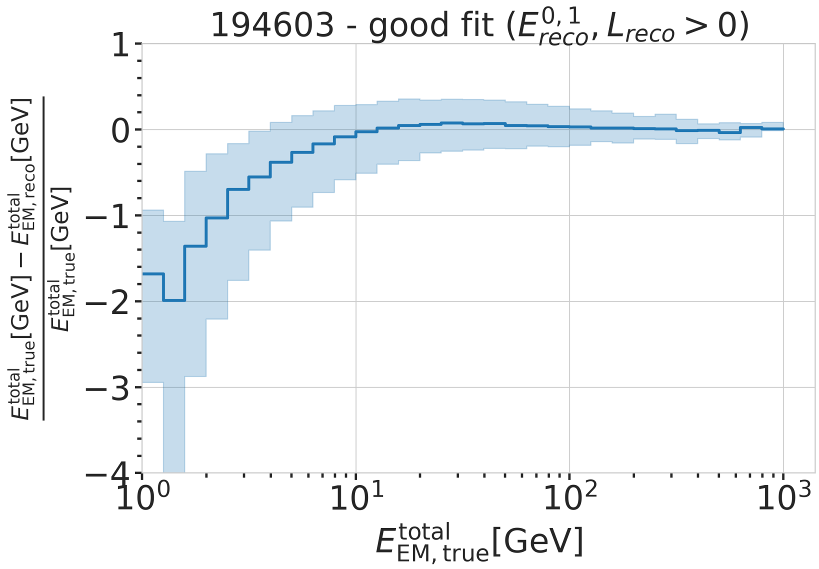


Figure 1.12

fix caption of this figure (RED)

Decay Length Resolution **Things to mention about the decay length resolution:**

- As already mentioned before, the decay length resolution is much worse than the energy resolutions. Figure 1.13 also shows that the median is below 0.0 for short true length and above 0.0 and approaching 1.0 for long true lengths.
- To investigate whether this is really due to the fact that one of the cascades is not observed, the decay length resolution was plotted against the total energy of the event and the minimum energy of the two cascades Figure 1.14.
- It can be seen that the median of the decay length resolution stabilizes at 0.0 for a total energy above 20 GeV, but the spread of the distribution is still quite large with a 1-sigma band of 80 % to 100 %.
- From the plot against the minimum energy it can be seen that the decay length resolution starts to be unbiased for a minimum energy of the cascades of 7 GeV, with an equivalently large spread.
- A preliminary takeaway from this is that the decay length reconstruction is not reliable at all for events with a total energy below 20 GeV or a minimum cascade energy below 7 GeV.

fix caption of this figure (RED)

fix caption of this figure (RED)

1.3.3 Low Energy Event Selection Efficiency

Make plot to show efficiency of the OscNext selection for HNL events. (ORANGE)

Plot energy (true total) and true decay length across the different levels (ORANGE)

Make table with the rates across the different levels for benchmark mass/mixing (ORANGE)

Discussion ideas:

- Show energy, length distribution across the different levels?
- Show efficiency as table across the different levels (MC events + fraction?)
- Compare this to BG efficiency? (maybe rather for the discussion)
- At which level does the selection reduce the HNL the most?

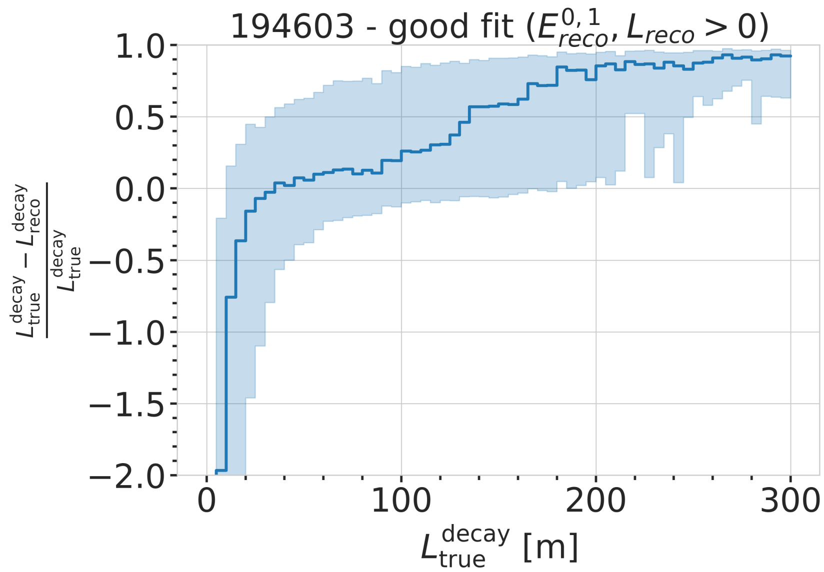


Figure 1.13

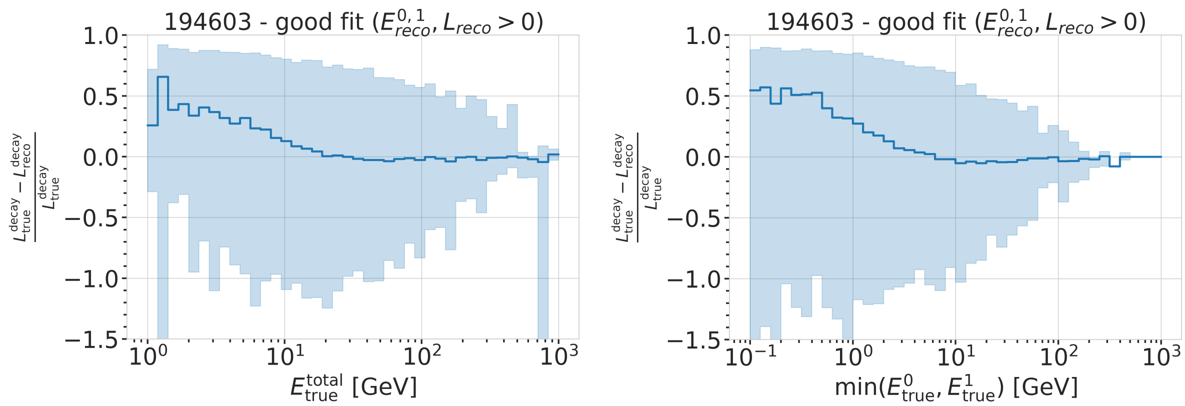


Figure 1.14

- Is there a place to improve the HNL selection? (Might have to factor in the BG efficiency, as well..)
- What of this might change with Upgrade? (maybe rather for the discussion)

Bibliography

Here are the references in citation order.

[1] R. Abbasi et al. "Measurement of Astrophysical Tau Neutrinos in IceCube's High-Energy Starting Events". In: *arXiv e-prints* (Nov. 2020) (cited on page 1).

[2] M. Usner. "Search for Astrophysical Tau-Neutrinos in Six Years of High-Energy Starting Events in the IceCube Detector". PhD thesis. Berlin, Germany: Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2018. doi: 10.18452/19458 (cited on page 1).

[3] P. Hallen. "On the Measurement of High-Energy Tau Neutrinos with IceCube". MA thesis. Aachen, Germany: RWTH Aachen Universisty, 2013 (cited on page 1).

[4] 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 1).

[5] M. G. Aartsen et al. "Energy Reconstruction Methods in the IceCube Neutrino Telescope". In: *JINST* 9 (2014), P03009. doi: 10.1088/1748-0221/9/03/P03009 (cited on page 1).

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