

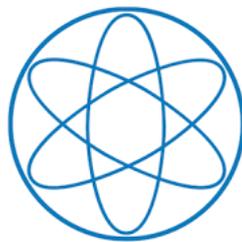
Technische Universität München  
Experimental Physics with Cosmic Particles

---

**The Search for High Energy Neutrino Sources  
at the IceCube Neutrino Observatory:**  
*Interplay Between Diffuse and Sub-Threshold Signal*

---

**Kristian Tchiorniy**



Thesis submitted within  
The Bachelor's Program for Physics

Supervising Professor: Prof. Dr. Elisa Resconi  
Second Examiner: PD Dr. rer. nat. Bela Majorovits

September 5, 2022

# Abstract

The IceCube Neutrino Observatory has made some significant discoveries in the field of neutrino astronomy, including the discovery of a diffuse astrophysical high energy neutrino flux. However it is still facing the challenge of pin-pointing the sources for these neutrinos in the sky. At IceCube, various analysis methods have been developed to try to improve on this, including the neutrino hotspot population analysis. This analysis searches for populations of sub-threshold neutrino sources or hotspots across the sky.

This thesis is a performance study of two aspects of the hotspot population analysis. The first investigated aspect is the search method for hotspots, in particular, the choice of minimum angular distance  $\Psi_{min}$  which two hotspots may occupy between one another in a map of the sky as seen by IceCube. It was found that a minimum angular distance of  $\Psi_{min} = 1^\circ$  is a reasonable choice, however this could still be chosen to be smaller, but the current tests made for smaller angles were inconclusive.

The second and main investigation looked into the influence of the diffuse astrophysical flux on the hotspot population analysis, in particular the simulation of neutrino sky maps for sub-threshold sources with additional injected neutrino sources. It was previously established that the simulation underestimated the diffuse flux in these pseudo sky scans, and so two amendments to the simulation method were investigated in this thesis. While setting a limit of the maximum number of hotspots saw little improvement, the second change provided more accurate estimations in sensitivity and discovery potential for the analysis method.

# Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>   | <b>1</b>  |
| <b>2</b> | <b>Neutrino Physics</b>                                     | <b>3</b>  |
| 2.0.1    | Neutrino Oscillations . . . . .                             | 4         |
| 2.1      | High Energy Neutrinos . . . . .                             | 5         |
| 2.1.1    | Generation . . . . .  | 5         |
| 2.1.2    | Interaction with Matter . . . . .                           | 6         |
| 2.1.3    | Atmospheric Neutrinos . . . . .                             | 7         |
| 2.2      | Astrophysical Messengers . . . . .                          | 9         |
| <b>3</b> | <b>The IceCube Neutrino Observatory</b>                     | <b>11</b> |
| 3.1      | Experimental Setup . . . . .                                | 12        |
| 3.2      | Detection Principle . . . . .                               | 13        |
| 3.2.1    | Cherenkov Effect . . . . .                                  | 13        |
| 3.2.2    | Event Topologies . . . . .                                  | 14        |
| <b>4</b> | <b>The Neutrino Point Source Analysis</b>                   | <b>18</b> |
| 4.1      | Neutrino Point Source Search . . . . .                      | 19        |
| 4.1.1    | Maximum Likelihood Search Method . . . . .                  | 19        |
| 4.1.2    | Significance, Sensitivity and Discovery Potential . . . . . | 21        |
| 4.2      | Neutrino Skymaps . . . . .                                  | 23        |
| <b>5</b> | <b>The Hotspot Population Analysis</b>                      | <b>25</b> |
| 5.1      | Analysis Method . . . . .                                   | 25        |
| 5.2      | Neutrino Hotspot Search Optimization . . . . .              | 27        |
| 5.2.1    | Current Neutrino Hotspot Search Method . . . . .            | 27        |
| 5.2.2    | Expectation, and KS and $\chi^2$ Tests . . . . .            | 28        |
| 5.2.3    | Results and Discussion . . . . .                            | 30        |

|          |  |           |
|----------|--|-----------|
| 5.2.4    | Conclusion . . . . .   | 34        |
| 5.3      | Diffuse Astrophysical Neutrino Flux Investigation . . . . .                        | 35        |
| 5.3.1    | Generation of Background Trials . . . . .  | 36        |
| 5.3.2    | Generation of Signal Trials . . . . .  | 36        |
| 5.3.3    | Influence of the Diffuse Astrophysical Neutrino Flux . .                           | 37        |
| 5.3.4    | Accounting for the Astrophysical Diffuse Neutrino Flux<br>in Simulations . . . . . | 40        |
| 5.3.5    | Results and Discussion . . . . .   | 43        |
| <b>6</b> | <b>Conclusion and Outlook</b>  | <b>46</b> |

# Chapter 1

## Introduction

The cosmos is a wild, outlandish, and explosive place. It is filled with unimaginably large things, that produce unimaginably small things, that then fly through space for unimaginably long distances, and then reach us here on Earth. In the 20th Century, great strides in physics were made, in order to better understand these small particles. Great strides were also made to explain these unimaginably large objects, stars, galaxies, black holes, supernova, blazars, and many other cosmic phenomena, in the fields of cosmology and astrophysics.

It became well understood that in order to best understand the large observed phenomena, the underlying physics at a subatomic level had to be understood. This included the study of cosmic rays, gamma rays, and, at some time later, neutrinos. These three categories of particles allow us today to make more and more precise statements about the astrophysical objects. This is because they are created and shot out from within these astrophysical objects, in an often awe-inspiring manor, such as through the collision of high energy particles in accretion discs near galaxy centers [1]. However there is still much to be understood from these particles or so-called astrophysical messengers, and the sources from which they originate. Neutrinos in particular are considered to be excellent candidates for astrophysical messengers. They are particles that interact exclusively via the weak interaction and thus they have ability to traverse through space, unhindered from obstacles which may present themselves along their way, such as intergalactic magnetic fields or dust clouds.

The IceCube Neutrino Observatory, deep within the antarctic ice of the South Pole, aims to detect such high energy neutrinos, and therefore learn more about the cosmos. The observatory is able to detect neutrinos from tens of GeV to above 100 PeV. It has already made some significant discoveries such as the detection of a diffuse astrophysical neutrino flux [2], and more recently, the discovery of evidence of neutrino emission from the TXS 0506+056 high energy blazar, with  $3\sigma$  significance [3]. However much is still not known about the exact origins of the high energy neutrinos that IceCube detects.

This thesis focuses on the influence of the diffuse astrophysical neutrino flux on a specific analysis, the aim of which being to look for a population of sub-threshold neutrino sources. Optimizing this analysis will allow us to make more accurate statements about the neutrino flux emitted by such sources and therefore better understand the signals that IceCube sees. This thesis begins with a review of neutrino physics, relevant for the neutrinos detected at IceCube. Chapter 3 details the IceCube Neutrino Observatory's setup and detection principle. Chapter 4 elaborates the neutrino point source analysis which is consequently of relevance for the main analysis of this thesis, the hotspot population analysis, discussed in Chapter 5. In Chapter 5 I also present the optimizations made to the analysis, which are then summarized and concluded in Chapter 6.

## Chapter 2

# Neutrino Physics

Neutrinos make up an essential part of the Standard Model and their properties have been widely studied in the past century, with the discovery of the weak force. The weak force is responsible for interactions occurring via exchange of charged  $W^\pm$  and neutral  $Z^0$  bosons. It was first theoretically developed by Enrico Fermi in 1934 [4], in order to explain the missing energy in  $\beta$ -decay spectrum in atomic nuclei, observed by Wolfgang Pauli. In 1930 Pauli had postulated that along with the electron, an almost massless neutral particle must also have been emitted, which was consequently named the "neutrino" by Fermi [5]. Figure 2.1 depicts the Feynman diagram of  $\beta$ -decay. The neutrino was experimentally confirmed by Cowan and Reines in 1956 [6].

Through the following years, it was established that the particles that interacted via the weak force are quarks, leptons, and the two aforementioned bosons. It was also established that neutrinos belong to the fermion class of particles (have a spin of  $\frac{1}{2}$ ) and furthermore make up three of the lepton flavors: the electron neutrino  $\nu_e$ , muon neutrino  $\nu_\mu$ , and tau neutrino  $\nu_\tau$ , and their anti-particle counterparts. Also they are chargeless in terms of the strong and electromagnetic forces [7].

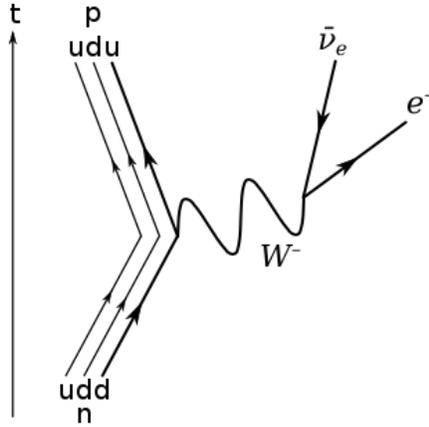


Figure 2.1: Feynman diagram describing the process of  $\beta$ -decay occurring in atomic nuclei, an important process of the weak force. Source: [8]

These three flavors were discovered as a consequence that each lepton, the electron, muon, and tau, have a corresponding neutrino, creating a so-called family. During interactions involving the weak force, lepton family number remains conserved, such as in beta-decay, which involves an electron and an anti electron neutrino. The  $W^\pm$  boson is responsible for the transition from charged lepton to neutrino within each family and vice versa (as well as up-type quarks to down-type quarks in their respective families and vice versa), and appears in so-called charged currents (CC). Meanwhile, the  $Z^0$  boson describes the interaction between same charge leptons (or quarks) in so called neutral currents (NC).

An important consequence of the interaction of neutrinos exclusively via the weak force is that they have a very low probability to interact with other particles. The typical cross-section of the weak force is many orders of magnitude smaller than any other force, staying under the range of  $\sigma_{weak} \lesssim 10^{-38} \text{cm}^2$  compared to a typical value for the strong force at  $\sigma_{strong} \approx 10^{-24} \text{cm}^2$  [7].

### 2.0.1 Neutrino Oscillations

Another important property of neutrinos is their very small rest masses. Although it was long believed that neutrinos are massless, it is now known that there are three discrete neutrino masses, thanks to the observation of neutrino

flavor oscillations in the atmosphere by the Super-Kamiokande Collaboration in 1998 [9]. It was proposed that there are flavor eigenstates, which correspond to the behaviour of neutrinos during the weak interaction, and then there are the mass eigenstates which correspond to the propagation of the neutrinos in vacuum. It was thus determined that the flavor eigenstates are in fact a superposition of the mass eigenstates, which can be described using the Pontecorvo–Maki–Nakagawa–Sakata matrix (PMNS matrix)  $U$  [10] such that:

$$\begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} = U \begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} \end{pmatrix} \cdot \begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} \quad (2.1)$$

where  $\nu_1, \nu_2, \nu_3$  are the mass eigenstates.  $U$  is also unitary such that  $U^{-1} = U^\dagger$ . Each entry of the PMNS matrix is described with a combination of 3 mixing angles ( $\theta_{12}, \theta_{23}, \theta_{13}$ ) and one complex phase  $\delta_{CP}$ . The idea of oscillating neutrinos was theorized in order to combat the solar neutrino problem [11], i.e. the discovery that the electron neutrino flux at Earth is only about one-third of that which is expected from energetic considerations of the fusion processes in the sun, a problem already observed in middle of the 20th century. Observations from the Super-Kamiokande Collaboration [9] helped to strengthen this newly proposed theory.

## 2.1 High Energy Neutrinos

### 2.1.1 Generation

Neutrinos with high energies (specifically GeV to PeV range) [12] can be emitted via collision of high energy protons with ambient matter near or at the acceleration sites of astrophysical objects, such as blazars or supermassive black holes. This mainly occurs via the interaction with other protons [13], with processes such as:

$$pp \longrightarrow \begin{cases} pn\pi^+ \longrightarrow pn\mu^+\nu_\mu \longrightarrow pne^+\nu_e\bar{\nu}_\mu\nu_\mu \\ pp\pi^0 \longrightarrow pp\gamma\gamma \end{cases} \quad (2.2)$$

whereas proton-photon interactions may also occur:

$$p\gamma \longrightarrow \Delta^+ \longrightarrow \begin{cases} n\pi^+ \longrightarrow n\mu^+\nu_\mu \longrightarrow ne^+\nu_e\bar{\nu}_\mu\nu_\mu \\ p\pi^0 \longrightarrow p\gamma\gamma \end{cases} \quad (2.3)$$

The same processes occur if one regards incident neutrons instead of protons, leading to the production of negatively charged pions  $\pi^-$ , which can occur by replacing all particles on the right side in the above processes with their antiparticles. At higher energies, also kaons can contribute to this decay spectrum [14]. As one can see neutrinos are typically produced in pion-decays, subsequent muon-decays, and neutron decays. Even if neutrinos are neutral fermions, magnetic fields can however still have an indirect effect on neutrinos, as they affect the cosmic rays, vital for their production. They cause synchrotron radiation which reduces their energy [13]. Thus, depending on the source properties, the dominant production mechanism can change the energy dependent flavor ratios, produced at the source [15].

Tau neutrinos have an extremely low rate of being produced due to their heavier masses, but due to neutrino oscillation, and the very long distances which the neutrinos must propagate in order to reach Earth, the flavors are mixed in such a way that we see a uniform distribution [13], with:

$$(\nu_e : \nu_\mu : \nu_\tau) = (\bar{\nu}_e : \bar{\nu}_\mu : \bar{\nu}_\tau) = (1 : 1 : 1) \quad (2.4)$$

### 2.1.2 Interaction with Matter

As previously mentioned, neutrinos interact exclusively through exchange of  $W^\pm$  and  $Z^0$  bosons, in CC and NC processes respectively. This allows neutrinos to interact with the atomic nuclei of ambient matter in both CC and NC processes. Such interactions include

$$\nu_l + N \longrightarrow \nu_l + N^* \quad (NC) \quad (2.5)$$

$$\nu_l + N \longrightarrow l + N^* \quad (CC) \quad (2.6)$$

Here,  $l$  is any lepton flavor and  $N$  is an atomic nucleus. In the NC process neutrinos exchange a virtual  $Z^0$  boson with a nucleon, depositing some fraction of their energy and initializing a hadronic cascade (denoted with  $N^*$ ). In the CC process, neutrinos decay into charged leptons of the same generation, via the exchange of a charged  $W^\pm$  boson with a nucleon (inverse beta decay) also creating an additional hadronic cascade.

These processes are energy dependant and therefore different neutrino energies can lead to different interaction sub-types dominating these two processes. For IceCube, a relevant energy range is  $E_\nu > 20$  GeV. At these higher energies deep-inelastic scattering becomes more prominent and the nucleus must now be treated at the quark level [14].

### 2.1.3 Atmospheric Neutrinos

Earth's atmosphere is abundant in ambient matter, which makes it prone to interactions with any incoming radiation, namely primary cosmic rays. Primary cosmic rays are composed mainly of protons and alpha particles, along with very small amounts of heavier nuclei and trace amounts of positrons and antiprotons [16]. During such interactions, air showers of secondary particles are produced. Pions and kaons are mainly produced, via proton-proton and proton-neutron collisions, which decay into one of various modes. Figure 2.2 depicts these different modes of cascades.

The muonic cascade mode sees the production of muon (and anti-muon) neutrinos, resulting from the decay of charged pions and kaons:

$$\pi^-(K^-) \longrightarrow \mu^+ + \nu_\mu \quad (2.7)$$

$$\pi^+(K^+) \longrightarrow \mu^- + \bar{\nu}_\mu \quad (2.8)$$

with pions decaying into this branch in 99% of cases, and kaons 67% [18]. Muons further decay into electrons during this cascade, and as per the weak force also further release electron neutrinos in the process:

$$\mu^- \longrightarrow e^- + \nu_\mu + \bar{\nu}_e \quad (2.9)$$

$$\mu^+ \longrightarrow e^+ + \bar{\nu}_\mu + \nu_e \quad (2.10)$$

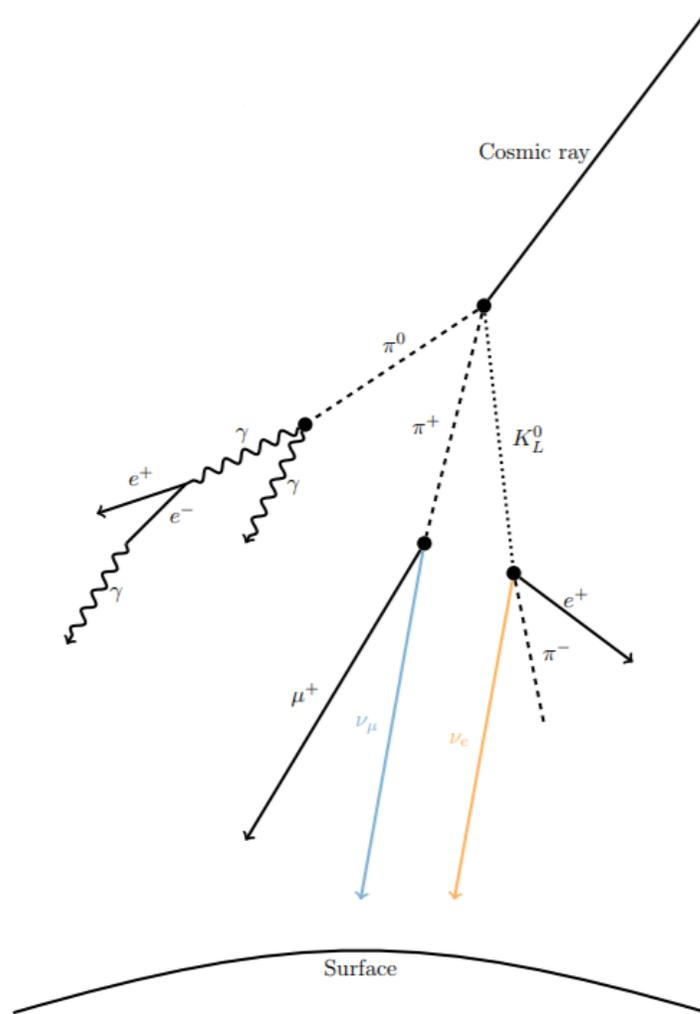


Figure 2.2: Air showers of kaons and pions originating from the interaction of primary cosmic rays with Earth's atmosphere. Source: [17]

## 2.2 Astrophysical Messengers

Astrophysical messengers are exactly what their names imply. They carry information about the source which had emitted them, namely the direction and energy with which they were produced. On an astrophysical scale, producers of such messengers can vary between supernova explosions, supermassive black holes in active galactic nuclei, neutron stars, or our very own Sun. One vital requirement for these particles is the ability to travel astronomically large distances while remaining unaffected on astrophysical time scales. The range that a relativistic particle with mass  $m$  and energy  $E$  can travel is  $r = c\tau\gamma = \frac{\tau E}{mc}$  where  $\tau$  is the lifetime of the particle in its rest frame,  $\gamma$  is the Lorentz factor and  $c$  is the speed of light in vacuum. The three particles that fulfil this criteria are cosmic rays (which include electrons, protons and other ionized nuclei), photons (mostly but not exclusively in the gamma ray spectrum) and of course neutrinos [1].

Oftentimes, astrophysical messengers like cosmic rays, which technically offer the most abundant information about their sources, can be considerably deflected in the interstellar magnetic field on their way from their source to the Earth. Thus their origin can be indiscernible. Therefore a combination of different astrophysical messengers often helps to paint a clearer picture, as many of the candidate astrophysical objects produces if not one but various different messengers. Protons and other cosmic rays produced at the site of origin containing a sufficient amount of energy can create photons and neutrinos in collisions with ambient matter close to the site of acceleration or accretion [1]. Photons however can still be absorbed by dust or softened by electron-positron pair production, which leaves high energy neutrinos as the next best candidate, due to their low cross-sections and inability to interact via any other force besides the weak interaction. Figure 2.3 gives a schematic overview of the previously mentioned processes. This thesis will only look at high energy neutrinos as astrophysical messengers.

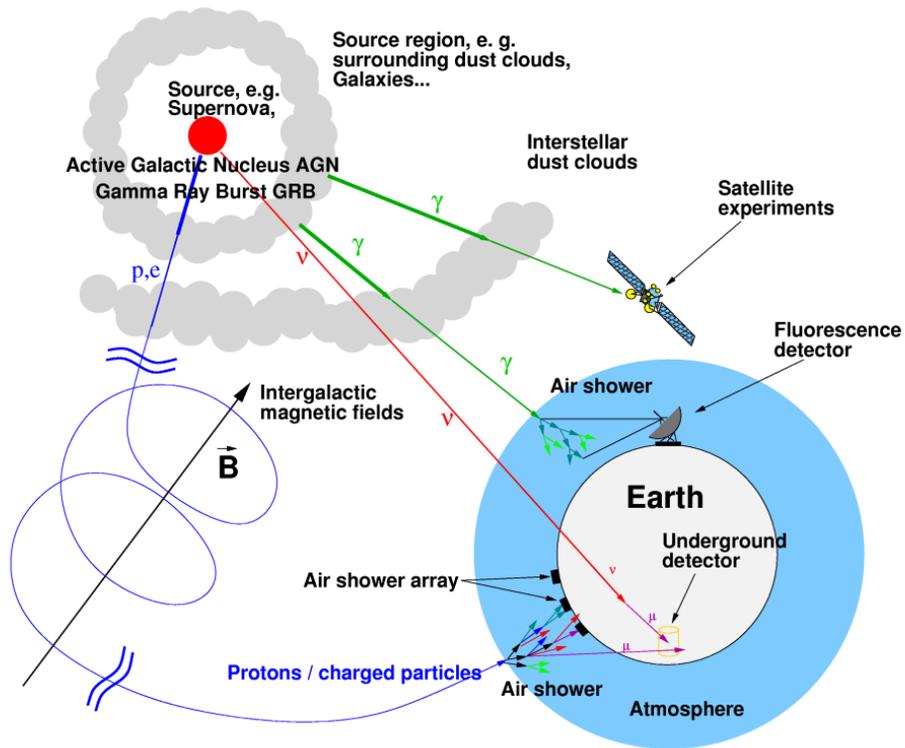


Figure 2.3: Schematic overview of the different astrophysical messenger particles (cosmic rays, photons and neutrinos) originating from an astrophysical source and reaching detectors on Earth, with cosmic rays deflecting off intergalactic magnetic fields, photons softening as they pass through dust clouds, and neutrinos passing through unaffected. Source: [19]

## Chapter 3

# The IceCube Neutrino Observatory

The previous chapter tries to make the point that neutrinos are one of the most elusive particles of the Standard Model due to their lack of charge and inability to interact with any other particle besides through the weak interaction. This makes them prime candidates for astrophysical messengers but it also makes detecting them difficult. However, at high energies, they still have the ability to interact with matter via NC and CC processes. An important outcome of these processes is the production of charged leptons which can then be much more easily observed. The cross section of this interaction increases for neutrinos of higher energies, and also if the ambient matter is denser and more massive. High energy neutrinos become, however, increasingly rarer with energy. A large and dense medium is needed in order to gain as much statistics as possible, while simultaneously having adequate shielding from particles of background sources.

This is where the IceCube Neutrino Observatory comes into place, the world's largest fully operational neutrino detector, located at the South Pole, in the antarctic ice. This chapter aims to discuss IceCube's setup and detection principle, as well as the types of signals it detects.

### 3.1 Experimental Setup

The IceCube Neutrino Observatory is the largest fully operational detector of its kind, and can be found deep underneath the antarctic ice at a depth of around 1500 m to 2500 m. It instruments a volume of around 1 km<sup>3</sup>. Extending down into the ice are 86 strings, each equipped with 60 light detecting digital optical modules (DOM). 8 strings in the center are more densely packed together, separated at 70m from each other, also with more densely packed DOMs with 7 m vertical separation along each string. This forms the DeepCore. The rest of the strings are situated around the DeepCore, each with a horizontal spacing of 125 m, and DOM to DOM separation of 17 m, forming a hexagonal shape. Above the underground structure lie a further 324 DOMs, distributed in 81 detection stations, all situated along the surface of the ice, forming the IceTop [20]. Figure 3.1 depicts a schematic overview of the entire detector setup.

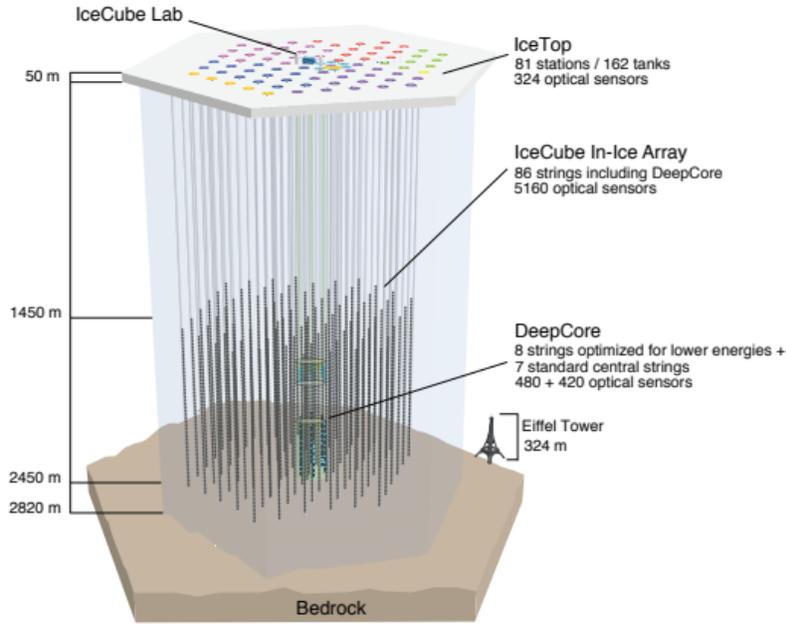


Figure 3.1: The IceCube Neutrino Observatory detector, consisting of 86 strings extending down into the antarctic ice, each equipped with 60 light sensors, forming a hexagonal shape. Source: [21]

Every section within the detector plays a role for different neutrino research purposes. The DeepCore and the surrounding detectors are mainly focused for detecting lower energy neutrinos at  $E_\nu < 10\text{GeV}$ , primarily through atmospheric muons, which plays a role in research in neutrino oscillation. For higher energy events, a larger volume is preferred to increase statistics, and thus the larger and less densely instrumented surrounding volume is used. The IceTop is mainly used for the detection of cosmic rays.

## 3.2 Detection Principle

As discussed in Section 2.1.2, at high enough energies, neutrinos can interact with ambient matter, such as that found in the antarctic ice occupied by IceCube, such that through the CC process, a charged lepton of the same flavor can be produced. These leptons are essential for the (indirect) detection of astrophysical and atmospheric neutrinos, through the so called Cherenkov effect, which will be discussed in the following section, along with the event topologies of each neutrino flavor.

### 3.2.1 Cherenkov Effect

The Cherenkov effect describes the phenomenon which occurs when a charged relativistic particle travels through a dielectric medium with a refractive index  $n > 1$ , at a velocity  $v$  faster than the speed of light in the same medium  $v = c/n$ . As this occurs, the particle emits light in the form of Cherenkov radiation, which occurs due to the asymmetric polarization of the medium in the front and in the rear side of the particle, which causes a change in electric dipole momentum [22]. This effect is the analog to the phenomenon of a sonic boom when an object travels faster than the speed of sound in a medium. This coherent light wave front is generated along the track of relativistic particle at the angle  $\theta_c$ , and like the Mach cone for supersonic movements, the superposition of all the spherical waves results in a Cherenkov cone, such as one depicted in Figure 3.2.

The Cherenkov angle  $\theta_c$  is defined as:

$$\cos \theta_c = \frac{1}{\beta c} \quad (3.1)$$

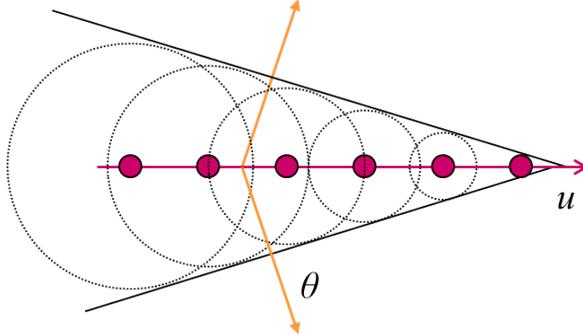


Figure 3.2: Geometry of the Cherenkov cone. A charged particle moving with a velocity  $u$  above the local phase velocity of light in the medium emits a sequence of circular waves. Source: [23]

where  $\beta = v/c$ . It is important to note that most of the relevant particles reaching the antarctic ice are already traveling at near the speed of light. Another important quality is the light yield from such an effect for a particular wavelength  $\lambda$  and distance  $x$ , which is described by the Frank-Tamm equation:

$$\frac{d^2N}{dx d\lambda} = \frac{2\pi\alpha}{\lambda^2} \sin^2 \theta_c \quad (3.2)$$

with the fine-structure constant  $\alpha = e^2/4\pi$ . Due to the  $1/\lambda^2$  dependence, one can see that the equation is maximized by lower wavelengths which is why one tends to see Cherenkov light appear mostly blue [23].

In the IceCube detector, this light is detected by the DOMs, which use high efficiency photomultiplier tubes (PMT), which are used to convert the detected light to an electrical signal. These signals are processed using various methods (for more details about the data filtering and processing see [24]), which are then used for energy and direction reconstruction.

### 3.2.2 Event Topologies

As previously discussed, neutrinos that can be effectively (indirectly) observed are those that interact with matter such as through CC processes, which in turn create charged leptons of the same flavor. These leptons then

have the ability themselves to react with the medium through which they traverse. With high enough energies/velocities they may emit Cherenkov radiation but also can cause electromagnetic cascades themselves. NC processes are of course also possible resulting in measurable hadronic cascades, but in this case, the neutrino flies then through the detector with little to no event signatures. This section serves to explain the typical event topology caused by each neutrino flavor,  $\nu_e$ ,  $\nu_\mu$ , and  $\nu_\tau$ , and their corresponding charged leptons, with a focus on muon neutrinos which are most relevant for this thesis.

### Electron Neutrino

Electron neutrinos  $\nu_e$  firstly can produce hadronic cascades through NC interactions. Then, at high enough energies may generate electrons, but after a critical energy  $E_c$  (in the case of the antarctic ice this energy sits at  $E_c \approx 90\text{MeV}$ ) electrons lose their energy through radiation or in this case Bremsstrahlung, rather than through ionization processes. This results in the emission of photons, which with a large enough given energy can in turn generate an electron-positron pair. Thus we see for most cases an electromagnetic cascade, consisting of electrons, positrons, and photons. Such events such as for a 10 TeV electron results in a maximum length of  $\approx 6.6 \text{ m}^1$  which is very small in comparison to the size of the detector and only allows a median reconstructed angular resolution of 10-15°. However due to this large deposit of neutrino energy in the form of electromagnetic cascades, accurate energy reconstruction is very much feasible for electron neutrinos [25], [26].

### Tau Neutrino

Tau leptons have very large masses compared to their lepton cousins, and are therefore not affected with energy losses in effects such as Bremsstrahlung. Instead they are only affected by ionization processes. Due to their very short lifetimes ( $\tau_\tau = 2.906 \cdot 10^{-13}\text{s}$ ) taus very quickly decay into hadrons at a rate of  $\sim 65 \%$  [27]. Thus the event topology for tau neutrinos takes the form of a so called double bang event, in which we see at first hadronic cascades from NC interactions, followed by a track of Cherenkov radiation caused by the propagation of the tau lepton in the ice, and ends with the final decay of the

---

<sup>1</sup>This was calculated under the assumption that at each case  $n$  of Bremsstrahlung, the emitted photon holds only  $1/2^n$  of the energy of the original electron [14]

lepton or second bang, in which the tau decays either in a electromagnetic or hadronic cascade, after losing considerable energy. Figure 3.3 depicts the summarized event topologies for the electron and tau neutrinos, as well as for the muon neutrino.

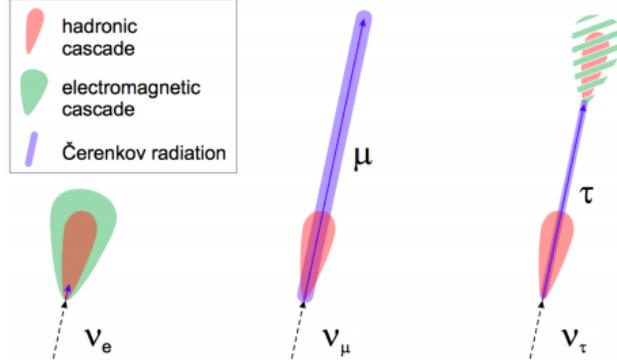


Figure 3.3: The event topologies for electron, muon, and tau neutrinos after interacting with the antarctic ice through either CC or NC processes. Source: [28]

### Muon Neutrino

Muon leptons are also considerably more massive than electrons and therefore before a critical energy of  $E_c \approx 500\text{GeV}$ , they are no longer prone to Bremsstrahlung but instead ionization processes. In this case, the effect is mostly independent of muon energy and therefore the muons can travel a considerably longer distance due to their longer lifetimes ( $\tau_\mu = 2.196 \cdot 10^{-6}\text{s}$ ) [27]. Above  $E_c$ , energy loss is dominated by Bremsstrahlung and other stochastic processes, with loss of energy over distance  $dx$  described by:

$$\frac{dE_\mu}{dx} = A + B \cdot E_\mu \quad (3.3)$$

where  $A = 2.4 \cdot 10^{-3}\text{GeVg}^{-1}\text{cm}^2$  and  $B = 3.2 \cdot 10^{-5}\text{g}^{-1}\text{cm}^2$  [29]. Using this, one can calculate the maximum length that a muon can travel through the antarctic ice, such as one with an energy of 10 TeV, resulting in a track length of 45 km, well beyond the detector volume.

For all muon energies, the muons will additionally emit Cherenkov radiation

along their tracks, which very minimally affects energy loss, but is of great relevance for IceCube. Through detection of the Cherenkov photons via the DOMs, and because of their long lifetime, muons provide the best directional reconstruction for astrophysical and atmospheric neutrinos. However they are still limited by an energy dependant mean angular difference between muon neutrino and muon, given by:

$$\langle \angle(\nu_\mu, \mu) \rangle = \frac{1.5^\circ}{\sqrt{E/\text{TeV}}} \quad (3.4)$$

which for energies  $E_\mu > 1$  TeV, provides angular accuracy already within  $1^\circ$ . Full energy reconstruction is however very difficult, due to many of the muons decaying only beyond the detection volume, thanks to their long lifetime [29]. The muon neutrino topology thus takes the form of a single initial bang in which NC interactions cause hadronic cascades and continues with a long luminous track caused by the secondary muon from CC interactions, as can be seen again in Figure 3.3.

# Chapter 4

## The Neutrino Point Source Analysis

The neutrinos that reach Earth originate from various sources. However, different sources dominate at different neutrino energies, with regions in energy spectrum that can overlap. Figure 4.1 depicts these different sources. IceCube measures energies above the 10 GeV range all the way up to around 100 PeV. It sees many atmospheric neutrinos in the 10 GeV to around 100 TeV ranges, but only starts seeing a dominance in astrophysical neutrinos (from active galactic nuclei) above 100 TeV. Over a decade, IceCube has however collected a large quantity of data which allowed the detection of an astrophysical diffuse neutrino flux with high statistical significance [2].

The main challenge that presents itself is the separation of neutrinos originating from an astrophysical source with high significance from background, i.e. those that come from interactions of cosmic rays with the Earth's atmosphere, and from the diffuse astrophysical flux. The two quantities at hand to do so are only the reconstructed direction and energy of an event induced by a neutrino interaction in the ice. IceCube therefore resolves to statistical methods in order to separate the two. One such method used by the point source analysis in IceCube is the maximum-likelihood search method [30], which will be elaborated in this chapter, along with the calculation of useful statistical quantities thereof, such as p-values, sensitivity, and discovery potential. Also the concept of the neutrino skymaps used in this analysis will be explained, which contain information about locations in the sky which are more likely to be associated with a source of astrophysical neutrinos.

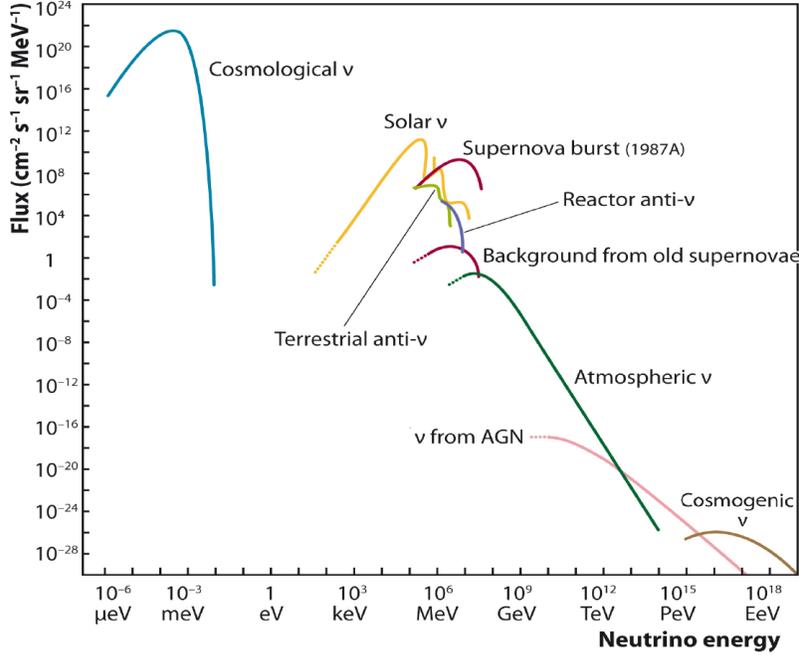


Figure 4.1: Measured and expected fluxes of neutrinos from artificial (reactor) and natural sources, for different neutrino energies. IceCube detects neutrinos from energies of tens of GeV to  $\sim 100$  PeV. Source: [12]

## 4.1 Neutrino Point Source Search

### 4.1.1 Maximum Likelihood Search Method

A direct method to evaluate the data received by the IceCube detectors would be to count events in bins in the size of the detector angular resolution and look for an excess of events over the background expectation, but this method would lead to a loss of information as, for example, events in the center of the bin would be weighted the same as events at the edge of the bin. Therefore a better method was adopted, namely the unbinned maximum likelihood search method [30], in which the data is treated on an event-by-event basis. In order to do this the IceCube data is modelled under two different hypotheses:

**Null Hypothesis  $H_0$ :** The data only consists of atmospheric and diffuse astrophysical (see Section 5.1) background signals.

**Signal Hypothesis  $H_s$ :** The data consists of both atmospheric and diffuse astrophysical background, and astrophysical neutrinos emitted from at least one source candidate.

In order to quantify how well the data fits to either hypothesis, a quantity known as the test statistic  $\mathcal{TS}$  is used, defined by the ratio of the probabilities of receiving the data under either hypothesis. This is more precisely defined as:

$$\mathcal{TS} = -2 \log \left[ \frac{P(\text{Data}|H_0)}{P(\text{Data}|H_s)} \right], \quad (4.1)$$

thus larger  $\mathcal{TS}$  would indicate that the data is less compatible with the background hypothesis.

In order to describe the two probability densities  $P(\text{Data}|H_0)$  and  $P(\text{Data}|H_s)$ , an unbinned likelihood function  $\mathcal{L}$  is defined:

$$\mathcal{L}(x_s, n_s, \gamma) = \prod_{i=1}^N \left( \frac{n_s}{N} S(x_i, \sigma_i, E_i | x_s, \gamma) + \left( 1 - \frac{n_s}{N} \right) B(x_i, E_i) \right). \quad (4.2)$$

We see that it is composed of two separate probability density functions (PDF),  $S_i$  and  $B_i$ , the former of which describes the probability of the  $i$ -th event coming from a direction  $x_i$ , with reconstruction uncertainty  $\sigma_i$ , and reconstructed energy  $E_i$  to originate from a true direction  $x_s$  with energy belonging to a distribution following a power law with spectral index  $\gamma$ .  $B_i$  describes the background distribution density, which is independent of source parameters. The term  $\frac{n_s}{N}$  describes the fraction of signal events relative to the total number of events  $N$ .

For  $n_s = 0$ , the likelihood function is reduced to  $\mathcal{L} = \prod_{i=1}^N B_i$ , which then describes the probability of background only events  $P(\text{Data}|H_0)$ , whereas for a number of expected events  $n_s = \hat{n}_s \geq 1$ , we then get a description for the signal probability  $P(\text{Data}|H_s)$ . Thus the  $\mathcal{TS}$  can be evaluated as:

$$\mathcal{TS} = 2 \sum_{i=1}^N \log \left[ \frac{\hat{n}_s}{N} \left( \frac{S_i}{B_i} - 1 \right) + 1 \right]. \quad (4.3)$$

The number of expected events  $\hat{n}_s$  and spectral index  $\gamma$  are then to be fitted in order to maximize the  $\mathcal{TS}$ , with the use of various maximizing algorithms, hence the name maximum likelihood search method.

### 4.1.2 Significance, Sensitivity and Discovery Potential

In order to understand and determine quantities that will later be of statistical relevance for analysis of IceCube data, in this case the sensitivity, and discovery potential, it is essential to elaborate the p-value. The calculation of the p-value requires the test statistic, and already provides a better understanding, answering the question of whether an astrophysical signal was detected or not.

The p-value  $p$  is a quantity which indicates how extreme a result is, given the background hypothesis  $H_0$ . It is the probability to obtain a result under the assumption of  $H_0$ . The smaller the value, the less compatible is the background hypothesis with the result [31]. It is important to note however that it does not give a direct indication that the opposing hypothesis, in our case the source hypothesis  $H_s$ , is true, but just the unlikelihood of receiving such a result under  $H_0$ . Using the previously defined test statistic  $\mathcal{TS}$  quantity, we can define the p-value  $p$  with:

$$p(\widetilde{\mathcal{TS}}) = P(\widetilde{\mathcal{TS}} \geq \mathcal{TS} | H_0) = 1 - \int_0^{\widetilde{\mathcal{TS}}} P(\mathcal{TS} | H_0) d\mathcal{TS}, \quad (4.4)$$

which would require the entire test statistic probability distribution for background data to be known. This can be achieved through the generation of Monte Carlo simulations.

Even a test to accept or reject a null hypothesis alone is not error-free, given its statistical nature, however, putting such errors into consideration and attempting to minimize them does help to make more reasonable claims. There are two such errors, namely the type I and II errors. The type I error

is the probability  $\alpha$  of falsely rejecting the null hypothesis  $H_0$ , despite it being true. It indicates the likelihood of making a false discovery claim of an analysis. The type II error corresponds to the opposite case, in this case the probability  $\beta$  of falsely accepting the null hypothesis, despite it being false. By minimizing  $\beta$ , one can decide better to accept the opposing hypothesis, and it is often said to be linked to the power of the test [32]. Table 4.1 summarizes the different outcomes and errors that can be made.

| Decision/Truth | $H_0$ is true                        | $H_0$ is false                      |
|----------------|--------------------------------------|-------------------------------------|
| $H_0$ is true  | Correct Decision<br>$p = 1 - \alpha$ | Type II Error<br>$p = \beta$        |
| $H_0$ is false | Type I error<br>$p = \alpha$         | Correct Decision<br>$p = 1 - \beta$ |

Table 4.1: Possible outcomes of a hypothesis test

Using these two errors, one can then define desired outcomes for such errors, in order to better evaluate outcomes of IceCube analyses. IceCube has three important defined thresholds for neutrino flux from simulations:

**Sensitivity:** p-value  $< 0.5$  ( $\alpha = 0.5$ ) for 90% of cases ( $\beta = 0.1$ )

**Discovery Potential:** p-value  $< 5\sigma$  ( $\alpha \approx 2.87 \cdot 10^{-7}$ ) for 50% of cases ( $\beta = 0.5$ )

**90% Upper-Limit:** p-value smaller than p-value obtained by analysing experimental data ( $\alpha = \text{p-value}$ ) for 90% of trials ( $\beta = 0.1$ )

The sensitivity flux is deemed as the threshold flux, at which a point source signal is visible, although the signal could still come from background, with a 10% chance of not seeing the signal. The discovery potential represents the flux needed to see a significant sign of a point source, with a 50% chance of not seeing the signal. The upper limit sets the very unlikely limit of not seeing a point source, representing the flux threshold at which one would measure at least the point source signal p-value with a 90% chance. Figure 4.2 shows an example sketch of a distribution which would pass the sensitivity threshold.

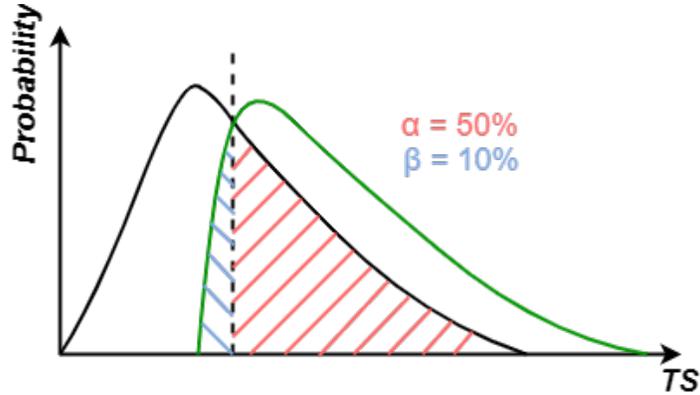


Figure 4.2: A sketch of a distribution (green) which would pass the sensitivity threshold given the background distribution (black), i.e. it contains test statistic values that produce a p-value of at least 50% ( $\alpha = 50\%$ ), for 90% of cases ( $\beta = 10\%$ ).

## 4.2 Neutrino Skymaps

When mapping the sky, the convention at IceCube is to use global spherical coordinates. Thus the angular coordinates of declination  $\delta$  and right ascension  $\alpha$  are used. For this thesis I look at the North Hemisphere, which spans  $-3^\circ \leq \delta \leq 81^\circ$  in declination and a full  $360^\circ$  (24h) in right ascension.

A skymap describes a map of the sky from the IceCube observatory point of view, with information about the probability that a point in the sky corresponds to the location of a neutrino point source. For the skymaps in this analysis, each point in the sky is allocated a p-value, which is calculated using the maximum likelihood search method. A python package called *healpy* [33] is used to subdivide the skymap into pixels. For example, using a resolution parameter of  $N_{side} = 256$  in *healpy*, the North Hemisphere skymap can be subdivided into  $\sim 4 \cdot 10^5$   $0.2^\circ \times 0.2^\circ$  pixels. Figure 4.3 depicts an exemplary full skymap.

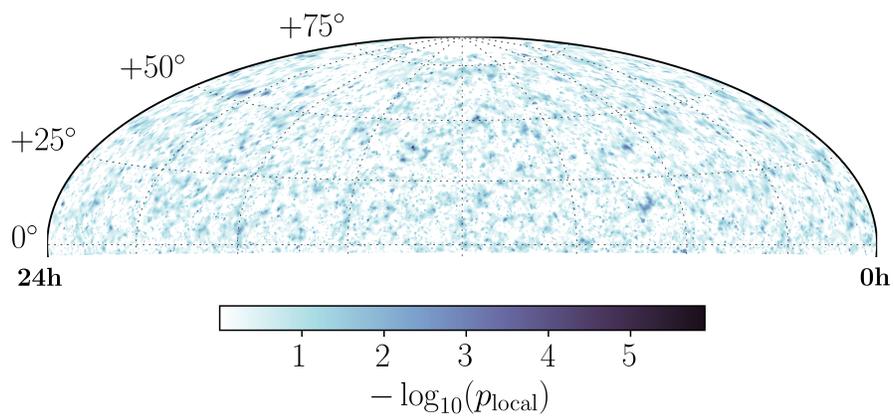


Figure 4.3: An example background skymap of the Northern sky, consisting of pixels which are each assigned a local p-value from the neutrino point source analysis. Plot courtesy of C. Bellenghi.

# Chapter 5

## The Hotspot Population Analysis

In the previous chapter, the neutrino point source analysis was introduced as an unbiased method to search for spots in skymaps with the strongest potential of neutrinos originating from astrophysical sources, i.e. points with very significant p-values. The analysis in this chapter, and main focus of this thesis, serves as a complimentary analysis to the neutrino point source analysis. In this analysis we observe the population of sub-threshold p-values, that is, less significant p-values that would otherwise be missed by the point source analysis. We refer to these p-values as hotspots, hence the name Hotspot Population Analysis (HPA). In this chapter, the details of the HPA method as well as the generation of pseudo-experiments thereof will be presented in order to then contextualize two of my investigations into the HPA, the first looking into the selection method of hotspots during a hotspot search of entire skymaps, the second then being a look into an issue regarding the diffuse astrophysical neutrino flux and the generation of pseudo experiments, as well as the methods attempted to overcome this obstacle.

### 5.1 Analysis Method

The hotspot population analysis presented in this section is based on the method presented in [34], [35].

The main aim of this analysis is to look for an excess of sub-threshold p-values in an entire skymap. This would be an indication of a population of

weak sources which are still producing a significant signal, but that would be missed by the analyses that only look for the strongest source in the sky, such as the point source analysis. Background is defined here as all neutrinos originating from the atmospheric air showers as well as the diffuse astrophysical neutrino flux. The diffuse astrophysical neutrino flux is a measure of the total flux of astrophysical neutrinos that IceCube has detected within a decade. It is assumed to be an isotropic flux across the sky, due to the fact that we do not know where exactly the sources for this signal originate. It is determined through the analysis of the power-law energy spectrum via comparison of Monte Carlo simulations with experimental data, with the latest analysis looking at through-going muon-neutrinos from the Northern Hemisphere [36].

In the case of a population of sub-threshold sources, there would then be a deviation from this uniform expectation. Since this is a counting method of typically large numbers of sub-threshold sources with a local p-value  $p_{local} < p_{thres}$  (i.e. hotspots), one would naturally utilize Poisson statistics. Accordingly, the poissonian p-value becomes the new relevant local p-value for the HPA, defined by:

$$p_{local}^{HPA} = p_{poisson} = \exp(-\lambda) \sum_{m=k}^{\infty} \frac{\lambda^m}{m!} \quad (5.1)$$

This is the probability to find  $k$  or more hotspots with a local p-value below a threshold of  $p_{thres}$ . The mean  $\lambda$  is a function of the selected p-value threshold  $p_{thres}$  and represents the expected number of hotspots from background. Consequently this new local p-value can be minimized by finding the optimal  $p_{thres}$ , which in turn defines the HPA test statistic:

$$\mathcal{TS}_{HPA} = \min_{\lambda} p_{local}^{HPA}(\lambda) \quad (5.2)$$

It is important to note that for the rest of this chapter, the  $-\log_{10}(p_{local}^{HPA})$  will often be used for simplification matters, in which case the  $\mathcal{TS}$  would aim to maximize this value. An extremely low  $\mathcal{TS}$  value might therefore be the indication of a population of sub-threshold sources that are too weak to be detected individually. Figure 5.1 depicts an exemplary output of this analysis, generated from a single simulated background skymap. The test statistic is then the the smallest Poisson p-value produced from the optimal

choice of threshold p-value, i.e. the threshold p-value at which the number of observed hotspots maximally deviates from the expected number.

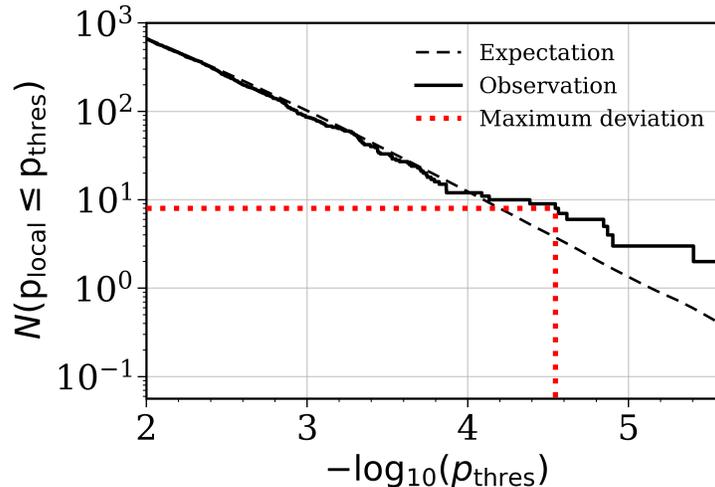


Figure 5.1: An example output of the hotspot population analysis performed on a single simulated background skymap, with the expected number of hotspots (black dashed), the observed number of hotspots from the background skymap (black solid), and the threshold p-value at which the observed count deviates from the expected number (red dotted) and thus produces the smallest Poisson p-value.

## 5.2 Neutrino Hotspot Search Optimization

The first investigation into the HPA aims to find a more optimized method to search for hotspots in a skymap. The following subsections serve to elaborate the current method for searching for hotspots, and the attempts I made to improve it.

### 5.2.1 Current Neutrino Hotspot Search Method

Once a p-value skymap is generated, a direct method would be to count all the p-values above a newly chosen lower threshold. This however does not agree with the expected poissonality as hotspots which lie too close to each other may be over-counted, despite the fact they cannot be differentiated as

two different sources given the angular resolution of the detector. Thus a more strenuous filtering must take place.

The devised search method works on the principle that a single source cannot necessarily be pinned down to one single pixel but rather a small area of pixels. However instead of accounting for all of the pixels in this area, we look at the pixel in the area with the strongest significance and center the local hotspot to this pixel. Thus the first additional criterion for the search method is to search for the lowest p-value among its 8 closest neighbouring pixels (identified using the healpy method `healpy.get_all_neighbours`). This was done in order to come closer the expected poissonality, however further filtering was still needed.

Thus after a skyscan, once all the potential hotspots have been counted, a second criterion is implemented. It compares two hotspots, which are both within a radial vicinity with an angular distance of  $\Psi_{min}$ , and discards the less significant hotspot (hotspot with larger local p-value). In order to optimize this method, and therefore come closer to the expected Poisson distribution, the I searched for the optimal minimum cutting angle  $\Psi_{min}$ . It should be noted that, in this analysis, a threshold p-value of  $-\log(p_{thres}) = 2$  was always used.

### 5.2.2 Expectation, and KS and $\chi^2$ Tests

A sensible approach to optimizing the hotspot search method was to work with pure background skymaps, as we want to use a Poisson p-value as a test statistic for this analysis. The choice of minimum cutting angle  $\Psi_{min}$  has a direct effect in the final count of hotspots. Thus a simple approach would be to create hotspot count distributions and compare them with a poissonian distribution, as done in Figure 5.2. This was indeed done with the generation of 1000 background skymaps through Monte Carlo simulations, and then counting spots at different thresholds, ranging from  $2 \leq -\log(p_{thres}) \leq 7$  in steps of 0.1. In order to generate the poissonian distribution, an expectation value  $\lambda$  is required, which was simply done by taking the mean of the 1000 hotspot counts at each threshold. This was done after once again filtering the hotspots with different  $p_{thres}$ . In order to compare the hotspot count distributions and the Poisson distribution, one can utilize goodness of fit tests, of which two were used:

### Kolmogorov-Smirnov Test

The KS test is mainly used to see how closely two distributions compare to each-other. The test is non-parametric, meaning that it does not assume any particular underlying distribution, and is mainly used for continuous one-dimensional distributions. Although our distributions aren't continuous, the large pool of samples allows for a good approximation as a continuous distribution, which still makes the KS test an effective measure for goodness of fit. The test statistic is given by:

$$\mathcal{TS}_{\text{KS}} = \sup_x |F_1(x) - F_2(x)| \quad (5.3)$$

where  $F_1(x)$  and  $F_2(x)$  are the two separate cumulative distribution functions (CDF) of the two statistical distributions being tested. In our case this just means looking for the two corresponding bins with the largest difference [32].

### $\chi^2$ Test

The  $\chi^2$  test (or more specifically Pearson's chi squared test) is similar to the KS test, but tests instead how well observed data fits to a hypothesized distribution, which is the Poisson distribution in our case. The test is meant to be used with frequency distributed data, which is exactly the case with the hotspot count distributions. The test statistic is given by:

$$\mathcal{TS}_{\chi^2} = \sum_{i=1}^N \frac{(O_i - E_i)^2}{E_i} \quad (5.4)$$

where  $O_i$  and  $E_i$  are the observed and expected frequencies respectively, and is summed over  $N$  bins. While the  $\chi^2$  test works with discontinuous distributions, because it sums over all bins, it is much more sensitive to excessively deviating values, which could then greatly change the outcome of the test statistic values, especially for smaller numbers of bins  $N$ . [32]

Smaller  $\mathcal{TS}$  values for either goodness of fit test would lead to bigger p-values, according to Equation 4.4, which in turn indicate that the two distributions fit closer together. Using the hotspots counts at each threshold obtained by running the search on 1000 background skymaps, I could consequently gain

some reasonable insight on the effect of changing the minimum cutting angle  $\Psi_{min}$ .

### 5.2.3 Results and Discussion

The 1000 background maps were searched for hotspots, starting with  $\Psi_{min} = 1^\circ$ , as this was the value used in the previous version of this analysis [37]. KS tests and  $\chi^2$  tests were then repeatedly performed for different p-value thresholds, in steps of 0.1 between  $2 \leq -\log(p_{thres}) \leq 7$ . Figure 5.2 shows the hotspot counts distributions for various  $-\log(p_{thres})$ . One can see that within errors, the distributions already fit quite well together, even for very large  $p_{thres}$  which is where the most deviations are expected.

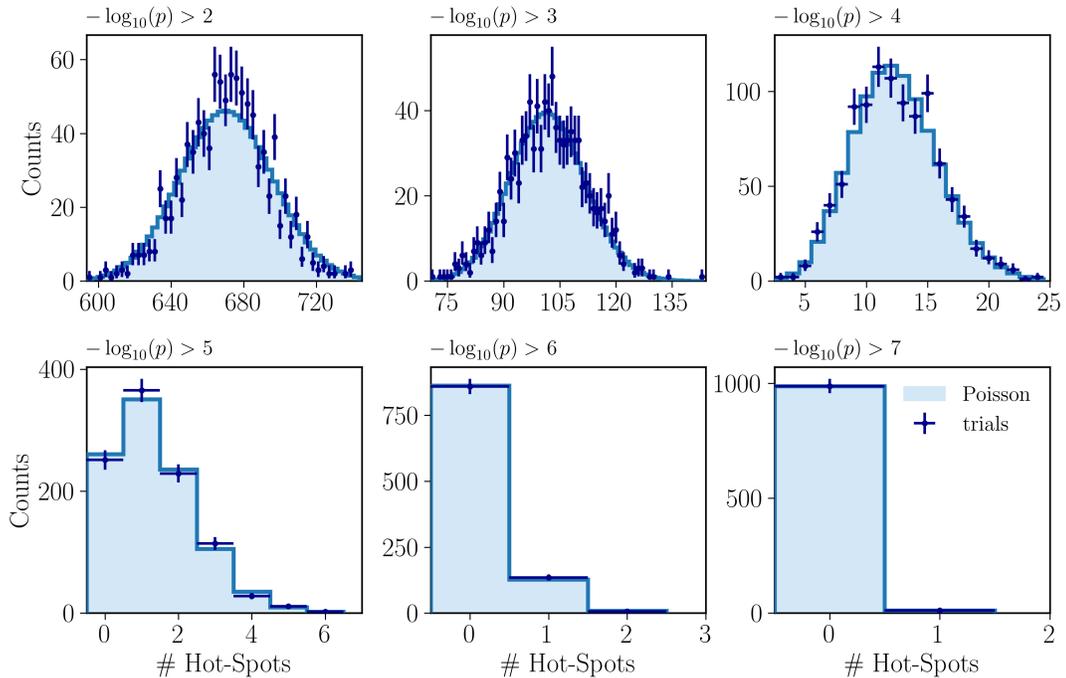


Figure 5.2: Hotspot count distribution histograms from 1000 background skymaps, at different p-value thresholds, with an minimum cutting angle of  $\Psi_{min} = 1^\circ$ . The dark blue error bars indicate the observed values, while the underlying shaded histograms indicate the Poisson distributions calculated using the mean of the trials

The same procedure was then performed for different  $\Psi_{min}$ , looking first at KS tests for various  $\Psi_{min} > 1$  and then separately testing for  $\Psi_{min} < 1$ , performing both KS and  $\chi^2$  tests.

### Test for $\Psi_{min} > 1^\circ$

This was tested for cutting angles all the way up to  $5^\circ$ . I firstly noticed that the values at the tails for the hotspot count distributions were much lower than the expected Poisson values. I came to the conclusion that this was largely due to the cutting of many hotspots since the criteria for a hotspot becomes stricter with larger  $\Psi_{min}$ . Figures 5.3 and 5.4 show exactly this behaviour already for distributions made with  $\Psi_{min} = 2^\circ$  and  $3^\circ$  respectively at a threshold p-value of  $-\log(p_{thres}) = 2$ . They also show how much the  $\chi^2$  residuals deviated at different numbers of hotspots per map.

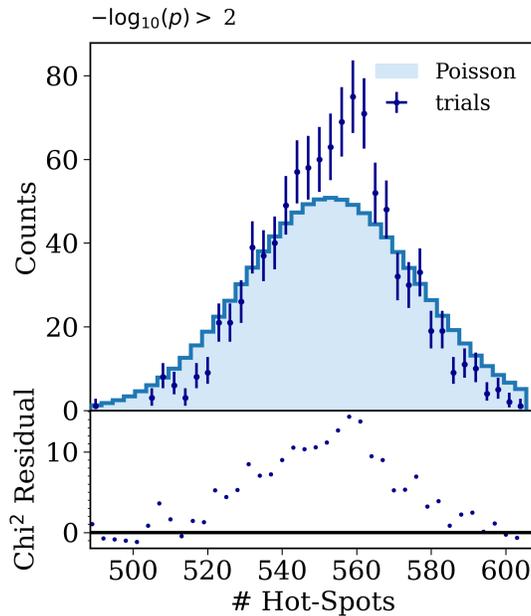


Figure 5.3: Top: Hotspot count distribution histograms from 1000 background skymaps, at a threshold p-value of  $-\log(p_{thres}) = 2$  for  $\Psi_{min} = 2^\circ$ . The dark blue error bars indicate the observed values, while the underlying shaded histograms indicate the Poisson distributions calculated using the mean of the trials. Bottom: residuals of the chi-squared test for different observed numbers of hotspots per map.

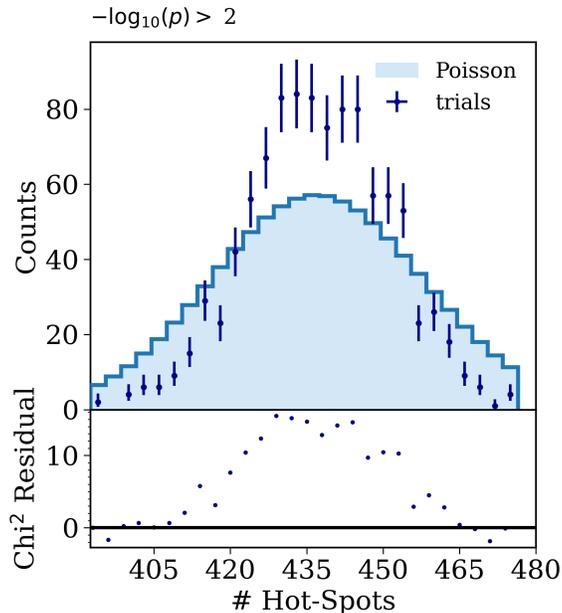


Figure 5.4: Top: Hotspot count distribution histograms from 1000 background skymaps, at a threshold p-value of  $-\log(p_{\text{thres}}) = 2$  for  $\Psi_{\text{min}} = 3^\circ$ . The dark blue error bars indicate the observed values, while the underlying shaded histograms indicate the Poisson distributions calculated using the mean of the trials. Bottom: residuals of the chi-squared test for different observed numbers of hotspots per map.

This is furthermore ascertained when looking at the distribution of KS p-values over different thresholds, which can be seen in Figure 5.5.

As previously mentioned, a larger p-value in this case would indicate a better fitting distribution, and from the KS tests it is clear that a minimum cutting angle of  $\Psi_{\text{min}} = 1^\circ$  performs best, especially around the range of  $-\log(p_{\text{thres}}) < 3.5$ . Beyond this, the KS tests output similar results which is expected as less cutting is likely, due to the larger improbability of finding two very significant hotspots so near each other so frequently.

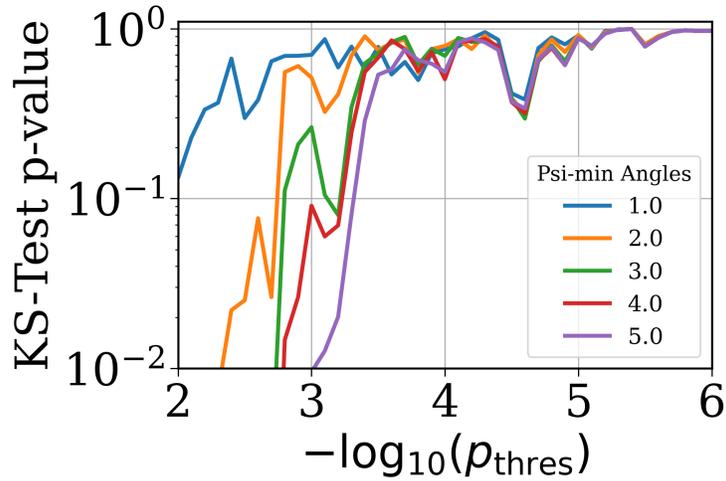


Figure 5.5: Distribution of KS p-values for different thresholds  $-\log(p_{thres})$  and different minimum cutting angles  $\Psi_{min} \geq 1$

**Test for  $\Psi_{min} < 1^\circ$**

This was more meticulously tested, testing a range of  $\Psi_{min}$  starting from  $0.5^\circ$  to  $1^\circ$  in steps of  $0.05^\circ$ . Both KS tests and  $\chi^2$  tests were performed, the results of which can be seen for selected  $\Psi_{min}$  in Figures 5.6 and 5.7.

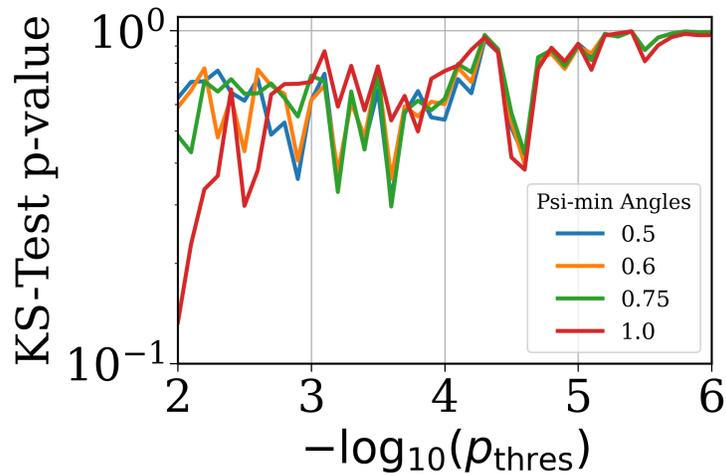


Figure 5.6: Distribution of KS p-values for different thresholds  $-\log(p_{thres})$  and different minimum cutting angles  $\Psi_{min} \leq 1$

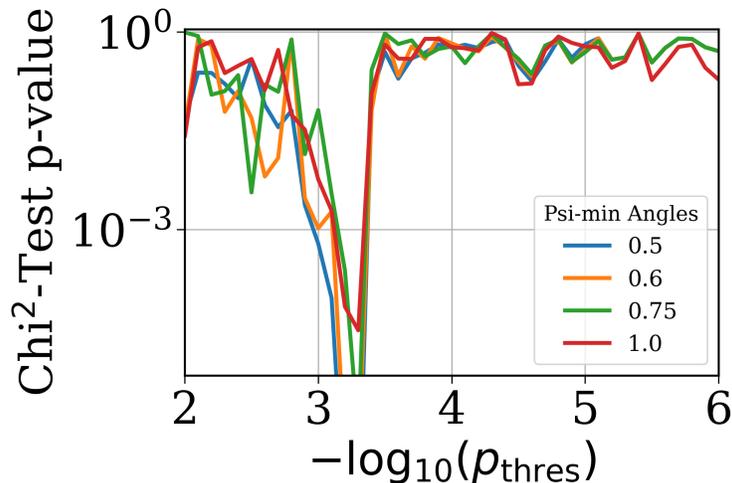


Figure 5.7: Distribution of  $\chi^2$  p-values for different thresholds  $-\log(p_{thres})$  and different minimum cutting angles  $\Psi_{min} \leq 1$

Here one can see that changing  $\Psi_{min}$  produces varied results, from which I could not draw straightforward conclusions. I attempted to parameterize the KS curves with exponential functions and logarithmic functions but this did not yield any useful results. Thus it was deemed that any further tests would lead to further inconclusive results. The large dip for many cutting angles by the  $\chi^2$  test at  $-\log(p_{thres}) \approx 3.2$  was concluded to be resulting from the contribution of a very large residual from a large deviating value, which is not surprising given the sensitivity of the goodness to fit test. Tests for  $\Psi_{min} < 0.5$  were deemed insensible since this would be nearing the median angular resolution that we have for muon tracks.

### 5.2.4 Conclusion

The test for the optimal minimum cutting angle was indeed an interesting investigation, that was deemed necessary for optimizing hotspot searches in terms of populations of weak astrophysical neutrino sources. Testing for  $\Psi_{min} > 1^\circ$  proved that smaller cutting angles are indeed preferred, however for cutting angles  $0.5^\circ < \Psi_{min} < 1^\circ$ , inconclusive results were produced, leading to the conclusion that  $\Psi_{min} = 1^\circ$  is indeed a reasonably safe choice for minimum cutting angle, as the median angular resolution for the single muon event from Monte Carlo simulations goes from  $0.1^\circ$  to more than  $1^\circ$

(near to the North Pole), although the possibility for a further optimized and smaller angular cut still remains. An minimum cutting angle of  $\Psi_{min} = 1^\circ$  will thus be used for the rest of this thesis.

### 5.3 Diffuse Astrophysical Neutrino Flux Investigation

In order to make accurate statements with the Hotspot Population Analysis when applying it to real IceCube data, the HPA must first be optimized and tested rigorously on simulated data, in order to iron out any discrepancies and understand how it reacts to different data. In order to do so, we generate pseudo data, which either tries to realistically mimic real data or provide extreme cases in order to understand how the HPA performs. This means the generation of pseudo skyscans or so called pseudo experiments or trials. It also allows us to make statements about the detection power of the HPA and in turn the conclusions we will eventually draw about the existence of a population of high energy neutrino sources, given our current set of analytical tools. In order to do this one must also try to realistically account for the background (for the HPA), including the diffuse astrophysical flux.

For the neutrino point source analysis, the diffuse astrophysical flux plays a very minimal roll in its outcomes. This is due to the fact that the point source analysis searches for very significant single point sources, which are associated with an astrophysical neutrino flux that is much stronger than the contribution coming from a diffuse flux. In the HPA, however, we look at much weaker sources, and in a certain regime, i.e. the large population regime, the astrophysical neutrino flux coming from the population becomes comparable with a diffuse flux of neutrinos.

There has been a known issue regarding the injection of sources in the pseudo trial generation process, pertaining to the unaccounted superfluous injection of the diffuse astrophysical flux of neutrinos. This section aims to explain the exact method used to generate pseudo background trials as well as signal trials, investigate the influence of diffuse astrophysical flux and the issue, and present the attempted methods to overcome this and also discuss its implications.

### 5.3.1 Generation of Background Trials

The most straightforward strategy to generate one trial, would be to generate an entire p-value skymap using Monte Carlo simulations. This however is very computationally demanding as the generation of a single skymap takes  $\sim 100$  CPU hours. Since tens of thousands of skymaps would be needed for various tests, this would altogether be extremely time and resource consuming. For the HPA, we work around the problem using the same strategy that was used in a previous works [35], [37].

This method first requires the generation of background trials. This is done with the use of skyscans from 2000 background skymaps of the Northern Hemisphere, extracting all of the hotspots with a p-value of  $-\log(p_{local}) > 2$ . These hotspots (which make up roughly 1.34 million independent hotspots) are then put into a background hotspot set. In order to then much more simply generate a pure background trial, one can then just select  $N$  random samples from this set, where  $N$  is a random number selected from the hotspot count distribution with the lowest p-value threshold (also  $-\log(p_{thres}) = 2$ ), calculated from a separate set of 1000 background skymaps.

### 5.3.2 Generation of Signal Trials

To generate signal trials, i.e. trials with populations of signals of weak sub-threshold sources, a slightly more complicated set of steps is required. In essence, the method used here is to simply inject sources onto an already existing generated background trial, generated using the method from the previous section. However there are a few things one must account for before simply injecting a signal at random.

Much like the background hotspot pool, a signal pool is generated from separate trials from the single point source search. Many trials are generated by injecting different signals at many different declinations chosen uniformly in  $\sin \delta$ , which are then extracted into the signal pool, this time additionally saving information about the true number of injected events  $n_{inj}$ , declination, and the local p-value calculated with the unbinned likelihood-ratio test.

Using this selection, one can complete a full signal trial generation:

1. A pseudo background trial is first generated, adding the randomly chosen background p-values to a hotspot count list.

2. A number of sources to be injected  $N$  is chosen, choosing  $N$  declinations and  $N$  neutrino fluxes corresponding to each source.
3. The fluxes are converted in a mean number of injected neutrinos for each source, using the detector effective area<sup>1</sup> and selected locations.
4. From the signal pool  $N$  trials are randomly chosen for the selected declinations and mean number of injected neutrinos.
5. The angular cut with  $\Psi_{min}$  is simulated when adding a signal to the hotspot count list, that was already used in [37]:
  - For each signal p-value, a random number  $r$  between 0 and 1 is drawn from a uniform distribution.
  - The p-value is added only if  $r > n \cdot \Omega_{spot}/\Omega_{NH}$  where  $n$  is the number of hotspots already in the list,  $\Omega_{spot} = \pi\Psi_{min}^2$  is the solid angle covered by a disc, and  $\Omega_{NH}$  is the solid angle covered by the Northern Hemisphere.
  - if  $r < n \cdot \Omega_{spot}/\Omega_{NH}$ , this corresponds to two hotspots being too close to one-another, in which case the hotspot with the smallest p-value is kept/added and the other discarded.

The method in the final step of this procedure, regarding the angular cut, accounts for the fact that the probability to find two hotspots that are too close to each other increases with the number of hotspots in the list. And thus a signal trial is created, mimicking the output of a skyscan, and therefore ready to be analyzed.

### 5.3.3 Influence of the Diffuse Astrophysical Neutrino Flux

The diffuse astrophysical flux of neutrinos takes into account the total amount of astrophysical neutrinos seen by IceCube since the start of operations, and, together with atmospheric neutrinos, is included in the background model for this analysis. It also forms the limit for the total flux of neutrinos that can be expected to be detected from the sky. Thus, in principle, it does not make

---

<sup>1</sup>this is a measure of how effectively neutrinos can be detected given the surface cross section area of the detector, [38]

sense to inject more than the total astrophysical flux in a simulated skymap. However for varying signal strength, the amount of diffuse astrophysical neutrinos in the background simulation would have to be adapted. It is, however, computationally not feasible to calculate the background expectation for different amounts of diffuse astrophysical neutrino.

The point source analysis includes both the atmospheric and diffuse astrophysical flux in its background model. Therefore when the fluxes for the single sources are calculated, a portion of the astrophysical diffuse flux in the injection location in the sky is also added to this flux. The contribution of the diffuse astrophysical flux for small amounts of injected sources is negligible, but for larger numbers of sources, the diffuse astrophysical flux plays a role, as the total flux from the sources should eventually reach the diffuse astrophysical flux. Figure 5.8 depicts the increase in the fraction of hotspots with a local p-value  $-\log(p_{local}) > 2$ , with and without the astrophysical diffuse flux, with a energy power law spectral index of  $\gamma = 2.28$  [36].

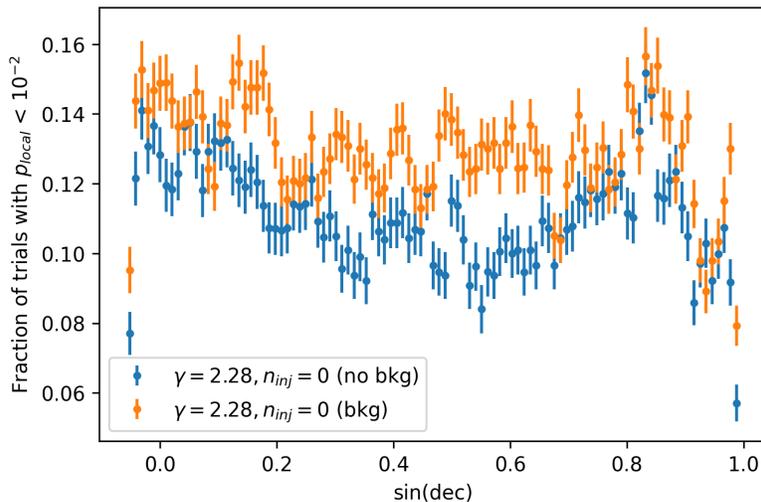


Figure 5.8: Fraction of hotspots with  $-\log(p_{local}) > 2$  in 2000 trials at different declinations, with (orange) and without (blue) the diffuse astrophysical neutrino flux contribution, without any injected neutrinos, with a spectral index of  $\gamma = 2.28$ .

We see what appears to be a general 20% increase at most declinations, showing that even with trials with no signal events, there is already an excess in small p-values above a threshold of  $-\log(p_{thres}) = 2$ . To understand the

impact of the diffuse astrophysical flux on the number of detected hotspots in a skymap, a further 100 background skymaps were performed, for both skymaps with double the diffuse astrophysical flux contribution and without any diffuse astrophysical flux contribution, Figure 5.9 depicts the results when calculating the expectation from background-only pseudo-experiments of the number of hotspots for different threshold p-values for both cases.

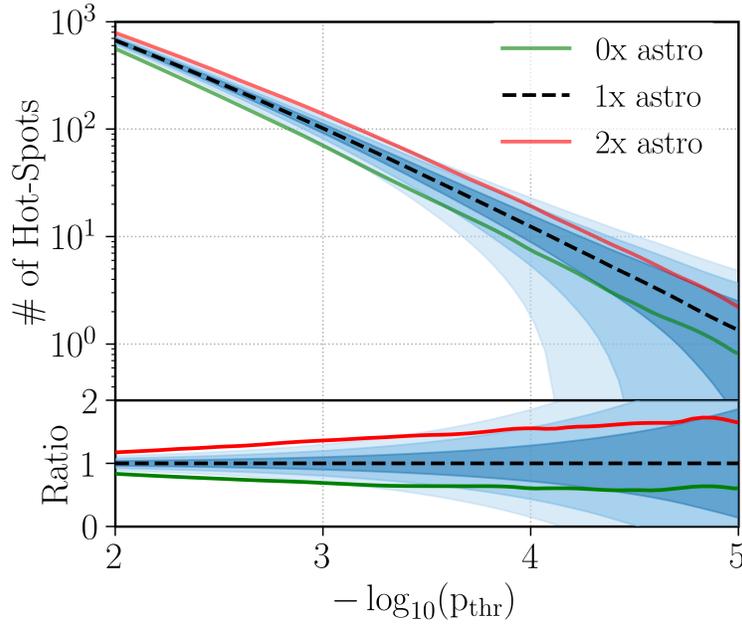


Figure 5.9: Top: Expected number of hotspots for background-only pseudo-experiments for different significance thresholds, from skymaps with 0x (green), 1x (black dashed), and 2x (red) the diffuse astrophysical neutrino flux contribution. The blue shaded areas correspond to the 1, 2, and 3 standard deviations from the expectation of 1x the background diffuse astrophysical flux contribution. Bottom: Ratio between the counts between 1x the diffuse astrophysical flux and both 0x and 2x (same color scheme). Plot courtesy of C. Bellenghi

One can see that the impact is indeed large as the difference in expected number of hotspots for both 0x and 2x the diffuse astrophysical flux contribution is already three standard deviations away from the standard expectation for  $-\log(p_{\text{thres}}) < 4$ .

### 5.3.4 Accounting for the Astrophysical Diffuse Neutrino Flux in Simulations

Further investigation into the impact of the diffuse astrophysical flux in the HPA revealed an important consideration that needed to be accounted for. This was made clear when calculating the fluxes needed for sensitivity and discovery potential from generated pseudo experiments with source injection, using the source injection method described in Section 5.3.2. Figure 5.10 depicts the different flux per source needed for sensitivity and  $5\sigma$  discovery potential.

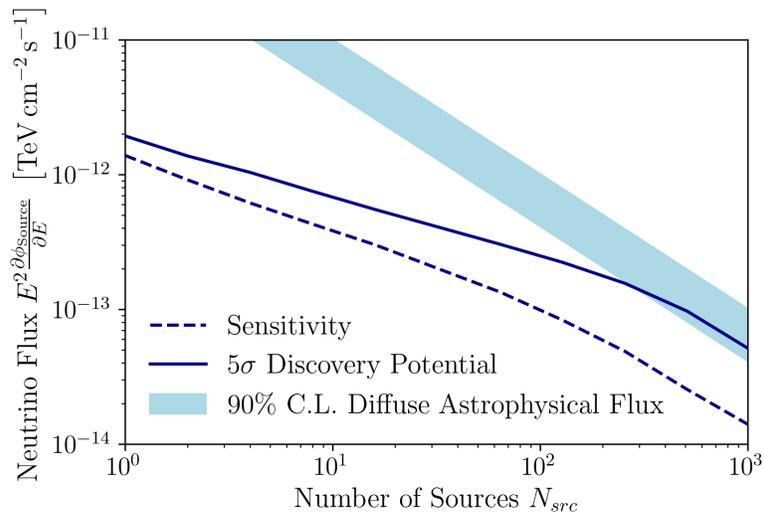


Figure 5.10: Sensitivity (dashed) and Discovery Potential (solid) flux per source for various numbers of injected sources of equal strength. The diffuse astrophysical neutrino flux is also depicted within 90 % confidence limits (light blue).

It should be noted that this was performed with injection of sources of equal strength (neutrino flux) at Earth in order to test the power of the analysis. The bending of the curves between 600 - 800 sources is a clear indicator of underestimations of the diffuse astrophysical flux, as around these numbers we should expect fluxes similar to the diffuse astrophysical flux as we reach total populations of hotspots around  $\sim 670$  as expected from pure background as can be seen in Figure 5.11, however the curves run parallel and below the background which would mean that we observe fluxes per source

less than background which is not possible.

This may be due to the injection method, which does not account for the fact that it also injects background flux alongside source flux on top of the background trials which already contained background flux. I devised two methods or amendments to the injection procedure to tackle this issue, which were the following:

### Method 1

This first method was devised in order to account for the fact that injection method continues to add hotspots to the hotspot count list, beyond the expected number of hotspots from simulated background skymaps. This leads to the underestimation of the diffuse flux for higher numbers of sources. In order to do so I refer to the previously calculated background hotspot count mean per skymap for the lowest threshold ( $-\log(p_{\text{thres}}) = 2$ ), which was  $\sim 670$ . This can be seen in the previously determined hotspot count distribution also seen in Figure 5.11. As this distribution is obtained from background neutrino skymaps which include both the atmospheric and astrophysical fluxes, this also indicates the number of hotspots that are generated on average by the astrophysical neutrino flux when it is randomly distributed across the sky.

I account for more extreme cases and look at the count for an additional 3 standard deviations:  $670 + 3 \cdot \sqrt{670}$  and round this number up to 750, as these numbers are anyway only used for testing purposes. Using this, I modify the signal injection method by adding an additional step to the end of the procedure:

6. Once a limit of 750 hotspots is reached in the hotspot count list, every time a new hotspot from the signal pool is chosen, it is compared with a randomly chosen hotspot in the hotspot count list, and the more significant hotspot is kept, and the other discarded.

This is done as it is clear that since the diffuse astrophysical flux serves as a limit for the total amount of sources that we may see in the sky, it would make little sense to inject any more sources beyond this, as it would mean these sources would have to have less than a single neutrino due to how low the flux per source would have to be to already achieve sensitivity.

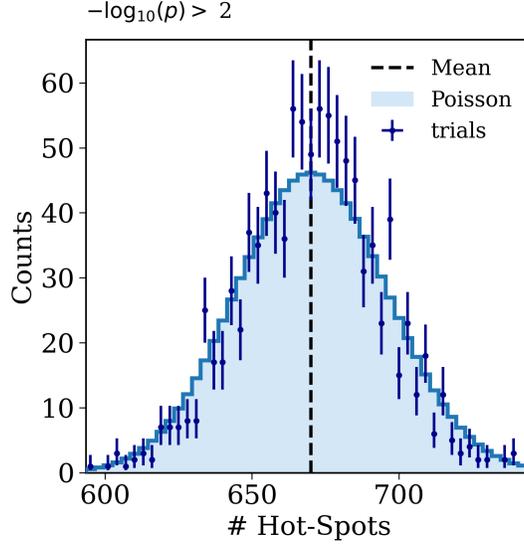


Figure 5.11: Hotspot count distribution histograms from 1000 background skymaps, at a threshold p-value of  $-\log(p_{\text{thres}}) = 2$  for  $\Psi_{\text{min}} = 1^\circ$ . The dark blue error bars indicate the observed values, while the underlying shaded histograms indicate the Poisson distributions calculated using the mean of the trials, which is also indicated by the black dashed line.

## Method 2

The second method was devised later as an indirect method to account for the background flux in the signal. In this case I keep the same injection method and I directly estimate the contribution of astrophysical diffuse neutrino flux or in this case the number of astrophysical background neutrinos per hotspot for every hotspot added from the set of signal p-values. To find this value, I use the differential diffuse astrophysical flux  $d\phi/dE$  which was estimated in [36]. It takes the form:

$$\frac{d\phi}{dE} = \Phi_{\text{astro}} \cdot \left(\frac{E_\nu}{E_0}\right)^{\gamma_{\text{astro}}} \quad (5.5)$$

where the flux normalization was estimated to be  $\Phi_{\text{astro}} = (1.44 \pm 0.25) \cdot 10^{-18} \text{GeV}^{-1} \text{cm}^{-2} \text{s}^{-1} \text{sr}^{-1}$ , the spectral index  $\gamma_{\text{astro}} = -2.28 \pm 0.09$ , also while using an energy normalization of 100 TeV, [36].

Taking the integrated neutrino flux in the Northern sky over nine years of

exposure, and over energies ranging from 200 GeV – 10 TeV [39], and then multiplying this with the fraction of a  $1^\circ$  disc in the entire sky  $\Omega_{\text{min}}/\Omega_{\text{NH}}$ , we get  $\sim 0.3$  neutrinos per disc from diffuse astrophysical flux contribution. This is an averaged number for the entire Northern Hemisphere which would not be very accurate for small populations of sources<sup>2</sup> but this is acceptable since we already know that the diffuse astrophysical flux contributes very little anyway at small populations. For larger populations, this inaccuracy becomes negligible as half the sources will be overestimated and the other half underestimated, which would on average lead to the correct calculation of the contribution.

### 5.3.5 Results and Discussion

New sensitivity and discovery potential fluxes were calculated using both methods, however it was clear that the best procedure would be to use both methods together, as they were both deemed soundly logical and did not interfere with each other directly. Method 1 alone did little to change the underestimation although there were very slight changes to the bending of the sensitivity curve, as can be seen in Figure 5.12, meanwhile Method 2 greatly changed the curve as expected. Figure 5.13 depicts the change in the sensitivity and  $5\sigma$  curves when using the second method.

We can see now the straightening of the bends of the curves from the previously revised injection method, as well as the intersection with the diffuse astrophysical flux, which would be expected if more and more significant sources are injected, thus increasing the overall flux. This also very much brings to attention the fact that it does not make sense to inject too many sources beyond the average of 670 number of sources produced from background simulations already including the diffuse astrophysical flux.

---

<sup>2</sup>this is due to the fact that the IceCube detector effective area is declination dependant

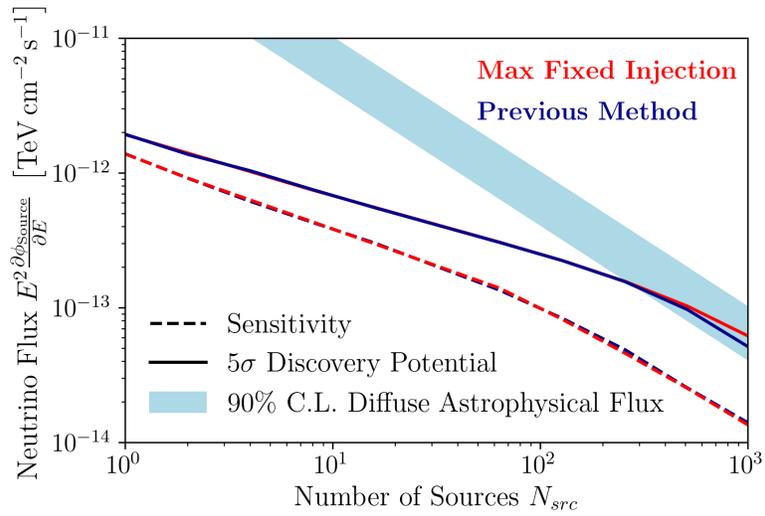


Figure 5.12: Sensitivity (dashed) and Discovery Potential (solid) flux per source for various numbers of injected sources of equal strength, comparing the fixed max number of injections method (red) to the original injection method (blue). The diffuse astrophysical neutrino flux is also depicted within 90 % confidence limits (light blue).

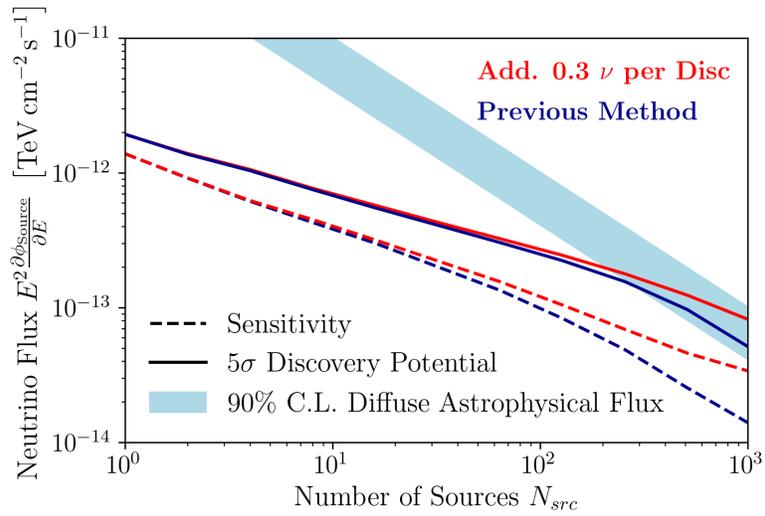


Figure 5.13: Sensitivity (dashed) and Discovery Potential (solid) flux per source for various numbers of injected sources of equal strength, comparing the additional 0.3 neutrinos per disc method (red) to the original injection method (blue). The diffuse astrophysical neutrino flux is also depicted within 90 % confidence limits (light blue).

# Chapter 6

## Conclusion and Outlook

The investigation into the optimization of the neutrino hotspot population analysis proved to be an interesting and insightful study of the search for sub-threshold high energy neutrino sources as well as of the diffuse astrophysical neutrino flux at the IceCube Neutrino Observatory. While the neutrino point source analysis aimed to search for single points in the sky of very high significance relating to neutrinos originating from astrophysical sources, the hotspot population analysis aims to generally look for populations of weaker neutrino sources, for any threshold, making it a powerful analysis indeed.

In this investigation I had first attempted to optimize the search for hotspots within the HPA. As it was established that background hotspot count distributions should ideally follow poissonian distributions, the search was optimized to maximize the poissonality of the hotspot count distributions it produces as an outcome. This was done by changing the defined minimum distance that hotspots may occupy between one another in a neutrino skymap, i.e. the minimum cutting angle  $\Psi_{min}$ . Testing for  $\Psi_{min} > 1^\circ$  proved that smaller cutting angles are indeed preferred as was expected. For cutting angles  $0.5^\circ < \Psi_{min} < 1^\circ$ , inconclusive results were produced, leading to the conclusion that  $\Psi_{min} = 1^\circ$  is indeed a reasonably safe choice for minimum cutting angle, as the median angular resolution for the single muon event from Monte Carlo simulations goes from  $0.1^\circ$  to more than  $1^\circ$  (near to the North Pole) although the possibility for a further optimized and smaller angular cut still remains.

For the second part of this investigation, I had studied the influence of the diffuse astrophysical neutrino flux on the HPA. In particular, the already known problem of the underestimation of the diffuse flux during the generation of pseudo trials with injected sources was studied. They are essential for the HPA in order to optimize it as a whole and understand how it would react to different data. This was observed by the sensitivity fluxes needed for different numbers of sources, estimated from these generated trials, which never intersected the diffuse astrophysical flux, where expected. Thus I devised two different methods or amendments to overcome this; the first method set a limit on the number of hotspots which a trial may have in total, taking the mean of the background hotspot count distribution plus three standard deviations as a limit. This saw slight improvements but the diffuse flux was still never crossed. The second method directly added the contribution of the diffuse flux per  $\Psi_{min} = 1^\circ$  disc containing the hotspot, which was estimated to be 0.3 neutrinos per disc. This led to considerable improvements to the underestimation and even saw to it that the discovery potential curve crossed the diffuse flux limit where expected.

Besides improving the underestimation, this investigation brought into focus the fact that testing for numbers of sources beyond the expected limit of the diffuse flux makes little sense, as it would mean that these sources produce less neutrinos than the background itself. The improvement in signal injection, however, now allows for much more accurate testing of the HPA, and also allows for better and more accurate simulations of populations of weak sub-threshold neutrino sources, in order to better compare with Ice-Cube data. In particular it will help to set an upper limit on the flux that might come from a different population of sources, in the case that we do not see any sources above  $3\sigma$  in significance.

Of course there is always room for improvement, such as by using declination based strengths for sources instead of equal strength at Earth, which would also then have to be applied to the additional background flux applied during injection. A look into star formation rates would, for example, help to provide a more accurate compensation for this. The *FIRESONG* tool could be used, which is meant to simulate such neutrino source populations according to different star formation models [40].

# Bibliography

- [1] T. K. Gaisser, R. Engel, E. Resconi, *Cosmic Rays and Particle Physics*, Cambridge University Press, 2016
- [2] IceCube Collaboration, *Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector*, *Science*, 2013, Vol 342, Issue 6161
- [3] The IceCube Collaboration et al, *Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A*, *Science*, 2018, Vol 361, Issue 6398
- [4] E. Fermi, *Versuch einer Theorie der  $\gamma$ -Strahlen. I*, *Zeitschrift für Physik*, 1934, volume 88, pages 161–177
- [5] G. Rajasekaran, *Fermi and the theory of weak interactions*, *Resonance* (Indian Academy of Sciences, Bangalore), Vol 19, No 1, p18-44, January 2014, arXiv:1403.3309 [physics.hist-ph]
- [6] C. L. Cowan, Jr., F. Reines, F. B. Harrison, H. W. Kruse and A. D. McGuire, *Detection of the Free Neutrino: A Confirmation*, 1956, *Science* 124, 103
- [7] S. Paul, N. Kaiser, W. Weise, et al, *Teilchen und Kerne*, lecture script at Technische Universität München
- [8] J. Holdsworth, *Weak Interaction*, Wikipedia website, [https://en.wikipedia.org/wiki/Weak\\_interaction](https://en.wikipedia.org/wiki/Weak_interaction)
- [9] Y. Fukuda et al (Super-Kamiokande Collaboration), *Evidence for Oscillation of Atmospheric Neutrinos*, *Phys. Rev. Lett.* 81, August 1998, 1562

- [10] Z. Maki, M. Nakagawa, S. Sakata, *Remarks on the Unified Model of Elementary Particles*, Progress of Theoretical Physics, Volume 28, Issue 5, November 1962, Pages 870–880
- [11] A. Yu. Smirnov, *Solar neutrinos: Oscillations or No-oscillations?*, September 2016, arXiv:1609.02386 [hep-ph]
- [12] IceCube Masterclass, *The Detection of Neutrinos in IceCube*, <https://masterclass.icecube.wisc.edu/en/learn/detecting-neutrinos>
- [13] J. K. Becker, *High-energy neutrinos in the context of multimessenger astrophysics*, Phys.Rept.458:173-246, January 2008, arXiv:0710.1557
- [14] M. J. Huber, *Sensitivity Studies for Blazar Stacking: Searches with the IceCube Neutrino Observatory*, Master Thesis, Technische Universität München, December 2015
- [15] S. Hümmer, M. Maltoni, W. Winter, and C. Yaguna, *Energy dependent neutrino flavor ratios from cosmic accelerators on the Hillas plot*, Astroparticle Physics 34, 205, June 2010, arXiv:1007.0006
- [16] C. Grupen, *Primary Cosmic Rays*, Astroparticle Physics. Undergraduate Texts in Physics. Springer, 2020
- [17] K. M. Krings, *Search for Galactic and Extra-Galactic Neutrino Emission with IceCube*, PhD thesis, Technische Universität München, March 2018
- [18] The IceCube Collaboration, *Blazar Population Analysis*, IC86-1, <https://wiki.icecube.wisc.edu>, 2014
- [19] J. Feintzeig, *Searches for Point-like Sources of Astrophysical Neutrinos with the IceCube Neutrino Observatory*, PhD thesis, University of Wisconsin–Madison, August 2014
- [20] M.G. Aartsen et al, *The IceCube Neutrino Observatory: instrumentation and online systems* 2017 JINST 12 P03012
- [21] The IceCube Collaboration, *IceCube Media Gallery*, <http://icecube.wisc.edu/gallery>
- [22] Jackson, J. D. *Classical Electrodynamics, 3rd Edition*, 1998

- [23] H. Alaeian, *An Introduction to Cherenkov Radiation*, Stanford University, March 2015
- [24] The IceCube Collaboration, *IceCube Data for Neutrino Point-Source Searches Years 2008-2018*, 2021, arXiv:2101.09836
- [25] M. G. Aartsen et al, *Energy reconstruction methods in the IceCube neutrino telescope*, 2014, JINST 9 P03009
- [26] M. G. Aartsen et al, *Search for Sources of Astrophysical Neutrinos Using Seven Years of IceCube Cascade Events* 2019, ApJ 886 12
- [27] M. Tanabashi et al (Particle Data Group), *Review of Particle Physics*, 2018, Physical Review D. 98 (3): 030001.
- [28] M. Wallraff, *Design, Implementation and Test of a New Feature Extractor for the IceCube Neutrino Observatory*, Master thesis, RWTH Aachen, 2010
- [29] R. Abbasi et al, *An improved method for measuring muon energy using the truncated mean of  $dE/dx$* , Nucl. Instrum., 2016, Meth. A703
- [30] J. Braun et al, *Methods for point source analysis in high energy neutrino telescopes*, Astroparticle Physics, Volume 29, Issue 4, May 2008, Pages 299-305
- [31] H. Niederhausen, *Hypothesis Testing*, Applied Multi-Messenger Astronomy course, Technische Universität München, 2020
- [32] J. S. Galpin, *Statistical research design and analysis*, course notes of University of the Witwatersrand, 2010
- [33] A. Zonca et al, *healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python*, Journal of Open Source Software, 2019, Vol. 4, Nr 35
- [34] M. G. Aartsen et al. (IceCube Collaboration), *All-sky Search for Time-integrated Neutrino Emission from Astrophysical Sources with 7 yr of IceCube Data*, The Astrophysical Journal 835, 151, 2017, arXiv:1609.04981

- [35] S. Coenders, *High-energy cosmic ray accelerators: searches with IceCube neutrinos*, PhD Thesis, Technische Universität München, 2016
- [36] The IceCube Collaboration, *Measurement of the Diffuse Astrophysical Muon-Neutrino Spectrum with Ten Years of IceCube Data*, 2019, PoS-ICRC2019-1017, arXiv:1908.09551
- [37] R. Reimann, *Search for the Sources of the Astrophysical High-Energy Muon-Neutrino Flux with the IceCube Neutrino Observatory*, PhD Thesis, RWTH Aachen University, 2019
- [38] The IceCube Collaboration, *Sensitivity of the IceCube Detector to Astrophysical Sources of High Energy Muon Neutrinos*, *Astropart.Phys.*20:507-532, 2004
- [39] The IceCube Collaboration, *Search for steady point-like sources in the astrophysical muon neutrino flux with 8 years of IceCube data*, *Eur. Phys. J. C* (2019) 79: 234
- [40] C. F. Tung et al, *FIRESONG: A python package to simulate populations of extragalactic neutrino sources*, *The Journal of Open Source Software*, 2021, 6(61), 3194