Technische Universität München
Fakultät für Physik

Masterarbeit in Physik

Sensitivity Studies for Blazar Stacking Searches with the IceCube Neutrino Observatory

Matthias Johannes Huber

05. Dezember 2015

Erstgutachter (Themensteller): Prof. E. Resconi
Zweitgutachter: Prof.

Abstract

Since its construction the IceCube Neutrino observatory experienced remarkable success. Besides the detection of the highest energy neutrinos worldwide IceCube was the first experiment to observe an astrophysical high-energy neutrino flux. Although in the meantime the collaboration detected nearly 100 high energy neutrino events, the origin of these events is still not identified. Blazars, being a subclass of active galactic nuclei and consequently one of the most powerful objects in the universe are supposed to be one of the most likely sources of high energy neutrinos.
During the course of this thesis the statistical significance of a neutrino flux coming from groups of blazars from the 1WHSP[a] catalog was investigated with the help of a stacking maximum likelihood approach. In order to improve this analysis process, primarily the behavior of the currently used track reconstruction methods was studied and a new approach for the determination of the uncertainties of these reconstruction was evolved. Afterwards the integrated sensitivities and discovery potentials for different subsets of the 1WHSP catalog were calculated for seven years of IceCube data. Ultimately the differential sensitivity and the discovery potential was evaluated for a subset of the 103 most interesting blazars from the full catalog and then compared to a predicted flux from all blazars under the assumption of a lepto-hadronic emission model. Since the predicted flux exceeds the discovery potential above energies of $\sim 560\,\mathrm{TeV}$ the IceCube data might predict a discovery once the theoretical model describes the actual situation in blazars.
Since the full stacking test on unblinded IceCube data is going to follow shortly after the completion of this thesis we might find a signification indication for neutrino point sources soon. Moreover a discovery of neutrinos from blazars would confirm the lepto-hadronic emission model in blazar jets.

[a]1WHSP: First Wide-field Infrared Survey Explorer (WISE) High Synchrotron Peaked blazar catalog

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the past century the field of physic experienced an tremendous amount of progress. After the constitution of special relativity and the first quantum field theories at the beginning of the 20th century, the Standard Model (SM) of particle physics was developed over the years, yielding in a successful collocation of the *zoo* of elementary particles and their fundamental interactions. Moreover, great achievements were made in the field of cosmology leading to a progress in the understanding of the origin of our universe and high energy physics on large scales.

However, while continuously enhancing both, physics on subatomic as well as on astronomic scales simultaneously it emerged that treating these two fields completely separately might not be the correct approach. Since it was pointed out that most of the largest phenomenas in the universe can only be fully explained by having the knowledge of the undergoing physics on the smallest level, the inevitable connection between both topics, astro- and particle physics was considered. On this account it seemed very promising to examine the field of astroparticle physics and in particular neutrino physics which both can be seen as cutting point between particle physics and astrophysics. The history of astroparticle physics most probably started with the discovery of Cosmic Rays, whose existence could not be explained by particle physicists without making making use of astrophysical phenomena. Since that time this field of physics yielded a huge progress in the understanding of astrophysical phenomena.

Moreover in the course of time the neutrino, which was first discovered in 1956 by studying the inverse beta decay, emerged to be one of the most important objects in order to prove connections between astrophysical phenomena and the SM. In the field of astrophysics the detection of neutrinos was first used in order to describe the nuclear fusion processes inside the sun. Moreover since the neutrino is only weakly interacting with its surrounding matter it can be perfectly used as a messenger for the original direction of their generation. Therefore the existence of neutrino can also give insight in the underlying processes, leading to the tremendous electromagnetic emission from one of the most powerful objects in the universe, the active galactic nuclei. While the generation of the electromagnetic radiation caused by these objects can currently not be confirmed to be either the result of fundamental leptonic or hadronic processes in the vicinity of these objects, the detection of astrophysical neutrinos coming from the direction of active galactic nuclei would confirm a lepto-hadronic scenario.

The IceCube Neutrino Observatory, currently being the largest neutrino telescope, was built in order to get insight in this field of astrophysical neutrinos. During the first years of operation IceCube detected several astrophysical neutrinos up to the PeV energies. Nevertheless, up to today the origin of these high energy neutrinos is not yet understood.

The work in this thesis concentrates on the search for the origin of astrophysical neutrinos coming from blazar candidates, which build a subclass of active galactic nuclei. Finding a significant neutrino signal from blazars would not only yield to the discovery of an astrophysical neutrino point source but would also confirm the lepto-hadronic emission model of active galactic nuclei. The thesis starts with an introduction to high energy Cosmic Ray physics and in particular astrophysical neutrinos. Afterwards in chapter 3 the IceCube detector is introduced. Chapter 4 and 5 present the event reconstruction and the selection of the final event sample, followed by the description of the analyses method used in this thesis (Chapter 6). Chapter 7 concentrates on the description of the first WISE[1] High Synchrotron Peaked (1WHSP) blazar catalog which is used in order to provide blazars as point source candidates for neutrinos. Finally the sensitivity and the discovery potential of the analysis are presented in chapter 8.

[1]WISE: Wide-field Infrared Survey Explorer [66]

Chapter 2

The High Energy Universe

The ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe.

(Philip W. Anderson)

The Standard Model of particle physics is a theory concerning a classification of all known subatomic components of the universe and their possible interactions. It was developed throughout the 20th century, as a collaborative effort of scientists around the whole world. The Standard Model is one of the most tested theories world-wide and based on its remarkable success in explaining the results of a wide variety of physical observations it is sometimes even denoted as "the theory of *almost* everything" [48].

Nevertheless there are still phenomena within our universe which cannot or only partially be described by this fundamental approach. One of those poorly decoded objects can be associated with the high energy Cosmic Rays (CRs) which bombard the atmosphere of the earth from outer space. The study of CRs had a special role in many areas of physics and still provides large potential for further discoveries.

The first part of this chapter briefly reviews the the composition of CR particles, their potential sources and their interaction with matter. Afterwards, in section 2.2 the role of high energy cosmic neutrinos is specified.

2.1 Cosmic Rays

The so-called CRs, or probably even more appropriate, cosmic particles were discovered in 1912 by Victor F. Hess, measuring the ionization rates of the air up to an altitude of around 5300 meter, using a hot air balloon. In contrast to the opinion at that time that the ionization rate was caused by radioactive elements on the earth, V. F. Hess was the first to discover the extraterrestrial origin of the radiation [56].

Charged cosmic particles primarily consist of ionized nuclei, composed of about 90 % protons, 9 % alpha particles and other nuclei from heavier elements. Being accelerated to sufficient energies, these CRs can produce high energy photons an neutrinos near the site of acceleration, in collisions with ambient baryons or photons (see subsection 2.2.1). Once the primary particles hit the Earth's atmosphere, they generate so-called secondary CRs in atmospheric air showers (refer to subsection 2.1.4) [52].

The primary particles span an energy range over 12 orders of magnitude, stretching up to 10^{20} eV. Therefore CRs can be many orders of magnitude more energetic than particles generated by current collider experiments, such as the Large Hadron Collider (LHC). The measured energy spectrum of these CRs is displayed in figure 2.1, whereas several special features can be observed, which might provide clues to the origin of these CRs. Starting from a few GeV up to the *knee* ($E_{knee} \sim 10^3 - 10^4$ TeV) the spectrum follows a simple power law:

$$N(E)\mathrm{d}E \propto E^{-\gamma}\mathrm{d}E \stackrel{!}{=} E^{-2.7}\mathrm{d}E, \qquad \text{for} \quad E < E_{knee} \tag{2.1}$$

Above this energy the spectrum becomes somewhat steeper with a spectral index γ of approximately 3.0, before reaching the region above the *ankle* ($E_{ankle} \sim 4 \cdot 10^6$ TeV), where it hardens again. The transition of the spectrum around the ankle is believed to represent the changeover from Galactic sources generating the measured events to even more powerful extragalactic CR origins (see subsection 2.1.2). Finally the flux steepens again above an energy of $E_{GZK} \sim 4 \cdot 10^7$ TeV. This fall-off is expected to be associated with the *GZK cut-off*, named after its discoverer K. Greisen, G. Zatsepin and W. Kuzmin, who pointed out that the universe could become opaque at such energies due to collisions of the primary particles with the cosmic microwave background radiation [52].

2.1.1 Acceleration of Cosmic Rays

Although the existence of CRs is verified since more than 100 years, the question of their origination is still not clarified. Among different theories, how these particles can obtain their enormous energies, the most accepted one is the theory of CR acceleration established by Enrico Fermi during the 1950s.

First order of Fermi acceleration implies that the gain of momentum and energy of cosmic particles arises from repeated reflections at astrophysical shock fronts propagating towards a plasma carrying a magnetic field. In the simplified one dimensional picture that is illustrated in figure 2.2, one can consider a relativistic particle, originating from a moving magnetized plasma, traveling in the positive x-direction towards a shock front, moving with velocity $-u$ in the negative x-direction. The shocked gas on the opposite site flows aways from the front, which then leads to a total velocity of $u_{gas} = -u + v = -\frac{3}{4}u$ in the laboratory frame, where v is the velocity of the gas relative to the shock front. Thus if one supposes the relativistic particle to be back-scattered by the gas, it travels downstream with velocity u_{gas}. If the magnetic cloud deflects the particle as well, it can redo this cycle once again [52].

The fractional energy gain the relativistic particle receives, can be easily calculated in this simplified picture, using the application of Lorentz transformations. Assuming the particle enters the gas behind the shock front with an incoming angle of θ_{in}, its energy $E^{'}_{in}$ in the rest frame of the gas can be calculated as

$$E^{'}_{in} \approx \gamma E_{in}(1 - \beta\cos(\theta_{in})), \tag{2.2}$$

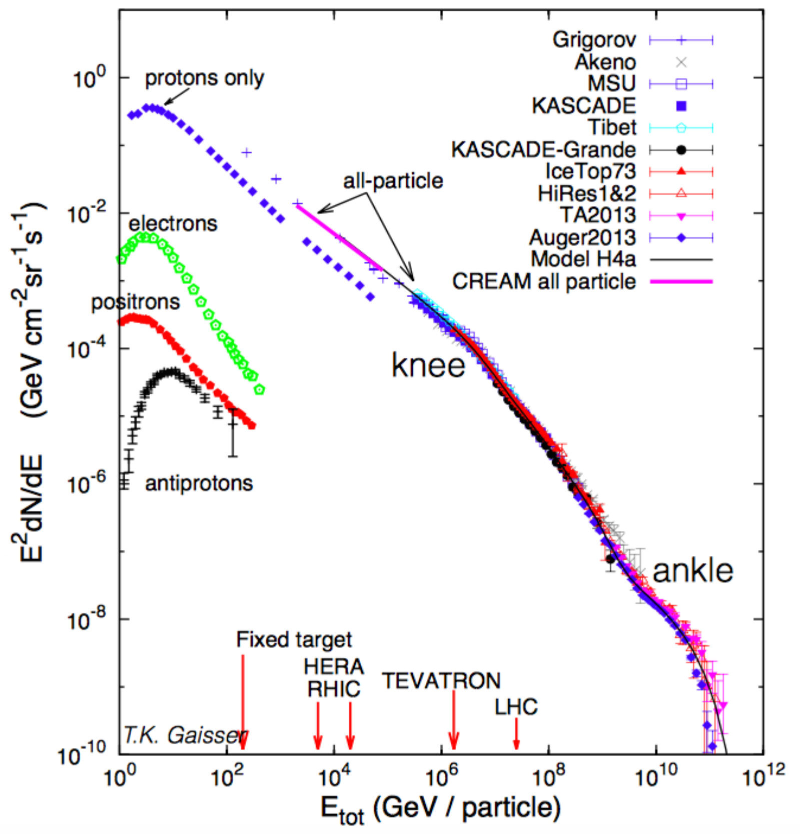


Figure 2.1: Energy spectrum of Cosmic Rays. At the x-axis, respectively energy axis, the maximum particle energies that can be generated by respective collider experiments is illustrated for comparison. The figure was taken from [36].

where c is the velocity of light, $\beta = u_{gas}/c$, γ corresponds to the Lorentz factor $1/\sqrt{1-\beta^2}$ and E_{in} is the energy of the incoming particle in the laboratory frame. After the particle is scattered backwards it passes the shock front again with the outgoing angle θ_{out}, whereas the outgoing energy $E_{out}^{'}$ in the rest frame of the gas is the same as the incoming energy $E_{in}^{'}$. Using the Lorentz back transformation into the laboratory frame one can compute the fractional energy gain of the particle:

$$\frac{\Delta E}{E_{in}} = \gamma^2 \left(1 - \beta \cos(\theta_{in}) + \beta \cos(\theta_{out}) - \beta^2 \cos(\theta_{in}) \cos(\theta_{out})\right) - 1 \sim \frac{u_{gas}}{c}. \tag{2.3}$$

Now that an energy gain in every cycle is stated, one can treat this acceleration as a statistical process. Under the assumption that the energy accretion in each cycle is $\Delta E = \alpha E$, the total energy of the particle after n encounters becomes

$$E_n = E_0(1+\alpha)^n, \tag{2.4}$$

where E_0 corresponds to the starting energy of the particle. At each stage of the acceleration there is of course a certain chance for the particle to escape further cycles. Assuming the

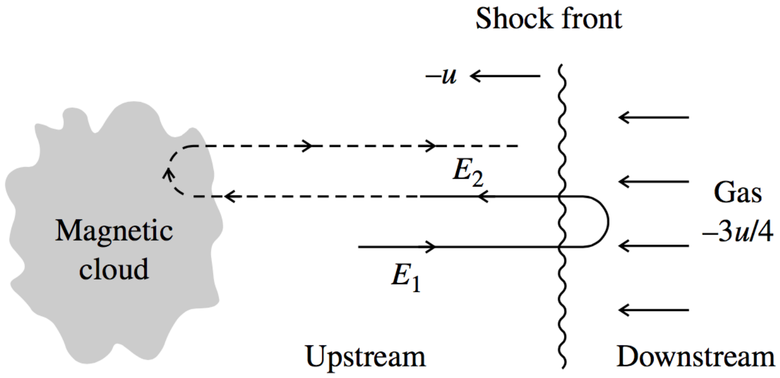


Figure 2.2: Illustration of the first order Fermi acceleration method. The figure was taken from [52].

probability to escape the acceleration loop is P_{esc},

$$N = N_0(1 - P_{esc})^n \tag{2.5}$$

particles, of the initially inserted N_0, are expected to be left in the acceleration circuit. Combining the two equations (2.4) and (2.5) one can derive the differential energy spectrum resulting from the first order Fermi acceleration method

$$\frac{dN(E)}{dE} = constant \cdot E^{-\gamma}, \qquad \gamma = \frac{\ln(1 - P_{esc})}{\ln(1 + \alpha)} \tag{2.6}$$

As visible in equation (2.6) this spectrum shows a power-law dependence with spectral index γ. Hence comparing this to the measured Cosmic Ray flux, shown at the beginning of section 2.1 and in figure 2.1, even this simplified model seems to fit very well for most energies. Nevertheless, especially at higher energies also other acceleration mechanism must be considered which leads to the fact that the acceleration of CRs is still an open, not fully understood field of research in astro-particle physics [52].

2.1.2 Potential Sources of Cosmic Rays

In the previous section 2.1.1 a promising method of energizing cosmic particles was presented. As a next step in our understanding, it is now necessary to figure out if there are objects in the universe that can fulfill all the requirements of the acceleration theories.

The energy range of CRs being measured up until now goes up to values of 10^{20} eV. There are certain limits physical objects have to satisfy in order to be able to generate such high-energy particles. Reminding oneself of the simplified model of shock acceleration from subsection 2.1.1, one can record that the charged moving particle is performing a circular motion due to the presence of the magnetic field $\mathbf{B}$, perpendicular to the moving direction of the particle and $\mathbf{B}$. The gyroradius r_g of this motion is

$$r_g = \frac{p}{|q|B_\perp} = \frac{E/c}{ZeB_\perp}, \tag{2.7}$$

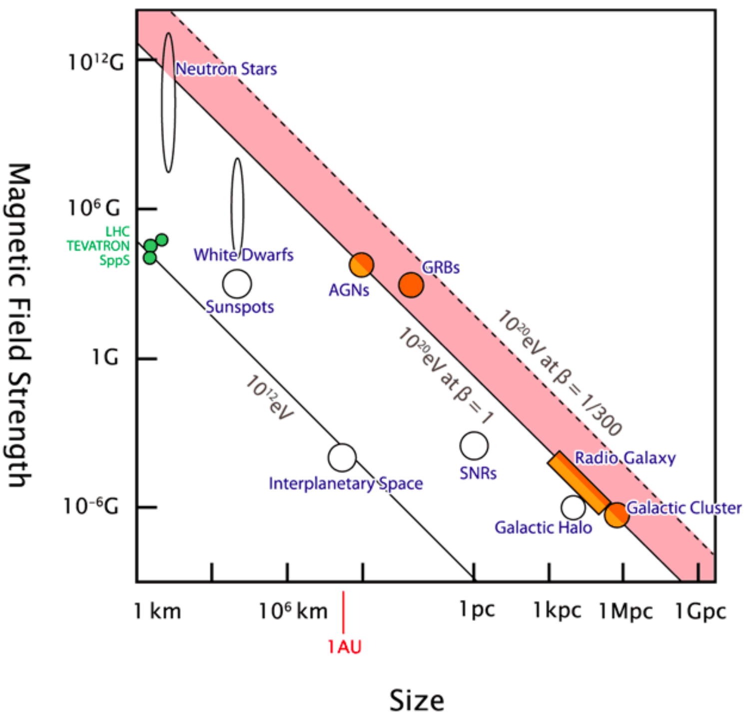


Figure 2.3: The Hillas diagram shows the magnetic field strength and the size of possible sites of acceleration. Objects above the black line in principle do have the theoretical potential to accelerate protons up to an energy of 10^{20} eV while for source below this line this scenario is not possible. The graphic was taken from [35].

where p corresponds to the momentum of the particle, q to its electric charge and $B_\perp$ to magnetic field strength perpendicular to the moving direction. In the second step the charge q is replaced by the factor Ze, where Z corresponds to the atomic number and e to the elementary charge. Moreover the momentum p is written in terms of the energy E. Thus, once the gyroradius r_g reaches the geometric size of the shock, the particle can no longer be confined in the vicinity of the moving front. This leads to a limitation of the maximum energy the particle can obtain, subject to the magnetic field and the geometric size of the accelerating object. Also taking into account the characteristic velocity βc of the scattering centers, the maximum energy is

$$E_{max} \sim 2 r_g \beta c Z e B_\perp . \tag{2.8}$$

In the Hillas diagram (figure 2.3), several classes of astrophysical objects that have the potential to generate very high energy cosmic particles are listed. The black line in this plot shows the minimum demand on the size and the magnetic field of the objects, in order to possess the capability to create 10^{20} eV protons. In the following some of the most important Galactic and extragalactic sources of high energy CRs are mentioned and shortly explained [35].

Galactic Sources

- **Supernova remnants (SNR):** A Supernova remnant is the resulting structure of a Supernova, a stellar explosion that briefly outshines an entire galaxy. Hence the remnant consists of the ejected material expanding into the space forming a shock wave. According to the Fermi acceleration mechanism from above, these shocks generate and accelerate cosmic particles which is why SNRs dominate the energy spectrum of CRs below the knee. Nevertheless from the observations of the physical properties, displayed in figure 2.3 it is recognizable that SNRs do not have the ability to generate CRs up to highest energies [42, 53].
- **Pulsar Wind Nebula (PWN):** A Pulsar is a rapidly rotating, highly magnetized neutron star. While radiating relativistic particles along its spin axis a pulsar wind is formed around the neutron star. Once this wind expands into the surrounding interstellar medium it creates a shock front [34].

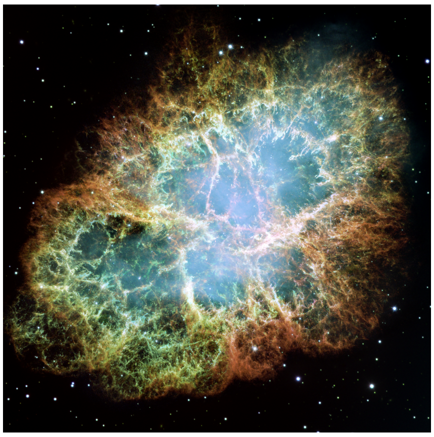
Extragalactic Sources

- **Active Galactic Nuclei (AGN):** AGNs consist of supermassive black hole that attracts the surrounding matter due to its gravitational potential. As a consequence an accretion disc arises around the centered black hole and two converse radio-emitting jets are formed perpendicular to the accretion disc. A schematic illustration of the structure of an AGN is shown in figure 2.4. Due to this structure they are one of the most luminous objects in the universe, being already detected in a wide range of the electromagnetic spectrum. Moreover in some particular theoretical scenarios they are promising candidates to be generators of high energy neutrinos. Looking at the Hillas diagram in figure 2.3 it is moreover visible that AGNs are one of the rare objects that are likely to fulfill the physical criteria in order to accelerate protons up to energies of 10^{20} eV. Additional information about AGNs are presented in more detail in section 7.1.
- **Gamma-Ray Bursts:** Gamma ray bursts are the most luminous objects known in the universe which appear as intense flashes of gamma rays, lasting approximately from milliseconds to a few minutes. The existence of a GRB is associated with extremely energetic explosion that might be caused by massive star collapse or compact objects collisions [28].

2.1.3 Interaction of Charged Particles with Matter

Charged high energy particles created in statistical acceleration processes may interact with matter in their vicinity, whereas the rate of the reactions in general depends on the density and the composition of the matter. Below the most important interaction types of CRs are discussed shortly [52]:

- **Energy loss by Ionization:** Charged particles lose some of their energy due to



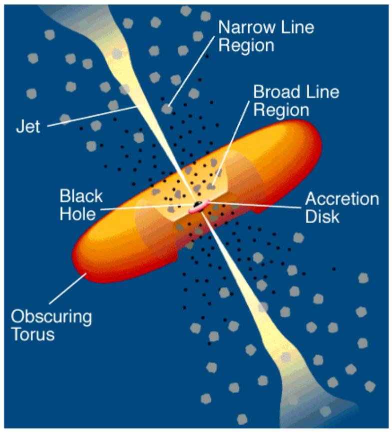


Figure 2.4: ***Left:*** Picture of the Crab Nebula SNR. ***Right:*** Schematic illustration of the structure of AGNs. This figure was taken from [64].

collisions with atomic electrons leading to the ionization of the atoms. This energy loss can be described by the Bethe Bloch equation [52].

- **Coulomb scattering:** Besides atomic electrons, charged CRs can also undergo scattering with the atomic nuclei of a medium. Due to the high mass of the nuclei, this process dominates the energy loss resulting from collisions with atomic electrons.
- **Energy loss by *Bremsstrahlung*:** Charged particles produce electromagnetic radiation once they are decelerated in the matter. This process is called *Bremsstrahlung*. It can for instance occur if the particle gets deflected by a magnetic field, caused by the medium. Since the radiative energy deficit depends on decreases with the square of the mass of the particle, this interaction type mainly only has to be considered for electrons.
- **Cherenkov radiation:** The Cherenkov effect will be explained in detail in section 3.1.2.1.

2.1.4 Atmospheric Air Showers

Once primary CRs enter the earth's atmosphere, they start to interact with the ambient matter and produce huge cascades, also called air showers of secondary particles. A schematic sketch of the dispersion of an atmospheric air shower, caused by a proton is displayed schematically in figure 2.5. In the collisions of the primary particles with molecules in the air mainly pions and kaons are generated with their decay products forming mainly three different subclasses of the shower cascade.

Due to their extremely short lifetime neutral pions π^0 decay nearly immediately into a pair of photons, which for their part then develop electromagnetic cascades consisting of electrons, positrons and photons. Since the absorption length of these cascades is very short, the electrons and photons from electromagnetic cascades are easily absorbed in the atmosphere.

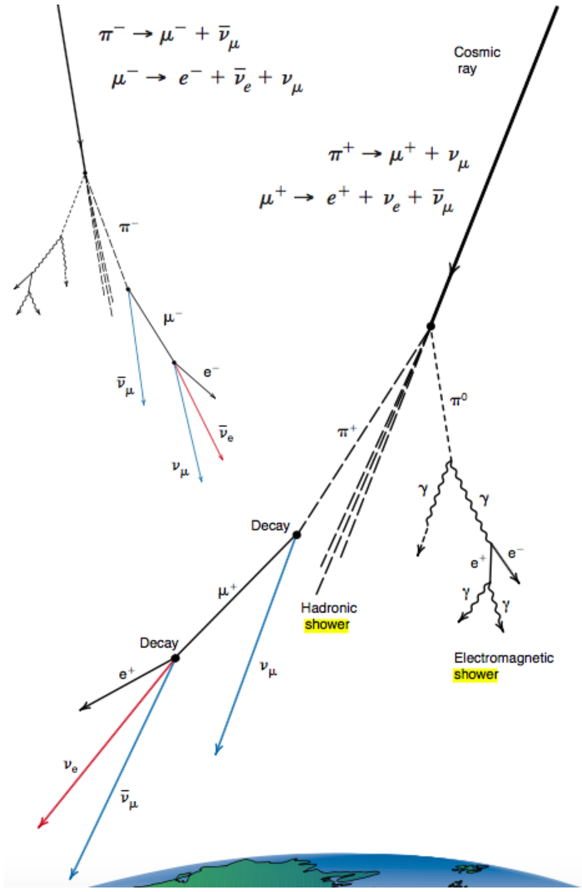


Figure 2.5: Sketch indicating the progress of a particle cascade in the atmosphere, induced by a primary CR proton. The figure was taken from [44].

Hence they constitute to the soft part of cosmic radiation [52]. Moreover hadronic cascades are produced in interactions of hadrons through the exchange of gluons.

The muonic component of air showers, consisting of muons and neutrinos depicts the third main part of the cascades. These muons are mainly produced in decays of charged pions $\pi^\pm$ and kaons $K^\pm$

$$\pi^+(K^+) \longrightarrow \mu^+ + \nu_\mu \tag{2.9}$$
$$\pi^-(K^-) \longrightarrow \mu^- + \bar{\nu}_\mu, \tag{2.10}$$

whereas the branching ratios of these decays is approximately 99.9 % for pions and 67 % for kaons [51]. The resulting muons can furthermore decay into electrons according to

$$\mu^\pm \longrightarrow e^\pm + \bar{\nu}_\mu(\nu_\mu) + \nu_e(\bar{\nu}_e) \tag{2.11}$$

Hence, in order to estimate the particular atmospheric flux of each particular neutrino flavour i, it seems reasonable to regard this magnitude as a convolution of the spectrum of the primary CRs ϕ_{CR} in the atmosphere and the yield of each neutrino flavour per primary particle $Y_{CR\rightarrow\nu_i}$:

$$\phi_{\nu_i} = \phi_{CR} \otimes Y_{CR\rightarrow\nu_i}. \tag{2.12}$$

Since the production of muons and neutrinos in the muonic component of air showers relies on the hadronic decay of pions, kaons and in case of electron neutrinos ν_e also on the decay of muons it is crucial that these particles can decay before they become part of different interactions. Hence in order to evaluate the neutrino yield $Y_{CR\rightarrow\nu_i}$ it is necessary to regard the decay length λ_{dec} of a relativistically moving particle which can be estimated as

$$\lambda_{dec} = \gamma c\tau = c\tau\frac{E}{mc^2} := h_0\frac{E}{\epsilon_{char}}, \tag{2.13}$$

where τ corresponds to the mean lifetime of the particle at rest and γ to the Lorentz factor which can be expressed as the ratio of the total energy E of the particle and its energy at rest mc^2. Based on the observed behavior from equation (2.13) it is visible that the mean decay length surmounts the atmospheric scale length h_0 once the energy E of the particle exceeds its characteristic energy ϵ_{char} [30].

Consequently for low energies nearly all pions and muons decay before they can interact with the surrounding matter yielding an observable ratio of muon and electron neutrinos at the surface of the Earth (refer to equations (2.11) and (2.10)) of

$$\frac{\nu_\mu + \bar{\nu}_\mu}{\nu_e + \bar{\nu}_e} \sim 2. \tag{2.14}$$

With increasing energy ($\sim$ GeV) especially muon interactions start to dominate over the decay process ($\epsilon_\mu \approx 1\,\mathrm{GeV}$), resulting in a suppression of the observed electron neutrino ν_e flux compared to one of muon neutrinos ν_μ. Moreover at even higher energies the decay of pions and kaons ($\epsilon_\pi = 115\,\mathrm{GeV}$, $\epsilon_K = 850\,\mathrm{GeV}$) is suppressed and they start to lose a fraction of their energy before decaying, yielding in a muon and neutrino spectrum following a power-law with a spectral index $\gamma \approx 3.7$ which is steeper than the spectral progress of the primary CRs ($\gamma \approx 2.7$, refer to section 2.1). Since the decay length $\lambda_{dec,\pi}$ of pions surmounts the atmospheric scale length h_0 at energies above about 100 GeV, the muon and muon neutrino generation at these energies is dominated by the decay of kaons [30, 51].

At the highest energies ($\sim$ 100 TeV) also the decay of kaons is highly suppressed and further processes, including the decays of heavy mesons ($D^\pm$, D^0, D_s, Λ_C) become important. Since these particles do have a maximum mean lifetime of approximately 10^{-12} s they decay instantaneously which is why the leptons and neutrinos directly produced in these decays are called *prompt* leptons and neutrinos. Other than the spectrum of neutrinos from pions and kaons these prompt particles, being created at an early stage of the air shower follow approximately the same $E^{-2.7}$ spectrum as the primary CRs [30, 51].

Moreover since tau neutrinos ν_τ are not created in the decay of pions and kaons they can only appear as a product of prompt decays. Nevertheless it must be noted that there has been no direct measurement of a *prompt* neutrino flux so far. The relative contributions of intermediate particles to the muon and the muon neutrino flux as functions of the particular energy are shown in figure 2.6 [30].

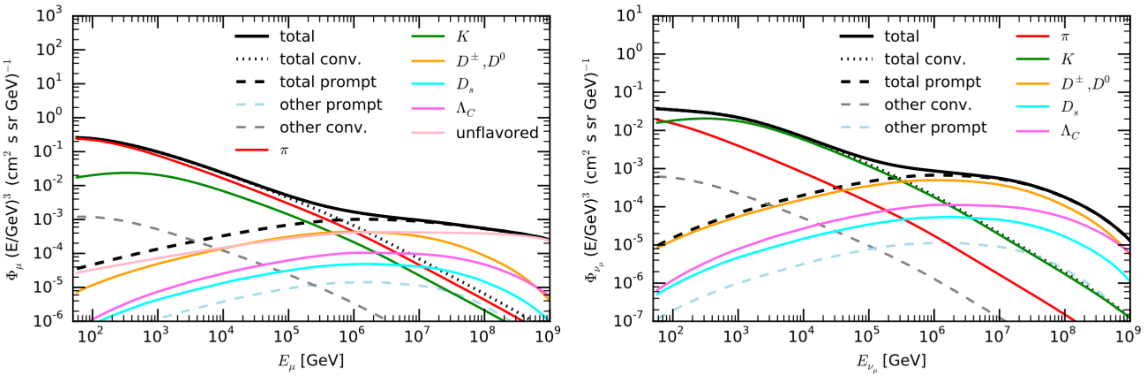


Figure 2.6: Contributions of intermediate particles to the atmospheric muon and muon neutrino flux as functions of the particular energy. ***Left:*** Muon ($\mu^+ + \mu^-$) flux. ***Right:*** Muon neutrino ($\nu_\mu + \bar{\nu}_\mu$) flux. Both figures were taken from [27].

2.2 High Energy Astrophysical Neutrinos

Although the current state of the science about CRs contains a plausible theoretical description of the measured electromagnetic energy spectrum, consisting of cosmic particles acquiring their energy in repeated reflections from astrophysical shock waves (see subsection 2.1.1), the exact origin of these particles still remains dubious. As mentioned before in this chapter primary CRs consist of protons and other heavier ionized nuclei, which are considerably deflected in the interstellar magnetic field on their way from their source to the earth. Therefore most of these primary cosmic particles do not contain any directional information about their site of acceleration [52].

Thus, to maintain unambiguous evidence for the origin of CRs one has to access information from different messengers. Luckily, protons, containing a sufficient amount of energy can create photons and neutrinos in collisions with ambient matter close to the site of acceleration. Compared to the charged primary CRs, photons and and neutrinos are not distracted from magnetic fields, which is why they both can give perfect clue to the scene of the generation of cosmic particles. Nevertheless photons can still be absorbed by dust or softened by electron-positron pair production, whereas the cross section for neutrinos to interact inside the interstellar medium is relatively small. The whole situation of different particles traveling from their source towards the earth is displayed schematically in figure 2.7. On these ground, in the following subsection we mainly concentrate on the generation and interactions of high energy neutrinos.

2.2.1 Generation of High Energy Neutrinos

As already mentioned above, high energy neutrinos can be created near the site of acceleration, as a secondary product in the collision of primary CR protons with ambient matter. The dominant reactions and their secondary products are generated in interactions with other

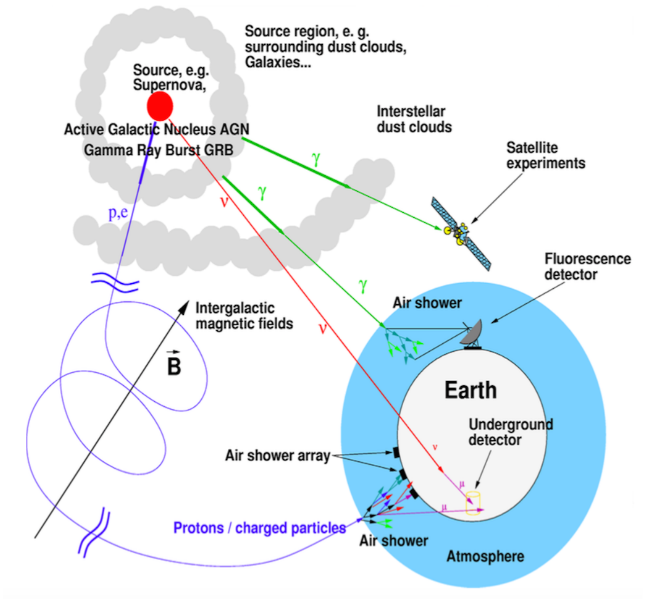


Figure 2.7: Schematic view of different astrophysical particles traveling towards the earth. Since the charged particles are strongly deflected by the Galactic magnetic field, only photons and neutrinos are in principle effectively suitable messengers to give insight to the origin of these particles. This figure was taken from [28].

protons

$$pp \longrightarrow \begin{cases} pn\pi^+ & \longrightarrow pn\mu^+\nu_\mu \longrightarrow pne^+\nu_e\bar{\nu}_\mu\nu_\mu \\ pp\pi^0 & \longrightarrow pp\gamma\gamma \end{cases} \tag{2.15}$$

or respectively with ambient photons

$$p\gamma \longrightarrow \Delta^+ \longrightarrow \begin{cases} n\pi^+ & \longrightarrow n\mu^+\nu_\mu \longrightarrow ne^+\nu_e\bar{\nu}_\mu\nu_\mu \\ p\pi^0 & \longrightarrow p\gamma\gamma \end{cases} . \tag{2.16}$$

The same processes occur if one regards incident neutrons instead of protons, leading to the production of negatively charged pions π^-. Looking at both equations above, these decay modes can be reached by replacing all particles on the right side with their antiparticles. At higher energies also kaons can contribute to this decay spectrum [17].

Under the assumption that the produced amount of negative and positive pions is the same, the ratio of the different neutrino flavors close to the source is

$$(\nu_e : \nu_\mu : \nu_\tau) = (\bar{\nu}_e : \bar{\nu}_\mu : \bar{\nu}_\tau) = (1 : 2 : 0) \tag{2.17}$$

Although no τ-neutrinos are produced at the source, one expects to observe their signal at the earth. Since neutrinos contain a mass different from zero, they oscillate between the different

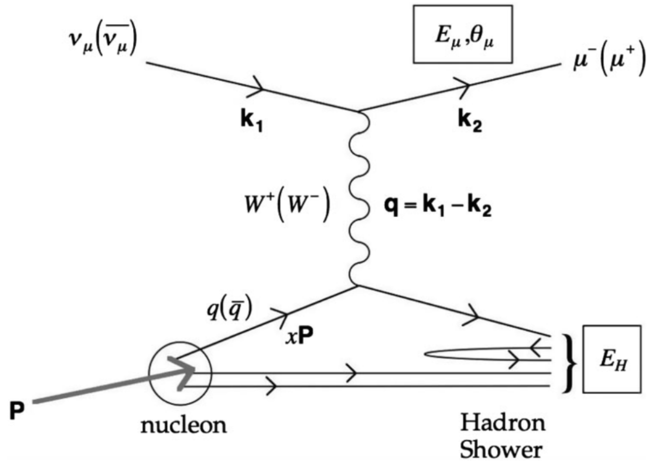


Figure 2.8: Illustration of the Feynman diagram for a CC deep inelastic interaction of a neutrino with a nucleon. The diagram was taken from [29].

flavor-eigenstates (e, μ, τ) on their way to the earth. Consequently, according to theory and previous measurements, the flavor ratio of astrophysical neutrinos that is expected to be discovered at the earth is [17]:

$$(\nu_e : \nu_\mu : \nu_\tau) = (\bar{\nu}_e : \bar{\nu}_\mu : \bar{\nu}_\tau) = (1 : 1 : 1) \tag{2.18}$$

Since tau neutrinos can only be produced in *prompt* decays of heavy mesons with charm contribution, and hence their production in the atmosphere is nearly completely suppressed (refer to subsection 2.1.4), the appearance of ν_τ could be interpreted to be almost surely of astrophysical origin.

2.2.2 Interactions of Neutrinos with Matter

Neutrinos are neutral leptons, which communicate with their environment exclusively through weak interaction processes. Thus, the cross section for interactions of neutrinos is much smaller than for photons or charged leptons, which also underly the electromagnetic force. This behavior provides on the one hand a perfect chance to use neutrinos as a messenger for searches for astrophysical sources of CRs, but on the other hand makes it also very difficult to detect such a signal.

Up to a certain range of the neutrino energy, there are mainly only two different interaction modes that have to be considered. Neutrinos can either undergo neutral current (NC) processes, where they exchange a virtual Z^0-Boson with a nucleon, depositing some fraction of its energy and initializing a hadronic cascade, or charged current (CC) processes, where the neutrino decays into the charged lepton l of the same generation, via the exchange of a charged $W^\pm$ boson with a nucleon (inverse beta decay) also creating an additional hadronic cascade.

$$\text{Neutral current:} \qquad \nu_l + N \longrightarrow \nu_l + N^* \tag{2.19}$$

$$\text{Charged current:} \qquad \nu_l + N \longrightarrow l + N^* \tag{2.20}$$

In both equations, N corresponds to the atomic nucleus, l to the lepton flavor (e, μ, τ) and N^* to the hadronic cascade. Moreover depending on the energy of the neutrino different interaction subtypes dominate these two processes. At low energies the nucleus N can be treated as a point-like object in the interaction with neutrinos, yielding a quasi-elastic scattering process. Moving up in energy to approximately $\sim 0.1\,\text{GeV}$-$20\,\text{GeV}$, resonant baryonic excitations of the nucleus, meaning the production of pions and kaons in the collision with the neutrino, have to be considered. At the highest energies ($E_\nu > 20\,\text{GeV}$) the nucleus must be treated on the parton-level since the neutrinos can resolve the structure of the individual quarks. These processes are considered as deep inelastic scattering [29].

The Feynman diagram for CC deep inelastic scattering (DIS) of a high energy muon neutrino ν_μ with a nucleon is visualized in figure 2.8.

The resulting charged particles in each of the CC reactions emit light along their track, excited by the Cherenkov effect (see subsection 3.1.2.1). These light cones can then be detected by large area neutrino telescopes, using photomultiplier tubes (PMTs) (see subsection 3.1.1).

Chapter 3

The IceCube Neutrino Observatory

Neutrinos are fermionic, electrically neutral elementary particles in the Standard Model (SM) of particle physics. Due to their capacity to interact exclusively through the exchange of the Z^0 and $W^\pm$ gauge bosons of the weak interaction, their cross sections are very small compared to other SM particles. In order to detect high energy neutrinos at adequate statistics, it is therefore of particular importance to know their behavior in surrounding matter.

As explained in section 2.2.2, high energy neutrinos up to certain energies mainly interact with matter via NC or CC collisions. The cross sections for both processes increase with the energy of the neutrino and the mass of the target particles. Thus large volume neutrino detectors including target material with very high masses are preferable in order to increase the probability to detect sufficient neutrino interactions. The appearance of neutrino events is mainly represented by Cherenkov photons emitted along the path of secondary charged leptons, created in the neutrino collisions (see subsection 2.2.2).

Although the small cross sections for neutrino interactions diminish the ability to record high statistic neutrino samples, this behavior can also be used to encapsulate these collisions against background reactions showing the same signature, by choosing an intelligent design and location for the detector. Due to the small cross sections neutrinos posses a very long mean free path in matter compared to other SM particles like photons or charged leptons. Thus, building neutrino detectors underground or underwater is an effective method to shield the data from a large part of background.

The IceCube Neutrino Observatory, located inside the Antarctic ice at the South Pole combines all the requirements listed above. In the following section the experimental setup of the detector (section 3.1) will be shortly explained. Afterwards, part 3.2 summarizes the data acquisition method, followed by a short discussion about the physics discovery potential of the IceCube experiment (section 3.3). Finally, in subsection 3.4 some typical IceCube conventions, which will be of constant practice in the following chapters, are explained.

3.1 Experimental Setup and Detection Technique

The IceCube detector is the world largest neutrino telescope, built inside the ice at a depth of about 1500 m to 2500 m. During the installation period between the years 2005 and 2011, 86 strings, each equipped with 60 light detecting digital optical modules (DOM, see subsection 3.1.1) were inserted into the Antarctic ice. Except from the 8 most inner, the strings are

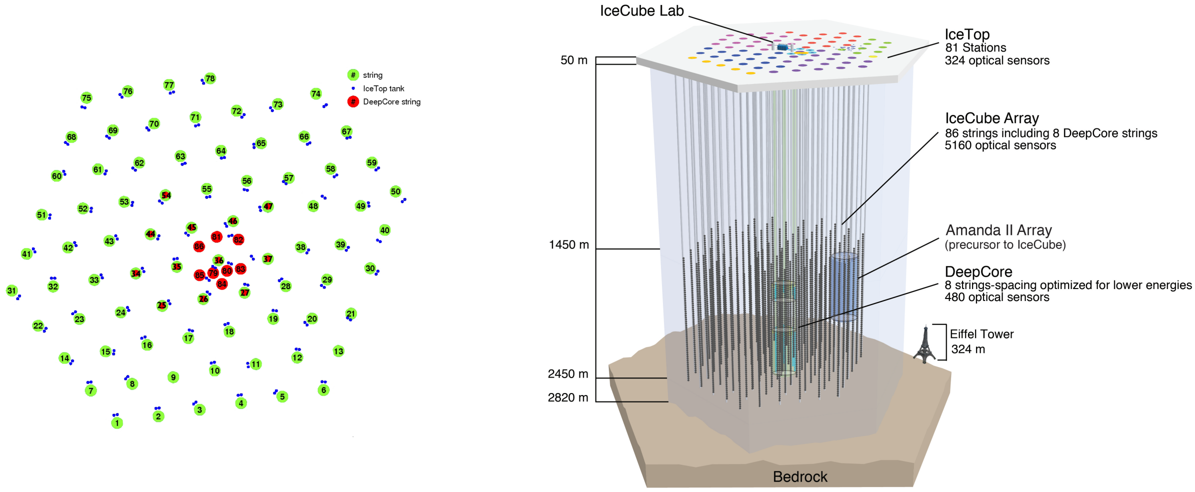


Figure 3.1: Illustration of the IceCube detector. ***Left:*** Top view of the detector. The red dots in the center illustrate the position of the denser spaced DeepCore strings. ***Right:*** Three dimensional model of the detector. This figure also displays the Amanda II Array, which was the precursor experiment to IceCube. Moreover the Eiffel Tower is illustrated to present the proportions of the whole setting. Both figures were taken from [59].

distributed in a hexagonal structure with a horizontal spacing of 125 m and a vertical distance of 17 m between each DOM, yielding a total detector volume of more than 1 km^3 inside the ice. The eight strings at the center of the detector (DeepCore [8]) are built in a denser configuration, having an average horizontal separation of only about 70 m and a DOM to DOM distance of around 7 m.

Above this hexagonal structure, further 324 DOMs are distributed in 81 detection stations on top of the ice (IceTop [7]). A schematic overview of this detector setup is shown in figure 3.1.

The different parts of the detector are designed in order to cover a wide area of neutrino research. The denser DeepCore strings, in addition to the surrounding detector, that can be used as a veto region for atmospheric muons, is configured to have a high sensitivity to low energy neutrino events ($E_\nu \gtrsim 10\,\mathrm{GeV}$), and therefore play a special role for searches in the field of neutrino oscillations. For events with higher energies, the small spacing is not that important anymore, but in order to increase statistics and to improve the accuracy of directional reconstructions of neutrino events, a large detector volume is preferable. Consequently the whole in-ice part of IceCube is used in the searches for the origin of astrophysical neutrinos [60].

Finally IceTop can be used as a cosmic ray detector or as a veto for atmospheric muons and neutrinos for the in-ice part.

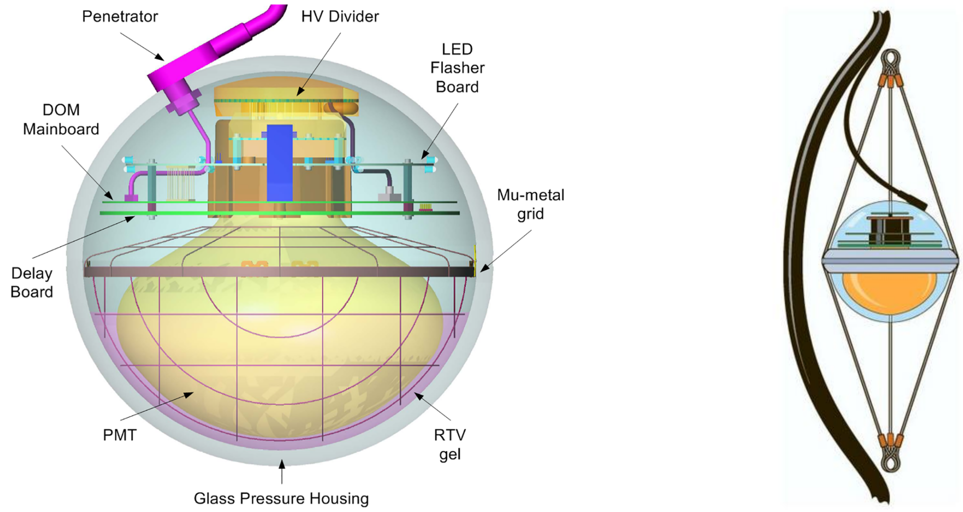


Figure 3.2: Representation of a Digital Optical Module, used in IceCube. ***Left:*** Display of the composition of an individual module. The picture is taken from [59]. ***Right:*** Attachment of a DOMs to the its main cable,whereas one of these cables is located in each drilled string hole. Both figures were taken from [61].

3.1.1 Digital Optical Module (DOM)

The centerpiece of the IceCube detector is represented by the light detecting digital optical modules inside the Antarctic ice. As visible in figure 3.2, each module mainly consists of a high efficiency photomultiplier tube (PMT), which is used to convert the detected light to an electrical signal (see section 3.2). This PMT as well as the hardware components are enclosed by a high pressure consistent glass sphere, in order to be shielded against the extreme conditions in the depths of the Antarctic ice. In addition, inside the glass sphere a DOM mainboard containing the analog and digital signal electronics is installed. To improve the performance of the photomultiplier a mu-metal grid, utilized to insulate against the terrestrial magnetic field, is placed on top of the PMT. Moreover a LED flasher-board containing 12 LEDs is used in the calibration process of the DOMs. All DOMs of a particular string are connected to their main cable (see figure 3.2), which is directing the recorded signal towards a data collecting point for all of the strings on top of the ice. To keep survey of the individual DOMs in the evaluation of the data, each DOM is equipped with a code, including a nickname, its position and further useful information [6].

3.1.2 Detection Technique

As mentioned in section 2.2.2, neutrinos interact with nucleons in neutral or charged current processes, creating hadronic cascades and in case of a CC interaction also an additional charged lepton, carrying the same flavor as the neutrino. According to the Cherenkov effect,

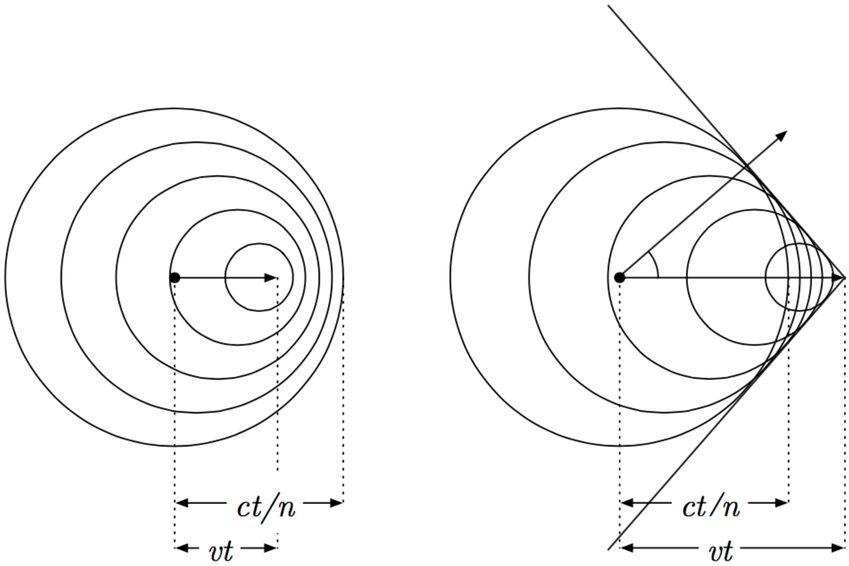


Figure 3.3: Illustration of the Cherenkov effect. This figure was taken from [37].

which will be explained in the following subsection, these relativistic charged leptons create photons along their path inside the detector. This light signature as well as the photons created by the hadronic cascades can be detected by the photomultiplier tubes (PMT), positioned in the DOMs inside the ice.

3.1.2.1 Cherenkov Effect

If a relativistic particle is traveling through a dielectric medium (refractive index $n > 1$) at a velocity v faster than the speed of light $c_{med} = c_{vac}/n$ in this medium, it emits Cherenkov radiation. This emission occurs due to the asymmetric polarization of the medium in the front and in the rear side of the particle, giving rise to a varying electric dipole momentum. Hence if a particle is fast enough, the polarization disturbance caused in consequence of the particle movement can not relax elastically to the mechanical equilibrium but the energy contained in this disturbance is emitted as light. This coherent light wave front is generated along the track of relativistic particle at the angle θ_C, also called Cherenkov angle, which is defined according to

$$\cos(\theta_C) = \frac{1}{\beta n}, \tag{3.1}$$

where β is the ratio of the velocity of the particle v and the speed of light in vacuum c_{vac}. An illustration of the Cherenkov effect is visible in figure 3.3 [21, 25].

Assuming that a particle is moving nearly at the speed of light ($\beta \sim 1$) inside the Antarctic ice, having a refractive index $n \approx 1.309$, one can approximately appoint the Cherenkov angle θ_C of a muon traveling through the IceCube detector to $\theta_C \approx 41°$. Moreover the light yield of Cherenkov radiation relative to the wavelength λ of the light and the distance dx along the track can be calculated using the Frank-Tamm-Equation

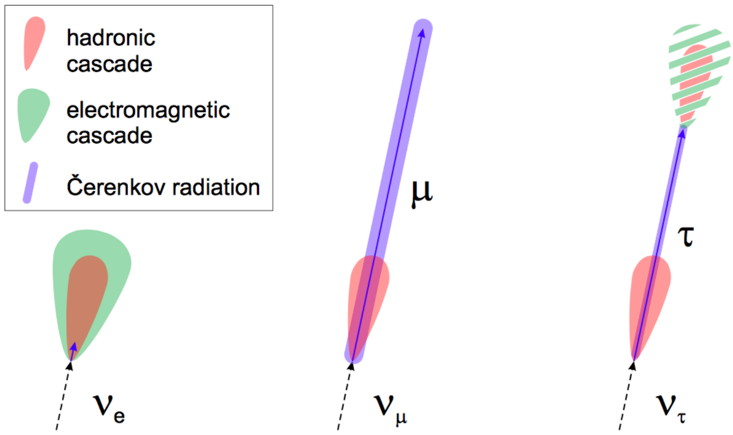


Figure 3.4: Illustration of the event signatures the different neutrino flavors produce in CC current interaction in the ice. Electron neutrinos ν_e appear as spherical cascades (subsection 3.1.3.1), whereas for muon neutrinos a long bright track is visible. The τ-neutrinos show up as a double-bang event with a luminous track in between. One should also always keep in mind that each hadronic cascade also has a small electromagnetic component. The figure was taken from [65].

$$\frac{d^2N}{dxd\lambda} = \frac{2\pi\alpha}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2(\lambda)}\right), \tag{3.2}$$

where N corresponds to the number of Cherenkov photons and $\alpha = 2\pi e^2/hc_{vac}$ to the fine structure constant. Since this function is proportional to $1/\lambda^2$, it is dominated by shorter wavelengths, which is why Cherenkov light appears mostly blue to our eyes [52, 25].

3.1.3 Event Topologies

Since neutrinos are only weak-interacting particles, one can not detect them directly but only via the observation of secondary particles resulting from their interactions with matter. Nevertheless in case of CC interactions of neutrinos in IceCube it is even possible to distinguish between the different flavors. Reminding oneself of section 2.2.2, it is easily notable that for both, neutral and charged current interactions the signal signature starts with the emission of light close to the interaction vertex, due to the generated hadronic cascades. In case of NC processes there are no more visible signatures inside the detector, given that the remaining neutrino can leave the detector without any further interactions.

However in case of CC interactions, one can distinguish between three different scenarios, according to the flavor of the primary neutrino. The three event categories are schematically shown in figure 3.4.

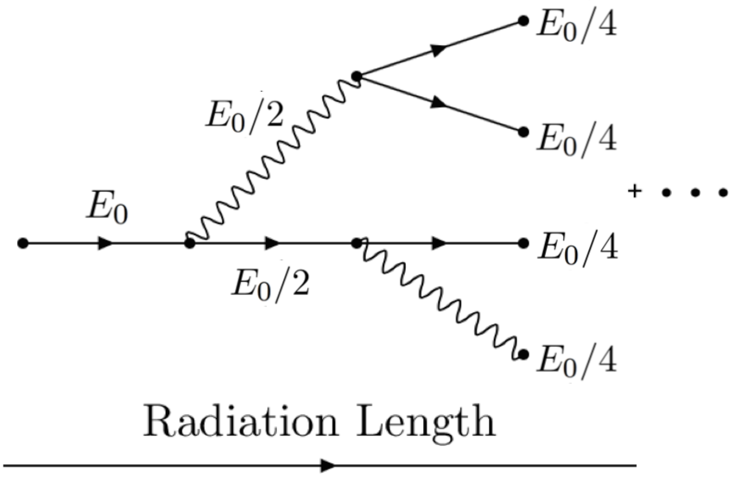


Figure 3.5: Simplified model of the energetics in an electromagnetic cascade. The representation of photons, electrons and positrons is carried out according to the common Feynman rules. Hence the straight lines represent electrons and positrons while the oscillating lines illustrate photons.

3.1.3.1 Electron Neutrinos ν_e

Next to the hadronic cascade, an electron neutrino ν_e interacting with a nucleus of the Antarctic ice in a CC process also generates an electron at collision vertex. Above a critical energy E_c (in ice $E_{c,ice} \approx 90\,\mathrm{MeV}$), radiative energy losses of the electron e dominate over ionization processes. Since electrons posses a relatively small mass, they mostly lose their energy due to the emission of *Bremsstrahlung* (see section 2.1.3),

$$\left\langle \frac{dE}{dx} \right\rangle_{rad} = -\frac{E}{X_0} \propto -\frac{E \cdot Z}{A \cdot m^2}, \tag{3.3}$$

where X_0 is the radiation length, Z the atomic and A the mass number of the nucleus [52]. In this case, m corresponds to the electron mass. Provided the energy of a radiated photon is bigger than two times the electron mass m, it can generate an electron-positron pair. At TeV energies the mean track length of a photon in a medium before it decays in an e^+e^- pair can be roughly approximated by the radiation length X_0. Consequently the primary electron evolves in an electromagnetic cascade, mainly consisting of electrons, positrons and photons. Once the electrons and positrons, participating in these cascades, pass the critical energy E_c, ionization energy losses become dominant and the evolution of the cascade breaks off.

Assuming a simplified model, as visible in figure 3.5, one can approximately calculate the maximum length of an electromagnetic cascade inside the IceCube detector. Assuming that, after traveling one radiation length, the primary electron with energy E_0 emits a photon carrying half of the primary energy. After another radiation length X_0 this photon generates an e^+e^- pair, where each particle is carrying the energy amount $E_0/4$. Moreover the remaining primary electron emits another photon with energy $E_0/4$. Hence after a depth of n radiation lengths, the energy of the particles inside the cascade is

$$E(n) = \frac{E_0}{2^n} \tag{3.4}$$

Consequently the maximum depth of the cascade evaluates to

$$n_{max} \cdot X_0 = \ln(\frac{E_0}{E_c})/\ln(2) \cdot X_0. \tag{3.5}$$

For a 10 TeV electron this corresponds approximately to a maximum length of $\sim 6.6\,\mathrm{m}$ in inside the antarctic ice ($X_{0,H_2O} \approx 36\,\mathrm{g\,cm^{-2}}$, $\rho_{ice} \approx 0.917\,\mathrm{g\,cm^{-3}}$) [52, 51].

Since this length is relatively small compared to the structure of the detector, the typical signature of an electron neutrino ν_e inside the detector is a nearly spherical shaped cascade composed by photons coming from the hadronic and the electromagnetic cascade (see figure 3.4). As a consequence of this, the direction of incoming neutrinos ν_e can only be identified with an accuracy of approximately 10°. In return their energy can be reconstructed quite good, since usually the neutrinos deposit most of their energy inside the detector [25].

3.1.3.2 Muon Neutrinos ν_μ

For muon neutrinos the situation inside the detector is different. Since the mass of a muon is relatively high compared to the mass of electrons ($m_\mu/m_e \approx 200$ [51]), the energy losses caused by stochastic effects like *Bremsstrahlung* are negligible up to a critical energy $E_{c,\mu}$ ($E_{c,\mu,ice} \approx 500\,\mathrm{GeV}$, see equation (3.3)). In this region the energy deficit is dominated by ionization. This effect is almost independent of the muon energy, which is the reason why muons lying in this energy sector and having a relatively long mean lifetime ($\tau_\mu = 2.197 \cdot 10^{-6}\,\mathrm{s}$) can travel long distances inside the ice [51].

With an energy greater than $E_{c,\mu}$ the stochastic effects become all-dominant and the average energy loss can be described by

$$\frac{dE_\mu}{dx} = A + B \cdot E_\mu, \tag{3.6}$$

with $A = 2.4 \cdot 10^{-3}\,\mathrm{GeV\,g^{-1}\,cm^2}$ and $B = 3.2 \cdot 10^{-5}\,\mathrm{g^{-1}\,cm^2}$. Moreover dx describes the distance along the muon track [5]. As a consequence of this equation, one can calculate the maximum length L_μ, a muon with starting energy E_μ can travel in ice until it loses all its energy

$$L_\mu(E_\mu) = \frac{\rho_{ice}}{B} \int_0^{E_\mu} \frac{1}{\frac{A}{B} + E} dE = \frac{\rho_{ice}}{B} \ln(1 + \frac{B}{A} E_\mu) \approx 4.55\,\mathrm{m\,GeV^{-1}} \cdot E_\mu. \tag{3.7}$$

Taking a muon with energy $E_\mu = 10\,\mathrm{TeV}$, this results in a track length $L_\mu(10\,\mathrm{TeV}) \approx 45\,\mathrm{km}$. Hence also in this very high energy regime the muon can move long distances through the ice. Additionally in both energy regimes the muon uniformly emits Cherenkov photons along the path of movement (see section 3.1.2.1). This light output only plays a minor role in consideration of energy losses, but it is the most important part of the identification inside the detector. These photons can be detected by the optical modules of IceCube and consequently the muon appears as a luminous track inside the detector. Summarized the signature of a CC muon neutrino ν_μ inside the ice consists of the light yield from the hadronic cascade at the interaction vertex in combination with a bright shining track, resulting from the secondary muon (see figure 3.4) [25, 5].

Since these muon show a long luminous lever arm inside the detector, usually the resolution of the muon direction in IceCube is very good. Nevertheless the determination of the neutrino direction is limited by the energy dependent angular difference of the neutrino and the muon, given by

$$\langle \sphericalangle(\nu_\mu, \mu) \rangle = \frac{1.5^\circ}{\sqrt{E/\mathrm{TeV}}}. \tag{3.8}$$

The direction of a high energy ($> 1\,\mathrm{TeV}$) muon neutrino can be pointed with an accuracy of less than 1°. Different than the directional reconstruction of muon neutrinos, the energy determination proves less accurate than the one of the cascade-like electron neutrinos. This is caused by the fact that muon neutrinos usually do not deposit all their energy inside the detector volume, since the moving path of the muon can exceed the size of the detector [25, 5].

3.1.3.3 Tau Neutrinos ν_τ

Due to their high mass ($m_\tau \approx 1.78\,\mathrm{GeV}$), tau leptons are nearly not affected by energy losses from stochastic effects such as *Bremsstrahlung*, pair production and photo-nuclear processes. Hence, their energy mainly decreases as an effect of ionization processes. Other than muons, the tau leptons have a very short mean lifetime ($\tau_\tau = 2.906 \cdot 10^{-13}\,\mathrm{s}$), leading to a rapid decay into mostly hadrons ($\sim 65\,\%$) [51]. Consequently the signature of tau neutrinos, as displayed in figure 3.4, is composed of a hadronic cascade at the interaction vertex, followed by a luminous track, created from the emitted Cherenkov light along the path of the tau and completed by a second hadronic and electromagnetic cascade, once the tau lepton finally decays. The appearance of a ν_τ is also called double-bang event.

3.2 Triggering and Data Acquisition

So far, in this chapter, we have specified the setup of the detector and the tools it provides. Moreover we also clarified the behavior and the type of bequeathed signal of neutrinos in the vicinity of ice. Nevertheless it is still not clear how to filter out the important information from the signals recorded by the detector.

First of all, one has to clarify what kind of signal is detected. Once one or more photons reach one of the PMTs, the module internally produces a digital output, also called *hit*, which basically consists of a timestamp and the digitized waveform information collected over a fixed

Trigger	**Event Rate** [Hz]
SMT8	2113
String	2240
Volume	3727
SLOP	13.3

Table 3.1: Common event triggers used in IceCube, in addition to their rates from Run 120029 [58].

time interval (presently $\sim 6.4\,\mu s$, [9]). Once the pulse of a *hit* exceeds a threshold of 0.25pe, where *pe* corresponds to a single photoelectron, the hit continues processing as a DOM *launch*.

Nevertheless for the search of neutrinos, single DOM launches are not very meaningful, since they are mostly generated from noise and independent of that do neither give any insight in the incoming direction nor in any other physical properties of the particle. Consequently at this point it seems reasonable to regard the whole detector at once, as a composition of 5160 DOMs. As already mentioned in subsection 3.1.1, each DOM is connected to its corresponding main cable, which is used to merge all information of the individual DOMs. Using this connection, one can deploy criteria that select the given information for physical events, while being resistant to noise. Due to the fact that the light emitted from particles going through the detector, clusters in specific regions in a certain time frame, the principle of local coincidences is introduced in the following. According to the definition, a local coincidence (LC) exists, if within a certain time window at least two neighboring or respectively next-to-neighboring DOMs of a string have a *launch* [9]. Consequently the entity of all DOM launches can be divided in two scenarios:

- **Hard Local Coincidence (HLC):** A HLC exists if a LC of at least two DOMs is observed within $2\,\mu s$. The DOM launches corresponding to the HLC are marked as HLC hits.
- **Soft Local Coincidence (SLC):** Every DOM launch, not corresponding to a HLC is regarded as a SLC hit [23].

The SLC hits are very likely to be from noise, whereas for HLC hits the probability to be a product of noise is highly diminished. Until now only hit criteria separately concerning the individual strings have been applied. Since one is interested in signatures of physical particles, further trigger thresholds, combining the whole detector structure are necessary. The most common trigger in the data acquisition (DAQ) of IceCube is the SMT8 (Simple Multiplicity 8) trigger, which requires at least 8 HLC hits within a time window of $5\,\mu s$. Some other common triggers, in company with the rate of events surviving the trigger cuts, are listed in table 3.1.

Once a trigger condition is fulfilled, the detector is read out for $10\,\mu s$. For SLC hits only the timestamp and the charge information are saved while for HLC hits the whole information is stored. Finally the summarization of all these information composes an entire event [23]. Following this procedure, IceCube produces almost 1 TB/day of raw DAQ output. Hence, before this data output is sent north, it runs through a filtering system, which already preselects the data online at the pole, subject to the different analysis purposes. During the filtering some fast reconstruction algorithms are already applied, which are then used to remove parts of the background events. The starting offline data sample used in the point source analysis in this thesis are MuonL3 data, which contain only events that already passed the muon filter (see chapter 5) [58].

3.3 Physics Potential

In 2004 the construction of the IceCube Neutrino Observatory at the South Pole started with the main ambition to find evidence for very high energy neutrinos reaching the Earth from outside the Solar System. Since the completion of the whole detector in December 2010, IceCube not only was the first experiment to observe an astrophysical high-energy neutrino flux at significance level of 5σ, but also had some further remarkable discoveries.

In 2012 the collaboration discovered the first two PeV neutrinos, expected to have a cosmological origin at a confidence level grater than 90 %. However IceCube is not only interested in the detection of astrophysical neutrinos, but also in the direction of their origin. Since neutrinos are not affected by magnetic fields, they are perfect messengers in order to find the position of their generation (see section 2.2). Hence several analyses such as all-sky point source surveys, as well as stacking searches using different source catalogs are performed on the recorded IceCube data. Up to now no evidence for the origin astrophysical neutrinos was found.

Moreover, next to the astrophysical neutrinos, IceCube also operates in other physical fields of interest. Since the detector consists not only of the in-ice part, but also IceTop (see section 3.1) it can also work in the field of Cosmic Rays (see subsection 2.1.4). The most inner part of the IceCube structure, DeepCore, is used to perform precision measurements on the values of the mixing angles, which are responsible for neutrino flavor oscillations [60].

Although the current version of the IceCube detector can be seen as a remarkable success, its ability to be an efficient instrument in neutrino astronomy is limited by the measured amount of cosmic neutrinos. During the previous years approximately 100 astrophysical neutrino events were recorded. Observations on the diffuse neutrino flux seem to suggest a much larger level of hadronic activity in the non-thermal universe than previously thought and suggest a rich discovery potential for a larger neutrino observatory. Hence an extension of the current detector structure, called *IceCube-Gen2* is planned for the future. Figure 3.6 shows the current idea of this extension [4].

3.4 IceCube Conventions

As mentioned before, the IceCube Neutrino Telescope is located at the celestial South Pole, which represents a very special geometrical position. Due to the rotation of the Earth the detector twines around its vertical axis once every day. Hence in order to describe the

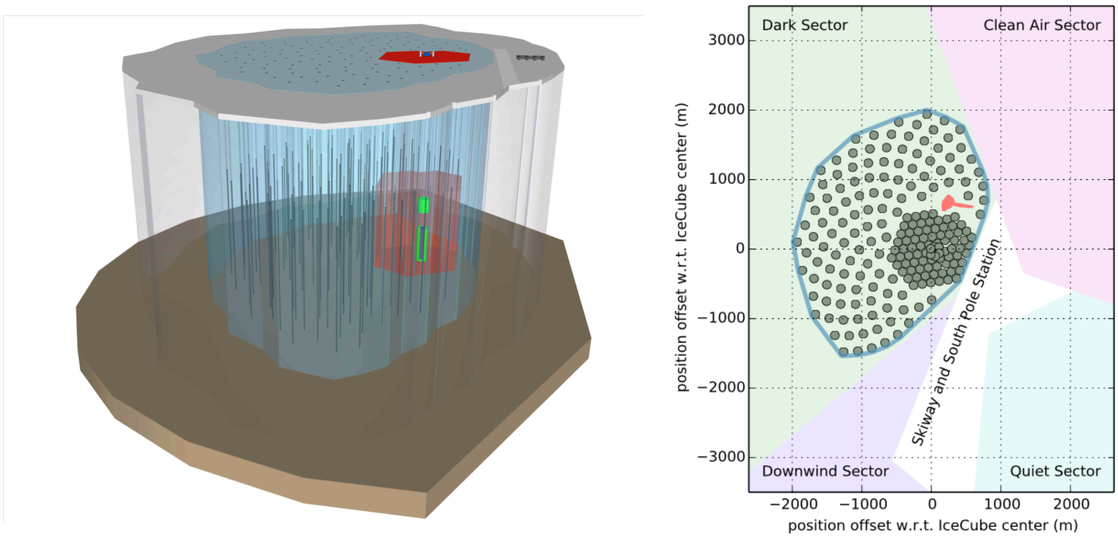


Figure 3.6: Overview of the present vision for an extension of the current IceCube detector. ***Left:*** Illustration of the planned area (blue) with additional strings around the current detector (red). The picture was taken from [60]. ***Right:*** Top view of the planned extension. This figure was taken from [4].

incoming direction of particles in IceCube it seems reasonable to use a spherical coordinate system which in principle specifies the direction of an incoming particle by two angles.

Since inside the IceCube collaboration several different fields of activities are treated there are two diverse variants of describing the incoming directions of particles. Since both coordinate systems are used throughout this whole thesis, it seems reasonable to shortly explain these conventions here.

In order to determine the track of a particle passing through the IceCube detector a *local* coordinate frame, being centered inside the detector at a depth of 1948 m below the surface is used. As visible in figure 3.7 the zenith angle θ illustrates the angular difference between the direction of the particle and the vertical detector z-axis, where $\theta = 0°$ corresponds to the South and $\theta = 180°$ to the North Pole. Moreover the azimuth angle ϕ describes the angle between the x-axis and the orthogonal projection of the particle track on the horizontal x-y-plane, having values in a range of $0°$ to $360°$. Hence this system can be interpreted as a geographical coordinate frame centered inside the detector at the South Pole, with the azimuth ϕ as longitude and the zenith θ as co-latitude angle. Since these coordinates perfectly describe the position of a particle inside the detector this convention is mainly used in the event reconstruction and selection (see chapters 4 and 5). Moreover the definition of the terms *northern* and *southern sky* that are commonly utilized inside the IceCube Collaboration are based on a partition of the zenith angle θ (refer to section 5.1).

Nevertheless since amongst others the IceCube experiment searches for the origin of astrophysical neutrino sources it seems furthermore reasonable to express the direction of the measured particles in *global*, commonly used coordinates in order to guarantee the comparability to other experiments. Consequently in addition to the local coordinates, the direction

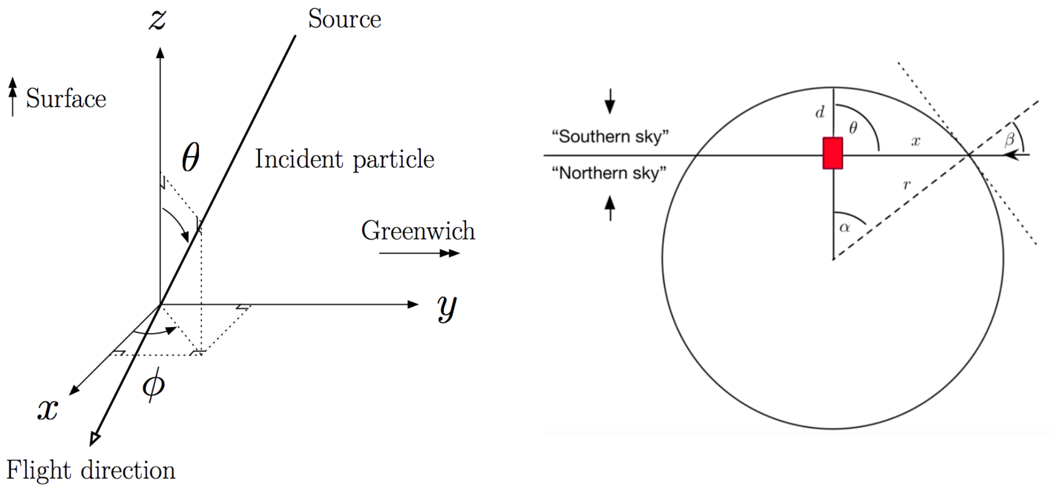


Figure 3.7: Illustration of the IceCube coordinate sytstems. ***Left:*** The *local* IceCube coordinate frame centered inside the detector. ***Right:*** Simplified two dimensional sketch of the Earth showing the relationship between the local and the global coordinate angles θ and δ. The red rectangle illustrates the IceCube detector. By knowing the depth d of the detector and the radius r of the Earth the local zenith angle θ can be converted to the global declination angle δ in this simplified model. Both figures were taken from [38].

of the incoming events in IceCube is also expressed in equatorial coordinates with origin at the geo-center, where the declination angle δ is comparable to the latitude and the right ascension α to the longitude in geographical coordinates. Hence $\delta = 90°$ illustrates the position of the North Pole, while $\delta = -90°$ corresponds to the South. The right ascension α is given in a range between $0°$ and $360°$ (respectively 0 h to 24 hour). Since this global coordinate convention is perfectly suitable to study the correspondence of IceCube data and certain points in the sky, it is applied in the chapters 6 to 9 of this thesis.

Moreover at this point it should also be noted that all neutrino fluxes that are calculated in this thesis are given in $E^2 d\phi/dE$ in order to guarantee comparability with other IceCube results.

Chapter 4

Event Reconstruction

In order to obtain evidence for the origin of astrophysical neutrinos it is important to identify the most meaningful characteristics of each event that enters the detector. Thereby the direction and the energy can be used to distinguish between astrophysical neutrino events and atmospheric background. A recorded event in IceCube consists of at least a few photons reaching different DOMs (see section 3.2). Extracting the time and charge information from these photons, a pattern of hits can be created for every event [28]. A visual representation of an exemplary track-like muon event in the IceCube detector is displayed in figure 4.1.

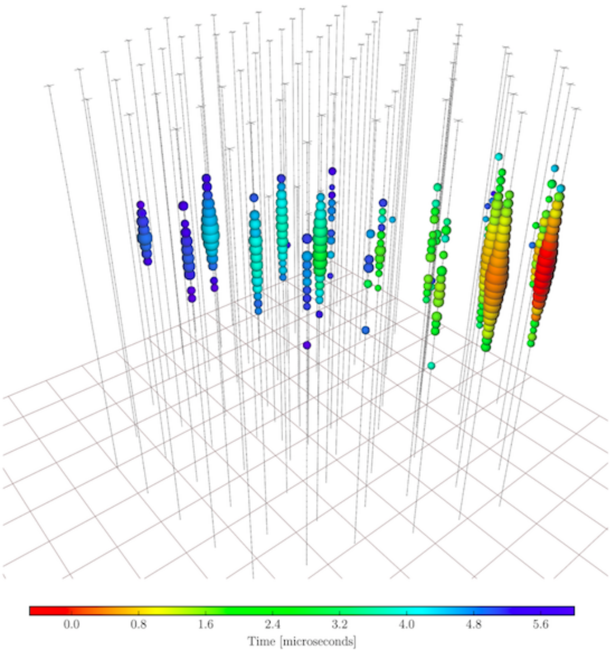


Figure 4.1: Image of a muon event inside the IceCube detector. The colors of the dots indicate the arrival time of the photons. In addition the information about the stored charge in each DOM is illustrated by the size of the dots at each DOM position. This figure was taken from [59].

Since for the point source search one is only interested in track-like appearances of events (see chapter 5), one mainly concentrates on reconstruction algorithms specialized on this case.

For both, directional and energy reconstructions, most algorithms in IceCube are applied consecutively in a chain, starting from simple *first-guess* methods followed by more accurate but also time consuming likelihood approximations, based on the results of the preceding algorithms.

The following chapter starts with the description of the track reconstruction methods. Hence physical results are in principle useless if one does not have any clue about their accuracy, it is also essential to determine the uncertainty of a reconstruction. Consequently section 4.1 is completed by a discussion about the uncertainty estimation of the track reconstruction algorithms. Finally, once the directional information can be evaluated for each event, some energy reconstruction methods will be explained in section 4.2.

4.1 Track Reconstruction

As already mentioned above, the directional reconstruction of events in IceCube takes place in different steps. In each step the result of the former algorithm is used as a seed, in order to receive even more accurate information. All of these methods use the information of the obtained photoelectrons (see section 3.2) to fit a linear track with direction (θ,ϕ) and vertex position $\mathbf{x} = (x, y, z)$ to each event.

4.1.1 LineFit

The LineFit is the first track reconstruction, which is applied to every IceCube event. It relies on the principle of least squares. First one assumes to observe N_{hit} photoelectron hits, whereas neither their positions $\mathbf{x}_i$ nor their arrival times t_i are known. Furthermore one defines the vertex of the track to be at $\mathbf{x}_0$, while the event moves with velocity $\mathbf{v}$. In the very simplified picture, in which the position of all hit DOMs is projected on the line of the muon track, one can approximate these DOM locations in terms of the variables $\mathbf{x}_0$, $\mathbf{v}$ and t_i:

$$\mathbf{x}_i \approx \mathbf{x}_0 + \mathbf{v} \cdot t_i \tag{4.1}$$

Using the difference between this estimate and the real value $\mathbf{x}_i$ one can estimate the most probable muon track. According to the least square method, one obtains the best fit by minimizing the sum of the squares of these differences with respect to the track variables $\mathbf{x}_0$ and $\mathbf{v}$ [2]:

$$\min_{\mathbf{x}_0,\mathbf{v}} \chi^2 := \min_{\mathbf{x}_0,\mathbf{v}} \sum_{i=1}^{N_{hit}} \left[\mathbf{x}_i - (\mathbf{x}_0 + \mathbf{v} \cdot t_i)\right]^2 \tag{4.2}$$

The main advantage of this method is that one does not need a computationally extensive minimizer, but can solve this optimization analytically using

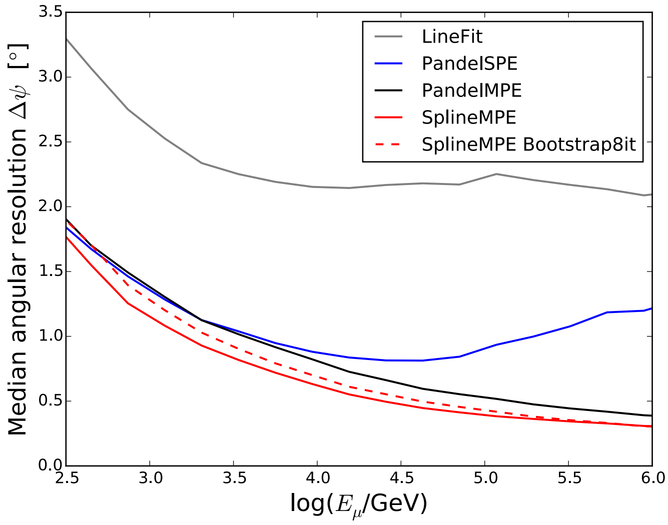


Figure 4.2: Median angular resolution of different track reconstructions as a function of the true event energy. The errors are calculated, using the true directional information of the Monte Carlo events.

$$\frac{d\chi^2}{d\mathbf{x}_0} \stackrel{!}{=} 0 \qquad \Rightarrow \qquad \mathbf{x}_0 = \langle \mathbf{x}_i \rangle - \mathbf{v}\langle t_i \rangle \tag{4.3}$$

$$\frac{d\chi^2}{d\mathbf{v}} \stackrel{!}{=} 0 \qquad \Rightarrow \qquad \mathbf{v} = \frac{\langle \mathbf{x}_i \cdot t_i \rangle - \langle \mathbf{x}_i \rangle \cdot \langle t_i \rangle}{\langle t_i^2 \rangle - \langle t_i \rangle^2}, \tag{4.4}$$

where $\langle \cdot \rangle$ corresponds to the arithmetic mean, defined as $\langle \mathbf{x}_i \rangle := \frac{1}{N_{hit}} \cdot \sum_{i=1}^{N_{hit}} \mathbf{x}_i$. Moreover this algorithm does not rely on a seed value, hence it represents a very good starting point [2]. Nevertheless the LineFit also implicates a lot of disadvantages. The algorithm does neither take care of any optical ice properties nor the Cherenkov cone, which defines the direction of the photons compared to the track. Consequently the angular error of the LineFit reconstruction (shown in figure 4.2), which can be evaluated from Monte Carlo data, is much worse than the resolution of the more advanced fits, which will be explained in the following subsections.

4.1.2 Maximum Likelihood Reconstructions

In order to increase the accuracy of the track determination, the direction of the light emission along the muon path, the fluctuating light absorption and scattering properties, resulting from different features of the Antarctic ice, have to be considered. Using a maximum likelihood approach these characteristics can be incorporated in the reconstruction. Assume one posses a set of measured data points $\mathbf{x}$ and the direction of the light emitting muon path, represented by a set of unknown parameters $\mathbf{a}$. As already mentioned at the beginning of section 4.1, the parameter space of a muon track in IceCube can be defined by its direction in addition to a vertex point, resulting in five different variables $\mathbf{a} = \{\theta, \phi, x, y, z\}$. These unknown values can be estimated by maximizing the likelihood of $\mathbf{x}$ given the parameters of the muon track. In case of statistically independent data points this likelihood function $\mathcal{L}$ can be written as

$$\mathcal{L}(\mathbf{x}|\mathbf{a}) = \prod_{x_i \in \mathbf{x}} p(x_i|\mathbf{a}). \tag{4.5}$$

The function $p(x_i|\mathbf{a})$ corresponds to the probability to measure an IceCube event x_i under the assumption of having a muon, traveling according to the parameters $\mathbf{a}$ inside the detector.

To describe the probability density $p(x_i|\mathbf{a})$ we assume the muon to travel in a straight line while continuously emitting photons along the Cherenkov cone. Since the hit time information of the single photoelectrons provide the most relevant knowledge, one derives a likelihood description based on the hit timestamps. Instead of just using the total arrival time of the photons, one concentrates on the time delay of the photons, caused by scattering inside the detector. The variable used in this context is called time residual (t_{res}) and is defined as the difference between measured time of the photons at the DOMs (t_{hit}) and the time the photons would have reached the DOM in the absence of scattering (t_{geo}):

$$t_{res} = t_{hit} - t_{geo} \tag{4.6}$$

The expected arrival time with influence of scattering effects can be evaluated from a geometrical consideration of the process. Looking at figure 4.3 it can be written in terms of measurable parameters in the following way

$$t_{geo} = t_0 + \frac{\hat{\mathbf{p}} \cdot (\mathbf{r}_i - \mathbf{r}_0) + d\tan(\theta_C)}{c_{vac}}, \tag{4.7}$$

where d corresponds to the minimal distance of the muon track to the hit DOM, θ_C describes the Cherenkov angle and $\mathbf{r}_i$ the position of the DOM. Reminding oneself of the likelihood description from equation (4.5), one notices that all photon hits, including multiple hits at one DOM, contribute to the likelihood. However due to the limited time resolution of the DOMs only the arrival time of the first pulse within a certain time window can be recorded [14].

Hence in practice simplifications of equation (4.5) are used. The Single-Photo-Electron (SPE) algorithm only considers the first photon at each DOM, which reduces the likelihood function to

$$\mathcal{L}(\mathbf{x}|\mathbf{a}) = \prod_{t_{res,i} \in \{i \in \mathbf{x} | \text{i first hit}\}} p_{\text{SPE}}(t_{res,i}|\mathbf{a}). \tag{4.8}$$

Nevertheless this simplification does not completely describe the real circumstances. Usually the first photon in one DOM is expected to be less scattered than the following ones at the

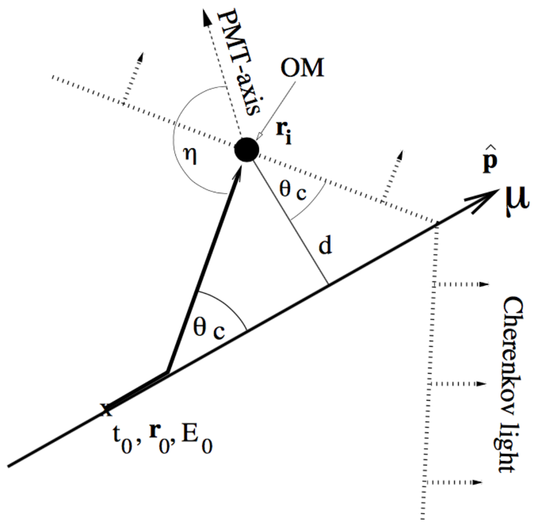


Figure 4.3: Schematic view of a muon going through the detector. The muon track is parametrized by the vertex position $\mathbf{r}_0$ and a direction $\hat{\mathbf{p}}$. Cherenkov photons are continously emitted along the track following the Cherenkov angle θ_C. One exemplary DOM is positioned at $\mathbf{r}_i$. This figure was taken from [14].

same module. Thus the first arriving photon at a DOM which detects multiple photons, is statistically expected to arrive earlier than a photon at a DOM, seeing only this photon. To take that aspect into account, the PDF is extended towards a cumulative function, considering also the number of photons N, arriving at each DOM

$$p_{SPE}(t_{res}) \quad \Longrightarrow \quad p_{MPE}(t_{res}) = N \cdot p_{SPE}(t_{res}) \cdot \left\{ \int_{t_{res}}^{\infty} p_{SPE}(t)dt \right\}^{N-1} . \tag{4.9}$$

This method is called Multi-Photo-Electron (MPE) reconstruction. For both methods the best fit values can be determined by maximizing the likelihood with respect to the muon track variables. For computational reasons one rather minimizes the negative logarithm of the likelihood, by making use of the *Simplex* minimization algorithm.

Of course, the accuracy of both methods strongly depends on the quality of the PDF of the photon arrival time. In IceCube there are in principle two different ways to calculate this PDF. One the one hand one can make use of an analytic function, called Pandel function, which is using the scattering and the absorption length as parameters, while on the other hand the PDF can be calculated from Monte Carlo simulations, using different spline tables to describe the properties of the Antarctic ice inside the detector. Looking at figure 4.2, it is easily visible that the reconstructions, considering all arriving photons, are more accurate than the ones, just looking at the first one. Moreover the algorithms using Spline Tables from simulations behave better than the ones using the analytic Pandel PDF [14].

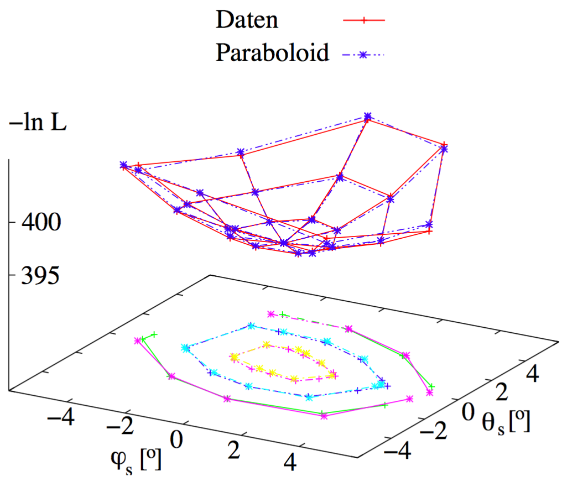


Figure 4.4: Exemplary representation of a two dimensional paraboloid fitted to the likelihood values of multiple base points in the φ-θ-plane. This picture was taken from [47]

4.1.3 Uncertainty Calculation of Track Reconstructions

In the point source search (see chapter 6) it is not only important to know the direction of an event, but also the knowledge about the accuracy of the directional reconstruction is a very crucial point since this uncertainty serves as the standard deviation of the gaussian PDF in the calculation of the signal likelihood of the events (compare to section 6.1.3). Again different methods exist in IceCube to evaluate the uncertainty of the track reconstruction algorithms. Both methods that are used in the point source analysis are explained subsequently.

4.1.3.1 Paraboloid Fit

The Paraboloid fit was used in former point source analysis to calculate the directional uncertainty for all selected events. In the likelihood reconstructions the algorithms approach the minimal value in the likelihood space step by step. The paraboloid fit uses this behaviour by scanning serveral likelihood points in the zenith and azimuth direction around the absolute minimum evaluated by the reconstruction. Using these points, a two dimensional paraboloid function is fitted to the likelihood distribution (see figure 4.4). The uncertainty of the reconstruction can then be calculated according to

$$\sigma_{paraboloid} = \sqrt{\frac{e_1^2 + e_2^2}{2}}, \tag{4.10}$$

using the paraboloid constants e_1 and e_2 from the fit [47].

4.1.3.2 Bootstrapping

The Bootstrapping algorithm uses a different approach. It uses resamples of single event information to make statistical statements about the original muon event. For the track reconstruction in IceCube the idea is the following. First regard a muon event as a set of pulses, created by photon hits at the DOMs. In order to obtain a new event, assembled from event information of the original one, one equips every pulse i with a weight w_i according to its relative charge

$$w_i = \frac{q_i}{q_{tot}}, \tag{4.11}$$

where q_i corresponds to the charge of the single pulse and p_{tot} to the charge of all pulses. In a next step multiple pulses are chosen randomly from a multinomial distribution, using the weights w_i and the charge of each of these selected pulses is set to $q_{i,new} = 1\text{pe}$. The selection goes on until the total charge of the resampled set of pulses is nearly the same as the charge of the original event, $q_{tot} \approx q_{tot,new}$. Finally this process is repeated n times and the track reconstruction is applied to each resampled set of pulses, resulting in n different directional information. If one assumes, that the results $\{\psi_i\}_{1\leq i\leq n}$ of the different resamples are gaussian distributed around their mean μ, the reconstruction uncertainty can be evaluated according to

$$\sigma_{bootstrap} = \sqrt{\frac{2}{\pi}} \cdot \frac{1}{n} \sum_{i=1}^{n} \sphericalangle(\psi_i, \mu). \tag{4.12}$$

4.1.3.3 Performance of the Track Uncertainty Calculators

The estimated uncertainties of an ideal method should be distributed close to the deviation of the reconstruction to the true direction of the neutrino event, $\Delta\psi$. Hence it seems reasonable to test the different error estimators by making use of simulated Monte Carlo events, including the true directions of the moving muon in addition to information about the starting neutrino events and the Pull correction variable, defined as

$$\text{Pull} = \frac{\Delta\psi}{\sigma}, \tag{4.13}$$

where σ corresponds to the uncertainty estimated by the fit. The closer this value is to 1, the better the estimated uncertainty σ approximates the true error $\Delta\psi$.

Before we start with the comparison of the Paraboloid and the Bootstrapping fit another application of the Pull value will be clarified. In IceCube this variable is not only a useful tool to compare different uncertainty fits, but it is also needed to correct the direction of the muon data events to their true neutrino direction. As already mentioned in subsection (3.1.2), the IceCube detector does not detect neutrinos directly, but only secondary particles created by neutrinos in the ice or the atmosphere. In case of track-like events, the detected particle is a muon resulting from a muon neutrino ν_μ. Unfortunately, since we are searching for astrophysical sources of neutrinos we are not interested in the direction of muons, but in the direction of the neutrinos creating these muons. The Pull correction variable can be used

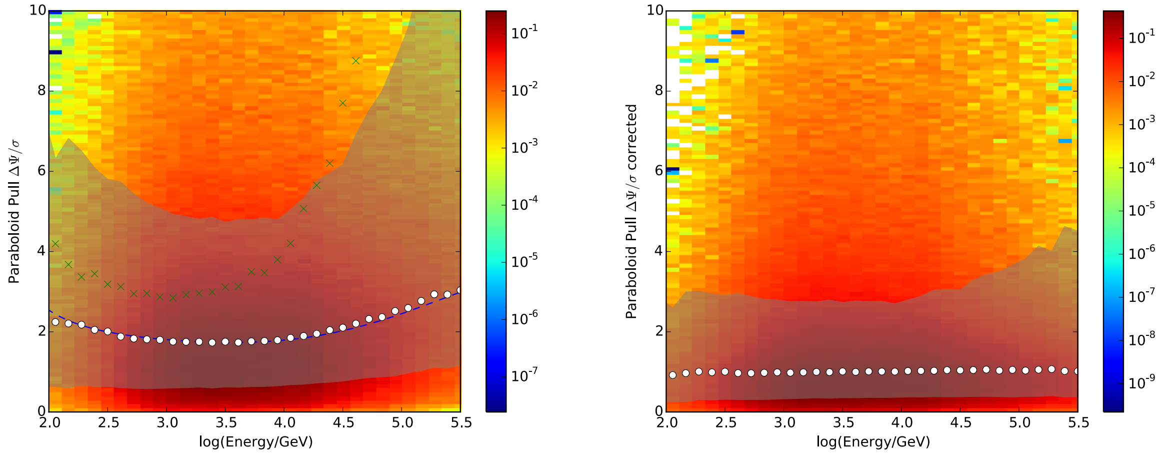


Figure 4.5: Correction of the estimated event uncertainties, using a Pull correction. ***Left:*** The Pull value for every simulated muon event is calculated and plotted in a two dimensional histogram versus the energy, evaluated by one of the energy estimators (see section 4.2), whereas each event is weighted with the corresponding event weight from the simulation. The white (green) points show the median (mean) values of the Pull in each energy bin and the grey shaded area illustrates the 80 % confidence belt, meaning the area which contains 80 % of all events. A polynonmial fit of order 6 is used to fit the median Pull values. ***Right:*** The polynomial fit is used to correct the estimated uncertainty σ towards the true deviation $\Delta\psi$. After the correction the median Pull values are perfectly distributed nearby 1, while also the confidence belt moves towards 1. The increase in area of the confidence belt at 10^5 GeV is only a statistical effect that can be explained by the much smaller number of events in this energy region.

to create a scaling factor, meaning a value to correct the uncertainty of the muon track to an uncertainty of the neutrino path. The full method is visualized and explained in figure 4.5.

Now that we know how to estimate the uncertainty of the directional reconstruction of muon neutrinos, we can compare the two possible algorithms explained above. Therefore one can make use of the corrected Pull distribution. The closer the corrected median values and the confidence belt are arranged around the unity the more accurately the fitted uncertainty describes the true situation. The left part of figure 4.6 shows the corrected pull distributions of the Paraboloid and the Bootstrapping fit.

Both fits seem to behave nearly similar, whereas if one examines more closely the 90 % belt one can see that for the Paraboloid method this area is slightly smaller and therefore the calculated uncertainties for most events are closer to the real errors.

Nevertheless the Paraboloid fit also has some disadvantages compared to the Bootstrapping method. If one tries to estimate the Paraboloid uncertainty of the SplineMPE reconstruction for all events in the final point source sample (see chapter 5), one notices that this algorithm fails for more than 10 % of the events, meaning that the two-dimensional parabola cannot be fitted correctly to the base points in the likelihood space. Mostly this failure happens for corner-clipper or badly reconstructed events, but also good-looking events are affected.

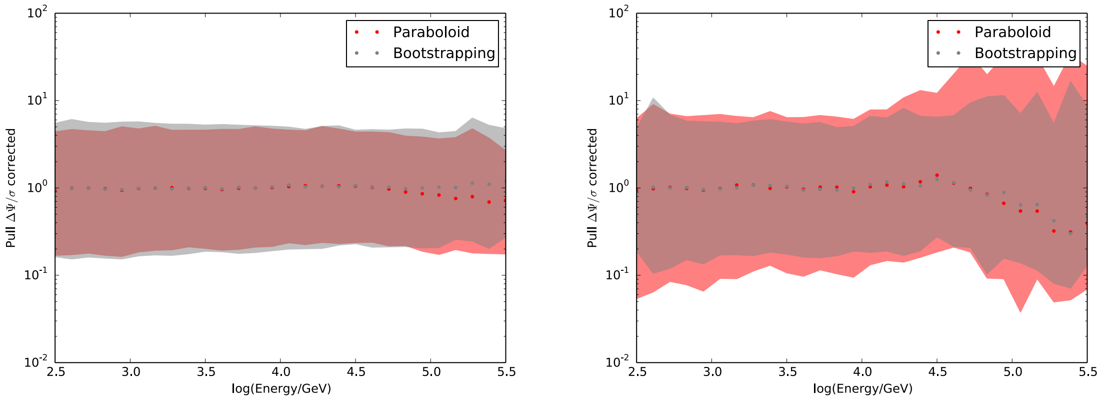


Figure 4.6: Corrected Pull distribution of the Paraboloid and the Bootstrapping uncertainty estimators. The dots represent the median values, while the shaded areas show the 90 % confidence belt. ***Left:*** All events are used. ***Right:*** Only events with a failed Paraboloid fit are used.

Although the paraboloid fit is not successful, it still yields a magnitude for the uncertainty. If one then compares the corrected Pull distributions of the Paraboloid and the Bootstrapping fit only for events, where the Paraboloid fit failed (see right part of figure 4.6) it is visible that the Bootstrapping fit is more accurate here for most events. The advantage of the Bootstrapping fit at this point is that the fit is successful for all events in the final analysis sample, since it only uses the SplineMPE method, which is required to be successful (see section 5.1).

Consequently the best method to estimate the uncertainty of the directional reconstruction consist of a combination of both methods, the results of the Paraboloid fit for all events where it ends successful and the results from the Bootstrapping method for the rest of the events. The corrected Pull distribution of this combined method is displayed in figure 4.7.

4.2 Energy Reconstruction

Next to the information about the direction of incoming neutrinos, their energy deposit inside the detector is the most interesting part. As already mentioned at the beginning of this chapter, the energy of the events is used in the point source analysis to distinguish between astrophysical neutrino events and atmospheric background. As well as for the track reconstruction, different methods exist in IceCube to reconstruct the energy deposit of the events.

4.2.1 MuEx

Since, in IceCube we do not only measure muon events whose tracks are located completely inside the detector but also events which exist outside the detection range of the detector,

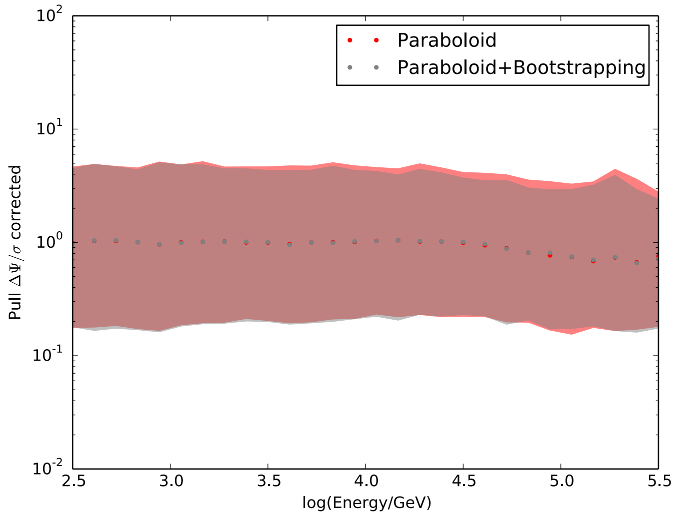


Figure 4.7: Corrected Pull distribution of the final track uncertainty reconstruction method, a combination of the Paraboloid and the Bootstrapping fit.

one cannot quantify the total energy of the events but just the visible energy deposit inside the detector. The MuEx[1] algorithm uses the the variable Λ, which represent the expected light yield per unit energy of a continuously light emitting, traveling muon to establish a connection between the energy deposit E of the muon and the expected number of photons λ inside the detector produced by the muon:

$$\lambda = \Lambda \cdot E. \tag{4.14}$$

Similar to the time residual PDFs (see section 4.1.2), Λ illustrates a function which is dependent on the distance of the track and the DOMs as well as of properties of the Antarctic ice. Within the MuEx algorithm this light yield is estimated by an analytic approximation of light emitted at the Cherenkov angle from a source moving in a straight line [28].

Assuming a Poissonian distribution of the measured number of photoelectron hits k_i at DOM i, with the expected number of hits λ_i at that DOM as the mean, one can define an overall likelihood function

$$\mathcal{L} = \prod_i \frac{\lambda_i^{k_i}}{k_i} \cdot e^{-\lambda_i} \tag{4.15}$$

Adding a noise level ρ to the mean values λ_i and inserting equation (4.14) one can rewrite the likelihood $\mathcal{L}$ as

$$\ln(\mathcal{L}) = \sum_{i=1}^{M} k_i \ln(\Lambda E + \rho) - (\Lambda E + \rho) - \ln(k_i!), \tag{4.16}$$

[1]MuEx: improved muon energy estimator

where M corresponds to the total number of IceCube DOMs. Maximizing this likelihood with respect to the only free parameter E, one obtains an estimate for the energy deposit of the muon [1, 46].

4.2.2 Millipede

In principle Millipede is using the same poissonian approach as MuEx (see equation (4.16)), but instead of calculating the number of expected events λ_i at each DOM from an analytic function, this method makes use of simulated data.

The basic concept of Millipede corresponds to a splitting of the muon track in a set of N equally spaced segments, where each of these segments k is assumed to be a point-like electromagnetic cascade with energy deposit E_k. This dispartment seems reasonable since the light output of high energy muons ($E_\mu > 100\,\mathrm{GeV}$) is dominated by stochastic energy losses, such as direct pair production, bremsstrahlung and photo-production which appear approximately as point-like particle showers along their path. The number of expected photons at DOM i resulting from segment k can then be written as

$$\lambda_i = \Lambda(\mathbf{r}_i) \cdot E_k, \tag{4.17}$$

or more general in case of regarding the total energy $\mathbf{E} = (E_1, \cdots, E_N)$ as

$$\lambda_i = \Lambda(\mathbf{r}_i) \cdot \mathbf{E}, \tag{4.18}$$

where $\Lambda(\mathbf{r}_i)$ is the light yield matrix. Inserting equation (4.18) in the likelihood function $\mathcal{L}$, one can again calculate the energy by maximizing $\mathcal{L}$ with respect to the energy $\mathbf{E}$. Due to the fact that the Millipede fit is fed with real physical information about scattering and absorption of photons, as well as information about the properties of the ice it can describe the energy deposit of muons in ice slightly better [33].

Chapter 5

Event Selection

The quality of statistical statements in a physical analysis strongly depends on the amount and the purity of the available data. Hence, it is beneficial to posses a large sample of events that consists on the one hand of as many signal events as possible while containing on the other hand as little background events as possible. In order to extract the optimal sample of events from the whole IceCube data for the analysis of astrophysical neutrino emission from point-like sources one only concentrates on charged current, track-like muon events because of their high directional resolution (see subsection 2.2.2).

As visible in figure 5.1 nearly all of the events recorded by the IceCube Neutrino Telescope are represented by atmospheric muon background while only a very small part of the sample results from astrophysical neutrinos. Thus it is necessary to develop methods which split our data set in a signal and a background part, without loosing too many signal events. For the selection of the final event sample we start with IceCube data at Muon Level3 (see section 3.2). This set is already reduced to an event rate of about 2.5 Hz, whereas still most of the events are atmospheric muons. In addition all of these events already contain the reconstruction results mentioned in chapter 4 [22].

Track-like signal neutrino events can be distinguished from atmospheric background by using their spatial and energy information. This behavior can be used to further remove background events in two different steps. First some (pre-)cuts on different reconstructed variables are applied to remove obvious background and bad reconstructed events (see section 5.1). These cuts are then followed by a more advanced boosted decision tree based event selection (see subsection 5.2).

5.1 Preliminary Data Reduction

Starting from Level3 muon events, the event selection continues with some global demands on different reconstructed magnitudes. The SplineMPE directional reconstruction algorithm (see section 4.1.2) is the main tool to determine the direction of an event in the point source analysis. Consequently we require this fit, as well as the the energy estimator SplineMPEMuEXDifferential, which is a MuEx energy reconstruction (see section 4.2.1), based on SplineMPE, to be successful for every event [22].

Apart from these global cuts the amount of the background strongly differs between the northern and the southern hemisphere. For events coming from the northern hemisphere

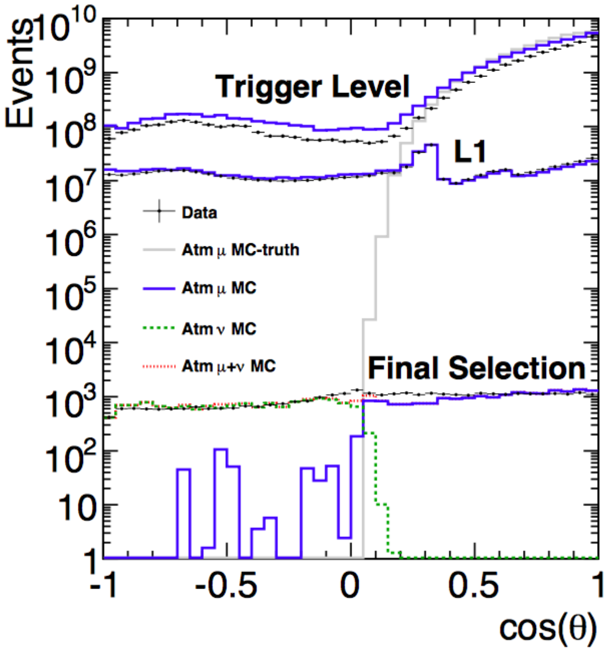


Figure 5.1: Event rate of different components inside the IceCube Neutrino Observatory for the IC40 data set (refer to section 8.3). On the one hand the rates are displayed on trigger level while on the other hand in addition they are illustrated for the final event sample. As visible in this figure, atmospheric muons dominate the background in the southern sky, while atmospheric neutrinos depict the leading part in the northern hemisphere. The figure was taken from [10].

(also called upgoing region) the earth serves as a shield against background events from atmospheric muons, whereas in the southern hemisphere these muons have a much bigger chance to pass the relatively thin layer of matter until they reach the detector. Hence it seems reasonable to apply different event selection criteria to the two different parts. The sky is thereby split according to the following rule:

- Southern hemisphere (downgoing region): $\theta \leq 85°$.
- Northern hemisphere (upgoing region): $\theta \geq 86°$.

Here θ corresponds to the zenith angle.

5.1.1 Cut Variables

Before one can start to remove a certain amount of data events from the sample it is necessary to think of appropriate cut variables. The parameters used in this event selection are listed and explained subsequently:

- $\sphericalangle(\mathbf{SplineMPE}, \mathbf{LineFit})$**:** Angular difference between the direction of the track estimated by the SplineMPE algorithm and the one from the LineFit. Since both algorithms use completely different assumptions in order to reconstruct the event track, large angular deviations indicate a low quality of the SplineMPE reconstruction. For instance large angular differences might indicate that a downgoing event is mis-reconstructed as upgoing.

- $\mathbf{n_{strings}}$**:** Number of strings that have at least one DOM with a photon hit. Having only a small number of $n_{strings}$ impedes the three dimensional reconstructability of an event track inside the detector.
- $\mathbf{n_{Dir}}$**:** Number of DOMs with a *direct* photon hit. A direct photon corresponds to a hit within a given time residual, meaning that the photon arrives at the DOM at a time close to its expected arrival time (see section 3.2). Since *direct* hits are supposed to be less scattered than other hits they contain a lot of precise directional information of the track. Hence a large number of n_{Dir} indicates a higher precision of the used directional reconstructions.
- **rlogL:** The variable rlogL is the reduced log likelihood of the track reconstruction fit. It is defined as $\text{rlogL} = \frac{\ln \mathcal{L}}{\text{d.o.f.}}$, where d.o.f. corresponds to the degrees of freedom in the track fit. Since the track reconstructions use the first hit at each DOM and fits for 5 parameters (see section 4.1), this number can be written as $\text{d.o.f.} = N_{Ch} - 5$, where N_{Ch} is the number of hit DOMs. Lower values of rLogL indicate a higher quality of the reconstruction.
- $\mathbf{l_{Dir}}$**:** Maximum distance between two *direct* hits along the particle track. A long distance l_{Dir} indicates a long visible lever arm of the track inside the detector which is favorable for a high quality of the directional reconstruction.
- $\mathbf{l_{empty}}$**:** Maximum distance between two *direct* hits, without having any other direct hit in between. Hence high value for l_{empty} might indicate that the measured hits are not just a result of one particle track inside the detector but are the composed signal of multiple particles.

5.1.2 PreCuts

Using the variables explained in the previous section, one can start to find selection criteria to obtain purer data sample. As mentioned above, the topology of background events strongly depends on the incoming direction of the events.

In the northern hemisphere the background is dominated by mis-reconstructed muons. To reject most of these while keeping the truly upgoing muon events, the precuts listed in the left column of table 5.1 are applied to the upgoing part of the data sample. Using these cuts, 98.75 % (97.65 %) of the well reconstructed (successful SplineMPE and SplineMPEMuEXDifferential fit) muon neutrinos, following a E^{-2} ($E^{-2.7}$) power law distribution, are kept while the background is reduced by a factor of two [22].

The background in the southern part of the sky mainly consists of a large amount of muon bundles, meaning the superposition of many atmospheric muon events to one high energy bundle. Since most of these events are well reconstructed and due to their abundance the

Upgoing	Downgoing
$\cos\left(\sphericalangle(\text{SplineMPE}, \text{LineFit})\right) > 0.5$	Successful SplineMPE Millipede Fit
$n_{strings} \geq 3$	$n_{strings} \geq 5$
SplineMPE rlogL < 12	SplineMPE rlogL < 9
$\sigma_{\text{MPEparaboloid}} \geq 15°$	$\sigma_{\text{MPEparaboloid}} \geq 5°$
SplineMPE $n^E_{Dir} \geq 6$	SplineMPE $n^E_{Dir} \geq 12$
SplineMPE $l^E_{Dir} \geq 75\,\text{m}$	SplineMPE $l^E_{Dir} \geq 250\,\text{m}$
SplineMPE $l_{empty} \leq 400\,\text{m}$	SplineMPE $l_{empty} \leq 400\,\text{m}$

Table 5.1: Preliminary event selection criteria for the upgoing and the downgoing region. Data taken from [22].

cuts in the southern region require harder thresholds than in the north. The cut values are listed in the right column of table 5.1. The cut efficiencies of some variables in the upgoing and the downgoing region are displayed in figure , attached to the appendix.

Besides the information we can get from of the InIce detector, also the data measured by IceTop can be used to reduce atmospheric background. Since IceTop detects atmospheric air showers, it can be used to test the InIce events for coincidences with on of these showers. If the criteria for coincidences are fulfilled these events will be removed from the sample [22].

5.2 Boosted Decision Tree based Event Selection

To further reduce the amount of background in the data sample a more advanced Boosted Decision Tree (BDT) selection method is used. As the name already implies, a decision tree consists of a series successive cuts, which divide the data sample in different classes. A schematic view of a decision tree in IceCube is displayed in figure 5.2.

Using a set of signal events, generated in Monte Carlo simulations and some part of the data as background, the variable cuts can be optimized in order to split background from signal. As a result the decision tree can determine the probability of a new unknown event to be a signal or background event. A decision tree (DT) is a highly efficient method in the classification of data, but nevertheless the method is not stable, since small differences in the data can produce large differences in the result [16].

This problem can be solved by *Boosting* the decision tree. As mentioned above a single DT is formed by using some training data (see figure 5.2), whereas at each final node the misclassified events are memorized. In the BDT method, these mis-classified events are adjusted with a certain weight to increase their importance and while using the training data with the new weights to focus on difficult event topologies the whole procedure is repeated. Altogether this process is performed multiple times. Compared to the probability of an unknown event to be signal or background like, that was determined by a single DT, the new likeliness consists of a weighted sum of all the results of all trees [16, 28]. Nevertheless the data separation method using BDTs works very accurate on the tested sample, it may suffer from overtraining, meaning that mechanism does not only split the data according

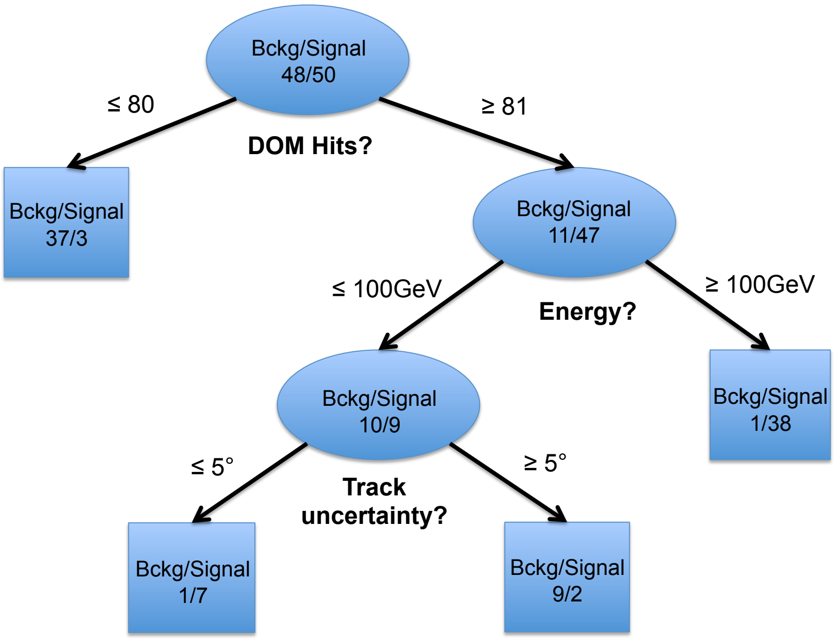


Figure 5.2: Overview of an exemplary decision tree. At each node the testing data (Signal + Background) is split according to a certain cut variable in order to optimize for the best discrimination power.

to signal and background properties but also due to statistical fluctuations in the sample. Hence the actual performance of the BDT becomes worse on the real data than on the tested sample. In order to avoid this behavior, the BDT is primarily trained only on some part of the sample (training sample), consisting of data and simulated events. Afterwards the BDT cuts are applied to the remaining part of the sample (test sample) in order to test if the BDT separation shows the same results as for the training sample.

Due to the same arguments as listed in the previous section, the BDT event selection in IceCube is split into two different parts, according to the declination of the events. Ultimately the event rate of the final sample is reduced to 10 mHz [22].

Chapter 6

Point Source Search Method

The observation of TeV neutrinos from astrophysical sources is a very hard task. The main challenge is to separate any small astrophysical neutrino signal from the exceeding background, resulting mainly from atmospheric air showers. Therefore, in IceCube one uses statistical methods to increase the potential of discovering evidence for such point sources. To maximize our capability of discovery, it is essential to establish criteria which separate the signal neutrino events from the atmospheric background. The two most important characteristics which can be used in IceCube are the angular resolution and the energy distribution of each event. Neutrino signals from astrophysical point sources are expected to be distributed in a certain area around the point source whereas events resulting from atmospheric background are locally uniformly spread in space, meaning that the frequency of their appearance is independent of their right ascension. Furthermore signal events show a different differential energy dependence. According to the theory of Fermi acceleration (see subsection 2.1.1) they follow an energy spectrum much harder than $E^{-2.7}$. At an energy range of $\sim 1\,\mathrm{TeV}$, neutrinos from atmospheric air showers follow instead a much softer $E^{-3.7}$ (prompt muons and neutrinos $E^{-2.7}$) differential energy spectrum (refer to subsection 2.1.4).

These information can be included in a search method to find statistical evidence for astrophysical point sources. Mainly there are two different ways to perform this analysis. A straight forward procedure could be realized by using angular bins in the size of the detector resolution followed by calculating the significance of an excess over the background expectation in each bin. Nevertheless this procedure is not using the full information. Once one makes use of a binary classification, the events either pass a bin cut (e.g. angular or energy) or not. Thus important event information can get lost, meaning that for instance events which are located in the center of a bin are weighted the same as events at the edge of the bin, resulting in a loss of information. Furthermore it is very hard to optimize the cuts, given a specific signal hypothesis and in case the hypothesis is not in perfect agreement with the signal the bin partition might still not be optimal [19, 54].

To avoid these handicaps the Point Source analysis in IceCube uses a different approach by performing a maximum likelihood search method which is using the spatial and the energy information on an event-by-event basis ("unbinned"). The general approach of this method will be explained in more detail in the following chapter. In the first part, the principles of the procedure of a single point source search in IceCube are explained. Section 6.2 then extends the previous analysis method in the case of using more than one point in the sky as signal source hypothesis. Finally the last section describes how interesting statistical variables (p-value, sensitivity, etc.) can be determined from the final point source data set.

6.1 Single Point Source Search

6.1.1 Maximum Likelihood Search Method

Regarding the neutrino point source search, the basic information that can be provided by the IceCube telescope is a data set of track-like muon events including information about the direction and the energy of each event. After having selected a final set of these muon events, considering different selection criteria to distinguish signal like from background like events (see chapter 5), it is desirable to discover statistical significant indications for one or more origins of these events. In order to find a local excess over the atmospheric neutrino and muon background an unbinned maximum likelihood test is performed, which uses the best information of each individual event. For this statistical purpose the IceCube data can be modeled by two different hypotheses at each specific point in the sky:

- H_0: The data consist of atmospheric background only (null hypothesis).
- H_S: The data consist of both, atmospheric background neutrinos and muons from air showers as well as astrophysical signal neutrinos emitted from one or multiple point source candidates (signal hypothesis).

The test ratio, also called test statistics, is a magnitude that tells us wether our data is better modeled by atmospheric background only or rather by background plus astrophysical signal events. Therefore the test statistic is defined as the ratio of the probability of obtaining our data under the assumption of the null hypothesis on the one hand and the signal hypothesis on the other hand. It is shown in equation (6.1):

$$\mathcal{TS} = -2\log\left[\frac{P(Data|H_0)}{P(Data|H_S)}\right] \tag{6.1}$$

Thus larger values of $\mathcal{TS}$ indicate that our data are less compatible with the background hypothesis H_0.

6.1.2 Unbinned Point Source Likelihood

As a next step one has to describe the probability densities $P(Data|H_0)$ and $P(Data|H_S)$ in terms of measurable variables given by our data events. Since we are only interested in the ratio of PDFs we do not have to take care of normalizations and therefore are allowed to use a poissonian likelihood approach. We can derive an unbinned likelihood function $\mathcal{L}$ for single source candidates starting from the case of a binned data set. Under the assumption that the number of events k_i one observes in each bin i is poissonian distributed around the mean number of events n_i that one expects to see, one can define the likelihood of a single source candidate as:

$$\mathcal{L} = \prod_i^{N_{bins}} \frac{e^{-n_i} \cdot n_i^{k_i}}{k_i!} = e^{-n_{tot}} \cdot \prod_i^{N_{bins}} \frac{n_i^{k_i}}{k_i!}. \tag{6.2}$$

Taking the limit $N_{bins} \to \infty$ one will either observe $k_i = 0$ or $k_i = 1$ events in each bin and therefore the likelihood reduces to

$$\mathcal{L} = e^{-n_{tot}} \cdot \prod_i^{N_{obs}} n_i = \lim_{dx \to 0} e^{-n_{tot}} \cdot \prod_i^{N_{obs}} (n_{tot} \cdot dx \cdot p_i), \tag{6.3}$$

where dx is the size of the bins, N_{obs} is the total number of observed events and p_i is the probability density function (PDF) of the events in bin i. Dividing the number of expected events in a signal and a background part ($n_{tot} = n_s + n_B$) and taking the approximation that the total number of events that one expects to see from background and signal sources is the same as the total number that is observed with the IceCube detector, one can further simplify equation (6.3). Moreover the probability density p_i can be expressed as a composite of background and signal PDFs, $\mathbf{B}_i$ and $\mathbf{S}_i$, with relative weights that add up to one:

$$p_i = \frac{n_s}{n_{tot}} \cdot \mathbf{S}_i + \frac{n_B}{n_{tot}} \cdot \mathbf{B}_i = \frac{n_s}{N_{obs}} \cdot \mathbf{S}_i + \frac{n_B}{N_{obs}} \cdot \mathbf{B}_i \tag{6.4}$$

Finally, after dropping all constant terms (e.g. terms including dx, N_{obs}, etc.) in equation (6.3), one can write the likelihood function for a point source, located at a certain position $\mathbf{x}_s = (\alpha, \delta)$, as

$$\mathcal{L}(n_s, \gamma | \mathbf{x}_s) = \prod_i \left[\frac{n_s}{N} \mathbf{S}_i(\mathbf{x}_i, \mathbf{x}_s, \sigma_i, E_i, \gamma) + \left(1 - \frac{n_s}{N}\right) \mathbf{B}_i(\delta, E_i) \right], \tag{6.5}$$

This function is now dependent on the number of signal events n_s that we expect to see as well as on the spectral index γ of the source. The probability densities in equation (6.1) can be interpreted by $\mathcal{L}$ setting $n_s = 0$ for the background hypothesis H_0 and $\hat{n}_s \geq 0$ for the signal hypothesis H_S . The test statistic then evaluates to

$$\mathcal{TS} = -2 \operatorname{sgn}(n_s) \log \left[\frac{\mathcal{L}(n_s = 0)}{\mathcal{L}(n_s = \hat{n}_s)} \right] = 2 \operatorname{sgn}(\hat{n}_s) \sum_i \log \left[\frac{\hat{n}_s}{N_{obs}} \left(\frac{\mathbf{S}_i}{\mathbf{B}_i} - 1 \right) + 1 \right]. \tag{6.6}$$

The number of expected signal events $\hat{n}_s$ and the spectral index γ are unknown parameters and must be determined by maximizing the $\mathcal{TS}$ with respect to both variables (see section 6.3) [20].

6.1.3 Signal and Background PDFs

The probability densities for events to be signal or background like, used in the Point Source likelihood (see equation (6.5)) are calculable quantities by means of using physical assumptions on both models (signal and background) and the data provided by the IceCube detector.

Suppose one wants to find evidence for an astrophysical source at position $\mathbf{x}_s = (\alpha, \delta)$, that emits signal events from an $E^{-\gamma}$ spectrum. The data sets we get from the IceCube detector consist of track-like muon events, each equipped with a reconstructed position $\mathbf{x}_i$ and a reconstructed energy estimation E_i. For each of these events one can determine a probability density to be a signal neutrino event resulting from a source at direction $\mathbf{x}_s$. These event PDFs are composed of a spatial and an energy part:

$$\mathbf{S}_i(\mathbf{x}_i, \mathbf{x}_s, \sigma_i, E_i, \delta, \gamma) = \mathcal{S}_i(\mathbf{x}_i, \mathbf{x}_s, \sigma_i) \cdot \mathcal{E}_i(E_i, \delta, \gamma) \tag{6.7}$$

For the the spatial signal probability we can assume that signal events cluster around the source position $\mathbf{x}_s$. Therefore an event that is close to the direction of the source is more likely to be signal-like than muons, detected far-off. Using the directional information of each event (position $\mathbf{x}_i$ and estimated uncertainty σ_i) that we receive from our data, we can define the spatial signal PDF as a two dimensional gaussian distribution around the source position $\mathbf{x}_s$:

$$\mathcal{S}_i(\mathbf{x}_i, \mathbf{x}_s, \sigma_i) = \frac{1}{2\pi\sigma_i^2} \cdot \exp\left(\frac{\|\mathbf{x}_i - \mathbf{x}_s\|^2}{2\pi\sigma_i^2}\right). \tag{6.8}$$

The likelihood of being background-like can be calculated in a similar way, as a product of directional and energy likeliness. As mentioned at the beginning of this section, background events are locally uniform distributed in space, meaning that they are randomly dispersed in right ascension. To identify the portion of background at different declination points we use scrambles of our data in right ascension, representing sets of background events. From these sets we can then obtain the spatial background distribution as a function of the declination. Equation (6.9) shows this spatial part of the background PDF of event i

$$\mathcal{B}_i(\mathbf{x}_i) = \frac{1}{2\pi} \cdot P(\delta_i), \tag{6.9}$$

where the factor $\frac{1}{2\pi}$ results from the uniform distribution in right ascension.

To calculate the test statistic $\mathcal{TS}$ (see equation (6.6)) one only needs to know the ratio of signal to background probability for each event. While the spatial event probabilities are determined independently for signal and background, we only compute the ratio of the energy part of the PDFs. This energy likelihood proportion is evaluated from simulation, using Monte Carlo neutrino events as signal data and scrambles of the experimental data as background.

6.1.4 Combination of Multiple Data Sets

The IceCube detector at the South Pole developped over the years from only a few strings, containing DOMs to a total of 86 strings with 60 DOMs each [60]. Therefore according to different run-times, the different data sets that were recorded over the years cannot be treated as being completely identical. Nevertheless it is necessary to combine all available data to increase our statistics and consequently also our chance of discovering the origin of astrophysical neutrinos.

Luckily unbinned likelihood methods are more or less thankful tools if one wants to use multiple different data sets. Each event carries its own signal and background probability density function which is determined by the corresponding data set. The source likelihood function $\mathcal{L}$ then becomes the product of the likelihood functions from each data set:

$$\begin{aligned}\mathcal{L}(n_s, \gamma|\delta) &= \prod_{j}^{N_{sample}} \mathcal{L}^j(n_s^j, \gamma|\delta) \\ &= \prod_{j}^{N_{sample}} \prod_{i\in j} \left[\frac{n_s^j}{N_{tot}^j} \mathbf{S}_i^j(\mathbf{x}_i, \mathbf{x}_s, \sigma_i, E_i, \delta, \gamma) + \left(1 - \frac{n_s^j}{N_{tot^j}}\right) \mathbf{B}_i^j(\delta, E_i) \right], \end{aligned} \tag{6.10}$$

where n_s^j corresponds to the mean number of signal events that we expect from sample j. The likelihood ratio $\mathcal{TS}$ is maximized globally, expecting the source to originate neutrinos from the same differential energy spectrum in all data sets. Therefore the likelihood $\mathcal{L}$ remains a function of only two parameters: the global spectral index γ and the total mean number of signal events n_s that we expect to see from all data sets combined [3].

In general the number of signals n_s^j is not identical for all the sets, but relies on different detector specific properties, such as the live-time and the detector acceptance. The fraction of signal events that results from each individual sample, $f^j(\gamma)$, can be evaluated using Monte Carlo simulation:

$$n_s^j = f^j(\gamma) \cdot n_s, \quad n_s = \sum_j n_s^j \tag{6.11}$$

Taking care of these changes, the test statistic $\mathcal{TS}$ from equation (6.6) expands to

$$\mathcal{TS} = 2\,\mathrm{sgn}(\hat{n}_s) \sum_{j}^{N_{sample}} \sum_{i\in j} \log \left[\frac{\hat{n}_s^j}{N_{tot}^j} \left(\frac{\mathbf{S}_i^j}{\mathbf{B}_i^j} - 1 \right) + 1 \right]. \tag{6.12}$$

Besides the two essential variables n_s and γ in this formula, also other physical parameters as for instance the arrival time of the events might be considered in order to further separate the signal from the background hypothesis. Nevertheless since in this thesis a time independent point source search is performed, the number of signal events n_s and the spectral index γ of the source are the only free parameters.

6.2 Stacking of Point Source Candidates

Until now we always considered the data to result from one single point source located at position $\mathbf{x}_s$. Another way to regard the origin of the IceCube data is to consider M source candidates which can emit astrophysical neutrinos at the same time. Therefore we regard our experimental data not only as a signal from one source plus atmospheric background, but now as a set of data events resulting either from one of the multiple sources or from atmospheric secondary particles.

To perform such a stacked analysis of multiple source candidates, only a few changes have to be implemented in the search method, compared to the one used for single source candidates. Mainly only the signal PDFs for all events have to be modified, while the background probabilities and the whole likelihood formalism stay the same. The likelihood of one event to be signal-like can now be treated as a weighted sum of all the signal PDFs, the event would have at each of the single source positions:

$$\mathbf{S}_i \to \mathbf{S}_i^{Stack} = \frac{\sum_{k=1}^{M} W^k R^k(\delta_k,\gamma)\cdot \mathcal{S}_i^k(\mathbf{x}_i,\mathbf{x}_{s,k},\sigma_i)\cdot \mathcal{E}_i(E_i,\gamma)}{\sum_{k=1}^{M} W^k R^k(\delta_k,\gamma)} \tag{6.13}$$

Each source candidate k is multiplied with two different kind of weights. On the one hand we consider a weight that represents the technical properties of the detector geometry. The efficiency of the IceCube telescope to detect neutrino events depends strongly on the direction of the event entering the telescope. Thus each source candidate gets a detector weight $R^k(\delta_k,\gamma)$, which represents the detection efficiency of IceCube for signal events originating from a source at position $\mathbf{x}_s = (\alpha,\delta)$, emitting neutrinos from a differential $E^{-\gamma}$ energy spectrum. These weights can be calculated using Monte Carlo simulations. On the other hand different properties of the individual sources (e.g. redshift, neutrino emission strength etc.) can also result in different chances to see signal events from each of those sources. These probabilities, mostly based on theoretical assumptions on the sources are merged in the second weight, the theoretical weight W^k, which is in particular independent of the position and spectral index of the sources [54].

Once one is going to use multiple years of data *again* (compare to section 6.1.4), the effective weight $f_j(\gamma)$ of each individual data set j has to be reconsidered as well. According to Bayes' theorem [43], the weight for each set extends to

$$f_j(\gamma) = \sum_{k=1}^{M} P(j|\delta_k,\gamma)\cdot P(\delta_k|\gamma), \tag{6.14}$$

where M corresponds to the number of sources. $P(\delta_k|\gamma)$ describes the relative likelihood of seeing signal events from source k compared to the remaining $M-1$ candidates, on the basis of the combination of all data sets. In case of just one point source candidate this value automatically becomes 100% and hence the whole weight $f_j(\gamma)$ immediately turns out to be the same as in equation (6.11), as expected. As before $P(j|\delta_k,\gamma)$ represents the feasibility of our signal, which originate from source k, to arise from data set j compared to the others. All PDFs appearing in equation (6.14) can be determined from Monte Carlo simulations, using the detection efficiency parameter of IceCube.

Using the stacking signal PDF and taking care of the data set weighting, the $\mathcal{TS}$ can be calculated in the exactly same way as for a single source, following the equations (6.5) - (6.6).

6.3 Significance, Sensitivity and Discovery Potential

In the whole previous chapter we have introduced the test statistic $\mathcal{TS}$, a magnitude calculable exclusively from information provided by the IceCube data and two free parameters n_s and γ, which provides an indication of the likeliness of the data to be a result from one or more declared points in the sky. Looking at equation (6.1) it is easily visible that the IceCube data is most likely compatible with the signal hypotheses H_S for large $\mathcal{TS}$ values. Therefore the test statistic is maximized with respect to the number of expected signal events n_s and the spectral index γ. Nevertheless the maximal test statistic value just gives an indication wether one statistically see a point source or not. To get a significant statement of a discovery one still has to test, wether the same outcome might also result from pure background.

In the following subsection the estimation of a p-value, as well as a sensitivity, discovery potential and upper Limits of a source hypothesis will be explained. Due to the fact that the calculation of the test statistic is the same also for multiple sources, these values can be evaluated in the exactly same way for multiple stacked source candidates.

6.3.1 P-Value

A p-value is a very common and powerful tool in the evaluation of statistical analyses. Having obtained a certain result, it describes the probability to get at least such an extreme outcome given the null hypothesis H_0. Therefore a p-value is in principle a measure of how unexpected a result is, under the assumption of H_0. The smaller the value, the less compatible is the null hypothesis with the result. The null hypothesis can be rejected once the p-value passes a certain fixed threshold level (typically 1.0 % or 0.1 %) and the result can be regarded as statistically significant. Nevertheless one has to keep in mind that significance cannot be equated with the prediction that signal hypothesis H_S corresponds to the truth but just that it is very unlikely to get this result under the assumption of H_0 [26].

In the case of the test statistic $\mathcal{TS}$ function in IceCube, the p-value is defined as

$$\text{p-value}(\tilde{\mathcal{TS}}) = P(\mathcal{TS} \geq \tilde{\mathcal{TS}}|H_0) = 1 - \int_0^{\tilde{\mathcal{TS}}} P(\mathcal{TS}|H_0)\, d\mathcal{TS}. \tag{6.15}$$

In order to calculate this p-value, the overall test statistic probability distribution $P(\mathcal{TS}|H_0)$ for background data has to be known. As already mentioned in subsection 6.1.3, background data can be created by scrambling the data events in right ascension. Performing a test statistic maximization for multiple background scrambles, the background probability distribution can be created. The p-value can then be calculated according to equation (6.15). The principle procedure is also presented schematically in figure 6.1 [26].

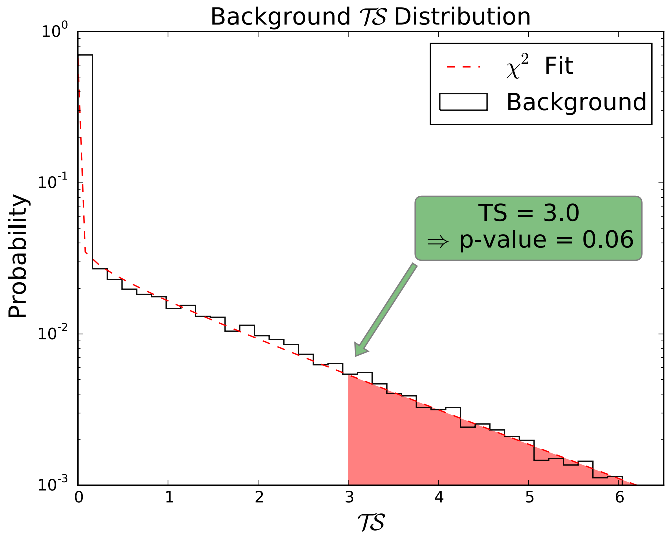


Figure 6.1: P-value evaluation, using the background $\mathcal{TS}$-distribution. The test statistic PDF is created from data scrambles and fitted by a χ^2-function. The p-value can then be estimated according to equation (6.15).

Decision \ **Truth**	H_0 is True	H_0 is False
H_0 is True	Correct Decision $p = 1 - \alpha$	**Type II error** $p = \beta$
H_0 is False	**Type I error** $p = \alpha$	Correct Decision (Power) $p = 1 - \beta$

Table 6.1: Possible outcomes of a hypothesis test.

6.3.2 Sensitivity, Discovery Potential and Upper Limits

Due to the statistical nature of a hypothesis test, the resulting implication, meaning the rejection or the acceptance of the null hypothesis, is in general not free of error. In general there are four different possible outcomes, listed in table 6.1.

If the decision corresponds to reality, then a correct judgement was made. Alternatively the decision was incorrect. In general one can divide the second scenario in two different ways. The so-called **type I error** α is the probability of rejecting the null hypothesis, although it is true. Therefore it represents a measure of a false discovery claim of an analysis. The second possible error that can occur, the **type II error** β, corresponds to the contrary setting. In this scenario the null hypothesis is not rejected although it does not match the truth. Thus it describes in some way the power of the test. The smaller β is, the more often you decide correctly to accept the signal hypothesis.

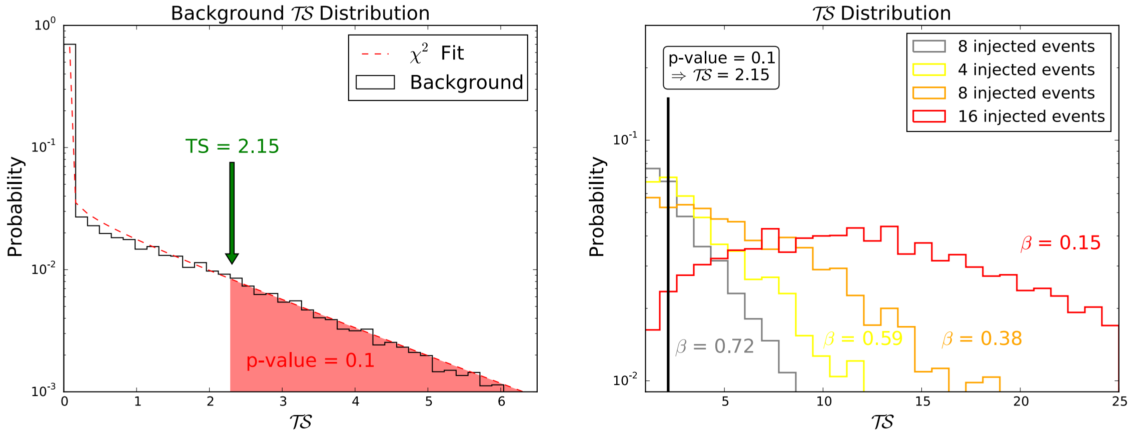


Figure 6.2: Schematic overview of the determination of a flux that is necessary to yield a p-value smaller than 0.1 in 90 % of the trials ($\rightarrow \alpha = 0.1$, $\beta = 0.1$). *Left:* Estimation of the test statistic value corresponding to the p-value of 0.1, using the $\mathcal{TS}$ PDF of scrambled background samples. A χ^2-distribution, in combination with a δ-distribution centered at 0 is used to fit the background probability curve. *Right:* Signal events are injected to the background samples in order reach a power of 90 % ($\beta = 0.1$), meaning that one reaches the given p-value in 90 % of all trials. The more events one injects, the closer one gets to the desired value of β. The desired flux corresponds in the end to the minimal number of events that fulfills the criteria on both errors, type I and type II.

Having the theoretical knowledge, it is now interesting how these information can be used to evaluate meaningful variables in IceCube. Three of the most common quantities that can be computed for IceCube analyses are defined below:

- **Sensitivity:** Flux, which yields a p-value smaller than 0.5 ($\rightarrow \alpha = 0.5$) in 90 % of the trials ($\rightarrow \beta = 0.1$).
- **Discovery Potential:** Flux, which yields a p-value smaller than 5σ ($\rightarrow \alpha \approx 2.87 \cdot 10^{-7}$) in 50 % of the trials ($\rightarrow \beta = 0.5$).
- **90%-Upper Limit:** Flux, which yields a p-value smaller than the p-value from the analysis with unscrambled data ($\rightarrow \alpha =$ p-value) in 90 % of the trials ($\rightarrow \beta = 0.1$).

The sensitivity declares the threshold flux starting from which it is very likely to fit a p-value smaller than 0.5 although the signal might be the result from background. Hence sources with a flux higher than the sensitivity flux should at least be visible, in the sense that they obtain a p-value smaller than 0.5. The discovery potential represents the flux that is necessary in order to be able to discover a significant sign of a neutrino point source with a 50 % chance of not seeing this signal. Hence sources with a flux above the discovery potential threshold are very likely to produce a statistically significant sign inside the IceCube detector. Ultimately in case of not measuring any significant results, the upper limit depicts the flux threshold starting from which one would detect at least the measured signal with a chance of 90 %.

Using the different specifications on the the type I and type II error, all these quantities can be determined in a uniform manner. Starting from a fixed value for α, the test statistic value corresponding to this type I error can be calculated in the same way as the p-value (see subsection 6.3.1), using data scrambled in right ascension to calculate the background probability distribution. To estimate the type II error we inject simulated signal events from Monte Carlo simulations following a predefined flux. The test statistic minimization is then repeated multiple times, with different number of injected signal events. Having also a predefined magnitude for β one can then calculate the number of signal events, and consequently also the flux, that are needed to reach a power of $1 - \beta$. A schematic overview of this procedure is also represented in figure 6.2.

In order to evaluate the Upper Limit the first step can be skipped and the maximal test statistic value from the real unscrambled data can be used directly to perform the second step.

Chapter 7

1WHSP Blazar Catalog

In the previous chapter, the analysis method used for the search for the origin of astrophysical neutrinos, was presented. As mentioned thereby, one can mainly distinguish two different scenarios. One the one hand, one can consider all recorded data to be either from one source point in the sky or from atmospheric background, while on the other hand it is also possible to assume multiple source candidates at different locations in the sky to be the generators of the neutrino signals. In the first scenario, in principle the whole sky is scanned along a certain grid, in order to find the most significant locations, meaning the locations where it is most probable to have neutrino generating sources. Thereby it is in principle not necessary to have any pre-knowledge about the scanned locations.

This turns out to be a major difference to the second setting. Once one starts to assume multiple neutrino sources, it is essential to start the analysis with a set, or catalog of source candidates which are likely to be high energy neutrino emitters. In principle theoretically this criteria is not mandatory, but since the amount of possible combinations of an arbitrary number of different sources nearly goes to infinity, it is computationally not possible to determine the significance of all these combinations. Assuming that the whole sky is split in a grid of 100 points, and 10 of these points are neutrino emitting source locations, this would already yield in a total of $\sim 1.7 \cdot 10^{13}$ possibilities to distribute the 10 sources on the 100 grid points. Consequently it seems exclusively reasonable to perform a stacking analysis on a set of source candidates, which all indicate to be neutrino sources.

Since the analysis performed in this thesis follows the principle of the second scenario from above, the following chapter starts with a description of Active Galactic Nuclei and in particular blazars, which are expected to contribute a major fraction to the amount of astrophysical neutrinos reaching the earth. Afterwards in section 7.2, the composition of the first WISE[1] High Synchrotron Peaked (1WHSP) blazar catalog, which is used as assemblage of neutrino sources in this analysis will be shortly presented. Finally the last part of this chapter, section 7.3 elaborates the differences of the 1WHSP catalog compared to other, previously used Blazar-Source-Catalogs in IceCube. Hence this section in addition illustrates a motivation for the point source stacking analysis, using the 1WHSP sources.

[1]WISE: Wide-field Infrared Survey Explorer [66]

7.1 Active Galactic Nuclei and Blazars

Active Galactic Nuclei (AGNs) are one of the most luminous objects in the entire universe, producing photons over a broad emission band, ranging from radio to TeV gamma ray energies. Consequently AGNs can be regarded as special laboratories for extreme physics, whereas they are not only useful probes of the Universe on large scales but moreover are promising candidates for high energy CRs [64, 54].

As already mentioned in subsection 2.1.2, the structure of an AGN consists of a rotating supermassive black hole attracting surrounding matter due to its gravitational potential energy. While being pulled towards the black hole, this material loses angular momentum in turbulent processes, building and heating up the so-called accretion disc (see figure 2.4). Clouds of gas rapidly moving in the potential of the black hole, close to the accretion disc produce strong optical and ultraviolet emission lines, whereas along some lines of sight, this radiation is obscured by a torus of gas and dust, which is formed far away of the accretion disc. The light emission, caused in this region is identified as broad-line emission. Beyond this region, and hence not being affected that strongly by the dust torus, narrow lines are emitted from slower moving clouds of gas in the narrow-line region. Perpendicular to the accretion disc and the dust torus, relativistic radio-emitting jets, consisting of particles at nearly the speed of light, are formed [64].

7.1.1 Classification of Active Galactic Nuclei

Since AGNs are axis-symmetric objects, their appearance on earth strongly depends on their observation angle. Moreover other characteristics, such as for instance the mass of the black hole or the mass accretion influence observed properties of an AGN. Hence historically, based on different observations, a confusing amount of different AGNs types and names were noted. Nevertheless, by using the most important empirical information of AGNs, these different types can be put in some order. An empirical classification of AGN, according to their radio loudness and the characteristics of their optical and ultraviolet spectra is shown in table 7.1. This classification scheme is explained in the following [64].

	Emission Line Characteristics		
	Type 0	**Type 1**	**Type 2**
Radio-Quiet:		Seyfert 1	Seyfert 2 NELG
Radio-Loud:	BL Lac FSRQ	BLRG SSRQ FSRQ	NLRG

Table 7.1: Empirical Classification of AGN.

AGNs can be divided into radio quiet and radio-loud objects, meaning that for the second scenario the radiated flux is dominated by radio emission (radio flux / optical flux > 10), whereas this effect is not regarded for the first type. Radio loudness of an object might be associated with the type of its host galaxy and the spin of the black hole, which might allow the creation of large scale radio jets. Approximately 15-20 % of all observed AGNs correspond to radio-loud objects.

Next to the radio loudness, AGNs can be classified according to the characteristics of their optical and ultraviolet spectra. Thereby AGNs can be split into three main groups [64]:

- **Type 1 AGN:** Objects corresponding to this group,are those with broad emission lines resulting from hot, fast moving clouds of gas located in the vicinity of the central black hole (refer to 2.1.2).
- **Type 2 AGN:** Type 2 AGNs, only show weak, narrow emission lines since emission from the hot, high velocity gas is concealed by the dust torus surrounding the center of the object (see section 7.1 and figure 2.4).
- **Type 0 AGN:** There are still AGN objects that show an unusual spectral behavior and hence do not fit the criteria of the first two types. These objects are identified with type 0 AGNs, whereas on assumes that the characteristics of these objects are linked to a very small angle of the jet to the line of sight.

As already mentioned above, different known objects participating in this classification are listed in table 7.1. Since the 1WHSP catalog only consists of blazar candidates, we mainly only examine characteristics of blazars in the rest of this chapter.

7.1.2 Blazar

Blazars display a subclass of radio-loud Active Galactic Nuclei. They are expected to be those types of AGN, whose relativistic jets point close to the direction of an observer on earth. Hence the spectral energy distribution (SED) of blazars is affected by mainly three different radiation types. In the optical band, the broadest hump (also called synchrotron peak ν^S_{peak}) is caused by non-thermal radiation, which can be associated with synchrotron radiation of relativistic electrons in the magnetic field of the jets. Moreover a thermal component, induced by the clouds of gas in the broad-line region as well as light output from the host galaxy can be visible in the optical energy region of the SED of a blazar. Additionally a second big hump, induced by non thermal radiation generated in the relativistic jets appears at γ-ray energies. Since there exists no established model for the emission in this energy region, two different scenarios, which are currently expected to be potentially responsible are introduced in subsection 7.1.2.1 below. Moreover typical SEDs of two different blazars are shown in figure 7.1 [32, 63].

Arising from different properties in the spectral energy distribution, blazars can be further classified. Using the same criteria as applied for classification of general AGN objects (see table 7.1), blazars can be split into flat-spectrum radio quasars (FSRQs) and BL Lacertae objects (BL Lacs). FSRQs show strong broad emission lines in the optical spectrum, while for BL Lacs mainly only weak emission lines are visible.

Moreover another partition of blazars also seams reasonable, and in case of the neutrino point source search in this thesis even more important. Since the synchrotron peak ν^S_{peak},

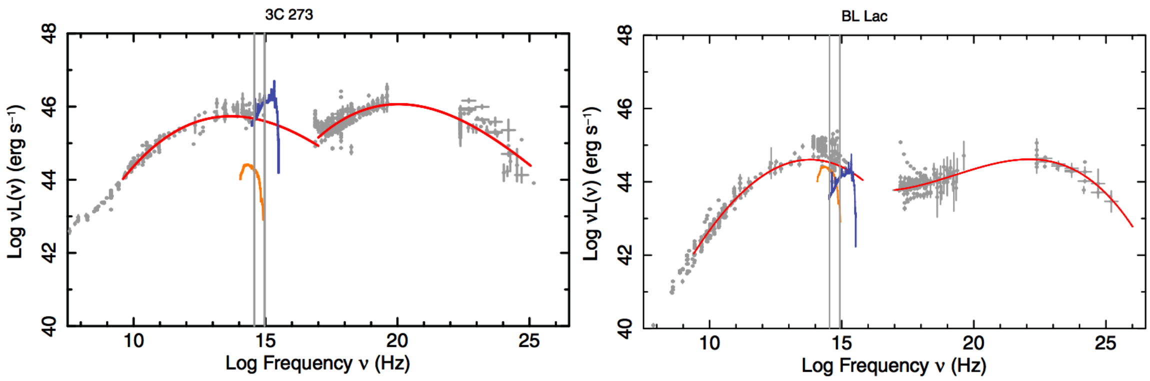


Figure 7.1: Both figures show the spectral energy distribution, meaning the differential flux as a function of the frequency of blazars. The colored lines signify the different components contributing to the distribution. The two red humps represent the non-thermal flux, caused by the jets, the blue line can be identified with the emission from the broad-line region and finally the orange curve displays flux caused by photons from the host galaxy. ***Left:*** SED of the FSRQ 3C273. The blue curve shows the strong, broad emission lines in the optical spectrum. ***Right:*** SED of a BL Lac. Compared to the FSRQ on the left, the flux of this blazar does not show such broad emission lines. Both pictures were taken from [32].

reflecting the maximum energy particles can be accelerated to inside the jets, can appear at different energies (respectively frequencies) blazars can be divided in low-frequency peaked, intermediate-synchrotron peaked and high-synchrotron peaked objects, according to the particular position of the synchrotron peak frequency ν^S_{peak} [32]:

- Low-Synchrotron Peaked (LSP): $\nu^S_{peak} < 10^{14}$ Hz
- Intermediate-Synchrotron Peaked (ISP): 10^{14} Hz $< \nu^S_{peak} < 10^{15}$ Hz
- High-Synchrotron Peaked (HSP): $\nu^S_{peak} > 10^{15}$ Hz

7.1.2.1 Blazar Emission Models

As already mentioned above, the origin of the photon flux in the high energy region is still an open field of study. According to the current state of knowledge, there are mainly two possible mechanism, contributing to this spectral component.

Leptonic Model

The leptonic effect, causing the photon flux in the high energy region is the inverse-Compton (IC) effect. Low energy photons are inverse-Compton scattered by the relativistic electron population inside the jets, gaining a fraction of the electron's energy each time. Hence these photons can be scattered up to very high energies. The primary photons can either arise locally from inside the jets, that is to say the synchrotron photons generated by relativistic moving electrons or they originate from external sources, such as for instance the accretion disc, the broad-line region or the dust torus [55].

Hadronic Model

Once relativistic protons inside a jet reach the energy threshold for pion production in photon-proton interactions, synchrotron-supported pair cascades will develop.

$$p\gamma \longrightarrow \Delta^+ \longrightarrow \begin{cases} n\pi^+ & \longrightarrow n\mu^+\nu_\mu \longrightarrow ne^+\nu_e\bar{\nu}_\mu\nu_\mu \\ p\pi^0 & \longrightarrow p\gamma\gamma \end{cases} . \tag{7.1}$$

Moreover, in hadronic models, the collision of protons with other ambient protons can lead to production of pions, according to

$$pp \longrightarrow \begin{cases} pn\pi^+ & \longrightarrow pn\mu^+\nu_\mu \longrightarrow pne^+\nu_e\bar{\nu}_\mu\nu_\mu \\ pp\pi^0 & \longrightarrow pp\gamma\gamma \end{cases} \tag{7.2}$$

In order to accelerate protons to the required ultra-relativistic energies, very high magnetic fields inside the jets are necessary (see section 2.1.2). Hence also the synchrotron emission of the primary protons and the secondary leptons has to be considered in order to describe a self-consistent synchrotron-proton blazar model (SPB) [18, 54].

7.1.2.2 Blazars as Neutrino Point Source Candidates

Next to the photons which are created in hadronic blazar models also high-energy neutrinos remain as a final product in the decay of charged pions (see equations (7.1) and (7.2)). Consequently, in case of hadronic emission in astrophysical objects one would expect the appearance of neutrinos in company with photons at very high energies, whereas the energy and the flux of the photons at the source is about a factor of two higher than for neutrinos [49, 40].

Hence, since the IceCube collaboration detected astrophysical neutrinos in an energy region between 30 TeV and a few PeV (see section 3.3), this would also imply the existence of very high energy (VHE) photons coming from the same direction and reaching energies in the range of $\gtrsim 60$ TeV. Unfortunately there are mainly two reasons obscuring the verification of this connection. One the one hand, current γ-ray detectors can only detect photons up to energies of ~ 40 TeV and consequently do not have the ability to detect the intended photons. On the other hand, depending on the properties of the astrophysical object and its distance to the earth, this photon signal might also be diluted due to absorption effects [49]. Once a VHE photon travel in the vicinity of extragalactic background light it gets absorbed, resulting in an electron-positron pair:

$$\gamma_{VHE} + \gamma_{EBL} \longrightarrow e^+ + e^- \tag{7.3}$$

These pairs then lose their energy in interactions with photons from cosmic microwave background, boosting them up to γ-ray energies [15].

Altogether one can note, that the direct detection of high energy photons being counterparts of the IceCube neutrino events is a very hard attempt at the moment. Nevertheless, since the creation of PeV neutrinos in hadronic processes, requires protons accelerated up to at least 10^{16} - 10^{17} eV, blazars are still some of the most likely sources of astrophysical neutrinos. Moreover the detection of the direct photon-neutrino connection would confirm the correctness of the hadronic emission model [49].

7.2 Composition of the Catalog

The first WISE[2] High Synchrotron Peaked (1WHSP) catalog consists of 998 sources, being either already confirmed or candidates of HSP γ-ray blazars. Thus it is currently the largest list of HSP blazars. According to current observations and the previous subsection, HSP blazars are bright sources of high energy TeV photons. Hence all of the objects in the catalog are expected to be origin of photons up to the highest γ-ray energies [15].

In order to create such a large sample of HSP blazars, it seems reasonable to use a large sample of multi-frequency data selected according to criteria which are specific for the SED of HSP blazars. Hence the starting point for the composition of the 1WHSP catalog is the AllWISE[3] catalog, including 747 million objects from all over the sky. Following different steps, these objects are selected for HSP blazars. Since the infrared spectral slope of the non-thermal radiation, resulting from the relativistic jets in blazars, typically represents an energy power law with index values between 0.4 and 0.8, blazars show the tendency to concentrate in a particular region of the infrared colour-colour diagram[4] (refer to figure 7.2). Hence taking only those objects from the AllWISE catalog, which are located in this region, reduces the size of the starting sample to approximately 4.8 million [15].

After removing further non-blazar objects, with the help of several additional constraints, based on the well-known broad-band spectral features, the size of the sample adds up to only 1347 objects. The spectral energy distributions of all of these objects are then analyzed individually, whereas an object is only kept if the position of its fitted synchrotron peak is bigger than $\nu^S_{peak} > 10^{15}$ Hz (see subsection 7.1.2). This yields a final set of 857 HSP blazar candidates from the AllWISE catalog [15].

Ultimately the 1WHSP catalog is completed by 141 additional objects, which did not fulfill all the selection criteria from above, but still illustrate likely HSP blazar candidates. An extract of the whole 1WHSP catalog is shown in figure 7.3 [15].

7.3 Properties of the 1WHSP Catalog

Since the composition of the 1WHSP catalog was clarified in the previous section, it is now time to discuss the properties of this HSP blazar sample. At first it seems reasonable to

[2]WISE: Wide-field Infrared Survey Explorer
[3]AllWISE: Latest version of the all-sky WISE catalog
[4]Colour-colour diagram: tool to compare the brightness of astrophysical objects at different wavelengths

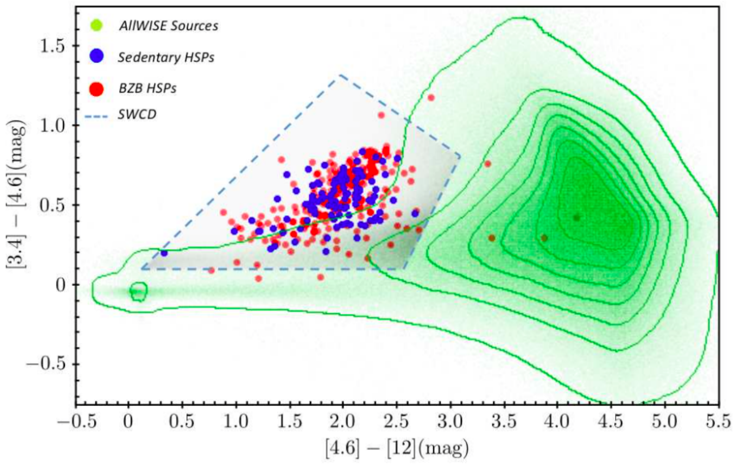


Figure 7.2: Illustration of the infrared colour-colour diagram of the AllWISE catalog. Moreover the infrared colours of several HSP blazars from the Sedentary Survey [31] and the BZCAT [45] catalog are displayed. As visible most of these blazars accumulate in the grey shaded area. Hence only the AllWISE objects being located inside the area which delimited by the blue dashed lines are selected to built the 1WHSP blazar sample. The figure was taken from [15].

1WHSPJ	Bz	Log(ν_p)	Log($\nu f_{(\nu_p)}$)	z	Fermi-LAT	Γ	Type	TeV	FOM
000116.3+293534	-	>17	>-12.8	0	-	-	1	-	0.03
000132.7-415525	-	15.8	-11.7	0	2FGL J0001.7-4159	2.1±0.19	1	-	0.40
000213.7-103816	-	16.6	-13.1	0	-	-	1	-	0.02
000319.5-524727	-	15.4*	-12.4*	0	3FGL J0003.2-5246	1.9±0.17	1	-	0.08
000513.7-261438	-	15.4	-12.6	0	-	-	1	-	0.05
000835.3-233927	y	>17	>-11.8	0.147	2FGL J0008.7-2344	1.6±0.25	3	-	0.31

Figure 7.3: Extract of the 1WHSP Blazar Catalog. This figure was taken from [15].

check the position of all of these objects. As visible in figure 7.4 the blazars are in the main uniformly spread over the whole sky, except from a certain band (±20° in galactic latitude b) around the position of the galactic plane.

Moreover, since only 35 of the 998 blazars have been detected in the TeV range, the WHSP catalog provides a quantitative measure of the potential visibility in this energy range for each object. This value is introduced as the Figure Of Merit (FOM) of each blazar candidate i in the catalog and is defined as

$$\mathrm{FOM}_i := \frac{(\nu f_\nu)_{peak,i}}{(\nu f_\nu)_{peak,0}}, \tag{7.4}$$

where $(\nu f_\nu)_{peak,i}$ corresponds to the synchrotron peak flux of each blazar i and $(\nu f_\nu)_{peak,0}$ to the one of the faintest blazar candidate, which was already detected in the TeV band. Hence an object, having a FOM ≥ 1.0 implies that its synchrotron peak flux is at least as bright as the flux of the faintest already known TeV blazar [15].

The reasonability of the FOM as a measure for the likeliness of a blazar to be a TeV γ-ray

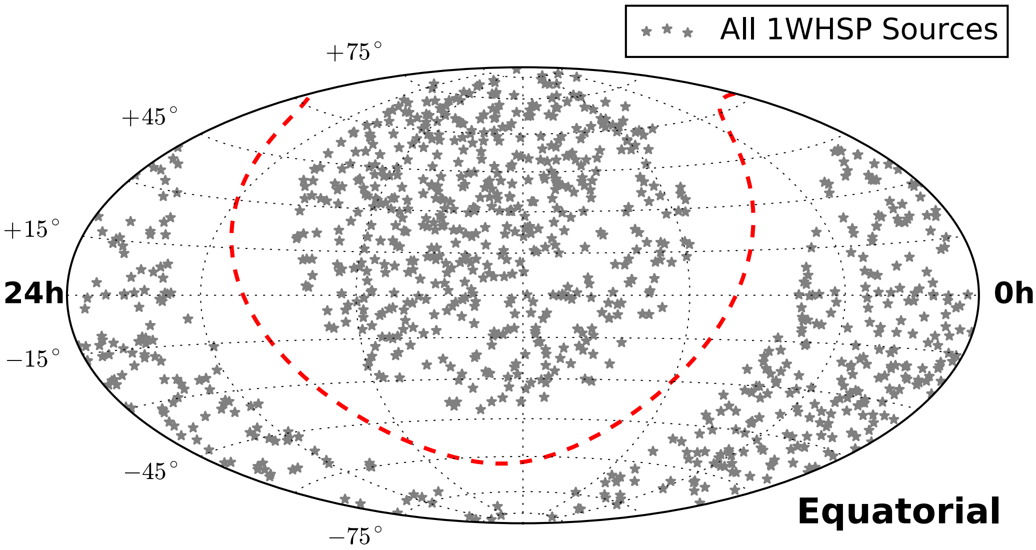


Figure 7.4: WHSP sources displayed on an equatorial skymap. The red dotted line illustrates the position of the Galactic Plane.

emitter, becomes clear if one regards the distributions of the synchrotron peak frequencies ν_{peak} and fluxes $(\nu f_\nu)_{peak}$ respectively for the WHSP candidates that are already detected in this energy band on the one hand and the ones that are not on the other hand (refer to figure 7.5). Since the ν_{peak} distribution of the 35 blazars, detected in the TeV band spans the entire range of the undetected distribution, in principle every blazar in the catalog has the chance to be detected in this high energy range [15].

From the $(\nu f_\nu)_{peak}$ distribution it is visible that so far the TeV detected sources are the brighter objects in the catalog, but nevertheless the peak flux of the undetected blazars is never smaller than about factor of ten than the flux of the faintest TeV detected source. Noting that for blazars a variability of about one order of magnitude is often observed in the X-ray and TeV band, yields the assumption that all of the WHSP blazar candidates may be detectable in the high energy regions by present γ-ray telescopes, whereas the FOM seems to be a reasonable value for the likeliness of that to happen [15].

Next to the FOM, the WHSP also contains information about the redshift z of each source in case it is available (see figure 7.3).

7.4 Differences to other Blazar Catalogs

In order to identify the importance of the 1WHSP catalog in different fields of high energy astrophysics, it is important to figure out the differences and innovations compared to other published source catalogs. Some of the most important samples of AGN-like objects are represented by the three catalogs published by the Fermi Large Area Telescope (Fermi-LAT).

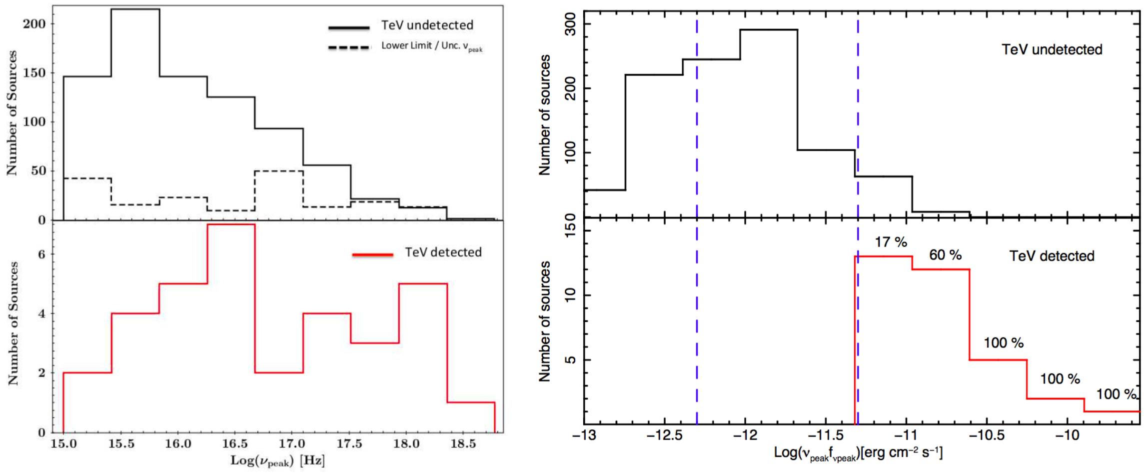


Figure 7.5: Distributions of the synchrotron peak frequencies and fluxes of respectively TeV undetected (upper plots) and detected (lower plots) blazars from the 1WHSP catalog. ***Left:*** Synchrotron peak frequency ν_{peak} distributions. ***Right:*** Distribution of the synchrotron peak fluxes $(\nu f_\nu)_{peak}$. Moreover for each $(\nu f_\nu)_{peak}$ bin the percentage of 1WHSP sources that are already TeV detected is noted. Both pictures were taken from [15].

Regarding only the first two Fermi-LAT catalogs, 1FGL[5] and 2FGL, it is visible that 189 objects from the 1WHSP catalog do have a counterpart in one of these catalogs. Moreover 107 further objects from the 1WHSP catalog can be identified with entries from the third Fermi-LAT catalog 3FGL, in the sense that their location is within a radius of 10 arc-minutes ($\approx 0.1667°$) of the position of one of the sources from the 3FGL data.

Consequently a total of 296 blazars from the 1WHSP catalog can be connected to an object of one of the Fermi-LAT catalogs, leaving a list of still 709 previously undetected γ-ray sources in the 1WHSP data set. Hence the 1WHSP catalog might be an important resource in order to investigate the properties of blazars, their jets and further characteristics of high energy γ-ray and neutrino physics [15, 11].

[5]FGL: Fermi Gamma-ray LAT

Chapter 8

Expected Performance of the WHSP Stacking Analysis

This chapter mainly describes the outcomes of the sensitivity studies of the point source stacking method, presented in chapter 6, using the blazars from the 1WHSP catalog as neutrino source hypothesis for seven years of data from the IceCube Neutrino Telescope. Before presenting the results, section 8.1 starts with a description of several tests, performed in order to verify the validity of the analysis method. Afterwards section 8.2 provides a short motivation for the usage of the 1WHSP catalog. In section 8.3 some characteristics of the final seven year point source event sample, which was generated from all IceCube data using the selection methods explained in chapter 5, are noted. Ultimately the sensitivities and the discovery potentials of the stacking analysis are presented in section 8.4.

8.1 Implementation of the Analysis Method

Before starting the performance of an analysis method on real data, it is necessary to verify the correctness of the applied technique step-by-step. Since the theoretical approach of the stacking analysis used in this thesis, is a commonly used and confirmed technique one only has to check the performance of the software executing this method. For this purpose, the main steps in the software, such as the minimization of the test statistic and the injection of simulated Monte Carlo events are tested. Finally the code was used to reproduce several results of former stacking analysis in IceCube, in order to cross-check its functionality.

8.1.1 Test Statistic Minimizer

As mentioned in section 6.1.1, the test statistic $\mathcal{TS}$ (refer to equation (6.1)) was introduced as a measure for the compatibility of the IceCube data with either the null hypothesis H_0, meaning that the data events are just a result from background, or the signal hypothesis H_S, assuming an additional astrophysical neutrino component. Since, this analysis is most interested in information, which are likely to correspond to the signal hypothesis, one is interested in the physical parameters giving the highest $\mathcal{TS}$ values (see equation (6.6)). Therefore in this analysis method, for every source position hypothesis, the test statistic formula presented in equation (6.6) is maximized with respect to the number of expected astrophysical neutrino events n_s and the spectral index γ. For computational simplification,

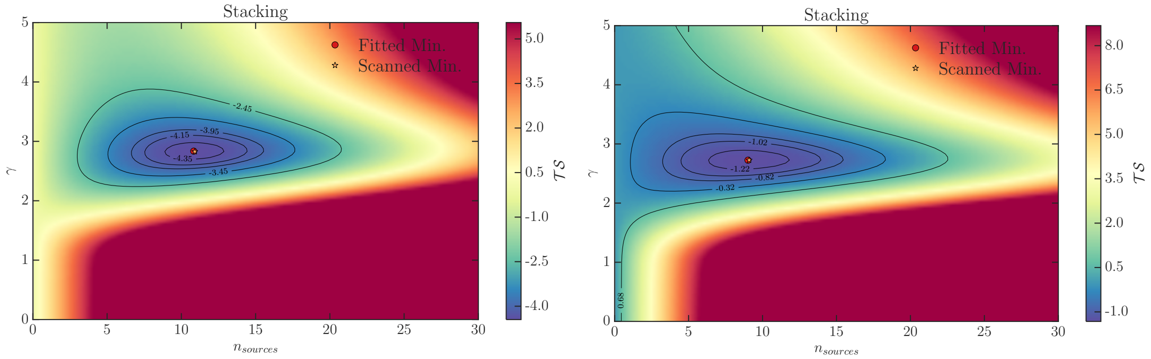


Figure 8.1: Both figures show test statistic (refer to equation (6.1)) values scanned along a certain fine grid in the n_s-γ plane. The yellow star displays the location of minimum evaluated by the *SciPy* minimizer. ***Left:*** The left picture assumes the 6 Milagro sources as the origin for the hypothetical neutrino signal. ***Right:*** In the right figure the neutrino signal is assumed to be a result of 233 Black Hole candidates.

this maximization process is executed as a minimization of the negative value of the $\mathcal{TS}$, using a *SciPy*[1]-minimizer, based on the L-BFGS algorithm [24].

Since this minimization is on of the central parts in the calculation of the sensitivity and the discovery potential (see section 6.3), it is crucial that this procedure performs as accurate as possible. In order to check the functionality of the minimizer, the test statistic is evaluated at every point in the n_s-γ space, whereas this scanned minimum is then compared to the result produced by the minimizer. A stacking point source search, as it is explained in section 6.2, assumes multiple sources in the sky as the origin of the neutrino signal. Since the calculations of the sensitivities and the discovery potentials, which are presented in the last section of this chapter, are based on different numbers of neutrino sources, it is reasonable to check the behavior of the minimization algorithm for different numbers as well. Some exemplary test statistic scans for 6 neutrino sources on the one hand and 233 on the other are illustrated in figure 8.1.

Looking at both figures, it is visible that the fitted minimum given by the minimizer nearly perfectly matches the result of the whole area scan. Comparing the exact values of both, the scan and the fit, one can see that the magnitude evaluated by the minimizer is even smaller. Hence the difference in both values can be traced back to the spacing of the scanned grid in the n_s-γ plane. The same result was observable in all other tests with different source hypothesis.

8.1.2 Injection of Monte Carlo Events

A second very important process in the code is the injection of neutrino signal events from Monte Carlo simulations, which is used to determine the sensitivity, the discovery potential and the upper limit flux (refer to subsection 6.3.2). As mentioned in subsection 6.3.2, it is

[1] SciPy: open source python based scientific library [39]

essential to insert signal neutrino events resulting from the position of one of the hypothetical source objects. Since in general, one does not want to simulate neutrino events for any possible source model weighting techniques are used to create a sample of injectable neutrino signal events. Since for a source hypothesis located at a certain declination one wants to inject events following a $E^{-\gamma}$ flux the idea of this method is to create a set of injectable neutrino events from the simulated signal events in a certain declination band around the source position. Afterwards the selected simulated events are rotated towards the exact source position and the actually injected events are randomly chosen from this sample according to their detection probability inside the detector. The exact process is explained in the following subsection.

In order to generate a set of signal events, resulting from a single source, all simulated Monte Carlo neutrino events $\mathbb{MC}$ in a certain declination band around the source position are rotated towards this point, in the sense that their true direction $\mathbf{x}_{i,true}$ is rotated exactly on the location of the source object $\mathbf{x}_s$:

$$\mathbf{x}_s^T \cdot M_R^i \cdot \mathbf{x}_{i,true} \stackrel{!}{=} 1 \qquad \forall i \in \mathbb{MC}, \tag{8.1}$$

where $M_R^i \in SO(3)^2$ corresponds to the rotation matrix, which is dependent on the specific position of each event i and can be determined according to [41]

$$M_R^i = \left\{ [1 - \cos(\alpha_i)]\, \hat{\mathbf{n}}_i \cdot \hat{\mathbf{n}}_i^T + \mathbb{1}\cos(\alpha_i) + \sin(\alpha_i) \left[\sum_{j=1}^{3} (\hat{\mathbf{n}}_i \times \mathbf{e}_j) \cdot \mathbf{e}_j^T \right] \right\}. \tag{8.2}$$

Since in this case one is only interested in the incoming direction of the events, all vectors $\mathbf{x}$ can be regarded as points on the three dimensional unit-sphere S^2, meaning that $\|\mathbf{x}\|_2 = 1$. In equation (8.2), the angle α_i corresponds to the angular difference between the true direction $\mathbf{x}_{i,true}$ of each individual event and the position of the source $\mathbf{x}_s$. It can be calculated as $\cos(\alpha) = \mathbf{x}_i \cdot \mathbf{x}_s$. Moreover since the vector $\hat{\mathbf{n}}$ describes the normalized rotation axis, it can be fixed as $\hat{\mathbf{n}}_i = \mathbf{x}_{i,true} \times \mathbf{x}_s$.

Besides, the reconstructed direction $\mathbf{x}_i$ of each signal event is rotated in an analogous way, using the rotation matrix created in the displacement of the true direction (equation (8.2)):

$$\mathbf{x}_{i,new} = M_R^i \cdot \mathbf{x}_i, \tag{8.3}$$

Altogether this procedure yields a set of events following a distribution of neutrino events reaching the IceCube detector and being generated at the location of the source. Moreover each event is equipped with a declination dependent weight w_i, corresponding to the detection efficiency of the detector (see section 6.2). In case of a stacking analysis, this process is performed for each individual source candidate. In the end, the events that are actually injected in the data sample, are randomly chosen from the set of the rotated Monte Carlo events, whereas the importance of each event is fixed according to a weight w_i coming from

[2] $SO(3)$: Special Orthogonal group in three dimensions. $M \in SO(3)$ illustrate orthogonal matrices with determinant $+1$.

its declination on the one hand and a weight resulting from the position of the emitting source.

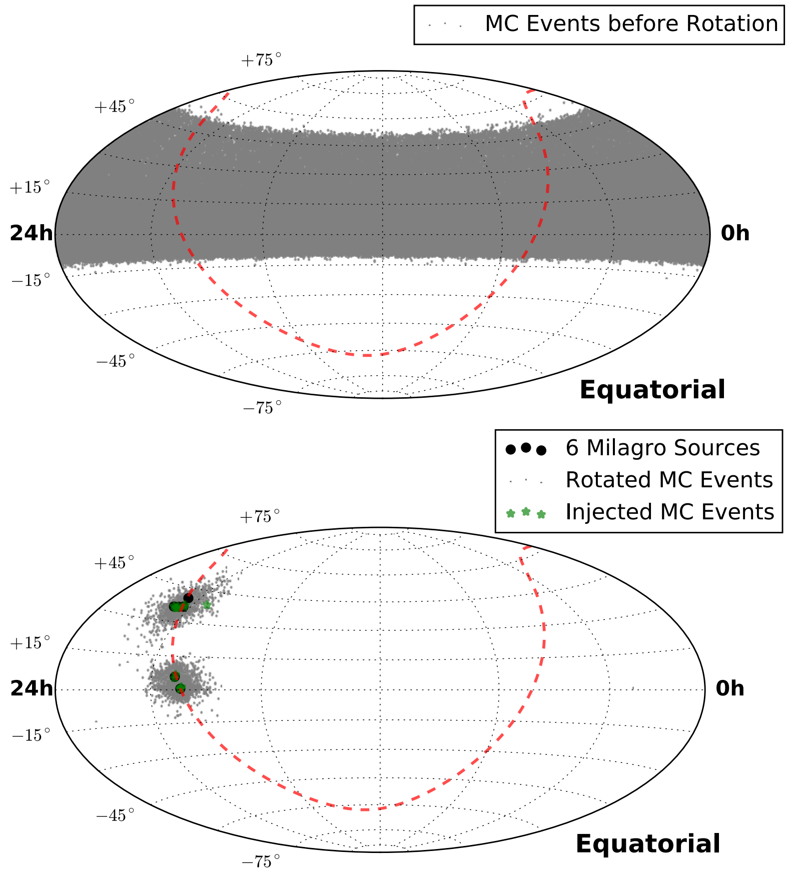


Figure 8.2: This figure describes the rotation and injection method of simulated neutrino events from selected sources. The red dotted line presents the location of the Galactic Plane. ***Top:*** The left picture shows the reconstructed position of all simulated MC events that are located at least in one declination band of one of the 6 Milagro sources. In case of these two figures each declination band contains an area of 20° around the source declination. ***Bottom:*** In the right figure the neutrino events are rotated towards the source positions. Moreover the green stars in this figure illustrate 15 events that would have actually been chosen for the injection to the data sample.

This whole rotation process is illustrated exemplary in figure 8.2, whereas in these plots the six Milagro sources are used as signal hypothesis for the IC40 (see section 8.3) IceCube dataset. Since the rotation M_R^i of the true direction of the events matches perfectly the source position and moreover the difference in the reconstruction error, meaning the angular distance between the true and the reconstructed position, before and after the rotation is equal except for machine precision errors. Consequently a sample of various source scenarios can be created using any feasible MC simulation.

IC40

Stacking	n_s	γ	p-value	Sensitivity [flux]	Disc. Potential [flux]
Milagro 17	9.0 (7.6)	2.73 (2.6)	0.278 (0.32)	6.80 (6.82)	22.4 (23.9)

IC40 + IC59 + IC79 + IC86

	n_s	γ	p-value	Upper Limit
Black Holes	16.8 (17.1)	3.12 (3.95)	0.44 (0.43)	$7.01 \cdot 10^{-12}$ ($6.88 \cdot 10^{-12}$)
Young SNRs	0 (0)	1.17 (-)	> 0.5 (> 0.5)	$4.81 \cdot 10^{-12}$ ($4.83 \cdot 10^{-12}$)
Young PWNs	0 (0)	2.15 (-)	> 0.5 (> 0.5)	$3.76 \cdot 10^{-12}$ ($3.12 \cdot 10^{-12}$)

Table 8.1: Listing of the results of previously completed stacking analyses calculated by the current analysis software. In parentheses the results evaluated the original analyses are noted.

8.1.3 Comparison to Former Analyses Results

After checking the functionality of mostly each individual step in the software, it seems reasonable to verify the correctness of the overall calculations, such as for instance the calculation of p-values and the sensitivities, by reproducing and comparing the results of former stacking analyses [10, 3]. For this purpose the p-value and the sensitivity, respectively the 90 % upper limit are calculated for four different different source hypothesis. First the analysis results of the "Milagro17" stacking search using the IC40 IceCube data set, and afterwards of the outcome of a stacking analysis based on four years of IceCube data, with 233 Black Hole Candidates, 30 Super Nova Remnants candidates and 10 Pulsar Wind Nebula candidates as hypothetical neutrino point sources respectively, were reproduced [3, 10]. The outcomes of these calculations are illustrated in table 8.1.

Comparing the calculated values to the results evaluated in the previous analyses, which are shown in parentheses in table 8.1, it is visible that both outcomes seem to agree quite well. The differences in some values might be explained by different settings, such as for instance in the binning of the energy in the calculation of the weights, in the particular code.

8.2 Motivation for the 1WHSP catalog

In the previous chapter the composition and the most important properties of the 1WHSP catalog were already presented. The catalog consists of 998 HSP blazar candidates (see subsection 7.1.2) and therefore currently constitutes the largest existing sample of HSP blazars. As mentioned in subsection 7.1.2 the SED of blazars mainly consists of two broad humps, whereas the low-energy one can be explained to be the result of synchrotron radiation from relativistic electrons in the jet. In contrast the origin of the high energy hump is still an open field of study, whereas different possible models exist that might explain this spectral appearance (see subsection 7.1.2.1). Since the acceleration of electrons up to relativistic energies inside the jets is a generally accepted concept and moreover essential to explain the low-energy synchrotron peak it seems reasonable to adopt the same concept for protons inside the jet. Once these protons exceed the threshold for pion production in interaction with ambient protons or photons, the high energy photon emission in blazars can be explained by a lepto-hadronic model. The existence of such a model would also imply the generation of high energy astrophysical neutrinos. In order the achieve the generation of the highest energy PeV neutrinos, as measured by the IceCube detector, protons accelerated up to energies of at least 10^{16} eV are required [50, 49].

8.2.1 High Synchrotron Peaked Blazars

Although the lepto-hadronic emission seems to be a promising scenario for the spectral emission in blazars, it is not a confirmed theory yet. Nevertheless, if one could discover a direct connection between blazar sources and high energy neutrinos it would be possible to confirm this scenario since the generation of neutrinos does not occur in the pure leptonic model. Moreover, since the lepto-hadronic acceleration in blazars is currently one of the few scenarios that could explain the appearance of the highest energy neutrinos, that were detected so far, blazars are probably one of the most promising sources of astrophysical neutrinos.

In the lepto-hadronic scenario, both photons and neutrinos are generated as secondary products in p-p and p-γ interactions. Consequently one expects the detection of astrophyiscal neutrinos in company with high energy photons. In case of the IceCube PeV neutrinos the energy of the demanded twin photons would be in the range of hundreds of TeV. Unfortunately the detection of photons in this energy range is quite hard, since the photon flux is attenuated due to pair production with ambient photons above 0.1 TeV (see subsection 7.1.2.2). Nevertheless the TeV photon emission in blazars is expected to be correlated to γ-flux in the GeV, meaning that powerful TeV emitters are strong GeV γ-ray sources as well. Previous observations have indicated that the HSP subclass of blazars are bright emitters in the GeV to TeV range and hence might be the dominant component of the extragalactic TeV background [50, 49, 15].

Taking all these arguments into account it seems more than reasonable that the 1WHSP catalog, being the largest existing sample of HSP blazars, might be used as a promising set of neutrino sources that can be utilized as signal hypothesis in a neutrino point source search.

8.2.2 Difference to Previous Blazars Stackings

Previous to the stacking analysis which is performed in this thesis, already some stacking searches (*IC86 Blazar Stacking* [54], *IC86-1 Blazar Population Analysis* [57]) using blazars as source hypothesis were performed in IceCube. These two stacking analyses were all based on information about blazars listed in the second Fermi AGN catalog (2LAC[3] [13]), whereas the first one uses only three different subsets of blazars, which were selected according to different characteristics of blazars, while the second analysis makes use of all blazars in this catalog. Hence at this point it seems reasonable to point out difference of these analyses, respectively the blazar catalogs compared to the one performed with the help of the 1WHSP catalog in this thesis. As already mentioned in section 7.4, 189 blazars from the 1WHSP

[3] 2LAC: AGN catalog based on the larger 2FGL catalog

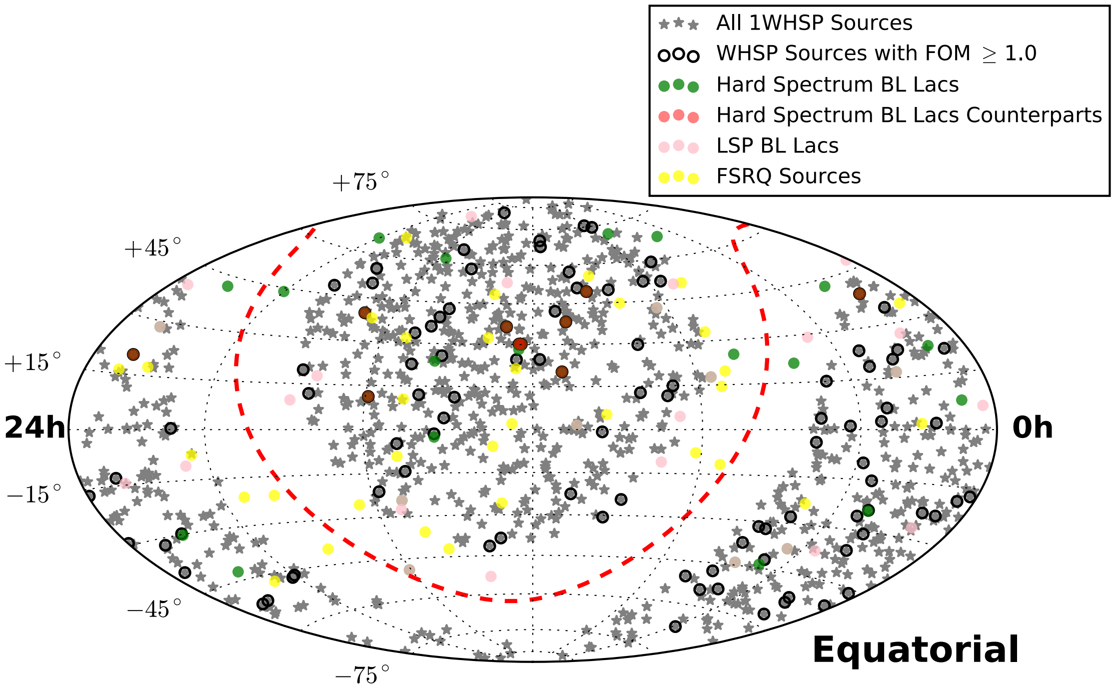


Figure 8.3: This figure shows the position of all WHSP blazars and the different blazar subsets from the second Fermi AGN catalog that were used in the *IC86 Blazar Stacking*. Since the FSRQ and the LSP BL Lac samples do not contain HSP blazars they do not have any counterpart in the 1WHSP catalog. Looking at the blazars from the Hard Spectrum BL Lacs it is visble that 10 of these overall 37 blazars do have a counterpart in the 1WHSP within a radius of 10′.

catalog do have a counterpart in the 2FGL catalog, leaving 809 new additional sources in the 1WHSP.

The correlation between the 1WHSP sources and the blazar sets used in the *IC86 Blazar Stacking* analysis are displayed in figure 8.3. From this sky-map it becomes clear that only 10 of the blazars used in this whole analysis do have a counterpart in the catalog of WHSP blazars.

Consequently the 1WHSP catalog definitely extends the list of previously used blazars in the search for neutrino point sources in IceCube and therefore opens a chance of discovering new physics.

8.3 Final Event Sample

For the stacking analysis performed in this thesis 7 years of IceCube data are available. The final sample that is used in order to calculate the p-values, the sensitivities and the discovery potentials is chosen from all measured IceCube events according to the selection method presented in chapter 5. Some important information about each subset of data that compose

Data Set	Livetime [d]	# Data Events	# MC Events
IC40	375.54	36900	621908
IC59	348.14	107011	1617012
IC79b	315.51	93133	1998573
IC86-I	332.61	136245	1242934
IC86-II,III,IV	1058.34	348786	6998483
MESE	1715.32	975	226625

Table 8.2: Information about the final event sample that is used in the WHSP Blazar Stacking analysis.

the final sample are listed in table 8.2. The name of each data subsample, respectively year in this table includes the number of strings that were installed inside the ice during the particular lifetime of the measurement. The supplement "b" in the IC79b dataset corresponds to the fact that for these events the improved SplineMPE muon track reconstruction was applied belated. Hence the resulting final sample of the seven years contains a total number of 722075 data events that were measured over a lifetime of $\sim 2430.13\,\mathrm{d}$.

In order to increase the chance of discovering statistical evidence for neutrino point sources, the final data set is completed by 975 MESE[4] events.

8.4 Sensitivity and Discovery Potential of the WHSP Stacking Analysis

In order to calculate the sensitivity and the discovery potential of a stacking analysis, the method specified in subsection 6.3.2 is operated on the final event sample. Therefore in the following section, first the partition of the 1WHSP catalog in different subsets, in order to find most likely set of neutrino point source candidates is explained. Ultimately the results of the sensitivity and discovery potential calculation of each of these blazar subsamples are presented in subsection 8.4.2.

8.4.1 Partitioning of the 1WHSP catalog

In section 8.2.1 the connection between the photon flux and the radiation of high energy energy neutrinos from blazars was emphasized. In lepto-hadronic emission scenarios the existence of very high energy neutrinos ($\sim$ PeV) is linked to the detection of a γ-ray flux in the TeV range. Since the 1WHSP catalog only consists of HSP blazars candidates, all of its sources are already very likely to be TeV photon emitters. However unfortunately only a few of these blazars in fact have been detected in this energy sector. Luckily the 1WHSP catalog provides a measure for the probability to be a potential TeV γ-ray emitter for each individual source. This value is called Figure Of Merit (see section 7.3 and equation (7.4))

[4]Medium Energy Starting Events

Subsample	FOM	# Blazars
1	≥ 1.00	103
2	≥ 0.50	207
3	≥ 0.25	376
4	≥ 0.10	737
5	≥ 0.00	998

Table 8.3: Partitioning of the 1WHSP catalog according to the FOM of each individual blazar.

FOM	# Sources	Sensitivity [flux]		Discovery Potential [flux]	
		Total	per Source	Total	per Source
≥ 1.00	103	2.70 (2.78)	0.026 (0.027)	7.27 (7.98)	0.071 (0.078)
≥ 0.50	207	2.78 (2.91)	0.013 (0.014)	9.36 (10.10)	0.045 (0.049)
≥ 0.25	376	3.31 (3.56)	0.008 (0.009)	11.64 (12.73)	0.031 (0.034)
≥ 0.10	737	4.57 (4.93)	0.006 (0.007)	15.98 (17.12)	0.022 (0.023)
≥ 0.00	998	5.11 (5.61)	0.005 (0.006)	16.40 (17.84)	0.016 (0.018)

Table 8.4: This table shows the sensitivities and the discovery potentials for each subset of the 1WHSP catalog. The results are based on the the final data sample generated from seven years of IceCube data, whereas in parentheses the results for the first six years are presented. All fluxes in this table are in given in units of $\left[E^2 d\phi/dE\right] = 10^{-12}\,\mathrm{TeV\,cm^{-2}\,s^{-1}}$.

and is defined as the ratio between the synchrotron peak flux of a blazar and the synchrotron peak flux of the faintest blazar that was already discovered in the TeV range.

Taking this into account it seems reasonable to split the blazars in multiple subsets according to their figure of merit in order to find the most promising set of neutrino generators. Since blazars with a high FOM are most likely to be also generators of high energy astrophysical neutrinos it does not make sense to observe subsets with only those blazars with a low FOM while ignoring the higher weighted sources. Consequently the partition of the 998 blazars of the 1WHSP catalog occurs in a cumulative manner, meaning that in each step the sample of high FOM blazars is extended by some blazars with a slightly lower FOM. The exact partition of the catalog is illustrated in table 8.3.

8.4.2 Sensitivity and Discovery Potential

In order to estimate the expected performance of the 1WHSP blazar stacking analysis on seven years of IceCube data, finally the sensitivity and the discovery potential are calculated for each of the selected subsamples listed in table 8.3. Therefore the simulated neutrino events that are injected to the scrambled data are assumed to be radiated from sources following an emission according to an E^{-2} energy power law (see subsection 6.3.2). The results of these

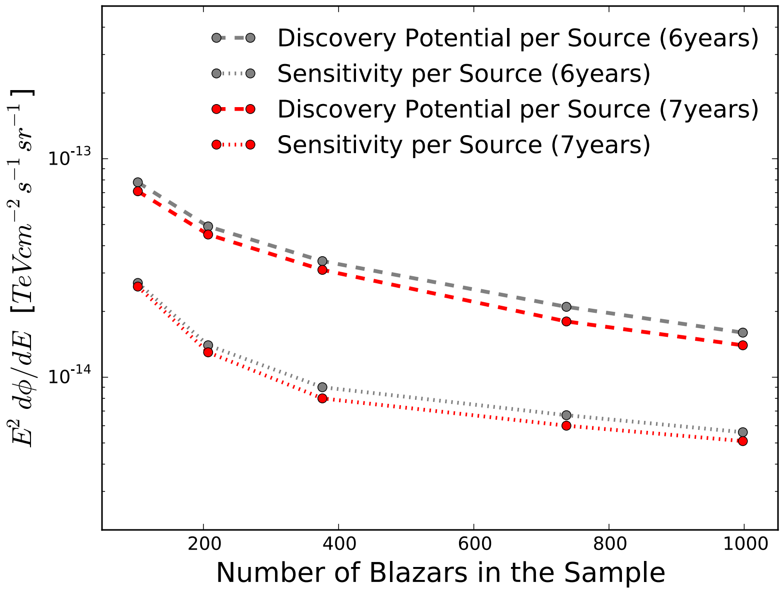


Figure 8.4: Illustration of the sensitivity and the discovery potential per source as a function of the number of sources included in the respective blazar subsample. The results are shown respectively for 6 years and 7 years of IceCube data plus MESE events.

estimations are listed in table 8.4, whereas next to results of the seven years of data also the previous results from the calculation for six years of IceCube data are noted.

While looking at these values, it is visible that all total sensitivity and the discovery potential fluxes increase as expected with the number of sources that are included in the particular subsample. Nevertheless the relative flux of each individual source that is necessary to fulfill the criteria of the sensitivity and the discovery potential definitions (see subsection 6.3.2) decreases with the total number of sources of the particular WHSP subsample. In addition this behavior is also illustrated in figure 8.4.

Moreover the differential sensitivity and discovery potential fluxes as a function of the energy are plotted in figure 8.5 for the first two blazar subsamples. These fluxes can be compared to neutrino background flux that is expected to be emitted from all existing blazars in the universe [50]. This comparison in particular as well as the interpretations of all further outcomes calculated in this section are discussed in chapter 9.

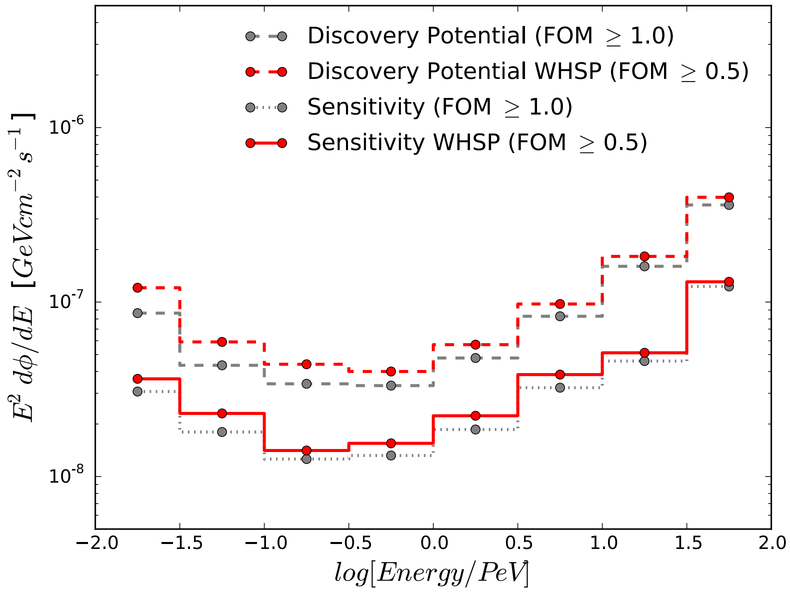


Figure 8.5: Representation of the differential sensitivity and discovery potential of different subsets of the 1WHSP catalog as a function of the neutrino energy. All fluxes illustrated in this figure are based on seven years of IceCube data.

Chapter 9

Conclusion and Outlook

So far in this thesis the sensitivities and the discovery potentials of several subsets of the 1WHSP catalog were calculated with the help of a stacking analysis approach (see section 6.2). In order to discuss the expected performances of this analysis, in this chapter my outcomes will be compared to the sensitivities and discovery potentials of the single all-sky point source search. Moreover, in section 9.2 my differential discovery potentials and sensitivities are compared to the predicted neutrino flux from all blazars [50]. Finally in section 9.3 an outlook is given, including the proximate procedure of this analysis as well as an improved version of the WHSP catalog.

9.1 Comparison to Single Point Source Analysis

In order to verify the additional chances that a stacking analysis can provide compared to the search for single point sources one can compare the sensitivity and the discovery potential of both analyses. The calculated values of the single point source search using six years of IceCube data are displayed in figure 9.1.

From this figure it is visible that the lowest sensitivity for the single point source search is reached for sources around the horizon, whereas the best sensitivity flux is approximately $4 \cdot 10^{-13}\,\mathrm{TeV\,cm^{-2}\,s^{-1}}$. Since the stacking analysis in this thesis was evaluated on five subsets of the 1WHSP blazar catalog (refer to subsection 8.4.1), all of these outcomes can be compared individually to the single search performance.

Looking at the measures evaluated for first subsample (listed in table 8.4), including only WHSP sources having a $FOM \geq 1$ for six years of IceCube data, it is observable that the required total flux of this subsample is approximately about a factor of 7 higher than the sensitivity flux at the optimal position of a single source. Since the total sensitivity flux of the stacking search increases once more sources are added to the observed blazar samples, this factor gets even higher for the other subsamples of the 1WHSP catalog. However once regarding the required flux that each individual source has to provide in the stacking search in order to fulfill the sensitivity criteria it is visible that for the first subset of the 1WHSP this value only amounts to approximately 6.7 % of the lowest flux value from the single point source search. Hence the required amount of neutrinos that each source has to provide is much less in this stacking approach. Since the flux per source decreases with the number of blazars in the observed subsample the difference to the single search increases even more

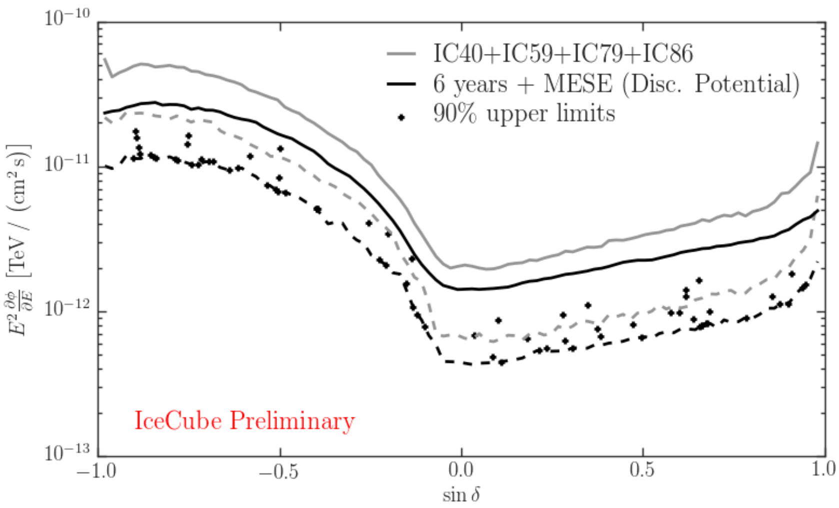


Figure 9.1: Sensitivity and discovery potential of the single Point Source search. The solid curves represent the discovery potential flux, while the dashed lines illustrate the sensitivity. The figure was taken from [62].

for the other subsamples. The sensitivity flux per source that was evaluated for the fifth subsample, meaning the whole 1WHSP blazar set as source hypothesis, amounts only 1.5 % of the best sensitivity from the single all-sky search.

Nevertheless this conclusion has to be handled with care, since a stacking analysis strongly depends on the input quality of neutrino sources, in the sense that the analysis can only be successful if the assumed sources are actually neutrino emitters. In the single point source search instead the whole sky can be scanned for the most significant source position. Hence it is not dependent on the quality of a certain neutrino source catalog.

In summary it can be recorded that the required sensitivity flux per source in the stacking approach, is approximately about a factor of 15 to 67 lower than in the single point source search, whereas the success of this analysis strongly depends on the predicted neutrino output of the selected blazars. Ultimately it can be noted that the exactly same behavior as for the sensitivity is also observable for the discovery potential.

9.2 Neutrino Background From Blazars

As mentioned in subsection 8.2.1 blazars have been suggested to be one of the most likely sources of high energy astrophysical neutrinos. Assuming proton acceleration to very high energies ($\sim 10^{16}$ eV), yields to the production of charged and neutral pions inside the jets. As a result of the pion decays a flux of high energy photons in the TeV range in addition to a flux of high energy neutrinos ($\sim$ PeV) is predicted. Since in this model the neutrino generation is directly connected to the emission rate of TeV photons, the expected neutrino flux can be theoretically estimated by Monte Carlo simulations based on information about the SED [50].

The theoretical modeling of the neutrino output from blazars was performed in [50], whereas the principle idea of the Monte Carlo simulation is shortly explained in the following. The Monte Carlo simulations in this paper characterize the entire γ-ray spectrum generated from BL Lacs, based on a number of specifications including for instance the distribution of the Doppler factor, a synchrotron model and an accretion disk component. In order to get information about the neutrino output an additional hadronic component based on the knowledge of the SED and the presumed neutrino spectra of a preselected sample of BL Lac blazars is added to this model. Based on these assumptions the expected observed neutrino flux per ν-flavour of a BL Lac object can be derived to be

$$E_\nu \cdot F_\nu(E_\nu) = \frac{1}{3} \cdot \frac{Y_{\nu\gamma} F_\gamma(> 10\,\mathrm{GeV})}{\int_{x_{min}}^{\infty} x^{-s} e^{-x} dx} \left(\frac{E_\nu}{E_{\nu,p}}\right)^{-s+1} \exp\left(-\frac{E_\nu}{E_{\nu,p}}\right), \tag{9.1}$$

where F_ν corresponds to the differential neutrino flux of all flavours and E_ν to the neutrino energy. The energy $E_{\nu,p}$ represents the peak energy of the neutrino spectrum which can be approximated from the values of the observed synchrotron peak frequency, the Doppler factor and the redshift of the source. Moreover the variable x is defined as $x \coloneqq E_\nu / E_{\nu,p}$, whereas x_{min} can be interpreted as the minimum normalized neutrino energy. The value s illustrates the power law index of the neutrino spectra which was shown to be approximately $\langle s \rangle \approx -0.35$ [50].

The variable $Y_{\nu\gamma}$ is defined as

$$Y_{\nu\gamma} \coloneqq \frac{F_\nu}{F_\gamma(> 10\,\mathrm{GeV})}, \tag{9.2}$$

where $F_{\nu,tot}$ corresponds to the total neutrino flux and $F_\gamma(> 10\,\mathrm{GeV})$ to the integrated photon flux above 10 GeV. Hence by construction $Y_{\nu\gamma}$ includes all the information on existence and strength of the putative photo-pion interactions inside the jets of a BL Lac blazar, meaning that very small values of $Y_{\nu\gamma}$ ($\ll 1$) would indicate the purely leptonic origin for the photon emission at γ-ray energies while values between 0.1-2 would allow the existence of semi-hadronic emission models in blazars. Finally the factor 1/3 in equation (9.1) arises due to the fact that the proportion of the three different neutrino flavours reaching the earth is expected to be uniform due to neutrino oscillations (see subsection 2.2.1) [50].

Looking again at equation (9.1) it is visible that except from the parameter $Y_{\nu\gamma}$ all values are determined from observations. Hence this variable can be used to test different assumptions on the model. In case of a semi-hadronic emission model from blazars this value is specified to $Y_{\nu\gamma} = 0.8$ as a benchmark. Finally the total neutrino background from all blazars can be estimated by integrating over these fluxes from all existing blazars [50].

The expected differential neutrino background flux under the assumption of a semi-hadronic emission model ($Y_{\nu\gamma} = 0.8$) is displayed in figure 9.2. In order to testify the expected performance of the 1WHSP stacking analysis the neutrino background flux can be compared

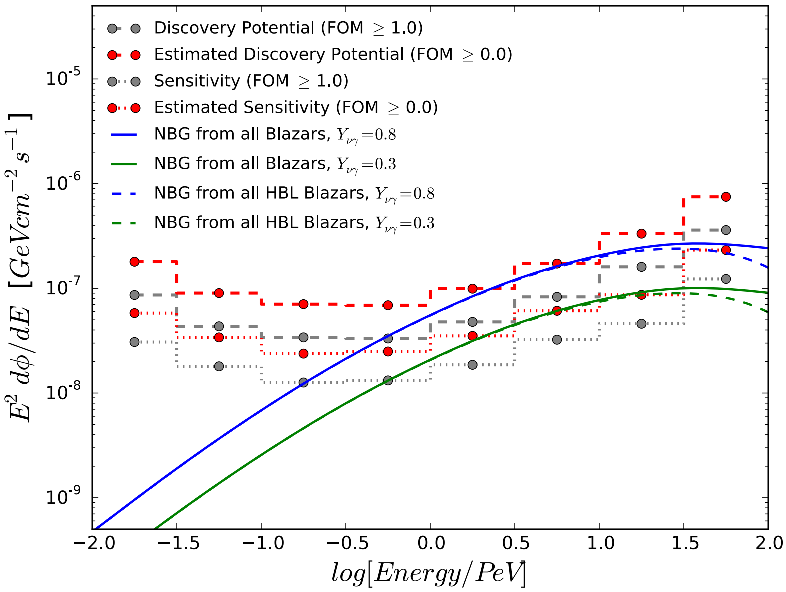


Figure 9.2: Representation of the differential sensitivity and discovery potential as a function of the neutrino energy. The fluxes for the blazar subset having a $FOM \geq 1.0$ is calculated directly. The differential sensitivity and discovery potential for the whole 1WHSP catalog is estimated by scaling these fluxes with the ratio of the of the particular total fluxes respectively for the sensitivity and the discovery potential. Moreover the expected neutrino background from all presumably existing blazars is displayed for $Y_{\nu\gamma} = 0.8$ and $Y_{\nu\gamma} = 0.3$ [50].

to the differential discovery potential and sensitivity flux. Both curves are also illustrated in figure 9.2.

Before trying to interpret this outcome, the meaning and characteristics of the discovery potential flux should be clarified first. In order to calculate the differential discovery potential in this stacking analysis the neutrino energy is split into different bins. For each individual bin separately simulated neutrino events having a true energy in the range of the particular bin are injected from an E^{-2} spectrum to the scrambled data sample until the criteria of the discovery potential are fulfilled (see subsection 6.3.2). From the required number of injected signal events N one can then calculate the discovery potential flux according to

$$E^2 \frac{d^3\phi}{dE d\Omega dt} = \frac{N}{\tau \int_\Omega \int_{E_{min}}^{E_{max}} E^{-2} A_{eff}(E,\Omega) dE d\Omega}, \tag{9.3}$$

where A_{eff} corresponds to the effective area which represents an energy and zenith dependent measure for the neutrino detection efficiency of the IceCube detector. In case of the differential flux E_{min} and E_{max} illustrate the lower and the upper bound of the energy bin. According to these observations the differential discovery potential flux in each individual bin can be interpreted as the neutrino flux that is required in this energy range in order to fulfill the discovery potential criteria without the necessity of the existence of any neutrino flux outside the bin range. Nevertheless this differential flux has to be handled with care since it is strongly dependent on the size of the binning. Assuming that the number of signal events N is nearly independent of the bin size it is visible from equation (9.3) that the flux $E^2 d\phi/dE$ decreases once the bin size increases.

Now that the meaning of the calculated differential flux is clarified it can be compared to the expected neutrino background flux from all blazars. As visible in figure 9.2 the discovery potential for the blazar subset of the 1WHSP catalog, including only sources with a $FOM \geq 1$ is below the neutrino background flux above an neutrino energy of approximately 560 TeV. This observation leads to promising presumption that the WHSP stacking analysis might be able to discover a statistically significant indication for astrophysical point sources under the assumption that the theoretically estimated blazar model describes the actual situation.

Nevertheless it must be considered that the neutrino background flux is calculated as the combined output of all presumable existing blazars. Since the 1WHSP catalog is currently the largest existing catalog of HSP blazars it can be approached as the entity of all HSP blazars. However the comparison of the neutrino background flux to the discovery potential of the first subset of the 1WHSP, only including 103 sources, seems only reasonable if this subsets contributes the bulk of the neutrino emission of all 1WHSP blazars. Hence it might be more appropriate to regard the differential discovery potential of the whole 1WHSP catalog. Since the calculation of this flux was not finished until the end of this thesis, it can only estimated here by scaling the differential flux of the first subset of the 1WHSP catalog with the ratio of the particular total fluxes (refer to table 8.4). This result is also illustrated in figure 9.2. Since this flux is still in the range of the expected neutrino background flux also in this scenario a neutrino point source discovery might be possible.

Ultimately it can be noted that the detection of neutrinos from blazars would also be a confirmation of the semi-hadronic emission model, since no high energy neutrinos are produced in purely leptonic models.

9.3 Outlook

Until now in this thesis we examined the stacking analysis with the 1WHSP catalog on blinded IceCube data, meaning that the right ascension value of all events were randomly chosen from the range between 0 and 2π. The sensitivity and the discovery potential of different subsets of the catalog were determined and ultimately compared to the expected neutrino background from all blazars, assuming proton acceleration to at least 10^{16} eV inside their jets. As mentioned to the previous section this modeled neutrino flux, provided that it is correct, predicts the discovery of a significant astrophysical neutrino signal resulting from the WHSP blazars.

Nevertheless so far only predictions about the expected performance of this stacking could have been made. Consequently the next step will be the inspection of the analysis on the real unblinded IceCube data. Not before the evaluation of the stacking on the real data it is possible to establish any conclusions.

Moreover even in case of not detecting any significant neutrino signal from the 1WHSP catalog, there are still several scenarios how to go on. Briefly a new, second version of the WHSP catalog including even more blazar, especially in the region around the galactic plane will be submitted. Furthermore besides the WHSP blazar samples, another catalog namely the 2FHL[1] exists that can also be tested for the seven years of IceCube data [12].

In addition also a theoretical weighting scheme of all sources in the catalog according to physical properties of the individual blazars, as mentioned in section 6.2 might be a conceivable facility in order to improve the behavior of the analysis. Assuming that the neutrino output of the individual WHSP blazars varies from source to source, the performance of a stacking analysis could be enhanced by providing each blazar k with a relative theoretical weight W^k according to the amount of its neutrino emission that we expect to see. Since the radiation of neutrinos in blazars is directly connected to the high energy TeV photon emission (refer to subsection 7.1.2.1) a plausible weighting scheme could for instance be realized by the use of the integrated γ-ray flux of the blazars. Hence high values of W^k, respectively the γ-ray flux, preferentially weight some sources over others, yielding a more realistic physical description of the neutrino source hypotheses that are used in the stacking analyses. Including these weights in the calculation of the stacking signal likelihood function (see equation (6.13)) might improve the performance of the analyses.

In summary one can say that the work in this thesis might represent a realistic chance of discovering neutrino point sources and consequently also confirm the correctness of a lepto-hadronic emission model in blazars, but for sure provides at least a promising starting point for further stacking searches.

[1] 2FHL: second catalog of hard Fermi Large Area Telescope (Fermi-LAT) sources

Bibliography

[1] M. G. Aartsen et al.
'Energy Reconstruction Methods in the IceCube Neutrino Telescope'.
In: *JINST* 9 (2014), P03009. arXiv: `1311.4767 [physics.ins-det]` (cit. on p. 39).

[2] M. G. Aartsen et al.
'Improvement in Fast Particle Track Reconstruction with Robust Statistics'.
In: *Nucl. Instrum. Meth.* A736 (2014), pp. 143–149. arXiv: `1308.5501 [astro-ph.IM]` (cit. on pp. 30 sq.).

[3] M. G. Aartsen et al. 'Searches for Extended and Point-like Neutrino Sources with Four Years of IceCube Data'. In: *Astrophys. J.* 796.2 (2014), p. 109. arXiv: `1406.6757 [astro-ph.HE]` (cit. on pp. 51, 71).

[4] M. Aartsen et al.
'IceCube-Gen2: A Vision for the Future of Neutrino Astronomy in Antartica'.
In: *arxiv e-Prints* (2014). eprint: `astro-ph/1412.5106v2` (cit. on pp. 26 sq.).

[5] R. Abbasi et al. 'An improved method for measuring muon energy using the truncated mean of dE/dx'. In: *Nucl. Instrum. Meth.* A703 (2013), pp. 190–198. arXiv: `1208.3430 [physics.data-an]` (cit. on pp. 23 sq.).

[6] R. Abbasi et al.
'Calibration and Characterization of the IceCube Photomultiplier Tube'.
In: *Nucl. Instrum. Meth.* A618 (2010), pp. 139–152. arXiv: `1002.2442 [astro-ph.IM]` (cit. on p. 19).

[7] R. Abbasi et al. 'IceTop: The surface component of IceCube'.
In: *Nucl. Instrum. Meth.* A700 (2013), pp. 188–220. arXiv: `1207.6326 [astro-ph.IM]` (cit. on p. 18).

[8] R. Abbasi et al. 'The Design and Performance of IceCube DeepCore'.
In: *Astropart. Phys.* 35 (2012), pp. 615–624. arXiv: `1109.6096 [astro-ph.IM]` (cit. on p. 18).

[9] R. Abbasi et al. 'The IceCube Data Acquisition System: Signal Capture, Digitization, and Timestamping'. In: *Nucl. Instrum. Meth.* A601 (2009), pp. 294–316. arXiv: `0810.4930 [physics.ins-det]` (cit. on p. 25).

[10] R. Abbasi et al. 'Time-Integrated Searches for Point-like Sources of Neutrinos with the 40-String IceCube Detector'. In: *Astrophys. J.* 732 (2011), p. 18. arXiv: `1012.2137 [astro-ph.HE]` (cit. on pp. 42, 71).

[11] F. Acero et al. 'Fermi Large Area Telescope Third Source Catalog'.
In: *arxiv e-Prints* (2015). arXiv: `1501.02003 [astro-ph.HE]` (cit. on p. 65).

[12] M. Ackermann et al. '2FHL: The Second Catalog of Hard Fermi-LAT Sources'.
In: *arxiv e-Prints* (2015). arXiv: `1508.04449 [astro-ph.HE]` (cit. on p. 84).

[13] M. Ackermann et al. 'The Second Catalog of Active Galactic Nuclei Detected by the Fermi Large Area Telescope'. In: *Astrophys. J.* 743 (2011), p. 171. arXiv: `1108.1420 [astro-ph.HE]` (cit. on p. 72).

[14] J. Ahrens et al. 'Muon track reconstruction and data selection techniques in AMANDA'. In: *Nucl. Instrum. Meth.* A524 (2004), pp. 169–194. arXiv: `astro-ph/0407044 [astro-ph]` (cit. on pp. 32 sq.).

[15] B. Arsioli et al. '1WHSP: an IR-based sample of naerly 1000 VHE gamma-ray blazar candidates'. In: *arxiv e-Prints* (2014) (cit. on pp. 61–65, 72).

[16] C. Autermann. 'Boosted Decision Trees'. In: (2007) (cit. on p. 44).

[17] J. K. Becker. 'High-energy neutrinos in the context of multimessenger astrophysics'. In: *arxiv e-Prints* (2008). eprint: `astro-ph/0710.1557v2` (cit. on pp. 13 sq.).

[18] M. Boettcher. 'Modeling the Emission Processes in Blazars'. In: *Astrophys. Space Sci.* 309 (2007), pp. 95–104. arXiv: `astro-ph/0608713 [astro-ph]` (cit. on p. 61).

[19] J. R. Braun. 'A Maximum-Likelihood Search for Neutrino Point Sources with the AMANDA-II Detector'. PhD thesis. University of Wisconsin–Madison, 2009 (cit. on p. 47).

[20] J. Braun et al. 'Methods for point source analysis in high energy neutrino telescopes'. In: *Astropart. Phys.* 29 (2008), pp. 299–305. arXiv: `0801.1604 [astro-ph]` (cit. on p. 49).

[21] P. A. Cherenkov. 'Visible emission of clean liquids by action of ? radiation'. In: (1934) (cit. on p. 20).

[22] S. Coenders. *IC86-II and IC86-III Point Source Full Sky Scan.* http://icecube.wisc.edu/ coenders/. Accessed: 2015-11-04. 2015 (cit. on pp. 41, 43 sqq.).

[23] S. Coenders. *IceCube: DOMs, DAQ, PnF, etc.* https://ecp.wiki.tum.de/Bootcamp. Accessed: 2015-11-04. 2014 (cit. on p. 25).

[24] S. Coenders. *Skylab.* `https://github.com/coenders/skylab/trunk/skylab`. 2015 (cit. on p. 68).

[25] S. Coenders. 'Capability of Earth Tomography Using Oscillations of Atmospheric Neutrinos with PINGU'. MA thesis. RWTH Aachen, 2013 (cit. on pp. 20 sq., 23 sq.).

[26] Fahrmeir et al. *Statistik - Der Weg zur Datenanalyse.* 2010 (cit. on p. 53).

[27] A. Fedynitch et al. 'Calculation of conventional and prompt lepton fluxes at very high energy'. In: *EPJ Web Conf.* 99 (2015), p. 08001. arXiv: `1503.00544 [hep-ph]` (cit. on p. 12).

[28] J. Feintzeig. 'Searches for Point-like Sources of Astrophysical Neutrinos with the IceCube Neutrino Observatory'. PhD thesis. University of Wisconsin–Madison, Aug. 2014 (cit. on pp. 8, 13, 29, 38, 44).

[29] J. A. Formaggio and G. P. Zeller. 'From eV to EeV: Neutrino Cross Sections Across Energy Scales'. In: *Rev. Mod. Phys.* 84 (2012), p. 1307. arXiv: `1305.7513 [hep-ex]` (cit. on pp. 14 sq.).

[30] T. Gaisser and M. Honda. 'Flux of Atmospheric Neutrinos'. In: *arxiv e-Prints* (2002). eprint: `astro-ph/0203272v2` (cit. on p. 11).

[31] P. Giommi et al. 'The sedentary multifrequency survey: Statistical identification and cosmological properties of high-energy peaked BL Lacs'. In: *arxiv e-Prints* (1999) (cit. on p. 63).

[32] P. Giommi et al. 'A simplified view of blazars: clearing the fog around long-standing selection effects'. In: *Mon. Not. Roy. Astron. Soc.* 420 (2012), p. 2899. arXiv: `1110.4706 [astro-ph.CO]` (cit. on pp. 59 sq.).

[33] D. Gora et al. 'Studies on Millipede'. In: (2014) (cit. on p. 39).

[34] J. W. Hewitt and M. Lemoine-Goumard. 'Observations of supernova remnants and pulsar wind nebulae at gamma-ray energies'. In: *Comptes Rendus Physique* 16 (2015), pp. 674–685. arXiv: `1510.01213 [astro-ph.HE]` (cit. on p. 8).

[35] A. M. Hillas. 'The Origin of Ultrahigh-Energy Cosmic Rays'. In: *Ann. Rev. Astron. Astrophys.* 22 (1984), pp. 425–444 (cit. on p. 7).

[36] H. Hu. 'Status of the EAS studies of cosmic rays with energy below 10**16 eV'. In: *arxiv e-Prints* (2009). arXiv: `0911.3034 [astro-ph.HE]` (cit. on p. 5).

[37] J. van Santen. 'Markov-chain monte-carlo reconstruction for cascade-like events in icecube'. MA thesis. Humboldt-Universitaet Berlin, 2010 (cit. on p. 20).

[38] J. van Santen. 'Surface zenith angles and overburden in IceCube'. In: (2014) (cit. on p. 28).

[39] E. Jones, T. Oliphant, P. Peterson et al. *SciPy: Open source scientific tools for Python.* [Online; accessed 2015-08-02]. 2001– (cit. on p. 68).

[40] S. R. Kelner, F. A. Aharonian and V. V. Bugayov. 'Energy spectra of gamma-rays, electrons and neutrinos produced at proton-proton interactions in the very high energy regime'. In: *Phys. Rev.* D74 (2006). [Erratum: Phys. Rev.D79,039901(2009)], p. 034018. arXiv: `astro-ph/0606058 [astro-ph]` (cit. on p. 61).

[41] M. Koecher. *Lineare Algebra und analytische Geometrie.* 4th ed. Springer Verlag, 2013 (cit. on p. 69).

[42] K. Koyama et al. 'Evidence for shock acceleration of high-energy electrons in the supernova remnant SN1006'. In: *Nature* 378 (1995), pp. 255–258 (cit. on p. 8).

[43] P. M. Lee. *Bayesian Statistics: An Introduction.* 4th ed. 2012 (cit. on p. 52).

[44] Los Alamos Science. *Celebrating the neutrino.* 25th ed. 1997 (cit. on p. 10).

[45] A. Maselli et al. 'VizieR Online Data Catalog: Blazars in the Swift-BAT hard X-ray sky'. In: *arxiv e-Prints* (2010) (cit. on p. 63).

[46] R. Nahnhauer et al. 'Energy reconstruction of muon neutrino events in DeepCore'. In: *arxiv e-Prints* (2013) (cit. on p. 39).

[47] T. Neunhöffer. 'Entwicklung eines neuen Verfahrens zur Suche nach kosmischen Neutrinopunktquellen mit dem AMANDA Neutrinoteleskop'. PhD thesis. Johannes Gutenberg University Mainz, Nov. 1999 (cit. on p. 34).

[48] R. Oerter. *The Theory of Almost Everything: The Standard Model, the Unsung Triumph of Modern Physics*. 2006 (cit. on p. 3).
[49] P. Padovani and E. Resconi. 'Are both BL Lacs and pulsar wind nebulae the astrophysical counterparts of IceCube neutrino events?' In: *Mon. Not. Roy. Astron. Soc.* 443.1 (2014), pp. 474–484. arXiv: `1406.0376 [astro-ph.HE]` (cit. on pp. 61 sq., 72).
[50] P. Padovani et al. 'A simplified view of blazars: the neutrino background'. In: *Mon. Not. Roy. Astron. Soc.* 452.2 (2015), pp. 1877–1887. arXiv: `1506.09135 [astro-ph.HE]` (cit. on pp. 72, 76, 79, 81 sq.).
[51] Particle Data Group. *Particle Physics Booklet*. 2012 (cit. on pp. 10 sq., 23 sq.).
[52] D. Perkins. *Particle Astrophysics*. 2nd ed. Oxford University Press, 2009 (cit. on pp. 4, 6, 8 sqq., 12, 21 sqq.).
[53] S. P. Reynolds. 'Supernova Remnants at High Energy'. In: *arxiv e-Prints* (2008) (cit. on p. 8).
[54] K. Schatto. 'Stacked searches for high-energy neutrinos from blazars with IceCube'. PhD thesis. Johannes Gutenberg-Universit"at Mainz, June 2014 (cit. on pp. 47, 52, 58, 61, 72).
[55] M. Sikora et al. 'Constraining Emission Models of Luminous Blazar Sources'. In: *Astrophys. J.* 704 (2009), pp. 38–50. arXiv: `0904.1414 [astro-ph.CO]` (cit. on p. 61).
[56] T. Stanev. *High Energy Cosmic Rays*. 2nd ed. 2009 (cit. on p. 3).
[57] The IceCube Collaboration. *Blazar Population Analysis IC86-1*. https://wiki.icecube.wisc.edu. Accessed: 2015-11-20. 2014 (cit. on p. 72).
[58] The IceCube Collaboration. *IceCube Hardware and Data Processing*. https://events.icecube.wisc.edu. Accessed: 2015-11-05. 2015 (cit. on pp. 25 sq.).
[59] The IceCube Collaboration. *IceCube Media Gallery*. http://icecube.wisc.edu/gallery. Accessed: 2015-11-04 (cit. on pp. 18 sq., 29).
[60] The IceCube Collaboration. *IceCube Research Highlights*. http://icecube.wisc.edu/highlights. Accessed: 2015-11-05 (cit. on pp. 18, 26 sq., 50).
[61] The IceCube Collaboration. *Introduction to the IceCube Hardware*. https://wiki.icecube.wisc.edu. Accessed: 2015-11-04. 2011 (cit. on p. 19).
[62] The IceCube Collaboration. *Plots for public release*. https://wiki.icecube.wisc.edu. Accessed: 2015-11-24. 2015 (cit. on p. 80).
[63] C. M. Urry. 'Bl lac objects and blazars: past, present, and future'. In: *ASP Conf. Ser.* 159 (1999), p. 3. arXiv: `astro-ph/9812420 [astro-ph]` (cit. on p. 59).
[64] C. M. Urry and P. Padovani. 'Unified schemes for radio-loud active galactic nuclei'. In: *Publ. Astron. Soc. Pac.* 107 (1995), p. 803. arXiv: `astro-ph/9506063 [astro-ph]` (cit. on pp. 9, 58 sq.).
[65] M. Wallraff. 'Design, Implementation and Test of a New Feature Extractor for the IceCube Neutrino Observatory'. MA thesis. RWTH Aachen, 2010 (cit. on p. 21).
[66] E. L. Wright et al. 'The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance'. In: *Astron. J.* 140 (2010), p. 1868. arXiv: `1008.0031 [astro-ph.IM]` (cit. on pp. 2, 57).

Acknowledgements

In the end, I want to thank everyone who supported me during the production of my thesis over the last year.

Foremost, I would like to express my sincere gratitude to Professor Elisa Resconi for giving me the opportunity to be a part of her likable working group at the TUM and in particular for giving me the chance to write my thesis about such an interesting, important and promising topic. Moreover I would like to thank her for the opportunity to participate in several conferences and work with scientists from all over the world.

Furthermore I would like to particularly thank my supervisor Stefan Coenders for his patience and and endurance in answering all my questions during the last months.

Besides I also want to thank all the other members of the IceCube group in Munich. It was a real pleasure to work and do coffee breaks with each of you.

Finally I would like to particularly express my thanks to Stefan Coenders, Elisa Resconi, Andrea Turcati and Joost Veenkamp for proof-reading my thesis.

Declaration

I, Matthias Huber, hereby certify that this document has been composed only by myself, and describes my own work, unless otherwise acknowledged in the text.

Munich, 05.12.2015