Bachelorarbeit
Muon track reconstructions in IceCube: Error estimations and their effects on point source sensitivity

Verfasser: Alexander Trettin (Matrikelnummer: 551 480)
Erstgutachter: Prof. Marek Kowalski
Zweitgutachter: Prof. Elisa Bernardini
Abgabe: 02.09.2016
Ich versichere, dass ich die Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, sowie die Zitate kenntlich gemacht habe.

.............................................
Datum

.............................................
Unterschrift
## Contents

1 Introduction
   1.1 Neutrino Astronomy ......................................................... 1
   1.2 The IceCube Neutrino telescope .......................................... 1
   1.3 Motivation for this work .................................................. 2

2 Reconstruction of muon events ............................................. 4
   2.1 Maximum Likelihood Method ............................................... 4
   2.2 Statistical model of the data ............................................. 4
   2.3 Error Estimation ............................................................. 6

3 Toy Simulations ........................................................................ 8
   3.1 Generating data for simulated muon events ........................... 8
   3.2 Reconstruction of simulated muon events ............................ 8

4 Toy-Monte-Carlo Point-Source analysis .................................... 12
   4.1 The Maximum-Likelihood method to skymaps with injected point source events .. 12
   4.2 Results .............................................................................. 14

5 Sensitivity .............................................................................. 21
   5.1 Definition and calculation .................................................. 21
   5.2 Sensitivity gain due to improved error estimation .................. 22

6 Summary .................................................................................. 24

References .................................................................................. 25
Abstract

IceCube is a 1 km$^3$ Čerenkov light detector located at the geographic South Pole, capable of detecting neutrinos at GeV energies and beyond via secondary Čerenkov emission. Focus of this work is the uncertainty on the reconstructed direction of muons, which dominantly stem from charged-current $\nu_{\mu}$-interactions in the ice. For this purpose, a simplified toy-simulation of muons with fixed stochastic components is produced and reconstructed with the so-called “SPE” likelihood. Two different hypotheses are used in the reconstruction: The first is a simplified one which assumes a constant energy loss of the muon in the detector, the other contains the exact time and magnitude of all energy losses as in the simulation. Finally, a search for a neutrino point source is simulated. The simulations take into account the reconstruction errors of the previously generated data. The errors of both reconstructions are rescaled to the same median value to isolate the effect of the spread of error estimates. Using both simulations, the two hypotheses can be compared with respect to their influence on the sensitivity of the telescope. The results show that assuming the correct stochastic energy losses in the reconstruction hypothesis improves the sensitivity by up to 20% for tracks where the stochastic losses make up around 90% of the total energy deposition. Such a fraction occurs regularly for realistic muon tracks at TeV energies and beyond. This sensitivity improvement due to the decreased spread of error estimates comes in addition to the decreased true reconstruction errors.

1 Introduction

1.1 Neutrino Astronomy

Neutrinos are nearly massless elementary particles which open up new observation avenues for astronomy besides electromagnetic and cosmic radiation. Since they only interact with the weak nuclear force, they aren’t deflected by electromagnetic fields and hardly ever absorbed along their path through space. This property gives them the potential to convey information about cosmic events near the edge of the observable universe. However, it also makes them hard to detect because they can traverse large amounts of matter without interacting with it at all. Furthermore, they can only be detected indirectly by observing particles which are produced on the rare occasion that they do interact with matter.

When neutrinos travel through matter, they often produce charged particles in the reaction

$$\nu_l + N \rightarrow l + X$$

where $l$ represents the flavor of the neutrino (e for electron, $\mu$ for muon, $\tau$ for tauon), $N$ is a nucleon and $X$ a shower of hadrons. This reaction is called a charged-current interaction. There are also neutral-current interactions, but those are not important for the purpose of this thesis. The charged lepton $l$ carries a part of the momentum of the incident neutrino. If the energy of the neutrino is high enough, the lepton will move along a path which is almost colinear with the path of the neutrino. The angular separation between them is less than one degree[1]. It can therefore be inferred that the neutrino which produced the lepton came from almost the same point in the sky and it is sufficient to reconstruct the direction of the lepton.

1.2 The IceCube Neutrino telescope

The IceCube Neutrino Observatory is a neutrino telescope located near the South Pole. It consists of approximately one cubic kilometer of Antarctic ice extending down to a depth between 1450 and 2450 meters. 5160 photon detectors (“digital optical modules”) have been distributed within its volume[2]. One way to observe the trajectories of charged particles is to let them pass through a
Figure 1: The IceCube Neutrino Observatory with the Eiffel Tower drawn for scale and the precursor to IceCube, the AMANDA II detector. Source:[2]

transparent medium such as water or ice. A charged particle, i.e. a lepton from a charged-current interaction, creates a cone of light behind itself when it travels through a medium at a speed greater than the speed of light in that medium. This light is known as Čerenkov radiation. This radiation is picked up by the photomultiplier tubes of the detector.

To finally detect point sources of neutrinos in the cosmos, the telescope has to observe many of these events. The majority of them are a background of atmospheric neutrinos which are created when cosmic rays interact with the atmosphere. To infer the existence of a neutrino source at a point in the sky, one must find a significant accumulation of neutrinos closely distributed around that point. The number of neutrinos needed to distinguish them from background noise determines the sensitivity of the telescope. The less observations are needed, the higher is the sensitivity.

This thesis focuses on the reconstruction of muons tracks. Muons are practical for a telescope because they produce long, straight tracks of Čerenkov light whose direction can be reconstructed with a relatively high accuracy. Electrons and taus, on the other hand, produce cascades for which the resolution in the direction is lower[1].

1.3 Motivation for this work

Muons traveling through ice lose their energy via Ionization, Bremsstrahlung, photo-nuclear interaction and pair production. At low energies (\(\sim 100\) GeV), ionization is the largest contributor to energy loss. At high energies (\(\sim 10\) TeV) the other processes begin to dominate.[3] While ionization
causes many small energy losses which can be modeled as a continuous loss-per-distance $dE/dx$, the other processes often create highly energetic secondary particles. Those energy losses have to be treated as a series of stochastic events. The fraction of energy that is lost in stochastic events relative to the total energy of the incident muon in this work is referred to as stochasticity$^1$:

\[
\text{stochasticity} = \frac{\text{energy of stochastic events}}{\text{total particle energy}}
\]

Muon events with low stochasticity can be approximated as tracks with a continuous energy deposition. This reconstruction hypothesis, however, becomes inaccurate with high particle energies ($> 10 \text{ TeV}$) when stochastic processes begin to dominate. In previous simulations, the estimated errors ($\sigma_{\text{ph}}$) of the reconstructed directions of high-energy particles were usually smaller than their true error ($\Psi$)$^4$. It is likely that this is due to the stochastic energy losses. The effect will be replicated in toy-simulations with variable stochasticities in section 3.2 of this work. In practice, the error estimate is corrected by an energy-dependent factor $s(E)$. It is an empirically determined function that scales the estimated error $\sigma_{\text{ph}}$ such that it agrees with the true error $\Psi$ in the mean over a large number of reconstructions.

\[
\frac{\Psi}{s\sigma_{\text{ph}}} = 1
\]

While the scaling can rectify the error estimation over many events, the error estimate for each individual event is still likely to deviate from its true value. This causes the error estimates for events with a high stochasticity to spread farther around their true value than those for events with a low stochasticity.

There are two signatures which are important for this analysis:

**Cascades:** When a muon loses a large fraction of its energy in one interaction, it creates a locally concentrated shower of secondary particles called a cascade. The Čerenkov radiation from a cascade forms a nearly circularly expanding light front.

**Tracks:** Muons also lose energy by the process of ionization. It behaves macroscopically like a relatively slow, continuous energy loss per distance. A track signature in the detector is a cone of Čerenkov radiation following a straight line$^2$.

These two signatures can be used in combination as hypotheses to explain the observations of the photomultiplier tubes. This work will deal with the reconstruction of muons, whose signature is a track with a number of cascades along its path.

Both effects, the reduced accuracy of the reconstructed direction and the larger spread of the error estimates, reduce the overall sensitivity of the detector when searching for neutrino sources in the universe. The aim of this work is to investigate the sensitivity loss specifically caused by the increased spread of the estimated errors, which results from modeling events with high stochasticities as simple tracks. This will be done by reconstructing a large number of simulated muon events once using the simplified track-only hypothesis for the muon and once using the entire signature including the track and cascades as hypothesis. This will generate pairs of estimated and true errors which will then be rescaled such that their median is the same for both reconstructions. This rescaling isolates the effect of the spread of the estimated errors. Using the rescaled data, a point-source search will be performed and the sensitivity flux estimated.

$^1$This definition is useful for this work since all simulated events will have a small, known number of cascades. It is not consistent if the number of cascades is unknown and possibly very large. For example, an event with 100% stochasticity and infinitely many cascades is indistinguishable from an event with a stochasticity of 0%.

$^2$See figure 2 on page 5 for an illustration of tracks and cascades.
2 Reconstruction of muon events

This work will focus on the reconstruction of muon events. Their signature is a long track, possibly with some cascades along its path. The goal of the reconstruction is to find the muon track whose direction best fits the data. This is accomplished using the maximum likelihood method.

2.1 Maximum Likelihood Method

The Maximum Likelihood (ML) method is a method to find the parameters of a statistical model which best fit the observed data. Let \( x_i \) be a set of observed data points and let \( f(x_i | \hat{\theta}) \) be a function which assigns each observation a probability density with the vector of parameters \( \hat{\theta} \). Then the goal of the method is to find the best fitting parameters, or maximum likelihood estimator \( \hat{\theta}_{\text{max}} \), which makes the observations most probable. The joint probability density of all observations is the product of all individual probability densities:

\[
f(x_1, x_2, \ldots, x_n | \hat{\theta}) = f(x_1 | \hat{\theta}) \times f(x_2 | \hat{\theta}) \times \ldots \times f(x_n | \hat{\theta})
\]

This function takes \( \hat{\theta} \) as a parameter and the observations as arguments. However, the maximum likelihood method assumes the observations to be fixed and attempts to optimize \( \hat{\theta} \). For this purpose, parameters and arguments are switched to define the likelihood as

\[
\text{likelihood} := \mathcal{L}(\hat{\theta} | x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i | \hat{\theta})
\]

The parameters \( \hat{\theta} \) are optimized to return the maximum likelihood.

This definition of the likelihood creates numerical difficulties because it returns extremely small numbers. For this reason, the calculations in this work will use the log-likelihood. Furthermore, the optimization algorithms which are readily available expect a function to be minimized, rather than maximized. Thus, the function used for reconstructions in this work is the negative log-likelihood:

\[
-\log \left( \mathcal{L}(\hat{\theta} | x_1, \ldots, x_n) \right) = - \sum_{i=1}^{n} \log \left( f(x_i | \hat{\theta}) \right)
\]

Because of the monotony of the logarithm, minimizing the negative log-likelihood will maximize the likelihood.

2.2 Statistical model of the data

Hypotheses with tracks and cascades

For the maximum likelihood method to work, one needs a statistical model for the data as a hypothesis with some parameters to optimize. The parameter to optimize is in this case the direction of the muon track. The first task is therefore to build a hypothesis for a muon inside the telescope which is parametrized only by two angles, its zenith and azimuth.

A muon hypothesis is built out of a track and optionally some cascades along its path. In this work, tracks are assumed to be infinite. This is a reasonable assumption since muons can travel several kilometers through ice and rock, much farther than the width and depth of the telescope. It is furthermore assumed that they are moving at the speed of light \( c \), which is a good approximation for the energies relevant to astrophysics.\(^3\)

\(^3\)For example, a muon (\( \mu^- : 105.65 \text{ MeV} / c^2 \)) at 10 GeV has a velocity of 0.99995c
A track is specified by:

- a vertex within the volume of the telescope
- its direction
- the time at which the muon passes by the vertex
- the energy the muon loses along the path through the detector

In realistic reconstructions, the position of the vertex also has to be optimized. This is neglected in this study. Instead, the vertex is always fixed to the “center of gravity” of all observed signals. Since the cascades are required to be positioned on the track of the muon, they are fully specified given only:

- their distance along the track from the vertex of the track
- the energy the muon loses via each cascade

Figure 2: Energy loss signatures used to model muons in the detector with at track on the left (a) and a cascade on the right (b). The images are not to scale.

**Probability Density Function**

Each observed event in this analysis corresponds to the arrival time of a photon at a photomultiplier tube. The probability density for these observations is described by a probability density function (PDF) \( f(t|n, \hat{\theta}) \). It returns the probability of counting a photon at time \( t \) given the index of the PMT and the track parameters \( \hat{\theta} \). In this work, the parameter of interest is the direction of the muon track, which implies \( \hat{\theta} = (\vartheta, \varphi) \) where \( \vartheta \) is the zenith and \( \varphi \) the azimuth of the direction.
The shape of the PDF depends on several factors such as the physical properties of the ice at the depth of the module, the time resolution of the electronics used, the orientation and distance of the module with respect to the muon track and so on.

**Combination of signatures in the “SPE” likelihood**

When reconstructing muons, the PDFs of the track \( f_{\text{Track}} \) and each cascade \( f_{c_i} \) have to be combined to yield the joint probability for each hit. Constant PMT noise is neglected in the simulation, but has to be used in the reconstruction for numerical reasons. This is achieved by weighing each probability density by their respective expected number of photons and dividing by the total number of expected photons. This yields

\[
f (t|n, \vartheta, \varphi) = \sum_{i=1}^{N} \frac{\mu_{n,c_i} f_{c_i} (t|n, \vartheta, \varphi)}{\mu_{n,\text{tot}}} + \frac{\mu_{n,\text{track}} f_{\text{track}} (t|n, \vartheta, \varphi)}{\mu_{n,\text{tot}}} + \frac{\xi}{\mu_{n,\text{tot}}}
\]

where \( N \) is the number of cascades, \( \mu_{n,c_i} \) is the number of expected photons in module \( n \) due to cascade \( i \). \( \mu_{n,\text{track}} \) and \( \mu_{n,\text{noise}} \) are the number of expected photons in module \( n \) due to the track and PMT noise, respectively. \( \xi \) is the constant rate of PMT noise. This is a special case of the so-called “SPE” (single photon electron) likelihood[1]. When the reconstruction hypothesis has no cascades, the sum in both expressions is omitted, and one obtains the standard SPE formula.

Plugging this statistical model into the log-likelihood function described in section 2.1, one arrives at the following function to be maximized:

\[
\log \left( L \left( \hat{\theta}|t_1, \ldots, t_n \right) \right) = \sum_{i=1}^{n} \log \left( f (t_i|n_i, \vartheta, \varphi) \right)
\]

where \( t_1, \ldots, t_n \) are the arrival times of all observed photons and \( n_i \) is the index of the PMT in which photon \( i \) has been detected.

**2.3 Error Estimation**

**Computation of the covariance matrix**

To compute error estimates for the reconstructed direction, the likelihood function in the region close to its maximum is modeled as a bivariate normal distribution. Because the direction \( (\phi', \vartheta') \) uses spherical rather than cartesian coordinates, two conditions need to be fulfilled to justify the approximation:

1. The coordinates are rotated such that the center of the distribution is close to where \( \vartheta = \frac{\pi}{2} \). The rotated coordinates \( (\phi, \vartheta) \) are approximately cartesian in the region of interest (see figure 3).
2. The standard deviation of the distribution is small relative to \( \pi \).
The coordinates $(\phi', \vartheta')$ are transformed to $(\phi, \vartheta)$ which can be used as cartesian coordinates in the region close to the direction of the track.

The likelihood function can be approximated as follows:

$$L(\hat{\theta}|t_1, \ldots, t_n) \approx f(\bar{x}|\bar{\mu}, \Sigma) = \frac{1}{2\pi \sqrt{\det(\Sigma)}} \exp \left( -\frac{1}{2} (\bar{x} - \bar{\mu})^T \Sigma^{-1} (\bar{x} - \bar{\mu}) \right)$$

$$\bar{x} = \begin{pmatrix} \phi \\ \vartheta \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \sigma_\phi^2 & \text{cov}(\vartheta, \phi) \\ \text{cov}(\vartheta, \phi) & \sigma_\vartheta^2 \end{pmatrix}$$

where $\vartheta$ and $\phi$ are zenith and azimuth, respectively, $\bar{\mu}$ is the position of the maximum and $\Sigma$ is the covariance matrix of the distribution. $\sigma_\phi^2$ is the variance of the zenith, $\sigma_\vartheta^2$ the variance of the azimuth and $\text{cov}(\vartheta, \phi)$ is the covariance of both coordinates. Since the reconstruction uses the negative log-likelihood instead of the likelihood function, one takes the negative logarithm of the model PDF:

$$g(\bar{x}|\bar{\mu}, \Sigma) = -\log(f(\bar{x}|\bar{\mu}, \Sigma)) = -\log \left( \frac{1}{2\pi \sqrt{\det(\Sigma)}} \right) + \frac{1}{2} (\bar{x} - \bar{\mu})^T \Sigma^{-1} (\bar{x} - \bar{\mu})$$

This corresponds to a paraboloid in two dimensions with a maximum at position $\bar{\mu}$. Such a paraboloid may also be represented as

$$g(\bar{x}|\bar{\mu}, \Sigma) = \delta + \bar{\beta} \cdot \bar{x} + \frac{1}{2} \bar{x}^T \Gamma \bar{x}$$

where $\Gamma = \Sigma^{-1}$. To find the parameters $\delta$, $\bar{\beta}$ and $\Gamma$ which best fit the data, the method of least squares is used, which can be done analytically according to [5]. Only the result for $\Gamma$ is needed to estimate the error of the reconstruction.

**Computation of the median error**

The resolution of the track reconstruction is determined by the median error of the direction. It is defined as the radius from the reconstructed direction which contains the true direction with a probability of 50%. To determine the median error margin, a bivariate normal distribution
is numerically integrated from its center outward until a cumulative probability of 0.5 has been reached.

\[ 0.5 = \int_0^R \int_0^{2\pi} \frac{r}{2\pi \sqrt{\det(\Sigma)}} \exp \left( -\frac{1}{2} (r \sin(\gamma), r \cos(\gamma)) \Sigma (r \sin(\gamma), r \cos(\gamma))^T \right) \, dr \, d\gamma \]

\( R \) is the median radius and \( \Sigma \) is the covariance matrix calculated in the paraboloid fit.

3 Toy Simulations

To do statistical analyses on event reconstructions, many (at least \( O(1000) \)) events are needed. Furthermore, the true direction of each reconstructed track has to be known to determine the true error of the reconstruction. For this reason, this work uses muon events which are generated in a toy simulation. Such a simulation generates data which behaves approximately like that of real events without actually simulating the physical processes involved, hence it is referred to as a “toy” simulation. The PDF for the photon generation are calculated using pre-computed spline tables, which are based on Monte Carlo simulations of many different source configurations[6, 7].

3.1 Generating data for simulated muon events

The process of data generation is essentially the opposite of the event reconstruction discussed in section 2. Instead of finding the track direction which best fits the observed data, the track parameters are chosen arbitrarily. The data is sampled from the probability density function corresponding to these parameters. This is done in four steps:

1. Choose the direction of the track to be generated at random and set its supporting vertex to some point within the geometry of the detector. If the event is also to contain cascades, set them to random positions along the track within 500 m of the vertex. Set the energy of the track and the energies of the cascades according to the desired stochasticity. This sets the event parameters \( \hat{\theta} \).

2. For each optical module \( n \), compute the expected number of photons \( \mu_n \) for the time frame of the event given the event parameters \( \hat{\theta} \).

3. For each optical module \( n \), set the number of actually observed photons by drawing from a Poisson distribution with an expectation value of \( \mu_n \).

4. Set the arrival time for each photon hit by drawing from the time dependent probability density function \( f(t|n, \hat{\theta}) \) for the respective module.

Steps 2 - 4 are repeated for each photon source i.e. the track and each cascade. One arrives at a dataset of photon hits which follows the probability density functions associated with the parameters of the event.

3.2 Reconstruction of simulated muon events

The reconstruction is applied to the data generated by the toy simulation as described in section 2. Every reconstruction is run twice. The first run uses the correct hypothesis (track and cascades) with which the data was first generated. The second run models the event as a simple track without cascades. This allows for a direct comparison of the results of both reconstruction methods. The minimum of the log-likelihood function was calculated using the MINUIT[8] optimization routine
which is readily available in the Python module \texttt{iminuit}. After the minimum was found by \texttt{MINUIT}, the covariance matrix of the direction coordinates is calculated using the parabola fit described in subsection 2.3. The median error $\sigma_{pb}$ is calculated by numeric integration of the gaussian which corresponds to the fitted covariance matrix. The true error $\Psi$ denotes the angle between the reconstructed track direction and the direction used to generate the data. The fraction $\Psi/\sigma_{pb}$ defines the \textit{pull} of the event:

$$\text{pull} := \frac{\Psi}{\sigma_{pb}}$$

Ideally, the median of the pull over a large number of events is 1, since $\sigma_{pb}$ was computed as the 50\% quantile of the gaussian.

**Results for events with highly stochastic energy losses**

In the following, all events consist of a muon with a total energy of 10 TeV, 90\% of which are lost in three cascades. Figure 4 shows the true errors and estimated errors using the simplified hypothesis and the complete event hypothesis including cascades. While the estimated errors for both reconstructions are of similar magnitude, the true errors are often larger than the estimated errors using the track-only hypothesis.

![Figure 4: True errors and estimated errors using the simplified track-only hypothesis and the complete event hypothesis including cascades.](image-url)
Figure 5: Normalized histogram of the logarithm of the pull using the complete and the simplified hypothesis. To compare the spread of both distributions, they were shifted in this diagram such that each median is zero. The shaded region, containing 26% of the events, represents the amount of events whose pull has been moved from the tails of the distribution towards the middle by using the full hypothesis reconstruction.

The pull distribution of the complete reconstruction has a median of 1.02, suggesting that the estimated errors are close to the true errors. The pull distribution of the track-only hypothesis has a median of 2.6, which suggests that the error is generally estimated too small. However, the median of the pull distributions can be corrected by rescaling the estimated errors, which is done in realistic reconstructions. The rescaled distributions are shown in figure 5. The figure shows that the pull distribution is tighter around the median for the full hypothesis than the pull distribution for the track-only hypothesis. This confirms what has already been stated in the motivation for this work: The estimated errors spread farther from their respective true errors when only the track hypothesis is used in the event reconstruction. This may be relevant for later point-source reconstructions, because the error estimates of each reconstructed track determine their weights in the reconstruction process. If the error of a track is estimated too small, its contribution to a point in the sky only a short angular distance away may be neglected, where it really shouldn’t be. Likewise, if the error of a track is estimated too large, its weight may also be too large for points far away and too small for points nearby. In any case, a wrong error estimate reduces the sensitivity of the point-source analysis.

One can thus improve a signal by moving it from the tails of the pull distribution towards the middle. In figure 5, the fraction of events whose error estimates would be improved this way by using the complete hypothesis reconstruction is 26%. This estimate only counts the number of signals which have been improved, but does not take into account how large the improvement for each signal is. It can be expected that the actual improvement in sensitivity will be in the same order of magnitude.
Results for variable stochasticities

The results above show that reconstructions using the full hypothesis including all cascades yield much more accurate error estimates than reconstructions using the simplified track hypothesis. Another test of the reconstructions is to vary the stochasticity between a low and a high value. For low stochasticities, the simple track hypothesis is a good approximation of the truth and should therefore yield accurate error estimates. As the stochasticity increases, the simple track becomes an increasingly inaccurate model and the error estimates become inexact. The full hypothesis including all cascades, however, is accurate for all stochasticities. The error estimates of reconstructions which use the full hypothesis should therefore not show any dependence on the stochasticity of the event. In figures 6 and 7, the dependence of the error estimates on the stochasticity is modeled using linear regression.

In the first figure, the hypothesis used for the reconstruction is an infinite muon track without cascades. The stochasticity has been set randomly between 0.1 and 0.9. A linear regression through ~5000 reconstructed events shows that the errors of events with a higher stochasticity tend to be underestimated more often than those of events with low stochasticities. The dependence on the stochasticity is statistically very significant with a p-value of the null-hypothesis $p = 1.1 \times 10^{-87}$.

In the second figure, the hypothesis used for the reconstruction is the full hypothesis including all cascades. A linear regression through ~5000 reconstructed events has a p-value for the null-hypothesis $p = 0.87$. This means that one can confidently assert that the pull-dependence on the stochasticity of the event has indeed disappeared.

Figure 6: Accuracy of error estimation as a function of stochasticity for the track-only reconstruction. $\Psi$ is the true error of an event while $\sigma_{pb}$ is the median error estimated by the reconstruction. Ideally $\langle \Psi/\sigma_{pb} \rangle = 1$ for a large number of events.
4 Toy-Monte-Carlo Point-Source analysis

The next step is to determine the sensitivity of the telescope. The challenge of detecting a neutrino source is that the signal events it generates have to be distinguished from background events which usually outnumber the signal events. To achieve this, one hypothesizes the existence of a source at a certain position in the sky and estimates the number of signals most likely emitted from that source using the Maximum Likelihood method. A hypothesis test then determines whether the result is significant enough to reject the null-hypothesis (i.e. the hypothesis that there is no neutrino source). In this section, a large number of these reconstructions will be performed to determine the sensitivity to a flux of muons.

4.1 The Maximum-Likelihood method to skymaps with injected point source events

The first step in determining the sensitivity of the telescope is to do a point source reconstruction using toy simulations of skymaps. They are generated by first inserting a background of events uniformly distributed over the entire solid angle of $4\pi$. Then, a point source is simulated at the north pole by injecting a certain number of events whose angular separation from the source has the same distribution as the reconstruction errors computed in the previous section. Putting the source at the north pole can be done without loss of generality, since the spherical coordinates can always be rotated to satisfy this condition. Hypothesizing the source’s existence at the north pole and applying the Maximum-Likelihood method, the injected number of signal events can be
reconstructed.

To recall section 2.1, the likelihood to be maximized for a number $N$ of observations is generally defined as

$$L(\hat{\theta}|x_1, ..., x_n) = \prod_{i=1}^{n} f(x_i|\hat{\theta})$$

with the estimator $\hat{\theta}$, the observations $x_i$ and the probability density function $f(x_i, \hat{\theta})$.

In the point source reconstruction is the estimator $\hat{\theta} = n_s$ is the number of signals emitted by the source. The observations $x_i$ are the angular distances of the observed signals from the hypothesized source. The probability density function for each event is the combination of the background and signal PDFs. The background PDF is constant for all events

$$f_b(x_i) = \frac{1}{4\pi}$$

while the signal PDF is modeled as a two-dimensional symmetric normal distribution

$$f_s(x_i) = \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{x_i^2}{2\sigma_i^2}\right)$$  \hspace{1cm} (2)

It is centered at the origin of coordinates. $x_i^2 = x_{i,1}^2 + x_{i,2}^2$ is the distance of the signal from the origin and $\sigma_i$ is the estimated error for the signal$^4$. This model can only be valid as long as $\sigma \ll \pi$, since otherwise the coordinates can no longer be approximated as being cartesian.

The signal- and background PDFs are weighed by their contribution to the total number of observes signals $N = n_s + n_b$. $n_s$ and $n_b$ are the number of signal and background events, respectively.

$$L(n_s|x_1, ..., x_n) = \prod_{i=1}^{N} \left(\frac{n_s}{N} f_s(x_i) + \frac{N-n_s}{N} f_b(x_i)\right)$$

As in the track reconstruction, the logarithm of the likelihood will be used for numerical reasons. In the logarithm, the division by $N$ yields only a constant offset which can be ignored for the maximum likelihood reconstruction. The function to be maximized is therefore:

$$\ln(L) = \sum_{i=1}^{N} \ln\left(\frac{n_s}{N} f_s(x_i) + \frac{N-n_s}{N} f_b(x_i)\right)$$

$^4$In this case, the estimated error corresponds to the 1-sigma region of a bivariate normal distribution which contains the 39.35% quantile, rather than the 50% quantile as in section 2.3. For a symmetric normal distribution, both are related by a constant scaling factor of 1.1774. The factor can be determined from the bivariate standard normal distribution (symmetric, $\sigma = 1$)

$$0.5 = \int_{0}^{R} \int_{0}^{2\pi} \frac{r}{2\pi} \exp\left(-\frac{r^2}{2}\right) \, dr \, d\phi$$

$$= \int_{0}^{R} r \exp\left(-\frac{r^2}{2}\right) \, dr$$

$$= 1 - \exp\left(-\frac{R^2}{2}\right)$$

$$\Leftrightarrow 0.5 = \exp\left(-\frac{R^2}{2}\right)$$

$$\Leftrightarrow \sqrt{-2\ln(1/2)} = R \simeq 1.1774$$

$$\Leftrightarrow R \simeq 1.1774\sigma$$
Although the maximum likelihood estimator is asymptotically unbiased for an infinite number of observations, it doesn’t necessarily have ideal properties for a finite set of observations\textsuperscript{[9]}. In the next section, tests will show that for finite observations, this estimator may underestimate $n_s$.

**Test Statistic**

After finding the Maximum Likelihood estimate $\hat{n}_s$, a hypothesis test is used to decide if the result is significant enough to be considered a discovery of a point source. The null-hypothesis for this test is the hypothesis that none of the signals originate from a neutrino source, i.e. $n_s = 0$. The test statistic $\lambda$ is defined as a function of the likelihood ratio:

$$\lambda = -2 \ln \left( \frac{L(n_s = 0)}{L(n_s = \hat{n}_s)} \right)$$

The hypothesis being tested is a nested hypothesis, meaning that the parameter $n_s$ is fixed for the null-hypothesis while it is free to vary otherwise. Because $\hat{n}_s$ is by definition the estimate with the largest likelihood, it is clear that $L(n_s = 0) \leq L(n_s = \hat{n}_s)$. This means that the test statistic is zero if $\hat{n}_s = 0$ and has a positive value otherwise. The significance of the result is determined by the probability that a reconstruction on a skymap with only background events would produce a test statistic value equal to or greater than the one observed. It is expressed in the p-value

$$p = \int_{\lambda_{obs}}^{\infty} g(\lambda, H_0) \, d\lambda$$

where $g(\lambda, H_0)$ is the distribution of test statistic for the null-hypothesis (only background events) and $\lambda_{obs}$ is the observed test statistic. The smaller the p-value is, the more significant is the result.

According to Wilk’s Theorem, the test statistic for a nested hypothesis follows a $\chi^2$-distribution with degrees of freedom equal to the number of constraints imposed by the null-hypothesis\textsuperscript{[9, p. 273]}, which is 1 in this case.

In this reconstruction, $n_s$ is bounded at zero, at which point the test statistic is also zero. All reconstructions where $n_s$ would be negative without the boundary (under-fluctuations) instead return $n_s = 0$ and a test statistic of $\lambda = 0$, which means that Wilk’s Theorem can’t be applied. Rather, the test statistic (in the large sample limit) is distributed\textsuperscript{[10, 4]} according to

$$f(\lambda)_{\text{lim N} \to \infty} = 0.5 \cdot \delta(x) + 0.5 \cdot \chi^2_1(x)$$

where the delta distribution captures all under-fluctuations. This distribution is referred to as half-chi-square distribution.

### 4.2 Results

First, tests using known distributions are performed to determine whether the maximum likelihood estimator is biased for the finite set of observations expected within a year’s worth of collected detector data. Then, the point-source reconstruction is performed using the results from the muon track reconstructions described in previous sections.

**Test using normal signal distribution**

A first test of the validity of the reconstruction method is to sample the signal events from a symmetric normal distribution. To make the conditions comparable to realistic point-source searches, the background consists of 100000 events\textsuperscript{[11]} uniformly distributed over the entire sky while the normal distribution used for the signal events has a standard deviation of $1^\circ$\textsuperscript{5}. This standard

\textsuperscript{5}This corresponds to the approximate median angular resolution of the IceCube detector for neutrinos with an energy of 1 TeV. The resolution is better than $1^\circ$ for event energies above 10 TeV\textsuperscript{[4]}.
deviation is used to compute $f_s(x_i)$ for all signals. The azimuth coordinate is sampled uniformly between 0 and $2\pi$. The zenith of the background events is distributed according to

$$Z = \arccos(X)$$

where $X$ is a variable distributed uniformly between -1 and 1. Figure 8 illustrates an example of a skymap with 20 injected source events. Figure 9 shows a histogram of 10000 reconstruction results for 10 injected source signals in each simulated skymap. The median of all reconstruction results is indeed close to the truth with a value of 9.79. The slight under-estimation is unlikely a statistical artifact, as producing the same histogram 20 times (with different seeds used by the random number generator) yields a mean result of $9.81 \pm 0.05$.

![Skymap with a source at the North Pole](image)

Figure 8: A skymap with uniformly distributed background signals and 20 signals from a source at the north pole. The standard deviation of the source PDF is $1^\circ$, the radius in this plot is limited to $12^\circ$. The density of background events is equivalent to the density achieved by sampling 100000 events uniformly over the entire sky.
Figure 9: Point source reconstruction results for 10 normally distributed source signals. The median of the reconstructed number of source signals in this histogram of 10000 reconstructions is 9.79, slightly less than the true value of 10. Since the number of source signals cannot be less than zero, the distribution of results is cut off at zero and results that would have otherwise been negative are put into the zero bin.

**Test of zero-fraction**

Another validity test is to run the reconstruction without any injected signals. If the median of all reconstructed $n_s$ converges to the truth, which is 0, it is to be expected that half of the reconstructions return a value of zero (since the algorithm is bounded by a minimum of $n_s = 0$) and half a value above zero. The same holds true for the test statistic, since it is zero if and only if $n_s$ is zero. It is useful to define the *zero-fraction* as:

$$\text{zero-fraction} := \frac{\text{number of reconstructions where } n_s = 0}{\text{number of reconstructions}}$$

To test this expectation, 20 histograms were produced, each one out of 10000 reconstructions with 100000 background events and no source signals. The mean and standard deviation of the zero-fraction of those histograms is $0.516 \pm 0.004$. This result suggests, again, that the number of source signals tends to be slightly underestimated.

One possible explanation for the under-estimation of signals is the limited density of background events and the relatively small area of the 1-sigma region of the signal PDF. With a lower density of background events and/or a smaller standard deviation of the source PDF, the odds increase that no background event lies within an area where the value of the source PDF isn’t negligibly small. This leads to the number of signals being under-estimated more often, such that the zero-fraction increases. Figure 11 demonstrates how increasing the density of background signals with a fixed source PDF moves the zero-fraction from almost 1 close to 0.5.
The zero-fraction also corresponds to the fraction of events which fall into the delta distribution of the test statistic distribution (see section 4.1). The results of this test suggest that the number of events is not enough to fulfill the half-chi distribution limit for the test statistic distribution. This effect has already been observed as the “under-sampling problem” in [4]. Histograms of the test statistic for those events where it is larger than zero show that they are still closely following a \( \chi^2 \)-distribution (see figure 10).

Figure 10: Test statistic distribution for 10000 skymaps with only background events in comparison to a \( \chi^2 \)-distribution for 1 degree of freedom. Only the reconstructions where the test statistic is greater than zero are included.
Figure 11: Zero-fraction as a function of the number of background events. For each data point, the zero-fraction was evaluated for 20 histograms, each containing the results of 10000 point-source reconstructions. All reconstructions use a symmetric normal distribution with a standard deviation of 1° as signal PDF. The value and error bar correspond to the mean and standard-deviation, respectively. It shows that the asymptotic limit where the median of the estimator of $n_s$ converges to the true value of 0.5 is only reached above $10^6$ observed background events.

Results using data from muon track reconstructions

After verifying that the point-source reconstruction is asymptotically accurate, it can now be applied using the data from the muon track reconstructions described in section 2. Instead of a normal distribution with a standard deviation that is constant for all signals, each signal now has an individual standard deviation, which is the estimated error of the reconstruction (see equation (2)). The true errors of the signals originating from the source (i.e. their distance from the source) now correspond to the true errors of the track reconstruction. This means in turn that, for every signal, a true error and an estimated error of a unique muon track reconstruction is needed. The computational effort to do a muon track reconstruction for every single signal, however, is prohibitive for the amount of signals in this point source reconstruction ($\mathcal{O}(10^5)$). The errors will instead be sampled from a distribution which models the reconstruction results. One way to sample pairs of true and estimated errors without assuming the shape of their underlying distribution is to use a kernel density estimator (KDE). The KDE, given a data set of $n$ true and estimated errors, assigns every point $\tilde{x}$ in the two-dimensional space of true and estimated errors a probability density by evaluating

$$f_h(\tilde{x}) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{\tilde{x} - \tilde{x}_i}{h}\right)$$
where \( \vec{x}_i = (\Psi_i, \sigma_i) \) is the 2-vector of estimated and true error of one event in the dataset, \( h \) is the bandwidth of the estimator and \( K(\vec{x}) \) is the kernel, in this case a bivariate standard normal distribution\(^6\). The gaussian KDE including automatic bandwidth estimation is conveniently implemented in Python’s scipy module which also contains a resampling function to produce data following the estimated probability density function. The KDE for both reconstructions (full hypothesis and infinite muon track) are shown in figure 12. At this point, it is important to remember that this work investigates the sensitivity gain only due to the reduced spread of the estimated errors, not the reduced true errors, when applying the full hypothesis reconstruction. To isolate that effect, the distribution of \( \log(\Psi, \sigma) \) is shifted such that the median of estimated and true error is the same for both reconstruction types.

Figure 12: KDE for the complete hypothesis reconstruction and the infinite muon reconstruction. \( \Psi \) is the true error and \( \sigma_{pb} \) is the estimated (median) error. Both were calculated in degrees. The points in the plot are the data used to compute the KDE. The KDE was produced using the logarithm of the errors for numerical reasons. The logarithm of the pull \( \log_{10}(\Psi/\sigma_{pb}) \) can be interpreted in these plots as the distance from the line where \( \Psi = \sigma_{pb} \) (shown as dashed line). While the data points for the full hypothesis reconstructions stay relatively close to that line, they significantly spread for the infinite muon reconstruction.

In summary, the positions of the background signals are sampled uniformly over the skymap according to (3). The estimated errors for those signals are sampled using the KDE. Then, a pair of an estimated and a true error is sampled from the KDE for each source signal event. The true error corresponds to the distance of the signal from the source. Finally, the number of source signals \( n_s \) is reconstructed using the maximum likelihood method described in the previous section.

Figure 13 shows the point-source reconstruction results for the full hypothesis and the infinite muon reconstructions after injecting 10 source signals into the skymap. The KDEs were shifted such that the median of estimated and true errors were 1° for both reconstruction methods. This makes the results comparable to the first test, where the signals were distributed according to a normal distribution with a standard deviation of also 1°. The results of the full hypothesis reconstruction differ only slightly from the previous results obtained by sampling the signals from a standard normal distribution. The amount of source signals is only slightly underestimated with

\(^6\)A more detailed description of kernel density estimators can be found in [12]
a median of 8.8. The distribution of the results is tighter around the median, which results in less events being in the zero-bin (∼320 for the full hypothesis reconstruction instead of ∼430 for the reconstruction of normally distributed events).

The results of the simplified infinite muon reconstructions, on the other hand, significantly underestimate the amount of signals with a median of 7.1. The distribution of results also spreads much stronger, resulting in more events in the zero-bin (∼670).

The distribution of the corresponding test statistic is shown in figure 14. Figure 15 shows the distribution of the test statistic for skymaps without any signal events, $g(\lambda, H_0)$. This distribution can be used to calculate the p-value for a skymap. If the value of the test statistic for a particular skymap is higher than the $3\sigma$-level (99.73% quantile of the test statistic distribution for background-only skymaps), it counts as a “$3\sigma$ discovery”. The $3\sigma$-level is

$$\lambda_{3\sigma, \text{Full}} = 7.46$$

for reconstructions using the full hypothesis and

$$\lambda_{3\sigma, \text{Track}} = 7.39$$

for reconstructions using the track-only hypothesis.

(a) Results with data from the full hypothesis reconstruction.

(b) Results with data from the infinite muon reconstruction.

Figure 13: Reconstruction results for $n_s$ for skymaps with 10 injected source events.
5 Sensitivity

5.1 Definition and calculation

The sensitivity flux is the neutrino flux of a source for which the observed test statistic is greater than the median of \( g(\lambda, H_0) \) in 90% of all reconstructions\cite{4}. The neutrino flux is defined as the expected number of neutrinos detected from a point-source within a year. Since the test statistic is zero for more than half of the reconstructions on background-only skymaps (see “zero-fraction” in section 4.2), the median of \( g(\lambda, H_0) \) is not well defined. In order to calculate a relative sensitivity gain, it is still sufficient to calculate the sensitivity flux assuming that the median of \( g(\lambda, H_0) \) is zero. The result is cross-checked by calculating the 3\( \sigma \)-discovery flux, which is the flux at
which the median of the test statistic distribution corresponds to a p-value of $\sim 0.003$. The corresponding value of the test statistic has been calculated for both reconstruction methods in the previous section.

Since all neutrinos are detected independently from one another, the actual number of detections follows a Poisson distribution, where the neutrino flux is the expectation value of that distribution. In the previous section, point source reconstructions were run using a known number $n_s$ of injected source signals. The procedure is repeated for all values of $n_s$ between 0 and 29 and the test statistic is calculated for all results. The resulting 30 histograms of the test statistic are normalized and stacked together, where each histogram $i$ (with a number of injected source signals $n_{s,i}$) is weighed by the poisson factor $p_i$. This factor corresponds to the probability of observing $n_{s,i}$ source signals given the neutrino flux $\Phi$.

$$p_i(\Phi) = \frac{\Phi^{n_{s,i}}e^{-\Phi}}{n_{s,i}!}$$

The stacked histogram for a given neutrino flux is calculated as follows: Let $N_i$ be the total number of reconstruction results in histogram $i$ and let $n_{k,i}$ be the number of reconstructions in bin $k$ in histogram $i$. Then, the content of bin $k$ in the stacked histogram $n_{k,\text{stack}}$ is:

$$n_{k,\text{stack}} = \sum_i p_i(\Phi) \frac{N_i}{n_{k,i}} n_{k,i}$$

For each histogram $i$, the number of results for which the test statistic is greater than zero is counted. Let that number be $n_{\lambda>0,i}$, then the sensitivity flux is calculated by solving

$$0.9 = \sum_i p_i(\Phi_{\text{sens}}) \frac{N_i}{n_{\lambda>0,i}}$$

The equation can be modified to calculate the 3$\sigma$-discovery flux. One counts the number of results in each histogram for which the test statistic is greater than the 3$\sigma$-level (see previous section). Let that number be $n_{3\sigma,i}$ for histogram $i$, then eq. (4) becomes

$$0.5 = \sum_i p_i(\Phi_{3\sigma}) \frac{N_i}{n_{3\sigma,i}}$$

### 5.2 Sensitivity gain due to improved error estimation

The sensitivity flux can now be calculated for the simplified track reconstruction and the full hypothesis reconstruction. First, the histograms of reconstructions were calculated for all possible numbers of injected source signals in 0, 1, ..., 29. Each contained 10000 reconstructions. The poisson factors for reconstructions with a greater number of injected source signals were considered small enough to be neglected. To calculate empirical error margins, the entire calculation was performed 6 times. The mean result with standard deviation is

$$\Phi_{\text{sens, Track}} = (9.22 \pm 0.06) \text{ Events/Year}$$

for the track-only reconstruction, while it is

$$\Phi_{\text{sens, Full}} = (7.52 \pm 0.04) \text{ Events/Year}$$

for the full hypothesis reconstruction. One arrives at the same mean result by stacking the histograms of all 6 runs, which gives 60000 reconstructions for each value of $n_s$. The sensitivity flux
is 18.4% lower for the full hypothesis reconstruction compared to the track-only reconstruction. The 3$\sigma$-detection flux is also calculated using the same histograms with a result of

$$\Phi_{3\sigma, \text{Track}} = (16.03 \pm 0.05) \text{ Events/Year}$$

for the track-only reconstruction and

$$\Phi_{3\sigma, \text{Full}} = (12.97 \pm 0.04) \text{ Events/Year}$$

for the full hypothesis reconstruction, which gives an improvement of 19.1%. The stacked histograms of the test statistic for the 3$\sigma$-detection flux is shown in figure 16. The median of the distribution corresponds to the 3$\sigma$-level which has previously been calculated in section 4.2.

This shows that one needs \( \sim 19\% \) less neutrino flux to detect a point-source of muon neutrinos in the cosmos using the full hypothesis including all cascades as the reconstruction hypothesis rather than the simplified track-only hypothesis. The result is compatible with the first estimate acquired earlier in figure 5 on page 10. It should be noted right away that this result can only be understood as an upper bound to the improvement in sensitivity. Firstly, it was computed from muon events where 90% of the muon’s energy is lost in three cascades. This is a rather extreme case where the minimally ionizing track hypothesis is a poor approximation of the truth. Figure 17 shows the sensitivity flux and the 3$\sigma$-detection flux of both reconstructions for different stochasticities between 0.1 and 0.9. Both flux constructions behave very similarly and yield approximately the same sensitivity improvement. The sensitivity gain is smaller for less extreme muon events where only 50% of the muon’s energy is lost in cascades.

Secondly, the full hypothesis contained the correct number of cascades, their true energies and their true distances from the track vertex. All of these variables would have to be approximated to the data in real muon track reconstructions, which increases the dimensionality of the problem and introduces more uncertainties. One such method which approximates the stochastic energy losses is described in [6].

Figure 16: Poisson stacked test statistic histograms for the track-only and the full hypothesis reconstructions at their respective 3$\sigma$-detection flux.
Figure 17: Sensitivity flux and 3σ-detection flux for stochasticities between 0.1 and 0.9 for reconstructions using the track-only signature and the complete signature containing the track and all cascades. The flux was calculated using histograms which each contain 10000 point-source reconstructions. The calculation was performed 6 times with different pseudo-random numbers for every data point to yield a mean and a standard deviation (shown as error bars).

6 Summary

The aim of this work was to investigate how the hypothesis used to reconstruct muon tracks in the IceCube neutrino telescope affects the error estimates of the track direction. It was furthermore studied how the error estimates in turn influence the sensitivity of the telescope when searching for neutrino sources in the cosmos. To study the effect, muons were simulated in toy-simulations whose event signatures consisted of an infinite track and three cascades at random positions along the track. From the data generated in those simulations, the track direction was reconstructed using two different hypothesis: One hypothesis contained only a track, while the other contained the track and all cascades, their energies and their position along the track. In section 3 it was shown that increasing the fraction of energy contained in the cascades causes the error estimates of the track-only reconstruction to become less accurate (see also figure 6 on page 11). If the cascade energies are included in the reconstruction hypothesis, however, the errors are estimated with the same accuracy independent from the energy contained in the cascades. This result confirms that the inaccuracy of the error estimates for highly energetic muons is indeed due to the model underlying the reconstruction.

It was then studied how the error estimates and specifically their spread affect the sensitivity of the detector when its used as a neutrino telescope to find point-sources in the cosmos. This was done using toy-simulations of skymaps with a uniform distribution of background events and a number of injected point-source events. The position of the point-source events and the error estimates of all events followed a distribution that approximates the results from the track reconstructions. As in current realistic point-source analyses, the estimated errors were rescaled by a factor such that the median of all estimated errors is equal to the median of all true errors. Using these skymaps, the toy-sensitivity of the telescope in the search for point-sources was estimated. The entire computation was performed twice, once where the underlying track reconstructions used the track-only hypothesis and once where the full hypothesis including all cascades was used. The results showed that, in the case that 90% of the energy of the muon is lost in three cascades
and only 10% in a quasi-continuous track, the overall sensitivity improved by 18% when using the full hypothesis compared to the use of the simplified track-only reconstruction. The 3σ-detection flux was also calculated for both reconstructions, yielding an improvement of 19%. This improvement comes in addition to any improvement of the true error due to the more accurate hypothesis. Reconstruction toolkits like Millipede, which model the energy losses of the muon more accurately[6], therefore possess an additional gain for point-source searches up to 20%, which is not captured in pure comparisons of the true error. The exact gain might be lower in real reconstructions, since the energy losses aren’t known a-priori and have to be approximated to the data.

Finally, the sensitivity was also calculated for different cascade energy fractions between 0.1 (only 10% of the muon’s energy is lost in cascades) and 0.9 (90% of the muon’s energy is lost in cascades). This showed that the improvement in sensitivity decreases quickly with an improvement of only $\sim 3\%$ when half of the total energy is contained in the cascades. However, realistic muons often have energies above 10 TeV. The most highly energetic particles originating from galactic sources are expected to have energies beyond 1 PeV, even greater orders of magnitude are expected for extragalactic sources[2]. One such extreme muon event, which deposited an energy of $(2.6 \pm 0.3$) PeV in the detector, has been detected on June 11th 2014[13]. At those energies, statistical losses dominate over the continuous track[3]. This shows that the additional sensitivity gain due to correct error estimation in the order of 20% might be a realistic expectation when searching for galactic and extragalactic neutrino sources.

References


