technische universität
dortmund

Arbeit zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
(Dr. rer. nat.)

Search for Astrophysical Tau Neutrinos using 7.5 years of IceCube Data

Maximilian Meier
geboren in Herdecke

2019

Lehrstuhl für Experimentelle Physik V
Fakultät Physik
Technische Universität Dortmund

Erstgutachter:	Prof. Dr. Dr. Wolfgang Rhode
Zweitgutachter:	Prof. Dr. Bernhard Spaan
Abgabedatum:	08. Juli 2019

Abstract

Astrophysical tau neutrinos are the last unidentified standard model messenger in astroparticle physics. Their identification can open new windows to neutrino physics, improve knowledge about cosmic neutrino sources and even test physics beyond the standard model. This work aims to constrain the tau neutrino component in the astrophysical neutrino flux observed by the IceCube Neutrino Observatory. Due to neutrino oscillations over cosmic baselines, a significant fraction of tau neutrinos is expected regardless of the exact neutrino production scenario at cosmic sources. The IceCube detector instruments a volume of $1\,\mathrm{km}^3$ to detect neutrinos interacting with the glacial ice at the South Pole at a depth between 1450 m and 2450 m. This is achieved by 5160 digital optical modules (DOMs), each equipped with a photomultiplier tube detecting Cherenkov light produced by secondary particles from neutrino interactions.

In this dissertation, a new tau neutrino identification method is developed using state-of-the-art machine learning techniques to increase the expected tau neutrino event rate by a factor of 2.5 over previous work. Tau neutrinos are identified by the so-called double pulse signature, where two charge depositions can be observed in the waveform recorded in a single IceCube DOM: the first from the hadronic cascade induced by the neutrino interaction; the second one from a non-muonic decay of the produced tau lepton. This signature can be resolved by IceCube at energies above roughly 100 TeV. IceCube data recorded from 2011 to 2018 is analyzed and two tau neutrino candidates are observed. The astrophysical tau neutrino flux normalization is measured with a binned Poisson likelihood fit and the flux is observed to be $0.44^{+0.78}_{-0.31}\,10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}$ at 100 TeV for an astrophysical spectral index of $\gamma = 2.19$. The observation is found to be incompatible with the non-observation of a tau neutrino flux at a significance of 1.9σ.

Kurzfassung

Astrophysikalische Tau-Neutrinos sind das letzte nicht identifizierte Botenteilchen des Standardmodells in der Astroteilchenphysik. Ihre Identifikation kann neue Türen zur Neutrinophysik öffnen, das Wissen über kosmische Neutrinoquellen maßgeblich verbessern und sogar Physik jenseits des Standardmodells testen. Ziel dieser Arbeit ist die Untersuchung der Tau-Neutrino-Komponente des vom IceCube-Neutrino-Observatoriums nachgewiesenen astrophysikalischen Neutrinoflusses. Unabhängig von den genauen Prozessen der Neutrinoproduktion in kosmischen Quellen wird auf der Erde aufgrund von Neutrino-Oszillationen über kosmische Distanzen ein signifikanter Anteil an Tau-Neutrinos erwartet. Der IceCube-Detektor instrumentiert ein Volumen von $1\,\mathrm{km}^3$ um Neutrinos zu detektieren, die im Eis am Südpol in einer Tiefe von 1450 m bis 2450 m wechselwirken. Die 5160 digitalen optischen Module (DOMs) sind jeweils mit einem Photomultiplier ausgestattet, die Cherenkov-Strahlung, die durch Sekundärteilchen der Neutrinowechselwirkung enstehen, detektieren.

In dieser Dissertation wurde eine neue Methode zur Identifizierung von Tau-Neutrino-Ereignissen entwickelt, die sich auf moderne Methoden des maschinellen Lernens stützt und die erwartete Tau-Neutrino-Ereignisrate im Vergleich zu vorherigen Arbeiten um einen Faktor 2.5 erhöht. Tau-Neutrinos werden über die sogenannte Doppel-Puls-Signatur identifiziert. Dabei werden zwei Ladungsdepositionen innerhalb eines IceCube-DOMs beobachtet: Die erste stammt von der hadronischen Kaskade der Neutrino-Wechselwirkung und die zweite wird durch den nicht-myonischen Zerfall des entstandenen Tau-Leptons erzeugt. Eine Signatur dieser Art kann von IceCube bei Energien oberhalb von 100 TeV aufgelöst werden. Die in dieser Arbeit analysierten Daten wurden zwischen 2011 und 2018 mit dem IceCube-Detektor aufgezeichnet und es wurden zwei Tau-Neutrino-Kandidaten gefunden. Die Normierung des astrophysikalischen Tau-Neutrino-Flusses wurde über einen gebinnten Poisson-Likelihood-Ansatz für einen spektralen Index von $\gamma = 2.19$ zu $0.44^{+0.78}_{-0.31}\,10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}$ bei 100 TeV bestimmt. Diese Beobachtung ist mit einer Signifikanz von 1.9σ inkompatibel mit einem nicht vorhandenen Fluss von Tau-Neutrinos.

Contents

1 Introduction

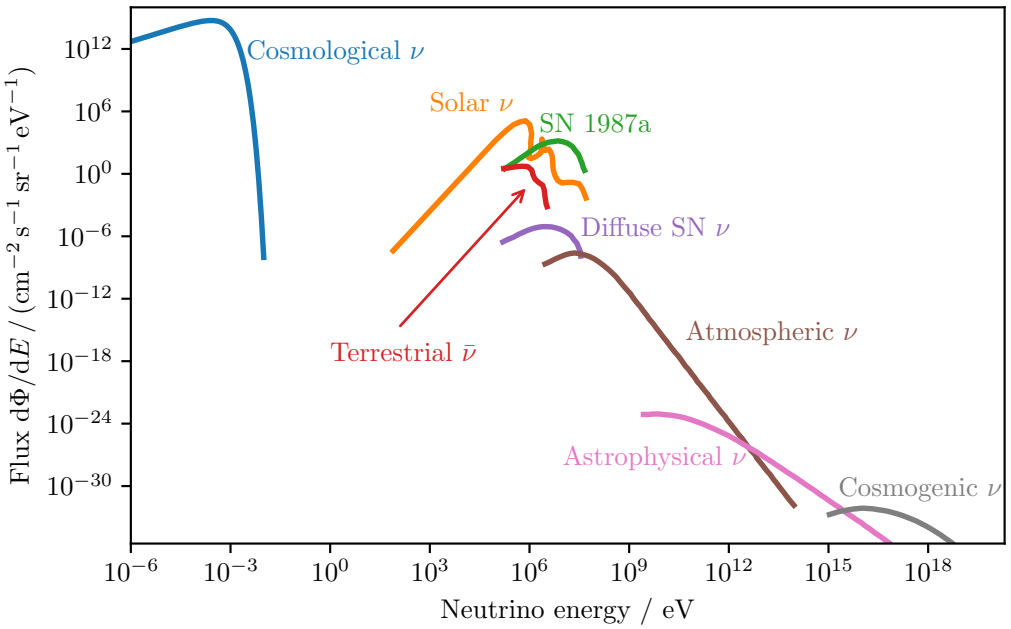


Figure 1.1: Measured and expected fluxes of natural sources of neutrinos. Shown are the energy spectra for cosmological neutrinos, solar neutrinos, terrestrial anti-neutrinos, the supernova 1987a, diffuse supernovae, atmospheric neutrinos, astrophysical neutrinos, and cosmogenic neutrinos. Figure adapted from [1].

High-energy neutrino astronomy is a very young field of astroparticle physics. The existence of high-energy neutrinos from extraterrestrial sources (cf. Fig. 1.1) was first observed by the IceCube collaboration in 2013 [2, 3]. After that discovery, many follow-up observations were conducted to precisely analyze the diffuse flux of astrophysical neutrinos via different detection channels [4, 5, 6, 7, 8]. The main goal of neutrino astronomy is to use astrophysical neutrinos to identify and examine the sources of high-energy cosmic rays. An important step in that direction was taken by the first discovery of a cosmic neutrino source TXS0506+056 in 2018 [9].

Neutrinos are uncharged particles only weakly interacting with matter, which makes them a unique messenger to examine high-energy cosmic ray sources. They are produced presumably by the decay of charged pions resulting in a flavor ratio of $\nu_e : \nu_\mu : \nu_\tau \simeq 1 : 2 : 0$ at the sources. Neutrino oscillations over cosmic baselines predict a flavor ratio at Earth of $\nu_e : \nu_\mu : \nu_\tau \simeq 1 : 1 : 1$. This astrophysical tau neutrino component is unique because tau neutrino production by cosmic ray air showers in the Earth's atmosphere is strongly suppressed by the high production threshold of heavy mesons that can decay into tau leptons and tau neutrinos. Astrophysical tau neutrinos are the last undetected standard model messenger particle and their detection would be the last missing confirmation for the astrophysical origin of the detected high-energy diffuse flux. Astrophysical tau neutrinos also allow us, in conjunction with a flavor ratio measurement, to explore neutrino oscillations at new scales, reaching higher energies and larger baselines than ever before.

Previously published IceCube analyses have found no tau neutrino candidates and only tau neutrino flux upper limits have been measured [10, 11]. This analysis aims to identify tau neutrino events with the so-called double pulse signature [12]. Above energies of roughly 100 TeV, the hadronic cascade from the neutrino interaction and the subsequent (non-muonic) decay of the tau lepton can produce two energy depositions observed as two distinct peaks in one of IceCube's optical light sensors. A new method to identify tau neutrino double pulse events is developed that relies on state-of-the-art machine learning techniques.

In this thesis, 7.5 years of IceCube data from May 2011 until December 2018 are analyzed. A binned Poisson likelihood fit is used to constrain the astrophysical tau neutrino flux based on different assumptions on the shape of the astrophysical flux obtained by recent IceCube measurements. Also, a differential upper limit is calculated to constrain the tau neutrino flux in a less model-dependent way. Finally, a new method to assign a p-value to tau neutrino candidate events is presented.

This dissertation is structured in the following way: chapter 2 gives a summary of some concepts of astroparticle physics, focusing on messenger particles relevant for this tau neutrino search. The next chapter, chapter 3, introduces the IceCube experiment and describes how the aforementioned messenger particles manifest themselves in the detector. The tau neutrino event selection developed in this thesis is explained in chapter 4. The applied analysis methods are presented and the resulting sensitivity to the astrophysical tau neutrino flux is evaluated in chapter 5. In chapter 6, the observed data sample and the resulting constraints on the astrophysical tau neutrino flux are presented and discussed in the context of recent IceCube measurements and current work-in-progress on astrophysical tau neutrinos. Chapter 7 summarizes the results and gives a brief outlook.

2 Astroparticle Physics

This chapter gives a short overview of some astroparticle physics topics, focusing on the relevant aspects for the analysis conducted in this dissertation.

Astrophysical events can be observed via the detection of cosmic messengers: cosmic rays, photons, neutrinos, and, only recently, gravitational waves. These messengers can convey information about the regions and objects that produce particles and accelerate them to the highest energies, like active galactic nuclei, gamma-ray bursts, and supernovae. Or, in the case of gravitational wave events, information about merger events of binary star systems can be obtained.

2.1 Cosmic Rays

The earth's atmosphere is steadily penetrated by a flux of charged particles from extra-terrestrial sources. The cosmic rays are composed of 90 % protons, 9 % α-particles and heavier ionized nuclei [13]. For extremely high energy cosmic rays above 1 EeV (10^{18} eV), the composition differs, which is subject of ongoing research [14, 15, 16].

The all-particle cosmic ray energy spectrum was measured from energies of a few GeV to energies beyond 100 EeV. An overview of the cosmic ray spectrum is given in Fig. 2.1. The whole energy spectrum can be approximately described by a series of power laws of the form $^{dN}/_{dE} \propto E^{-\gamma}$ covering different energy regions. The spectral index γ in this approximative model is given by

$$\gamma \simeq \begin{cases} 2.7 & \text{for } 10\,\text{GeV} \leq E \leq 1\,\text{PeV} \\ 3.1 & \text{for } 10\,\text{PeV} \leq E \leq 1\,\text{EeV} \\ 2.6 & \text{for } 10\,\text{EeV} \leq E, \end{cases}$$

for each energy region where the transition regions are called the "knee" (around 3 PeV) and the "ankle" (around 3 EeV) [13]. Charged cosmic rays are only suited in a limited way to study their sources because they get deflected by interstellar magnetic fields. The deflection is energy-dependent, such that only protons with energies above 10 EeV could be used to detect close cosmic ray sources. Unfortunately, the

flux in this energy regime is very small, due to the steep spectrum, and, additionally, protons with energies above 6×10^{19} eV (60 EeV) are expected to be suppressed by the Greisen-Zatsepin-Kuzmin (GZK) cutoff [17, 18]. The GZK-cutoff predicts a limited range for ultra high energy protons due to interactions with the cosmic microwave background (CMB). Above the GZK-cutoff energy, the center-of-mass energy of the cosmic ray proton and a CMB photon is large enough to produce a Δ-resonance. In this process, the protons lose some of their energy and the Δ^+ decays into a nucleon and a pion, which can also further decay into so-called cosmogenic neutrinos [19]. Cosmogenic neutrinos have yet to be observed, but their flux has been constrained by IceCube [20].

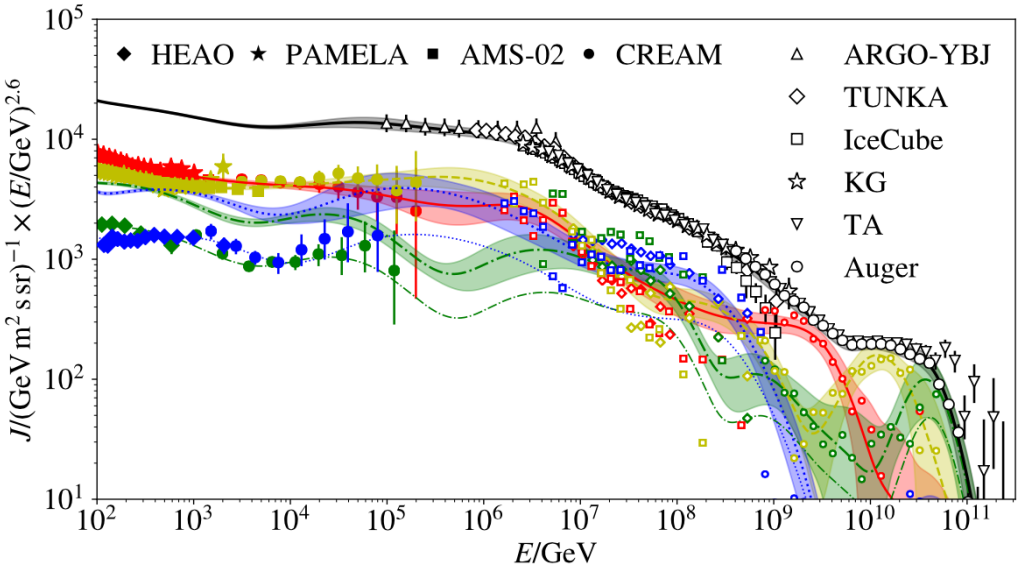


Figure 2.1: All-particle cosmic ray spectrum (black line) broken down into contributions from protons (red line), helium (dashed yellow line), the oxygen group (green dash-dotted line) and the iron group (dotted blue line). For the oxygen and iron group, the elementary flux is also shown without error bars. The shaded regions represent the combination of statistical and systematic uncertainties. This plot combines measurements from the HEAO satellite [21], PAMELA [22, 23], AMS-02 [24, 25], CREAM-I [26], CREAM-II [27], ARGO-YBJ [28], TUNKA [29, 30], IceCube [31], KASKADE-Grande [32, 33], Telescope Array [34] and the Pierre Auger Observatory [35, 14]. Figure is taken from [36].

The details about the origins of cosmic rays and about the processes by which particles are accelerated to the highest energies are an active area of research (a recent review is given in [37]). The currently most popular model for particle

acceleration in astrophysical environments is called diffusive shock acceleration, which was first introduced by Fermi [38]. This is also commonly referred to as first-order Fermi acceleration. Ultimately, the diffusive shock acceleration predicts particle spectra of the form $\mathrm{d}N/\mathrm{d}E \propto E^{-2}$. The softer spectrum observed for high energy cosmic rays can be attributed to magnetic suppression [39] as well as interactions with CMB photons and extragalactic background light (EBL) during cosmic-ray propagation [40].

2.2 Atmospheric Neutrinos and Muons

When high-energy cosmic rays penetrate the Earth's atmosphere, they collide with air nuclei. As a result of the primary interaction, a multitude of particles is produced in subsequent collisions and interactions. This process creates a cosmic-ray induced extensive air shower. The air shower consists of a hadronic and an electromagnetic component. The hadronic component also generates a muonic subcomponent and gives rise to the production of neutrinos from decays of e.g. charged pions and kaons.

2.2.1 Atmospheric Muons

Conventional atmospheric muons are produced by decays of charged pions and kaons. These hadrons inherit the primary cosmic ray spectrum $\mathrm{d}N/\mathrm{d}E \propto E^{-2.7}$. Due to their rather large lifetimes of $\sim 10^{-8}\,\mathrm{s}$, they interact with the atmosphere before decaying and losing energy in the process, which results in a much softer spectral index of $\gamma \simeq 3.7$. Conventional atmospheric muons are often produced in large amounts in a single shower resulting in muon bundles of up to thousands of muons. The number of muons in an air shower is also often referred to as the muon multiplicity.

Heavier, charmed mesons (D, D_s, Λ_c) and unflavored mesons (e.g. η, ρ^0) are also produced in cosmic-ray air showers. The lifetime of charmed and unflavored mesons is approximately four orders of magnitude smaller than the lifetime of charged pions and kaons. This reduces the probability of energy losses drastically before decaying into muons. As a result, the so-called prompt atmospheric muons are produced with a spectral index of $\gamma \simeq 2.7$. The decay of unflavored mesons into muon pairs is rather rare but counteracted by the absolute frequency with which these particles are produced. Above PeV energies, unflavored mesons start to be the dominant parent particle for atmospheric muons. The energy integrated flux of prompt muons is much lower compared to conventional muons, due to the higher energy threshold

required to produce charmed and heavy unflavored mesons. But, because of the harder energy spectrum, the prompt flux becomes dominant around 600 TeV.

Atmospheric muons are an important background to neutrino events for underground neutrino detectors. Muons trigger the IceCube detector at a rate of about 2700 events per second. The muon flux can accordingly be measured with high precision (cf. [41]). The results are consistent with an $E^{-3.7}$ spectrum and show an indication for a prompt component at the highest energies, albeit with large uncertainties.

2.2.2 Atmospheric Neutrinos

Similarly to atmospheric muons, atmospheric neutrinos can be divided into a conventional and a prompt component. Conventional atmospheric neutrinos originate from the decays of kaons, pions, and subsequent muon decays.

$$K^+, \pi^+ \rightarrow \mu^+ + \nu_\mu$$
$$\mu^+ \rightarrow \nu_\mu + e^+ + \nu_e$$

This list of decays is not comprehensive, and charge conjugated processes are to be considered as well. These neutrinos follow a similar spectrum as conventional atmospheric muons of $E^{-3.7}$ due to energy losses undergone by the pions and kaons before their decay. The muon decay becomes less relevant at high energies, as the fraction of muons decaying before reaching the surface gets smaller. Thus the majority of electron neutrinos are produced in kaon decays.

In contrast to prompt atmospheric muons, prompt atmospheric neutrinos are only produced in the decay of charmed mesons ($D^0, D^\pm$). Due to similar branching ratios of charmed mesons decaying into electron and muon neutrinos the expected flavor ratio is roughly $\nu_e : \nu_\mu \simeq 1 : 1$ [42].

D_s mesons can also decay into a tau lepton and a tau neutrino but these are only rarely produced, resulting in a prompt tau neutrino flux of smaller than 10 % of the prompt muon or electron neutrino flux at energies above 1 PeV (cf. Figure 2.2). The prompt atmospheric muon and electron neutrino fluxes are already extremely hard to detect because they are dominated by astrophysical neutrinos, which will be introduced in section 2.3. The flux of prompt electron and muon neutrinos has not yet been observed [6, 4], so the prompt tau neutrino flux is usually deemed negligible.

Measurements of the neutrino flux over seven orders of magnitude are presented in Figure 2.3. It shows mostly measurements of the atmospheric muon neutrino flux. At energies above 100 TeV, the fluxes show a change in the spectrum, which is described by an additional component: astrophysical neutrinos.

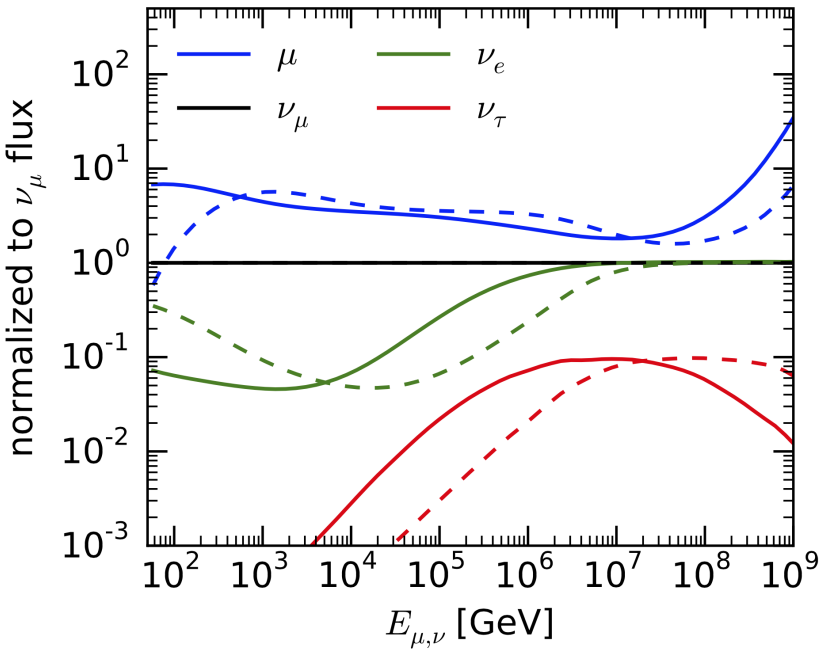


Figure 2.2: Atmospheric lepton flavor ratios at Earth's surface relative to the muon neutrino flux for vertically down-going (solid) and horizontal particles (dashed). Figure is taken from [43].

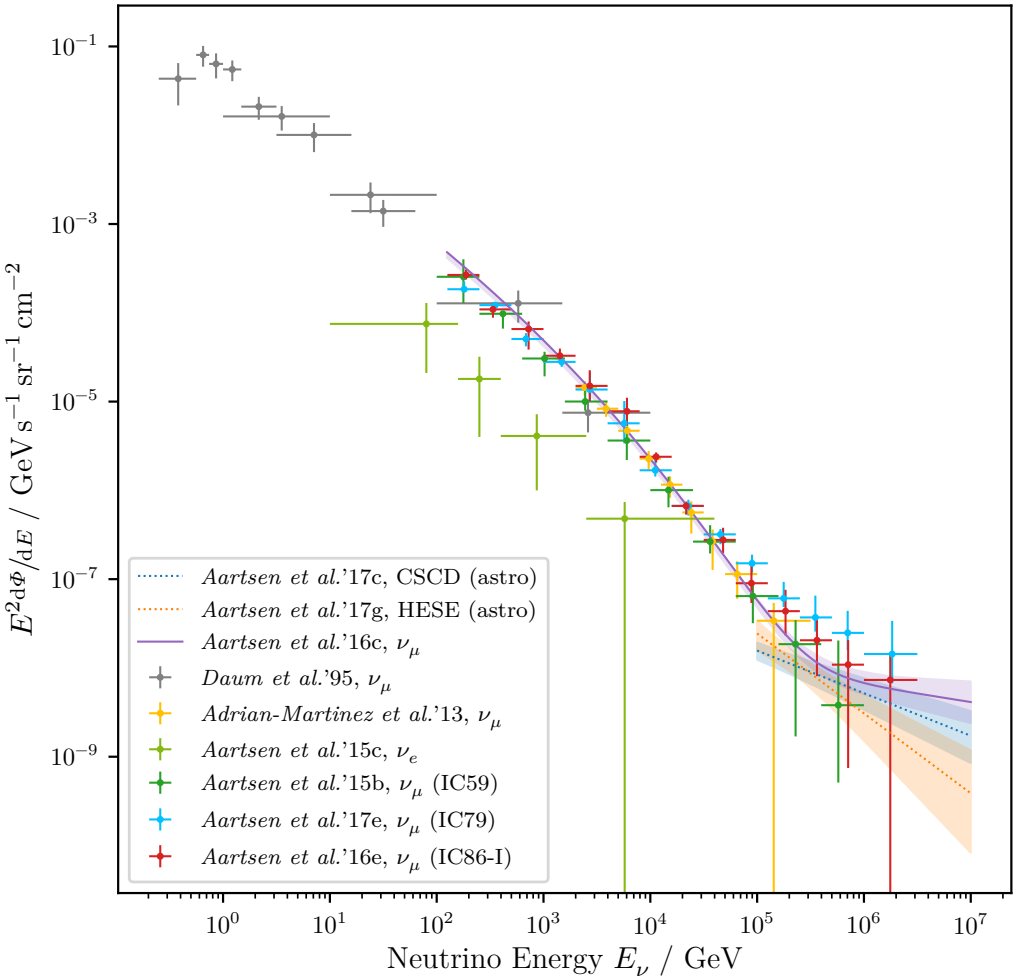


Figure 2.3: Measurements of neutrino fluxes over seven orders of magnitude in neutrino energy. Data points with error bars show unfolding flux measurements from Fréjus [44], ANTARES [45] and IceCube for muon neutrinos [46, 47, 48] and electron neutrinos [49]. Solid and dashed lines with error bands depict different forward folding measurements. The forward folding measurements are either combining a measurement of the atmospheric neutrino flux with the astrophysical neutrino flux (cf. section 2.3) [50] or measure the astrophysical flux only [51, 52]. Figure is taken from [53].

2.3 Astrophysical Neutrinos

Neutrinos from astrophysical sources of cosmic rays can be produced by the same mechanisms that produce neutrinos in the Earth's atmosphere. Typically, neutrinos are either produced by mesons created in inelastic hadronuclear collisions (pp-interactions) or by interactions of cosmic rays with radiation fields ($p\gamma$-interactions). Possible candidates for neutrino sources are numerous and will not be discussed here. An overview of possible neutrino sources is given in [54].

In the case of pp-interactions, a high energy proton accelerated in the source or its vicinity scatters inelastically with ambient hadronic matter to produce mesons, similarly to the processes described for the meson production in cosmic ray air showers.

$$p + p \rightarrow [\pi^0, \pi^+, \pi^-] + X \tag{2.1}$$

In contrast, for photohadronic interactions (with CMB photons, EBL photons or photons created by the source itself) the most dominant process of pion production, via the Δ^+-resonance, does not produce negatively charged pions:

$$p + \gamma \rightarrow \Delta^+ \rightarrow \begin{cases} \pi^+ + n \\ \pi^0 + p \end{cases} \tag{2.2}$$

Both neutrino production scenarios also produce high energy photons through $\pi^0 \rightarrow \gamma + \gamma$. Neutrinos produced by pion decays like this carry an average energy of approx. 5 % of the primary proton ($E_\nu \simeq 0.05 E_p$). In general, the source density is expected to be small compared to the density of the Earth's atmosphere and thus pions are not expected to lose a significant amount of their energy before decaying. As a result, the neutrino flux is expected to follow a power law similar to the expected cosmic ray flux at the source (albeit shifted to lower energies). For sources that predominantly produce neutrinos via photohadronic processes (e.g. radio-loud AGN), the resulting neutrino spectrum might be steeper than the proton spectrum [55].

Usually, neutrinos and anti-neutrinos are expected to be created equally by assuming a similar production rate of positive and negative charged pions. This only holds for pp-interactions, while for $p\gamma$-interactions a larger fraction of electron neutrinos compared to electron anti-neutrinos is produced. In general, IceCube can not discriminate between neutrinos and anti-neutrinos. The exception is the Glashow resonance [56], which only occurs for electron anti-neutrinos at energies around 6.3 PeV. This process theoretically allows to probe neutrino production scenarios, but significant conclusions are expected to require large exposures [57].

2.3.1 Flavor Ratio and Neutrino Oscillations

The neutrino flavor ratio predicted by neutrino production through pion decays at the source $\nu_e : \nu_\mu : \nu_\tau = 1 : 2 : 0$. This implicitly assumes that the muons produced in pion decays quickly decay before losing a significant amount of their initial energy. If the muon interacts with matter or strong magnetic fields before decaying, the muon decay will only produce a low energetic pair electron and muon neutrino resulting in a flavor ratio of $\nu_e : \nu_\mu : \nu_\tau = 0 : 1 : 0$. This is often referred to as the muon-damping scenario [58]. The next possible scenario is assuming a beam of high-energy neutrons from the source that will also produce a flux of electron anti-neutrinos through β-decays and thus a flavor ratio of $\nu_e : \nu_\mu : \nu_\tau = 1 : 0 : 0$ [59]. In the charm production scenario, equal amounts of electron and muon neutrinos are produced by the decay of charmed mesons: $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 0$. This is comparable to the production of prompt neutrinos in Earth's atmosphere. This scenario might be relevant at the highest energies if the flux of conventional neutrinos is suppressed by hadronic and radiative cooling [60]. An important implication of all these production scenarios is that neither case predicts any tau neutrino production at astrophysical sources.

However, the neutrinos produced at astrophysical sources oscillate between flavor eigenstates on their way to Earth due to non-vanishing mass eigenstates. This paragraph follows the explanations and definitions from [61]. Flavor and mass eigenstates of neutrinos are connected via the Pontecorvo-Maki-Nakagawa-Sakata mixing matrix (PMNS matrix) U

$$|\nu_\alpha\rangle = \sum_j U^*_{\alpha,j} |\nu_j\rangle \tag{2.3}$$

with the three flavor eigenstates $\alpha = e, \mu, \tau$, the mass eigenstates $j = 1, 2, 3$, and elements of the PMNS matrix $U_{\alpha,j} = \langle \nu_\alpha | \nu_j \rangle$. For a fixed propagation length L the flavor eigenstate takes the form $|\nu_\alpha(L)\rangle = \sum_k e^{-iE_k L} U^*_{\alpha,k} |\nu_k\rangle$. The probability of a neutrino being produced in flavor state α and being observed in flavor state β after propagating the distance L is given as $P_{\alpha,\beta} = \left| \langle \nu_\beta | \nu_\alpha(L) \rangle \right|^2$.

Over distances larger than the size of the solar system the oscillating interference terms average out to give the propagation matrix

$$\mathscr{P}_{\alpha,\beta} = \sum_j \left| U_{\alpha,j} \right|^2 \left| U_{\beta,j} \right|^2 . \tag{2.4}$$

This matrix can be used to propagate flavor ratios injected at the source to the corresponding flavor ratio that can be observed on Earth $\vec{\Phi}^f = \mathscr{P} \vec{\Phi}^i$, where the flavor ratio at the source $\vec{\Phi}^i$ and the flavor ratio at Earth $\vec{\Phi}^f$ are normalized such

that $||\vec{\Phi}||_1 = 1$. In the tri-bimaximal mixing (TBM) model[1] [63] the propagation matrix is given by

$$\mathscr{P}_{\mathrm{TBM}} = \frac{1}{18}\begin{pmatrix} 10 & 4 & 4 \\ 4 & 7 & 7 \\ 4 & 7 & 7 \end{pmatrix}. \tag{2.5}$$

The expected flavor composition at Earth, given the previously discussed neutrino production scenarios at astrophysical sources and the propagation matrix $\mathscr{P}_{\mathrm{TBM}}$, are presented in Table 2.1. Although no scenario predicts tau neutrino production at the source, due to neutrino oscillations, regardless of the scenario, a significant tau neutrino fraction is expected at Earth. The pion production scenario is usually considered as the benchmark model, predicting an almost equal flavor ratio of $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$. Measurements of the astrophysical neutrino flux with IceCube usually assume an equal flavor ratio and quote a per-flavor flux normalization.

Table 2.1: Flavor ratio of astrophysical neutrinos injected at the source $\vec{\Phi}^i$ and the resulting observable flavor ratio at Earth $\vec{\Phi}^f$ based on the TBM model for different neutrino production scenarios.

Scenario	$\vec{\Phi}^i = (\Phi^i_e, \Phi^i_\mu, \Phi^i_\tau)$	$\vec{\Phi}^f = (\Phi^f_e, \Phi^f_\mu, \Phi^f_\tau)$
Pion decay	$0.33 : 0.67 : 0.00$	$0.33 : 0.33 : 0.33$
Muon damping	$0.00 : 1.00 : 0.00$	$0.22 : 0.39 : 0.39$
Neutron decay	$1.00 : 0.00 : 0.00$	$0.56 : 0.22 : 0.22$
Charm production	$0.50 : 0.50 : 0.00$	$0.39 : 0.31 : 0.31$

2.3.2 Astrophysical Neutrino Flux Measurements

The first evidence for the proposed astrophysical neutrino flux was observed with the IceCube detector in 2013 [2], in the form of a diffuse flux measurement. This observation has been followed up by many analyses utilizing different detection channels (muon neutrinos, cascades, and starting events) to analyze the spectrum of astrophysical neutrinos with increasing precision. So far, all measurements are consistently described with an unbroken power law ($\Phi_\nu(E) = \Phi_0(E/100\,\mathrm{TeV})^{-\gamma}$). However, recent measurements still have found a range of different spectral indexes, with $\gamma = 2.19$ for upgoing muon neutrinos [4], $\gamma = 2.53$ for contained cascade events [6] and $\gamma = 2.92$ for high energy starting events [52] (all of these measurements

[1] The TBM model is an approximation to flavor mixing, which does not do current precision measurements of oscillation parameters justice, but it is sufficient to illustrate the effects of neutrino oscillations over astronomical distances in the context of this dissertation. More precise calculations regarding flavor ratios of astrophysical neutrinos can be found in [62].

can be compared in Fig. 2.3). The origin of this discrepancy between the different detection channels remains to be understood fully, along with a complete picture of the production of astrophysical neutrinos.

The astrophysical flux is largely observed as a diffuse flux in the sense that the arrival directions of astrophysical neutrinos are found to be isotropic. The sources of astrophysical neutrinos have not yet been identified with one exception. In 2018, the IceCube collaboration presented 3.5σ evidence for neutrino emission from the direction of the blazar TXS0506+056, which was prompted by observing a high-energy neutrino event coincident in time and direction with a gamma-ray flare from the same position in the sky [9].

3 The IceCube Neutrino Observatory

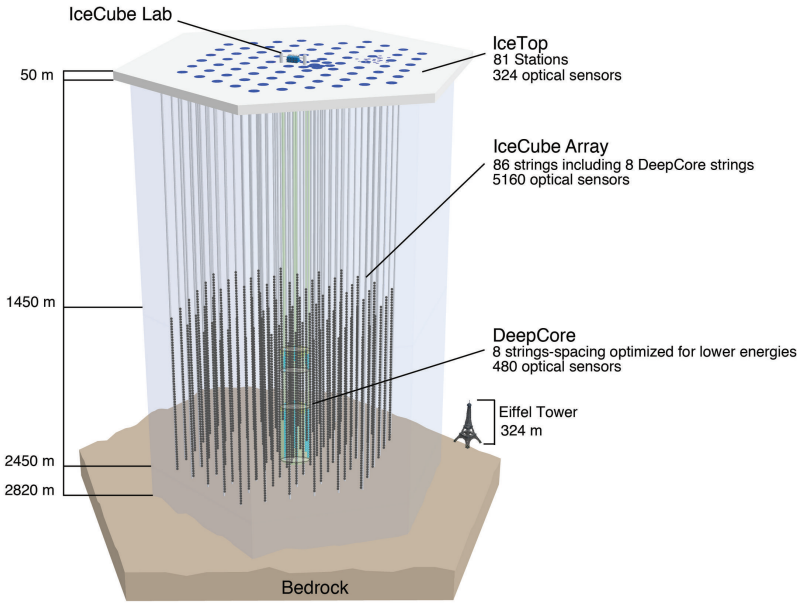


Figure 3.1: Schematic view of the IceCube Neutrino Observatory.

This chapter will briefly describe the components of the IceCube Neutrino Observatory: the South Pole ice used as the detection medium (section 3.1), the process of data acquisition (section 3.2), and the event topologies relevant for this work (section 3.4).

The IceCube Neutrino Observatory is located at the geographic South Pole and can be divided into three parts: the IceCube detector [64], DeepCore [65] and IceTop [66] (see Figure 3.1). The main part, the IceCube detector, is deployed at a depth between 1.5 km and 2.5 km and consists of 4860 digital optical modules (DOMs). The full detector configuration encompasses 86 strings (including 8 DeepCore strings) with 60 DOMs on each string. The strings are arranged in a triangular grid with a hexagonal footprint, a string-to-string distance of about 125 m and a vertical spacing of 17 m between DOMs, instrumenting a total volume of 1 km^3 of glacial ice. DeepCore is

a small infill array of 8 strings with a denser string, and a denser DOM spacing to mainly target neutrinos with energies down to a couple GeV to increase the sensitivity to e.g. atmospheric neutrino oscillations [67]. Each DOM contains a 10 inch Photomultiplier (PMT) and digitizing electronics protected from the pressure of the surrounding ice by a glass sphere.

3.1 Optical Properties of the South Pole Ice

IceCube's detection medium is the South Pole ice. The South Pole ice (SPICE) model [68, 69, 70] is the most frequently used ice model within the IceCube collaboration. The optical properties of the ice are modeled with depth-dependent absorption coefficients and effective scattering coefficients. These parameters are measured with in-situ calibration light sources. Three versions of this ice model are commonly used: SPICE Mie [68], SPICE Lea [70] and the latest version SPICE 3.2. The first iterations of the SPICE model described the individual ice layers as completely symmetric and isotropic. SPICE Mie extended the ice model to include calculations based on Mie scattering [71]. Additionally, it was found that the ice layers are not perfectly horizontal but rather slightly tilted [69], which is also included in this model. The model SPICE Lea incorporates evidence that the ice is not isotropic with respect to the direction in which the photons propagate. The latest version of the ice model, SPICE 3.2, mostly improves the model by using additional data of the in-situ calibration light sources and an updated description of the ice tilt. The latest model also allows for variations in the relative DOM-to-DOM efficiencies.

At a depth near 2000 m the absorption and scattering coefficients are increased due to an increased dust concentration. This region is referred to as the dust layer. This can be seen in Fig. 3.2 alongside the general depth dependence of the absorption and scattering coefficients.

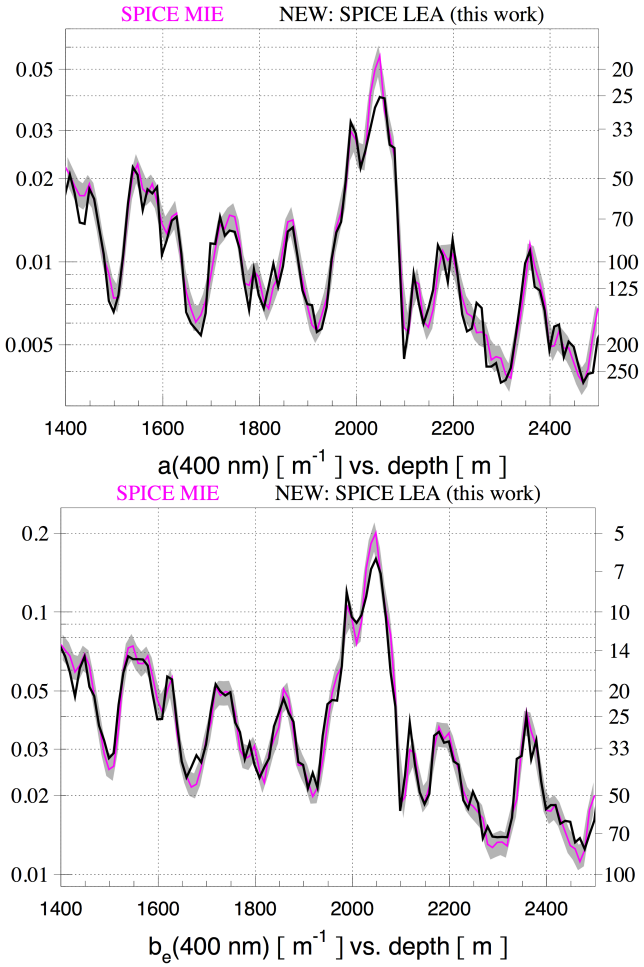


Figure 3.2: Fitted values of the depth-dependent absorption coefficients (top) and effective scattering coefficients (bottom) of the ice models SPICE Mie and SPICE Lea for photons with a 400 nm wavelength. The y-axis on the right side of the plots shows the absorption length and the effective scattering length in meters respectively. Figure is taken from [70].

3.2 Detection Principle and Data Acquisition

The detection of particles within the IceCube detector happens indirectly via the Cherenkov effect [72]. A high-energy charged particle produces so-called Cherenkov radiation when propagating through a transparent dielectric medium with a velocity that is larger than the speed of light in this medium. The produced radiation falls within the visible spectrum and can be detected by IceCube DOMs.

The basic detection unit in IceCube is called a hit, which is defined as the photomultiplier signal exceeding a trigger threshold of 0.25 PE (photoelectrons). The pulse shape caused by a hit is called a waveform. The digitization of signals (triggered hits) is done with the waveform digitizers on the DOM mainboard. Each DOM has two types of digitizers: the *Analog Transient Waveform Digitizer* (ATWD) and the *fast Analog to Digital Converter* (fADC). To minimize the DOMs dead time each DOM contains two ATWDs so that the first one can be busy digitizing signals while the other one records new signals. The ATWD produces waveforms with a duration of 422 ns and a time resolution of 3.3 ns resulting in a waveform with 128 time bins. The fADC has a smaller time resolution of 25 ns but is capable of recording waveforms with a duration of 6.4 µs. The ATWD also has a dynamic range which is controlled by three different gain channels (multipliers of 16, 2 and 0.25).

If a pair of nearest DOMs or next-to-nearest DOMs on a string is hit within a time frame of 1 µs, the hits are considered local coincidence hits (HLC). For these HLC hits, the full waveform information is sent to the IceCube Lab on the surface. An IceCube event is defined via the simple multiplicity trigger (SMT). The SMT requires at least 8 HLC hits within a time period of 5 µs anywhere in the detector. The event rate at trigger level is determined by the atmospheric muon flux, resulting in a median rate of ~ 2700 events per second. This data is processed further with filter algorithms selecting events based on fast reconstruction for, e.g., arrival direction and energy, which reduce the data volume from about 1 TB/day at trigger level to a volume of about 100 GB/day, which can be transmitted with the satellite bandwidth available to the IceCube collaboration to the corresponding data centers. [73, 74]

3.3 Neutrino Interactions

Neutrinos cannot be detected directly, but only via secondary particles produced in neutrino interactions with matter via the weak force. For neutrino energies relevant in this work, $E_\nu > 1\,\mathrm{TeV}$, the interaction is described as deep inelastic neutrino-nucleon scattering. The neutrino interacts with a nucleon either by exchanging

a charged $W^{\pm}$-boson (charged-current (CC) interaction) or a neutral Z^0-boson (neutral-current (NC) interaction):

$$\nu_l + N \to l + X \tag{3.1}$$
$$\nu_l + N \to \nu_l + X. \tag{3.2}$$

In a NC interaction (eq. 3.2) the neutrino transfers a fraction of its energy to the nucleon producing a hadronic cascade X. CC interactions (eq. 3.1) produce the same hadronic cascade at the interaction vertex. Additionally, the neutrino ν_l is transformed into a charged lepton l of the same flavor. Interactions of different neutrino flavors thus produce leptons of the corresponding flavor in CC interactions, which results in distinguishable signatures in the IceCube detector (cf. section 3.4).

This work relies on the deep inelastic scattering cross section calculated in [75]. The cross section is presented in Figure 3.3 separately for CC and NC interactions and for neutrinos and antineutrinos. The cross section rises with energy, which results in an increased shielding effect by the Earth for up-going neutrinos at very high energies. This effect is also zenith-angle dependent, as an increased zenith angle also increases the overburden. The survival probability for a zenith angle of $\theta = 180°$ and a neutrino energy of $E_\nu = 1\,\mathrm{PeV}$ is reduced to about 1 % [76]. This effect is slightly mitigated for tau neutrinos due to the fact that the dominating energy loss for a tau is the decay, so each tau neutrino interaction just produces a less energetic tau neutrino (and in leptonic decays also an electron neutrino or a muon neutrino) instead of being fully absorbed by the Earth.

As shown in Figure 3.3, between $\sim 5\,\mathrm{PeV}$ and $\sim 7\,\mathrm{PeV}$, the deep inelastic scattering cross section for antineutrinos is dominated by the interaction via the Glashow resonance (GR) [56]. This process describes the resonant scattering of electron antineutrinos with stationary electrons. This happens when the center of mass energy of electron and neutrino coincide with the rest mass of the W^--boson, which happens at a neutrino energy of $E_\nu \simeq 6.3\,\mathrm{PeV}$.

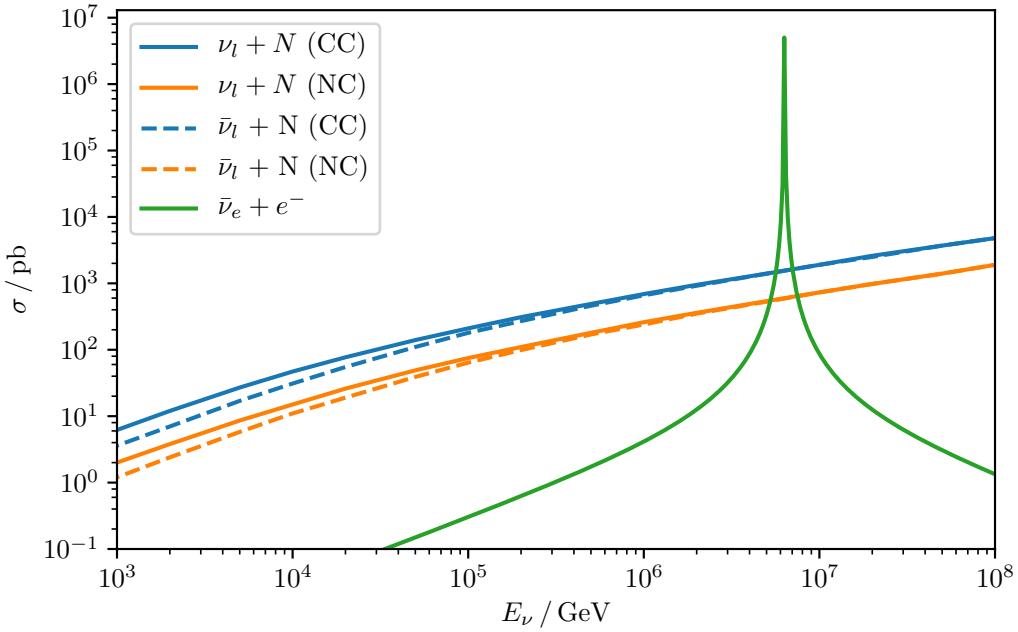
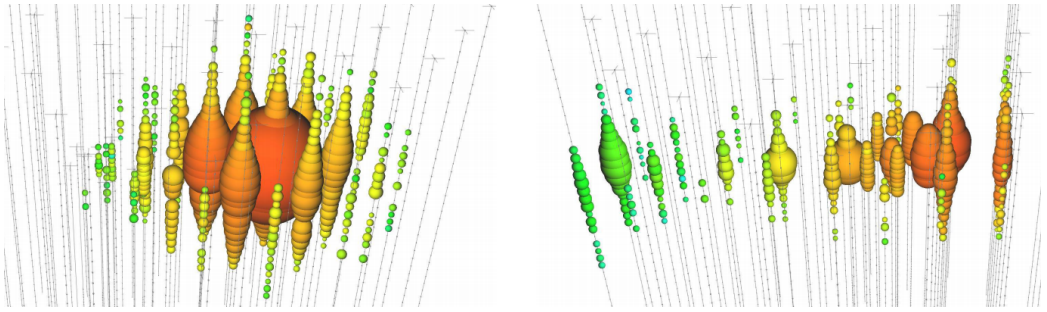


Figure 3.3: High energy (anti-)neutrino cross section as a function of energy. Values for CC and NC interactions are taken from [75]. The cross section for resonant $\bar{\nu}_e + e$ scattering (Glashow resonance) is described in [56].

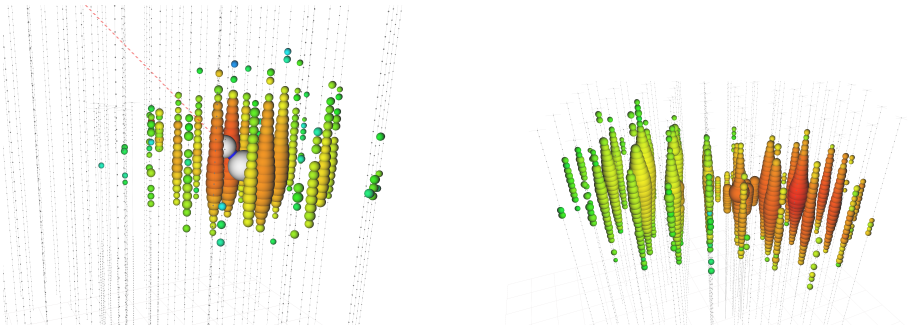
3.4 Event Topologies

IceCube events can be grouped into different event topologies based on the hit patterns they produce in the detector. Events can be visualized as shown in Figure 3.4. Each hit is depicted as a sphere around the DOM that was hit with a size corresponding to the amount of detected light. The color indicates the arrival time of photons ranging from early (red) to late (green/blue).



(a) Single cascade. (b) Track.

Figure 3.4: Examples of events for a single cascade and a starting muon track topology in IceCube data. Event views are taken from [2].



(a) Simulated double cascade event with a tau neutrino energy around 500 TeV and a tau length of 63 m. The gray spheres indicate the neutrino interaction vertex and the tau decay vertex.

(b) Simulated double bang event with a tau neutrino energy of approximately 14 PeV and a tau length of 346 m.

Figure 3.5: Examples of events for simulated charged current tau neutrino interactions.

Single cascade events are produced by NC neutrino interactions from all flavors and by electron neutrino CC interactions. In any of those cases, a hadronic cascade is produced in the interaction. In the case of an electron neutrino CC interaction

an additional electron (or positron if an antineutrino is interacting) is produced. Electrons and positrons have very short radiation lengths and thus deposit their energy over a distance of a few meters. Therefore, hadronic and electromagnetic cascades are effectively point-like light sources in the detector when compared to the inter string spacing of 125 m. Ultimately, these single cascade events result in an almost spherical light pattern in the detector (cf. Fig. 3.4a).

High energy muons can travel up to several kilometers through media such as air or ice [77]. Thus, they can traverse the IceCube detector while emitting Cherenkov light and producing stochastic energy losses along the way. This results in a light pattern resembling the muon track (cf. Fig. 3.4b). Tracks can originate from two types of events: atmospheric muons penetrating the detector or muon neutrinos undergoing a CC interaction and producing a muon in the process. In the latter case, if the interaction happens inside the detector, an additional hadronic cascade is produced at the interaction vertex producing a hybrid topology called a "starting track".

Finally, tau neutrinos can also produce a number of signatures that can theoretically be distinguished from the usual single cascades and tracks. After the initial hadronic cascade in a tau neutrino CC interaction, a tau is produced. The tau propagates through the detector until it decays after an average tau length of around $\ell_\tau \simeq 50\,\mathrm{m} \cdot {}^{E_\tau}/_{\mathrm{PeV}}$. Due to the large mass of the tau of 1.777 GeV ([42]), the most dominant energy loss is its decay and thus the actual tau track is not as bright compared to a muon with similar energy [77]. When decaying the tau produces either a track or an additional cascade based on the decay products. The relevant decay channels are

$$\begin{aligned} \tau^- &\to \mu^- + \bar{\nu}_\mu + \nu_\tau && B = 17.39\,\% \\ \tau^- &\to e^- + \bar{\nu}_e + \nu_\tau && B = 17.82\,\% \\ \tau^- &\to h^- + \nu_\tau && B = 64.79\,\% \end{aligned}$$

where $B_i = {}^{\Gamma_i}/_{\Gamma}$ is the respective branching ratio and h^- denotes the decay into any combination of hadrons. In the case where the tau decays into a muon, the overall signature will just be a starting track changing its brightness after the decay. The more interesting case for this work is the decay into either hadrons or into an electron, which produces a second cascade. The signature produced by these double cascades strongly depends on the tau length. At energies below $\sim 100\,\mathrm{TeV}$, the double cascades are not resolvable as such and just manifest themselves as a single cascade in the detector, such that it is impossible to distinguish them from NC neutrino interactions and CC electron neutrino interactions. When looking at higher energies, the probability either observing a double pulse in a waveform (due to light from both cascades hitting one DOM with sufficient time separation) or being able to reconstruct the event as a double cascade (cf. section 4.3.1) rises.

When both cascades are separated spatially in a way such that two cascades are visible, the signature is referred to as a double bang. Figure 3.5 shows two sample events created by a charged current tau neutrino interaction followed by either a hadronic or an electronic decay. Technically, both events are double cascade events, but in the case of the 14 PeV event (Fig. 3.5b), a double bang is clearly visible. A simulated, double pulse waveform is presented in Figure 3.6 and compared to a waveform from a single cascade event.

A larger variety of possible tau neutrino signatures in IceCube is described in detail in [12].

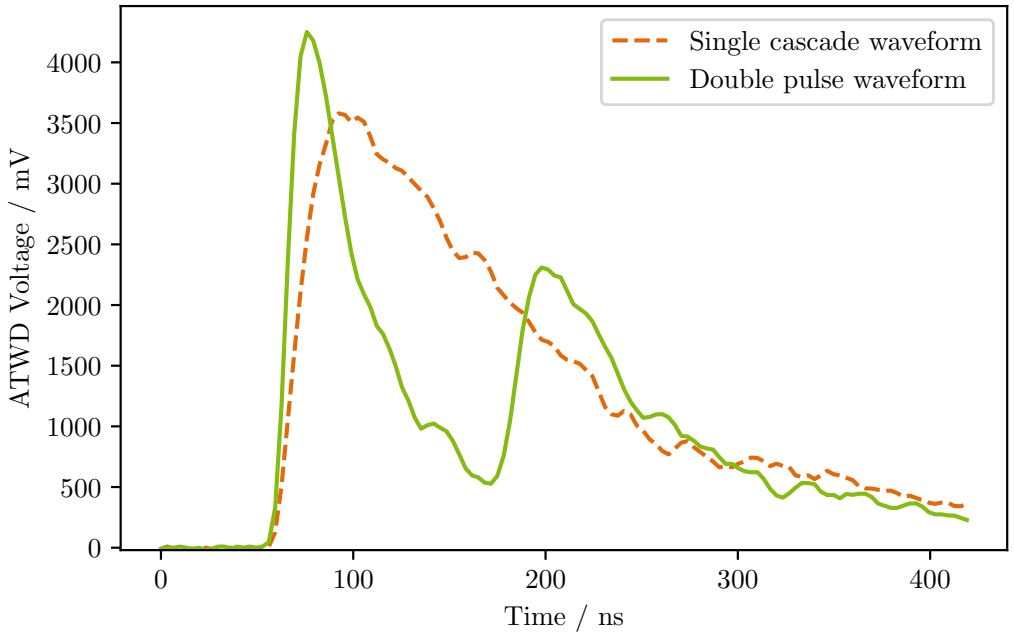


Figure 3.6: Comparison of a double pulse waveform created by a simulated charged current tau neutrino event (with a tau length of approx. 30 m) and a single cascade waveform from a simulated neutral current neutrino interaction.

4 Event Selection

Previous chapters were used to motivate the event selection and analysis conducted in this thesis, while also providing the necessary physics concepts to understand the reasoning applied here. The goal of this event selection is to obtain a tau neutrino-enriched sample, containing a high fraction of charged current tau neutrino events in which the tau decay produces a second cascade.

In this dissertation, data recorded by the IceCube detector in a time period between May 2011 and December 2018 is analyzed. During the whole period, the detector was operating in the 86-string configuration. The event selection criteria are developed and optimized with simulations only. Data is only used during the development of the event selection to check for disagreements between data and simulations. Only a 10 % subset of the data (burnsample) is used to avoid the introduction of a selection bias from the analyzer. This blind analysis policy is common in the IceCube collaboration and part of the internal review process. Most Figures in this chapter show the full amount of data used in this analysis, although, during the development of the event selection only the burnsample was available.

This event selection is divided into two parts. The first part describes the selection of double pulse waveforms, which is done first on low-level data (section 4.2). Events, that can produce double pulse signatures, are selected, while single cascade events originating from charged current electron neutrino interactions or neutral current neutrino interactions, in general, are removed from the data sample. After this step, the sample consists of mostly atmospheric muons, or muon and tau neutrinos interacting via a charged current interaction. Step two, described in section 4.3, aims to select the remaining charged current tau neutrino events based on their topological signature in the detector. The subset of remaining, dominant background events at this stage produces track-like events, while the sought after signal events are cascade-like. After this step, the resulting tau neutrino-enriched data sample is inspected and key characteristics are highlighted (section 4.4).

4.1 IceCube Simulations

All event selection steps require, besides the data recorded with the IceCube detector, labeled events from simulations. A simulated event factors in all processes from particle generation over propagation to the detector and particle interactions inside the detector to the signal recorded by detector electronics. A detailed explanation of IceCube's simulation chain is given in section A.1. This analysis uses dedicated simulations for each neutrino flavor, as well as `CORSIKA` simulation [78] for atmospheric muons from cosmic ray air showers, and `MuonGun` [79], which allows to replace full air-shower simulations by faster and therefore more efficient single-muon simulations.

The simulations are usually generated following a simple unbroken power law, but they can later be reweighted to any desired energy spectrum. The spectral index can also be chosen in a way to maximize statistics at, for example, higher energies. Throughout the event selection, the components are weighted in the following way, if not stated differently. Atmospheric muons (from `CORSIKA` simulations) are weighted to the spectrum described in [80]. The muons simulated with `MuonGun` use a parametrized spectrum based on `CORSIKA` simulations using the sum of single muons from the decays of light hadrons predicted by `SIBYLL` [81] and from charmed hadron decays predicted by `DPMJET` [82]. The expected flux for conventional atmospheric electron and muon neutrinos is based on the same spectrum for cosmic rays [80] and uses a modified version of the hadronic interaction model `DPMJET` to calculate a neutrino flux [83]. The prompt atmospheric component is modeled after [84]. Tau neutrino production in the atmosphere is very rare and will thus be neglected [43, 76]. For the astrophysical neutrino flux the measurement [50] is used, which measured an unbroken power law with a normalization of $\Phi_\nu = 0.9^{+0.3}_{-0.27} \cdot 10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}$ at 100 TeV and a hard spectral index of $\gamma = 2.13 \pm 0.13$. An astrophysical flavor ratio of $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$ is assumed.

4.2 Double Pulse Selection

This section describes the selection of double pulse waveforms recorded at a respective IceCube DOM. The starting point for this part of the event selection is called "Level 2". IceCube data at this level contains all triggered events and information about which of the several online filters the respective event passed. These filters sort events into different categories used by different types of analyses based on fast online reconstruction algorithms. This analysis does not rely on selecting events from a specific filter since there are no filters specifically designed to search for high energy

tau neutrinos. As explained earlier though (see sec. 3.4), the energy required to resolve double cascades as two separate pulses with IceCubes time resolution is above 10 TeV. This led to the introduction of a total charge cut of $Q_{\text{tot}} \geq 1 \times 10^{3.3}$ PE, which roughly corresponds to events with that deposited energy. Each event at energies above 10 TeV contains a variety of waveforms observed by different DOMs. Double pulses can only be measured in DOMs close to both the charged current interaction vertex and the decay vertex of the tau, therefore waveforms with few deposited photoelectrons ($Q_{\text{tot,WF}} \geq 432$ PE) are discarded [10].

4.2.1 Definition and Characterization of Signal and Background

Observables

The majority of observables used to describe the waveforms to possibly identify double pulse waveforms are based on the successive characterization of the derivative. At first, the beginning of the waveform is determined by searching for a monotonic increase of the waveform amplitude for a time period of 19.8 ns (6 ATWD bins). After defining the beginning of the waveform, it is divided into segments of 13.2 ns (4 ATWD bins) and the derivative is calculated for each segment. The derivative is scanned for a first pulse rising edge, a first pulse trailing edge and a second pulse rising edge by looking for $n_1 \geq 2$, $n_2 \geq 2$ and $n_3 \geq 3$ consecutive segments with a positive, negative and again positive derivative. Then the duration of these edges (number of segments) and their integrals are recorded and used as observables. These observables were originally developed in [10] and used for the first search for astrophysical tau neutrinos via the double pulse channel.

In addition to these derivative-based variables, new observables are generated to increase the separation power between single and double pulse waveforms.

- **Summary statistics**: Describes the waveform with summary statistics, in particular, the mean, variance, skewness, and kurtosis.
- **Number of local maxima**: Number of relative maxima compared with 15 ATWD bins to each side.
- **Smoothed distance**: Euclidean distance between a smoothed and an unsmoothed waveform (normed in a way that all bins sum to 1). The waveforms are smoothed with a sliding average width of $n = 7$.
- **Fit quality**: To estimate the compatibility with a typical single cascade the waveform is fitted with an exponential decay function. The χ^2/ndof and the Euclidean distance between the fit result and the waveform are recorded.

The most important of these new observables turned out to be the fit quality and the mean of the waveform. The fit quality variables are helpful to distinguish between an actual second pulse from an energy loss and a bumpy trailing edge due to e.g. late pulses from multiple scattering. The mean of the waveform is proportional to the total charge deposited and can thus enable energy-dependent cuts.

Signal

There is no obvious label for double pulse waveforms in IceCube simulations, so a set of waveforms has to be defined as signal waveforms manually. A double pulse waveform produced by a tau neutrino crucially depends on the deposited energies and vertex positions of the tau neutrino interaction vertex and the subsequent tau decay. As a first cut, all ATWD waveforms are selected that fulfill the following conditions (adapted from [85]):

$$\begin{aligned} \Delta t = |(t_1 + t_\tau) - t_2| &\geq 100\,\text{ns} \\ t_1 &\leq 200\,\text{ns} \\ t_2 &\leq 200\,\text{ns}, \end{aligned}$$

where t_1 represents the light travel time from the neutrino interaction vertex to the respective DOM, t_2 the light travel time from the tau decay vertex to the DOM and t_τ the decay time of the tau lepton.

These geometrical cuts select 13 000 waveforms, of which roughly 50 % show recognizable double pulses. To obtain a signal sample with a higher purity of double pulse waveforms a set of cuts is applied to the derivative features. An earlier iteration of the Double Pulse Algorithm described in [10], was optimized to pick up as many double pulse waveforms as possible with very little focus on the rejection of background waveforms [86]. Here the cuts are chosen to pick up the waveforms with a substantial double pulse feature (6874 waveforms), which are used as the sample of signal waveforms in the following classification.

Background

For the double pulse selection, two different types of backgrounds exist. The first type originates from single cascade events (ν_e CC and ν_X NC interactions), which mostly produce single pulse waveforms, except if heavy mesons (containing charm or bottom quarks) are produced in the hadronic shower. The production of heavy

mesons is not considered in IceCube simulations yet, but this effect is expected to be negligible because the single cascade component is subdominant at the final level of the event selection. The overall goal of this part of the analysis is the suppression of these single cascade backgrounds.

The second type of background is irreducible because double pulses can be easily mimicked by muons produced in high energy cosmic ray interactions and from charged current muon neutrino interactions. Usually, the first pulse is produced by direct light from the muon or from the initial hadronic cascade of the neutrino interaction and the second pulse is the result of a high energy stochastic loss from the muon (mostly Bremsstrahlung). This component will be ignored in the training to provide a clear goal for the classification to distinguish between double pulses waveforms and waveforms from single cascade events. The track-like component will therefore only be used for validation purposes and comparisons with data at this stage.

Background waveforms from single cascade events are abundant, so only a subset of those is used in the classification. For the ν_e CC interaction 1.2×10^6 waveforms and for the NC interactions 2×10^5 waveforms are randomly sampled for each neutrino flavor, which roughly resemble the relative occurrences of these event types. This also shows that the ratio between signal and background waveforms is heavily in favor of the background.

4.2.2 Classification

The classification of double pulse waveforms against the background of single cascade waveforms is carried out with a Random Forest (see section A.2, the parameters used for the Random Forest are presented in Table A.1). The training was carried out in a 10-fold cross validation so that every training example gets assigned a classification score. After this, a new model is built with all available training examples, which is used to assign a classification score to the remaining simulation events and to all data events.

The distribution of the classification score referred to from here on as “Double Pulse score”, is depicted in Fig. 4.1. Events are weighted to the fluxes described at the beginning of chapter 4, individual waveforms inherit the weight of the event they belong to. The double pulse waveforms start dominating the distribution above 0.6.

After this classification, only events that contain at least one waveform that survives a cut on the double pulse classification score are kept for further analysis. To choose such a cut, the remaining number of events per year for ν_τ CC events and single

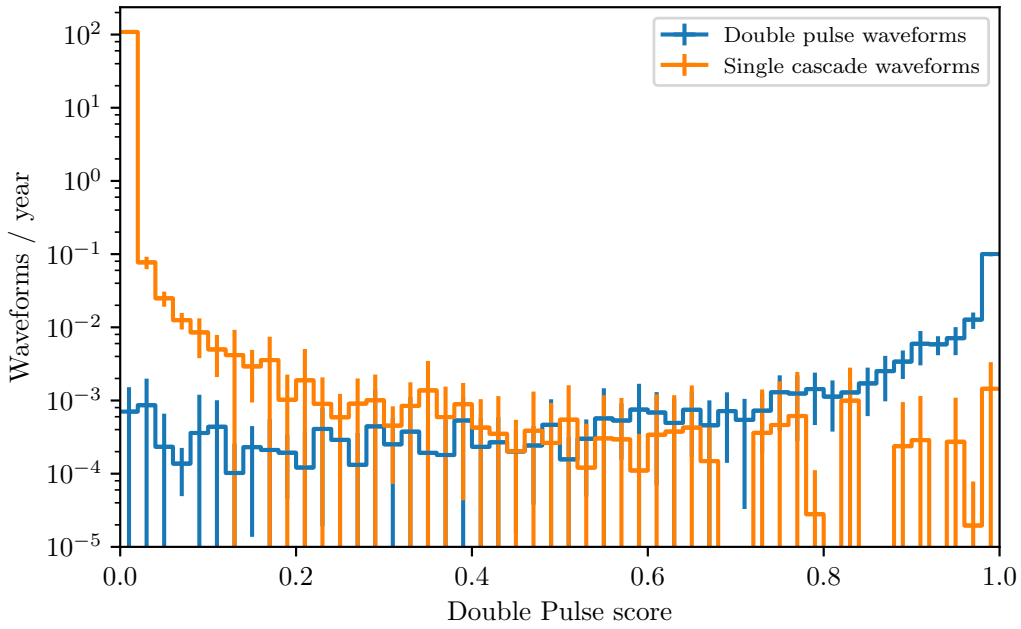


Figure 4.1: Distribution of the double pulse classification scores of the waveforms used in the Random Forest training. The standard deviation in each bin is determined from the fluctuation in each bin during the cross validation.

cascade events has to be considered, since, on the one hand, an event can have multiple double pulse waveforms. On the other hand, only a very limited amount of the total ν_τ CC waveforms were actually used as the signal to train the Random Forest and it is possible to retrieve a fraction of these waveforms by applying the final model to all ν_τ CC waveforms. Figure 4.2 shows these event expectations as well as the purity as a function of the Double Pulse score cut.

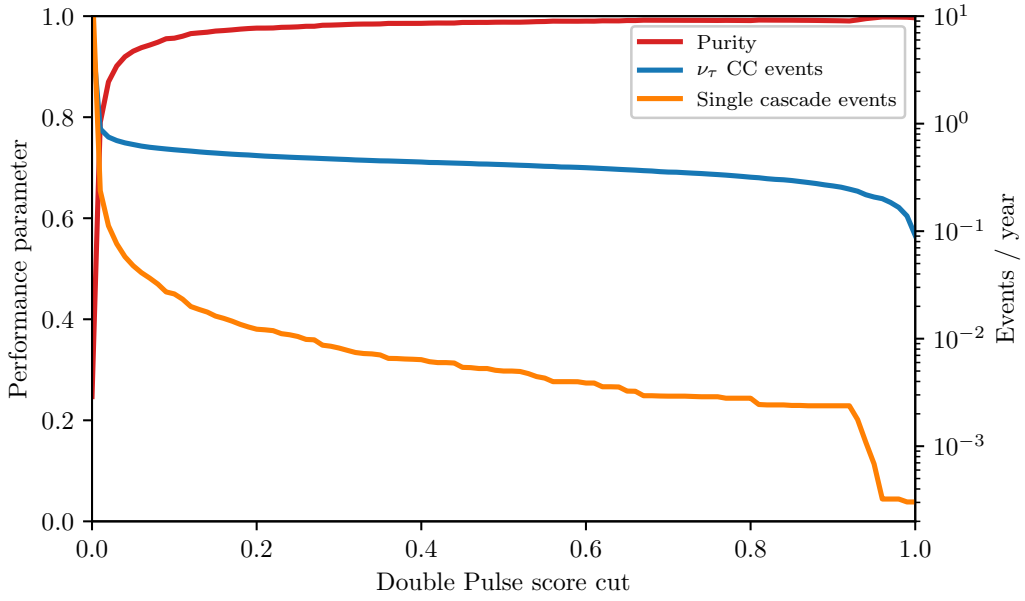


Figure 4.2: Influence of the Double Pulse score cut value on expected events per year for charged current tau neutrino and single cascade events (right y-axis) and the purity with respect to the two shown components (left y-axis).

At this point, a rather soft threshold of 0.2 is chosen. This way, the majority of single cascade backgrounds are already suppressed quite well, while increasing the retained signal by 90 % compared to [85], ending up with 0.63 events per year. The purity at this step is also increased from 90 % to 97 %.

While single cascade events get removed quite well with this cut, the signal is still dominated by background from atmospheric muons and ν_μ CC (see Figure 4.3). This figure also shows a comparison between the burnsample of 2012 data and the sum of simulations. The error bars shown use a modified Poisson likelihood, which also takes Monte Carlo uncertainties from limited statistics into account [87][1].

[1] This style of presenting binned comparisons between data and simulation is suggested in [88].

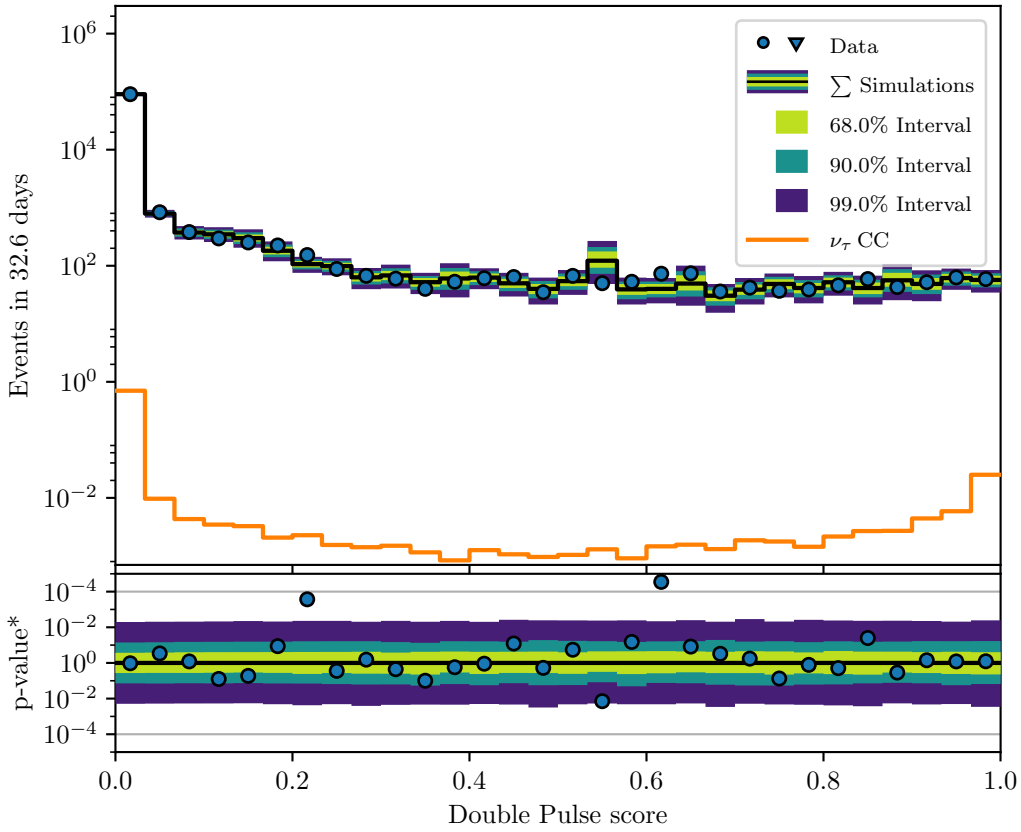


Figure 4.3: Comparison of the double pulse classification score distribution for simulations and the burnsample from the year 2012. In addition to the sum of simulations, the ν_τ CC component is depicted, which shows that the data at this point is still dominated by atmospheric muons roughly five orders of magnitude more abundant than the ν_τ CC component.

4.3 Cascade Selection

It was shown in the previous section that track-like events from atmospheric muons and charged current muon neutrinos make up an irreducible background when selecting events that contain double pulse waveforms. Fortunately, these events can be distinguished due to their distinctive event topologies. The previous analysis step was applied on very low-level data ("Level 2") containing only a minimum of low-level reconstructions. As a first measure to improve the quality of available reconstructions the IceCube "Level 3" cascade filter stream[2] reconstructions are applied to the data. Additional observables, that focus on the rejection of muon background and on characteristics from the double pulse selection, are generated as well. All observables used in the classification, or relevant to this analysis in any other manner, are described below.

4.3.1 Data Processing and Feature Generation

Reconstructions and observables realized in Cascade "Level 3"

Tensor of Inertia[3]
The pattern of hits (charges collected by each DOM) can be interpreted as a body with masses and thus used to determine a first guess for cascade directions and as an analytic seed for further reconstructions. The center of gravity marks the pivot point for this reconstruction

$$\vec{x}_{\mathrm{CoG}} = \frac{\sum_i (q_i \vec{x}_i)}{\sum_i q_i}, \tag{4.1}$$

where q_i is the total charged collected by the i-th DOM and $\vec{x}_i$ the DOM position. The tensor of inertia elements can then be calculated via

$$I^{jk} = \frac{\sum_i q_i \left(\delta^{jk}(\vec{r}_i)^2 - \vec{r}_i^{\,j} \cdot \vec{r}_i^{\,k}\right)}{\sum_i q_i} \tag{4.2}$$

with the distance $\vec{r}_i$ of DOM i to the center of gravity $\vec{x}_{\mathrm{CoG}}$. This tensor (in this case a 3×3 matrix) has an eigenvalue e for each of its main axes. The eigenvalue ratio can be used to measure the sphericity of the observed charge pattern.

$$q_{\mathrm{ToI}} = \frac{\min_{k\in[0,2]} e_k}{\sum_k e_k} \tag{4.3}$$

[2]The source code for this project can be found at code.icecube.wisc.edu/projects/icecube/browser/IceCube/projects/level3-filter-cascade.
[3]This reconstruction is referred to as `CLast` in the IceCube code base.

The value of q_{ToI} approaches ⅓ for a perfect sphere and reaches 0 if all hits lie on a straight line.

CascadeLlhVertexFit
The CascadeLlhVertexFit performs a likelihood reconstruction of a neutrino induced cascade to reconstruct the cascade vertex and vertex time. For each event, the likelihood

$$-\ln \mathscr{L}(t_0, \vec{x}_0 | t_{\mathrm{obs}}, \vec{x}_{\mathrm{obs}}) = -\sum_i \ln p_i(t_{\mathrm{obs},i}, \vec{x}_{\mathrm{obs},i} | t_0, \vec{x}_0), \tag{4.4}$$

is minimized, where the PDF $p(t_{\mathrm{obs},i}, \vec{x}_{\mathrm{obs},i} | t_0, \vec{x}_0)$ describes the probability that a DOM's first hit at time t_{obs} and position $\vec{x}_{\mathrm{obs}}$ would be the result of a cascade characterized by the set of parameters $t_0, \vec{x}_0$. The index i then runs over all DOM hits in a particular event. The PDF (called the Pandel-PDF [89]) models the photon arrival time distribution as an analytic function, which approximates the south pole ice as a homogeneous medium, so that time delays in the arrival time from photon propagation through the ice are considered.

All likelihood-based reconstructions also provide estimates for the fit quality with respect to the used hypotheses: the reduced log-likelihood (rlogl). This value is calculated by normalizing the best fit log-likelihood with the number of DOMs hit in an event. Smaller values generally indicate better agreement with the hypothesis used in the fit.

CascadeFillRatio
CascadeFillRatio uses a reconstructed vertex and a given set of hits to calculate different radii and a so-called fill ratio. The radii are based on the RMS and the mean of the charge distribution.

$$r_{\mathrm{RMS}} = \sqrt{\left(\sum \vec{r} \cdot \vec{x}\right) - \left(\sum \vec{r}\right)^2} \tag{4.5}$$

$$r_{\mathrm{mean}} = \frac{1}{N} \sum |\vec{r}| \tag{4.6}$$

$$r_{\mathrm{mean+rms}} = r_{\mathrm{mean}} + r_{\mathrm{RMS}}, \tag{4.7}$$

where the vertex position $\vec{x}$ and the distances between vertex and hit $\vec{r}$. The fill ratio then is just the ratio of the number of DOMs hit inside the sphere defined by the reconstructed vertex and the calculated radius to the total number of DOMs inside the sphere.

LineFit, SPEFit, and MPEFit
Because track-like events are a major background in this analysis, dedicated track reconstructions can help successfully identifying these as background events.

The LineFit [90] is a first guess track reconstruction with an analytic solution. The reconstruction assumes photons traveling on a wavefront perpendicular to the track. This approximation can be described by

$$\vec{x}_i = \vec{x} + t_i \cdot \vec{v}, \tag{4.8}$$

where $\vec{x}$ and $\vec{v}$ are a position and a velocity vector, used as the free parameters of this reconstruction. $\vec{x}_i$ and t_i are the position of DOM i and the arrival time of the first pulse at that DOM. This is a linear problem, with the analytic solution

$$\vec{x} = \langle \vec{x}_i \rangle - \vec{v} \langle t_i \rangle \tag{4.9}$$

$$\vec{v} = \frac{\langle t_i \cdot \vec{x}_i \rangle - \langle \vec{x}_i \rangle \langle t_i \rangle}{\langle t_i^2 \rangle - \langle t_i \rangle^2}, \tag{4.10}$$

where $\langle\ \rangle$ denotes the mean with respect to all hits. The absolute value of the velocity vector $v = |v|$ can be used to e.g. distinguish cascades ($v \simeq 0$) from minimally ionizing muons ($v \simeq c$). LineFit is commonly used as a seed for the likelihood-based track reconstructions SPEFit and MPEFit.

SPEFit (single-photoelectron) and MPEFit (multi-photoelectron) are, in general, very similar to CascadeLlhVertexFit, in the sense that the likelihood uses the Pandel-PDF as the photon arrival time distribution. The difference between the cascade fit and the track reconstructions is the hypothesis used to model the light propagation. For a cascade, a stationary light source is assumed, while track reconstruction uses a moving light source on an infinite track and also considers the angle of emitted Cherenkov light. SPEFit only relies on the first hit measured in each DOM, while MPEFit uses a more thorough description by calculating the arrival time distribution for the first N photons [91].

DipoleFit
DipoleFit [90] is a first guess algorithm, that calculates the dipole moment by averaging over individual dipole moments of hits (at position $\vec{x}_i$) adjacent in time. The performance of this reconstruction can be improved by not using the two closest hits in time but hits with a distance of N hits between them.

$$\vec{M} = \frac{1}{N_{\text{hits}} - N} \sum_{i=N+1}^{N_{\text{hits}}} \frac{\vec{x}_i - \vec{x}_{i-N}}{|\vec{x}_i - \vec{x}_{i-N}|}, \tag{4.11}$$

The IceCube “Level 3” cascade filter uses $N = 4$.

BayesianFit
This fit [92] is a slight modification that can be applied to any of the available likelihood-based reconstructions. It makes use of Bayes’ theorem

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)} \tag{4.12}$$

to add a prior $P(H)$ in form of a zenith PDF to track reconstructions. $P(E|H)$ is the likelihood of observing an event E given the track hypothesis H. $P(E)$ is just a normalization constant and thus it can be omitted from this calculation. These biased reconstructions usually do not improve the directional reconstructions but it can greatly improve the ability to distinguish between down-going and up-going events.

NChannel
NChannel describes the number of DOMs hit in an event.

Monopod
Monopod is the most accurate method used by IceCube to reconstruct cascades. To achieve this, Monopod employs a maximum likelihood approach to match a simulated template cascade with the observed data. The templates are generated by running Monte Carlo simulations for mono-energetic cascades at different positions and with different incident directions in the detector and make use of the fact that the light yield at the DOMs scales linearly with energy. Due to the complete Monte Carlo simulation of the detector and the glacial ice in Antarctica, all effects introduced by detector electronics and photon propagation in the ice are automatically included with this approach. The simulations are realized with the software Photonics [93] on a discrete grid and then parametrized by a multi-dimensional B-spline surface [94]. After binning the observed number of photons at each DOM, the observed event can be described by a Poisson Likelihood

$$\ln \mathscr{L} = k \ln(E\Lambda + \rho) - (E\Lambda + \rho) - \ln(k!), \tag{4.13}$$

where the expected number of photons in a time bin is given by the expected number of noise photons ρ in addition to the template Λ scaled by the energy E. This equation does not have an analytical solution for the parameters describing the cascade, however, solutions can be found by using numerical minimization. [95]

The IceCube cascade "Level 3" uses two techniques to split up the recorded hits, which will be explained in the following paragraphs.

CoreRemoval Fits
Observed hits are split into two halves: the case, which includes the hits close to a given vertex (core); and the corona, which is made up of hits further out. Hits are associated with the core if their distance to the reconstructed vertex position is smaller than a certain radius, and associated with the corona otherwise (for a more detailed description see [96]).

TimeSplit Fits
Recorded hits are split based on their recorded hit times. The point in time of the split is calculated as the charge weighted mean time

$$t_{\text{split}} = \frac{\sum t_i \cdot q_i}{\sum q_i}. \tag{4.14}$$

These two splitting strategies are used to rerun some of the described reconstructions (CscdLlhVertexFit, SPEFit, and DipoleFit) on both halves of the recorded hits.

Additional observables

I3VetoModule
This algorithm performs first guess calculations to estimate how contained a certain event is in the detector volume. Of interest for this analysis specifically are the depth of the first hit, which can give a hint if a cascade might be associated with a penetrating muon. Since this calculation can get easily messed up by noise hits, a cleaning algorithm to remove noise pulses is applied beforehand. The vertical layer of strings containing the DOM that registered the highest charge of the event and an observable called "earliestContainment", an integer value that combines information from the aforementioned vertical layer containing the first hit and the number of the DOM with the earliest hit, are calculated as well. Thus this variable combines information about the position of the earliest hit in the x-y-plane with a rough estimate of the z-position.

VHESelfVeto
The Very High Energy Self Veto is a technique used in the event selection, that first reported evidence for high energy astrophysical neutrinos [2, 3]. This analysis implements a detector edge veto to identify high energy starting events (HESE) with more than 6000 PE over a background of penetrating atmospheric muons. The veto decision is made by calculating the time at which the deposited charge in the detector exceeds a configured detection threshold (250 PE) and a charge weighted average position of the light that contributed to exceeding this threshold. If the charge deposited within the outer layer of the detector by that time crosses 3 PE the event is vetoed.

Veto Track
Penetrating muons produced by cosmic rays in the Earth's atmosphere are the most challenging background for this analysis (which will be illustrated in section 4.3.3). They are especially difficult to identify if they produce only a dim track and then lose most of their remaining energy in a catastrophic stochastic loss mimicking a cascade-like event. The VHESelfVeto was primarily designed for very high energetic

events (the passing fraction of atmospheric muons rises quickly with decreasing energy). To extend the veto approach to lower energies, some modifications are necessary. The approach taken here uses the reconstructed Monopod vertex as a starting point. All hits are considered for the veto that are not consistent with the reconstructed cascade, but consistent in time and space with a hypothetical downgoing track. The direction of this possible dim downgoing track is unknown and it is very unlikely to be reconstructed correctly by the available algorithms for track reconstruction. To make sure possible veto hits are picked up, the veto calculation is repeated for of 104 track hypotheses, covering all of the northern hemisphere. The finally selected "Veto Track" is the track with the highest veto charge among those tested tracks. The recorded charge and the number of hits on that track can be a useful quantity to reject atmospheric muons. The method is described in full detail in [79].

Starting Track
Another possible background, especially for cascade analyses, are starting tracks originating from charged current muon neutrino interactions. To reject those events a strategy similar to the one described for the "Veto Track" can be used, except in this case the incoming track has to be replaced by an outgoing one. Again this method is searching for hits inconsistent with the reconstructed cascade, but consistent with an outgoing track. It is necessary here to scan the whole sky instead of the northern hemisphere only, since muon neutrinos, in contrast to muons, are not absorbed by earth entirely (at energies relevant for this analysis). A subclass of the "Starting Track" is the "Upgoing Track", which is restricted to the southern hemisphere. Again, method and implementation are described in full detail in [79].

Dustyness and Borderness
The Dustyness and Borderness characterize events based on the relative charge deposited and the relative number of DOMs hit in the dust layer ($z \in [-137\,\mathrm{m}, -217\,\mathrm{m}]$) and the detector border (outer layer of DOMs) [53].

ProjectedQ
This reconstruction was developed in [53] and is inspired by the Hillas parameters [97], which is an established method for feature generation for ground-based IACTs (e.g. FACT [98], MAGIC [99], and HESS [100]) and widely used in imaging gamma-ray astronomy. The Hillas parameters are describing the two-dimensional light distribution observed in their camera induced by the Cherenkov radiation emitted from extended air showers in the atmosphere. Comparable events can be constructed for IceCube by projecting the three-dimensional charge distribution of the event onto a plane, which is defined perpendicular to the reconstructed track.

The charge distribution can be described by a two-dimensional Gaussian, which is quantified via the empirical covariance matrix. The Hillas parameters characterize the 1-σ ellipse of the Gaussian with e.g. its width, length or size.

Additional interesting observables used here were found to be the radius of a sphere containing 100 % of the hits normalized by the median distances to the center of gravity in the x-y-plane, the radius containing a percentage of the total charge, and the total charge contained in a sphere with that radius.

Double Pulse Selection
The previous step selects events with at least one double pulse waveform. Additional information about this selection can be useful for the final identification of tau neutrinos because for example the probability for a background event from single cascade background greatly depends on the highest recorded double pulse score in an event. Also the number of waveforms over the charge threshold of $Q_{\mathrm{tot,WF}} \geq 432\,\mathrm{PE}$, the number of survived waveforms, and mean, median and standard deviation of all classified waveforms and of all waveforms that survived the double pulse score cut are computed.

DelayTimeDPTrack
This parameter calculates the time difference between the geometrically expected arrival time from unscattered photons of the reconstructed track (SPEFit) and the actual start time of the waveforms classified as double pulse waveforms. If an event has more than one of those waveforms, the mean delay time is calculated.

Double Cascade Reconstruction

Tau neutrino events produce light at two distinct positions in the detector, the tau neutrino interaction vertex and the tau decay vertex (in case the tau decays into an electron/positron or hadrons). Events can be reconstructed with this double cascade hypothesis by extending the approach described for the Monopod reconstruction for single cascade events. This hypothesis only adds two new parameters to the fit resulting in nine fit parameters, the vertex of the neutrino interaction $\vec{x}_1$, the interaction time t_1, the direction (θ, ϕ) the energy deposited in the first cascade E_1 and finally the two additional parameters: the tau length L and the energy deposited as a result of the tau decay E_2. Both cascades are assumed to be connected by a tau lepton traveling at the speed of light, which inherits its direction from the incoming neutrino. This modification to Monopod was developed in [101]. The performance of the so-called Taupede reconstruction strongly depends on initial values for especially the tau length. A robust procedure for a more reliable reconstruction with Taupede is laid out in [102].

This reconstruction is not used in the event selection of this analysis but only for the description of final tau neutrino candidates due to several reasons: this analysis can be used as a cross-check for [102, 103] as long as analysis and event selection are mostly independent. This is important because although the direct double cascade reconstruction is based on a very compelling hypothesis, it relies heavily on precise knowledge of the ice model and numerical minimization in a multi-dimensional parameter space. Another reason is the computational cost of the double cascade reconstruction that was made mostly feasible for the analysis in [102, 103] by using the established event selection of high energy starting events as a starting point.

4.3.2 Feature Selection

After running all these reconstructions, each event is described by 700 observables in total. As a first step features that contain only one constant value or exclusively "Not a Number" (NaN) values are removed as they do not provide any information for separating signal from background events.

Additional information carried by features with a very high correlation to each other is marginal. To spot these features the Pearson correlation is calculated for each feature pair. In case multiple feature combinations have high correlations only one of the involved features is kept. The threshold to remove features is chosen to be $\rho \geq 0.98$. These two steps reduce the number of interesting features to 276.

An important action before using multivariate machine learning algorithms is also to make sure that experimental data is described sufficiently well by the simulations at hand. Usually, this is evaluated by comparing one-dimensional histograms of simulation to data either by eye or through a statistical test (e.g. Kolmogorov-Smirnov-Test). When using multivariate methods, this might not be sufficient as also correlations between different features can in principle be used to tell apart simulations from data. A different approach is taken in this work, where the multivariate model itself is used to measure the agreement between data and simulations and to identify features that are responsible for potential disagreements. This method was developed in collaboration with [53] (cf. [104]) and will be described in detail in the following paragraphs.

Multivariate Classification for Identification of Disagreements between Data and Simulations

The basic idea in this approach is to use the same algorithm for the identification of disagreements between data and simulations, and for the classification between

signal and background (which will be described in section 4.3.3). This procedure will ensure that if the algorithm is not able to identify significant mismatches even when it is specifically trained to do so. It will be safe to use it for the separation of signal from background without expecting artifacts from imperfect simulations.

In this work, a Random Forest is trained in a 10-fold cross validation. Simulations are treated as the signal (label: 1) and data is treated as the background (label: 0). The Random Forest will assign a classification score to each event: events that are identified as more "simulation-like" receive scores closer to 1, whereas "data-like" events will be assigned scores closer to 0. In the ideal case of a perfect simulation, the classifier will be guessing randomly and it will be impossible to distinguish these types of events so that both distributions will look similar being centered around values of 0.5. The Random Forest is trained with all remaining 276 features and with a ratio between data and simulations of $1 : 1$. To obtain this ratio a subset of simulation events is sampled according to their physical weights[4]. The top panel of figure 4.4 shows the resulting classification score distribution, which shows that a clear separation between data and simulations is possible at this stage. In general, the performance of a classifier can be measured and compared via the area under the curve (AUC) of the Receiver Operating Characteristic curve (ROC curve, both values are explained in detail in [105]). For this classifier, the AUC is 0.953 ± 0.001[5].

The feature importance can be used to identify features that have a significant contribution to this separation. These values measure how frequent and how much (based on an impurity measure) a feature contributes to the decision process (a more detailed explanation can be found in section A.2.2). All feature importances sum to 1. For a set of uninformative features, the distribution of feature importances is roughly expected to be Gaussian with a mean value of $1/N_{\text{features}}$. This can be exploited for the identification of outliers which will be features that allow the classifier to abuse disagreements between data and simulation, which can then be removed from the feature set so that it will not be considered in the signal background separation.

A feature is considered as an outlier if

$$f_i > k_{\text{Gaussian}}(\alpha) \cdot \text{MAD}, \tag{4.15}$$

[4] These events might be simulated with a different spectral index than expected from state of the art flux models. The events are weighted to the fluxes described in the beginning of this chapter.

[5] Although details can be found in [105], a bit of knowledge about the AUC is important to understand this section without further reading. The AUC is a performance measure for a classifier, that is independent of balance between classes in training and a specific score threshold. It ranges between values of 0 and 1, while 1 would describe a perfect classifier and 0.5 a classifier that is guessing randomly.

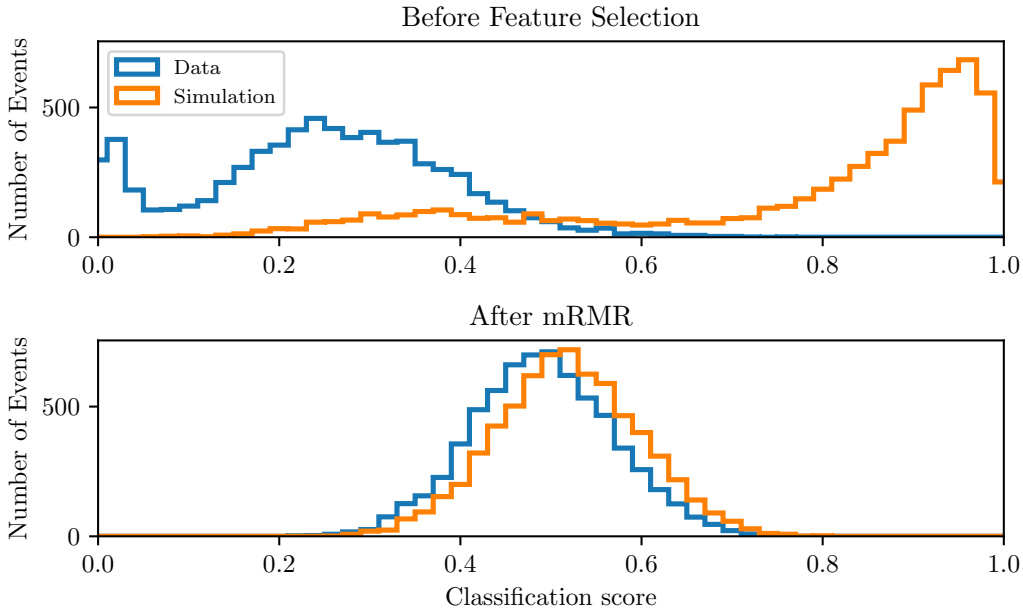


Figure 4.4: Classification score distribution for data and simulation. Top: distributions before the main part of the feature selection (276 features). Only information-less features and features with high correlations are removed before this classification. Both distributions can be clearly separated (AUC = 0.95). Bottom: distributions after removing features with the method explained in this section and then selecting a subset of the remaining features with the mRMR algorithm [106] (98 features). Both distributions get closer to the expected optimum for a purely guessing classifier (AUC = 0.6).

where f_i denotes the feature importance of feature i, k_{Gaussian} a scaling factor and the MAD = median($|f_i - \text{median}(f)|$) the Median Absolute Deviation [107]. The MAD is a measure of the variability of a sample that is robust to outliers. The scaling factor

$$k_{\text{Gaussian}}(\alpha) = 1.4826 \cdot \Phi^{-1}(1 - \alpha/2) \tag{4.16}$$

uses the inverse cumulative distribution function of a Gaussian Φ^{-1} to scale the MAD to cover a $1 - \alpha$ central interval (this outlier definition is adopted from [53]). In this work, a threshold of $\alpha = 0.1$ is chosen. Since the training is done with a 10-fold cross validation the information from the different folds can be used to mitigate some of the statistical fluctuations in the feature importances that result from choosing a specific realization of both data and simulations for each fold. A feature is only considered a significant source of mismatches if the condition 4.15 is satisfied in at least eight of the ten folds. The process of removing features is

done iteratively here with $n_{\text{it}} = 3$ because this yields better results in terms of a smaller AUC of the final feature set compared to e.g. choosing a different α. 80 features are removed within these iterations. The largest group of features (17) are reconstructed event times, that are not properly simulated by construction (IceCube simulation only aims to get relative timing right).

The final result of this method will be evaluated after reducing the dimensionality in another step of selecting features with the mRMR algorithm with the remaining 196 features, which will be described in the next paragraph.

Feature Selection with the mRMR Algorithm

After removing features that can contribute significantly to a separation between data and simulations due to disagreements either in that feature alone or by also exploiting correlations with other features, the remaining set of features can be considered for the classification task at hand. Now a subset of these features is selected to optimize the result of the classification between track-like and cascade-like events. This is done with the minimum redundancy maximum relevance (mRMR) algorithm [106]. The algorithm iteratively selects features by maximizing the relevance of feature x_j to the target class y while minimizing the redundancy between x_j and all other selected features $x_k \in S$, where S is the set of so far selected features. So at step j, each feature is scored by

$$q_j = I(x_j, y) - \frac{1}{|S|} \sum_{x_k \in S} I(x_j, x_k) \tag{4.17}$$

$$I(x, y) = -\frac{1}{2} \ln(1 - \rho(x, y)^2) \tag{4.18}$$

The function $I(x, y)$ is the mutual information between two features x and y based on a correlation coefficient $\rho(x, y)$. The implementation used in this work [108] uses the Pearson correlation for continuous features and Cramer's V for discrete features. The scoring value q_j is maximized at each step and the feature with index $\text{argmax}(q_j)$ is added to the set S. This can, in general, be repeated until the desired number of features is reached.

The choice of the number of features, that will be used for the classification can be quite arbitrary. Here, for each set of k features, a Random Forest is trained on a subset of the simulation (to reduce computational costs) to evaluate the performance at each step. This is done in a 5-fold cross validation and, finally, the feature set that maximizes the mean AUC is picked. Figure 4.5 shows the AUC and the Kuncheva index as a function of the number of selected features. The AUC is rising mostly monotonically with additional features, but at around 50 features it is running into

a plateau. The set maximizing the AUC consists of 96 features. There are two reasons for not selecting fewer features (e.g. 50 where the plateau is reached): On the one hand, having a fraction of features not containing much information for the classification does not have a meaningful impact on the performance of a Random Forest. The reason for that is at each node, where a cut is chosen, the Random Forest internally selects the best feature for the cut. The result of this thus might just be the inclusion of features, that will rarely or never be used. On the other hand, a small improvement in the AUC can yield a big improvement in background rejection for the finally chosen classification score cut.

As the result of the feature selection also depends on the specific realization of simulated events, different sequences of features could be selected for multiple simulation runs. The influence of this effect is characterized by the Kuncheva index here [109]. A mathematical description of the Kuncheva index is given in section A.4.1. The Kuncheva index calculates a stability value between 0 and 1 based on different runs of the feature selection on subsets of the simulations (in this case five sub runs each bootstrapping 20 % of all available simulation). It is based on a consistency index between two sets chosen such that it is increasing monotonically with the intersection between the two sets. It is constructed to be independent of the size of the sets and the total number of features and such that the index is constant for randomly drawn sets of the same size.

At this point, it is time to evaluate the effect of the method described in section 4.3.2 and the mRMR selection on the AUC for the classification between data and simulations. The result for this is shown in the bottom panel of Figure 4.4. The resulting AUC is 0.598 ± 0.005 and can now be compared to the previous result in the top panel. There is a significant improvement from the initial feature set, which can be seen especially by comparing the change in the distributions between the top and bottom panel. A question that might arise is, why the AUC still significantly differs from the optimal value of 0.5. One reason is, that the feature set at this point still contains features that show worrying disagreements when compared to data as one-dimensional histograms (c.f. Figure A.1). This shows a limitation of the method, that is designed to remove features with such levels of disagreement. The limitation most likely arises as a consequence of limited data statistics at this level (even when already using the burnsample of six years of data).

To solve this problem all 98 features are inspected individually and 20 features that still have visible regions of disagreement are removed manually. As a final modification to the feature set the two features were added describing the depth of the first hit and the horizontal layer of DOMs containing the DOM that recorded the highest charge of the event. These were identified as potentially important for the most problematic atmospheric muon background events by investigating

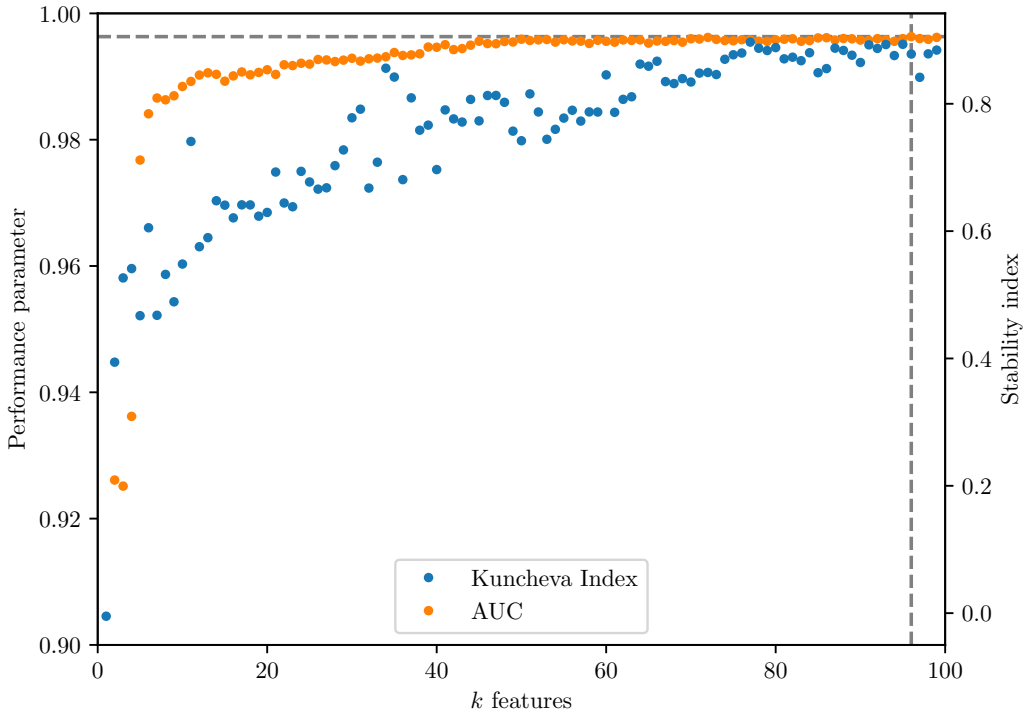


Figure 4.5: Performance of each set of k features selected with the mRMR algorithm. The performance is quantified by the AUC for a Random Forest trying to distinguish the signal from the background (in this case an AUC of 1 would be desirable). At the same time, the Kuncheva stability index is used to characterize the stability of each feature selection step.

high scoring background events in earlier iterations of the classification described in section 4.3.3.

The AUC for the classification between data and simulations is depicted in Figure 4.6 as a function of the number of features selected with the mRMR algorithm. The points where this function is shown are chosen as multiples of 20 except for two special cases: the set of 96 features selected with the mRMR algorithm and the final set of 78. The uncertainty is estimated from a 10-fold cross validation. This suggests that the manual removal of features has a stronger influence on the AUC than the removal of the same amount of random features.

Although the final AUC of 0.577 ± 0.010 is still not a perfect result, it is much smaller than the AUC at the beginning of the feature selection and also the AUC of 0.96, which can be achieved for a separation between signal and background. The

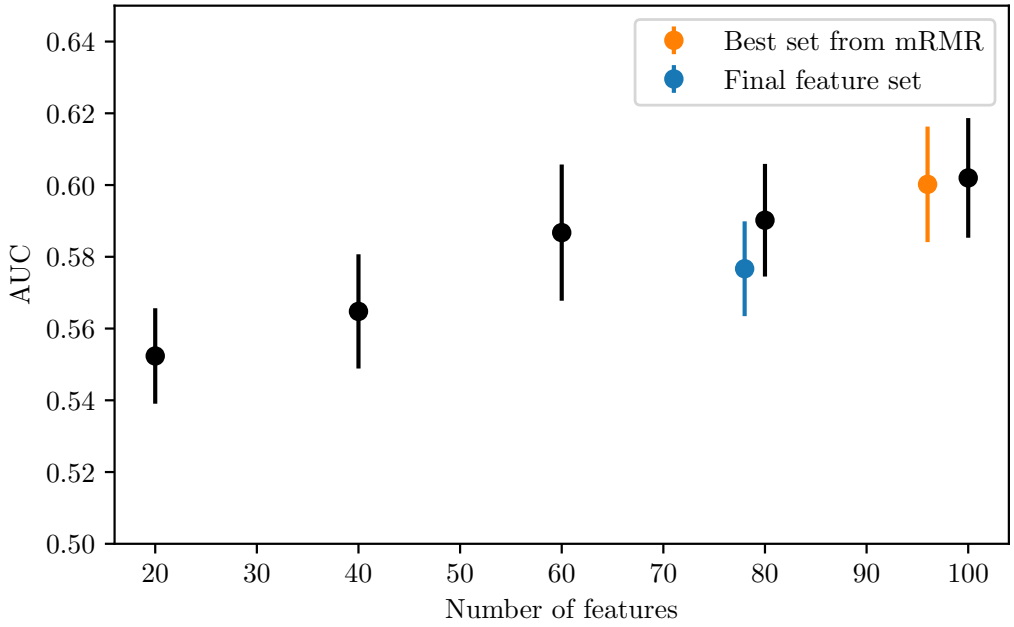


Figure 4.6: AUC for the classification between data and simulations as a function of the number of features. Two special cases are shown: the best set of features selected with the mRMR algorithm and the final feature set that is used in the following classification.

remaining disagreement might arise from the fact that the data at this point of the analysis is on a rather low level as no "quality cuts" are applied that focus on e.g. quality of reconstructions. Additionally, the comparison always inherently has to assume energy spectra for all relevant components, which are only known with limited precision. Other possible sources of disagreement are imperfect knowledge of systematic uncertainties and detector parameters.

Distributions for the highest-ranked features from the Random Forest described in section 4.3.3 are shown in Figure A.2. A way to inspect all removed and selected distributions is described in section A.4.2.

4.3.3 Classification

To obtain a final sample, that is dominated by events originating from charged current tau neutrino interactions another event classification has to be employed to distinguish cascade-like signal from track-like background events. Usually, in similar

analyses ([10], [102]) at this stage, a veto algorithm is applied, which can help to drastically reduce the amount of atmospheric muon background, thus limiting the analysis to starting events only and reducing the fiducial volume quite significantly as well. This work refrains from the explicit use of a veto region as a major step of the event selection to increase the fiducial volume and to also retain the possibility of selecting partially contained events. To lower the difficulty of the classification task with respect to atmospheric muons without an explicit veto, two measures are taken. First, only events where both the neutrino interaction vertex and the tau decay vertex are contained within the detector boundaries given by the outer layer of strings are considered to be signal events in the classification. For evaluation purposes, all partially contained events are considered as part of the signal again. Second, a cut is applied, removing all events where the reconstructed Monopod vertex is at least 60 m outside the outer boundary of the detector, which removes mis-reconstructed events and events far outside the detector.

The background that is supposed to be removed in this classification comes from track-like charged current muon neutrino and atmospheric muon events. In general atmospheric muons are simulated using extensive air shower simulations of cosmic ray primaries (with `CORSIKA`). This process is computationally expensive, so it is impossible with current implementations and resources to produce enough `CORSIKA` simulations to perform a reasonable background estimation for the atmospheric muon component. A solution to this problem can be `MuonGun` simulations, which provide a fast way of simulating single muons by using parameterizations of single muon fluxes estimated from previous `CORSIKA` simulations to reweight these events. The next paragraphs show how `CORSIKA` and `MuonGun` simulations can be combined in this analysis to describe the background of atmospheric muons.

The importance of different components of muon bundles with different multiplicities is depicted in Figure 4.7. It shows the two-dimensional distribution for the used `CORSIKA` simulations. The x-axis shows the classification score for the used model, referred to as “Cascade score” in this thesis, which will be introduced in a bit more detail later and the y-axis the relative energy of the most energetic muon in a bundle to the sum of all muon energies in that bundle. The bottom part of the plot shows the contribution of large multiplicity muon bundles from ultra high energy cosmic rays (UHECRs), which are distributed only at very low Cascade scores even though the effective livetime of the simulation in that regime is very high. The top part illustrates the most challenging type of background from cosmic ray air shower muons. These are muon bundles with one relatively high energy leading muon that reach the highest Cascade scores even though they are simulated scarcely due to just occurring rarely in those air showers. Fortunately, these can effectively be described by `MuonGun` simulations.

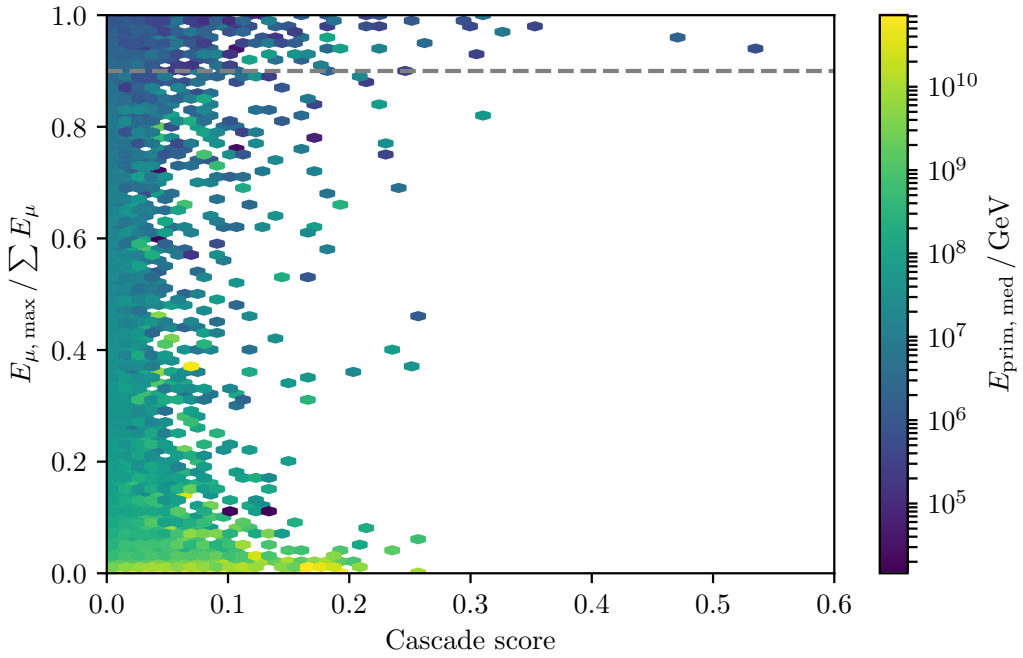


Figure 4.7: Two-dimensional distribution of the relative energy of the most energetic muon in a muon bundle to the total bundle energy and the Cascade score for the used full air shower `CORSIKA` simulation. The bins are colored based on the median energy of the primary nucleus from low energies (blue) to high energies (yellow).

This small amount of leading muons/single muons can be substituted by `MuonGun` simulations to effectively increase the number of simulated events of this type by a factor of 100. A comparison between the Cascade score distribution of the components that are exchanged here for a better description of the atmospheric muon background is shown in Figure 4.8. For lower Cascade scores both distributions agree reasonably well, which leads to the assumption that this combination is also suited well to describe data at higher Cascade scores. This comparison can be seen in Figure 4.9b and will be discussed in a later paragraph. The value of 90 % for the relative energy of the leading muon was chosen so that on the one hand both components agree well with each other in Cascade score and some additional observables and on the other hand the new combination describes experimental data as well as possible.

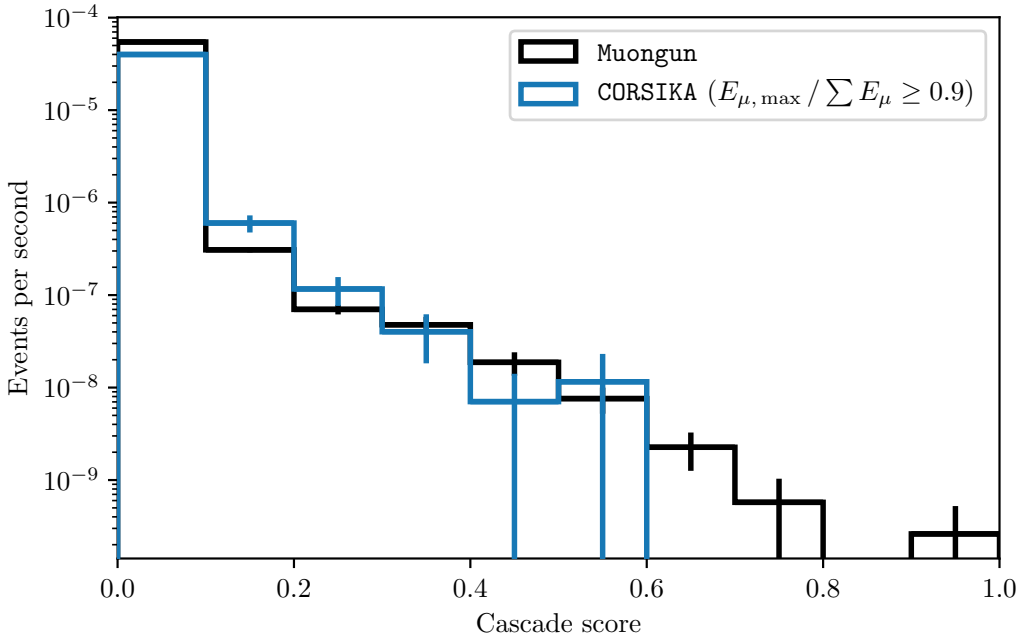


Figure 4.8: Comparison of the Cascade score distribution of `MuonGun` simulations (∼464 000 events) with events from `CORSIKA` simulations (∼4000 events), where the highest energy muon in a bundle holds at least 90 % of the total bundle energy.

Having the description of atmospheric muons with reasonable simulation statistics established, the Random Forest[6] can be trained. The training is conducted with a 10-fold cross validation. Considered as the signal are, as mentioned earlier, charged current tau neutrino events that have both a contained tau neutrino interaction vertex and a contained tau decay vertex. The background components included in the training are charged current muon neutrino events and atmospheric muon events (both from `CORSIKA` and `MuonGun` simulations).

The resulting Cascade score distribution is shown in Figure 4.9. The top panel shows the simulation broken down into individual components. The signal, depicted in green, is dominating the other components above Cascade scores of 0.65. The component of uncontained signal events has a similar magnitude as both background components in that region. The bottom panel can be used to compare data and simulations. The distribution of data events shows no artifacts or disagreements from the simulated distribution. The same comparison which was used during optimization

[6] The exact settings used can be found in Tab. A.1.

with only 10 % of the data can be seen in Figure A.6. The feature importances for the features used in the Random Forest are depicted in the Figures A.4 and A.5.

After training a classifier a threshold can be chosen above which events are kept for the final sample. This cut value is optimized here using the model rejection factor [110]

$$\mathrm{MRF} = \frac{\bar{\mu}_{90}}{n_s} \tag{4.19}$$

$$\bar{\mu}_{90} = \sum_{n_{\mathrm{obs}}=0}^{\infty} \mu_{90}(n_{\mathrm{obs}}, n_b) P_{n_b}(n_{\mathrm{obs}}), \tag{4.20}$$

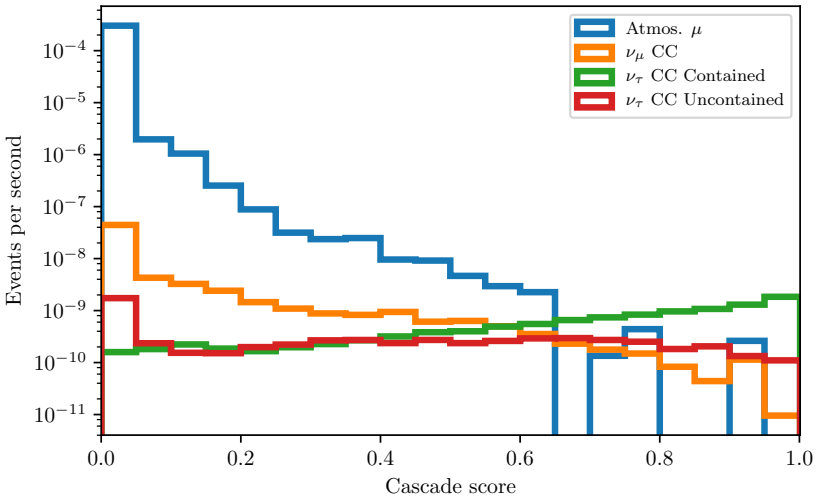
where $\bar{\mu}_{90}$ is the average upper limit over all possible observable event counts n_{obs} weighted with their respective Poisson probabilities of occurring $P_{n_b}(n_{\mathrm{obs}})$ under the assumption that there is no signal. This average upper limit is often considered as a sensitivity for analyses in the regime of small signals. The upper limits for this calculation are constructed with the unified approach suggested by Feldman and Cousins [111]. This method makes sure that physical restrictions on parameters (e.g. $\mu \geq 0$) are taken into account. The unified treatment also solves a problem, in which the choice between a one-sided interval (upper limit) or a two-sided interval based on the observed data can invalidate the confidence interval.

The model rejection factor (MRF) can be used to avoid a potential bias that can arise by choosing a final cut based on the observed data events. Instead, the MRF uses the expectation from simulations to find a threshold that results in the optimal upper limit (for a counting experiment) only based on the expectations and without considering the result of the final observation.

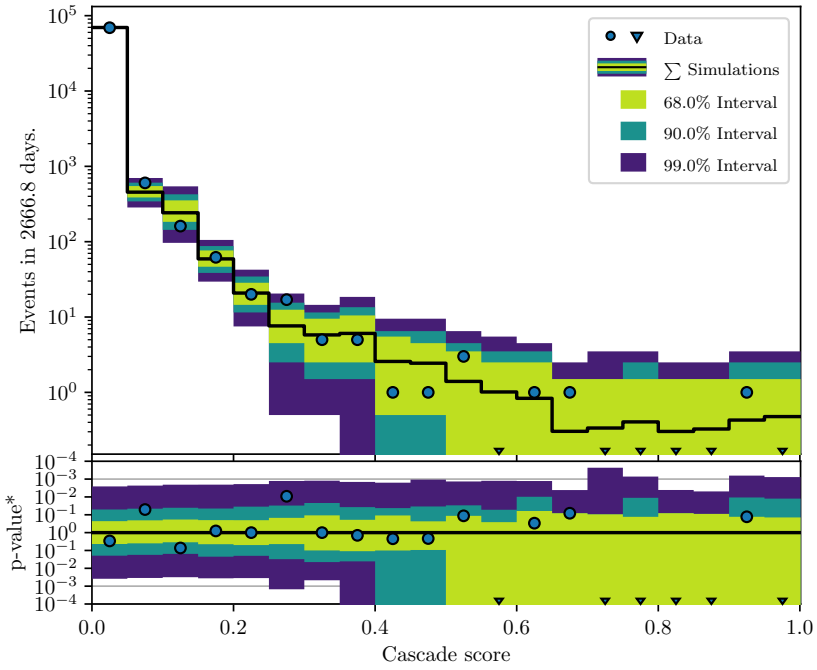
The MRF for this classification is shown in Figure 4.10 as a function of the Cascade score threshold. This curve suggests keeping all events with a Cascade score above 0.62, resulting in a MRF of 1.39 ± 0.20. This value can be interpreted by

$$\bar{\Phi}_{90} = \Phi_{\mathrm{Model}} \cdot \mathrm{MRF}, \tag{4.21}$$

in a sense that the MRF can be easily transformed to the average upper limit on the tau neutrino flux normalization $\bar{\Phi}_{90}$ for this kind of counting experiment.



(a) Cascade score distribution of different simulated components.



(b) Comparison between the Cascade score for simulations and data from more than seven years of data from May 2011 until December 2018.

Figure 4.9: Distributions of the Cascade score for different simulated components (top panel) and a comparison between data and simulations (bottom panel). The charged current tau neutrinos start to become the dominant component for Cascade scores above ~ 0.65. In that Cascade score region two events are found in 7.3 years of data, which will be described in detail in chapter 6.

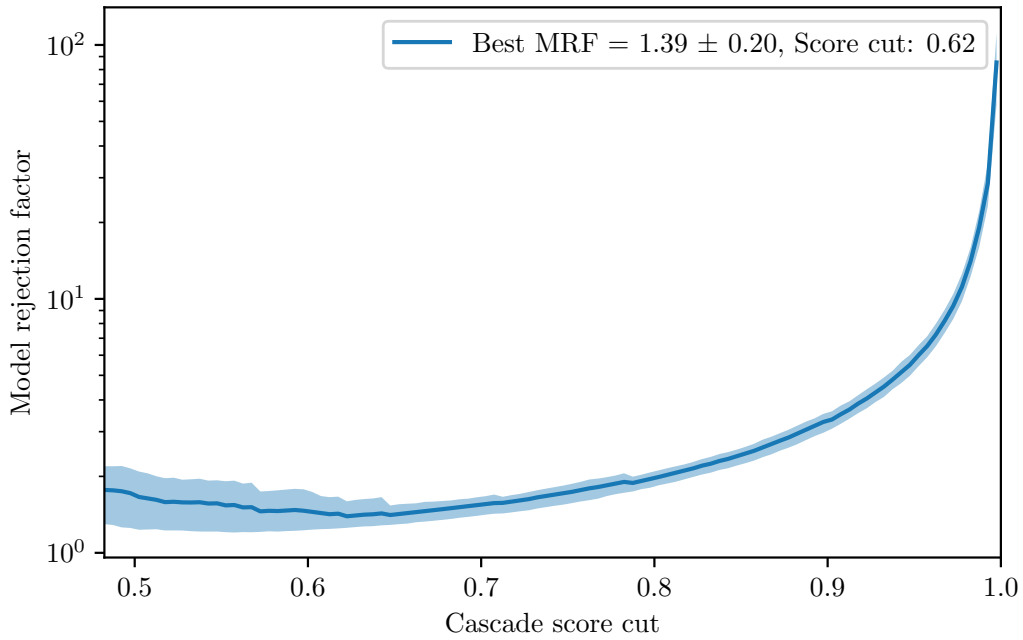


Figure 4.10: Model rejection factor as a function of the Cascade score cut (for the Cascade score distribution depicted in Fig. 4.9). The uncertainty is determined by a combination of fluctuations due to limited simulation statistics and variations from the different cross validation folds.

4.4 Final Sample

With the chosen Cascade score cut the remaining event expectations for the different simulated components are listed in Table 4.1 where all numbers are scaled to the analyzed amount of data of 7.3 years. Two data events survive all selection criteria in the analyzed time period (cf. Fig. 4.9b). These events are described in detail in section 6.1.

The background of the final sample is still dominated by track-like events with a majority of atmospheric muons. But, of all available simulations describing atmospheric muons only a total of 4 simulated events survive all selection criteria. This leads to large uncertainties in the background rate estimation. To solve this problem and to achieve a better description of the most important background component of this analysis additional simulations were produced. The available simulation was investigated to develop a scheme for a more efficient biased simulation specifically targeting this event selection.

For an incoming muon to produce both a double pulse and a signature that can not be easily identified as a track the muon has to undergo a large stochastic loss (most likely due to bremsstrahlung) near the edge of the detector, in which the muon loses most of its energy. This can result in a double pulse from the general losses of the muon followed by a large stochastic loss near a DOM. The large stochastic loss is important to produce a cascade-like signature. In case the muon is facing inwards the detector the energy of the incoming muon also has to be reduced heavily so that the track after the loss can not or can only barely be seen. This behavior is depicted in Figure 4.11a which characterizes the largest energy loss for single muon events with high Cascade scores (above 0.4). It shows both the distance of the largest energy loss to the detector boundary and the energy lost relative to the muon energy when it was generated outside of the detector.

This observation led to a strategy of focusing on events that satisfy the conditions:

$$d_{\text{edge}} \geq -150\,\text{m} \tag{4.22}$$

$$\frac{E_{\text{loss}}}{E_{\mu}} \geq 0.3, \tag{4.23}$$

which are chosen conservatively so that no interesting events are expected to be missed. This was verified by comparing the Cascade score distributions for old and new simulations between values of 0.4 and 0.62. To evaluate these conditions only knowledge about the energy losses and their positions is required. This helps to reduce the computation time of the simulation because events can be removed after muon propagation, but before photon propagation which is the computationally

most expensive step of the simulation. Due to these constraints over 80 % of the generated events do not have to be simulated fully.

A comparison between old and new simulation is depicted in Figure 4.11b. For the following analysis, both datasets are used as a weighted combination, which is shown as the green histogram. The increase in expected events from the new simulation is already included in Table 4.1.

Table 4.1: Event expectations and relative fractions at the final level for data and all simulated components in 7.3 years.

Component	Events in 7.3 years	Fraction
Data	2	-
ν_τ CC	2.14	66.5 %
ν_μ CC	0.23	7.1 %
ν_e CC + GR	0.13	4.0 %
$\sum_X \nu_X$ NC	0.05	1.6 %
Atmos. μ	0.67	20.8 %

4.4.1 Effective Area

The effective area is a common quantity to measure the quality of an event selection for both signal acceptance and background rejection. The effective area interprets the efficiency of the detector and the event selection as the surface area of a detector that would collect the same number of neutrinos with a theoretical efficiency of 100 %. The definition of the effective area in an energy bin ΔE is given by

$$A_{\text{eff}}(\Delta E) = A_{\text{gen}} \frac{N_{\text{sel}}(\Delta E)}{N_{\text{gen}}(\Delta E)} \tag{4.24}$$

where $N_{\text{gen}}(E)$ is the number of generated events as a function of the energy E, A_{gen} is the surface area these events were generated on (assumed to be energy independent here), and $N_{\text{sel}}(E)$ is the number of selected events. Figure 4.12 shows the average effective area for neutrinos and anti-neutrinos for the signal component compared to the neutrino background components from ν_μ CC and ν_e CC + GR events. Additionally, the event selection published in [10] is shown as a comparison. The signal effective area is increased over almost all energies, resulting in an improvement by a factor of 2.5 between 200 TeV and 1 PeV with more similar results outside of this energy region. For the ν_μ CC component a strong decrease in effective area, especially at the highest energies, can be observed.

4.4.2 Tau Neutrino Identification Efficiency

In the following paragraphs, the signal component of the final sample is investigated with respect to simulated tau neutrino/tau quantities like the neutrino energy as well as tau energy and tau decay length.

From both the neutrino energy distribution and the tau length distribution at different stages of the event selection (Figure 4.13), it can be seen that only a small fraction of tau neutrino events can be selected with the double pulse criterion. This results from the fact that only a small subset of events is eligible to produce a resolvable double pulse feature in one DOM due to the geometric requirements for such an event. First, both vertexes producing the respective pulses need to be close enough to a DOM and/or to deposit a sufficient amount of energy to produce two visible peaks in a single waveform. Second, the tau length (which is correlated linearly to the tau energy) needs to be large enough to allow for a significant separation of both pulses at the same time. This results in a very low efficiency especially at neutrino energies below 100 TeV or events with tau lengths below 10 m respectively.

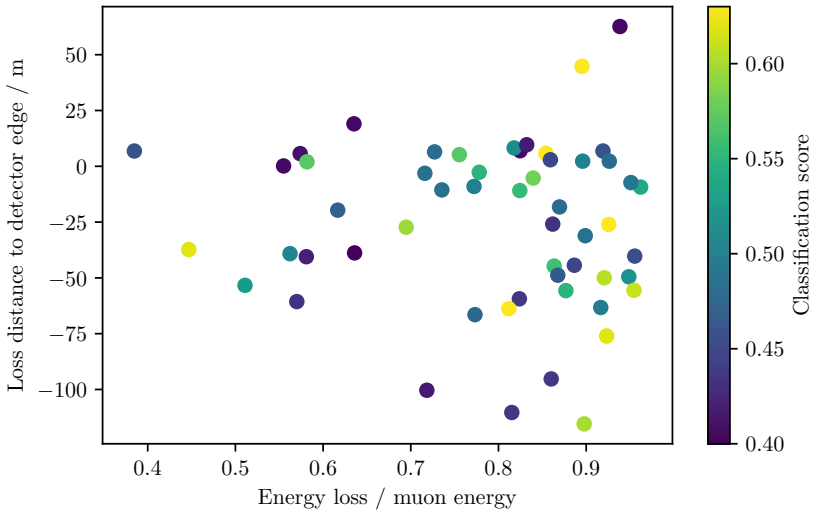
The shaded regions in Figure 4.13 indicate the central interval which includes 90 % of all events. The 90 % interval for the neutrino energy spans values between 126 TeV and 6310 TeV with a median energy of 890 TeV. The length interval is defined by a lower limit of 8.8 m and an upper limit of 93.6 m. The median length is 34.2 m (all of these values are based on the underlying assumption for the astrophysical spectrum with a spectral index of $\gamma = -2.13$).

Figure 4.14 shows the correlation between tau lengths and tau energies in the final sample. Without the effects of an event selection, the average tau length is given by

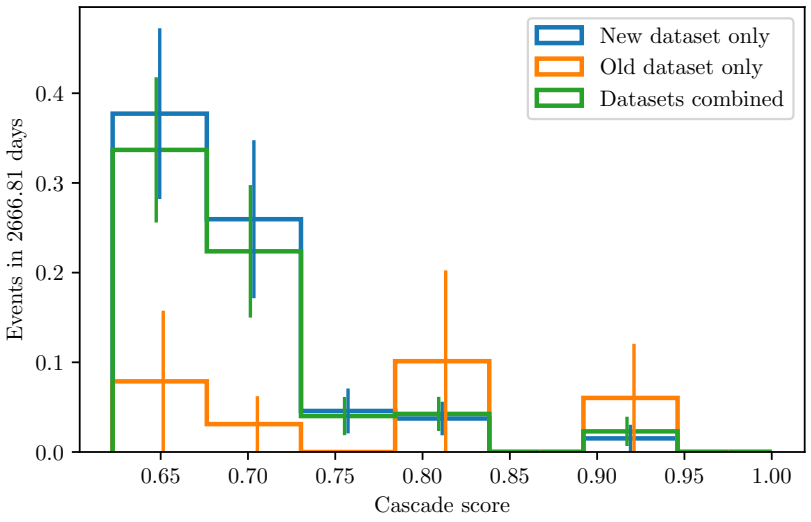
$$\ell_\tau \simeq 50\,\mathrm{m}\left(\frac{E_\tau}{1\,\mathrm{PeV}}\right) \tag{4.25}$$

and thus a linear function of the tau energy. This is depicted as a solid black line. The dashed line shows the average tau length present in the sample calculated for small energy bins. The behavior of the average length in the sample shows a clear selection bias. For tau energies below 1 PeV, the average tau length for selected events is systematically higher than the average expected tau length from eq. 4.25. This is a result of the increasing difficulty to resolve two energy depositions as a double pulse when the energy depositions are closer and closer together. Given the timing resolution of IceCube, these events will appear as single cascade events and thus removed during the event selection. Higher energy taus show the same behavior but in the opposite direction. In this energy region, the selected taus from the tau neutrino events are biased towards lower tau lengths because at some point

it gets more and more difficult to observe two cascades that are up to hundreds of meters apart within one DOM while still being able to reject waveforms with bumpy trailing edges from single cascade events. Another factor enhancing this is the fact, that signal events with large tau lengths are harder to select from a track-like background compared to the more single cascade-like events with shorter tau lengths.



(a) Scatter plot of simulated atmospheric single muon events for Cascade scores above 0.4. The y-axis shows the distance of the largest energy loss to the detector edge (positive values are outside of the detector boundary). The x-axis denotes the energy lost in the largest energy loss relative to the muon energy at its generation outside of the detector. Both quantities are true values from simulation. The color scale displays the Cascade score of each event and it is normalized in a way that all events surviving the event selection are colored in yellow.



(b) Comparison of the Cascade score distribution for the `MuonGun` simulation used for the event selection and the biased simulation as well as their weighted combination (binned in the same way as in section 5.4).

Figure 4.11: Characteristics of the most signal-like atmospheric muon background events (top) and a comparison between different `MuonGun` datasets (bottom).

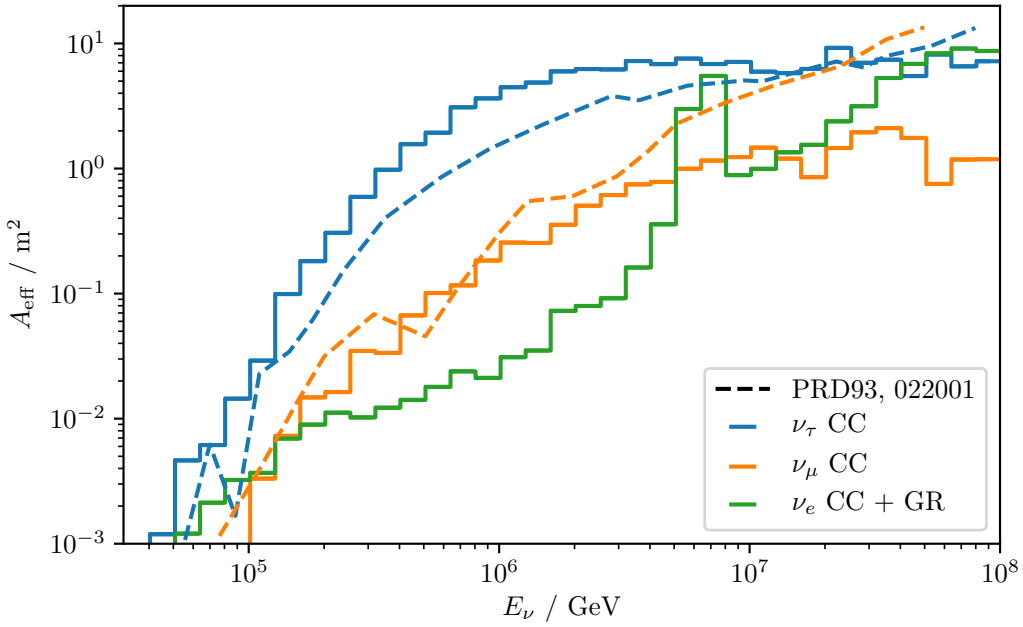
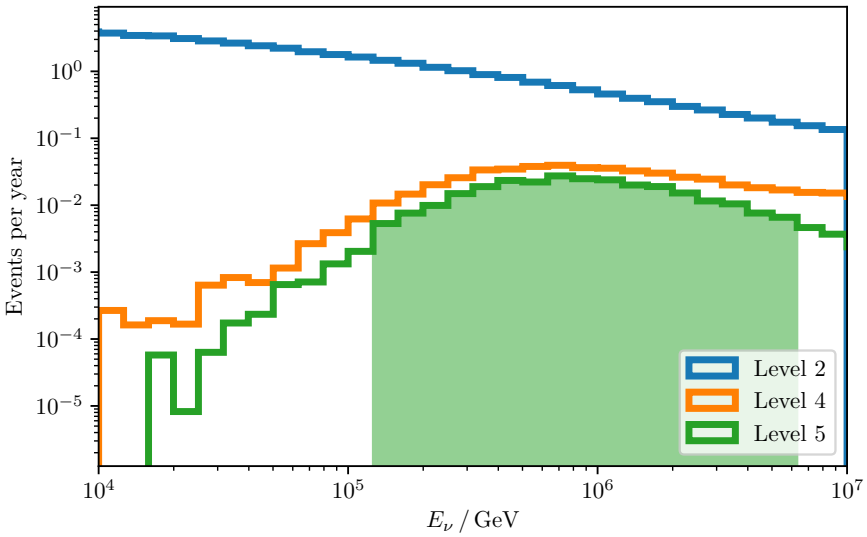
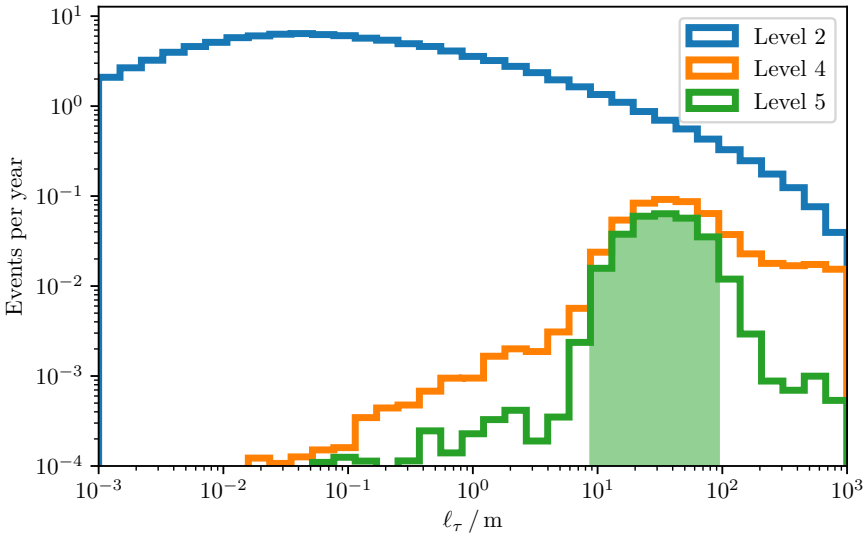


Figure 4.12: The effective area of the final event selection for charged current interactions of different neutrino flavors. The electron neutrino component also includes events from Glashow resonance interactions resulting in the peak near 6.3 PeV. The dashed lines show a comparison with the event selection from [10] for charged current tau neutrino and muon neutrino events.



(a) Neutrino energy distribution at different stages of the event selection.



(b) Tau decay length distribution at different stages of the event selection.

Figure 4.13: Comparison of the charged current tau neutrino component at different stages of the analysis. Level 2 is the starting point of the analysis and contains all triggered events. Level 4 contains all events that pass the criteria of the double pulse selection (section 4.2) and Level 5 is the final sample. The shaded regions indicate the central 90 % interval.

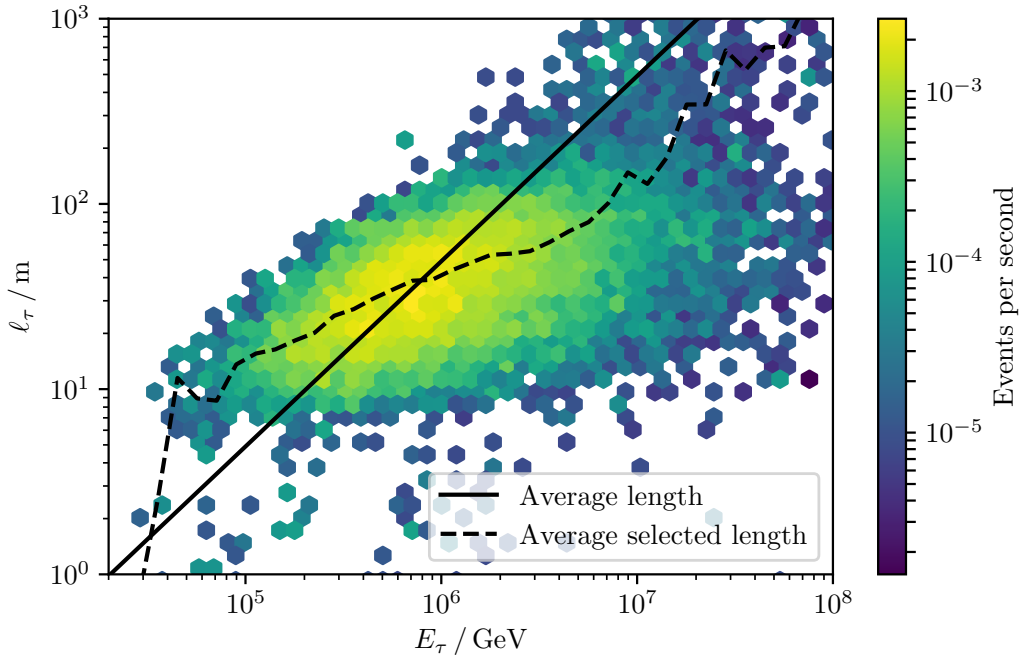


Figure 4.14: Two-dimensional distribution of tau length and tau energy. The average tau length as a function of tau energy is shown as a solid black line which shows a linear increase with the tau energy ($\ell_\tau \simeq \frac{50\,\mathrm{m}}{\mathrm{PeV}} E_\tau$). The dashed black line depicts the average tau length among the selected events in an energy bin.

5 Analysis Method

The astrophysical tau neutrino flux has been constrained by various IceCube analyses. The first group aims to measure the tau neutrino flux itself with the double pulse topology or via the double cascade channel using the direct reconstruction described in section 4.3.1 [10, 112, 113]. The second group aims to measure the astrophysical flavor ratio looking to combine different neutrino signatures from all flavors [5]. The approach used in [102] uses results from the double cascade reconstruction as additional observables in the analysis to increase the flavor ratio sensitivity to the tau neutrino component.

The goal in this thesis is to measure the astrophysical tau neutrino flux directly with the event selection based on the double pulse signature. In this chapter, the analysis methods utilized to constrain the astrophysical tau neutrino flux are described and their sensitivity based on the event sample, developed in chapter 4, is discussed. These methods have been developed on simulations only and they are applied to 7.5 years of IceCube data in chapter 6.

This chapter provides an introduction to the relevant statistical methods and concepts in sections 5.1 and 5.2. Next, the modeling of fluxes for neutrinos and atmospheric muons is described in section 5.3 and the observables used in the likelihood analysis get introduced in section 5.4. Section 5.5 quantifies the sensitivity of the analysis based on the analysis method and the utilized simulations. The considered systematic uncertainties and their incorporation into the analysis are described in section 5.6. Finally, section 5.7 outlines a method to assign each observed event a p-value, a probability of how likely it originates from the background probability density function (PDF).

5.1 Likelihood Description

The likelihood function provides information about how likely given, observed values are, based on a set of parameters θ. The likelihood function is defined as

$$\mathscr{L}(\theta|\vec{x}) = f(\vec{x}|\theta) \tag{5.1}$$

where $f(\vec{x}|\theta)$ is a known joint PDF of a realization of observed values $\vec{x}$ from the sample $\vec{X} = (X_1, \dots, X_n)$. Another thing to keep in mind is that $\mathscr{L}$ is not normalized over the parameter space, and thus is not a PDF. [114]

To formulate the likelihood for this analysis an ideal choice would be to use the set of observed events with all their observables (e.g. all observables described in chapter 4) to construct the joint PDF $f(\vec{x}|\theta)$. This would make use of all available information, but this construction would not be feasible. One reason for that is the limited amount of simulated events, that makes it difficult to describe unbinned high-dimensional PDFs. To simplify this, the observable space can be reduced to a small number of observables that are well suited for the analysis at hand (e.g. an energy estimator for the measurement of an energy spectrum). Another form of simplification is to discretize the observable space by binning the observables. The binning choice and the observables used in this analysis are discussed in section 5.4.

After discretizing the observable space the analysis transforms into a counting experiment for each bin that can be described by a Poisson distribution.

$$\mathscr{L}(\theta|\vec{n}) = f(\vec{n}|\vec{\mu}(\theta)) = \prod_{i=1}^{m} p(n_i|\mu_i(\theta)) \tag{5.2}$$

$$p(n_i|\mu_i(\theta)) = \frac{(\mu_i(\theta))^{n_i}}{n_i!} \exp(-\mu_i(\theta)) \tag{5.3}$$

where the product runs over all m bins and each bin i is described by a Poisson expectation $\mu_i(\theta)$ which depends on the parameters of interest. The observed event properties $\vec{x}$ are substituted by an event count in each bin $\vec{n}$. [6]

The so-called log-likelihood is commonly used instead of the likelihood since it is often easier to handle, especially concerning numerical stability:

$$\ln \mathscr{L}(\theta|\vec{n}) = \sum_{i=1}^{m} n_i \ln(\mu_i(\theta)) - \mu_i(\theta). \tag{5.4}$$

Terms that do not depend on θ can be omitted from the likelihood since they do not carry helpful information to favor a parameter set over another.

5.2 Point and Interval Estimation

The likelihood explained in the previous section can be used for inference on a parameter θ via point and interval estimation [114]. One method that can be used for point estimation is a maximum likelihood estimator. The maximum likelihood estimate $\hat{\theta}$ for a sample point $\vec{x}$ is the value for θ that maximizes the likelihood

$\mathscr{L}(\theta|\vec{x})$ (defined in eq. 5.1). This best fit result is an estimate of a point in the parameter space for the free parameters so that the model used in the likelihood describes the measured data $\vec{x}$ best.

However, the probability that the estimated value $\hat{\theta}$ is the true one $P(\hat{\theta} = \theta_t)$ is 0 (where θ_t is the true value of θ). This problem can be bypassed by using an interval estimator. This method is used to construct an interval in the parameter space, which has a non-zero probability of containing the true parameter θ_t.

Interval estimation is closely related to hypothesis tests. The overall goal of a hypothesis test is to infer something about an underlying parameter via two complementary hypotheses, where a hypothesis is a general statement about a parameter. A test features two hypotheses, a null hypothesis H_0 and an alternative hypothesis H_1. The null hypothesis states that the parameter $\theta \in \Theta_0$ is part of a subset of the full parameter space Θ, while the alternative hypothesis describes the complement $\theta \in \Theta_0^C$. [114]

To specify a test, a test statistic needs to be defined. A test statistic provides a measure to objectively quantify the outcome of a test. One possible test statistic is the likelihood ratio test statistic related to the maximum likelihood estimation. The test statistic for the null hypothesis $H_0 : \theta \in \Theta_0$ and the alternative hypothesis $H_1 : \theta \in \Theta_0^C$ is given by

$$\lambda(\vec{x}) = \frac{\sup_{\Theta_0} \mathscr{L}(\theta|\vec{x})}{\sup_{\Theta} \mathscr{L}(\theta|\vec{x})}. \tag{5.5}$$

The numerator maximizes the likelihood within the sub-parameter space allowed for the null hypothesis $\theta \in \Theta_0$ while the denominator maximizes the likelihood over the whole parameter space $\theta \in \Theta$. The test statistic value is a function of the data $\vec{x}$ that can be used to decide whether to reject or to not reject the null hypothesis. The bigger the value of this test statistic the larger the probability that the null hypothesis is true. [114]

The null hypothesis can be rejected with confidence level $1 - \alpha$ if

$$\lambda(\vec{x}) < k_\alpha \tag{5.6}$$

holds (this corresponds to a hypothesis test of size α). The critical value k_α has to be chosen in a way that $\sup_{\Theta_0} P(\lambda(\vec{x}) < k_\alpha) = \alpha$ applies. The critical value can either be computed if the test statistic $\lambda(\vec{x})$ follows a known distribution or by sampling the test statistic distribution from pseudo experiments. In the latter case, pseudo experiments are generated by injecting a parameter set θ_t into the statistical description and to calculate the likelihood ratio test statistic (eq. 5.5) from the drawn realization. This can be repeated for a number of pseudo experiments N to

obtain a good description of the test statistic distribution, which can, in turn, be used to empirically calculate k_α.

In addition to the statement for a hypothesis test at a certain confidence level if the H_0 is rejected or not, a p-value p can be constructed that quantifies the incompatibility of an observed test statistic value with the test statistic distribution of the null hypothesis:

$$p := \int_0^x \mathrm{d}x' f_{H_0}(\lambda(x')). \tag{5.7}$$

This computes the probability to find a test statistic value as or even more extreme than the observed one ($P(\lambda(x') \leq \lambda(x))$) under the assumption that the null hypothesis is true. Small p-values, therefore, indicate a low probability that the observed data originates from H_0.

An interval estimate for a parameter θ is a pair of functions, which satisfies the condition $\theta_L(\vec{x}) \leq \theta_U(\vec{x})\ \forall \vec{x}$. After observing $\vec{x}$ the inference $\theta_L(\vec{x}) \leq \theta \leq \theta_U(\vec{x})$ with a certain confidence level α is made. The statistical interpretation of such an interval estimate is that the interval covers the true parameter θ_t in a fraction α of independent measurements. An interval estimator in combination with a confidence level α is often referred to as a confidence interval.

Confidence intervals can be constructed from the inversion of a hypothesis test (in this case a likelihood ratio test). For each null hypothesis $H_0 : \theta = \theta_0$ with the acceptance region $\mathscr{A}(\theta_0)$ for a test with confidence level $1 - \alpha$ the confidence interval[1] $C(\vec{x})$ can be defined as

$$C(\vec{x}) = \{\theta_0 : \vec{x} \in \mathscr{A}(\theta_0)\}. \tag{5.8}$$

To ensure coverage of the confidence interval, the acceptance region has to fulfill the following condition for each tested null hypothesis $H_0 : \theta = \theta_0$:

$$P_{\theta_0}(\vec{X} \in \mathscr{A}(\theta_0)) \geq \alpha \tag{5.9}$$

where $\vec{X}$ is the random sample that yields realizations $\vec{x}$. [114]

In practice, this construction is performed via a likelihood ratio test for all physically relevant null hypotheses for a parameter of interest and returns the set of tests that did not reject its null hypothesis as the confidence interval. Physical bounds can be incorporated in this construction by restricting the parameter space in unphysical

[1] Note that this construction does not force the obtained set $C(\vec{x})$ to be an interval, so technically this would be a confidence set. The terminology confidence interval is more common in physics, so this technicality will be overlooked.

regions. For example, a Poisson expectation value μ could be restricted to positive values only.

$$H_0 : 0 \leq \mu = \mu_0 \tag{5.10}$$
$$H_1 : 0 \leq \mu \neq \mu_0 \tag{5.11}$$

These restrictions in the parameter space can lead to one-sided intervals even if performing a two-sided test (as suggested in eq. 5.10 - 5.11). This behavior, also used by Feldman and Cousins, was called a "unified approach" to unify the treatment of upper limits (one-sided intervals) and two-sided confidence intervals [111].

5.3 Modeling of Neutrino Fluxes and Atmospheric Muons

The description of both conventional and prompt atmospheric neutrino fluxes, as well as atmospheric muon fluxes, is handled in the same way as illustrated at the beginning of chapter 4. For the astrophysical component, an $E^{-2.13}$ energy spectrum was assumed from [50] (see Table 5.1 for the updated parameters). Since the development of the event selection, an updated result of the same analysis with an additional two years of IceCube data was released which is used in the following analysis as a baseline [4]. Overall, this is only a small change, because the event selection is robust against small changes in the spectral index (compare Tab. 4.1 and Tab. A.2). Additionally, in chapter 4 the astrophysical flavor ratio was assumed to be fixed at $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$. In this section, the tau neutrino component of the astrophysical flux will be measured, so it will be a free parameter in the likelihood.

Due to the low number of expected signal events in this analysis (~ 2), the astrophysical spectral index (which is assumed to be the same for all neutrino flavors) will be fixed throughout the analysis so that the tau neutrino flux normalization will be the only free parameter. Since the spectral index might have a big influence on the tau neutrino flux measurement, several different spectral indexes are compared in this analysis. The different assumptions for the astrophysical neutrino flux originate from IceCube measurements making use of different detection channels or combinations of them. The chosen analyses are presented in Table 5.1 and are selected to cover a large range of spectral indexes. The chosen measurements are the most recent results from the diffuse ν_μ analysis [4] with a very hard spectrum, the high energy starting event analysis (HESE) [7][2] with a softer spectrum and

[2] The current and most recent version of this analysis is used here, which is not yet published. Consider the presented values as preliminary as they still might change in the future. More information about the analysis in general can be found in [52], which describes the most recent publish iteration of the analysis.

the global fit analysis which combines different event samples for a global IceCube measurement of the astrophysical flux [5]. An equivalent to Table 4.1 for the flux measurements presented here can be found in section A.6.1.

Table 5.1: Measurements of the per flavor astrophysical neutrino flux from different IceCube analyses. The astrophysical flux is described as an unbroken power law with the form $\Phi_\nu(E) = \Phi_0 \cdot \left(\frac{E}{100\,\mathrm{TeV}}\right)^{-\gamma}$.

Analysis		Φ_0 at 100 TeV / $(10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1})$	γ
Diffuse ν_μ	[4]	$1.01 \pm^{0.26}_{0.23}$	$2.19 \pm^{0.10}_{0.10}$
Global Fit	[5]	$2.23 \pm^{0.37}_{0.40}$	$2.50 \pm^{0.09}_{0.09}$
HESE	[7]	$2.15 \pm^{0.77}_{0.75}$	$2.89 \pm^{0.31}_{0.28}$

5.4 Observables and Binning

Previous IceCube measurements targeting the astrophysical energy spectrum rely either on an unfolding approach [53] reconstructing the energy spectrum in several energy bins in a model-independent way or they use a binned Poisson likelihood modeling all components and then fitting their parameters to match data and simulation in a chosen discretized observable space [4, 5, 52, 6]. These analyses rely on large data samples of through-going tracks, cascades or starting events in general. They are making use of the unfolded neutrino energy or some form of reconstructed energy and/or the reconstructed direction (mostly the zenith angle) as these quantities hold the most information for a measurement of the energy spectrum and thus give the best sensitivity.

For this analysis, the number of expected events does not justify a measurement of the tau neutrino energy spectrum (especially while trying to constrain background components or detector systematics as nuisance parameters). Instead, this analysis focuses on a measurement of the astrophysical tau neutrino flux normalization, so the reconstructed energy and reconstructed zenith angle are less important observables. Since the data sample also only has a purity between 55 % and 69 %, depending on the astrophysical flux assumption observables that show a strong separation between signal and background increase the sensitivity of the flux normalization measurement. For the same reason, these observables are also later used to construct a p-value for an individual event to originate from the background distribution.

For this analysis, the observables are chosen to be the Double Pulse score and the Cascade score introduced in chapter 4. In the case of the Double Pulse score, only the highest score per event is used. These two observables are specifically constructed to distinguish signal and background. Both score values itself describe an important topological characteristic of the tau double pulse event signature. The binning is mostly chosen based on the limitations given the expected size of the data sample and the number of simulated events left after the event selection. Another consideration, especially for the binning in Cascade score, is that the score is not continuous, because it is limited by the number of trees used to train the Random Forest for the event selection. So the binning is also chosen in a way that each bin contains as close to the same amount of Random Forest decisions as possible. Thus, for the Cascade score seven equidistant between 0.6225 and 1.0, and for the Double Pulse score, only the two non-equidistant bins $[0.2, 0.8, 1.0]$ are used. The resulting two-dimensional observable distribution is presented in Fig. 5.1 for the signal component (top) and the sum of all background components (bottom) assuming an astrophysical $E^{-2.19}$ spectrum and a flavor ratio of $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$. For a better comparison, the one dimensional Cascade score distributions are shown for the two Double Pulse score slices in Figure 5.2.

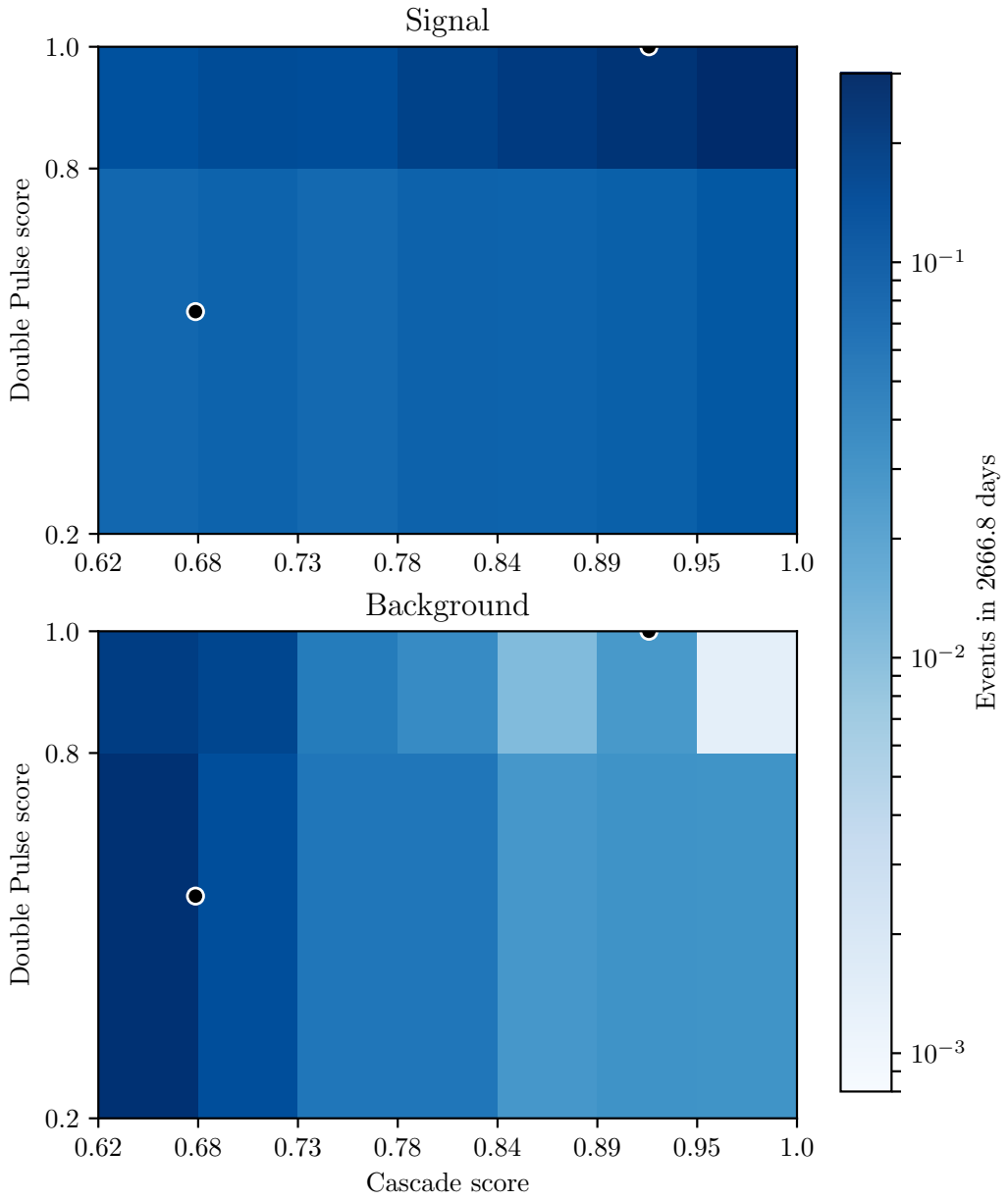


Figure 5.1: Binned 2D-distribution of the Double Pulse score and the Cascade score. The top panel shows the signal component and the bottom panel the sum of all considered background components. For the depicted distributions an astrophysical $E^{-2.19}$ spectrum is assumed. Color encodes the expected number of events in 7.3 years of data. The two observed events are indicated by black dots and are described in more detail in chapter 6.

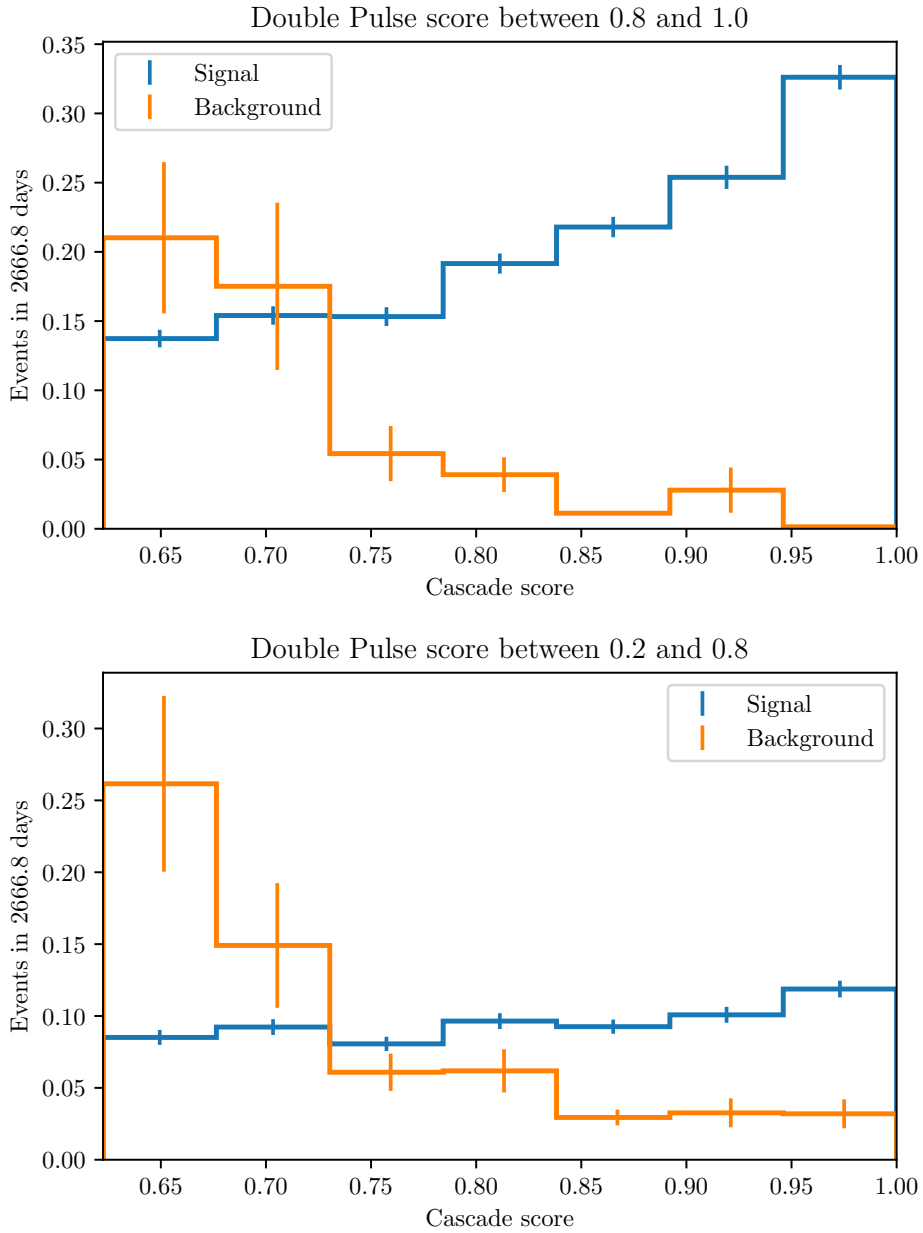


Figure 5.2: Slices of the 2D observable distribution in Double Pulse score. The top panel shows the Cascade score distribution for Double Pulse scores between 0.8 and 1.0 and the bottom panel for Double Pulse scores between 0.2 and 0.8. The error bars are shifted away from the bin centers slightly for better visibility only.

5.5 Likelihood Fit and Sensitivity

The goal of this work is to constrain or measure the astrophysical tau neutrino flux normalization. This is realized using the binned Poisson log-likelihood introduced in eq. 5.4. Since the tau neutrino flux normalization is the only free parameter in the fit, the log-likelihood takes the following form

$$\ln \mathscr{L}(\lambda|\vec{n}) = \sum_{i=1}^{m} n_i \ln(\mu_{b,i} + \lambda\mu_{s,i}) - (\mu_{b,i} + \lambda\mu_{s,i}) \tag{5.12}$$

$$\lambda = \frac{\Phi_{\nu_\tau}}{\Phi_{\nu_\tau,\mathrm{ref}}} = \frac{\mu_s}{\mu_{s,\mathrm{ref}}}. \tag{5.13}$$

The parameter λ is a scaling factor from reference value $\Phi_{\nu_\tau,\mathrm{ref}}$ to the tau neutrino flux Φ_{ν_τ} or respectively from the reference signal expectation $\mu_{s,\mathrm{ref}}$ to the number of signal events μ_s, where the reference is chosen to be the baseline expectation for the analysis assuming the flavor ratio of $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$.

To construct confidence intervals with the method described in section 5.2 different values for the signal expectation μ_s are injected with a step size of 0.05. For each value of μ_s, $N = 100\,000$ pseudo experiments are generated. A pseudo experiment is done by drawing background events from the background distribution with the Poisson expectation μ_b, and signal events from the signal distribution with the Poisson expectation μ_s. The obtained sample is then used to perform the likelihood fit. Each pseudo experiment is used to calculate the log-likelihood ratio

$$\mathrm{TS} = -\ln\left(\frac{\mathscr{L}(\mu_{s,\mathrm{inj}})}{\mathscr{L}(\hat{\mu}_s)}\right) = \ln(\mathscr{L}(\hat{\mu}_s)) - \ln(\mathscr{L}(\mu_{s,\mathrm{inj}})), \tag{5.14}$$

which can be used as a test statistic (TS) for the two-sided test with the hypotheses

$$H_0 : 0 \leq \mu_s = \mu_{s,\mathrm{inj}} \tag{5.15}$$

$$H_1 : 0 \leq \mu_s \neq \mu_{s,\mathrm{inj}}. \tag{5.16}$$

Here, $\hat{\mu}_s$ is the value of parameter μ_s that maximizes the likelihood described in eq. 5.12 while being restricted to the physical region of positive values $\mu_s \in [0, \infty)$. The obtained test statistic distribution can be used to obtain the acceptance region for each $\mu_{s,\mathrm{inj}}$ and thus also a critical value k_α.

This procedure is illustrated in Fig. 5.3. For each scanned value of μ_s the binned test statistic distribution is encoded by color. The red line shows the critical values for each test statistic distribution at a 90 % confidence level. Additionally, two likelihood scans for exemplarily selected background-only pseudo experiments (where

for the dashed line zero events and for the solid line three events were sampled) are presented. These likelihood scans are used in combination with the critical values to construct confidence intervals on the parameter μ_s as introduced in section 5.2.

Another property of the likelihood ratio can be observed in Figure 5.3, which shows the asymptotic behavior described by Wilks' theorem [115]. The theorem states that as the sample size n becomes large, the test statistic $-2\ln(\lambda(\vec{x}))$ (with λ as it is defined in eq. 5.5) for a nested model follows a χ^2 distribution with k degrees of freedom, where $k = \dim(\Theta) - \dim(\Theta_0)$. The critical value for larger values of μ_s ($\mu_s \gtrsim 10$) in Figure 5.3 is asymptotically approaching the critical value corresponding to the 90 % confidence level of a χ^2 distribution with one degree of freedom of 2.71 after multiplying it with the missing factor of 2 to match the test statistic definition in Wilks' theorem. Overall this shows two things, expected asymptotic behavior and the invalidity of Wilks' theorem in the most important region of the parameter space for this analysis, which further motivates the determination of the test statistic distributions via pseudo experiments. The thin lines visible in the test statistic distributions are a consequence of the discreteness of the Poisson distribution and the injection of small event counts.

The sensitivity to the astrophysical tau neutrino flux is here defined as the average upper limit that is obtained by background only pseudo experiments (by injecting no true signal, $\Phi_{\nu_\tau} = 0\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}$ or $\mu_s = 0$). To estimate the sensitivity to different astrophysical spectra for each spectrum, the observable histograms are constructed by reweighting all events to the corresponding astrophysical flux. The average upper limit, as well as the distribution of upper limits from 1000 pseudo experiments, is shown in Figure 5.4. To test the robustness of the sensitivity to different spectra for each fit pseudo experiments are injected for different astrophysical spectra, which show stable results in all cases. The resulting sensitivities are presented in Table 5.2 along with their ratio to the corresponding flux normalization Φ_0. The uncertainty is the standard deviation of the distribution of the upper limits. The relative average limit to the $E^{-2.89}$ spectrum is weaker compared to the harder spectra because the soft spectrum reduces the number of expected signal events significantly (cf. Tab. A.2 - A.4). The results based on the analyzed data sample are presented in section 6.2.

5.5.1 Differential Upper Limit

The model-dependent constraints described in the previous paragraphs are not able to indicate which energy region contributes most to the achieved constraint. A differential limit can be used to characterize the energy dependence of the analysis.

Table 5.2: Sensitivities to the tau neutrino flux normalization for different astrophysical spectra. The third column shows the ratio between the sensitivity and the corresponding astrophysical flux normalization.

Astrophysical spectrum	$\bar{\Phi}^{\mathrm{UL},90}_{\nu_\tau}$ at 100 TeV / $(10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1})$	$\bar{\Phi}^{\mathrm{UL},90}_{\nu_\tau}/\Phi_0$
$E^{-2.19}$	1.24 ± 0.62	1.23
$E^{-2.50}$	2.32 ± 1.13	1.04
$E^{-2.89}$	4.38 ± 2.05	2.04

For a simple counting experiment and a null observation the differential limit can be calculated via

$$\Phi^{\mathrm{UL},90}_{\nu_\tau}(E_\nu) = \frac{N_{90}}{4\pi\, E_\nu\, T \log(10)\, A_{\mathrm{eff}}(E_\nu)} \tag{5.17}$$

with the 90 % confidence level upper limit for zero observed events in the case of negligible background N_{90}, the total observation time T and the energy-dependent effective area A_{eff} [116]. This equation can be derived from

$$\frac{\mathrm{d}N}{\mathrm{d}t\mathrm{d}\Omega} = \int_{E_0/\omega}^{E_0\omega} A_{\mathrm{eff}}(E) \frac{\mathrm{d}\Phi}{\mathrm{d}E\mathrm{d}A} \quad \mathrm{d}E \tag{5.18}$$

which describes how a differential flux can be translated to an event rate. Equation 5.17 can be recovered by inserting the hypothesis of an E^{-1} signal box flux defined over one decade in energy ($\omega = \sqrt{10}$) and by assuming a constant effective area within this decade of energy. This method is used in several other IceCube analyses (e.g. [10]), partly with generalizations to not be limited to either null observations or counting experiments [20].

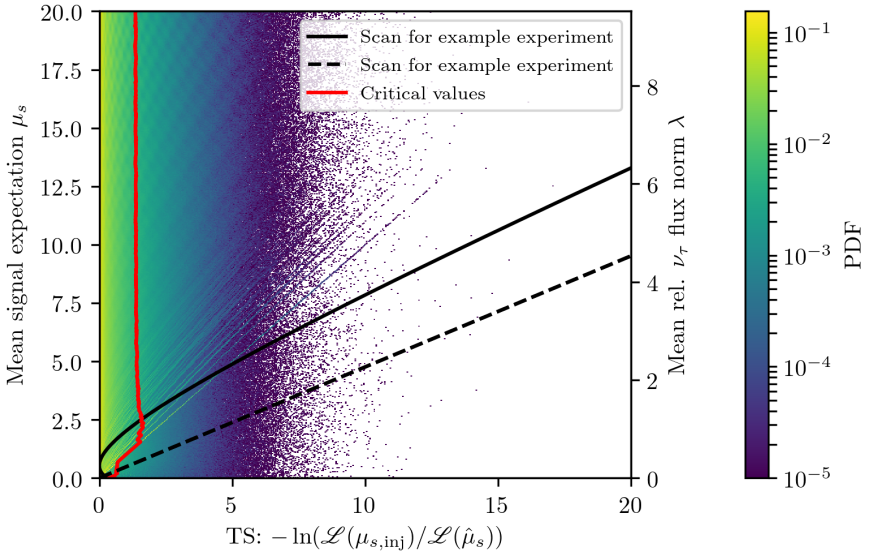
This approach can be combined with the likelihood fit described in this section by replacing μ_s with μ_{diff} in eq. 5.12 based on the underlying flux

$$\Phi_{\mathrm{diff}}(E, E_c) = \begin{cases} \Phi_0 E^{-\gamma} & \forall E \in [E_c/\omega, E_c\omega] \\ 0 & \text{otherwise} \end{cases} \tag{5.19}$$

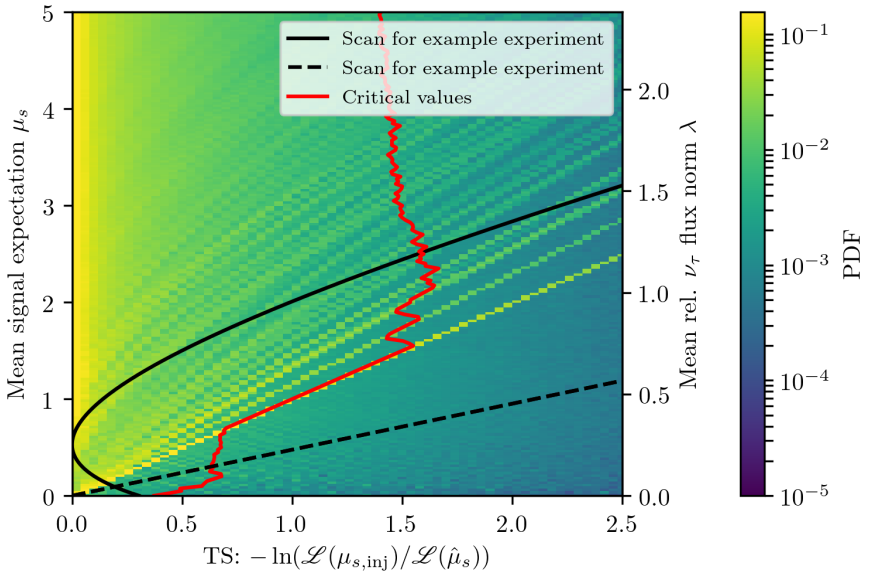
where E_c is the central energy of the box flux and ω defines the width of the interval the flux as $[E_c/\omega, E_c\omega]$. For this method, in general, the choice of both γ and ω is arbitrary, but to achieve comparable results especially with [10] in this work $\gamma = 1$ and $\omega = \sqrt{10}$ are chosen.

The most stringent limit on the tau neutrino flux will be placed in the case of a null observation. This case can be used to illustrate the differential upper limit for this

analysis and to compare it to astrophysical flux predictions. The differential limit assuming a null observation is depicted in Fig. 5.5. It is calculated for four energies E_c per decade between 100 TeV and 10 PeV. The figure shows a comparison to different models (exemplary selected) predicting the astrophysical neutrino flux from chocked jets and low luminosity gamma-ray bursts [117], low luminosity AGNs [118] and a multicomponent model [119]. The multicomponent model includes atmospheric backgrounds, a galactic contribution from cosmic ray interaction and two different types of extra-galactic components from pp interactions (in e.g. starburst galaxies) and $p\gamma$ interactions from AGN or tidal disruption events. The differential limit highlights the most sensitive energy region for this analysis from 300 TeV to 1 PeV, while also showing that in case of a null observation some of the chosen models can be excluded in certain energy regions even by means of this less model-dependent method. The differential limit presented in this section is the limit that would be obtained for a null observation. This case is used here to illustrate the sensitivity of the differential limit, whereas the observed differential limit is presented in section 6.2.2.

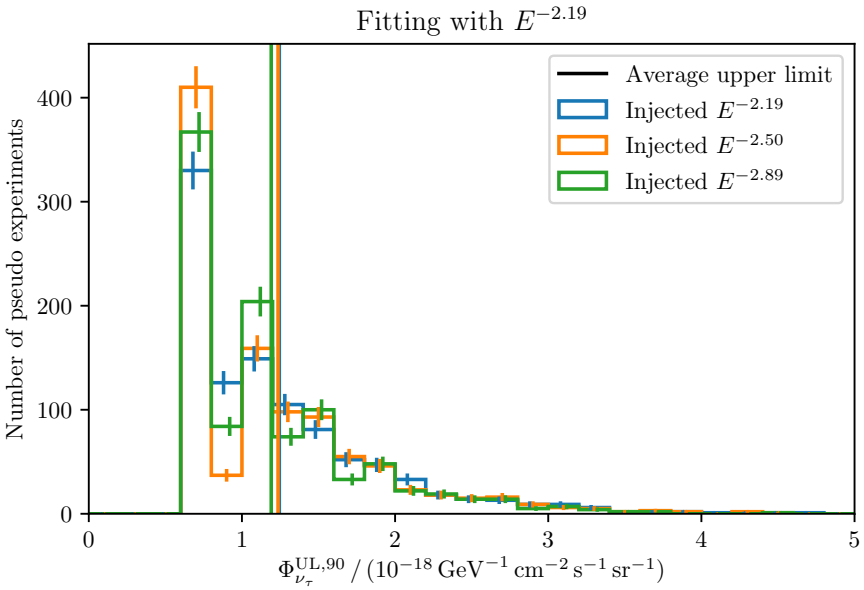


(a) Test statistic distributions for $\mu_{s,\mathrm{inj}} \in [0, 20]$.

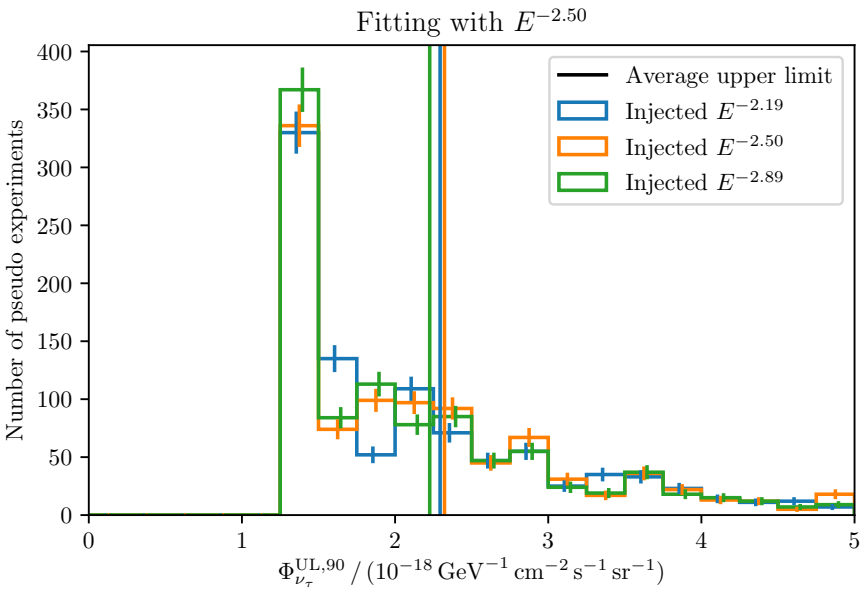


(b) Test statistic distributions for $\mu_{s,\mathrm{inj}} \in [0, 5]$.

Figure 5.3: Test statistic distributions for various values of $\mu_{s,\mathrm{inj}}$. The black lines show scans for exemplary pseudo experiments and the red line shows the critical value obtained for each test at a 90 % confidence level. If the observed test statistic value is higher than the critical value the null hypothesis will be rejected and the corresponding value of $\mu_{s,\mathrm{inj}}$ will not be part of the constructed confidence interval.



(a) Distribution of upper limits $\Phi_{\nu_\tau}^{\mathrm{UL},90}$ fitting with an astrophysical $E^{-2.19}$ spectrum.



(b) Distribution of upper limits $\Phi_{\nu_\tau}^{\mathrm{UL},90}$ fitting with an astrophysical $E^{-2.50}$ spectrum.

Figure 5.4: 90 % upper limit distributions for 1000 pseudo experiments. Each plot shows a different astrophysical flux assumption used in the likelihood fit. The different colors indicate the injection of pseudo experiments based on different astrophysical assumptions. For this estimation of the sensitivity the injected signal is set to $\Phi_{\nu_\tau} = 0\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}$. The error bars have been slightly shifted from the bin centers for better visibility only.

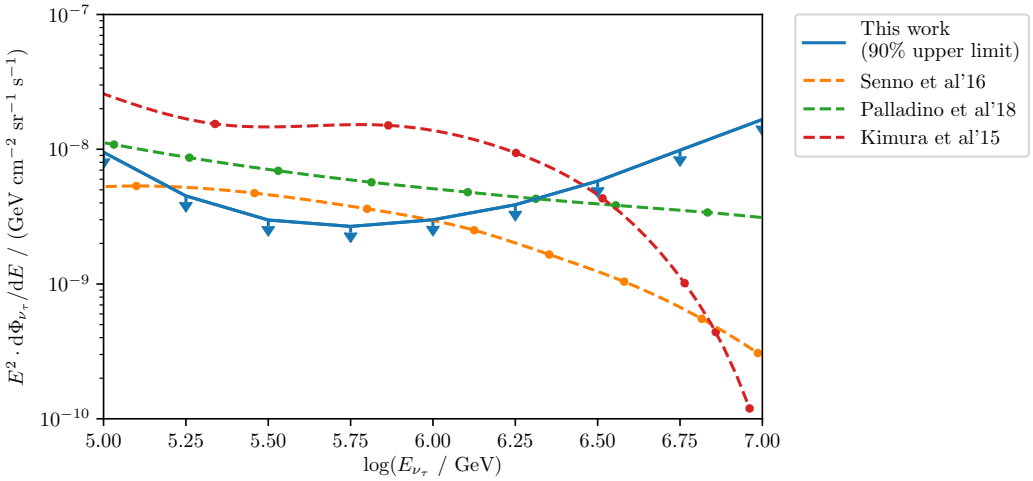


Figure 5.5: Sensitivity to the differential upper limit on the tau neutrino flux between 100 TeV and 10 PeV (Differential upper limit assuming a null observation). Models for astrophysical neutrino fluxes from choked jets and low luminosity gamma-ray bursts [117] (orange), low luminosity AGNs [118] (red) and from a multicomponent model [119] (green) are compared.

5.6 Systematic Uncertainties

Systematic uncertainties in this analysis can be divided into two relevant subgroups: uncertainties arising from imperfectly known detector properties (detector systematics), and uncertainties associated with limited knowledge about modeled flux and particle interactions (flux systematics). The category of detector systematics may change the detection threshold or the shape of observable distributions obtained after the analysis. On the other hand, flux systematics only change the expectation of certain types of events based on the change in e.g. incident particle flux. An overview of the considered sources of systematic uncertainties is given in Tab. 5.3. Each source of systematic uncertainties is described in more detail in the following subsections.

Due to the small number of expected events in the data sample, it is not beneficial to incorporate the number of sources of systematic uncertainties presented here into the likelihood fit as nuisance parameters (as this is usually done in analyses with larger data samples (e.g. [4, 5, 7])). This treatment would allow the consideration of correlations between systematic parameters, which are *a priori* unknown (for the most parameter combinations, for example, the correlation between uncertainties of the astrophysical spectral index and the astrophysical normalization are known from their simultaneous measurements). In this analysis, the influence of systematic variations is investigated by individually reporting their impact on measured quantities (e.g. the upper and lower bound of the confidence intervals). A conservative, worst-case estimate for the total influence can be obtained by adding the relative changes due to the systematic uncertainties in quadrature, which effectively approximates all sources of systematic uncertainties to be uncorrelated. This section gives an overview of the sources of systematic uncertainties. Their influence on the tau neutrino flux measurement is presented in section 6.2.1.

DOM Efficiency

The first considered parameter is the so-called DOM efficiency, which combines several systematic effects. The DOM efficiency defines how many photons are detected as a function of the deposited energy in the detector. Thus it combines both effects that influence the production and the actual detection of photons.

The detection uncertainty is mainly described by the quantum efficiency of the photomultiplier (PMT), the transmittance of the DOMs glass and gel and the area that is shadowed by various cables outside the DOM [120, 73]. The systematic uncertainty for the PMT efficiency calibration is 7.7 % [121]. The dominant contributions arise from the laser beam energy calibration uncertainty and other systematic effects

Table 5.3: Overview of the considered sources of systematic uncertainties. They are grouped into detector systematics and flux systematics. Astrophysical flux normalizations are given in units of $10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}$. The atmospheric (single) muon flux is described by a parametrization of `CORSIKA` simulations, which is described in more detail at the beginning of chapter 4.

Type	Description	Nominal value / Fiducial model	Syst. variation
Detector	DOM efficiency	0.99	$\pm 10\,\%$
Detector	bulk ice scattering	Spice 3.2	$^{+10\,\%}_{-7\,\%}$
Detector	bulk ice absorption	Spice 3.2	$^{+10\,\%}_{-7\,\%}$
Detector	hole ice	$p_2 = 0$	$p_2 \in [-3, -1, 1]$
Flux	$\Phi_{\nu,\mathrm{conv}}$	HKKMS06 [83]	$\pm 30\,\%$
Flux	$\Phi_{\nu,\mathrm{prompt}}$	BERSS [84]	cf. [84]
Flux	$\Phi_{\nu,\mathrm{astro}}$	1.01, 2.23, 2.15	$^{+0.26}_{-0.23}$, $^{+0.37}_{-0.40}$, $^{+0.77}_{-0.75}$
Flux	γ_{astro}	2.19, 2.50, 2.89	± 0.1, ± 0.09, $^{+0.31}_{-0.28}$
Flux	Φ_{μ}	CORSIKA param.	$\pm 15\,\%$

originating from the measurement setup. The overall uncertainty of the mentioned effects is estimated to be $\pm 10\,\%$ [95, 73].

Besides the baseline simulation with a DOM efficiency of 0.99, simulated datasets with DOM efficiencies 0.9 and 1.08 are available for neutrino simulation.

Overall the DOM efficiency shifts the energy threshold of an analysis and also has an impact on the energy reconstruction. For this analysis though the most relevant effect is the energy threshold shift, since energy information is not directly included in the likelihood fit.

Optical Properties of the South Pole Ice

Another source of systematic uncertainty are the optical properties of the South Pole ice that IceCube is embedded in. The uncertainty arises from the model uncertainties on scattering and absorption coefficients. The current way of simulating systematic uncertainties (by reproducing a whole dataset with different simulation parameters) does not provide a feasible solution to treat all of the depth-dependent scattering and absorption coefficients (c.f. [69]) as individual parameters. Instead, the estimated global uncertainty on the coefficients of $10\,\%$ is used to model the imperfect knowledge of the optical parameter of the ice.

Three different systematic variations are simulated to describe the effects of this global uncertainty: increased scattering by $10\,\%$, increased absorption by $10\,\%$, and simultaneously decreased absorption and scattering by $7\,\%$ each. These choices were made to roughly model the 1σ contour of the fit performed in [69]. This also takes into account that the uncertainties of the scattering coefficients and the absorption coefficients are strongly correlated.

Hole Ice

The term hole ice is used to describe the refrozen column of ice that is created during the deployment of a string. The drill holes have an approximate diameter of 60 cm. The refrozen ice contains air bubbles which increase the amount of scattering within the bubble column compared to the remaining bulk ice [122]. The change in scattering effectively changes the angular acceptance of the deployed IceCube DOM, which is illustrated in Figure 5.6. The angular sensitivity of the DOM within the hole ice column is parametrized with

$$0.34(1+1.5\cos\eta-0.5\cos^3\eta)+p_1\cos\eta(\cos^2\eta-1)^3+p_2\exp(10(\cos\eta-1.2)) \quad (5.20)$$

where η is the incident photon zenith angle. Thus $\cos\eta = -1$ describes vertically down-going photons, and $\cos\eta = 1$ up-going photons. The angular acceptance is modeled to describe flasher data [69] and the shown variations of the parameters p_1 and p_2 roughly cover the 1σ region of the obtained best-fit model. Overall, compared to a bare DOM, hole ice reduces the sensitivity to frontal photons (up-going), but also allows more photons to hit the PMT due to the additional scattering that would otherwise be missed, therefore, making the angular acceptance a bit more isotropic.

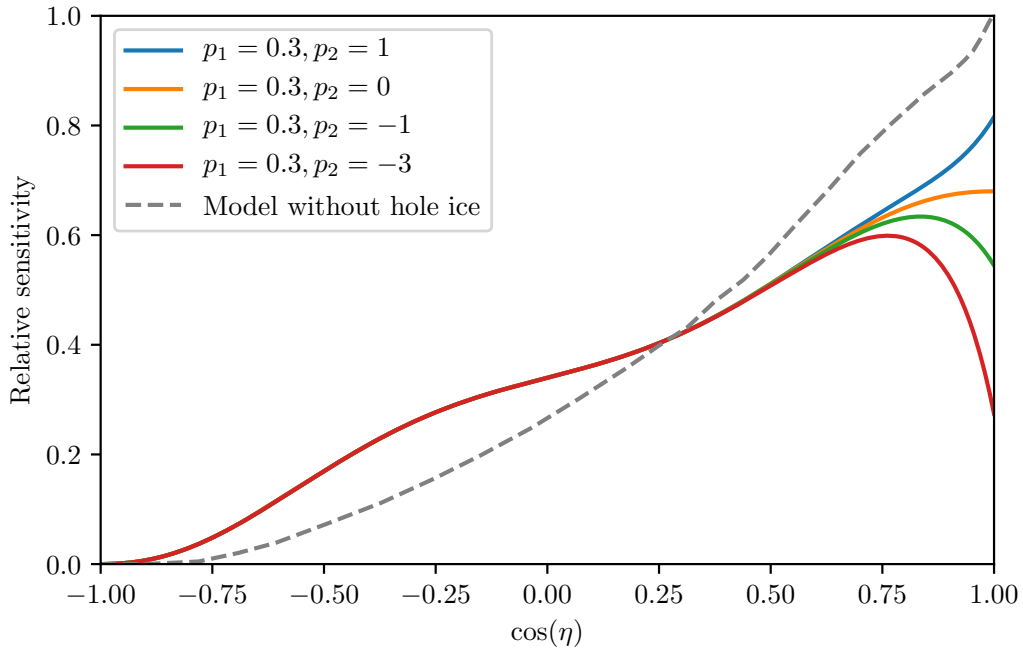


Figure 5.6: Relative sensitivity of the IceCube DOM as a function of the incident photon angle to the PMT. Shown are different values for the hole ice parameters p_1 and p_2 and the parametrization in eq. 5.20. The dashed gray line describes the angular sensitivity of a bare IceCube DOM.

Atmospheric Neutrinos

The uncertainties on the conventional atmospheric neutrino flux models are dominated by the uncertainty on cosmic ray parameterizations and imperfect knowledge about hadronic interaction models (which is estimated here to be roughly $\pm 30\,\%$ for both electron and muon neutrinos) [123].

For the calculation of prompt atmospheric neutrino fluxes, the uncertainties related to the primary cosmic ray flux are the same as for conventional atmospheric neutrinos. The used prompt flux model supplies their estimation of uncertainties due to variations of QCD parameters responsible for the production of heavy quarks, which are used in this analysis [84].

Astrophysical Neutrinos

The models for astrophysical neutrino fluxes are taken from IceCube measurements (as presented in Table 5.1). The measurement uncertainties on the astrophysical normalization and the spectral index for the astrophysical background of electron and muon neutrinos are also treated as a systematic uncertainty.

Atmospheric Muons

The uncertainty on the flux of atmospheric muons has the same origin as the uncertainty for atmospheric neutrinos. Their overall uncertainty is estimated to be $\pm 15\,\%$ [123].

5.7 Event-wise Classification

In case any events are observed in the final sample, although they passed all selection criteria of the event selection they are still not guaranteed to be signal events. An interesting question is, to ask for each event, how likely it is for an event to be a background event. This can be answered with the observed p-value given the likelihood and the test statistic

$$\mathscr{L}(\lambda|i) = \frac{P_{b,i} + \lambda P_{s,i}}{1+\lambda} \tag{5.21}$$

$$\mathrm{TS} = -\ln\left(\frac{\mathscr{L}(\lambda=0)}{\mathscr{L}(\hat{\lambda})}\right) \tag{5.22}$$

where P_b and P_s are probability distribution functions describing the simulated events in some observable space, λ is bound in the interval $[0,1]$ and is used to fit the "signalness" in the corresponding observable bin i. The null hypothesis here is given by the background-only hypothesis $\lambda = 0$, and for the alternative hypothesis λ is free to float within its boundaries. Since this procedure only inspects one event at a time, the fit result will be either $\hat{\lambda} = 0$ or $\hat{\lambda} = 1$ depending on whether the signal

or background PDF is larger in the respective observable bin. A background only test statistic can be produced by sampling pseudo experiments from the background distribution, which will then be used to calculate the p-value that the observed event originates from the background PDF.

Here, the same observables, as used for the measurement of the tau neutrino flux normalization (they are described in detail in section 5.4), are utilized. The observables are specifically chosen here because they are designed to achieve a good separation between signal and background (cf. chapter 4). The binning is also kept the same as described in section 5.4. For this type of statistical question, the expected event rates are not relevant, but the distributions have to be PDFs. To achieve this, they are normalized such that

$$P_{s,i} = \frac{\mu_{s,i}}{\sum_i^m \mu_{s,i} \cdot \mathrm{vol}_i}, \tag{5.23}$$

where vol_i is the volume of bin i.

The test statistic distribution for an $E^{-2.19}$ spectrum is shown exemplarily in Figure 5.7 (distributions for the other spectra are presented in Figures A.10 and A.11). The plot shows the background only TS distribution, which is used for the analysis as well as a signal TS distribution used to estimate the sensitivity of the p-value calculation to events originating from the signal PDF. The sensitivity is measured in terms of the median p-value of events sampled from the signal PDF that is 0.099, 0.094 and 0.114 for the three analyzed astrophysical spectra $E^{-2.19}$, $E^{-2.50}$ and $E^{-2.89}$ respectively. The distributions are discrete as each bin can only produce one TS value and all bins with a background PDF higher than the signal PDF give a TS value of 0 due to fitting $\hat{\lambda} = 0$.

The observed event candidates and their p-values are discussed in section 6.1. The method, developed in this work, was also applied to calculate p-values for tau neutrino candidates in [112].

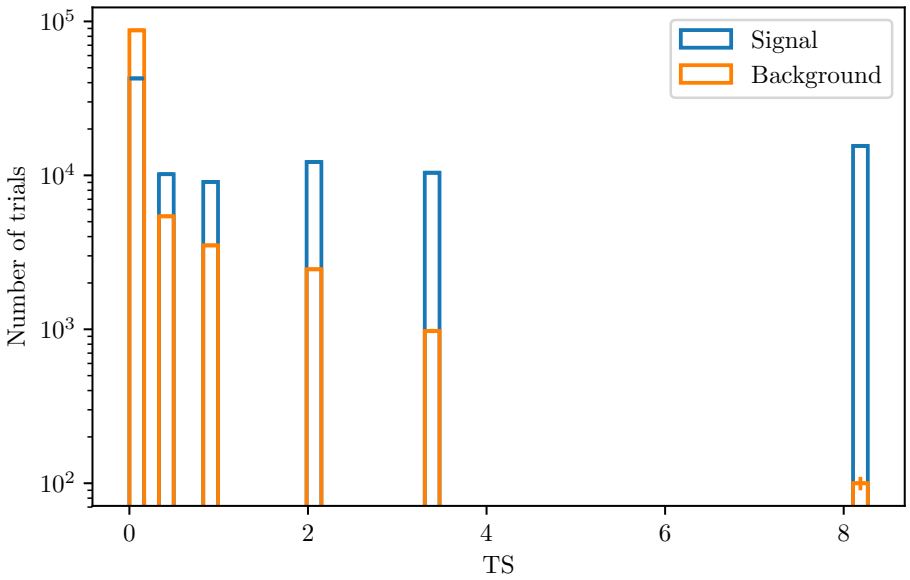


Figure 5.7: Test statistic distribution for the p-value calculation for events sampled from the background PDF (orange) and events sampled from the signal PDF (blue) assuming an astrophysical $E^{-2.19}$ spectrum.

6 Results

This chapter will describe the results obtained with the methods described in chapter 5 on IceCube data recorded between May 2011 and December 2018. All analysis methods were either developed on simulations only or on simulations with 10 % of the available data to check for agreement between data and simulations. In section 6.1 the events found in the final sample are presented and characterized in detail. Section 6.2 describes the result of the likelihood fit and the resulting constraints on the astrophysical tau neutrino flux. Lastly, in sections 6.3 and 6.4 the obtained results are compared to other IceCube analyses and its astrophysical implications are discussed.

6.1 Data Sample

After applying the event selection to the full sample of 7.5 years of IceCube data, two tau neutrino candidate events were found, as already indicated in Figures 4.9b and 5.1 which already show the full data sample. The events were found in the 2014 and 2015 season respectively[1].

6.1.1 2014 Event

An event view of the tau neutrino candidate observed in the 2014 season is shown in Figure 6.1. Three waveforms, depicted in Figure 6.2, passed the double pulse selection criterion. The waveforms are found on three neighboring DOMs, where the one observed in the central DOM has the highest possible Double Pulse score of 1.0 (orange). For the other two waveforms, Double Pulse scores of 0.45 (blue) and 0.81 (green) are found respectively. The Cascade score to determine if this event has a cascade-like topology is 0.92, which makes it the most interesting event in the sample.

This event is a starting event, which is supported by the fact that it passes the VHESelfVeto (which is described in sec. 4.3). The observed charge for this event

[1] A pole season starts in spring and ends in spring in the following year.

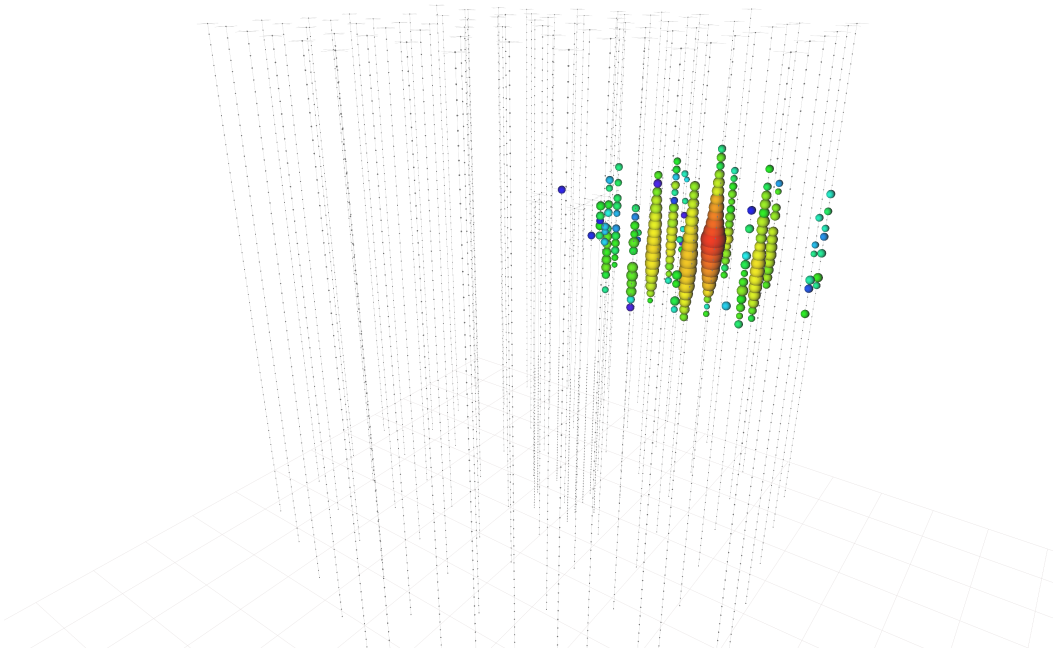


Figure 6.1: Event recorded on April 22nd 2015 (`RunID`: 126283, `EventID`: 47286594). Deposited charge $Q_{\text{tot}} = 9001\,\text{PE}$, reconstructed (EM-equivalent deposited) energy $E_{\text{Monopod}} = 91\,\text{TeV}$.

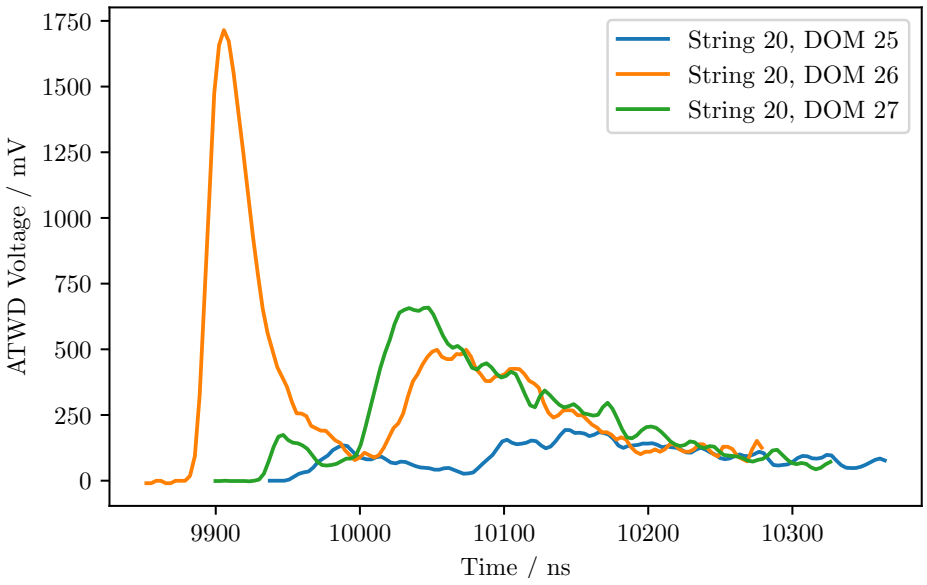


Figure 6.2: Double Pulse waveforms of the event recorded on April 22nd 2015 (`RunID`: 126283, `EventID`: 47286594). The corresponding Double Pulse scores are 0.45 (blue), 1.0 (orange) and 0.81 (green).

is $Q_{\text{tot}} = 9001\,\text{PE}$ making it a suitable candidate for the VHESelfVeto as it is optimized for events above 6000 PE. This makes it a very likely neutrino candidate as opposed to the possibility of being an atmospheric muon. The position in which this event occurred is just above the dust layer which could, combined with the reconstructed zenith of 49°, result in obscured light from a low energetic outgoing muon track originating from a charged current muon neutrino interaction. But, the background rejection for this event selection results only in a fraction of 7 % to 9 % for charged current muon neutrino events. The p-value based on the observed Double Pulse score and Cascade score for this event of being a background event is

$$p_{\gamma=2.19} = 0.035$$
$$p_{\gamma=2.50} = 0.035$$
$$p_{\gamma=2.89} = 0.031$$

for the respective astrophysical spectra.

The same event was found in [103] but with a completely different approach used in the event selection and for the identification of tau neutrinos as double cascades. The approach used there applies a direct double cascade reconstruction (described in section 4.3.1) aiming to reconstruct the sought after event topology of a tau neutrino interaction and a subsequent tau decay (excluding the tau decaying into a muon). Their approach yields a reconstructed energy of $E_1 = 9\,\text{TeV}$ for the first cascade, $E_2 = 80\,\text{TeV}$ for the second cascade, and a tau length of $\ell_\tau = 17\,\text{m}$.

Although the reconstructed energy of the second cascade is only a lower bound for the tau energy, the expected average tau length for an 80 TeV tau is roughly 4 m. This translates to a probability of an 80 TeV tau to propagate more than 17 m of 1.3 %. The effect of observing smoking gun signatures first for which observed quantities might be far away from the mean, and thus the probability of such an event happening might be small, is expected. This mostly results from a selection bias, which is introduced due to small experimental sensitivities in certain regions of the observable space. The smaller the tau length, the smaller the probability gets to either be able to correctly reconstruct the tau length or to be able to identify a double pulse. This results in a much larger sensitivity for outliers to higher tau lengths in the length distribution, and finally in larger average tau lengths selected by the analysis compared to the expected average tau length from the exponential decay distribution. This behavior is illustrated for this analysis in Figure 4.14.

Another interesting consideration is that both this double pulse analysis and the double cascade analysis [103] have different types of dominant backgrounds. For this analysis, the dominant backgrounds are still track-like events, namely atmospheric

muons and charged current muon neutrino events, as they can easily produce double pulses and can be hard to distinguish from cascades when producing high energetic losses close to the detector border. For the double cascade analysis, the dominant background stems from mis-reconstructed single cascade events.

Here, a study about the overlap of both analyses was conducted and is presented in Table A.5. Both the rather small overlap in the number of signal events and the even smaller overlap in the background components suggests, that the combination of both tau search strategies would be beneficial for the overall sensitivity to the tau neutrino flux and the identification of individual tau neutrino candidates.

6.1.2 2015 Event

An event view of the event observed in the 2015 season is presented in Figure 6.3. For this event, one waveform, shown in Figure 6.4, passed the double pulse selection criterion with a Double Pulse score of 0.565. The Cascade score is found to be 0.675, being much closer to the selection threshold than the other event and thus in a background dominated region of the observable space. Therefore, the resulting p-value for this event is

$$p_{\gamma=2.19} = 1.0$$
$$p_{\gamma=2.50} = 1.0$$
$$p_{\gamma=2.89} = 1.0$$

which is independent of the underlying astrophysical spectrum.

The event has a lower total charge of $Q_{\text{tot}} = 7165\,\text{PE}$, but a higher reconstructed energy $E_{\text{Monopod}} = 119\,\text{TeV}$. This lightly up-going event (with a reconstructed zenith angle of roughly 96°) is a starting event as well, also passing the VHESelfVeto. Despite starting with a cascade contained inside the detector volume, the event also shows horizontal development afterward indicating a muon leaving the detector volume.

These observations make the double pulse much more likely to be originating from Cherenkov light from the muon and its secondaries hitting the double pulse DOM before the light from the hadronic cascade due to the speed of the muon being larger than the speed of light in the medium.

Recently an analysis with a similar goal was conducted [112] using straight cuts for the event selection utilizing almost the same period of IceCube data. This comparable event selection found three events, containing the events found in this analysis as well. The third event found there turned out to originate from two coincident muons.

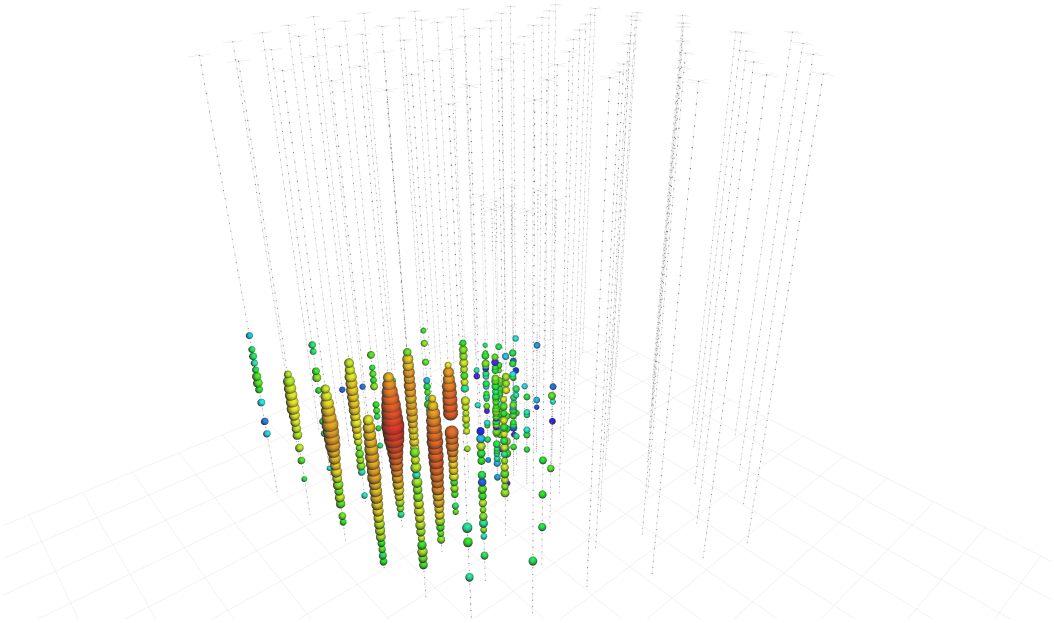


Figure 6.3: Event recorded on April 2nd 2016 (`RunID`: 127762, `EventID`: 66008498). Deposited charge $Q_{\text{tot}} = 7165\,\text{PE}$, reconstructed (EM-equivalent deposited) energy $E_{\text{Monopod}} = 119\,\text{TeV}$.

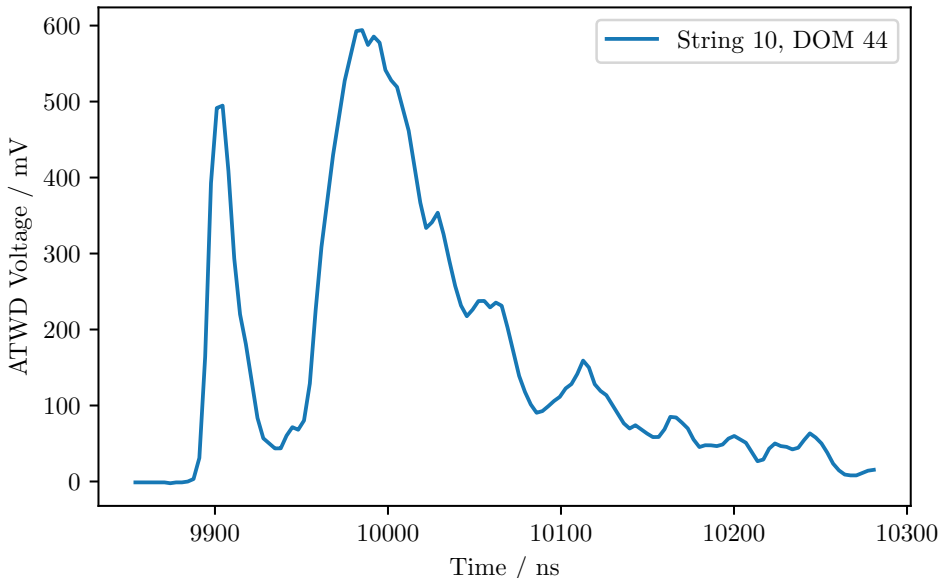


Figure 6.4: Double Pulse waveforms of the event recorded on April 2nd 2016 (`RunID`: 127762, `EventID`: 66008498). The corresponding Double Pulse score is 0.565.

6.2 Astrophysical Tau Neutrino Flux Normalization

In addition to the detailed description of each found event, the data sample introduced in chapter 4 was used to perform the likelihood fit introduced in chapter 5. The fit measures the astrophysical tau neutrino flux normalization for three different assumptions on the astrophysical spectral index.

$$\Phi_{\nu_\tau}(E_\nu) = 0.44^{+0.78}_{-0.31} \cdot 10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1} \left(\frac{E_\nu}{100\,\mathrm{TeV}}\right)^{-2.19} \tag{6.1}$$

$$\Phi_{\nu_\tau}(E_\nu) = 0.83^{+1.46}_{-0.59} \cdot 10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1} \left(\frac{E_\nu}{100\,\mathrm{TeV}}\right)^{-2.50} \tag{6.2}$$

$$\Phi_{\nu_\tau}(E_\nu) = 1.62^{+2.78}_{-1.11} \cdot 10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1} \left(\frac{E_\nu}{100\,\mathrm{TeV}}\right)^{-2.89} \tag{6.3}$$

The quoted statistical uncertainties are at a 68 % confidence level. The statistical uncertainties at a 90 % confidence level are shown in Table A.7. The likelihood scans and the critical values for the observed data sample are presented in Figures 6.5, A.12a and A.12b.

To examine the significance that the observed data sample does not contain a tau neutrino component a likelihood ratio test (cf. eq. (5.5)) with the hypotheses

$$H_0 : \Phi_{\nu_\tau} = 0$$
$$H_1 : \Phi_{\nu_\tau} > 0$$

is conducted. This is accessed by sampling the likelihood ratio test statistic distribution through pseudo experiments. The test statistic distribution and the observed test statistic value for an $E^{-2.19}$ spectrum are illustrated in Fig. 6.6. The observed test statistic value corresponds to a p-value of 0.056 (1.91σ). The p-value only slightly depends on the assumed astrophysical spectrum as the p-value for the $E^{-2.50}$ spectrum is 0.055 (1.92σ) and for the $E^{-2.89}$ spectrum is 0.052 (1.95σ).

6.2.1 Impact of Systematic Uncertainties

To study the impact of systematic parameters on the fit result the fit is redone with systematically altered observable distributions. To investigate flux systematics the weights of the simulated events can be recalculated according to the desired flux models. Detector systematics are studied using dedicated simulations where the corresponding detector parameters are different from the baseline simulation.

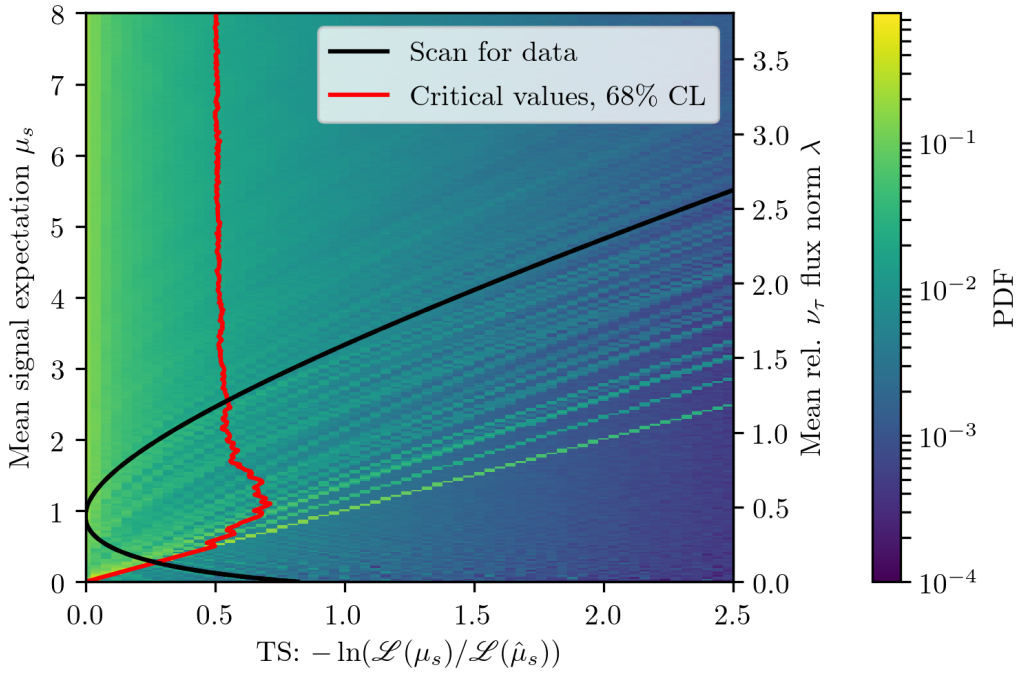


Figure 6.5: Likelihood scan for the observed data sample for an astrophysical spectrum of $E^{-2.19}$. The red line shows critical values at 68 % confidence level obtained from the underlying test statistic distributions for each value of μ_s.

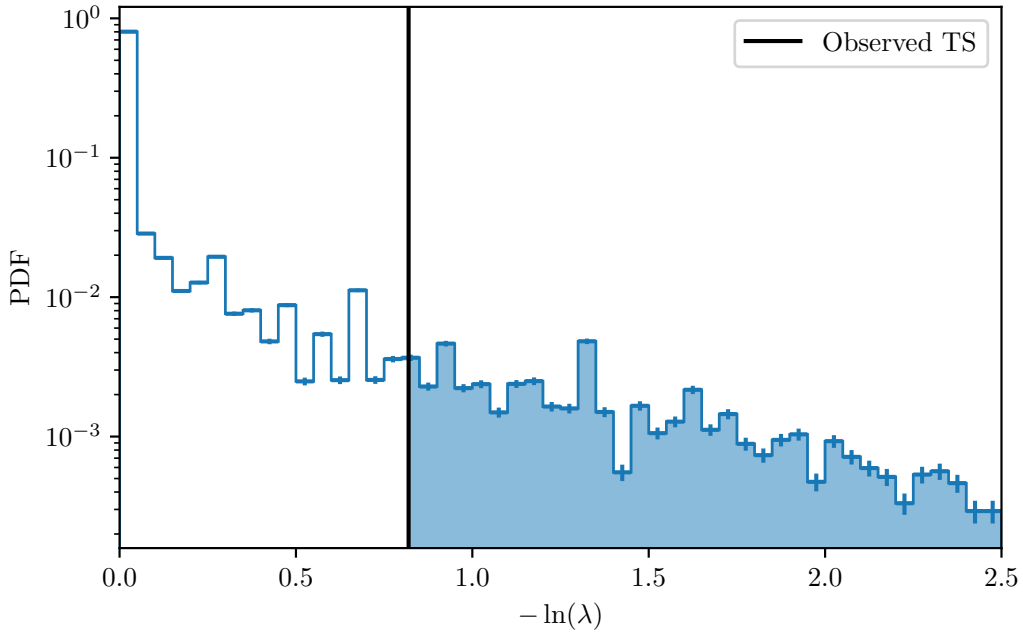


Figure 6.6: Test statistic distribution of likelihood ratio test (cf. eq. 5.5) to calculate the significance with respect to a zero tau neutrino flux normalization. This test statistic distribution is the same as the one shown in the first column of Fig. 6.5 (for $\mu_s = 0$).

The calculations here have two important caveats. First, the systematic datasets only cover an energy range up to 10^7 GeV, while the baseline simulations go up to 10^8 GeV. This is circumvented by comparing the obtained upper and lower limits for the flux to the limits obtained when the baseline simulation is truncated at 10^7 GeV and scaling the systematic boundaries accordingly. Second, due to limited computing resources, systematic datasets for the atmospheric muon simulation do not exist. This is handled by always using the baseline simulation for the atmospheric muon simulation while going through the systematic datasets for the neutrino simulation.

The resulting changes[2] due to systematic variations are presented in Table 6.1 for an astrophysical $E^{-2.19}$ spectrum (values for the other spectra can be found in Tab. A.8 and A.9). A conservative estimate on the influence on the lower and upper bound of the confidence interval can be obtained by adding up changes in quadrature, that push the boundaries of the interval outwards. To perform this calculation, for each source of systematic uncertainties only the worst-case variation is considered. This results in a relative reduction of the lower limit by 5.2 %, and a relative increase of the upper limit by 18.1 % in total. The p-value is estimated to increase by 32.3 % due to the systematic uncertainties based on the assumptions explained above.

Overall, these systematic variations show a negligible impact on the analysis compared to the uncertainties arising from the Poisson fluctuation due to the small expected number of events, which was presented in Tab. 5.2. For the lower limit of the confidence intervals most systematic variations improve the lower limit. Exceptions are increased DOM efficiency, certain hole ice parameters, and an increased atmospheric muon background flux. These also show the strongest decline in the significance to reject the hypothesis of no present tau neutrino flux.

6.2.2 Differential Upper Limit

The differential upper limit described in section 5.5.1 is presented in Figure 6.7 for the observed data sample. In addition to the theoretical models already presented in Figure 5.5, the differential upper limit observed in [10] is shown. The analysis is also searching for tau double pulse events with IceCube and uses the same method presented here, which calculates the differential limit based on the hypothesis of an E^{-1} signal flux[3] defined over one decade in energy. The differential upper limit from this work improves the limit presented in [10] by roughly a factor of 3 over the whole energy range. This is achieved by the combination of an increased detector

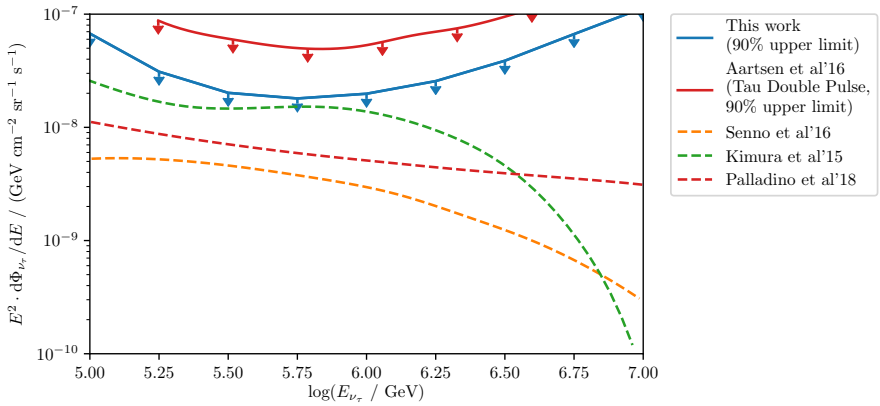
[2] Note that the possible values for the relative changes are limited by the step size used for the injected signal mean for μ_s, which restricts the possible confidence intervals in μ_s to the used step size of 0.05.

[3] The paper actually states that an E^{-2} signal flux is used, which turned out to be a typing error.

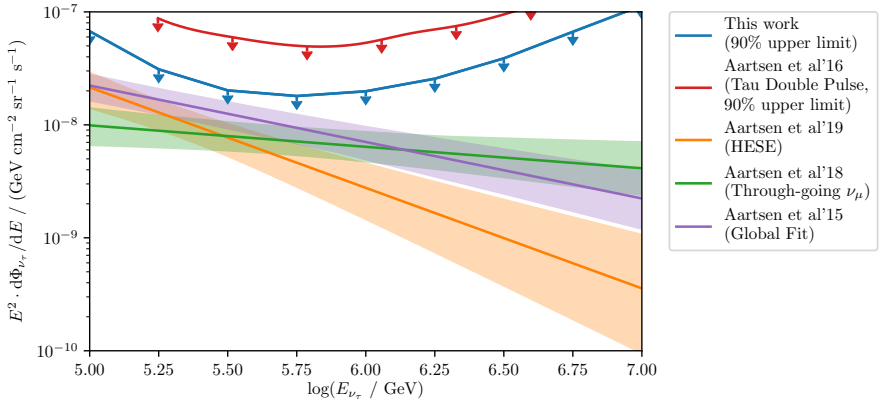
Table 6.1: Relative influence of the tested sources of systematic uncertainties on the lower and upper bound ($\Phi^{l,68}_{\nu_\tau}$ and $\Phi^{u,68}_{\nu_\tau}$) of the tau flux normalization. The last column quantifies the impact on the p-value to reject a tau neutrino flux normalization of $\Phi_{\nu_\tau} = 0$. The values shown here are calculated for an astrophysical $E^{-2.19}$ spectrum. Nominal values are shown in Tab. 5.3.

Description	Variation	$\Delta\Phi^{l,68}_{\nu_\tau}/\Phi^{l,68}_{\nu_\tau}$	$\Delta\Phi^{u,68}_{\nu_\tau}/\Phi^{u,68}_{\nu_\tau}$	$\Delta p_{\Phi_{\nu_\tau}>0}/p_{\Phi_{\nu_\tau}>0}$
scattering (bulk ice)	+10 %	+27.2 %	+17.7 %	+0.8 %
absorption (bulk ice)	+10 %	+22.9 %	+10.6 %	−2.3 %
scat. and abs.	−7 %	+19.6 %	+7.6 %	+9.4 %
hole ice	$p_2 = -3$	−1.9 %	−5.6 %	+10.8 %
	$p_2 = -1$	+10.4 %	+0.3 %	+0.4 %
	$p_2 = +1$	+0.1 %	−2.7 %	+26.9 %
DOM efficiency	+10 %	−4.8 %	−6.6 %	+11.6 %
	−10 %	+3.6 %	+3.6 %	+4.4 %
$\Phi_{\nu,\text{conv}}$	+30 %	0.0 %	0.0 %	−4.4 %
	−30 %	0.0 %	0.0 %	+3.7 %
$\Phi_{\nu,\text{prompt}}$	BERSS_u [84]	0.0 %	−1.0 %	−1.6 %
	BERSS_l [84]	0.0 %	0.0 %	0.0 %
$\Phi_{\nu,\text{astro}}$	+0.26	0.0 %	0.0 %	0.0 %
	−0.23	+9.1 %	+1.0 %	+6.0 %
γ_{astro}	+0.1	+9.1 %	+1.0 %	+6.7 %
	−0.1	0.0 %	0.0 %	−2.4 %
Φ_μ	+15 %	0.0 %	−2.0 %	−2.1 %
	−15 %	+9.1 %	+1.0 %	+1.6 %

livetime, an improved event selection and by using the binned Poisson likelihood fit as a more sensitive analysis method compared to a simple counting experiment. The observed differential limit is also compatible with the theoretical models shown in Fig. 6.7 and with current IceCube measurement of the astrophysical neutrino flux.



(a) Comparison of the differential upper limit with models for astrophysical neutrino fluxes from choked jets and low luminosity gamma-ray bursts [117] (orange), low luminosity AGNs [118] (green) and from a multicomponent model [119] (red).



(b) Comparison of the differential upper limit with the astrophysical neutrino flux measurements used for the model dependent constraints. The through-going diffuse ν_μ result is shown in green [4], the latest HESE result in orange [7] and the Global Fit in purple [5]. The shaded regions indicate the 1σ regions.

Figure 6.7: Differential upper limit on the tau neutrino flux for the observed data sample between 100 TeV and 10 PeV. The red solid line shows the differential upper limit observed by [10].

6.3 Comparison with other IceCube Measurements

Up to the date of this work, two different strategies are used within the IceCube collaboration to measure the flux of astrophysical tau neutrinos. Event selections focusing on extracting events showing a double pulse topology, which is accessed by simultaneously searching for double pulses in waveforms and cascade-like event topologies. This type of event selection is mostly sensitive to tau neutrinos while the most dominant backgrounds are track-like events (this work, [10], [112]). Event selections based on the double bang topology use extensive likelihood-based reconstructions to reconstruct the tau length amongst other physical parameters and use these observables to identify tau neutrinos. So far, since these reconstructions are computationally expensive, this type of event selection is based on high energy starting events (HESE) [2, 3, 52, 7] containing mostly high energy neutrinos of all flavors, thus with a high probability of being astrophysical. These high energy neutrinos are classified into three classes, cascades, tracks, and double cascades and then used to measure the astrophysical flavor ratio, while the sensitivity to the tau neutrino component is coming from the double cascade events [102, 103]. The dominant backgrounds are mis-reconstructed single cascade events.

Due to the limitations of the HESE selection, these analyses have a slightly larger energy threshold compared to the double pulse analyses and they have to rely on the precise knowledge of systematic effects, for example, knowledge of the south pole ice model as it is an important part of the double cascade reconstruction. In contrast, the presence of astrophysical electron and muon neutrinos in the HESE data sample allows the simultaneous measurement of the astrophysical neutrino flux of all flavors, hence, the astrophysical flavor ratio. It also allows a more sophisticated integration of nuisance parameters into the measurement and thus can provide better estimates on the influence of systematic effects.

The results obtained in this work are compatible with the other double pulse analyses conducted with IceCube data. The differential upper limit (as shown in the previous section), as well as the model-dependent 90 % upper limit of $\Phi_{\nu_\tau}(E_\nu) < 5.1 \cdot 10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}\,(E_\nu\,/\,100\,\mathrm{TeV})^{-2}$, based on the null observation made in [10] is larger than the measured tau neutrino flux here. The successor of this analysis [112] also follows the goal of measuring the tau neutrino flux normalization with a binned Poisson likelihood fit as it is done in this work. It improves the event selection presented in [10] significantly and finds three tau neutrino candidates. Two of those events are also found in this analysis (cf. sec. 6.1). There a best fit tau normalization of 0 is found and the 90 % upper limits (set on for the same astrophysical flux assumptions of $E^{-2.19}$, $E^{-2.50}$ and $E^{-2.89}$ as analyzed in this thesis) are comparable with the upper boundaries of the 68 % confidence intervals.

Overall, even though different methods are used in most parts of both analyses, the good agreement is not surprising because both event selections have a large overlap and the data samples are not statistically independent.

Comparing the results in this work to those obtained by the double cascade analyses is slightly more tricky since their results are presented in the form of a flavor triangle. Two flavor composition measurements for the double cascade analysis using the HESE sample based on six years of data [102] and the successor of this analysis based on 7.5 years of data [103] (which is still referred to as work in progress) exist. The result based on 7.5 years of data is shown in Figure 6.8. The flavor composition is presented as a flavor triangle, where the likelihood ratio at each point is evaluated by fitting the flux normalizations for each flavor but constraining their sum to the total observed flux. Both show the best-fit point and the 68 % and 90 % confidence interval contours.

The analysis presented in [102] found no double cascade events in six years of IceCube data in the HESE sample [52] and set a 90 % upper limit $\Phi_{\nu_\tau}(E_\nu) < 2.68 \cdot 10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}\,(E_\nu\,/\,100\,\mathrm{TeV})^{-2.97}$ on the tau neutrino flux. This sets an upper limit on a spectrum even softer than the $E^{-2.89}$ spectrum tested in this analysis. The second double cascade analysis [103] uses the same strategy to fit the flavor ratio, but increases the amount of data to 7.5 years of the HESE selection [7] and partly improves the used analysis methods. Two double cascade events were found resulting in an observed flavor ratio of $\nu_e : \nu_\mu : \nu_\tau = 0.29 : 0.50 : 0.21$. The limited amount of statistics in both analyses still entails large contours in the flavor triangle.

When combining the assumption of equal flavor contributions for ν_e and ν_μ in this work with the observed tau neutrino flux normalization the resulting flavor ratio is contained within the 1σ-contours of both double cascade analyses[4]. Figure 6.8 presents the measurement of the tau neutrino flux normalization obtained in this work for the $E^{-2.89}$ spectrum measured with the HESE analysis using 7.5 years of IceCube data [7] in the context of the astrophysical flavor composition measurement [103]. The observed consistencies are hardly surprising here because, although the event selections follow a different strategy to identify tau neutrino events, there is a significant overlap in both the event selection and the observed data sample (as already discussed in sec. 6.1).

[4] It should be kept in mind that the soft spectrum of $E^{-2.97}$ in [102] does not allow a fair comparison in a statistical meaningful way without further analysis.

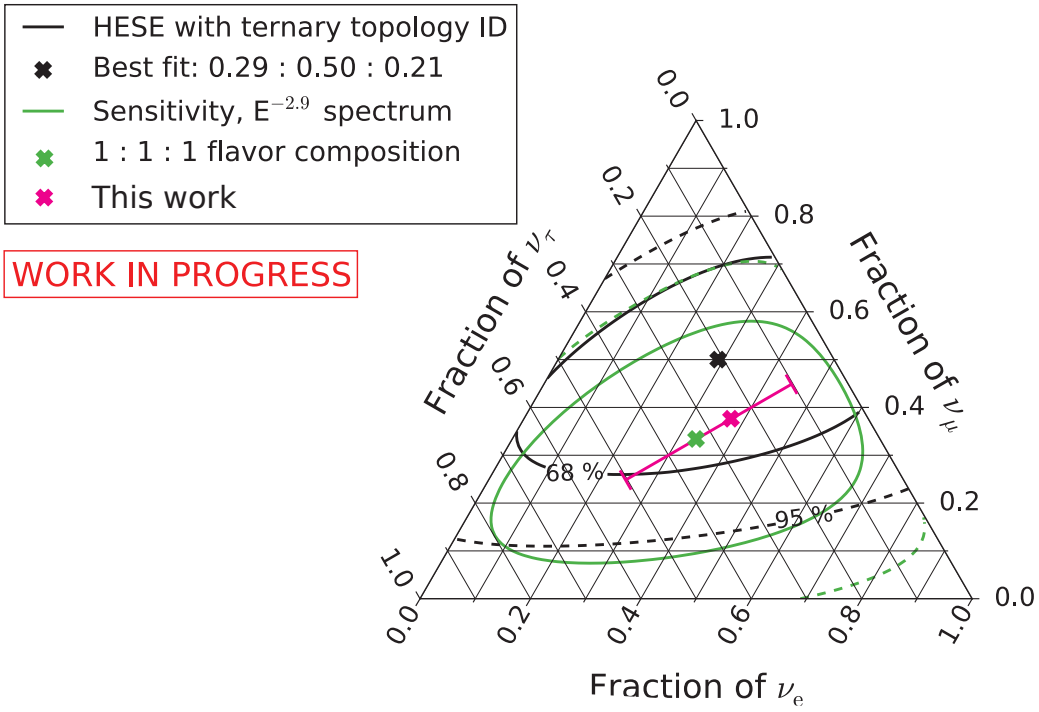


Figure 6.8: Astrophysical flavor composition measurement obtained by IceCube's double cascade analysis [103] based on 7.5 years of IceCube data using the HESE event selection [52, 7]. The green lines show the 68 % and 90 % contour of the analysis' sensitivity and the black lines show the contours of the measurement. The best fit and the error bar in magenta depict the result obtained in this work. The contributions from ν_e and ν_μ are fixed in this analysis, and only the ν_τ fraction is varied. The strong assumptions on the ν_e and ν_μ contributions lead to, in general, smaller uncertainties compared to a simultaneous measurement of three flavors. Figure adapted from [103].

6.4 Discussion

The event selection developed in this thesis and presented in chapter 4 significantly improves the properties of the final data sample compared to the previous double pulse analysis conducted with IceCube [10]. The properties can be compared most easily for an $E^{-2.50}$ spectrum because event expectations for that spectrum are presented in the publication. The number of expected events per detector livetime is increased by a factor of 2.5. The tau neutrino purity of the final sample is also increased from 59 % to 69 %. Additionally, the analyzed data taking period has been extended from 914 days to 2667 days.

The identification of double pulse waveforms and the rejection of single cascade waveforms has been improved by favoring a machine learning based approach over previously used straight cuts. New observables for the double pulse identification have been added to further increase the separating power achieved by the Random Forest. The double pulse identification can potentially be improved by the development of a more elaborate method to define signal waveforms for the classification. Another possibility, investigated as a part of this thesis, is the application of a machine learning model capable of classifying waveforms without the explicit construction of observables. This was tested on simulations with a variety of neural network architectures, where a convolutional neural network showed the most promising results [124].

The cascade selection (cf. section 4.3) developed in this thesis does not rely on a containment cut. As a consequence, the atmospheric muon background is the dominant background contribution and it has been investigated carefully. The classification of atmospheric muons as background events can be improved by more efficient reconstructions to identify small amounts of outgoing charge after a catastrophic energy loss of a muon. Another point of improvement could be the development of observables to distinguish between events entering and leaving the detector for partially contained events.

Overall, the incorporation of a machine learning approach into the event selection increases its reliability by using validation techniques such as cross validation. This also improves the ability to verify the event selection compared to event selections based on straight cuts, especially in regions close to the signal region, with observed data via the classification score distributions.

In this dissertation, a binned Poisson likelihood fit is used to measure the tau neutrino flux normalization for different astrophysical spectral indexes. This is an improvement over the previously used method of a counting experiment [85]. For the likelihood fit, observables developed in this thesis are used, increasing the

sensitivity of the analysis compared to conventional observables like the reconstructed energy [112]. Using the same observables, a method has been developed to assign p-values to potential tau neutrino candidates.

This analysis has found two tau neutrino candidate events, that were identified via the double pulse event topology. One is found to be a background event, while the other event has a higher probability of being a tau neutrino with a p-value of approximately $p \simeq 0.035$ with only a slight dependence on the astrophysical flux assumption. A major fraction of the sensitivity comes from events with energies between 100 TeV and 6 PeV. For the soft spectrum, the signal expectation is therefore lowered by 40 %. With a total event expectation between 2.19 and 3.47 in the analyzed amount of 7.5 years of IceCube data finding two events is not unexpected, but a sign for either a small under-fluctuation or a very soft spectrum.

As a result, this is, in combination with [112], the first observation of tau neutrino candidates with the double pulse topology. Additionally, this is the first non-zero measurement of the tau neutrino flux using the double pulse method as a tau neutrino identifier, giving a hint at the existence of an astrophysical tau neutrino flux with a significance of roughly 1.9σ (independent of the assumed astrophysical spectrum). The measured tau neutrino fluxes are also consistent with the expected benchmark flavor ratio of $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$.

7 Summary and Outlook

The main goal of this dissertation is the identification of high-energy astrophysical tau neutrino candidate events within data from the IceCube detector, and a subsequent measurement of the astrophysical tau neutrino flux. To achieve this, a new tau neutrino identification method has been developed to increase the expected tau neutrino event rate by a factor of 2.5 compared to a previous IceCube analysis which used the same event signature for tau neutrino identification [10]. This event selection is the first (amongst double pulse analyses) to renounce the explicit use of a veto technique to reject atmospheric muons and thus extend the available fiducial volume of the detector. As a result, the modeling of the atmospheric muon background has been given particular attention. The applied methods to constrain the tau neutrino flux have also been improved by employing a binned Poisson likelihood fit and a new strategy to characterize the compatibility of observed event candidates has been defined before the measurement.

Data recorded with the IceCube detector between May 2011 and December 2018 (resulting in 2667 days of detector uptime) has been analyzed and two tau neutrino candidate events have been found. One event appears to be background-like with a p-value of $p = 1.0$ and the other one showed very signal-like characteristics resulting in a p-value of $p \simeq 0.035$ (almost independent of the assumed astrophysical spectrum). For the first time, current iterations of IceCube's astrophysical tau neutrino searches have overcome the null-observation regime (cf. [112, 103]).

The observed data sample results in the following measurement of the astrophysical tau neutrino flux based on the assumed astrophysical shape:

$$\Phi_{\nu_\tau}(E_\nu) = 0.44^{+0.78}_{-0.31} \cdot 10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1} \left(\frac{E_\nu}{100\,\mathrm{TeV}}\right)^{-2.19}$$

$$\Phi_{\nu_\tau}(E_\nu) = 0.83^{+1.46}_{-0.59} \cdot 10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1} \left(\frac{E_\nu}{100\,\mathrm{TeV}}\right)^{-2.50}$$

$$\Phi_{\nu_\tau}(E_\nu) = 1.62^{+2.78}_{-1.11} \cdot 10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1} \left(\frac{E_\nu}{100\,\mathrm{TeV}}\right)^{-2.89}$$

These results are not in conflict with the expected flavor ratio of $\nu_e : \nu_\mu : \nu_\tau \simeq 1 : 1 : 1$, but exclude the non-observation of an astrophysical tau neutrino flux with a significance of 1.9σ (also almost independent on the assumed astrophysical

spectrum). In addition, a differential limit has been derived to also constrain the astrophysical tau neutrino flux in a more model-independent way.

The signal-like tau neutrino candidate is also observed by [103] and an *a posteriori* analysis to access the probability of this candidate is currently in progress [125]. Both analyses have found this event with complementary identification methods, which both have different dominant sources of background. Thus, a combination of both strategies might be able to attribute this event to a charged current tau neutrino interaction with a much larger significance.

In the future, improvements to the statistics of tau neutrino analyses are necessary to conduct precise measurements of the astrophysical flavor ratio, or, for example, to search for neutrino sources based on a tau neutrino enriched sample. To achieve this, additional observation time and improvements to both the analysis methods for tau neutrino analyses and the available detectors would be desirable.

An improved version of IceCube's global fit [5] combining different event selections is currently in development [126]. It could potentially be improved further by the inclusion of the event selection developed in this dissertation with respect to the tau neutrino flavor sensitivity. This would also have great benefits over this analysis as the additional events from other event selections could enable a simultaneous measurement of the flavor ratio and the inclusion of systematic uncertainties as nuisance parameters.

In the future, two major experimental improvements are considered for very high energy neutrino detection. The next-generation neutrino observatory IceCube-Gen2 [127] is planned to increase the detector volume of IceCube to $\sim 10\,\mathrm{km}^3$ with less dense instrumentation. This is expected to increase the sensitivity to tau neutrinos above PeV energies by an energy-dependent factor of 5 - 10, when considering identification via the double bang signature [127]. The improvement for double pulse events will be smaller, as both energy depositions are required to be close to a DOM, thus less dense instrumentation does not scale in the same way as for double bang searches. Another prospect for observing neutrinos at higher energies is the detection of radio signals produced in electromagnetic showers. This detection technique is used by e.g. the Askaryan Radio Array (ARA) [128] built around the IceCube detector, which can detect neutrinos with energies beyond 10 PeV. This detection principle will significantly improve the sensitivity to cosmogenic neutrinos.

A Appendix

A.1 IceCube Simulations

The IceCube event simulation considers all relevant physical processes from particle generation to the actual recorded signal by detector electronics. Simulated IceCube events can be divided into three subgroups by the used event generator: neutrino events, muons and muon bundles from cosmic ray air showers, and single muons.

The software `NuGen` (neutrino generator) generates neutrino events at the Earth's surface following a PDF based on the assumed neutrino flux (usually a power-law). `NuGen` is a modified version of the neutrino generator `ANIS` [129]. It also propagates neutrinos from the surface to the IceCube detector describing the Earth with the Preliminary Reference Earth Model (PREM) [130] and handles the initial neutrino interaction.

Cosmic ray primaries and the resulting extended air showers are generated with `CORSIKA` [78]. The resulting output of an air shower that is interesting for IceCube are the muons that reach the surface.

Simulating the full development of air showers is computationally expensive. A solution to that is the so-called `MuonGun` simulation [79] (the same techniques are used in the software `MUPAGE` [131]), which parametrizes the muon yield close to the IceCube detector derived from the output of `CORSIKA` simulations, propagated to the detector with `PROPOSAL` [77].

After the event generation, all events are treated in the same way. Charged leptons are propagated with `PROPOSAL` [77, 132]. `PROPOSAL` simulates the energy losses and decays of these leptons. Hadronic cascades are simulated by the software `cmc`, which creates photons with a direction and their time of production at the position of the hadronic cascade. At energies above 1 TeV, the cascade elongation is taken into account.

Computationally the most challenging step is the propagation of photons from their origin through the south pole ice. The software `clsim` generated photons from the leptonic energy losses, which is handled in a similar way as in `cmc`. The cascades get parametrized and summarized in `Photonics` tables [93] that can be used to sample

produced photon counts. After that, the photons are propagated individually via ray-tracing through the inhomogeneous south pole ice until they either hit a DOM or get absorbed.

Finally, the whole detector response to the photon hits is simulated. This includes detector noise, physical processes in the PMT and effects of the digitizing electronics. After this step simulations can be treated the same way as data.

A.2 Tree-based Learning Algorithms

Tree-based algorithms are supervised learning algorithms that can be applied to classification and regression problems. The most basic unit of a tree-based algorithm is a decision tree. First, a decision tree is trained with labeled training data, where for a classification task the true class and for a regression task the value to be estimated has to be known. The decision tree splits the training set by a consecutive set of straight cuts in available observables. Each cut is optimized by a criterion depending on the task. In this thesis, the Gini impurity is used exclusively. The Gini impurity is defined by

$$I_G = \sum_{i=1}^{N} p_i \cdot (1 - p_i) \tag{A.1}$$

for a set of examples with N classes, and the probability for an example associated with class i p_i of being selected at random. To choose the best cut during the training process all possible cuts are evaluated on each observable and the cut that maximizes the Information gain/Gini gain IG_G is chosen.

$$IG_G = I_{G,\text{parent}} - \sum_{i \in \{<,\geq\}} I_{G,i}. \tag{A.2}$$

$I_{G,\text{parent}}$ is the Gini impurity at the parent node before a split is chosen. The sum goes over both subsamples at the nodes obtained by the split.

After selecting a split the procedure is repeated for each new node until a stopping condition is met. The usual stopping condition is that no information gain $IG_G > 0$ can be achieved. Other conditions as the number of examples at a node or a maximal depth of the tree can be used as well. This algorithm is also often referred to as the CART-algorithm (Classification and regression trees) [133].

After this the training process is complete and the decision tree can be used for classification. This is done by applying cuts to an example until it falls into a terminal node, where the example is classified as the dominant class at that node.

A.2.1 Random Forest

A Random Forest [134] is an ensemble method combining several decision trees. To obtain an ensemble of statistically independent decision trees the training process for the decision trees needs to be modified. For each tree, the training set is bootstrapped from the original training set (by drawing the same number of samples with replacement). At each node that is built only a subset of observables is considered. A common choice for the available number of observables is $\sqrt{n_{\text{Observables}}}$. After training a decision is made by majority vote. Each tree assigns a class to an example. These votes can be combined to a classification score, where the score for class i is the number of trees that voted for class i over the total number of trees.

A.2.2 Feature Importance

After the training the importance of an observable X_j for predicting the label Y can be evaluated by adding up weighted impurity decreases $p(t)\Delta I_G(s_t, t)$ for all nodes t that use the observable X_j:

$$f(X_j) = f_j = \frac{1}{N_T} \sum_T \sum_{t \in T : v(s_t) = X_j} p(t) \Delta I_G(s_t, t) \tag{A.3}$$

where N_T is the number of trees in the forest, $p(t)$ is the fraction of examples reaching node t and $v(s_t)$ is the variable used in the split s_t [135]. The values of f_j are normed in a way that $\sum_j f_j = 1$.

A.3 Cross-validation

Cross-validation is a technique to validate model performance in a statistical analysis used multiple times throughout this thesis. It is used to obtain a generalized performance of a statistical model (in this thesis often a Random Forest) by exposing it to statistical fluctuations in the training and test set. This creates a reliable estimate of the performance for the application on experimental data.

A n-fold cross-validation splits a data sample into n disjoint subsets. The analysis is repeated n times using $n-1$ subsets as the training set and the remaining subset as the test set. This way, each example in the data sample will be used once for testing and thus get an unbiased estimate of e.g. the class prediction or a reconstructed value as it is not part of the training set.

A.4 Feature Selection

A.4.1 Kuncheva Index

The Kuncheva stability index is given by

$$\mathscr{I}_S(\mathscr{A}) = \frac{2}{K(K-1)} \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} I_C(S_i, S_j), \tag{A.4}$$

with a set of K feature sets $\mathscr{A} = \{S_1, \dots, S_K\}$. The consistency index I_C is

$$I_C(A, B) = \frac{rn - k^2}{k(n-k)}, \tag{A.5}$$

where k denotes the size of the sets A and B, n denotes the total number of features considered in the feature selection and r is the size of the intersection between A and B. This consistency index satisfies the three conditions explained in the main matter – monotonicity, limits, and correction for chance. [109]

A.4.2 Observable distributions

This section shows a subset of the observable distributions discussed in 4.3.2. The remaining distributions can be downloaded: features removed by feature importance[1], manually removed features[2] and the final feature set[3]. For the observables "Tensor of Inertia Eigenvalue ratio q_{ToI}" and the "CascadeFillRatio (based on $r_{\mathrm{mean+RMS}}$)" presented in Figure A.2 (top right and bottom right), worrying disagreements can still be observed for small values. Fortunately, these disagreements are in a region of these observables where the signal-to-background ratio is extremely small. It can be expected that regions like this are classified as very background-like by the Random Forest classifier. This behavior was checked after the classification by applying a classification score threshold of 0.01, which already removed the disagreements in these observables.

[1] `https://icecube.wisc.edu/~mmeier/analysis_docu/html/_downloads/removed_features.zip`
[2] `https://icecube.wisc.edu/~mmeier/analysis_docu/html/_downloads/manually_removed_features.zip`
[3] `https://icecube.wisc.edu/~mmeier/analysis_docu/html/_downloads/features_final.zip`

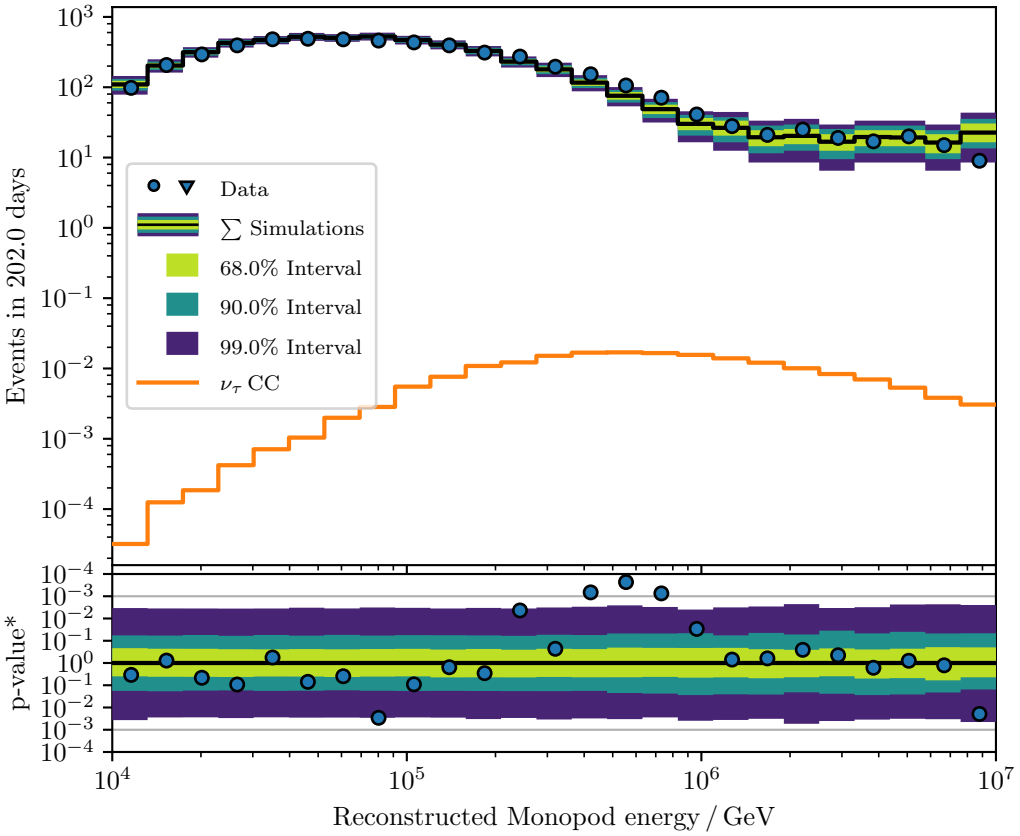


Figure A.1: Comparison of data and simulations after the double pulse selection. Distribution of the energy reconstructed with Monopod. The energy reconstructed with this algorithm still shows clearly visible disagreements especially in an energy region interesting for this analysis (between 200 TeV and 1 PeV).

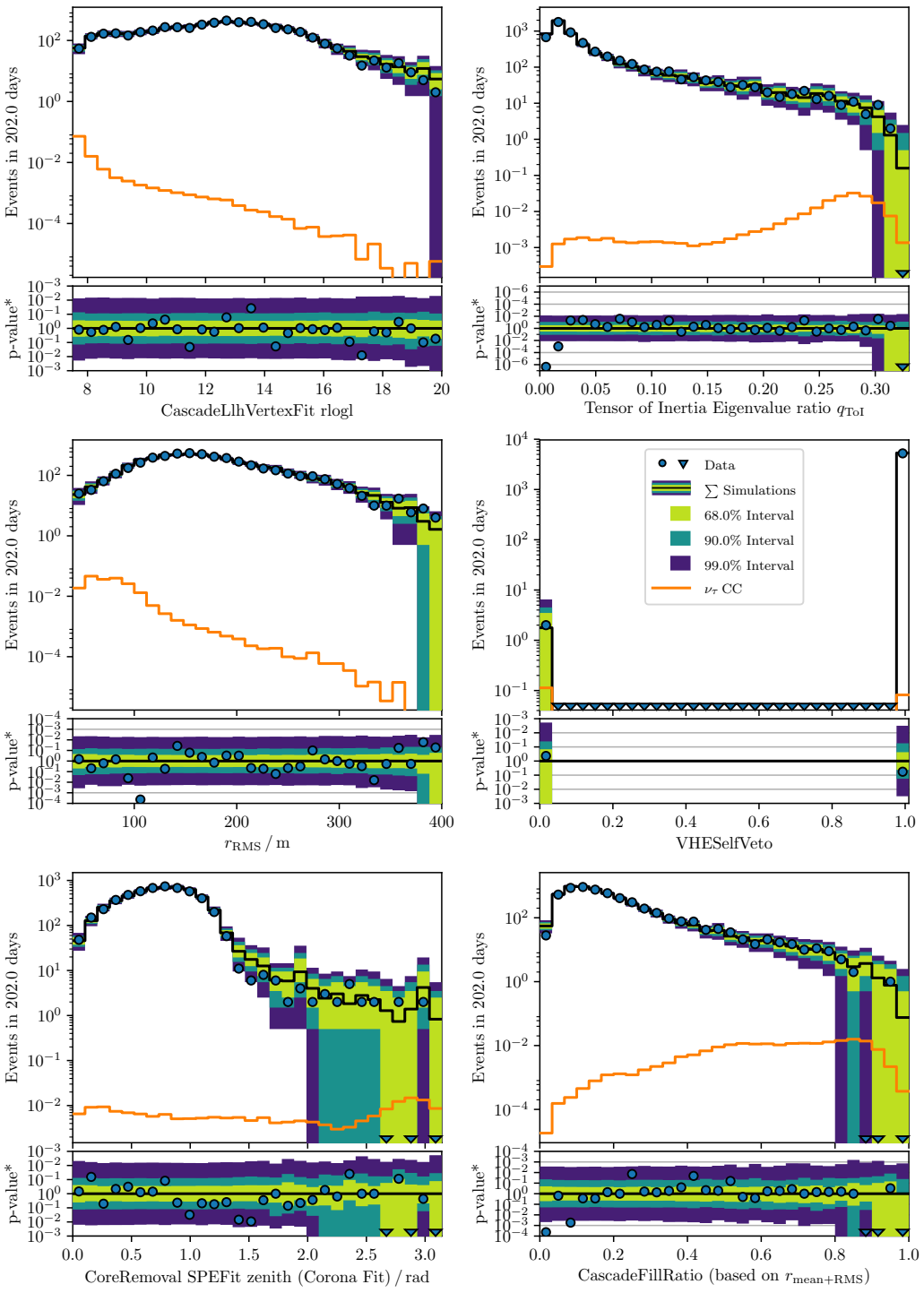


Figure A.2: Comparison of data and simulations after the double pulse selection. Distributions of the six observables with the highest feature importance in the Random Forest (from top left to bottom right).

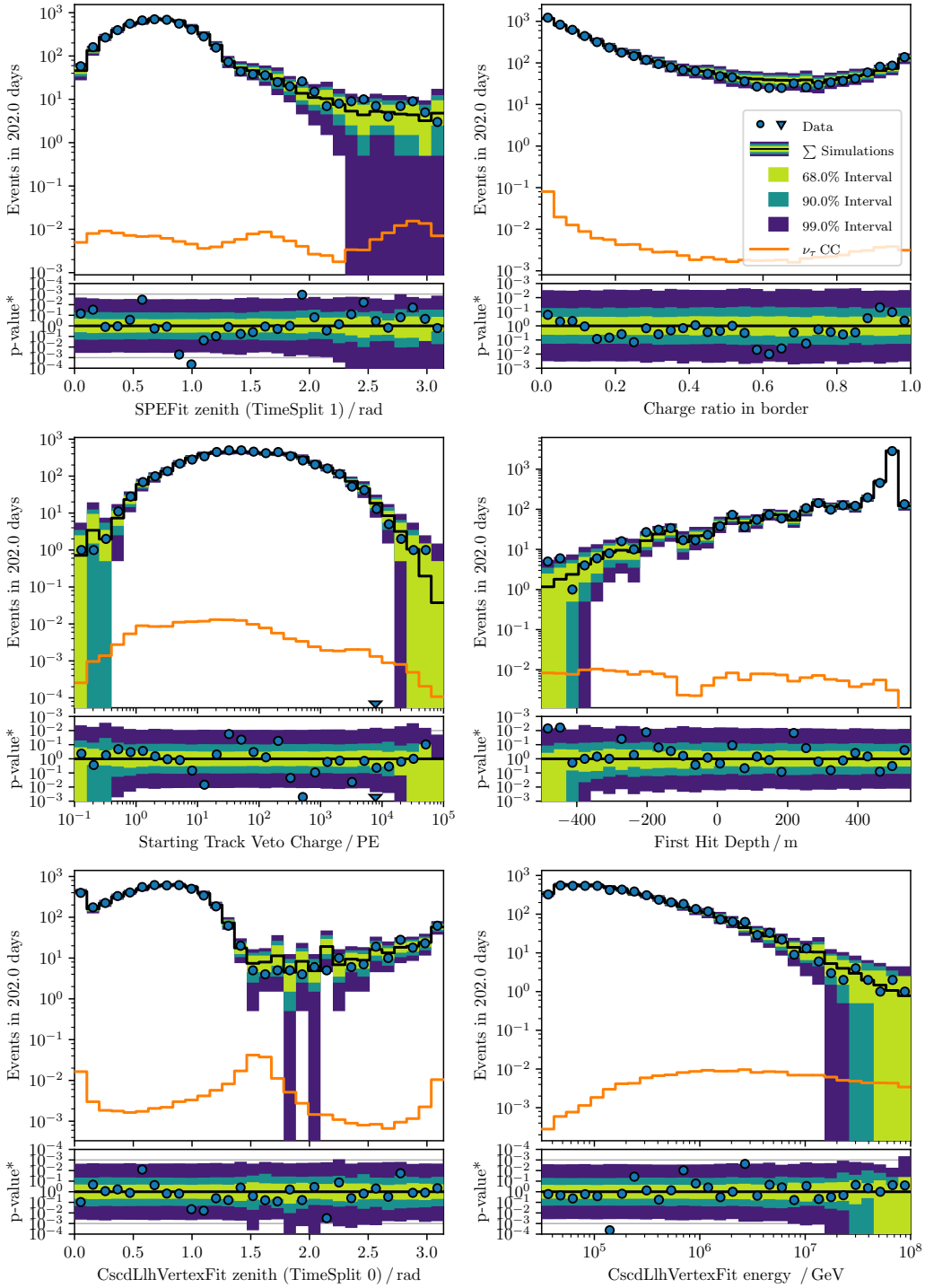


Figure A.3: Comparison of data and simulations after the double pulse selection. Distributions of the six observables ranked seventh to twelfth based on their feature importance in the Random Forest (from top left to bottom right).

A.5 Classification

Table A.1: Settings used for the Random Forest classification in sections 4.2 and 4.3.3.

Parameter	Value	Comment
`n_estimators`	200	Number of trees.
`criterion`	`'gini'`	Criterion to measure the quality of a split. `'gini'` selects the Gini impurity as a measure.
`max_depth`	`None`	Maximum depth of the trees. `None` results in an unlimited depth.
`min_samples_split`	2	Minimum number of samples required to perform a split at a node.
`min_samples_leaf`	1	Minimum number of samples to arrive at a leaf node.
`min_weight_fraction_leaf`	0	Minimum fraction of total weights required to be at a leaf node.
`max_features`	`'sqrt'`	Number of features considered at each node. Here: `max_features`= $\sqrt{n_{\text{features}}}$.
`max_leaf_nodes`	`None`	Maximum number of leaf nodes. `None` results in an unlimited number of lead nodes.
`min_impurity_decrease`	0	Minimum decrease in impurity required to induce a split.
`bootstrap`	`True`	Bootstrap samples when building trees.

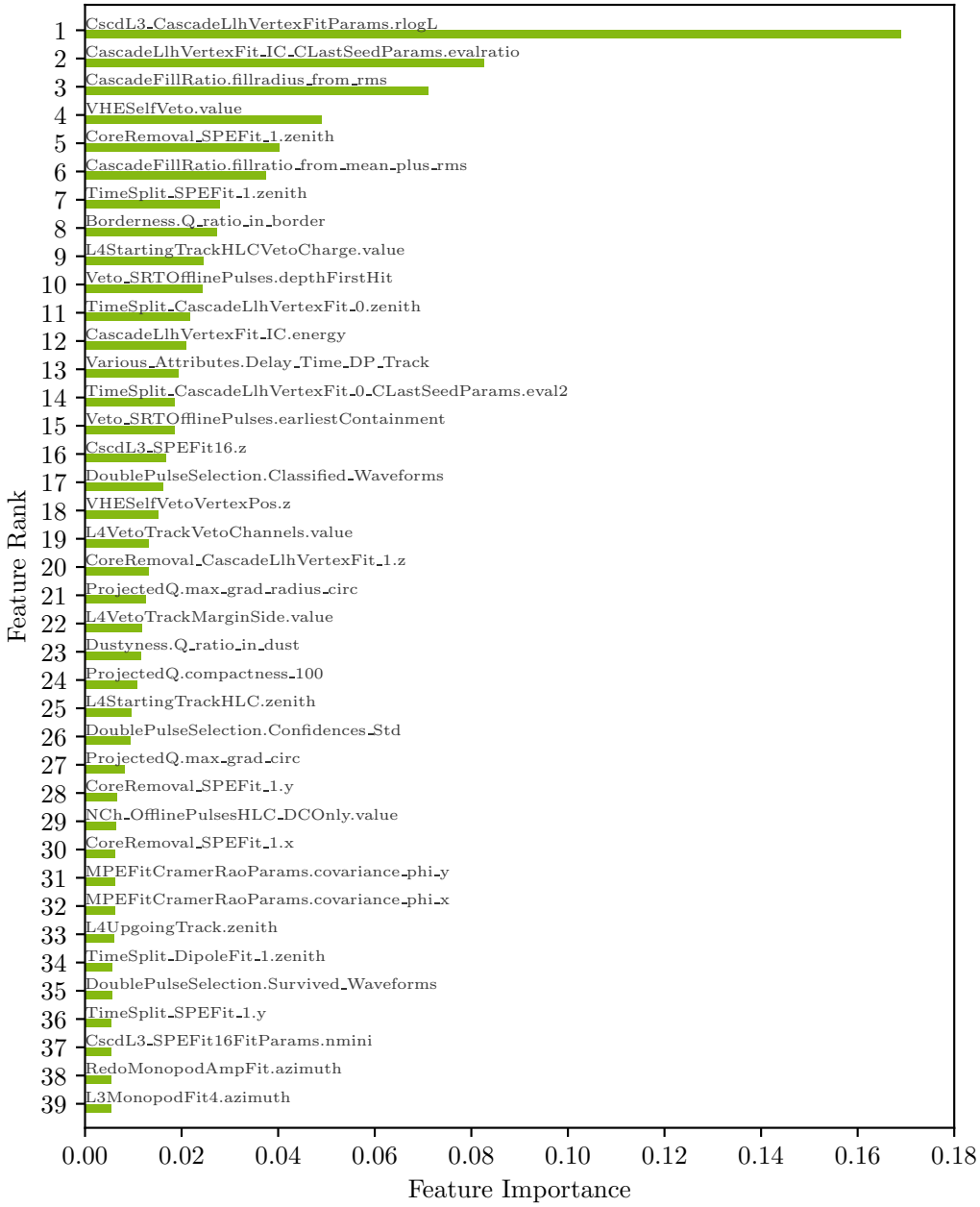
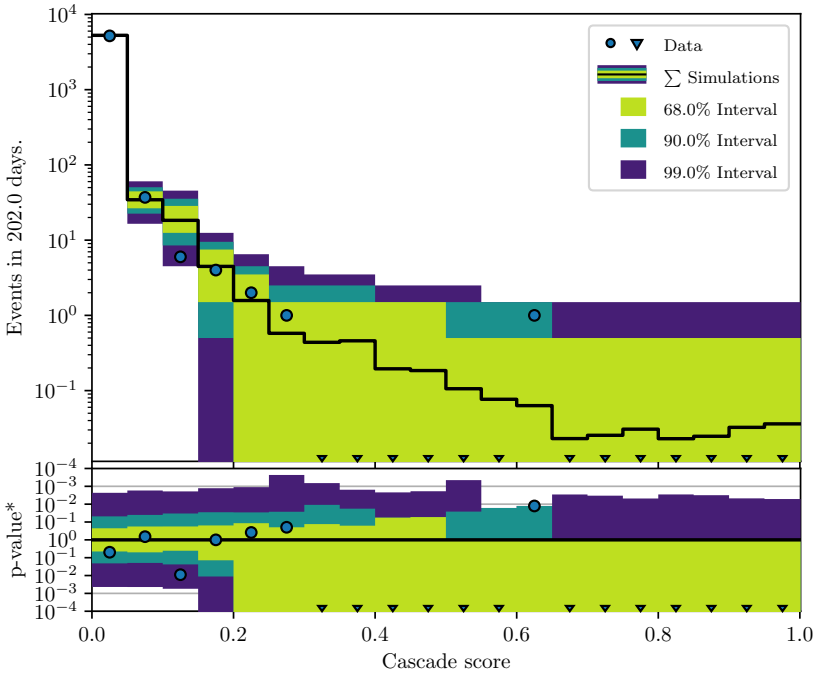
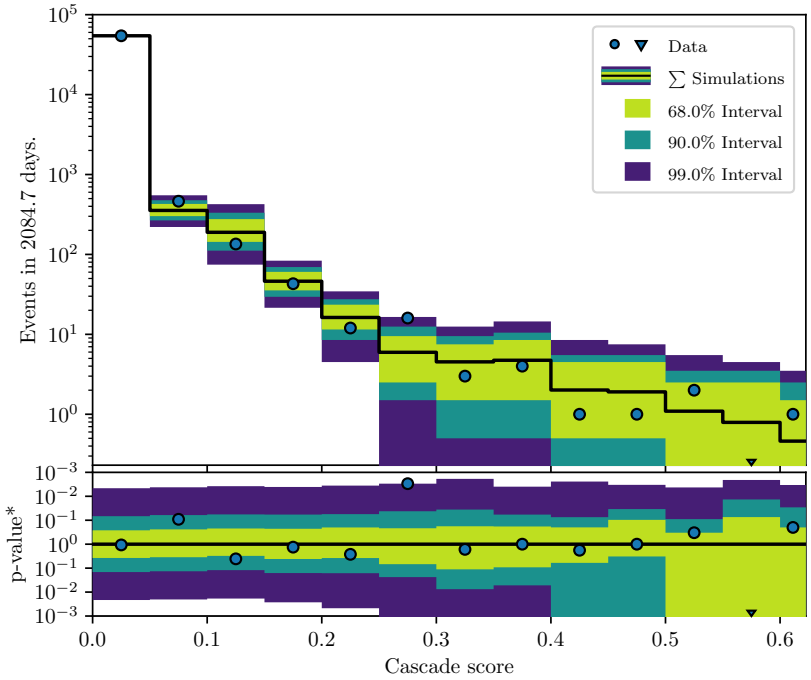


Figure A.4: Feature importances for the classification between cascade-like and track-like events for the first half of the features (cf. Figure A.5).

Figure A.5: Feature importances for the classification between cascade-like and track-like events for the second half of the features (cf. Figure A.4).



(a) Comparison between the classification score for simulations and the burn-sample from six years of data from 2011 until 2017.



(b) Comparison between the classification score for simulations and data from six years of operation between 2011 and 2017 in the background region.

Figure A.6: Comparisons between data and simulations that were used during the optimization of the event selection. Originally the event selection was developed with 10 % of the data from six years to check for good agreement between data and simulation (top). As an additional check, the background region (Classification score ≤ 0.62) was used to make sure the region close to the signal region also behaves as expected (bottom).

A.5.1 Comparison of Individual Seasons of Data Taking

The Cascade score distributions for individual seasons of data taking are presented in Figures A.7 and A.8. They have been analyzed to identify potential cross-year inconsistencies. The most remarkable feature in these distributions is a large apparent under-fluctuation in the first bin of the last season, IC86-VIII. This can be explained by seasonal variations in the atmospheric muon flux. The production of muons in the atmosphere depends on the atmospheric density, and thus on its temperature.

Figure A.9 shows the observed event rates before the cascade selection for the individual seasons, IC86-I to IC86-VIII. The event rates are split up into two time periods, January to May and June to December. For the last season, IC86-VIII, where the apparent under-fluctuation is observed, only data from June to December is used in this analysis. This time period consistently shows a lower event rate over all seasons compared to the other time period. In this thesis, all atmospheric fluxes are weighted to their yearly average, resulting in an apparent under-fluctuation for the season IC86-VIII.

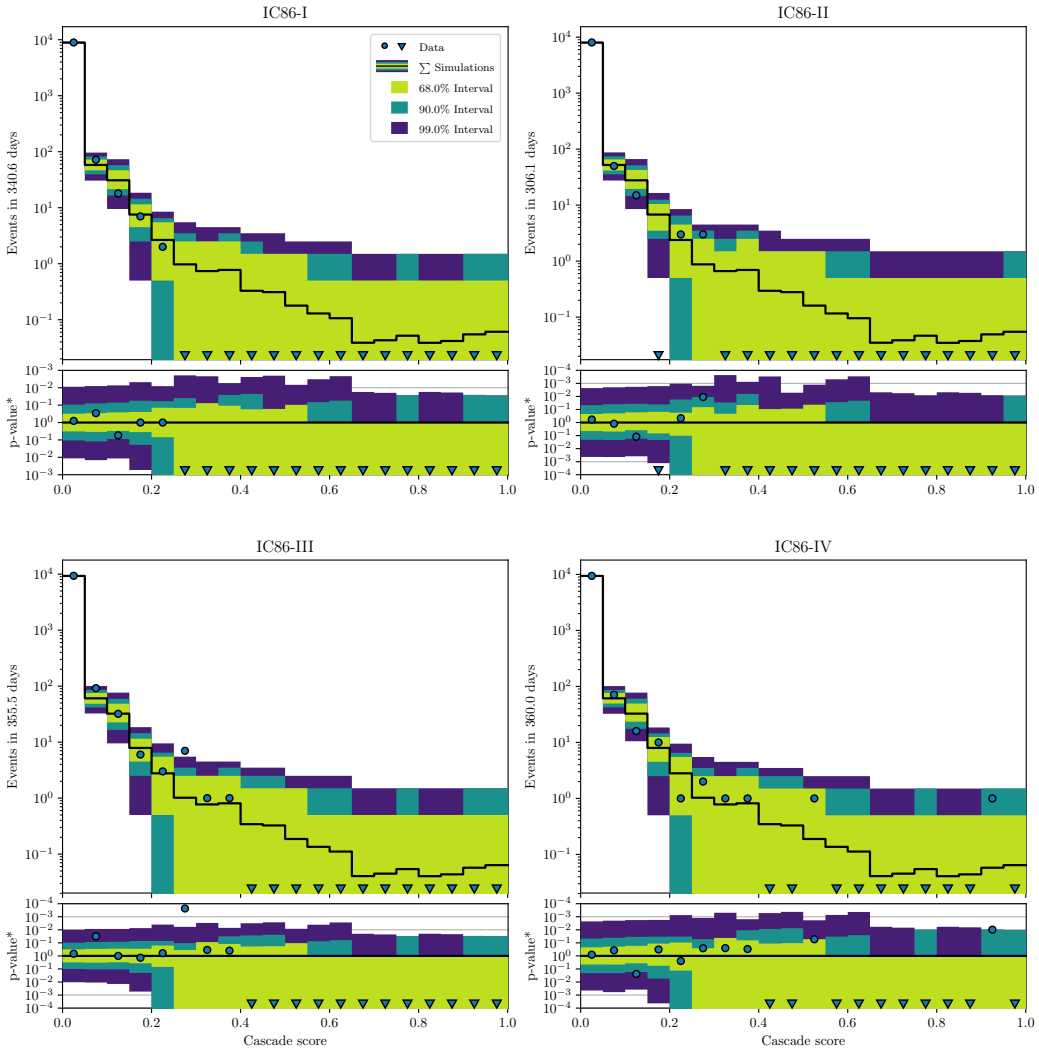


Figure A.7: Cascade score distributions for individual seasons of data taking (IC86-I - IC86-IV).

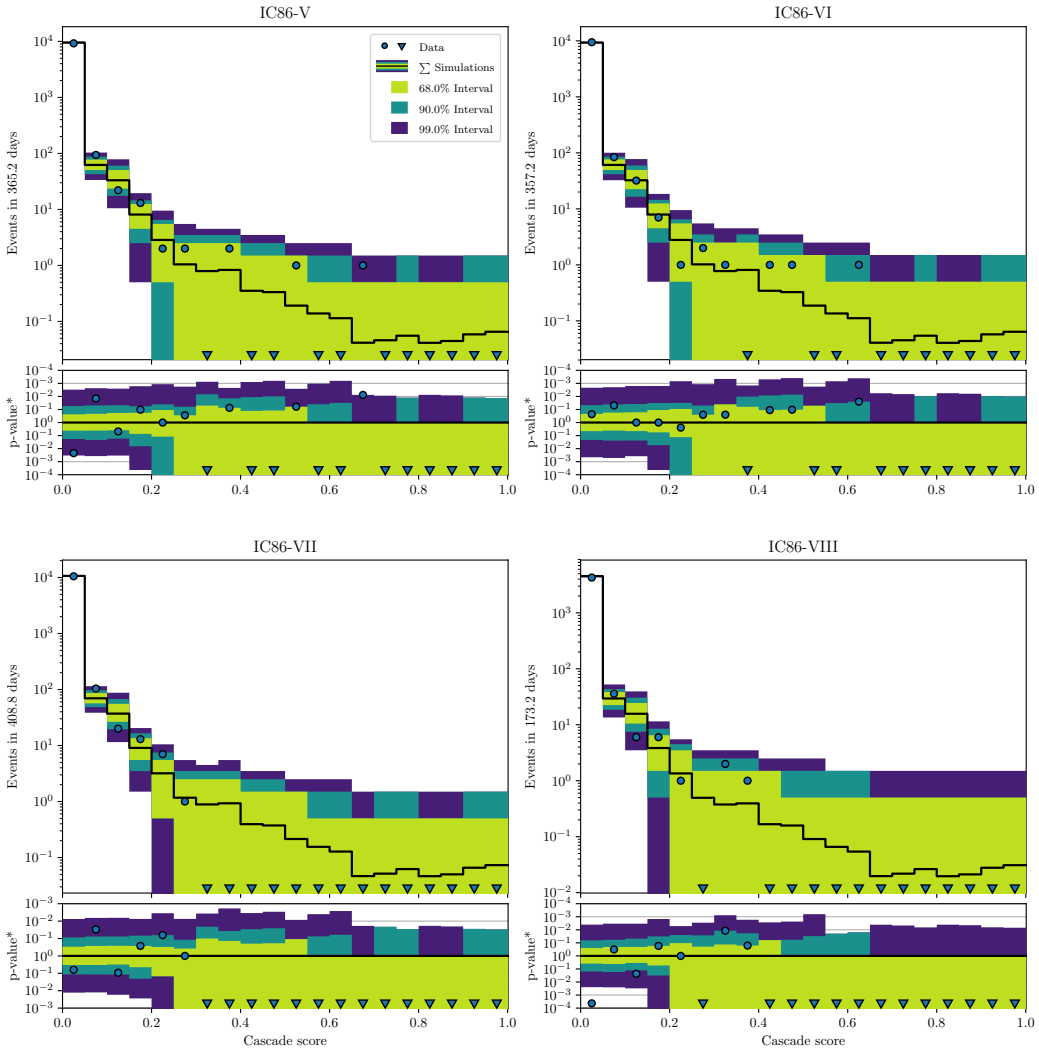


Figure A.8: Cascade score distributions for individual seasons of data taking (IC86-V - IC86-VIII).

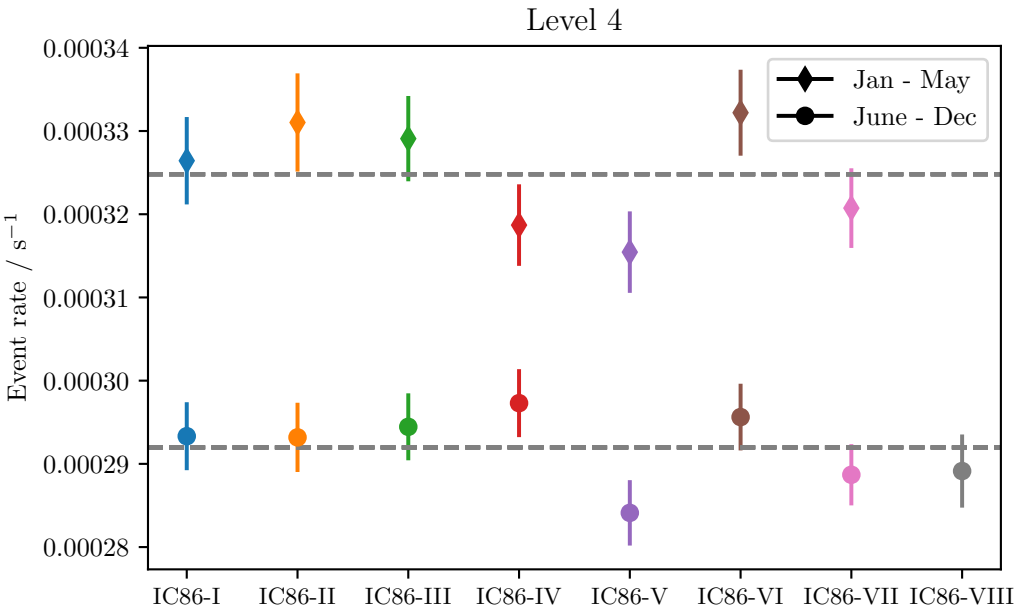


Figure A.9: Observed event rates before the cascade selection (at Level 4) for the individual seasons of data taking (IC86-I - IC86-VIII) broken down into the two time periods, January to May and June to December. For the last year, IC86-VIII, only data from June to December is used. The dashed gray lines show the mean event rate for the respective time period.

A.6 Final Sample

A.6.1 Expected Event Rates

Table A.2: Event expectations and relative fractions at the final level for data and all simulated components in 7.3 years for the Diffuse ν_μ analysis as stated in Table 5.1.

Component	Events in 7.3 years	Fraction
Data	2	-
ν_τ CC	2.10	65.8 %
ν_μ CC	0.24	7.5 %
ν_e CC + GR	0.13	4.1 %
$\sum_X \nu_X$ NC	0.05	1.6 %
Atmos. μ	0.67	21.0 %

Table A.3: Event expectations and relative fractions at the final level for data and all simulated components in 7.3 years for the Global Fit analysis as stated in Table 5.1.

Component	Events in 7.3 years	Fraction
Data	2	-
ν_τ CC	2.48	69.3 %
ν_μ CC	0.25	7.0 %
ν_e CC + GR	0.13	3.6 %
$\sum_X \nu_X$ NC	0.05	1.4 %
Atmos. μ	0.67	18.7 %

Table A.4: Event expectations and relative fractions at the final level for data and all simulated components in 7.3 years for the HESE analysis as stated in Table 5.1.

Component	Events in 7.3 years	Fraction
Data	2	-
ν_τ CC	1.26	54.8 %
ν_μ CC	0.20	8.7 %
ν_e CC + GR	0.12	5.2 %
$\sum_X \nu_X$ NC	0.05	2.2 %
Atmos. μ	0.67	29.1 %

A.7 Analysis Method

A.7.1 Event-wise classification

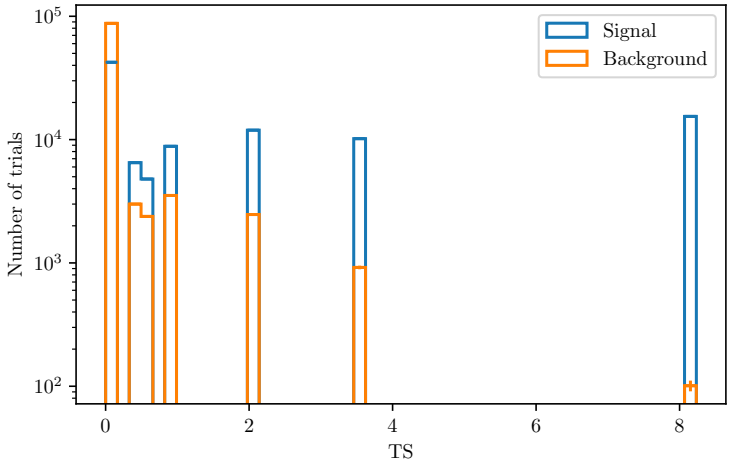


Figure A.10: Test statistic distributions for the p-value calculation for events sampled from the background PDF (orange) and events sampled from the signal PDF (blue) assuming an astrophysical $E^{-2.50}$ spectrum.

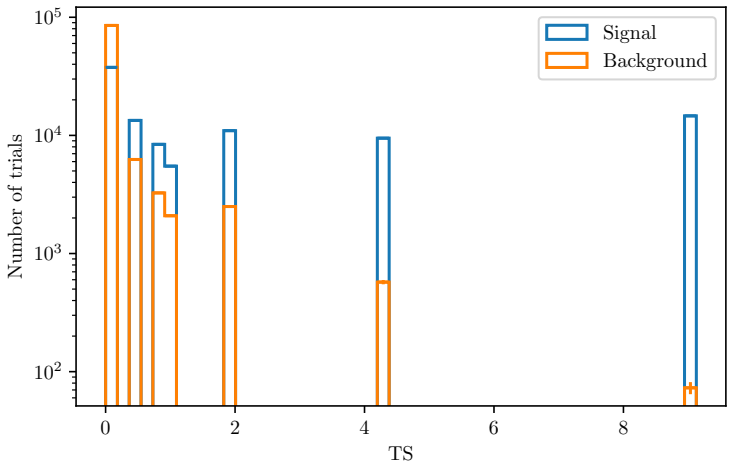


Figure A.11: Test statistic distributions for the p-value calculation for events sampled from the background PDF (orange) and events sampled from the signal PDF (blue) assuming an astrophysical $E^{-2.89}$ spectrum.

A.8 Results

A.8.1 Astrophysical Tau Neutrino Flux Normalization

Influence of Classification Score Cut Variations on the Tau Neutrino Flux Normalization

The event selection for this analysis was optimized with only a fraction of the atmospheric muon simulation available at the end. The test presented here accesses the possible influence of the additional simulation on the significance of the result. And, in general, the robustness of the result with regards to changes in the classification score cut.

To conduct this test, the observable binning, described in section 5.4, is modified and the Poisson likelihood fit, described in section 5.5, is repeated with three different configurations for the observable binning. The equidistant bin width for the classification score described in the main matter is kept, and one bin is added and removed respectively. The result of the three fits, all assuming an astrophysical $E^{-2.19}$ spectrum, are compared in Table A.6. The observed result is almost independent of the chosen classification score cut, and thus, of the number of chosen classification score bins.

Table A.5: Study of the overlap between the double pulse analysis (this work) and the double cascade analysis [103]. These numbers are obtained by comparing the effect of both event selections on the same set of simulated events assuming astrophysical neutrinos with a spectrum of $E^{-2.50}$ only (neglecting the contributions from atmospheric neutrinos). The relative overlap is calculated by dividing the overlapping event expectation with the respective event expectation for the analysis in the 2nd and 3rd column. Atmospheric muons are not included in this study, because both analyses can not be compared fairly since this analysis uses atmospheric muon simulations specifically biased for the event selection presented here. Overall atmospheric muons are a major background component in this analysis, while they are rather negligible in the double cascade analysis.

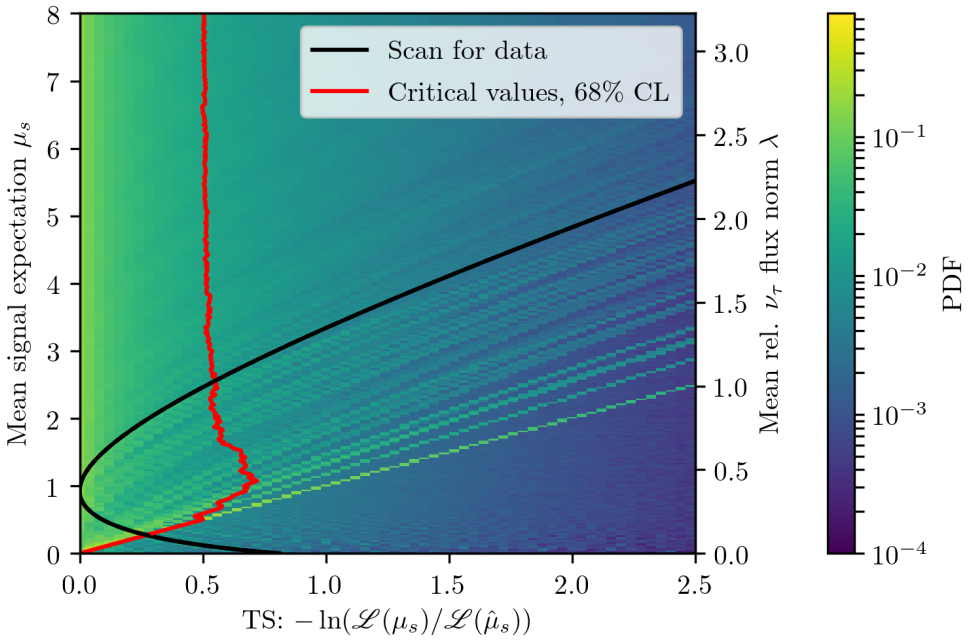
Component	Overlap w.r.t. this analysis	Overlap w.r.t [103]
ν_τ CC	36.2 %	27.1 %
ν_μ CC	4.0 %	0.4 %
ν_e CC	9.5 %	5.5 %
ν_e GR	1.6 %	0.2 %
$\sum_X \nu_X$ NC	0.1 %	0.0 %

Table A.6: Measurement of the astrophysical tau neutrino flux normalization for different classification score cuts, and thus different binning choices. The astrophysical flux is assumed to be $E^{-2.19}$. The uncertainties are given at the 68 % CL.

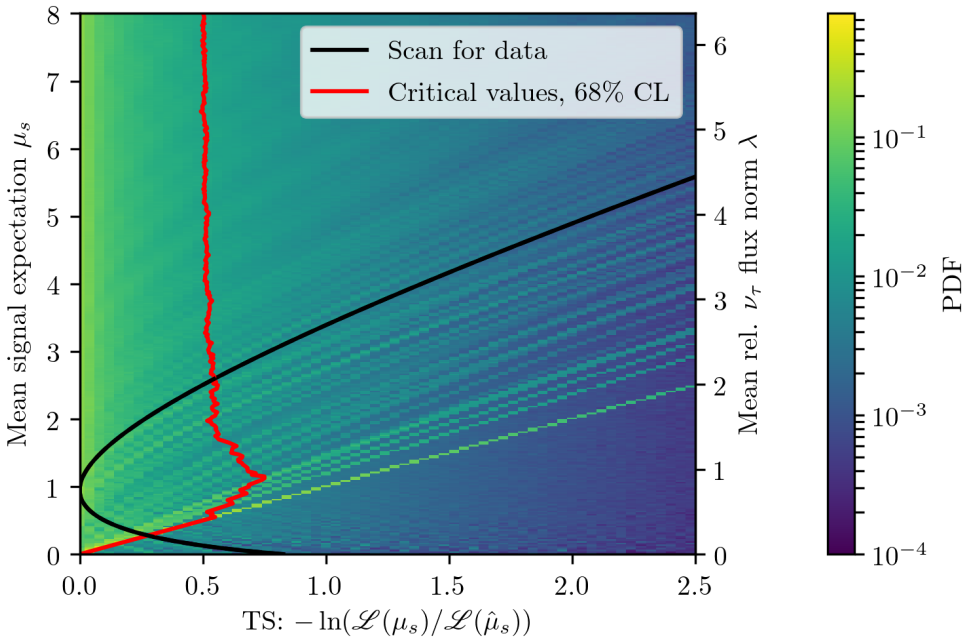
Number of bins	Φ_{0,ν_τ} at 100 TeV ($10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}$)	Significance / σ
8 bins	$0.42^{+0.74}_{-0.28}$	1.90
7 bins	$0.44^{+0.78}_{-0.31}$	1.91
6 bins	$0.42^{+0.79}_{-0.29}$	1.83

Table A.7: Measurement of the astrophysical tau neutrino flux normalization presented in chapter 6 for the different astrophysical spectra. Statistical uncertainties are given at the 68 % and 90 % confidence level.

γ	Φ_{0,ν_τ} at 100 TeV / $(10^{-18}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1})$	Uncertainties at 68 % CL	Uncertainties at 90 % CL
2.19	0.44	$^{+0.78}_{-0.31}$	$^{+1.48}_{-0.40}$
2.50	0.83	$^{+1.46}_{-0.59}$	$^{+2.76}_{-0.74}$
2.89	1.62	$^{+2.78}_{-1.11}$	$^{+5.26}_{-1.45}$



(a) Likelihood scan for an astrophysical spectrum of $E^{-2.5}$.



(b) Likelihood scan for an astrophysical spectrum of $E^{-2.89}$.

Figure A.12: Likelihood scans for the observed data sample. The red line shows critical values at the 68 % confidence level obtained from the underlying test statistic distributions for each value of μ_s.

A.8.2 Impact of Systematic Uncertainties

Table A.8: Relative influence of the tested sources of systematic uncertainties on the lower and upper bound ($\Phi^{l,68}_{\nu_\tau}$ and $\Phi^{u,68}_{\nu_\tau}$) of the 68 % confidence interval. The last column quantifies the impact on the p-value to reject a tau neutrino flux normalization of $\Phi_{\nu_\tau} = 0$. The values shown here are calculated for an astrophysical $E^{-2.50}$ spectrum. Nominal values are shown in Tab. 5.3.

Description	Variation	$\Delta\Phi^{l,68}_{\nu_\tau}/\Phi^{l,68}_{\nu_\tau}$	$\Delta\Phi^{u,68}_{\nu_\tau}/\Phi^{u,68}_{\nu_\tau}$	$\Delta p_{\Phi_{\nu_\tau}>0}/p_{\Phi_{\nu_\tau}>0}$
scattering (bulk ice)	+10 %	+27.2 %	+18.9 %	−10.2 %
absorption (bulk ice)	+10 %	+22.9 %	+11.7 %	+0.1 %
scat. and abs.	−7 %	+19.6 %	+9.7 %	−9.4 %
hole ice	$p_2 = -3$	−1.9 %	−3.8 %	+2.5 %
	$p_2 = -1$	+10.4 %	+1.2 %	−2.1 %
	$p_2 = +1$	+0.1 %	−0.8 %	+17.5 %
DOM efficiency	+10 %	−4.8 %	−5.7 %	+14.0 %
	−10 %	+3.6 %	+3.6 %	+15.1 %
$\Phi_{\nu,\text{conv}}$	+30 %	0.0 %	−1.0 %	−4.3 %
	−30 %	0.0 %	0.0 %	+3.9 %
$\Phi_{\nu,\text{prompt}}$	BERSS_u [84]	0.0 %	0.0 %	−2.0 %
	BERSS_l [84]	0.0 %	0.0 %	+0.2 %
$\Phi_{\nu,\text{astro}}$	+0.37	0.0 %	0.0 %	0.0 %
	−0.40	+9.1 %	0.0 %	+6.7 %
γ_{astro}	+0.09	+9.1 %	0.0 %	+4.3 %
	−0.09	0.0 %	0.0 %	+1.1 %
Φ_μ	+15 %	0.0 %	−1.0 %	−1.8 %
	−15 %	+9.1 %	+1.0 %	+3.1 %

Table A.9: Relative influence of the tested sources of systematic uncertainties on the lower and upper bound ($\Phi^{l,68}_{\nu_\tau}$ and $\Phi^{u,68}_{\nu_\tau}$) of the 68 % confidence interval. The last column quantifies the impact on the p-value to reject a tau neutrino flux normalization of $\Phi_{\nu_\tau} = 0$. The values shown here are calculated for an astrophysical $E^{-2.89}$ spectrum. Nominal values are shown in Tab. 5.3.

Description	Variation	$\Delta\Phi^{l,68}_{\nu_\tau}/\Phi^{l,68}_{\nu_\tau}$	$\Delta\Phi^{u,68}_{\nu_\tau}/\Phi^{u,68}_{\nu_\tau}$	$\Delta p_{\Phi_{\nu_\tau}>0}/p_{\Phi_{\nu_\tau}>0}$
scattering (bulk ice)	+10 %	+27.2 %	+18.9 %	+9.4 %
absorption (bulk ice)	+10 %	+22.9 %	+12.8 %	+1.3 %
scat. and abs.	−7 %	+19.6 %	+8.7 %	+15.3 %
hole ice	$p_2 = -3$	−1.9 %	−2.9 %	+23.9 %
	$p_2 = -1$	+10.4 %	+1.3 %	+17.9 %
	$p_2 = +1$	+0.1 %	−0.8 %	+19.6 %
DOM efficiency	+10 %	−4.8 %	−5.7 %	+23.3 %
	−10 %	+3.6 %	+3.6 %	+10.2 %
$\Phi_{\nu,\text{conv}}$	+30 %	−8.3 %	−1.0 %	−2.6 %
	−30 %	0.0 %	0.0 %	+4.9 %
$\Phi_{\nu,\text{prompt}}$	BERSS_u [84]	−8.3 %	0.0 %	−1.7 %
	BERSS_l [84]	0.0 %	0.0 %	+0.3 %
$\Phi_{\nu,\text{astro}}$	+0.26	0.0 %	0.0 %	0.0 %
	−0.23	0.0 %	+1.9 %	+12.6 %
γ_{astro}	+0.31	0.0 %	+1.0 %	+0.4 %
	−0.28	−8.3 %	0.0 %	+5.5 %
Φ_μ	+15 %	−8.3 %	−1.9 %	−1.7 %
	−15 %	0.0 %	+2.9 %	+0.9 %

B Acknowledgments

Firstly, I would like to thank my advisor Dr. Dr. Wolfgang Rhode for his continuous support of my research. Further, I would like to thank the members of my dissertation committee: Dr. Bernhard Spaan, Dr. Dmitri Yakovlev, and Dr. Gerald Schmidt. Similarly, I appreciate the time invested by Dr. Shigeru Yoshida and Dr. Lu Lu for their invaluable feedback during the internal IceCube review process of my analysis. Both were also great hosts during my stay at Chiba University, which was generously funded by the DAAD. I learned a lot and had a great time in Chiba thanks to all members of the Neutrino Astronomy Group.

I also want to thank my colleagues from the astroparticle physics group at TU Dortmund, who helped me in many enlightening discussions. Especially, I want to thank the people, who directly contributed to this thesis by proof-reading: Thorben Menne, Jens Stücker, Alexander Sandrock, Mathis Börner, Mirco Hünnefeld, Philipp Schlunder, Simone Mender, Jan Soedingrekso, Alicia Fattorini, Kevin Schmidt and William Martin.

I would also like to thank the entire IceCube collaboration for their constructive feedback and support during the last years. I am especially thankful for a lot of helpful discussions with Dr. Hans Niederhausen.

Lastly, I want to thank my family, especially my parents, for supporting all of my decisions during my doctoral study and all of my academic education. I also want to thank Verena Klein for supporting me and having my back during most of the time of my research.

Bibliography

[1] U. F. Katz and Ch. Spiering. "High-Energy Neutrino Astrophysics: Status and Perspectives". In: *Prog. Part. Nucl. Phys.* 67 (2012), pp. 651–704. DOI: 10.1016/j.ppnp.2011.12.001.

[2] M. G. Aartsen et al. "Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector". In: *Science* 342 (2013). DOI: 10.1126/science.1242856.

[3] M. G. Aartsen et al. "Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data". In: *Phys. Rev. Lett.* 113 (2014), p. 101101. DOI: 10.1103/PhysRevLett.113.101101.

[4] M. G. Aartsen et al. "A Measurement of the Diffuse Astrophysical Muon Neutrino Flux Using Eight Years of IceCube Data". In: *PoS(ICRC2017)* 1005 (2017). DOI: 10.22323/1.301.1005.

[5] M. G. Aartsen et al. "A combined maximum-likelihood analysis of the high-energy astrophysical neutrino flux measured with IceCube". In: *Astrophys. J.* 809.1 (2015), p. 98. DOI: 10.1088/0004-637X/809/1/98.

[6] Hans Niederhausen. "Measurement of the High Energy Astrophysical Neutrino Flux Using Electron and Tau Neutrinos Observed in Four Years of IceCube Data". PhD thesis. Stony Brook University, 2018.

[7] M. G. Aartsen et al. "Characterization of the Astrophysical Diffuse Neutrino Flux with High-Energy Starting Events and Prospects for Future Measurements with IceCube". In: *PoS(ICRC2019)* 1004 (2019). DOI: TBD.

[8] M. G. Aartsen et al. "Atmospheric and astrophysical neutrinos above 1 TeV interacting in IceCube". In: *Phys. Rev.* D91.2 (2015), p. 022001. DOI: 10.1103/PhysRevD.91.022001.

[9] IceCube Collaboration et al. "Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert". In: *Science* 361.6398 (July 2018), pp. 147–151. DOI: 10.1126/science.aat2890.

[10] M. G. Aartsen et al. "Search for Astrophysical Tau Neutrinos in Three Years of IceCube Data". In: *Phys. Rev.* D93.2 (2016), p. 022001. DOI: 10.1103/PhysRevD.93.022001.

[11] M. G. Aartsen et al. “Search for Astrophysical Tau Neutrinos in Six Years of High-Energy Starting Events in IceCube”. In: *PoS(ICRC2017)* 974 (2018). DOI: 10.22323/1.301.0974.

[12] D F Cowen and the IceCube Collaboration. “Tau Neutrinos in IceCube”. In: *Journal of Physics: Conference Series* 60.1 (2007), p. 227. DOI: 10.1088/1742-6596/60/1/048.

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

[14] A. Aab et al. “Depth of maximum of air-shower profiles at the Pierre Auger Observatory. II. Composition implications”. In: *Phys. Rev. D* 90 (12 Dec. 2014), p. 122006. DOI: 10.1103/PhysRevD.90.122006.

[15] R. U. Abbasi et al. “Study of Ultra-High Energy Cosmic Ray composition using Telescope Array’s Middle Drum detector and surface array in hybrid mode”. In: *Astroparticle Physics* 64 (Apr. 2015), pp. 49–62. DOI: 10.1016/j.astropartphys.2014.11.004.

[16] R. U. Abbasi et al. “Indications of Proton-Dominated Cosmic-Ray Composition above 1.6 EeV”. In: *Phys. Rev. Lett.* 104 (16 Apr. 2010), p. 161101. DOI: 10.1103/PhysRevLett.104.161101.

[17] Kenneth Greisen. “End to the Cosmic-Ray Spectrum?” In: *Phys. Rev. Lett.* 16 (17 Apr. 1966), pp. 748–750. DOI: 10.1103/PhysRevLett.16.748.

[18] G. T. Zatsepin and V. A. Kuz’min. “Upper Limit of the Spectrum of Cosmic Rays”. In: *Soviet Journal of Experimental and Theoretical Physics Letters* 4 (Aug. 1966), p. 78.

[19] V.S. Beresinsky and G.T. Zatsepin. “Cosmic rays at ultra high energies (neutrino?)” In: *Physics Letters B* 28.6 (1969), pp. 423–424. DOI: 10.1016/0370-2693(69)90341-4.

[20] M. G. Aartsen et al. “Differential limit on the extremely-high-energy cosmic neutrino flux in the presence of astrophysical background from nine years of IceCube data”. In: *Phys. Rev.* D98.6 (2018), p. 062003. DOI: 10.1103/PhysRevD.98.062003.

[21] J. J. Engelmann et al. “Charge composition and energy spectra of cosmic-ray nuclei for elements from Be to NI - Results from HEAO-3-C2”. In: *Astronomy and Astrophysics* 233 (July 1990), pp. 96–111.

[22] O. Adriani et al. “PAMELA Measurements of Cosmic-Ray Proton and Helium Spectra”. In: *Science* 332.6025 (Apr. 2011), p. 69. DOI: 10.1126/science.1199172.

[23] O. Adriani et al. "Time Dependence of the Proton Flux Measured by PAMELA during the 2006 July-2009 December Solar Minimum". In: *Astrophysical Journal* 765.2, 91 (Mar. 2013). DOI: 10.1088/0004-637X/765/2/91.

[24] M. Aguilar et al. "Precision Measurement of the Proton Flux in Primary Cosmic Rays from Rigidity 1 GV to 1.8 TV with the Alpha Magnetic Spectrometer on the International Space Station". In: *Phys. Rev. Lett.* 114 (17 Apr. 2015), p. 171103. DOI: 10.1103/PhysRevLett.114.171103.

[25] M. Aguilar et al. "Precision Measurement of the Helium Flux in Primary Cosmic Rays of Rigidities 1.9 GV to 3 TV with the Alpha Magnetic Spectrometer on the International Space Station". In: *Phys. Rev. Lett.* 115 (21 Nov. 2015), p. 211101. DOI: 10.1103/PhysRevLett.115.211101.

[26] H. S. Ahn et al. "Energy Spectra of Cosmic-ray Nuclei at High Energies". In: *Astrophysical Journal* 707.1 (Dec. 2009), pp. 593–603. DOI: 10.1088/0004-637X/707/1/593.

[27] Y. S. Yoon et al. "Cosmic-ray Proton and Helium Spectra from the First CREAM Flight". In: *Astrophysical Journal* 728.2, 122 (Feb. 2011). DOI: 10.1088/0004-637X/728/2/122.

[28] P. Montini and S. M. Mari. "The bending of the proton plus helium flux in primary cosmic rays measured by the ARGO-YBJ experiment in the energy range from 20 TeV to 5 PeV". In: (Aug. 2016). arXiv: 1608.01389.

[29] V.V. Prosin et al. "Tunka-133: Results of 3 year operation". In: *Nucl. Instrum. Methods Phys. Res. A* 756 (2014), pp. 94–101. DOI: 10.1016/j.nima.2013.09.018.

[30] E.E. Korosteleva et al. "Measurement of Cosmic Ray Primary Energy with the Atmospheric Cherenkov Light Technique in Extensive Air Showers". In: *Nuclear Physics B - Proceedings Supplements* 165 (2007). Proceedings of the Cosmic Ray International Seminars, pp. 74–80. DOI: 10.1016/j.nuclphysbps.2006.11.012.

[31] Katherine Rawlins and T. Feusels. "Latest Results on Cosmic Ray Spectrum and Composition from Three Years of IceTop and IceCube". In: *PoS(ICRC2015)* (2016), p. 334. DOI: 10.22323/1.236.0334.

[32] The KASCADE-Grande Collaboration et al. "The spectrum of high-energy cosmic rays measured with KASCADE-Grande". In: (June 2012). arXiv: 1206.3834.

[33] S. Schoo et al. "The energy spectrum of cosmic rays in the range from 10^{14} to 10^{18} eV". In: *PoS(ICRC2015)* 263 (2015). DOI: 10.22323/1.236.0263.

[34] Dmitri Ivanov. "TA Spectrum Summary". In: *PoS(ICRC2015)* (2016), p. 349. DOI: 10.22323/1.236.0349.

[35] Ines Valino. "The flux of ultra-high energy cosmic rays after ten years of operation of the Pierre Auger Observatory". In: *PoS(ICRC2015)* 271 (2016). DOI: 10.22323/1.236.0271.

[36] Hans Peter Dembinski et al. "Data-driven model of the cosmic-ray flux and mass composition from 10 GeV to 10^{11} GeV". In: *PoS(ICRC2017)* 533 (2018). DOI: 10.22323/1.301.0533.

[37] Luis A. Anchordoqui. "Ultra-high-energy cosmic rays". In: *Physics Reports* 801 (Apr. 2019), pp. 1–93. DOI: 10.1016/j.physrep.2019.01.002.

[38] E. Fermi. "Galactic Magnetic Fields and the Origin of Cosmic Radiation." In: *Astrophysical Journal* 119 (Jan. 1954), p. 1. DOI: 10.1086/145789.

[39] Rafael Alves Batista and Günter Sigl. "Diffusion of cosmic rays at EeV energies in inhomogeneous extragalactic magnetic fields". In: *Journal of Cosmology and Astro-Particle Physics* 2014.11, 031 (Nov. 2014). DOI: 10.1088/1475-7516/2014/11/031.

[40] Roberto Aloisio. "Acceleration and propagation of ultra-high energy cosmic rays". In: *Progress of Theoretical and Experimental Physics* 2017.12 (Nov. 2017). DOI: 10.1093/ptep/ptx115.

[41] Tomasz Fuchs. "Charmante Myonen im Eis. Messung des hochenergetischen atmosphärischen Myon-Energiespektrums mit IceCube in der Detektorkonfiguration IC86-I". PhD thesis. TU Dortmund, 2016.

[42] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.

[43] Fedynitch, Anatoli et al. "Calculation of conventional and prompt lepton fluxes at very high energy". In: *EPJ Web of Conferences* 99 (2015), p. 08001. DOI: 10.1051/epjconf/20159908001.

[44] Fréjus Collaboration et al. "Determination of the atmospheric neutrino spectra with the Fréjus detector". In: *Zeitschrift für Physik C Particles and Fields* 66.3 (Sept. 1995), pp. 417–428. DOI: 10.1007/BF01556368.

[45] S. Adrián-Martínez et al. "Measurement of the atmospheric ν_μ energy spectrum from 100 GeV to 200 TeV with the ANTARES telescope". In: *European Physical Journal C* 73, 2606 (Oct. 2013). DOI: 10.1140/epjc/s10052-013-2606-4.

[46] M. G. Aartsen et al. "Development of a general analysis and unfolding scheme and its application to measure the energy spectrum of atmospheric neutrinos with IceCube". In: *The European Physical Journal C* 75.3 (Mar. 2015), p. 116. DOI: 10.1140/epjc/s10052-015-3330-z.

[47] M. G. Aartsen et al. "Measurement of the ν_μ energy spectrum with IceCube-79". In: *The European Physical Journal C* 77.10 (Oct. 2017), p. 692. DOI: 10.1140/epjc/s10052-017-5261-3.

[48] Mathis Börner. "Unfolding measurement of the Atmospheric Neutrino Spectrum using IceCube-79/86". In: Aug. 2016, p. 1098. DOI: 10.22323/1.236.1098.

[49] M. G. Aartsen et al. "Measurement of the Atmospheric ν_e Spectrum with IceCube". In: *Phys. Rev. D* 91 (12 June 2015), p. 122004. DOI: 10.1103/PhysRevD.91.122004.

[50] 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.

[51] Hans Niederhausen and Yiqian Xu. "High Energy Astrophysical Neutrino Flux Measurement Using Neutrino-induced Cascades Observed in 4 Years of IceCube Data". In: *PoS(ICRC2017)* 968 (2018). DOI: 10.22323/1.301.0968.

[52] M. G. Aartsen et al. "Observation of Astrophysical Neutrinos in Six Years of IceCube Data". In: *PoS(ICRC2017)* 981 (2017). DOI: 10.22323/1.301.0981.

[53] Mathis Börner. "Bestimmung des Energiespektrums von atmosphärischen Myonneutrinos mit 3 Jahren Daten des IceCube-Detektors". PhD thesis. TU Dortmund, 2018.

[54] P. Mészáros. "Astrophysical Sources of High-Energy Neutrinos in the IceCube Era". In: *Annual Review of Nuclear and Particle Science* 67 (Oct. 2017), pp. 45–67. DOI: 10.1146/annurev-nucl-101916-123304.

[55] Kohta Murase, Yoshiyuki Inoue, and Charles D. Dermer. "Diffuse neutrino intensity from the inner jets of active galactic nuclei: Impacts of external photon fields and the blazar sequence". In: *Physical Review D* 90.2, 023007 (July 2014). DOI: 10.1103/PhysRevD.90.023007.

[56] Sheldon L. Glashow. "Resonant Scattering of Antineutrinos". In: *Phys. Rev.* 118 (1 Apr. 1960), pp. 316–317. DOI: 10.1103/PhysRev.118.316.

[57] Daniel Biehl et al. "Astrophysical neutrino production diagnostics with the Glashow resonance". In: *Journal of Cosmology and Astro-Particle Physics* 2017.1, 033 (Jan. 2017), p. 033. DOI: 10.1088/1475-7516/2017/01/033.

[58] Jörg P. Rachen and P. Mészáros. "Photohadronic neutrinos from transients in astrophysical sources". In: *Physical Review D* 58.12, 123005 (Dec. 1998). DOI: 10.1103/PhysRevD.58.123005.

[59] Luis A Anchordoqui et al. “Galactic point sources of TeV antineutrinos”. In: *Physics Letters B* 593.1 (2004), pp. 42–47. DOI: https://doi.org/10.1016/j.physletb.2004.04.054.

[60] Rikard Enberg, Mary Hall Reno, and Ina Sarcevic. “High energy neutrinos from charm in astrophysical sources”. In: *Physical Review D* 79.5 (2009), p. 053006. DOI: 10.1103/PhysRevD.79.053006.

[61] Lingjun Fu, Chiu Man Ho, and Thomas J. Weiler. “Aspects of the flavor triangle for cosmic neutrino propagation”. In: *Physical Review D* 91.5, 053001 (Mar. 2015). DOI: 10.1103/PhysRevD.91.053001.

[62] Sandip Pakvasa, Werner Rodejohann, and Thomas J. Weiler. “Flavor ratios of astrophysical neutrinos: implications for precision measurements”. In: *Journal of High Energy Physics* 2008.2, 005 (Feb. 2008). DOI: 10.1088/1126-6708/2008/02/005.

[63] P. F. Harrison, D. H. Perkins, and W. G. Scott. “Tri-bimaximal mixing and the neutrino oscillation data”. In: *Physics Letters B* 530.1-4 (Mar. 2002), pp. 167–173. DOI: 10.1016/S0370-2693(02)01336-9.

[64] A. Achterberg et al. “First year performance of the IceCube neutrino telescope”. In: *Astroparticle Physics* 26.3 (2006), pp. 155–173. DOI: 10.1016/j.astropartphys.2006.06.007.

[65] R. Abbasi et al. “The design and performance of IceCube DeepCore”. In: *Astroparticle Physics* 35.10 (2012), pp. 615–624. DOI: 10.1016/j.astropartphys.2012.01.004.

[66] R. Abbasi et al. “IceTop: The surface component of IceCube”. In: *Nucl. Instrum. Methods Phys. Res. A* 700 (2013), pp. 188–220. DOI: 10.1016/j.nima.2012.10.067.

[67] Christopher Wiebusch and for the IceCube Collaboration. “Physics Capabilities of the IceCube DeepCore Detector”. In: (July 2009). arXiv: 0907.2263.

[68] M. Ackermann et al. “Optical properties of deep glacial ice at the South Pole”. In: *Journal of Geophysical Research: Atmospheres* 111.D13 (2006). DOI: 10.1029/2005JD006687.

[69] M. G. Aartsen et al. “Measurement of South Pole ice transparency with the IceCube LED calibration system”. In: *Nucl. Instrum. Meth.* A711 (2013), pp. 73–89. DOI: 10.1016/j.nima.2013.01.054.

[70] M. G. Aartsen et al. “The IceCube Neutrino Observatory Part VI: Ice Properties, Reconstruction and Future Developments”. In: 2013. arXiv: 1309.7010.

[71] Craig F. Bohren and Donal R. Huffman. *Absorption and Scattering of Light by Small Particles*. Wiley VCH Verlag, 1998. DOI: 10.1002/9783527618156.

[72] P. A. Cherenkov. "Visible emission of clean liquids by action of radiation". In: *Doklady Akademii Nauk SSSR* 2 (1934), p. 451.

[73] M. G. Aartsen et al. "The IceCube Neutrino Observatory: Instrumentation and Online Systems". In: *JINST* 12.03 (2017), P03012. DOI: 10.1088/1748-0221/12/03/P03012.

[74] R. Abbasi et al. "The IceCube data acquisition system: Signal capture, digitization, and timestamping". In: *Nucl. Instrum. Methods Phys. Res. A* 601.3 (2009), pp. 294–316. DOI: 10.1016/j.nima.2009.01.001.

[75] Amanda Cooper-Sarkar, Philipp Mertsch, and Subir Sarkar. "The high energy neutrino cross-section in the Standard Model and its uncertainty". In: *Journal of High Energy Physics* 2011, 42 (Aug. 2011). DOI: 10.1007/JHEP08(2011)042.

[76] Maximilian Meier. "Berechnung von Tau-Neutrino-Eereignisraten für den IceCube Detektor". Master thesis. TU Dortmund, 2015.

[77] J.-H. Koehne et al. "PROPOSAL: A tool for propagation of charged leptons". In: *Computer Physics Communications* 184.9 (2013), pp. 2070–2090. DOI: 10.1016/j.cpc.2013.04.001.

[78] D. Heck et al. *CORSIKA: a Monte Carlo code to simulate extensive air showers*. Feb. 1998.

[79] Jakob van Santen. "Neutrino Interactions in IceCube above 1 TeV". PhD thesis. University of Wisconsin-Madison, 2014.

[80] Thomas K. Gaisser. "Spectrum of cosmic-ray nucleons, kaon production, and the atmospheric muon charge ratio". In: *Astroparticle Physics* 35.12 (2012), pp. 801–806. DOI: 10.1016/j.astropartphys.2012.02.010.

[81] Eun-Joo Ahn et al. "Cosmic ray interaction event generator SIBYLL 2.1". In: *Phys. Rev. D* 80 (9 Nov. 2009), p. 094003. DOI: 10.1103/PhysRevD.80.094003.

[82] P Berghaus, T Montaruli, and J Ranft. "Charm production in DPMJET". In: *Journal of Cosmology and Astroparticle Physics* 2008.06 (June 2008), p. 003. DOI: 10.1088/1475-7516/2008/06/003.

[83] M. Honda et al. "Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data". In: *Phys. Rev. D* 75 (4 Feb. 2007), p. 043006. DOI: 10.1103/PhysRevD.75.043006.

[84] Atri Bhattacharya et al. "Perturbative charm production and the prompt atmospheric neutrino flux in light of RHIC and LHC". In: *Journal of High Energy Physics* 2015.6 (June 2015), p. 110. DOI: 10.1007/JHEP06(2015)110.

[85] Donglian Xu. "Search for Astrophysical Tau Neutrinos in Three Years of IceCube Data". PhD thesis. University of Alabama, 2015.

[86] Dawn Williams, Donglian Xu, and Pavel Zarzhitsky. "Detecting Tau Neutrinos in IceCube with Double Pulses". In: *PoS(ICRC2013)* 0643 (2013).

[87] Carlos A. Argüelles, Austin Schneider, and Tianlu Yuan. "A binned likelihood for stochastic models". In: (2019). arXiv: 1901.04645.

[88] R. Aggarwal and A. Caldwell. "Error Bars for Distributions of Numbers of Events". In: *Eur. Phys. J. Plus* 127 (2012), p. 24. DOI: 10.1140/epjp/i2012-12024-0.

[89] Dirk Pandel. "Bestimmung von Wasser- und Detektorparametern und Rekonstruktion von Myonen bis 100 TeV mit dem Baikal-Neutrinoteleskop NT-72". Diploma thesis. Humboldt-Universität zu Berlin, 1996.

[90] J. Ahrens et al. "Muon track reconstruction and data selection techniques in AMANDA". In: *Nucl. Instrum. Methods Phys. Res. A* 524 (2004), pp. 169–194. DOI: 10.1016/j.nima.2004.01.065.

[91] Robert Johannes Lauer. "Extending the search for cosmic point sources of neutrinos with IceCube beyond PeV energies and above the horizon". PhD thesis. Humboldt-Universität zu Berlin, 2010.

[92] V. M. Aynutdinov et al. "Technique of Neutrino-Induced Muon Detection on the Earth Surface". In: *Physics of Particles and Nuclei Letters* 109 (2001), p. 43.

[93] Johan Lundberg et al. "Light tracking for glaciers and oceans: Scattering and absorption in heterogeneous media with Photonics". In: *Nucl. Instrum. Methods Phys. Res. A* 581 (2007), pp. 619–631. DOI: 10.1016/j.nima.2007.07.143.

[94] Nathan Whitehorn, Jakob van Santen, and Sven Lafebre. "Penalized Splines for Smooth Representation of High-dimensional Monte Carlo Datasets". In: *Comput. Phys. Commun.* 184 (2013), pp. 2214–2220. DOI: 10.1016/j.cpc.2013.04.008.

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

[96] Douglas Lowery Rutledge. "A Search for Neutrino-Induced Electromagnetic Showers in the 2008 Combined IceCube and AMANDA Detectors". PhD thesis. Pennsylvania State University, 2011.

[97] A. M. Hillas. "Cerenkov light images of EAS produced by primary gamma". In: *International Cosmic Ray Conference* 3 (Aug. 1985).

[98] H. Anderhub et al. "Design and operation of FACT – the first G-APD Cherenkov telescope". In: *Journal of Instrumentation* 8.06 (2013), p. 06008. DOI: 10.1088/1748-0221/8/06/P06008.

[99] J. Aleksić et al. "Performance of the MAGIC stereo system obtained with Crab Nebula data". In: *Astroparticle Physics* 35.7 (2012), pp. 435–448. DOI: 10.1016/j.astropartphys.2011.11.007.

[100] J.A Hinton. "The status of the HESS project". In: *New Astronomy Reviews* 48.5–6 (2004), pp. 331–337. DOI: 10.1016/j.newar.2003.12.004.

[101] Patrick Hallen. "On the Measurement of High-Energy Tau Neutrinos with IceCube". Master thesis. RWTH Aachen University, 2013.

[102] Marcel Usner. "Search for Astrophysical Tau-Neutrinos in Six Years of High-Energy Starting Events in the IceCube Detector". PhD thesis. Humboldt-Universität zu Berlin, 2018.

[103] M. G. Aartsen et al. "IceCube High Energy Starting Events at 7.5 Years - New Measurements of Flux and Flavor". In: *EPJ Web Conf.* 207 (2019), p. 02005. DOI: 10.1051/epjconf/201920702005.

[104] M. Börner et al. "Measurement/Simulation Mismatches and Multivariate Data Discretization in the Machine Learning Era". In: *Astronomical Data Analysis Software and Systems (ADASS XXVII)*. Accepted for publication. 2017.

[105] Tom Fawcett. "An introduction to ROC analysis". In: *Pattern Recognition Letters* 27.8 (2006). ROC Analysis in Pattern Recognition, pp. 861–874. DOI: 10.1016/j.patrec.2005.10.010.

[106] C. Ding and H. Peng. "Minimum Redundancy Feature Selection from Microarray Gene Expression Data". In: *Journal of Bioinformatics and Computational Biology* 03.02 (Apr. 2005), pp. 185–205. DOI: 10.1142/S0219720005001004.

[107] Carl Friedrich Gauss. "Bestimmung der Genauigkeit von Beobachtungen". In: *Zeitschrift für Astronomnie und Verwandte Wissenschaften* 1 (1816), pp. 185–197.

[108] Nicolas De Jay et al. "mRMRe: an R package for parallelized mRMR ensemble feature selection". In: *Bioinformatics* 29.18 (2013), pp. 2365–2368. DOI: 10.1093/bioinformatics/btt383.

[109] Ludmila I. Kuncheva. "A Stability Index for Feature Selection". In: *Proceedings of the 25th Conference on Proceedings of the 25th IASTED International Multi-Conference: Artificial Intelligence and Applications*. Innsbruck, Austria: ACTA Press, 2007, pp. 390–395. URL: http://dl.acm.org/citation.cfm?id=1295303.1295370.

[110] Gary C. Hill and Katherine Rawlins. "Unbiased cut selection for optimal upper limits in neutrino detectors: The Model rejection potential technique". In: *Astropart. Phys.* 19 (2003), pp. 393–402. DOI: 10.1016/S0927-6505(02)00240-2.

[111] Gary J. Feldman and Robert D. Cousins. "A Unified approach to the classical statistical analysis of small signals". In: *Phys. Rev.* D57 (1998), pp. 3873–3889. DOI: 10.1103/PhysRevD.57.3873.

[112] Logan James Wille. "The Search for High Energy Tau Neutrinos using the IceCube Neutrino Observatory". PhD thesis. University of Wisconsin-Madison, 2019.

[113] Simon De Kockere. "A Search for Astrophysical Tau Neutrinos with IceCube using an Advanced Statistical Approach". Master thesis. Ghent University, 2018.

[114] George Casella and Roger L. Berger. *Statistical Inference*. 2nd. Duxbury, 2002.

[115] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *Annals Math. Statist.* 9.1 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360.

[116] Luis A. Anchordoqui et al. "Neutrino bounds on astrophysical sources and new physics". In: *Phys. Rev. D* 66 (10 Nov. 2002), p. 103002. DOI: 10.1103/PhysRevD.66.103002.

[117] Nicholas Senno, Kohta Murase, and Peter Meszaros. "Choked Jets and Low-Luminosity Gamma-Ray Bursts as Hidden Neutrino Sources". In: *Phys. Rev.* D93.8 (2016), p. 083003. DOI: 10.1103/PhysRevD.93.083003.

[118] Shigeo S. Kimura, Kohta Murase, and Kenji Toma. "Neutrino and Cosmic-Ray Emission and Cumulative Background from Radiatively Inefficient Accretion Flows in Low-Luminosity Active Galactic Nuclei". In: *Astrophys. J.* 806 (2015), p. 159. DOI: 10.1088/0004-637X/806/2/159.

[119] Andrea Palladino and Walter Winter. "A multi-component model for observed astrophysical neutrinos". In: *Astron. Astrophys.* 615 (2018), A168. DOI: 10.1051/0004-6361/201832731.

[120] A. Karle and T. Makuluni. *Optical transmission and sensitivity data of IceCube Digital Optical Modules*. Tech. rep. 2007.

[121] R. Abbasi et al. "Calibration and characterization of the IceCube photomultiplier tube". In: *Nucl. Instrum. Methods Phys. Res. A* 618.1 (2010), pp. 139–152. DOI: 10.1016/j.nima.2010.03.102.

[122] Rongen, Martin. "Measuring the optical properties of IceCube drill holes". In: *EPJ Web of Conferences* 116 (2016), p. 06011. DOI: 10.1051/epjconf/201611606011.

[123] Anatoli Fedynitch, Julia Becker Tjus, and Paolo Desiati. "Influence of hadronic interaction models and the cosmic ray spectrum on the high energy atmospheric muon and neutrino flux". In: *Phys. Rev. D* 86.11, 114024 (Dec. 2012). DOI: 10.1103/PhysRevD.86.114024.

[124] W. Martin. *RISE Germany Internship: Applying Deep Learning Methods to the Search for Astrophysical Tau Neutrinos*. Tech. rep. TU Dortmund, 2017. DOI: 10.17877/DE290R-19186.

[125] M. G. Aartsen et al. "First Double Cascade Tau Neutrino Candidates in IceCube and a New Measurement of the Flavor Composition". In: *PoS(ICRC2019)* 1015 (2019). DOI: TBD.

[126] M. G. Aartsen et al. "All-flavor Multi-Channel Analysis of the Astrophysical Neutrino Spectrum with IceCube". In: *PoS(ICRC2017)* 976 (2018). DOI: 10.22323/1.301.0976.

[127] M. G. Aartsen et al. "IceCube-Gen2: A Vision for the Future of Neutrino Astronomy in Antarctica". In: (2014). arXiv: 1412.5106.

[128] P. Allison et al. "Performance of two Askaryan Radio Array stations and first results in the search for ultrahigh energy neutrinos". In: *Phys. Rev. D* 93 (8 Apr. 2016), p. 082003. DOI: 10.1103/PhysRevD.93.082003.

[129] A. Gazizov and M. Kowalski. "ANIS: High energy neutrino generator for neutrino telescopes". In: *Computer Physics Communications* 172.3 (Nov. 2005), pp. 203–213. DOI: 10.1016/j.cpc.2005.03.113.

[130] Adam M. Dziewonski and Don L. Anderson. "Preliminary reference Earth model". In: *Physics of the Earth and Planetary Interiors* 25.4 (1981), pp. 297–356. DOI: 10.1016/0031-9201(81)90046-7.

[131] G. Carminati et al. "MUPAGE: a fast atmospheric MUon GEnerator for neutrino telescopes based on PArametric formulas". In: (July 2009). arXiv: 0907.5563.

[132] Mario Dunsch et al. "Recent Improvements for the Lepton Propagator PROPOSAL". In: *Computer Physics Communications* 242 (2019), pp. 132–144. DOI: 10.1016/j.cpc.2019.03.021.

[133] Leo Breiman et al. *Classification and regression trees. Florida*. 1984.

[134] Leo Breiman. “Random Forests”. In: *Machine Learning* 45.1 (2001), pp. 5–32. DOI: 10.1023/A:1010933404324.

[135] Gilles Louppe et al. “Understanding variable importances in forests of randomized trees”. In: *Advances in neural information processing systems*. 2013, pp. 431–439.

List of Figures

List of Tables