Search for High Energetic Neutrinos from Core Collapse Supernovae using the IceCube Neutrino Telescope

D i s s e r t a t i o n

zur Erlangung des akademischen Grades

d o c t o r r e r u m n a t u r a l i u m

(Dr. rer. nat.)

im Fach Physik

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

M.Sc. Alexander Johannes Stasik

Präsidentin der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät:
Prof. Dr. Elmar Kulke

Gutachter/innen:
1. Marek Kowalski
2. Anna Franckowiak
3. Michael Kachelrieß

Tag der mündlichen Prüfung: 14.12.2017

Abstract

The recent discovery of a high energy flux of astrophysical neutrinos was one of the breakthroughs of the last years. However, the origin of these neutrinos remains still unknown. Also, the search for the sources of high-energy cosmic rays is closely connected to neutrinos since neutrinos are produced in hadronic interactions, and thus the detection of a neutrino source would be a *smoking gun* signature for cosmic rays. Many potential neutrino source classes have been discussed, among these are core-collapse supernovae.

In this thesis, seven years of data from the IceCube neutrino observatory are tested for correlation with the direction of hundreds of core-collapse supernovae. The analysis benefits from the good angular reconstruction of the order of one degree and below of the about 700000 muon track events and an extensive database of optical observations of supernovae. Using a time-dependent likelihood method, the sensitivity of the analysis is increased by stacking the sources in a combined analysis.

No significant clustering of neutrino events around the position of core-collapse supernovae is found. Upper limits of different neutrino light curve models are computed, and the contribution of core-collapse supernovae to the measured diffuse high energetic neutrino background is constrained. These limits allow excluding certain types of core-collapse supernovae as the dominant source of the observed high energetic astrophysical neutrino flux.

Zusammenfassung

Die Entdeckung eines hochenergetischen Flusses astrophysikalischer Neutrinos stellt einen wesentlichen physikalischen Durchbruch der letzten Jahre dar. Trotz allem ist der Ursprung dieser Neutrinos immer noch unbekannt. Die Suche nach den Quellen der hochenergetischen kosmischen Strahlung ist direkt verbunden mit der Suche nach Neutrinos, da diese in den gleichen hadronischen Prozessen erzeugt werden und eine Neutrinoquelle deshalb einen direkten Hinweis auf eine Quelle der kosmischen Strahlung darstellen würde. Viele potentielle Quellen der Neutrinos werden diskutiert, darunter Kern-Kollaps Supernovae.

In dieser Arbeit werden sieben Jahre Daten des IceCube Neutrinoteleskopes mit der Richtung mehreren Hundert Kernkollaps-Supernovae auf Korrelation getestet. Die Analyse gewinnt dabei durch die gute Richtungsrekonstruktion der 700000 Muonspurdaten und der großen Datenbank optische beobachteter Supernovae. Die Sensitivität der zeitabhängigen Likelihood-Analyse wird durch die Kombination mehrere Quellen in einer einzigen Analyse gesteigert.

Es wurde kein statistisch signifikantes Cluster von Neutrinos an den Positionen der Supernovae gefunden. Daraus wurden obere Grenzen für verschiedene Modelle berechnet und der Beitrag von Kernkollaps-Supernovae zum diffusen Neutrinofluss eingeschränkt. Daraus können bestimmte Typen von Supernovae als dominate Quelle der diffusen hochenergetischen astrophysikalischen Neutrinos ausgeschlossen werden.

Contents

1. Introduction

The night sky has fascinated people ever since. Astronomical records of observations date back up to 10000 years in the past [1]. For most of this period, observations were made using the naked eye. This changed in 1608 when the first telescopes were built and used first by Galileo [2]. This development and subsequent technical innovations have boosted the possibilities of observational astronomy, and this trend will likely continue in the future. Along with observational improvements, developments in physics allowed a better understanding of observed phenomena. Starting from Kepler's Laws and continuing to present day modeling of supernova explosions based on particle and relativistic plasma physics and gravitational wave physics as well as observation of expanding universe as consequence of general relativity.

Classical astronomy is performed in the optical part of the electromagnetic spectrum. In the last century, due to technical innovations, the spectrum has been extended from optical to include the radio, infra-red, ultra-violet and gamma-ray range, starting the era of multi-wavelength astronomy. This new data revealed new features of already known objects, entirely new objects, but also challenged the previous understanding and interpretation of these new observations.

The latest step in this evolution is the beginning of multi-messenger astronomy. While multi-wavelength astronomy is limited to photons, multi-messenger astronomy utilizes all accessible messenger particles, including charged particles or cosmic rays [3], gravitational waves [4] and neutrinos [5]. All these messengers have different characteristics, are produced in different processes, suffer from different absorption processes and are also detected by different techniques, as discussed in [6]. This also promoted the use of the term astroparticle physics since multi-messenger astronomy utilizes detection techniques and data analysis methods from particle physics.

Among these messengers, neutrinos are the ideal particles for astronomy. They are electrically neutral, so are not deflected by magnetic fields. They, therefore, travel on straight lines, pointing back to their origin. Neutrinos are also only weakly inter-

acting, allowing them to pass through gas clouds and other astrophysical obstacles without being absorbed or scattered. One challenge of neutrino astronomy is that they are hard to detect. The expected neutrino signal from a potential source is thus generally very weak.

A flux of neutrinos with astrophysical origin was discovered by the IceCube collaboration [7]. The origin of this diffuse flux remains a mystery. Many scenarios and potential source classes have been proposed, but no correlation has yet been found. This thesis aims to search for neutrinos from supernovae, the spectacular explosion at the end of a stars life, and to test the hypothesis that neutrinos from supernovae can explain the observed diffuse astrophysical neutrino flux. Highly energetic neutrinos are also a smoking gun signature of the source of cosmic rays, an open question since their discovery in 1912 by Victor Hess [3]. The detection of a neutrino source should directly point to a cosmic ray source as well.

In this thesis, a catalog of supernovae, discovered through optical detection is tested for correlation with neutrinos measured with the IceCube neutrino telescope [8]. The analysis benefits from the excellent spatial localization of the supernovae due to optical observations. The analysis utilizes time information as well as a technique called stacking, the simultaneous analysis of many potential sources in one single analysis to increase the total signal strength in the analysis.

This thesis is organized as follows: Chapter 2 introduces cosmic rays, their production and the generation of neutrinos. Chapter 3 describes stars and supernovae and introduces models for neutrino production in supernova explosions. Chapter 4 introduces the techniques for neutrino detection and the IceCube neutrino detector. A different approach for detecting neutrino sources in real time is discussed in chapter 5. Chapter 6 discusses the supernova catalog used in this thesis and how it was compiled. In chapter 7 the derivation of a diffuse astrophysical flux from a source model and cosmological assumptions are presented. Chapters 8 and 9 discuss the likelihood method used in this thesis. In chapter 10 the implementation and unblinding of the analysis are discussed, and chapter 11 presents the interpretation of the results. Chapter 12 gives a summary and an outlook.

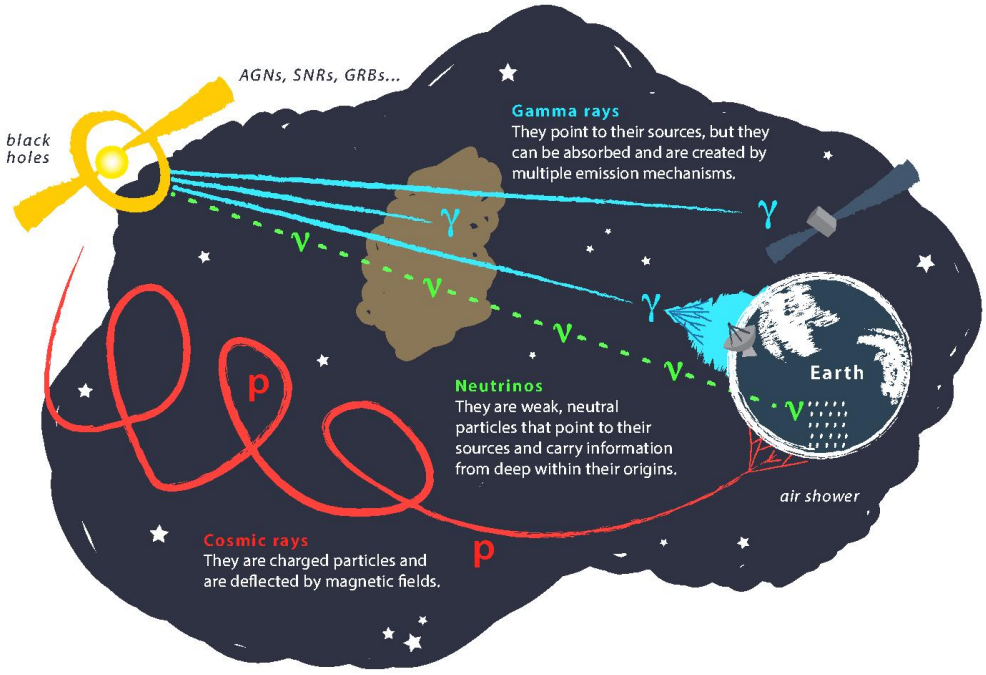

Figure 1.1.: The figure shows the multi-messenger approach, utilizing different observation channels, figure taken from the IceCube collaboration `www.icecube.wisc.edu`.

2. Cosmic Messenger Particles

Cosmic messengers are particles produced in sources of cosmological distances. They travel over large distances and can eventually be detected on Earth. Cosmic messengers are the only way how we can learn about the universe were are part of. The cosmological distances are enormous compared to human scales and even with the speed of light, the typical travel distance from the sources to Earth take up to millions or billions of years. Thus, cosmic messengers have to be stable to survive their journey. This leaves photons, electrons, stable nuclei and neutrinos as the cosmic messengers particles[1]. Cosmic messenger particles are observed up to energies of $10^{21}\,\mathrm{eV}$ [9]. The sources of these extreme high energetic particles are still unknown and one of the main challenges for today's astroparticle physics. This chapter discusses properties of the high energetic cosmic messenger particles with a special focus on neutrinos.

2.1. Cosmic Rays

In 1912 Victor Hess discovered an ionizing radiation in a set of balloon experiments [3]. The observed radiation intensity increased with height, which was surprising at that time since the only known sources of ionizing radiation where radioactive elements mainly present in rocks in the Earth. Since the Earth could be excluded as the source of this radiation and only the direction of the origin, the cosmos, was known, it was simply called *cosmic rays*. Cosmic rays have been studied since then. Nowadays the spectrum is well measured up to about $10^{21}\,\mathrm{eV}$ as shown in figure 2.1. The cosmic rays contain mainly protons and heavier nuclei [9, figure 29.1 and table 29.1]. The contribution of electrons and positrons can be neglected in the energy regimes above

[1] Neutrons can also be called cosmic messenger particles since at large energies, they can still travel astrophysical distances before they decay. For cosmological distances, the possible distances are too small. For the context of this work, neutrons are not discussed. Gravitational waves are also considered as a cosmic messenger, even if the corresponding particle, the graviton, has not been discovered yet.

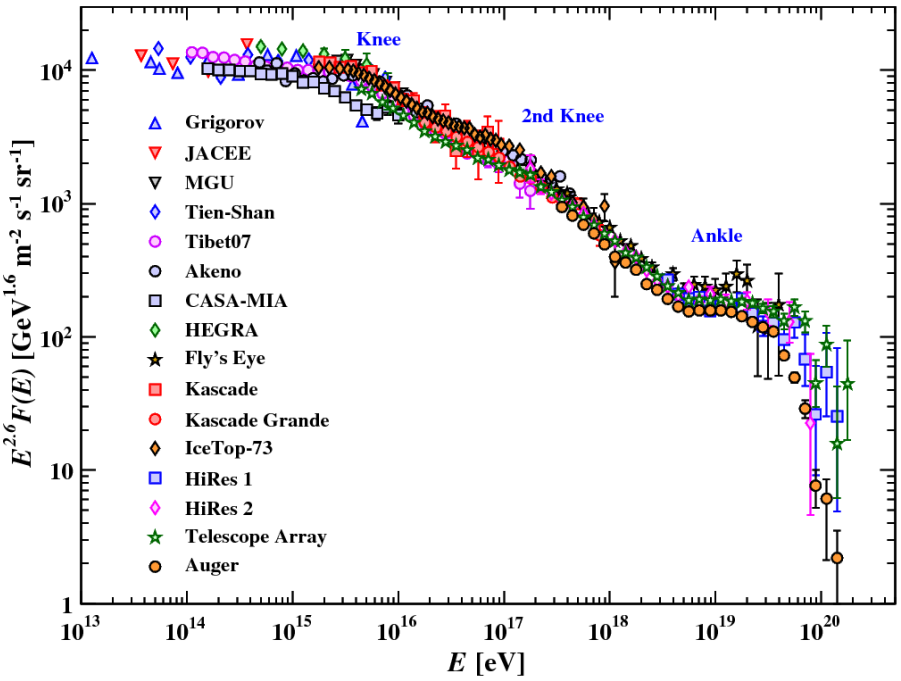

Figure 2.1.: Primary cosmic-ray energy spectrum multiplied by $E^{2.6}$ to highlight the changes of spectral index. Figure taken from [9].

a few GeV, see [9, figure 29.2]. The energy spectrum of cosmic rays can be described by a set of power laws of the form $E^{-\gamma}$ where γ is the spectral index. The changes in the spectral index are potentially connected with the transition between different source classes, e.g. the transition from galactic to extra-galactic sources.

After over a hundred years of studying the cosmic rays, their origin is still unknown. Due to magnetic fields present in space and the electric charge of cosmic rays, their measured direction does not point back to their sources. Identifying the sources is one of the big open challenges in modern astroparticle physics. The extreme energies which are observed in cosmic rays require special environments for the production of cosmic rays. The potential sources of the highest energetic cosmic rays are naturally

linked to extreme objects like exploding stars, active galaxies, gamma-ray bursts and black holes [10]. A connection between the size of a potential source and the strength of magnetic fields to accelerate charged particles was derived by Hillas already in 1984 [11]. This Hillas criterion is based on the idea that for a charged particle to be efficiently accelerated, the size of the accelerator has to be larger than the Larmor radius. This sets an upper limit on the maximum energy which can be reached by a particle accelerator in the source, given a certain magnetic field strength. Figure 2.2 shows the original famous Hillas plot.

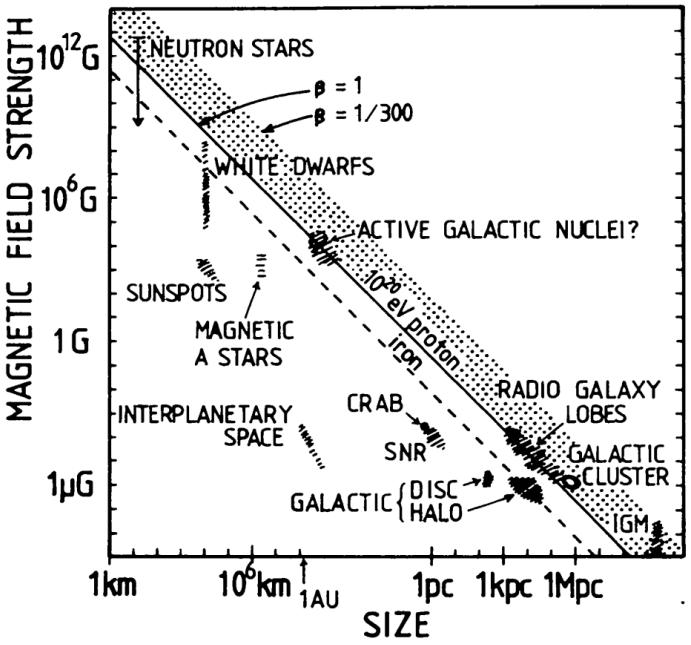

Figure 2.2.: The original Hillas plot [11]. Shown is the Hillas criterion to reach an energy of 10^{20} eV in protons and iron. Some potential source classes are shown in the plot as well.

2.2. Diffuse Shock Acceleration

The mechanism to accelerate particles to this extreme high energies is not jet understood and subject to current research. One potential mechanism is diffuse shock acceleration (DSA), a realization of Fermi acceleration [12, 13, 14]. DSA naturally

produces particle spectra following a power law similar to the observed energy spectra in cosmic rays. A requirement for DSA is the presence of shocks. Shocks occur when an object, e.g. a plasma cloud, moves supersonically through the surrounding medium. This medium is a fluid and in astrophysical scenarios typically a plasma. The front of the shock forms a discontinuity in density and velocity in the medium since any information in the medium can only be transmitted with the speed of sound in that medium. At the shock front, efficient acceleration of charged particles might be possible if diffusive shock acceleration is realized by nature.
The idea of DSA is that a charged particle gains a small amount of energy every time it crosses the shock front and is scattered back. To reach high energies it has to cross the shock front several times. Thus, a magnetic field behind and in front of the shock front is required to confine the particle. If the particles gain a fixed fraction of its current energy in every cycle and also has a certain probability of escaping the accelerator, the resulting differential spectrum has a power law shape.
In the rest frame of the shock front, there is an upstream region from where the plasma is moving with u_1 towards the shock front and a downstream region where the plasma is moving away from the shock with $u_2 < u_1$. The distribution function of particles can be described by a differential equation $f(x, p, t)$ in space x, pressure p and time t. It can be shown that a power law

$$f = f_0 \cdot p^{-q}$$

can solve the differential equation with $q = \frac{3r}{r-1}$ and $r = u_1/u_2$ as the *compression ratio*, the ratio between up and down-stream velocities [15]. Typical values for r are three to four and thus typical values for q are 4 to 4.5. The result for f can be translated into [15]

$$\frac{\mathrm{d}N}{\mathrm{d}p} \propto f_0 \cdot p^{-q+2}. \tag{2.1}$$

At this energy, the rest mass can be neglected and the momentum is equal to the energy. The energy spectrum can be described by power law spectra E^{-2} to $E^{-2.5}$. This derivation ignores energy losses in the source or on the way to the observer. When including these, resulting spectra soften and are compatible with observations, see figure 2.1.
More realistic treatments of the acceleration scenario include non-planar shock fronts,

turbulence in the context of magnetohydrodynamics, relativistic shock velocities, and energy loss processes. The treatment of these problems is typically done with numerical simulation [16, 17].
It is believed that the cosmic rays can be accelerated by DSA processes. Shocks are common phenomena in astrophysics and are observed on various scales, from our local stellar system up to the size of galaxies. DSA is potentially a very general mechanism and might be realized in some of these shock environments.

2.3. Particle Interaction of Cosmic Rays

High-energetic cosmic-ray particles can either interact at the source or on their way to Earth. In this interaction, they might produce secondary particles which are also cosmic messengers them self like γ photons or neutrinos. Two types of interaction are discussed here, the photo hadron interaction and the hadron interaction. The following discussion is restricted to the interaction of protons but works similarly with heavier nuclei.

Hadronuclear Interactions Cosmic-ray protons ray can interact with gas, located in the source or in interstellar gas clouds. For simplicity, the gas is assumed to consist of protons only. In the proton-proton interaction, many unstable hadrons are produced. The decay of the unstable particles will eventually happen via the lightest hadron, the pion, as shown in equation 2.2.

$$p + p \longrightarrow \begin{cases} \pi^{\pm} + X \\ \hookrightarrow \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu}) \\ \quad \hookrightarrow e^{\pm} + \nu_e(\bar{\nu}_e) + \bar{\nu}_{\mu}(\nu_{\mu}) \\ \pi^0 + X \\ \hookrightarrow 2\gamma \end{cases} \tag{2.2}$$

X is indicating all potential secondary particles heavier than pions. These secondaries might also decay via pions. Neutral pions will dominantly decay into two γ, positiv (negative) charged pions decay into a muon (anti-muon) and a neutrino (anti-neutrino) [9]. The muon then continues to decay into two additional neutrinos and an electron

or positron, depending on the charge of the muon. The decay products follow the initial proton spectrum, so the expected neutrino spectrum will follow the initial $E^{-\gamma}$ cosmic ray spectrum with $\gamma \approx 2 - 2.5$.

Photo-Meson Interaction Cosmic ray protons can also interact with ambient photons. The photons can be produced in the same source or be present as infrared or CMB photons. The interactions happens in the simplest case by the production of a Δ^+ resonance which then either decays into a proton or a neutron and the corresponding pion (figure 2.3).

$$p + \gamma \longrightarrow \Delta^+ \longrightarrow \begin{cases} \pi^\pm + n \to \pi^\pm + p^+ + e^- + \bar{\nu_e} \\ \qquad\qquad \hookrightarrow \mu^\pm + \nu_\mu(\bar{\nu}_\mu) \\ \qquad\qquad\qquad \hookrightarrow e^\pm + \nu_e(\bar{\nu}_e) + \bar{\nu}_\mu(\nu_\mu) \\ \pi^0 + p \\ \hookrightarrow 2\gamma \end{cases} \tag{2.3}$$

The neutrino spectrum depends both on the cosmic ray spectrum and on the photon spectrum. Harder photon spectra result in a higher interaction probability and also in a harder neutrino spectrum [18]. Both photo-meson interaction and hadronuclear interaction work very differently, but the resulting neutrino spectrum both have a high energetic component and also the flavor ratio is equal. The same is also true for the produced γ particles from π^0 decay, but absorption of γ-photons can be quiet different due to the different environments of pp and $p\gamma$ interaction, so the observed γ-ray signal might be very different.

2.4. Neutrinos

An astrophysical neutrino flux has first been discovered by IceCube in 2013 [7] at energies between $\mathcal{O}(100\,\mathrm{TeV})$ and $\mathcal{O}(1\,\mathrm{PeV})$. A follow-up analysis combining several data sets then decreased the lower energy bound to about 10 TeV. An energy spectrum of $E^{-2.5}$ was fitted to the data [19], see figure 2.3. Several searches for spatial clustering of the direction of neutrinos have been performed so far, but no significant clustering has been found [20]. Due to the absence of any point-like source, the astrophysical

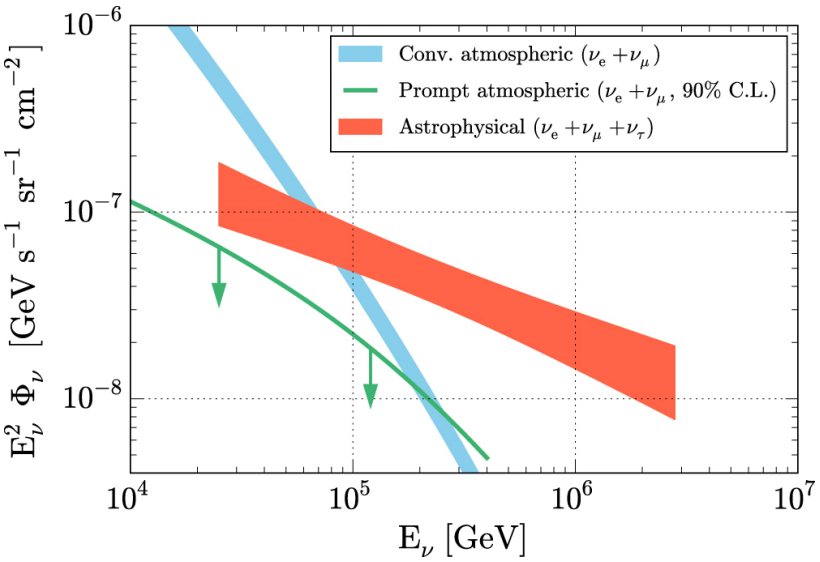

Figure 2.3.: Best-fit neutrino spectra for a single power law model (all flavors combined). The blue and red shaded areas correspond to 68% C.L. allowed regions for the conventional atmospheric and astrophysical neutrino flux, respectively. The prompt atmospheric flux is fitted to zero; shown is the 90% C.L. upper limit on this component instead (green line) [19].

flux is called diffuse. Since the previous discussion already suggested small, compact accelerators as sources of cosmic rays and thus also neutrinos, the nondetection of sources of the astrophysical neutrino flux points to a large population of dim sources to account for the measured diffuse neutrino flux. It has been estimated that at least order of 100 point sources is needed to produce the observed, unresolved astrophysical neutrino flux [21]. A contribution of neutrinos from core-collapse supernovae to the measured diffuse neutrino flux is discussed in this thesis.

3. Stellar Evolution and Supernovae

This chapter gives a brief description of the evolution of stars from their formation to their spectacular eventual end, in various types of supernova explosions. The supernova explosion mechanism and the associated generation of high energetic particles are described. Supernova models which predict production of high-energy neutrinos and their connection to observed supernovae classes are also discussed since they motivate the neutrino-supernovae correlation search.

3.1. Star Formation and Evolution

Stars form from cold, massive gas clouds. The self-gravity of the cloud causes gas to fall in towards the center of gravity of the cloud. The compression increases temperature and pressure in the center, which then counteracts the further in-fall of gas. Depending on the initial conditions of the gas cloud, pressure and temperature in the center can become large enough for nuclei to overcome the Coulomb barrier and initiate nuclear fusion. This initial fusion reaction is typically hydrogen to helium [22].
The fusion process generates radiation. This radiation acts as counter pressure against the gravitational pressure from outer shells of the gas cloud. Since the fusion rate in the core increases with density and temperature, it can stabilize the system. Gravitational pressure and radiation pressure are in equilibrium at that point. Such an object is defined as a star. In this stage, the star is performing fusion at a constant rate. It is also typically the longest stage in a star's life. The ongoing fusion process of hydrogen will enrich the stellar core with helium. At this stage helium normally[1] cannot undergo fusion, since temperature and pressure are too low to overcome the Coulomb barrier of the helium nuclei[2]. As the amount of helium in the core grows,

[1] If the star is very heavy, helium fusion in the core can also be possible at early stages of the star, while hydrogen is undergoing fusion in a shell around it.

[2] As a simplified rule of thumb, heavier nuclei contain more protons, have a stronger electric field and thus higher Coulomb barrier. Thus larger temperature and pressure are required for fusion.

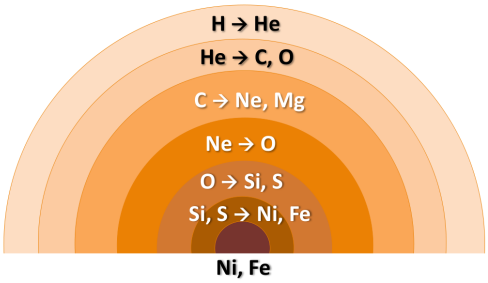

Figure 3.1.: Sketch of shell burning in a massive star ($20M_{\odot}$) [22].

it starts to replace the hydrogen. Since temperature and pressure are only large enough for fusion of hydrogen but not for helium, the region where fusion is possible decreases, and consequently the total fusion rate also decreases. The total radiation counter-pressure decreases and gravity compress the star even further. This process further increases temperature and pressure in the core and can eventually initiate helium fusion. As helium fusion begins, it generates radiation which again balances the gravitational pressure. The star moves to a new equilibrium.
Enrichment of the core with products of previous fusion reactions and initiation of subsequent fusion reactions can happen several times and characterizes the evolution of the star. A star typically spends most of its lifetime in the stage of hydrogen and helium fusion. Fusion of heavier elements usually happens on much shorter time scales but also releases more energy. As a consequence, stars tend to undergo many different stages of evolution towards the end of their life. An overview of the properties of various fusion processes is discussed in [23, 24]. These can happen simultaneously at different shells of the star, a process called shell burning. A sketch of a shell burning core can be seen in figure 3.1. It is worth noting that figure 3.1 only represents the inner core of the star. The majority of the star still contains plasma which is not undergoing any fusion. During the later stages of stellar evolution, with the fusion of heavier elements, the radiation pressure can dramatically increase. The increase of pressure expands the diameter of the star since the core expands, displacing outer layers. In the most extreme cases, the outer shell of the star is blown away. Typical examples of this effect are Wolf-Rayet stars [25]. As a result of the increase in size, the optical luminosity also increases due to the larger surface of the star. Initiation of heavier element fusion continues until the gravitational pressure is insufficient to

provide enough pressure and temperature in the core to initiate the next fusion stage or if the core has reached iron. Iron has the highest binding energy per nucleon, and thus fusion processes terminate here. At this point, the radiative pressure vanishes and the star collapse under the gravitational pressure. Depending on the initial mass of the star and thus on its evolution, it can result in one of the several possible scenarios.

3.2. White Dwarfs

If the final mass of the star at the end of the fusion process is below the so-called *Chandrasekhar mass,* of $1.44 M_{\odot}$ [26], it will most likely end as a white dwarf. When fusion stops, the star collapses under its gravitational pressure leading to core compression. The pressure in the core is ultimately balanced by electron degeneration pressure, stabilizing the white dwarf[3]. No fusion processes happen in white dwarf anymore. White dwarfs are small, hot objects since they contain most of the gravitational energy of the former star. Due to their high temperature, they appear white in the optical observations. Their low luminosity indicates a small surface, giving them the name dwarf. A white dwarf cools slowly over long time scales with decreasing brightness and temperature. No plausible mechanism is known for these very static objects to accelerate cosmic rays to the highest energies. White dwarfs are typically not considered as potential sources of the high energetic astrophysical neutrinos. Therefore this work instead focuses more on the heavier stars and their later evolution.

3.3. Supernovae

If the mass of the star core is larger than the Chandrasekhar mass [26], the stop of fusion reaction can result in a spectacular event, a so-called supernova. There are two known types of supernovae, thermonuclear supernovae and core-collapse supernovae (CCSN). This work focuses on core-collapse supernovae, so the mechanism of thermonuclear supernovae is only briefly described here.

[3]It is interesting to note that the classical electrostatic force is not dominant, but the quantum physical Pauli principle.

3.3.1. Thermonuclear Supernovae

Thermonuclear supernovae are believed to emerge from binary systems of a white dwarf and a companion star. Due to tidal forces, the white dwarf accretes matter from the companion star. It becomes heavier until it eventually exceeds the Chandrasekhar mass [26]. At this point, the gravitational pressure can no longer be compensated by the Pauli degeneration pressure of electrons, so the white draws stars to collapse. The temperature and pressure in the core suddenly increase and triggers an almost instantaneous fusion process. The rapid release of energy causes the star literally to explode. Since the mass limit where the white dwarf explodes is determined by the Chandrasekhar mass [26] and their intrinsic brightness is known and also very large, thermonuclear supernovae make good standard candles for cosmology. The known intrinsic luminosity of these supernovae allows measuring the distance as a function of the redshift, which is one of the main observables modern cosmology. For further discussion, see [27, 28]. Thermonuclear supernovae happen in a 'clean' environment, meaning that not much circumstellar material is around. The formation of a shock and the associated shock acceleration of charged particles is very unlikely. Because of this, thermonuclear supernovae are not considered as a source of high energetic neutrinos in this thesis.

3.3.2. Core Collapse Supernovae

Core Collapse supernovae are most likely the final stage of the evolution of massive stars. They are of special interest for neutrino astronomy, since they are believed to be able to accelerate charged particles to high energies. They would thus produce high-energy neutrinos and other messenger particles. The physics of core-collapse supernovae is still a topic of active research and not fully understood. Here a summary of the process is given. For a more detailed discussion of core collapse supernovae, see the excellent review [29] and further literature referenced there.
Again, core-collapse supernovae (CCSN) are believed to happen at the end of the life of massive stars. When fusion reactions finally stop, the star consists of shells of the remnants of the previous fusion processes. It starts to contract under its gravitational pressure. If the iron core exceeds the Chandrasekhar mass [26] of $1.44 M_{\odot}$, electron degeneration pressure cannot longer stabilize the core, and it starts to collapse. This is the beginning of the so-called core-collapse supernova (figure 3.2, upper left panel).

At the initial stage of the core collapse, the iron atoms in the core undergo electron capture

$$\mathrm{p}^+ + \mathrm{e}^- \rightarrow \mathrm{n} + \nu_e$$

which reduces the electron pressure. It also shifts the composition from iron to more neutron-heavy nuclei which are unstable and undergo β decays. Electron capture, β decay and also photodisintegration of iron nuclei to helium cool the core and reduce the electron density. As a consequence, the core collapse accelerates. At a density of $\rho_{\mathrm{trap}} \approx 10^{12}\,\mathrm{g/cm^3}$ the core becomes opaque for neutrinos. They become trapped since their escape time is larger than the collapse time (figure 3.2, upper right panel). At this stage, the core is essentially homogeneous. The in-falling and compression continues until the core reaches nuclear density of $\rho_{\mathrm{nuc}} \approx 10^{14}\,\mathrm{g/cm^3}$. The nuclear matter is much less compressible than the previous plasma, halting any further continue to fall-in and thus any further compression stops. The outer layers of the core are still in-falling on the in-compressible inner core. The core rebounds back creating a shock wave traveling outwards through the still in-falling outer shells of the star (figure 3.2, middle left panel). This sets the stage for the final supernova explosion: If the rebounding shock is strong, it does not only stop the outer layers from in-falling on the core but also moves outwards and blows away the outer shells. The result would be observed as a supernova. This mechanism is called *prompt* mechanism. Current research and modeling show that supernova shocks are most likely not energetic enough for this process since the shock loses much of its energy by dissociation of heavier nuclei in the outer shell [29]. This dissociation increases the cooling rate again since electron capture is more efficient for protons than for heavier nuclei. Nevertheless, the neutrinos produced in the electron capture processes leave the star and form the so-called *prompt neutrino burst* which carries away energy. This leads to even more electron capture. The shock stalls and material downstream resumes in-fall on the core again (figure 3.2, middle right panel).

After the core bounce, additional material will fall in on the core and form a compact remnant. This *proto-neutron star* will then evolve into a neutron star or a black hole. The limit for the formation of a black hole is the Tolman–Oppenheimer–Volkoff limit of about $2.5 M_\odot$ [30, 24]. The proto-neutron star is still opaque to neutrinos which remain trapped inside. Instead, they diffuse out of the proto-neutron star (figure 3.2, lower left panel). The neutrinos deposit their energy by interactions with the outer

shells, mainly via the interactions:

$$\nu_e + n \rightarrow e^- + p$$
$$\bar{\nu}_e + p \rightarrow e^+ + n$$

as shown in figure 3.2 on the lower right panel. This deposition of energy revises the shock. It starts to move outwards and can finally cause the supernova explosion. The mechanism is called delayed neutrino-heating mechanism.
As mentioned before, the details of the core collapse mechanism are still subject to current research and not fully understood. It is remarkable that neutrinos, despite being weakly interacting particles, are the driving factor behind core-collapse supernova explosions. The physical properties of the supernova depend on many parameters, such as the initial condition of the star as well as the supernova mechanism itself. While there are correlations between observed supernovae and progenitor stars, it is currently impossible to determine all properties of the progenitor by just observing the supernova explosion.

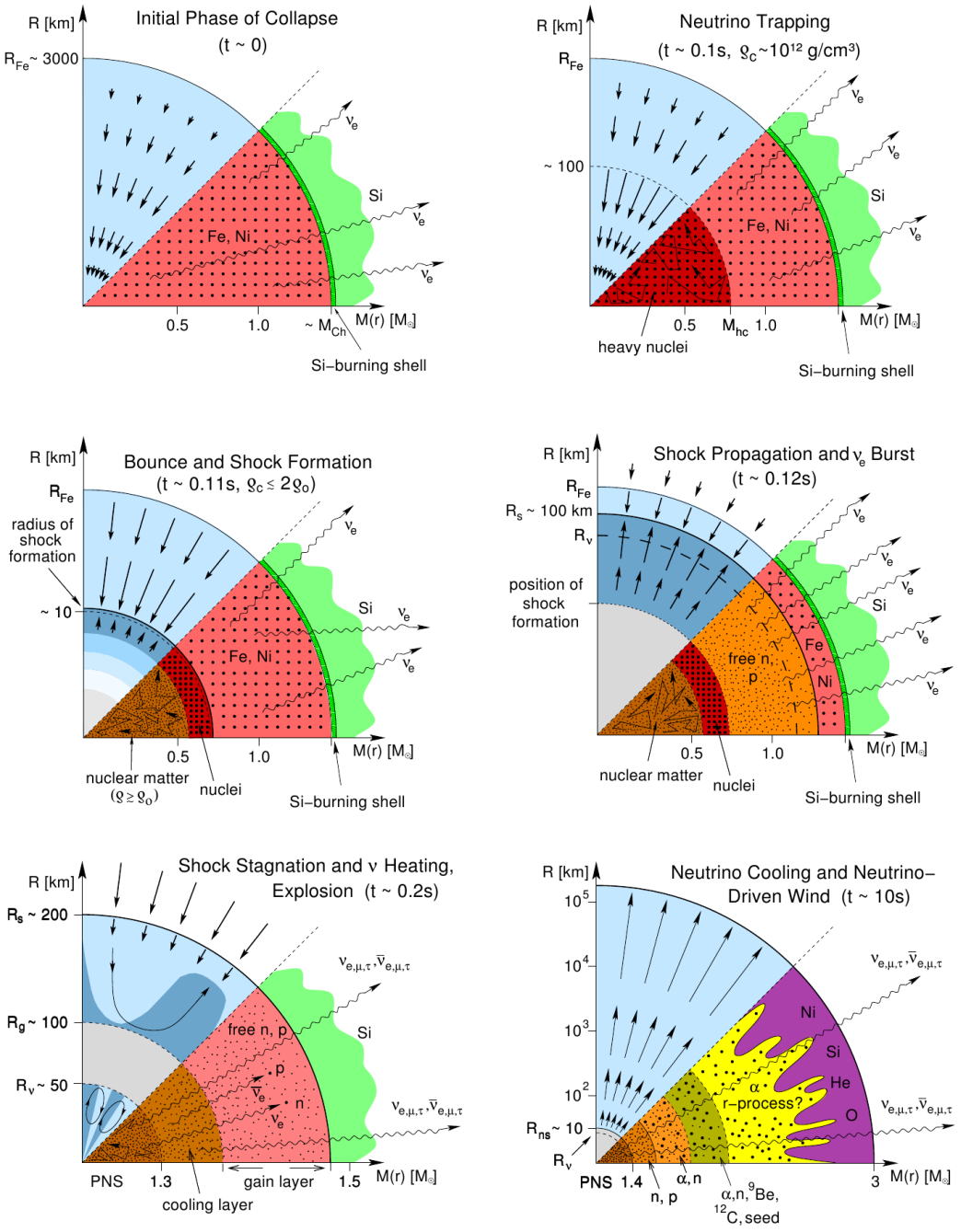

Figure 3.2.: Sketch of the states of a core-collapse supernova as discussed in the text. Figure taken from [29].

3.4. Supernova Classification

Supernovae are classified by observed spectral features. The presence or absence of certain atomic lines, in combination with the shape of the light curve, defines the supernova class. This classification is purely observational and does not easily connect to the explosion mechanism of the supernova. A schematic view of the classification is shown in figure 3.3 and discussed in [31]. Figure 3.3 only shows a simplified picture. The supernova spectrum can change during the temporal evolution. The presence or absence of a line is not a binary feature, but a relative measure. A more detailed classification of supernovae into several sub-classes is described in a recent publication [32].

The mapping from observed spectral classes of supernovae to their physical classes is challenging. While it is generally believed that type Ia supernovae are of thermonuclear origin and all other types are the result of a core-collapse supernova, a further separation into different progenitors of the core collapse supernovae is still subject of current research. As discussed in [31] and [32], there seems to be a connection between massive circumstellar medium and type IIn supernovae. There are also claims of a connection between a massive, high-loss stars and type Ib, Ic, Ibc and IIb supernovae. Type Ic supernovae are also believed to be connected with GRBs [33] supporting the assumption of a jet present in these type of supernovae.

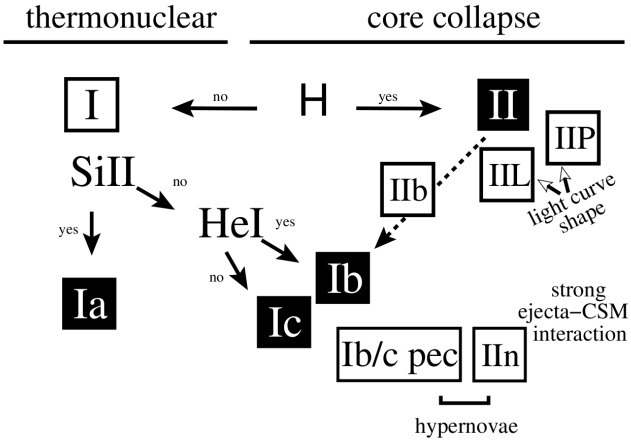

Figure 3.3.: Classification schema for supernovae, figure taken from [31].

3.5. Neutrino Production in Core-Collapse Supernovae

Core collapse supernovae are a potential source class for high energetic neutrinos ($E_\nu > 100\,\mathrm{GeV}$) and might be able to explain the diffuse astrophysical neutrino flux discovered by IceCube [7, 19]. The high energetic neutrinos are not produced by the core collapse itself, but rather by diffusive shock acceleration of charged particles and subsequent production of neutrinos as discussed in section 2.2. Supernovae are promising sources since they provide the required environment for efficient shock acceleration: Fast shocks in a dense, potentially magnetized medium which provides scattering of charged particles and can also contain the accelerated particles at higher energies. There are two potential scenarios for neutrino production which are studied in this thesis, the circumstellar medium supernovae, and the choked jet supernovae.

3.5.1. Circumstellar Medium Supernovae

Circumstellar medium (CSM) supernovae are supernovae which have a large and massive medium surrounding the star (hence the name). Most likely, this is a result of a strong mass loss in the later stages of the stellar evolution. This could either be due to strong stellar winds or small outburst before the final supernovae [34]. The CSM could also be the entire outer shell of the star itself, blown away by strong radiation pressure (as in a Wolf-Rayet star [25].)
When the core collapse supernova explodes, the ejecta works as a piston, compressing the circumstellar medium and forming shocks. These shocks then move through the circumstellar medium and provide an environment for potential diffuse shock acceleration. Around the shock front, turbulence and compression of the plasma are expected. Due to the flux-freezing theorem [35], also a strong amplification of the local magnetic field is expected. This leads to trapping and scattering of charged particles close to the shock and thus to a fast and efficient acceleration of charged particles. The scenario is very similar to the standard supernova remnant evolution but occurs on much shorter time scales [36]. The CSM supernova has been modeled, and the potential neutrino emission has been studied independently by Murase et al. (model I) [37] and Zirakashvili and Ptuskin (model II) [38].

Model I Murase et al. performed a modeling of particle acceleration of supernovae in dense circumstellar media supernovae with a special focus on neutrino and γ emission

[37]. This work estimates the expected neutrino emission from a supernova based on an energetic argument. The assumption is a spherical shell of constant density into which the supernova ejecta crashes. The model does not take into account temporal evolution of the neutrino signal, but provides the total integrated flux (called fluence). The duration of neutrino emission is expected to last about 10^7 to 10^8 seconds, depending on the parameter of the supernova and the circumstellar medium. The estimated energy spectrum follows a power law ($E^{-\gamma}$) with spectral index of $\gamma = 2$. The kinetic explosion energy of the supernova, the ejecta mass as well as the circumstellar medium density and its radius determine fluence and duration of the expected neutrino emission. The fluence of muon neutrinos Φ_ν can be estimated by

$$E_\nu^2 \Phi_\nu \approx 6 \cdot 10^{-2}\,\mathrm{GeV cm^{-2}}\,\min(1, f_{pp})\epsilon_{\mathrm{cr},-1}\mathcal{E}_{\mathrm{ej},51}d_1^{-2} \tag{3.1}$$

where f_{pp} is the efficiency for the pp hadro-nuclear interactions, $\epsilon_{\mathrm{cr},-1}$ is the efficiency of conversion from kinetic energy of the ejecta to cosmic rays in units of 0.1, $\mathcal{E}_{\mathrm{ej},51}$ is the kinetic ejecta energy in units of 10^{51} erg and d_1 is the distance to the source in units of 10 Mpc. Murase et al. discuss two models (A and B), which are supposed to span the range of potential circumstellar medium supernovae. Model A assumes a shell density of $n_{\mathrm{sh}} = 10^{11}\,\mathrm{cm^{-3}}$, a distance and thickness of the shell of $R_{\mathrm{sh}} = \Delta R_{\mathrm{sh}} = 10^{15}\,\mathrm{cm}$ and shock velocities of $V_f = 10^{3.5}\,\mathrm{kms^{-1}}$ and $V_r = 10^4\,\mathrm{kms^{-1}}$ for forward and reverse shock. Model B assumes $n_{\mathrm{sh}} = 10^{7.5}\,\mathrm{cm^{-3}}$, $R_{\mathrm{sh}} = \Delta R_{\mathrm{sh}} = 10^{16.5}\,\mathrm{cm}$ and $V_f = 10^{3.7}\,\mathrm{kms^{-1}}$ and $V_r = 10^{3.9}\,\mathrm{kms^{-1}}$. Model A is designed to mimic short, bright supernovae like SN 2006gy (radiation energy $\mathcal{E}_{\mathrm{ph}} \approx 10^{51}\,\mathrm{erg}$ and peak luminosity $L_{\mathrm{ph}} \approx 10^{44}\,\mathrm{ergs^{-1}}$) and model B models dimmer, longer lasting supernovae like SN 2008iy ($\mathcal{E}_{\mathrm{ph}} \approx 10^{50}\,\mathrm{erg}$ and $L_{\mathrm{ph}} \approx 10^{42.5}\,\mathrm{ergs^{-1}}$). The shape of the shell does not have a strong influence on the fluence [37], and also the two different models produce similar fluence, see figure 3.4

Model II Ptuskin and Zirakashvili also study the potential neutrino emission of circumstellar medium interaction supernovae, especially of type IIn supernovae [38]. They take the temporal evolution into account, use a Monte Carlo simulation and then parameterize the outcome of the simulation. For the circumstellar medium, a continuous strong wind is assumed with the typical $\rho \propto r^{-2}$ density profile is assumed as potentially present in high mass loss stars like Wolf-Rayet stars [25]. Simulations are terminated after 30 years. At this point the flux has decreased to a neglectible

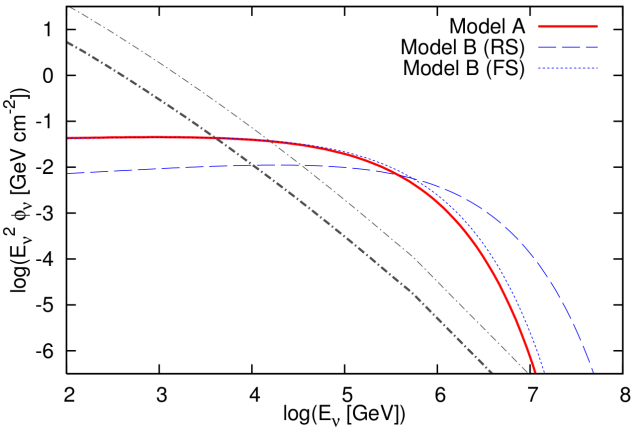

Figure 3.4.: Energy fluences of muon neutrinos from a SN crashing into dense CSM, where $\epsilon_B = 10^{2.5}$, $\epsilon_{cr} = 0.1$ and $d = 10\,\mathrm{Mpc}$ are assumed. Thick and thin curves represent Model A and Model B, respectively. The dotted-dashed curves show the zenith-angle-averaged ANB within a circle of radius 1°; we use $\Delta t = 10^7\,\mathrm{s}$ for Model A (thick line) and $\Delta t = 10^{7.8}\,\mathrm{s}$ for Model B (thin line) [37].

value. The muon neutrino flux expectation at a distance D and time t is given by

$$f(E_\nu)E_\nu^2 = 10^{-8}\,\frac{\mathrm{erg}}{\mathrm{cm^2 s}}\left(1+\frac{t}{t_{pp}}\right)^{-1} D_{\mathrm{Mpc}}^{-2}\xi_{CR}\left(\frac{\dot{M}}{10^{-2}\,\mathrm{M_\odot yr^{-1}}}\right) \times \left(\frac{u_w}{100\,\mathrm{kms^{-1}}}\right)^{-1}\left(\frac{E_{SN}}{10^{52}\,\mathrm{erg}}\right)^{3/2}\left(\frac{M_{ej}}{10M_\odot}\right)^{-3/2} \tag{3.2}$$

where D_{Mpc} is the distance of the source, $\dot{M}$ is the mass loss rate of the star, u_w is the wind velocity and E_{SN} and M_{ej} are supernova energy and ejecta mass. The time parameter t_{pp} is given by

$$t_{pp} = 0.2\,\mathrm{y}\left(\frac{\dot{M}}{10^{-2}\,\mathrm{M_\odot yr^{-1}}}\right)\left(\frac{u_w}{100\,\mathrm{kms^{-1}}}\right)^{-1}\left(\frac{E_{SN}}{10^{52}\,\mathrm{erg}}\right)^{3/2}\left(\frac{M_{ej}}{10M_\odot}\right)^{-3/2}. \tag{3.3}$$

The parameterization is chosen such that it refers to the typically assumed values. The time evolution is sketched in figure 3.5 for different values of t_{pp}. Note that about $50-75\%$ of the total flux is emitted within the first year of the explosion. The time scale is similar to the model I discussed in the previous paragraph.

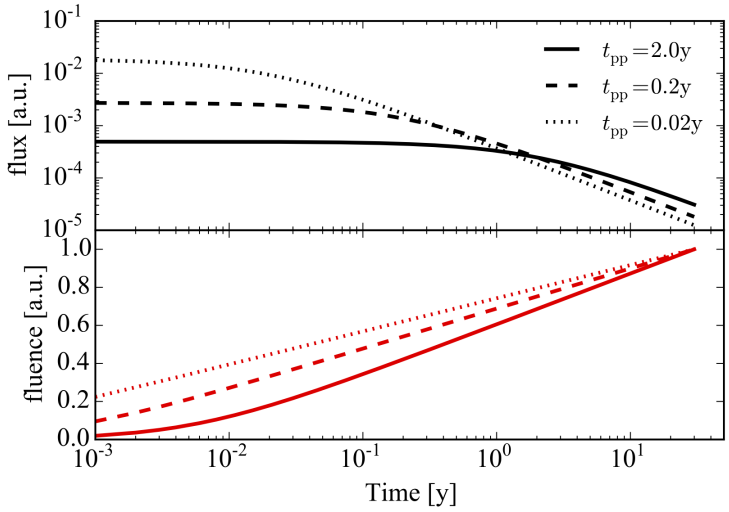

Figure 3.5.: Time evolution and cumulative flux of neutrinos for different time parameters from [38], all normalized to unity. The upper plot shows that the flux stays broadly constant for roughly $t_{pp}/2$. The lower plot shows the fraction of the total energy emitted up to a certain time. As discussed in the text, about $50-75\%$ of the total flux is emitted within the first year.

Comparing both models To compare both models (Murase et al. [37]) and (Zirakashvili and Ptuskin [38]), some estimates are made since the two models have different parameter assumptions.

The scenario to compare both models is chosen to be the one with the parameters of scenario A of Murase et al. It assumes a homogeneous CSM with radius and thickness of $10^{15.5}$ cm and a number density of $10^{11}\,\mathrm{cm}^{-3}$. Assuming that the CSM consists only of protons, the total mass would be about $11\,\mathrm{M}_\odot$. Translating this to the wind case in model II by Zirakashvili and Ptuskin where the parameters are stellar wind velocity and wind mass loss rate and assuming a wind velocity of $100\,\mathrm{kms}^{-1}$, the mass loss rate is about $1.1\,\mathrm{M}_\odot\mathrm{y}^{-1}$. The value is very large, but not unrealistic, assuming that such a strong wind only happens for the last years before the supernova. Furthermore, model I assumes a kinetic energy of the ejecta of 10^{51} erg and an ejecta velocity of $10^4\,\mathrm{kms}^{-1}$. This leads to an ejecta mass of about $1\,\mathrm{M}_\odot$.

These values are used to compute the time constant t_{pp} in model II using equation 3.2. The outcome is $t_{pp} = 10^{7.4}$ s. The value t_{pp} is the timescale during which a significant

fraction of the total neutrino energy is emitted, see figure 3.5. This value should be compared to the 10^7 s assumed in the model I [37]. Even if t_{pp} cannot easily be compared to the rough estimate on the total duration time in model I, both parameters end up in the same order of magnitude, even using very different approaches.
To compare the fluence estimate from both models, a source at a distance of 10 Mpc is assumed. The fluence prediction of model I is $\Phi E^2 = 6 \cdot 10^{-2}\,\mathrm{GeV/cm^2}$. Injecting the previously computed parameters into equation 3.2 and integrating over a time period of 30 years, the fluence estimate is $\phi E^2 = 2\,\mathrm{GeV/cm^2}$, about a factor 20 higher. If the time integration is restricted to 10^7 s, which is the time estimate of model I, the fluence output of model II is $\phi E^2 = 0.4\,\mathrm{GeV/cm^2}$, so only a factor three difference.
To summarize the comparison, both models predict a hard neutrino E^{-2} power spectrum. The typical timescales are the same if assuming similar supernova and CSM parameters. Though the models assume different scenarios (homogeneous CSM shell versus a wind like profile) and utilize different methods (semi-analytic calculation in model I [37] versus Monte Carlo simulation and parameterization of the results in model II [38]), results agree. Since the details of supernova parameters are generally not known from observations, these details are not of great importance for this work. Both models point into the same direction motivates a search for such hard neutrino spectra from individual supernovae. The main difference in the search for neutrinos is with the time regime. To cover this, a variety of parameters and the neutrino light curve models will be tested to cover the parameter space predicted by the two models.

3.5.2. Choked Jet Supernovae

The choked jet scenario is a model aiming to explain the connection of gamma-ray bursts and supernovae in a broader, unified picture. The basic idea is that a massive star is producing two anti-parallel jets when it undergoes core collapse. These jets then move outwards through the star envelope. Efficient shock acceleration is expected, both at the head of the jets and also in internal shocks within the jets [39, 40].
Depending on the properties of the star envelope and the jet itself, the jet eventually penetrates the photo-sphere and emits a strong gamma-ray signal from inside the jet together with neutrinos. The result is a GRB with prompt neutrinos (see figure 3.6, right panel). For a less-energetic jet or a more massive outer star shell, the penetration of the outer shell might ultimately not happen. In this case, the jet stalls inside the

star and the gamma-ray signal appears only when the energy from the jet reaches the photosphere. This process takes time, and the structure of the initial jet is destroyed in that process. The first neutrinos may have already left the star before the jet stalls. Thus they might appear before the gamma-ray signal. This mechanism is the proposed scenario for low-luminosity GRBs, see figure 3.6 again. In the case of an even denser star shell or less energetic jet, the jet is stalled or choked far inside the star. Neutrinos can still leave the star, but the gamma ray signal does not. Therefore the neutrinos are called orphan neutrinos. It eventually thermalizes and appears as a hypernova, an extremely bright supernova. For all cases, expected energy spectrum of neutrinos is potentially very hard [40].

As mentioned in [40], the expected duration of the neutrino emission in the choked jet scenario is about $10^{1.5}$ s, which is orders of magnitude shorter than the expected electromagnetic emission. Thus, in the case of a correlation search for these neutrinos based on observed optical counterparts due to large uncertainties in the time when the burst occurs because of the large associated uncertainty in the electromagnetic signal, the search window would be much larger than the expected neutrino emission duration.

The spectral classes of supernova expected to be connected with the choked jet scenario are mainly supernovae of types Ib/c where the progenitor is expected to be very massive. Thus the neutrinos produced by choked jets might contribute significantly to the observed diffuse high energetic neutrino flux. Many parameters of this model are uncertain, so there is no clear prediction for flux. Rather, we have order-of-magnitude estimates. Still, it is worth and an essential part of this thesis to test this proposed source of high energetic neutrinos. Therefore, very general and model-independent test will be applied.

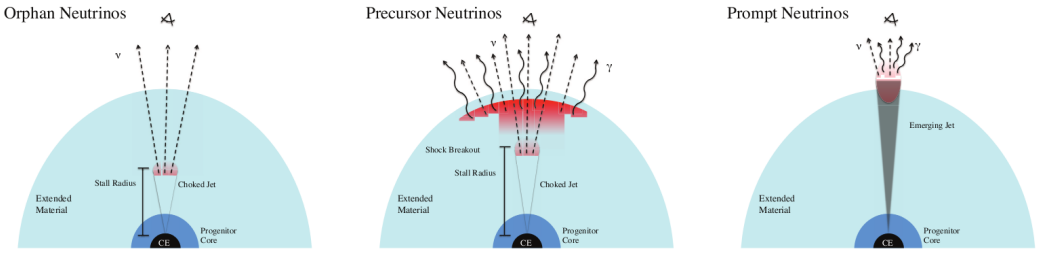

Figure 3.6.: Left panel: The choked jet model for jet-driven SNe. Orphan neutrinos (orphan in the sense that there is no correlated gamma signal) are expected since electromagnetic emission from the jet is hidden, and such objects may be observed as hypernovae. Middle panel: The shock breakout model for low luminosity (LL) GRBs, where trans-relativistic shocks are driven by choked jets. A precursor neutrino signal is expected since the gamma-ray emission from the shock breakout occurs significantly after the jet stalls. Right panel: The emerging jet model for GRBs and LL GRBs. Both neutrinos and gamma-rays are produced by the successful jet, and both messengers can be observed as prompt emission. Figure and caption taken from [40]

4. The IceCube Neutrino Telescope

This chapter describes the IceCube Neutrino Observatory and how neutrino detection works. IceCube is a $1\,\mathrm{km}^3$ detector located at the geographical south pole in the Antarctic glacier.

4.1. Neutrino Detection

Neutrinos are elementary leptons in the standard model. They were first proposed by Pauli to solve the problem of the energy spectrum observed in β decays: The β decay was believed to be a two-body-decay at that time ($n \to p^+ + e^-$), but the observed energy spectrum matched a three-body-decay. Pauli postulated the existence of a third, invisible particle present in the decay ($n \to p^+ + e^- + \bar{\nu}_e$), later called neutrino [41]. This neutrino then takes away part of the energy and explains the three-body decay spectrum observed in the electron. Since the neutrino has not been discovered at that point, it had at least to be electrically neutral.
Similar to the electron which has two heavier companions (the μ and τ leptons), there are the corresponding ν_e, ν_μ and ν_τ neutrinos in the accordance with the three generations, see figure (4.1). Today it is known that one main characteristic of the neutrinos is the lack of electrical and color charge. Thus neutrinos only interact via the weak force. This makes neutrinos interesting as cosmic messenger particles since they are not deflected or absorbed on their way to the observer, but they are also challenging to detect for the same reason.

4.1.1. Neutrino Interactions

Neutrinos interact only via the weak force and hence by the exchange of the charged $W^\pm$ and neutral Z^0 bosons. Interactions involving a neutral Z^0 boson are called neutral current interactions (NC), interactions involving a charged $W^\pm$ boson are

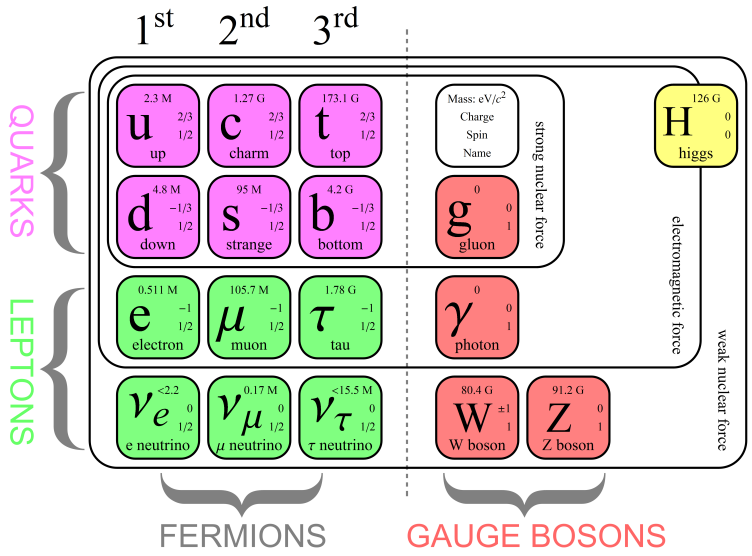

Figure 4.1.: Sketch of the standard model of particle physics, figure is taken from `http://www.physik.uzh.ch/groups/serra/StandardModel.html`.

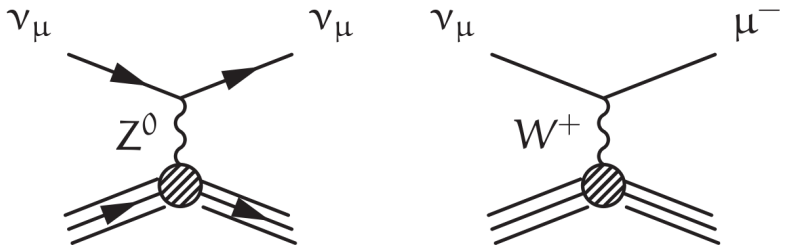

Figure 4.2.: Feynman diagram of a muon neutrino undergoing a neutral current (left plot) and charged current (right plot) interaction [42].

called charged current interactions (CC) (see figure 4.2). NC interactions only transfer momentum between the neutrino and the target particle. CC interactions involve the conversion of the neutrino into the corresponding charged lepton (e, μ, τ). At energy scales above 10 GeV, the energy is larger then the typical nuclei binding energy ($\mathcal{O}(1\,\mathrm{MeV})$) and the neutrino interacts with a single nucleon without regarding other constituents of the nucleus. The dominant interaction at this energy and above is deep inelastic scattering (DIS) [43]. The cross section is shown in figure 4.3. The

large mass of the $W^{\pm}$ (80.4 GeV) and Z^0 bosons (91.2 GeV) [9] lead to the small cross section of neutrino interactions, of the order of one pico barn. The cross section shows a general trend to increase with energy. The ratio of cross section for CC interactions to NC interactions is roughly three, independent of energy.

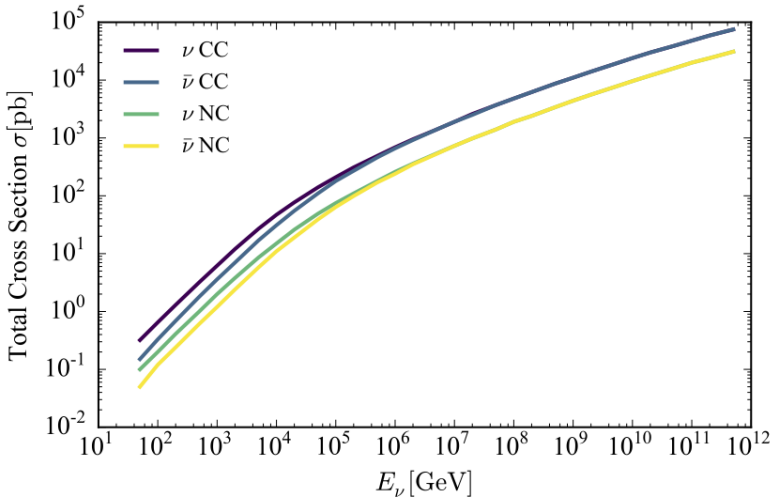

Figure 4.3.: Deep inelastic scattering cross section for energies above 10 GeV [44]. The cross section increases with energy.

4.1.2. Cherenkov Effect

In the deep inelastic scattering processes (both CC and NC), the scattering nucleus is typically destroyed, and many unstable particles are produced during the hadronization, seen in the lower right of the two Feynman graphs on figure 4.2. If the interaction is a charged current interaction, the corresponding charged lepton of the same flavor as the initial neutrino is produced. These particles are typically high energetic and have velocities close to the speed of light c. In the presence of an optical medium, the velocity is also typically larger than the speed of light in the medium $c' = c_0/n$ where c_0 is the vacuum speed of light, and n is the refractive index of the medium. The charged particles emit Cherenkov light [45] which is the main detection signal for neutrinos in IceCube.
When the charged particle passes through an optical medium, it polarizes the medium.

The polarized medium then emits a coherent light front, the Cherenkov light. The opening angle θ of the light cone is connected with the speed of the particle β by

$$\cos(\theta) = \frac{1}{n\beta}$$

where n is the refractive index and $\beta = v/c$ the velocity of the particle. Figure 4.4 shows a sketch of the generation of Cherenkov light. The spectrum of the emitted Cherenkov radiation is given by

$$\frac{\mathrm{d}^2 N}{\mathrm{d}\lambda \mathrm{d}x} = \frac{2\pi\alpha z^2}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2(\lambda)}\right) \tag{4.1}$$

where λ is the emitted wavelength, x the distance, α the fine structure constant, z is the charge of the particle, and n is the refractive index [9]. Cherenkov radiation has no clear peak but increases with frequency. Since $n(\lambda) \to 1$ for $\lambda \to \infty$, there is a natural cut off and the total energy output by Cherenkov radiation stays finite. The strongest Cherenkov light is typically present in the blue and ultraviolet regime. Therefore classical light detection techniques like photomultiplier tubes can be used for detection. Measurements of the Cherenkov light can be used to reconstruct the properties of the charged leptons and thus of the initial neutrino. Since the neutrino cross section is very low and the detection relies on Cherenkov photons, the natural choice for a detector would be to build an instrument from a large and transparent medium with very low background.

4.2. The IceCube Neutrino Detector

IceCube consists of several sub-detectors. This thesis focuses only on the in-ice detector. Other parts like the surface array IceTop are not used and also not discussed here.

4.2.1. Design

The IceCube neutrino detector utilizes the Antarctic glacier ice at the South Pole as a medium for Cherenkov neutrino detection [8]. There are 5160 photo sensors known as digital optical modules (DOMs), which are deployed between 1450 m and 2450 m

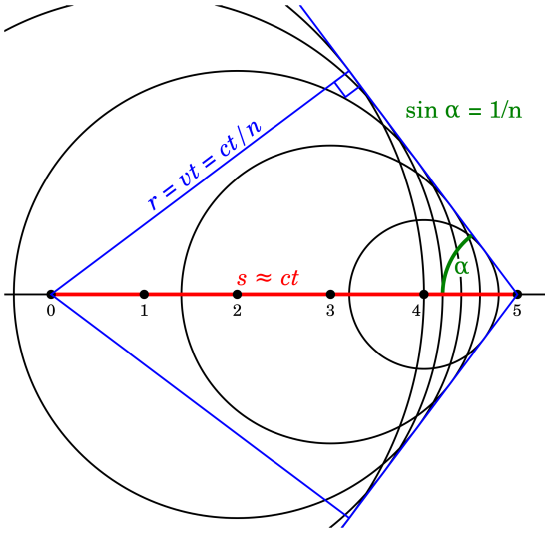

Figure 4.4.: Sketch of the geometry of the Cherenkov light. Spherical waves are produced at the point zero to five and the envelope, the Cherenkov cone, is indicated with the blue lines. Figure is taken from `https://commons.wikimedia.org/wiki/File:Cherenkov_Wavefront.svg`.

deep in the Antarctic glacier. The total instrumented volume is about $1\,\mathrm{km}^3$. Figure 4.5 shows a sketch of the detector setup. The detector is built by melting holes in the ice and deploying a so-called string before re-freezing of the water. DOMs are attached to the string with a vertical distance of about 17 m. The string itself consists of several wires providing power supply and communication for the DOMs. Once the string is frozen in the ice, it is operational. IceCube consists of 86 strings in total, with a horizontal spacing of typical 125 m in a hexagonal shape. Construction was finished in 2011, and the detector is fully operational since then. Data was also taken while the detector was still under construction with a fewer number of strings operational. The IceCube internal notation for datasets is ICXX where XX denotes the number of involving strings in the data taking. From 2011 on, the data sets are called IC86-I, IC86-II and so on, counting the years of full detector configuration. IceCube uses a spherical coordinate system to describe directions in detector coordinates. The two angles azimuth ϕ and zenith θ are used. The azimuth angle ϕ is defined between 0 and 2π and the zenith angle θ between 0 and π. A vector with $\theta = 0$ is pointing directly at the sky where $\theta = \pi$ corresponds to pointing towards the center of the Earth. Since

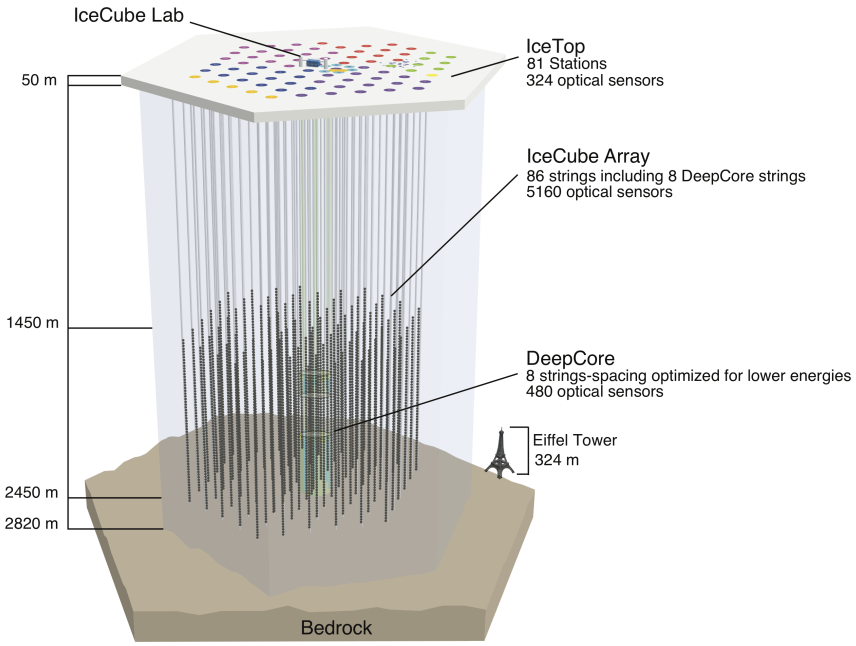

Figure 4.5.: The IceCube neutrino observatory as build in the Antarctic glacier. This thesis utilizes data from the in-ice array. The colors at the surface indicate different deployment seasons [8].

the θ axis is parallel to the Earth axis, the detector is spinning around its zenith axis with a period of one day. Thus objects in the sky essentially keep their zenith angle while their azimuth angle is constantly changing. Any effect on absorption in the Earth or atmosphere is thus only a function of zenith.

4.2.2. The Digital Optical Modul

The digital optical module (DOM) is the central building block of IceCube. A sketch can be seen in figure 4.6. The DOM consists of a 10 inch photomultiplier tube (PMT) [46] and a circuit board to control and handle the data output from the PMT. The electronics allow recording of the PMT pulses with nanosecond resolution. The PMT

and the circuit board are protected from external pressure by a 0.5-inch-thick glass sphere. The sphere is capable of enduring a pressure of 690 bar during the refreezing process during deployment of the string. An optical gel is applied to match the optical properties of glass sphere and PMT also provide mechanical support, see figure 4.6. The DOMs send their data to the surface along the strings which also provide power supply and a GPS clock signal for timing.

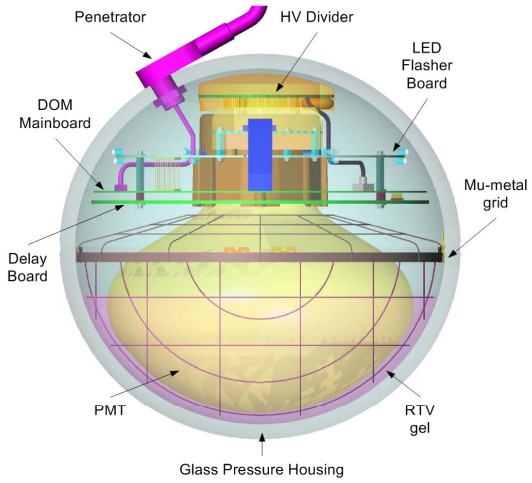

Figure 4.6.: Sketch of the IceCube digital optical module (DOM) [8].

4.2.3. Particle Detection with IceCube

When a charged particle moves through the detector, it emits Cherenkov light. The photons are scattered and absorbed in the ice of the Antarctic glacier in and around the IceCube detector. Some of the photons will eventually reach the photocathode of one of the DOMs and deposit a signal. Due to the excellent optical properties in terms of absorption length of 200 m and an effective scattering length of 70 m, photons can travel quite far in the ice and still be detected over large distances [47], see figure 4.7. The actual measurement of the PMT is voltage over time. This is called the waveform and is the start of all event reconstructions.
If a few nearby DOMs each detect photons within a short time window, a trigger is generated, and the waveform from all DOMs are sent to a computer cluster located in the IceCube Lab located on the surface. The triggering techniques are described in

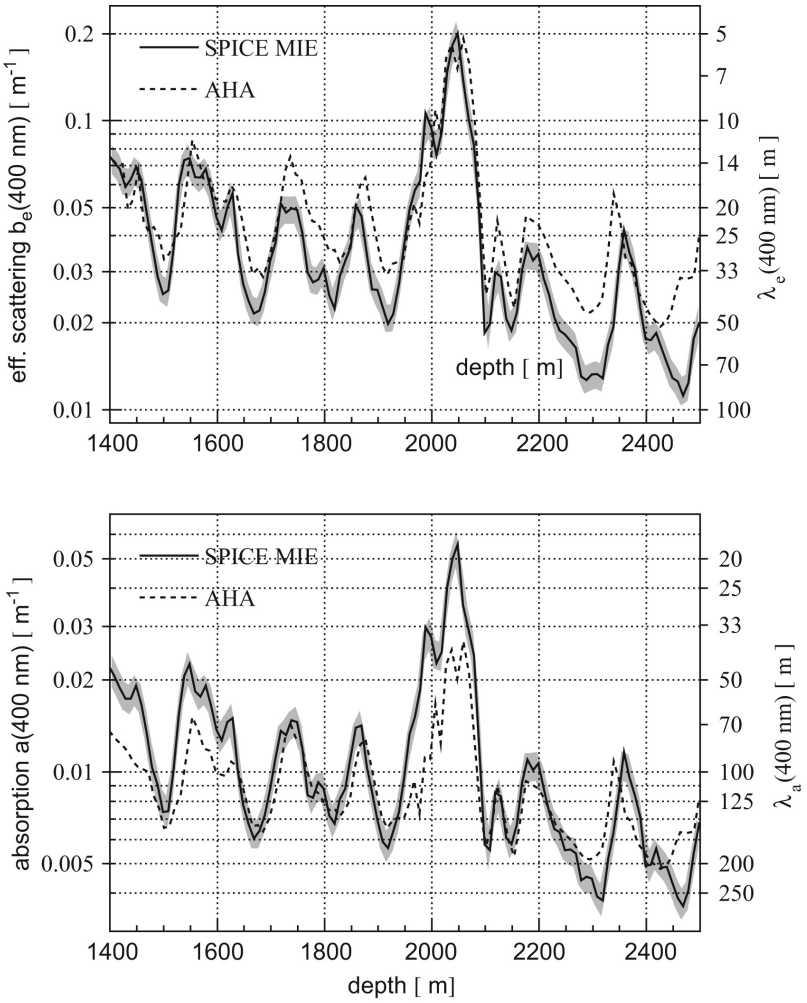

Figure 4.7.: Ice properties as a function of depth [47]. The dust layer, contamination in the ice at a depth of about 2000 m from volcanic activity can be seen.

detail in [8]. At the IceCube Lab, a computer cluster combines the different waveform to create an *event*. Every event is a potential neutrino interaction in or around the detector. These events are the elementary parts of IceCube data. The rest of this chapter describes how the events are classified, filtered and reconstructed.

4.3. Event Topologies

Depending on the neutrino flavor and the interaction type, there are two main event topologies. Track-like events are caused by muons which travel up to several kilometers through the ice. Cascades are blob-like signatures from electron, tau or NC interactions where the produced particles deposit their energy on very short distances. This is mostly below the spatial resolution of the detector. IceCube has a vertical spacing of 17 m and a horizontal spacing of ≈ 125 m, so these values somehow set an order of magnitude limit for the spatial resolution.

Cascades If a neutrino undergoes a neutral current (NC) interaction, the observed signal comes purely from the fragments of the interaction partner, a nucleus of the ice. The deep inelastic scattering reaction breaks of the nucleus (figure 4.2). In that process, high energetic particles are produced which subsequently hadronize and decay. Many particles are produced, and the resulting light signal is the combined signal from all these particles. The typical length scale on which the particles lose their energy is below the typical spacing of DOMs in IceCube. Subsequently, the single particles cannot easily be resolved, and the entire signature of the neutrino interactions appears spherical in the detector. Such a signature is called a cascade. Since the cascades are very localized, their total light yield which is connected to the neutrino energy can be estimated very well. The direction of the initial neutrino is very hard to estimate since the cascade is spherically shaped.
If the neutrino undergoes a charged current interaction, the corresponding charged lepton is generated in the interaction. The charged lepton carries away a significant fraction of the neutrino energy and due to relativistic boost, also has almost the same direction. If the initial neutrino was a ν_e, the resulting electron loses its energy on typically a few meters and thus is covered in the light by the hadronization processes. Figure 4.8 shows an event view of a cascade event.

Double Bangs If a ν_τ neutrino undergoes a charged current interaction, it produces a τ lepton. A τ can on average travel a distance of roughly 50 m per PeV through ice [48] before it decays. The energy dependence is a result of relativistic time dilatation. The signature is a cascade at the neutrino interaction and a second cascade when the τ lepton decays. Therefore such an event topology is called a double bang. Below 20

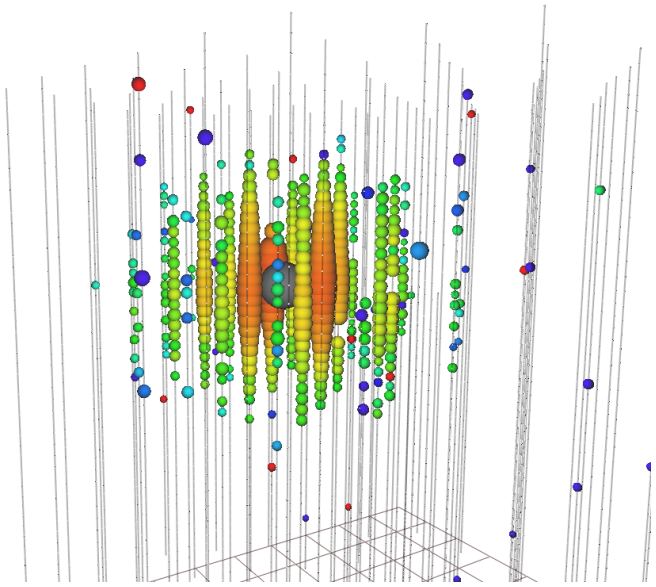

Figure 4.8.: Event view of a simulated cascade event. The spheres mark hit DOMs; the size represents the accumulated amount of light and the color the time of the first detection of light going from red to blue.

meters or correspondingly some few of 100 TeV, the two cascades cannot be separated with the resolution available in IceCube. No double bang signature has never been observed in IceCube so far [48].

Muon Tracks If a muon neutrino undergoes a CC interaction, the out-coming muons direction is very close to the direction of the initial neutrino. The average angular difference is given by

$$\Delta\phi \approx 0.7^\circ \left(\frac{E_\nu}{\mathrm{TeV}}\right)^{-0.7} \tag{4.2}$$

with $\Delta\phi$ as the angular difference between the neutrino and the muon and E_ν being the neutrino energy [49]. For energies above 600 GeV , the intrinsic angular error is thus below 1°. This sets the lower limit how accurate a muon neutrino direction can be reconstructed by measuring the muon. Muons can travel up to several kilometers in ice and produce Cherenkov light while they do so. In the IceCube detector, the path of a muon can be resolved. These events are called track-like events (or simply tracks) just by their signature in the detector. Figure 4.9 shows an event view of a track event.

The direction of the track can be reconstructed very well due to the long lever arm in the detector. Interactions can happen outside of the detector so that the muon enters the instrumented volume of IceCube from outside. It is also possible that the muon leaves the instrumented volume and continues its path. Thus the total interaction is not fully contained in the detector, makes the energy estimation challenging.
Charged current muon neutrino interactions are the event signature that is utilized in this thesis to perform a point source analysis. The following discussion is restricted only to this type of events since they are the one of interest for this analysis.

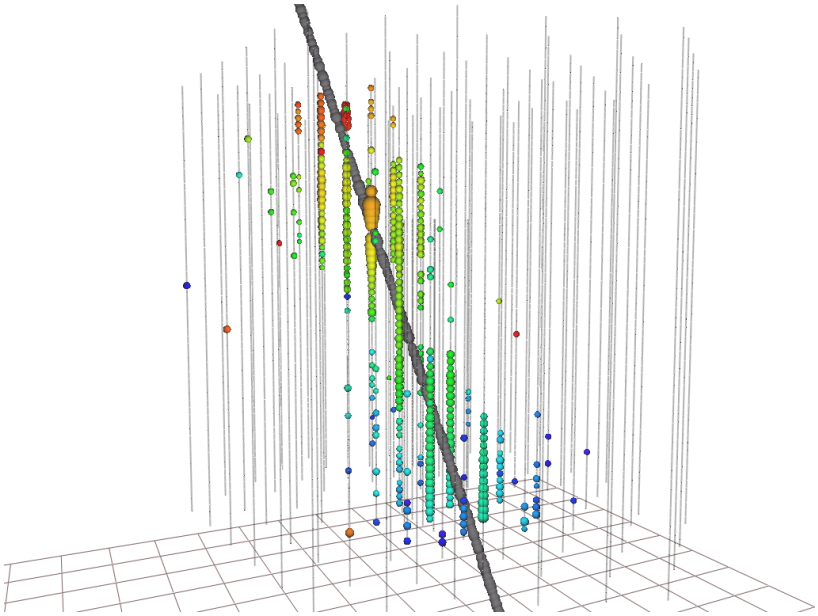

Figure 4.9.: Event view of a simulated track event. The spheres mark hit DOMs; the size represents the accumulated amount of light and the color the time of the first detection of light going from red to blue.

4.4. Background Events

The signatures of astrophysical muon neutrino μ_ν interactions are buried in several types of background processes. A key challenge in neutrino astronomy is to extract the astrophysical neutrino interactions from the background. Since the development of this event selection is a common problem in point source analysis, there is a dedicated

event selection developed inside the IceCube collaboration for point source analysis to select atmospheric muon neutrinos and suppress any background events [20]. The development of this sample was not part of this work and thus is only briefly discussed in the following sections.
The atmosphere is constantly bombarded by cosmic ray particles. The interaction of a highly energetic cosmic ray particle with a nucleus of the atmosphere produces particle showers of mainly unstable particles. Most particles decay in the upper atmosphere, leaving only a small fraction reaches the surface. The reaction is similar to hadronuclear interactions of cosmic rays with the difference that the atmosphere is typically much thicker than typical astrophysical cosmic ray sources, see section 2.3. The largest population of particles which reach the Earth surface (and eventually can also travel several kilometers below the surface and reach IceCube) are neutrinos and muons. Thus the main background events in IceCube are atmospheric muon and neutrino-induced events. Both are discussed in the following. Figure 4.10 shows a sketch of an atmospheric particle shower.

4.4.1. Atmospheric Muons

Muons are unstable particles with a lifetime of about $2.2\,\mu$s [9]. Mouns only reach earth and the IceCube detector if they do not decay before. If they are highly energetic, they benefit from relativistic time dilatation $\Delta t' = \gamma \Delta t$ with $\gamma = 1/\sqrt{1-v^2/c^2}$ and survive the journey from their production side at the height of about 20 to 30 kilometers to the surface and even further. Thus, the atmospheric muons reaching IceCube are typically high energetic. If such an atmospheric muons reach the detector, it mimics the signature of a muon produced in a muon neutrino interaction. The two are not distinguishable since they are both just muons.
To suppress the atmospheric muon background, there are several strategies in place depending on the direction of the muons. The most energetic muons can reach the surface and also travel few kilometers below the surface, but not hundreds of kilometers. This fact is used in IceCube as a directional cut: Muons that are reconstructed to originate from below the detector must have traveled through the entire earth to reach the detector from this direction. The probability is neglectable for muons to do this. Neutrinos, on the other hand, are capable of traveling through the Earth without interaction and then undergo a CC interaction close to the IceCube detector.

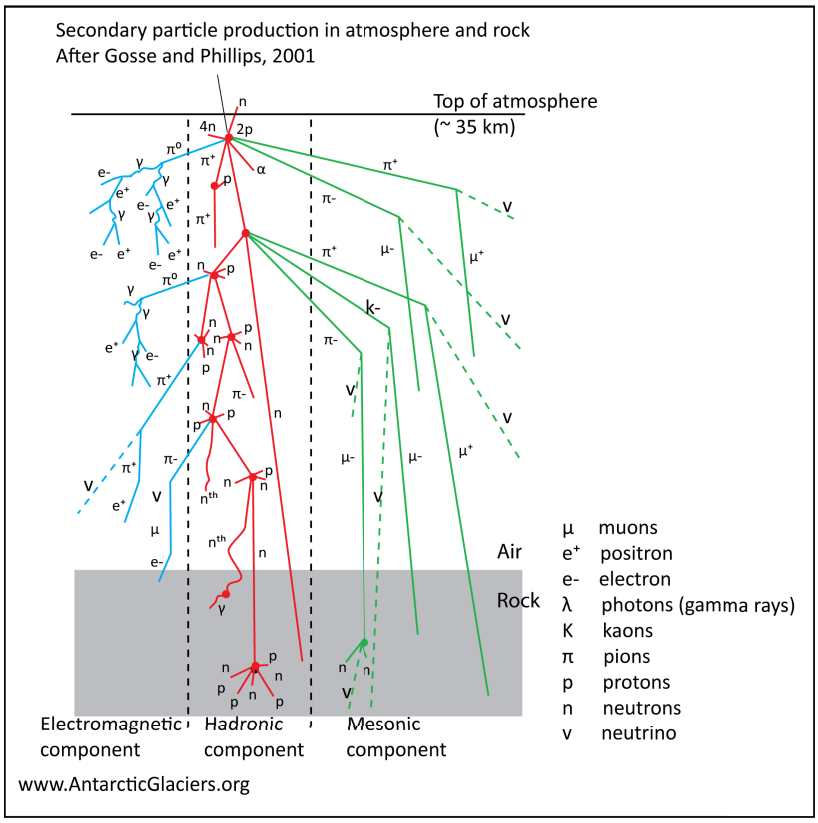

Figure 4.10.: Atmospheric particle shower with the electromagnetic, hadronic and mesonic components. Figure is taken from `http://www.antarcticglaciers.org/glacial-geology/dating-glacial-sediments-2/cosmic-rays/`.

In this interaction, a muon is produced which then enters the detector from below (or is produced inside of the detector).

The result is a strong background suppression as a function of material that has to be passed on the way to the detector and thus as a function of zenith angle θ of the event. For muons originating below the horizon, the chance of atmospheric origin is negligible. The detector is split into an up-going region (direction coming from below of the detector) and down-going region (events originating above the horizon). The up-going region is essentially free of atmospheric muons, whereas in the down-going region they dominate the sample.

The up-going region (up-going is defined in this analysis everything up above the

horizon ($\delta > 0°$) and down-going accordingly) is dominated by misreconstructed down-going events. To reduce this background, strict requirements on the quality of event reconstruction is required. These cuts remove badly reconstructed and thus miss reconstructed downgoing events from the sample. This is done utilizing a set of straight cuts and a boosted decision trees (BDT), see [20]. The remaining irreducible background consists of atmospheric neutrinos which are discussed in the next section. For the down-going region, there are many high energetic and well reconstructed atmospheric muons. Quality cuts on the reconstruction quality and energy are applied, again utilizing a BDT in the final step. Still, cosmic ray interactions in the atmosphere can produce bundles of multiple muons which are moving in parallel due to the large relativistic boosting if the initial cosmic ray particle was highly energetic. These muon bundles can mimic the signature of single, highly energetic neutrino-induced muon. It has been shown that muon bundles lose energy more constantly than single, highly energetic muons. They undergo so-called stochastic losses when losing a large amount of energy in a very short distance [50]. The difference is thus given by the smoothness of light yield along the track. This light yield smoothness is used as an additional parameter in the BDT.
The final event rate as function of zenith angle θ is shown in figure 4.11. One can see that the down-going region is dominated by atmospheric muons and the contribution of atmospheric neutrinos can be neglected. In the up-going region, the opposite is the case: Atmospheric muons are negligible, and most of the background comes from atmospheric neutrinos. Also, the expected rate of astrophysical neutrinos is larger in the up-going region.
Since the total event rate is essentially constant over all declination bands, the signal to background ratio in the up-going region is much better than in the down-going region. This will also be discussed later regarding point source sensitivity.

4.4.2. Atmospheric Neutrinos

Similar to the production of production of atmospheric muons, neutrinos are also produced in cosmic ray showers in the upper atmosphere. The process is similar to hadronuclear interaction expected at astrophysical neutrino sources (see section 2.3). The atmospheric neutrinos are, except for the highest energies, not absorbed by the earth and should thus be isotropic in the detector. There is no way to suppress the

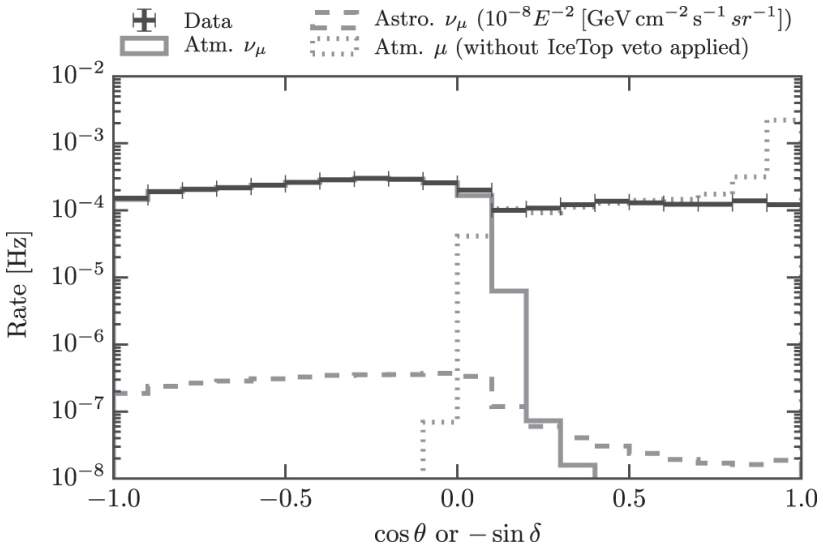

Figure 4.11.: Zenith ($\cos(\theta)$) or declination ($-\sin(\delta)$) distribution of the through-going track sample after event selection (2012 to 2015 data). Values of -1 correspond to vertically up-going events. Shown is the experimental data (black), compared to the atmospheric $\nu_\mu + \bar{\nu}_\mu$ expectation of conventional atmospheric (solid gray) and astrophysical neutrinos (dashed gray), and atmospheric muons (dotted gray) from Monte Carlo simulation [50].

atmospheric neutrino background.

Conventional Atmospheric Neutrinos In section 2.3, it was claimed that the neutrinos follow the primary cosmic ray spectrum. The difference between astrophysical neutrino source environments and the atmosphere is the thickness, as the atmosphere is several orders of magnitude thicker. Thus charged pions $\pi^\pm$ and muons $\mu^\pm$ undergo energy losses according to the Bethe-Bloch formula [51] before they decay. In the high-energy regime above 10^4 GeV, pair production, and bremsstrahlung dominate the energy losses. The energy loss can be approximated by

$$\frac{\mathrm{d}E}{\mathrm{d}x} \approx A \cdot E + B \tag{4.3}$$

where E is the energy and $A(E)$ and $B(E)$ are parameterizations of the energy loss. The energy loss increases linearly with energy (the energy dependence of A and B

can be neglected to first order). The initial cosmic ray spectrum with $E^{-2.7}$ thus transforms into an $E^{-3.7}$ energy spectrum of atmospheric neutrinos. This spectrum is much softer than the expected astrophysical neutrino spectrum (E^{-2} to $E^{-2.5}$).
Most of atmospheric neutrinos are produced with the decay of the relatively long living $\pi^{\pm}$ ($2.2 \cdot 10^{-8}$ s) and $\mu^{\pm}$ ($2.2 \cdot 10^{-6}$ s). They are called the conventional atmospheric neutrinos and make up the dominant part of the atmospheric neutrino background.

Prompt Atmospheric Neutrinos Atmospheric neutrinos can also be produced by the decay of heavier mesons, especially charmed mesons. Charmed mesons have a typical lifetime of 10^{-13} s. In contrast to the production of conventional atmospheric neutrinos, charmed mesons decay before they can lose significant fractions of their energy. Therefore, the neutrino spectrum from the decay of charmed mesons follows the initial cosmic ray spectrum with $E^{-2.7}$. Since the meson decay happens quasi-instantaneously, the neutrinos are called prompt. The expected prompt atmospheric neutrino flux is about two orders of magnitude lower than the conventional one at energies of 10^4 GeV (see [52, discussion and especially figure 3]).
Prediction of prompt atmospheric neutrino flux is strongly model dependent. The flux can be confused with a diffuse astrophysical neutrino flux since both would have a similar energy spectrum and isotropic characteristics.

4.5. Event Reconstruction

Event reconstructions are required to estimate properties of the initial neutrino from the measured photon deposition in the various DOMs of IceCube. The main parameters of interest are the direction of the origin of the neutrino, an error estimate on this direction, an energy estimate E_ν and the event time t. The event time is typical knows with the much higher precision than the other parameters since IceCube provides nanosecond timing information. There is thus no dedicated event time estimator required. Directional and energy estimation is described in the following selection.

4.5.1. Neutrino Direction Reconstruction

The reconstruction of event directions is done in several steps. The more elaborate algorithms usually need a starting point close to the actual direction, called a *seed*.

Basic algorithms usually do not need a seed, so their outcome is often used as a seed for more advanced algorithms.

Line Fit The line fit is a very basic algorithm based on the χ^2 method to fit a track to the observed photon arrival times [53]. No properties of the ice nor the Cherenkov cone are used. Instead, it is assumed that the light travels on a straight line $\mathbf{v}$ through the detector. The PMTs are located at $\mathbf{r}_i$ and are hit at times t_i. They are connected to the light track via

$$\mathbf{r}_i \approx \mathbf{r} + \mathbf{v}t_i \tag{4.4}$$

and the χ^2 is defined as

$$\chi^2 \equiv \sum_{i=1}^{N_{\text{hits}}} (\mathbf{r}_i - (\mathbf{r} + \mathbf{v}t_i))^2 \tag{4.5}$$

where N_{hits} is the total number of photon hits [53]. The algorithm allows multiple hits on a single DOM. The χ^2-problem can be solved analytically. Still, this algorithm is not very robust against noise hits and does also not take into account any knowledge about the ice. The line fit is used as a fast first guess algorithm to seed more sophisticated reconstruction methods.

Likelihood Reconstruction Likelihood reconstructions use the likelihood method, an estimation technique widely used in statistics [54]. Based on a model of a physical process which provides probability density functions (PDFs), it gives an estimate on model parameters that fit best to the measured data. The maximum likelihood method is extensively discussed in chapter 8.
Given the measured data $\mathbf{x}$ which is a set of measurements of individual wave forms from the DOMs, the optimal track hypothesis $\mathbf{a}$ is found. The track $\mathbf{a} = (\vec{r}_0, t_0, \vec{v}, E_0)$ is described by anchor point $\vec{r}_0$, direction $\vec{v}$, time t_0 and energy E_0, all with respect to $\vec{r}_0$. The muon is assumed to be highly relativistic ($|\vec{v}| = c$).
The likelihood function $\mathcal{L}(\mathbf{a}|\mathbf{x})$ is defined as the product of evaluated PDFs $p(x_i|\mathbf{a})$:

$$\mathcal{L}(\mathbf{a}|\mathbf{x}) = \prod_i p(x_i|\mathbf{a}). \tag{4.6}$$

Regarding time, the interesting quantity is not the absolute time t_i, but the time difference between the measured time and the theoretical travel time of an un-scattered

photon t_{geo} [53]. The value of t_{geo} thus is the lower limit on the arrival time. The residual is given by

$$t_{\text{res}} \equiv t_i - t_{\text{geo}}. \tag{4.7}$$

A negative value of t_{res} thus corresponds to a noise photon since the photon arrived before the minimal theoretical arrival time, while a positive value $t_{\text{res}} > 0$ usually corresponds to scattered photons. The scattering probability is an essential part of the ice model and is reflected in the PDFs. The likelihood function $\mathcal{L}(\mathbf{a}|\mathbf{x})$ is then maximized with respect to the track hypothesis $\mathbf{a}$ to find the track that describe the observation $\mathbf{x}$ best. This general procedure can be performed with various PDFs trying to describe photon propagation in the detector as accurate as possible. The more advanced models typically require larger computational time to evaluate the PDFs. Simpler models are applied first to provide a good seed value for the more time consuming algorithms.

Single PhotoElectron (SPE) Likelihood Fit The SPE likelihood is built from PDFs $p_1(\mathbf{a}|t_{\text{res}})$ which describe the arrival time of individual photons in the DOMs [53]. Thus signal from a single DOM can contribute several times to the likelihood function if it is hit several times. For all photon hits, we have:

$$\mathcal{L}_{\text{SPE}}(\mathbf{a}|\mathbf{x}) = \prod_{i}^{N_{\text{hits}}} p_1(x_{\text{res},i}|\mathbf{a}). \tag{4.8}$$

The PDFs which describe the time residuals are given by the Pandel function [53], an analytic approximation:

$$p(t_{\text{res}}) \equiv \frac{1}{N(d)} \frac{\tau^{-(d/\lambda)} \cdot t_{\text{res}}^{(d/\lambda-1)}}{\Gamma(d/\lambda)} \cdot e^{\left(t_{\text{res}} \cdot \left(\frac{1}{\tau} + \frac{t_{\text{medium}}}{\lambda_a} + \frac{d}{\lambda_a}\right)\right)} \tag{4.9}$$

$$N(d) = e^{-/\lambda_a} \cdot \left(1 + \frac{\tau \cdot c_{\text{medium}}}{\lambda_a}\right)^{-d/\lambda} \tag{4.10}$$

where $c_{\text{medium}} = c/n$ is the speed of light in ice, λ_a is the absorption length, $\Gamma(d/\lambda)$ the Gamma function and $D(d)$ is a normalization factor. λ and τ are determined by Monte Carlo simulations of the photon propagation in ice. The main advantage of the Pandel function is that it is very fast to evaluate (see the discussion in [53, section

3.2.1]). Thus the likelihood maximization becomes very fast as well.

Spline MPE Fit The SPE fit can be improved in several ways. Moving away from PDFs for isolated individual photons, one can study the time distribution of all arriving photons. The first of the N photos which hit a certain DOM will arrive at a certain time t_{res} with a probability

$$p_N^1(t_{\text{res}}) = N \cdot p_1(t_{\text{res}}) \cdot \left(\int_{t_{\text{res}}}^{\infty} p_1(t) \mathrm{d}t \right)^{N-1} = N \cdot p_1(t_{\text{res}}) \cdot (1 - P_1(t_{\text{res}}))^{N-1} \quad (4.11)$$

where $p_1(t_{\text{res}})$ is the SPE PDF and $P_1(t_{\text{res}})$ is its cumulative. In the MPE (multi photo electrode) likelihood fit equation 4.11 is used as PDF for the first photon hit in each DOM. The total number of photon hits is also taken into account, but not their individual arrival times. The MPE likelihood function thus can be written as

$$\mathcal{L}_{\text{MPE}} = \prod_{i}^{N_{\text{Ch}}} p_N^1(t_{\text{res}}|\mathbf{a}) \quad (4.12)$$

where N_{Ch} is the number of DOMs in the detector.
The likelihood estimator can further be improved by using a more detailed description of the light propagation in the ice. The Pandel function used in the SPE fit is fast to evaluate and can describe MonteCarlo simulation of photon propagation, but the description is not perfect [53]. The Pandel function assume a homogeneous and isotropic ice, which is not actually the case for IceCube, see figure 4.7. Thus full Monte Carlo simulation of photon propagation in the ice have been interpolated by spline functions [55]. Using these spline functions provide a significant improvement compared to the Pandel functions in terms of reconstruction accuracy. This leads to the final directional reconstruction method, the Spline MPE reconstruction that is utilized in this thesis. If not stated explicitly, all reconstructions are spline MPE fits for the rest of this work.

Uncertainty on Directional Estimate The error estimate of the angular reconstruction is an essential ingredient for the point source analysis (see chapter 8). Note that the reconstruction error only estimates the error on the muon direction reconstruction, the error between muon and muon neutrino μ_ν (equation 4.2) is not taken into

account by this estimator. The error is computed on an event by event basis and is used as an individual weighting of events in the analysis. The angular error is called σ. Several error estimators are present in IceCube. This thesis uses the *Paraboloid error* estimator [56]. It is estimated from the likelihood landscape of the Spline MPE. When the minimum of the likelihood function $-\mathcal{L}$ has been found, a two dimensional paraboloid in the sub-manifold of the event direction is fitted around the likelihood minimum to estimate how sharp the minimum is. The one-sigma standard error is located at the contour where the likelihood function has dropped by a factor of a half [54].

$$-\log \mathcal{L}_\sigma = -\log \mathcal{L}_{\max} + \frac{1}{2}, \tag{4.13}$$

see figure 4.12. With σ_1 and σ_2 as the major and minor axis of the ellipsis, the combined error estimate is

$$\sigma = \sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}}. \tag{4.14}$$

The σ is then used as the angular reconstruction error.

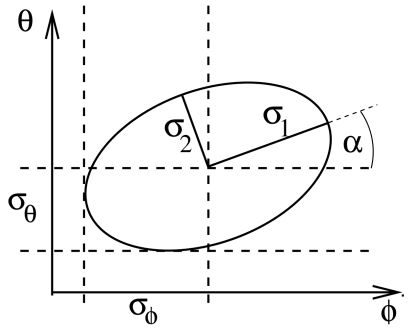

Figure 4.12.: The one-sigma paraboloid fit in the azimuth zenith parameter plane of the likelihood. This fit is used as error estimator on the angular reconstruction. The two main axis of the paraboloid are plotted, figure taken by [56].

4.5.2. Energy Estimator

The neutrino energy cannot be measured directly, so instead, we measure the energy of particles which were created in the neutrino interaction. In the optimal case, all secondary particles are produced and decay inside of the instrumented volume of Ice-

Cube, so that the total deposited energy in the detector can be measured and used as a lower limit on the neutrino energy. In case of muon neutrino CC interactions where tracks are produced, the production vertex lies typically outside of the instrumented volume or, in the opposite case where the neutrino interaction happens inside the detector, the muon leaves the detector. Of course, it is also possible that both production outside the detector and leaving the instrumented volume occurs. In any of these scenarios, not even the total energy of the secondary particles is deposited inside the detector, but the measured energy can still be used as a lower limit.
For large energies, the differential energy loss of a muon is linear to its energy

$$\frac{\mathrm{d}E_\mu}{\mathrm{d}x} \propto E_\mu + \text{const} \tag{4.15}$$

and thus measuring the differential energy loss can be used to estimate the muon energy. Furthermore, the rate of produces Cherenkov photons is also proportional to $\mathrm{d}E_\mu/\mathrm{d}x$, so the measured number of photons has a direct proportionality to the muon energy E_μ [57].
The photon density close to the track is proportional to $1/r$ and for larger distances can be described by a random walk where the photon density is given by $\frac{1}{\sqrt{r}}e^{-r\lambda_\mathrm{p}}$ with $\lambda_\mathrm{p} = \sqrt{\frac{\lambda_\mathrm{e}\lambda_\mathrm{a}}{3}}$ and λ_e as the effective scattering length and λ_a the absorption length [58]. Both expressions can be combined by

$$\mu(r) = l_0 A \cdot \frac{1}{2\pi \sin\theta_\mathrm{C}} e^{-r/\lambda_\mathrm{p}} \frac{1}{\lambda_\mathrm{C} r \tanh(r/\lambda_\mathrm{C})} \tag{4.16}$$

where

$$\sqrt{\lambda_\mu} \equiv \frac{\lambda_\mathrm{C}}{\sin\theta_\mathrm{C}} \sqrt{\frac{2}{\pi\lambda_\mathrm{p}}} \tag{4.17}$$

and

$$\lambda_\mathrm{C} \equiv \frac{\lambda_\mathrm{e}}{3} e^{\lambda_\mathrm{e}/\lambda_\mathrm{a}} \tag{4.18}$$

where l_0 is the number of emitted photons per distance, A is the size of effective photo-collective area of the sensor and θ_C is the Cherenkov angle [59]. Assuming that the light yield is proportional to energy, the expectation value of detected photons is given by

$$\lambda = \mu(E_\mu) + \lambda_\mathrm{BG} \tag{4.19}$$

whith

$$\mu(E_m u) = \frac{E_\mu}{E_{\text{ref}}} \cdot \mu(r) \tag{4.20}$$

where E_{ref} is the reference energy for which equation 4.16 was evaluated and λ_{BG} is the rate of background photons. Assuming that the number of photons is given by a Poisson distribution

$$p(k|\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}, \tag{4.21}$$

it can be used as a PDF. This leads then to the definition of the likelihood function

$$\mathcal{L}(E_\mu) = \prod_{i=1}^{N_{\text{Ch}}} \frac{(\lambda_i(E_\mu))^k}{k!} e^{-\lambda_i(E_\mu)}$$

which is used as the energy estimator and called *MuE* (short for muon energy). The expression above does assume a constant and continuous light yield. In reality, highly energetic muons tend to undergo so called stochastic energy losses, with a lot of energy lost over a short distance (see also section 4.4). This stochastic energy loss is added by a probability distribution for the light yield λ [59]:

$$G_\mu(\lambda) = \frac{\text{const.}}{\lambda} \left(e^{-wy} + (y/\sigma)^2\right)^{-1}$$

where $y \equiv \ln \frac{\lambda}{\mu}$ and w is the *skewness* parameter to describe large over fluctuation by processes such as bremsstrahlungs-losses. The energy reconstruction utilizing this extension G is called *MuEx* in IceCube and is also the energy reconstruction which is utilized in this work. Figure 4.13 shows true and reconstructed neutrino energy for the IC86-I data set.

4.5.3. Pull Correction

The angular error of the directional reconstruction has been estimated from the error ellipse of the likelihood function, see section 4.5.1. The circularized average σ of the two main axes of the error ellipse was computed. For a point source analysis, the median error between the true and the reconstructed direction σ_{true} is relevant.
The error estimator (section 4.5.1) and the true error, the distance between true muon

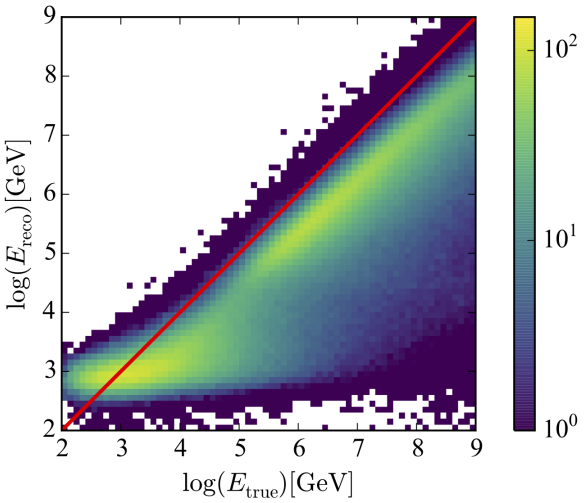

Figure 4.13.: Reconstructed energy versus true neutrino energy for the IC86-I data set, simulated events. For most cases, the energy estimator underestimates the true neutrino energy (below the red line). In some cases, the extrapolation of the nergy losses leads to an overstimation of the energy losses. The color scale is in arbitrary units of flux.

direction and reconstructed muon direction are connected by the pull:

$$P = \frac{\sigma_{\text{true}}}{\sigma_{\text{reco}}}.$$

It turns out that the ratio is not unity, but instead a function of energy. To correct for this effect, the so-called *pull correction* is applied. Based on Monte Carlo simulation weighted with an E^{-2} spectrum to simulate a potential signal spectrum, the median pull is computed as a function of energy. This pull ratio is then interpolated with a spline function, and this spline function is used to correct the estimated error σ_{reco} also for experimental data. Note that the assumed spectrum has no strong impact since the correction is performed in energy bins.
Counterintuitively, the pull is not corrected to be unity, but to be ≈ 1.177. In one dimension, the one σ range of a normal distribution contains about 68%. At higher dimensions, this value decreases. In two dimensions, the one-sigma range only contains about 40% [56]. When correcting the pull to 1.177, the one-sigma area then contains 50% of the distribution. The pull correction is shown in figures 4.14 and 4.15 for a

Monte Carlo data set, see the caption.

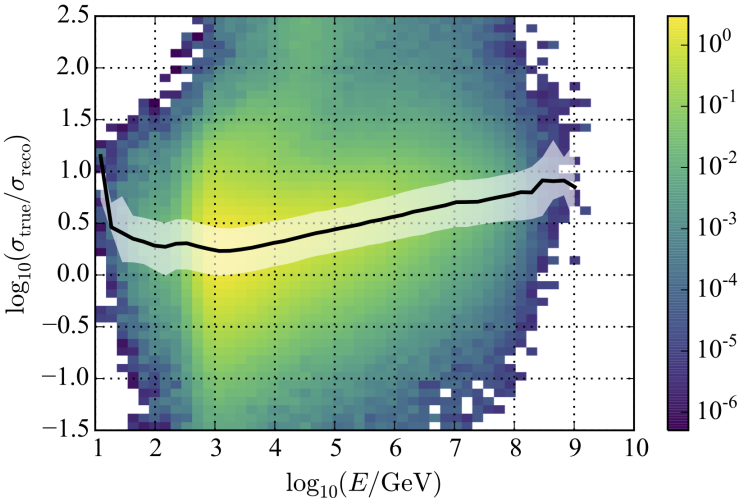

Figure 4.14.: Visualization of the pull correction with the IC86-II data set for the down-going region. The histogram shows the distribution of pulls for an E^{-2} spectrum. The weights are in arbitrary units. The black line shows the median of the pulls and the white area the central 50% of the distribution.

4.6. Utilized Muon Track Datasets

This analysis uses seven years of IceCube data. The data was collected with different detector configurations because the detector was still under construction when data-taking began. Construction and larger maintenance at the South Pole are only possible during the Antarctic summer because of rough weather conditions and lack of daylight otherwise. IceCube data is taken in *seasons*, the periods between two main maintenance sessions. During a season, the detector is unchanged. The different seasons were initially named by the number of operating strings and subsequently by counting the seasons of operation in this specific configuration. An overview of the different seasons is given in table 4.1.

Livetime is defined as the time during which the detector was fully operational. Part of the data was excluded because of unstable data taking, detector off-time or other

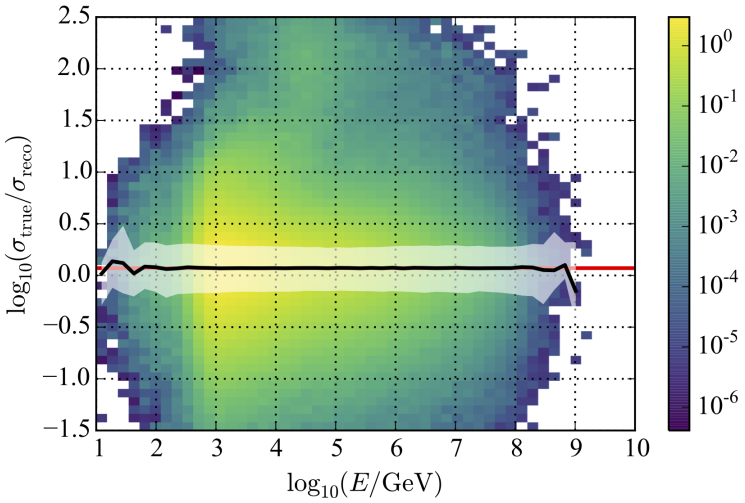

Figure 4.15.: Visualization of the pull correction with the IC86-II data set for the down-going region. The histogram shows the distribution of pulls for an E^{-2} spectrum after pull correction. The red line guides the eye for the value of $\log_{10}(1.177)$.

technical problems. The evolution from trigger level to the final analysis sample involves several levels of filtering and reconstruction. Since the development of the final data sample was not part of this work, the steps of generating the final sample are not discussed. The data sample was generated by Stefan Coenders as part of his Ph.D. thesis [21, chapter 6] to search for static point sources with IceCube [20]. The difference in requirements for datasets for his time integrated search and for this time-dependent analysis on time scales of several to hundreds of days is minimal. Thus an additional optimization is not needed.

The data from trigger level is first searched for track-like signatures. On these track like events, line fit reconstructions and simple likelihood reconstructions are performed (see section 4.5.1). Straight cuts on the quality of these reconstructions are used to remove bad reconstructed events. Finally, a machine learning algorithm (a boosted decision tree) is used to create an almost pure neutrino sample. The BDT was optimized for the relatively hard power-law spectra expected from astrophysical neutrino sources.

The effective area is a quantity to describe the efficiency of an event selection. It

Season	Start	End	Livetime [d]	Events
IC40	2008-04-05	2009-05-13	375.5	36900
IC59	2009-05-13	2010-05-31	348.1	107569
IC79	2010-05-31	2011-05-13	315.5	93842
IC86-I	2011-05-13	2012-05-15	332.6	138324
IC86-II	2012-05-15	2013-05-02	330.4	105300
IC86-III	2013-05-02	2014-05-06	360.0	114834
IC86-IV	2014-05-06	2015-05-18	367.2	118468

Table 4.1.: Overview over properties of the different IceCube seasons used in this work.

describes the size of a hypothetical area in which each neutrino would be detected, taken all detection efficiencies and cuts of the data into account. The effective area as a function of energy is illustrated in figure 4.16. The plot shows a general improvement in later IceCube seasons due to the bigger detector. It should be noted that at lower energies, the down-going region has a much smaller effective area than the up-going region. This is a result of harder energy cuts in the down-going region required to remove contamination of atmospheric muons. At about 10^6 GeV the down-going effective area becomes larger than the up-going effective area. At these energies, the Earth becomes opaque for neutrinos, and thus up-going neutrinos do not reach the detector anymore. The impact of different seasons is illustrated in figure A.2. Assuming a static source with a neutrino spectrum of E^{-2}, the relative signal strength of a source is plotted as a function of declination. The contribution is a result of both the lifetime as well as the filter efficiencies.

Distribution of Data Seven years of IceCube data is used in this analysis. Figures A.1 and 4.17 show the distribution of the experimental data in right ascension and declination. As discussed before, the right ascension distribution is flat since the background is isotropic and any geometrical acceptance effect along the hexagonal axis are averaged out over longer timescales. Because of this, the right ascension distribution is typically not considered when discussing detector effects.
The declination distribution can be divided into an up-going region ($\sin(\delta) > 0$) and a down-going region ($\sin(\delta) < 0$). Since the background is different in both regions (atmospheric muons in the down-going region, miss-reconstructed atmospheric muons in the up-going region), typically both regions are filtered separately which results

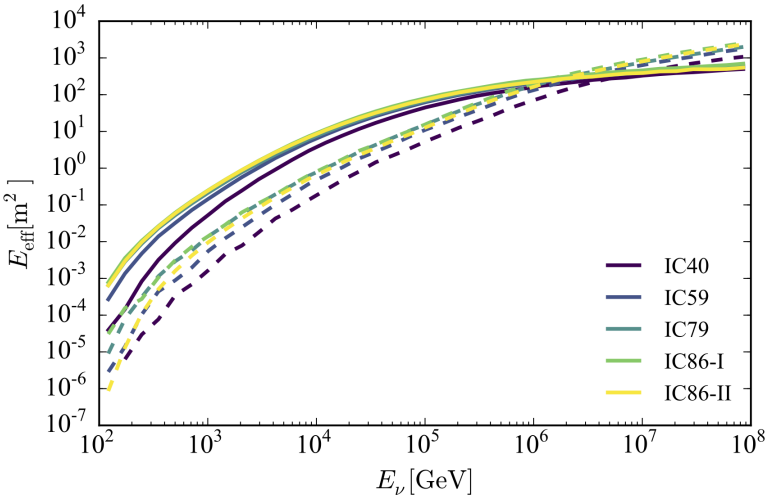

Figure 4.16.: Effective area of muon neutrinos for the different IceCube seasons. The solid lines indicate the up-going region with declination larger than zero; the dashed line indicates the down-going region with declination smaller than zero. Note that the effective area is significantly smaller than the physical size of the detector in the order of kilometers.

in different signatures. The rate in the region $\sin(\delta) \to 1$ decreases because below the detector there is the rock where muons undergo larger energy losses than in ice. The effective detector volume is larger in the horizontal direction than in vertical direction. Figure 4.18 shows the pull corrected reconstruction errors concerning the initial neutrino direction as a function of energy is shown. The reconstruction error decreases with energy since more photons are deposited in the detector, and thus more information is available. Above an energy of 1 TeV, the reconstruction error is below 1°.

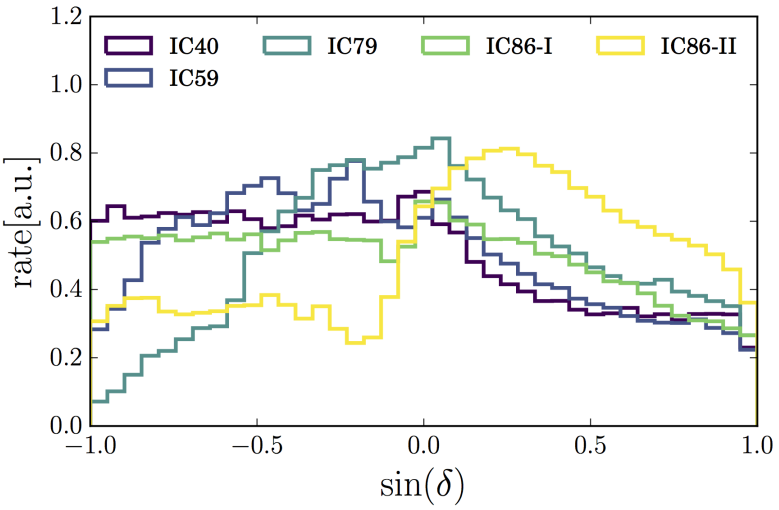

Figure 4.17.: Declination distribution of experimental data for all used seasons.

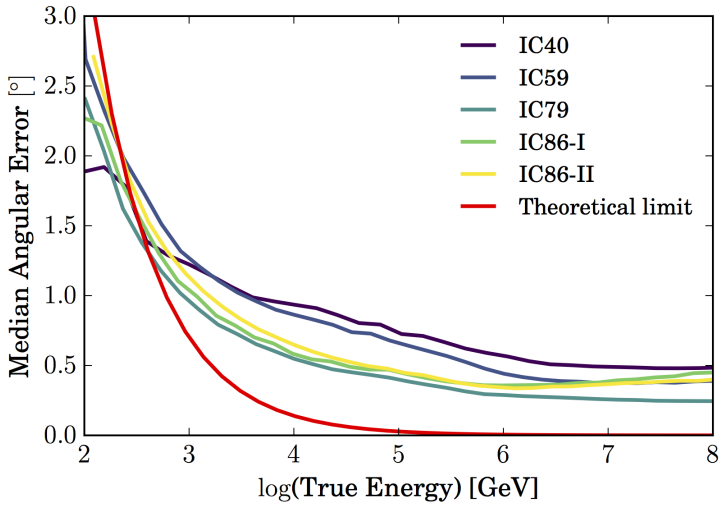

Figure 4.18.: Median reconstruction error of different IceCube seasons as a function of energy. The red line indicates the average angle between muon and neutrino and is the lower limit of the possible error.

5. Realtime Multi Messenger Astronomy

The idea of Multi-Messenger Astronomy is to combine data from different messenger particles like photons, neutrinos, gravitational waves or charged messenger particles. The combination can lead to a deeper understanding of the physical mechanisms in the object of interest [6]. It can also lower the detection threshold or likewise increase the significance by observation in different detection channels.
When performing multi-messenger astronomy on transient objects, there are two main strategies. This thesis describes a multi-messenger study using archival data. A second approach is to perform a multi-messenger analysis in real-time. As part of the Ph.D. project, the technical framework for real-time analysis in IceCube was developed. This chapter will give a brief description of real-time multi-messenger astronomy and the IceCube real-time system.

Archival Searches The first strategy is the archival search in multi-messenger data. Recorded data sets of different type of messenger particles are analyzed for temporal and spatial coincidence. This can be done in several ways:

1. In a blind correlation study at least two data sets are used. They are scanned for clustering between the different datasets. An experiment to perform this type of tests is the AMON network [60]. The procedure requires that both data sets have sufficient coverage in space and time. Ideally, the coverage is full sky and continuous in time. This is possible with neutrino telescopes like IceCube [8] and gravitational wave telescopes like LIGO [4] since they do not depend on the weather, day and night nor is their field of view shielded by the Earth. Also the gamma-ray telescope HAWC has some capabilities in this regard [61]. This method is fairly model-independent since it only assumes a spatial and temporal correlation between the different messengers.

2. Catalog based correlation studies use a set of detected sources from a certain

messenger channel and searches for correlations in another channel, typically an all-sky channel. A typical example would a correlation study of optically identified sources with neutrino events as done in this thesis. This study would benefit from the excellent spatial localization of the sources.

Triggered Searches in Real Time In triggered real-time searches, a telescope with a wide field of view does continuously observe the sky. It scans for clustering in its data in real-time. The clusters are typically below the significance required for a detection in this single channel. These sub-threshold clusters are then used as triggers for follow-up observatories. The follow-up observatories then start to observe the corresponding region in the sky to eventually find a counterpart. The follow-up telescopes typically have a small field of view, but sufficient to cover the error circle of the trigger. Also, their detection threshold is eventually lower, so they can make a significant detection if they are guided by the trigger. A typical setup would be a neutrino telescope like IceCube that provides triggers for optical telescopes [62] or gamma-ray telescopes [63]. The main challenges for real-time systems on the triggering side are that data analysis, generation of triggers and notification of follow-up observatories has to be done with minimum delay to cover fast transients. A part of this work was the development of parts of the IceCube real-time system as described in [64]. This chapter gives a summary of this paper as well as a highlight of the contribution to the system by the author.

5.1. The IceCube Real-Time System

The IceCube real-time system is a framework implemented at the South Pole and on a computer cluster in at the University of Madison, Wisconsin, to reconstruct, analyze and distribute the data of IceCube in real-time to follow up partners. The location of IceCube at the geographical South Pole results in two challenges: The first challenge is a limited computational power at the detector due to restrictions in power consumption at the remote place of South Pole. Thus only less time-consuming event reconstructions can be performed on the side. The most advanced reconstruction in IceCube cannot be used or only with reduced quality parameters in real-time. The second challenge is the limited bandwidth of data transfer to the North. It is impossible to transmit the raw data to a computer cluster in the North to reconstruct

it there. A cruel selection of which event information is sent has to be made.
The real-time system works in three steps:

1. Reconstruction and filtering of events at the detector at the South Pole
2. Transmission of the reconstructed events to the North
3. Clustering analysis and generation of triggers
4. Distribution of triggers to the follow-up observatories

Event Reconstruction at Pole The event reconstruction for the real-time data stream is done together with the standard event processing at Pole. Also for the standard off-line data streams, basic event and reconstructions are already performed at Pole. The events are then typically written on disk and are sent to the North later to run more advanced reconstructions and check the data quality. At this stage at Pole, an event selection is performed to select events for the real-time data stream. This is done using a boosted decision tree [65]. Compared with the offline-data stream, the offline-date stream is worse because of two main reasons:

1. In the offline-event sample, the most advanced and time-consuming event reconstructions are performed. In the real-time stream, only fast event reconstructions can be performed. Because of this, the data stream is expected to be worse regarding reconstruction quality.
2. The event selection and filtering are based on event reconstructions. The better the event reconstructions, the better the event selection can be. In offline-analyses, this is done in an iterative process, but the real-time event selection is restricted to the available reconstructions at Pole.

Nevertheless, the real-time event selection can achieve similar sensitivity as the offline sample, see figure 5.1. It is to some degree remarkable that the real-time event stream comes so close to the off-line event selection, utilizing much more time and computation power.

Transmission and Real-Time Analysis The events are then sent to the North via the IRIDIUM satellite system (`www.iridium.com`). In the North, they are received by the

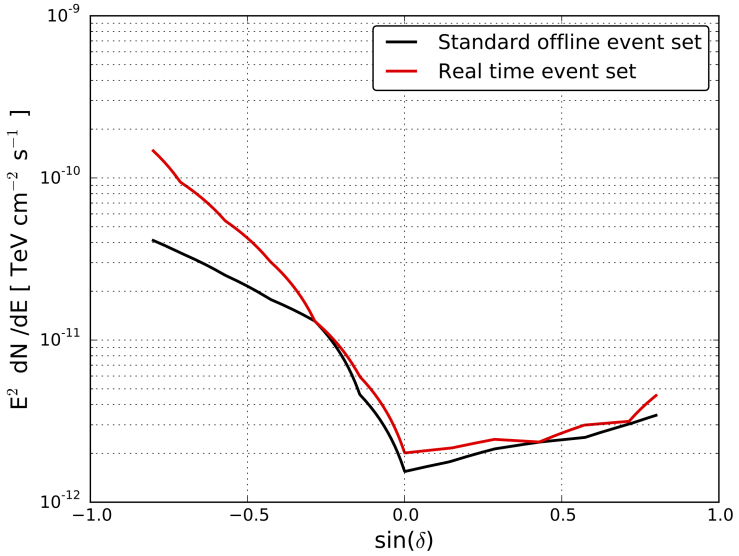

Figure 5.1.: Time independent point source sensitivities for the online data stream compared to an offline event selection, 322 days livetime. The sensitivity was studied by the author of this thesis, see [66].

IceCube computer center located at the University of Madison. The median latency between the neutrino interaction and arrival of data in Madison is about 22 seconds. The data transfer itself is realized within the IceCubeLive system, the transmission, and monitoring system of IceCube. In the North, a software system was built to receive, archive and analyze these events in real-time. Therefore it provides standard analysis tools like general stacking routines, access to the internal database of previous events and a large variety of communication tools to notify follow-up observatories. The software is designed and tested for maximal reliability and stability. The system is used in all current IceCube real-time streams as discussed in [64]. A sketch of the system layout is shown in figure 5.2. Note that the IceCube real-time system is under current improvement. This description corresponds to the status of 2016 as it was implemented by the author of this thesis.

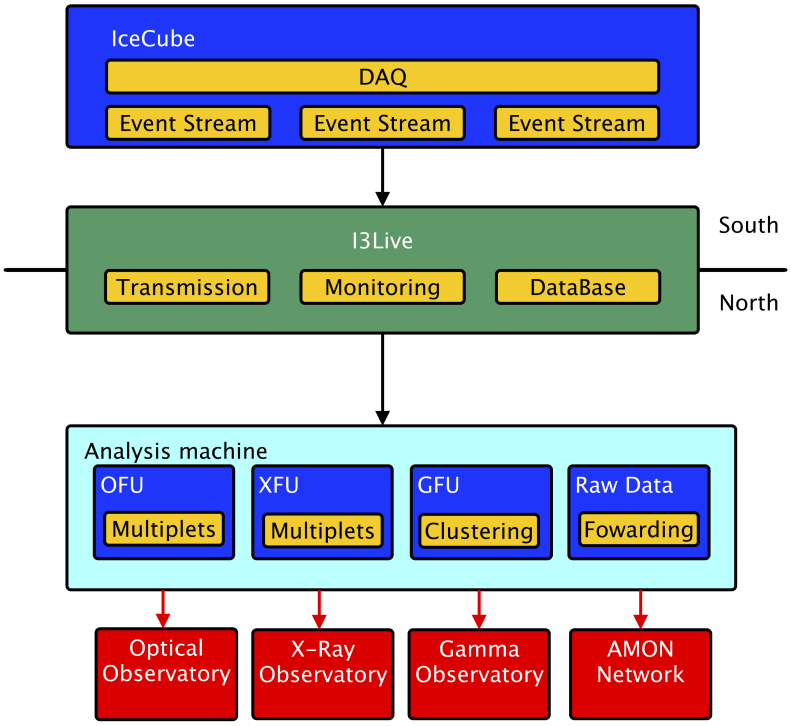

Figure 5.2.: Design of the IceCube real-time system. OFU, GFU, and XFU are different real-time analysis channels, see the discussion in [64]. The plot is taken from [66]

6. Supernova Catalog

Supernovae are one of the most prominent transient astrophysical objects. They are discovered almost exclusively in the optical regime either by chance or by dedicated transient surveys. The number of transient programs and therefore the number of detected supernovae has increased significantly in the last years. This chapter describes how the supernovae used in this thesis are compiled from various surveys and discusses the main properties of the resulting supernova catalog.

6.1. Optical Transient Surveys

Traditionally optical telescopes have a small field of view and can therefore only monitor a small fraction of the sky per night. Dedicated optical transient surveys aim to detect all transient objects in the sky as early as possible. Available instruments are small or medium-sized telescopes with a large field of view to cover a larger fraction of the sky in a single observation. Still, several observations are required to cover a significant fraction of the sky, and thus only short exposures can be performed. Combined with the usage of smaller telescopes this results in a relatively small limiting magnitude of the observation compared to classical optical astronomy using large telescopes and long exposures, like the Hubble Ultra Deep Field [67] is one of the most extreme examples.

Detection of optical transients is done via comparison to a reference image. This comparison is made using automated pipelines utilizing image subtraction and advanced image analyzing methods, see [68, 69]. Interestingly, still, the most advanced analysis pipelines normally require a human in the loop at the stage of final decision between transient and noise. Comparisons to reference observations of a certain region in the sky are required to identify transients. These reference images are either older images from the same instrument or other archival observations.

Optical transient surveys are always a compromise between observation frequency,

observation depth, and observation area. Different observation strategies are:

- Focus on the brightest (and potentially most nearby) transients only. Thus high-frequency scans (preferably every night) are performed over the entire sky with only a small optical depth. This strategy aims to ensure discovery of all transients up to a certain limiting magnitude as early as possible. An example of this strategy is the All-Sky Automated Survey for Supernovae (ASAS-SN) survey [70]. ASAS-SN utilizes commercially available $14cm$ photographic lenses and CCD cameras. It achieves a coverage of 20,000 square degrees each clear night with a limiting V-band magnitude of 17[1]. Thus ASAS-SN provides an almost complete coverage of nearby, bright transients.

- Compromise between sky coverage and survey depth. Most surveys utilize medium-size telescopes with a better limiting magnitude and smaller field of view. They have several observation modes with different scanning frequencies and depth. Often pre-selected areas in the sky are an object of intensive observation, but also unbiased scans are performed. A typical example is the Palomar Transient Facility (PTF) located at the Palomar Observatory, California. PTF uses a 1.26-meter telescope and has a field of view of 7.78 square degrees and a limiting magnitude of 21.5. It covers about 1000 square degrees every night [71].
  Another example is the Panoramic Survey Telescope And Rapid Response System (Pan-STARRS). Pan-STARRS uses a 1.8-meter telescope located at Haleakala, Hawaii. In its final stage, it will consist of four telescopes of that type. Pan-STARRS performed a 3π sky survey and different transient targeting operations. Pan-STARRS has compiled high-quality reference images for any position in the sky with a single instrument resulting in similar systematic errors for each image. Pan-STARRS has a coverage of 6000 square degrees in a single observation and has a cadence of about seven days [72].

- Some surveys focus on a relatively small part of the sky which is then monitored regularly with a small cadence. Other parts of the sky are ignored. This provides an almost complete coverage of all transients in this field at the cost of low rate

[1] `http://www.astronomy.ohio-state.edu/~assassin/index.shtml`

and minimal completion. An example for this is the OGLE program, mainly searching for microlensing induced transients [73].

An overview of current transient surveys can be found in table 1 in [71]. It is also worth mentioning the upcoming Zwicky Transient Facility which will increase the number of detected sources by about an order of magnitude [74].
After the discovery of a transient via comparison with reference images, typically a follow-up observation is performed. The object is observed in the subsequent nights to obtain an optical light curve. Optical transient surveys typically perform photometric measurements. To identify a transient, optical spectra are taken from the object at different times and eventually multi-wavelength observations, e.g., in radio, X-ray or IR can be performed. The number of detected transients normally exceeds the capabilities of follow-up observations. Identifying the most interesting follow-up candidates is a key challenge in modern wide-field optical surveys [75].

6.2. Supernovae Date Bases

Discoveries of supernovae and other transient objects are usually made public via Astronomer's Telegrams (ATels [76]). ATels are short web-based notifications of transient objects to alert the community and to eventually trigger follow-up observations. In many cases, the triggering survey has more information, for example, the full light curve, which is not enclosed in the original ATel. Several catalogs try to collect all publically available information about reported supernovae including first discovery and position, but also light curves and spectra. One of the most complete supernova catalogs is the WiseRep catalog [77]. WiseReo is operated by members of the PTF team, making it a primary source for PTF discoveries. There is also to mention the recently upcoming Open Supernova Catalog [78] which tries to merge all publically available sources. Many surveys or follow-up facilities still keep part of their data private, therefore private communication is often required to access the full data set. The supernovae catalogs in this works are mainly based on the Open Supernova Catalog [78] and is cross-referenced with the WiseRep catalog [77] as well as private information from PTF, OGLE, Nearby Supernova Factory (SNF) and selected publications on few or single supernovae.

6.3. Catalog Merging

The final supernova catalog for this analysis is compiled in several steps from the previously discussed sources:

Catalog Sources Supernovae from the WiseRep catalog [77], ASAS-SN survey data [70] and the publically available Open Supernova Catalog [78] are queried and merged to the same format. Various formats are used in the astronomy community. In this work, we adopt the modified Julian day (MJD) format for time and equatorial coordinates in J2000 format for the position. The WiseRep catalog should be fully included in the Open Supernova Catalog which is meant to be the successor. Nevertheless, additional information not present in the full public Open Supernova Catalog is available in the WiseRep catalog. ASAS-SN discoveries should also be included in the Open Supernova Catalog via ATels, but in several cases, only parts of the original information are included. The strategy to generate the catalog is to use primary sources if possible. Thus the WiseRep catalog is included as the primary source for PTF.

Catalog Merging The different catalogs are merged and tested for double counts. Supernovae can be named several times, first by the discoverer, e.g., the survey and second by the International Astronomical Union (IAU) which uses a naming scheme like SN2016a for the first supernova in 2016. Sometimes a supernova is also discovered by different observers at different times and reported independently. Associations are not always performed, especially in case of delayed publication. Doubled entries in the catalog are removed by first scanning for the same name and then for catalog entries with a reported angular distance smaller than 0.1° and a difference between the detection dates of less than 50 days. If two entries fulfill this criterion, the entries are merged. Of the two positions are slightly offset, the average is used in this work. The first observation is assumed as the better estimate of the grue explosion time. If distance and redshift are both available and distance has been evaluated redshift-independent, the direct measurement of distance is taken. If only redshifts were observed, the larger value is taken to be conservative. All doubling with a different name but very similar spatial and temporal position have been scanned manually. The reported differences in redshift were always $< 10\%$.

6.4. Distance Estimator

The distance d of a supernova is an important quantity since the flux from the source is expected to decrease with $\propto 1/d^2$. The impact in the neutrino search of a source thus directly depends on distance. Distance is not uniquely defined in astrophysics, see the full discussion in chapter 7. The important quantity for the optical brightness is the luminosity distance $D_{\mathrm{L}}(z)$ which is a function of the cosmological redshift and the cosmological model. The redshift z is defined as the observed shift of the electromagnetic spectrum to larger wavelength

$$z = \frac{\Delta\lambda}{\lambda_0} = \frac{\lambda_{\mathrm{obs}}}{\lambda_0} - 1$$

where λ_0 is the wavelength in the lab system and λ_{obs} is the observed wavelength. Redshift is measured by comparing patterns in the optical spectrum of certain observed atomic lines with laboratory measurements. Redshift (and also blue-shift) of astrophysical objects can be caused by several reasons: The peculiar motion of the source or the observer can cause the relativistic Doppler effect. Stars have a certain movement within their host galaxies; the galaxies move, and also entire galaxy clusters show movement. This results in a redshift due to peculiar motion z_{pec}. Assuming an average peculiar velocity of $v_{\mathrm{pec}} = 300\frac{km}{s}$, see discussion in [79], the dispersion in redshift is given by

$$z_{\mathrm{pec}} = \sqrt{\frac{1+\beta}{1-\beta}} - 1 \approx 0.001 \tag{6.1}$$

with $\beta = v/c$ and c as the velocity of light. A second effect is gravitational red-shifting. It happens when a photon moves out of a gravitational potential. This effect is not considered here. The third cause of redshift is cosmological expansion. When space expands, the wave of a photon is stretched, and thus the photon is redshifted. This cosmological redshift z_{cosmo} is connected to the luminosity distance. The problem is that the observed redshift is the combination of all these effects.
For supernovae which do not have a redshift-independent distance measurement, it is necessary to estimate the distance based on the cosmological redshift. Following the derivation in [80], the cosmological redshift z_{cosmo} is then given by

$$z_{\mathrm{cosmo}} = \frac{1+z_{\mathrm{obs}}}{1-z_{\mathrm{pec}}} - 1$$

with z_{obs} as the observed redshift and z_{pec} as the redshift caused by peculiar motion. Note that the decomposition into the different contributions to the redshift is impossible since their components are not measured.
If the assumption is made that a star is moving away from the observer and thus redshifts the source, the cosmological redshift and thus the distance of the source would be under-estimated. If it is assumed that the source is moving towards the observer and is thus blue-shifted, the cosmological redshift and thus the distance would be over-estimated. To be conservative, the larger cosmological redshift is assumed, see figure 6.1. For nearby sources often direct distance measurements are available. If only redshift distance is given, the upper border of the blue shaded region is assumed for the cosmological redshift to be conservative.
In a next step, the luminosity distance $D_{\mathrm{L}}(z)$ is computed based on the estimated cosmological redshift if the distance has not been measured independent of the redshift. The luminosity distance is the ratio between intrinsic brightness of a source L and the measured flux S [81]

$$D_{\mathrm{L}} = \sqrt{\frac{L}{4\pi S}}.$$

Computation of luminosity distance from given redshift is done using the `astropy` software package [82] and the cosmological parameter measured by the Planck experiment [83] ($H_0 = 67.4\,{}^{\mathrm{km}}/_{\mathrm{Mpcs}}$, $\Omega_{\mathrm{M}} = 0.308$, $\Omega_{\lambda} = 0.691$). The distance based on the redshift is only assumed if no other distance measure is available since at short distances it is not very reliable, especially because of the previously discussed peculiar motion effect, see figure 6.1.

6.5. Catalog Properties

This section describes the final compiled catalog which is used in this analysis. The focus is on core-collapse supernovae. Thermonuclear supernovae (type Ia) were also collected during the catalog compilation, but are not used for the analysis. They are used for a cross check with the core collapse supernovae catalog at the end of this chapter.
Figure 6.2 shows the spectral class distribution of the supernovae catalog. The spectral classes are combined to match the potential scenarios for neutrino production, see chapter 3. Circumstellar medium (CSM) supernovae are expected to be connected

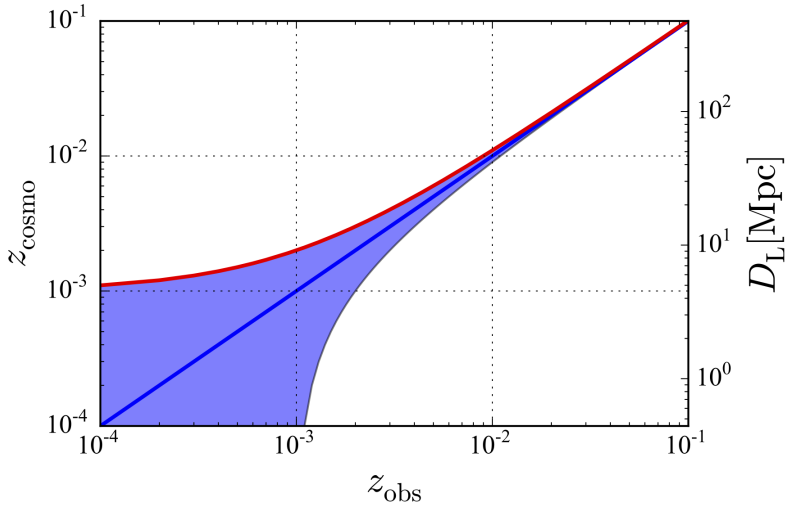

Figure 6.1.: Possible region of cosmological redshift for a given observed redshift assuming a peculiar motion of $v_{\mathrm{pec}} = 300\frac{km}{s}$. The upper red line is the conservative assumption. Note that the effect of peculiar motion is most important for nearby sources.

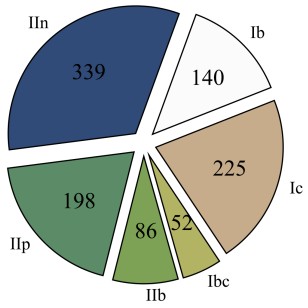

Figure 6.2.: Distribution of the spectral supernova classes.

with supernovae type IIn [37]. A connection between circumstellar medium and supernovae of type IIp is also discussed [84, 85]. The main argument is that the plateau-like light curve can eventually be explained by the CSM, since it is heated by the supernova and radiates the energy away with a different, more constant (plateau-like) time profile. The connection of CSM and type IIp supernova is much weaker than to type IIn supernova. To avoid a contamination of the type IIn scenario, type IIn and IIp

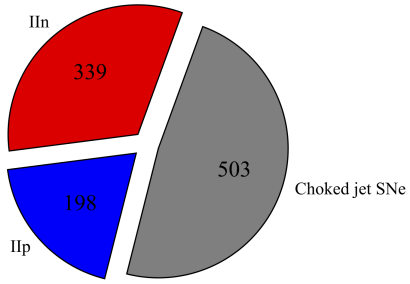

Figure 6.3.: Distribution of the final analysis sub-classes.

supernovae are handled independently in this analysis.
Choked-jet supernovae are expected to be connected with type Ic [40, 32], but also with supernovae type Ib, Ib/c and IIb [86, 32]. Since these types of supernovae are potentially all connected with the choked jet scenario, they are merged in a sub-catalog for this analysis. The final sample thus consists of three sub-catalogs:

1. Choked jet supernovae (type Ib, Ic, Ib/c and IIb)
2. Type IIn supernovae
3. Type IIp supernovae

The division of the catalog in different spectral classes is shown in figure 6.3. The spatial distribution of the catalog is shown in figures 6.4 and 6.5. For illustrative purposes, the type IIn and the type IIp sample are merged in the spatial display. The galactic plane is visible in both plots.

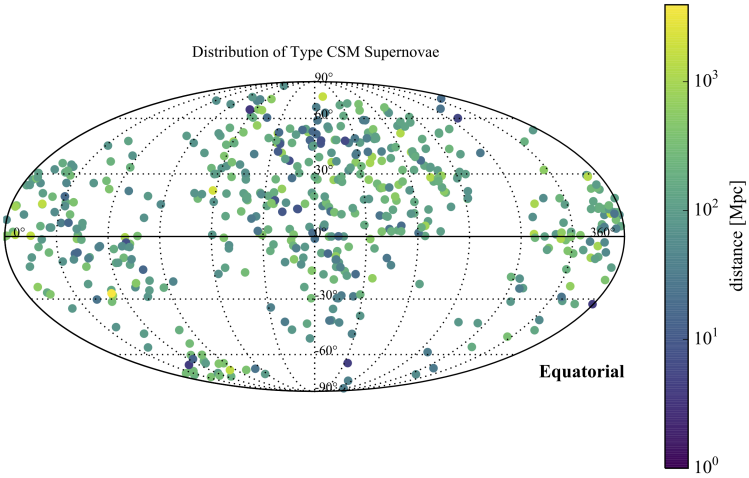

Figure 6.4.: Equatorial distribution of type IIn and type IIp supernovae. The galactic plane is visible since most optical surveys avoid the galactic plane, where background stars and dust make the detection of extragalactic SNe difficult.

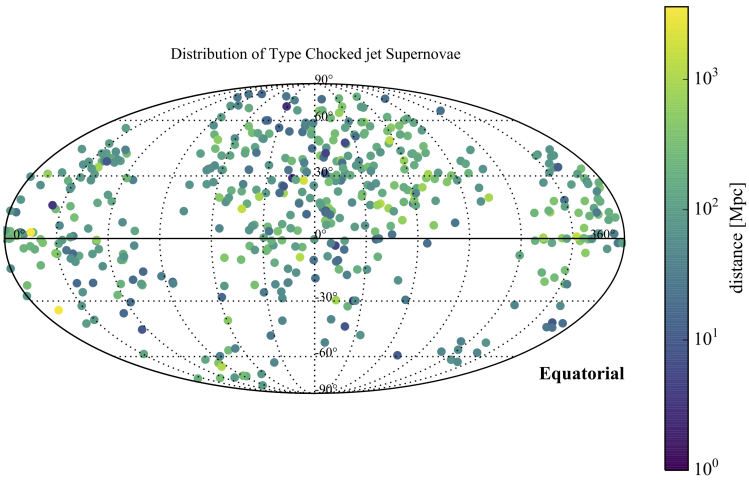

Figure 6.5.: Equatorial distribution of potential choked jet supernovae The galactic plane is visible since most optical surveys avoid the galactic plane, where background stars and dust make the detection of extragalactic SNe difficult.

6.6. Catalog Comparison to Star Formation Rate

This section describes the comparison of the supernova catalog with theoretical expectation. Therefore the overall supernova rate and the detection efficiency of a supernovae survey are evaluated.

Cosmological Supernovae Rate The rate of supernovae in a certain redshift bin depends on the volume enclosed in that redshift bin and the local SN density. The supernova density $\rho_{\mathrm{SN}}(z)$ is the result of the formation and evolution of stars during the evolution of the cosmos. The volume of the redshift bin depends on the cosmological evolution model. So the rate of supernovae is the product of comoving volume [81, 83] and the supernovae density both as a function of redshift:

$$\frac{\mathrm{d}N_{\mathrm{SN}}}{\mathrm{d}z}(z) = \rho_{\mathrm{SN}}(z) \times \frac{\mathrm{d}V}{\mathrm{d}z}(z) \tag{6.2}$$

with $\frac{\mathrm{d}V}{\mathrm{d}z}(z)$ as the differential comoving volume. The comoving volume as function of redshift is shown in figure 6.6. The rate of supernovae as a function of redshift has

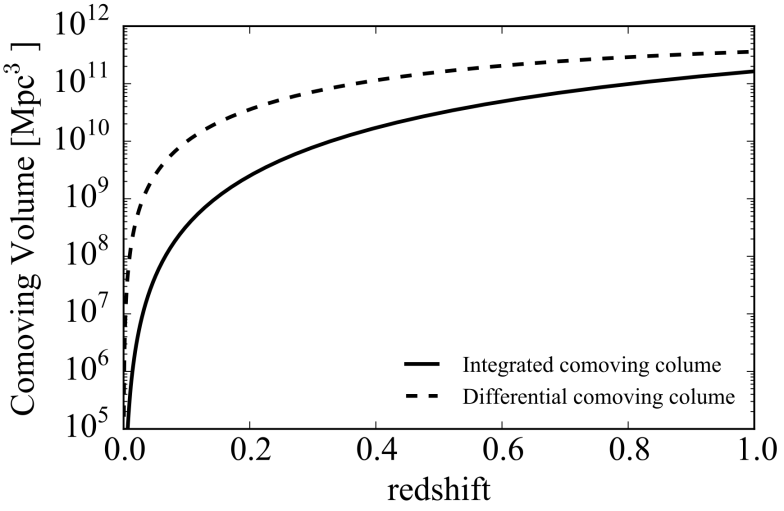

Figure 6.6.: The plot shows differential and integrated comoving volume as function of redshift, computed using `astropy` [82] and cosmological parameter from the latest Planck results [83].

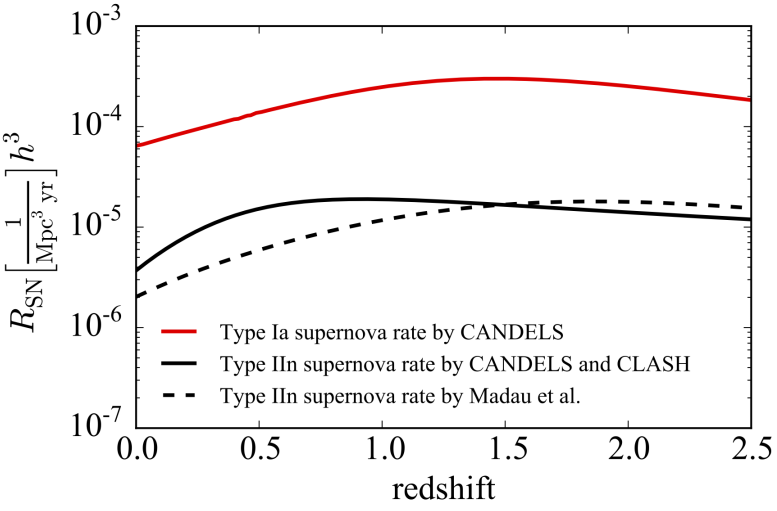

Figure 6.7.: The plot shows the redshift evolution type of type Ia and IIn supernova rate. The rate of type Ia supernova is given in [88] and the type IIn supernova rate in [87]. For comparison, the same supernova rate by Madau et al. [89] is shown. Other type of core collapse supernovae differ only by a constant factor from the type IIn supernova rate [87].

been measured by the CANDELS and CLASH supernova surveys [87, 88] as well as by Maudau et al [89]. Type Ia supernova show a different redshift dependence than core collapse supernovae (CCSN)[2]. All classes of CCSNs (Ib, Ic, Ibc, IIn and IIp) are expected to follow the same redshift dependence and their fraction of the total CCSN rate is constant [87]. The redshift evolution is shown for type Ia and type IIn supernovae in figure 6.7.

Optical Detection Chance This part discusses the chances to detect a supernova due to instrumental limitations. Combined with the total rate of supernovae, this gives a prediction for the number of detected supernovae.
Every survey has a certain limiting magnitude up to which sources can be detected. In a very idealized scenario, this limiting magnitude is constant. In that case, every object which is brighter than the limiting magnitude will be detected if it is in the field of view of the survey. The brightness of an object (apparent magnitude m) depends on

[2]The difference is also due to the fact that type Ia or thermo-nuclear supernovae require a binary companion which sets different constraints compared to CCSNs.

both its intrinsic brightness (absolute magnitude M) and the distance to the observer. The absolute magnitude is defined as the apparent magnitude at a distance of 10 pc. The distance modulus μ is defined as

$$\mu = m - M. \tag{6.3}$$

Note that the magnitude system is logarithmic in brightness and also inverted: The smaller the magnitude, the brighter is the source. The distance modulus is connected with luminosity distance D_{L} via

$$\mu = 5 \log_{10}(D_{\mathrm{L}}) - 5 \tag{6.4}$$

where D_{L} is given in units of parsec. Using equation 6.3 and 6.4, the apparent magnitude for an object at given distance can be computed.
The peak magnitude of a supernova is believed to follow a normal distribution (center and width for the different types of supernovae are given in [90]). To compute the detection chance at a certain distance, the absolute peak magnitude M is converted to the apparent magnitude m using equation 6.3. To be detected, the source apparent magnitude m_{app} must be brighter than the limiting magnitude of the survey. For a brightness distribution, detection efficiency ϵ_{det} is given by

$$\epsilon_{\mathrm{det}} = \int_{-\infty}^{M_{\mathrm{lim}}} \mathcal{N}(m, \sigma) \mathrm{d}m \tag{6.5}$$

with $\mathcal{N}(m_{\mathrm{app}}, \sigma)$ as a normal function and m and σ as center and width of the apparent brightness distribution. Thus the detection efficiency ϵ_{det} is the fraction of the brightness distribution that appears brighter than the limiting magnitude M_{lim} of the survey.
An additional effect which needs to be included is host extinction. Supernovae are typically located inside galaxies which contain gas and other absorbing material. The supernova thus appears dimmer. A host extinction correction on the apparent magnitude is applied using the parameters from [88]. Instead of performing the convolution with the full host extinction probability only the average of the distributions is used as suggested in [88]. The detection efficiencies as a function of redshift for different types of supernovae are shown in figure 6.8. A limiting magnitude of 21 for the instrument

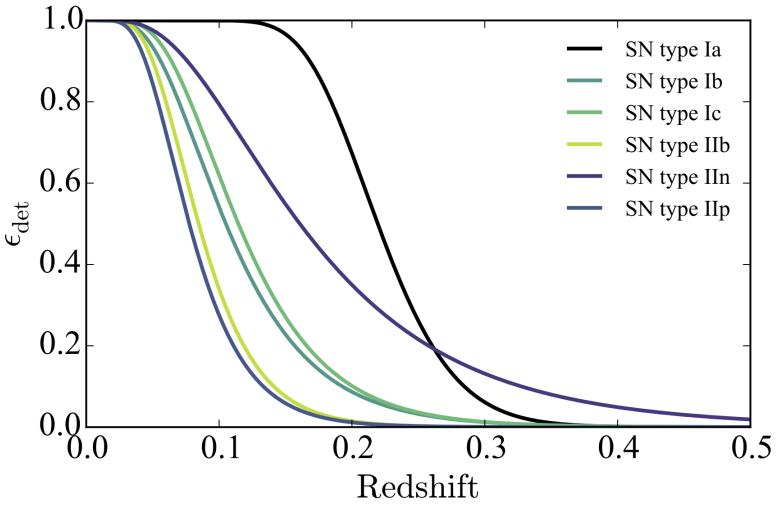

Figure 6.8.: Detection efficiency for different types of supernovae with an assumed limiting magnitude of 21 as realized in the Palomar Transient Factory [71].

is chosen. This corresponds to the properties of the Palomar Transient Facility [71]. Supernovae of type Ia are intrinsically brighter than core-collapse supernovae. They can thus be detected up to larger redshifts compared to the most core-collapse SNe. Due to their large spread in luminosity, the chance to find a type IIn supernova at large redshifts is larger than for supernovae type Ia. Nevertheless, the total rate of type Ia supernovae is significantly larger than of supernovae type IIn (see figure 6.7). Thus the majority of discovered supernovae are still of type Ia supernova at larger redshifts. The rate of the estimated supernova as a function of redshift is shown in figure 6.9. There are some limitations to the previous calculation. They are mainly caused by limitations of optical surveys:

- First of all, optical surveys are not able to cover the entire sky of 4π. The field of view, even on wide-field telescopes, is very small compared to the full sky. Thus only a small fraction of the sky can be observed at one time. Performing a full sky scan takes a considerable amount of time. Also, since most optical telescopes are located on Earth, there are certain parts of the other hemisphere of the sky which can never be observed since the direction is blocked by Earth. This can be solved by using several telescopes located in both hemispheres.

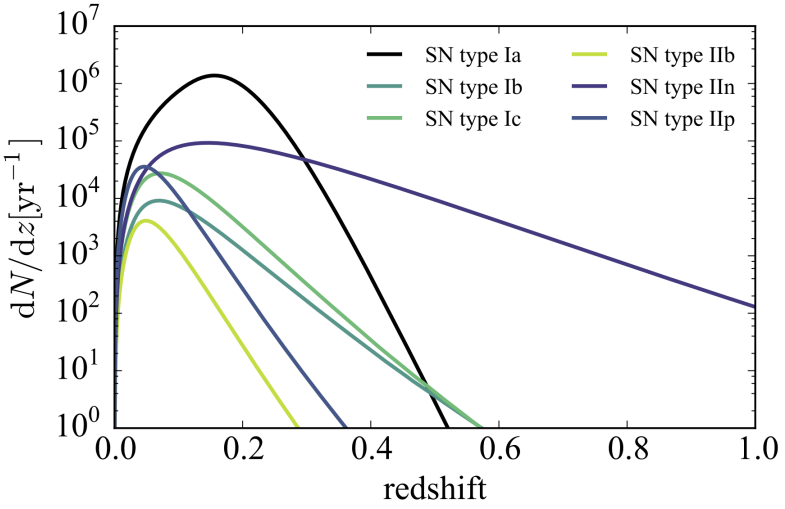

Figure 6.9.: Differential expectation rate of supernovae as a function of redshift. Detection efficiency as shown in figure 6.8 is included.

- The limiting magnitude does vary due to different weather conditions on the telescope side. The limiting magnitude is also not only limited by external conditions but a result of the observation strategy. The exposure length determines the limiting magnitude and any survey strategy is a compromise between depth and coverage of the survey.

- Supernovae are transient objects. The above computation was done using the supernova peak magnitude. If a supernova is detectable only at the peak magnitude, it requires that the telescope points exactly at the location when the supernova reaches peak magnitude. The time the supernova is detectable depends on how long it is brighter than the limiting magnitude of the telescope. Depending on the frequency of how often a certain region of the sky is visited, many supernovae will be missed.

- Selection bias can also reduce the number of detected supernovae. To identify a supernova, spectroscopic follow-up observations have to be performed to classify the object. The available spectroscopical time is limited, and selection has to be done which transient objects to follow up and which not. Since surveys are often targeted to a specific transient class, follow-up observations of other transients

might not happen and will result in a large number of unclassified transients.

These effects can cause differences between expected and observed rates of supernovae. They are hard to model, since the actual survey observation schedules, their image subtraction algorithms, weather effects and similar effects have to be taken into account. Surveys can perform this efficiency estimates for their own data, but for compiled catalogs from many different surveys and also accidental discoveries, this is basically impossible.
In this work the rate of discovered supernovae of type Ia is used to estimate this overall efficiency of the compiled core-collapse supernova catalogs caused by the previously discussed effects. The type Ia supernova sample is independent of the core collapse sample and therefore suitable for the test. The number of observed type Ia supernova is compared with the expectation, assuming a full sky coverage of 24 hours per day. The ratio as a function of redshift is then used as a correction function. This function is shown in figure 6.10. The correction factor becomes larger for larger redshift, implying that the coverage becomes worse at larger distances. The number of measured

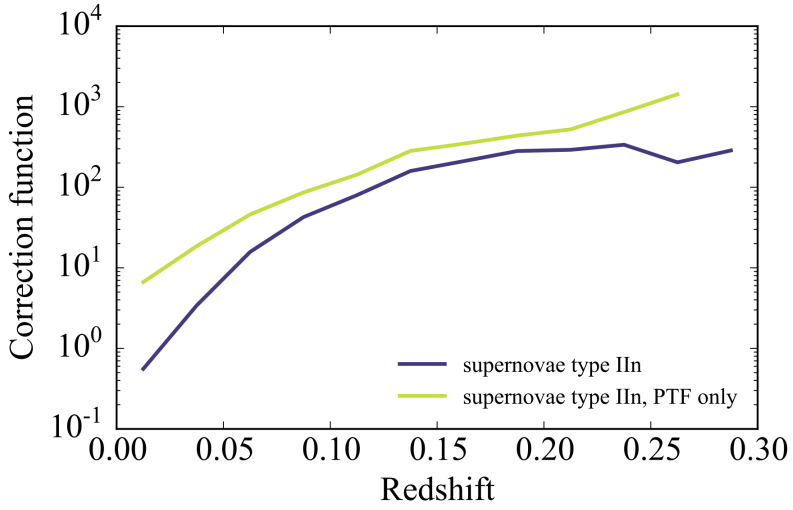

Figure 6.10.: The plot shows the correction function to account for the efficiency of the surveys, both for PTF and also for all detected supernovae combined.

type IIn supernovae and the expectation number are shown in figure 6.11. The distributions of the other types of core collapse supernova look similar. Overall, the rough

estimate can reproduce the distribution of supernovae as a function of redshift to some degree. This estimate presented here did also not aim to fully recover the properties of the surveys, but it provides a sanity check for the general assumptions made. These assumptions are later on used to compute the diffuse astrophysical neutrino flux, see chapter 7.

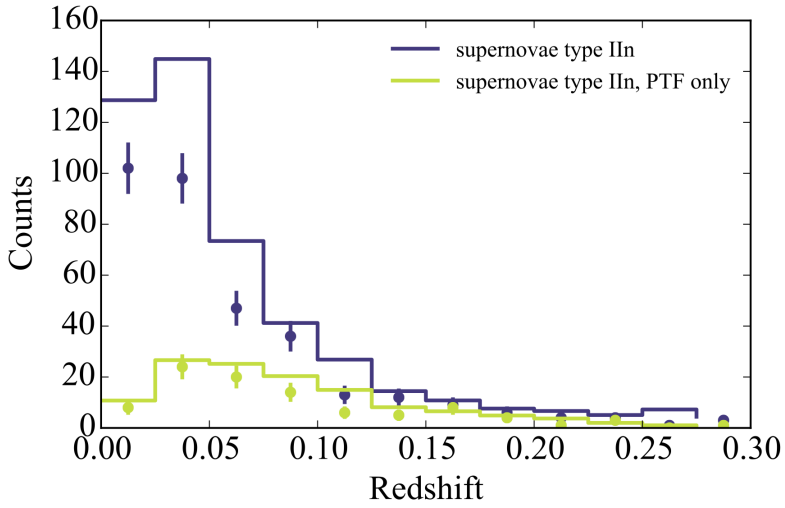

Figure 6.11.: Comparisons of type IIn supernovae in the final sample and expected number after correction of estimated efficiency. The error bars only include counting statistics, no systematic effects have been included.

7. Diffuse Neutrino Flux

The diffuse astrophysical neutrino flux is defined as the flux of neutrinos which has an astrophysical origin and appears isotropic so that no individual sources can be resolved. It is the combined flux from all neutrino sources distributed over the universe. Since each source is most likely very dim[1] and the sources are distributed isotropically; also the diffuse flux appears isotropic.
This chapter describes how the diffuse flux is computed. It is assumed that the universe is isotropically filled with a certain source population which makes up a diffuse neutrino flux. Assuming each source has the same intrinsic brightness[2] and a cosmological evolution model, the cumulative flux is computed. The cosmological parameters from the latest Planck result [83] are used in this analysis. Cosmological computation is utilizing the `astropy` software package [82].
First, it is studied how the neutrino flux from a single source with given intrinsic power spectrum appears to an observer at a cosmological distance. Then the cosmological evolution of the population of the source is discussed. In the last step, both are combined to compute the diffuse flux.

7.1. Flux from a Single Source

A transient source emits a certain amount of neutrinos over its livetime. Restricting to power-law spectra with a cut-off, the differential number spectrum is given by

$$\frac{\mathrm{d}N}{\mathrm{d}E'} = \Phi_0 \left(\frac{E'}{E_{\mathrm{ref}}}\right)^{-\gamma} \mathrm{e}^{-\frac{E'}{E_{\mathrm{cutoff}}}} \tag{7.1}$$

[1] No single neutrino point source in the high-energy regime has been discovered yet. Current lower limits on the number of neutrino sources are about 100 [20].
[2] This is called the standard candle assumption.

where Φ_0 is the the fluence normalization, E' the energy, E_{ref} the anchor for the power law energy spectrum and E_{cutoff} the energy cutoff.
If the source is observed at a certain distance r and isotropic emission of the source is assumed, the differential number of particles per area is given by

$$\frac{\mathrm{d}N}{\mathrm{d}E'\mathrm{d}A} = \Phi_0 \left(\frac{E'}{E_{\mathrm{ref}}}\right)^{-\gamma} \frac{\mathrm{e}^{-\frac{E\prime}{E_{\mathrm{cutoff}}}}}{4\pi r^2}. \tag{7.2}$$

So far, a static, Eucledian universe was assumed. But in an expanding universe, redshift dependent energy losses have to be considered: A neutrino observed with an energy E has been emitted with a larger energy $E' = (1+z)E$ where z is the redshift [81]. Taking the redshift energy loss into account, the spectrum of a source at redshift z is then given by

$$\frac{\mathrm{d}N}{\mathrm{d}E} = \Phi_0(1+z) \left(\frac{(1+z)E}{E_{\mathrm{ref}}}\right)^{-\gamma} \mathrm{e}^{-\frac{(1+z)E}{E_{\mathrm{cutoff}}}} \tag{7.3}$$

since the spectrum has to be evaluated at the energy in the source frame. The additional factor of $(1+z)$ results from the transformation from the source to the observer frame $\mathrm{d}E' \to \mathrm{d}E$. The differential number fluence is thus given by [91]

$$\frac{\mathrm{d}N}{\mathrm{d}E\mathrm{d}A} = \Phi_0(1+z) \left(\frac{(1+z)E}{E_{\mathrm{ref}}}\right)^{-\gamma} \frac{\mathrm{e}^{-\frac{(1+z)E}{E_{\mathrm{cutoff}}}}}{4\pi D_{\mathrm{p}}^2(z)} \tag{7.4}$$

where $D_{\mathrm{p}}(z)$ is the proper distance [81].

7.2. Flux from a Redshift Shell

The contribution to the total diffuse neutrino flux from a certain redshift shell $\mathrm{d}z$ is the number of sources present in the redshift shell times the neutrinos fluence of each individual source. The number of sources is given by the volume of the shell $\mathrm{d}V$ times the local source density $\rho(z)$. A constant source rate would correspond to $\rho(z) = \rho_0$, otherwise the evolution of the source density can be scaled with the local ($z = 0$) rate ρ_0 and a cosmological evolution function $f(z)$ by $\rho(z) = \rho_0 f(z)$. Thus the diffuse

differential neutrino number density $\frac{\mathrm{d}n}{\mathrm{d}E}$ can be written as

$$\frac{\mathrm{d}n}{\mathrm{d}E} = \rho(z) \times \frac{\mathrm{d}N}{\mathrm{d}E\mathrm{d}A}\mathrm{d}V. \tag{7.5}$$

Static Euclidean Universe In case of a static, Euclidean universe, the differential volume of a shell at radius r is given by

$$\mathrm{d}V(r) = 4\pi r^2 \mathrm{d}r. \tag{7.6}$$

Combining with equation 7.2 and inserting into equation 7.5, all dependencies on r cancel out

$$\frac{\mathrm{d}n}{\mathrm{d}E} = \rho_0 \times \frac{\mathrm{d}N}{\mathrm{d}E}\mathrm{d}r. \tag{7.7}$$

The contribution from each $\mathrm{d}r$ shell is equal. Assuming an infinite, static universe, the integral over it would give an infinite flux. This divergence is known as Olber's paradox [92].

Realistic Universe The previous computation can be repeated in the observed, evolving universe. The comoving volume $\mathrm{d}V_\mathrm{C}(z)$ is given by [81]

$$\begin{aligned}\frac{\mathrm{d}V_\mathrm{C}}{\mathrm{d}z} &= D_\mathrm{H}\frac{4\pi(1+z)^2 D_\mathrm{A}^2(z)}{\sqrt{\Omega_\Lambda + \Omega_k(1+z)^2 + \Omega_m(1+z)^3}} \\ &= \frac{c}{H_0}\frac{1}{(1+z)^2}\frac{4\pi D_\mathrm{L}^2(z)}{\sqrt{\Omega_\Lambda + \Omega_k(1+z)^2 + \Omega_m(1+z)^3}}\end{aligned} \tag{7.8}$$

with D_H as the Hubble distance, H_0 as the Hubble constant, $D_\mathrm{A}(z)$ as the angular distance and Ω_Λ, Ω_k and Ω_m as the cosmological parameter. The rate of sources[3] can thus be written as [91]

$$\frac{\mathrm{d}\dot{N}_\mathrm{sources}}{\mathrm{d}z} = \frac{\mathrm{d}V_\mathrm{C}}{\mathrm{d}z}\frac{\rho(z)}{(1+z)} \tag{7.9}$$

where the factor $(1+z)$ results from the redshift effect on time. The supernova rate $\rho(z)$ is observed from Earth. The rate at redshift z therefore has to be corrected by the $(1+z)$ factor to get the supernova rate in the local frame. The differential diffuse

[3] Typical transient sources are supernovae or gamma ray bursts.

flux from a redshift shell is then given by

$$\begin{aligned}\frac{\mathrm{d}\Phi(E,z)}{\mathrm{d}z} &= \frac{\mathrm{d}n}{\mathrm{d}E\mathrm{d}z} = \frac{\mathrm{d}\dot{N}_{\mathrm{sources}}}{\mathrm{d}z} \times \frac{\mathrm{d}N}{\mathrm{d}E\mathrm{d}A} \\ &= \frac{c}{H_0}\frac{1}{(1+z)^2}\frac{4\pi D_{\mathrm{L}}^2(z)}{\sqrt{\Omega_\Lambda + \Omega_k(1+z)^2 + \Omega_m(1+z)^3}}\frac{\rho(z)}{(1+z)} \\ &\times \frac{\mathrm{d}N}{\mathrm{d}E}\frac{1}{4\pi D_{\mathrm{p}}^2(z)} \\ &= \frac{c}{H_0}\frac{\mathrm{d}N}{\mathrm{d}E}\frac{\rho(z)}{(1+z)}\frac{1}{\sqrt{\Omega_\Lambda + \Omega_k(1+z)^2 + \Omega_m(1+z)^3}} \end{aligned} \tag{7.10}$$

where in the last step it was used that $D_{\mathrm{p}} = {}^{D_{\mathrm{L}}}/_{(1+z)}$. Similar results are given in the literatue by Murase [93, appendix B] and Ahlers and Halzen [94].
The total diffuse flux is given by integration over the full redshift range:

$$\begin{aligned}\Phi(E) &= \int_0^\infty \frac{\mathrm{d}\Phi(E,z)}{\mathrm{d}z}\mathrm{d}z = \int_0^\infty \frac{\mathrm{d}n}{\mathrm{d}E\mathrm{d}z}\mathrm{d}z \\ &= \frac{c}{H_0}\int_0^\infty \frac{\mathrm{d}N}{\mathrm{d}E}\frac{\rho(z)}{(1+z)}\frac{1}{\sqrt{\Omega_\Lambda + \Omega_k(1+z)^2 + \Omega_m(1+z)^3}}\mathrm{d}z. \end{aligned} \tag{7.11}$$

Comparing different cosmological Models Figure 7.1 summarizes previous discussion. As already mentioned, the static Euclidean universe leads to constant contribution from each redshift shell and thus to Olber's paradox. When including red-shifting of the energy ($E \to {}^{E}/_{(1+z)}$), the contribution of further distant sources is decreased since the energy spectrum is falling with E^{-2} and this factor translates into a $(1+z)^{-2}$ dependence of the flux.
Moving from a constant source evolution $\rho(z) = \text{const}$ to the actual observed source rate $\rho(z)$ partially compensates the suppression of large redshift and leads to a maximum of contribution to the diffuse flux at about $z = 0.4$, see figure 7.1.

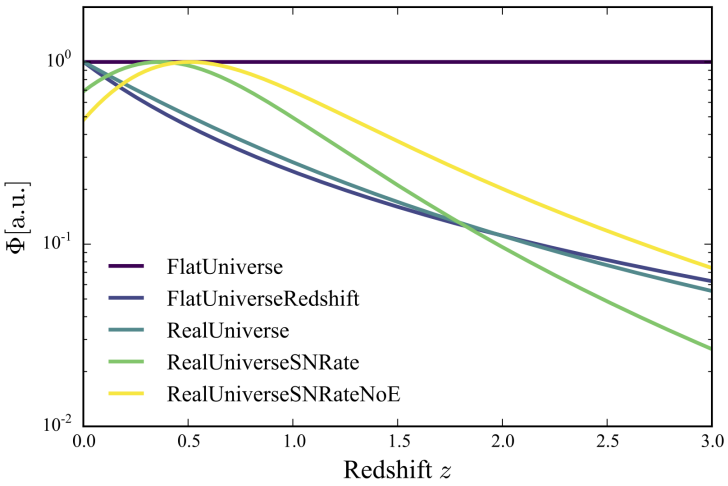

Figure 7.1.: The plot shows differential flux contribution from different models per shell for an E^{-2} spectrum at 100 GeV. All lines are normalized to their maximum. The line indicated with *FlatUniverse* describes a static, Euclidean universe with constant source density. The constant contribution of each redshift shell leads to Olber's paradox when assuming an infinite universe. The line indicated with *FlatUniverseRedshift* already accounts for the energy shift, but keeps all previous assumptions. The third line *RealUniverse* now in cooperated the current cosmological model (see equation 7.10). The line indicated with *RealUniverseSNRate* goes away from the assumption of constant source density and shifts to the measured supernovae formation rate, see also the discussion in section 3. The last line *RealUniverseSNRateNoE* is only for illustrative purposes. It switches off the redshift effect on energy to show the impact on the final computation. Comparing the two last lines, one can see the importance of the redshift on the energy.

8. Likelihood Point Source Search

This chapter describes the likelihood method which is used to perform a search for neutrinos from supernova explosions. First, an introduction to the likelihood method and the general use in IceCube point source searches is given. Then the implementation of the method is discussed. Finally, the extension to the analysis of many sources simultaneously and the multi-component fit which has never been used in point source searches before is discussed.

8.1. The Likelihood Function

The outcome of an experiment is typically a set of measured data points $\{x_1, x_2, \ldots, x_N\}$. Assuming the data originates from a probability density function (PDF) $P(x|a)$ which depends on a parameter a, the likelihood function [54] is defined as

$$\mathcal{L}(x_1, x_2, \ldots, x_N|a) = P(x_1|a)P(x_2|a)\cdots P(x_N|a) \tag{8.1}$$

$$= \prod_{i=1}^{N} P(x_i|a). \tag{8.2}$$

The better the data fits to the distribution $P(x|a)$, the larger is the value of the likelihood function $\mathcal{L}$. This is used in the maximum likelihood method to estimate the parameter a from a measurement. Instead of interpreting $\mathcal{L}$ as a function of the data $\{x_1, x_2, \ldots, x_N\}$ for given a, $\mathcal{L}$ can be interpreted as a function of a for given data $\mathcal{L}(a|x_1, x_2, \ldots, x_N)$. The likelihood function can be used to find the parameter a_0 that gives the best description of the data and thus the largest value of $\mathcal{L}$:

$$\left.\frac{\mathrm{d}\mathcal{L}(a)}{\mathrm{d}a}\right|_{a=a_0} \stackrel{!}{=} 0. \tag{8.3}$$

This equation is typically solved numerically, only in special cases an analytic solution is possible. It can be shown that the maximum likelihood method is unbiased and is efficient in the sense that it reaches the minimum variance bound [54] for large N. In this regime it is fair to call it the optimal estimator. However, for small N the maximum likelihood method can have a bias, as discussed in [54].
The likelihood method is very sensitivity to the correct PDF $P(a|x_i)$. Whether the PDF is accurate cannot be assessed from the maximum likelihood method.

8.2. Point Source Likelihood

The point source likelihood is constructed to estimate the strength of a potential point-like neutrino source covered in background events. The method is common in IceCube and used in many previous analyzes, see [95, 96] for reviews. The point source likelihood is defined as

$$\mathcal{L}(n_\mathrm{s}, \gamma) = \prod_{i=1}^{N} \left(\frac{n_\mathrm{s}}{N} \mathcal{S}(\nu_i|\gamma) + \left(1 - \frac{n_\mathrm{s}}{N}\right) \mathcal{B}(\nu_i) \right) \tag{8.4}$$

where the product i runs over all N neutrino events in the data set. $\mathcal{S}(\nu_i|\gamma)$ and $\mathcal{B}(\nu_i)$ are the signal and background PDF evaluated for the ith neutrino ν_i respectively. n_s is the signalness parameter and γ is the spectral index of the signal energy spectrum. The signalness n_s describes the estimated strength of signal in the data set. Spectral index and signalness are both parameters of the likelihood function since neither the strength nor the spectral index are known and they are both to be estimated by the likelihood function. n_s is restricted to the intervall $[0, N]$ since a negative signalness is un-physical. It would correspond to a negative source or a sink of neutrinos. If the best fit would lead to $n_\mathrm{s} < 0$, this has to be interpreted as a statistical under-fluctuation and not as a neutrino sink. The effect of the semi-positive restriction on n_s is discussed later, see section 9.2. The spectral index γ is restricted to the intervall $[1, 4]$ since this is the range where energy spectra are in general expected from supernovae as a result of diffusive shock acceleration, see section 2.2.
For numerical reasons the logarithm of the likelihood function is used instead of the likelihood itself. The position of the maximum um $\mathcal{L}$ does not change when changing

to its logarithm:

$$\begin{aligned}\log(\mathcal{L}(n_\mathrm{s},\gamma)) &= \log\left(\prod_{i=1}^{N}\left(\frac{n_\mathrm{s}}{N}\mathcal{S}(\nu_i|\gamma)+\left(1-\frac{n_\mathrm{s}}{N}\right)\mathcal{B}(\nu_i)\right)\right)\\ &= \sum_{i=1}^{N}\log\left(\left(\frac{n_\mathrm{s}}{N}\mathcal{S}(\nu_i|\gamma)+\left(1-\frac{n_\mathrm{s}}{N}\right)\mathcal{B}(\nu_i)\right)\right). \end{aligned} \tag{8.5}$$

The test statistic λ is defined as the likelihood ratio

$$\lambda = 2\log\left(\frac{\mathcal{L}(\hat{n}_\mathrm{s},\hat{\gamma})}{\mathcal{L}(0)}\right) \tag{8.6}$$

where $\hat{n}_\mathrm{s}$ and $\hat{\gamma}$ are the parameters that maximize equation 8.4[1] and $\mathcal{L}(0)$ corresponds to the null hypothesis where $n_\mathrm{s} = 0$. The value of the spectral index γ becomes meaningless if $n_\mathrm{s} = 0$.
The likelihood ratio can directly be expressed as a function of the PDFs. Plugging equation 8.4 into equation 8.6 gives

$$\begin{aligned}\lambda &= 2\log\left(\frac{\mathcal{L}(\hat{n}_\mathrm{s},\hat{\gamma})}{\mathcal{L}(0)}\right)\\ &= 2\log\left(\frac{\prod_{i=1}^{N}\left(\frac{n_\mathrm{s}}{N}\mathcal{S}+\left(1-\frac{n_\mathrm{s}}{N}\right)\mathcal{B}\right)}{\prod_{i=1}^{N}\mathcal{B}}\right)\\ &= 2\sum_{i=1}^{N}\log\left(\frac{n_\mathrm{s}}{N}\frac{\mathcal{S}}{\mathcal{B}}+1-\frac{n_\mathrm{s}}{N}\right)\\ &= 2\sum_{i=1}^{N}\log\left(n_\mathrm{s}\chi+1\right)\end{aligned} \tag{8.7}$$

with $\chi = \frac{1}{N}\left(\frac{\mathcal{S}}{\mathcal{B}}-1\right)$. The variable χ now contains the signal over background ratio of the PDFs and gives a direct estimate how signal-like a single event is. The equation 8.7 is used in the implementation of this analysis. Instead of maximizing the likelihood function equation 8.4 the test statistic equation 8.7 is maximized. The position of the maximum does not shift since $\mathcal{L}(0)$ is only a constant. For numerical reasons $-\lambda$ is minimized instead of maximizing λ.

[1] Equation 8.5 is likewise maximized with the same parameters since the logarithm is a strictly monotonous function and $\mathcal{L}$ is positive definite.

8.3. The Probability Density Functions

A probability density function (PDF) describes the probability to measure a certain value. It is the keystone of the maximum likelihood method. The PDF is normalized to unity and can be discrete or continuous. In the context of neutrino point source likelihood analysis (equation 8.4) there are two important PDFs, the signal PDF $\mathcal{S}$ and the background PDF $\mathcal{B}$. They describe the expected distribution of signal and background events and are used in the likelihood analysis to discriminate between them and to estimate the amount of signal in the data set. Both signal and background PDF can be written as a product

$$\mathcal{S}(\nu_i) = \mathcal{N}_{\mathcal{S}}(\alpha_i, \delta_i) \times \mathcal{E}_{\mathcal{S}}(E_i, \delta_i|\gamma) \times \mathcal{T}_{\mathcal{S}}(t_i) \tag{8.8}$$

$$\mathcal{B}(\nu_i) = \mathcal{N}_{\mathcal{B}}(\alpha_i, \delta_i) \times \mathcal{E}_{\mathcal{B}}(E_i, \delta_i) \times \mathcal{T}_{\mathcal{B}}(t_i). \tag{8.9}$$

where $\mathcal{N}$ is the spacial PDF, $\mathcal{E}$ is the energy PDF and $\mathcal{T}$ is the time PDF. α_i and δ_i are right ascension and declination of an event i, E_i is its energy and t_i its time. The energy PDF $\mathcal{E}$ is not only a function of the energy, but also of declination since the detector has a declination dependent energy acceptance, see section 4.2. The quantity γ describes the assumed spectral index of the signal.

8.4. Background PDF

The background PDF $\mathcal{B}$ is constructed to describe the distribution of background events. A most accurate description is required since a potential discovery is defined as a point source like fluctuation above the background described by this PDF. To construct the background PDF, there are two methods available:

1. The first strategy is to simulate the background. One starts from the assumption that the background consists of atmospheric muons, atmospheric neutrinos, and diffuse astrophysical neutrinos. A simulation of the full detector response and event reconstruction is performed to all these kinds of background. This procedure is very complicated and computationally expensive. If one has a full understanding of the physics of the background processes and the detector behavior, this is the best way to handle the background. Unfortunately, the simulation of background still includes large uncertainties both due to unknowns in

the underlying physical processes and the simulation of detector response.

2. The second approach is to generate the background PDF by using real data. The assumption is that the data is contaminated by only a negligible amount of signal events. Since no point source has been discovered so far, this assumption is very reliable. As discussed in the next section, the spatial background PDF does not contain full $2\mathbb{D}$ spatial information, only a declination dependence. The influence of a (yet undiscovered) point source would thus still be smeared out in the background PDF and not spoil the analysis. This approach already includes all detector effects and is, therefore, model independent. The measured data is then used to generate the background PDF.

8.4.1. Spatial Background PDF

The spatial background PDF describes the chance for a background event to be reconstructed at a certain position in the sky. The background events originate from cosmic ray interactions in the atmosphere which produce muons and atmospheric neutrinos and from the diffuse astrophysical neutrino background[2]. The sources of the background are isotropic. Any directional dependence of the measured background events is thus a result of detector effects or shielding of the detector by the Earth.
IceCube has a hexagonal shape; the detector acceptance depends on the azimuth angle. The azimuth angle is a detector coordinate; it is fixed on the detector. Since the Earth and IceCube with it are spinning around the azimuth axis, the effects are averaged out when changing to equatorial coordinates. In equatorial coordinates, the sky is fixed, and the Earth is spinning. So in equatorial coordinates, there is no dependence on the right ascension any more[3]. All changes of the background rate can described as a function of declination only $\mathcal{N}_{\mathcal{B}}(\alpha, \delta) \rightarrow \mathcal{N}_{\mathcal{B}}(\delta)$. The background PDF is generated from experimental data in several steps:

1. First the measured data is put into a one-dimensional histogram of sinus declination angles bins ($(\sin(\delta))$). The histogram is then normalized to unity. Bins of sinus declination are used for equally sized bins on the sphere instead of binning in declination.

[2]The diffuse astrophysical neutrino background is the combined flux from all astrophysical neutrino sources. No source can be resolved in this flux. For a detailed discussion, see chapter 7.
[3]Right ascension is the corresponding angle to the azimuth.

2. In a second step, the histogram is interpolated by a spline function to get a smooth function.

3. In the final step the normalization is changed by an additional factor $\frac{1}{2\pi}$ to obey correct normalization over the entire sphere

$$\int_{\Omega} \mathcal{N}_{\mathcal{B}}(\alpha, \delta) \mathrm{d}\Omega = \int_{0}^{2\pi} \int_{-1}^{1} \mathcal{N}_{\mathcal{B}}(\sin(\delta)) \mathrm{d}\alpha \mathrm{d}\sin(\delta) = 1. \tag{8.10}$$

8.4.2. Time Background PDF

The background time PDF is defined by the inverse livetime of the data set

$$\mathcal{T}_{\mathcal{B}}(t_i) = \frac{1}{\text{livetime}}. \tag{8.11}$$

The background is expected to be constant in time. Any small scale fluctuations due to detector down times or partial detector failures of the order of several minutesare averaged out over the typical time scales of the analysis which is at least 20 days, see chapter 10.

8.5. Signal PDF

The signal PDF describes the distribution of potential signal neutrinos. The hypothesis is that the source of neutrino is point-like (point source) and follows a power law energy spectrum. Furthermore, it is assumed that the signal follows a certain neutrino light curve in time. Space, energy and time PDF are constructed accordingly.

8.5.1. Signal Space PDF

The space part of the signal PDF $\mathcal{N}_{\mathcal{S}}$ describes the point spread function of a certain source. The point spread function describes the distribution of reconstructed events around a source location. An ideal neutrino telescope would have an infinite resolution, and the spatial PDF would just be a two-dimensional delta distribution. A real neutrino telescope like IceCube has only a limited resolution. The point spread

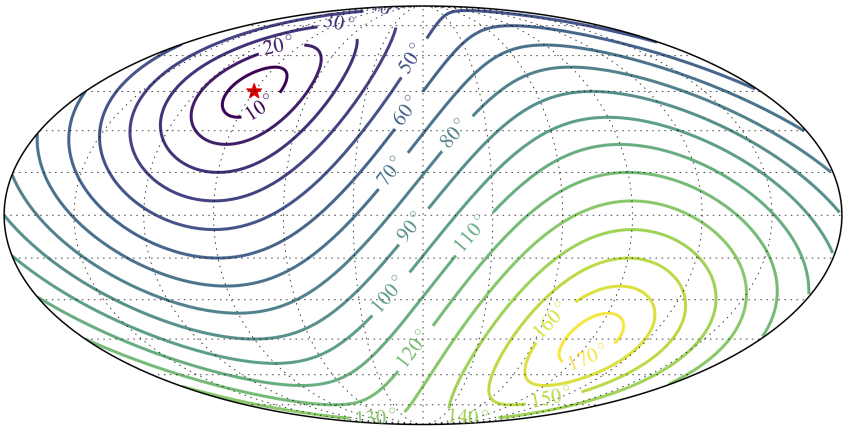

Figure 8.1.: The plot shows the angular distance from the red point with contour lines in 10° steps. The sphere is projected with the Mollweide projection.

function of IceCube is assumed to be a two-dimensional normal distribution:

$$\mathcal{N}_{\mathcal{S}}(\alpha_i, \delta_i, \sigma_i) = \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{1}{2}\frac{d_\phi^2(\alpha_i, \delta_i, \alpha_S, \delta_S)}{\sigma_i^2}\right). \tag{8.12}$$

α_i and δ_i are the reconstructed right ascension and declination of the event, and σ_i is the estimated error on this reconstructed direction. α_S and δ_S are the directions of a potential source in the sky. The quantity $d_\phi(\alpha_i, \delta_i, \alpha_S, \delta_S)$ is the angular distance of the neutrino direction and the source position on the sphere. An illustration of this so-called great circle distance is shown in figure 8.1. Equation 8.12 contains two assumptions:

1. First it is assumed that the estimate of the reconstruction error follows a normal distribution.

2. The second approximation is concerning the geometry. The geometrical framework for this analysis is the unit sphere. The equivalent to a multidimensional Gaussian function on flat space is the so-called Von Mises-Fisher distribution [97]. The sky plane corresponds to the 2-sphere. For narrow peaked Gaussian, the sphere locally looks flat and can be approximated by the classical Gaussian

function. This assumption is valid for track-like events used in this analysis with a typical resolution of 1°. A more detailed study of this can be found in appendix D in [42].

The spatial information is the most important part of the signal PDF. Most separation power from background comes from it.

8.5.2. Signal Time PDF

This analysis utilizes time information to further separate signal and background events. Several models for the *neutrino light curve* $LC_\nu(t)$ are assumed, see section 3.5.1. The function $LC_\nu(t)$ describes how a hypothetical source evolves over time. The light-curve can be transferred into a signal time PDF $\mathcal{T}_\mathcal{S}$ by just ensuring proper normalization:

$$\mathcal{T}_\mathcal{S}(t) = \frac{LC_\nu(t)}{\int_{t_\text{start}}^{t_\text{end}} LC_\nu(t')\mathrm{d}t'} \tag{8.13}$$

where t_start and t_end are beginning and end of the corresponding IceCube season.

8.6. Energy Weighting

The energy weight of an event is defined as the ratio of signal energy PDF $\mathcal{E}_\mathcal{S}$ to background energy PDF $\mathcal{E}_\mathcal{B}$

$$\frac{\mathcal{E}_\mathcal{S}(E_i, \delta_i|\gamma)}{\mathcal{E}_\mathcal{B}(E_i, \delta_i)}. \tag{8.14}$$

This ratio contributes to the likelihood ratio, see equation 8.7. It weights how signal-like an event looks based on its energy. Since different energy spectra are expected for signal and background neutrinos, this weighing provides an additional separation of signal and background events. IceCube has strong declination dependence, so the energy PDFs are functions of both energy E and declination angle δ.
To generate the energy PDFs $\mathcal{E}_\mathcal{S}(E)$ and $\mathcal{E}_\mathcal{B}(E)$, signal and background histograms of energy and declination angle is generated. The background histogram is generated using real data, signal histogram by using simulated data, see figure 8.2. Each declination slice of the histograms is normalized to unity. By this normalization, each

declination slice itself is now a correctly normalized PDF

$$\int_{10^1\,\mathrm{GeV}}^{10^{10}\,\mathrm{GeV}} \mathcal{E}(E,\delta)\mathrm{d}E = 1. \tag{8.15}$$

Note that the normalization holds for the energy, not for the declination[4]. The energy

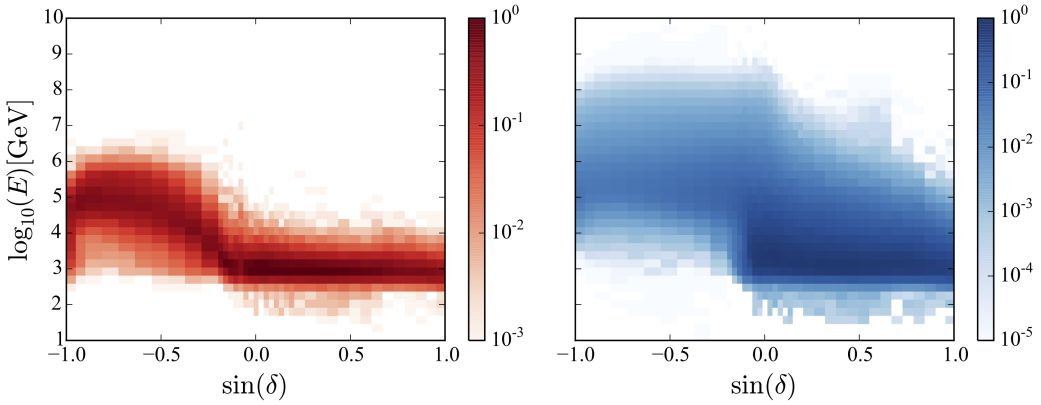

Figure 8.2.: Distribution of events in declination and energy for the IC86-I season. The left panel shows experimental data, and the right panel simulated signal events for an E^{-2} energy spectrum.

weight is defined as the ratio of signal to background histogram. Due to limited statistics some of the histogram entries are empty, and so the ratio is not well defined. If both signal and experimental histogram are empty, this is not problematic since no evaluation is required in this bin ever. If only the background PDF is zero, this would mean division by zero. The region is most likely not background free but appears to be due to limited statistics. The conservative extrapolation is to use the largest well-defined signal to background ratio of that declination slice in the non-defined regions. These empty background bins are only a problem in simulation cases where signal events are artificially injected. When performing the analysis on real data, the background PDF is generated from the full data set and thus is not required to be evaluated at empty regions in the background histogram. The ratio plot and thus the energy weighting function is shown in figure 8.3. The histogram in figure 8.3 is then interpolated with a two-dimensional spline function. This interpolation is used as the

[4]The normalization in the declination is already realized in the spatial PDFs.

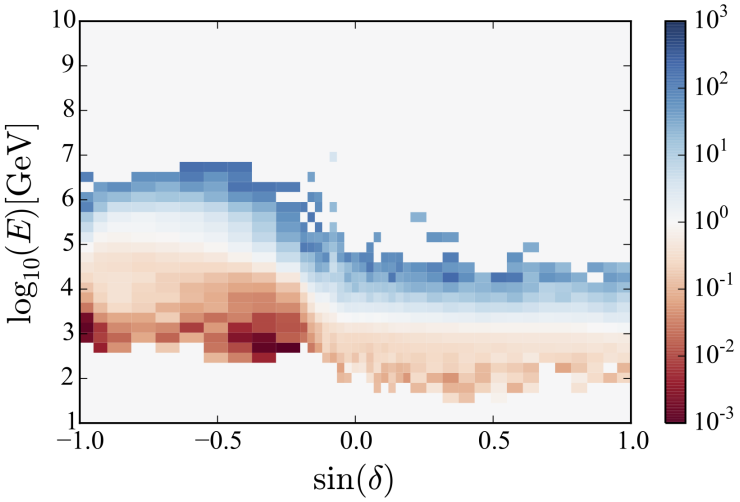

Figure 8.3.: Energy weights, defined as the ratio of signal and background PDFs for the IC86-I data set. An E^{-2} energy spectrum was assumed for the signal. The color scale indicates the weighting as defined in equation 8.14.

energy weighting function in the analysis.
So far a fixed energy spectrum $E^{-\gamma}$ was assumed. Since the spectral index γ is a free parameter in the likelihood function (equation 8.4), the energy weights have to be generated for all potential values of γ. Instead of repeating the generation of histograms for each γ as described previously, the energy weighting is pre-evaluated between $\gamma = 1$ and $\gamma = 4$ in steps of 0.25. Between this grid points, a Taylor expansion up to second order is used as an interpolation in the γ regime of spectral indexes to speed up the likelihood evaluation.

8.7. Combination of Different Data Sets

This analysis utilized data from different IceCube seasons. Each season has its own event selection and partly different detector configuration. This section describes how different data sets are combined in a single likelihood function. For the following discussion it is assumed that the data sets have no overlap and are independent of each other. This is true for the IceCube seasons used in this analysis since the data has been taken at different times.

The likelihood function was introduced as the product of probability density functions for each measured event in the data set (equation 8.4). The first guess how to combine different data sets is thus just the product of the corresponding likelihood functions

$$\mathcal{L}^{\star}_{\mathrm{comb}}(n_{\mathrm{s}}, \gamma) = \prod_{j=1}^{M} \mathcal{L}_j(n_{\mathrm{s}}, \gamma) \tag{8.16}$$

where the index j runs over all M data sets. Equation 8.16 is only true if all data sets have the same expectation of signal. If each data set has a different signal expectation, the total n_{s} has to be distributed over the different data sets accordingly. n_j is the expected number of signal events in a certain data set j for a given model and an assumed spectral index γ. A weight for each data set is defined by

$$w_j(\gamma) = \frac{n_j(\gamma)}{\sum_{l=1}^{M} n_l(\gamma)}. \tag{8.17}$$

where the denominator just ensures proper normalization. Equation 8.16 then becomes

$$\mathcal{L}_{\mathrm{comb}}(n_{\mathrm{s}}, \gamma) = \prod_{j=1}^{M} \mathcal{L}_j(w_j(\gamma) n_{\mathrm{s}}, \gamma). \tag{8.18}$$

This definition has the required properties. The total n_{s} events are distributed over the data sets accordingly to their expectation values. To understand this behavior, a Gedankenexperiments is performed: Assume two data sets where the second set has a two times larger signal expectation than the first set ($w_1 = \frac{1}{3}$, $w_2 = \frac{2}{3}$). If the combined likelihood function is evaluated with $n_{\mathrm{s}} = 3$, the correct description is to expect one event in data set one ($3 \times \frac{1}{3} = 1$) and two events in data set two ($3 \times \frac{2}{3} = 2$).

The correct weighting is ensured by using this method. It is always used in this analysis whenever different data sets are combined. Note that the likelihood becomes more complicated with respect to γ. γ does not only appear in the signal PDF $\mathcal{S}(\nu_i|\gamma)$, but also as a scaling factor for $n_{\mathrm{s}}(\gamma)$ (equation 8.17).

8.8. Stacking

Stacking is a method to boost the sensitivity of an analysis in the case of more than one potential source. By stacking, many potential sources are combined (stacked) and analyzed simultaneously such that their signal adds up. As a downside also the background from each potential source is stacked, so the total background is also increased.
In the point source likelihood function, stacking is realized by summing up signal PDFs from different sources

$$\mathcal{S} = \sum_{j=1}^{M} w_j \mathcal{S}_j \tag{8.19}$$

where the index j is summing over all M sources, $\mathcal{S}_j$ are the individual PDFs of each source and w_j is a weighting term. Since the individual PDFs $\mathcal{S}_j$ are normalized to unity each, the weights have to obey

$$\sum_{j=1}^{M} w_j = 1 \tag{8.20}$$

to ensure the correct normalization of the stacked signal PDF. Following the same argument as in section 8.7, the weights are constructed to be proportional to the signal expectation of each source. In the case of a transient sources the signal expectation of a source j is given by

$$n_j = \underbrace{\frac{\Phi_0^j}{D_\mathrm{p}^2}}_{\text{Source Properties}} \times \underbrace{\int_{t_\mathrm{start}}^{t_\mathrm{end}} \int_{E_\mathrm{min}}^{E_\mathrm{max}} LC_j^\nu(t) E^{-\gamma} \times Acc(t, \delta_j, E) \mathrm{d}t \mathrm{d}E}_{\text{Time Dependence}} . \tag{8.21}$$

Φ_0 is the intrinsic power of the source, D_p the propper distance, $LC^\nu(t)$ the expected neutrino light curve, $E^{-\gamma}$ is neutrino source spectrum with assuned spectral index γ. $Acc(t, \delta, E)$ is the effective area. δ is the declination of the source and γ the assumed spectral index of the signal energy spectrum. The boundaries of the integral t_start and t_end are the times of beginning and end of the data set and the energy range of the analysis E_min and E_max.
The first term in equation 8.21 depend only on properties of the source. The flux scales linearly with the intrinsic power of the source and decreases quadratically with

the distance. The second term depends both on the light curve, the spectrum and on the detector acceptance. It describes which part of the neutrino light curve was sampled by which detector configuration.
This analysis uses several seasons of IceCube data. Within an IceCube season the detector acceptance is constant and can be taken out of the integral. If an analysis is performed with more than one data set of IceCube season, signal expectation values are computed for each source and each season. As discussed in section 8.7, different data sets can be treated as being independent since they have no overlap. The computed expectation values for all sources and seasons can be written as

$$\begin{pmatrix} n_1^1 & n_1^2 & \dots & n_1^M \\ n_2^1 & n_2^2 & \dots & n_2^M \\ \vdots & \vdots & \ddots & \vdots \\ n_K^1 & n_K^2 & \dots & n_K^M \end{pmatrix} \tag{8.22}$$

where the column are the K seasons and the rows are the M sources. The total number of events in a certain season l is given by the sum of the corresponding row

$$n_l^{\mathrm{tot}} = \sum_{j=1}^{M} n_l^j, \tag{8.23}$$

the total number of events from a source j is given by sum over the corresponding column

$$n_{\mathrm{tot}}^j = \sum_{l=1}^{K} n_l^j, \tag{8.24}$$

and the total expectation number simply by

$$n_{\mathrm{tot}} = \sum_{j=1}^{M} \sum_{l=1}^{K} n_l^j. \tag{8.25}$$

So at this point there are two different weightings:

1. Data set weighting: Each data set has a weight assigned, proportional to the total number of expected signal events in this data set n_l^{tot}. The signalness parameter n_{s} (see equation 8.4) is distributed on the different data sets according

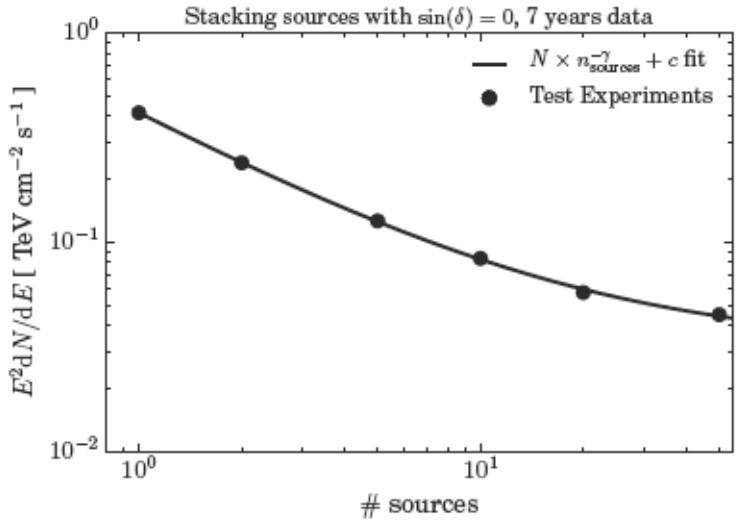

Figure 8.4.: The plot shows the gain of sensitivity as an effect of stacking. An increasing number of sources of equal brightness is simulated and the sensitivity is computed. The sources are located at the same declination to cancel out detector effects, but at different right ascensions to prevent overlap.

to this weighting, see section 8.7.

2. Source weights: Each source gets a weight in each data set. It describes how to distribute the n_l^{tot} for that specific season on the different sources, see equation 8.19. The weights are given by

$$w_l^j = \frac{n_l^j}{\sum_{k=1}^{M} n_l^k}. \tag{8.26}$$

Figure 8.4 shows the gain in sensitivity as a result of stacking. Equally bright sources have been simulated at the same declination. The right ascension is different so there is no source overlap. The sensitivity follows as falling power law. By stacking only 10 sources a gain of a factor of 4 in sensitivity is possible. The test setup of equally bright and equally distant sources is unrealistic, but is shows the effect of stacking clearly. Of course, in reality the sources do not have equal distance, but it was chosen here to demonstrate the effect of stacking.

8.9. Multi-Component-Fit

The stacking method requires an assumption to estimate the signal expectation. Since no neutrino signal has been detected from any of the stacked sources, any assumption about the weights is model dependent. In this analysis, the weights depend on the intrinsic luminosity, the distance, the energy spectrum, the explosion time and the light curve of the supernovae, see equation 8.21. The dimming of a source with distance $\propto D_{\mathrm{p}}^{-2}$ assumes isotropic emission, but the distance measures of supernovae typically have uncertainties. Neutrino light curves are model assumptions since no neutrino signal from a point source has ever been measured in the energy range relevant here. Typically there is also not enough optical data available (e.g., full optical light curve) to apply more advanced neutrino emission models. The standard candle ansatz, assuming equal intrinsic neutrino brightness of all sources is probably wrong, but more detailed modeling is impossible. Due to these uncertainties, it is very likely that the weights do not correspond to the actual ratio of fluxes.
Nonoptimal weighting can decrease the sensitivity of a stacked analysis dramatically. Dim sources that get a high weight assigned only add background to the analysis. Bright sources with accidentally low weights thus do not contribute to the analysis since they are suppressed by their low weights.
Even if the correct weights (and thus the expectation number of events for each source) is known, these weights do not necessarily give the best data description. The actual number of signal events in a data sample is a Poisson random number based on the expectation values. So even if on average the weights correspond to the ratio of expectation values, they do most likely not match the number of events from the stacked sources in the actual data set.

Formulation of Multi-Component-Fit

To overcome the problem stated above, a fitting of the weights is applied in this analysis. This is the first time in an IceCube analysis that this is done. The weights are interpreted as additional parameters of the likelihood function. The likelihood function is maximized also with respect to the weights. Fitting the weights w_j and the total value of n_{s} is mathematically equivalent to fitting the parameter n_j for each

source j independently:

$$\log \mathcal{L}(n_\mathrm{s}, \gamma, w_1, w_2, \dots) = \log \sum_{i=1}^{N} \left(\frac{n_\mathrm{s}}{N} \sum_{j=1}^{M} w_j \mathcal{S}_j + \left(1 - \frac{n_\mathrm{s}}{N}\right) \mathcal{B} \right) \tag{8.27}$$

$$= \log \sum_{i=1}^{N} \left(\sum_{j=1}^{M} \frac{w_j n_\mathrm{s}}{N} \mathcal{S}_j + \left(1 - \frac{n_\mathrm{s}}{N}\right) \mathcal{B} \right) \tag{8.28}$$

$$\iff \log \mathcal{L}(n_1, n_2, \dots, \gamma) = \log \sum_{i=1}^{N} \left(\sum_{j=1}^{M} \frac{n_j}{N} \mathcal{S}_j + \left(1 - \frac{\sum_{j=1}^{M} n_j}{N}\right) \mathcal{B} \right) \tag{8.29}$$

with $n_j = w_j n_\mathrm{s}$. The introduction of the additional fit parameters does at first decrease the sensitivity compared to the analysis using fixed weights: A multi component fit can better mimic local background over-fluctuations at a single source position and confuse this with an actual signal. In the fixed weights approach one assumes a constant ratio between the potential sources. On average, a local over-fluctiation at a signal source position is compensated by another under-fluctuation. This makes the fit more robust against local background fluctuations. But this feature is also responsible for the better performance of the multi component fit compared to the fixed weights ansatz. A single bright neutrino source which is not known to be bright in neutrinos can drive the entire likelihood fit without influencing the fit of other, dim sources.
The multi-component fit requires a likelihood maximization in an $M + 1$ dimensional parameter space[5] compared to a 2 dimensional space when assuming fixed weights. This maximization is computationally very expensive and only feasible for a few tens of sources.

Performance of Multi-Component Fit

Comparison between the fixed weights method and the multi-component fit is made within a test setup. Five sources of the same intrinsic brightness are equally distributed over the sky. The sources do not appear equally bright in the detector since the detector acceptance has a strong dependence on the declination of the source. To mimic the effect of weighting errors, the intrinsic brightness is then varied with a log-

[5] The number of parameters are one degree of freedom for each of the M sources plus the global spectral index γ.

normal distribution for each source. After the variation, the total intrinsic brightness is rescaled to the original one, so the total power of the sources stays the same. Sensitivity is computed for each of the different values of σ in the lognormal distribution for fixed weights and the multi-component fit. The result is shown in figure 8.5. A

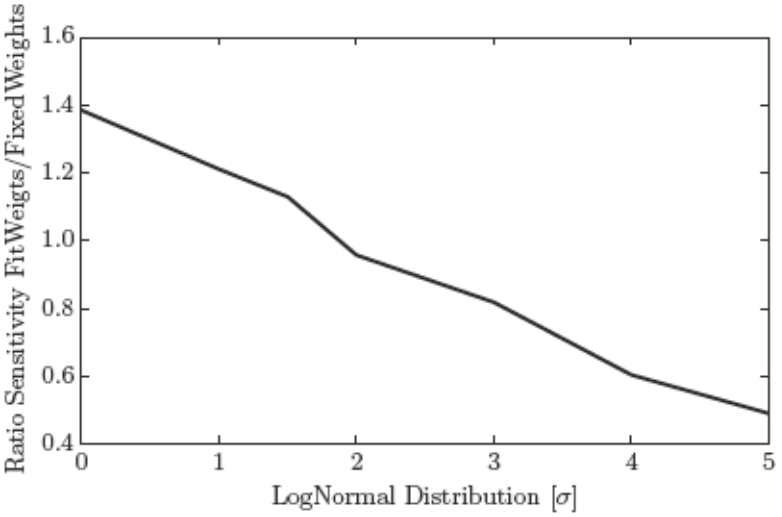

Figure 8.5.: The plot shows the ratio of sensitives for five stacked sources of same intrinsic luminosity. The injected flux is then smeared out by a log-normal distribution and then re-normalized to the same total brightness. For larger intrinsic spreads, the multi-component fit tends to perform better than the assumption of fixed weights.

better performance of the multi-component fit can be seen for larger values of σ. The fixed weights method performs better for small σ. If the weights differ more and more from the true value, the fitting of weights starts to perform better as one would expect from the previous discussion. Another advantage of the fitting of weights is the model independence. This allows giving more general statements than if certain weights are assumed. So the interpretation of results using this analysis are much more general.

Combining Data Sets in the Multi-Component Fit

When performing the multi-component fit on more than one data set simultaneously, weights have to be computed accordingly. For each source the total number of expected events n_j has to be distributed on the different datasets following the descrip-

tion in section 8.7. Starting from the expectation values (equation 8.22), the season weight of each source is computed from the expectation value for this specific source overall season. The weight of the jth source concerning the lth season is given by

$$w_j^l = \frac{n_j^l}{\sum_{k=1}^{K} n_j^k} \tag{8.30}$$

and the expectation value in this season by $n_j^l = w_j^l n_j$. Comparing equation 8.26 and equation 8.30, the difference appears in the normalization: The weights in the fixed weight case are normalized per data set, the weights in the weights fitting case are normalized per source.
Even if the weights are fitted in the multi-component fit, the distribution over the different IceCube seasons is still determined by the ratio of expectation values. Fitting the weights does only remove assumptions about the intrinsic brightness of the source, but still, contains the model dependence on the neutrino light curve model.

8.10. Hypothesis Testing

The goal of hypothesis testing is to measure how likely or unlikely the outcome of an experiment is assuming a certain hypothesis H_0 [54]. To quantify the result of a statistical test, the scalar quantity called test statistic λ is defined. In the context of the likelihood method (see discussion in section 8.2), the test statistic can be defined as

$$\lambda = 2 \log \left(\frac{\mathcal{L}(\hat{n}_\mathrm{s}, \hat{\gamma})}{\mathcal{L}(0)} \right)$$

where $\mathcal{L}(\hat{n}_\mathrm{s}, \hat{\gamma})$ is the maximized likelihood function, $\hat{n}_\mathrm{s}$ and $\hat{\gamma}$ are the parameters at maximum and $\mathcal{L}(0)$ is the likelihood function evaluated under the assumption of the hypothesis to be tested (H_0). In the context of this analysis, H_0 always describes the background only hypothesis, the data set does not contain any signal from a point source[6].
The more consistent the outcome of the likelihood maximization $\mathcal{L}(\hat{n}_\mathrm{s}, \hat{\gamma})$ is to the null hypothesis H_0, the closer the value of λ is to zero. Due to statistical fluctuations of real date, even if H_0 is true, the λ distribution will not result in δ-peak at $\lambda = 0$

[6] If $n_\mathrm{s} = 0$, then the value of the spectral index γ is degenerated and thus not of interest.

only, but in a distribution around it.

p-Values

The p-value is a scalar which quantifies the consistency of an experimental outcome with a hypothesis H_0. If λ_{exp} is the experimental outcome of an experiment and $H_0(\lambda)$ is the distribution of background test statistic, the measured p-value is defined as

$$\begin{aligned} p &= 1 - \int_0^{\lambda_{\text{exp}}} H_0(\lambda)\mathrm{d}\lambda \\ &= \int_{\lambda_{\text{exp}}}^{\infty} H_0(\lambda)\mathrm{d}\lambda. \end{aligned} \tag{8.31}$$

The p-value is the probability to get an experimental result which is equally or more inconsistent with the expectation from null hypothesis H_0 under the assumption that H_0 is true. The smaller the p-value, the larger the inconsistency.
The p-value is also called the significance. Significance can also be measured in units of standard deviations σ of a normal distribution. The typical requirement for discovery in particle physics is a 5σ discovery, which corresponds to a p-value of roughly $6 \cdot 10^{-7}$. The requirement means that the null hypothesis H_0 is discarded only if the measured outcome looks so different to the expectation from H_0 that a result as the measured one or an even more different extreme result will only happen because of fluctuations of H_0 in about one out of ten million trials. At this point, one might consider an alternative hypothesis H_1 and claim a discovery.

Hypothesis Testing

In hypothesis testing, one want to test two hypotheses against each other. H_0 is called the null hypothesis and typically corresponds to the established model where the alternative hypothesis H_1 typically incorporated H_0 plus an additional, so far unknown component. This unknown component can be a new resonance in a spectrum or a neutrino point source. In the context of this work, H_0 is also called the background hypothesis and H_1 the signal hypothesis[7].
Before the experiment is performed, the required significance for discarding H_0 is selected. This could be 90%, 5σ or any other value. It has to be chosen before the

[7] The signal hypothesis contains the background hypotheses plus a point source signal on top.

experiment is actually performed. The sensitivity level then corresponds to a threshold value λ_{thres} for the test statistic, see equation 8.31. Thus, if the experimental outcome λ_{exp} is below the threshold value λ_{thres}, the null hypothesis is accepted and H_1 rejected. If $\lambda_{\text{exp}} > \lambda_{\text{thres}}$, the null hypothesis H_0 is rejected and the alternative hypothesis H_1 is accepted. When testing the null hypothesis H_0 against the signal hypothesis H_1, there are four scenarios:

	accept H_0	accept H_1
H_0 true	✓	Error Type II
H_1 true	Error Type I	✓

Table 8.1.: Possible outcome of a hypothesis testing involving two hypothesis.

Correct Hypothesis Selected

The hypothesis testing identified the correct hypothesis. This is the ideal case and indicated with the check marks in table 8.1.

Error Type I An error of type I is the probability to reject the null hypothesis H_0 given that it is true. It is also called false positive. The chance for error type I α is is given by

$$\alpha = \int_{\lambda_{\text{thres}}}^{\infty} H_0(\lambda)\mathrm{d}\lambda$$

and is identical to the p-value.

Error Type II The error of type II, also called false negative, is the chance to reject the signal hypothesis H_1 even if it is true. It is given by

$$\beta = \int_{0}^{\lambda_{\text{thres}}} H_1(\lambda)\mathrm{d}\lambda.$$

Choice of Test Statistic Threshold

The selection of λ_{thres} is a compromise between an error of type I and an error of type II. If λ_{thres} is set to a large value, the chance of falsely claiming a discovery is very low, but the chance to identify H_1 is very low, too. In contrast does the choice

of a small λ_{thres} lead to more likely discoveries, but also to more false claims just as a result of background fluctuations. There is no optimal choice of λ_{thres} since it is a matter of statistical interpretation. A smaller value of the error of first type α is often called conservative. In the context of potential discoveries and fundamental claims, analyses are often designed to be more on the conservative side.

Characterizing the Performance of an Analysis

Studying both errors of type I and II of a certain analysis allow to give a measure on the strength of the analysis. Typically this is done by computing the required signal strength that leads to a signal hypothesis H_1 which then fulfills certain selections of α and β.

Sensitivity

In IceCube the sensitivity is defined as the required flux for which $\alpha = 0.5$ and $\beta = 0.1$. This is also called *median sensitivity at* 90% *confidence level*. For this work, sensitivity is the main quantity to describe the performance. The sensitivity is given in a unit of signal strength; this can be the neutrino flux, the neutrino fluence or a similar unit.

Discovery Potential

The discovery potential is defined as the required flux for a type I error of 5σ in 50% of the experiments ($\beta = 0.5$). Compared to the sensitivity, the discovery potential is much more sensitive to the background. Also, the flux required to fulfill the discovery potential requirement is typically larger than for sensitivity. A practical challenge when computing the discovery potential is that the background test statics distribution $H_0(\lambda)$ has to be known to an accuracy of about the 5σ regime. If generated by simulation, this requires about 10 million test experiments which can be computational challenging. Figure 8.6 illustrates both sensitivity and discovery potential.

Computing Sensitivity The sensitivity is estimated from a simulation. First, several hundred simulations with only background events are performed. The likelihood function is maximized on this data set and the outcoming test statistic values λ are stored. They are then ordered. This leads to the inverse cumulative distribution of

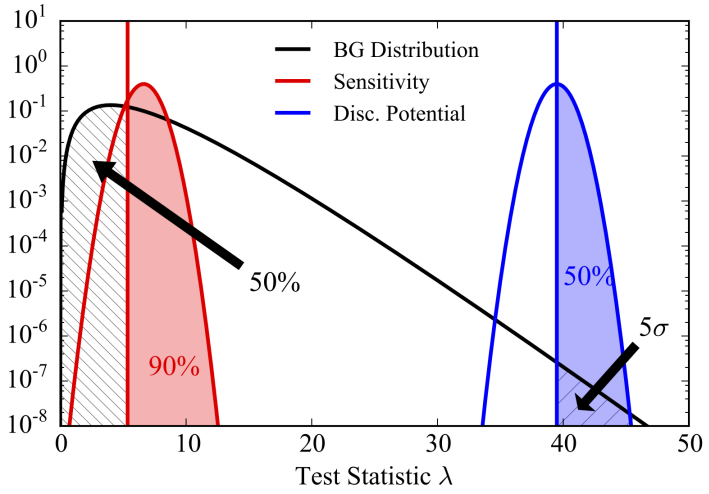

Figure 8.6.: Schematic plot describing definition of sensitivity and discovery potential. The black line is the background distribution $H_0\lambda$ and red and blue line the signal hypothesis $H_1(\lambda)$ for sensitivity and discovery potential.

H_0

$$1 - \int_0^{\lambda} H_0(\lambda')\mathrm{d}\lambda'$$

where $H_0(\lambda)$ is the background test statistic distribution. To compute the sensitivity, the median on the distribution $H_0(\lambda)$ is computed. Since the likelihood maximization is restricted to $n_\mathrm{s} \geq 0$, any under fluctuation is fitted to $n_\mathrm{s} = 0$ and thus there is a pile up at $\lambda = 0$, so the typical median of the background is typically zero or close to. Then the procedure is repeated with additional signal events from a simulated point source with a certain strength i. This generates the distribution $H_i(\lambda)$ where the subscript i denotes the injected flux of strength i in arbitrary units. The integral

$$P_\mathrm{detect} = \int_{\lambda_\mathrm{median}}^{\infty} H_i(\lambda')\mathrm{d}\lambda' \tag{8.32}$$

is computed. It describes the chance to measure a test static value $\lambda > lambda_\mathrm{median}$. Several scenarios with different signal strength i are computed, and the injected signal strength is plotted against the corresponding outcome of equation 8.32. The distribution is interpolated to find the required flux to have a chance to be above the

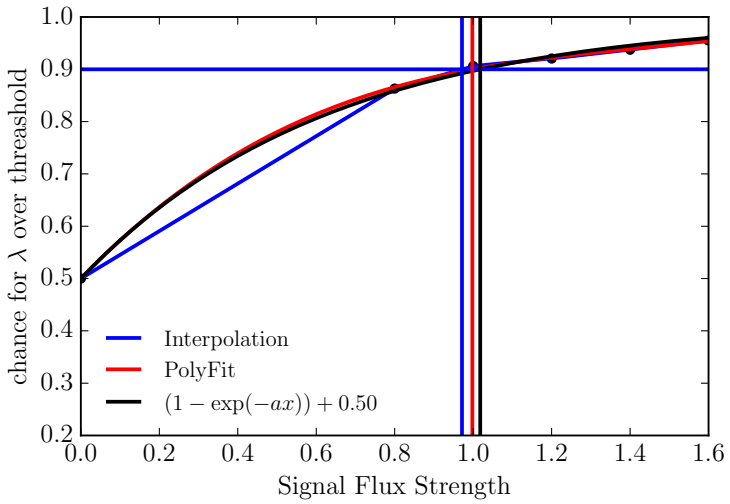

Figure 8.7.: The plot shows the interpolation to find the sensitivity flux. Compared are three methods for interpolation.

threshold in 90% of the cases. Figure 8.7 shows the process of interpolation to find the required sensitivity flux. Several interpolations are performed for testing, a linear interpolation, a polynomial fit and a dedicated fit function $f(x) = (1 - \mathrm{e}^{-ax}) + 1/2$. Computing the sensitivity requires many simulations with various signal strengths and is computational very demanding. Therefore, also the discovery potential has not been computed in this thesis since it would have required even more computational effort. The limit for claiming a discovery has been set to 5σ consistent with the standard in particle physics. For performance studies, sensitivity has been used to quantify the analysis performance instead of discovery potential.

9. Behavior of the Likelihood Function

This chapter discusses the behavior of the point source likelihood method. The effect of under- and overestimation of the angular error, fitting negative signal strengths, and influence of energy information in the fit are discussed.

9.1. Testing Setup

A simplified test framework is used to study the behavior of the point source likelihood method. In contrast to using the actual data set, this method allows full control of all the parameters of both signal and background events. Since this study address the features of the point source likelihood analysis in general, this simplification is well justified. The test setup is defined as follows:

- The analysis is carried out on a flat Euclidean $2\mathbb{D}$ plan of 10×10. It is assumed that the background is flat. Thus the space background PDF is given by

$$\mathcal{N}_{\mathcal{B}} = \frac{1}{A_{\text{tot}}} = \frac{1}{10 \times 10}. \tag{9.1}$$

  where A_{tot} is the total area of the region where the analysis is performed. For each realization of the data, a set of 10^3 events is drawn randomly from this flat distribution.

- It is assumed that the signal is point like, so all signal events are assumed to originate from a single point. It is located right at the center of the plane at $(0, 0)$. The point spread function is assumed to be Gaussian. The PDF is thus given by the product of two Gaussian functions

$$\mathcal{N}_{\mathcal{S}}(x, y) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{\left(-\frac{1}{2}\frac{(x-x_0)^2}{\sigma^2}\right)} \times \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{\left(-\frac{1}{2}\frac{(y-y_0)^2}{\sigma^2}\right)} \tag{9.2}$$

where x and y are the two coordinates in the plane, x_0 and y_0 are the coordinates of the assumed position of the source $(0, 0)$ and σ is the width of the distributions. This expression is a simplified version of equation 8.12 where the distance on the sphere has been replaced the by Euclidean metric. The value of σ is set fixed to 0.5 for all signal events if not quoted otherwise.

- An energy PDF is simulated by a power law

$$\mathcal{E}(E, \gamma) \propto E^{-\gamma} \tag{9.3}$$

defined and normalized in the interval $[10^0, 10^5]$. For the background $\gamma = 4$ and for the signal $\gamma = 2$ is chosen if not marked otherwise. Signal and background events are drawn from the PDF (equation 9.3). The background spectral index γ is fixed. The spectral index of the signal is a free parameter in the likelihood function.

- The likelihood function $\mathcal{L}$ and test statistic λ are defined as described in equation 8.4 and equation 8.7. For each trial, signal and background events are drawn from the corresponding PDFs. The test statistics is maximized for each trial to find $\hat{n}_\mathrm{s}$. Distributions of $\hat{n}_\mathrm{s}$ and corresponding λ as well as the sensitivity are studied. The sensitivity is defined in the usual way as median sensitivity at 90% confidence level, see chapter 8.10.

9.2. Testing Spatial PDF only

First only the spatial PDFs (equations 9.1 and 9.2) are used in the likelihood function, so the analysis is sensitive to the spatial distribution of the events. First, 10000 test experiments with only background events are performed without the injection of any signal events. In the optimization process n_s is limited to non-negative values.
Then. signal events are injected and the performance of the analysis is studied again.

Background Events only The outcome of the trials without injection of any signal event is shown in figure 9.1. About 50% of the trials are fitted to zero as expected for an equal distribution of over and under-fluctuations. All these cases end up in the first bin. When removing the lower bound of n_s, the test statistics equation 8.6 is

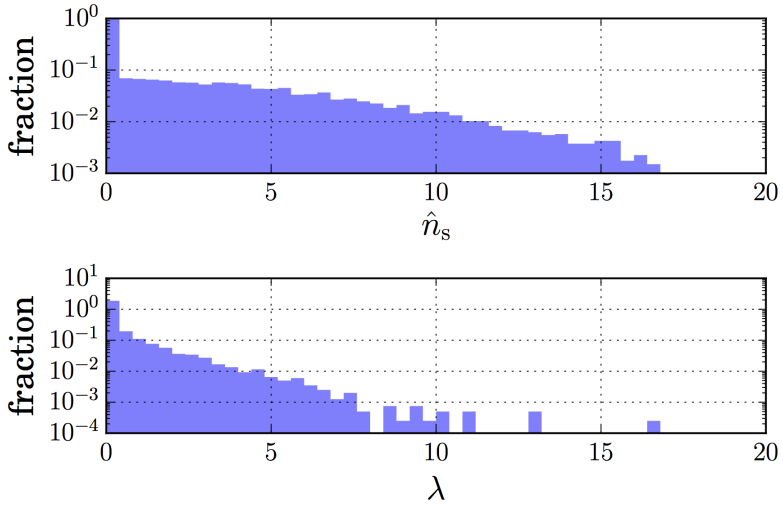

Figure 9.1.: Distribution of estimated n_s (upper panel) and test statistics values λ for 1000 background events without any signal injected. The maximization has a lower bound of $n_s \geq 0$.

extended to

$$\lambda = 2 \log \left(\frac{\mathcal{L}(\hat{n}_s, \hat{\gamma})}{\mathcal{L}(0)} \right) \times \text{sign}(\hat{n}_s) \tag{9.4}$$

to allow for negative values of n_s. sign is the signum function which return -1 for an argument smaller than zero and $+1$ otherwise. Allowing for negative values of n_s shows the symmetry of under fluctuations even clearer, see figure A.3.

Background and Signal Events In a next step 20 signal events are injected on top of the background events according to the signal space PDF (equation 9.2) and the experiment is repeated 10^4 times. The outcome is shown in figure 9.2. The likelihood method is (on average) able to recover the true number of injected events. The distribution of best fitted values of $\hat{n}_s$ follows a normal distribution. When decreasing the width of the point spread function σ, the width of this normal distribution becomes smaller, see figure A.4. A smaller value of σ corresponds to a better separation of signal and background events. Therefore the test statistics distribution λ gets shifted to larger values, since the trials with injected signal look more and more unlikely compared to the null hypothesis with no signal events.

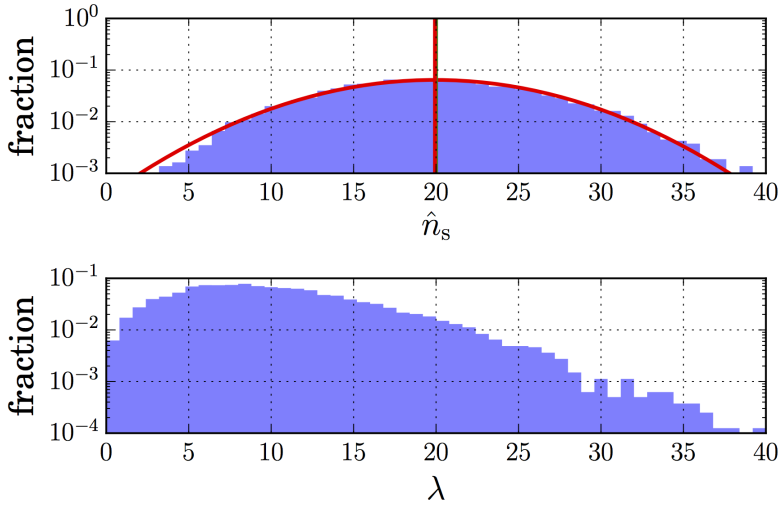

Figure 9.2.: Distribution of estimated n_s (upper panel) and test statistics values λ for 1000 background events and 20 injected signal events for default $\sigma = 0.5$. The red line is a Gaussian fit to the distribution and the green line the true value of injected events.

Over and Underestimating the Reconstruction Error Now the effect of under- or overestimation of the actual reconstruction error is tested. The error estimator is scaled by a factor f with the true error:

$$\sigma_{\text{est}} = f \times \sigma \tag{9.5}$$

where σ is the width of the normal distribution from which the reconstruction errors are drawn. The result is shown in figures A.5 and A.6 for $f = 0.5$ and $f = 2$.

The wrong estimation of the actual reconstruction error introduces a bias in the distribution n_s. This bias works in both directions, see figure 9.3. The effect on the test statistics distribution λ is always a shift to lower test statistics values. The result is a worse separation from the background only distribution. The wrong estimation of the reconstruction error σ will therefore always decrease the sensitivity of the analysis, see figure 9.3 again. The reason is that the assumption about the data (in the case the distribution of σ) is wrong, thus the wrong PDF is tested. Therefore this effect will also occur if the distribution of reconstruction errors will differ from the normal

distribution or otherwise not reflect the actual data.

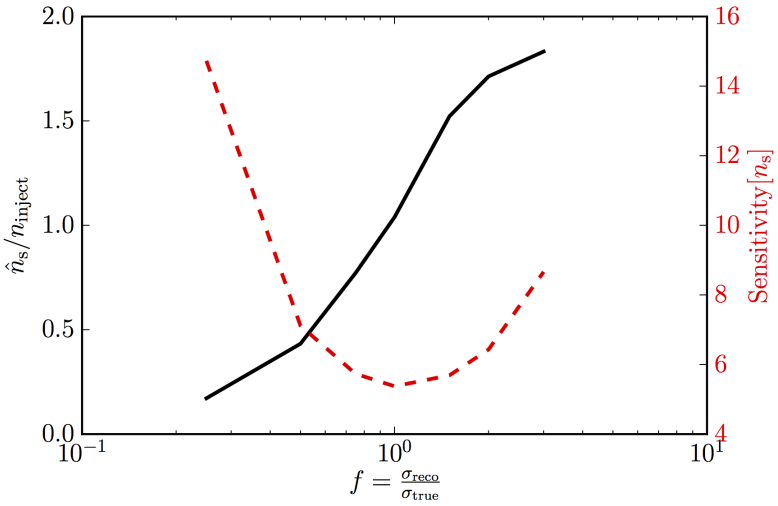

Figure 9.3.: Effect of wrong estimation of the reconstruction error σ on $\hat{n}_s$ and the sensitivity. The black (left) axis and the bold line shows the center of a Gaussian fit to the $\hat{n}_s$ distribution normalized with the true injected number of events. The red (right) axis and the dashed line shows the sensitivity in units of required signal events. 10^4 trials with 1000 background events and different values of signal events have been performed.

9.3. Space and Energy PDF

The test setup is now extended by an energy PDF in the likelihood function:

$$\mathcal{N}(x_i, y_i, \sigma_i) \longmapsto \mathcal{N}(x_i, y_i, \sigma_i) \times \mathcal{E}(E_i|\gamma)$$

The energy PDF mimics a power law energy spectrum $E^{-\gamma}$, see equation 9.3. This introduces a new parameter γ to the likelihood which describes the shape of the assumed signal spectrum. The spectral index γ is a free parameter which is maximized

simultaneously with n_s. Thus the likelihood function becomes two dimensional:

$$\mathcal{L}(n_s, \gamma) = \prod_{i=1}^{N} \left[\frac{n_s}{N} \mathcal{S}(\nu_i|\gamma) + \left(1 - \frac{n_s}{N} \mathcal{B}(\nu_i)\right) \right]. \tag{9.6}$$

Contour plots of the likelihood function are shown in figure 9.4 for background events only and in figure 9.5 for 20 injected signal events. The two parameters n_s and γ are

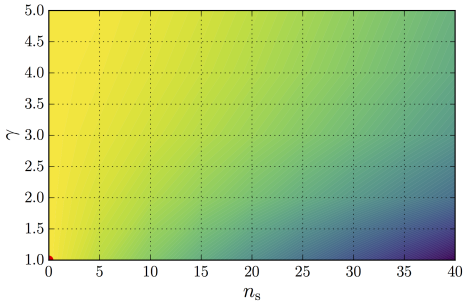

Figure 9.4.: Contur plot of two dimensional likelihood function with 1000 background events. The red dot marks the maximum of the function.

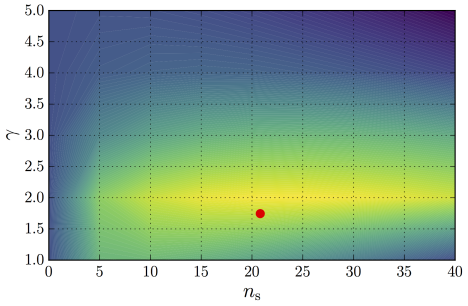

Figure 9.5.: Same plot as figure 9.4, but with 20 signal events injected. The injection spectrum is E^{-2}.

correlated. If $n_s = 0$ then the parameter γ is degenerate since any factor containing the spectral index γ vanishes. This can also be seen in figures 9.4 and 9.5.
The distribution of best fits in the two dimensional parameter is shown in figure 9.6. The distributions in n_s and γ are both Gaussian, but show a tail towards larger n_s and softer spectra.
The correlation coefficient [54] between n_s and γ was computed to to be 0.4 for an E^{-2} injection spectrum and 20 injecte signal events. A spatial over-fluctuation seem to be correlated with a softer signal spectrum which looks more like background. This behavior is expected since the likelihood is preferring a spatial under-fluctuation if it has a very hard spectrum or a softer spectrum when combined with a spatial over-fluctuation. In other words, fitting a spatial over-fluctuation can result in a large likelihood value even if the energy spectrum is not very signal like. On the other hand, a very signal like energy spectrum can still lead to the best fit, even in case of

lower number of events in spatial correlation.

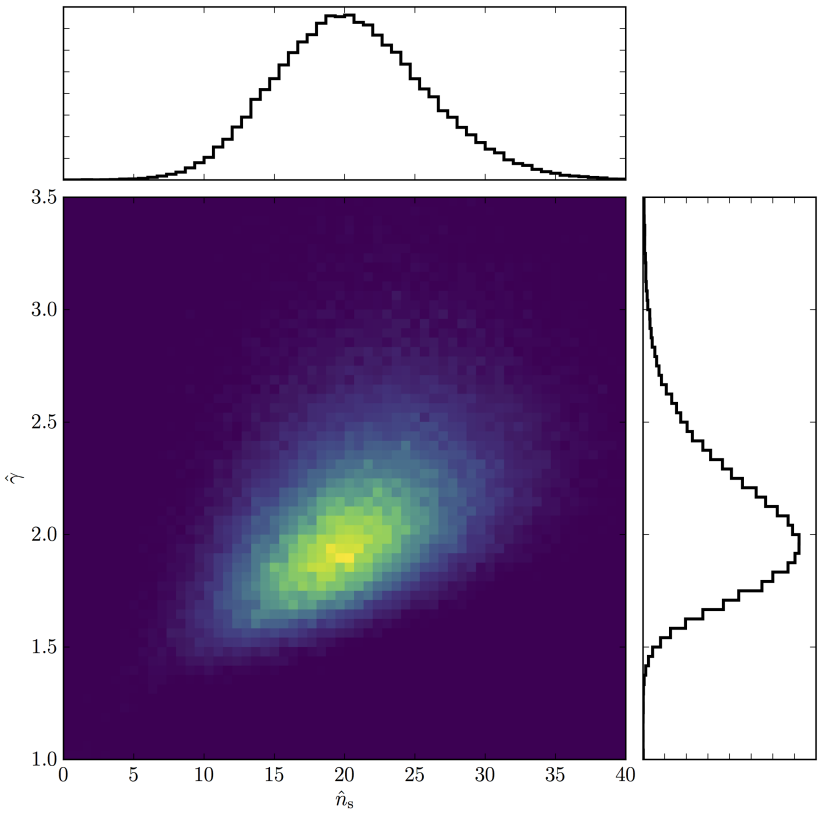

Figure 9.6.: Best fit parameter distribution for a two parameter fit with 1000 background events and 20 signal events. 10^4 trials have been performed.

Background Test Statistic Distribution The test statistic distribution λ from likelihood functions follows a χ^2 distribution first discussed by Wilks in 1938 if the likelihood function fulfills certain requirements [98]. The main requirement is that the likelihood function behaves Gaussian in each parameter. The degrees of freedom of the χ^2 distribution depend on the number of parameters in the likelihood function. Since no negative values of n_s are allowed, the requirements of Wilks theorem are violated [98, equation 3]. Only the test statistics values corresponding to $n_s > 0$ are thus expected to fulfill Wilks theorem. To fit the entire λ distribution a combination of χ^2

distribution and a δ function located at zero are fitted simultaneously, both with free normalization. The result is shown in figure 9.7. The fitted χ^2 distribution describes the but some disagreement can be seen. The fitted number of degrees of freedom is 1.51. The value is smaller than two since n_s and γ are correlated quantities. It also shows that the background test statistic distribution can be approximated by a χ^2 fit if needed for larger significance. The χ^2 fit is not perfect because Wilk's theorem is violated. When repeating the same with a one-dimensional likelihood test statistic distribution as shown in figure 9.8, a much better agreement with the χ^2 distribution can be seen. Therefore, in the actual analysis, the Monte Carlo generated distribution is used instead of the χ^2 distribution.

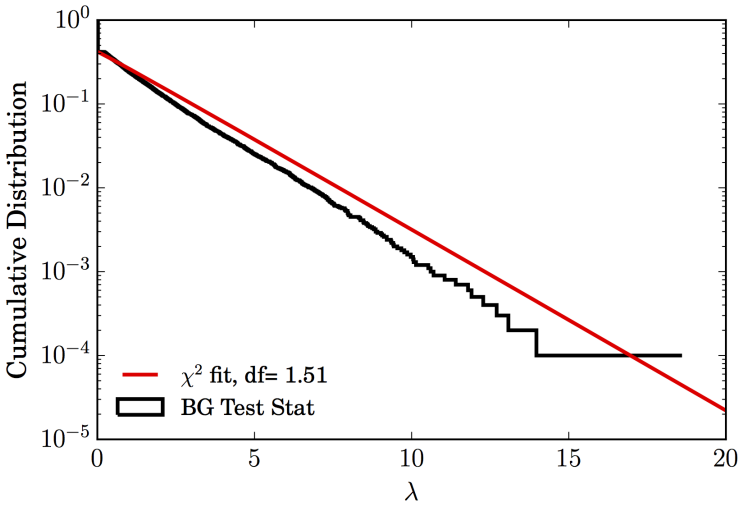

Figure 9.7.: The cumulative test statistic distribution for a two-parameter likelihood fit of 10^5 trials in black and a χ^2 plus δ peak fit is shown in the plot.

Sensitivities The sensitivity is defined in the usual way (see section 10) as 90% median sensitivity. Sensitivities are computed for three scenarios: First using only spatial information, second using spatial and energy information, but assuming a fixed spectral index and third using space and energy and fitting both signalness n_s and the spectral index γ. The sensitivity is given in units of signal events and are shown in figure 9.9.

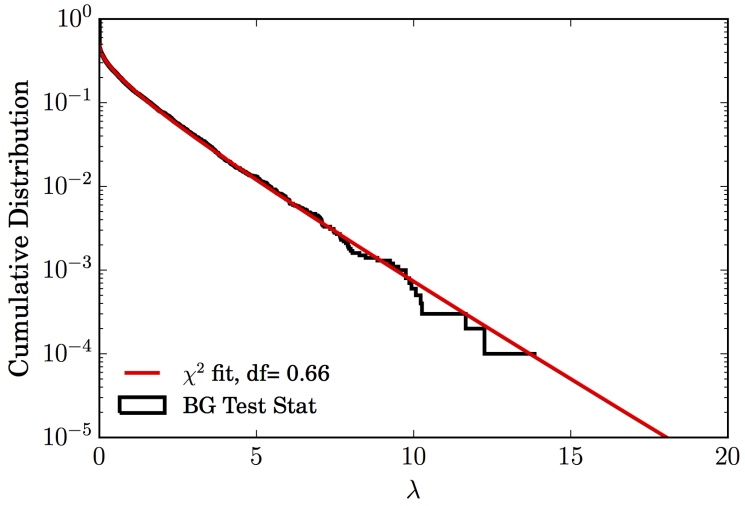

Figure 9.8.: The cumulative test statistic distribution for a one parameter likelihood fit of 10^5 trials in black and a χ^2 plus δ peak fit is shown in the plot.

The sensitivity for space only case (red curve) is independent of the signal spectrum since it does not have any influence on the likelihood function. The sensitivity of the likelihood function including the energy term, but not fitting the spectral index γ (blue curve) performs better than space only case for signal spectrum with a spectral index harder than ~ 2.2. For softer energy spectra ($\gamma > 2.2$) it performs worse. Naively one would expect the best sensitivity at $\gamma = 2$ where the hypothesis corresponds to the actually injected signal spectrum. To understand this behavior, one has to keep in mind that the likelihood method depends on the ratio of signal and background energy PDFs, see equation 8.7. Due to this feature, a signal spectrum $\gamma_{\text{inject}} < 2$ makes the assumed signal hypothesis $\gamma = 2$ wrong, but the background hypothesis ($\gamma = 4$) becomes even more incorrect. If the spectrum turns even harder (γ becomes smaller), the separation of signal and background events in the energy regime becomes more pronounced and the sensitivity thus becomes better. If on the other hand, γ becomes softer than the assumed spectral index of 2, both signal and background PDFs start to look more and more similar. The sensitivity becomes worse than space only case. Signal events are weighted as background events due to their low energy and thus sensitivity becomes worse.

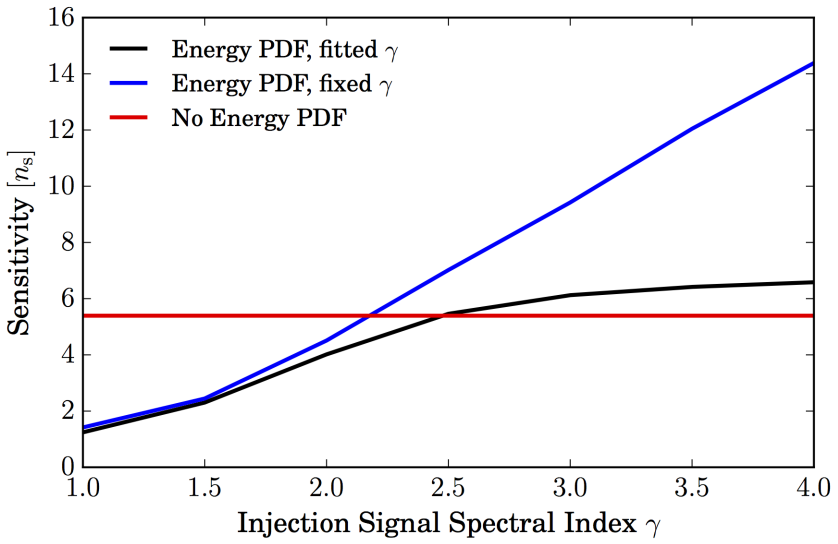

Figure 9.9.: The plot shows sensitivities as a function of the spectral index of the signal spectrum γ. The injected spectra for signal is shown on the x-axis and $\gamma = 4$ for the background. For the fixed signal spectrum $\gamma = 2$ is assumed. The sensitivity is given in units of number of signal events.

When also fitting the spectral index γ of the energy spectrum (black curve), the analysis performs better for spectra harder than about $\gamma_{\text{inject}} < 2.5$, compared with space only case. Comparing with the fixed energy assumption, the fitting always performs better, but at extreme hard spectra γ_{inject} both methods converge to the same sensitivity. Compared with space only analysis, the fitting of the spectrum performs better since it utilizes more information. It performs slightly worse than space only method for softer spectra where the signal and background energy PDFs look very similar. Here the fitting of the spectral index allows to better mimic background fluctuations which are then confused with signal and make the sensitivity worse.
Counterintuitively the spectral index fitting analysis does perform better than the one using a fixed spectral index, even if the true value ($\gamma_{\text{inject}} = 2$) is used. So fitting the spectral index results in a better performance than using the truth, as will be discussed next. Similar results were also found when using the real IceCube data and simulation, see figure A.8 in the appendix.

Interpretation of Fitted Spectral Index The use of the full likelihood function (equation 9.6) allows fitting the spectral index of the potential source. Events that contribute to this fit are only events which are spatially close by the source and are considered to be signal like[1]. Typically, these are only a few events. A general problem of likelihood methods is that they do not work very well at low statistics, see the discussion in [54]. Among other effects, low statistics can introduce bias. To quantify this bias, another test setup has been defined: Energies are randomly drawn from an E^{-2} distribution and a likelihood method is then used again to estimate the spectral index γ. This setup mimics the energy fit without any spatial PDF, so only the fitting of an energy spectrum with perfect signal-background separation is simulated. Figure 9.10 shows the average estimated spectral index as a function of the number of injected events. A bias towards softer spectral indices is clearly visible. So when interpreting the spectral index of a potential point source, one should always keep in mind that the spectral index was estimated by the likelihood method which does not perform well in the case of low statistics. Note also that the estimation of the signalness parameter $\hat{n}_\mathrm{s}$ is not problematic since it is estimated using the full data set also with all background events. So the estimate of the signal strength is not problematic, but the interpretation of the spectrum of a point source is critical.

This bias can also be understood in a descriptive way: The PDF is given by

$$\mathcal{E}(E|\gamma) = \frac{E^{-\gamma}}{\int_{E_\mathrm{min}}^{E_\mathrm{max}} E'^{-\gamma} \mathrm{d}E'}$$

and in case of a single event, the likelihood function is just the PDF. The likelihood maximisation then just implies

$$\frac{\mathrm{d}\mathcal{E}(E|\gamma)}{\mathrm{d}\gamma}\Big|_{\gamma=\gamma_0} = 0.$$

This has been done numerically and the result is shown in figure 9.11. To compute the expectation value for $\hat{\gamma}$ the function shown in figure 9.11 is weighted with the PDF for an injection spectrum for $\gamma_\mathrm{inject} = 2$. The expectation value for the estimated spectral index is $\langle\hat{\gamma}\rangle \approx 12.7$ which is consistent with the numerical experiment, see figure 9.10. The reason for this bias is that the distribution of energies and thus also the PDF is

[1] This is only true in a statistical sense. All events contribute to the fit of the signal energy spectrum, but with a weight depending on their spatial distance.

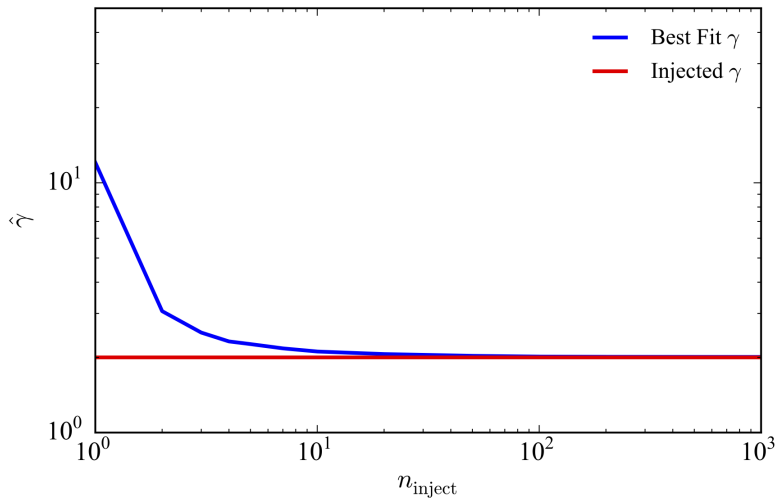

Figure 9.10.: Average estimated spectral index γ as a function of the number of injected signal events. Only the energy PDF is used without any background events.

extremely asymmetric. In case of only one event, it is much more likely to get a low energy event than a high-energy event. This finally causes the bias in the estimator of the spectral index. This is also illustrated in figure A.7 where PDFs for different spectral indexes are plotted. Note that the previous discussion was based on a single event for illustrative purposes, but the same argument holds for more than one event. This bias also explains why the overall sensitivity of point source likelihood analyzes perform better when the spectral index of the energy is fitted, compared with using the injection truth: Due to the few events at the sensitivity level, the signal spectrum is determined by only a small effective number of signal events. The likelihood estimator for the signal spectrum γ (not for the signalness n_{s}) is therefore in the low statistics regime where the likelihood estimator is biased as shown in this section. Because of this, using the true injection spectrum does decrease the sensitivity, compared to the fitting of the weights. In the test case here, it would be preferable to use a softer spectral index than the injection one to improve sensitivity.

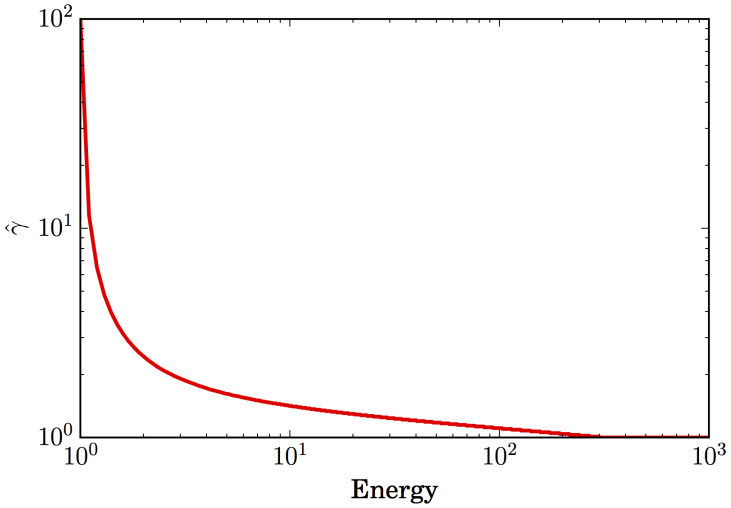

Figure 9.11.: Estimated spectral index γ for a single event with a certain energy.

9.4. Summary

This chapter discussed general features of point source likelihood analysis as used in this thesis. A simplified setup was used to study properties and behavior of this method. It was discussed how an under- or overestimation of the angular error introduces a bias and decreases the sensitivity. The analysis was extended by using energy information. Two methods were studied where one was fitting the spectral index of the energy spectrum and the other one was assuming a fixed spectrum. It was found that when using energy information, fitting the spectrum performs better than using the true value which is first confusing. Finally, this could be explained by a bias of the likelihood estimator in the energy regime due to low signal statistics.

10. Analysis and Unblinding

This thesis is testing supernovae as a potential source of high energetic neutrinos. This chapter describes the details of the actual neutrino search, including the implementation of the analysis, performance of different data sets, sub-catalogs to be tested, different neutrino light curve models and weight schemes. The outcome of the statistical test (unblinding) is shown.

10.1. Analysis Software

An essential part of this thesis was the development of a software package to perform the likelihood point source search described in the previous chapters. As this analysis utilizes many techniques not common in previous IceCube analyzes, in particular stacking analysis with variable weighting and time dependence, it was not possible to use existing software packages. The analysis software developed here is designed to become the standard software in the IceCube Collaboration for this type of time-dependent stacking analysis. One key feature of the software is its modular design, which is easy to maintain. The second key feature was performance; the usage of 7 years of IceCube neutrino data with the computationally demanding likelihood maximization, requires an optimization of the software regarding both computation time and memory consumption.
The software has been written in the programming language `python`[1] and intensive use of the `NumPy` [99], `SciPy`[2] and `AstroPy` [82] packages. `NumPy` provides fast, `Fortran` and `C` based routines for algebraic operations. `SciPy` provides functions for interpolation, root finding, minimization and numerical integration. `AstroPy` is mainly used to perform astrophysical and cosmological computation since many standard functions such as the conversion from redshift to luminosity distance are already implemented.

[1] `www.python.org`
[2] `www.scipy.org`

No further software dependencies are required for the analysis software and in particular no need for IceCube internal software. The software is also capable of performing analyses with different multi-messenger data, e.g., gamma-ray observations.

Generation of Simulated Background Data Sets To test the analysis and to perform simulated trials to compute the sensitivity, a procedure to simulate background event samples is required. These background event samples have the same characteristics as the real neutrino data but are guaranteed to have no signal contribution from a point source. Thus they correspond to the background hypothesis H_0. This is done by so-called *scrambling* of experimental data. The measured data sample is expected to contain mainly background events, so the influence of point source signal events on the observed energy spectrum or the rate of events as function of declination is negligible.
In the scrambling process, the right ascension of each event is changed to a random value between 0 to 2π. The event time is also changed to a random value within the livetime of the data set.
As previously discussed, the background event distribution is uniform in right ascension and time. After scrambling, any point source originally present in the measured data is lost, since the typical point-like clustering right ascension, declination and time in smeared out over the entire declination band and the full data taking period. As long as the point source is not bright enough to dominate the rate in the corresponding declination band, the scrambled data have the same underlying distribution as the original dataset, but without any point source (If any point source were bright enough, it would already have been seen in previous point source searches performing all-sky scans). This procedure is only possible due to the rotational symmetry of IceCube in right ascension and integration over time scales larger than few days[3].
Scrambling is used in this thesis to randomly generate background data sets that have the same characteristics as the real data but are statistically independent. The scrambling procedure is also the standard method used in IceCube for point source searches.

[3] On timescales smaller than about a day the detector also has a sensitivity dependence on right ascension due to the hexagonal shape of IceCube. But when observing over longer timescales, this effect is averaging out.

Signal Injector The injection of signal events is required for general testing and computing the sensitivity of the analysis. The challenge is to generate fake signal event sets which have the same characteristics as expected from a real neutrino point source given, a certain flux model.
Simulated point source signal events are generated from a Monte Carlo generated simulation data set. Signal simulation in IceCube is done by simulating neutrino events from all directions and force an interaction in or around the detector. The probability for this forces interaction is then converted into a quantity called `OneWeight`. `OneWeight` is the probability of detecting this certain event given a certain isotropic neutrino flux.
To simulate a point source, the full sample of simulated events coming from all directions is firstly restricted to only those events originating from a 10° declination band centered on the position of the source. This is done to only select simulated events with the same characteristics as those expected for true signal neutrino events from a real source. A declination band is sufficient since all detector effects depend on declination, rather than on right ascension. Any right ascension effects are averaged out due to the rotation of the detector concerning the equatorial coordinate system.
A correction factor $4\pi/A_{\text{band}}$ with A_{band} as the area of the declination band and 4π as the full sky is used to account for the declination band cut. The `OneWeight` values are multiplied with the assumed flux and spectrum

$$p_{\text{event}} = \texttt{OneWeight} \times \frac{4\pi}{A_{\text{band}}} \times \Phi_0 \times d_{\text{L}}^{-2} \times E_{\text{MCTruth}}^{-\gamma} \times T_{\text{DataTaking}}$$

where Φ_0 is the flux normalization, d_{L} is the distance of the source, γ is the spectral index, E_{MCTruth} is the Monte Carlo neutrino energy which was simulated and $T_{\text{DataTaking}}$ is the total time of data taking. This leads to a probability of each simulated event, p_{event}, to be measured for the simulated source.
The sum over all probabilities $n_{\text{exp}} = \sum_i p_i$ gives the expectation number of events from the simulated source. To generate signal event samples for a source, a Poissonian random number of expectation value $\lambda = n_{\text{exp}}$ is first drawn. In the second step, the corresponding number of signal events is randomly drawn from the simulated events according to their individual probabilities p_i. In a third step, the selected simulation events are shifted in the direction of the source. The true direction of the simulated event is shifted to the position of the source. The reconstructed direction is shifted

in parallel, maintaining the offset due to the reconstruction error. In the final step, simulated signal events are then added to the scrambled background events. This combined sample is then used as a test dataset with injected signal events.

Energy Spectrum Fit The analysis code performs a fit of the spectral index γ of the signal energy spectrum $E^{-\gamma}$. Therefore the spectral index γ has to be found for which the signal energy PDF $\mathcal{E}_{\mathcal{S}}(\gamma)$ maximizes the likelihood function. Since generation and evaluation of $\mathcal{E}_{\mathcal{S}}(\gamma)$ are computationally intensive, it is only evaluated once, on a grid in γ and stored during the likelihood maximization. The grid in the spectral index $E^{-\gamma}$ is done between $\gamma = 1$ and $\gamma = 4$ in steps of 0.1. Spectral indices γ that lie between those grid points are interpolated by Taylor expansion to second order. This procedure significantly speeds up the maximization process by about a factor of five.

Likelihood Maximization Instead of maximizing the likelihood function, the negative of the test statistic λ is minimized (see equation 8.7). Minimization is done using the L-BFGS algorithm [100, 101] from the `SciPy` package. The L-BFGS algorithm is optimized for minimization of larger dimensional functions with box-like bounding conditions and a low-dimensional correlation between the optimization parameters. It was found that this algorithm best fits the requirements of the analysis. The minimizer was also tested for stability and robustness. A measurable dependence on the seed was discovered. To overcome this, a brute force grid scan of the parameter space is first performed to find the initial seed and to avoid falling into a local minimum.

10.2. Sensitivity Test for Static Sources

The performance of the analysis code is compared with a previous analysis, a 7-year time independent point source search [20]. The two analyses do not share any code, so this test provides a full and independent cross check. The sensitivity for a single time-independent point source is tested as a function of declination of the source with seven years of IceCube data. The result is shown in figure 10.1. The best sensitivity is achieved at the horizon ($\sin(\delta) = 0$) and stays relatively constant in the up-going region. In the down-going region, it becomes much worse due to the larger background of atmospheric muons. This analysis performs slightly better than the reference analysis [20] in the down-going region ($\sin(\delta) < 0$). The reason for this is a finer binning

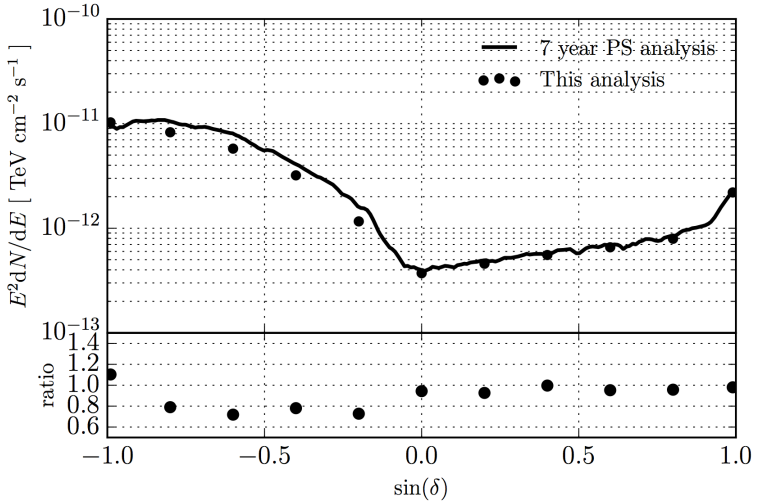

Figure 10.1.: Comparing point source sensitivities for a single static point source at different declination directions of the potential source assuming an E^{-2} spectrum. The reference analysis is the 7-year point source all-sky search [20].

in the pull correction as well as an improved seed for the final stage of the likelihood maximization.

The comparison shows that this analysis can reproduce previous results and also performs slightly better. Sensitivities for different scenarios, including energy information and knowledge about the spectral index of the spectrum is shown in the appendix, see figure A.8. The best sensitivity is achieved by including energy information in the likelihood and also leaving the spectral index free as a free parameter. This also provides a model-independence, since no signal spectrum has to be assumed. So in this thesis, an analysis method with a free-floating index of the signal energy spectrum is used.

10.3. Neutrino Light Curve Models

Adding time dependence to the likelihood analysis requires an assumption about the neutrino light curve since the signal time PDF $\mathcal{T}_\mathrm{S}(t)$ is generated from it. The neutrino light curve itself is, however, unknown. In principle, it could be modeled from the

observed optical light curve using model assumptions about the connection between optical and neutrino production.
Unfortunately, the optical light curves for most of the supernovae in the catalog (see chapter 6) have not been measured. Also, a connection between optical and neutrino light curve is not known. Therefore, several neutrino light curve scenarios, both model independent and also according to physical models, are tested.

Box Function Neutrino Light Curve To cover a broad set of potential neutrino light curves, a box-like shaped neutrino light curve is first assumed. A box light curve is unrealistic since it is unlikely that the source will suddenly start to produce neutrinos a constant rate and then also suddenly stop production. Nevertheless, the box can be used as a conservative and model-independent test. The light curve box function is defined by

$$LC_{\text{box}}(t) = \frac{\Theta(t - t_{\text{start}}) - \Theta(t - t_{\text{end}})}{\Delta T} \tag{10.1}$$

where Θ is the Heaviside function, t_{start} and t_{end} are beginning and end of the box function and $\Delta T = t_{\text{end}} - t_{\text{start}}$ is a normalization. If the box lies partially outside the data taking period t_1 and t_2, the normalization is changed to guarantee correct normalization

$$\int_{t_1}^{t_2} LC_{\text{box}}(t)\mathrm{d}t = 1. \tag{10.2}$$

In the following, the test scenario using a box-shaped neutrino light curve is simply called *box scenario*. The three width of box functions which are used as model-independent tests are 100, 300 and 1000 days. The first time windows of 100 and 300 days are motivated by expected duration of neutrino emission from type IIn supernova [37]. The additional 1000 days time window is used to cover also the emission from slowly evolving sources as discussed in [38]. Figure 10.2 shows the box time PDFs.

The test of the three box function light curve models provides an essential model-independent test for neutrino emission at these timescales. Any internal structure of the light-curve is unimportant since the box function is only sensitive to the integrated flux. The gain in model independence comes together at the cost of a decrease of sensitivity since an additional amount of background is accumulated. The box function light curve test can be understood as a time-independent analysis restricted

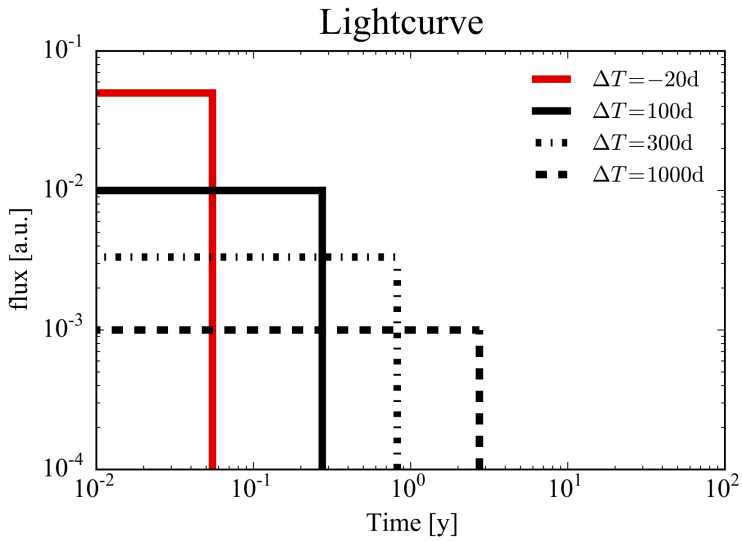

Figure 10.2.: Illustration of box-shaped time PDFs. The red box is used to indicate a search 20 days *before* the first optical detection of the supernova. The height of the different boxes is the result of normalization.

to the assumed extent of the box function[4]. Extending the length of the box function thus increases the amount of background events by the same proportion. This has been studied in figure 10.3. For short time windows, the sensitivity does not change when increasing the width of the search window and therefore the amount of background events. In this time regime, the sensitivity is completely signal dominated. The importance of background becomes more prominent when changing the sensitivity definition to a more background dependent measure as discussed in the caption. If the sensitivity requirement is instead defined to be above 90% of the background distribution instead of 50%, the background rate becomes very important. Thus, for an analysis such as this one, which aims for a discovery and thus p-value better than 50% a shorter time window is more optimal.

Search for Choked Jet Neutrinos A potential scenario for the production of high energetic neutrinos are choked jets [40]. In this scenario, neutrino production is expected in a jet inside of the supernova. Thus, the neutrino signal is expected before

[4]This of course only holds in case of a single source and not stacking where each source has a different assumed staring end of the box function

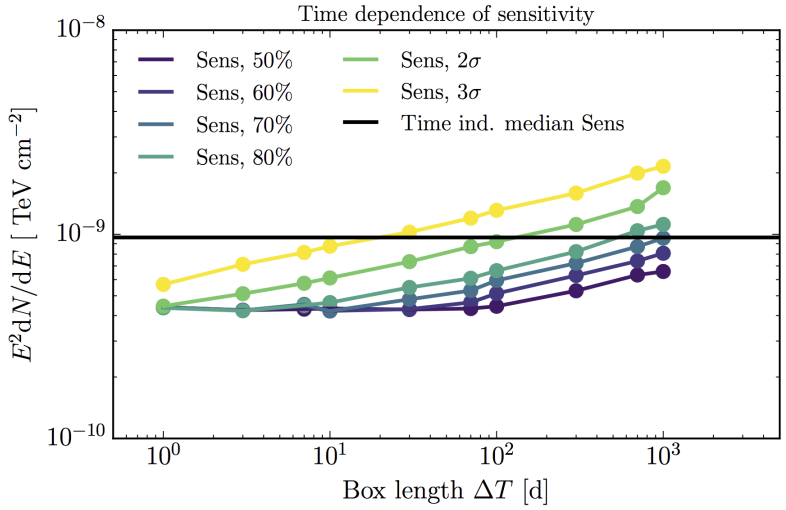

Figure 10.3.: Study of the sensitivity of a single source as a function of the time window for a box search. The sensitivity is quantified by the fluence, so time-integrated flux proportional to the total number of required signal events. The background is a function of the length of the time search window on the x-axis. Different definitions of sensitivity have been introduced here: The sensitivity is the required flux to measure a test statistic larger than the $X\%$ of the background distribution in 90% of the cases. Sens, 50% thus corresponds to the default sensitivity definition and for example Sens, 2σ is the required flux to be above the 2σ p-value of the background distribution in 90% of the cases. For the time-independent search seven years were assumed.

any optical signal is emitted. The duration of the neutrino emission should depend on the internal structure of the internal jet, but its exact duration is unknown. To be able to detect corresponding neutrinos without further knowledge, a time window from 20 days prior up to the first observation of the supernova is chosen as the time PDF (see figure 10.2). The duration of the actual neutrino bust is expected to be much shorter, but the width of 20 days is chosen to incorporate the uncertainty of the delay between neutrino burst and optical emission. Also, a supernova is typically not detected at the time of the explosion, but rather with some delay, because optical surveys only scan the night sky with a certain frequency and certain depth in brightness (see the discussion in chapter 6). The 20 days are chosen to be conservative in this

regard and to make sure to cover any potential neutrino signal.

Neutrino Light Curve Model for Supernovae Type IIn A recent work [38] has simulated the emission of neutrinos from type IIn supernovae with respect to the temporal evolution. The temporal evolution is parameterized as

$$LC_{\text{decay}}(t) \propto \left(1 + \frac{t}{t_{\text{pp}}}\right)^{-1} \tag{10.3}$$

where t is the time since the explosion and the parameter t_{pp} is connected with the physical properties of the supernova by equation 3.3 as discussed in chapter 3.5.1. The time PDF is again generated by normalizing the light curve function, ensuring that

$$\int_{t_1}^{t_2} LC_{\text{decay}}(t)\mathrm{d}t = 1. \tag{10.4}$$

A plot of this model can be seen in figure 3.5. To scan the parameter space of potential physics parameters, the values tested in this analysis are $t_{\text{pp}} = 0.02\,\text{yr}$, $0.2\,\text{yr}$ and $2\,\text{yr}$. This choice is made to cover the typically assumed values for supernovae type IIn. This scenario is called *decay* scenario for the rest of this thesis.

10.4. Splitting of Supernova Catalogs

The supernova catalogs of different classes of supernovae (chapter 6) are further split up into two sub-catalogs each. The idea is to have a small sample of nearby and potentially bright supernovae (*high quality sample*) and a larger sample of many potentially faint supernovae (*large sample*).
For the *high quality sample*, many properties such as explosion time and distance are likely to be well measured, since nearby objects will appear brighter making optical measurements easier. A weakness of the nearby, *high quality sample* is the lack of statistics, so it might not be a representative sample of the whole supernova class.
This is the key advantage of the *large sample* with stability against fluctuations of individual sources being traded off for a worse signal to noise ratio.It is important to note that the issue with deviations from the standard candle assumption in the *high quality sample* is dealt with by fitting individual weights (see chapter 8), but this method is only feasible for small samples due to the large computational demand.

So the *large sample* tests a much larger population, assuming they all have the same intrinsic brightness.
The catalogs of supernova types are split into the *high quality sample* and the *large sample* according to their expected signal strength. For each source j, the expected number of signal neutrinos n_j is computed using the following equation:

$$n_j = \underbrace{\frac{\Phi_0^j}{D_\mathrm{p}^2}}_{\text{Spacial Dependence}} \times \underbrace{\int_{t_\mathrm{start}}^{t_\mathrm{end}} \int_{E_\mathrm{min}}^{E_\mathrm{max}} LC_j^\nu(t) E^{-\gamma} \times Acc(t, \delta_j, E) \mathrm{d}t \mathrm{d}E}_{\text{Time Dependence}} . \tag{10.5}$$

Φ_0 is the intrinsic power of the source, D_p the proper distance, $LC^\nu(t)$ the expected neutrino light curve, $E^{-\gamma}$ is neutrino source spectrum, $Acc(t, \delta, E)$ the detector acceptance function or effective area, δ is the declination of the source and γ the assumed spectral index of the signal energy spectrum. The boundaries of the integral t_start and t_end are the times of beginning and end of the data set and the energy range of the analysis E_min and E_max.
Under the standard candle assumption, all Φ_0^j are the same. The first term (spatial dependence) takes the distance to the source into account. From a source further away, the expected number of neutrinos will naturally be lower. The second term, temporal dependence, computes the overlap of the assumed neutrino light curve $LC_j^\nu(t)$ with the detector acceptance function $Acc(t, \delta_j, E)$. The term can be understood as a measure of the fraction of the neutrino light curve that was sampled with which part of the detector. The time dependence is mainly a result of different IceCube seasons, so $Acc(t, \delta_j, E)$ is constant in time during each IceCube season. Since a neutrino light curve has to be assumed for this process, the following procedure is model dependent. The source candidates are ordered by their expected signal strength from strongest to weakest. The cumulative fraction of total expected flux is computed. This is shown with the blue lines in figure 10.4. The cut between the *high quality sample* and the *large sample* is fixed when about 70% of the expected total flux comes from the *high quality sample*. The cut cannot be set to exactly 70% because it goes in steps with the (integer) number of supernovae. Instead, the closest possible cut is made, independent if it is above or below. Therefore, the position of the cut is different for each time model since the assumed light curves are different.

So for each type of supernovae (type IIn, type IIp and type Ib/c), there are two

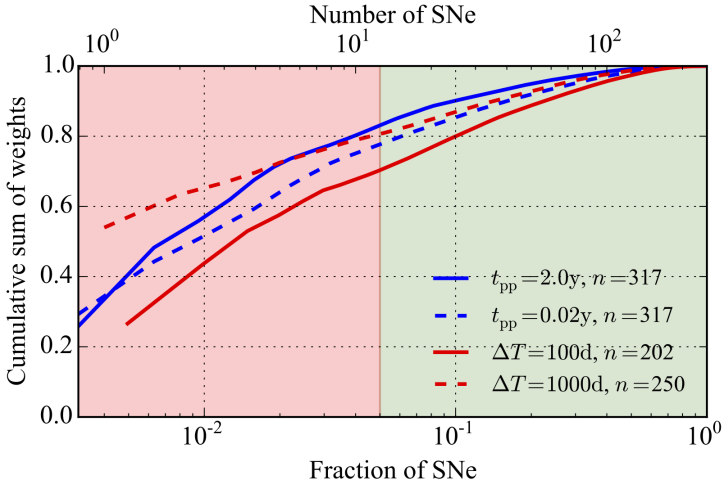

Figure 10.4.: Demonstration of the catalog splitting the type IIn supernovae. The *high quality sample* is indicated by the red shaded area and the *large sample* by the green shaded area. The blue lines indicate the fraction of total emitted flux as function of number of sources and fraction of total catalog.

independent sub-catalogs where the *high quality sample* typically contains about 5% of all detected supernovae and the *large sample* contains the remaining 95%.

10.5. Unblinding Procedure

To prevent bias during the development of analyses, the IceCube Collaboration follows the procedure of blind analysis. The analysis is developed blindly, meaning that only simulated and scrambled data is used to test it. The real experimental data is not allowed to be touched at all. The analysis is then reviewed internally by the collaboration. When the review is finished, the analysis is *unblinded*. *Unblinding* means that the analysis method is fixed now and may not be changed at all. The analysis is then performed once with the real experimental data, and the result is reported to the IceCube Collaboration. As part of the unblinding preparation, all background distributions are generated from scrambled background data sets to be able to immediately compute the p-value after unblinding. Figure 10.5 shows an

example of such a distribution with a χ^2 fit. Note that half of the distribution is located in the first bin at a $\lambda = 0$. This is because under-fluctuations are always fitted to zero, as previously discussed.

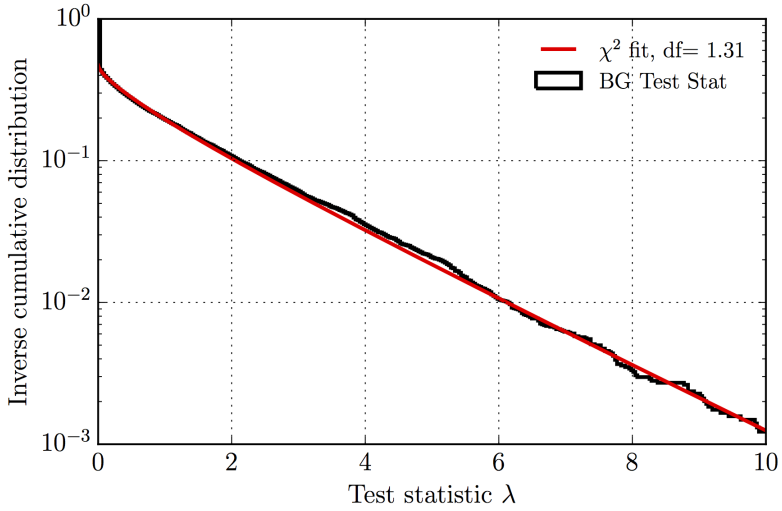

Figure 10.5.: Test statistic distribution λ computed before unblinding to estimate the sensitivity and be able to compute the p-value after unblinding. This distribution was computed for the test scenario of the choked jet scenario with a 20 days box function, *high quality sample*, fixed weights. Note the peak in the first bin where 52% of the distribution is located.

The analysis was performed with several scenarios shown in table 10.1. For each of the scenarios, the analysis has been executed with both catalog samples (*high quality sample* and *large sample*). For the *high quality sample* case, both a test with fixed weights assuming standard candles as well as a fitting of weights (see chapter 8) has been performed. The *large sample* was only tested with the fixed weight method because of computational limitations.

10.6. Unblinding Results

The analysis was approved for unblinding by the IceCube collaboration. The proposed scenarios described in the previous section were tested with the experimental data and a p-value was measured for each of the scenarios. The p-values are shown in table

		Choked jet SNe (Ib, Ic, Ib/c, IIb)	CSM SNe IIn	CSM SNe IIp
Choked jet model	$\Delta T = -20\,\mathrm{d}$	✓		
Generic box function	$\Delta T = 100\,\mathrm{d}$	✓	✓	✓
	$\Delta T = 300\,\mathrm{d}$	✓	✓	✓
	$\Delta T = 1000\,\mathrm{d}$	✓	✓	✓
CSM model $\sim \left(1 + \frac{t}{t_{pp}}\right)^{-1}$	$t_{\mathrm{pp}} = 0.02\,\mathrm{yr}$		✓	✓
	$t_{\mathrm{pp}} = 0.2\,\mathrm{yr}$		✓	✓
	$t_{\mathrm{pp}} = 2\,\mathrm{yr}$		✓	✓

Table 10.1.: Overview of the final set of tested scenarios and parameters in this analysis.

10.2.

A p-value larger than 50% is quoted if the maximum test static is $\lambda = 0$. Because under-fluctuations were fitted to zero (since negative signal-strength was not allowed in the likelihood maximization), the background distribution shows a pile up at zero (see figure 10.5). About 50% of the distribution is fitted to $\lambda = 0$. The p-value is thus not well defined for an outcome of zero. Therefore in table 10.2, these cases are marked with $> 50\%$.

In 27 of the total 48 tested scenarios, a p-value of $> 50\%$ is measured. One expects this from the background only hypotheis is H_0 purely due to background fluctuations. The most significant p-value is 0.62% for the case of the type IIp supernovae, the large sample and a box search with a duration of 1000 days. Small p-values are also seen for type IIn supernovae in the *high quality sample* on short time scales, both in the box light curve fit of 100 days (fixed weights an fitting of weights) as well as in the fitting of weights assuming the declining light curve CSM model with the shortest time scale.

None of the measured p-values is particularly low. Because many different scenarios have been tested, one would expect some larger p-values just due to background fluctuation. This will be discussed in the next chapter.

p-value	Box [days]				Decay [years]		
	100	300	1000	-20	0.02	0.2	2.0
IIn fixed	8.34	> 50	> 50		48.7	> 50	> 50
IIn fit	6.4	47.3	> 50		1.6	42.6	> 50
IIn large	> 50	> 50	> 50		> 50	> 50	30.7
IIp fixed	46.8	> 50	> 50		45.2	> 50	> 50
IIp fit	31.7	> 50	> 50		> 50	> 50	> 50
IIp large	49.4	> 50	0.62		37.5	10.0	45.4
Ib fixed	> 50	> 50	> 50	36.2			
Ib fit	> 50	> 50	34.8	> 50			
Ib large	> 50	> 50	6.66	41.9			

Table 10.2.: The table show pre-trial p-values of the tested scenarios in %. The label *box* refers to a box shaped neutrino light curve and *decay* to the CSM model. *Fixed* and *fit* refer to the *high quality sample* analyzed with fixed weights (standard candle assumption) and with fitting of weights. The term *large* refers to the *large sample*, tested with the fixed weights standard candle assumption.

11. Interpretation

In the previous chapter, the unblinding result of the analysis was shown. This chapter firstly performs a trial factor correction to account for the multiple trials that have been performed and then discusses the statistical interpretation. Finally, upper limits on the neutrino fluence of core-collapse supernovae are computed. Additionally, this chapter will discuss the implications of the result on the global picture of core-collapse supernovae as sources of high-energy neutrinos and their contribution to the diffuse flux of astrophysical neutrinos.

11.1. Trial-Factor Correction

So far, only the individual p-values of each tested scenario have been discussed. Due to the number of test scenarios, a trial-factor correction has to be applied before interpreting the unblinding result [102, 103]. The trial-factor takes into account that performing many experiments which individually have a low chance of success is likely to contain at least one successful result, purely due to the large number of trials. The computation of the trial factor is easy in case of independent experiments but requires more effort in the case where the different experiments are correlated, as for this thesis. In this work, the trial factor is computed by generating 500 data sets with scrambled background events and analyzing every data with all test scenarios. Through this procedure, correlations between the different test scenarios are correctly accounted for.

The distribution of the lowest p-values from each trial for different scenarios is then compared with the lowest p-value measured in the unblinding of the experimental data. The distribution of the lowest p-values is shifted to lower p-values compared with the distribution of a single test scenario. The final post-trial p-value is then computed based on the distribution of lowest p-values, see figure 11.1. The smallest pre-trial p-value of 0.6% then corresponds to a post-trial p-value of 19.5%, which is

consistent with the background only hypothesis H_0. An experimental result like the measured one is expected in one of five trials just by background fluctuation.

The smallest p-value is not the only relevant result, but also the distribution of

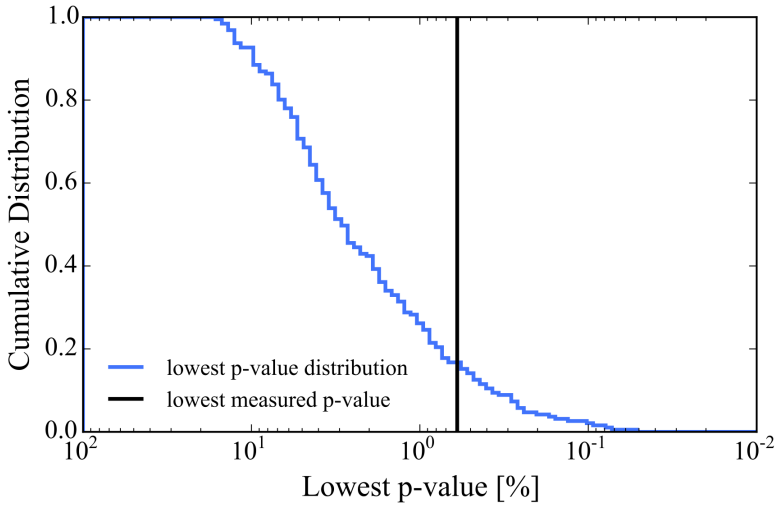

Figure 11.1.: Distribution of lowest p-values of all test scenarios from background trials and the lowest measured p-value from the unblinding of the analysis.

p-values of all test scenarios. If a small p-value is measured more often than expected from the background distribution, this would still be a significant difference in the background expectation, even if the smallest p-value is not significant alone. To test this, the background (null hypothesis H_0) distribution of p-values is tested against the measured distribution of p-values (see figure 11.2). A Kolmogorov-Smirnov-test [54] was performed to test consistency between the background distribution and the unblinding result. The outcome is a p-value of 29%.

Neither the smallest p-value nor the full distribution of the unblinded p-values shows a significant derivation from the background hypothesis H_0. P-values of 19.5% and 29% are small over-fluctuations from the background. One should check if these over-fluctuations grow with additional data in the future. At the current level, they are far below the 5σ level from where a discovery would be claimed. Rather than claiming a discovery, the upper limits of the experimental results are studied.

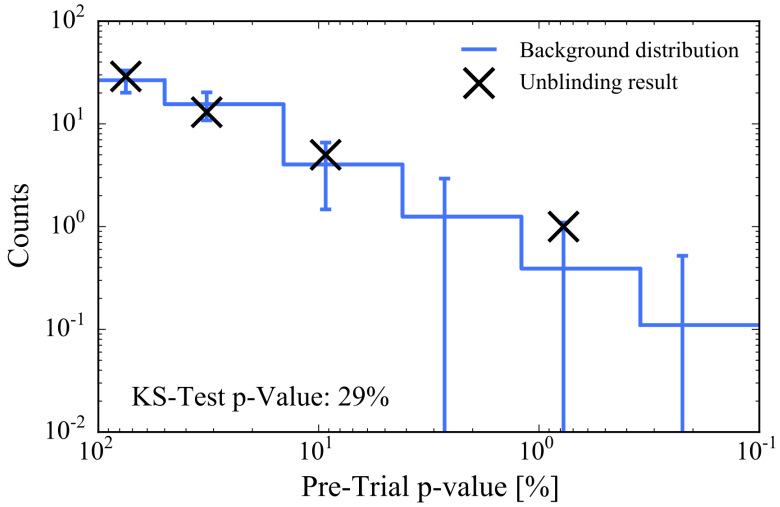

Figure 11.2.: Distribution of p-values of all test scenarios from background trials and the p-values from the unblinding. Data and background expectation agree, as indicated by the KS test.

11.2. Upper Limit on Single Source Fluence

Since no signifianct excess was found in the unblinding of the data, upper limits on the neutrino fluence of supernovae are computed. The upper limit at a certain confidence level is defined as the strengh a source can have while still remaining below the measured strength with the given confidence level. In case of the 90% confidence upper limit, it is the signal strength which leads to an experimetal result as strong as the observed one in only $100\% - 90\% = 10\%$ of the cases. In this thesis, the upper limit on the fluence is either given by the 90% upper limit on the measured excess, or, if no excess was measured ($\lambda = 0$), the 90% upper limit is the sensitivity.
The upper limit is given in units for a total emitted neutrino energy of a single source. To compute this, isotropic emission, a power spectrum ($E^{-\gamma}$) and a neutrino light curve as discussed in the previous chapters, are assumed. The total neutrino energy is given by

$$E_\nu^{\mathrm{tot}} = \Phi_{\text{Upper limit}} \times 4\pi d_{\mathrm{ref}}^2 \times \int_{-\infty}^{\infty} LC_\nu(t)\mathrm{d}t \times \int_{E_{\mathrm{low}}}^{E_{\mathrm{high}}} E \cdot E^{-\gamma}\mathrm{d}E$$

where $E_{\text{low}} = 10^2\,\text{GeV}$ and $E_{\text{high}} = 10^7\,\text{GeV}$ are the borders of the energy integration range. The measured upper limit flux is given by $\Phi_{\text{Upper limit}}$. The total neutrino energy now still depends on the assumed spectral index. Figure 11.3 shows the effect of different assumed spectral indices on the sensitivity in terms of total neutrino energy. The total neutrino energy increases if a softer energy spectrum is assumed. This is because the difference from signal to background energy spectrum becomes smaller and thus signal background seperation no longer works well anymore for softer signal spectra, so the analysis is less sensitive. For the following discussion and plots,

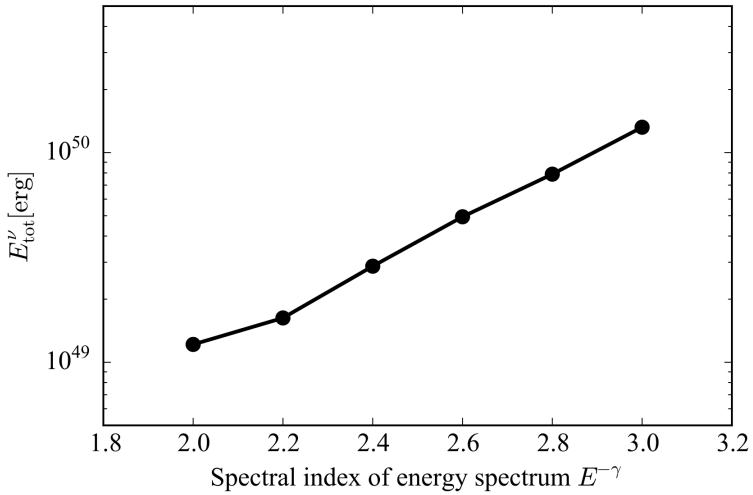

Figure 11.3.: Sensitivity for the analysis of type IIn supernovae, using the small catalog sample, fixed weights and a time window of 300 d as a function of assumed spectral index γ of the sources. A positive correlation is observed. By changing the spectral index from 2 to 3, the required total neutrino energy increases by about an order of magnitude.

an energy spectrum of E^{-2} is assumed unless otherwise stated.
The unblinding of the data resulted in upper limits for each catalog, scenario and time model parameter, as well as for fitting weights and assuming fixed weights. Since the *large sample* is statistically independent of the *high quality sample*, we can select the stronger upper limit of the two catalogs for each corresponding analysis scenario. It turns out that the upper limits from the *high quality sample* always gives a better upper limit than the *large sample* catalog. This is not surprising since the

signal expectation of the *high quality sample* is roughly two times larger with lower background.

The *high quality sample* was analysed both using fixed weights and fitting of the weights. Fitting of weights and the fixed weight assumption result in similar upper limits. The quoted upper limit is always the weaker upper limit of the both to be conservative. Table 11.1 shows the upper limits for the different analysis scenarios assuming an E^{-2} spectrum. For comparison, the upper limits are also illustrated in figures 11.4 and 11.5 for the two light curve models. Here only the upper limits from the *high quality sample* are shown. Upper limits are also computed on individual sources, results can be found in the appendix (A.4).

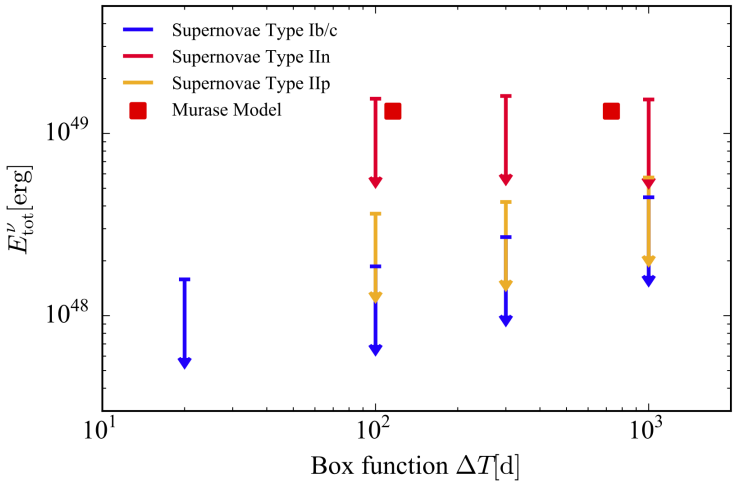

Figure 11.4.: Upper limit on total neutrino energy assuming a box-like neutrino light curve. The model prediction by Murase et al. [37] is also shown for comparison.

For the case of supernovae of type IIn, the upper limits are consistent with both of the theoretical predictions [37]. The energy range of the order of $10^{48} - 10^{49}$ erg. This is consistent with a typical assumed total kinetic energy of 10^{51} erg in a supernova and an acceleration efficiency of about 10%. Further on, the protons undergo hadronic interactions and transfer about 10% of their energy to neutrinos [37]. Thus the outcome of the unblinding is not in tension with standard assumptions about supernova emmission mechansim.

		Upper limit $E_\nu^{\mathrm{tot}}[10^{49}\,\mathrm{erg}]$		
		Small sample		Large sample
		Fixed	Fit	Fixed
Supernova Type IIn				
Box function	$\Delta T = 100\,\mathrm{d}$	1.29	1.55	5.49
	$\Delta T = 300\,\mathrm{d}$	1.21	1.60	4.94
	$\Delta T = 1000\,\mathrm{d}$	1.33	1.53	5.32
CSM model $\sim \left(1 + \frac{t}{t_{pp}}\right)^{-1}$	$t_{\mathrm{pp}} = 0.02\,\mathrm{yr}$	1.78	2.45	5.25
	$t_{\mathrm{pp}} = 0.2\,\mathrm{yr}$	1.53	1.96	4.95
	$t_{\mathrm{pp}} = 2\,\mathrm{yr}$	1.49	3.10	6.37
Supernova Type IIp				
Box function	$\Delta T = 100\,\mathrm{d}$	0.327	0.363	1.22
	$\Delta T = 300\,\mathrm{d}$	0.421	0.284	1.35
	$\Delta T = 1000\,\mathrm{d}$	0.573	0.401	2.46
CSM model $\sim \left(1 + \frac{t}{t_{pp}}\right)^{-1}$	$t_{\mathrm{pp}} = 0.02\,\mathrm{yr}$	0.578	0.355	2.16
	$t_{\mathrm{pp}} = 0.2\,\mathrm{yr}$	0.614	0.263	2.70
	$t_{\mathrm{pp}} = 2\,\mathrm{yr}$	0.644	1.38	2.77
Supernova Type Ib/c				
Box function	$\Delta T = 100\,\mathrm{d}$	0.185	0.168	0.719
	$\Delta T = 300\,\mathrm{d}$	0.248	0.270	1.35
	$\Delta T = 1000\,\mathrm{d}$	0.355	0.446	1.67
	$\Delta T = -20\,\mathrm{d}$	0.159	0.126	0.594

Table 11.1.: The upper limits in terms of total energy emitted in neutrinos of the different tested scenarios, the weights were fitted. The energy is computed assuming an E^{-2} spectrum and integration between $10^2\,\mathrm{GeV}$ and $10^7\,\mathrm{GeV}$. Note that the upper limit from the large sample catalog are significantly worse.

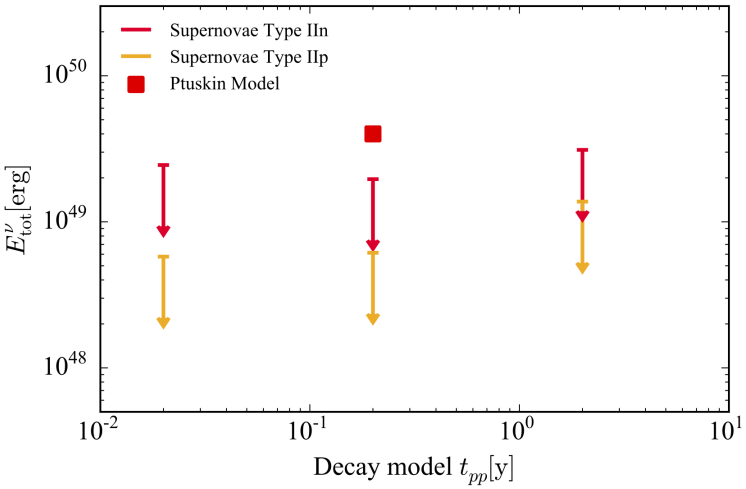

Figure 11.5.: Upper limit on total neutrino energy assuming a $LC_\nu \propto (1 + t/t_{\rm pp})^{-1}$ neutrino light curve as predicted by [38].

11.3. Contribution to the Diffuse Astrophysical Neutrino Flux

This section discusses the limits on the contribution of supernovae to the total diffuse astrophysical neutrino flux. The computation of the diffuse flux was discussed in chapter 7. The computation is repeated with the upper fluence limits for individual sources discussed in the previous section. The ΛCDM cosmological model with the latest Planck results [83] is used.
The main source of uncertainty of the diffuse flux is the supernova rate. CANDELS and CLASH [87] and Madau et al. [89] are dedicated studies aiming to measure the rate of core-collapse supernovae. The difference between the two dedicated studies CANDLES and CLASH [87] and Madau et al. [89] is highest up to about a factor of 2. This is illustrated in figure 6.11. The largest core collapse rate by the CANDELS and CLASH survey [87] is chosen as the conservative upper limit for the following calculations. The fraction of different types of core-collapse supernovae on the total rate was taken from the CANDELS and CLASH survey [87].
The fluence upper limits for a single source in the different tested scenarios are very similar. The following diffuse flux upper limit correspond to the box function with length $\Delta T = 100\,\mathrm{d}$ for supernovae types IIn and IIp and the box function of $\Delta T = -20\,\mathrm{d}$ for supernovae type Ib/c. It should be noted that the diffuse flux does not depend on the time profile of the source, but only on the total neutrino fluence. The results of the diffuse flux computation are shown in figure 11.6 for an $E^{-2.5}$ spectrum and in figure A.9 in the appendix for an E^{-2} spectrum. Both plots are compared with the diffuse astrophysical neutrino flux measured by IceCube [19]. The spectrum of $E^{-2.5}$ is chosen to mimic the observed diffuse flux and to study the potential supernova contribution. The second spectrum E^{-2} is chosen to study a very hard model spectrum. The effect of different spectral index assumptions is shown in the appendix in figure A.10 as well. This energy range was chosen to cover the central 90% region of the analysis.

The 90% upper limits contribution on the diffuse flux is 12.8% for type Ib/c supernovae, 27.5% for type IIn supernovae and 96.2% for type IIp supernovae. It should be noted that all upper limits drop by 49% when changing to the supernova rate estimate by Madau [89]. The weakest upper limit to the diffuse flux comes from the supernovae type IIp. Type IIp supernovae are the most prominent sub-population of core-collapse supernovae (about 50% of all CCSN [87]) and therefore can result in a

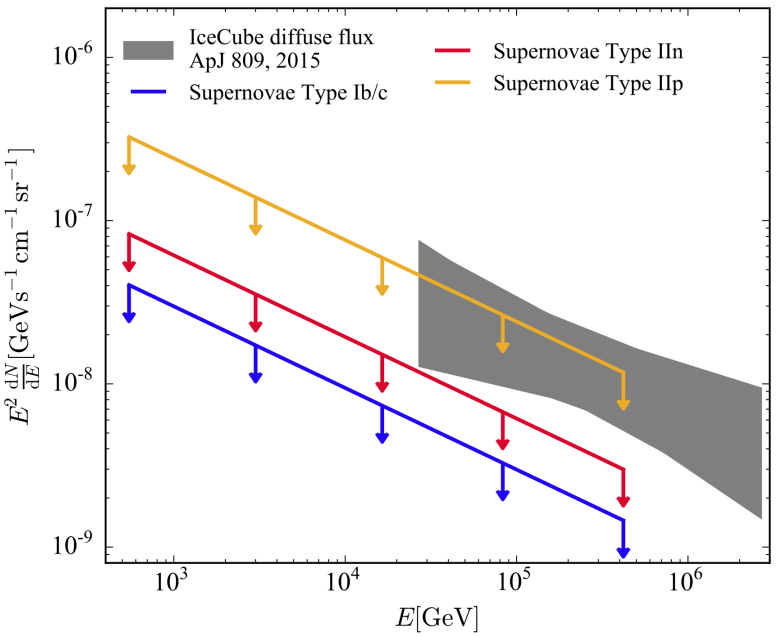

Figure 11.6.: Diffuse flux upper limits for the different supernova types assuming an $E^{-2.5}$ energy spectrum compared with the measured diffuse astrophysical neutrino flux [19]. The energy range plotted here is the central 90% energy region.

large diffuse flux, even if the limit of an individual supernova is quiet strong. Thus, type IIp supernovae can still explain the observed diffuse neutrino flux. The fraction of the total core-collapse supernovae of type IIn supernovae is about 6.4% and the fraction of type Ib/c supernovae about 25% [87]. Therefore, the upper limits on the contribution to the diffuse flux are quiet strong.

This thesis puts limits on the maximal contribution of certain classes of core-collapse supernovae to the diffuse astrophysical neutrino flux. While the maximal contribution of type Ib/c and IIn are 12.8% and 27.5%, supernovae type IIp cannot be ruled out as the source of the entire diffuse flux.

The potential source classes were already discussed by Kowalski [5] in 2014, where an overview plot of potential source classes was presented. The plot is updated with the upper limits from this thesis, excluding supernovae type IIn and Ib/c (see figure 11.7). The long-term strategy of neutrino astronomy is to shrink the parameter space

of source classes in figure 11.7. This can be done by untargeted searches like the IceCube point source search [20] which shrink the parameter space or by catalog searches which exclude certain classes of sources like this thesis. Figure 11.7 also shows that there are still many possibilities for source classes which have not been ruled out. Catalog-based correlation studies should be performed to further shrink the potential source classes space and finally to identify the sources of the astrophysical neutrinos and also potentially the sources of cosmic rays.

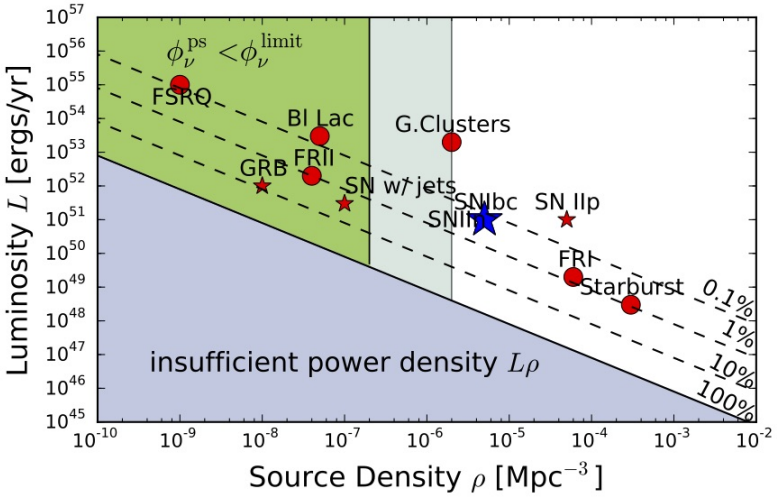

Figure 11.7.: Figure taken from [5]. The potential source classes of the diffuse neutrino flux are shown as a function of source density and energy output per source. Further, constrains from point source searches are also shown. This thesis excluded supernovae of type IIn and Ib/c, indicated by the blue star, as potential sources.

12. Conclusion and Outlook

This thesis aims to test the correlation between core-collapse supernovae and high energetic neutrinos. A supernova catalog was compiled from various optical observations to be correlated with seven years of IceCube neutrino data. The existing stacked likelihood point source method, the current standard in neutrino astronomy, was further developed by the author of the thesis. The main improvement was the fitting of weights in the stacking process. This improves the sensitivity in case of limited information about the stacked sources and introduces a general model independence about the weights. Fitting of weights has never been done before in neutrino astronomy.
In this thesis, several sub-classes of core collapse supernovae have been tested for their neutrino emission. The final results show an overall p-value of 20% deviation from the background. This deviation is too low to claim a discovery, so upper limits have been computed. The contribution of core-collapse supernovae to the measured diffuse astrophysical neutrino flux has been estimated as sub-dominant for supernovae type Ib/c and type IIn. Supernovae type IIp is not yet ruled out as the source of the diffuse flux. This means that the main source of the observed high energetic neutrinos is still to be discovered.
In the future, this analysis can be improved in several ways. First, a further optimization of the analysis code combined with additional computational resources can increase the number of sources which can be tested using floating weights. This might remove the catalog splitting into two samples, utilizing more of the existing data in the most advanced analysis method. Second, the accumulation of data over time will increase the number of observed nearby supernovae and the corresponding neutrino data, thus increasing potential signal. Last, the upcoming all-sky optical sources with low cadence will potentially improve this analysis the most. Discovering supernovae immediately after their explosion and the potential availability of optical lightcurves will allow to model their potential neutrino emission on a source by source base instead of assuming a generic light curve for all supernovae.

Appendix A.

Additional Material and Plots

A.1. Additional Plots of Distribution of IceCube Data

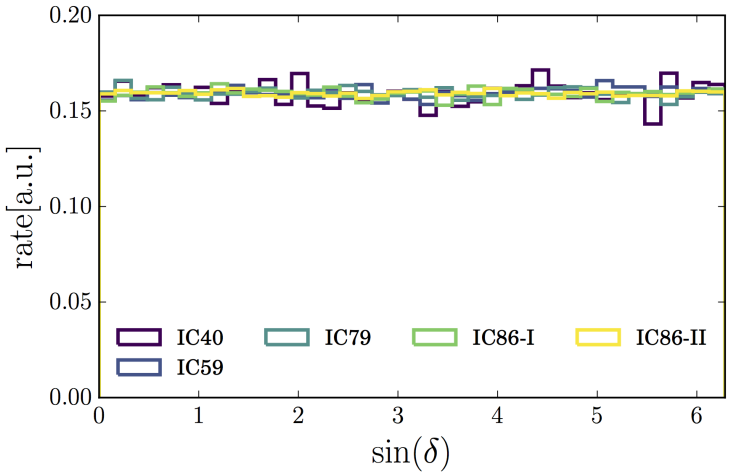

Figure A.1.: Right ascention distribution of experimental data for all used seasons.

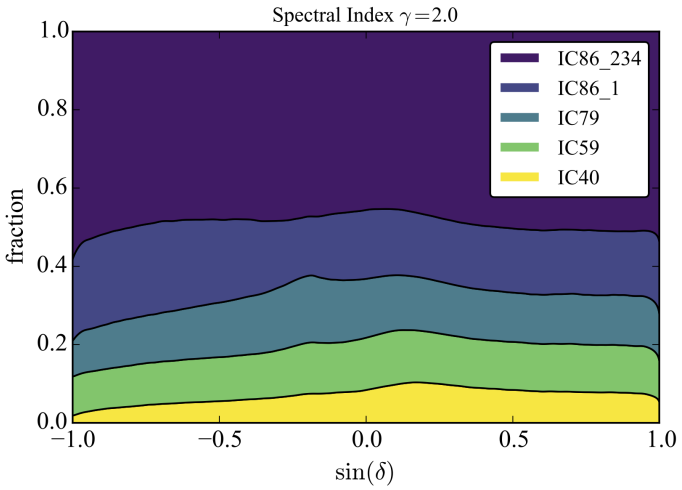

Figure A.2.: The relative signal strength of a time independent E^{-2} signal spectrum in the different seasons. The later IceCube seasons contribute about 50% to the total signal expectation.

A.2. Additional Plots of Likelihood Behaviour

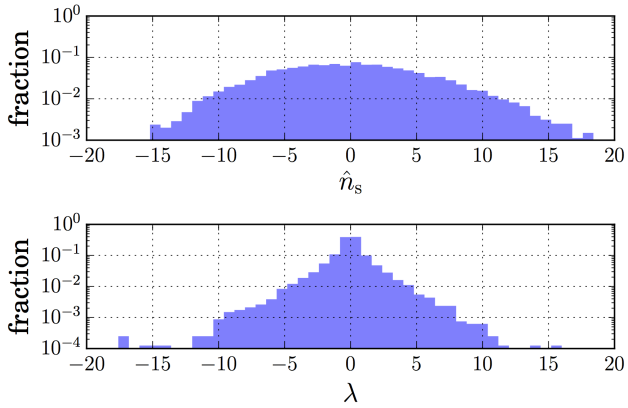

Figure A.3.: Distribution of estimated n_s (upper panel) and test statistics values λ for 1000 background events without any signal injected. The mimization is unbound, any value of n_s is allowed, unlike figure 9.1.

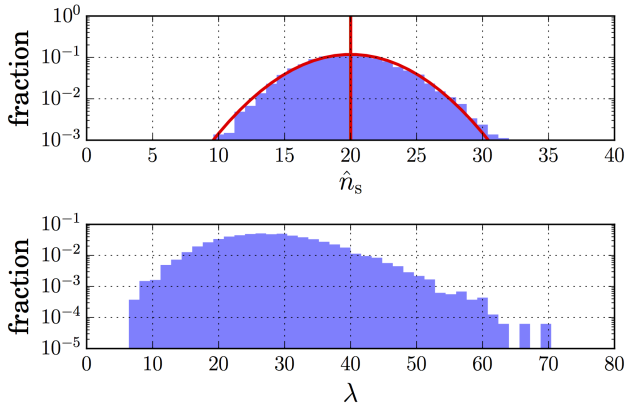

Figure A.4.: Distribution of estimated n_s (upper panel) and test statistics values λ for 1000 background events and 20 injected signal events. The red line is a Gaussian fit to the distribution and the green line the true value of injected events. A $\sigma = 0.25$ has been used, compared to the default settings, see figure 9.2.

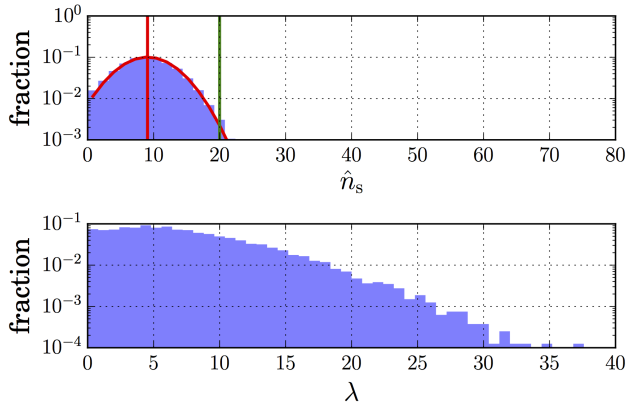

Figure A.5.: Distribution of estimated n_s (upper panel) and test statistics values λ for 1000 background events and 20 injected signal events for default $\sigma = 0.5$. The red line is a Gaussian fit to the distribution and the green line the true value of injected events. The angular reconstruction error is underestimated with a factor $f = 0.5$, see equation 9.5.

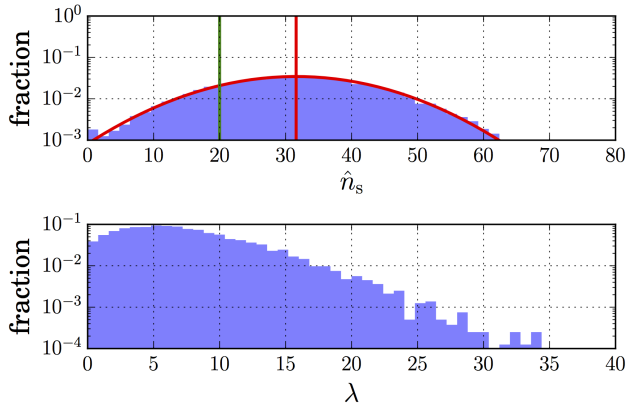

Figure A.6.: Distribution of estimated n_s (upper panel) and test statistics values λ for 1000 background events and 20 injected signal events for default $\sigma = 0.5$. The red line is a Gaussian fit to the distribution and the green line the true value of injected events. The angular reconstruction error is overestimated with a factor $f = 2$, see equation 9.5.

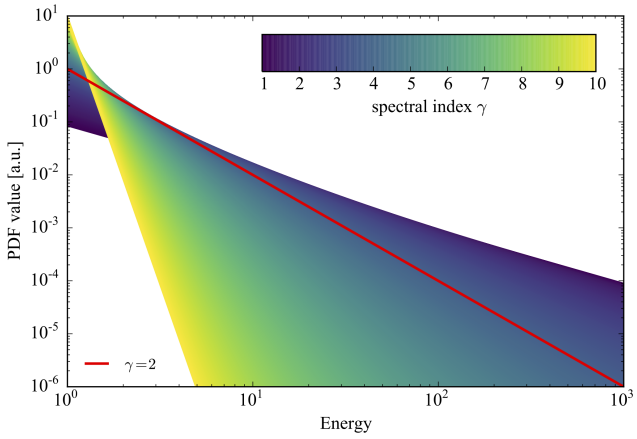

Figure A.7.: Plot of PDFs for $E^{-\gamma}$ spectra normalized for the range between 10^0 and 10^3. The injection index of E^{-2} is highlighted.

A.3. Point Source Sensitivity with and without Energy

Figure A.8 shows the point source sensitivites comparison for ana anlysis utilizing spacial information only or using energy information (both with fitting the spectrum and using the injected truth). The interesting behaviour that in the upgoing region ($\sin(\delta > 0)$ the utilizing of spacial information under the true spectrum assuption performs worse than not utilizing energy information at all was discussed in section 9.3. In the donw-going region, this effect cannot be observed any more. This is due to the harder energy cuts, where the discussed effect (section 9.3) does not apply.

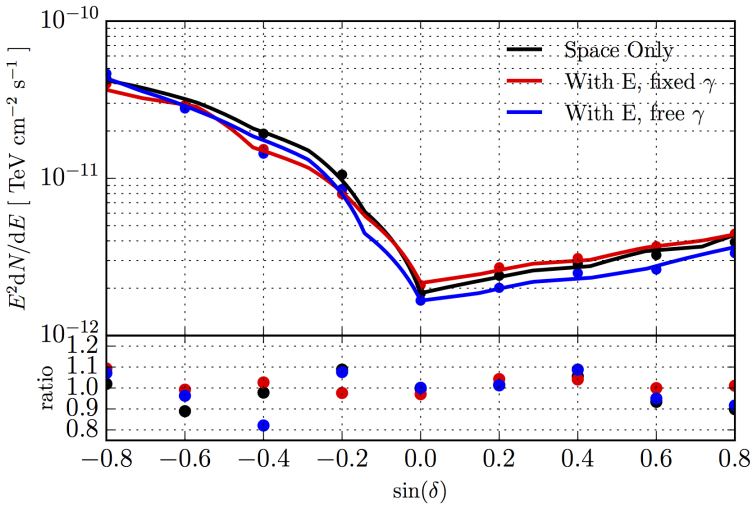

Figure A.8.: Comparing of point source sensitivities for a single static point source at different declination directions of the potential source assuming an E^{-2} spectrum. Tested are the analysis using only spacial information, fitting the energy spectrum and also assuming a fixed energy spectrum (also E^{-2}). The reference analysis is the 1-year point source all-sky search [20], but only the IC86-I data set was used. The reference is always shown with respect to the sime analysis setup in the reference analysis.

A.4. Upper Limits on Individual Sources

This section shows upper limits on individual supernovae as it was computed post unblinding. The assumed spectrum is always E^{-2}.

A.4.1. Supernovae Type IIn

Name	ra [rad]	dec [rad]	time	Flux Upper Limit [10^{49} erg]
CSS140111:060437-123740	1.59	-0.22	2013-12-24	49.8
iPTF13cjz	0.52	0.332	2013-08-02	11.7
PSN J13522411+3941286	3.63	0.693	2015-01-09	16.8
PSN J14041297-0938168	3.68	-0.168	2013-12-20	4.8
PTF10aaxf	2.54	0.166	2010-11-03	29.5
PTF10fqs	3.22	0.252	2010-04-16	22.2
SN2008S	5.39	1.049	2008-02-01	5.3
SN2009kr	1.36	-0.274	2009-11-06	19.1
SN2011an	2.09	0.287	2011-03-01	65.3
SN2011ht	2.65	0.905	2011-09-29	6.6
SN2012ab	3.24	0.098	2012-01-31	64.18
SN2013gc	2.13	-0.49	2013-11-07	28.4

A.4.2. Supernovae Type IIp

Name	ra [rad]	dec [rad]	time	Flux Upper Limit [10^{49} erg]
iPTF13aaz	2.96	0.228	2013-03-21	1.0
SN2012A	2.73	0.299	2012-01-07	1.0
SN2012aw	2.81	0.204	2012-03-16	1.0
SN2014bc	3.22	0.826	2014-05-19	3.0

A.4.3. Supernovae Type Ib/c

Name	ra [rad]	dec [rad]	time	Flux Upper Limit [10^{49} erg]
iPTF13bvn	3.93	0.033	2013-06-17	4.0
iPTF15afv	5.92	0.601	2014-04-30	3.8
MASTER OT J120451.50	3.16	0.471	2014-10-28	1.0
PTF11eon	3.53	0.823	2011-06-01	1.1
SN2008ax	3.28	0.727	2008-03-03	1.6
SN2008dv	0.95	1.267	2008-07-01	1.2
SN2010br	3.16	0.777	2010-04-10	4.1
SN2011jm	3.38	0.046	2011-12-24	1.8
SN2012cw	2.68	0.06	2012-06-14	4.3
SN2012fh	2.81	0.434	2012-10-18	1.1
SN2013df	3.26	0.545	2013-06-07	1.7
SN2014C	5.92	0.601	2014-01-05	2.3

A.5. Diffuse Upper Limits

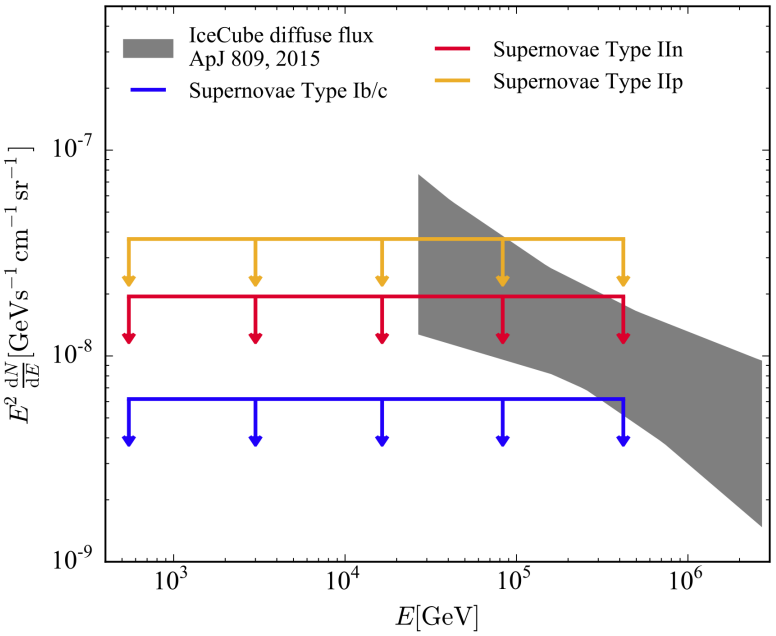

Figure A.9.: Diffuse flux upper limits for the different supernova types assuming an E^{-2} energy spectrum compared with the measured diffuse astrophysical neutrino flux [19]. The energy range plotted here is the central 90% energy region.

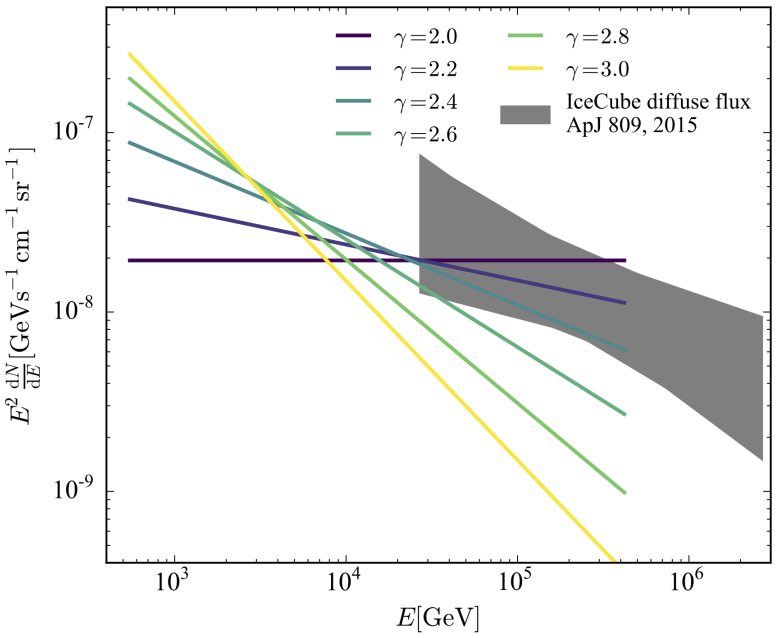

Figure A.10.: 90% Upper limit on the diffuse astrophysical flux for different assumed spectral indices γ of the source. The upper limits are shown with respect to the supernova type IIn results. The energy range is again the central 90% range.

Bibliography

[1] V. Gaffney, S. Fitch, E. Ramsey, et al. *Time and a place: a luni-solar 'time-reckoner' from 8th millennium BC Scotland. Internet Archaeology*, 34, 7 2013. doi:10.11141/ia.34.1. 5

[2] C. M. Graney. *On the Accuracy of Galileo's Observations. Baltic Astronomy*, 16:443–449, 2007. 0802.1095. 5

[3] V. F. Hess. *Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten. Physikalische Zeitschrift*, 13:1084–1091, November 1912. 5, 6, 9

[4] B. P. Abbott et al. *LIGO: The Laser interferometer gravitational-wave observatory. Rept. Prog. Phys.*, 72:076901, 2009. doi:10.1088/0034-4885/72/7/076901. 0711.3041. 5, 61

[5] M. Kowalski. *Status of High-Energy Neutrino Astronomy. J. Phys. Conf. Ser.*, 632(1):012039, 2015. doi:10.1088/1742-6596/632/1/012039. 1411.4385. 5, 149, 150

[6] M. Santander. *The Dawn of Multi-Messenger Astronomy*. 2016. 1606.09335. 5, 61

[7] M. G. Aartsen et al. *Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector. Science*, 342:1242856, 2013. doi:10.1126/science.1242856. 1311.5238. 6, 14, 25

[8] M. G. Aartsen et al. *The IceCube Neutrino Observatory: Instrumentation and Online Systems. JINST*, 12(03):P03012, 2017. doi:10.1088/1748-0221/12/03/P03012. 1612.05093. 6, 36, 38, 39, 40, 61

[9] K. A. Olive et al. *Review of Particle Physics. Chin. Phys.*, C38:090001, 2014. doi:10.1088/1674-1137/38/9/090001. 9, 10, 13, 35, 36, 44

[10] L. O'C. Drury. *Origin(s) of Cosmic Rays. PoS*, CRISM2014:014, 2015. 1412.1376. 11

[11] A. M. Hillas. *The Origin of Ultra-High-Energy Cosmic Rays. araa*, 22:425–444, 1984. doi:10.1146/annurev.aa.22.090184.002233. 11

[12] E. Fermi. *On the Origin of the Cosmic Radiation. Phys. Rev.*, 75:1169–1174, Apr 1949. doi:10.1103/PhysRev.75.1169. 11

[13] A. R. Bell. *The acceleration of cosmic rays in shock fronts. I. mnras*, 182:147–156, January 1978. doi:10.1093/mnras/182.2.147. 11

[14] A. R. Bell. *The acceleration of cosmic rays in shock fronts II. Monthly Notices of the Royal Astronomical Society*, 182(3):443, 1978. doi:10.1093/mnras/182.3.443. /oup/backfile/content_public/journal/mnras/182/3/10.1093/mnras/182.3.443/2/mnras182-0443.pdf. 11

[15] R. D. Blandford and J. P. Ostriker. *Particle Acceleration by Astrophysical Shocks. Astrophys. J.*, 221:L29–L32, 1978. doi:10.1086/182658. 12

[16] M. A. Malkov and L. O. Drury. *Nonlinear theory of diffusive acceleration of particles by shock waves. Reports on Progress in Physics*, 64(4):429, 2001. 13

[17] V. Zirakashvili and V. Ptuskin. *Numerical simulations of diffusive shock acceleration in SNRs. Astroparticle Physics*, 39–40:12 – 21, 2012. doi:https://doi.org/10.1016/j.astropartphys.2011.09.003. Cosmic Rays Topical Issue. 13

[18] S. R. Kelner and F. A. Aharonian. *Energy spectra of gamma-rays, electrons and neutrinos produced at interactions of relativistic protons with low energy radiation. Phys. Rev.*, D78:034013, 2008. doi:10.1103/PhysRevD.82.099901, 10.1103/PhysRevD.78.034013. [Erratum: Phys. Rev.D82,099901(2010)], 0803.0688. 14

[19] M. G. Aartsen et al. *A combined maximum-likelihood analysis of the high-energy astrophysical neutrino flux measured with IceCube. Astrophys. J.*, 809(1):98, 2015. doi:10.1088/0004-637X/809/1/98. 1507.03991. 14, 15, 25, 148, 149, 161

[20] M. G. Aartsen et al. *All-sky Search for Time-integrated Neutrino Emission from Astrophysical Sources with 7 yr of IceCube Data. Astrophys. J.*, 835(2):151, 2017. doi:10.3847/1538-4357/835/2/151. 1609.04981. 14, 44, 46, 57, 83, 130, 131, 150, 158

[21] S. Coenders. *High-energy cosmic ray accelerators: searches with IceCube neutrinos.* Dissertation, Technische Universität München, München, 2016. 15, 57

[22] C. A. Bertulani and T. Kajino. *Frontiers in Nuclear Astrophysics. Prog. Part. Nucl. Phys.*, 89:56–100, 2016. doi:10.1016/j.ppnp.2016.04.001. 1604.03197. 17, 18

[23] S. Woosley and T. Janka. *The physics of core-collapse supernovae. Nature Phys.*, 1:147, 2005. doi:10.1038/nphys172. astro-ph/0601261. 18

[24] S. E. Woosley, A. Heger, and T. A. Weaver. *The evolution and explosion of massive stars. Rev. Mod. Phys.*, 74:1015–1071, Nov 2002. doi:10.1103/RevModPhys.74.1015. 18, 21

[25] P. A. Crowther. *Physical Properties of Wolf-Rayet Stars. Ann. Rev. Astron. Astrophys.*, 45:177–219, 2007. doi:10.1146/annurev.astro.45.051806.110615. astro-ph/0610356. 18, 25, 26

[26] S. Chandrasekhar. *The Maximum Mass of Ideal White Dwarfs. Astrophysical Journal*, 1931. 19, 20

[27] E. Bravo, C. Badenes, and D. Garcia-Senz. *Thermonuclear supernova models, and observations of type Ia supernovae. AIP Conf. Proc.*, 797:453–462, 2005. doi:10.1063/1.2130267. [,453(2004)], astro-ph/0412155. 20

[28] W. Hillebrandt and J. C. Niemeyer. *Type Ia supernova explosion models. Ann. Rev. Astron. Astrophys.*, 38:191–230, 2000. doi:10.1146/annurev.astro.38.1.191. astro-ph/0006305. 20

[29] H.-T. Janka, K. Langanke, A. Marek, et al. *Theory of Core-Collapse Supernovae. Phys. Rept.*, 442:38–74, 2007. doi:10.1016/j.physrep.2007.02.002. astro-ph/0612072. 20, 21, 23

[30] J. R. Oppenheimer and G. M. Volkoff. *On Massive Neutron Cores. Phys. Rev.*, 55:374–381, Feb 1939. doi:10.1103/PhysRev.55.374. 21

[31] M. Turatto. *Classification of supernovae. Lect. Notes Phys.*, 598:21, 2003. doi:10.1007/3-540-45863-8_3. astro-ph/0301107. 24

[32] A. Gal-Yam. *Observational and Physical Classification of Supernovae.* 2016. doi:10.1007/978-3-319-20794-0_35-1. 1611.09353. 24, 74

[33] J. Hjorth and J. S. Bloom. *The Gamma-Ray Burst - Supernova Connection*, pages 169–190. November 2012. 24

[34] E. O. Ofek et al. *Precursors prior to Type IIn supernova explosions are common: precursor rates, properties, and correlations. Astrophys. J.*, 789(2):104, 2014. doi:10.1088/0004-637X/789/2/104. 1401.5468. 25

[35] H. C. Spruit. *Essential Magnetohydrodynamics for Astrophysics. ArXiv e-prints*, January 2013. 1301.5572. 25

[36] A. R. Bell. *Particle acceleration by shocks in supernova remnants. Braz. J. Phys.*, 44:415–425, 2014. doi:10.1007/s13538-014-0219-5. [,1286(2013)], 1311.5779. 25

[37] K. Murase, T. A. Thompson, B. C. Lacki, et al. *New Class of High-Energy Transients from Crashes of Supernova Ejecta with Massive Circumstellar Material Shells. Phys. Rev.*, D84:043003, 2011. doi:10.1103/PhysRevD.84.043003. 1012.2834. 25, 26, 27, 28, 29, 73, 132, 145

[38] V. N. Zirakashvili and V. S. Ptuskin. *Type IIn supernovae as sources of high energy astrophysical neutrinos. Astropart. Phys.*, 78:28–34, 2016. doi:10.1016/j.astropartphys.2016.02.004. 1510.08387. 25, 26, 28, 29, 132, 135, 147

[39] K. Murase and K. Ioka. *TeV–PeV Neutrinos from Low-Power Gamma-Ray Burst Jets inside Stars. Phys. Rev. Lett.*, 111(12):121102, 2013. doi:10.1103/PhysRevLett.111.121102. 1306.2274. 29

[40] N. Senno, K. Murase, and P. Meszaros. *Choked Jets and Low-Luminosity Gamma-Ray Bursts as Hidden Neutrino Sources. Phys. Rev.*, D93(8):083003, 2016. doi:10.1103/PhysRevD.93.083003. 1512.08513. 29, 30, 31, 74, 133

[41] G. Rajasekaran. *The Story of the Neutrino*. 2016. 1606.08715. 33

[42] T. Glüsenkamp. *Search for a cumulative neutrino flux from 2LAC-blazar populations using 3 years of IceCube data*. Ph.D. thesis, 2016. 34, 96

[43] J. A. Formaggio and G. P. Zeller. *From eV to EeV: Neutrino Cross Sections Across Energy Scales. Rev. Mod. Phys.*, 84:1307–1341, 2012. doi: 10.1103/RevModPhys.84.1307. 1305.7513. 34

[44] A. Cooper-Sarkar, P. Mertsch, and S. Sarkar. *The high energy neutrino cross-section in the Standard Model and its uncertainty. JHEP*, 08:042, 2011. doi: 10.1007/JHEP08(2011)042. 1106.3723. 35

[45] P. A. Čerenkov. *Visible Radiation Produced by Electrons Moving in a Medium with Velocities Exceeding that of Light. Phys. Rev.*, 52:378–379, Aug 1937. doi: 10.1103/PhysRev.52.378. 35

[46] R. Abbasi, Y. Abdou, T. Abu-Zayyad, et al. *Calibration and characterization of the IceCube photomultiplier tube. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 618(1–3):139 – 152, 2010. doi:https://doi.org/10.1016/j.nima.2010.03.102. 38

[47] M. Aartsen, R. Abbasi, Y. Abdou, et al. *Measurement of South Pole ice transparency with the IceCube {LED} calibration system. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 711:73 – 89, 2013. doi:https://doi.org/10.1016/j.nima.2013.01.054. 39, 40

[48] M. G. Aartsen et al. *Search for Astrophysical Tau Neutrinos in Three Years of IceCube Data. Phys. Rev.*, D93(2):022001, 2016. doi:10.1103/PhysRevD.93.022001. 1509.06212. 41, 42

[49] J. G. Learned and K. Mannheim. *High-Energy Neutrino Astrophysics. Annual Review of Nuclear and Particle Science*, 50(1):679–749, 2000. doi:10.1146/annurev.nucl.50.1.679. https://doi.org/10.1146/annurev.nucl.50.1.679. 42

[50] M. G. Aartsen et al. *Searches for Extended and Point-like Neutrino Sources with Four Years of IceCube Data. Astrophys. J.*, 796(2):109, 2014. doi:10.1088/0004-637X/796/2/109. 1406.6757. 46, 47

[51] C. Grupen. *Physics of particle detection.* In *American Institute of Physics Conference Series*, volume 536 of *American Institute of Physics Conference Series*, pages 3–34. September 2000. doi:10.1063/1.1361756. physics/9906063. 47

[52] A. Schukraft. *A view of prompt atmospheric neutrinos with IceCube. Nucl. Phys. Proc. Suppl.*, 237-238:266–268, 2013. doi:10.1016/j.nuclphysbps.2013.04.105. 1302.0127. 48

[53] J. Ahrens et al. *Muon track reconstruction and data selection techniques in AMANDA. Nucl. Instrum. Meth.*, A524:169–194, 2004. doi:10.1016/j.nima.2004.01.065. astro-ph/0407044. 49, 50, 51

[54] R. J. Barlow. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences.* Wiley, 1993. ISBN 0471922951. 49, 52, 89, 90, 106, 118, 123, 142

[55] N. Whitehorn, J. van Santen, and S. Lafebre. *Penalized Splines for Smooth Representation of High-dimensional Monte Carlo Datasets. Comput. Phys. Commun.*, 184:2214–2220, 2013. doi:10.1016/j.cpc.2013.04.008. 1301.2184. 51

[56] T. Neunhoffer. *Estimating the angular resolution of tracks in neutrino telescopes based on a likelihood analysis. Astropart. Phys.*, 25:220–225, 2006. doi:10.1016/j.astropartphys.2006.01.002. astro-ph/0403367. 52, 55

[57] R. Abbasi et al. *An improved method for measuring muon energy using the truncated mean of dE/dx. Nucl. Instrum. Meth.*, A703:190–198, 2013. doi:10.1016/j.nima.2012.11.081. 1208.3430. 53

[58] M. Ackermann et al. *Optical properties of deep glacial ice at the South Pole. J. Geophys. Res. Atmos.*, 111(D13):D13203, 2006. doi:10.1029/2005JD006687. 53

[59] M. G. Aartsen et al. *Energy Reconstruction Methods in the IceCube Neutrino Telescope. JINST*, 9:P03009, 2014. doi:10.1088/1748-0221/9/03/P03009. 1311.4767. 53, 54

[60] M. W. E. Smith et al. *The Astrophysical Multimessenger Observatory Network (AMON). Astropart. Phys.*, 45:56–70, 2013. doi:10.1016/j.astropartphys.2013.03.003. 1211.5602. 61

[61] R. J. Lauer. *Monitoring the Variable Gamma-Ray Sky with HAWC. AIP Conf. Proc.*, 1792(1):070013, 2017. doi:10.1063/1.4969010. 1610.05172. 61

[62] A. Franckowiak, C. Akerlof, D. F. Cowen, et al. *Optical follow-up of high-energy neutrinos detected by IceCube. ArXiv e-prints*, September 2009. 0909.0631. 62

[63] M. G. Aartsen et al. *Very High-Energy Gamma-Ray Follow-Up Program Using Neutrino Triggers from IceCube. JINST*, 11(11):P11009, 2016. doi:10.1088/1748-0221/11/11/P11009. 1610.01814. 62

[64] M. G. Aartsen et al. *The IceCube Realtime Alert System. Astropart. Phys.*, 92:30–41, 2017. doi:10.1016/j.astropartphys.2017.05.002. 1612.06028. 62, 64, 65

[65] B. P. Roe, H.-J. Yang, J. Zhu, et al. *Boosted decision trees, an alternative to artificial neural networks. Nucl. Instrum. Meth.*, A543(2-3):577–584, 2005. doi:10.1016/j.nima.2004.12.018. physics/0408124. 63

[66] A. Stasik, T. Kintscher, M. Kowalski, et al. *The Online Follow-Up Framework for Neutrino-Triggered Alerts from IceCube. PoS*, ICRC2015:1069, 2016. 64, 65

[67] S. V. W. Beckwith et al. *The Hubble Ultra Deep Field. Astron. J.*, 132:1729–1755, 2006. doi:10.1086/507302. astro-ph/0607632. 67

[68] Y. Cao, P. E. Nugent, and M. M. Kasliwal. *Intermediate Palomar Transient Factory: Realtime Image Subtraction Pipeline. Publications of the Astronomical Society of the Pacific*, 128(969):114502, 2016. 67

[69] F. J. Masci, R. R. Laher, U. D. Rebbapragada, et al. *The IPAC Image Subtraction and Discovery Pipeline for the Intermediate Palomar Transient Factory. Publications of the Astronomical Society of the Pacific*, 129(971):014002, 2017. 67

[70] T. W. S. Holoien et al. *The ASAS-SN Bright Supernova Catalog – I. 2013–2014. Mon. Not. Roy. Astron. Soc.*, 464(3):2672–2686–2686, 2017. doi:10.1093/mnras/stw2273. 1604.00396. 68, 70

[71] A. Rau, S. R. Kulkarni, N. M. Law, et al. *Exploring the Optical Transient Sky with the Palomar Transient Factory. Publications of the Astronomical Society of the Pacific*, 121(886):1334, 2009. 68, 69, 79

[72] K. C. Chambers, E. A. Magnier, N. Metcalfe, et al. *The Pan-STARRS1 Surveys. ArXiv e-prints*, December 2016. 1612.05560. 68

[73] L. Wyrzykowski, Z. Kostrzewa-Rutkowska, J. Klencki, et al. *OGLE-IV Transient Search summary of season 2015b. The Astronomer's Telegram*, 8484, December 2015. 69

[74] E. Bellm. *The Zwicky Transient Facility.* In P. R. Wozniak, M. J. Graham, A. A. Mahabal, et al., editors, *The Third Hot-wiring the Transient Universe Workshop*, pages 27–33. 2014. 1410.8185. 69

[75] J. S. Bloom, J. W. Richards, P. E. Nugent, et al. *Automating Discovery and Classification of Transients and Variable Stars in the Synoptic Survey Era. Publications of the Astronomical Society of the Pacific*, 124(921):1175, 2012. 69

[76] R. Rutledge. *The Astronomer's Telegram: A Web-based Short-Notice Publication System for the Professional Astronomical Community. Publ. Astron. Soc. Pac.*, 110:754, 1998. doi:10.1086/316184. astro-ph/9802256. 69

[77] O. Yaron and A. Gal-Yam. *WISeREP—An Interactive Supernova Data Repository. Publications of the Astronomical Society of the Pacific*, 124(917):668, 2012. 69, 70

[78] J. Guillochon, J. Parrent, and R. Margutti. *An Open Catalog for Supernova Data.* 2016. 1605.01054. 69, 70

[79] C. G. Tsagas and M. I. Kadiltzoglou. *Peculiar Raychaudhuri equation. Phys. Rev.*, D88(8):083501, 2013. doi:10.1103/PhysRevD.88.083501. 1306.6501. 71

[80] E. R. Harrison. *Interpretation of Redshifts of Galaxies in Clusters. Astron. J.*, 191:L51, July 1974. doi:10.1086/181545. 71

[81] D. W. Hogg. *Distance measures in cosmology.* 1999. astro-ph/9905116. 72, 76, 84, 85

[82] Astropy Collaboration, T. P. Robitaille, E. J. Tollerud, et al. *Astropy: A community Python package for astronomy. aap*, 558:A33, October 2013. doi: 10.1051/0004-6361/201322068. 1307.6212. 72, 76, 83, 127

[83] P. A. R. Ade et al. *Planck 2015 results. XIII. Cosmological parameters. Astron. Astrophys.*, 594:A13, 2016. doi:10.1051/0004-6361/201525830. 1502.01589. 72, 76, 83, 148

[84] J. E. Andrews, J. S. Gallagher, G. C. Clayton, et al. *SN 2007od: A Type IIP Supernova with Circumstellar Interaction. The Astrophysical Journal*, 715(1):541, 2010. 73

[85] T. Moriya, N. Tominaga, S. I. Blinnikov, et al. *Supernovae from red supergiants with extensive mass loss. Monthly Notices of the Royal Astronomical Society*, 415(1):199–213, 2011. doi:10.1111/j.1365-2966.2011.18689.x. 73

[86] T. Piran, E. Nakar, P. Mazzali, et al. *Relativistic Jets in Core Collapse Supernovae.* 2017. 1704.08298. 74

[87] L.-G. Strolger, T. Dahlen, S. A. Rodney, et al. *The Rate of Core Collapse Supernovae to Redshift 2.5 From The CANDELS and CLASH Supernova Surveys. Astrophys. J.*, 813(2):93, 2015. doi:10.1088/0004-637X/813/2/93. 1509.06574. 77, 148, 149

[88] S. A. Rodney et al. *Type Ia Supernova Rate Measurements to Redshift 2.5 from CANDELS : Searching for Prompt Explosions in the Early Universe. Astron. J.*, 148:13, 2014. doi:10.1088/0004-6256/148/1/13. 1401.7978. 77, 78

[89] P. Madau and M. Dickinson. *Cosmic Star Formation History. Ann. Rev. Astron. Astrophys.*, 52:415–486, 2014. doi:10.1146/annurev-astro-081811-125615. 1403.0007. 77, 148

[90] D. Richardson, R. L. Jenkins, J. Wright, et al. *Absolute-Magnitude Distributions of Supernovae. Astron. J.*, 147:118, 2014. doi:10.1088/0004-6256/147/5/118. 1403.5755. 78

[91] K. Murase, K. Asano, and S. Nagataki. *Effects of Cosmic Infrared Background on High Energy Delayed Gamma-Rays from Gamma-Ray Bursts. Astrophys. J.*, 671:1886–1895, 2007. doi:10.1086/523031. astro-ph/0703759. 84, 85

[92] G. Couture. *Olber's paradox revisited in a static and finite Universe. European Journal of Physics*, 33(3):479, 2012. 85

[93] K. Murase. *High energy neutrino early afterglows gamma-ray bursts revisited. Phys. Rev.*, D76:123001, 2007. doi:10.1103/PhysRevD.76.123001. 0707.1140. 86

[94] M. Ahlers and F. Halzen. *Pinpointing Extragalactic Neutrino Sources in Light of Recent IceCube Observations. Phys. Rev.*, D90(4):043005, 2014. doi:10.1103/PhysRevD.90.043005. 1406.2160. 86

[95] J. Braun, J. Dumm, F. De Palma, et al. *Methods for point source analysis in high energy neutrino telescopes. Astropart. Phys.*, 29:299–305, 2008. doi:10.1016/j.astropartphys.2008.02.007. 0801.1604. 90

[96] J. Braun, M. Baker, J. Dumm, et al. *Time-dependent point source search methods in high energy neutrino astronomy. Astroparticle Physics*, 33:175–181, April 2010. doi:10.1016/j.astropartphys.2010.01.005. 0912.1572. 90

[97] M. Evans. *Statistical distributions.* Wiley, New York, 2000. ISBN 978-0471371243. 95

[98] S. S. Wilks. *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. Ann. Math. Statist.*, 9(1):60–62, 03 1938. doi:10.1214/aoms/1177732360. 119

[99] S. Van Der Walt, S. C. Colbert, and G. Varoquaux. *The NumPy array: a structure for efficient numerical computation. ArXiv e-prints*, February 2011. 1102.1523. 127

[100] D. F. Shanno. *Conditioning of Quasi-Newton Methods for Function Minimization. Mathematics of Computation*, 24(111):647–656, 1970. 130

[101] R. Fletcher. *A new approach to variable metric algorithms. The Computer Journal*, 13(3):317–322, 1970. doi:10.1093/comjnl/13.3.317.

/oup/backfile/content_public/journal/comjnl/13/3/10.1093/comjnl/13.3.317/2/130317.pdf. 130

[102] E. Gross and O. Vitells. *Trial factors or the look elsewhere effect in high energy physics. Eur. Phys. J.*, C70:525–530, 2010. doi:10.1140/epjc/s10052-010-1470-8. 1005.1891. 141

[103] G. Choudalakis. *On hypothesis testing, trials factor, hypertests and the BumpHunter.* In *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011.* 2011. 1101.0390. 141

Acknowledgements

I would like to thank the following people who helped me along the way in the last years. First, Marek Kolwaski for the possibility for this project and for his support and supervision in the last years. I also want to thank Anna Franckowiak for the endless discussions and the running. Further on, Elisa Bernardini for her insights on statistical problems. I want to thank Konstancja Satalecka for her motivation and support. A special thanks goes to Thomas Kintscher who suffered from endless discussions, his help in all kind of computer-related issues and beeing my office colleague in the last years. I would also to thank Robert Stein, Juliana Stachurska and Anna Frankowiak for proof-reading this thesis.
Thanks to Carla Schenker for being her. Finally, I would like to thank my family for their endless support and encouragement.

Erklärung

Ich erkläre, dass ich die vorliegende Dissertation selbständig und nur unter Verwendung der von mir gemäß § 7 Abs. 3 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin
Nr. 126/2014 am 18.11.2014, angegebenen Hilfsmittel angefertigt habe.

Berlin, den 14.12.2017 .....

Alexander Stasik