ECAP-2023-011

Measuring the Diffuse Astrophysical Neutrino Flux with Public IceCube Data Using the PLEnuM Framework

Bachelorarbeit aus der Physik

Vorgelegt von

Anton Seißler

01 September 2023

Erlangen Centre for Astroparticle Physics
Friedrich-Alexander-Universität Erlangen-Nürnberg

Betreuer: Prof. Dr. Claudio Kopper

Contents

1 Introduction

Neutrino astronomy is an emerging field that aims to observe neutrinos produced by galactic and extra-galactic astronomical objects. Neutrinos are elementary particles that only interact via the weak force and gravity. This characteristic yields some interesting advantages over astronomy using photons. Since the cross-section for an interaction of neutrinos with matter is very small, they can penetrate matter that would be opaque to light. Furthermore, neutrinos are not affected by the electromagnetic force. This lets them travel through electromagnetic fields without being deflected. Observing astrophysical neutrinos can help to learn about the production and acceleration mechanisms of neutrinos and other astroparticles in astrophysical objects such as supernova remnants and active galactic nuclei.
The weakly interacting nature of neutrinos also creates practical difficulties. Since the interaction probability of neutrinos is so low, huge neutrino detectors are needed to make significant observations. One such neutrino detector is the IceCube Neutrino Observatory which was completed in December of 2010. This observatory uses the natural ice at the South Pole to measure the Cherenkov light that is produced by neutrino interactions. The neutrino measurements by IceCube could confirm an extraterrestrial neutrino flux in 2013 [1]. Since then further analysis of this neutrino flux has been conducted ([2],[3]). After these first successes, IceCube now aims to identify point sources emitting neutrinos.
The goal of this thesis is to measure the diffuse astrophysical neutrino flux with two public releases of IceCube data. The analysis uses a binned maximum likelihood fit and closely follows the methodology used by PLEnuM [4], a framework for diffuse neutrino flux analysis and point source search.
In Sec. 2 the production mechanisms of astrophysical neutrinos and the main background for this analysis, atmospheric neutrinos and muons, are discussed. Sec. 3 describes the composition of the IceCube Observatory and explains its detection principles. After that, the contents of the two data sets used in this analysis are shown in Sec. 4. In Sec. 5 the theoretical neutrino and muon flux models required for the analysis are introduced. Next, in Sec. 6 the binned maximum likelihood method that is applied in this analysis is explained. Finally, the results of the likelihood fits for the two data sets are presented in Sec. 7. Additionally, a combined fit utilizing both samples is described and the results are discussed.

2 Neutrino and Muon Sources

In this chapter, the sources of the diffuse astrophysical neutrino flux analyzed in this thesis are discussed. Furthermore, the main source of background information for this analysis, atmospheric air showers, is described.

2.1 Astrophysical Neutrinos

Astrophysical neutrinos with energies in the TeV to PeV range are produced from interactions of cosmic rays with surrounding matter [5].
Cosmic rays are charged particles like protons and alpha particles that are accelerated to high energies. One standard mechanism to accelerate cosmic rays to the highest energies is called Fermi acceleration. This acceleration occurs in shock waves produced by cosmic accelerators such as supernova remnants or active galactic nuclei. By repeatedly traversing the moving shock front of the shock waves, particles receive a net gain of energy until they escape. This acceleration mechanism produces an energy spectrum of high-energy particles with a power-law shape. [6]
When these particles interact with photons or matter near the acceleration site they produce pions and muons. These pions and muons then decay into electron and muon neutrinos (see (2.1), (2.2), (2.5) and (2.6)). [5]
Due to an effect called neutrino oscillation neutrinos can change their flavor when propagating through space [7]. After propagation, a neutrino can be measured as a different flavor than the flavor it was originally produced as. Even though the neutrino production mechanism described above only includes electron and muon flavors, neutrino oscillation leads to an expected ratio of electron, muon, and tau neutrinos reaching Earth of roughly 1:1:1 [5].

2.2 Atmospheric Neutrinos and Muons from Particle Showers

The main background data for this analysis comes from atmospheric particle showers. These particle showers are caused by cosmic rays with high energies that reach Earth. They consist of 90% protons, 9% alpha particles, and other ionized nuclei but also

electrons and uncharged particles [7]. When these high-energy particles hit Earth's atmosphere they interact strongly with the nuclei of the molecules in the atmosphere. This induces a cascade of subatomic particles, with mostly pions and kaons, called air showers [6]. The main decay channel for charged pions and kaons is the decay into muons and muon (anti-) neutrinos [8]:

$$\pi^- \to \mu^- + \bar{\nu}_\mu \tag{2.1}$$

$$\pi^+ \to \mu^+ + \nu_\mu \tag{2.2}$$

$$K^- \to \mu^- + \bar{\nu}_\mu \tag{2.3}$$

$$K^+ \to \mu^+ + \nu_\mu \tag{2.4}$$

This mechanism produces atmospheric muon neutrinos and muon antineutrinos that are detected by IceCube (see Sec. 3).
At energies below 2.5 GeV the decay length of the produced muons is shorter than their production height in the atmosphere of about 15 km [7]. Such muons can further decay into muon neutrinos, electron neutrinos and electrons:

$$\mu^- \to e^- + \bar{\nu}_e + \nu_\mu \tag{2.5}$$

$$\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu \tag{2.6}$$

This is another source of muon neutrinos as well as a production mechanism of electron neutrinos.
For muon energies above 2.5 GeV the decay probability of the muons decreases. The muons that do not decay reach the surface of the Earth as atmospheric muons. With enough energy, these muons can travel through the ice into the IceCube detector volume and be detected by IceCube as discussed in Sec. 3.

3 IceCube Detector

IceCube is a neutrino observatory located at the South Pole. One of its main goals is the measurement of astrophysical neutrinos in order to search for potential neutrino sources. Since the interaction cross-section of neutrinos is very small the IceCube Observatory uses a one cubic kilometer large volume of natural ice as the detector [9]. When a neutrino enters the ice there is a chance of an interaction with the ice atoms. Since neutrinos interact through the weak force either a charged $W^{\pm}$ boson or a neutral Z^0 boson is exchanged. These interactions are called charged-current and neutral-current interactions. During a charged-current interaction a lepton of the same flavor as the interacting neutrino is created. During a neutral-current interaction only a momentum transfer occurs. In both cases, the neutrino transfers a large amount of its momentum to its interaction partner. This leads to a hadronic cascade at the interaction vertex [10]. The neutrino interactions can be divided into two main morphologies, cascade-like events, and track-like events:

- **Cascade-like events:**
  Cascade-like events are caused by all neutral-current interactions. The particle in the ice that received the momentum from the neutrino causes a hadronic cascade. Secondly, charged-current interactions with electron neutrinos and tau neutrinos also produce cascade-like events. The produced lepton quickly loses its energy in the ice and produces an electromagnetic cascade. [10]

- **Track-like events:**
  Charged-current interactions of muon neutrinos produce high-energy muons. These muons propagate through the ice in a straight line, a track. These muon tracks can also be produced by atmospheric muons traveling through the detector.
  More rarely tau neutrinos can also lead to track-like events. When a tau neutrino interacts through charged-current a tau particle is produced. These tau particles have an extremely short lifetime of approximately 10^{-15} s [8] and decay quickly. This decay can lead to secondary muons with a branching ratio of about 17% [8] which produce the track-like morphology. [10]

The hadronic cascades and muon tracks are detected using an electromagnetic effect called the Cherenkov effect. When the charged particles created in the neutrino interactions (hadronic cascades and muons) travel through the ice, they polarize the ice

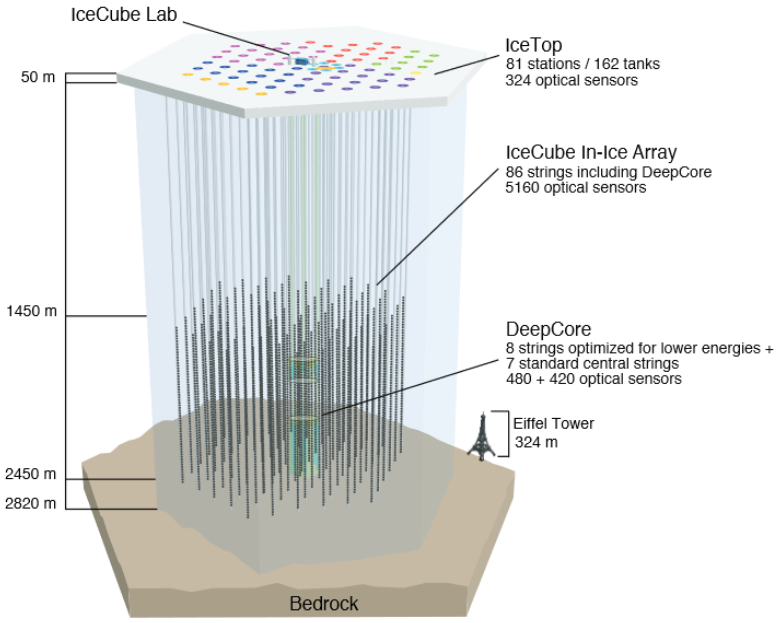


Figure 3.1: Illustration showing the IceCube Observatory. The 86 strings holding the DOMs form the IceCube In-Ice Array. IceTop, an array of 162 ice tanks, is located at the top of the detector. It is used to detect air showers [9]. (Figure taken from [9])

molecules, which leads to an emission of electromagnetic waves. Normally these waves annihilate each other through interference. However, when the particle travels at a speed faster than the emitted light, a wavefront emerges, called Cherenkov light. This is possible since the speed of light in a medium like ice is only a fraction of the speed of light in vacuum. [10]
To detect the light produced by the secondary particles IceCube uses light sensors called Digital Optical Modules (DOMs). DOMs are spherical sensors made from a glass housing and a Photo Multiplier Tube (PMT). The PMT detects the emitted light by converting the incoming photons into a charge (often measured in photoelectrons) via the photoelectric effect. The DOMs are arranged on 86 vertical strings that are deployed between 1450 m and 2450 m depth in the ice. Each string holds 60 DOMs. The strings are distributed evenly in the detector volume. Fig. 3.1 shows a sketch of the IceCube Observatory. [9]
Using the light measurements by the DOMs during an event, the properties of the interacting neutrino are reconstructed. By using different criteria such as the reconstruction quality or the event morphology events are selected from the measurements to form a data set.

4 Data Sets

The analysis in this thesis is done using two samples of public IceCube data, the "HESE 7.5 year data release" [2] and the "All-sky point-source IceCube data: years 2008-2018" [11]. The contents and event selection of these data sets are discussed in this chapter.

4.1 HESE 7.5 Year Data Release

The first data set used in this analysis is the IceCube high-energy starting event sample (HESE). The sample consists of a selection of events from IceCube data of seven and a half years. The goal of the event selection is to produce a sample of astrophysical neutrino events and reduce the amount of background events included. This is done by aiming to select only events starting within the detector volume (starting events) with high deposited energy.
As described in Sec. 2.2 a large amount of the background flux is atmospheric muons. Muons produced in air showers in the southern atmosphere have to travel through the ice to reach the detector volume. The tracks of these muons start outside the detector. By only including events with interaction vertices inside the detector, atmospheric muons can be filtered out. In order to do so, the HESE sample defines a veto region at the edges of the detector volume and at the layer of ice with the highest concentration of dust. The DOMs in this region are called veto DOMs. Fig. 4.1 shows the selected veto region and veto DOMs in the detector. If the total observed charge in the veto region at the starting time of an event exceeds a set threshold, the event is likely not a starting event and is rejected from the sample. Using this selection method a majority of atmospheric muon events can be removed from the data set. [2]
Furthermore, only events with a total observed charge over 6000 PE (photoelectrons) are included. This ensures that only events with high energies are selected. Since the astrophysical neutrino flux dominates at higher energies [5], this decreases the amount of atmospheric neutrino events in the sample. [2]
After the selection, a total of 102 events remain. The relevant information for the diffuse flux analysis are the reconstructed deposited energy and zenith angle of the events. The HESE data release includes the values of those reconstructed quantities for every measured event. For this analysis, the experimental data is represented as a two-dimensional histogram as shown in Fig. 4.2.

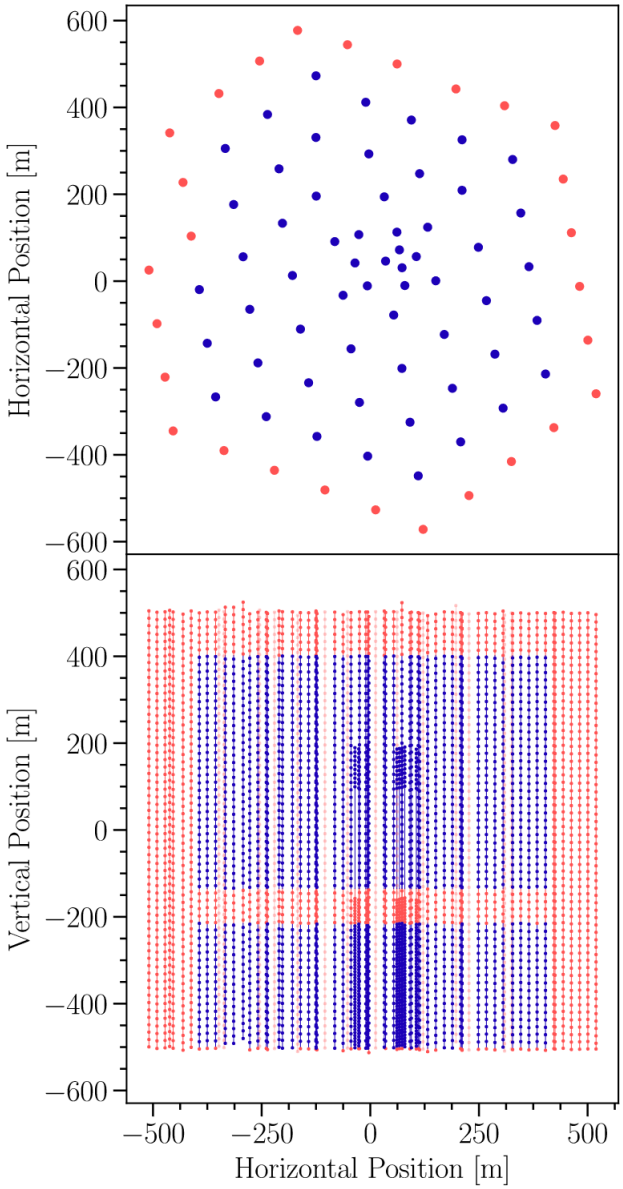


Figure 4.1: The veto region used in the event selection of the HESE sample. Veto DOMs and strings are shown in red and the non veto ones are shown in blue. The upper figure shows a top view of the detector while the lower figure shows a side view. The outer layer of stings, the top 90 m and lower 10 m of the detector, and a region with a high concentration of dust are defined as the veto region. (Figure taken from [2])

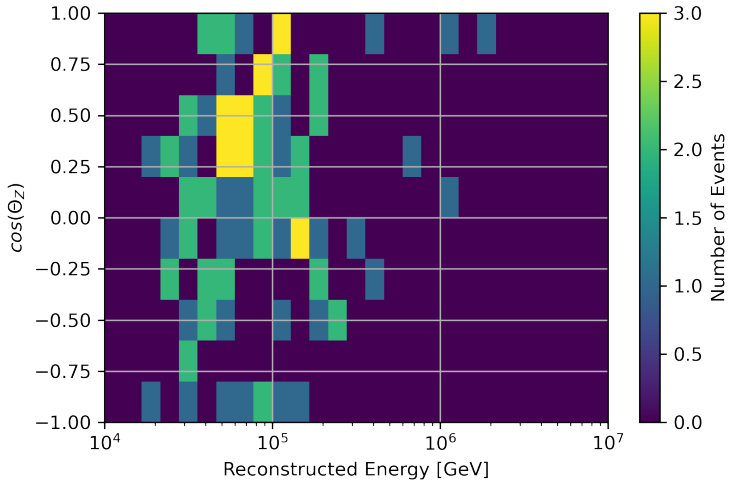


Figure 4.2: The distribution of all 102 HESE events as a function of reconstructed Energy and zenith angle θ_Z. The data is represented as a two-dimensional histogram. The x-axis shows the reconstructed energy on a logarithmic scale while the y-axis shows the cosine of the reconstructed zenith angle. Events with $\cos(\theta_Z) > 0$ are down-going events coming from the southern sky and events with $\cos(\theta_Z) < 0$ are up-going events coming from the northern sky.

4.2 All-Sky Point-Source IceCube Data: Years 2008-2018

The second data set is the sample used in IceCube's time-integrated point source search with 10 years of data [11]. This sample includes muon track events from 2008 to 2018. The data release consists of data sets IC40, IC59, IC79, and IC86_I to IC86_VII. Those sets contain data obtained while the deployment of the strings in the detector was still in progress. The samples include data observed with 40, 59, 79, and 86 strings of DOMs respectively. In this analysis only data measured after all 86 strings were deployed in 2010 is used, to limit the variance in detector response.
For the data sets IC86_II to IC86_VII a Boosted Decision Tree (BDT) trained on simulated data is used for event selection. The BDT aims to reduce the contamination of atmospheric muons and cascades in the data. For the Northern Sky, the BDT categorizes events as neutrino track-like events, muon tracks produced by atmospheric muons, and cascade-like events. Only neutrino-induced events are selected. Data set IC86_I still uses a likelihood-based selection process [12]. Using this selection only about 0.1% of muon tracks remain in the sample. [11]
For this analysis only events from the Northern sky with declination $\delta > -5°$ are used. This reduces the atmospheric muon contamination significantly since muons are stopped by the earth when traveling from the Northern sky to the detector [3]. The selected sam-

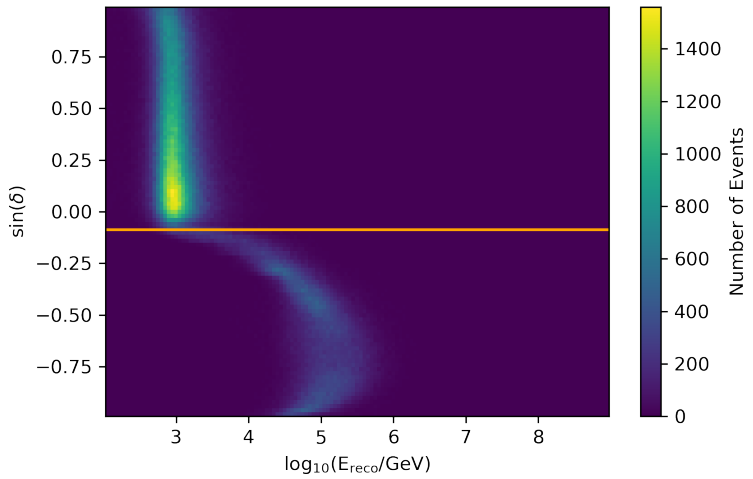


Figure 4.3: Distribution of events from the muon track data-set as a function of reconstructed energy and declination angle. The horizontal line shows the separation between the Northern and Southern sky at $\delta = -5°$. Above the separating line, the observed events from the Northern sky are shown. They consist of mostly astrophysical and atmospheric neutrino events. Towards the horizon at $\delta = 0°$ the intensity increases, since the neutrinos have to travel through a shorter path through the earth. This decreases the probability of interaction before reaching the detector region. For the Southern sky harsher cuts in the event selection were needed which leaves mostly events with energies above 10 TeV [11].

ples include a total of (761162) events. Fig. 4.3 shows the distribution of those events over the reconstructed energy and the declination angle.
Compared to the HESE data set, the muon track sample contains a much higher event count, which could lead to more significant analysis results. However, the restriction to starting events of the HESE data set allows for an easier analysis of the whole sky, while in this analysis only the Northern sky events from the muon track data set are included.

5 Neutrino and Muon Flux Models

The observed neutrino flux consists of three parts, the atmospheric neutrino flux, the atmospheric muon flux, and the astrophysical neutrino flux. This chapter characterizes the models used to describe those flux components.

5.1 Astrophysical Neutrino Flux

For the astrophysical flux, a single power-law is assumed:

$$\frac{d\Phi_\nu}{dE_\nu}_{\text{astro}} = \Phi_{\text{astro}} \left(E_\nu/10^5\,\text{GeV}\right)^{-\gamma_{\text{astro}}} 10^{-18}\,\text{GeV}^-1\text{cm}^{-2}\text{s}^{-1}\text{sr}^{-1} \tag{5.1}$$

Here $\frac{d\Phi_\nu}{dE_\nu}_{\text{astro}}$ is the all-flavor astrophysical neutrino flux, E_ν the neutrino energy, Φ_{astro} the normalization parameter and γ_{astro} the spectral index. The normalization and spectral index are the parameters that will be determined by the analysis. Fig. 5.1 shows an example of the astrophysical flux using best-fit values of the HESE data release [2] for Φ_{astro} and γ_{astro} [2].

5.2 Atmospheric Neutrino Flux

A large portion of events measured with IceCube comes from atmospheric air showers induced by cosmic rays. These air-showers produce muons and neutrinos which are detected by IceCube. To analyze the astrophysical flux a model for the atmospheric background is needed. The matrix cascade equations framework MCEq [13] calculates the muon and neutrino flux produced by air showers. The particles in an air shower are characterized by cascade equations. These equations describe the development of the particle fluxes in a particle cascade traveling through the atmosphere [7]. MCEq converts the cascade equations into a matrix form and solves them numerically. [13]
This analysis uses the atmospheric flux model produced by MCEq. The MCEq muon and neutrino fluxes are shown in Fig. 5.2.

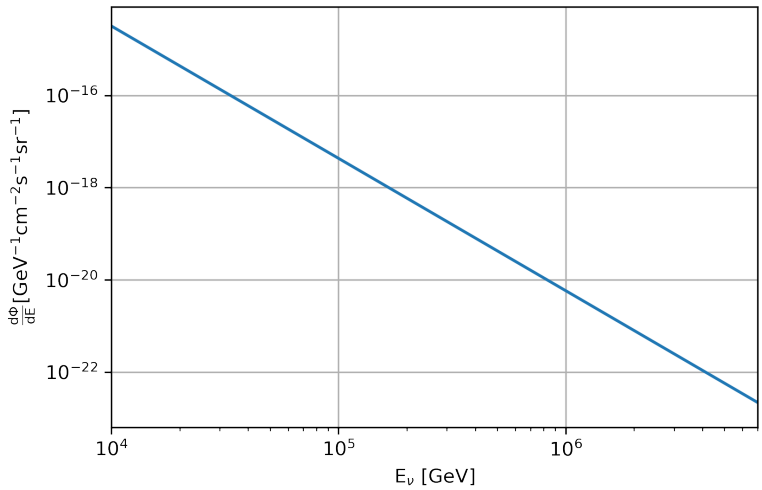
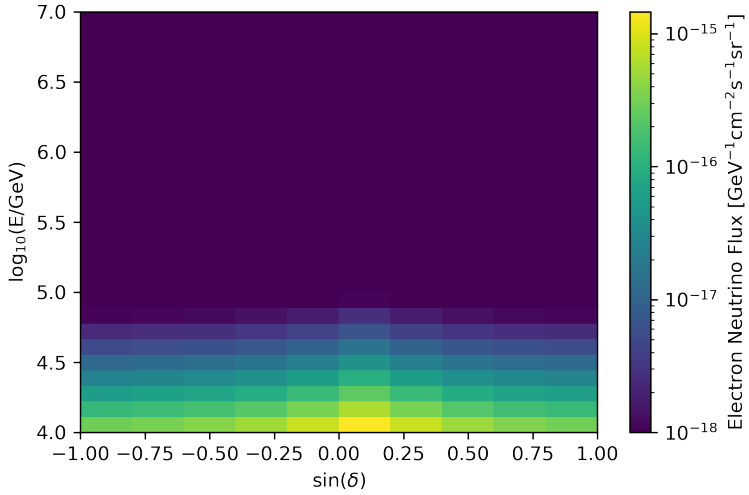
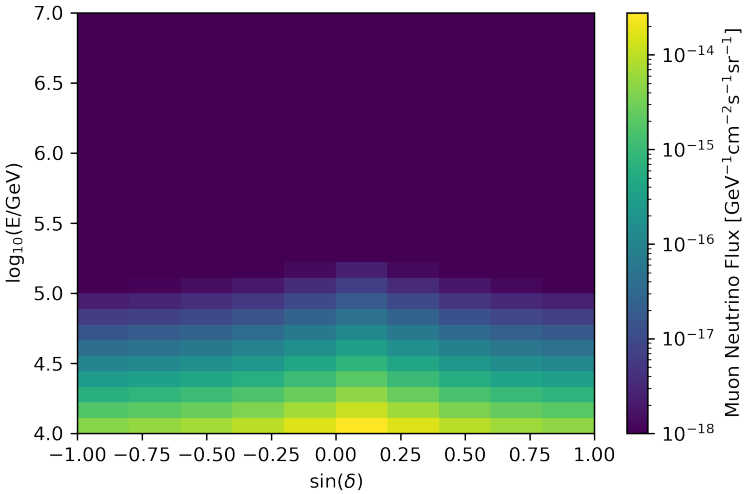


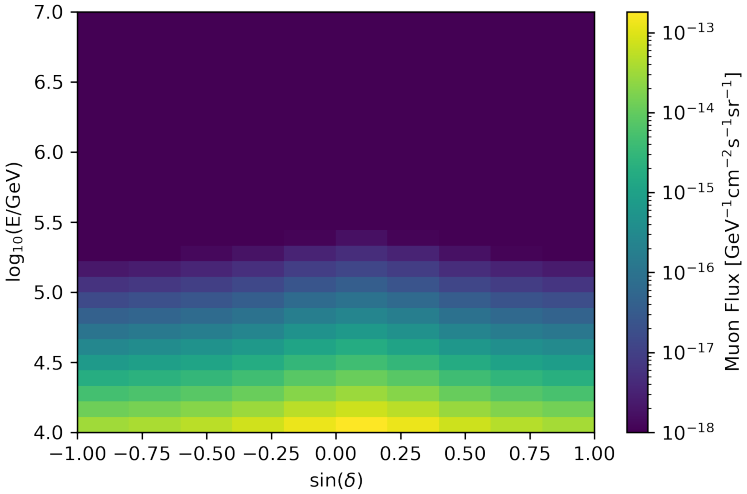
Figure 5.1: Astrophysical neutrino flux assuming a power-law distribution and best-fit parameters from the HESE data release Φ_{astro}=6.37 and $\gamma_{\text{astro}} = 2.87$ [2].



(a) MCEq atmospheric electron neutrino flux.



(b) MCEq atmospheric muon neutrino flux.



(c) MCEq atmospheric muon flux.

Figure 5.2: Atmospheric flux models as calculated by MCEq [13]. Figures Fig. 5.2a through Fig. 5.2c show the theoretical atmospheric electron neutrino, muon neutrino, and muon flux as a function of energy and declination.

6 Binned Maximum Likelihood Method

6.1 Likelihood Function and Fit

A binned maximum likelihood method is used to fit the models described in Sec. 5. The data sets and flux models are represented as two-dimensional histograms. The likelihood to measure the data k for a hypothesis μ in a single bin is given by the Poisson distribution

$$\mathcal{L}_{bin}(k|\mu) = \frac{\mu_{bin}^{k_{bin}}}{k_{bin}!} \exp(-\mu_{bin}). \tag{6.1}$$

The likelihood for the entire data set is obtained by multiplying the likelihoods of each bin:

$$\mathcal{L}(k|\mu) = \prod_{bin} \frac{\mu_{bin}^{k_{bin}}}{k_{bin}!} \exp(-\mu_{bin}) \tag{6.2}$$

In this analysis, the data k_{bin} are the counts of the bins in the 2D-histograms of the data sets shown in Fig. 4.2 and Fig. 4.3. The hypothesis $\mu_{bin}(\vec{\Theta})$ are the counts of the bin in the 2D-histograms of the expected event counts, which depend on the free parameters $\vec{\Theta}$. The calculation of these histograms is described in Sec. 7.1. The best-fit parameters for the model and data are found by maximizing the likelihood function. For better numerical stability and convenience the negative log-likelihood, $-2\log\mathcal{L}(k|\mu)$, is minimized instead. This results in the same best-fit values, since the logarithm of the likelihood function has the same maxima as the function itself.

6.2 Hypothesis Test

A hypothesis test is conducted to determine if the observed data significantly supports the existence of a diffuse astrophysical neutrino flux. The null hypothesis H_0 states that only an atmospheric background flux exists. The signal hypothesis H_1 assumes a combination of atmospheric background flux and astrophysical neutrino flux. For both hypotheses, the best-fit parameters are determined and the corresponding likelihood

values are calculated. The test statistic TS is defined as the ratio of the likelihoods:

$$TS = -2\log\left(\frac{\mathcal{L}(k|\mu(H_0))}{\mathcal{L}(k|\mu(H_1))}\right) \tag{6.3}$$

Wilks' Theorem states that the test statistic approaches a chi-squared distribution χ^2_l for large sample sizes under the assumption of the null hypothesis. l is equal to the difference in degrees of freedom of the null hypothesis H_0 and the signal hypothesis H_1. The p-value is defined as the probability of observing data with a calculated test statistic TS if the null hypothesis is true. It is used to determine whether or not the null hypothesis should be rejected. The p-value is calculated by integrating over the corresponding chi-squared distribution:

$$\mathrm{p} = \int_{TS}^{\infty} dx \chi^2_l \tag{6.4}$$

To express the significance of the hypothesis test the p-value can be determined using the same integral over the Gaussian distribution:

$$N(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{6.5}$$

$$\int_a^{\infty} dx N(x, 0, 1) = \mathrm{p} \tag{6.6}$$

By calculating the inverse of the integral the significance in units of Gaussian standard deviations is found.

7 Fitting Process and Results

In this chapter, the results of the likelihood fits and the hypothesis tests for the two data sets are described. First, each sample is analyzed separately. The results of a combined fit of the two data sets are presented in Sec. 7.4.

7.1 Calculating the Expected Events

The expected number of measured events for a given flux has to be calculated to compare the experimental data with the flux models described in Sec. 5. The number of observed events depends on the response of the detector to incoming particles. The detector response is determined by Monte Carlo simulations. This is done by generating Monte Carlo events, simulating the propagation and interaction of the particles and emitted light in the IceCube detector. This information can then be represented as an effective area A_{eff}. The effective area is a measure for the sensitivity of a detector to particles with a certain energy and direction. Geometrically this means, in order to detect all particles with energies E from direction θ, ϕ the detector would have to be the size of the corresponding effective area.
The expected number of events N_ν is calculated by integrating over the product of the effective area and the assumed flux [11]:

$$N_\nu = \int_{T_{\text{live}}} dt \int_{\Delta\Omega} d\Omega \int_{\Delta E} dE A_{\text{eff}}(E, \theta, \phi) \frac{d\Phi}{dE}(E, \theta, \phi) \tag{7.1}$$

For the analysis, the expected number for each bin i, k of the 2D-histograms has to be calculated. By assuming a constant effective area and flux in each bin we can approximate as follows:

$$N_{\nu_{bin}} = \int_{T_{\text{live}}} dt \int_{\Delta\Omega} d\Omega \int_{\Delta E} dE A_{\text{eff}_{bin}} \frac{d\Phi}{dE}_{bin} \tag{7.2}$$

With $d\Omega = \sin\theta d\theta d\phi$ this can be calculated as:

$$\begin{aligned} N_{\nu_{bin}} &= \int_{T_{\text{live}}} dt \int_{\Delta\theta} d\theta \sin\theta \int_{2\pi} d\phi \int_{\Delta E} dE A_{\text{eff}_{bin}} \frac{d\Phi}{dE}_{bin} \\ &= T_{\text{live}} \cdot 2\pi \cdot \Delta \sin\theta \cdot \Delta E \cdot A_{\text{eff}_{bin}} \cdot \frac{d\Phi}{dE}_{bin} \end{aligned} \tag{7.3}$$

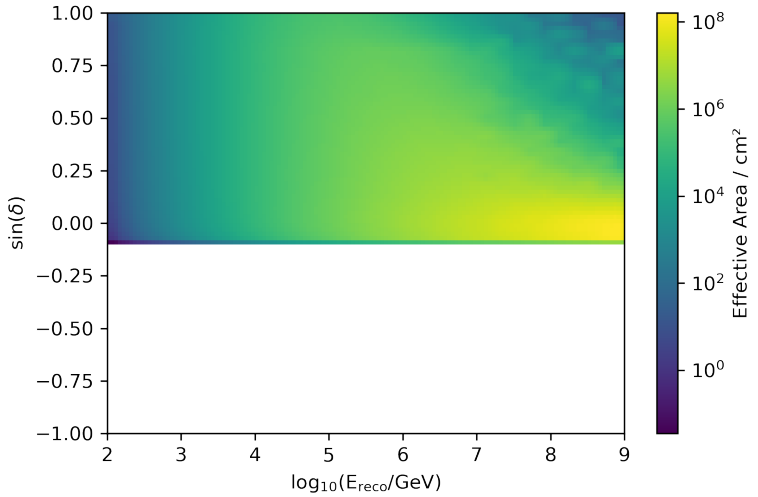


Figure 7.1: The effective area of the IceCube detector for the Northern sky from the muon track data set [11] as a function of energy and declination. The effective area generally increases with the particle energy, since higher energy particles have a higher interaction probability and are more likely to be detected.

The muon track sample data release includes the effective area determined by the simulation. Fig. 7.1 shows the effective area as a function of energy and zenith angle. The expected number of measured events are calculated using (7.3), the 2D-histograms for the effective area, the astrophysical and atmospheric flux models as well as the bin widths $\Delta \sin\theta$ and ΔE and the detector live-time T_{live}. Fig. 7.2 shows the expected events for the astrophysical and atmospheric flux components.

The HESE data release does not include an effective area but instead comes with the data of the Monte Carlo simulation. Each Monte Carlo event has a value for its particle energy and zenith angle, the particle type and a Monte Carlo weight called the OneWeight. The OneWeight is a measure for the interaction and detection probability of the simulated particle and contains the same information as the effective area. To obtain the weight of an event for a given neutrino flux model, the OneWeight must be multiplied by the flux. By summing all weighted Monte Carlo events in a bin the expected event rate is obtained. The expected number of events is calculated as:

$$N_{\nu_{bin}} = T_{\text{live}} \cdot \sum_i OneWeight_i \cdot \frac{d\Phi}{dE}(E_i, \theta_i, \phi_i) \tag{7.4}$$

When calculating the event rate for a specific particle flux only Monte Carlo events with the matching particle type must be included. For example, the atmospheric muon flux is

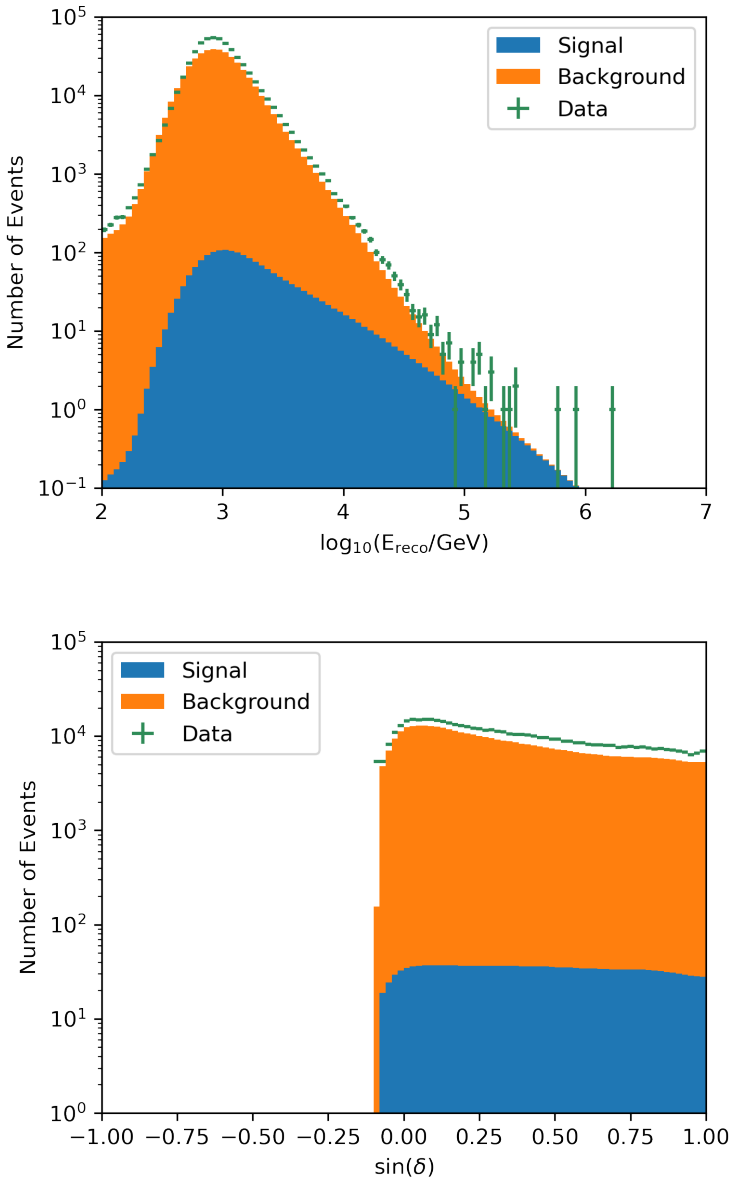


Figure 7.2: The expected number of events for the muon track data set for the Northern sky. The atmospheric flux (orange) and astrophysical flux (blue) are shown as stacked histograms. For this histogram best-fit values of $\Phi_{\text{astro}} = 1.44$ and $\gamma_{\text{astro}} = 2.37$ from [3] are used for the astrophysical component. The atmospheric normalization is set to $\Phi_{\text{atmo}} = 1$. The upper figure shows the energy spectrum while the lower one shows the distribution over the declination angle. The expected amount of events decreases with higher energies, which matches the flux models discussed in Sec. 5. At lower energies, the number of events falls since the corresponding effective area is much smaller (see Fig. 7.1).

calculated with the muon Monte Carlo events. Fig. 7.3 shows the expected atmospheric and astrophysical fluxes for the HESE data sample.

7.2 HESE Data Set Analysis

7.2.1 Likelihood Fit

First, the results of the diffuse flux analysis using the IceCube high-energy starting event sample described in Sec. 4 are discussed. For the analysis, only events with reconstructed energies over 60 TeV are used. This further decreases the fraction of muons in the sample, since most muon events are expected at energies below this cut (see Fig. 7.3) [2]. After the cut 60 events of the original 102 remain.
For the likelihood function, the models for the astrophysical and atmospheric flux from Sec. 5 are used. The astrophysical flux $\frac{d\Phi_\nu}{dE_\nu}_{\text{astro}}$ depends on the two fit parameters Φ_{astro} and γ_{astro}. The atmospheric flux is obtained by adding the three components of the MCEq models from Sec. 5.2:

$$\frac{d\Phi}{dE}_{\text{atmo}} = \frac{d\Phi_{\nu_e}}{dE_{\nu_e}} + \frac{d\Phi_{\nu_\mu}}{dE_{\nu_\mu}} + \frac{d\Phi_\mu}{dE_\mu} \tag{7.5}$$

To also fit the atmospheric background to the data a normalization parameter for the neutrino flux Φ_{atmo} is introduced. This allows the background flux to be scaled to better match the data. Since the atmospheric muon flux only contributes to a very small part of the background, the normalization for the muon component is set to the best-fit value from the HESE analysis of 1.19 [2]. This results in the atmospheric background flux model

$$\frac{d\Phi}{dE}_{\text{atmo}}(\Phi_{\text{atmo}}) = \Phi_{\text{atmo}} \cdot \frac{d\Phi_{\nu_e}}{dE_{\nu_e}} + \Phi_{\text{atmo}} \cdot \frac{d\Phi_{\nu_\mu}}{dE_{\nu_\mu}} + 1.19 \cdot \frac{d\Phi_\mu}{dE_\mu}. \tag{7.6}$$

From those models, the expected number of events per bin are calculated as described in Sec. 7.1. Adding the expected events from the atmospheric and the astrophysical component leads to the model $\mu(\Phi_{\text{atmo}}, \Phi_{\text{astro}}, \gamma_{\text{astro}})$ needed for the likelihood fit. Together with the 2D-histogram of the data set from Sec. 4 the likelihood function is formed with (6.2). The fit parameters are Φ_{atmo}, Φ_{astro} and γ_{astro}.
When fitting this model to the data the found minimum of the log-likelihood function was very wide in the atmospheric norm dimension. This would result in very uncertain fit results for Φ_{atmo}. To combat this, a prior is introduced. The prior follows a Gaussian distribution with mean $\mu_{\text{Gauss}} = 1$ and standard deviation $\sigma = 0.4$. These values are taken from the prior used in the HESE analysis [2]. The prior is implemented by multiplying the Gaussian distribution by the likelihood function:

$$\mathcal{L}(k|\mu) = \exp\left(-\frac{(\Phi_{\text{atmo}} - \mu_{\text{Gauss}})^2}{2\sigma^2}\right) \cdot \prod_{bin} \frac{\mu_{bin}^{k_{bin}}}{k_{bin}!} \exp(-\mu_{bin}) \tag{7.7}$$

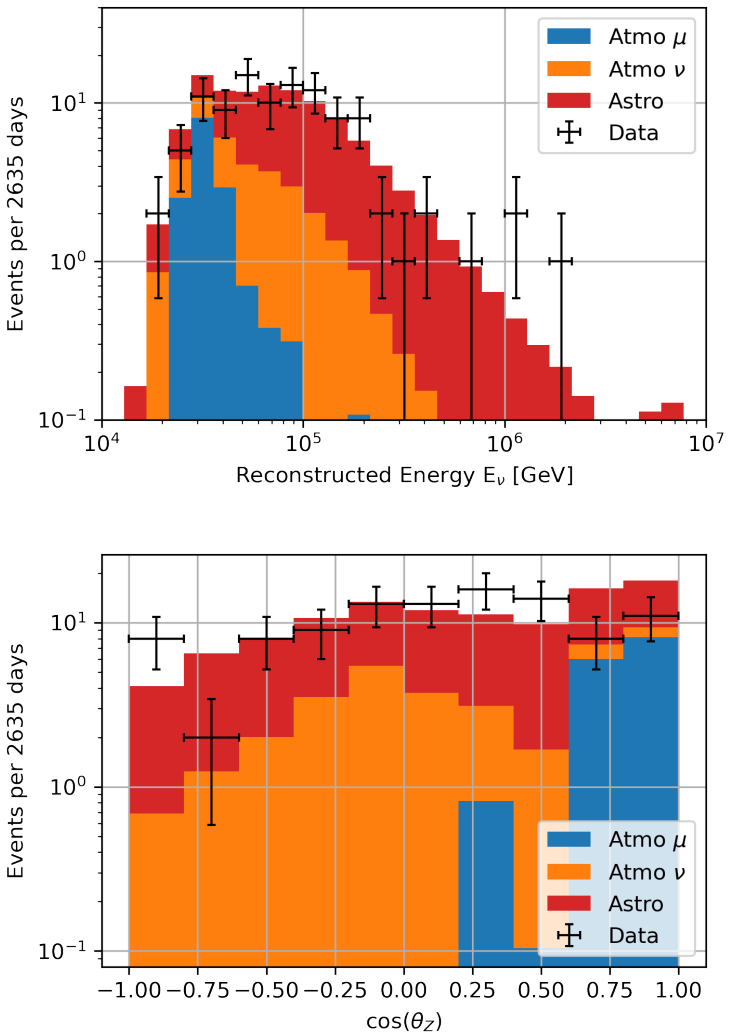


Figure 7.3: The expected number of events as calculated from the HESE Monte Carlo Simulations represented as stacked histograms. The first figure shows the distribution over the reconstructed energy and the second one shows the distribution over the reconstructed zenith angle. Atmospheric muon events are represented in blue, atmospheric neutrino events in orange, and the astrophysical flux in red. The experimental data is shown in black. For the calculation of the expected number of events, best-fit values from the HESE data release have been used. The expected event rate decreases with higher energies as expected from the flux models in Sec. 5. The distribution is mostly flat as a function of the zenith angle. Atmospheric muons are only expected in the down-going direction ($\cos(\theta_Z) > 0$) since muons from the up-going direction ($\cos(\theta_Z) < 0$) are absorbed by the earth. Neutrinos from the Northern sky have a chance to interact while traveling through Earth. For this reason, the neutrino fluxes in the up-going direction are slightly smaller than in the down-going direction.

Φ_{atmo}	0.91
Φ_{astro}	6.34
γ_{astro}	3.22

Table 7.1: The best-fit values for the HESE data set as determined by the likelihood fit.

By minimizing the negative log-likelihood the best fit parameters are found. The results are listed in Tab. 7.1. Fig. 7.4 shows the theoretical astrophysical and atmospheric fluxes following the newfound best fit values.

Next, some 2D likelihood scans are conducted to analyze the uncertainty of the best-fit results. First the likelihood function as a function of the astrophysical norm Φ_{astro} and the spectral index γ_{astro} is evaluated. This is done by setting a grid of fixed values for the two parameters and minimizing the likelihood function with respect to the third parameter, the atmospheric normalization Φ_{astro}. For every point $(\Phi_{\text{astro}i}, \gamma_{\text{astro}_j})$ on the grid the corresponding likelihood is calculated. Next, the likelihood ratio of each point and the best fit are determined:

$$-2\log\left(\frac{\mathcal{L}(\Phi_{\text{astro}i}, \gamma_{\text{astro}_j})}{\log\mathcal{L}(\text{bestfit})}\right) = -2(\log\mathcal{L}(\Phi_{\text{astro}i}, \gamma_{\text{astro}_j}) - \log\mathcal{L}(\text{bestfit})) = -2\Delta\log\mathcal{L} \tag{7.8}$$

Following Wilks' Theorem (Sec. 6) the 68% and 95% confidence regions are calculated. Fig. 7.5 shows the resulting likelihood landscape and the confidence regions in comparison to the analysis done in [2].
Both confidence regions show a similar shape and size when compared to the HESE analysis. The best-fit values of both analyses lie inside the 95% confidence regions of the other fit and the 68% confidence regions show a slight overlap. This suggests a compatibility of the two fit results. The main difference is in the determined spectral index. While this analysis calculated a value of 3.22 the HESE analysis results in a value of 2.87. This difference can be explained by the exclusion of systematic uncertainties in this analysis. The HESE analysis includes parameters for the detector systematics such as the DOM efficiency. Furthermore, a more complex atmospheric neutrino flux model is used that is split into a conventional and prompt component and includes systematic parameters like the Kaon-Pion ratio. [2]
The same likelihood scan is repeated for the two normalization parameters of the signal and background flux. The results are shown in Fig. 7.5. The shape of the confidence contours shows a slight degeneracy in the atmospheric and astrophysical normalization parameters. At the high fit value for the spectral index of γ_{astro} the atmospheric and astrophysical fluxes have similar shapes (see Fig. 7.4). This means that lower values for the atmospheric normalization are counteracted by a higher astrophysical normalization in the likelihood function and vice versa.

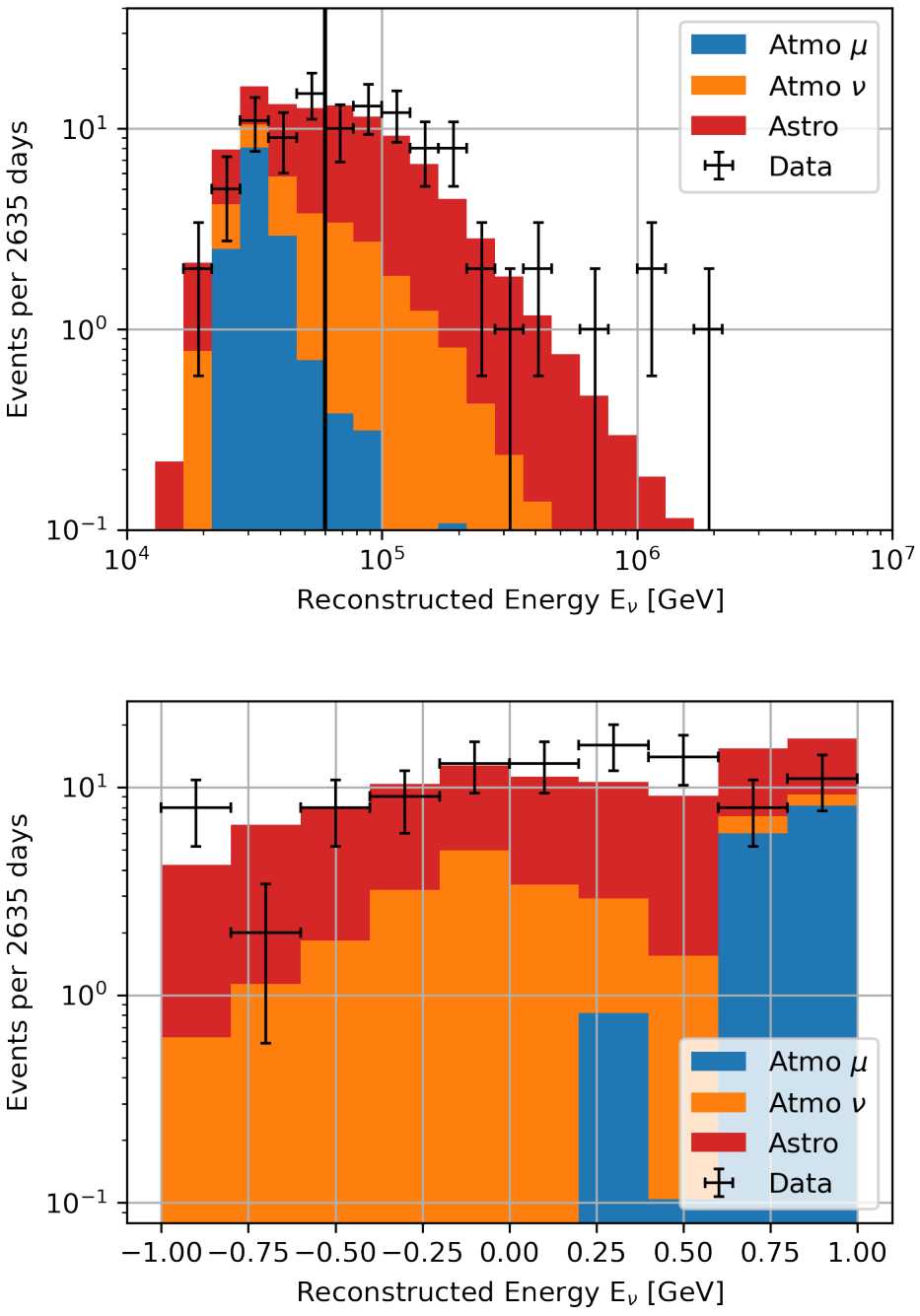


Figure 7.4: The best fit of the atmospheric and astrophysical fluxes compared to the HESE data set. The upper figure shows the fluxes energy spectrum. The vertical line shows the 60 TeV cutoff. Only events over this energy are included in the fitting process. The lower figure shows the distribution over the reconstructed zenith angle.

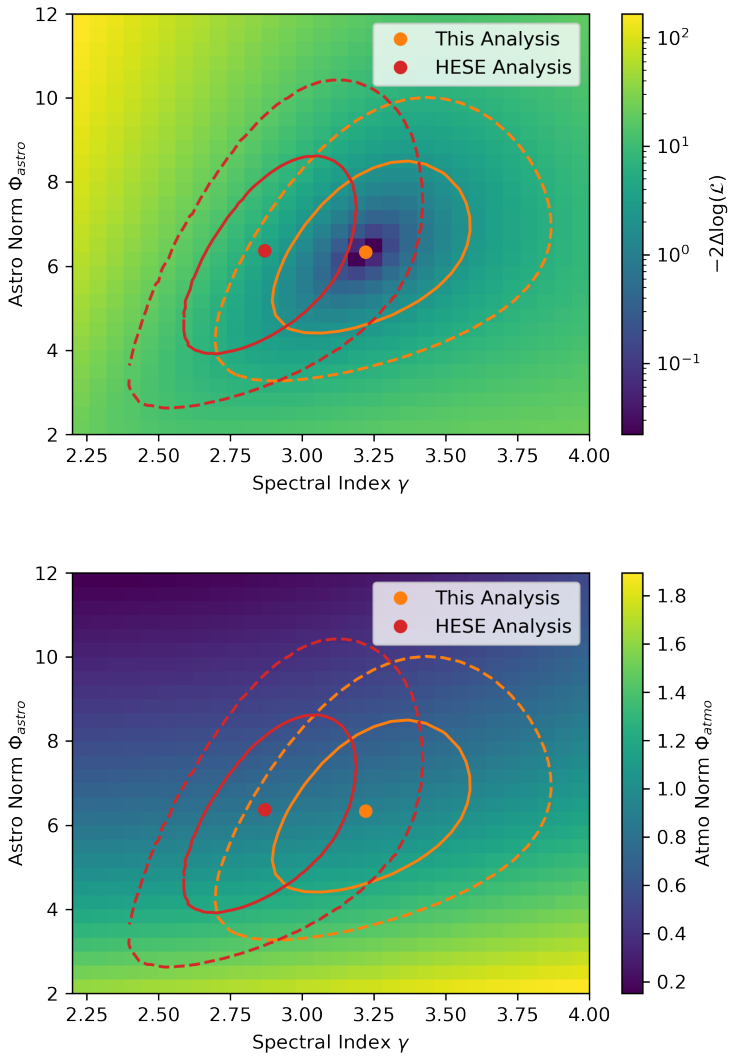


Figure 7.5: 2D likelihood scan over the astrophysical normalization Φ_{astro} and spectral index γ_{astro}. The first plot shows the likelihood ratio for each grid point, while the lower plot shows the corresponding best-fit value for the atmospheric normalization Φ_{atmo}. Orange shows the best fit and confidence contours obtained in this analysis. The solid lines enclose the 68% confidence region and the dotted lines the 95% confidence region. For comparison the fit values and contours from the HESE analysis [2] are displayed in red.

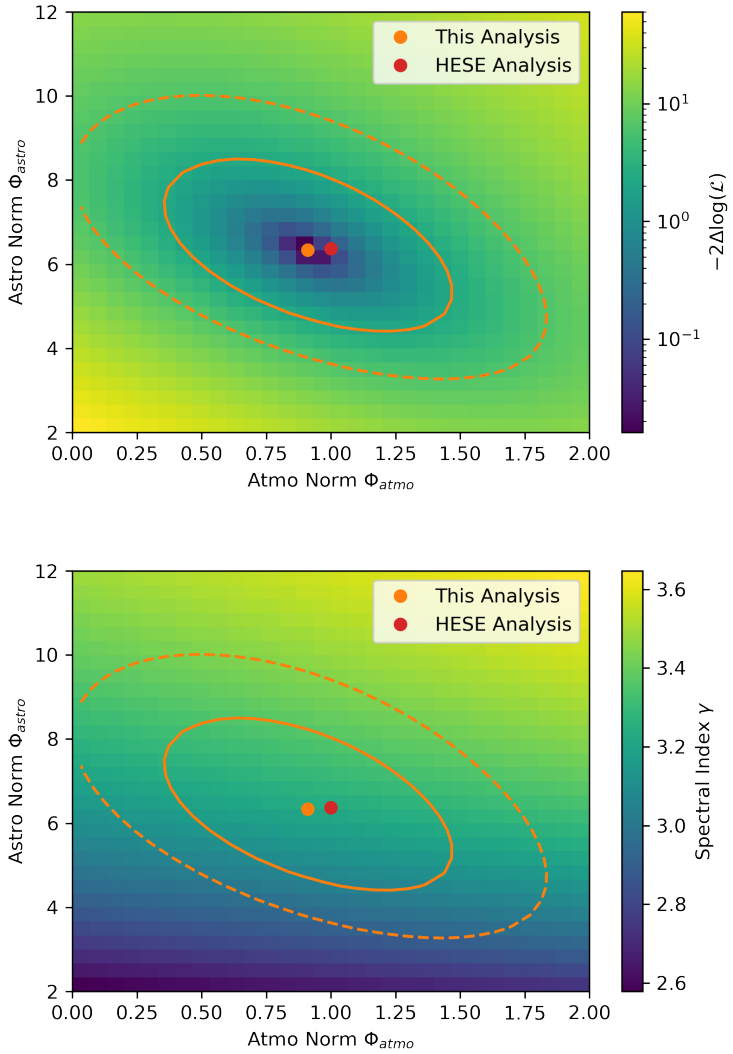


Figure 7.6: 2D likelihood scan over the two normalization parameters Φ_{astro} and Φ_{atmo}. The first plot shows the likelihood ratio, while the lower plot shows the corresponding best-fit value for the spectral index γ_{astro}. The solid lines show the 68% confidence region and the dotted lines the 95% confidence region.

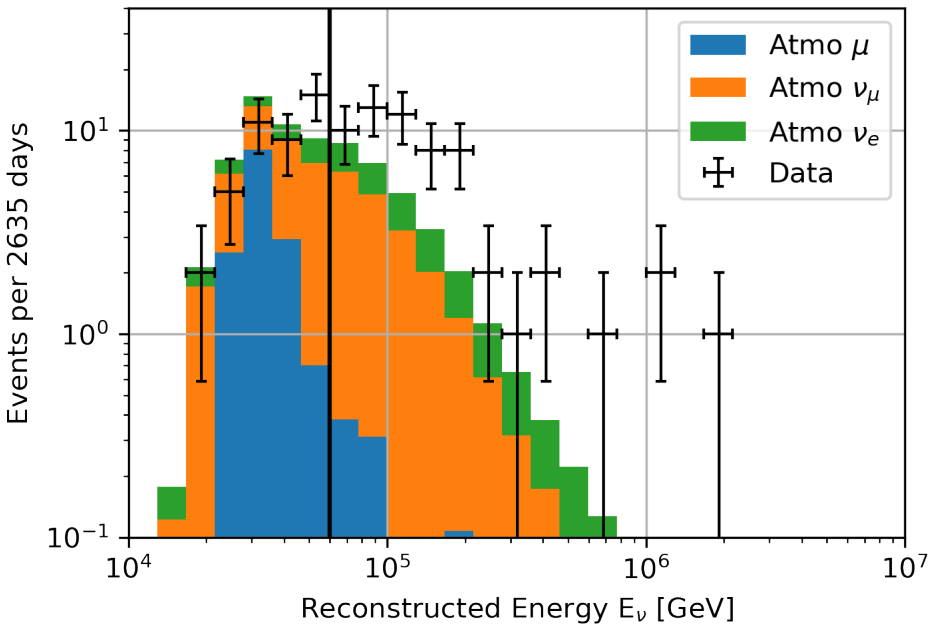


Figure 7.7: Comparison of the HESE data set and the fitted background hypothesis.

7.2.2 Hypothesis Test

For the hypothesis test described in Sec. 6.2 the fit has to be repeated for the null hypothesis, which assumes that only an atmospheric neutrino flux exists. This is done by setting the astrophysical normalization to zero and then minimizing the negative log-likelihood. The minimization results in a best-fit value for the atmospheric normalization of $\Phi_{\text{atmo}} = 2.50$. Fig. 7.7 shows the comparison of the fitted background hypothesis to the data. For the test statistic defined in (6.3) the likelihood ratio of the null hypothesis and the signal hypothesis is calculated. Using (6.4), a p-value of 2.64×10^{-10} is determined. As described in Sec. 6.2 this p-value is represented in units of Gaussian standard deviations. The resulting significance is 6.32 σ. This means the null hypothesis can be rejected with a confidence of 6.32 σ.

Φ_{atmo}	1.24
Φ_{astro}	1.53
γ_{astro}	2.27

Table 7.2: The best-fit values for the muon track data set as determined by the likelihood fit.

7.3 Muon Track Data Set Analysis

7.3.1 Likelihood Fit

The fitting procedure is repeated for the muon track data set. For this sample no priors are used. Since the background data dominates in the 1 TeV range (see Fig. 7.8) a energy threshold of $10^{3.5}$ GeV is used. To focus more on the astrophysical neutrino flux, only events above the cut are included in the fitting process. Minimizing the log-likelihood function results in the best-fit parameters shown in Tab. 7.2. Compared to the results from the HESE fit, the muon track sample results in a harder spectral index for the astrophysical flux. As expected the astrophysical normalization is much smaller than the HESE fit result, since the HESE sample contains all flavors of neutrinos while the muon track data set aims to only include the muon flavor. The neutrino flux models following the best-fit parameters are shown in comparison to the data in Fig. 7.8.
The fit result for the astrophysical normalization $\Phi_{\text{astro}} = 1.53$ lies inside the uncertainties of the result of an IceCube muon track data set analysis by the IceCube Collaboration of $\Phi_{\text{astro}} = 1.44^{+0.25}_{-0.26}$ [3]. The fitted spectral index $\gamma_{\text{astro}} = 2.27$ lies just barely outside the IceCube Collaboration results of $\gamma_{\text{astro}} = 2.37^{+0.09}_{-0.09}$ [3]. This shows that the fit conducted in this thesis is a good reproduction of official fit results.
The 2D likelihood scans explained in Sec. 7.2.1 are repeated for this fit. The resulting likelihood landscapes and confidence regions are shown in Fig. 7.9 and Fig. 7.10. Compared to the fit using the HESE data set, the confidence regions are much smaller in this likelihood scan. This can be explained with the much higher number of events in the sample.

7.3.2 Hypothesis Test

To conduct the hypothesis test for the muon track data set the background hypothesis is fitted. The resulting atmospheric normalization is $\Phi_{atmo} = 1.28$. The fitted background model is shown in Fig. 7.11. Calculating the test statistic results in a p-value of 1.57×10^{-26} which equals a significance of 10.60 σ.

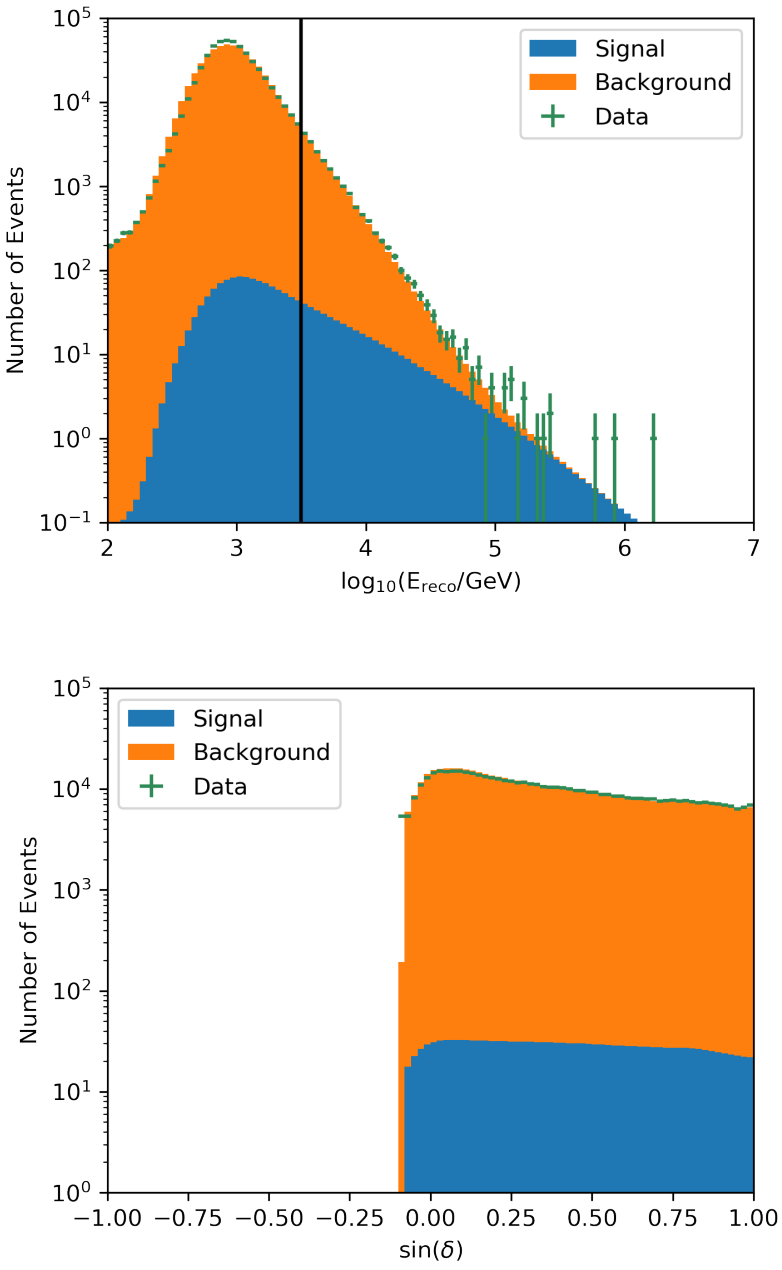


Figure 7.8: Comparison of the best-fit models to the data for the muon track data set. The vertical line shows the $10^{3.5}$ GeV energy cut.

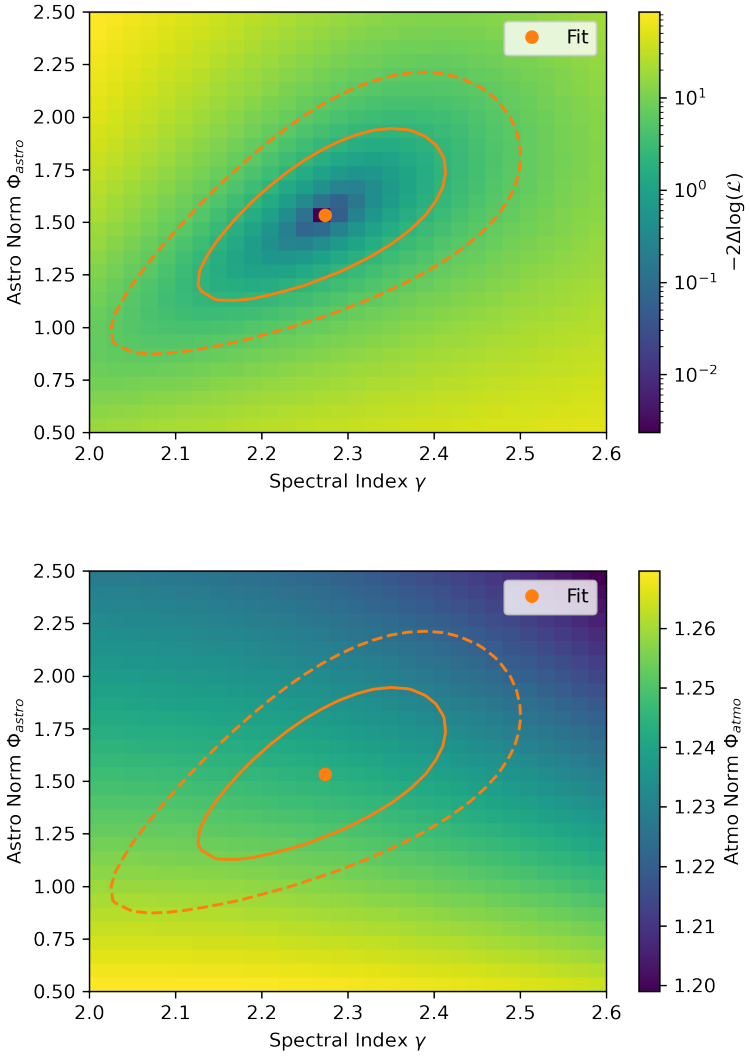


Figure 7.9: 2D likelihood scan over the astrophysical normalization Φ_{astro} and spectral index γ_{astro}. The first plot shows the likelihood ratio, while the lower plot shows the corresponding best-fit value for the atmospheric normalization Φ_{atmo}. The solid lines enclose the 68% confidence region and the dotted lines the 95% confidence region.

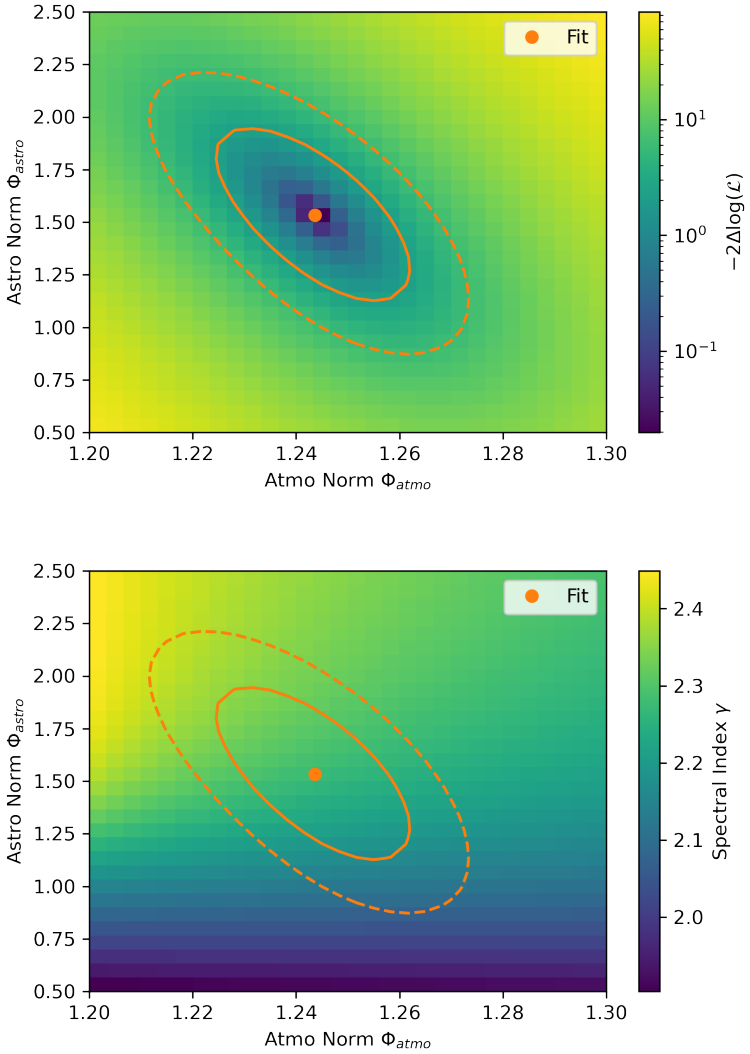


Figure 7.10: 2D likelihood scan over the two normalization parameters Φ_{astro} and Φ_{atmo}. The first plot shows the likelihood ratio, while the lower plot shows the corresponding best-fit value for the spectral index γ_{astro}. The solid lines show the 68% confidence region and the dotted lines the 95% confidence region.

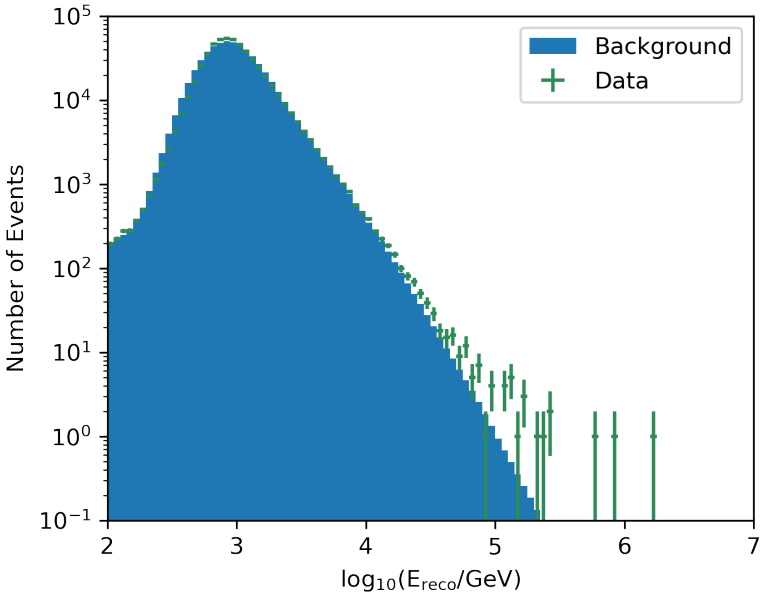


Figure 7.11: Fitting the background hypothesis to the muon track data. The data and model match best at lower energies since the much higher amount of measurements has a larger effect on the likelihood function.

7.4 Analysis of the Combined Data Sets

Next, the observed events of the two data sets are analyzed together. In this section, the analysis of the combined data is discussed.

7.4.1 Combined Likelihood Fit

To fit the flux models to the combined data a new likelihood function is needed. The combined likelihood function is constructed as the product of the two likelihood functions used for the muon track data fit and the HESE fit:

$$\mathcal{L}_{\text{combined}}(\vec{\Theta}_{\text{HESE}}, \vec{\Theta}_{\text{MuonTrack}}) = \mathcal{L}_{\text{HESE}}(\vec{\Theta}_{\text{HESE}}) \cdot \mathcal{L}_{\text{MuonTrack}}(\vec{\Theta}_{\text{MuonTrack}}) \tag{7.9}$$

This likelihood function is dependent on the parameters $\vec{\Theta}_{\text{HESE}}$ of the HESE fit and the parameters $\vec{\Theta}_{\text{MuonTrack}}$. Keeping the two sets of parameters independent of each other would give the same results as the two separate fits from Sec. 7.2.1 and Sec. 7.3.1. Since the two previous analyses assume the same models for the atmospheric and astrophysical fluxes the fit parameters are set to be equal for the combined fit.
For the astrophysical neutrino flux, the HESE fit assumes an all-flavor neutrino flux. Since the muon track data set selection aims to only include track-like events induced by muon neutrinos, a flux consisting of only muon flavor is assumed. Because of neutrino oscillation astrophysical neutrinos are expected to come in a ratio of roughly 1:1:1 (see Sec. 2.1). This means the parameter for the astrophysical normalization for the HESE data should be about three times as high as the same parameter for the muon track data.
For the atmospheric flux model only the components of the MCEq model (Sec. 5.2) with the correct flavor are included. This means the atmospheric normalization does not need to be modified.
Considering this leads to the following fit parameters for the combined fit:

$$\Phi_{\text{atmo,combined}} = \Phi_{\text{atmo,HESE}} = \Phi_{\text{atmo,MuonTrack}} \tag{7.10}$$

$$\Phi_{\text{astro,combined}} = \Phi_{\text{astro,HESE}} = 3 \cdot \Phi_{\text{astro,MuonTrack}} \tag{7.11}$$

$$\gamma_{\text{combined}} = \gamma_{\text{astro,HESE}} = \gamma_{\text{astro,MuonTrack}} \tag{7.12}$$

Using this definition for the fit parameters the likelihood function for the combined data sets is minimized. The combined fit uses the same energy cuts as the separate fits (60 TeV for the HESE data and $10^{3.5}$ GeV for the muon track data). Tab. 7.3 shows the resulting best-fit values in comparison to the values of the separate fits from Sec. 7.3.1 and Sec. 7.2.1.

Next, the fit results are analyzed using a 2D likelihood scan over the astrophysical normalization and the spectral index. Fig. 7.12 shows the result of the likelihood scan and

	Combined	HESE	Muon Track
Φ_{atmo}	1.24	0.91	1.24
Φ_{astro}	4.48	6.34	1.53 (4.59)
γ_{astro}	2.49	3.22	2.27

Table 7.3: Best-fit values of the combined fit compared to the separate fit results.

the confidence regions in comparison to the HESE fit and the muon track data fit.
The scan shows that both confidence contours of the separate fits do not overlap. This shows that the two results are not compatible under the assumptions of this analysis. However, as shown in Fig. 7.5 including systematic uncertainties like in the HESE analysis has a significant effect on the fit results. Including systematic in the analysis of both data sets might show better compatibility. Furthermore, the analysis of the HESE data set included higher energy events (over 60 TeV) while the muon track data set includes events with lower energies (over $10^{3.5}$ GeV). The different fit results could hint at a more complex astrophysical neutrino flux model, that shows different spectral indices depending on the energy.
The results of the combined fit show a spectral index that lies between the two separate best-fit values. The best-fit results of the combined fit are much closer to the muon track fit since the muon track data set consists of much more observed events than the HESE data set. With the combined fit a smaller confidence region could be obtained which shows that a combination of different data sets can lead to a better certainty in the obtained fit values.

7.4.2 Combined Hypothesis Test

Similarly to the separate analyses, a hypothesis test is conducted. For the null hypothesis only the atmospheric component is fitted and the astrophysical normalization $\Phi_{\text{astro,combined}}$ is set to zero. The minimization results in the best-fit value of $\Phi_{\text{atmo,combined}} = 1.29$ for the background hypothesis. After the calculation of the test statistic, a p-value of 1.39×10^{-37} is determined. This equals a significance of 12.76σ. With the larger amount of data compared to the separate fits, the fit results in a higher significance.

7.4.3 Astrophysical Normalization Ratio

As discussed in Sec. 7.4.1 the HESE and muon track data set contain different portions of the three neutrino flavors. While the HESE data set includes all three flavors the muon track data set aims to only include neutrinos of the muon flavor. Therefore the astrophysical normalization parameters of the two data sets differ by a factor

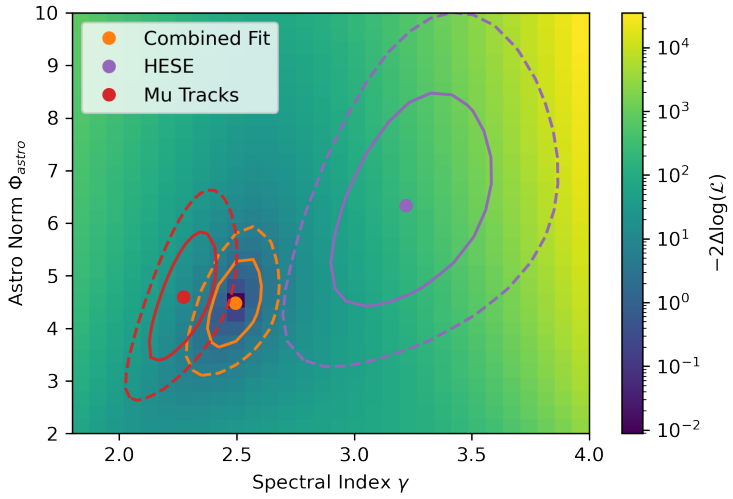


Figure 7.12: 2D likelihood scan over the astrophysical normalization Φ_{astro} and spectral index γ_{astro}. The best-fit results and confidence regions are compared to the separate fits. The solid lines show the 68% confidence region and the dotted lines the 95% confidence region.

$\Phi_{\text{astro,mu}}/\Phi_{\text{astro,HESE}}$.
In the combined fit from Sec. 7.4.1 a flavor ratio of 1:1:1 is assumed. Therefore a factor of 1/3 was chosen. However, the muon track data set does not purely consist of muon flavor events. As explained in Sec. 2.1 tau neutrinos can also lead to a track-like morphology via muonic decay. This leads to a 5% portion of non-muon flavor events in the muon track data set [14].
Taking this into account the fit can be repeated with independent astrophysical normalization parameters for the two data sets. The likelihood function for this fit is:

$$\begin{aligned}&\mathcal{L}_{\text{combined}}(\Phi_{\text{atmo,combined}}, \Phi_{\text{astro,HESE}}, \Phi_{\text{astro,MuonTrack}}, \gamma_{\text{combined}})\\ &= \mathcal{L}_{\text{HESE}}(\Phi_{\text{atmo,combined}}, \Phi_{\text{astro,HESE}}, \gamma_{\text{astro,combined}})\\ &\cdot \mathcal{L}_{\text{MuonTrack}}(\Phi_{\text{atmo,combined}}, \Phi_{\text{astro,MuonTrack}}, \gamma_{\text{combined}})\end{aligned} \tag{7.13}$$

Tab. 7.4 shows the best-fit values found through minimization in comparison to the results of Sec. 7.4.1. As expected the fitted ratio of the astrophysical normalization parameters of 0.57 is higher than 1/3. To analyze the uncertainty of this ratio a 2D likelihood scan is done. For this scan, the two astrophysical normalization parameters are varied while the best-fit likelihood is calculated. The likelihood landscape and the 68% and 95% containment regions are shown in Fig. 7.13. The likelihood scan shows that a ratio of 1/3 for the normalization parameters lies just barely inside the 95% confidence region. This shows that the difference in astrophysical normalization parameters for the two data sets can not be explained by just the different flavor contents of the samples.

	Independent	Dependent
$\Phi_{\text{atmo,combined}}$	1.23	1.24
$\Phi_{\text{astro,HESE}}$	3.09	4.48
$\Phi_{\text{astro,MuonTrack}}$	1.77	1.49
$\gamma_{\text{astro,combined}}$	2.44	2.49
$\Phi_{\text{astro,MuonTrack}}/\Phi_{\text{astro,HESE}}$	0.57	0.33

Table 7.4: Comparison of the best-fit values for a fit with independent astrophysical normalization parameters and a fit with a fixed ratio of 1/3.

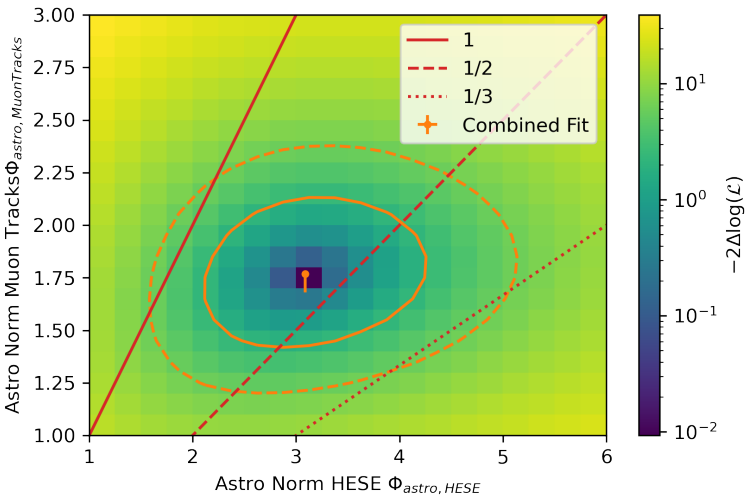


Figure 7.13: Likelihood scan over the two astrophysical normalization parameters. The best-fit value is plotted with an error bar showing the 5% contamination of non-muon flavor events in the track data set. The three diagonal lines show the points with a normalization parameter ratio of 1, 1/2, and 1/3.

8 Conclusion

In this thesis, two public IceCube data sets were used to measure the diffuse astrophysical neutrino flux. The fit results for a single power law astrophysical model show good compatibility with official IceCube analyses. However, the difference in the spectral index of the two fits indicates that a simple power law does not accurately describe the data of both samples simultaneously.
The combination of both data sets into a combined fit resulted in a higher significance and lower uncertainties for the fitted parameters. This shows that combining multiple data sets can lead to better results when measuring the diffuse astrophysical neutrino flux. In the future, adding more data sets to the analysis, possibly from other neutrino detectors, could improve the results even further.
The analysis in this thesis did not include any systematic uncertainties in the fitting process. As a next step, implementing detector systematics, like the DOM efficiency, would lead to a more accurate modeling of the detector response. This could allow more precise fit results.
Lastly, the simple power law model for the astrophysical neutrino flux could be replaced with more complex models that aim to better describe the observed neutrino flux.

Bibliography

[1] "Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector". In: *Science* 342.6161 (Nov. 2013). DOI: 10.1126/science.1242856. URL: https://doi.org/10.1126%2Fscience.1242856.

[2] R. Abbasi et al. "IceCube high-energy starting event sample: Description and flux characterization with 7.5 years of data". In: *Phys. Rev. D* 104 (2 July 2021), p. 022002. DOI: 10.1103/PhysRevD.104.022002. URL: https://link.aps.org/doi/10.1103/PhysRevD.104.022002.

[3] R. Abbasi et al. "Improved Characterization of the Astrophysical Muon–neutrino Flux with 9.5 Years of IceCube Data". In: *The Astrophysical Journal* 928.1 (Mar. 2022), p. 50. DOI: 10.3847/1538-4357/ac4d29. URL: https://dx.doi.org/10.3847/1538-4357/ac4d29.

[4] Lisa Johanna Schumacher et al. "PLE*nnu*M: A global and distributed monitoring system of high-energy astrophysical neutrinos". In: *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Sissa Medialab, July 2021. DOI: 10.22323/1.395.1185. URL: https://doi.org/10.22323%2F1.395.1185.

[5] Edoardo Vitagliano, Irene Tamborra, and Georg Raffelt. "Grand unified neutrino spectrum at Earth: Sources and spectral components". In: *Reviews of Modern Physics* 92.4 (Dec. 2020). DOI: 10.1103/revmodphys.92.045006. URL: https://doi.org/10.1103%2Frevmodphys.92.045006.

[6] John G. Learned and Karl Mannheim. "High-Energy Neutrino Astrophysics". In: *Annual Review of Nuclear and Particle Science* 50.1 (2000), pp. 679–749. DOI: 10.1146/annurev.nucl.50.1.679. eprint: https://doi.org/10.1146/annurev.nucl.50.1.679. URL: https://doi.org/10.1146/annurev.nucl.50.1.679.

[7] Thomas K. Gaisser, Ralph Engel, and Elisa Resconi. *Cosmic Rays and Particle Physics*. 2nd ed. Cambridge University Press, 2016. DOI: 10.1017/CBO9781139192194.

[8] R. L. Workman et al. "Review of Particle Physics". In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[9] M.G. Aartsen et al. "The IceCube Neutrino Observatory: instrumentation and online systems". In: *Journal of Instrumentation* 12.03 (Mar. 2017), P03012–P03012. DOI: 10.1088/1748-0221/12/03/p03012. URL: https://doi.org/10.1088%2F1748-0221%2F12%2F03%2Fp03012.

[10] Jöran Benjamin Stettner. "Measurement of the energy spectrum of astrophysical muon-neutrinos with the IceCube Observatory". Veröffentlicht auf dem Publikationsserver der RWTH Aachen University; Dissertation, RWTH Aachen University, 2021. Dissertation. Aachen: RWTH Aachen University, 2021, 1 Online–Ressource : Illustrationen, Diagramme. DOI: 10.18154/RWTH-2021-01139. URL: https://publications.rwth-aachen.de/record/811376.

[11] IceCube Collaboration. *IceCube Data for Neutrino Point-Source Searches Years 2008-2018*. 2021. DOI: 10.21234/CPKQ-K003. URL: https://arxiv.org/abs/2101.09836.

[12] M. G. Aartsen et al. "Searches for Extended and Point-Like Neutrino Sources With Four Years Of IceCube Data". In: *The Astrophysical Journal* 796.2 (Nov. 2014), p. 109. DOI: 10.1088/0004-637x/796/2/109. URL: https://doi.org/10.1088%2F0004-637x%2F796%2F2%2F109.

[13] Anatoli Fedynitch et al. *Calculation of conventional and prompt lepton fluxes at very high energy*. 2015. arXiv: 1503.00544 [hep-ph].

[14] M. G. Aartsen et al. "OBSERVATION AND CHARACTERIZATION OF A COSMIC MUON NEUTRINO FLUX FROM THE NORTHERN HEMISPHERE USING SIX YEARS OF ICECUBE DATA". In: *The Astrophysical Journal* 833.1 (Dec. 2016), p. 3. DOI: 10.3847/0004-637x/833/1/3. URL: https://doi.org/10.3847%2F0004-637x%2F833%2F1%2F3.

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst und keine anderen Quellen und Hilfsmittel als die angegebenen verwendet habe.

Ort, Datum	Unterschrift