



Master Thesis

# Steps Towards an Improved Time-Integrated Point Source Search for the IceCube Neutrino Observatory

Tomas Kontrimas (03698597)

19.12.2019

Supervisor: Prof. Dr. E. Resconi  
Second Assessor: PD Dr. T. Enßlin

# Contents

<b>Abstract</b> . . . . .	v
<b>1 Introduction to Neutrino Astronomy</b> . . . . .	1
<b>2 The IceCube Neutrino Observatory</b> . . . . .	7
2.1 Neutrino Detection . . . . .	7
2.2 Event Reconstruction . . . . .	12
<b>3 Point Source Search</b> . . . . .	19
3.1 Likelihood Formalism . . . . .	19
3.2 Full-Sky Scan and Trial Factor . . . . .	23
<b>4 Towards an Improved Time-Integrated Point Source Search</b> . . . . .	25
4.1 SkyLLH . . . . .	25
4.2 Kernel Density Estimation Tool . . . . .	30
4.3 Results . . . . .	37
<b>5 Summary and Conclusion</b> . . . . .	43
<b>A List of figures and tables</b> . . . . .	45
<b>Bibliography</b> . . . . .	51
<b>Acknowledgements</b> . . . . .	59





# Abstract

The IceCube Neutrino Observatory searches for neutrino point sources in order to explain the origin of the diffuse high-energy neutrino flux detected in 2013. The maximum likelihood method of the standard point source search analysis is introduced. An improvement to the approach of this point source search analysis, by using new software tools and a kernel density estimation (KDE) for the generation of the likelihood function's probability density functions (PDFs), is presented in this thesis. It summarizes my contribution to the SkyLLH framework and the development of the KDE Tool project. The implemented adaptive KDE method shows a significantly better spatial PDF description, which is especially pronounced towards lower energies ( $\sim 10$  TeV) and soft spectrum sources. The likelihood fit bias for the number of injected signal events is reduced substantially for soft spectrum sources. Furthermore, the next steps for potential further improvements are discussed.



# Chapter 1

## Introduction to Neutrino Astronomy

The physics branch of elementary particle physics started with J. J. Thomson's discovery of the electron in 1897 [1]. It laid the foundations for figuring out the structure of an atom and a particle-based composition of matter. In 1911, Lord Rutherford performed an experiment of  $\alpha$  particles scattering through a thin foil of gold and concluded that the atom consists of a small massive nucleus and light electrons around it. In 1932, J. Chadwick explained that the origin of radiation from beryllium bombarded by  $\alpha$  particles is caused by a neutral particle with a great penetration power – the neutron [2]. Shortly afterwards E. Fermi developed the theory of  $\beta$ -decay, naming the missing third particle *neutrino*. The existence of neutrinos was confirmed experimentally through the studies by Cowan and Reines of the inverse  $\beta$ -decay [3].

At the same time, measurements of ionization in air led to the discovery of cosmic rays (CRs). In 1912, Victor Hess measured the ionization rate dependency on altitude by flying in a balloon up to 5 km [4]. He found that the rate increases with altitude and concluded that ionizing particles are extraterrestrial. Due to the development of particle detectors, e.g. Wilson-cloud chamber, Geiger tube, and photographic emulsion chambers, many new particles were discovered by analyzing trajectories of tracks from CR interactions. Efforts to systematize the newly discovered particle *zoo* led to the standard model of particle physics (figure 1.1).

Cosmic ray particles are ionized nuclei and consist of about 90% protons, 9% alpha particles, and other heavier elements [6]. One can see the cosmic ray spectra in figure 1.2. It shows the measured CR flux up to  $10^{11}$  GeV. In comparison, the Large Hadron Collider (LHC) [7] achieves proton-proton collisions with a centre-of-mass energy of 14 TeV, which is equivalent to  $\sim 10^8$  GeV cosmic ray energy. The energy spectrum can be described by a single power law with two transition regions: the *knee* at  $\sim 3$  PeV and the *ankle* at  $\sim 3$  EeV [6].

In the last century, a spectacular astronomical event – Supernova 1987A (SN 1987A), located in the Large Magellanic Cloud (LMC), 50 kpc away [19], could be observed at all wavelengths and in neutrinos. The first detection of neutrinos, which do not originate from the Sun, marked the birth of neutrino astrophysics [20]. The Kamiokande II detector observed a burst of neutrino events  $\sim 18$ h before the optical

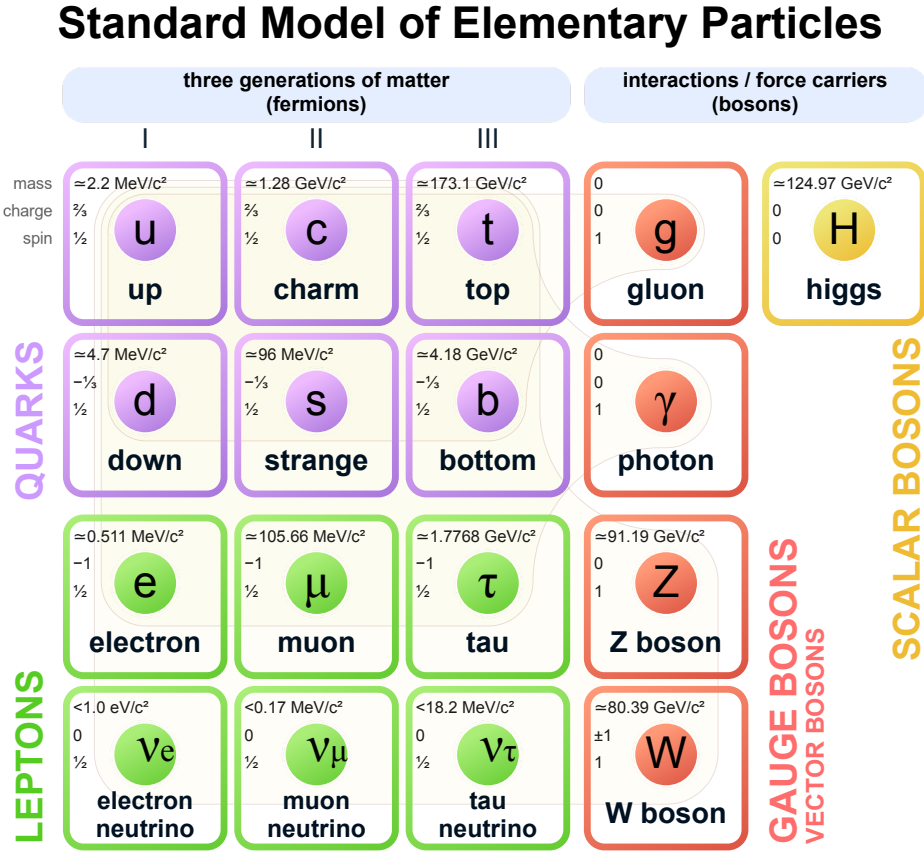


Figure 1.1: Standard model of elementary particles. Figure taken from [5].

sighting [21]. In addition, the IMB and the Baksan detectors also observed a burst of neutrino events around the same time [19].

One class of potential cosmic rays and neutrino sources are Active Galactic Nuclei (AGNs). An AGN is an extremely luminous object at the center of a galaxy. The approximate structure of an AGN consists of a super-massive black hole (SMBH) surrounded by an accretion disk and a dusty torus [22]. The gravitational potential of a SMBH pulls matter towards the center, which in turn loses its angular momentum in turbulent processes and heats up the accretion disk. A fraction of the AGNs have relativistic jets perpendicular to the accretion disk. Due to the asymmetrical structure, the observed spectrum depends on the relative position of an observer. AGNs with the jet pointing close to the Earth form the subclass called blazars.

Because the cosmic rays are charged, they get deflected in interstellar magnetic

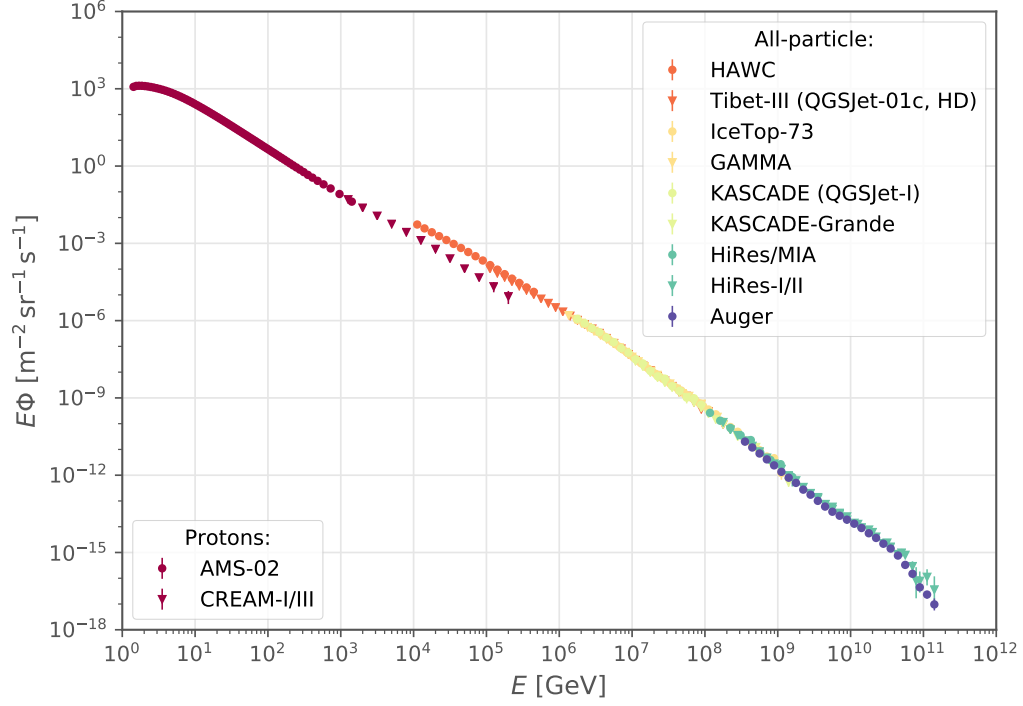


Figure 1.2: Cosmic ray proton and all-particle spectra. Measured data is taken from [8–17]. The plot was generated using [18].

fields and hence their production and acceleration site location can not be reconstructed. One could use neutral particles (e.g. photons and neutrinos), created by CR interactions with matter near the site location, in order to reconstruct the site location, as they travel in a straight line. The possible types of high energy CR production sites are summarized in the Hillas plot (figure 1.3).

Using large neutrino detectors, like the IceCube Neutrino Observatory (section 2), one can reconstruct the direction of the neutrino source and infer properties of the production mechanism. Because neutrinos travel unobscured, they can also act as an alert for an interesting astronomical phenomena for other observatories.

High energy cosmic rays can produce neutrinos and photons in proton-proton and

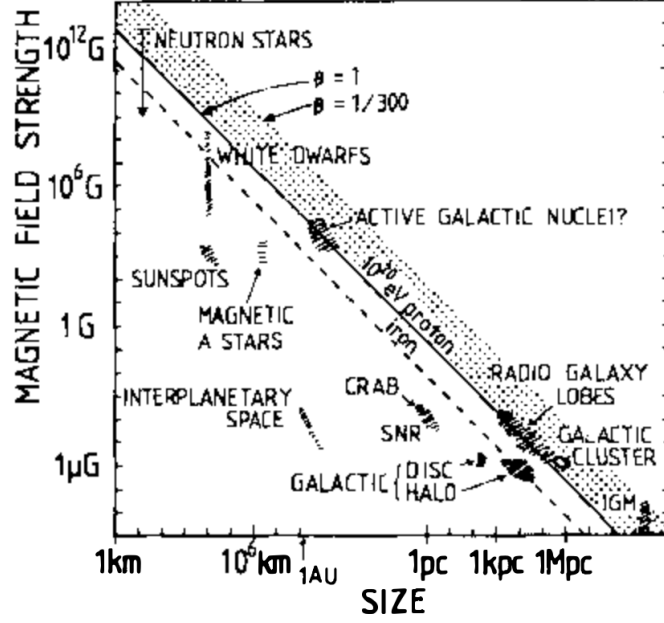


Figure 1.3: The Hillas plot shows the size and the magnetic field strength of possible acceleration sites. Objects above the diagonal line could accelerate protons up to  $10^{20}$  eV. Figure taken from [23].

proton-photon collisions [24]. These mentioned collisions result in pion  $\pi$  generation, that decay into electrons  $e$ , photons  $\gamma$ , and neutrinos  $\nu$ :

$$\begin{aligned}
 \pi^+ &\rightarrow \mu^+ \nu_\mu \rightarrow e^+ \nu_\mu \bar{\nu}_\mu \nu_e \\
 \pi^- &\rightarrow \mu^- \bar{\nu}_\mu \rightarrow e^- \bar{\nu}_\mu \nu_\mu \bar{\nu}_e \\
 \pi^0 &\rightarrow \gamma\gamma,
 \end{aligned} \tag{1.1}$$

where the bar symbol over a neutrino denotes an antineutrino. Assuming an equal distribution of pions, the ratio of produced neutrino flavors close to the source is  $(\nu_e : \nu_\mu : \nu_\tau) = (\bar{\nu}_e : \bar{\nu}_\mu : \bar{\nu}_\tau) = (1 : 2 : 0)$ . Due to the long travel path and non-vanishing neutrino mass eigenstates, neutrinos oscillate between flavors and arrive at the Earth with  $(\nu_e : \nu_\mu : \nu_\tau) = (\bar{\nu}_e : \bar{\nu}_\mu : \bar{\nu}_\tau) = (1 : 1 : 1)$  flavor ratio [25].

Multi-messenger astronomy strives to combine received information from photons, cosmic rays, neutrinos, and gravitational waves [26]. It could lead to a better understanding of their production sites. The messengers (excluding gravitational waves) and their general propagation routes to the Earth are shown in figure 1.4.

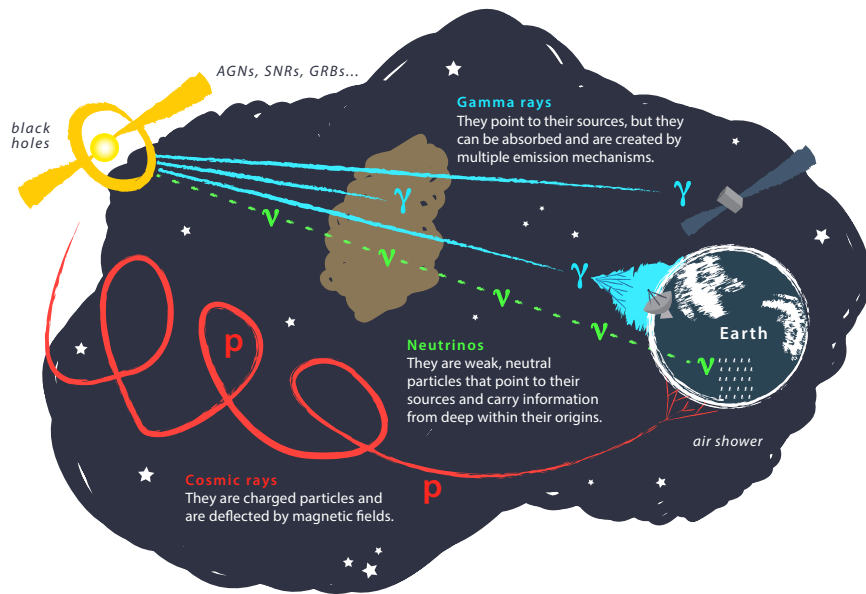


Figure 1.4: Propagation and detection scheme of different messengers from the source.  
Figure taken from [27]





## Chapter 2

# The IceCube Neutrino Observatory

The IceCube Neutrino Observatory is located at the geographic South Pole near the Amundsen–Scott South Pole Station. It is an in-ice array of 5160 digital optical modules (DOMs), that instruments a  $1\text{ km}^3$  volume of Antarctic glacial ice. The detector construction was completed on December 18, 2010. The structure of the IceCube detector is shown in figure 2.1. The in-ice array consists of 86 vertical strings, each string containing 60 DOMs, deployed between 1450 m and 2450 m below the surface. 81 stations at the surface of the ice form the IceTop array, which detects cosmic ray initiated air showers and provides a veto region for the downward-going neutrino detection. Each IceTop station contains 2 ice-filled tanks with 2 DOMs per tank.

The observation of neutrinos relies on the fact that when neutrinos interact with ice they create charged particles, moving faster than the phase velocity of light in ice and emitting Cherenkov radiation [29]. Resulting Cherenkov photons are detected in the in-ice array of DOMs. Such a huge detector volume is needed due to the very low neutrino interaction cross sections. The reconstruction of the incident neutrino direction and energy uses the recorded photon data in the whole array with the timing information and the model of light propagation in ice.

### 2.1 Neutrino Detection

Three generations of neutrinos are present in the standard model of particle physics (figure 1.1). They interact with matter (e.g. ice) via the weak force in two interaction channels, called “charged current” (CC) and “neutral current” (NC). An example of the CC interaction channel can be seen in figure 2.2, where  $l$  is a lepton (electron, muon, or tau) and  $\nu_l$  is a respective flavor neutrino. The charged vector boson  $W^\pm$  is the CC interaction mediator. A similar weak force interaction can also happen with quarks ( $q$ ). The NC interaction channel is mediated via the  $Z^0$  vector boson (figure 2.3). It does not change the flavor of the interacting particle.

IceCube detects neutrinos with an incident neutrino energy above  $\sim 100\text{ GeV}$  [30].

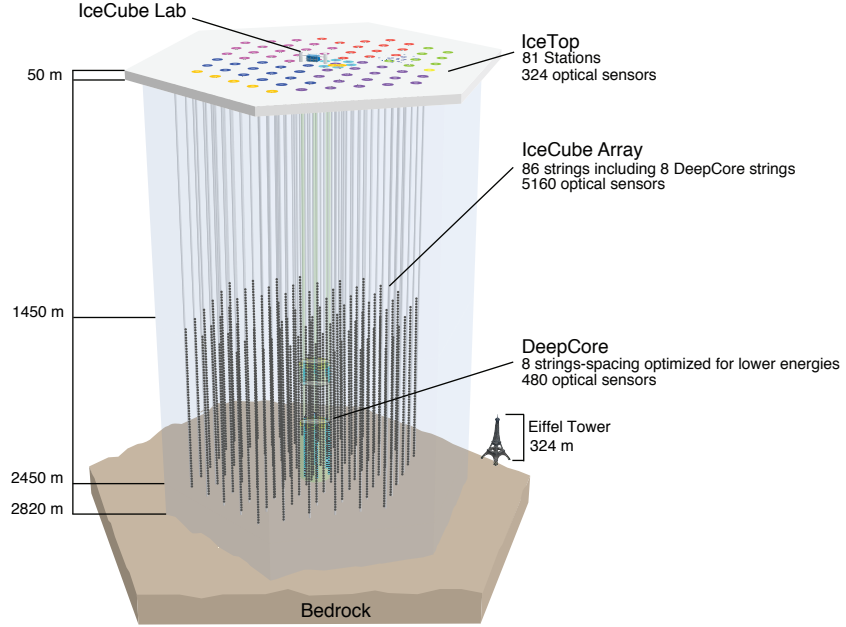


Figure 2.1: The hexagonal structure of the IceCube detector. Deployment seasons of the strings are indicated by different colors at the surface of the ice. The strings form a triangular grid with a horizontal spacing of about 125 m. Figure taken from [28].

The dominant scattering process in this energy regime is deep-inelastic scattering (DIS) [31]. Neutrino interactions with the nucleus of the ice can be summed up to

$$\begin{aligned}
 \nu_l + N &\rightarrow l^- + X & (\text{CC}) \\
 \bar{\nu}_l + N &\rightarrow l^+ + X & (\text{CC}) \\
 \nu_l + N &\rightarrow \nu'_l + X & (\text{NC}) \\
 \bar{\nu}_l + N &\rightarrow \bar{\nu}'_l + X & (\text{NC}),
 \end{aligned} \tag{2.1}$$

where  $N$  is a nucleus and  $X$  is the remnant of  $N$ . The remnant  $X$  can produce a hadronic cascade, while the charged lepton can propagate through the detector material, emitting Cherenkov radiation. The neutrino interaction cross section with

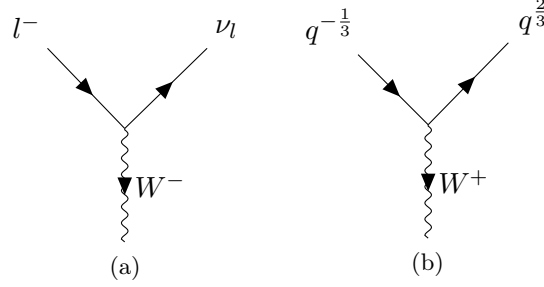


Figure 2.2: An example of “charged current” interaction vertices.

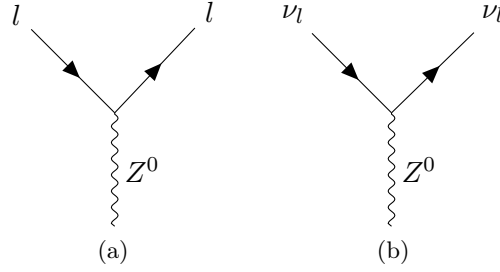


Figure 2.3: An example of “neutral current” interaction vertices.

electrons in the target material is usually very small and can be neglected. It only has to be taken into account in the Glashow resonance energy range. The Glashow resonance [32] describes a steep increase of the electron antineutrino cross section with the electron as a target (equation 2.2) at around 6.3 PeV neutrino energy.

$$\bar{\nu}_e + e^- \rightarrow W^- \quad (2.2)$$

Cross sections of the before mentioned neutrino interactions are shown in figure 2.4.

The most interesting neutrino interaction for the point source search is the CC interaction, producing a propagating muon, as it leaves a track like signature inside the detector. Using the track as a long lever arm, the reconstruction algorithm can reconstruct the incident neutrino direction with an angular resolution of  $1^\circ$  and better [33]. This allows to pinpoint the possible astronomical object, which produced the reconstructed neutrino. The produced electron leaves a cascade like signature, which direction can be reconstructed with an  $\mathcal{O}(10^\circ)$  median angular resolution [34]. The produced tau leaves a double bang like signature, with the angular resolution lying between the cascade and the track resolution, depending on the separation length between the two cascades. The first cascade forms at the tau production vertex. It

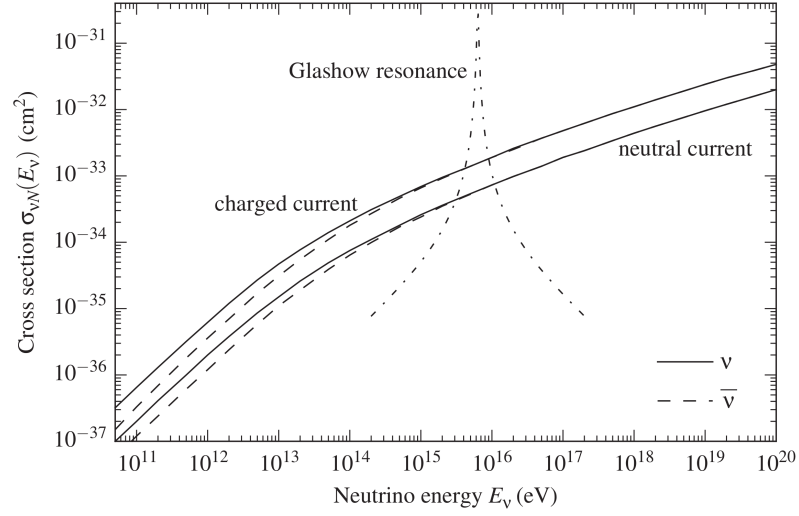


Figure 2.4: Neutrino- and antineutrino-nucleon cross sections for CC and NC interactions. In addition, the Glashow resonance of the electron antineutrino-electron cross section is shown. Figure taken from [6].

is followed by a track and the second cascade at the tau decay vertex. Signatures of propagating leptons are shown in figure 2.5.

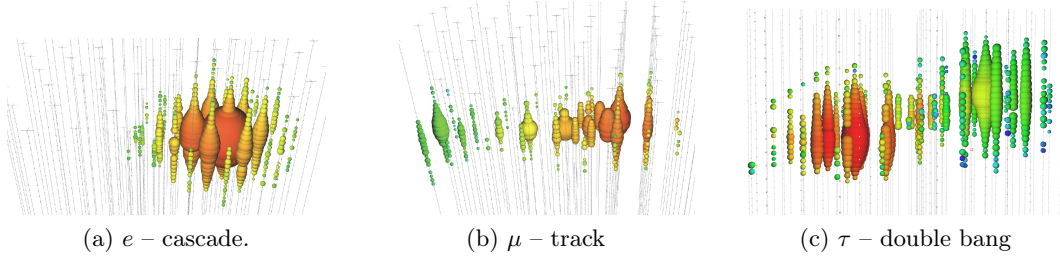


Figure 2.5: Signatures of three different propagating leptons. The color scale ranges from early hits (red) to late hits (blue). The bubble size is proportional to the logarithm of the deposited energy inside the DOM. Figures taken from [35].

The Cherenkov radiation emission angle,  $\theta_C$ , from the moving charged particle is

$$\cos(\theta_C) = \frac{1}{n\beta}, \quad (2.3)$$

where  $n$  is the refractive index of the medium and  $\beta$  is the ratio between the particle's velocity and the speed of light. In the IceCube detector  $\theta_C$  is about  $41^\circ$  [36]. The

number of Cherenkov photons emitted by a single charged particle is given by Frank-Tamm formula [37]:

$$\frac{dN}{dx d\lambda} = \frac{2\pi\alpha}{\lambda^2} \sin^2(\theta_C), \quad (2.4)$$

where  $\alpha$  is the fine-structure constant,  $d\lambda$  is a wavelength range, and  $dx$  is a path length. The radiation appears blue, because its intensity is inversely proportional to the square of the wavelength.

### 2.1.1 Digital Optical Module

The Cherenkov radiation photons in ice are collected using digital optical modules (DOMs). The main part of a DOM is the photo-multiplier tube (PMT), which can detect even single photons. DOMs use the 10'' diameter Hamamatsu R7081-02 PMT, facing downwards in the bottom glass hemisphere. It is secured with a high-strength silicone gel, which provides a mechanical support to the whole internal structure of the DOM components and a good optical coupling between the PMT and the glass sphere. The PMT is specified for the wavelength range from 300 nm to 650 nm, with a peak quantum efficiency around 25% near 390 nm [38]. A schematic view of a DOM is shown in figure 2.6.

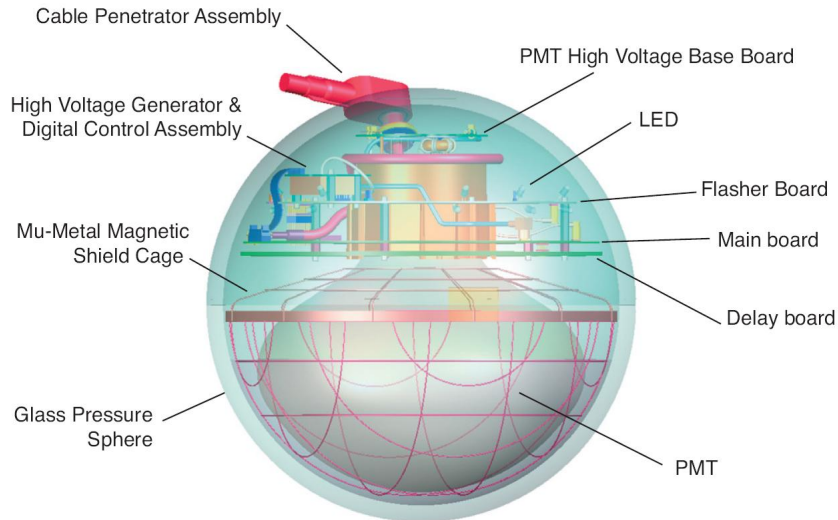


Figure 2.6: A schematic view of the IceCube digital optical module. Figure taken from [39].

The main board is located around the PMT neck. It embeds a data-acquisition computer, controlling all devices inside the DOM and communicating with the data

acquisition system on the surface. An important main board function is the digitization of recorded PMT waveforms. It uses two types of custom integrated circuits: an analog transient waveform digitizer (ATWD) [40] and a fast analog-to-digital converter (fADC). They cover the recording durations of 427 ns and 6.4  $\mu$ s, respectively [38]. Two sets of ATWDs are used to reduce deadtime.

The information level sent out to the surface depends on the local coincidence condition. In the case when only a single DOM exceeds 0.25 photoelectron (PE) threshold, a time stamp and a brief charge summary are sent to the surface. If a local coincidence event (when a nearest or next-to-nearest neighbor DOM also exceeds the threshold within  $\pm 1 \mu$ s) is registered, the full digitized waveform is sent out.

### 2.1.2 Ice Properties

In order to accurately reconstruct the incident neutrino direction and energy from the registered photons, one needs to know the optical ice properties, e.g. effective scattering and absorption lengths. The in situ measurements of light scattering and absorption started since the AMANDA [41] experiment, which precedes IceCube. They found that all air bubbles in ice have transformed into the solid phase at around 1500 m below the surface [42] and no longer influence the optical ice properties. The optical ice properties below 1500 m depend almost only on the dust concentration in ice.

The six-parameter ice model was introduced in [36]. It describes the global fit of the likelihood function between simulated and recorded events. As each DOM contains an LED flasher board, a brief flash of photons can be produced. The resulting light distribution can be recorded with other DOMs. The thickness of the ice layer parametrization was arbitrary chosen to be 10 m.

The currently used ice model is the South Pole Ice (SPICE) 3.2.1 model [43] (the optical ice properties are shown in figure 2.7). It adds the ice anisotropy [44], an LED beam as a 2D Gaussian approximation, and a relative DOM efficiency as additional fit parameters of the likelihood function. To measure photon distributions, LEDs of all DOMs on 85 strings were flashed and recorded on a single string [45]. The simulation of a direct photon propagation in ice was performed using the photon propagation code (PPC) [46]. It was used to fit the ice model parameters against the measured data.

## 2.2 Event Reconstruction

The event can be reconstructed after the information is received from all DOMs. Using a *feature extraction* pulses, containing the time and charge information, are

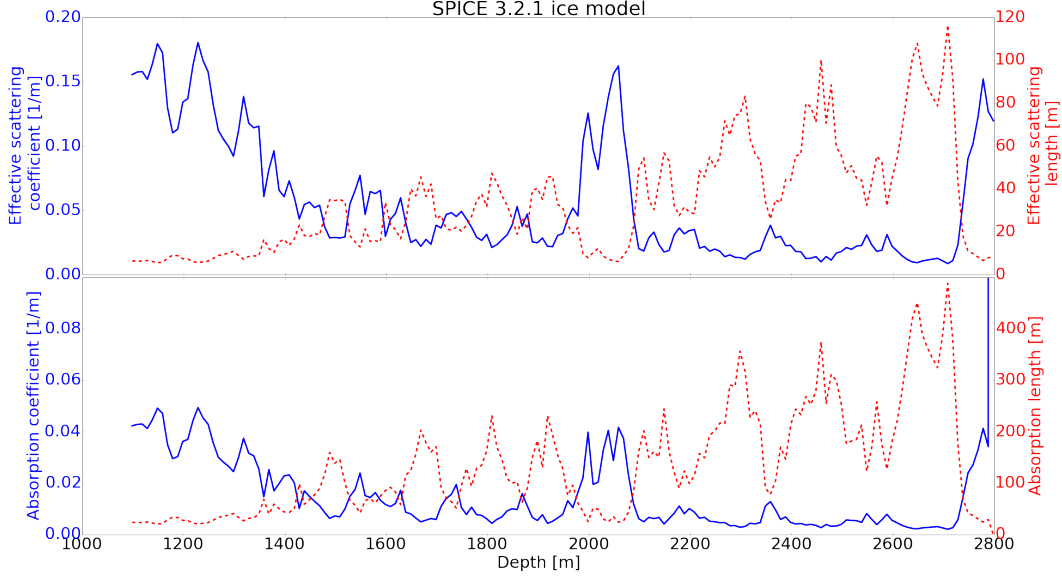


Figure 2.7: The SPICE 3.2.1 model [43]. Dashed red lines show the effective scattering (top) and absorption (bottom) lengths dependency on depth. One can see a dust layer impact between 2000 m and 2100 m. Figure taken from [47].

extracted from the registered waveforms. A short summary of the event direction and energy reconstruction algorithms is presented in the following sections.

### 2.2.1 Direction and Angular Error

The simplest and fastest algorithm of the event direction reconstruction is *Linefit* [48]. It ignores the geometry of the Cherenkov cone and the optical ice properties. The position vector of the muon at time  $t$  is written as

$$\mathbf{r} = \mathbf{r}_0 + \mathbf{v}t, \quad (2.5)$$

where  $\mathbf{v}$  is the muon velocity vector and  $\mathbf{r}_0$  is the initial vertex position at time  $t_0$ . A  $\chi^2$  function is defined as

$$\chi^2 = \sum_{i=1}^N (\mathbf{r}_i - \mathbf{r}_0 - \mathbf{v}(t_i - t_0))^2, \quad (2.6)$$

where  $\mathbf{r}_i$  and  $t_i$  are the position and the pulse time of the  $i$ th DOM, respectively, and  $N$  is the total number of pulses. The minimization of  $\chi^2$  has an analytic solution, yielding a vertex position and a muon velocity fit. It is usually used to seed more sophisticated reconstruction methods [49].

Methods like the single photo electron (SPE) fit and the multi photo electron (MPE) fit [50] take into account the Cherenkov angle and optical ice properties. They assume a single infinitely long muon track and introduce a relative arrival time (time residual) variable as

$$t_{\text{res}} \equiv t_{\text{hit}} - t_{\text{geo}}, \quad (2.7)$$

where the geometrical photon arrival time is

$$t_{\text{geo}} = t_0 + \frac{\hat{\mathbf{p}} \cdot (\mathbf{r}_i - \mathbf{r}_0) + d \tan \theta_C}{c} \quad (2.8)$$

Here  $\hat{\mathbf{p}}$  is a muon direction and  $d$  is a perpendicular distance between the track and a DOM position. The  $t_{\text{geo}}$  assumes a Cherenkov photon that travels undelayed and unscattered. Ideally, the  $t_{\text{res}}$  distribution would be a delta function, but due to the PMT jitter, noise, stochastic cascade production, and scattering it broadens and shifts.

Given the track parameters as  $\mathbf{a} = (\mathbf{r}_0, t_0, \hat{\mathbf{p}}, E_0)$ , the likelihood function of experimentally measured independent values  $x_i$  can be expressed as

$$\mathcal{L}(\mathbf{a}|\mathbf{x}) = \prod_i p(x_i|\mathbf{a}) \quad (2.9)$$

The SPE likelihood function, where each photon has an equal contribution, is

$$\mathcal{L}_{\text{SPE}}(\mathbf{a}|t_{\text{res}}) = \prod_i^N p_{\text{SPE}}(t_{\text{res},i}|\mathbf{a}) \quad (2.10)$$

The probability distribution function  $p_{\text{SPE}}$  can be obtained from the simulation of photon propagation in ice. The minimized negative SPE likelihood returns a track parameters fit.

Because of the scattering induced time delay, the first arrived photon to the DOM is more likely less scattered than the following photons and carries more accurate directional information. In order to account for this effect, the MPE method introduces a modification to the SPE time probability distribution. It only uses the first photon  $t_{\text{res}}$  information and the number of registered hits in the DOM. The time distribution of the first of  $N$  arriving photons is given by

$$p_{\text{MPE}}(t_{\text{res}}|\mathbf{a}) = N p_{\text{SPE}}(t_{\text{res}}|\mathbf{a}) \left( \int_{t_{\text{res}}}^{\infty} dt p_{\text{SPE}}(t|\mathbf{a}) \right)^{(N-1)} \quad (2.11)$$



The MPE likelihood function is

$$\mathcal{L}_{\text{MPE}}(\mathbf{a}|t_{\text{res}}) = \prod_i^{N_{\text{DOM}}} p_{\text{MPE}}(t_{\text{res},i}|\mathbf{a}) \quad (2.12)$$

The updated MPE fit version is called *SplineMPE*. It uses multi-dimensional penalized splines to model the time probability distribution function [51]. The splines are fit to the tabulated data of photon propagation simulation. The account of stochastic muon energy losses at energies above 1 TeV is implemented in the *SegmentedSplineMPE* [52] reconstruction algorithm

A basic angular error estimation in IceCube is estimated using *paraboloid fit* [53]. It transforms the coordinate system of the track to the one shown in figure 2.8. The relative coordinates  $\theta$  and  $\varphi$ , as the approximate coordinates in the tangent plane, are introduced around the best fit track direction.

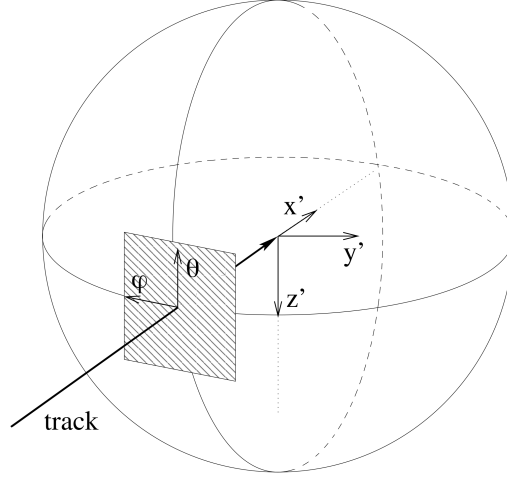


Figure 2.8: Paraboloid fit coordinate system. Figure taken from [53].

The likelihood space near the minimum of the direction fit is usually shaped as a normal distribution, therefore, the log likelihood space is shaped as a paraboloid. The paraboloid fit scans the likelihood space around the best fit position on a defined grid, given by  $\theta$  and  $\varphi$ , maximizing the log likelihood with the vertex position  $\mathbf{r}$  as a parameter. An ellipse that describes a  $1\sigma$  error region is fit to the resulting likelihood space. The ellipse axis  $\sigma_1$  and  $\sigma_2$  can be combined to the average circular error

$$\sigma_{\text{paraboloid}} = \sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}} \quad (2.13)$$

The  $\sigma_{\text{paraboloid}}$  gives the angular error estimation on the muon direction. At the low energy range the kinematic angle introduces an additional angular error smearing,

while at the high energies the angular error is underestimated in comparison to the Monte Carlo (MC) simulation. In order to fix these effects the paraboloid sigma is calibrated to match the Monte Carlo simulation data. The *pull correction* factor is introduced as a function of energy, so that the median of  $\frac{\Delta\Psi}{\sigma_{\text{paraboloid}}^{\text{corrected}}}$  distribution lies at a 1.1774 value, where  $\Delta\Psi$  is a great circle distance between the true and the reconstructed neutrino directions in MC simulation. The 1.1774 value comes from the fact that the 50% containment of a 2D normal distribution is at  $r = 1.1774\sigma$ .

### 2.2.2 Energy

The lower limit of the neutrino energy can be estimated from the collected Cherenkov light [54]. The produced light from electromagnetic showers has a low variance with respect to the deposited energy [55]. If the neutrino charged current interaction vertex and the lepton track are contained inside the detector, the energy reconstruction accuracy is only limited by the detector resolution. In the case of a neutral current interaction, the resulting neutrino carries a fraction of the initial energy outside the detector, which can not be detected and can only be estimated statistically. Both electromagnetic and hadronic showers have similar light deposition profiles, except that hadronic showers have a suppressed light yield due to the presence of more neutral particles.

The vertex of registered high energy muon neutrino interaction is often outside the detector due to the long muon track length. The average muon range in ice was estimated to be  $x_f = \log(1 + E_i \frac{b}{a})/b$ , with  $a = 0.268 \frac{\text{GeV}}{\text{mwe}}$  and  $b = 0.470 \frac{10^{-3}}{\text{mwe}}$  fit for stochastic energy losses and  $E_i$  as an initial muon energy [56]. For example, the average range of 300 GeV (10 TeV) energy muon is  $\sim 1$  km ( $\sim 6$  km), respectively.

The muon energy can be determined from the track signature by observing the specific energy loss  $dE/dx$  [57]. It is modeled as

$$\frac{dE_\mu}{dx} = A + BE_\mu, \quad (2.14)$$

where  $A$  is the energy loss due to the ionization and  $B$  is due to the stochastic energy loss. The conventional muon energy loss estimation method sums up the detected charge and then compares it to the simulated expected charge for a fixed energy loss with the same track trajectory. As the photoelectron yield is expected to be directly proportional to the energy loss rate (above the muon energy of  $\sim 1$  TeV) [57], the  $dE_\mu/dx$  can be estimated and used to infer the muon energy  $E_\mu$ .

An improvement to the conventional muon energy loss estimation can be made by using the truncated mean  $dE_\mu/dx$  method [57]. It splits the muon track and the detector into bins, whose edges are defined by perpendicular planes to the track. The energy loss rate is calculated for each bin. The truncation discards some part of the highest energy loss rate bins in order to reduce statistical fluctuations due

to stochastic energy losses. Then the new energy loss rate is calculated using the information from the remaining bins, following the conventional method algorithm.



## Chapter 3

# Point Source Search

The IceCube Neutrino Observatory discovered the existence of a diffuse flux of high-energy astrophysical neutrinos in 2013 [58]. Since then, the search for a neutrino point source has been an ongoing effort. The most significant result comes from the blazar TXS 0506+056 with  $3.5\sigma$  evidence for neutrino emission [59]. The currently used likelihood formalism and point source search methods are presented in the following sections.

### 3.1 Likelihood Formalism

The likelihood function is used to infer model parameters from the observed data. It can also be used in a hypothesis testing. The current point source search analyses use the unbinned maximum likelihood ratio test method [60].

#### 3.1.1 Hypothesis Testing

The first step of the unbinned maximum likelihood ratio test is the definition of the hypothesis. In general, the hypothesis is defined by making a statement about a model and its parameters. The hypothesis testing requires a definition of the two complementary hypotheses. They are called the null hypothesis,  $H_0$ , and the alternative hypothesis,  $H_1$  [61]. The general notation of the null and alternative hypotheses is  $H_0 : \theta \in \Theta_0$  and  $H_1 : \theta \in \Theta_0^c$ , where  $\Theta_0$  is some subset of the parameter space,  $\Theta_0^c$  is its complement, and  $\theta$  is a population parameter.

In order to be able to make a decision between the two hypotheses, a hypothesis test has to be defined. It specifies for which sample values accept  $H_0$  as true and for which sample values reject  $H_0$  and accept  $H_1$  as true [61]. This divides the parameter space into the acceptance and rejection regions. The hypothesis test is specified in terms of a test statistic, expressed as

$$\lambda(X_1, \dots, X_n) : \mathbf{X} \rightarrow \mathbb{R}, \quad (3.1)$$

where  $\lambda$  is a function that maps an outcome into a scalar number. A very general and widely applicable hypothesis test is the likelihood ratio test (LRT).

### 3.1.2 Likelihood Ratio Test

The data in the point source search dataset is composed of independent events from the signal and background component classes [60]. Therefore, a two-component likelihood function can be defined as

$$\mathcal{L}(n_s, \mathbf{p}_s | D) = \prod_{i=1}^N \left[ \frac{n_s}{N} S_i(\mathbf{p}_s) + \left(1 - \frac{n_s}{N}\right) B_i \right], \quad (3.2)$$

where  $n_s$  is the mean number of signal events in the data sample  $D$  of  $N$  total events and  $\mathbf{p}_s$  is a set of source model parameters. The source model parameters include the source position  $\mathbf{x}_s = (\alpha_s, \delta_s)$  and the spectral index  $\gamma$  of the source flux. The  $S_i(\mathbf{p}_s)$  and  $B_i$  are the values of the  $i$ th data event of the signal and background probability density functions (PDFs), respectively.

The time-integrated signal PDF  $S_i$  can be expressed as a joint probability of the spatial and energy PDFs

$$S_i(\mathbf{p}_s) \equiv S_i(\mathbf{x}_i, E_i | \mathbf{p}_s) = \mathcal{S}_S(\mathbf{x}_i | \mathbf{p}_s) \mathcal{E}_S(E_i | \mathbf{x}_i, \mathbf{p}_s) \quad (3.3)$$

By assuming that the detector introduces a Gaussian point spread function, which is convoluted with the source position, the signal spatial PDF can be express as

$$\mathcal{S}_S(\mathbf{x}_i | \mathbf{p}_s) \equiv \mathcal{S}_S(\Delta\psi_i, \sigma_i | \mathbf{p}_s) = \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{\Delta\psi_i^2}{2\sigma_i^2}\right), \quad (3.4)$$

where  $\Delta\psi_i$  is the spatial angle between the  $\mathbf{x}_s$  and  $\mathbf{x}_i$ , source and event positions, respectively, and  $\sigma_i$  is the event angular resolution uncertainty.

Because the IceCube detector is located at the South Pole, the azimuth (right ascension) dependency in the signal energy PDF, caused by the detector's geometry, averages out over a day, leaving that the energy PDF depends only on the zenith (declination) of the event. Therefore, separate energy PDFs for a predefined list of zenith bands are created and a suitable one is picked for the evaluation. This leads to the following energy PDF form

$$\mathcal{E}_S(E_i | \mathbf{x}_i, \mathbf{p}_s) = \mathcal{E}_S(E_i | \delta_i, \mathbf{p}_s) \quad (3.5)$$

The background PDF expression follows the same derivation as the signal PDF, leading to

$$B_i \equiv \mathcal{S}_B(\mathbf{x}_i) \mathcal{E}_B(E_i | \delta_i) \quad (3.6)$$

The spatial and energy components of the background PDF can be derived either from the experimental or Monte Carlo simulation data.

A general likelihood ratio test form for testing the null hypothesis against the alternative hypothesis is

$$\lambda(x) = \frac{\sup_{\theta \in \Theta_0} \mathcal{L}(\theta|x)}{\sup_{\theta \in \Theta} \mathcal{L}(\theta|x)}, \quad (3.7)$$

where  $\Theta$  is the full parameter space and  $\mathcal{L}$  is a likelihood function. In the time-integrated point source search analysis the null hypothesis,  $H_0$ , is usually defined as a hypothesis that the data consists solely of the background neutrino events. The alternative hypothesis,  $H_1$ , is defined as a hypothesis that the data consists of the background neutrino events and the astrophysical neutrino events, produced by a point source [60]. The likelihood ratio test statistic, using the constructed likelihood function (equation 3.2) and zero signal events for the null hypothesis, is defined as

$$TS = -2 \log(\Lambda) = -2 \log \left[ \frac{\mathcal{L}(n_s = 0)}{\mathcal{L}(\hat{n}_s, \hat{\mathbf{p}}_s)} \right], \quad (3.8)$$

where  $\hat{n}_s$  and  $\hat{\mathbf{p}}_s$  are the values which maximize the likelihood function.

A hypothesis testing can produce two types of errors. They are called ‘‘Type I Error’’ and ‘‘Type II Error’’. The possible outcomes are summarized in table 3.1.

Table 3.1: The two types of errors in a hypothesis testing.

		Decision	
		accept $H_0$	reject $H_0$
Truth	$H_0$	Correct decision	Type I Error
	$H_1$	Type II Error	Correct decision

We want to minimize the Type 1 Error to avoid a false  $H_0$  hypothesis rejection, in other words, making a decision that there is an astrophysical neutrino clustering, while in reality there is only a fluctuation of the background neutrino events. It is done by specifying a level of the test –  $\alpha$ . A level  $\alpha$  likelihood ratio test is constructed by choosing  $k$  such that the  $\sup_{\theta \in \Theta_0} P_\theta(TS(\mathbf{X}) \geq k) \leq \alpha$  inequality holds [61].

The analysis limits can be estimated without using the measured data. It is done by simulating an experiment with a mean number of injected signal events  $\mu$  in the pseudo dataset. The sensitivity is commonly defined as the flux required to get a test statistic value  $TS_\mu$  that is higher than the median of the null hypothesis test statistic distribution in 90% of performed experiments. The discovery potential is defined as the flux required to get a test statistic value  $TS_\mu$  that is higher than the  $1 - 5\sigma$  quantile of the null hypothesis test statistic distribution in 50% of performed experiments [62]. A graphical example of the sensitivity and discovery potential test

statistics distributions definition is shown in figure 3.1. The returned sensitivity or discovery potential  $\mu$  value relation to the flux is defined in the following section.

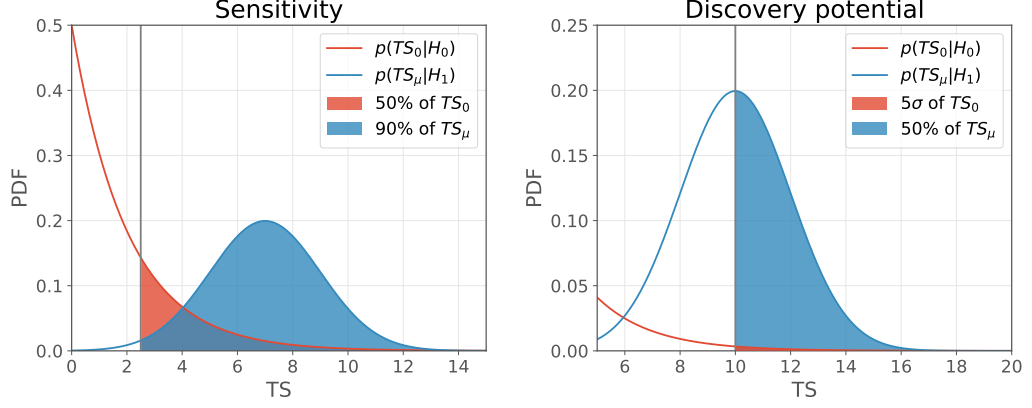


Figure 3.1: Illustration for the definition of sensitivity and discovery potential.

### 3.1.3 Source Flux

A generic differential factorized flux of an astrophysical neutrino source, as used in traditional IceCube point source searches, can be expressed as

$$\frac{d\Psi(\alpha, \delta, E, t|\mathbf{p}_s)}{dE d\Omega dA dt} = \Phi_0 \Psi_S(\alpha, \delta|\mathbf{p}_s) \mathcal{E}_S(E|\mathbf{p}_s) \mathcal{T}_S(t|\mathbf{p}_s), \quad (3.9)$$

where  $A$  is an area,  $\Omega$  is a solid angle,  $\Phi_0$  is the flux normalization, and  $\Psi_S$ ,  $\mathcal{E}$ , and  $\mathcal{T}_S$  are the spatial, energy, and time profiles respectively. The spatial profile of a point like source is defined as a Dirac  $\delta$  function

$$\Psi_S(\mathbf{r}|\mathbf{p}_s) = \delta(\mathbf{r} - \mathbf{r}_s), \quad (3.10)$$

where  $\mathbf{r} = (\alpha, \delta)$  is the celestial coordinate vector. The energy profile is usually expressed by a power law function

$$\mathcal{E}_S(E|\mathbf{p}_s) = \left( \frac{E}{E_0} \right)^{-\gamma}, \quad (3.11)$$

where  $\gamma$  is a spectral index and  $E_0$  is a reference energy (usually 100 TeV). For a time-integrated search the time profile  $\mathcal{T}_S(t|\mathbf{p}_s)$  is a box shaped function spanning the entire dataset livetime – unity.



The flux relation to the mean number of signal events is

$$\mu = T \int d\Omega \int dE A_{\text{eff}}(E, \Omega) \Psi(E, \Omega), \quad (3.12)$$

where  $T$  is the dataset livetime and  $A_{\text{eff}}$  is the detector effective area. The effective area depends on the neutrino energy and the observation angle.

### 3.1.4 p-Value

Instead of setting the confidence level  $\alpha$  and only reporting whether to accept or reject the null hypothesis after the hypothesis testing is done, the another value of the test statistic can be introduced. It is called *p-value*. The p-value allows to have an independent measure of how strongly the null hypothesis is accepted or rejected. A small p-value indicates a strong evidence against the null hypothesis. A graphical example of the p-value definition is shown in figure 3.2.

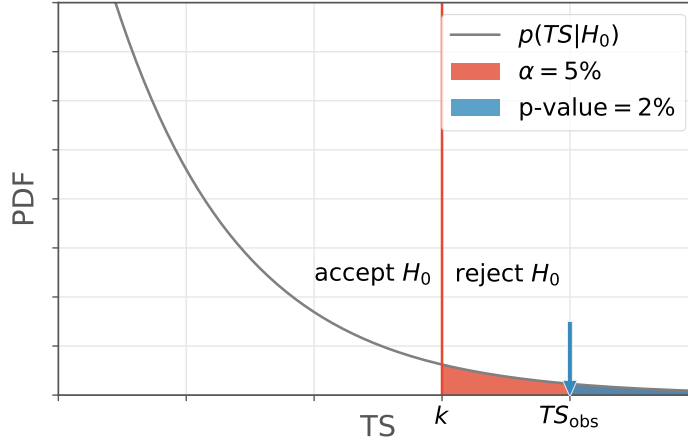


Figure 3.2: An  $\alpha$  level test and a p-value example. Given the  $\alpha = 5\%$ , one can see that the  $TS_{\text{obs}}$  rejects the null hypothesis. The p-value of observed test statistic is  $p = 2\%$ , which also indicates a strong evidence against the null hypothesis.

## 3.2 Full-Sky Scan and Trial Factor

A full-sky scan is an extension of the point source maximum likelihood ratio test. It is performed by scanning the whole sky on a fine grid and maximizing the log likelihood function (equation 3.8) with  $n_s$  and  $\gamma$  as free parameters at each point on a grid in

order to get the maximum test statistic distribution map. Because of the completely different muon background from the northern and southern hemispheres, the sky is divided into two parts at the  $\delta = -5^\circ$  declination boundary [63] and each hemisphere is scanned separately. In order to calculate a local pre-trial p-value, the background TS distribution is generated from the data itself by randomly scrambling its right ascension values for many trials. The scrambling removes any possible clustering of the signal events. Finally, the two hotspots in each hemisphere with the lowest p-value are chosen.

A trial factor has to be introduced because of a *look elsewhere* effect. There can be a clustering of events in the sky from random fluctuations, even for the pure background data. The grid scan can be seen as the number  $N$  of independent observations, where  $N$  is called a *trial factor*. It is used to compensate for the possible random clustering. The trial corrected probability to observe no p-value lower than  $p_{\min} \equiv x$  in  $N$  independent observations is

$$P\left(\min_i\{p_i\} < x\right) = 1 - (1 - x)^N \quad (3.13)$$

The usual trial factor  $N$  for the full-sky scan is in the order of  $10^5$ .

To get a lower post-trial p-value a scan can be designed to look only at the interesting positions in the sky, lowering the number of observation and making the trial factor and its penalty smaller. For example, this is done in the source catalog searches by scanning the locations of already registered possible neutrino sources in  $\gamma$ -rays [63]. One of the possible catalog choices is the Fermi-LAT 4FGL catalog [64] including 5065 sources above  $4\sigma$  significance in  $\gamma$ -rays.

## Chapter 4

# Towards an Improved Time-Integrated Point Source Search

The most recent time-integrated point source search analysis was done by performing a full-sky scan, a search within a selected source catalog, a catalog population study, and three stacked Galactic catalog searches in 10 years of IceCube data [63]. No discovery of a point source was made. The IceCube detector sensitivity is limited by the collected data, which is proportional to the data collection time and effective area of the detector, and the analysis method capabilities. Each new year of the IceCube detector livetime makes up a smaller and smaller part of the combined dataset. Because of that, we decided to see what could be gained by improving already existing point source search analysis methods. A new software and mathematical formalism, based on a kernel density estimation (KDE) approach, were developed.

### 4.1 SkyLLH

SkyLLH [65] is a successor of the SkyLab [66] project. SkyLab is a Python based framework, designed to perform unbinned likelihood maximization. Due to the lack of development guidelines, the framework became difficult to maintain and to implement new likelihood analyses. To solve these issues a new framework was written from scratch.

The SkyLLH framework is also based on Python. Within this framework the user can develop and perform general likelihood ratio hypothesis testing. The likelihood function can be defined using provided classes, which are based on its mathematical objects, e.g. PDFs, PDF ratio, etc. The classes are implemented using object oriented programming technique. It allows to have a hierarchical class structure and to create common code interfaces. Hence, mathematical objects of the same kind have the same interfaces and can be swapped.

This means that every customized mathematical object will have implementation of at least base object methods.

#### 4.1.1 Structure

The SkyLLH framework contains two projects, `skyllh` and `i3skyllh`:

- The `skyllh` project is publicly available on the open source GitHub repository<sup>1</sup>. It defines the core framework functionality and common code interfaces by utilizing the abstract base class (ABC) module.
- The `i3skyllh` project provides proprietary code, which extends the core framework functionality and is tailored to the IceCube Neutrino Observatory detector and ongoing likelihood analyses.

The `skyllh` project code is structured into six modules:

- `core` module holds all the abstract base classes, which define the core framework, and detector independent extended classes and functions.
- `i3` module extends the core framework with classes specifically customized to the IceCube detector and its data format. Additional experiments, like Fermi, Antares, etc., can be implemented into their own modules and be available for a combined likelihood analysis calculation.
- `physics` module provides classes with generic source flux models.
- `plotting` module contains common plotting utilities of generated objects.
- `utils` module holds functions extending standard Python module capabilities.

The distinct division of the framework structure fulfills one of SkyLLH design goals – having a detector independent framework suitable for different projects. It also allows an easy development of likelihood analyses with the combined data from different detectors. The modular approach and defined core interfaces make the framework more transparent and easier to maintain in the future.

#### 4.1.2 Sample Analysis Definition

The analysis definition process in SkyLLH is realized by creating an `Analysis` object with desired properties. The user can develop its own or utilize already predefined `create_analysis` functions, e.g. in the `diffuse_ps` module the `create_analysis` function creates a point source analysis object using multi-dimensional PDFs for the signal and background likelihood function construction, generated from the Monte Carlo data. The important `create_analysis` function parameters are `datasets`,

---

<sup>1</sup><https://github.com/IceCubeOpenSource/SkyLLH>

`source`, flux normalization, energy, and spectral index of the reference power law flux model, and `minimizer_impl`.

The `datasets` parameter contains a list of `Dataset` instances. SkyLLH has predefined functions for creating `DatasetCollection` instances of datasets, which are currently used in most of the ongoing IceCube analyses. The desired dataset collection, containing a list of the same version datasets, can simply be loaded from the provided `data_samples` dictionary. Predefined datasets can also be customized easily by adding additional experimental, Monte Carlo, and (or) auxiliary data definitions. In addition to the data definitions, the `Dataset` object can contain data preparation functions, which modify loaded data, and data field renaming dictionaries for the experimental and Monte Carlo datasets.

The `source` is defined by e.g. a `PointLikeSource` instance at a given location in the sky. If the likelihood developer wants to change the source position after the `Analysis` object is created, the `Analysis` class has a `change_source` method, which applies necessary changes to all the source dependent objects of the `Analysis`. For the source flux model the `PowerLawFlux` is created with the given flux normalization, energy, and spectral index values.

The likelihood developer can choose the minimizer implementation from `LBFGSMinimizerImpl` and `NR1dNsMinimizerImpl` objects. L-BFGS is a limited-memory BFGS (Broyden–Fletcher–Goldfarb–Shanno) algorithm, used for solving large nonlinear optimization problems subject to simple bounds on the variables [67]. If the likelihood developer wants to minimize only the number of signal events  $n_s$ , the one dimensional Newton-Raphson minimization method [68] is provided.

### 4.1.3 My contributions

My main contributions to the SkyLLH framework are presented in the following sections. Other contributions include setting up the unit tests framework, investigating the influence of the dataset  $\sin(\delta)$  binning to the spline PDF creation, and fixes of the existing code functionality and documentation.

#### 4.1.3.1 Data samples

As mentioned earlier, in order to simplify the `Dataset` object creation, SkyLLH provides predefined data samples. The samples, previously defined in SkyLab and being in the approved neutrino sources datasets list, were adapted to the SkyLLH dataset collection structure. Each dataset collection holds the short definition, describing properties of the used dataset, and its version. Versioning allows the likelihood developer to work using the most current dataset version, or to reproduce older analysis results with a specific dataset version.

The datasets from the same dataset collection version are defined by providing

name, livetime or path of a good run list file, and paths of experimental and Monte Carlo data files. For the energy PDF creation  $\sin(\delta)$  and  $\log(E)$  binning definitions are added. Auxiliary data files (e.g. kernel density estimation PDFs) can also be added by providing the paths of data files to the `add_aux_data_definition` method.

#### 4.1.3.2 Trials Calculation

The generation of a test statistic distribution requires to perform many analysis trials with pseudo data samples. Therefore, the `Analysis` class was extended with the `do_trials` method. It simply executes the `do_trial` method  $N$  times, using the given parameters, and collects returned results. For an efficient trials calculation the possibility to utilize multi-processing was added.

The `do_trial` method performs an analysis trial by generating a pseudo data sample. It contains scrambled background events and possible signal events. The number of injected signal events is usually determined from a Poisson statistic for a given mean number of signal events. If one is only interested in the test statistic distribution of the null hypothesis, the mean number of signal events can be set to zero. After the pseudo data sample is generated, the likelihood ratio function, at a given source location, is maximized. The likelihood maximization returns the test statistic value and the corresponding parameters, e.g.  $\hat{n}_s$  and  $\hat{\gamma}$ .

#### 4.1.3.3 Sensitivity and Discovery Potential<sup>2</sup>

The definition of the sensitivity and discovery potential was presented in section 3.1.2. Their estimation needs the null hypothesis (background only) test statistic distribution and test statistic distributions for a given mean number of signal events  $-\mu$ . The previously presented `do_trials` method can be utilized to generate these distributions.

The analysis trials calculation is a computationally expensive method. In order to minimize unnecessary calculations, we chose to implement a binary search method. For a given  $\mu$  range the test statistic distributions are calculated at the lower and upper range values. If the desired probability of the signal test statistic does not exceed the 50% (for the sensitivity) or the  $1 - 5\sigma$  (for the discovery potential) part of the background test statistic threshold at the upper range, the lower range value gets assigned the upper range value, while the upper range value doubles. This procedure continues until the  $\mu$  range, covering the desired probability, is found.

After the covering  $\mu$  range is found, it is minimized by calculating the test statistic distribution at the middle point of the range, and setting the lower or upper bound of the  $\mu$  range to the middle point, whichever gives a range still covering the desired

---

<sup>2</sup>The currently used code implementation was optimized from the one described here by reusing calculated trials and implementing a more rigid interval search method.

probability. This procedure is repeated until the  $\mu$  range size gets smaller than the given  $\mu$  tolerance. The final  $\mu$  value is linearly interpolated from the  $\mu$  range bounds at the desired probability point. To get the sensitivity or discovery potential, the final  $\mu$  value is converted to the flux using the `mu2flux` method, which returns the flux in  $(\text{GeV cm}^2 \text{s})^{-1}$  units.

#### 4.1.3.4 Logging

Logging is an important feature allowing the framework to be more transparent to its user. The slightly modified version of the standard Python logging module [69] was implemented. The `Logger` and `Handler` objects have logging and handling levels, e.g. “Info”, “Warning”, “Error”, etc., defined. The developed `debugging` module provides a functionality for setting up the desired logging level logger and the desired handling level stream and (or) file handler.

The standard logging flow is summarized in figure 4.1. The `Logger` object creates a `LogRecord` object and, if the record is not filtered out, it is passed to the appropriate `Handler`. The `Handler` then checks if the record is not filtered out at the handling level and emits the `LogRecord` to the stream and (or) file. By setting the appropriate logging and handling levels one can avoid the unnecessary formatting of the log message. This approach minimizes the additional computational costs introduced by the logging feature.

Because the SkyLLH framework utilizes multiple processes for the parallelized calculations, the standard Python logging flow was not supported. This issue was solved by introducing the `QueueHandler` class. It collects the logged `LogRecord` objects from the worker processes in a queue and passes them to the main process. After all worker processes are finished, the main process can handle the queued records by utilizing the standard `Logger.handle` method.

#### 4.1.3.5 Documentation

In order to have a self contained documentation, the SkyLLH framework utilizes Python docstrings of every class and method. The docstings follow the format of numpydoc docstring guide [71]. We chose the Sphinx tool [72] for writing and generating additional HTML documentation. The Sphinx `autodoc` extension allows to extract, format, and link docstrings from the code. It can automatically generate the documentation of SkyLLH API. Using this approach the framework developer has to maintain only the one code documentation source.

Besides the API, the HTML documentation also includes a short introduction to the framework and the tutorial of multi-dataset time-integrated single source analysis. The useful guides for the framework developers are separated into their own section.

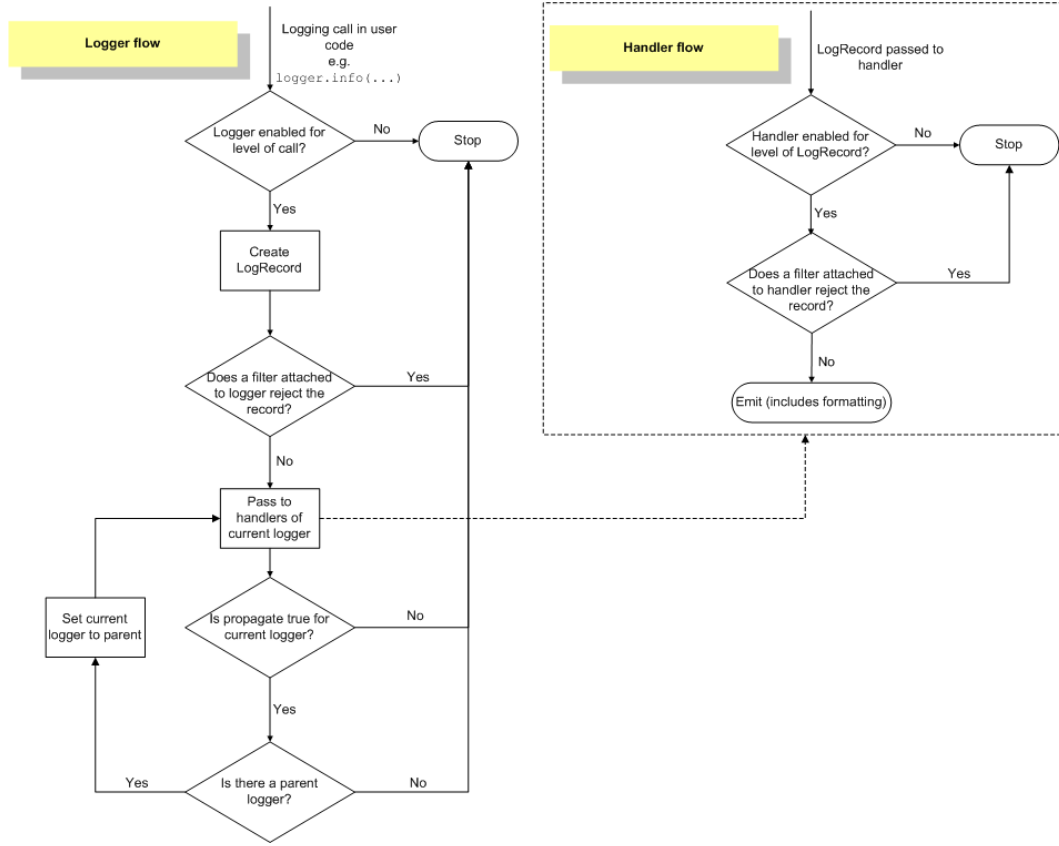


Figure 4.1: A flowchart of the logging procedure in Python. Figure taken from [70].

## 4.2 Kernel Density Estimation Tool

KDE Tool is a newly developed Python software package, which main function is to replace the current point source likelihood signal and background PDF generation with the use of a kernel density estimation method in order to better describe them. The tool provides a simple to use interface for the desired PDF generation, using either a binned or an adaptive KDE method. The mathematical formalism behind the PDF generation and the main KDE Tool methods are explained in following sections.



### 4.2.1 Kernel Density Estimation

The one dimensional probability density function of a  $\{x_i\}$  dataset, using a kernel density estimation method, is written as [73]

$$P_{\text{KDE}}(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - x_i}{h}\right), \quad (4.1)$$

where  $K(x)$  is a kernel function, which is centered at 0 and integrates to 1,  $N$  is the number of samples in the dataset, and  $h$  is a smoothing parameter – bandwidth, which would usually tend to 0 as  $N$  goes to  $\infty$ . The normalization constant before the sum ensures that  $\int P_{\text{KDE}}(x)dx = 1$ . The  $D$  dimensional probability density function of the  $\{\mathbf{x}_i\}$  dataset can be expressed using the multivariate kernel function as

$$P_{\text{KDE}}(\mathbf{x}) = \frac{1}{Nh^D} \sum_{i=1}^N K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right) \quad (4.2)$$

There exist many different alternative kernel forms (shown in table 4.1), although the quality of the PDF description depends much more on the chosen bandwidth, than the kernel form [74]. If not stated specifically, the Epanechnikov kernel is used for the following calculations, as it theoretically optimizes the asymptotic mean integrated squared error (AMISE) [75].

Table 4.1: One dimensional kernel forms with argument ranges.

Kernel name	$K(x)$
Uniform	$\frac{1}{2}$ for $x \in [-1, 1]$
Gaussian	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$
Epanechnikov	$\frac{3}{4}(1 - x^2)$ for $x \in [-1, 1]$
Triangle	$1 -  x $ for $x \in [-1, 1]$
Cosine	$\frac{\pi}{4} \cos\left(\frac{\pi}{2}x\right)$ for $x \in [-1, 1]$

The proper description of the probability density function using a narrow kernel needs large data samples, otherwise the KDE method just overfits the given data. On the other hand, if the kernel is too wide, boundaries and narrow structures will be smeared. This outcome can be described as a bias versus variance problem [74]. By expanding the kernel function using Taylor series one finds that the integrated square bias size is  $\propto h^4$ , while the integrated variance size is  $\propto h^{-1}$ . Because of a

finite sample size, the compromise between the introduced bias and variance has to be made. An example of the bandwidth size influence can be seen in figure 4.2.

There exists different methods for the optimal bandwidth estimation. If the data distribution is following a Gaussian function, one can use the *Silverman's rule of thumb* [74] estimation of the optimal bandwidth  $h_{\text{opt}} = \left(\frac{4\hat{\sigma}^5}{3N}\right)^{\frac{1}{5}}$ , where  $\hat{\sigma}$  is the standard deviation of the sample. If the data distribution form is unknown, one can either use the *plug-in* [76] or cross validation [77] method. Both methods use the initially generated  $P_{\text{KDE}}$  itself for the bandwidth estimation procedure.

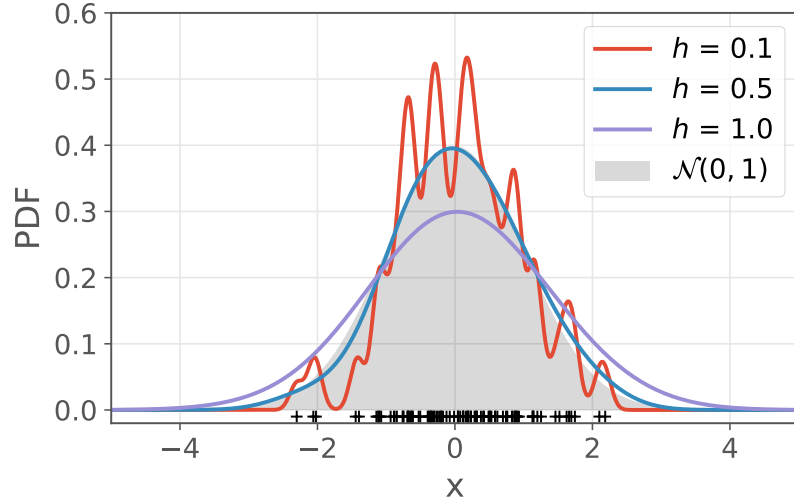


Figure 4.2: A kernel density estimation with different bandwidth  $h$  values. The data points (shown as black crosses) are drawn randomly from the standard normal distribution.

#### 4.2.2 Mathematical formalism

A new mathematical framework of a point source likelihood construction from the Monte Carlo data, using a kernel density estimation method, was developed in our group. It is completely general – compatible with any event reconstruction method and event angular uncertainty estimator. It strives to improve the description of the signal and background probability density functions in equation 3.2.

The general point source likelihood can be written as

$$\mathcal{L}(r_s, r_b | \{x\}, N) = \frac{(r_s + r_b)^N}{N!} e^{-(r_s + r_b)} \prod_i \left\{ \frac{r_s}{r_s + r_b} f(x_i | s) + \frac{r_b}{r_s + r_b} f(x_i | b) \right\}, \quad (4.3)$$

where  $r_s$  and  $r_b$  are the generation rates of the signal and background events, respectively. The data sample point  $x_i = (\mathbf{d}_{\text{rec}}^i, E_{\text{rec}}^i, \sigma^i)$  is defined by the observables:  $\mathbf{d}_{\text{rec}}$  is the reconstructed direction using the *SplineMPE* method,  $E_{\text{rec}}$  is the reconstructed energy using the truncated mean  $dE_\mu/dx$  method, and  $\sigma$  is the pull corrected paraboloid angular uncertainty estimation  $\sigma_{\text{paraboloid}}^{\text{corrected}}$ . The mentioned reconstruction methods are presented in section 2.2.

The joint PDF of the data sample  $\{x\}$  is

$$f(\{x\} | t = \text{sig, bkg}) = \prod_i f_t(\mathbf{d}_{\text{rec}}^i | E_{\text{rec}}^i, \sigma^i) f_t(E_{\text{rec}}^i, \sigma^i) \quad (4.4)$$

One can obtain the 4D joint PDF of all observables from the Monte Carlo dataset by using the 3D PDFs. The signal PDF is defined as

$$f_{\text{sig}}(\mathbf{d}_{\text{rec}}, E_{\text{rec}}, \sigma | \sin \delta_s, \gamma) \approx \frac{1}{2\pi \sin(\psi)} f_s(\psi | E_{\text{rec}}, \sigma, \gamma) f_e(E_{\text{rec}}, \sigma | \sin \delta_s, \gamma), \quad (4.5)$$

where the spatial PDF is defined as

$$f_s(\psi | E_{\text{rec}}, \sigma, \gamma) = \boxed{f_s(\psi, E_{\text{rec}}, \sigma | \gamma)} / \boxed{f_s(E_{\text{rec}}, \sigma | \gamma)} \quad (4.6)$$

and the energy PDF is defined as

$$f_e(E_{\text{rec}}, \sigma | \sin \delta_s, \gamma) = \boxed{f_e(E_{\text{rec}}, \sigma, \sin \delta_s | \gamma)} / \boxed{f_e(\sin \delta_s | \gamma)} \quad (4.7)$$

The background PDF is expressed as

$$f_{\text{bkg}}(\mathbf{d}_{\text{rec}}, E_{\text{rec}}, \sigma) = \frac{1}{2\pi} \boxed{f_b(\sin \delta_{\text{rec}}, E_{\text{rec}}, \sigma)} \quad (4.8)$$

The PDFs marked with a blue box are generated using a kernel density estimation method. The spatial PDF (equation 4.6) can be assumed to be rotationally symmetric, therefore, it can be defined in the source centered coordinates. The transformation to the source centered coordinates reduces the 2D direction observable  $\mathbf{d}_{\text{rec}}$  to the 1D  $\psi$  observable, where  $\psi$  is the spatial angle between the source direction and the reconstructed direction. The code is implemented to work on the log scale internally in order to improve the numerical stability.

### 4.2.3 KDE Tool

KDE Tool is a Python framework extending the Meerkat software package functionality. It allows the user to conveniently create a KDE object, which has already implemented methods for the generation of multi-dimensional PDFs. The properties of a multi-dimensional PDF, for each dimension, are defined in a model file:

- **values** parameter describes how to generate data. The user can choose which dataset field to use and apply additional mathematical functions to the chosen data.
- **bandwidth** parameter defines a grid of a kernel bandwidth  $h$ , which is used for the cross validation.
- **nbins** parameter defines the number of bins to use in the binned PDF approximation.
- **range** parameter can modify the PDF's phase space. The default phase space strictly covers the range of generated **values**.

The most important parts of the KDE Tool are presented in the following sections.

#### 4.2.3.1 Meerkat

Meerkat [78] (Multidimensional Efficiency Estimation using the Relative Kernel Approximation Technique) is a software package providing a set of C++ classes for the kernel density estimation of a multi-dimensional PDF over a given phase space. The code implementation depends on the ROOT [79] framework. Meerkat is the main KDE Tool dependency. It provides the implementation of the binned and adaptive KDE generation techniques, which will be presented in sections 4.2.3.3 and 4.2.3.4.

#### 4.2.3.2 Cross Validation

In order to be independent from a model of the data distribution in our sample, we chose to use the likelihood cross validation technique [74] for finding the optimal bandwidth. In general, the cross validation methods divide the sample into a training and a validation part. The leave-one-out cross validation (LOOCV) is an exhaustive cross validation method. It tests every possible way to divide the sample by splitting it into  $N$  subsamples, where  $N$  is the number of data points in the sample. Each subsample is then used once for the validation, while the remaining  $N - 1$  subsamples are used for the generation of the PDF. This requires to do  $N$  passes over the sample and is computationally expensive.

A non-exhaustive cross validation method is the K-fold cross validation. It splits the sample into  $K$  subsamples, where  $K < N$ . The LOOCV method can be seen

as a special case of the K-fold cross validation method with  $K = N$ . Choosing a  $K$  value allows to compromise between the computation time and the cross validation stability.

A figure of merit of the K-fold cross validation is defined as a likelihood

$$\mathcal{L}_{\text{CV}}(\mathbf{h}) = \frac{\sum_{k=1}^K \left[ \sum_{i=1}^{N/K} [\log(p_k(\mathbf{x}_i))w_{\mathbf{x}_i}] \right]}{K}, \quad (4.9)$$

where  $p_k$  is the generated PDF using a kernel density estimation on  $K - 1$  training subsamples (with bandwidth  $\mathbf{h}$ ),  $\mathbf{x}_i$  is a point from the validation subsample, and  $w_{\mathbf{x}_i}$  is a weight of  $\mathbf{x}_i$ . Because the  $\mathcal{L}_{\text{CV}}(\mathbf{h})$  function is not differentiable, it is evaluated at every point of the bandwidth grid, which is defined in the PDF model file. The final PDF is generated using all data points and the  $\mathbf{h}_{\text{opt}}$  bandwidth value, which maximizes the  $\mathcal{L}_{\text{CV}}(\mathbf{h})$  likelihood. An example of the 5-fold cross validation bandwidth scan is shown in figure 4.3.

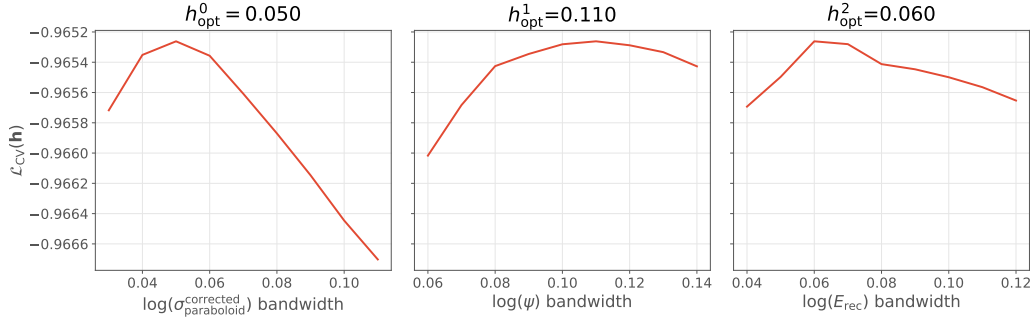


Figure 4.3: Slices of the 5-fold cross validation bandwidth scan at the  $\mathbf{h}_{\text{opt}}$  bandwidth.

#### 4.2.3.3 Binned KDE

The following mathematical derivation follows [78]. In the limit of large data samples ( $N \rightarrow \infty$ ), the constant bandwidth  $P_{\text{KDE}}(x)$  description of a dataset suffers from the boundary effect (underestimation of the true PDF at the dataset edges). One can notice that the  $P_{\text{KDE}}(x)$  approximates a convolution of the  $P_{\text{true}}(x)$  with the kernel  $K(x)$ . The paper proposes a simple solution for the boundary effect correction.

The corrected  $P_{\text{corr}}(x)$  probability density function is defined as

$$P_{\text{corr}}(x) = \begin{cases} \frac{\sum_{i=1}^N K(x-x_i)}{(U \otimes K)(x)} & \text{for } x \in X, \\ 0 & \text{otherwise.} \end{cases} \quad (4.10)$$

where the  $U(x)$  function equals 1 in  $x \in X$  and 0 otherwise. The  $\otimes$  symbol denotes a convolution operator. If the  $P_{\text{true}}(x)$  is flat at the boundary, the  $P_{\text{corr}}(x)$  expression returns an asymptotically unbiased estimation for small kernel bandwidths. In order to incorporate a known  $P_{\text{true}}(x)$  behavior at the boundary, the  $U(x)$  function can be changed into the approximation function  $F(x)$ . This leads to the approximate PDF expression

$$P_{\text{approx}}(x) = \frac{\sum_{i=1}^N K\left(\frac{x-x_i}{h}\right)}{(F \otimes K)(x)} \times F(x) \quad (4.11)$$

The previously mentioned  $P_{\text{KDE}}(x)$  approximation holds near the boundaries  $\sum_{i=1}^N K(x - x_i) \simeq (F \otimes K)(x)$ , leaving  $P_{\text{approx}}(x) \simeq F(x)$ . In the case the approximation function changes slowly in the inner region, the  $(F \otimes K)(x) \simeq F(x)$  approximation holds and  $P_{\text{approx}} \simeq P_{\text{KDE}}$ .

Because the convolution part has to be calculated for each data point, the evaluation of equation 4.11 can become rather slow for large datasets. One can use the interpolated approach, where  $P_{\text{approx}}(x)$  is evaluated on multi-dimensional grid nodes and the function value is linearly interpolated

$$P_{\text{interp}}(x) = \text{Bin}(P_{\text{approx}}(x)) = \frac{\text{Bin}(\sum_{i=1}^N K(\frac{x-x_i}{h}))}{\text{Bin}((F \otimes K)(x))} \times F(x), \quad (4.12)$$

where  $\text{Bin}(f)$  denotes a multi-linear interpolation of the function  $f$ . The binned KDE method evaluates equation 4.12 with the approximate function  $F(x) = U(x)$  and a constant bandwidth.

#### 4.2.3.4 Adaptive KDE

The adaptive KDE method modifies equation 4.1 by allowing the bandwidth  $h$  to vary. The main idea of this method is to increase the bandwidth at the low data density regions and decrease it in the regions, where data is more concentrated. This allows to minimize bumps in the distribution tail and describe the given data more accurately. There are two main adaptive KDE methods, called *balloon estimator* and *sample smoothing estimator* [80, 81].

The balloon estimator is expressed as

$$\hat{f}_1(\mathbf{y}) = \frac{1}{Nh(\mathbf{y})^D} \sum_{i=1}^N K\left(\frac{\mathbf{x}_i - \mathbf{y}}{h(\mathbf{y})}\right), \quad (4.13)$$

where the bandwidth depends on a point at which the kernel is estimated. The main problem of this estimator is that it typically does not integrate to 1 when applied globally.

The sample smoothing estimator is expressed as

$$\hat{f}_2(\mathbf{y}) = \frac{1}{N} \sum_{i=1}^N \frac{1}{h(\mathbf{x}_i)^D} K\left(\frac{\mathbf{x}_i - \mathbf{y}}{h(\mathbf{x}_i)}\right), \quad (4.14)$$

where the kernel size is individually scaled. This estimator does not have the previously mentioned estimator problem –  $\hat{f}_2(\mathbf{y})$  returns a probability density function. Its only disadvantage is that the estimator can be influenced by the data points lying far away from the estimation location.

In Meerkat, the adaptive KDE PDF generation uses the sample smoothing estimator, where the bandwidth  $h$  in equation 4.12 kernel is expressed as

$$h(\mathbf{x}_i) = h_0 P(\mathbf{x}_i)^{-1/D}, \quad (4.15)$$

where  $P(\mathbf{x}_i)$  is an approximated local data density at the point  $\mathbf{x}_i$ ,  $h_0$  is the initial bandwidth, and  $D$  is the number of dimensions. The local data density is unknown, therefore, one has to use another method for the initial estimation. We chose to use the binned KDE method to estimate the initial PDF. The optimal bandwidth value for its generation is picked using the bandwidth cross validation (equation 4.9). After that, the initial PDF is used in the adaptive KDE method as an approximation function  $F(x)$  and as a local data density approximation in equation 4.15. Finally, the initial bandwidth  $h_0$  in the adaptive KDE method is cross validated, too.

### 4.3 Results

We start by investigating the difference between the maximum cross validation likelihood (equation 4.9) of the binned KDE and adaptive KDE methods. Later we compare properties of the point source search, using the adaptive KDE method to generate likelihood PDFs, with the standard point source search. All the following results were generated using the Northern Tracks version-002-p05 Monte Carlo dataset for the IC86, 2012-2016 season.

The components of the signal PDF (equation 4.5) depend on the source spectral index  $\gamma$ . We investigate the spectral index influence by generating PDFs with  $\gamma$  values of 2.0 and 3.5, hard and soft spectra, respectively. The background PDF component (equation 4.8) does not depend on the spectral index. The maximum cross validation results between the binned and adaptive KDE methods are presented in table 4.2.

From the maximum cross validation results we see that the adaptive KDE method gives up to a few percent improvement in comparison to the binned KDE method and is larger for higher dimensional PDFs. The limit of model independent kernel density estimation method is reached, because the used Monte Carlo dataset does not sample the whole parameter space sufficiently. In order to gain a larger improvement,

Table 4.2: Maximum cross validation results between the binned KDE and adaptive KDE methods. “Binned KDE” and “Adaptive KDE” columns show the cross validation figure of merit (equation 4.9) at respective  $\mathbf{h}_{\text{opt}}$  bandwidths. Relative improvement is defined as  $\frac{\mathcal{L}_{\text{CV}}^{\text{adaptive}}(\mathbf{h}_{\text{opt}}) - \mathcal{L}_{\text{CV}}^{\text{binned}}(\mathbf{h}_{\text{opt}})}{|\mathcal{L}_{\text{CV}}^{\text{binned}}(\mathbf{h}_{\text{opt}})|}$ .

$\gamma = 2.0$	Binned KDE	Adaptive KDE	Relative improvement, %
$f_s(\psi, E_{\text{rec}}, \sigma \gamma)$	-1.7292814	-1.7123579	0.98
$f_s(E_{\text{rec}}, \sigma \gamma)$	-1.3554718	-1.35133	0.31
$f_e(E_{\text{rec}}, \sigma, \sin \delta_s \gamma)$	-1.3440044	-1.3262637	1.32
$f_e(\sin \delta_s \gamma)$	-0.04730884	-0.04719831	0.23
$\gamma = 3.5$	Binned KDE	Adaptive KDE	Relative improvement, %
$f_s(\psi, E_{\text{rec}}, \sigma \gamma)$	-0.96526176	-0.96014804	0.53
$f_s(E_{\text{rec}}, \sigma \gamma)$	-0.5605257	-0.5600165	0.09
$f_e(E_{\text{rec}}, \sigma, \sin \delta_s \gamma)$	-0.5599171	-0.55274665	1.28
$f_e(\sin \delta_s \gamma)$	-0.0507674	-0.05074739	0.04
$f_b(\sin \delta_{\text{rec}}, E_{\text{rec}}, \sigma)$	-0.5393583	-0.52820015	2.07

one could construct the approximation function  $F(x)$  using the fact, that the  $\psi$  distribution should approximately follow the Rayleigh distribution.

For the one dimensional  $f_e(\sin \delta_s|\gamma)$  PDF one can do a simple check between the Monte Carlo data and the generated PDF using the binned KDE and adaptive KDE methods. It is shown in figures 4.4 and 4.5. Both PDFs are almost identical and closely follow the binned Monte Carlo data. Such a direct check for the higher dimensional PDFs gets difficult to interpret, therefore, we move on to the checks, where generated PDFs are used as an input parameter. As the adaptive KDE method behaves as expected and gives a slight improvement in comparison to the binned KDE method, we only check how its behavior compares to the standard point source analysis methods.

In order to check how well the spatial part of the signal PDF compares to the standard paraboloid approximation and the Monte Carlo data, we pick an ensemble of events with the similar paraboloid description. This means that we need to define



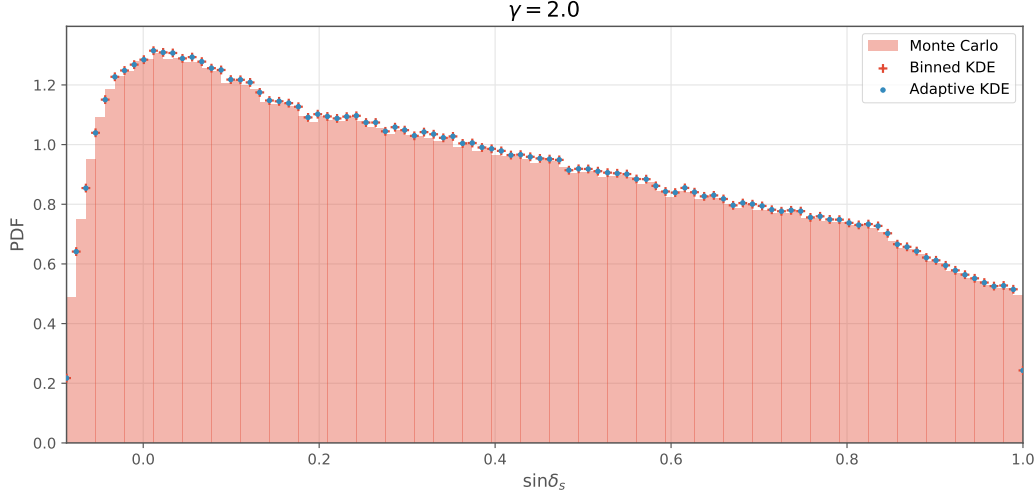


Figure 4.4:  $f_e(\sin \delta_s | \gamma)$  PDF distribution comparison of the Monte Carlo data with the binned KDE and adaptive KDE methods for  $\gamma = 2.0$ .

a “bin” in the reconstructed energy and the corrected paraboloid angular uncertainty values.

$$f(\psi | E_{\text{rec}} \in \Delta E_{\text{rec}}, \sigma_{\text{paraboloid}}^{\text{corrected}} \in \Delta \sigma_{\text{paraboloid}}^{\text{corrected}}) = \sum_{i=1}^N f(\psi | E_{\text{rec}}^i, \sigma_{\text{paraboloid}}^{\text{corrected}, i}) \times \frac{w_i}{W}, \quad (4.16)$$

where  $w_i$  is the weight of the Monte Carlo event. The  $f(\psi | E_{\text{rec}}, \sigma_{\text{paraboloid}}^{\text{corrected}})$  term can be expressed either as a Rayleigh PDF or the spatial PDF (equation 4.6). The results of different “bins” are shown in figures 4.6 and 4.7, for the spectral index values of 2.0 and 3.5, respectively. We see that the signal spatial PDF, generated using the adaptive KDE method, improves the description of the Monte Carlo events at the low energy regime and the distribution tails. The improvement for the soft spectra is even larger.

The SkyLLH framework provides a native support of the KDE Tool generated PDFs by expanding the predefined Northern Tracks dataset with auxiliary PDF files. It also provides `create_analysis` function, which is designed for building the likelihood ratio object from multi-dimensional PDFs. Therefore, using the SkyLLH framework, one can implement the whole analysis. In order to see how well the KDE method improves the PDFs for the likelihood function, we implemented a  $n_{\text{sig, fit}}$  bias test. It injects the number of signal events (described by the mean of Poisson statistics) into a pseudo dataset and checks the number of fit signal events by the

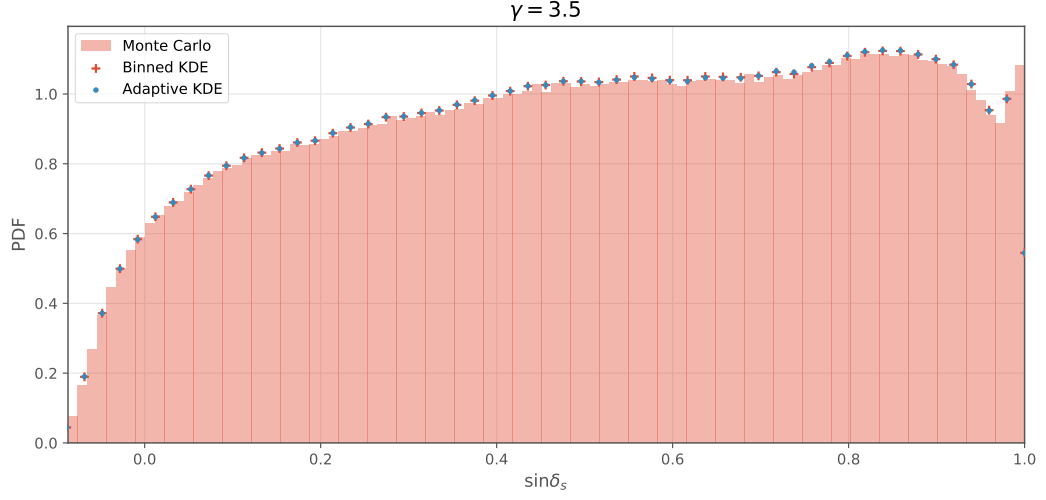
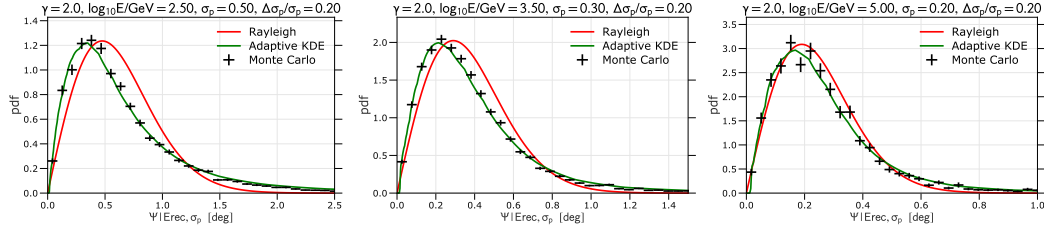


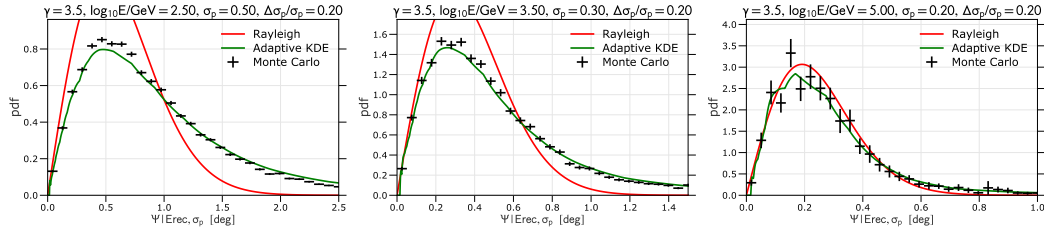
Figure 4.5:  $f_e(\sin \delta_s | \gamma)$  PDF distribution comparison of the Monte Carlo data with the binned KDE and adaptive KDE methods for  $\gamma = 3.5$ .

analysis. The point source position is set at  $(\alpha_s = 77.358^\circ, \delta_s = 5.693^\circ)$ . The results can be seen in figures 4.8 and 4.9, for  $\gamma$  values equal to 2.0 and 3.5, respectively. Ideally, the median of the number of fit signal events distribution should follow the diagonal line, otherwise the analysis likelihood function description introduces a bias term. The adaptive KDE method shows almost no bias for  $\gamma = 2.0$ . For  $\gamma = 3.5$  it introduces a slight bias, due to the overestimation of the spatial PDF (figure 4.7) tails. It could possibly be improved by using better reconstructed observables.



(a)  $\log(E)/\text{GeV} = 2.5$ ,  $\sigma_p = 0.5$  (b)  $\log(E)/\text{GeV} = 3.5$ ,  $\sigma_p = 0.3$  (c)  $\log(E)/\text{GeV} = 5.0$ ,  $\sigma_p = 0.2$

Figure 4.6: Spatial PDF distribution for  $\gamma = 2.0$  and different reconstructed energy and corrected paraboloid angular uncertainty bins.



(a)  $\log(E)/\text{GeV} = 2.5$ ,  $\sigma_p = 0.5$  (b)  $\log(E)/\text{GeV} = 3.5$ ,  $\sigma_p = 0.3$  (c)  $\log(E)/\text{GeV} = 5.0$ ,  $\sigma_p = 0.2$

Figure 4.7: Spatial PDF distribution for  $\gamma = 3.5$  and different reconstructed energy and corrected paraboloid angular uncertainty bins.

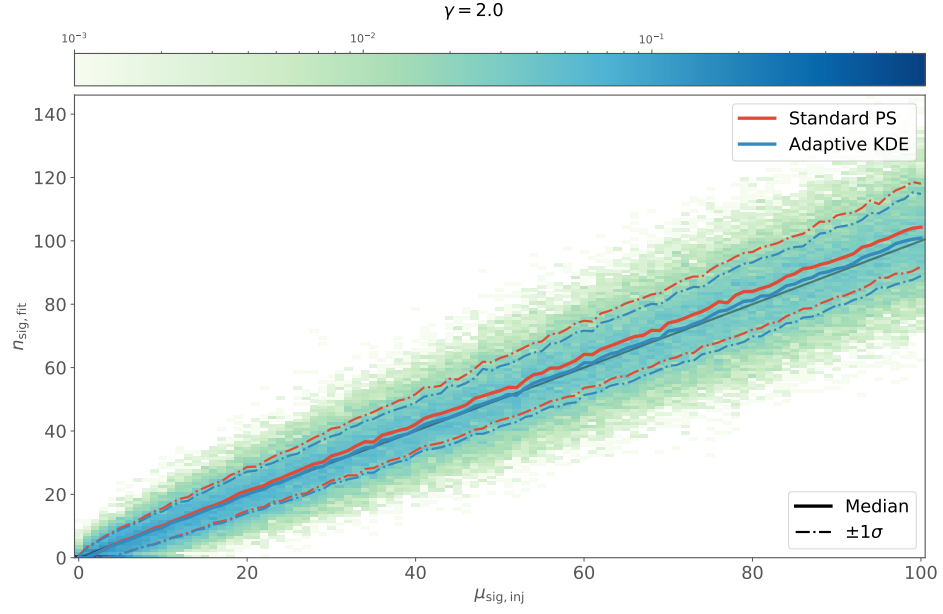


Figure 4.8: The number of injected signal events fit for  $\gamma = 2.0$  spectral index.

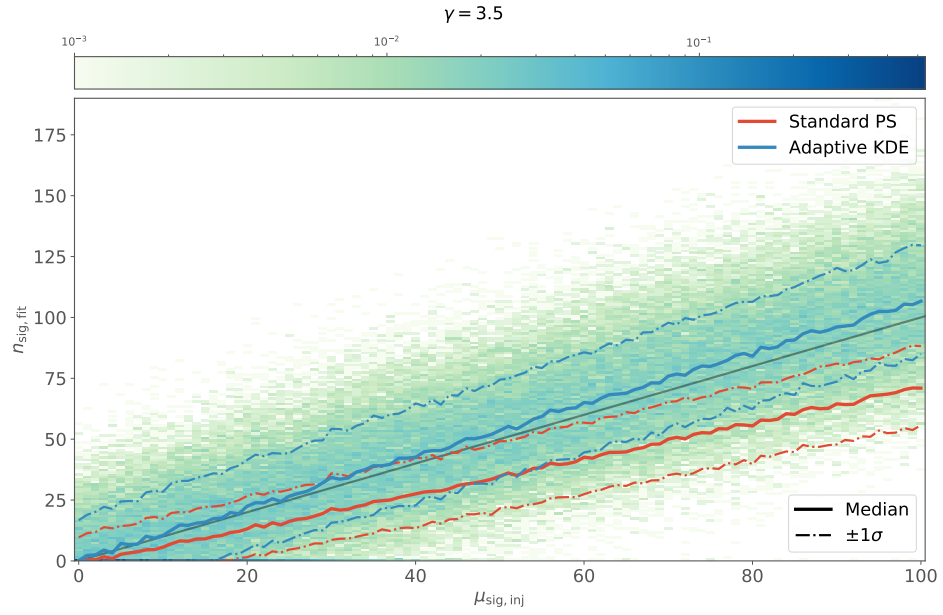


Figure 4.9: The number of injected signal events fit for  $\gamma = 3.5$  spectral index.

## Chapter 5

# Summary and Conclusion

The origin of the detected astrophysical neutrino flux is still a mystery. This thesis presented steps to improve the time-integrated neutrino point source search using newly developed software tools. The introduction of the adaptive kernel density estimation method for the generation of the likelihood function's signal and background probability density functions (PDFs) shows an improvement in the spatial PDF description and in the parameter fit result for the number of signal events. These improvements are particularly pronounced for softer spectrum sources. The SkyLLH framework allows a straightforward analysis implementation and the inclusion of the PDFs, generated using the KDE Tool. The improvement of the PDF description should lead to better analysis sensitivities and discovery potentials.

The KDE Tool will be made publicly available. Because it is analysis independent, it could be used in other IceCube analyses to achieve the same improvements and to solve the  $n_{\text{sig,fit}}$  vs.  $\mu_{\text{sig,inj}}$  bias problem.

In order to achieve even larger improvements in the future, one could use a Rayleigh distribution for the construction of the adaptive KDE method approximation function. It would help to better describe the spatial PDF in the parameter space regions, where the coverage of the Monte Carlo dataset is limited. Another potential improvement could be the utilization of improved event reconstruction observables as input for the KDE Tool. In this regard, machine learning techniques applied to the IceCube data have shown promising results in the enhanced reconstruction of the event energy and the event angular uncertainty [82].



## Appendix A

### List of figures and tables





# List of Figures

1.1	Standard model of elementary particles. Figure taken from [5]. . . . .	2
1.2	Cosmic ray proton and all-particle spectra. Measured data is taken from [8–17]. The plot was generated using [18]. . . . .	3
1.3	The Hillas plot shows the size and the magnetic field strength of possible acceleration sites. Objects above the diagonal line could accelerate protons up to $10^{20}$ eV. Figure taken from [23]. . . . .	4
1.4	Propagation and detection scheme of different messengers from the source. Figure taken from [27] . . . . .	5
2.1	The hexagonal structure of the IceCube detector. Deployment seasons of the strings are indicated by different colors at the surface of the ice. The strings form a triangular grid with a horizontal spacing of about 125 m. Figure taken from [28]. . . . .	8
2.2	An example of “charged current” interaction vertices. . . . .	9
2.3	An example of “neutral current” interaction vertices. . . . .	9
2.4	Neutrino- and antineutrino-nucleon cross sections for CC and NC interactions. In addition, the Glashow resonance of the electron antineutrino-electron cross section is shown. Figure taken from [6]. . .	10
2.5	Signatures of three different propagating leptons. The color scale ranges from early hits (red) to late hits (blue). The bubble size is proportional to the logarithm of the deposited energy inside the DOM. Figures taken from [35]. . . . .	10
2.6	A schematic view of the IceCube digital optical module. Figure taken from [39]. . . . .	11
2.7	The SPICE 3.2.1 model [43]. Dashed red lines show the effective scattering (top) and absorption (bottom) lengths dependency on depth. One can see a dust layer impact between 2000 m and 2100 m. Figure taken from [47]. . . . .	13
2.8	Paraboloid fit coordinate system. Figure taken from [53]. . . . .	15
3.1	Illustration for the definition of sensitivity and discovery potential. . .	22

3.2 An $\alpha$ level test and a p-value example. Given the $\alpha = 5\%$ , one can see that the $TS_{\text{obs}}$ rejects the null hypothesis. The p-value of observed test statistic is $p = 2\%$ , which also indicates a strong evidence against the null hypothesis. . . . .	23
4.1 A flowchart of the logging procedure in Python. Figure taken from [70].	30
4.2 A kernel density estimation with different bandwidth $h$ values. The data points (shown as black crosses) are drawn randomly from the standard normal distribution. . . . .	32
4.3 Slices of the 5-fold cross validation bandwidth scan at the $\mathbf{h}_{\text{opt}}$ bandwidth. . . . .	35
4.4 $f_e(\sin \delta_s   \gamma)$ PDF distribution comparison of the Monte Carlo data with the binned KDE and adaptive KDE methods for $\gamma = 2.0$ . . . . .	39
4.5 $f_e(\sin \delta_s   \gamma)$ PDF distribution comparison of the Monte Carlo data with the binned KDE and adaptive KDE methods for $\gamma = 3.5$ . . . . .	40
4.6 Spatial PDF distribution for $\gamma = 2.0$ and different reconstructed energy and corrected paraboloid angular uncertainty bins. . . . .	41
4.7 Spatial PDF distribution for $\gamma = 3.5$ and different reconstructed energy and corrected paraboloid angular uncertainty bins. . . . .	41
4.8 The number of injected signal events fit for $\gamma = 2.0$ spectral index. . .	42
4.9 The number of injected signal events fit for $\gamma = 3.5$ spectral index. . .	42

# List of Tables

3.1	The two types of errors in a hypothesis testing. . . . .	21
4.1	One dimensional kernel forms with argument ranges. . . . .	31
4.2	Maximum cross validation results between the binned KDE and adaptive KDE methods. “Binned KDE” and “Adaptive KDE” columns show the cross validation figure of merit (equation 4.9) at respective $\mathbf{h}_{\text{opt}}$ bandwidths. Relative improvement is defined as $\frac{\mathcal{L}_{\text{CV}}^{\text{adaptive}}(\mathbf{h}_{\text{opt}}) - \mathcal{L}_{\text{CV}}^{\text{binned}}(\mathbf{h}_{\text{opt}})}{ \mathcal{L}_{\text{CV}}^{\text{binned}}(\mathbf{h}_{\text{opt}}) } . . . . .$	38



# Bibliography

- [1] Chen Ning Yang. *Elementary Particles*. Apr. 18, 2016. ISBN: 978-0-691-65201-6.
- [2] J. Chadwick. “The Existence of a Neutron”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 136.830 (June 1, 1932), pp. 692–708. ISSN: 1364-5021, 1471-2946. DOI: 10.1098/rspa.1932.0112.
- [3] C. L. Cowan et al. “Detection of the Free Neutrino: A Confirmation”. In: *Science* 124.3212 (July 20, 1956), pp. 103–104. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.124.3212.103.
- [4] Todor Stanev. *High Energy Cosmic Rays*. 2nd ed. Springer-Praxis Books in Astronomy and Planetary Sciences. Berlin ; New York : Chichester, UK: Springer ; Praxis, 2009. 333 pp. ISBN: 978-3-540-85147-9.
- [5] *Standard Model*. In: *Wikipedia*. Page Version ID: 920221506. Oct. 8, 2019. URL: [https://en.wikipedia.org/w/index.php?title=Standard\\_Model&oldid=920221506](https://en.wikipedia.org/w/index.php?title=Standard_Model&oldid=920221506) (visited on 10/15/2019).
- [6] Thomas K. Gaisser, Ralph Engel, and Elisa Resconi. *Cosmic Rays and Particle Physics*. 2nd ed. Cambridge: Cambridge University Press, 2016. ISBN: 978-1-139-19219-4. DOI: 10.1017/CB09781139192194.
- [7] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *Journal of Instrumentation* 3.08 (Aug. 14, 2008), S08001–S08001. ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08001.
- [8] The Pierre Auger Collaboration. “The cosmic ray energy spectrum measured using the Pierre Auger Observatory”. In: *Proceedings of the 35th International Cosmic Ray Conference*. Vol. ICRC2017. Proceedings of Science, 2017, p. 486. arXiv: 1708.06592.
- [9] Y. S. Yoon et al. “Proton and Helium Spectra from the CREAM-III Flight”. In: *The Astrophysical Journal* 839.1 (2017), p. 5. DOI: 10.3847/1538-4357/a68e4. arXiv: 1704.02512.
- [10] S. Ter-Antonyan. “Sharp knee phenomenon of primary cosmic ray energy spectrum”. In: *Physical Review D* 89.12 (2014), p. 123003. DOI: 10.1103/PhysRevD.89.123003. arXiv: 1405.5472.

- [11] R. Alfaro et al. “All-particle cosmic ray energy spectrum measured by the HAWC experiment from 10 to 500 TeV”. In: *Physical Review D* 96.12 (2017), p. 122001. DOI: 10.1103/PhysRevD.96.122001. arXiv: 1710.00890.
- [12] T. Abu-Zayyad et al. “Measurement of the cosmic ray energy spectrum and composition from  $10^{17}$  eV to  $10^{18.3}$  eV using a hybrid fluorescence technique”. In: *The Astrophysical Journal* 557 (2001), pp. 686–699. DOI: 10.1086/322240. arXiv: astro-ph/0010652.
- [13] R. U. Abbasi et al. “First observation of the Greisen-Zatsepin-Kuzmin suppression”. In: *Physical Review Letters* 100 (2008), p. 101101. DOI: 10.1103/PhysRevLett.100.101101. arXiv: astro-ph/0703099.
- [14] M. G. Aartsen et al. “Measurement of the cosmic ray energy spectrum with IceTop-73”. In: *Physical Review D* 88.4 (2013), p. 042004. DOI: 10.1103/PhysRevD.88.042004. arXiv: 1307.3795.
- [15] T. Antoni et al. “KASCADE measurements of energy spectra for elemental groups of cosmic rays: Results and open problems”. In: *Astroparticle Physics* 24 (2005), pp. 1–25. DOI: 10.1016/j.astropartphys.2005.04.001. arXiv: astro-ph/0505413.
- [16] W. D. Apel et al. “The spectrum of high-energy cosmic rays measured with KASCADE-Grande”. In: *The Astrophysical J.* 36 (2012), pp. 183–194. DOI: 10.1016/j.astropartphys.2012.05.023. arXiv: 1206.3834.
- [17] M. Amenomori et al. “The all-particle spectrum of primary cosmic rays in the wide energy range from  $10^{14}$  eV to  $10^{17}$  eV observed with the Tibet-III air-shower array”. In: *The Astrophysical Journal* 678 (2008), pp. 1165–1179. DOI: 10.1086/529514. arXiv: 0801.1803.
- [18] Kai Krings. *Kkrings/Crspectra*. Sept. 19, 2019. URL: <https://github.com/kkrings/crspectra> (visited on 12/17/2019).
- [19] W. David Arnett et al. “Supernova 1987A”. In: *Annual Review of Astronomy and Astrophysics* 27.1 (Sept. 1989), pp. 629–700. ISSN: 0066-4146, 1545-4282. DOI: 10.1146/annurev.aa.27.090189.003213.
- [20] Kai Zuber. *Neutrino Physics*. 1st ed. Series in High Energy Physics, Cosmology and Gravitation. Taylor & Francis, 2003. ISBN: 0-7503-0750-1.
- [21] K. Hirata et al. “Observation of a Neutrino Burst from the Supernova SN1987A”. In: *Physical Review Letters* 58.14 (Apr. 6, 1987), pp. 1490–1493. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.58.1490.
- [22] C. Megan Urry and Paolo Padovani. “Unified Schemes for Radio-Loud Active Galactic Nuclei”. In: *Publications of the Astronomical Society of the Pacific* 107 (Sept. 1995), p. 803. ISSN: 0004-6280, 1538-3873. DOI: 10.1086/133630. arXiv: astro-ph/9506063.

- [23] A. M. Hillas. “The Origin of Ultra-High-Energy Cosmic Rays”. In: *Annual Review of Astronomy and Astrophysics* 22.1 (Sept. 1984), pp. 425–444. ISSN: 0066-4146, 1545-4282. DOI: 10.1146/annurev.aa.22.090184.002233.
- [24] Luis A. Anchordoqui et al. “Cosmic Neutrino Pevatrons: A Brand New Pathway to Astronomy, Astrophysics, and Particle Physics”. In: *Journal of High Energy Astrophysics* 1-2 (May 2014), pp. 1–30. ISSN: 22144048. DOI: 10.1016/j.jheap.2014.01.001. arXiv: 1312.6587.
- [25] Julia K. Becker. “High-Energy Neutrinos in the Context of Multimessenger Physics”. In: *Physics Reports* 458.4-5 (Mar. 2008), pp. 173–246. ISSN: 03701573. DOI: 10.1016/j.physrep.2007.10.006. arXiv: 0710.1557.
- [26] B. P. Abbott et al. “Observation of Gravitational Waves from a Binary Black Hole Merger”. In: *Physical Review Letters* 116.6 (Feb. 11, 2016), p. 061102. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.116.061102.
- [27] *Physicus - IceCube Gallery*. URL: <https://gallery.icecube.wisc.edu/internal/v/GraphicRe/promo/Particle+Physics/physicus.pdf.html> (visited on 10/29/2019).
- [28] *ArrayWSeasonsLabels - Icecube Gallery*. URL: <https://gallery.icecube.wisc.edu/internal/v/GraphicRe/visuals/detector/arraygraphics2011/other/pdf/ArrayWSeasonsLabels.pdf.html> (visited on 11/28/2019).
- [29] P. A. Čerenkov. “Visible Radiation Produced by Electrons Moving in a Medium with Velocities Exceeding That of Light”. In: *Physical Review* 52.4 (Aug. 15, 1937), pp. 378–379. ISSN: 0031-899X. DOI: 10.1103/PhysRev.52.378.
- [30] IceCube Collaboration et al. “Measurement of the Atmospheric Neutrino Energy Spectrum from 100 GeV to 400 TeV with IceCube”. In: *Physical Review D* 83.1 (Jan. 5, 2011), p. 012001. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.83.012001. arXiv: 1010.3980.
- [31] J. A. Formaggio and G. P. Zeller. “From eV to EeV: Neutrino Cross Sections Across Energy Scales”. In: *Reviews of Modern Physics* 84.3 (Sept. 24, 2012), pp. 1307–1341. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.84.1307. arXiv: 1305.7513.
- [32] Sheldon L. Glashow. “Resonant Scattering of Antineutrinos”. In: *Physical Review* 118.1 (Apr. 1, 1960), pp. 316–317. ISSN: 0031-899X. DOI: 10.1103/PhysRev.118.316.
- [33] IceCube Collaboration et al. “Search for Steady Point-like Sources in the Astrophysical Muon Neutrino Flux with 8 Years of IceCube Data”. In: *The European Physical Journal C* 79.3 (Mar. 2019), p. 234. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-019-6680-0.

- [34] IceCube Collaboration et al. “Search for Astrophysical Sources of Neutrinos Using Cascade Events in IceCube”. In: *The Astrophysical Journal* 846.2 (Sept. 11, 2017), p. 136. ISSN: 1538-4357. DOI: 10.3847/1538-4357/aa8508. arXiv: 1705.02383.
- [35] *Detecting Neutrinos / IceCube Masterclass*. URL: <https://masterclass.icecube.wisc.edu/en/learn/detecting-neutrinos> (visited on 12/01/2019).
- [36] IceCube Collaboration et al. “Measurement of South Pole Ice Transparency with the IceCube LED Calibration System”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 711 (May 2013), pp. 73–89. ISSN: 01689002. DOI: 10.1016/j.nima.2013.01.054. arXiv: 1301.5361.
- [37] I. Frank and Ig. Tamm. “Coherent Visible Radiation of Fast Electrons Passing Through Matter”. In: I. E. Tamm. *Selected Papers*. Ed. by Boris M. Bolotovskii, Victor Ya. Frenkel, and Rudolf Peierls. Berlin, Heidelberg: Springer Berlin Heidelberg, 1991, pp. 29–35. ISBN: 978-3-642-74628-4 978-3-642-74626-0. DOI: 10.1007/978-3-642-74626-0\_2.
- [38] IceCube Collaboration et al. “The IceCube Neutrino Observatory: Instrumentation and Online Systems”. In: *Journal of Instrumentation* 12.03 (Mar. 14, 2017), P03012–P03012. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/03/P03012. arXiv: 1612.05093.
- [39] Ulrich F. Katz and Christian Spiering. “High-Energy Neutrino Astrophysics: Status and Perspectives”. In: *Progress in Particle and Nuclear Physics* 67.3 (July 2012), pp. 651–704. ISSN: 01466410. DOI: 10.1016/j.ppnp.2011.12.001. arXiv: 1111.0507.
- [40] R. Abbasi et al. “The IceCube Data Acquisition System: Signal Capture, Digitization, and Timestamping”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 601.3 (Apr. 2009), pp. 294–316. ISSN: 01689002. DOI: 10.1016/j.nima.2009.01.001.
- [41] P. O. Hulth and The AMANDA Collaboration. “The Amanda Experiment”. In: (Dec. 6, 1996). arXiv: astro-ph/9612068.
- [42] M. Ackermann et al. “Optical Properties of Deep Glacial Ice at the South Pole”. In: *Journal of Geophysical Research* 111.D13 (2006), p. D13203. ISSN: 0148-0227. DOI: 10.1029/2005JD006687.
- [43] *Ice Models - IceCubeWiki*. URL: [https://wiki.icecube.wisc.edu/index.php/Ice\\_models#Current\\_recommended\\_ice\\_model\\_and\\_settings](https://wiki.icecube.wisc.edu/index.php/Ice_models#Current_recommended_ice_model_and_settings) (visited on 12/01/2019).



- [44] IceCube Collaboration et al. “The IceCube Neutrino Observatory Part VI: Ice Properties, Reconstruction and Future Developments”. In: (Sept. 27, 2013). arXiv: 1309.7010 [astro-ph].
- [45] Dmitry Chirkin. “Ice Model Update”. URL: <https://events.icecube.wisc.edu/event/75/contributions/4869/attachments/3947/4325/icetalk.pdf> (visited on 12/01/2019).
- [46] Dmitry Chirkin. “Photon Tracking with GPUs in IceCube”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 725 (Oct. 2013), pp. 141–143. ISSN: 01689002. DOI: 10.1016/j.nima.2012.11.170.
- [47] *Spice3.2.1 - IceCubeWiki*. URL: [https://wiki.icecube.wisc.edu/index.php/File:Spice3.2.1\\_layered\\_scatt\\_abs\\_withlength.png#filelinks](https://wiki.icecube.wisc.edu/index.php/File:Spice3.2.1_layered_scatt_abs_withlength.png#filelinks) (visited on 12/08/2019).
- [48] V. J. Stenger. *Track Fitting For Dumand II Octagon Array*. External Report HDC-1-90. University of Hawai’i at Manoa, Manoa, Hawaii, USA, 1990.
- [49] M. G. Aartsen et al. “Improvement in Fast Particle Track Reconstruction with Robust Statistics”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 736 (Feb. 2014), pp. 143–149. ISSN: 01689002. DOI: 10.1016/j.nima.2013.10.074. arXiv: 1308.5501.
- [50] J. Ahrens et al. “Muon Track Reconstruction and Data Selection Techniques in AMANDA”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 524.1-3 (May 2004), pp. 169–194. ISSN: 01689002. DOI: 10.1016/j.nima.2004.01.065.
- [51] Nathan Whitehorn, Jakob van Santen, and Sven Lafebre. “Penalized Splines for Smooth Representation of High-Dimensional Monte Carlo Datasets”. In: *Computer Physics Communications* 184.9 (Sept. 2013), pp. 2214–2220. ISSN: 00104655. DOI: 10.1016/j.cpc.2013.04.008. arXiv: 1301.2184.
- [52] Federica Bradascio and Thorsten Glüsenskamp. “Improving the Muon Track Reconstruction of IceCube and IceCube-Gen2”. In: *EPJ Web of Conferences* 207 (2019), p. 05002. ISSN: 2100-014X. DOI: 10.1051/epjconf/201920705002. arXiv: 1905.09612. URL: <http://arxiv.org/abs/1905.09612> (visited on 12/05/2019).
- [53] Till Neunhöffer. “Estimating the Angular Resolution of Tracks in Neutrino Telescopes Based on a Likelihood Analysis”. In: *Astroparticle Physics* 25.3 (Apr. 2006), pp. 220–225. ISSN: 09276505. DOI: 10.1016/j.astropartphys.2006.01.002. arXiv: astro-ph/0403367.

- [54] IceCube Collaboration et al. “Energy Reconstruction Methods in the IceCube Neutrino Telescope”. In: *Journal of Instrumentation* 9.03 (Mar. 17, 2014), P03009–P03009. ISSN: 1748-0221. DOI: 10.1088/1748-0221/9/03/P03009. arXiv: 1311.4767.
- [55] Leif Rädcl and Christopher Wiebusch. “Calculation of the Cherenkov Light Yield from Electromagnetic Cascades in Ice with Geant4”. In: *Astroparticle Physics* 44 (Apr. 2013), pp. 102–113. ISSN: 09276505. DOI: 10.1016/j.astropartphys.2013.01.015. arXiv: 1210.5140.
- [56] Dmitry Chirkin and Wolfgang Rhode. “Propagating Leptons through Matter with Muon Monte Carlo (MMC)”. In: (Aug. 3, 2016). arXiv: hep-ph/0407075.
- [57] IceCube collaboration et al. “An Improved Method for Measuring Muon Energy Using the Truncated Mean of dE/Dx”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 703 (Mar. 2013), pp. 190–198. ISSN: 01689002. DOI: 10.1016/j.nima.2012.11.081. arXiv: 1208.3430.
- [58] IceCube Collaboration et al. “Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector”. In: *Science* 342.6161 (Nov. 22, 2013), pp. 1242856–1242856. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1242856. arXiv: 1311.5238.
- [59] IceCube Collaboration. “Neutrino Emission from the Direction of the Blazar TXS 0506+056 Prior to the IceCube-170922A Alert”. In: *Science* (July 12, 2018), eaat2890. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aat2890. arXiv: 1807.08794.
- [60] Jim Braun et al. “Methods for Point Source Analysis in High Energy Neutrino Telescopes”. In: *Astroparticle Physics* 29.4 (May 2008), pp. 299–305. ISSN: 09276505. DOI: 10.1016/j.astropartphys.2008.02.007. arXiv: 0801.1604.
- [61] George Casella and Roger L. Berger. *Statistical Inference*. 2nd. ISBN: 0-534-24312-6.
- [62] Rene Reimann. “Differential Sensitivity and Limits in Point-Source Analyses”. In: *IceCube Internal Report* (Oct. 23, 2019), p. 11.
- [63] IceCube Collaboration et al. “Time-Integrated Neutrino Source Searches with 10 Years of IceCube Data”. In: (Oct. 18, 2019). arXiv: 1910.08488 [astro-ph].
- [64] The Fermi-LAT collaboration. “Fermi Large Area Telescope Fourth Source Catalog”. In: (Sept. 27, 2019). arXiv: 1902.10045 [astro-ph].
- [65] Martin Wolf. “SkyLLH – A Generalized Python-Based Tool for Log-Likelihood Analyses in Multi-Messenger Astronomy”. In: (Aug. 14, 2019). arXiv: 1908.05181 [astro-ph, physics:hep-ex].

- [66] IceCube Collaboration. *IceCube SVN, Skylab*. URL: <https://code.icecube.wisc.edu/svn/sandbox/skylab/> (visited on 11/24/2019).
- [67] Ciyu Zhu et al. “Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization”. In: *ACM Transactions on Mathematical Software* 23.4 (Dec. 1, 1997), pp. 550–560. ISSN: 00983500. DOI: 10.1145/279232.279236.
- [68] Rubin H. Landau, Manuel José Páez Mejía, and Cristian C. Bordeianu. *Computational Physics: Problem Solving with Computers*. 2nd revised and enlarged edition. OCLC: ocn163615850. Weinheim : Chichester: Wiley-VCH ; John Wiley [distributor], 2007. 593 pp. ISBN: 978-3-527-40626-5.
- [69] *Logging HOWTO — Python 3.8.0 Documentation*. URL: <https://docs.python.org/3/howto/logging.html#logging-advanced-tutorial> (visited on 11/11/2019).
- [70] *Logging Facility for Python — Python 3.8.0 Documentation*. URL: <https://docs.python.org/3/library/logging.html> (visited on 11/11/2019).
- [71] *Numpydoc Docstring Guide — Numpydoc v1.0.Dev0 Manual*. URL: <https://numpydoc.readthedocs.io/en/latest/format.html> (visited on 11/11/2019).
- [72] *Sphinx Documentation*. URL: <http://www.sphinx-doc.org/en/master/> (visited on 11/04/2019).
- [73] V. A. Epanechnikov. “Non-Parametric Estimation of a Multivariate Probability Density”. In: *Theory of Probability & Its Applications* 14.1 (Jan. 1969), pp. 153–158. ISSN: 0040-585X, 1095-7219. DOI: 10.1137/1114019.
- [74] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Monographs on Statistics and Applied Probability 26. Boca Raton: Chapman & Hall/CRC, 1998. 175 pp. ISBN: 978-0-412-24620-3.
- [75] J. S. Marron and M. P. Wand. “Exact Mean Integrated Squared Error”. In: *The Annals of Statistics* 20.2 (June 1992), pp. 712–736. ISSN: 0090-5364. DOI: 10.1214/aos/1176348653.
- [76] S. J. Sheather and M. C. Jones. “A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation”. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 53.3 (July 1991), pp. 683–690. ISSN: 00359246. DOI: 10.1111/j.2517-6161.1991.tb01857.x.
- [77] Peter Hall, J. S. Marron, and Byeong U. Park. “Smoothed Cross-Validation”. In: *Probability Theory and Related Fields* 92.1 (Mar. 1992), pp. 1–20. ISSN: 0178-8051, 1432-2064. DOI: 10.1007/BF01205233.

- [78] A. Poluektov. “Kernel Density Estimation of a Multidimensional Efficiency Profile”. In: *Journal of Instrumentation* 10.02 (Feb. 26, 2015), P02011–P02011. ISSN: 1748-0221. DOI: 10.1088/1748-0221/10/02/P02011.
- [79] Rene Brun and Fons Rademakers. “ROOT — An Object Oriented Data Analysis Framework”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1-2 (Apr. 1997), pp. 81–86. ISSN: 01689002. DOI: 10.1016/S0168-9002(97)00048-X.
- [80] George R. Terrell and David W. Scott. “Variable Kernel Density Estimation”. In: *The Annals of Statistics* 20.3 (Sept. 1992), pp. 1236–1265. ISSN: 0090-5364. DOI: 10.1214/aos/1176348768.
- [81] M.C. Jones. “VARIABLE KERNEL DENSITY ESTIMATES AND VARIABLE KERNEL DENSITY ESTIMATES”. In: *Australian Journal of Statistics* 32.3 (Sept. 1990), pp. 361–371. ISSN: 00049581. DOI: 10.1111/j.1467-842X.1990.tb01031.x.
- [82] Mirco Huennefeld. “Reconstruction Techniques in IceCube Using Convolutional and Generative Neural Networks”. In: *EPJ Web of Conferences* 207 (2019). Ed. by C. Spiering, p. 05005. ISSN: 2100-014X. DOI: 10.1051/epjconf/201920705005.

# Acknowledgements

This thesis would not have been possible without the generous support of many people.

First, I want to thank Prof. Elisa Resconi, who gave me an opportunity to work on such an interesting topic and be a part of her amazing working group at the Technical University of Munich. Moreover, I would like to thank her for the opportunity to participate in several conferences and meet different scientists from the whole IceCube collaboration.

Furthermore, I want to thank my direct supervisor Dr. Martin Wolf, who supported me during the whole thesis and patiently answered all my questions. In addition, thanks to Dr. Hans Niederhausen and Theo Glauch for the support and many suggestions, which improved my work and the software tools for the better.

A special thanks to all colleagues in the ECP group for a warm welcome and the great time I had with you.