Modellunabhängige Suche nach Neutrinoquellen mit dem ANTARES Neutrinoteleskop

Model-independent search for neutrino sources
with the ANTARES neutrino telescope

Der naturwissenschaftlichen Fakultät
der Friedrich-Alexander-Universität Erlangen-Nürnberg
zur Erlangung des Doktorgrades Dr. rer. nat.

vorgelegt von
Stefan Geißelsöder
aus Ansbach

Als Dissertation genehmigt von der naturwissenschaftlichen Fakultät
der Friedrich-Alexander-Universität Erlangen-Nürnberg

Tag der mündlichen Prüfung:	26.01.2016
Vorsitzender des Promotionsorgans:	Prof. Dr. Jörn Wilms
Gutachter:	Prof. Dr. Gisela Anton PD Dr. Alexander Kappes

Kurzzusammenfassung

Der Ursprung der hochenergetischen kosmischen Strahlung, die wir auf der Erde beobachten können, ist trotz jahrzehntelanger intensiver Forschung immer noch nicht eindeutig bekannt. Viele Theorien, die versuchen die Entstehung dieser Teilchen zu erklären, sagen auch einen Fluss von hochenergetischen Neutrinos aus dem Kosmos vorher. Zwar wurde vor kurzem die Existenz dieser Neutrinos bestätigt, die Quellen konnten dabei aber nicht identifiziert werden. Das ANTARES Neutrinoteleskop wurde im Mittelmeer 40 km von Toulon in einer Tiefe von 2475 m gebaut, um diese und andere Fragen zu klären. Es besteht aus 885 Photovervielfachern, die in einer dreidimensionalen Struktur angeordnet sind, um die Tscherenkovstrahlung zu detektieren, die von Teilchen ausgesendet wird, die bei der Interaktion von Neutrinos mit dem Meerwasser entstehen.

Die Interaktionen von Neutrinos in den ANTARES Daten zu identifizieren und zu rekonstruieren stellt herausfordernde Aufgaben dar. Die ersten beiden neuen Methoden, die in dieser Arbeit vorgestellt werden, versuchen diese Aufgaben möglichst effizient durch Mustererkennung zu lösen. Der erste Algorithmus klassifiziert die Daten in von unterhalb des Detektors kommende Neutrinosignale und von oberhalb kommende Teilchen, um einen Großteil der von oben kommenden atmosphärischen Muonen ausfiltern zu können. Der zweite Algorithmus verbessert die Richtungsrekonstruktion, indem er für jedes Neutrino den jeweils genausten Rekonstruktionsalgorithmus auswählt.

Der Hauptfokus dieser Arbeit liegt auf einer neuen Methodik, die die Verteilung der rekonstruierten Neutrinos auswertet. Im Gegensatz zu den meisten Analysen in ANTARES versucht diese Suche nicht, ein spezifisches, theoriebasiertes Modell einer Neutrinoquelle zu detektieren, sondern zielt darauf ab, die deutlichste Überfluktuation mit beliebiger Größe, Form und Verteilung der Neutrinos an einer beliebigen Position im Himmel zu finden. Um dieses Ziel zu erreichen wird die Dichte der Neutrinoereignisse in 180 voneinander unabhängigen Skalen ausgewertet. Wenn eine Region mit unerwartet hoher Neutrinodichte identifiziert worden ist, wird das potentielle Neutrinosignal mit Pseudoexperimenten mit zufälligen Neutrinoverteilungen verglichen, um seine Signifikanz zu bestimmen. Die Stärken einer derart flexiblen modellunabhängigen Suche liegen üblicherweise nicht in einer hohen Sensitivität bezüglich einer bestimmten Quellannahme, sondern im Entdecken einer Hypothese, die dann genauer ausgewertet werden kann.

In den Daten, die zwischen 2007 und 2012 aufgezeichnet wurden, wurde eine sehr große, ausgedehnte Struktur in Richtung des Zentrums unserer Galaxie gefunden. Die Signifikanz beträgt 2.52σ, was als konsistent mit zufälligen Fluktuationen betrachtet werden kann. Um eine unabhängige Überprüfung der Ergebnisse vorzunehmen wurde der IC40 Datensatz analysiert, der von der IceCube Kollaboration veröffentlicht wurde. Bei dieser Auswertung wurde tatsächlich in dem Bereich, an dem die Struktur aus den ANTARES Daten mit dem für IceCube sichtbaren Bereich überlappt, ebenfalls eine Überfluktuation festgestellt. Die so identifizierte Struktur hat eine Signifikanz von 2.14σ.

Da trotz diverser weiterer Studien anhand der vorliegenden Daten keine eindeutige Erklärung der Ergebnisse möglich ist, wird eine weitere Analyse empfohlen, die auf die gefundenen Ergebnisse spezialisiert sein kann und daher deutlich eindeutigere Erkenntnisse liefern könnte.

Unabhängig davon ist das erhaltene Ergebnis gegenwärtig die signifikanteste, räumlich begrenzte Hypothese für einen erhöhten Fluss hochenergetischer astrophysikalischer Neutrinos.

Abstract

The origin of high energetic cosmic rays has been puzzling since their discovery. Many theories about the sources of these cosmic rays also predict a flux of high energetic cosmic neutrinos. Recently, the existence of such a high energetic neutrino flux has been confirmed, but the location and nature of its sources remains unknown. The ANTARES neutrino telescope was built in the Mediterranean Sea, 40 km off the French coast near Toulon in a depth of 2475 meters to help answer this and other questions. It consists of a three dimensional array of 885 photomultiplier tubes that detect the Cherenkov light emitted by secondary particles, which are produced in interactions between neutrinos and nuclei in the water.

The identification and reconstruction of the observed neutrino events constitute challenging tasks. Parts of this thesis deal with algorithmic approaches to improve these tasks using pattern recognition. The first application is the suppression of undesired background by a classification algorithm. The second approach is the selection of the best available direction reconstruction for each neutrino.

The main focus of this thesis lies on a new method to evaluate the spatial distribution of the observed neutrinos. While most approaches test one specific hypothesis for a specific source, derived from theory or other measurements, this search refrains from optimizing for individual source hypotheses and tries to detect the most pronounced density fluctuation in the spatial distribution, regardless of its specific position, size, shape or internal distribution as unbiased as possible. To achieve this, the statistical likelihood for the observed neutrino density is evaluated in multiple scales up to distances between events of 180°. To recognize a potential cosmic neutrino signal, regions with the most pronounced deviations are identified and compared to the expectations from a random background hypothesis. The strength of such a flexible, model-independent search is not the sensitivity for a specific source hypothesis, but instead to detect also unexpected hypotheses that can then be analyzed in more detail.

In the data recorded from 2007 to 2012 this search found a very large structure close to the direction of the center of our galaxy with a post-trial significance of 2.52σ. It can therefore be explained best by a statistical fluctuation. As a simple crosscheck this method has been applied to a publicly available data sample recorded independently by the neutrino telescope IceCube. This evaluation also resulted in an overfluctuation at the location where the most significant structure from ANTARES data overlaps with the field of view of IceCube. With the devised analysis method the found structure in the IC40 data has a significance of 2.14σ.

While this is intriguing, ultimately, a dedicated follow-up analysis that is optimized for the derived hypothesis is necessary to find unambiguous evidence for its true nature.

Since, despite further studies, no unambiguous explanation could be found for the obtained results, a follow-up analysis is recommended, that can be adapted specifically to the results and therefore has a higher chance to provide unambiguous insights.

Nevertheless this result constitutes the most significant spatially resolved hypothesis for the sources of high energetic astrophysical neutrinos so far.

Contents

Part I

Introductions

This part of the thesis introduces various basic concepts useful for the understanding of the methods and results. It covers aspects of neutrinos in general, the ANTARES experiment, the underlying physics and of the algorithms used.

1 Motivation

The sky has always been fascinating for mankind, making us wonder about the mechanisms behind the universe. Until the 20^{th} century, observing the sky was limited to the visible part of the electromagnetic spectrum. This began to change in 1931 when Karl Guthe Jansky noticed that a part of the background noise in transatlantic radio communications originated from the Milky Way [1]. Over the course of the following decades the sky was studied not only in radio wavelength, but also in all other parts of the electromagnetic spectrum like microwaves, infrared, ultraviolet, X-ray or γ-ray. An example for the different aspects revealed by different parts of the electromagnetic spectrum is shown in Figure 1. Another good example for a multiwavelength observation revealing different aspects of the same object can be seen in Figure 5. With every new analyzed energy range new phenomena were discovered, deepening our understanding of the processes that shape our universe.

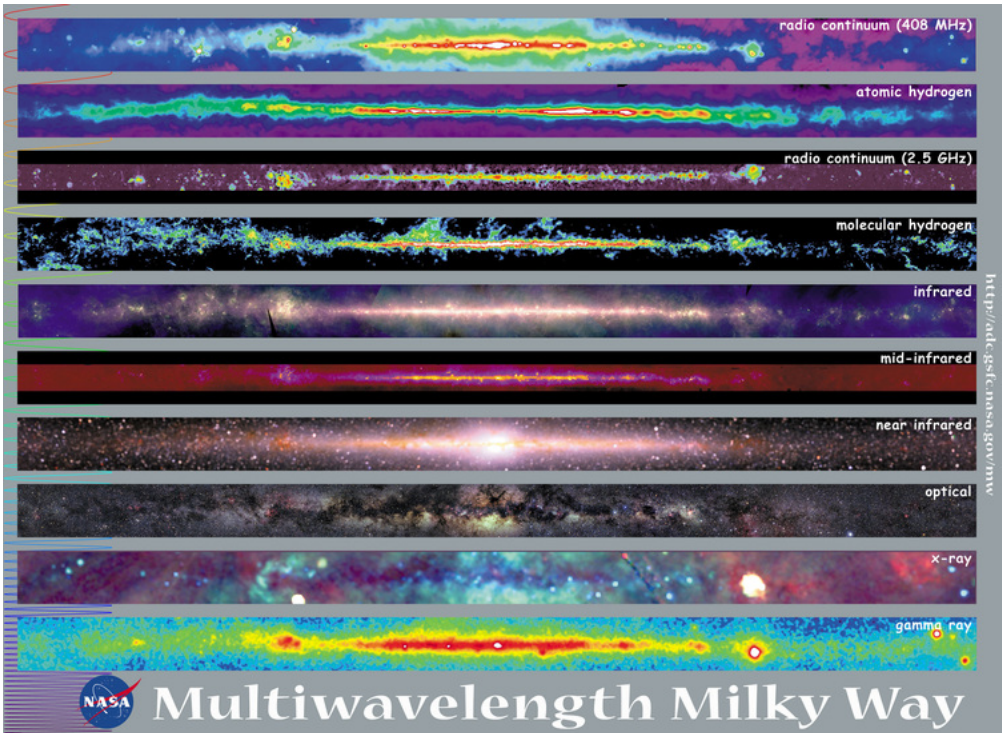


Figure 1: Our galaxy as seen in different wavelength bands. Taken from [2].

Unfortunately for astronomy, electromagnetic waves are absorbed if dense matter is between us and the source. This prevents a glimpse into some of the most interesting regions, for example the center of our galaxy.

A completely different approach in astronomy is to rely on other particles instead of photons. One candidate are so called cosmic rays. These particles have first been described by Victor Hess in 1912 in [3]. They consist mainly out of protons and alpha particles and their flux extends to very high energies as seen in Figure 2. It shows the flux of cosmic rays versus energy, measured by a multitude of independent experiments. For the low energy range, which is not covered in this plot, direct measurements of cosmic rays can be performed, for example by experiments on balloons, satellites or the International Space Station. The shown high energy range is mainly covered by Cherenkov telescopes, for instance the Pierre Auger Observatory [4]. But cosmic rays are charged particles and therefore they're influenced during their propagation by galactic and extragalactic magnetic fields. Hence only the most energetic particles are deflected little enough that they approximately point back to their origin.

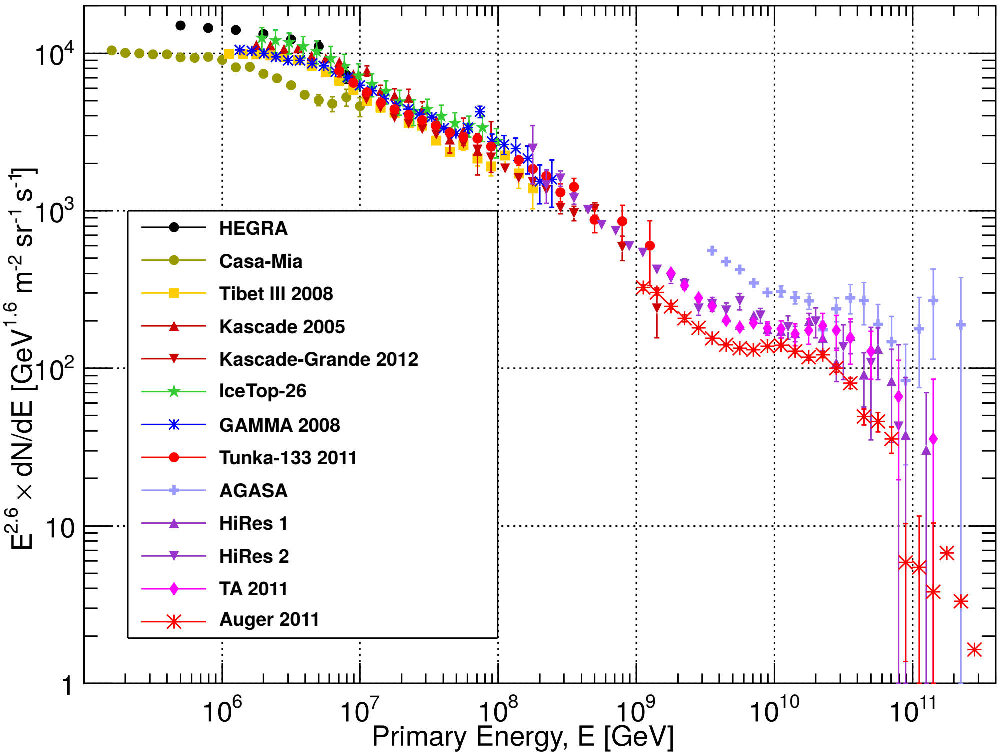


Figure 2: The energy spectrum of cosmic rays. From [5]. Multiple experiments have measured different energy ranges with different techniques. The features that can be observed in this energy distribution are the so called "knee" at around 10^8GeV and the "ankle" around 10^{10}GeV. The distribution has different spectral indices between them. These features suggest that different mechanisms could be involved.

The relatively new field of neutrino astronomy is a promising candidate to overcome these drawbacks. Neutrinos are electrically neutral particles, so they are not deflected by magnetic fields. They interact only by gravity and the short ranged weak force, allowing them to travel large distances even through matter. Because neutrinos cannot be observed directly, their detection must rely on secondary particles, which are generated once a neutrino interacts with matter. These secondary particles can emit Cherenkov radiation, which occurs when charged particles travel faster than the speed of light in the medium they are currently traversing [6]. This radiation can then be measured for instance by highly sensitive photomultipliers.

The “Astronomy with a Neutrino Telescope and Abyss environmental RESearch” project (ANTARES, described in [7]) is a Cherenkov radiation based neutrino telescope, located in the Mediterranean Sea, 40 km off the French coast. Other collaborations, for instance BAIKAL [8] and IceCube [9], are also doing research in the promising field of neutrino astronomy, with ANTARES being the largest operational telescope in the Northern Hemisphere.

For the success of every current, complex experiment sophisticated simulation and data analysis software is absolutely essential. This thesis illustrates improvements of several steps in the data evaluation chain of ANTARES. These improvements are not limited to this experiment as only minor details are application specific, but the ideas behind them are general and can be applied to a multitude of scenarios.

The first improvement is an enhanced classification to separate neutrino signals from undesired background. The second presented idea is to efficiently combine multiple independent reconstruction algorithms using a classification algorithm to obtain a more accurate reconstruction result. The main part of this thesis is a new strategy to detect structures of arbitrary position, size, shape and neutrino distribution in the observed neutrino signal which are incompatible with a random background assumption. This method does not require a theoretical model of the neutrino emission.

Chapters 2, 3 and 4 sketch an overview over neutrinos, neutrino detection with ANTARES and pattern recognition, covering properties of neutrinos, possible scenarios for neutrino emission, the architecture of the neutrino telescope itself, its data taking and the software used for data processing and simulations.

The new methods that have been developed for this thesis are explained in detail in chapters 5, 6 and 8.

The results of the created data processing chain and their explanations are presented in chapter 9, while the interpretation of these results can be found in chapter 10.

2 Neutrinos

The existence of neutrinos has first been postulated by Wolfgang Pauli in 1930 to explain the continuous energy distribution of electrons in beta decays, see [10]. The first proof of their existence was found in 1956 with the Cowan-Reines experiment, see [11], when electron-antineutrinos produced by a nuclear reactor were observed by inverse beta decay. Since then many other experiments have investigated the properties of neutrinos. They are electrically neutral and therefore only interact via gravity and the weak force. This is also the reason why they are so hard to detect, as they do not interact with magnetic fields, not by strong force interactions and, similar to to photons, hardly by gravity, see [12]. Neutrinos are fermions as they have a half-integer spin. There are three known types of neutrinos, the so called flavors. Together with the three leptons e, μ and τ, the three neutrinos (and the respective antiparticles of each particle) form the leptons, a group of light elementary particles. An overview over the standard model of particle physics with leptons in green is shown in Figure 3. Each neutrino flavor is named after the lepton the neutrino produces in charged current interactions. Maybe the most remarkable characteristic of neutrinos is that they oscillate between the different flavor states, alternating between electron, muon and tau flavor. Therefore neutrinos can be expected to have a mass greater than zero, since from theory it is known that oscillation between the neutrino flavors isn't possible for massless neutrinos, see for instance [13], [14], [15] and [16]. Nevertheless their exact masses are still unknown. The mass of the electron-neutrino ν_e for instance is known to be below 2 eV from measurements of the beta decay of tritium [17]. The standard model of cosmology together with the interpretations of different observations give upper limits for the sum of the masses of all neutrino flavors of below 1 eV [18] or even below 0.23 eV [19].

Since no difference between neutrinos and antineutrinos will be relevant for any part of this thesis, antineutrinos will not be addressed separately, but are included when referring to neutrinos from here on.

2.1 Sources

Neutrinos of different energy scales are produced by different objects and processes. The first observed electron-(anti)neutrinos from radioactive decays in nuclear reactors for instance had energies of a few MeV. A well established extraterrestrial source of neutrinos is the sun [21], which is also emitting neutrinos with energies up to few MeV [22]. Until now the only identification of a neutrino source outside of our solar system has been possible during the outburst of the supernova 1987A. It occurred on 23. February 1987 in the Large Magellanic Cloud, a nearby galaxy. Three neutrino experiments that have been active at that time, Kamiokande II [23], IMB [24] and Baksan [25] in total observed 25 neutrinos from the supernova, also with energies of few MeV. But there are neutrino sources that reach orders of magnitude higher energies.

As already mentioned in chapter 1, Victor Hess discovered cosmic rays in 1912. On further investigation it was discovered that this flux of charged par-

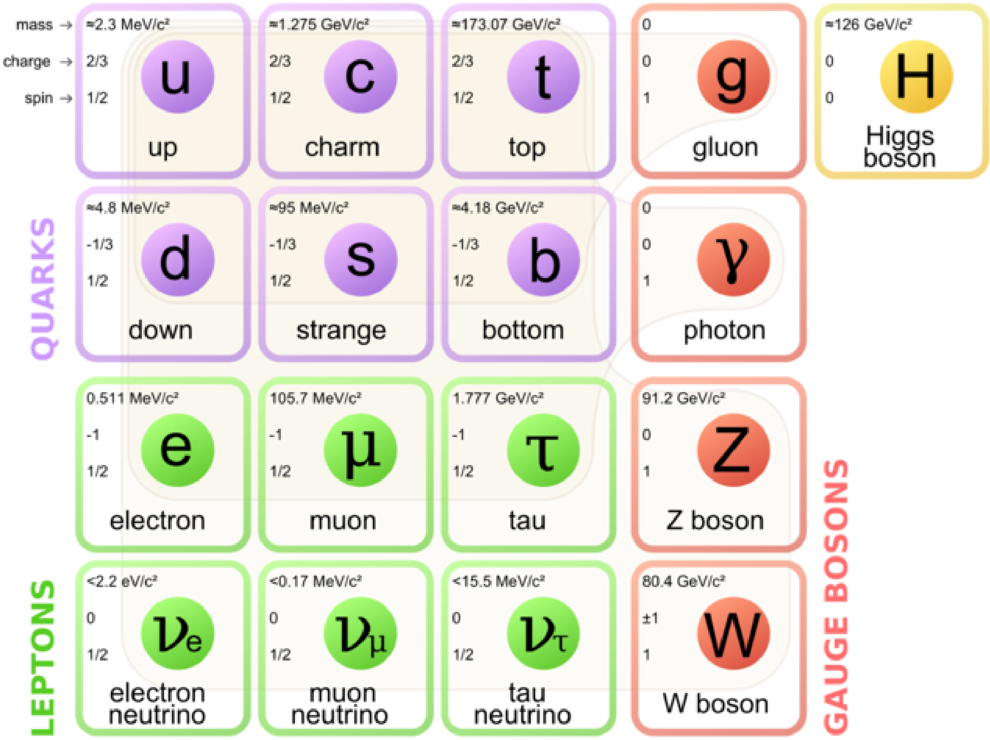


Figure 3: The standard model of particle physics. Taken from [20].

ticles extends to very high energies as seen in Figure 2. But even today the sources of the highest energetic cosmic rays are still unclear. There are models for a variety of galactic and extragalactic objects that predict the acceleration of hadrons to highest energies, making these objects potential candidates for sources of high energy cosmic rays. If these theories are true, the same hadronic processes must also generate very high energy neutrinos. Once an accelerated proton produces a hadronic shower, the resulting charged pions can produce neutrinos as described in equation 1 (Branching ratio >99.9%) followed by equation 2. Charged kaons which are also produced in hadronic showers can produce neutrinos either by decaying to charged pions (28%), resulting again in equation 1, or directly to muons (64%), resulting in equation 2.

$$\begin{aligned} \pi^- &\rightarrow \mu^- + \bar{\nu}_\mu \\ \pi^+ &\rightarrow \mu^+ + \nu_\mu \end{aligned} \tag{1}$$

$$\begin{aligned} \mu^- &\rightarrow e^- + \nu_\mu + \bar{\nu}_e \\ \mu^+ &\rightarrow e^+ + \bar{\nu}_\mu + \nu_e \end{aligned} \tag{2}$$

The existence of such a high energy neutrino flux has recently been confirmed by the IceCube neutrino telescope, published in [26] and [27], but it has not been possible to identify a source of the observed neutrinos.

In our galaxy there are several candidates for sources of high energetic neutrinos. A promising candidate would for instance be supernova remnants (SNRs). The shock front generated by the immense release of energy during a super nova explosion propagates into the surrounding interstellar medium. By repeated scattering in front of and behind this shock front, particles from the interstellar medium can be accelerated to high energies. This acceleration process is known as Fermi-acceleration. It is explained in detail in [28]. Probably the most famous SNR, the crab nebula, is shown in Figure 4.

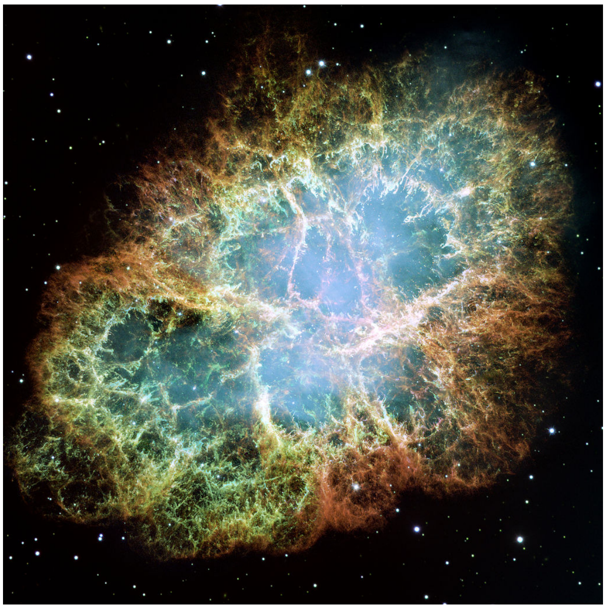


Figure 4: Supernova remnants like the crab nebula could be a source of cosmic neutrinos. Taken from [29].

Apart from SNRs, also pulsars could be possible sources of high energetic neutrinos. Pulsars are fast rotating neutron stars which have strong magnetic fields, allowing them to accelerate charged particles powerfully.

Another possibility for galactic neutrino sources could be the jets of so called micro-quasars, stellar black holes of a few solar masses accreting matter for instance from a companion star.

The Fermi Bubbles are large extended regions in our galaxy that reach 50° above and below the Galactic Center, with a width of up to 40°. They were first described in 2010 in [30]. As their origin is not completely clear, the detection of a neutrino flux originating from these bubbles would help clarify their formation. A possible scenario could for instance be a jet produced at the Galactic Center, which could also accelerate hadrons and thus produce neutrinos.

On the other hand models of these potential galactic sources show that these phenomena can't produce the most energetic observed cosmic rays, as explained for instance in [31]. Therefore even more powerful mechanisms must be involved. As there are no such mechanisms known within our galaxy and the gyroradius of these particles would already be larger than the extension of our galaxy, the highest energetic particles are believed to be of extragalactic origin.

One of the most promising candidates would for instance be active galactic nuclei (AGNs), see [32]. AGNs have a massive black hole in the center of the host galaxy. When matter is accreted and consumed by the black hole, jets are formed orthogonal to the plane of the accretion disk. These jets could be an excellent method to accelerate particles to highest energies. An image of such an AGN can be seen in Figure 5.

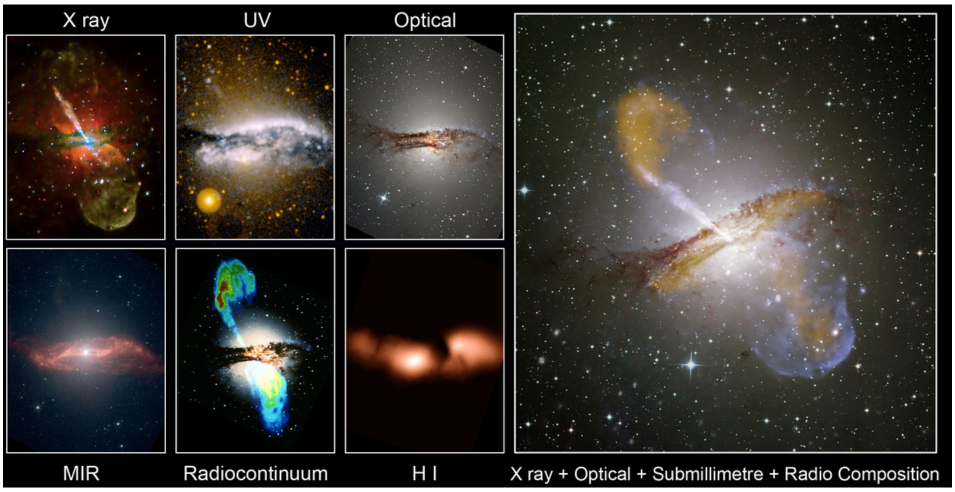


Figure 5: The AGN Centaurus A as overlay of multiple wavelength measurements. Taken from [33].

Another interesting possible source could be gamma-ray bursts (GRBs). They are short but very intense outbursts of highly energetic γ-radiation. As they are distributed isotropically over the sky, it was realized quickly that they must be of extragalactic origin, see for instance [34]. The typical timescale of GRBs is rather short, with two different timescales being observed. Short GRBs have a lifetime of below 2 seconds, whereas typical long GRBs last up to minutes. After both types a longer afterglow can be observed. Possible explanations for GRBs include for instance the collapse of a massive star to a neutron star or a black hole. More insight on GRBs can be found for instance in [35] or [36].

There are several other extragalactic objects which might contribute to the production of high energetic neutrinos. A more complete list is given for instance in [31]. But besides the already suspected candidates, of course there is also the possibility that the involved mechanisms are still unknown or not fully understood yet.

2.2 Propagation

Once a neutrino is generated at its source, it is hardly influenced during its propagation through space. As they are electrically neutral, they cannot be deflected for instance by galactic or extragalactic magnetic fields along their path. As explained in [12], they are not influenced by gravitational effects significantly more than photons due to their small mass and high velocity. This makes them ideal candidates for astronomy, since, unlike photons, they can traverse dense matter. Cosmic rays on the other hand are charged particles and therefore, except for extremely energetic ones, are deflected by magnetic fields along their way, obscuring their origin.

Since neutrinos oscillate between the three flavor states and the typical traveling distance is far longer than the distances required for oscillation to occur at the considered energies, the expected flavor ratio of cosmic neutrinos on earth is 1:1:1, regardless of the original flavor(s) that have been produced at the source, see for instance [37].

2.3 Detection

The biggest advantage of neutrinos for astronomy, their low probability to interact between the source and the Earth, becomes the most relevant disadvantage when trying to detect them with a neutrino telescope. A direct detection of neutrinos is beyond the capabilities of todays technology. Therefore the detection of neutrinos has to rely on the secondary particles created when a neutrino interacts via the weak force with a nucleus of the matter it traverses. The deviation between the direction of the neutrino and of the produced secondary particle is small at the energies relevant for this thesis as shown in [38], page 43. For example the median angular difference betweeen the direction of a muon-neutrino with an energy of 1 TeV and the muon resulting from a CC interaction is below 1°.

There are two possibilities for interactions between neutrinos and matter. The interaction can occur as a neutral current (NC) interaction with a Z_0 boson as a mediator, shown in Figure 6, or as a charged current (CC) interaction, with a W^+ or W^- as mediator and the charged lepton corresponding to the flavor of the neutrino as resulting secondary particle, shown in Figure 7.

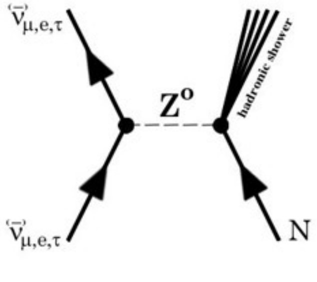


Figure 6: Feynman graph for the deep inelastic neutral current neutrino interactions of all flavors. From [39].

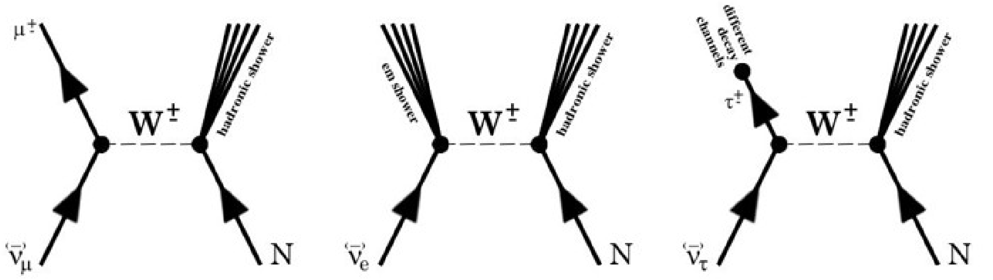


Figure 7: Feynman graphs for the deep inelastic charged current neutrino interactions. From [39].

If the interaction occurs in a dielectric medium and the produced charged secondary particles move faster than the speed of light in this medium the so called Cherenkov radiation, described for instance in [6] and [40], is emitted at a characteristic angle Θ_E.

$$\Theta_E = \arccos \frac{1}{n\beta} \tag{3}$$

Equation 3 shows the dependence of Θ_E on the refractive index of the medium n and the velocity v of the traversing particle with $\beta = v/c$. c denotes the speed of light in vacuum. For the considered energies and seawater this angle is about 42°. A schematic of the emission of Cherenkov light is shown in Figure 8. This light can be measured to detect the neutrino and to reconstruct its properties.

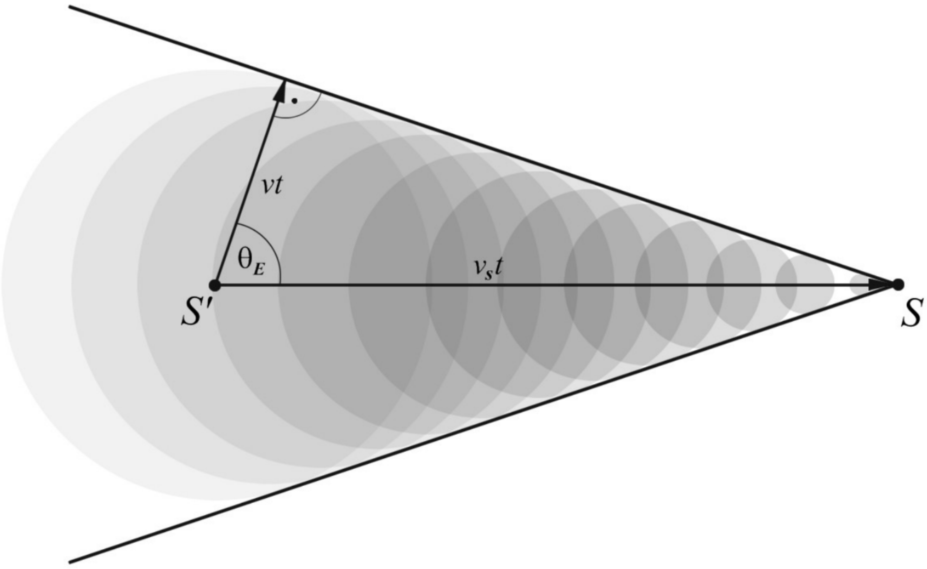


Figure 8: Schematic of the emission of Cherenkov radiation along the track of a particle. S' marks the starting point where the interaction took place, S is the current position of the particle. Due to the movement of the particle, which is faster than the speed of the light in the medium, a cone shaped light front is formed. Image taken from [39].

3 ANTARES

As explained in chapter 2.3, secondary particles produced by neutrino interactions emit Cherenkov radiation, which can be detected. Photomultiplier tubes (PMTs) can be used to measure these photons. Since the probability that a neutrino interacts close to a PMT is low, a large volume covered by a high number of PMTs is favorable. Of course this volume has to be filled with an optically transparent medium to allow the photons to propagate to the PMTs. Since a neutrino interaction only produces few photons compared to a visible light source, as few other light sources as possible should be interfering with the measurement, making the deep sea or deep in the antarctic ice favorable options for the location of a neutrino telescope. From the position, time and amplitude information measured by the PMTs, the direction of the secondary particle and therefore of the neutrino can be reconstructed and its energy can be estimated . All modern neutrino telescopes follow this general scheme. At the time this thesis is written, ANTARES ("Astronomy with a Neutrino Telescope and Abyss environmental RESearch") is the largest operating neutrino telescope in the Northern Hemisphere. It is located in the Mediterranean Sea[1], 40 km from the French coast near Toulon at a depth of 2475 meters to partly shield it against muons produced in the atmosphere. It uses the water of the deep sea as optical medium for neutrino detection.

The international ANTARES collaboration consists of 32 institutes from France, Italy, the Netherlands, Germany, Spain, Russia and Morocco.

3.1 Detector layout

The PMTs measuring the Cherenkov radiation in ANTARES have a diameter of 25.4 cm. They are enclosed in glass spheres together with some necessary electronics. These spheres are called optical modules (OMs). Three OMs and some additional electronics are grouped together in so called "storeys". The OMs of each storey have a spacing of 120° between them to optimally cover the whole surrounding. The structure of a storey can be seen in Figure 9. Each OM is oriented downwards by 45° because about 10^6 times more muons generated by cosmic ray interactions in the atmosphere than from neutrino interactions reach the telescope from above, resulting in a large background. On the other hand, only neutrinos reach the telescope from below, as they are able to traverse the earth. Therefore ANTARES has been optimized for upgoing muons. A schematic visualizing the possible paths of muons that can reach ANTARES is shown in Figure 10.

Twenty-five storeys[2] are attached to each line with a distance of 14.5 m between the storeys. The layout of ANTARES contains 12 of these lines that are anchored at the sea bottom, each held upright by the buoyancy forces of a buoy attached at the top of the line and of the OMs along the line. The lines have a mean distance of about 70 meters between them. The footprint of these lines can be seen in Figure 11. All together this setup results in an extension of

---

[1] 42°48'N, 6°10'E

[2] One line has only 20 instead of 25 storeys

Figure 9: An ANTARES storey with the three OMs. Taken from [41].

about 210 by 210 by 350 meters and an instrumented volume of about $0.01 km^3$. An artists impression of the overall structure can be seen in Figure 12.

Since the whole setup is exposed to sea currents, the lines are not always aligned exactly vertically. This effect is compensated by a calibration of the telescope geometry. The measurement of the current geometry is performed every two minutes using acoustic emitters and hydrophones located at every fifth storey. With the help of other techniques, these measurements allow position calibrations by triangulation with a precision of $<$ 10 cm, see [44]. The obtainable time resolution, also achieved by sophisticated calibrations, is at a level of 1 ns.

Many more details about the components, structure, calibration and operation of ANTARES can be found for instance in [7].

3.2 Event processing

ANTARES follows the “all data to shore” concept, meaning that all measurements above a certain low threshold are sent to a computer farm at a shore station via optical cables. The measurements are then called a “hit”. Every hit has a timestamp assigned by the electronics. Therefore the hits generated by the same physical event at similar times can be grouped together by the data processing chain to be recognized as one ”event”, sometimes also referenced as a “frame”.

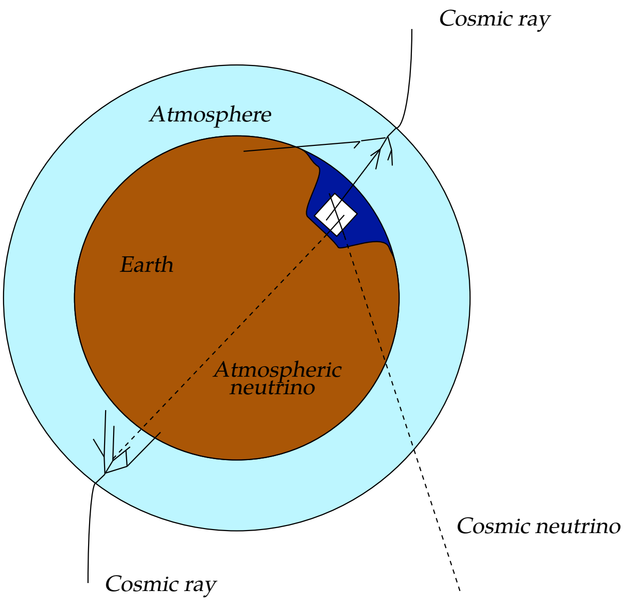


Figure 10: Sources of muons in ANTARES. Atmospheric muons from cosmic ray interactions in the atmosphere reach the detector from above, but only from interactions of muon-neutrinos can muons reach it from below. Image taken from [42].

To filter out events that contain only noise, caused for instance by bioluminescence, algorithms analyze the incoming data before they are stored. There are many ideas about how these so called triggering algorithms or “triggers” can search for potentially interesting events. In ANTARES they rely on a preselection of the measured hits, where only hits above a threshold or in coincidence with other hits on the same storey are analyzed further. The trigger algorithms then analyze these preselected, so called L1 hits. If any of the simultaneously active trigger algorithms identifies an event as potentially interesting, it is written to disk together with the hits contained in a time window of ± 2200 ns before and after the first and last triggered L1 hit. In this thesis two trigger algorithms are used for data selection.

- 3N trigger:
  It searches for L1 hits that fulfill equation 4:

  $$|\Delta t| < \frac{|\Delta x|}{c_{medium}} \tag{4}$$

  with Δt being the time difference between two hits, Δx the distance between the position of these two hits and c_{medium} the speed of light in seawater.

  If a cluster of at least five connected hits that fulfill this criterion is found, the hits are checked if they are compatible with a muon trajectory. If this is the case, the event is triggered by the 3N trigger.

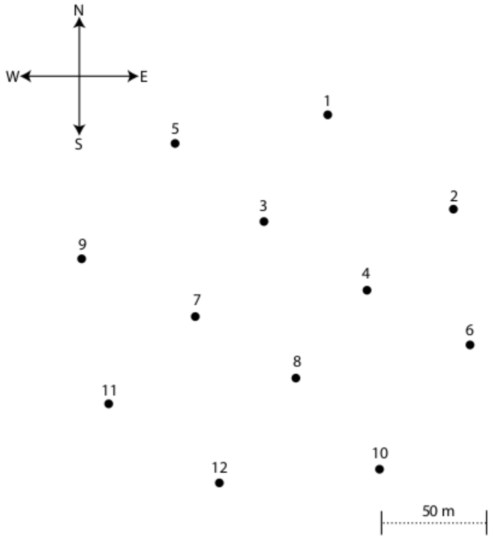


Figure 11: Footprint of the lines in the ANTARES telescope. Taken from [42].

- T3 trigger:
  At least two of the L1 hits are detected on neighboring storeys within 100 ns, or on next to neighboring storeys within 200 ns. If an event contains two of these coincidences, the T3 trigger keeps the event.

More information on the trigger algorithms can be found for instance in [45], [46] and [47].

The data taking in ANTARES is organized in so called “runs”. A run typically lasts several hours and is intended to be a data taking period of similar conditions. The runs are identified using a runnumber, which is a six digit number. Until today for every run the first digit of the runnumber is zero. This leading zero will be omitted when listing runnumbers in this thesis.

Since one storey doesn’t provide enough information to reconstruct an event decently, but low-energetic events often do not produce enough light to be observed by multiple storeys, a neutrino telescope has a lower limit for the energy of reconstructable events depending on the geometry, the data taking conditions and the applied reconstructions. For ANTARES this lower limit lies somewhere around 10 GeV, depending on the exact evaluation methods.

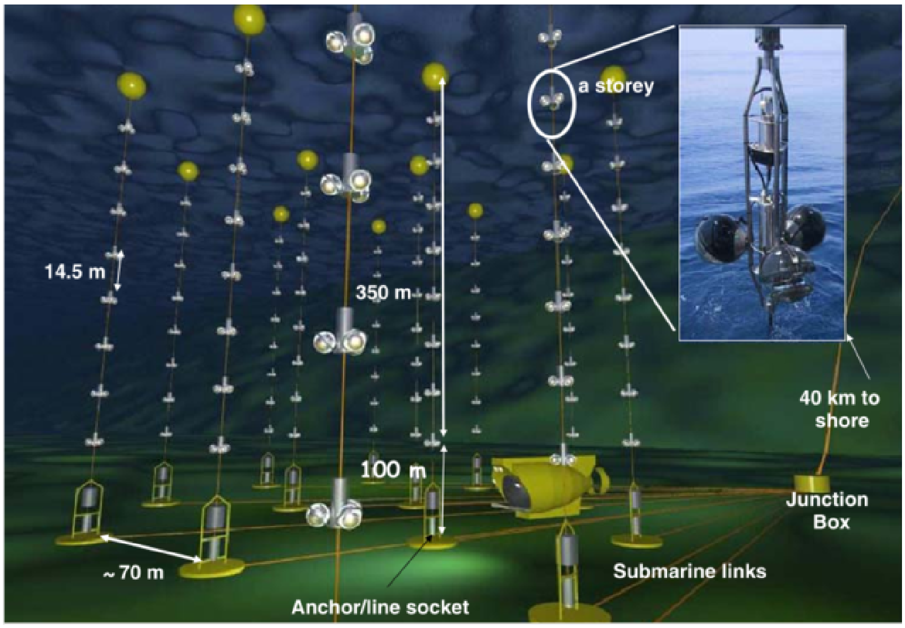


Figure 12: Overview over the ANTARES neutrino telescope. Based on a depiction from [43].

3.3 Event signatures

With the achievable resolution of ANTARES one cannot distinguish all possible types of neutrino interactions from each other. With three neutrino flavors, two possible interactions per flavor and one antiparticle per particle, in principle there are twelve different interactions. Since there is no way to measure the charge of the generated particles, the interactions of particles and antiparticles look alike. From the remaining six interactions all neural current (NC) interactions produce indistinguishable hadronic cascades, also called hadronic showers. This is indicated in the overview in Figure 13, with only four different signatures remaining. For energies relevant in ANTARES the signature of all particle cascades is almost pointlike. This can be seen in Figure 14, where the traveling distances for showerlike events are short, keeping in mind that the distance between two lines is about 70 meters in ANTARES. To resolve those structures a much denser detector is necessary, like for instance Super-Kamiokande, see [48], but these detectors do not yet cover such large volumes. With the dense instrumentation of future projects like ORCA, described in [49], it will be possible to resolve such small structures even for large volumes.

The charged current (CC) interactions of neutrinos produce the lepton corresponding to the flavor of the neutrino, accompanied by a hadronic shower. The electron created by ν_e CC interactions produces an electromagnetic cascade with a pointlike signature that is so similar to the ones generated by hadronic cascades that these cannot be distinguished with ANTARES for individual events. Therefore the interactions c) and d) in Figure 13 can be considered to have the same signature.

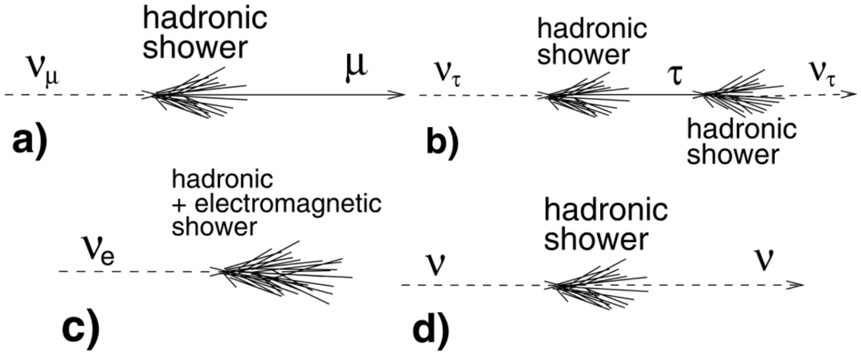


Figure 13: Possible event signatures for a neutrino telescope. Taken from [50].

In contrast to that, muons created by ν_μ CC interactions travel considerable distances depending on their energy as seen in Figure 14. With a typical energy loss, the track length of muons between 100 GeV and 1 TeV can already surpass the instrumented volume of ANTARES.

Since they emit light along their path, their signature is elongated and tracklike. Therefore muon-neutrinos can be distinguished from the showerlike event signatures. The signature of ν_τ is highly energy dependent. For energies below 100 TeV these events look pointlike since the τ has a lifetime of $(290.3 \pm 0.5) \cdot 10^{-15}$ s (see [17]) and therefore decays quickly, producing a second cascade within or next to the first hadronic one, making it similar to the ν_e signature. The expected traveling distance increases rapidly with energy at about 50 m/PeV, theoretically resulting in a tracklike event signature for highest energies. In between these two cases, a so called “double bang” signature is expected. Due to the size of ANTARES, it is unlikely that both cascades can be observed within the detector volume and can be distinguished from an ordinary track event with an additional energy loss due to a shower along the track. Furthermore, a nu_τ can also result in a track if the produced τ decays to a μ, which occurs in about 17.4% of all events, see [17].

This leaves us with two distinguishable signature types, the pointlike shower events and the elongated tracklike events as seen in Figure 15.

Due to its elongation the tracklike signature contains more information about the direction and therefore also allows a more precise direction reconstruction. The small size and the symmetrical shape of pointlike cascades render a direction reconstruction especially tricky. Although there has been substantial progress on this topic in recent years, see for instance [39], no competitive level is reached yet. Therefore this analysis uses tracklike signatures only.

3.4 Backgrounds

The desired signals are highly relativistic upgoing muons generated by cosmic muon-neutrinos. Besides a background level of about 3 kHz per PMT by elec-

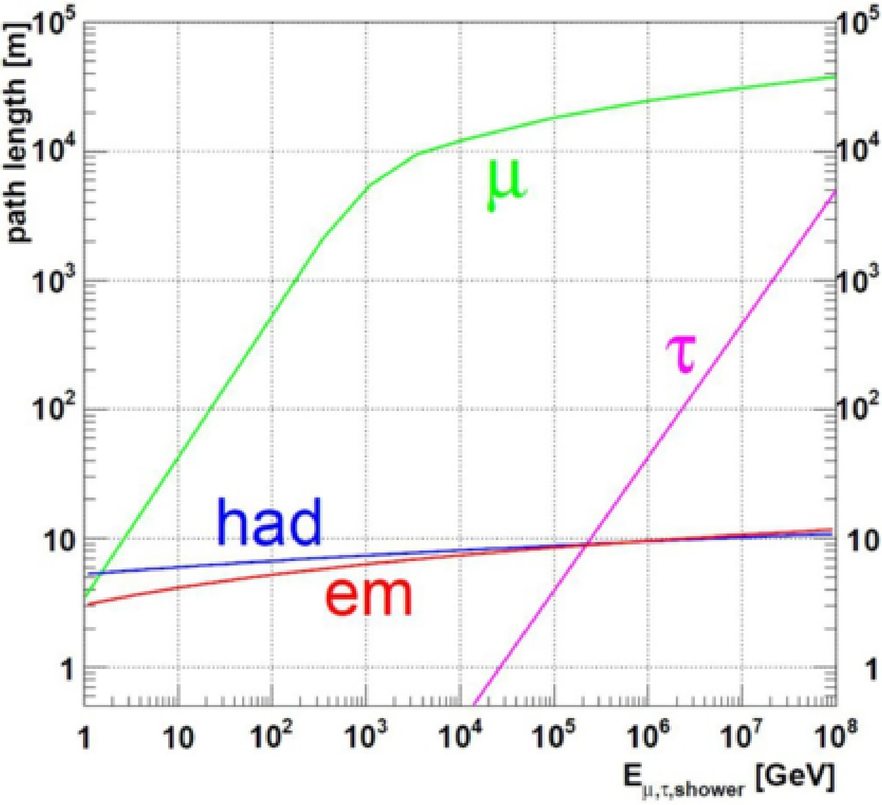


Figure 14: Travelling distance for different events in water as a function of energy. Taken from [50].

tronics [42], an omnipresent background for this measurement in ANTARES are photons emitted by the radioactive decay of ${}^{40}_{19}\mathrm{K}$. This isotope is contained in natural sea water and, by the decays explained in equations 5 to 7, produces electrons with energies above the threshold for Cherenkov radiation.

$$ {}^{40}_{19}\mathrm{K} \rightarrow {}^{40}_{20}\mathrm{Ca} + e^- + \bar{\nu}_e \qquad (89\%) \tag{5}$$

$$ {}^{40}_{19}\mathrm{K} + e^- \rightarrow {}^{40}_{18}\mathrm{Ar} + \nu_e \qquad (11\%) \tag{6}$$

$$ {}^{40}_{19}\mathrm{K} \rightarrow {}^{40}_{18}\mathrm{Ar} + e^+ + \nu_e \qquad (0.001\%) \tag{7}$$

This decay produces rates of about 37 kHz per PMT. As it can only be detected in the direct vicinity of OMs, this process generates uncorrelated background patterns and is therefore suppressed by the requirement of correlations on more than one storey.

Another omnipresent source of background is bioluminescence caused by multiple species of microorganisms as well as larger animals in the deep sea. An example for one of the larger culprits can be seen on the left in Figure 16. While the amount of light produced by the decay of ${}^{40}_{19}\mathrm{K}$ is very steady, the

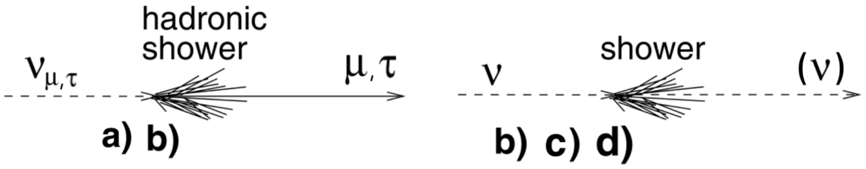


Figure 15: Actually distinguishable event signatures for a ANTARES. Modified from [50].

amount of bioluminescence is highly variable and influenced by different factors for instance the sea current, as shown in Figure 17. In some cases the light pattern caused by bioluminescence range over multiple storeys, but the pattern and the timing is clearly different from what is expected for tracklike event signatures. More information on the involved bioluminescence phenomena can be found in [51].

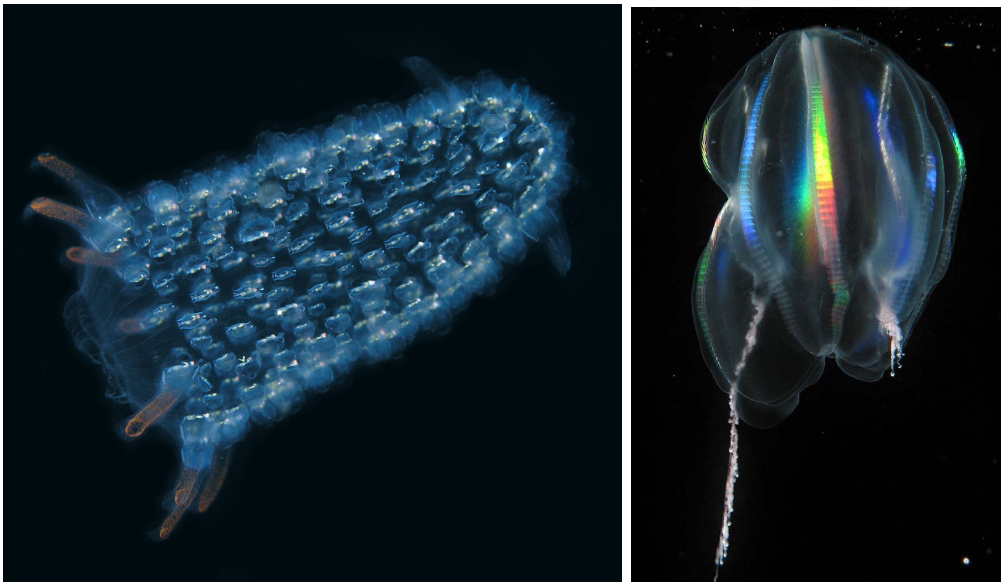


Figure 16: Examples for bioluminescence in macroscopic animals, here a comb jelly (left, from [52]) and an arctic comb jelly (right, from [53]).

Muons from interactions of cosmic rays in the atmosphere produce the largest amount of triggered events, as they are orders of magnitude more frequent than muons from neutrino events, see Figure 18. This problem may seem like it can be solved entirely by only considering upgoing muon tracks with a zenith angle[3] $\Theta_Z > 90°$ (or $\cos(\Theta_Z) < 0$ in Figure 18).

Unfortunately there is a small chance that the hit pattern produced by downgoing muons looks similar to an upgoing muon-neutrino event and there-

---

[3] Chapter 3.7 gives a brief introduction of the coordinate systems.

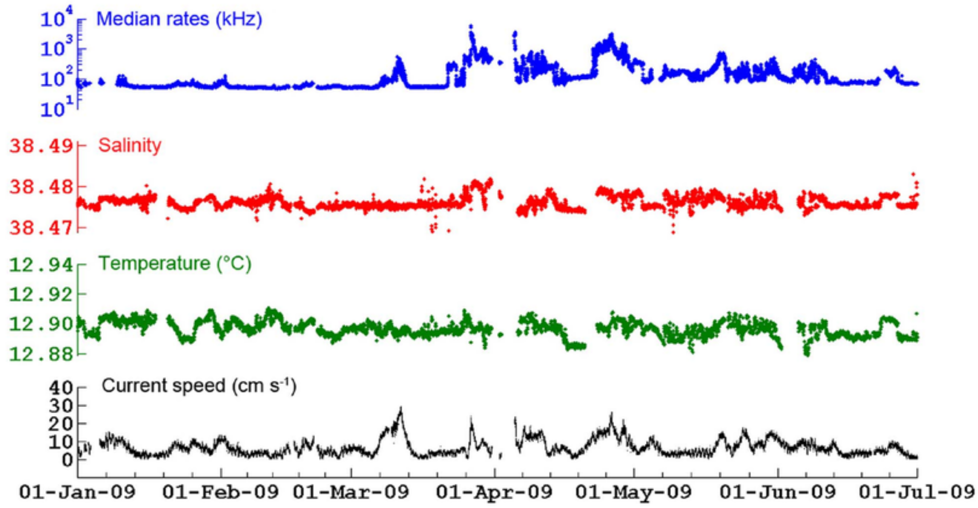


Figure 17: Environment measurements from the deep sea and background rates in ANTARES. There is a correlation observed between the optical rate in blue and the sea current in black. Taken from [51].

fore they happen to be misreconstructed by the reconstruction algorithms. Although the probability for this is tiny, the shear number of those events makes misreconstructed downgoing muons one of the main backgrounds for data analysis. Two of the approaches that are explained later, especially in chapter 5, deal with this problem. Depending on the accuracy of the filtering, the remaining sample contains mostly upgoing muon-neutrino events.

This selection is dominated by upgoing neutrinos from interactions of cosmic rays in the atmosphere on the other side of the Earth, which constitute an irreducible background for the search of cosmic neutrinos. As a first order approximation, these remaining background events are distributed isotropically in each direction. While deviations from this isotropy are known, overall these are too small to be relevant in this thesis. Furthermore, in the evaluation method described in chapter 8, the numbers are derived from recorded data and hence, to a certain extent, can take potential anisotropies into account as explained in chapter 9.7. An analytical model for the flux of neutrinos can be found for instance in [55] and [56].

Since the number of events expected from interactions in the atmosphere is considerably larger than the number expected from any cosmic source, an evaluation is required that recognizes the cosmic origin of neutrinos, for example based on features like their energy or spatial distribution. This task is the main focus of this thesis and will be addressed from chapter 8 onwards.

3.5 Software

The official software framework of ANTARES is called Seatray, described e.g. in [57] and [58]. It is derived from the software framework Icetray, which is

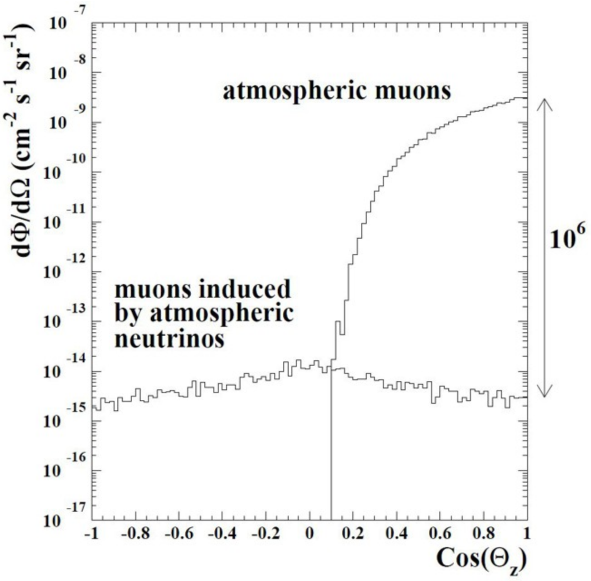


Figure 18: Angular distribution of the flux of muons above 1 TeV at the ANTARES location and depth. Atmospheric muons dominate the downwards direction almost up to the horizon. Taken from [54].

used by the IceCube collaboration. The framework is composed of several core components which provide the necessary basic functionality and multiple so called modules. These modules can be developed completely independent from each other by different users, each for an own analysis or other purpose. Since all modules are required to implement the interface defined by the Seatray core, they can be shared and combined at will. The core is written in C++, while users are free to write their modules either in C++ or in Python.

The feature extraction and classification used in this thesis is contained in a module called antares-rdfclassify. The functionality which is behind this module is also available as a stand-alone version called SGClassify. The multiscale source search is contained in a module called signal-first, which has been the name of this search during development. These modules can be found in the svn version control system of Seatray.

3.6 Simulations

To be able to evaluate the performance of data analysis algorithms in detail one would need a large sample of events for which all relevant information (Particle type, interaction position, energy, direction, for some evaluations even the produced secondary particles, ...) has to be known. Since it is virtually impossible to obtain this without any errors for recorded events, precise simulations have to be conducted, generating events where this information is available. This is achieved by stepwise Monte Carlo simulation (MC) of all involved processes. More detailed information on this topic than what is presented here can be found in [42].

1. Event generation
   A tool called GENHEN is used to generate neutrino events according to the expected flux. Documentation on the internals of GENHEN can be found in [59] and [60]. The outcome contains the trajectory of the primary particle (the neutrino) as well as long lived secondary particles (e.g. muons). The energy distribution of the events is expected to follow a power law with a spectral index of approximately -2.0 for cosmic neutrinos. As the simulation of these energy distributions would result in little available statistics for events with high energies, a harder energy spectrum is simulated and the events are assigned a numerical weight to be able to reweight each event according to any desired flux afterwards. The details of this weighting process are described for instance in [61] or [42].

   The background of atmospheric muons with a spectral index of -3.7 is simulated using a tool called MUPAGE, see for instance [62]. The propagation of the particles to the telescope is simulated by a code called MUSIC [63].

2. Photon emission and propagation
   The Cherenkov light emitted by all particles generated by the previous step is calculated by a program called KM3. As the tracking of each generated photon would be too intense computationally, a lookup table is used to model the generation, propagation and scattering of the photons between the track of the particle and the PMTs detecting them. This lookup table is based on fullscale simulations conducted with GEANT.

   In addition to that, there is light produced by particle showers which occur along the tracks of the high energetic particles. To simulate these particle showers a tool called GEASIM is used. Internally it relies on GEANT for the simulation of the generated light, too.

3. The detector response
   The outcome of the previous step are all photons that arrive at the detector. The next step is to simulate the behavior of ANTARES. The tool used for this task is called TriggerEfficiency. Information about the internals of this tool can be found in [64]. It simulates the response of the PMTs to the incoming photons and the effect of the currently active calibration on the electronics. The detected simulated hits are then processed the same way as measured hits as described in chapter 3.2.

To obtain simulations that better reflect the actual data taking and detector conditions, the simulation chain has been adapted to include the exact calibrations of the ANTARES neutrino telescope which were actually used during the simulated data taking periods. This also includes the duration of individual runs, conditions like e.g. the bioluminescence rate and the software setups. These individualized simulations are called run-by-run simulations (RbRs). There are two different versions of these simulations used in this thesis. The difference between RbR 2.2 and RbR 3.0 are the included versions of the used simulation tools. This thesis is written during the transition phase from version 2.2 to version 3.0. While in principle RbR 3.0 offers the best available simulation, RbR 2.2 are still used whenever there were not enough RbR 3.0 simulations available to guarantee a reliable evaluation.

3.7 Coordinate systems

Three different coordinate systems are especially useful for the understanding of the argumentations and results presented in this thesis.

First of all, there is the local coordinate system of ANTARES. It is defined by two angles, zenith Θ and azimuth Φ. The zenith angle ranges from 0 to 180 degrees. A zenith angle of 0° denotes a downgoing direction, directly from above ANTARES. A zenith of 90° is a horizontal direction and a zenith angle of 180° corresponds to a straight upgoing direction from below the telescope. The azimuth ranges from 0° up to 360° and specifies the rotation around the z-axis. An azimuth of 0° denotes the easting direction, 90° northing. Event reconstructions for example give their results in this coordinate system. Figure 19 shows the final event sample, obtained as explained in chapter 7, in local coordinates.

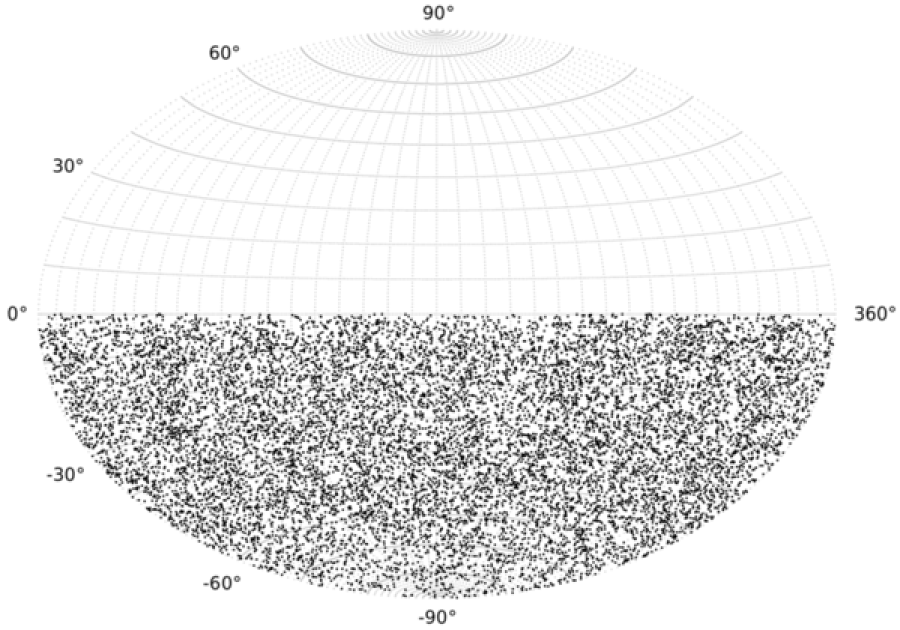


Figure 19: Distribution of the selection of neutrinos described in chapter 7 in local coordinates of ANTARES.

The second relevant coordinate system in this thesis is the equatorial coordinate system. In contrast to the local coordinate system, which revolves with the Earth, a fixed point in the sky, for instance a distant star, has fixed, unique coordinates in the equatorial coordinate system. The two angles that identify a point in this system are called declination δ (from -90° to +90°) and right ascension α (from 0° to +360°). Equatorial coordinates are constructed such that all points on a projection of the equator of the Earth on the sky have a declination of 0°. The direction straight above the north pole has a declination of +90°, straight "below" the south pole is -90°. Analogous to the azimuth in local coordinates, the right ascension defines the rotation around the z-axis. Using equatorial coordinates to map the sky has the useful property that the visibility of ANTARES is approximately constant for all regions that have the same declination, making it the coordinate system of choice for most of the evaluations in this thesis. The distribution of the final event sample in equatorial coordinates is shown in Figure 20. Instead of the right ascension in degrees, this degree of freedom is often given as the so called hour angle, in hours, minutes and seconds, ranging from 0 to 24 hours. The exact coordinates of an object slightly vary over time due to the nutation and precession of the rotation of the Earth, and over larger time scales also due to the movement of the objects themselves. Considering the resolution of current neutrino telescopes, these tiny changes aren't relevant for neutrino astronomy (yet).

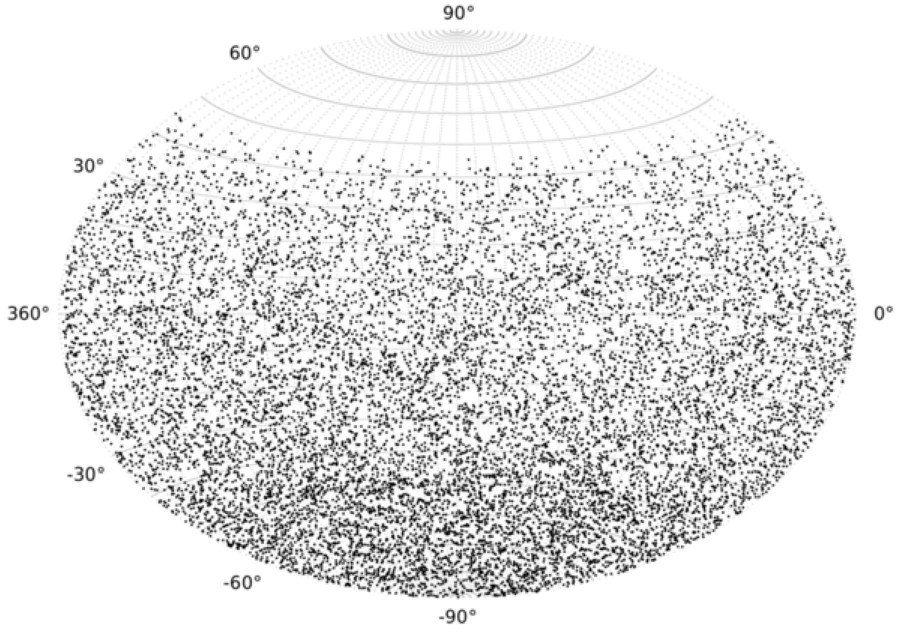


Figure 20: Distribution of the selection of neutrinos described in chapter 7 in equatorial coordinates.

The third important coordinate system is the galactic coordinate system. Similar to equatorial coordinates, a fixed object in the sky has unique fixed galactic coordinates, too. The two angles are called galactic latitude b (-90° to +90°) and longitude l (-180° to +180°). Instead of being oriented at the equatorial plane of the Earth like equatorial coordinates, galactic coordinates are based on the galactic plane, which is the plane the galactic disk revolves in. An object with a galactic latitude of 0° is located exactly in this plane, so many galactic objects have a latitude close to 0°. A positive latitude defines objects "above" the galactic plane ("above" corresponding to north on earth), negative latitudes are below the galactic plane. A galactic longitude of 0° corresponds to the direction of the galactic center, ± 180° defines the rotation to the left and to the right. A schematic of the construction of galactic coordinates can be seen in Figure 21. Figure 22 shows the night sky in galactic coordinates. One can see nicely the benefit of this coordinate system for galactic astronomy, as the plane of our galaxy separates the upper and the lower part of the skymap. The distribution of the final sample of events in galactic coordinates is shown in Figure 23. A reference for the location of some nearby galaxies in both coordinate systems can be seen in Figure 24.

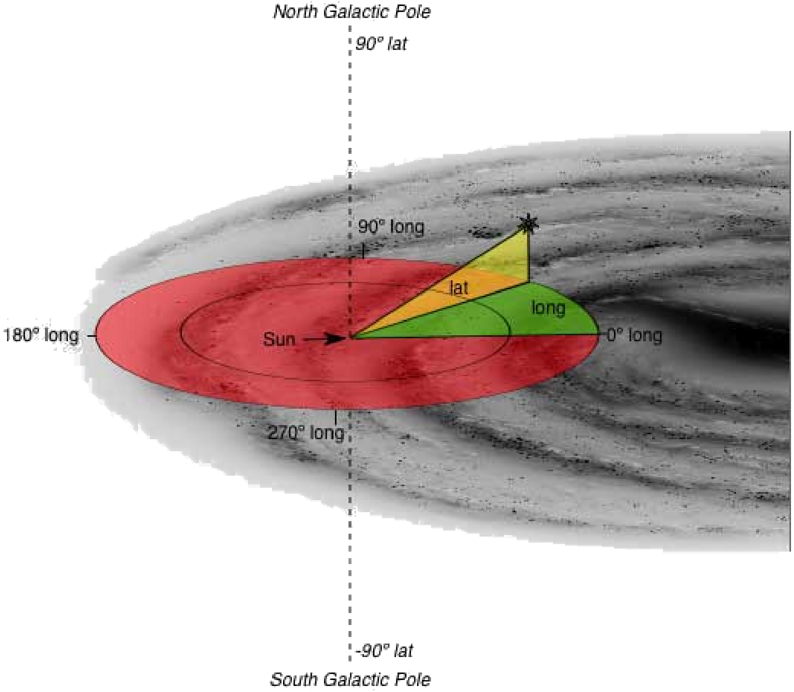
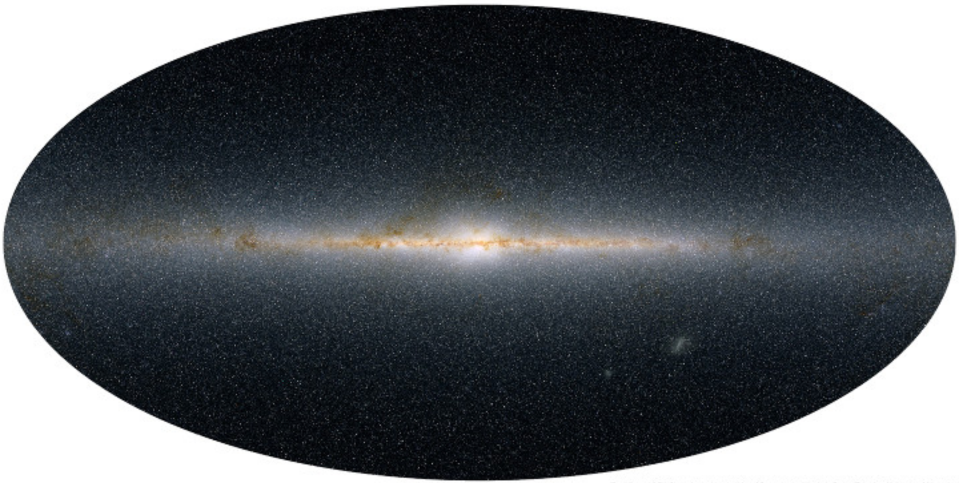


Figure 21: Concept of galactic coordinates. Modified from [65].



Two Micron All Sky Survey Image Mosaic: Infrared Processing and Analysis Center/Caltech & University of Massachusetts

Figure 22: The location of the galactic plane in galactic coordinates. Modified from [66].

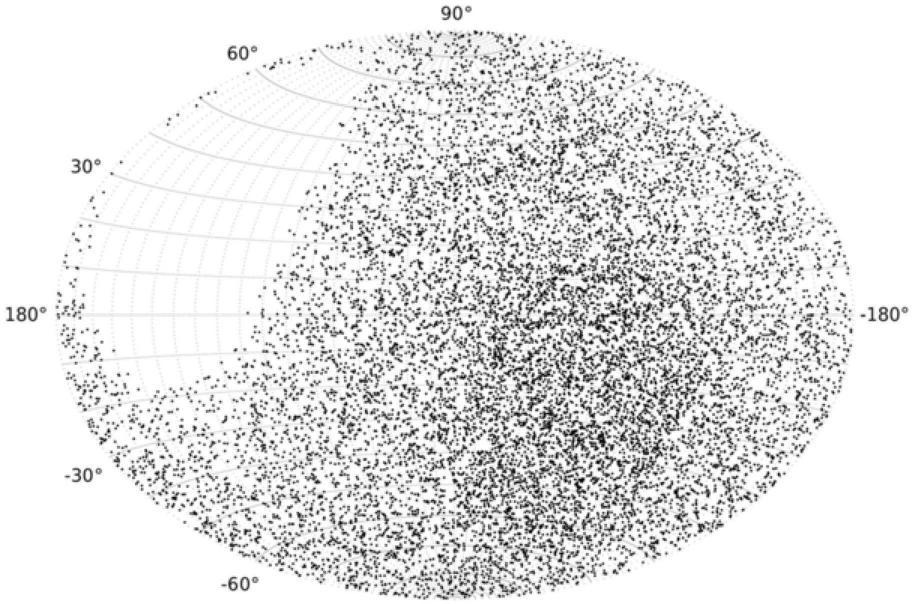
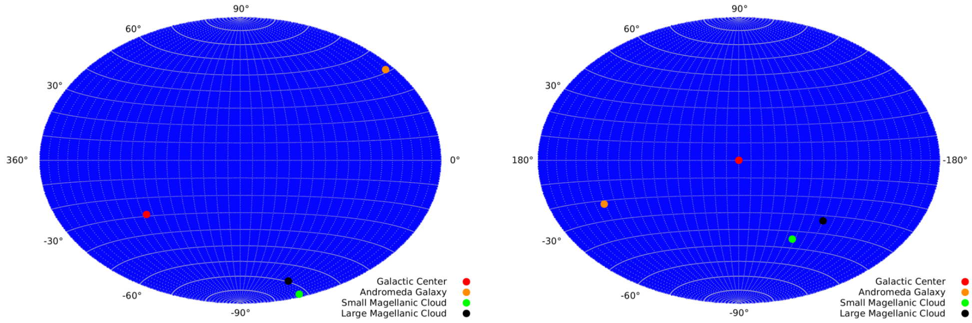


Figure 23: Distribution of the selection of neutrinos described in chapter 7 in galactic coordinates.



(a) (b)

Figure 24: Locations of some nearby galaxies. a) in equatorial and b) in galactic coordinates.

4 Pattern recognition

Pattern recognition in a nutshell is the science of finding interesting structures in data. One of the simplest approaches to perform such a recognition is to apply a threshold to a variable, where events are considered signal above the threshold and background below. Of course in many applications, especially with complex data or more differentiated goals, one can improve considerably if a more sophisticated approach is used. This chapter introduces the basics of pattern recognition which are helpful for the understanding of the solutions presented in chapters 5 and 6.

4.1 Pattern recognition overview

Classical pattern recognition usually involves the individual steps as depicted in Figure 25:

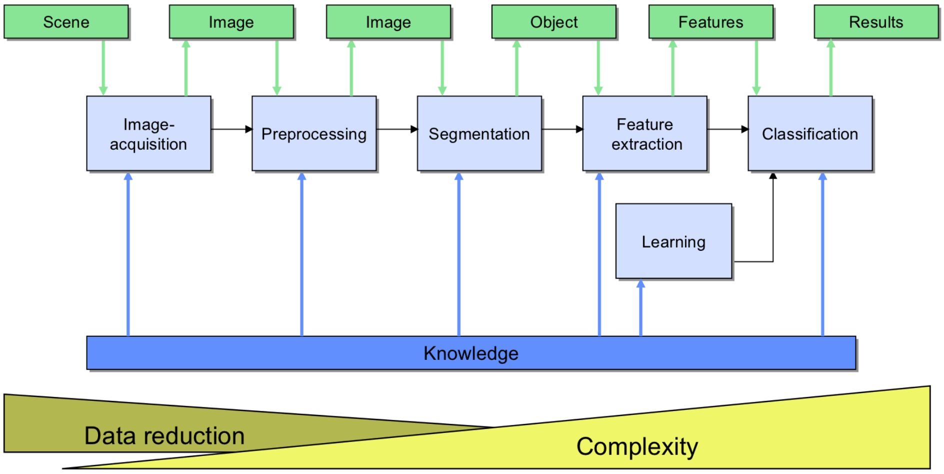


Figure 25: The pattern recognition pipeline (here for images). These are the steps commonly involved in a pattern recognition solution. Taken from [67], original from [68]. Image credit: Thomas Wittenberg

1. Data acquisition
   The first step in data analysis is the acquisition of the data. For a digital processing the data has to be sampled and quantized. Every day examples for this step are digital cameras or cell phones for images and speech. In the context of this work the PMTs and the electronics of the ANTARES neutrino telescope perform this task.

2. Preprocessing
   Measurements usually contain noise or artifacts. Often these can be reduced at least partially, for instance by the application of filters to the data. Common examples include Gaussian and median filters, normalizations or edge detections.

3. Segmentation
   Segmentation means identifying the possibly interesting parts of the data for further analysis. While this may seem like a simple task, in many applications it turns out to be surprisingly difficult, yet also essential. The variety of methods ranges from simple thresholds in black and white images to complex algorithms, for instance the Color-Structure-Code [69]. For ANTARES multiple steps can be seen to fulfill parts of this duty, for instance hit selections or the triggers.

4. Feature extraction
   Features are numbers describing the measured data in a way which is comparable between multiple recorded events. For images this could for instance be the size of an object found by segmentation. In many cases transformations can also be used to describe the data, for instance Fourier or Wavelet transformation. A typical example for a feature for ANTARES data would be the outcome of a reconstruction algorithm.

5. Feature selection
   It is important to use meaningful features, since numbers which are uncorrelated to the current task don't benefit the classification and can even decrease the achieved performance. A tricky aspect is that sometimes features are hardly meaningful on their own, but allow to exploit valuable correlations in combination with other features. If the number of features permits an exhaustive search this would be a favorable option as it is guaranteed to find the global optimum. In cases where this is not feasible other approaches can be used, for instance feature transformations like the Principle Component Analysis (PCA) or the Linear Discriminant Analysis (LDA). In cases where a clear and broad global optimum exists without other pronounced (local) optima, even a simple greedy search approach can lead to the desired result. A more robust approach than greedy feature selection, that is also applicable to high dimensional feature spaces, is described in this thesis in chapter 5.4.

6. Training
   Labeled data is data for which the desired outcome of the task is known. For all supervised learning algorithms it is required to train a model that maps the observed feature input to the desired output. The outputs are so called classes, integer numbers that represent one group of events that share a common property. A simple example that is used in this thesis is the separation in upgoing or downgoing events, considered signal and background. But also more complex assignments are possible like an estimation which is the best reconstruction algorithm for an event. Examples for such algorithms are artificial neural nets, decision trees or support vector machines.

   Regression is an approach similar to classification, but the outcome is not an integer class number but a continuous value. This is especially useful for applications which try to estimate missing variables, but it is not used in this thesis.

A different branch of methods would be unsupervised learning. These methods use similarities in the data to automatically identify classes. A widely known example is nearest neighbor clustering. Due to the requirement to identify certain, a-priori defined event signatures, only supervised learning strategies are considered in this thesis.

7. Classification
   The trained model is applied to new data to estimate the outcome. A classification of labeled data which have not been used for the training can be used to evaluate the quality of the model and therefore to estimate the accuracy of the outcome on unlabeled data.

The preprocessing and feature computation are application specific for the different tasks in this thesis, but the classification algorithm is the same for all tasks and therefore its concept is briefly introduced here.

Much more information about various aspects of pattern recognition and the mentioned methods and concepts can be found for instance in [70], [71] and [72].

4.2 Random Decision Forests

There is a vast variety of classification algorithms available today, each with specific advantages and disadvantages. The selection of the classification algorithm for this thesis has been done based on a comparison of several algorithms as described already in [67]. The best performing method at that time has been the Random Decision Forest (RDF). Due to its favorable properties like a high classification accuracy, robust behavior, tolerance for correlated features and relatively fast execution times, the RDF algorithm has been used for all classification purposes in this thesis.

The concept of a RDF was first introduced in [73]. It is based on the decision tree model. A decision tree consists of so called “nodes”. In the simplest and also most common case one node performs one decision on one feature. For instance it can test the value of feature number seven to be greater than 1.0. Depending on the outcome of this decision, a different linked node is traversed next. These linked follow-up nodes are called children nodes. Trees which are limited to two children per node are called binary trees. The first node in a tree is called root. Nodes without children are called leaves. These leaves can either simply contain one clearly assigned output (the class the event belongs to) or probabilities for different outputs. This structure allows to assign one output to one set of input variables. A simplistic example for a decision tree with two features F1 and F2 and two classes C1 and C2 is shown in Figure 26.

While this concept on its own is working well, many attempts to improve its performance have been made. The RDF tries to make use of the internal variation included in the data used for training the model. This is achieved by training an ensemble of multiple decision trees. But instead of using exactly the same data for each training, which would result in the same tree again and again, only a random subset of the training data and a random subset of the features is used for each tree. This produces decision trees that are able

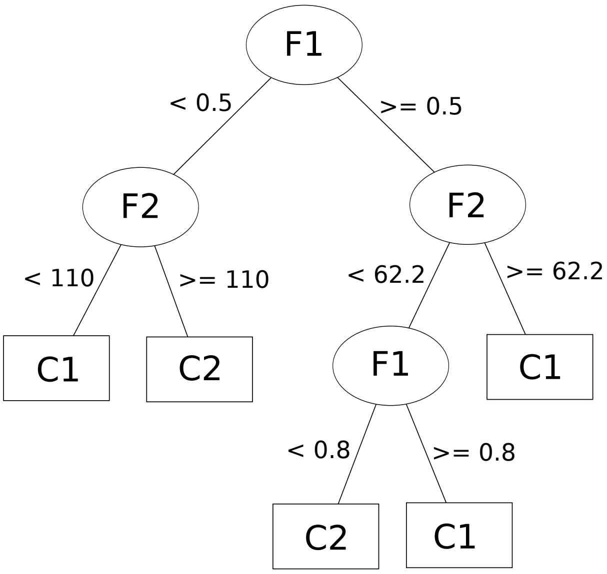


Figure 26: A very simple example for a decision tree with two features F1 and F2 and two classes C1 and C2. The evaluation starts at the top. Each node represents one decision on one feature. The outcome of this decision determines which node to visit next. The rectangular final nodes, so called leaves, contain the result.

to generalize the contained information better than one single tree. The final decision which class is chosen for an event is obtained by a majority vote of the trees. The percentage how many trees agree on the class output can to a certain extent be used as a quality parameter. The implementation has been forked from the 2011 open source version of the alglib library, see [74]. If not stated otherwise, the parameters used for the training of the RDFs are set to 101 trees and for each individual tree a random subset of 60% of the events and of $2 \cdot \sqrt{n}$ features is used for the training, with n being the number of available features here. For the case of 804 features which is explained in chapter 5, this results in 57 features for each training. These values have first been optimized in a parameter study in [67] and since then some have been fine tuned further by observations on new datasets.

There are many algorithms for the training of a decision tree. Famous ones include for instance CART [75] or ID3 [76]. In general the optimization task is to determine which variable to use for the next split and where to set the threshold for this split. Since a globally optimal solution is very hard to compute (np-complex, see for instance [77]), most algorithms aim to generate "sufficiently optimal" trees by an iterative greedy selection. The training algorithm used in the implementation also follows this greedy selection scheme. To determine the next optimal split several metrics are commonly used, for instance the Gini impurity or the information gain. Further information on the properties and details of various training algorithms in general can be found in [75] and [76].

The RDF algorithm has some favorable properties for an application in this thesis. For instance the highly popular class of boosted, tree-based algorithms also trains multiple trees, but increases the weights of events based on the classification result of previous trees. While this often achieves an even more accurate classification result in cases where the available labeled data represent the data in the application sufficiently well, it's performance is less stable if there are systematic deviations between the datasets. These can potentially occur when a classifier trained on simulations is applied to recorded ANTARES data. Compared to properly trained artificial neural networks (ANNs) the performance can be expected to be similar, but most tree based algorithms are less sensitive to imperfectly selected features. Strong correlations between the features or features without useful information for the current classification task can in the worst case prevent the training of an ANN from reaching a stable optimum (in a reasonable amount of time). Based on the achieved performance in many other applications[4], the recent development of deep neural networks is considered a promising candidate to further enhance the solutions in this thesis involving classification.

4.3 Evaluation

When it comes to the evaluation of the performance of a classification one simple rule is to be followed at all times: Never evaluate on the same data that has been used for training. Otherwise the obtained numbers will be unrealistically good, but they do not at all represent the performance of the classification in a realistic application on unseen data. This can be avoided by a separation of the available data into a disjunct training and test set. The drawback of this method is, that events which are in the test set are never used for training and events in the training set are not evaluated. This can easily be solved by performing a second training and evaluation with interchanged datasets. These results already give a first impression of systematic effects influencing the model, but only with a "statistic" of two numbers. Additionally most algorithms benefit from more data used for training, so a model trained on the full dataset is likely to show a better performance than one trained only on half of the dataset. These considerations lead to the default evaluation method, cross-validation. It is the standard way to evaluate the performance of a classifier introduced in almost any book on classification, for example also in [71]. N-fold cross-validation splits the available dataset into n random, disjunct subsets. Then the first $n-1$ parts are used for training and the nth part is evaluated. The process is repeated n times, each time with one different subset used for evaluation and all remaining subsets combined as the training set. This guarantees, that every event has been considered in the evaluation, while a large fraction of the available dataset can be used for the training of the model each time. The ideal, extreme case would be leave-one-out cross-validation, where in each fold only one single event is evaluated. But this also means, that for a dataset with m entries also m repetitions of the training process have to be performed. While this constitutes the optimal case for smaller datasets, it

---

[4] See for instance comparisons at [78]

usually is neither computationally feasible nor required for large datasets. In this case a number of folds $n < m$ is chosen such that the evaluation fits the requirements.

To assess the performance of the presented classifications some specific terms are useful. In this thesis the term efficiency of class X denotes the fraction of events that actually belong to class X and are correctly classified to be class X. If this efficiency is 1.0, all events of this class have been correctly identified. 0.0 means that no event of this class has been classified correctly. Purity of class X denotes the fraction how many of the events that are believed to be of class X actually are from this class. A purity of 1.0 means that no events from other classes were classified to be class X. 0.0 means that only events that actually are from other classes have been classified to be class X and none that actually belonged to this class.

In the case of only two classes, signal and background, purity and efficiency by default refer to the purity and efficiency of the signal class if not explicitly stated otherwise.

Part II

Methods

This part presents the relevant aspects of the algorithms that have been developed for this thesis[5].

All comparison plots are generated by discrete cuts. Therefore only marked data points in the Figures correspond to computed results. Lines connecting the points in some Figures are intended for visualization purpose only. Error bars show statistical errors if not stated otherwise.

---

[5] Chapters 6 and 8 are based on ANTARES internal notes ANTARES-PHYS-2015-001 and ANTARES-SOFT-2015-001, which are neither published nor publicly available.

5 Up / Down classification

5.1 Previous work

As discussed in chapter 3.4, the desired signal of upgoing muon-neutrinos is accompanied by a high number of background events of misreconstructed downgoing atmospheric muons. Common practice to get rid of these is to only include events with an upgoing reconstructed direction and to require a very good reconstruction accuracy. The idea explained here uses a prefiltering of events based on a classifier that distinguishes upgoing and downgoing signatures in the telescope. The intention of this classification is to allow a less strict cut on the required reconstruction quality, resulting in more neutrinos for analysis. The foundation for this classification has already been described in [67], but since then several optimizations have been introduced. The status was a classification using a RDF with two classes (up and downgoing) based on 137 features designed for this task. This achieved a mean suppression of downgoing muons of 90.8% and a mean efficiency for upgoing muon-neutrino events of 93%. The main problem with this result was that the number of misreconstructed atmospheric muons is orders of magnitude higher than the number of upgoing events. Therefore a suppression of 90.8% is not enough, especially since many of these events would also be suppressed by other cuts which are required for an analysis.

A modified version of this classification software has been used to investigate the possibility of a coarse energy reconstruction in [79]. It showed that the RDF classifier, together with the already designed and some task specific features, can successfully be applied to other tasks in ANTARES. This property is also observed for the classification approach described in chapter 6.

5.2 Optimizations

One of the first optimizations has been to extend the feature space. The features described in [67] all rely on one binning of the information belonging to an event in 100 time bins. The length of these bins varies as the duration of the events does, reducing the accuracy with which features based on this binning can be compared between different events. The feature computation has been extended to include the same computations as before in multiple different, fixed time binnings. The median time of all hits of one event is used as the center of these time windows, with a fixed time binning of ± 250 ns and ± 1000 ns for the feature computation. In addition to that, the number of hit selections used in the feature computation has been extended. In [67] the features were calculated twice, once on all hits and once on L1 hits only. A reminder of the meaning of these hit selections can be found in chapter 3.2. The updated version performs a third computation of all features on all triggered hits, so all hits that contributed to the activation of one of the triggering algorithms. New features from intermediate steps of the previous feature computations were also added for each of the nine feature computations (three different hit selections and three different timing windows). In total this results in 783 features (87 instead of 68 features per computation). Besides these, externally computed features are also incorporated. These are the outputs of two standard ANTARES direction reconstruction algorithms, namely “Aafit”, described in [38], and “BbFit”, described in [80]. For each reconstruction algorithm the reconstructed direction and the various quality parameters are used as features. For the up/down classification used in this thesis the final feature vector for each event contains 804 features.

To address the fact that the number of signal and background events is vastly unbalanced (10^6 times more downgoing atmospheric muons), at one point during the optimization of this classification a cut was introduced that requires 80% of all decision trees of the RDF to identify an event as upgoing to actually accept it as upgoing signal. Events with an agreement below 80% were considered “probably still a downgoing muon” and therefore discarded. This variable is called “RDFSafety” in the data processing.

In total these changes improved the accuracy of the classification to reach a muon suppression of 97% while still preserving a muon-neutrino efficiency of 80%.

5.3 RDF cascade

As further optimization, this classification is now used as a two-step process to reach a higher muon suppression while preserving roughly the same amount of upgoing neutrino events. The second classification step can exploit the changed distribution of parameters due to the changed ratio of downgoing atmospheric muons to upgoing neutrinos and the specific properties of those downgoing muons that survived the first step. A further improvement was observed when the RDFs are not trained with the same number of events for both classes. In general the distribution of instances between the classes should mimic the distribution encountered in an application. But if the deviation between the classes, depending on the actual classification task, becomes too large[6], the training can become unstable and tends to ignore underrepresented classes. As the number of downgoing atmospheric muons in this task is tremendously larger than the number of upgoing neutrinos, the actual distribution was replaced. Asymmetric distributions for downgoing versus upgoing of up to 10:1 turned out to give stable results for the first classification step. To include a safety margin for later application, a ratio of 3:1 has been used for the training of the first RDF step. Therefore the first step identifies unclear events as downgoing

---

[6] Too large in this case means that completely ignoring the class with fewer events leads to a smaller error. For example, in a two class classification where 99.9% of all events belong to class 0 and only 0.1% to class 1, the classification reaches 99.9% accuracy by labeling every event to be class 0. Therefore the training algorithm will optimize in that direction if the fraction of correctly classifiable events ($0.999 \cdot E_0 + 0.001 \cdot E_1$) is lower. ($E_x$ denotes the efficiency for class x, the probability that events from class x are recognized to be from that class.) So E_0 should be greater than $1 - \frac{0.001}{0.999} \cdot E_1$. If we optimistically assume that both efficiencies are equal (usually the efficiency of the smaller class is lower), then E_0 should already be greater than 99.9%. Since the efficiencies are more or less fixed for a task with given classes and features, there is a maximal stable deviation in the number of events between the classes. In reality the effect is a continuous trend, not a sudden change in behavior.

muons and can be seen as a course filtering to remove most of the muons. The training of the second step was performed with a ratio of 1:2, so with more neutrinos than muons. The effect is that the second step is very careful not to lose neutrino signal and is intended to only remove muons that can be identified easily due to the shifted distributions of the classes after the first step. For the given classification task the cascade with these distributions between the classes outperforms cascades with homogeneous or inverted distributions. This two-step cascade of RDFs allowed to drop the previously required cut of 80% tree agreement, but of course this can still be applied if needed.

The final achieved accuracy of this improved classification reaches a mean muon suppression of $99.85\% \pm 0.1\%(\text{stat.}) \pm 0.02\%(\text{syst.})$ and a mean neutrino efficiency of $81.7\% \pm 0.5\%(stat.) \pm 0.8\%(syst.)$. The systematic errors here are estimated based on cross-validation results. They only account for the systematics of the classification, not those introduced by the detector, the simulations or the rest of the data analysis. It may not be intuitive to interpret the difference between the final RDF cascade result and the intermediate step. The 2.85% gain in muon suppression (from 97% to 99.85%) may not seem like much, but it results in a remaining number of atmospheric muons of 0.15% instead of 3%, which in the end means 20 times less atmospheric muons.

The result of this RDF cascade is shown energy resolved in Figure 27. To generate this plot, Run-by-Run Monte Carlo simulations version 2.2 have been used (See chapter 3.6 for more information on the simulations). It includes muon-neutrinos and atmospheric muons that were triggered by the triggering algorithms described in chapter 3.2. One can see that the efficiency with which muon-neutrinos are identified is relatively constant above an energy of 100 GeV. The suppression of atmospheric muons is very high for lower energies. The drop at higher energies is caused by the low statistics in the simulations for high-energetic events.

A comparison of cascade depths (numbers of RDF steps) and quality cuts showed that a second step is more efficient than requiring a higher RDFSafety value, but adding a third step is less efficient for this task.

5.4 Feature selection

An issue for the up/down classification in [67] was found to be the optimal selection of the features to use. The approach had been to use genetic algorithms for this optimization task. See for instance [81] or [82] for more information on their usage in optimization. Although in theory they could be flexible and directed enough to find an optimal solution in a reasonable amount of time, this approach turned out not to converge to any stable optimum. Since the feature space is high dimensional and the classification method is highly nonlinear, a greedy search that iteratively selects only the one feature that gives the highest gain in performance did not provide a stable solution either. In this context greedy searches starting with empty, full and partially filled feature subsets have been tested for addition and removal of features. An exhaustive search approach that simply tests all possible combinations of features is no option for computing time reasons here, as the feature space is large and the possible

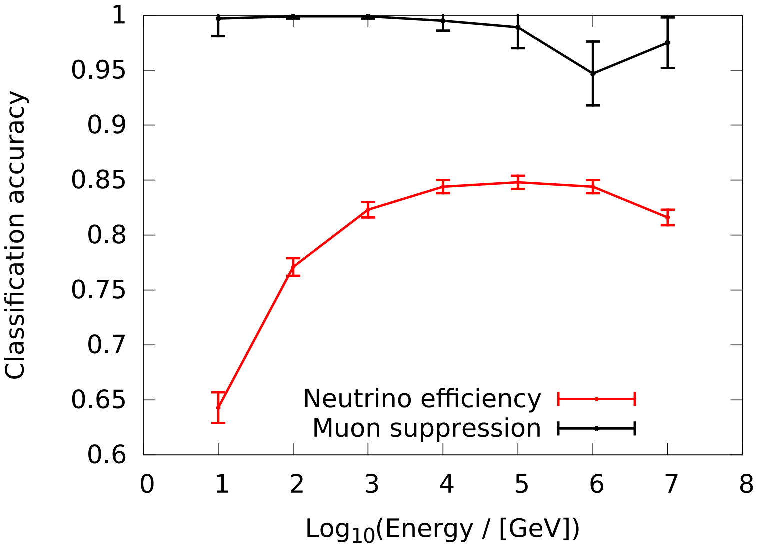


Figure 27: Neutrino efficiency and muon suppression versus energy for the two step RDF cascade on RbR 2.2. The simulated events are weighted with an E^{-2} spectrum.

combinations are too numerous.

Therefore a different method of optimization one might call a “greedy ensemble” has been designed to solve this problem. The idea is to heuristically combine some of the selection criteria from genetic algorithms with the focused direction and short runtime of a greedy search algorithm.

The search starts with an empty solution. In the first iteration the empty solution is extended by one single feature that yields the best performance on its own. Since this decision must be based on a certain data sample, but might change for a different sample, the search for the optimal feature is performed N times, each time on a random subset of the training data. All different solutions that are generated by this process are kept. The number of new solutions per old solution cannot be greater than the number of tested subsamples N, and it is at least one in cases where all evaluations agree on the same feature. The result is an ensemble of M solutions, depending on how many features have been found to be the best possible addition for one of the previous subsets. M denotes the current number of solutions here. As an example the first iteration might result in $M = 2$ solutions with solution S1 containing feature F5 and solution S2 containing feature F8.

In the second iteration the same process is repeated for each of the M solutions that were found in the first iteration. For the given example the extension of S1 might lead to a new solution S3, which contains F5 and F8, a solution S4, that contains F5 and F6, and a solution S5, which contains F5 and F12. The same is done with the old solution S2, which could result e.g. in S6 with F8 and F5 and S7 with F8 and F1.

Although equal solutions like S3 and S6 can be treated as one solution, due to the exponential growth of the number of solutions this scheme quickly reaches a very large number of solutions to process in each iteration. Since testing tens of thousands of solutions in each step is not computationally feasible, a cutoff is applied at the end of each iteration. The set of all currently found solutions is reevaluated again on a larger fraction of the full data sample and only the K best performing solutions are kept for the next iteration. For solutions that have already been reevaluated multiple times, the mean of all evaluations is used, increasing the accuracy of the estimate over time.

This pruning of the list of found solutions keeps the runtime of each iteration almost constant also for large feature spaces. If K is set to 1, the scheme is mostly equivalent to a greedy search algorithm. In this case this approach inherits the main drawback of a greedy search, the tendency to get stuck at a local optimum. With a larger but fixed number of solutions per iteration this tendency gets reduced, but not eliminated. On the other hand, apart from effects due to hardware limitations like cache sizes, the runtime per iteration and therefore of the whole process scales linearly with the number of kept solutions.

This iterative scheme stops either when there are no more features to add, after a fixed number of iterations, or, and this constitutes the common case, until the obtained solutions did not improve the performance of the classification during the last L iterations. One could argue that no improvement in one new iteration should be sufficient to stop the optimization, but this would neglect the possibility that in some cases a gain in discrimination power can be achieved only by the correlation of multiple features.

As a further tweak the stopping of iterations is not done simultaneously for all solutions once the performance of the best solution doesn't improve anymore, but individually for each solution. The benefit of this strategy is that new solutions which are not yet as good as the current best one can still evolve further at least L iterations on their own, while solutions which are kept due to their good performance, but which do not improve any further can already be excluded from the iteration scheme to save time. As a result the optimization will continue to search for better solutions as long as any of the M solutions still improved during the last L iterations.

This algorithm shares traits with the genetic algorithms that have been used in [67], as it also uses generations of multiple solutions and tries to iteratively find improved new solutions derived from the old, but it also shares the fast and highly focused aspect of greedily searching for new features with a plain greedy search algorithm.

In the tested application scenario of up/down classification with a high dimensional feature space of 804 features, the described parameters by default are set to $N = 3$, $K = 30$ and $L = 6$. The number of currently found solutions M varies, but is smaller (in the beginning) or equal to K after each completed iteration.

The best solution for the full feature set on RbR 2.2 simulations has been obtained in a detailed study by Thomas Kittler [83]. The numbers for accuracy in this study do not consider the weights of the events in the simulation.

The best solution[7] found in this study for a plain up/down RDF classification reached a classification accuracy of 94.87% ± 0.42% compared to 94.02% ± 0.42% for the full feature vector. The relatively small gain can be explained by the observation that RDFs can handle strongly correlated features exceptionally well, a property they share with many other tree-based classification algorithms.

For the actually used two step RDF cascade as described in chapter 5.3, no solution with a statistically significant improvement of the performance could be obtained at all. The best solution[8] in terms of a solution with few features but a good accuracy reached 96.67% ± 0.16% in comparison to 96.74% ± 0.16% for the full feature vector.

Therefore the RDF cascade in this thesis is used with the full feature vector as this classification scheme gives the highest accuracy that has been achieved by any means.

5.5 Possible improvements

An important improvement for an analysis would be to adapt the classification to the individual data selection. While the trained RDF cascade is a powerful tool to reduce the number of background events independent of the actual analysis goal and of other applied cuts, it could improve considerably if other cuts which are additionally applied to data in an analysis were also applied before the training. The reason for this is that many of the events which are filtered out by the classification would have also been rejected by other cuts, reducing the observed benefit of the classification. A training with accurately prefiltered datasets would also allow the classification to adapt to the desired application in the best possible way, giving the best final result.

Another idea how to potentially improve this classification would be to evaluate a relatively new branch of pattern recognition methods, often called deep neutral networks (DNNs). This contains for instance the so called deep belief networks, see for instance [84] or deep convolutional neutral networks, see [85] and various others. Comparisons[9] between RDFs, several other classification techniques and DNNs for different tasks indicate the potential for a substantial gain in accuracy by DNNs.

---

[7] Using the features: 784,786,788,164,99,50,24,176,670,78,550,221,394,360,306,745,791,382, 686,295,748,271,101,495,111,192,320,124,424,265,91,755,737,622,290,425,52,611,609,679,510, 584,616,276,585,283,250,9,304,90,446,628,199,214

[8] Using the features: 784,788,786,252,581,438,494,27,12,36,50,791,342,686,78,513,360,285,75, 694,138,236,390,304

[9] For instance at [78]

6 Selectfit

There are multiple direction reconstruction algorithms for tracklike muon events in ANTARES. The most common way to make use of these track reconstructions in a physics analysis is to evaluate the results of one favorable algorithm on all events and to ignore all others. For some events however the selected algorithm may not perform well or even fails, although these events would be reconstructed well with another algorithm. In these cases adding a different reconstruction avoids loosing these events and therefore benefits the final analysis. A first approach to combine algorithms could be to use thresholds in the form of:

1. If reconstruction A has a better quality parameter than a certain threshold, use reconstruction A.
2. If A has a poor quality parameter but B has a better quality parameter than a certain threshold, take reconstruction B.
3. Otherwise, do not use this event.

With some parameter scanning for the thresholds this approach can already give some benefit compared to a single algorithm. But on the other hand it is a time consuming process to manually tune the thresholds for every new application scenario. It also becomes increasingly complex if we want to use more than two reconstruction algorithms and sometimes it does not use the best reconstruction available for an event. An example for this non-optimal performance would be an event where reconstruction A has a quality parameter slightly above the threshold, but reconstruction B gives an outstandingly good reconstruction. The simple threshold approach would still use A here.

The idea of Selectfit is to have an automated process that cannot only combine many reconstruction algorithms, but can also choose the most accurate available reconstruction for each event as reliably as possible. Additionally, a new quality parameter for error estimation is introduced to allow the identification of well reconstructed events, regardless of which specific reconstruction was chosen.

6.1 Overview

Selectfit uses a classification to find the best reconstruction for a given event and also to estimate the error of the chosen reconstruction. The current version is designed to combine four reconstructions for tracklike events, namely Aafit [38], Bbfit [80], Bbfit with MEstimator [80] and Gridfit [86]. Another version exists which combines these four and the Dusj reconstruction, see [87] and [88], and the Q Strategy for shower events as explained in [89]. The first scenario is intended to enhance analyses using the muon-neutrino channel, whereas the second version explores the potential of this approach to combine tracklike and showerlike events in one single reconstruction. This would allow analyses to use events of both signature types at the same time. The results shown in this thesis are for the first, tracklike case only, because the handling of shower events in ANTARES is not yet at the same level of experience and precision as for track events.

The same strategy can also be used to combine energy estimation algorithms, but these ideas are pursued by other members of the ANTARES collaboration now and since energy information is not evaluated in this thesis, these attempts are also not covered here.

Detailed example scripts on how to use Selectfit for direction reconstruction can be found in Appendix A.1. Obviously, the logic behind Selectfit can also be used with any other classification algorithm, software package and similar application scenario.

6.2 Features for Selectfit

The current version of Selectfit uses the outputs of the algorithms it combines and the number of pulses as an energy correlated variable as input. For each reconstruction strategy it includes:

- The zenith value
- The azimuth value
- All available quality parameters

There are at least two quality parameters for each algorithm, but this is not necessary. One would be sufficient, but the more useful information an algorithm outputs about its result, the more accurate the selection and error estimation can be in the end. A detailed list of features can be found in appendix B.

6.3 Classifier for Selectfit

The classifier used here is again the versatile Random Decision Forest as already described in chapter 4.2. The classes in this task are not defined as upgoing and downgoing like they have been in chapter 5, but instead as the best direction reconstruction algorithm for each event for the first RDF and for a second RDF as the angular error of that reconstruction that has been selected in the first step. Selectfit consists out of these two steps.

6.4 Selecting the best reconstruction

The first and major step of Selectfit is to find the best direction reconstruction algorithm for each event. It is not necessary for all reconstructions to be available for an event. An algorithm may not output a result for instance if an internal optimization doesn't converge. If one or more of the algorithm results cannot be found, the angular difference of the missing reconstructions is internally considered to be 2π, hence this algorithm is not selected. The two predefined possibilities, four track reconstructions or four track and two shower reconstructions can be used in the current implementation as described in appendix A. The default outcome is a number ranging from 0 to the number of considered reconstructions:

- Class 0 : “No fit was available at all”
- Class 1 : “Aafit is the best available fit”
- Class 2 : “Bbfit is the best available fit”
- Class 3 : “Bbfit with MEstimator is the best available fit”
- Class 4 : “Gridfit is the best available fit”
- Class 5 : “Dusj is the best available fit”
- Class 6 : “Q Strategy is the best available fit”

The zenith and azimuth of Selectfit are set to the zenith and azimuth of the selected reconstruction.

Since the algorithm sometimes misclassifies an event, it can happen that the selected reconstruction is not actually the best that was available for this event. To achieve a better overall performance compared to the best single reconstruction algorithm, the error introduced by misclassification may not become larger than the benefit for events where a different reconstruction was chosen correctly. It has been observed that the misclassifications occur mainly for those events, where the best and the selected algorithm have a similar angular error. Therefore the majority of misclassified events introduces only small additional errors. Nevertheless, for the configuration with four track reconstructions, this prevented the addition of other, similarly designed reconstructions. If a reconstruction is too similar to an already considered algorithm or has a too poor performance, it will in total introduce a higher error due to more misclassification than what can be gained by the few events it reconstructs better.

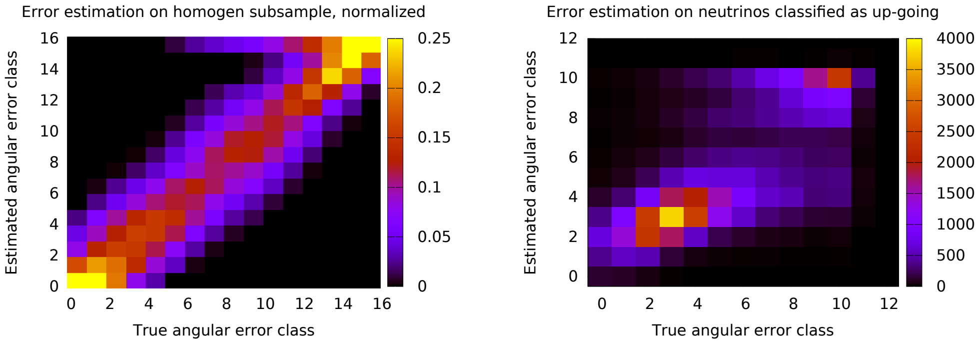
An alternative to selecting one of the direction reconstruction results would be to merge all results, for instance with weights resembling how likely each reconstruction seems to be the best. This approach has been tested and typically produced a lot worse results than simply selecting the best. The reason for this observed behavior is that a good reconstruction with a low angular error will usually become worse by merging it with less accurate directions, even if the worse direction has a small weight. The case where two algorithms give a similar error, but lie on opposing sides of the true direction, is rare. Therefore merging of results is not contained anymore in the current version of Selectfit.

6.5 Error Estimation

If only one track reconstruction is used for an analysis, one would typically cut on the main quality parameters of this reconstruction in order to obtain a sample of accurately reconstructed events. But when combining multiple algorithms, cutting only on the quality parameter of the dominant algorithm will remove exactly those events, where the addition of other algorithms improved the result most. Combining the quality parameters by hand for each algorithm is the same task as combining the fits themselves. Therefore the chosen approach to solve this is again a classification. The main purpose of a cut on a quality variable of a direction reconstruction is to cut away those reconstructions which have a large angular error. To mimic this behavior with a classification, the events are put into bins (classes) of similar angular errors during training. Some experiments on how these bins should be formed covered examples of linear inter-class distances (0°-1°, 1°-2°, ..., 179°-180°), exponential inter-class distances (0°-1°, 1°-2°, 2°-4°, 4°-8°, ...) and some manually hard-coded schemes. Since an ideal distribution for the training of this classification would have an equal number of events in all classes, also a dynamic computation of class boundaries with equal number of events was tested. The problem with this approach is that the class boundaries change for every training, especially if many classes are to be trained. Therefore this idea had to be discarded. For all remaining possibilities also the width of a class bin was varied as well as the start and end points (e.g. 0°-0.5°, 0.5°-1.5°, 1.5°-2.5°, ..., 44.5°-45.5°, greater than 45.5°). The experiments showed that an exponential distribution yields the best results. The exponential factor has some effect on the result, but its possible range is limited. It may not be too large because otherwise there are only few classes, which would only allow very coarse selection of events. It may also not be too small because then the number of classes grows quickly up to a point where not enough events for a stable training are contained in the less likely classes anymore. The final class distribution for the Selectfit error estimate was fixed to:

- Class 0 : 0.0° - 0.1°
- Class 1 : 0.1° - 0.2°
- Class 2 : 0.2° - 0.4°
- Class 3 : 0.4° - 0.8°
- Class 4 : 0.8° - 1.6°
- Class 5 : 1.6° - 3.2°
- Class 6 : 3.2° - 6.4°
- Class 7 : 6.4° - 12.8°
- Class 8 : 12.8° - 25.6°
- Class 9 : 25.6° - 51.2°
- Class 10 : 51.2° - 102.4°
- Class 11 : 102.4° - 180.0°
- Class 12 : “No reconstruction available”

The quality parameter of Selectfit is set to the error estimate class determined by this classification. As for selecting the best fit, also for the error estimation some events are misclassified. Figure 28 shows the misclassification during error estimation. Ideally this would be a diagonal line. Figure 28a shows how well the estimation can work. In this test the same number of events has been used during training for each of the (in this study still 16 instead of 12) classes. The result of an optimization for the actual distributions is shown in Figure 28b. As can be seen in both cases, misclassifications occur dominantly between neighboring and next-to-neighboring classes. An increased probability for misclassifications between neighboring classes is not surprising, as this is for instance also caused by events with an actual angular error of 1.55° being identified as class 5 (starting from 1.6°) instead of the correct class 4 (up to 1.6°).



(a)

(b)

Figure 28: a) Normalized distribution for equal number of events per class during training, 16 classes. This scenario demonstrates how well the error estimation can be adapted in the the idealized case where each class has the same number of events. b) Distribution for a realistic, precut sample in absolute numbers, 12 classes, optimized to minimize errors for realistic distributions.

One idea to improve the result of the error estimation would be to use a regression, a similar method to estimate missing values, instead of a classification. The investigation of this possibility is pursued by another member of the ANTARES collaboration.

6.6 Accuracy of Selectfit

Combining the fit selection and the error estimation gives Selectfit. The results which are achieved for charged current neutrino events on the small available sample of RbR 3.0 simulations is shown in Figure 29. The x-axis shows the median angular error between the true simulated direction and the reconstructed direction. The different points in the figure are the results for different quality cuts, with stricter cuts being to left. The y-axis shows the cumulative efficiency, stating what fraction of neutrinos survive a certain cut level. The goal for a direction reconstruction is to reconstruct as many neutrinos as possible with an angular error as small as possible, so the best, only theoretically achievable point in this plot is to the top left. The red curve shows the performance for Aafit, the best and most often used single reconstruction algorithm. To obtain the different values of the curve a cut on λ, the main quality parameter of Aafit, was varied from -4.0 to -9.0 in steps of 0.1, which covers more than the commonly used range. The blue curve is the performance of Selectfit on the same data set, with each point being a different cut on the estimated error class from 0 to 12. One can see that regardless of the actual cutting value, Selectfit does improve the performance of the direction reconstruction. The green curve is obtained by the "perfect Selectfit". For each event it chooses the best available reconstruction and it also uses the exact angular error of the selected reconstruction as error estimate. Therefore this is an upper limit on what could theoretically be achieved with the Selectfit approach for muon-neutrinos in ANTARES. For loose cut values and hence less accurate reconstruction results, the gain in the number of neutrinos by Selectfit is around 11%. For intermediate cuts, which is where e.g. a cut on λ greater -5.2 or -5.4 would be located, the gain is between 8% and 9%. This is between the 4.th and 5.th datapoint of Selectfit from the left in Figure 29. For the most accurate reconstructions both reconstruction methods deteriorate quickly, but Selectfit reaches about half the angular error of Aafit for the tiny, most precisely reconstructed neutrino samples. In total Selectfit shows a gain of at least 8% in the number of neutrinos for the whole range of possible quality cuts. Since the available RbR 3.0 sample has only limited statistics and has not yet been validated at the same level of detail, Figure 30 shows the same comparison for a larger sample of RbR 2.2 simulations. All plots showing results for RbR 2.2 do not include statistical error bars as they are indistinguishable from zero. This plot shows all events weighted according to an assumed signal flux of $E^{-2.0}$ and also includes Aafit with a cut on a second quality parameter β, which was set to the default optimized cut $\beta < 1^\circ$ as used in many analyses. One can also note that, even though the difference between the reconstructions is very similar, there is a difference observable for the absolute values of the median angular errors for both reconstructions. The better performance is achieved for RbR 2.2, which seems plausible since overall RbR 3.0 contains more realistic simulations, for instance scattered light from interaction vertices. Still it is not absolutely clear that the magnitude of this difference is fully covered by this effect, but since it can be observed similarly for both reconstructions, it is not a feature introduced by Selectfit and it will be investigated by future work on the simulations.

Figure 31 shows the individual causes of the improvement. It includes a curve in purple for always using Aafit as reconstruction but with a RDF classification as quality parameter and a curve in black which uses Selectfit to choose a reconstruction, but then again Aafit λ and β as quality cuts. This clearly shows that for high efficiencies and worse median angular errors the performance improvement is due to the selection of different fits. On the other hand for very accurate reconstructions, Selectfit and Aafit using an RDF error estimate have approximately the same performance. This shows that the gain for small angular errors is mainly due to the RDF error estimation.

A similar behavior can be seen in Figure 32, where the effect of changing the selection or the error estimation from Selectfit to the respective "perfect" step from the perfect fit is shown.

The contributions of the individual fits to the Selectfit results derived from recorded data without any cut on the error estimation are:

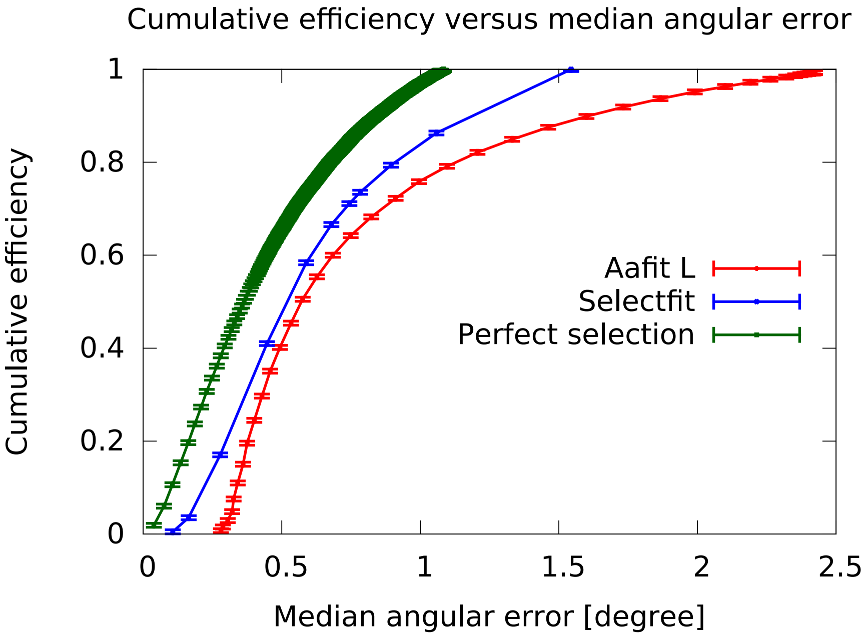


Figure 29: Comparison of Aafit with a varying cut on λ, Selectfit with varying error estimation cut, and the perfect selection of the best available fit with a varying cut on the exact angular error on RbR 3.0. Error bars indicate statistical errors.

- Aafit: 27.0 ± 0.02 %
- Bbfit: 2.0 ± 0.006 %
- Bbfit with MEstimator: 2.0 ± 0.006 %
- Gridfit: 69.0 ± 0.04 %

The errors for these numbers are statistical errors only.

The contributions of the individual fits to the Selectfit results in the final data sample for this thesis as described in chapter 7 are:

- Aafit: 48.2 ± 0.6 %
- Bbfit: 0.5 ± 0.07 %
- Bbfit with MEstimator: 4.7 ± 0.2 %
- Gridfit: 46.5 ± 0.6 %

As a further improvement for this application scenario one could also drop Bbfit without MEstimator, as it would reduce the probability for misclassifications, resulting in a more accurate final result.

Another idea to improve the reconstruction beyond what is shown in this thesis would be to adapt the training of Selectfit to the analysis it is used for.

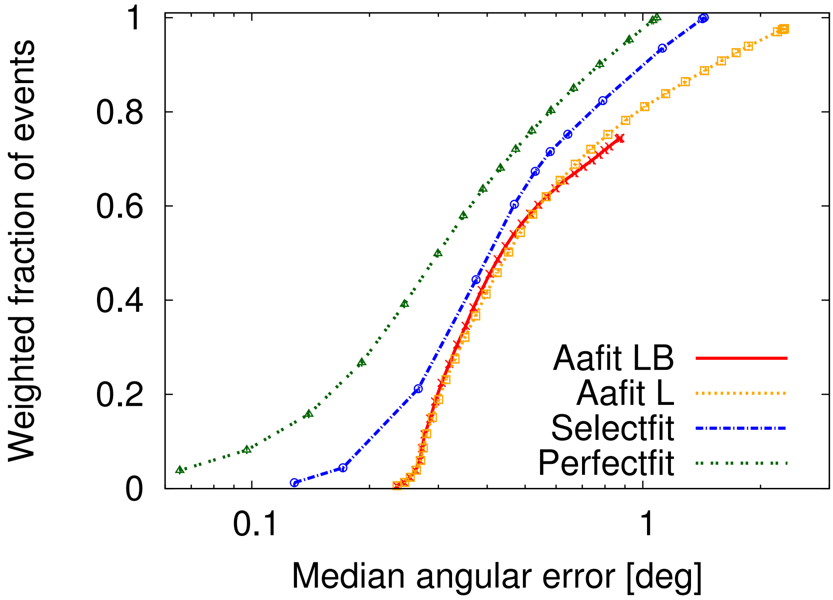


Figure 30: Comparison of Aafit with λ and β cuts, Aafit with λ cut only, Selectfit and a theoretic, perfect Selectfit on RbR 2.2

If any other cuts that can be applied before Selectfit in the data selection of the analysis are also applied in exactly the same way to the training data, the resulting Selectfit models can be expected to achieve better results for an analysis using these cuts than without this adaption.

Summarizing the observations for the developed and tested version, Selectfit for the direction reconstruction of tracklike events outperforms the best available single track reconstruction algorithm for any compared accuracy and neutrino efficiency.

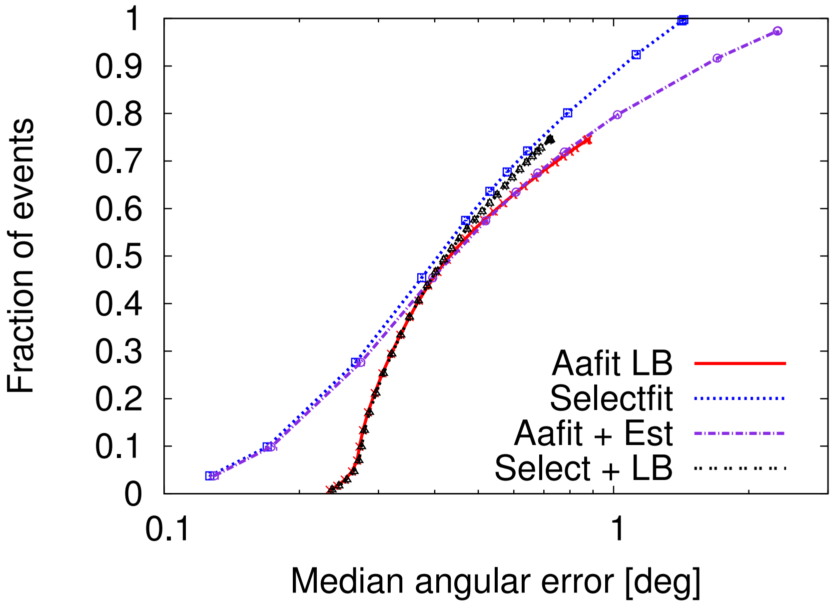


Figure 31: Improvement from the single steps of Selectfit in blue compared to Aafit with λ and β cuts in red. Using Selectfit to combine all algorithms but using λ and β to identify the most accurate reconstructions gives the black dotted line. There is only a gain for less strict cuts. The purple dotted line shows Aafit with a RDF to identify accurate reconstructions. A gain is present for very accurate cut levels only.

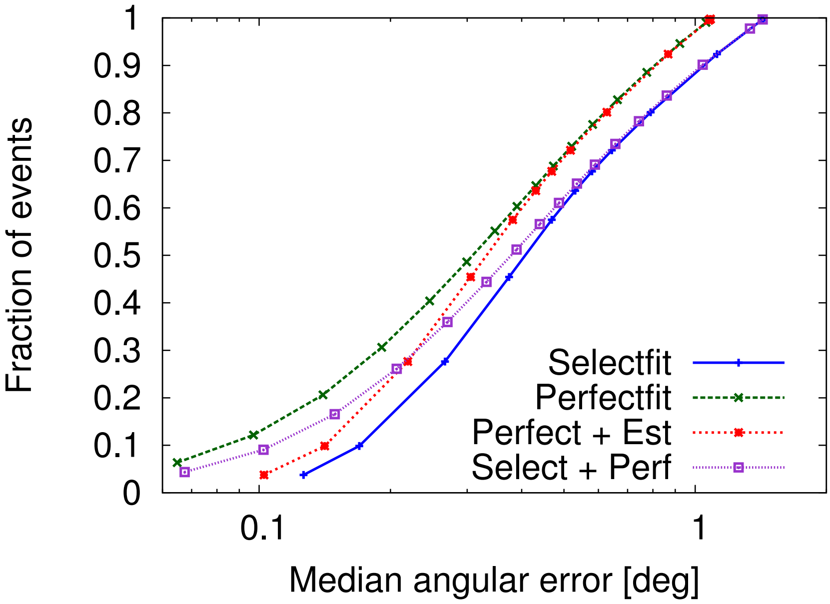


Figure 32: Theoretical improvement from perfectly error-free steps instead of trained Selectfit RDFs. The actual Selectfit performance is shown in blue. Replacing only the RDF selection of the best algorithm with an error-free selection gives the red dotted line. Replacing only the estimation of the angular error with an error-free estimation yields the purple dotted curve. Replacing both steps by the error free versions gives the green curve.

7 Data selection

7.1 Run selection

The considered time period is from the beginning of the data taking with ANTARES on 29.01.2007 until on 31.12.2012. To obtain a reliable and well understood data set, several criteria were required for data taking runs to be included in this analysis:

- The mean optical rate must be below 300 kHz. The purity and efficiency of some reconstruction algorithms drops considerably for even higher background rates.
- The run setup must be a physics setup. This is to exclude e.g. data with artificially emitted light, which was intended for calibration.
- The data quality category must be 1 or better. This category is derived as a standard quality assessment in ANTARES for all runs. More information on the exact procedure can be found in [90].
- The run may not be identified to contain so called “sparks”. These are seldom malfunctions where light is produced by an electric discharge.
- The data file must have been processed correctly by the standard ANTARES data production and have been accessible on the mass data storage.

Most of these requirements are de facto standard for ANTARES data analysis. A list of all included runs can be found in appendix H. In total this corresponds to an effective lifetime of the detector of 1332.8 days.

7.2 Event selection

The data in the selected runs does not only contain upgoing muon-neutrino events, but mostly atmospheric muons. To identify the desired signal some filtering of the events is performed:

- Application of the up/down classification as described in chapter 5 to get rid of a large fraction of downgoing atmospheric muon events. Only events which are classified as upgoing (Class 1) are processed further.
- Selectfit as described in chapter 6 is applied to each event. Only events with a reconstructed zenith angle of more than 90° (upgoing) are kept.
- The estimated angular error from Selectfit must be below 1.6° (only error classes 0, 1, 2, 3 and 4 as defined in chapter 6.5 are kept).

Even after these cuts the resulting muon-neutrino sample is expected to contain a contamination due to the vast number of atmospheric muons. This chapter presents a comparison of the actual measurements with the event numbers and composition expected from simulations. But as explained in chapter 8, absolutely no information derived from simulations is used in this analysis.

7.3 Comparison of simulations with data

To check whether the expectations derived from simulations are valid for recorded data, a small sample of data runs has been analyzed together with the corresponding RbR simulations. The subset of data files used for this comparison is called "burn sample". It is excluded from the actual evaluation to avoid a potential bias. Some of the simulated contributions had to be scaled up to compensate for missing or empty files. This introduces an uncertainty that is not contained in the presented error bars.

Figure 33 shows the comparison of the small available sample of RbR 3.0 simulations with the corresponding data files. On the x-axis is the error estimation class as defined in chapter 6.5. More accurate direction reconstructions are to the left. On the y-axis is the number of events. The simulated events in these plots have been weighted according to the expected atmospheric flux. The plot is cumulative as the values shown e.g. at error class 2 contain all events of classes 0, 1 and 2. Therefore these cumulative curves show the composition an actual sample would show for different cuts on the estimated error class.

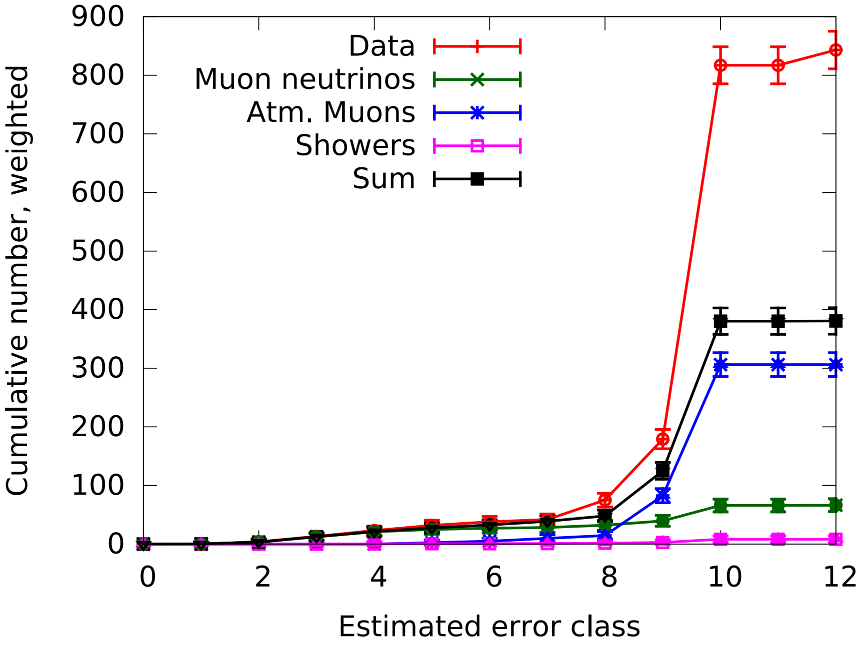


Figure 33: Data Monte-Carlo comparison for RbR 3.0.

Ideally the red curve, representing the event numbers from data, and the black curve, containing the sum of all simulated contributions, should match. Obviously this is not true for very poor reconstruction qualities as seen for large values of the error class in Figure 33.

This is not surprising since class 12 for instance means that there was no reconstruction result available at all, while classes 10 and 11 denote events with reconstruction errors between 51.2 and 180.0 degrees. Events that have such tremendous reconstruction errors are likely triggered due to pure noise without any particle interaction, which is not included in the simulations.

Since no spatially resolved analysis could work with errors of up to 180 degrees, Figure 34 shows a zoom to the more interesting region of, depending on the goal of an analysis, potentially acceptable reconstruction errors. What can be seen here is that the agreement between data and Monte Carlo simulations is rather good for ANTARES standards. What is also visible is that for the used quality cut value of class 4 or better, this comparison suggests that there is hardly any remaining contamination of atmospheric muons left.

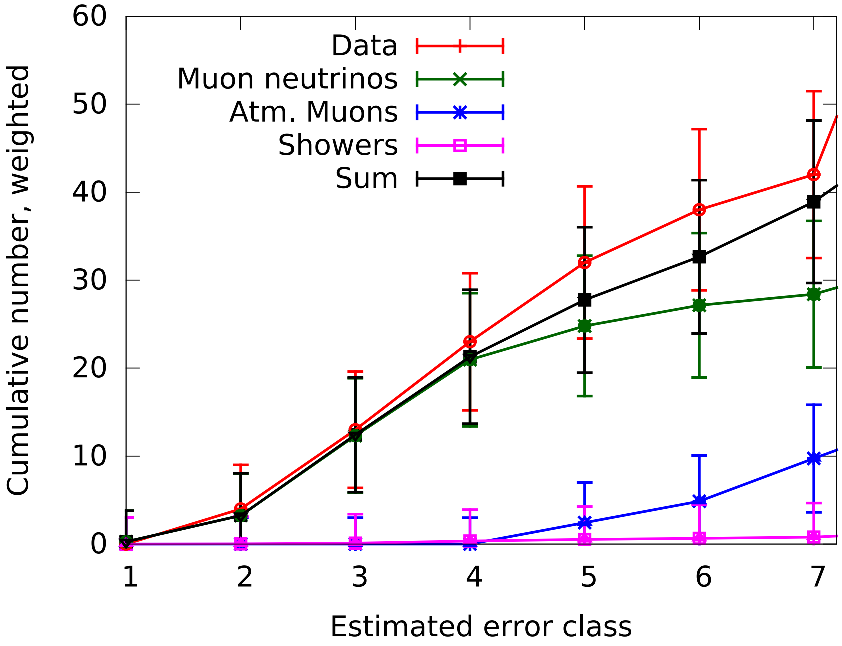


Figure 34: Data Monte-Carlo comparison for RbR 3.0. Zoom to realistic cut values.

But as the statistical uncertainty for the RbR 3.0 sample is large, another crosscheck with RbR 2.2 has been conducted. Figure 35 shows the same comparison as Figure 33 did, this time for a different sample of RbR 2.2 runs. The observed behavior is similar to the previous comparison, but the gap for very poor angular errors is even larger. Again following the reasoning that arbitrarily large angular errors can't be used for a spatially resolved analysis, the zoom to realistic values in Figure 36 also shows a good agreement between MC and data and mostly similar distributions as before on RbR 3.0. But it's important to note that in this comparison more atmospheric muons survive harder cuts, resulting in an expected contamination on the order of 29.7% for the remaining sample after a quality cut for error class 4. The applied Selectfit has been trained using RbR 3.0 and can be expected to give more accurate results on the more recent simulations. In the comparison using RbR 2.2, the remaining muons may as well be muons that have not been simulated correctly in this

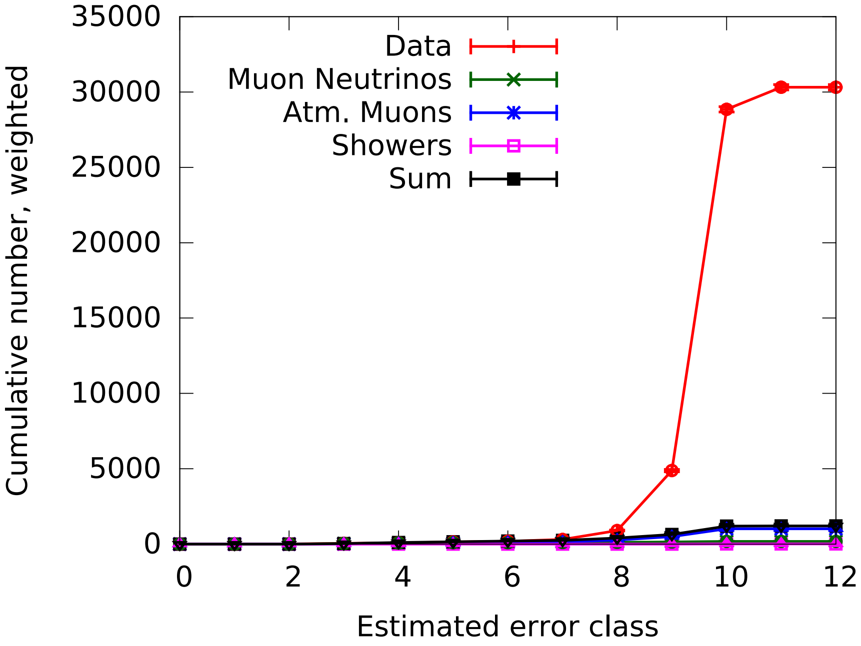


Figure 35: Data Monte-Carlo comparison for RbR 2.2.

older version. Nevertheless, since a conservative estimate is preferred here and because the comparison on RbR 2.2 has higher statistics, this is also considered the contamination rate for the obtained final sample. Based on the numbers for a cut on error class 4 in Figure 36, the unblinding of the full data sample has been expected to yield 13078 ± 114 events.

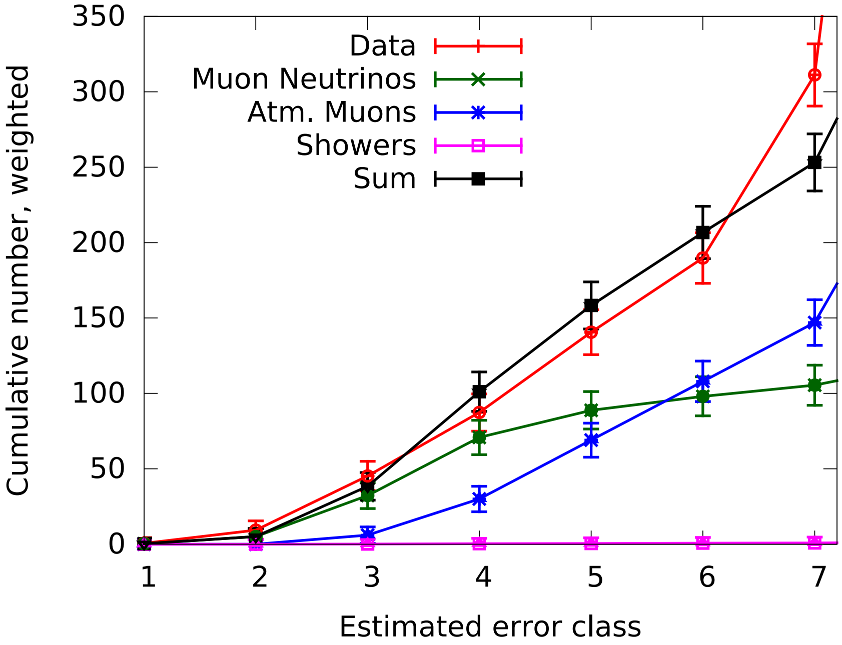


Figure 36: Data Monte-Carlo comparison for RbR 2.2. Zoom to realistic cut values.

8 Multiscale source search

The data recorded with the ANTARES neutrino telescope allows to address diverse physics questions. Many sophisticated analyses have been performed, each optimized specifically for one of the different possibly detectable physical phenomena, see for instance [91], [92], [93], [94] or [95].

Most searches for spatially confined sources of cosmic neutrinos rely on a model assumption derived from theory describing one specific potential source. These searches evaluate the distribution of reconstructed neutrinos to detect or reject this specific hypothesis. This usually means searching for an excess beyond the expectation by atmospheric neutrinos at a known position in the sky with a predefined size and shape.

The idea behind the model-independent multiscale source search is to approach this task the other way around. Instead of optimizing highly for one specific prediction, this approach is intended to reveal the most significant excess of neutrinos within the recorded data in any region in the sky. A dedicated analysis can then be optimized for the hypothesis derived from the findings of this search.

The main advantage of this approach is that, due to its unbiasedness and flexibility, it could potentially also reveal neutrino sources which would not be tested in a realistic time scale by other searches. Apart from that there is a highly reduced risk that this search suffers from deviations between assumptions, simulations and the reality.

Of course there are disadvantages which come with an evaluation like this. The main drawback is the high number of trials when evaluating for sources of arbitrary sizes, shapes and distributions on all positions in the sky. This inevitably lowers the sensitivity for any single specific hypothesis. But as already explained, the main objective of this search is to identify a promising hypothesis which can then be tested by a dedicated follow up analysis. Another negative aspect of such a strategy is that, compared to a search for one specific hypothesis, the interpretation of the result is less straightforward. Multiple plausible interpretations can be found, which may have to be tested in more detail. Finally, instead of identifying a source, an unspecific approach like this may also just reveal a yet uncompensated systematic effect in the measurement. While in terms of knowledge gain on cosmic neutrinos this could be seen as a disappointing result, it would definitely be useful for other upcoming analyses in ANTARES and maybe also similar experiments like KM3NeT [96].

Taking everything into consideration, the model-independent multiscale source search offers a good chance to identify the most dominant sourcelike structure in the data, while, on its own, it's unlikely to confirm such a structure with a high significance.

8.1 Input

Once the events are selected as described in chapter 7, their direction information is converted from local coordinates to equatorial coordinates (see chapter

3.7). This is done with the astro package from Seatray (see chapter 3.5), using only the reconstructed zenith, azimuth and the time the event was recorded. Everything that is actually required as input for the multiscale source search are the declination and the right ascension of each event. Evaluating energy and time too is possible in general, but not included in this search.

8.2 The search space: spherical grids

To evaluate the distribution of neutrino positions a discrete grid is used. This grid approximates a sphere by 165016 gridpoints which have a distance of $\approx 0.5°$ to their next neighbors in declination as well as right ascension. The distribution of the gridpoints is 720 points at a declination of 0°, with a reduced number of gridpoints in other declinations according to the cosine of the angle between the declination and the horizon. This results in rows of gridpoints with less points per row towards the poles. This setup could be replaced by a more optimal scheme, which distributes the gridpoints more evenly, but the current implementation allows to exploit the regularity of the indices of gridpoints to speed up some lengthy computations considerably. The spacing is chosen to be 0.5° because a spacing of 0.25° drastically increases the runtime of the evaluation such that pseudo-experiments cannot be done in a reasonable amount of time any more. Furthermore this spacing corresponds nicely to the expected median angular resolution of about 0.44° achieved with Selectfit and the chosen quality cut. On the other hand using a wider spacing, for instance 1°, unnecessarily reduces the accuracy of the method. The radius of the spherical grid is set to 1.0 to simplify the computations. Figure 37 shows a spherical grid with gridpoints in blue and random events in white[10]. The tiles at each gridpoint have a radius of about 0.5°, hence they overlap and form a closed sphere. The upper part without neutrinos is north (declination of +90°) in equatorial coordinates. The distribution of the events follows the visibility for and acceptance of ANTARES as described in chapter 8.3.3. There will be one of these spheres used separately for the evaluation of each of the 180 distance scales discussed in chapter 8.3.2. According to the sampling theorem as discussed in [100] or [71], using a grid with a spacing of 0.5° theoretically allows to reconstruct information of the scale of 1° or larger without error. This limit is shifted for this grid due to the discontinuities of the gridpoint spacing to around 1.5°. This doesn't mean that smaller scales can't be resolved in this search, but that they can't be guaranteed to be resolved without discretization errors due to the binning.

The distance of two points on the sphere is computed as the great-circle distance, the shortest distance between two points on the surface. To minimize numerical errors the more common equation 8 was substituted by equation 9 for computing the great-circle distance. Here ϕ denotes the declination, λ the right ascension and $\Delta\sigma$ the angle between point 1 and point 2. Since the radius

---

[10] For the visualization of the search spheres and the events on them, the vtk file format [97] is used. The visualization of the vtk files is done with Paraview 4.1 [98]. The skymaps in later chapters are plotted with Gnuplot, mainly Version 4.6 patchlevel 5, see [99].

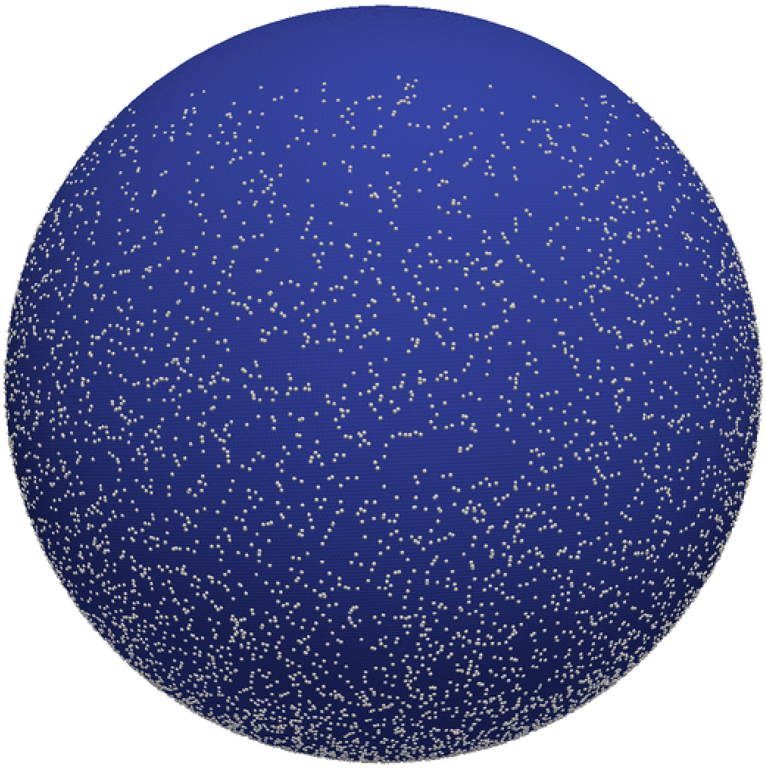


Figure 37: A spherical grid with gridpoints in blue and random events in white. The gridpoints are rendered with a radius of 0.5°, hence they overlap and form a closed sphere. Since the sphere is a three dimensional model, only the front part facing the observer is visible here.

of the spheres is set to 1, $\Delta\sigma$ is also directly the distance measure.

$$\Delta\sigma = \arccos(\sin\phi_1 \sin\phi_2 + \cos\phi_1 \cos\phi_2 \Delta\lambda) \tag{8}$$

$$\Delta\sigma = 2\arcsin(\sqrt{\sin^2(0.5\Delta\phi) + \cos\phi_1 \cos\phi_2 \sin^2(0.5\Delta\lambda)}) \tag{9}$$

8.3 From neutrinos to search spheres

To demonstrate the steps of the multiscale search we look at a distribution of random events with two artificial point sources at a declination of -70°. The setup is shown in Figure 38. To illustrate the steps more clearly the sources are unrealistically strong, one with 18 neutrinos within an area of 0.5° by 0.5°, the other one with 12 neutrinos.

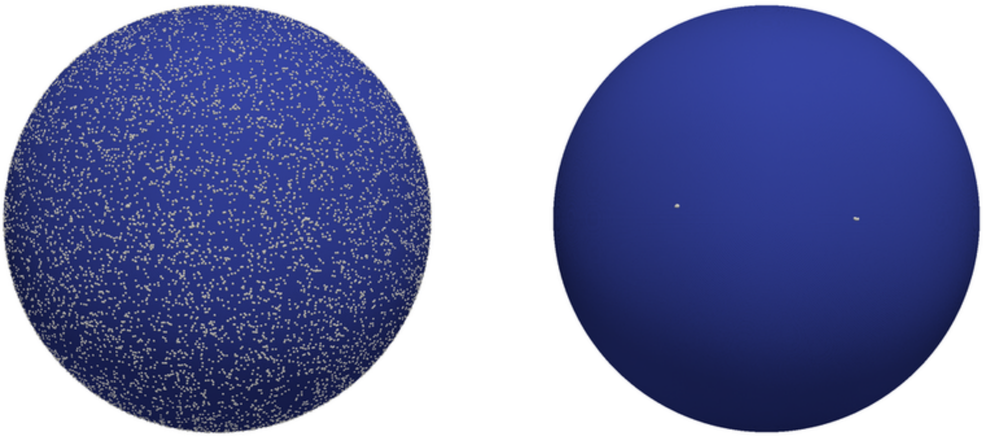

(a) (b)

Figure 38: a) A spherical grid with 12000 random events and two point-like sources. View from below to the south pole (declination of -90°). b) The same setup displayed without the random events.

8.3.1 Counting neutrinos

The search starts by counting the number of neutrinos located in a ring around each gridpoint with a radius corresponding to the current search scale. The first distance scale is from 0.0° to 0.5°.

```
for each gridpoint in grid:
  for each neutrino:
    if 0.0 < distance < 0.5:
      counter = counter + 1
```

For an illustration of this counting scheme see Figure 39. It shows the counting for a search scale of 1.0° to 1.5°. The result of this evaluation is one number for each gridpoint of the search sphere. The results for three scales from the example with two point sources from Figure 38 can be seen in Figure 40.

8.3.2 180 spheres

The process described in chapters 8.3.1 to 8.3.6 is performed for 180 different search scales (from 0° to 90°) in steps of 0.5°. Each iteration i evaluates only the neutrinos in a ring (not the whole circle area) between $i \cdot 0.5°$ and $(i+1) \cdot 0.5°$ around each gridpoint, with index i starting at 0. The result of each iteration is computed independently (except for the smoothing between search distances in chapter 8.3.4) and stored in an independent spherical grid.

8.3.3 Expectation from pseudo-experiments

The next step will be to compute the Poisson probability for the observed number of neutrinos at each gridpoint. To compute this probability one needs the expected number of neutrinos. The value how many neutrinos are expected

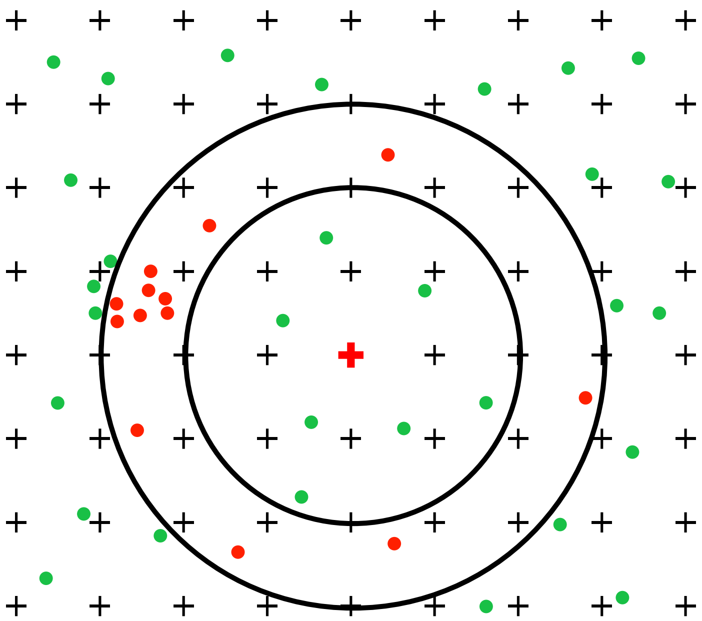


Figure 39: Scheme of the neutrino counting. Crosses mark the gridpoints with a distance of 0.5° between them. Green and red dots are neutrinos. The red cross is the gridpoint that is being evaluated. The current search scale is between the black circles. In this example it is between 1.0° (inner circle) and 1.5° (outer circle). Neutrinos which are found in the current search scale at the current searchpoint are shown in red. In this example the outcome of the evaluation of this scale at the red gridpoint is 13.

around each gridpoint is defined by the visibility of a point in equatorial coordinates for ANTARES, folded with the acceptance of ANTARES for the local coordinates that contribute to these equatorial coordinates, folded with the efficiency of the software reconstruction chain for these local coordinates. Even a complex analytical model is only an approximation for the effects of all these contributions. As an alternative the expectation can be approximated directly from the measured data, which is done in this search. This automatically includes the properties of the acceptance and efficiency of the whole reconstruction chain by design. Additionally, such a strategy doesn't rely on simulations to be "precise enough", nor does it need any manual tuning if for example parts of the reconstruction chain are exchanged.

To obtain the expectation the set of all neutrinos as described in chapter 8.1 is used again. If we were to have abundant statistics and only background neutrinos without any neutrinos from sources, the measurement itself would already be an approximation of the expectation. Since there are only few neutrinos and some of them could originate from spatially confined sources, the data is scrambled first using random times within the data taking period when

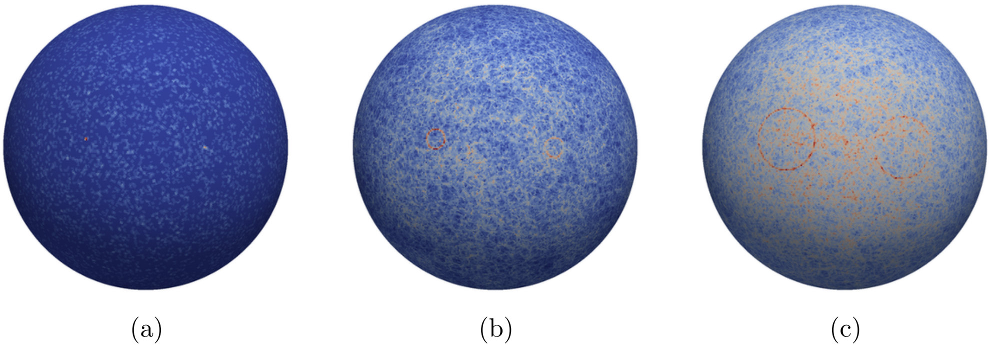

(a) (b) (c)

Figure 40: a) The spherical search grid with the number of events counted in a ring between 0.0° and 0.5° around each gridpoint. b) Number of events between 3.0° and 3.5°. c) Number of events between 10.0° and 10.5°.

computing the equatorial from the local coordinates. Potential source structures in the right ascension are removed by that, placing all events randomly in a declination band. Then the numbers of neutrinos are counted as described in section 8.3.1.

The problem remains that there isn't enough statistics to cover the whole sky with the 165016 gridpoints with the few thousand events. To improve this one can simply compute the mean of the counted values in one declination band. This is justified because in theory the expected number of events in equatorial coordinates should be the same at each point within one declination band, since the rotation of the Earth is supposed to obliterate any possible inhomogeneities. This assumption is by far exact enough at this point to estimate the expectation, but, as described in chapter 9.7, it is not 100% exact. The result is shown as the red curve in Figure 41. Still, these mean values per declination band show a high fluctuation due to the low statistics, especially near the poles where only few gridpoints belong to one band. To reduce the fluctuations of these values, several filters are applied to the vector. These filters are well known for instance from the field of digital image processing as explained in [100], [101] and [71]. The concept of filtering to reduce fluctuations from noise has been used before its application to digital data. The effects of filters on signals of all kinds have been studied extensively in the more general field of signal processing, explained for example in [102] and [103]. The values of the first two bins are ignored by the filtering algorithm in this case due to the large unphysical fluctuations around the poles.

The first applied filter is a so called median filter, a non-linear filter that for each value of the vector vec_i at the index i computes the median of the values $vec_{i-1}, vec_i, vec_{i+1}$ and stores it as the new value for vec_i. This is a common approach for instance in image processing to remove artifacts. The result of this processing can be seen in the green curve in Figure 41.

The resulting expectation still contains an unreasonable amount of statistical fluctuation with a high frequency. Therefore the next step is to use a linear low-pass filter. For linear filters the stencil notation will be used. This notation is the common way to write down discrete linear filters, see for instance [100].

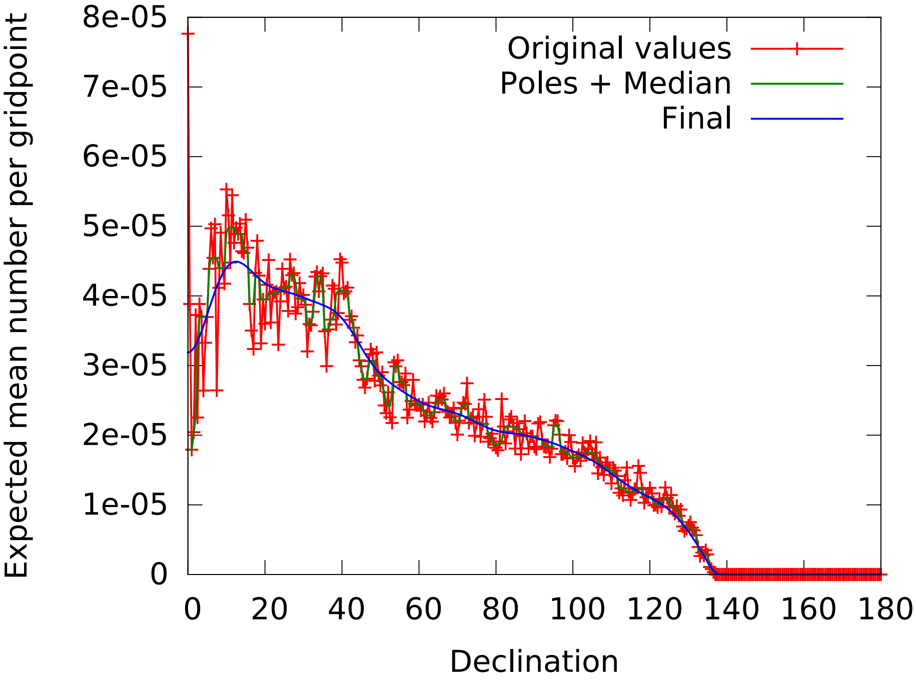


Figure 41: Approximation of the acceptance of ANTARES and the reconstruction chain for the search distance between 0.0° and 0.5°.

For the vector of means and one dimensional data in general, a generic stencil looks like this:

```
weight [w c e]
```

Applying this stencil to a value of the vector at index i, with i being the index of the declination band here, means computing equation 10:

$$value_{(i)} = weight \cdot (w \cdot value_{(i-1)} + c \cdot value_{(i)} + e \cdot value_{(i+1)}) \tag{10}$$

Using such a separable linear filter multiple times equals the one-time usage of a larger filter, for example:

```
twice 1/4 [1 2 1] equals 1/16 [1 4 6 4 1]
```

A brief explanation for this example can be found in appendix C. The benefit is that the larger filter doesn't have to be implemented manually.

In the two dimensional case, for instance when the points of the search sphere are considered to be pixels of an image, a linear filter looks like this:

```
       [nw n ne]
weight [w  c e ]
       [sw s se]
```

The value is computed for the gridpoint at the center of the filter, again denoted with c here. Applying the stencil to a gridpoint at index i, j, with i being the index of the declination band and j the index of the right ascension within the declination band, means computing equation 11:

$$\begin{aligned} value_{(i,j)} = weight \cdot (nw \cdot value_{(i+1,j-1)} + n \cdot value_{(i+1,j)} + ne \cdot value_{(i+1,j+1)} \\ +w \cdot value_{(i,j-1)} + c \cdot value_{(i,j)} + e \cdot value_{(i,j+1)} \\ +sw \cdot value_{(i-1,j-1)} + s \cdot value_{(i-1,j)} + se \cdot value_{(i-1,j+1)}) \end{aligned} \tag{11}$$

The filter which is used here to reduce the statistical fluctuations of the vector of expected mean values is one of the simplest one-dimensional low-pass filters:

```
1/3  [1  1  1]
```

This corresponds to a simple mean of the three values. All of these filters need to treat the borders in a special way, simply because there is no value left of index 0. Due to the spherical nature of this task the best value at the index -1 (which would be one before the first index 0) is the value of index 1, because this is also the value corresponding to the position of the other side of the pole. The stencil for index 0 is then

```
1/3  [0  1  2]
```

For the last index it is the other way around. Simply using this filter would efficiently reduce the statistical fluctuations, but it treats all values as if they were produced with the same statistics. To compensate for this effect an additional tweak is used here. For indices where one of the values is 0.0 the low-pass filter is not applied. This preserves the location of the boundary where the visibility and efficiency drop to zero and therefore prevents distortions which would occur at the tail otherwise. This filter is applied 60 times to obtain a continuous approximation of the expected mean. As explained earlier this corresponds to a one-time application of a larger lowpass filter. The filter is also applied once without the treatment for zero values to include the uncertainty for the measurement for values close to zero. The result of this procedure is shown in Figure 41 as the blue curve. To further reduce fluctuations which could be introduced by the scrambling of the data, the whole process of obtaining an expectation value for each gridpoint is conducted 30 times and for each gridpoint the mean of the 30 computed values is used.

Since this expectation is different for every search distance that is evaluated, this process is performed separately for each of the 180 search distances. The results for 10° to 10.5° are shown in Figure 42. One can clearly see that, due to the increased statistics for larger search rings, the statistical fluctuations are a lot smaller and the approximation is more accurate.

The obtained values of these approximated expectations are then set for every gridpoint according to its declination, resulting in expectations for the spherical grids as shown in Figure 43 for 0, 10 and 50 degrees. The magnitude of the expected values differs as already seen in Figures 41 and 42, therefore the color scale (corresponding to the y-value in the two-dimensional plots) is readjusted here. Red are high values corresponding to an expectation of more event, blue are low values indicating the expectation of fewer events from this direction.

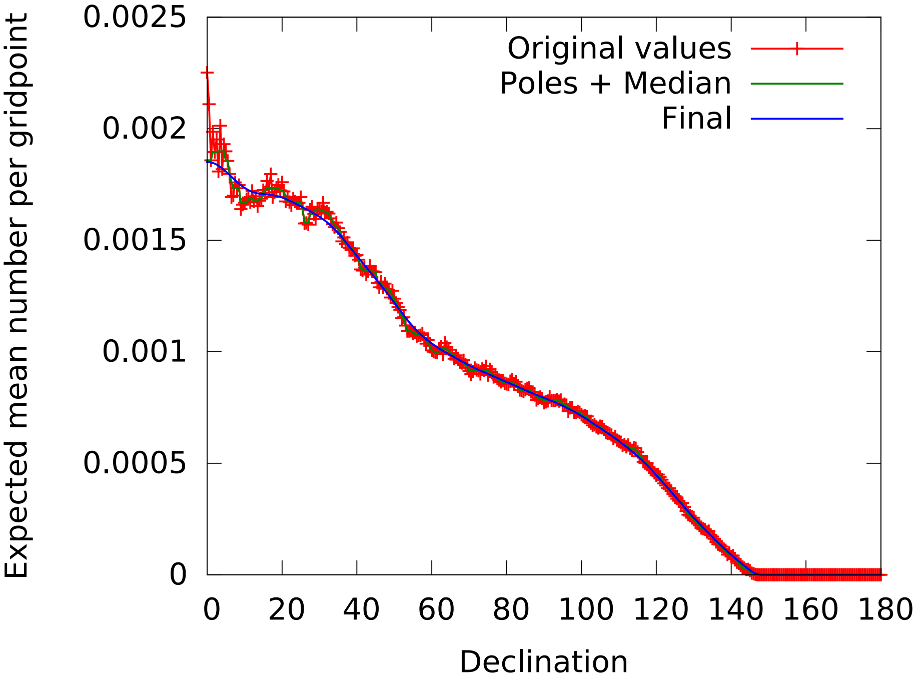


Figure 42: Approximation of the acceptance of ANTARES and the reconstruction chain for a search distance between 10.0° and 10.5°.

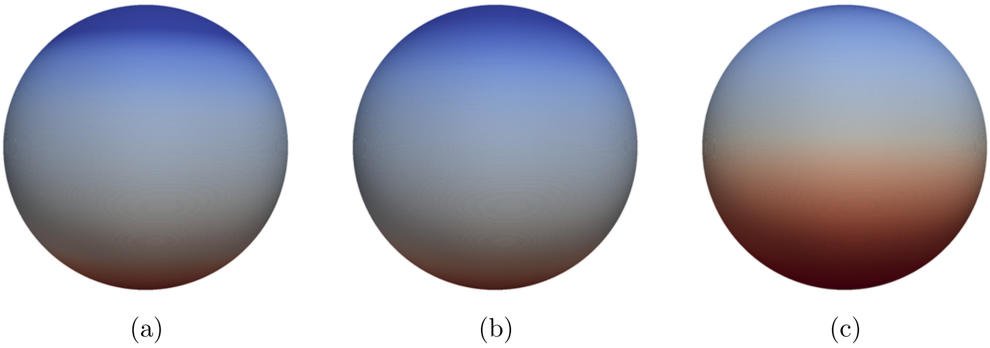

(a) (b) (c)

Figure 43: The distribution of the expected number of neutrinos. a) 0.0° to 0.5° around each gridpoint. b) 10.0° to 10.5°. c) 50.0° to 50.5°. The color scale is rescaled.

8.3.4 Poisson probabilities

The algorithm has counted the number of neutrinos n around each gridpoint as described in chapter 8.3.1 and has an estimate for the expected mean number λ from chapter 8.3.3. With this information the Poisson probability P for n could be computed using equation 12, where e is Eulers number.

$$P(n) = \frac{\lambda^n e^{-\lambda}}{n!} \tag{12}$$

The problem with this computation is that it is numerically unstable for large values of n or λ. This instability not only leads to slightly wrong probabilities, but even to the failure of the computation. A possible solution to this is to take care of the exponent and mantissa separately as in equation 13. Another idea is to use the datatype long double[11] and the corresponding library functions with a higher precision for the computations. But both of these approaches only shift the range of the instability to higher values, they don't get rid of it entirely. Combining the two optimizations shifts the value where the computation of $P(n = \lambda)$ breaks down from 27 for float and 144 for double precision to 1547. Although this is enough in most cases of this search,

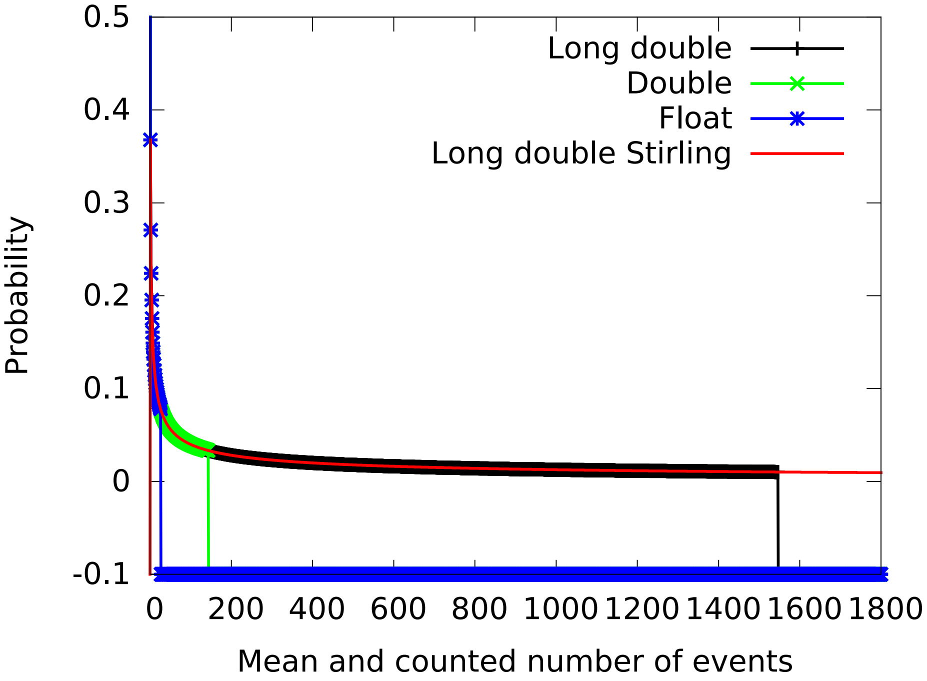


Figure 44: Comparison of the implementations and approximations of equation 12 for $P(n = \lambda)$. A probability of -0.1 denotes a failed computation.

---

[11] Using g++ (SUSE Linux) 4.3.4 on the Woodcrest cluster in Erlangen long double gives 128 bit precision.

an approximation using the so called “Stirling formula” as denoted in equation 14 is introduced for larger values.

$$P(n) = \frac{\frac{e^{-\lambda}}{\log_{10}(e^{-\lambda})}\frac{\lambda^n}{\log_{10}(\lambda^n)}}{\frac{n!}{\log_{10}(n!)}} \cdot 10^{\log_{10}(e^{-\lambda})+\log_{10}(\lambda^n)+\log_{10}(n!)} \tag{13}$$

$$P(n) = \frac{e^{n(1+log_{10}(\frac{\lambda}{n}))-\lambda}}{\sqrt{2\pi(n+\frac{1}{6})}} \tag{14}$$

Since the deviation of this approximation from the true values vanishes already around $P(n = \lambda = 20)$ it is used in all cases where either n or λ is above 120. The behavior of the different implementations is shown in Figure 44. The final result using long double precision for computation gives exact results in terms of double precision and has been confirmed numerically stable until $P(n = \lambda = 10^7)$, which is definitely enough for much larger searches.

The algorithm then computes the probability for the observed or a greater number of neutrinos using equation 15 at each gridpoint.

$$P(x \geq n) = 1 - \sum_{i=0}^{n-1} P(i) \tag{15}$$

For historical and technical reasons the Poisson probabilities are then rescaled as in equation 16.

$$R = log_{10}(\frac{1}{P(x \geq n)}) \tag{16}$$

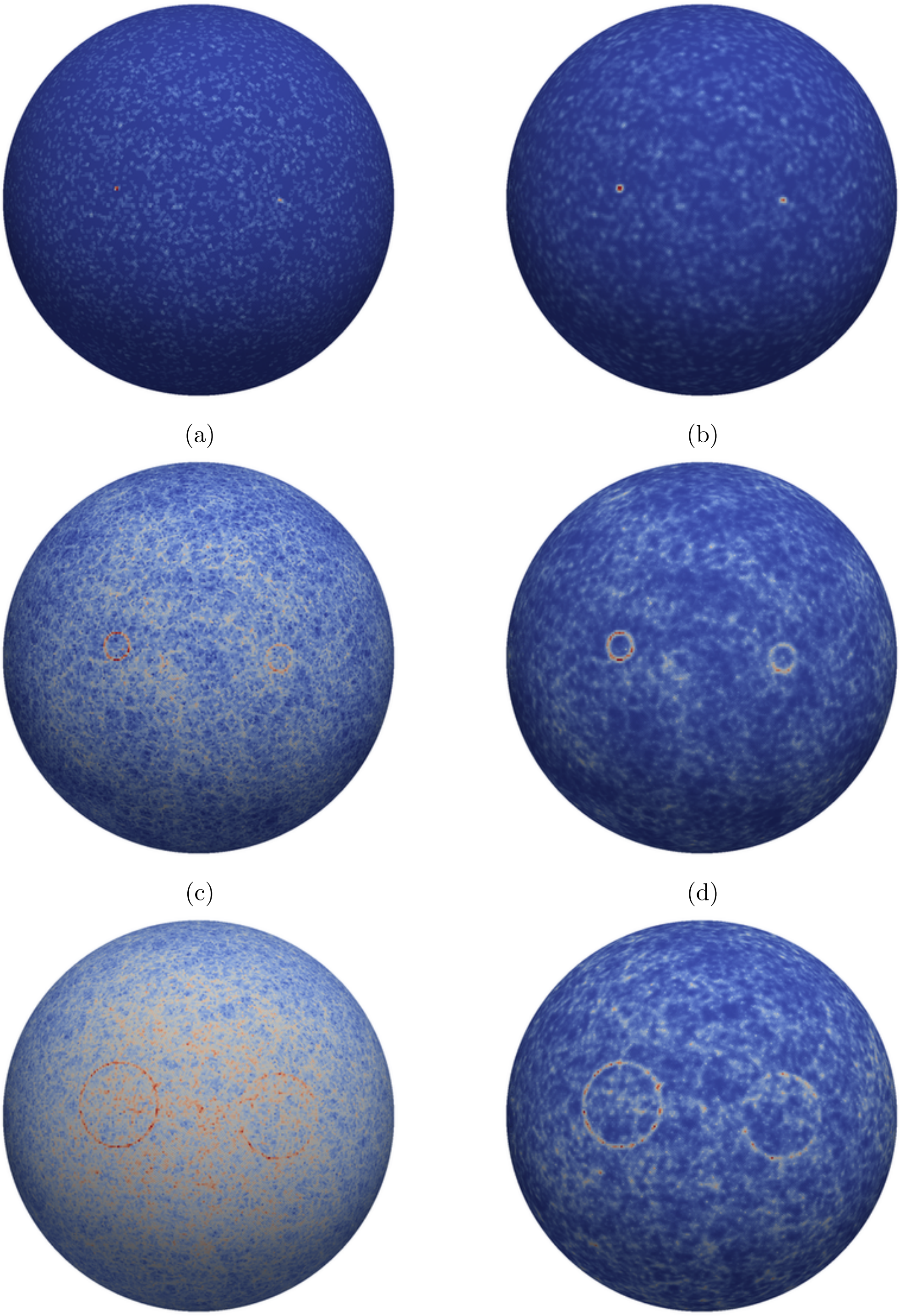
To reduce statistical (only few events) and systematical (event reconstructions have errors) fluctuations, the R values are low-pass filtered on the whole sphere. The assumption is that “real” sources are more likely to extend to neighboring gridpoints, whereas for random fluctuations the probability to have another random fluctuation at a neighboring gridpoint is low. The stencil used for this filtering is:

```
     [1   1   1]
1/20 [1   12  1]
     [1   1   1]
```

Two dimensional filters have been explained in chapter 8.3.3. This filter focuses strongly on the central value for each gridpoint, adding only 40% of neighboring information. This is motivated heuristically, because it preserved the signal in various test cases. Finally a filter is applied to smooth between the different search distances. The motivation for this is that a source which is prominent in a search distance of for example 1.0° will likely be observed when searching for 1.5° too, while random fluctuations are less likely to extend to neighboring search distances. For each gridpoint the neighboring values are from the searches with a 0.5° smaller and 0.5° larger search ring. The stencil used is:

```
1/6 [1 4 1]
```

The effect of these computations on the search spheres is shown in Figure 45. Red means high values, blue means low values. The color scale is readjusted between the different search scales.



(a) (b)

(c) (d)

(e) (f)

Figure 45: a) The spherical search grid from Figure 40a.
b) The grid after the computations described in chapter 8.3.4.
c) and d) analog for the search radius between 3.0° and 3.5°.
e) and f) analog for the search radius between 10.0° and 10.5°.

8.3.5 Segmentation

At this point of the analysis spheres have been computed where high R values at a gridpoint indicate more neutrinos in a certain distance around the gridpoint. Therefore potential source regions containing more neutrinos should be linked to high values on the spheres. We can assume that low values are not linked to detectable sources. In pattern recognition separating background from potentially relevant information is called segmentation. A lot of information on segmentation can be found in [100]. The chosen solution for this task in this analysis is probably the simplest possible solution, applying a threshold to all R values. Similar to the segmentation of a black and white image, a histogram of the values of all gridpoints is analyzed to obtain the threshold. The distribution as shown in red in Figure 46 has a long tail of non-zero values, which is where potentially signal-related information is contained. If only the maximal values are considered, the search still shows a good potential to find the most dominant source, but a lot of information on the shape of the source and on other less dominant sources is lost. To extract the tail of the distribution robustly, the first step is to cut away the lower 50%. The result is shown in green in Figure 46. Only very low values close to 0 are discarded by this. The next step is a smoothing of the histogram 50 times with the stencil

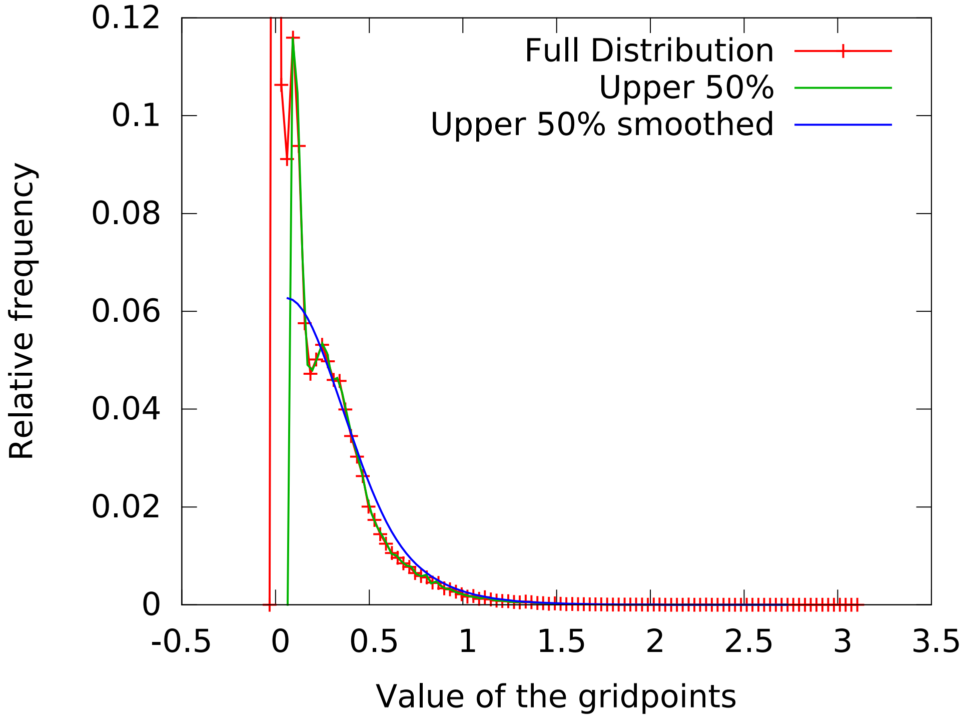


Figure 46: The histogram of the R values on the grid, the distribution of only high R values and the smoothed distribution of high values.

$1/3 \quad [1 \quad 1 \quad 1]$

The result of this filtering is shown as the blue curve in Figure 46.

In the next step a line is calculated from the point where the smoothed distribution of the upper 50% has its maximum value $maxVal$ to the first point on the right side where the y-value is below $0.5 \cdot maxVal$. The x-value where this line intersects y=0 is used as a threshold θ. This is illustrated in Figure 47. Only values in the tail right of this threshold are kept, lower values are set to zero. The advantage of computing the threshold like this compared to a fixed threshold or a fixed percentage is that the threshold adapts to the distribution. The effect of the segmentations on the search spheres is shown in Figure 48.

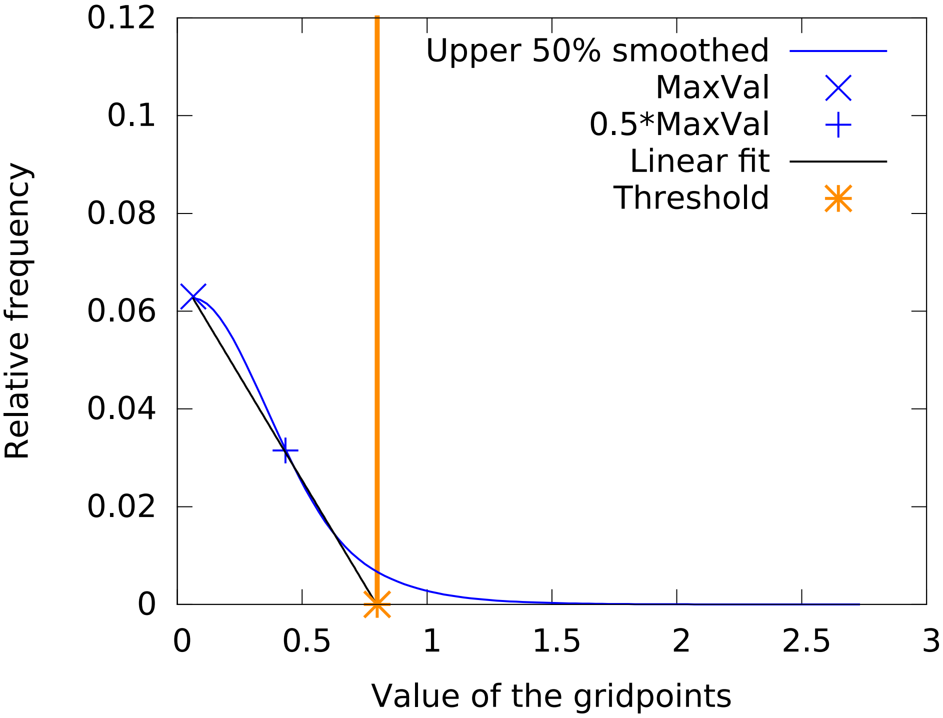
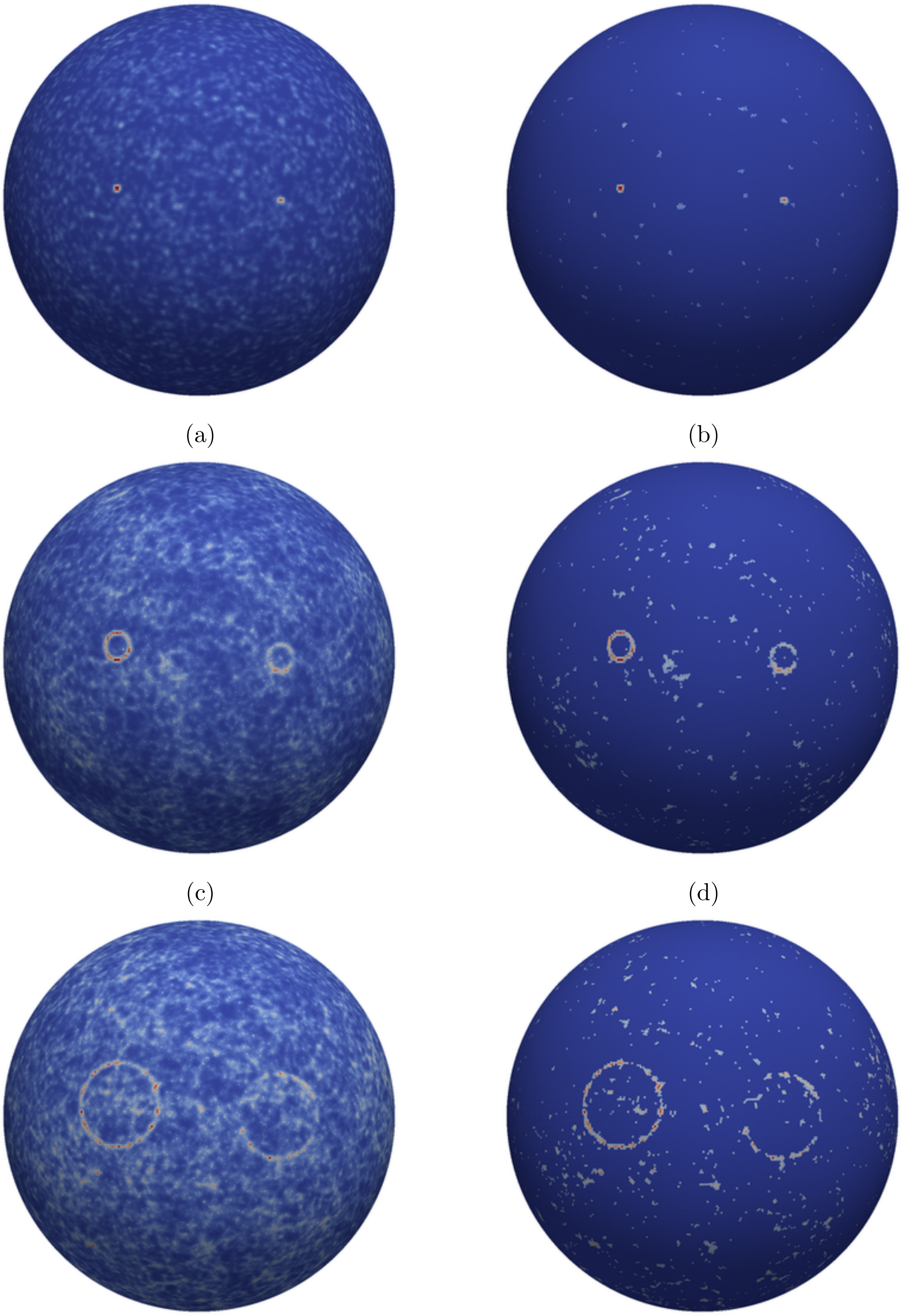


Figure 47: A fit from the maximum trough the half-maximum point determines the threshold.

An additional option, which will be used later, is to scale the distance between the previous minimum value x_{min_old} and θ, using a factor of α. The new threshold is then given by equation 17.

$$\theta_{final} = (\alpha \cdot x_{min\_old} + (1 - \alpha) \cdot \theta) \tag{17}$$

For now *alpha* is set to 0 and therefore $\theta_{final} = \theta$.



(a) (b)

(c) (d)

(e) (f)

Figure 48: a) The spherical search grid from Figure 45b.
b) The grid after the segmentation described in chapter 8.3.5.
c) and d) analog for the search radius between 3.0° and 3.5°.
e) and f) analog for the search radius between 10.0° and 10.5°.

8.3.6 Remapping

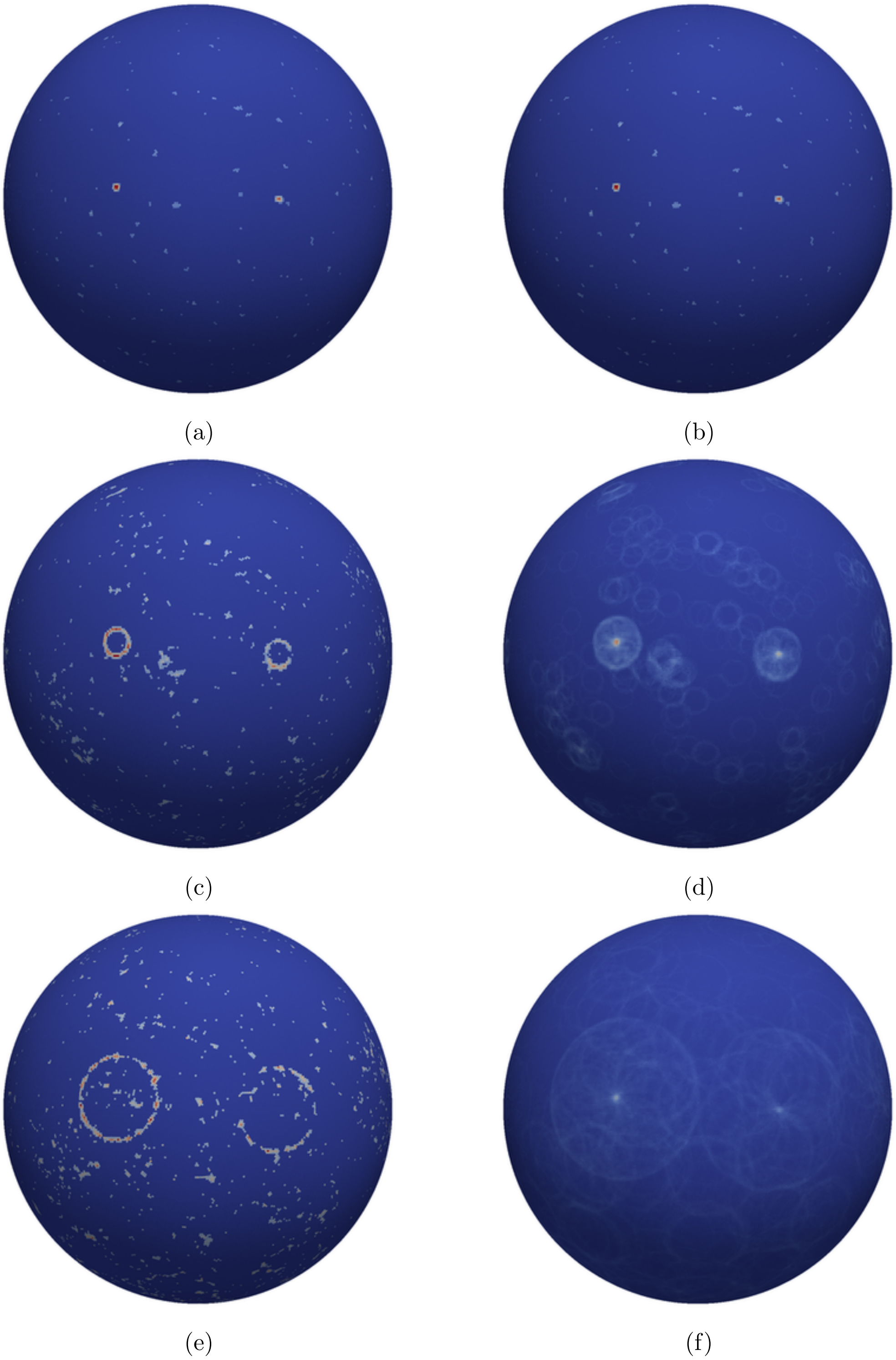
At this point the search has obtained spheres where only potentially signal related gridpoints have values different from 0.0. The next step is to reconstruct the location of the neutrinos that caused the detected overfluctuations. For the search sphere with a search distance between 0.0° and 0.5° the situation doesn't change, since the information up to 0.5° is already located at the exact gridpoint it originated from when counting the neutrinos. For all other spheres with search distances $d > 0.5°$, the information stored at a gridpoint originated from counting neutrinos that are d degrees away from the gridpoint. To evaluate where the overfluctuation actually occurred, this information must be mapped back to the origin. This is done using a copy of the current spherical grid with all gridpoints initially set to 0.0. Then the following steps are performed[12]:

```
for each gridpoint $p_{(i,j)}$ in the original grid:
  find the set $p_d$ of all gridpoints in distance $d$ around $p_{(i,j)}$
  for each gridpoint $p$ of $p_d$ in the new grid:
    $value_p = value_p + \frac{value_{p_{(i,j)}}}{size(p_d)}$
```

Afterwards the new grid contains the corresponding fractions of the overfluctuations mapped back to their origin and this grid is used from there on. The effect of these computations on the search spheres is shown in Figure 49. The remapping distributes the information of an overfluctuation at one gridpoint back to all gridpoints where the neutrinos causing the overfluctuation could have been. The information where the neutrino distribution had a higher density is therefore automatically encoded in the pattern how the remapped circles overlap, see for example Figure 49d. As already stated in the beginning, only marginal differences are observed for the first grid as seen in Figure 49b. The color scale for the remapped plots on the right is not readjusted between the search scales. One can see that the point sources are detected most dominantly by the contributions from smaller scales.

---

[12] The naive implementation of this step, as in the explained scheme, has an extremely long runtime. Therefore the actual implementation differs slightly from the explained scheme to exploit a much more cache-coherent access pattern. In other words: The optimized remapping performs the same operations in a more efficient ordering.



(a) (b)

(c) (d)

(e) (f)

Figure 49: a) The spherical search grid from Figure 48b. b) The grid after the remapping described in chapter 8.3.6. Nothing changes up to 0.5°.
c) and d) analog for the search radius between 3.0° and 3.5°.
e) and f) analog for the search radius between 10.0° and 10.5°.

8.4 180 to one search sphere

The current status of the search are 180 spheres, each representing the search for overfluctuations in the number of neutrinos at a different scale. Each of these searches is left with only the most relevant results for its distance.

A reasonable next step would be to search for structures in these 180 spheres, taking into account the neighbors within a search sphere as well as the neighboring search spheres. This has been implemented with two different clustering approaches, but none could be optimized well enough to robustly and reliably find the test sources. There were multiple issues with these approaches, for instance when to optimally stop a cluster from growing in which direction (within one sphere must be treated differently than between spheres), how to map the three-dimensional clusters to a two dimensional skymap in the end and several more. For all of these problems there were solutions found, but the search didn't reach a satisfactory performance.

The better and feasible solution turned out to be to fuse the 180 search spheres into one. It is not a priori clear what value to assign to each gridpoint in the resulting sphere when we have a vector of 180 values. Several methods how to do this have been tested, for instance assigning the

- maximum value
- median value
- mean value
- sum of all values (L1 norm)
- L2 norm of all values (see equation 18)
- L3 norm of all values
- number of connected non-zero values
- and several others.

The Ln norm for a vector v is defined as:

$$L_n = \sqrt[n]{\sum_{i=0}^{N} v_i^n} \tag{18}$$

with N being the number of entries in the vector. Although taking the maximum value might seem like a reasonable idea, this is not robust as many gridpoints were observed to have a random overfluctuation somewhere in the 180 values. What worked best in terms of robustness and source detection for the scenarios tested during development was the simple sum and the L2 norm of all values. The reason is that significant sources tend to be found on multiple scales, resulting in many relatively high non-zero values in the sum. Random fluctuations on the other hand can often be observed on a single scale only. The ordinary sum was chosen for highest robustness and a good ability to identify source locations. During this optimization it was discovered that not including the first evaluation bin (from 0.0° to 0.5°) in the sum further enhances the stability of the overall search. Therefore it is not included.

It should be noted that the potential of this step could be exploited even better, either by evaluating the full 180 spheres or by a more sophisticated way of combining the 180 to one. But the devised algorithm is the best achieved solution. The result of the summation can be seen in Figure 50.

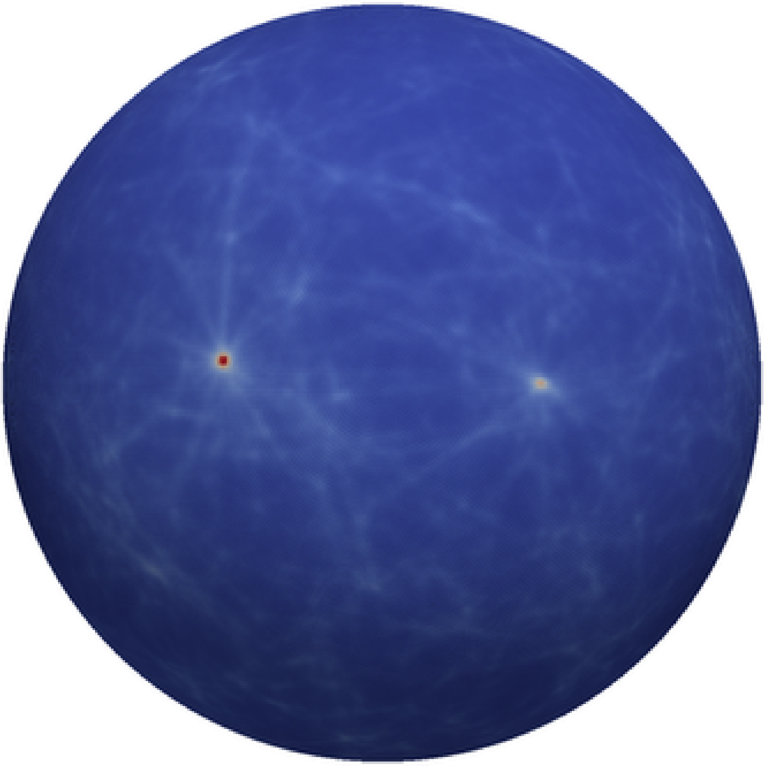


Figure 50: The sum of the 180 spheres

8.5 From one search sphere to clusters

Now there is only one search sphere left containing the combined information on neutrino density from all search scales. A source could now either be a very high value on that sphere or a larger region with "higher-than-usual" values or a combination of both. To identify these regions where "higher-than-usual" values are present, a segmentation can be performed exactly as described in chapter 8.3.5. This time α is varied, resulting in different numbers of remaining gridpoints. $\alpha = 1$ means not cutting away anything at all, a lower α means a harder cut. $\alpha = 0$ is the same threshold as used for the individual spheres, but also harder cuts with negative values are possible. One can see the effect of different thresholds for segmentation in Figure 51. Higher positive values introduce additional artifacts but also allow extended source structures to be found. Hard negative cuts only preserve the highest peaks, which works well for the presented unrealistically strong point sources. Another illustrative example for the effect of different segmentation thresholds can be found in appendix D.

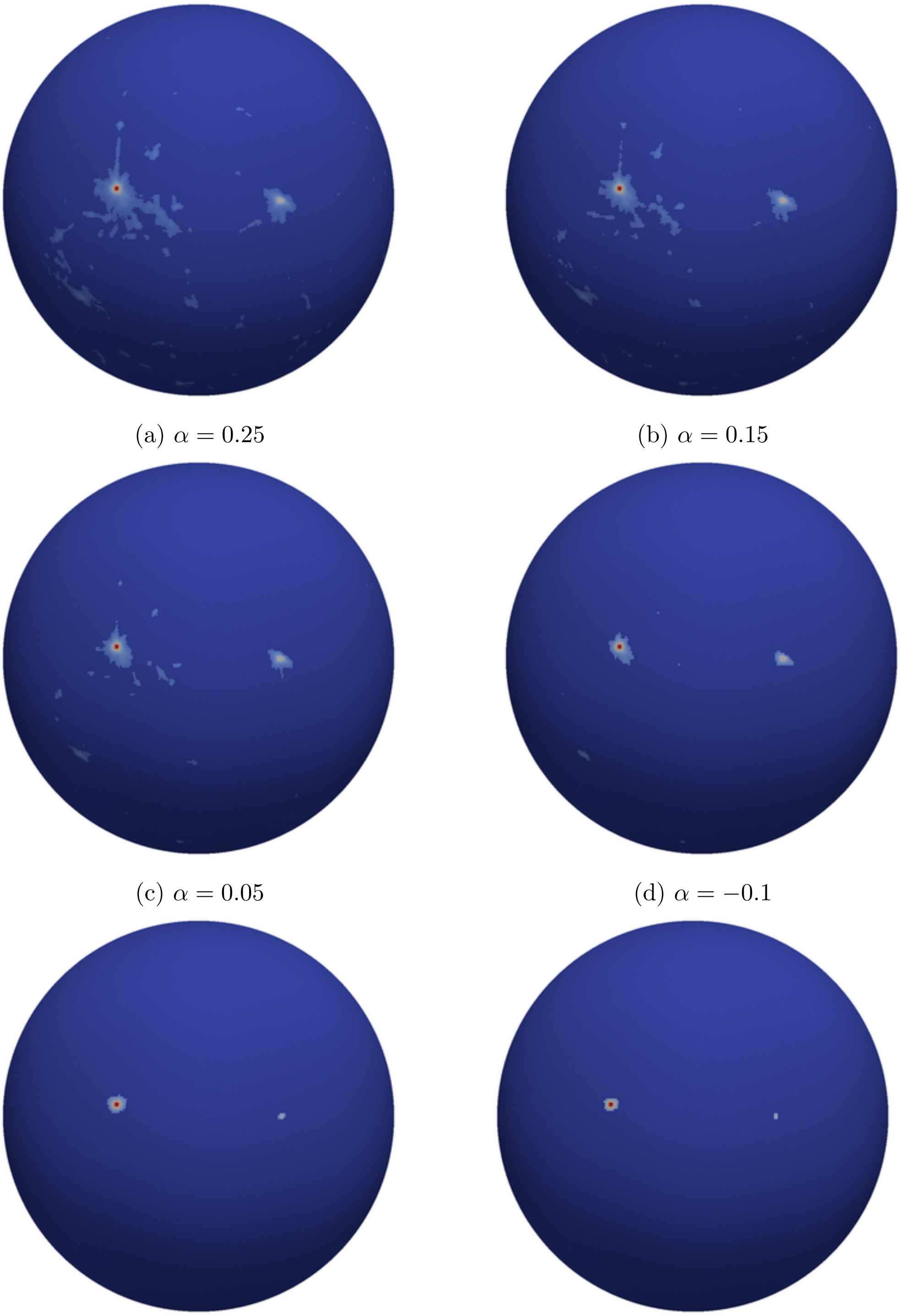
Because all values of the whole sphere are taken into account for the segmentation, this is a global approach. In contrast to that also local approaches that focus on the change of the value next to a cluster (for example via edge detection and heuristics) or the mean and variance of the values from within a cluster and the surrounding neighborhood of the cluster have been tested. The problem with these local methods was that faint random clusters in neutrino-poor regions had an unreasonably high chance of surviving the segmentation, whereas relatively obvious extended sources without clear contours were overlooked often. The segmentation could certainly benefit from exploiting the local structures in a reliable way instead of only using the global distribution, but the global approach is the best found solution that performs well and is guaranteed to show robust behavior.

This step can introduce a bias to the search. Cutting softly results in keeping many gridpoints for the rest of the search, which prefers extended sources. Cutting harshly on the other hand biases towards smaller sources and removes faint but large sources completely from the remaining sample. Therefore multiple segmentations are used and evaluated independently. This introduces an additional trial factor that has to be accounted for in chapter 8.6.2, but it also allows the search to continue as unbiased as possible. Nevertheless, due to the trial factor, not too many different segmentations can be kept for this analysis.

The next step applied to the segmented grid is a non-linear median filter. This is a common approach in image processing to remove artifacts as explained in [100]. Similar to the linear two dimensional filters already introduced in chapter 8.3.3, the median filter also operates on the current gridpoint and its direct neighbors. But instead of writing a weighted linear combination of the considered values to the central gridpoint, it writes the median of the nine values. The effect of this filtering is shown in Figure 52.

The gridpoints of the resulting segmented spherical grid are then checked for connectedness. A connected group of gridpoints is called a segment in the general context of segmentation and in this work it is also called a cluster. All gridpoints are checked in the order from highest to lowest value and each time the considered gridpoint is directly connected to an already found cluster, it is added to this cluster. Connected here means that one of its eight neighbors is already part of the cluster. If it is not connected to any cluster, a new cluster is started. Once all gridpoints are assigned to clusters, all found clusters which have directly neighboring points are fused together.

After this is done all remaining gridpoints are contained in clusters. They have the same extensions as the segments in Figure 52.



(a) $\alpha = 0.25$

(b) $\alpha = 0.15$

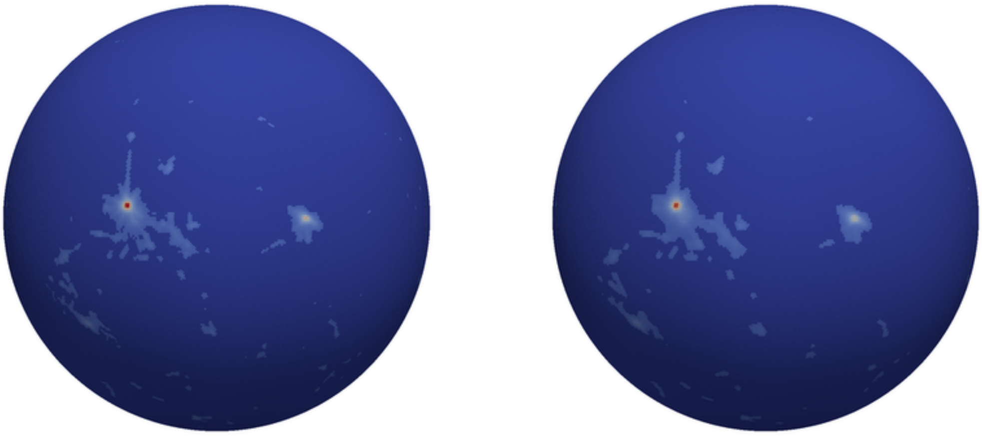
(c) $\alpha = 0.05$

(d) $\alpha = -0.1$

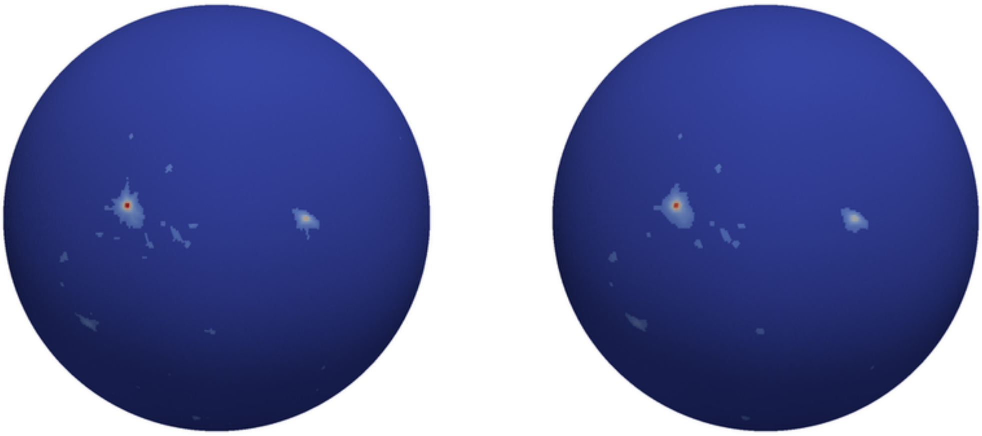
(e) $\alpha = -0.4$

(f) $\alpha = -0.9$

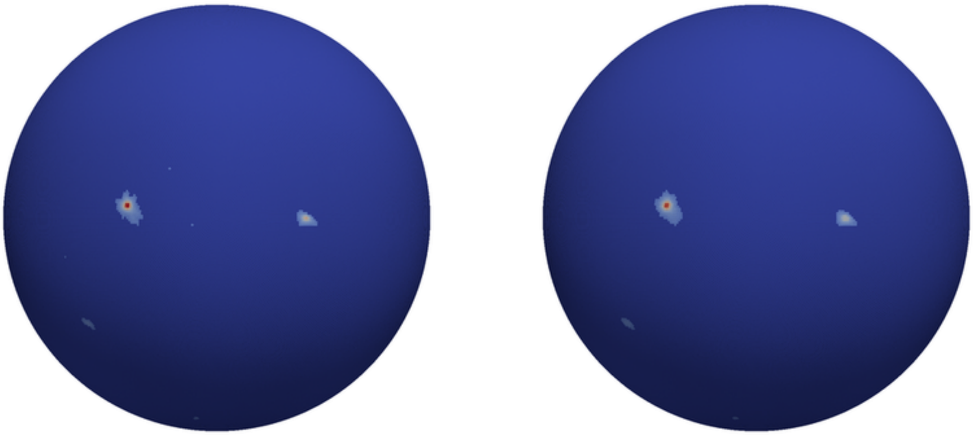
Figure 51: The effect of different α values on the segmentation result.



(a) $\alpha = 0.25$ (b) $\alpha = 0.25$ and median filter



(c) $\alpha = 0.05$ (d) $\alpha = 0.05$ and median filter



(e) $\alpha = -0.1$ (f) $\alpha = -0.1$ and median filter

Figure 52: The effect of the median filter on the segmented grids.

8.6 From clusters to significances

At this stage the information is condensed to clusters. Each cluster is composed of a vector of the gridpoints that are part of it. The next step is to distinguish the potentially significant sources from random accumulations. To do this one needs to know the probability how likely a cluster could have been generated by random events. This probability could in theory be easily determined by pseudo-experiments using scrambled data. The problem with this approach is, that requiring *exactly* the same cluster at the same position with the same gridpoints and the same values for each gridpoint is certainly not going to happen in a bearable amount of time. And to compute a probability we would need many occurrences of exactly this cluster. So this is not an option.

8.6.1 Relevance of a cluster

We can however describe the clusters using various evaluation metrics. For instance one could compute the probability for a cluster of the same size or larger, with size measured by the number of gridpoints. With a fair amount of pseudo-experiments there will likely be a sufficient amount of smaller and larger clusters to estimate the probability for the size of an observed cluster. To make sure that this probability is not confused with the significance of a cluster, we call the value computed in such a way “relevance”. It can be interpreted as an estimation of the pre-trial significance, but it uses approximations and implicitly already contains some trial factors for the search in multiple scales and locations. The purpose of the relevance metrics is to provide a measure with which the clusters can be compared with all other clusters from all segmentations.

Only considering size to measure the relevance neglects smaller or even point-like sources. A good metric to find these could be the maximal value of any gridpoint of the cluster. But also many other metrics can be defined to identify outstanding clusters, each with a certain, sometimes not obvious bias. An incomplete list of the tested metrics can be found in appendix E.

Apart from basic, directly computed relevance metrics one can also use meta-metrics that rely on other relevance metrics. This can be done for instance by direct linear combination of basic metrics. The idea behind these combinations is that a dominant source could stand out in multiple certain metrics. Taking this idea one step further leads to even more general meta-metrics, but at some point these computations become so variable that sources without a clear common trait can become hard to distinguish from random clusters.

If all relevance metrics were to be used for the evaluation of the found clusters, the potential to achieve a high significance would drop considerably, because for the metric that might actually identify a source with a high significance there is a large number of other metrics and each metric has a chance to overfluctuate for a random accumulation. But on the other hand if we chose too few metrics we might introduce a large bias and maybe miss a source in the data. Since the intention of this search had been not to optimize for a specific source model, the selection had to be done heuristically, loosely optimizing for all sources that have been included in the pseudo-experiments during the development.

While a combination of two metrics, metrics 0 and 1 in appendix E, was found to be performing very well for all kinds of sources as shown in chapter 8.7, only metric 0, the size of the cluster in gridpoints, has been used for the data analysis of ANTARES data to reduce the resulting trial factor. By this choice the search becomes biased towards larger extended sources[13], but this has been a deliberate choice, justified by the fact that ANTARES has already conducted sophisticated searches for small and pointlike sources, see for instance [91]. It is relevant to note that the choice of suited metrics is influenced by the choice of segmentation thresholds from chapter 8.3.5 and vice versa. Together with this explicit bias towards larger clusters and the resulting choice for this metric, the segmentation thresholds α have been fixed to -0.11 and +0.25. A list of the tested setups that have been used to derive these choices is contained in Appendix G.

For this selection of metric and segmentations several thousand pseudo-experiments with scrambled data are conducted. Each pseudo-experiment computes the whole process as described in chapter 8 up to this point. Since one pseudo-experiment typically contains many clusters, this results in tens to hundreds of thousands of comparison values for each metric. But only the maximum value for any cluster in any metric is considered in each pseudo-experiment, so that in the final analysis a cluster will be compared to the distribution of the achieved maxima instead of the distribution of all values. This proved to be more reliable than comparing against the whole distribution, because there are different systematics for the highest relevance value of a cluster in each pseudo-experiment than for the second or third highest, since their possible relevance values can be influenced by the highest scoring cluster in this pseudo-experiment. The comparison with the distribution of the maximal values leads to a correct estimation for the most relevant cluster in each pseudo-experiment, but to an underestimation for clusters with lower relevance values in the same pseudo-experiment. This solution is used because the aim of this search is the most significant structure in the sky. The found clusters can now be compared in each metric with a distribution of values from pseudo-experiments.

In this context a few definitions are useful. The p-value denotes the probability how likely something occurs. At this stage for example a p-value of 0.2 for a cluster is interpreted that in 20% of the pseudo-experiments a cluster with the same or a higher relevance value has been found in the same metric. Significances are given in σ, which is derived from the standard deviation of a Gaussian distribution. Therefore $\approx$ 68.27% of all observed values are expected within a 1.0σ interval around the mean value for a Gaussian distribution, about 95.45% within a 2.0σ interval and in general a fraction of $erf(\frac{N}{\sqrt{2}})$ within Nσ. erf is the (Gauss) error function here. In the two-sided interpretation it is irrelevant whether a value is higher or lower than the mean value. Since this

---

[13] The search would still be able to detect unexpectedly strong small sources, but with a considerably reduced sensitivity. The evaluation of multiple scales results in an extended cluster area around the actual small location (see Figures 49d and 52b) for strong sources.

search is aiming for overfluctuations only, the significances are interpreted in a one-sided way, meaning that low values are ignored and only high relevance values result in high significances. Hence 1.0σ instead means that in 68.27% of the pseudo-experiments no relevance value as high as the observed one has been found. For example 2.52σ corresponds to a p-value of 0.0117, approximately one in 85, 2.14σ to 0.0324 or about one in 31.

Directly comparing a cluster to the exact obtained distribution of relevance values to compute the p-value only allows certain, discrete p-values to occur. This means that two clusters with different relevance values may be mapped to the same p-value if no value is between them in the distribution used for comparison, making their relevances indistinguishable. The possible minimal p-value would also be limited by the number of pseudo-experiments.

Therefore the tail of the distribution of relevance values is substituted by a fit, following an exponential decay in the form of $a \cdot e^{bx}$, which is the tail of a Poisson distribution. The result can be seen in Figure 53. The fit is computed based on the part of the tail of the distribution where the green fit is also shown in this plot, but it is only used for high values where the actual statistic is not sufficient anymore, indicated by the threshold value in black. Up to this value the actual distribution is used to evaluate the significance. The threshold in black, from where on the fit instead of the plain histogram is used, is calculated as $0.8 \cdot beginning\_of\_the\_tail + 0.2 \cdot value\_where\_third\_to\_last\_zero\_gap\_occured$. If there are less than three gaps in the histogram with a zero value, the beginning of the first present gap is used instead. This heuristic is used to make sure that a range of the distribution with stable statistics is used to fit the tail, regardless

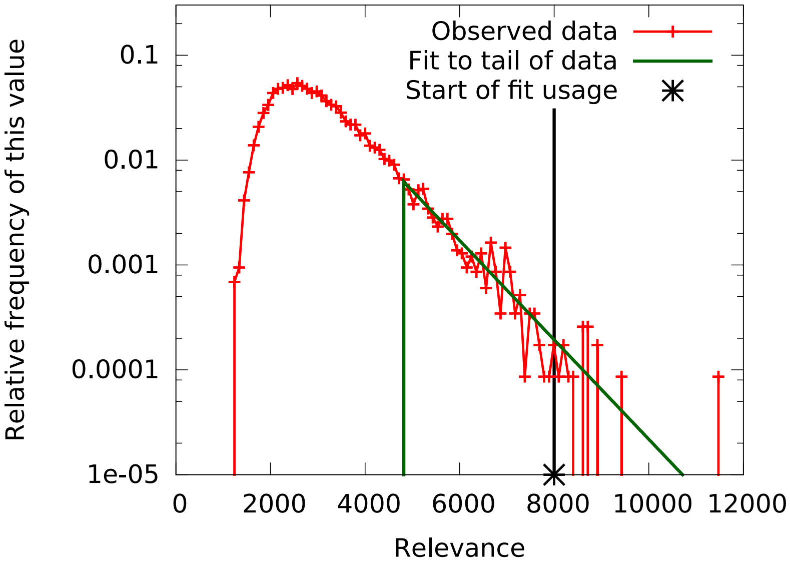


Figure 53: An example for the distribution of relevance values and the fit to this distribution. The fit is computed based on the whole tail of the distribution, but it is only used for values greater than the threshold marked in black.

of binning effects and poor statistics.

After the comparison of the relevance values with these distributions each cluster has one p-value for every metric, expressing how likely it could have been produced by chance according to this metric. The overall p-value for a cluster is the minimum of the p-values from all metrics. From that p-value a renormalized relevance (similar to a significance) is computed using the inverse error function[14].

8.6.2 Significance of a cluster

At this stage we have clusters on the search sphere, each cluster with a value called renormalized relevance which states how likely it could have been produced by a random neutrino distribution. The relevance would equal the significance if we had used only one relevance metric and only one segmentation. But there is a trial factor for the additional possibilities that the observed renormalized relevance value can be produced in any segmentation (and by any metric if more than one is used). The corresponding trial factor could simply be applied now, but since it is a conservative upper bound and the different segmentations are highly correlated, we can give a more accurate estimate by conducting pseudo-experiments. A large number of pseudo-experiments with scrambled data would have to be conducted again. Each experiment would have to compute the whole process described in chapter 8 up to this point again. But almost the whole process has already been computed in chapter 8.6.1 to compare the relevance values to the distributions from pseudo-experiments. The only additional step is the computation of the renormalized relevance values for each cluster. So the results from the previous pseudo-experiments are reused, the few missing computations are added and we have a large distribution of renormalized relevance values from pseudo-experiments.

Just as for the relevance values in chapter 8.6.1, in a direct comparison with the observed distribution to compute the p-value, the number of pseudo-experiments limits the possible maximal significance. Therefore the tail of the distribution of relevance values is again substituted by a fit following the decay of the tail of a Poisson distribution. The result can be seen in Figure 54. Just like for the fit of the plain relevance values in Figure 53, the actual histogram is used up to the threshold shown in black. The fit is only used for values above this threshold and hence does not at all influence any of the significances quoted in this thesis. The strong fluctuations in the left of the histogram in red aren't statistical fluctuations but systematic binning and combination effects. They are caused by the combination of the, in this case two, independent distributions from the two used segmentations, since for every entry in the histogram, the maximum renormalized relevance value found for this pseudo-experiment is used.

Finally, the distribution with the fitted tail from Figure 54 is used to compute a p-value for a found cluster. It is a one-sided p-value, since only overfluctuations are considered in this search. The same method could also search for

---

[14] Using the boost library, see [104]. A scaling factor of $\sqrt{2}$ has to be applied to the results of this function to obtain the correct results.

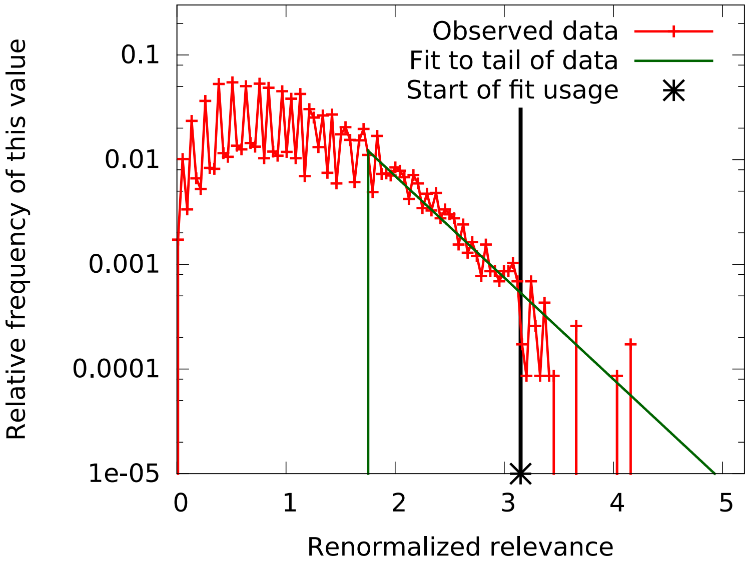


Figure 54: Fit to the distribution of renormalized relevance values used for the computation of significances.

underfluctuations if several maximum operations were substituted by minimum operations. The obtained p-value is converted to a significance for this cluster using the inverse error function from [104] again. The resulting significance values correspond to the standard way of converting p-values to significances in Astroparticle physics, assuming a Gaussian distribution and a two-tailed evaluation.

8.7 Sensitivity

In this chapter the sensitivity of this method is illustrated for various source scenarios. Of course by design not all possible source morphologies that can be detected by the method can be covered, but instead several basic setups are shown to give an idea of the performance. Two relevance metrics, the size of the cluster N (Metric 0) and the average of the highest $\sqrt{N}$ values in the cluster (Metric 1), are used for these plots. The sensitivities are calculated for pseudo-experiments with 13000 random background events with the quoted number of source neutrinos artificially added. Each of the spherical grids in this chapter is rotated such that it presents a reasonable view on the current source(s). The sensitivity curves are smoothed, since only 20 repetitions per data point could be conducted due to time constraints. The visualizations of the setups are shown with the maximum number of events per source tested in the corresponding sensitivity plots. To clarify explicitly, in this chapter the given sizes of the benchmarked sources are their radius, not their diameter.

8.7.1 Small and pointlike sources

Since dedicated, model optimized searches didn't find clear indications for pointlike sources, it is not to be expected for this search either. Figure 55 shows a setup with one almost pointlike source with a radius of 0.5° at a declination of -70°. The corresponding sensitivity curve in Figure 56 shows, how likely it is to detect the point source with a significance of 3σ, depending on the number of events randomly distributed within the source. This means that a point source of this size and location needs 12 events to have a chance of 50% to be detected with a significance of at least 3σ. Figures 57 and 58 show the corresponding information for a source with a radius of 1.0°.

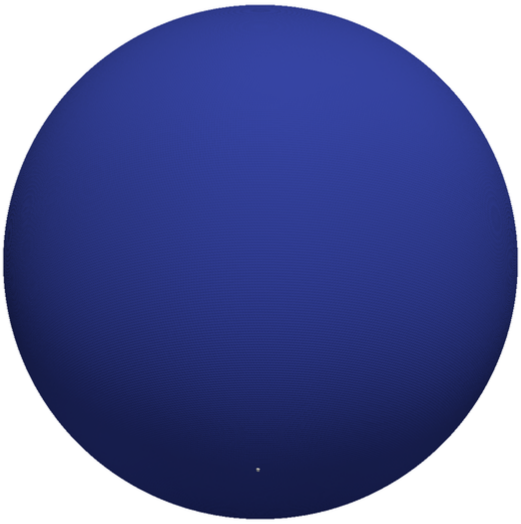


Figure 55: Point source smaller 0.5° at a declination of -70°.

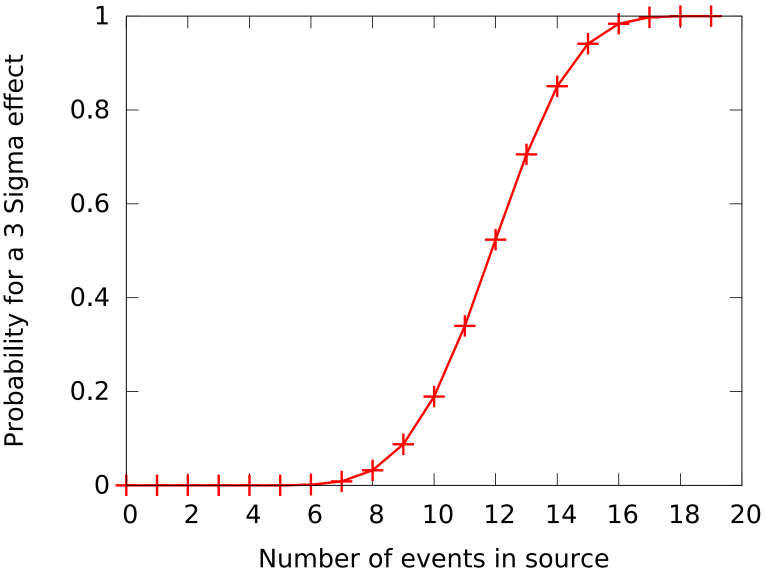


Figure 56: Sensitivity for a point source smaller 0.5° at a declination of -70°.

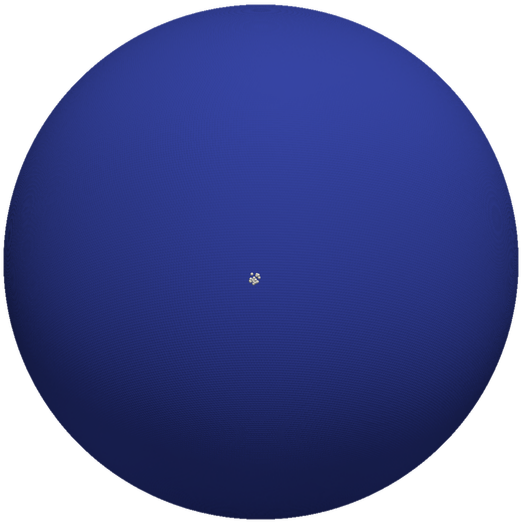


Figure 57: Point source smaller 1° at a declination of -29°.

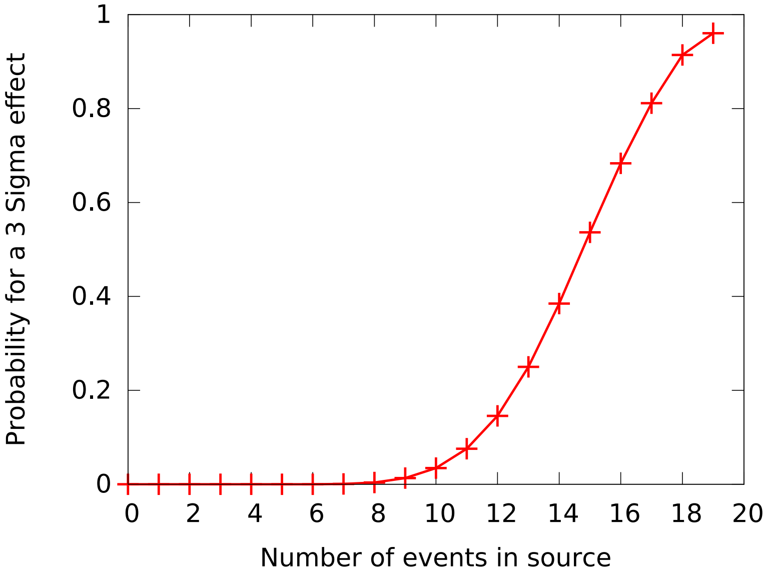


Figure 58: Sensitivity for a point source smaller 1° at a declination of -29°.

8.7.2 Extended sources

The extended sources here are added using random positions for the events within the source. These tests show the performance for scenarios with extended sources like the Fermi Bubbles or the hotspot as seen by IceCube in [105]. The declination of the Galactic Center is at -29°, the hotspot at the time this is written is believed to be centered at -23° with a possible extension between 8° and 20°. Similar extended setups are shown in Figures 59 and 61. Analog to the point sources, the probability for a 3σ effect depending on the number of events within the source for these setups is shown in Figures 60 and 62. To give an impression how the result of the developed analysis looks like for extended sources, Figure 63 shows the result of the method for this scenario. This plot can be compared to Figure 50, which shows the same stage of the computation for the example with two point sources. Figure 64 shows the effect of a segmentation with α = -0.11 on Figure 63. The corresponding plot for the two pointlike sources is Figure 52f. Every connected region here is one cluster. One can see that filaments are extending the cluster. This happens at locations where random accumulations of background events are close to the actual structure. The individual significance of each cluster is evaluated according to the chosen metrics. The cluster containing the artificial source events is detected with more than 5σ here. All other clusters on this sphere do not reach significances above 0.5σ.

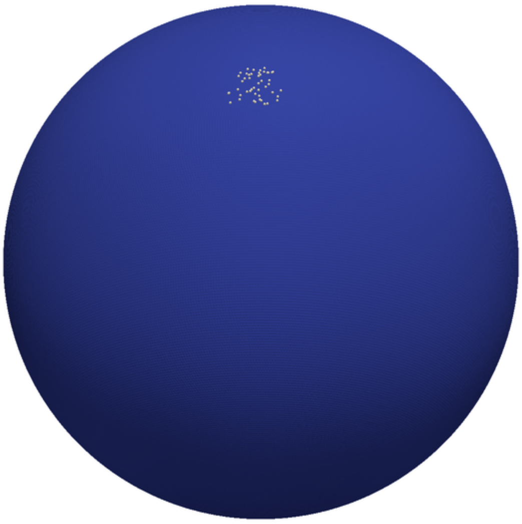


Figure 59: Extended source smaller 5° at a declination of +10°.

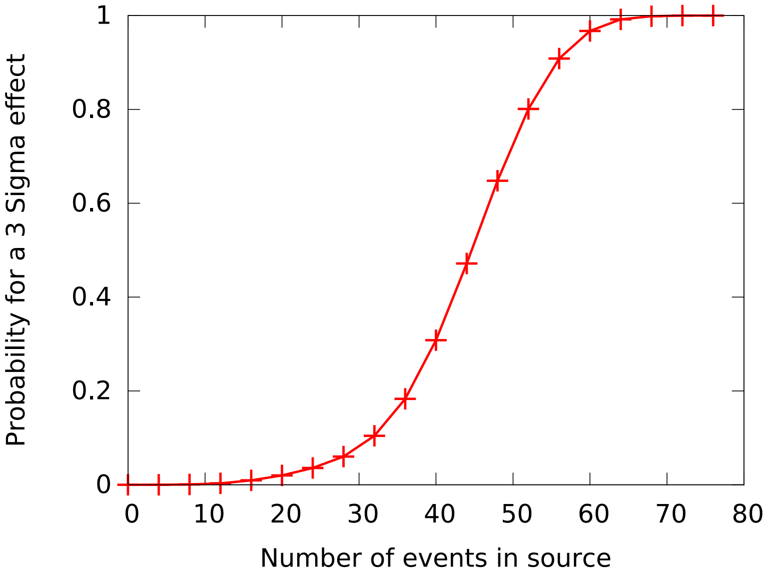


Figure 60: Sensitivity for an extended source smaller 5° at a declination of +10°.

Figure 61: Extended source smaller 10° at a declination of -29°.

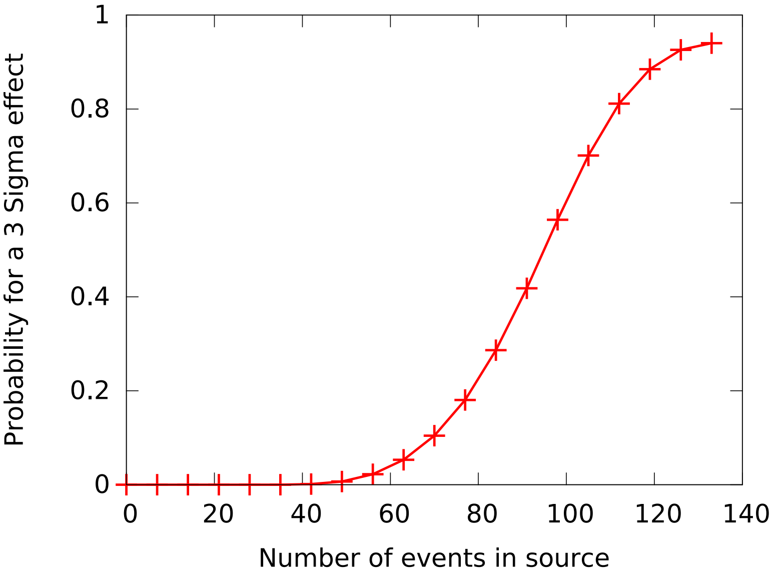


Figure 62: Sensitivity for an extended source smaller 10° at a declination of -29°.

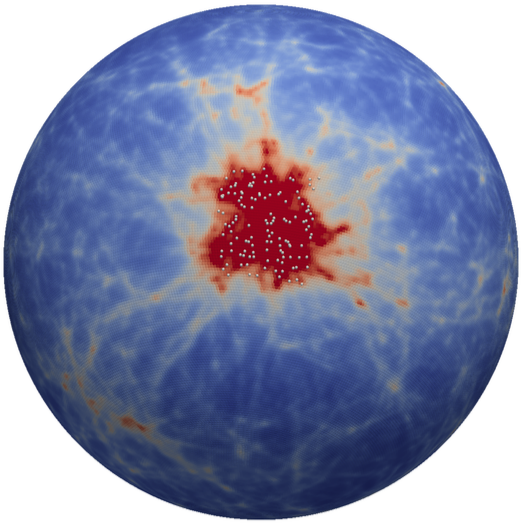


Figure 63: Result for a setup like in Figure 61. This is after the summation of all search scales described in chapter 8.4, but before any segmentation cut is applied as described in chapter 8.5. The color scale can be considered arbitrary units, red means higher event density.

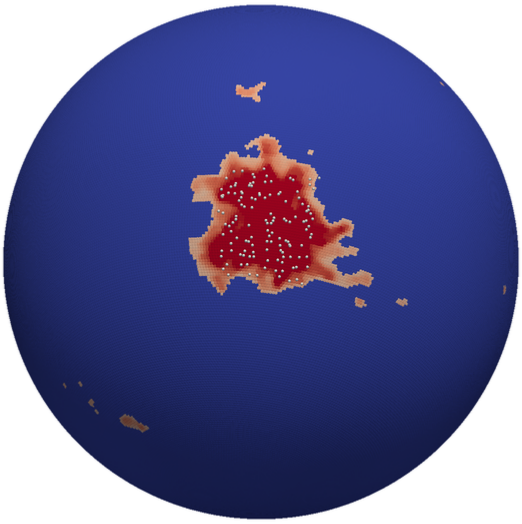


Figure 64: The effect of a segmentation with α = -0.11 on the result shown in Figure 61.

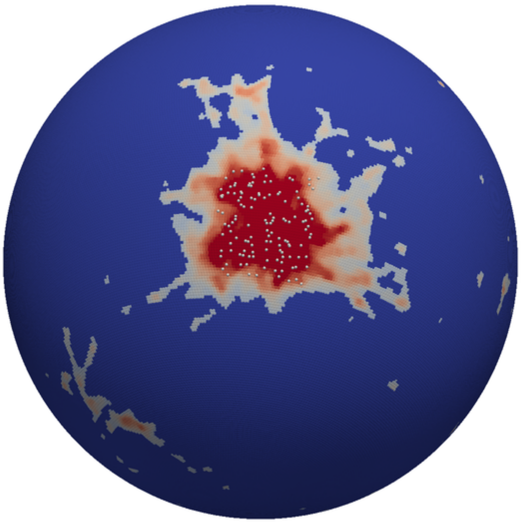


Figure 65: The effect of a segmentation with $\alpha = +0.25$ on the result shown in Figure 61.

8.7.3 Unexpected shapes

A common idea of a source is a structure with an approximately spherical shape. This search does not require such a shape to find a source. The artificial setup of a straight line in Figure 66 is used to demonstrate that arbitrary shapes can be found. The sensitivity for this scenario is shown in Figure 67. In the scenario shown in Figure 68, the even more arbitrary shape of a line with a gap is shown. Whether this shape is reconstructed as one connected or two separate lines depends strongly on the number of events in the line and random fluctuations in the gap. The sensitivity for this setup, independent of whether a connected or two separate sources are identified, is shown in Figure 69. Figure 70 shows the detailed result obtained for this scenario before segmentation. One has to keep in mind, that this is for a scenario with unrealistically many events within the source. This result after the segmentation with α = -0.11 is shown in Figure 71. The two parts have been detected as one long line in this example. Just like for the results for the extended sources, these results can be compared to the corresponding ones shown for the scenario with two point sources in Figure 50 and Figure 52f.

To demonstrate that the ability to identify arbitrary shapes is not limited to spherical and straight shapes, the setup shown in Figure 72 contains a curved line. The corresponding sensitivity can be seen in Figure 73. The results are shown in Figure 74 without segmentation and in Figure 74 after the segmentation using α = -0.11. The general shape follows the desired directions, but there are considerable filaments extending beyond the region where the artificial events have been added. These are caused by random fluctuations near the actual source region, but there is no way to distinguish them from "real source events" in a real application.

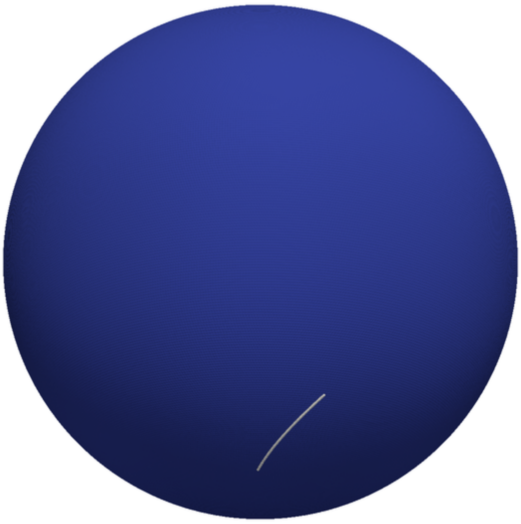


Figure 66: Diagonal line, 20° length, at a declination of -70°.

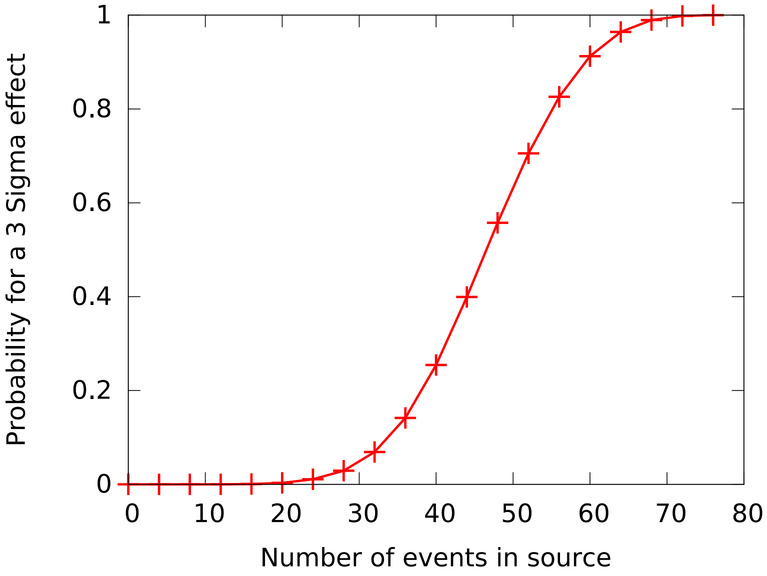


Figure 67: Sensitivity for a diagonal line, 20° length, at a declination of -70°.

Figure 68: Diagonal line with interruption, 20° length, at a declination of -29°.

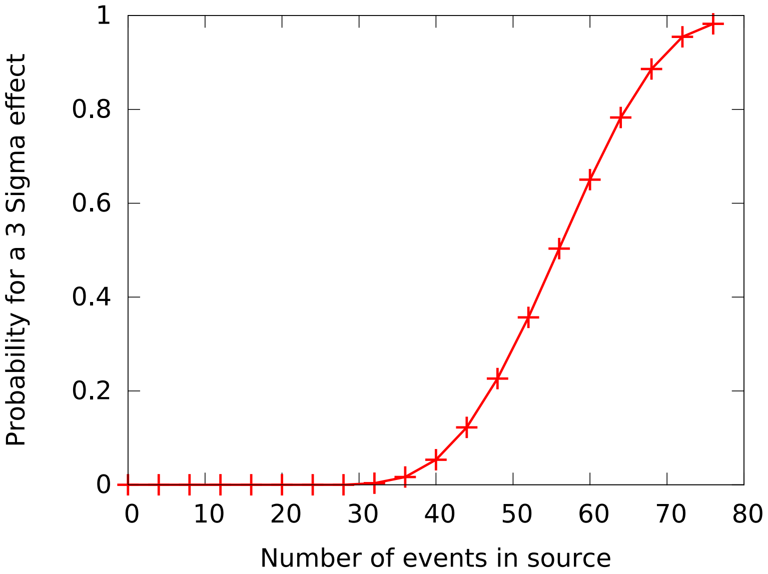


Figure 69: Sensitivity for a diagonal line with interruption, 20° length, at a declination of -29°.

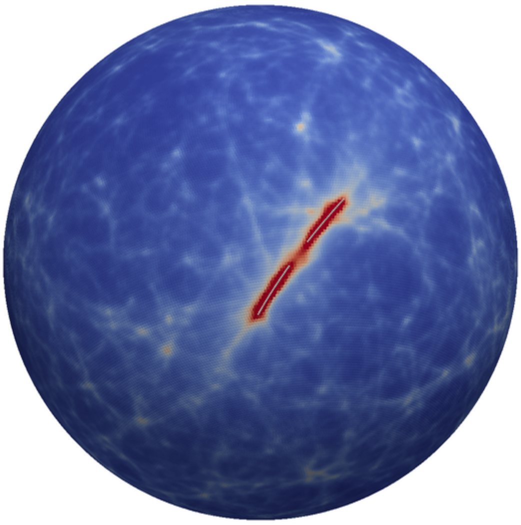


Figure 70: Result for the setup in Figure 68. This is after the summation of all search scales described in chapter 8.4, but before any segmentation cut is applied as described in chapter 8.5. The color scale can be considered arbitrary units, red means higher event density.

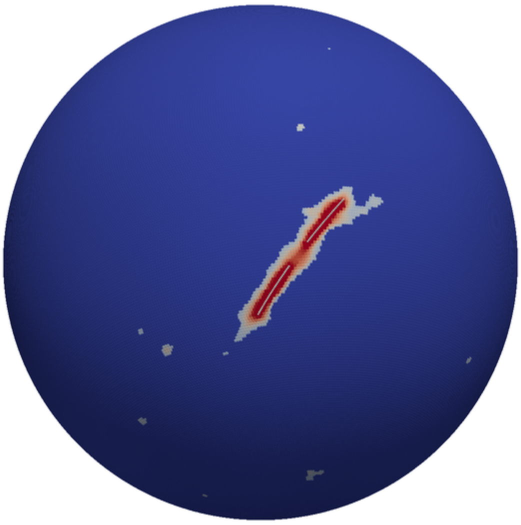


Figure 71: The effect of a segmentation with α = -0.11 on the result shown in Figure 68. The two parts of the line are connected in this example, but this is not the case for all pseudo-experiments of this setup, especially with less events.

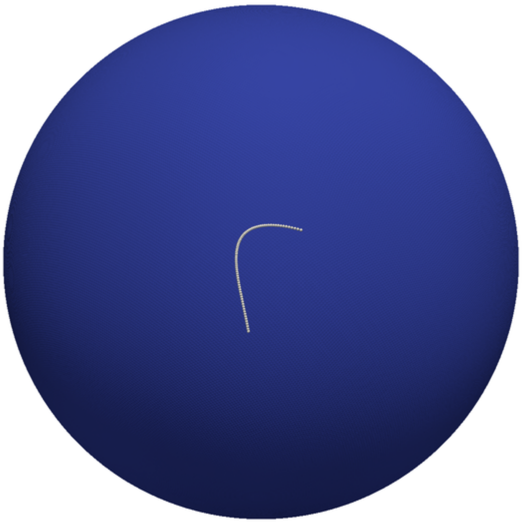


Figure 72: Curved line, 20° length, at a declination of -29°.

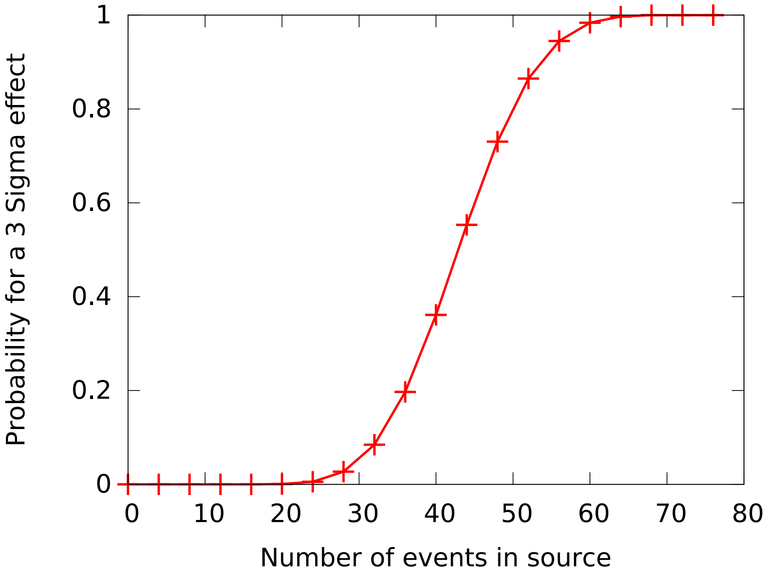


Figure 73: Sensitivity for a curved line, 20° length, at a declination of -29°.

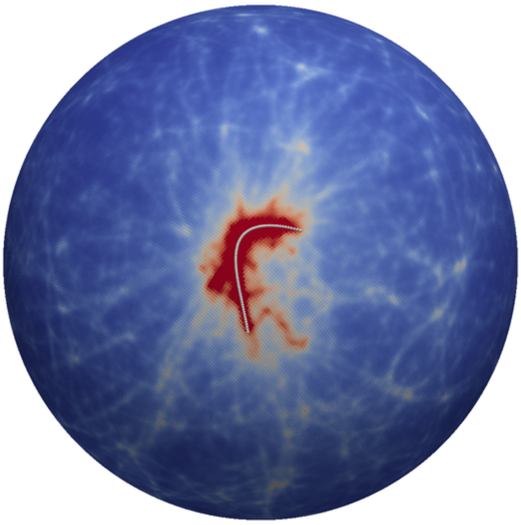


Figure 74: Result for the setup in Figure 72. This is after the summation of all search scales described in chapter 8.4, but before any segmentation cut is applied as described in chapter 8.5. The color scale can be considered arbitrary units, red means higher event density.

Figure 75: The effect of a segmentation with α = -0.11 on the result shown in Figure 72. The filaments in this example are caused by random background fluctuations near the source.

8.7.4 Neighboring sources

If other sources are located close to a source, the probability to detect the presence of the whole structure is higher than the sum of the probabilities to detect the sources independently. This is demonstrated by the setup as seen in Figure 76. The sensitivity for this scenario is shown in Figure 77. The numbers of events are per source here. What can be seen is that the number of events required for each point source to obtain a 3σ effect with a certain probability is lower than for one source of the same extension and location alone. For comparison: The sensitivity for a single source of the same size and location has been presented in Figure 58.

8.7.5 Diffuse flux

This search is not at all sensitive to a completely diffuse flux. It adapts to the data in the beginning of the search in such a way that the overall excess is automatically compensated, making it completely impossible to detect anything that doesn't have a spatial structure.

Figure 76: Three point sources smaller 1° at declinations of -25°, -29° and -34°.

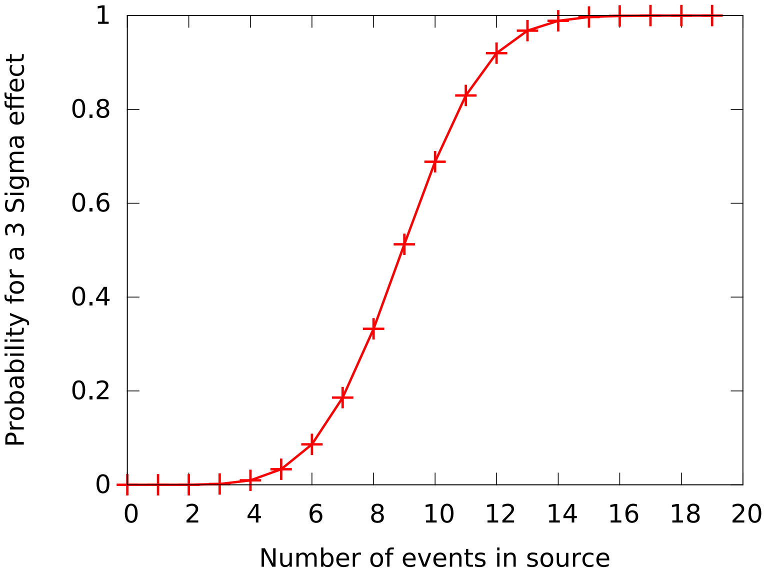


Figure 77: Sensitivity for three point sources smaller 1° at declinations of -25°, -29° and -34°.

8.8 Modifications for IceCube data

8.8.1 Necessary modifications

Although no explicit assumption has been made derived from properties of the ANTARES neutrino telescope for this search, this happened implicitly at some points when the algorithm was optimized with the specific application in mind. If the same algorithm is to be applied to the public IC40 dataset [106], released by the IceCube collaboration, some details should be reconsidered. What was realized to be the relevant difference between the application to ANTARES and IceCube data is the approximation of the expectation of the number of neutrinos for a certain declination as described in chapter 8.3.3. While the probability to detect a neutrino changes slowly with the declination for ANTARES, see Figure 41, there is a leap for IceCube at a declination of about 0° as seen in Figure 78. The consequence of this discrepancy is that the estimated expectation derived from scrambled data cannot be smoothed as much for IceCube as for ANTARES, because otherwise the estimated expectation would strongly smear the observed jump. The number of applied lowpass filter operations is therefore reduced from 60 to 15 to better preserve this leap. The result of the acceptance estimation for the IC40 data can be seen in blue in Figure 78 for up to 0.5° and in Figure 79 for 10°.

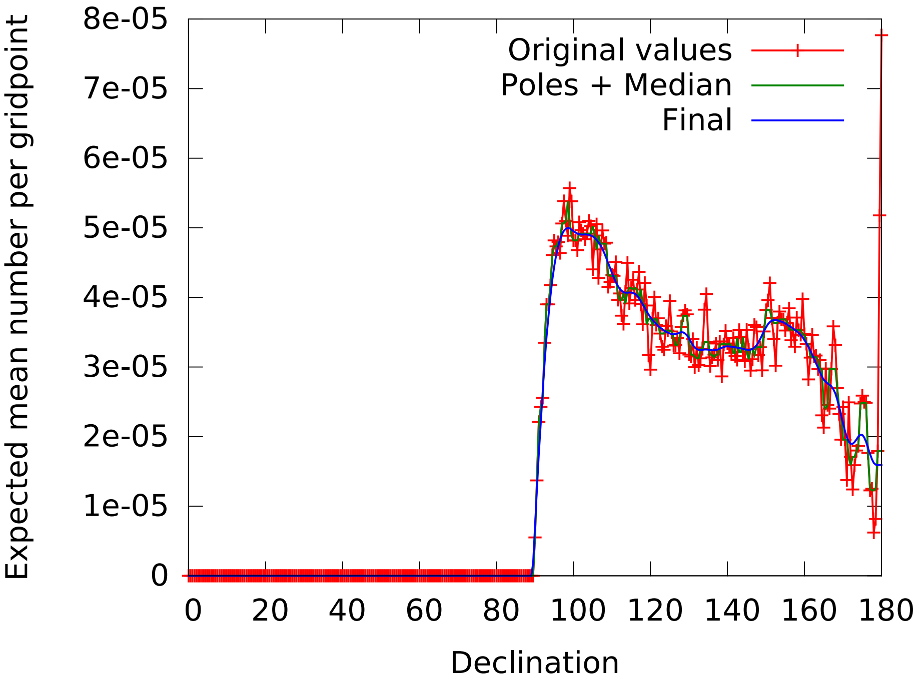


Figure 78: Approximation of the acceptance of the IceCube detector and the reconstruction chain for a search distance up to 0.5°, based on the IC40 data sample.

Apart from this adaption, no part of the algorithm needs to be modified to

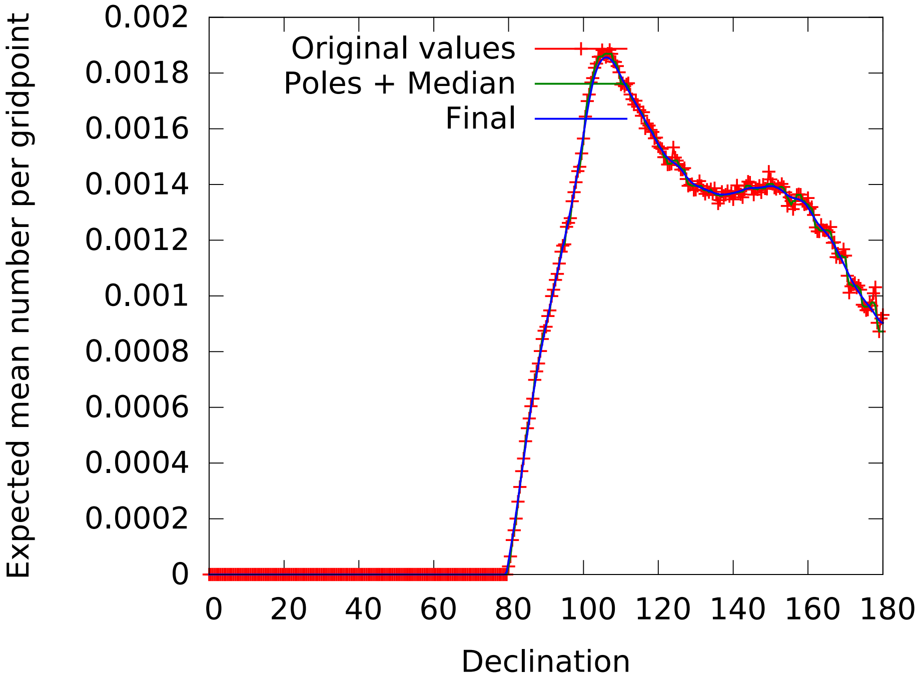


Figure 79: Approximation of the acceptance of the IceCube detector and the reconstruction chain for a search distance between 10.0° and 10.5°, based on the IC40 data sample.

apply it to the public IC40 dataset.

8.8.2 Optional modifications

Since the ANTARES and IceCube datasets are completely independent from each other but both should contain the same physical observation of the neutrino sky, one can be used to check a hypothesis derived from the other in the region where both datasets overlap without a trial factor.

A simple way to achieve this with this search method would be to analyze the IceCube data with the hypothesis generated by the ANTARES data and compare the results by eye. Since the outcome is not trivial to interpret, this approach could easily be misleading. To perform this check on a more solid basis, the idea is to predefine some requirements for the comparison of the hypothesis from ANTARES data and potential clusters in the IceCube data. First of all, the same segmentation threshold is used to evaluate the IC40 data as was used to find the most significant structure in the ANTARES data. Furthermore, a cluster has to overlap with the template of the ANTARES structure. Since large structures, which barely overlap with the template by only one pixel, certainly are not the aim of such a search, a minimum overlap is required for a cluster to be taken into account. This overlap requirement is set to 51% of the size of the new cluster to require it to be more related to the template than to other regions. Since requiring an overlap with a template implicitly means searching in a confined region, the metric to evaluate the significance cannot be size, as it would prefer clusters which exactly have a fraction of 51% of their size inside and 49% outside of the template. Therefore the metric is changed to metric number 1, topSqrt, as described in Appendix E. This metric uses the mean value of the $\sqrt{N}$ highest pixel-values in the cluster, with N being the size of the cluster. Judging from pseudo-experiments, this is the second best tested metric.

Part III

Results and discussion

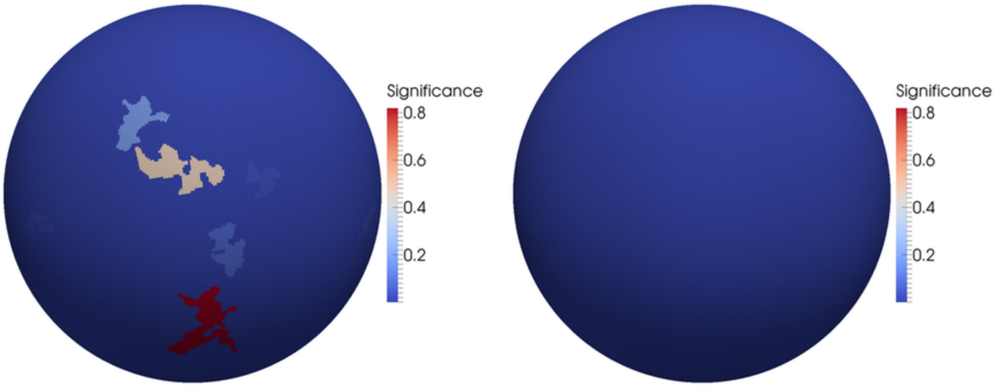
This part of the thesis presents the results obtained using the described methods and gives an interpretation of these results.

9 Results

Only after the development of the methods is finished and the analysis procedure and all cuts are fixed, the ANTARES data are processed for the first time, the so called “unblinding”. This mandatory blinding policy prevents that a bias can influence the development of the methods or analyses. After filtering the recorded data as described in chapter 7, the unblinding for this analysis resulted in 13283 neutrino event candidates. The spatial distribution of these events has been analyzed using the techniques described in chapter 8. The result of this analysis are clusters, each with an independent significance. As explained in chapter 8.6.1, only the size of the clusters is used as relevance metric to assess their significance. Since two different segmentation thresholds have been used, there are two different resulting skymaps. The trial factors for all evaluations have been incorporated into the post-trial significance of each individual cluster.

9.1 Found clusters

The segmentation with a threshold of α = -0.11, see chapter 8.5, is the harder of the two applied segmentation cuts. It detects multiple extended clusters which are close together. These clusters are not significant on their own. The result as computed on the search sphere can be seen in Figure 80. Figure 81 shows the same results as a skymap in equatorial coordinates, Figure 82 as skymap in galactic coordinates. All skymaps are presented in Hammer-Aitoff projection. The highest significance of 0.81σ is found for the cluster at the bottom, colored in red on the sphere and in white in the skymaps.



(a)

(b)

Figure 80: Result of the unblinding with a segmentation threshold of α = -0.11 on the search sphere. Multiple extended clusters have been found, but no structure is significant on its own. a) View centered on the clusters. b) View of the opposite side.

The threshold of $\alpha = +0.25$ is the less strict segmentation cut. Its intention is to find very large structures or to allow neighboring structures to merge. The clusters found using α = -0.11 fuse in this segmentation, resulting in one large

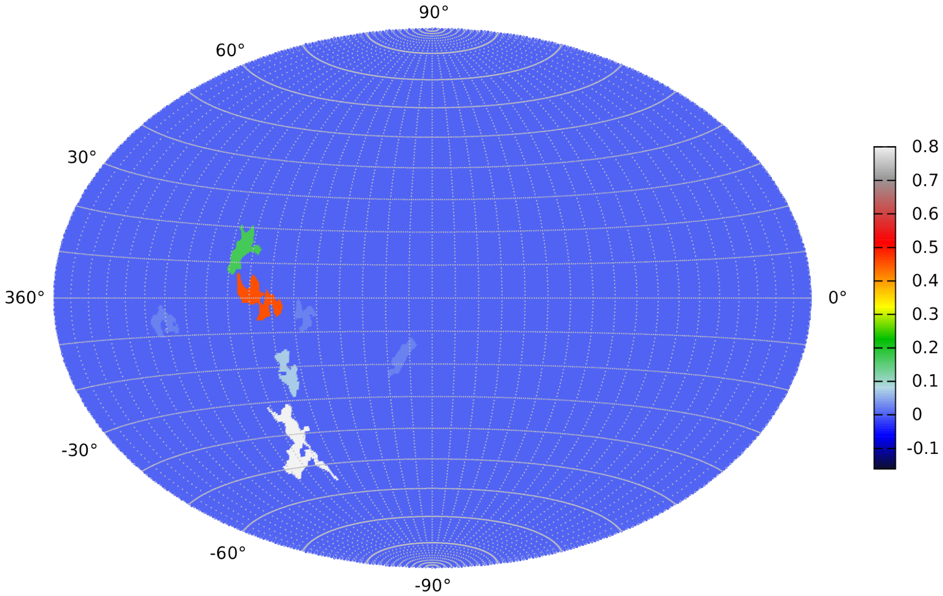


Figure 81: Result of the unblinding with a segmentation threshold of α = -0.11 in equatorial coordinates. The colorcode is the significance in σ.

cluster with a post-trial significance of 2.85σ. Due to the compensation of a systematic effect investigated in chapter 9.7, this significance gets reduced to 2.52σ. The significance of the clusters observed for α = -0.11 changes are only slightly reduced by this effect to 0.79σ. The very large structure can be seen on the search sphere in Figure 83. Figure 84 shows the result as skymap in equatorial coordinates, Figure 85 in galactic coordinates. This cluster encloses the Galactic Center, but it is larger than the largest extended structures that are currently suspected to emit neutrinos, for instance the Fermi Bubbles [94].

Since the obtained results aren't trivial to interpret, several further studies are presented in the following chapters.

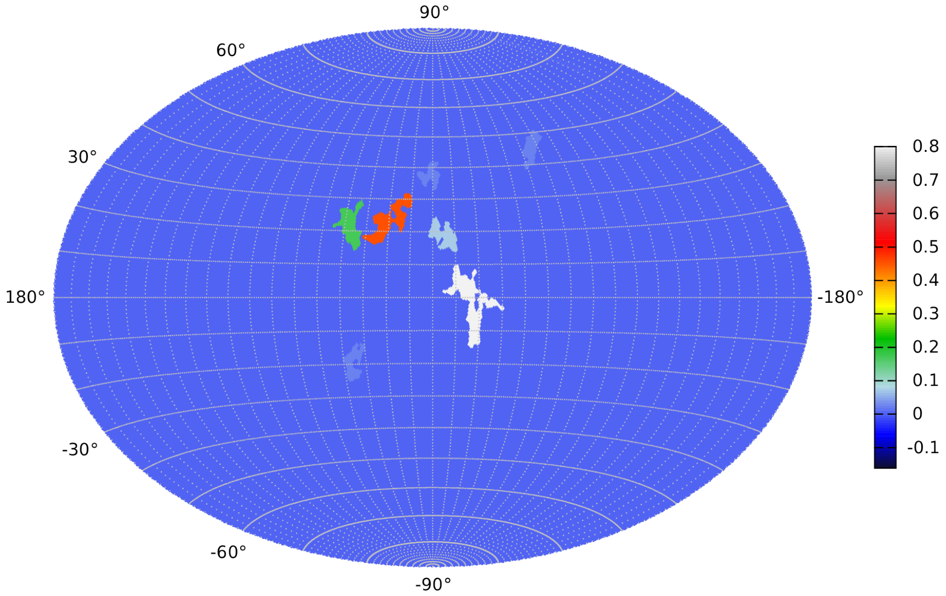
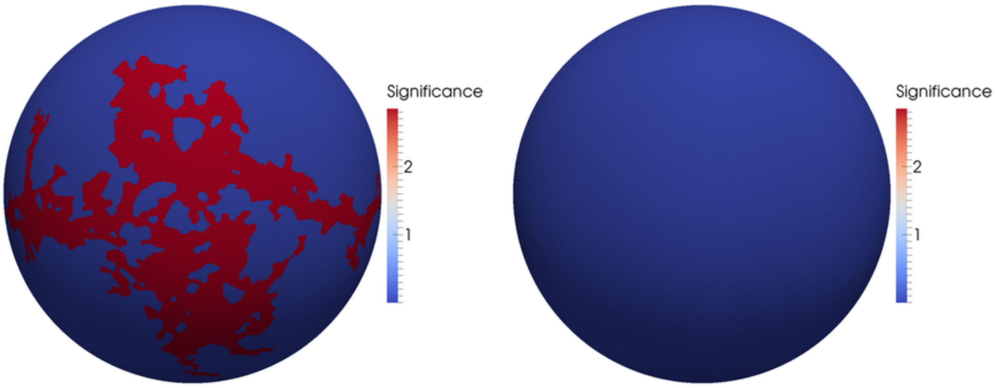


Figure 82: Result of the unblinding with a segmentation threshold of α = -0.11 in galactic coordinates.



(a) (b)

Figure 83: Result of the unblinding with a segmentation threshold of $\alpha = +0.25$. A very large structure with a post-trial significance of 2.85σ is found. a) View centered on the cluster region. b) View of the opposite side.

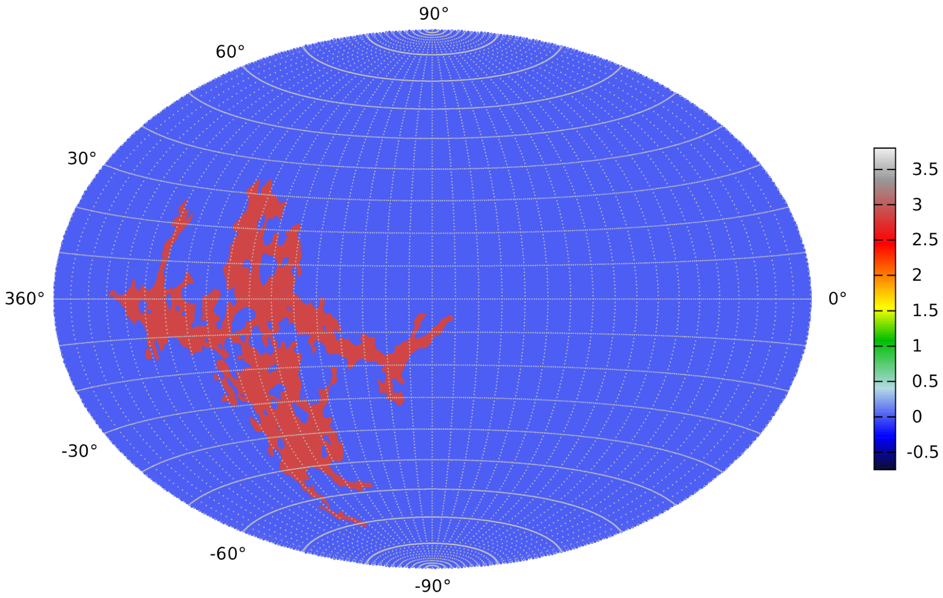


Figure 84: Result of the unblinding with a segmentation threshold of $\alpha = +0.25$ in equatorial coordinates. The colorcode is the significance in σ.

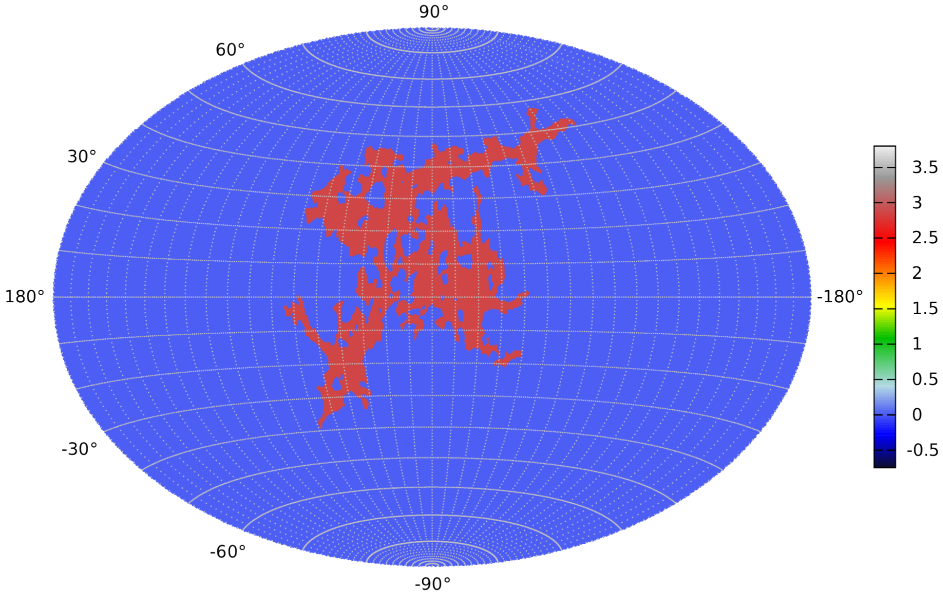
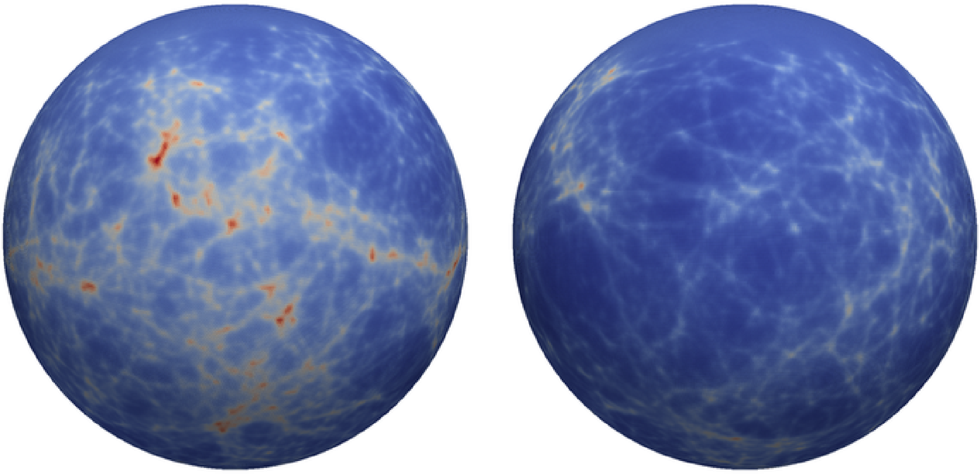


Figure 85: Result of the unblinding with a segmentation threshold of $\alpha = +0.25$ in galactic coordinates.

9.2 Result before segmentation

Figures 86, 87 and 88 show the structures that have been found by the multiscale search after summation of the individual scales as described in chapter 8.4, but before the segmentations are applied as described in chapter 8.5. What can be seen is that the clusters are located where the highest overfluctations on all scales are observed, so this relation works as expected also for the final data analysis. Figures 89, 90 and 91 show the same plots, but this time with the position of the neutrino events as overlay. They allow to validate that these overfluctuation regions actually correspond to regions where many events are contained.

It has to be noted that the plotted variable is not directly the flux at that point, but it is influenced by all scales up to 90° around this point. One of the most relevant differences to a flux map is that the same cluster of events near other overfluctuations might stand out with a high value, while surrounded by underfluctuations, the same cluster can have a much lower, unremarkable value.



(a) (b)

Figure 86: Result before the segmentation. a) View centered on the cluster region. b) View of the opposite side.

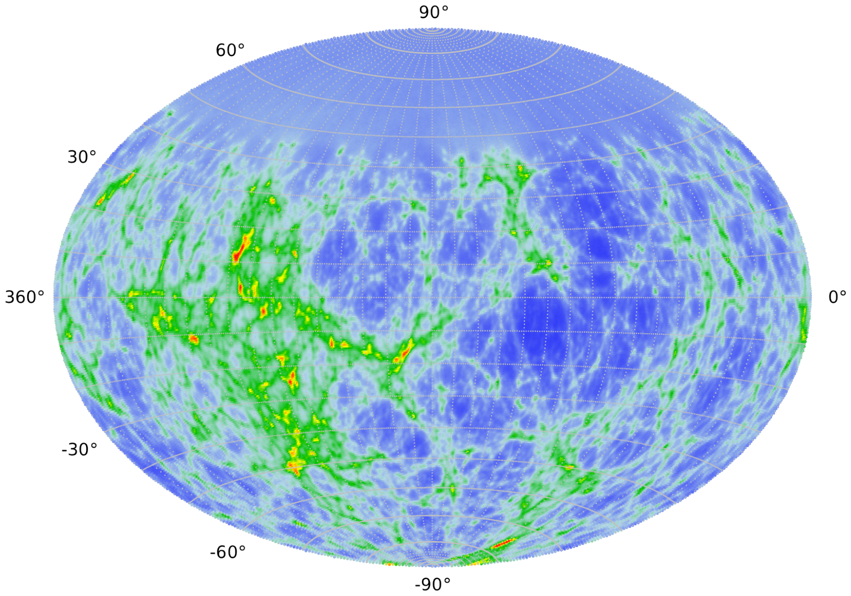


Figure 87: Skymap before the segmentation in equatorial coordinates.

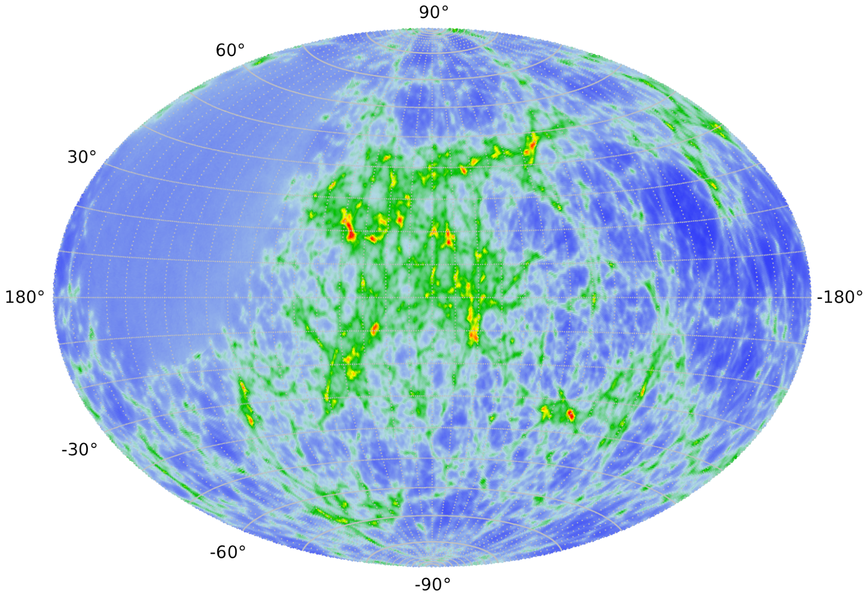


Figure 88: Skymap before the segmentation in galactic coordinates.

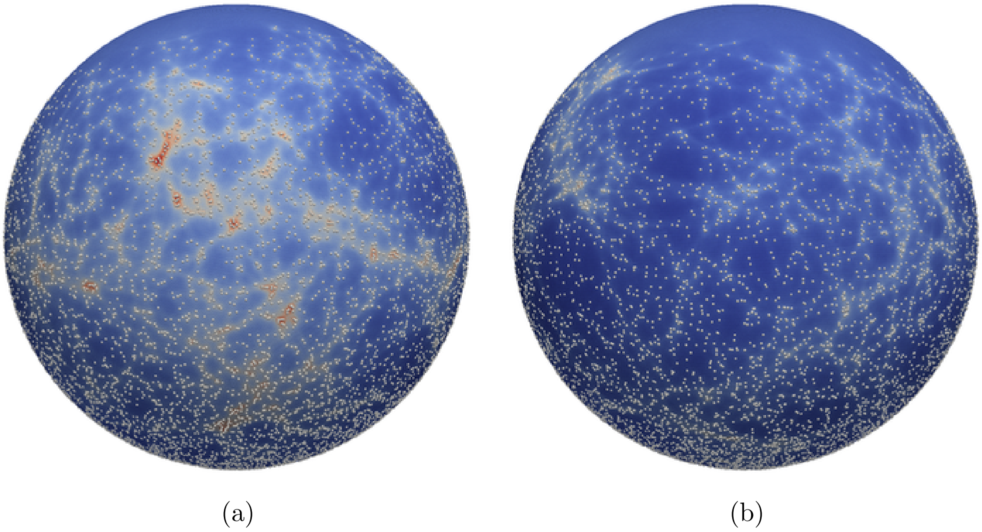

(a) (b)

Figure 89: Result before the segmentation with events in white. a) View centered on the cluster region. b) View of the opposite side.

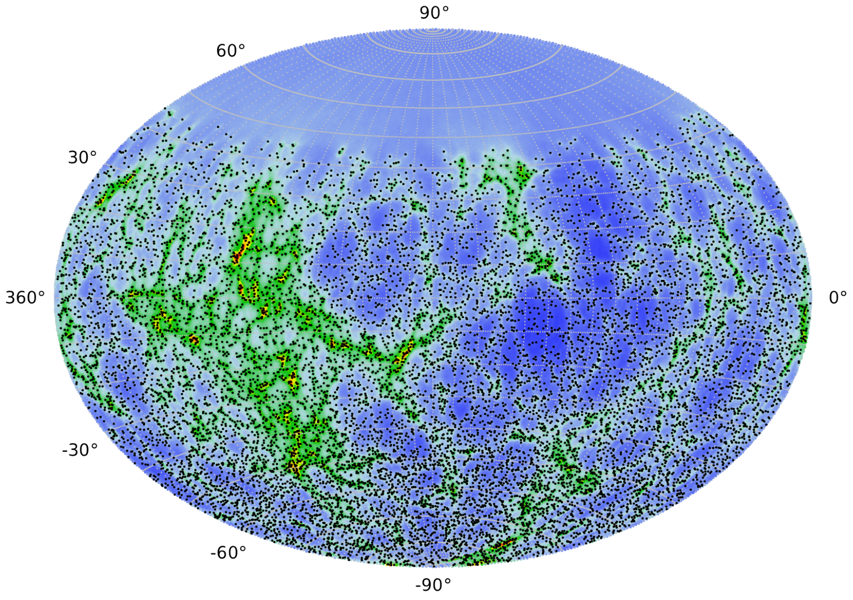


Figure 90: Skymap before the segmentation in equatorial coordinates with the events in black.

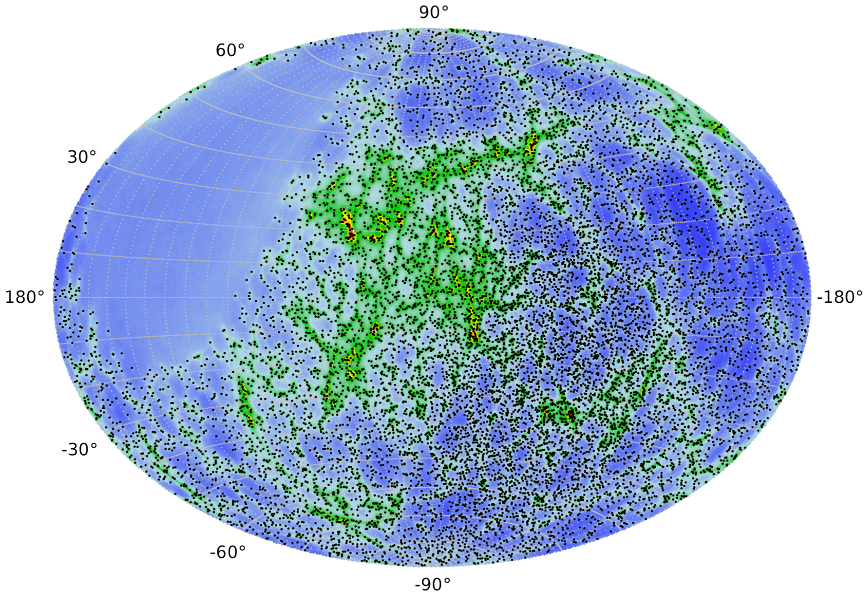


Figure 91: Skymap before the segmentation in galactic coordinates with the events in black.

9.3 Comparison with results for scrambled data

To be able to compare the obtained results to skymaps without any sources, pseudo-experiments with randomly scrambled data have been conducted. As examples three randomly selected skymaps are presented here in Figures 92, 93 and 94 using exactly the same color code that has already been used for the results of the actual data analysis, for instance in Figure 87. They show that the individual overfluctuations observed in the recorded data are common, but their distribution across the sky is not.

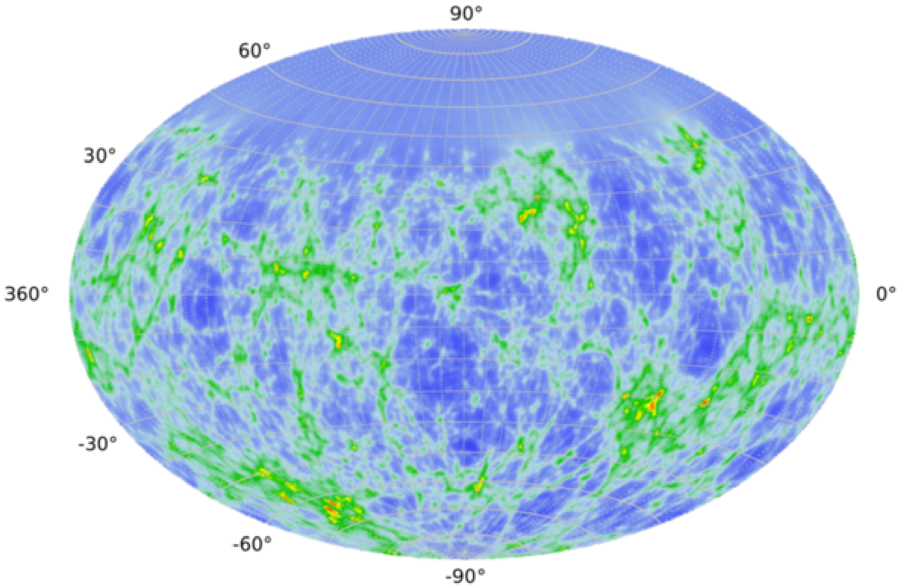


Figure 92: Result for scrambled data set I in equatorial coordinates.

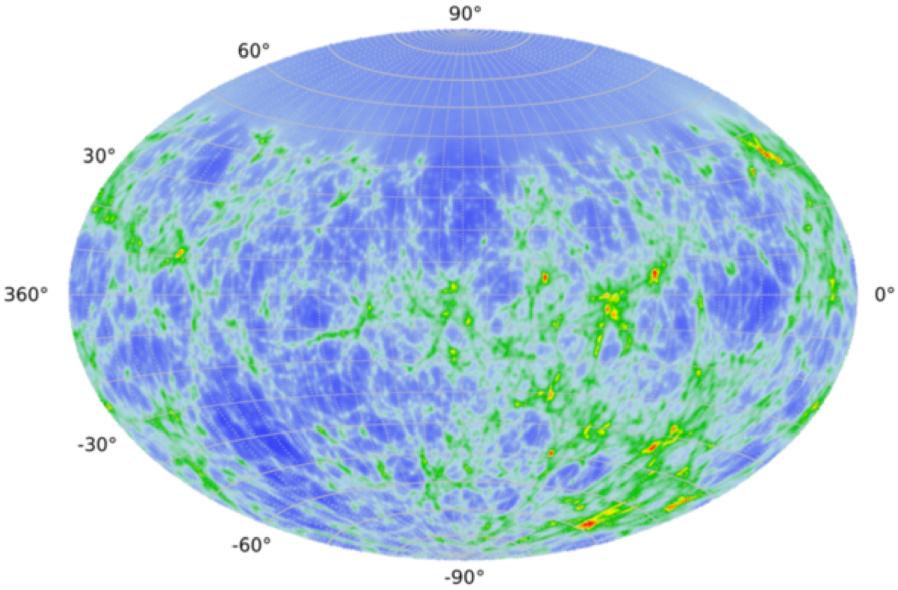


Figure 93: Result for scrambled data set II in equatorial coordinates.

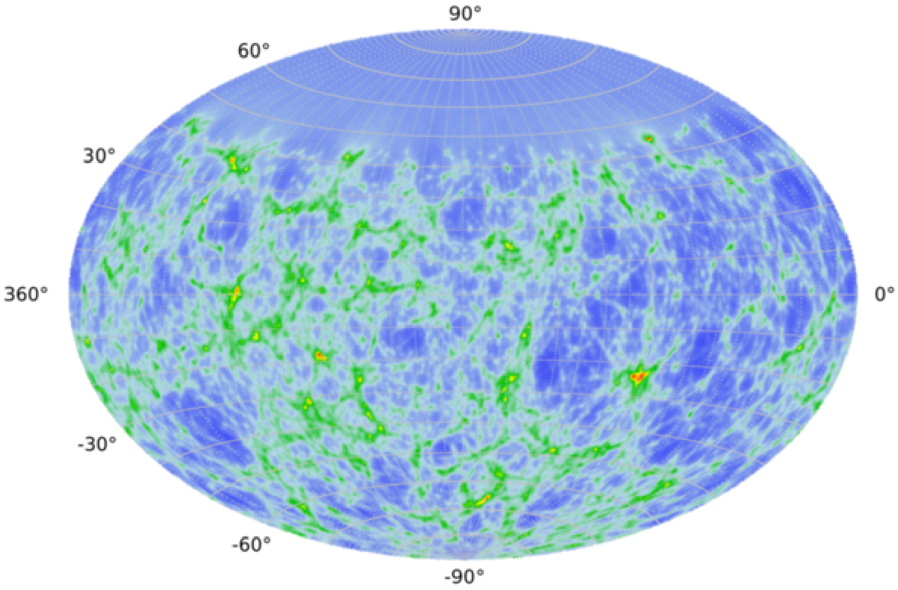


Figure 94: Result for scrambled data set III in equatorial coordinates.

9.4 Subsampling

To check if the method requires exactly the 13283 selected events to give the observed result, tests with randomly subsampled datasets have been conducted. They were done with 11000 randomly chosen events out of the 13283, but also with a relatively small subsample of 5516 events to evaluate the potential behavior of this method on an event sample of the same size which has been used for the most recent ANTARES point source search, see [91]. This procedure has not been repeated more often due to time constraints (generating reasonable statistics for the pseudo-experiments takes most of the time) and also since the stability of the method has been analyzed in a different setup in chapter 9.5.

9.4.1 11000 events, random subsample I

A random subsampling of 11000 out of 13283 events is expected to find similar structures. Due to the nonlinearity of the used metric "Size" to assess the relevance of a cluster, the significance can be expected to vary considerably. The presented plots show the results in galactic coordinates.

The detailed structure resulting from this evaluation can be seen in Figure 95. The most significant cluster, found with a segmentation cut of +0.25, reaches 2.94σ and is shown in Figure 97. So in this case the random subselection even surpasses the significance of the full dataset. For the evaluation using a segmentation cut of -0.11, shown in Figure 96, the most significant cluster has 1.87σ, which also is higher than what was observed for the full dataset.

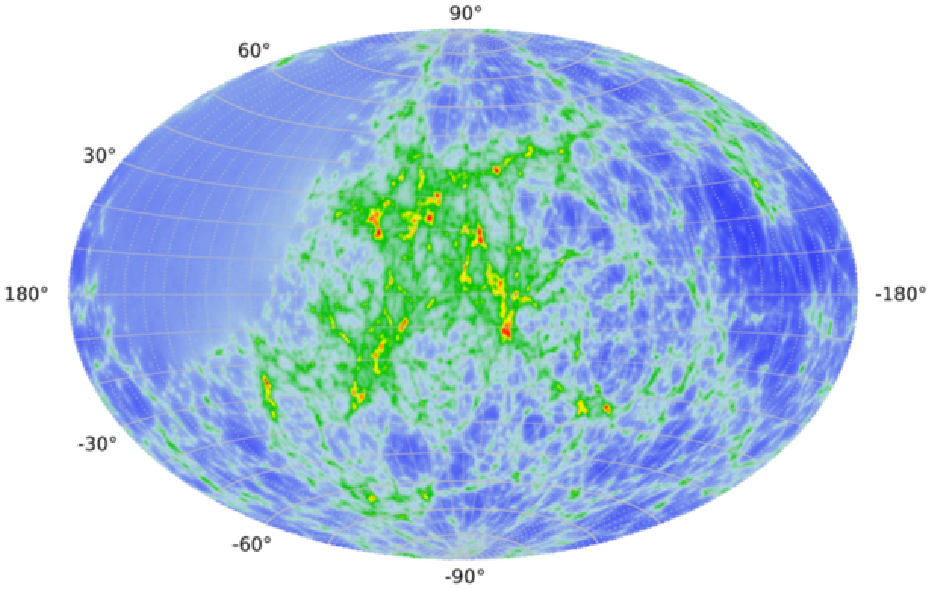


Figure 95: Result of random subsampling I of 11000 events before segmentation in galactic coordinates.

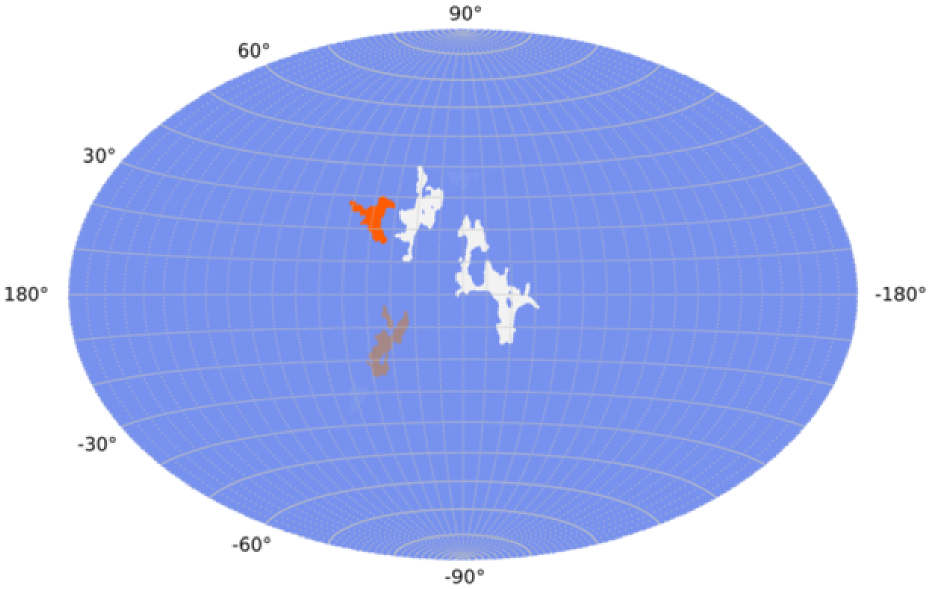


Figure 96: Result of random subsampling I of 11000 events with a segmentation threshold of α = -0.11 in galactic coordinates.

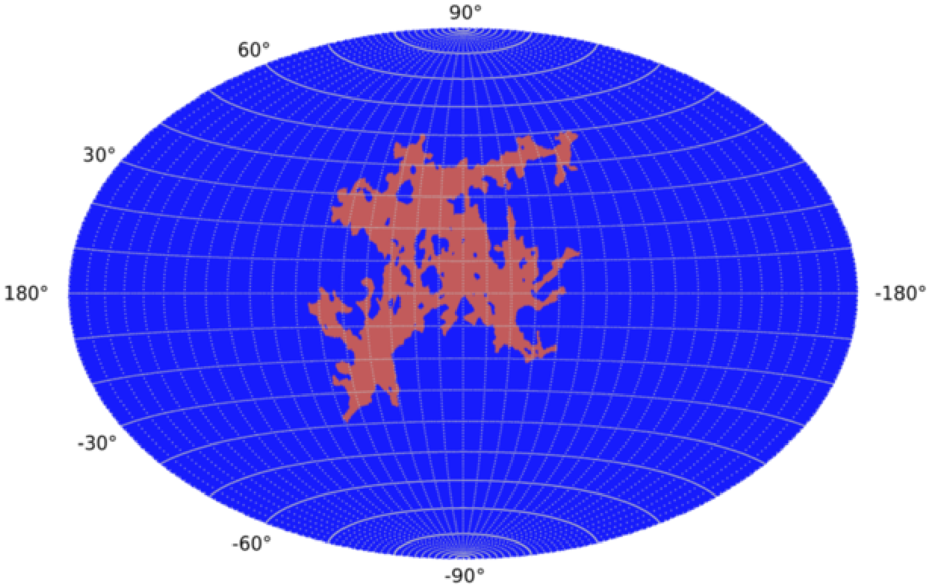


Figure 97: Result of random subsampling I of 11000 events with a segmentation threshold of α = 0.25 in galactic coordinates.

9.4.2 11000 events, random subsample II

A second random subselection of 11000 events shows a different picture. Although the skymap containing the detailed structure in Figure 98 looks similar to Figure 95, Figure 100 shows that not all overfluctuation regions merged for a segmentation threshold of $\alpha = 0.25$. The most significant cluster for this segmentation cut only reached 0.53σ. The most significant cluster in the evaluation using a segmentation cut of -0.11 shown in Figure 99 has 1.11σ, again more than what has been observed for this segmentation with the full dataset.

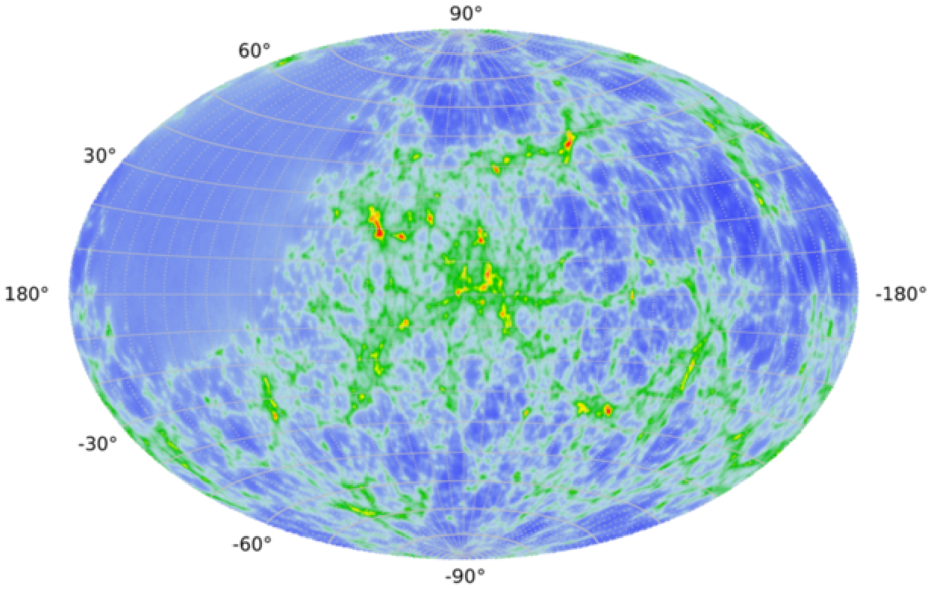


Figure 98: Result of random subsampling II of 11000 events before segmentation in galactic coordinates.

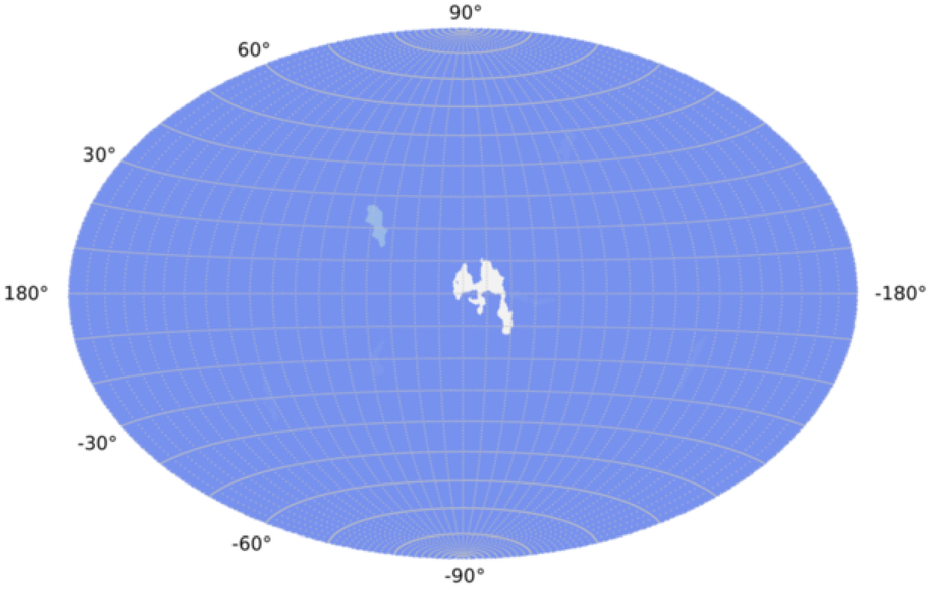


Figure 99: Result of random subsampling II of 11000 events with a segmentation threshold of α = -0.11 in galactic coordinates.

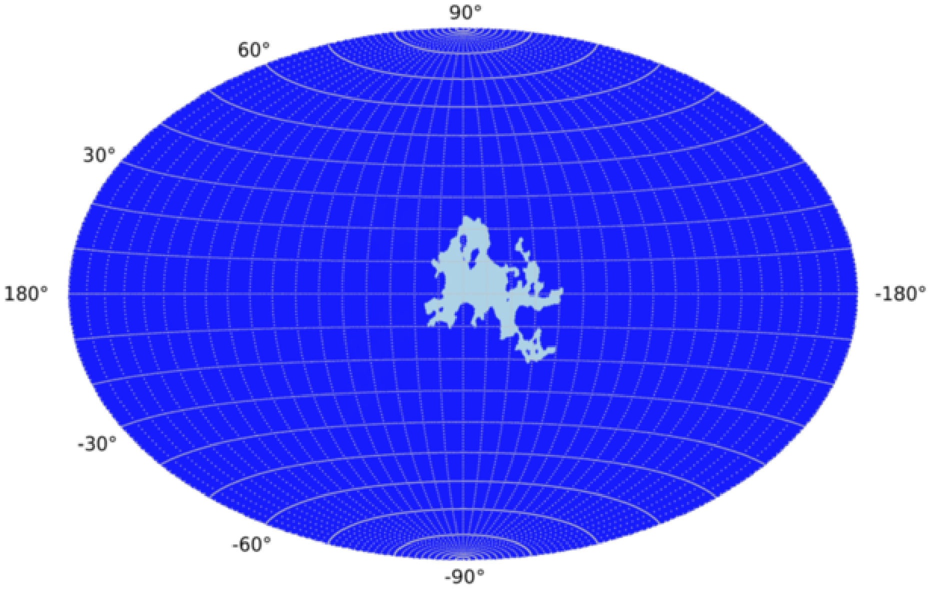


Figure 100: Result of random subsampling II of 11000 events with a segmentation threshold of $\alpha = 0.25$ in galactic coordinates.

9.4.3 5516 events, random subsample I

Random subsamples of 5516 out of 13283 events can be expected to differ considerably more from the full data sample than samples with 11000 events. Also the significance of found clusters can be expected to vary even more. The motivation for the exact number of 5516 events is that this is the number of neutrino candidates used in the most recent ANTARES point source search.

The resulting detailed skymap is shown in Figure 101. The obtained significances are 0.81σ for $\alpha = 0.25$ and 0.60σ with $\alpha =$ -0.11.

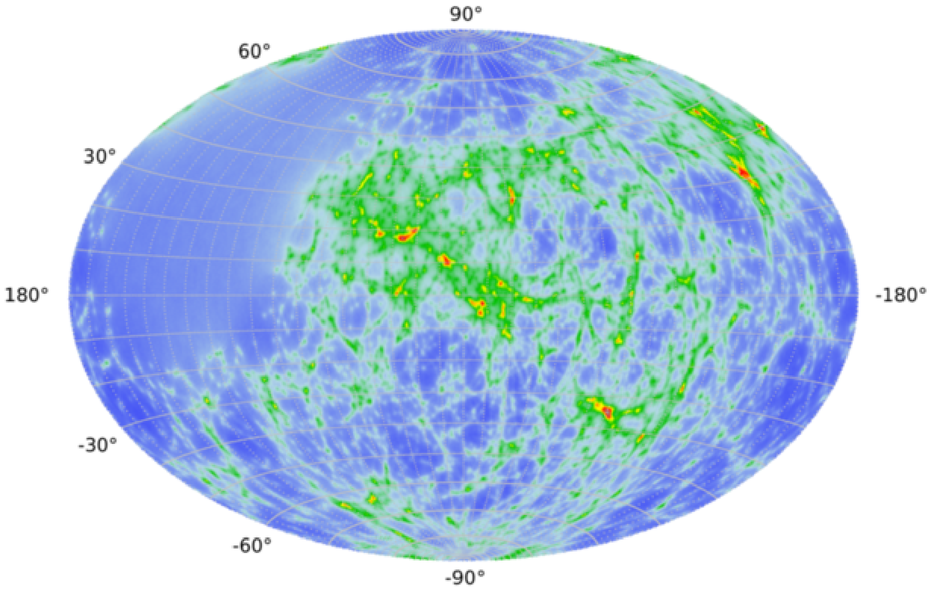


Figure 101: Result of random subsampling I of 5516 events before segmentation in galactic coordinates.

9.4.4 5516 events, random subsample II

A second random subsample with 5516 event has been processed. The resulting detailed skymap is shown in Figure 102. The obtained significances in this case were 0.10σ for $\alpha = 0.25$ and 0.07σ with $\alpha = 0.11$.

9.4.5 Result of random subsampling

Since most comparisons with subsampled datasets show a trend to obtain considerably less significant results with less events, the question arises if this is the normal effect of more statistics allowing a more precise measurement. Theoretically speaking, it could be the case that this is not due to statistical fluctuations, but that the method itself is not stable. If this were to be true, the outcome wouldn't stabilize even for a clearer source with more events.

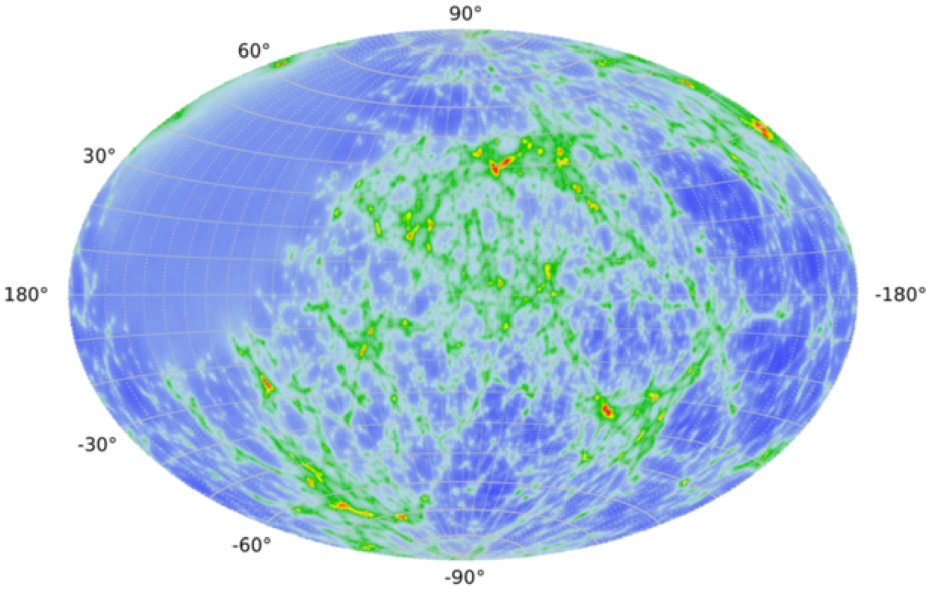


Figure 102: Result of random subsampling II of 5516 events before segmentation in galactic coordinates.

9.5 Recreating the cluster in pseudo-experiments

This chapter shows the results of pseudo-experiments conducted to test if the outcome of the method stabilizes for strong sources. This helps to distinguish whether the method itself might be unstable for a structure like the observed cluster, or if the method stabilizes for strong sources and the observed behavior is best explained by the assumption that the structure is just at the edge of the detectable strength.

Pseudo-experiments have been simulated with a varying number of events M placed artificially within the template of the cluster. The location of the artificial events within the template is random, but according to the visibility of ANTARES. A random background of $13283 - M$ randomly scrambled events is added for each experiment to in total obtain the same number of events as in the actual data sample. This allows to generate a coarse curve of the obtained significance for varying number of events in the cluster, shown in Figure 103. For each of the data points 30 repetitions have been conducted.

Due to the strong nonlinearity of the method, the error bars are huge before the method stabilizes for clear signals. Unfortunately these large fluctuation in the transition region prevent an answer to where the observation might be located on that function. It has to be noted that the generated curve is not exact, since the events have been randomly distributed within the cluster. If they were to be distributed with an internal structure, less events would be needed to achieve the significances. But modeling these fluctuations can hardly be done without further speculative assumptions about the composition of the

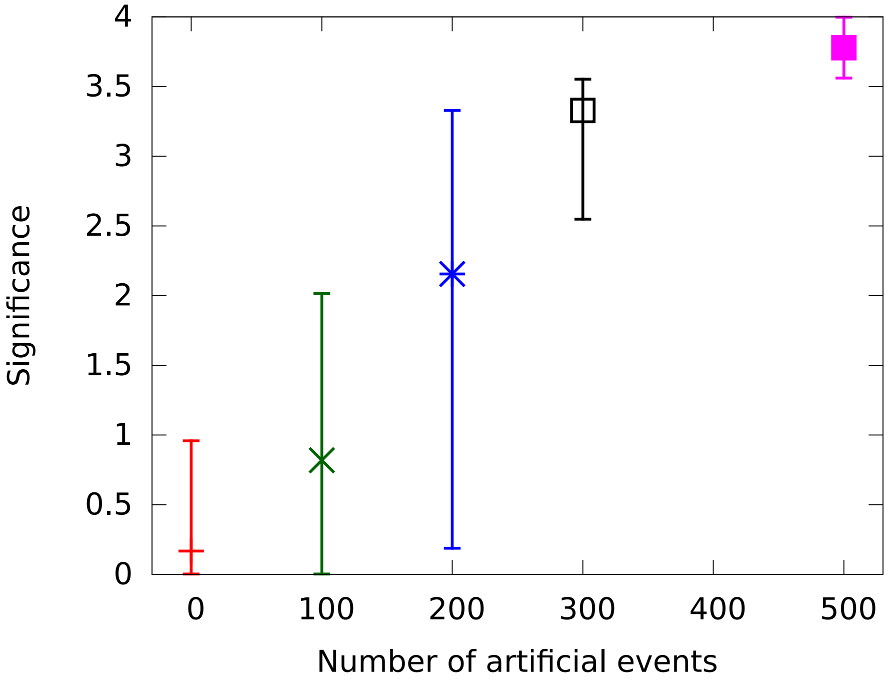


Figure 103: Detected significance against the number of artificial events in the structure. The values are median values, the error bars represent 1σ ranges.

observed structures. Anyways it would only shift the stabilization to different numbers of events, not reduce the fluctuations themselves. Therefore the result of this check, that the method does give stable results for this cluster if it is strong enough, wouldn't change.

Another result of these tests is that the the cluster can be explained best by an excess between 200 and 300 events. Of course this test alone is no reliable estimate due to the large spread of the results around 200 added events. Another idea how to obtain the number of excess events in the cluster has been to estimate the number of events expected within the shape of the cluster. Pseudo-experiments yield a mean expected number of 1279 events, while in the measured data 1533 events are contained within the cluster shape. This would result in 254 excess events. Since these pseudo experiments do not account for the systematic effect presented in chapter 9.7, the actual number should be slightly lower. Based on all available information the best estimate for the number of excess events in the cluster shape is about 220 to 230.

9.6 Distributions

The intention of the checks of various distributions in this chapter is to either detect uncompensated systematic effects, which could have artificially created or fostered the observed cluster, or to notice potentially interesting deviations between the cluster region and the background.

The first presented check of a distribution is the composition of the track reconstruction algorithms as selected by Selectfit seen in Figure 104. It shows how frequent each algorithm was selected for the events located within the cluster in red compared the events in the rest of the sky outside of the cluster in blue. A discrepancy between the two regions in this check could either reveal instabilities of Selectfit or show deviations in the properties of the events in and outside of the cluster. This could occur because the algorithms behave differently e.g. for events with different energies. Since not many of the events inside the cluster can be signal events even if there were to be very strong sources within the cluster, no deviation is to be expected.

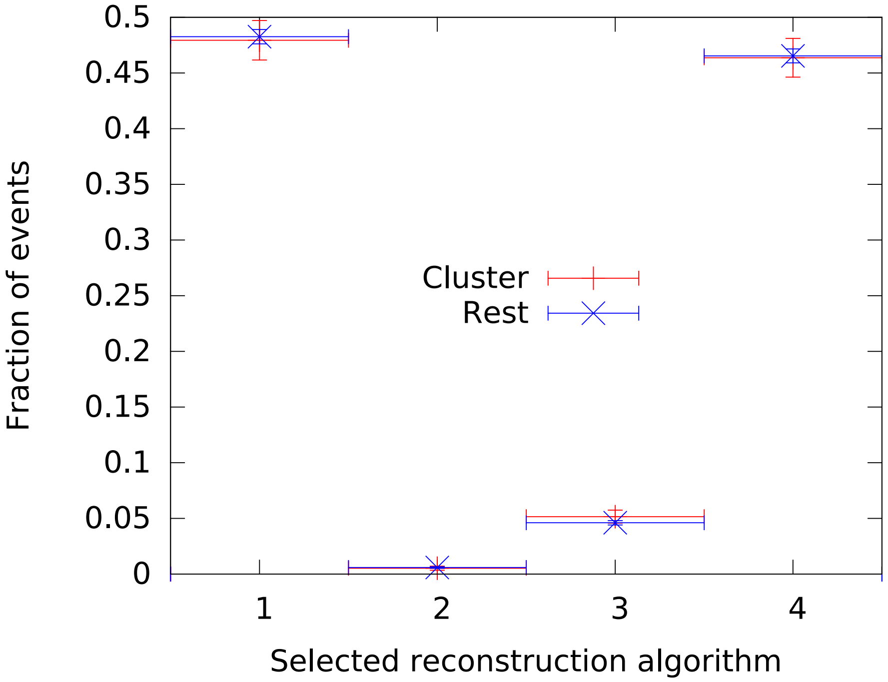


Figure 104: Normalized fractions of the reconstruction algorithms selected by Selectfit in and outside of the cluster. 1=Aafit, 2=Bbfit, 3=Bbfit with M-Estimator, 4=Gridfit. No significant deviation between the regions is observed.

The next check in Figure 105 shows a comparison of the number of events in each month, accumulated over the years. What meets the eye is that the data taking efficiency drops considerably during spring and summer. This is mainly caused by increased bioluminescence reducing the number of clearly recognizable events. This effect is well known within the ANTARES collaboration and is not caused by any methods specific to this analysis. What can also be seen is that the trend is the same for the events within the cluster and outside of it, hence no additional systematic becomes obvious here.

Figure 106 shows a time resolved comparison too, but this time the number of events is plotted against the hour in local time at the ANTARES location (UTC+1, neglecting swapping between summer and winter time). While the

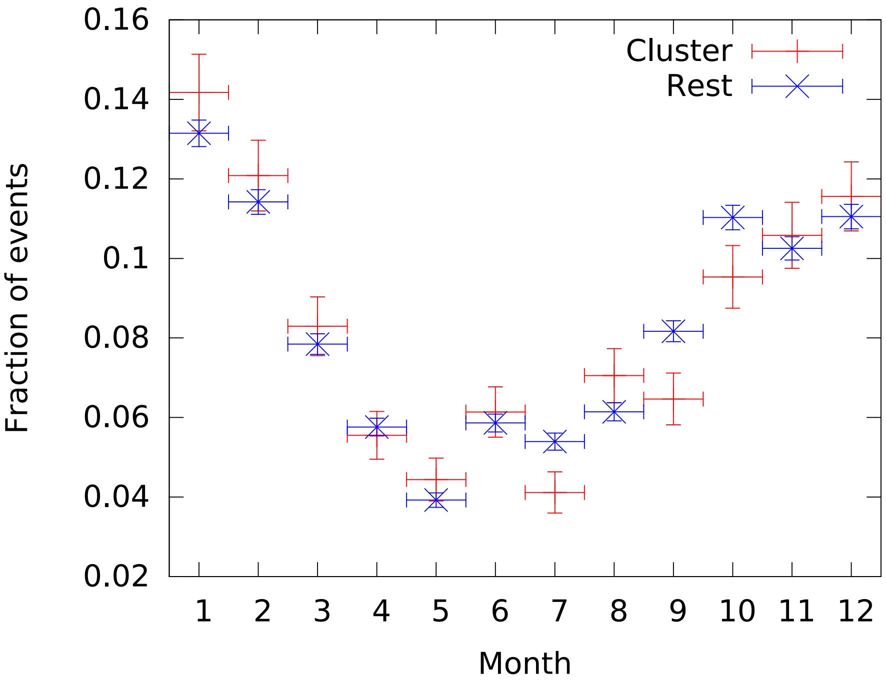


Figure 105: Comparison of the normalized fraction of events recorded in and outside of the cluster against the time of the year. There is a large seasonal variation observed in the ANTARES data taking, mainly due to different bioluminescence rates. No systematic deviations are obvious here between the cluster and the rest of the skymap.

number of events in the rest of the sky seems rather continuous, there are considerable deviations visible between the number of events detected within the cluster during day and night. There are multiple effects that can plausibly explain this behavior. First of all, obviously a different region of the sky is visible during the day than during the following night, simply due to the ≈ 180° rotation of the Earth during the 12 hour period. If the data were taken during one single day and night only, the cluster with its maximal extension of up to 160° in right ascension (the core structure is about 40° wide) would be expected to be fully visible only for about 10.67 hours (2.67 hours for the core). These periods would be extended by a smooth transition of partial visibility for those times where parts of the cluster are not visible straight through the Earth, but with a different efficiency under a zenith angle between 180° (straight upgoing) and 90° (exactly horizontal).

Over the course of a year one would assume this visibility effect to be compensated by the rotation of the Earth around the Sun, because objects which are best visible during the daytime in summer are best visible during nighttime in winter. If one takes the observed effect of seasonal variations as seen in Figure 105 into account, it becomes clear that one expects a variation like the one seen in Figure 106 for a region of the sky which is best visible at night in the

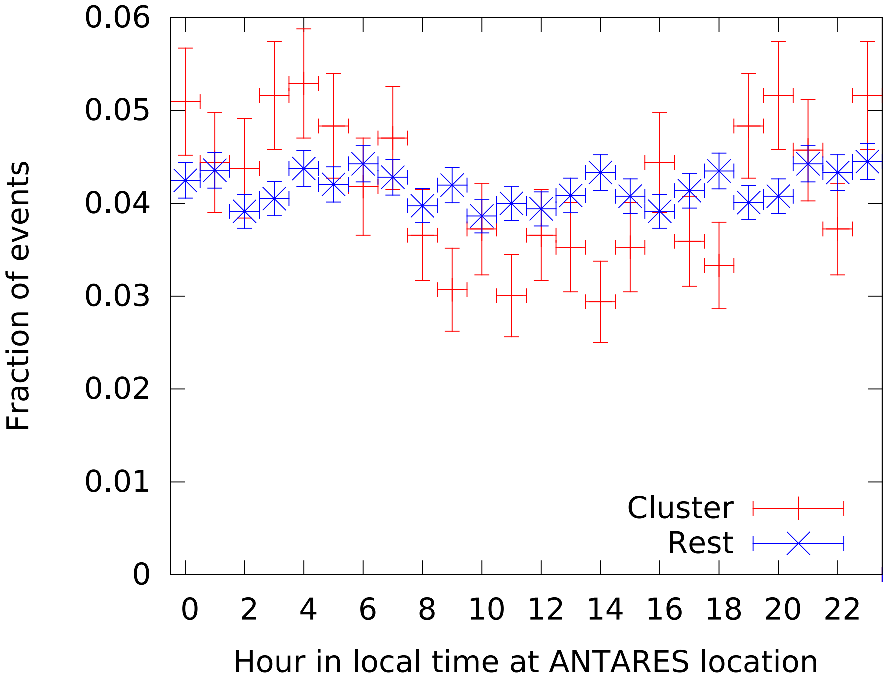


Figure 106: Comparison of the normalized distribution of events in and outside of the cluster against the daytime the events were recorded in local time.

winter and at day in the summer.

As a matter of fact, the cluster is close to the region best visible during winter nights. The exact date and time with direct orientation on the cluster may not be perfectly clear from the Figures alone, but based on Figure 105, beginning of January seems like a reasonable guess, and based on Figure 106, the center of the overfluctuations could be around 1 am. On January 1st, at 1 am, above the opposite side of the Earth, the direct visibility of ANTARES is centered approximately at a right ascension of 287° and a declination of -19°. This is relatively close to the approximated center of the cluster region with a right ascension of 255° and declination of -15°, especially when the extensions of the cluster are considered.

A completely different effect, which could lead to a drop during daytime, is the operation of the telescope by humans. Reboots of unresponsive hardware are performed, experimental configuration setups are tested and calibrations of the detector are conducted, all during working hours, all introducing short periods of time where no useful data for a physics analysis can be taken. This explanation is also very plausible since the drop is observed between 8 am and 9 am, about the designated starting time of shift duties in ANTARES. It also ends between 17 pm and 18 pm, the reasonable estimate for the end of most shifts. If this effect is actually dominant, the drop should also be clearly visible in the distribution for all events against the local time, not just for the cluster. This comparison is shown as a zoom to the relevant range of the y-axis in Figure

109. Indeed, the drop is clearly visible, although not as pronounced as for the cluster.

Why is the drop not visible for the rest of the sky in Figure 106 then? From the first plausible explanation, the combination of efficiency variations throughout the year and differences in visibility, one would expect a weak overfluctuation for the rest during daytime. It cannot be as pronounced as the underfluctuation for the cluster, since the rest of the sky is bigger than the cluster region and therefore the variations in visibility must be less pronounced, but there should still be an overfluctuation there. Depending on the exact strength of the effect of human shifter interaction pulling in the opposite direction, anything between a slight overfluctuation (only the visibility effect contributes) and the same underfluctuation as observed for the cluster region (only the shifter interaction contributes) is compatible. The zoom to the relevant range of the y-axis for distribution of the rest of the sky alone in Figure 107 reveals, that in fact there is a small underfluctuation. Therefore the interaction of the shifters seems to be the dominant effect.

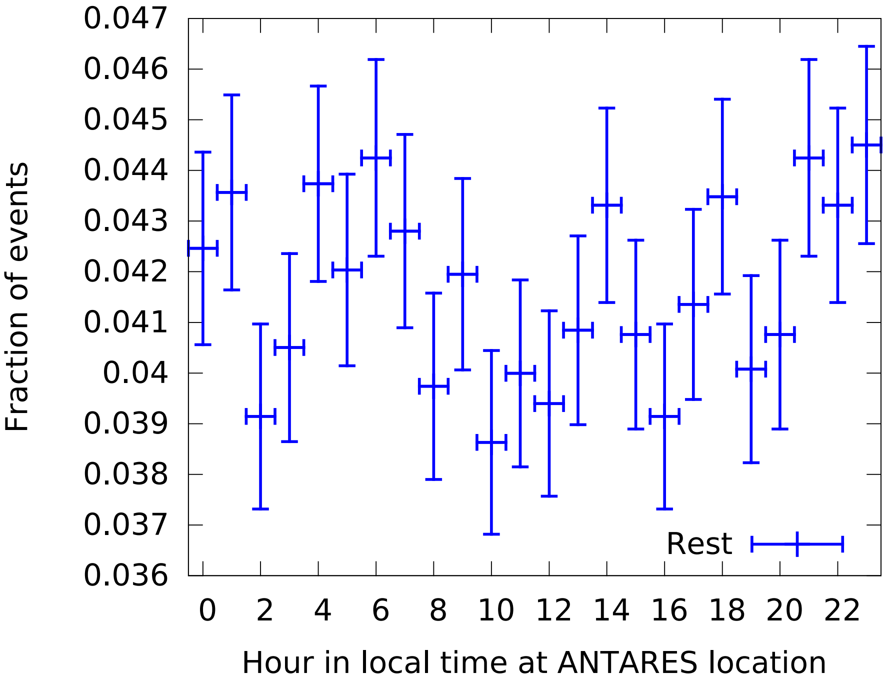


Figure 107: Zoom to the normalized distribution of events outside of the cluster against the hour the events were recorded in local time.

It should be noted that other effects could possibly contribute to these variations, too. For instance a seasonal variation can be induced by the change of the effective temperature of the atmosphere on the other side of the Earth as described in [107]. It could possibly be noticed in a plot of the zenith angle of the events against the month. A seasonal effect on the number of neutrinos resulting from interactions in the atmosphere on the other side of the Earth can be expected to move in zenith with the seasons. If on the other hand it is caused by seasonal variations at the ANTARES site, likely these wouldn't show a variation in zenith distribution. Unfortunately there is not enough statistics to answer this without a doubt.

To investigate a completely different topic, the number of events for each hour of data taking is presented in Figure 108. This is meant to check if a large number of events are recorded almost at the same time, for instance due to an unknown malfunction of the detector or a extremely unlucky configuration of bioluminescent macroorganisms or other unprecedented phenomena. If many events are recorded during such a short time period with approximately the same hit pattern, in theory they might be reconstructed as many neutrino events from similar directions. As one can see, there is no such accumulation in time.

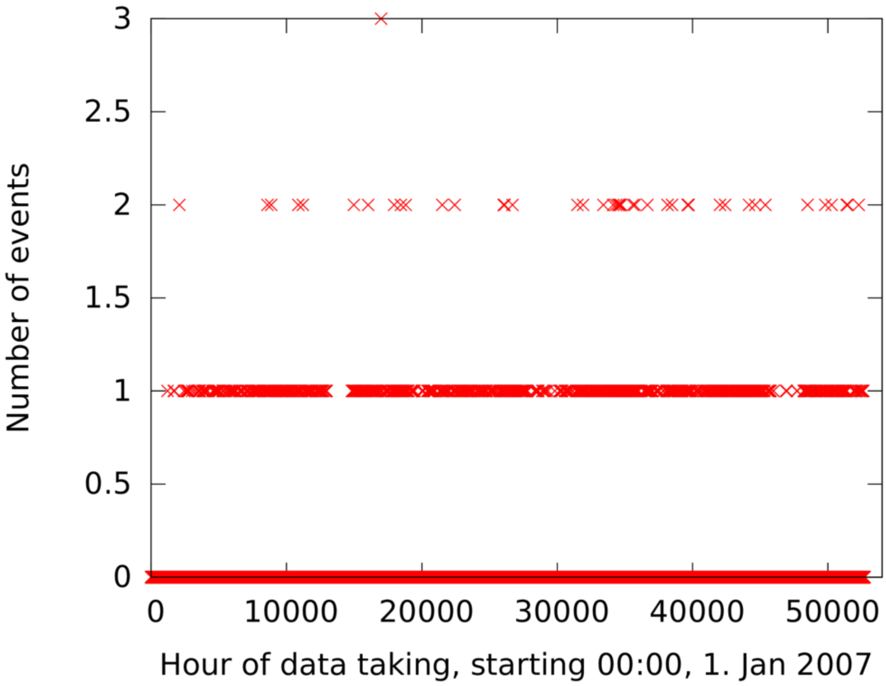


Figure 108: Number of events inside the cluster against the hour the events were recorded. The aim of this plot is to detect spikes which could for instance indicate problems with the data taking at a certain time. No significant spike is observed.

Several further checks that showed no unexpected deviations beyond what can be presented here have been conducted. Some other checks, for instance a detailed evaluation of the energy spectrum, would be interesting at this point. But studies of plain physical properties would introduce a bias and therefore a trial factor for follow-up analyses, which should include these physical properties in a detailed evaluation to maximize their discovery potential. Since such a trial factor reduces their potential for a significant result, no properties which are likely to contribute to the significance of such an analysis are evaluated here prematurely.

9.7 Systematic effect of variable data taking efficiency

The combination of the variations observed in the data taking efficiency over the course of a day as seen in Figures 106 and 109, and of a year, seen in Figures 105 and 110, can potentially introduce an asymmetry in the data taking efficiency even in equatorial coordinates.

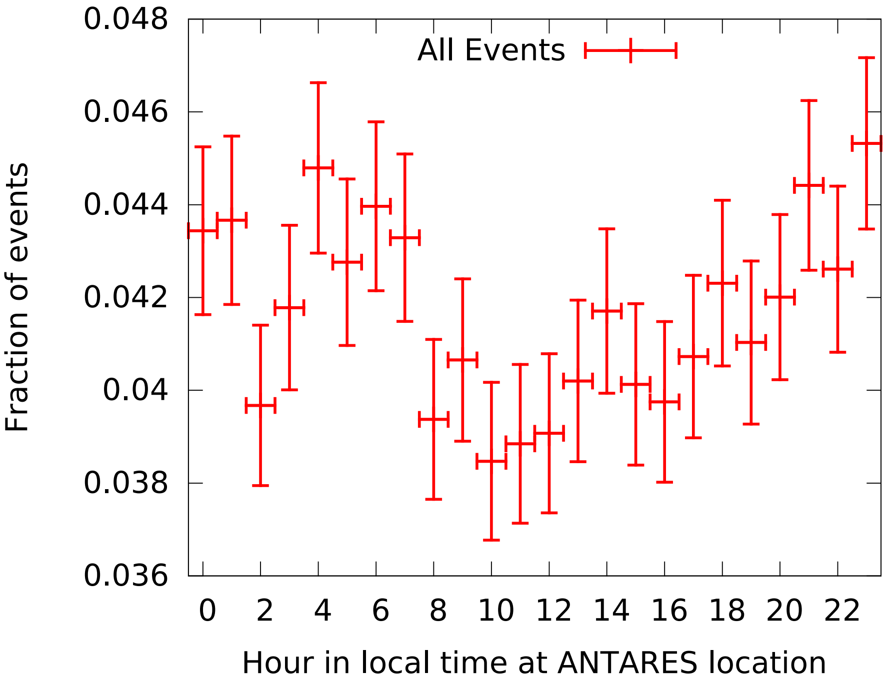


Figure 109: The distribution of all observed events against the hour of the day.

Pseudo-experiments have been conducted which include unrealistically strong variations for these distributions, resulting in an increased data taking efficiency near the location of the observed cluster. The deviations used have been 100% reduced efficiency during April, May and June and ≈ 50% reduced efficiency during day time. The effect in equatorial coordinates is shown in Figure 111. The center of the created overfluctuation is located precisely at the estimated maximum of this effect as given already in the explanation of Figure 106 in chapter 9.6. Without the systematic effect the distribution of events would be expected to be similar to Figure 117.

As a next step, pseudo-experiments have been designed to model the realistic, actually observed variations in data taking efficiency. The resulting distributions for 30 combined pseudo-experiments, shown in Figures 112 and 113, match the actual variations nicely.

The resulting skymaps of the 30 pseudo-experiments can be seen in Figures

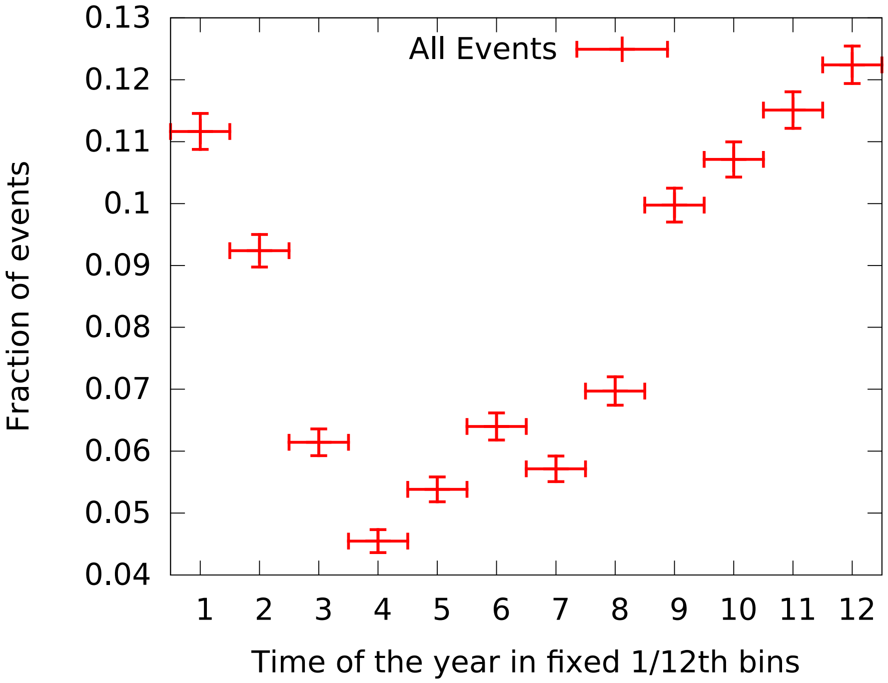


Figure 110: The distribution of all observed events against $\frac{1}{12}$th of the year (Time periods of fixed length give a more stable impression than months).

114 and 115. They show only a weak trend to create or foster an extended structure near the observed cluster. The most significant result of the 30, shown in Figure 116, even shows a result uncorrelated to the result observed for the ANTARES data.

To compare the realistic effect on the number of events, a combination of 100 pseudo-experiments resulted in the distribution of events as shown in Figure 117. In contrast to the unrealistically strong variations from Figure 111, the effect is not visible anymore.

To quantitatively study the dipole-like impact of the systematic effect on the number of events, a projection of the events to one dimension, right ascension, is performed. As seen in Figure 118, the effect of the systematic on the normalized number of events is on the order of ±1%, while the statistical fluctuations for each individual pseudo-experiment with 13283 ± 115 events are on the order of ±10%.

Nevertheless there is a systematic effect increasing the probability to observe a cluster at the location found in the ANTARES data. A solution to compensate for it is to randomly redistribute the exact event times when scrambling the data instead of generating new random times anywhere within the data taking period as performed in chapter 8.3.3. This preserves the variations observed throughout the year and the day in the pseudo-experiments, hence they include the systematic effect. With this change for the pseudo-experiments, the significance of the observed cluster described in chapter 9.1 drops to 2.52σ (from

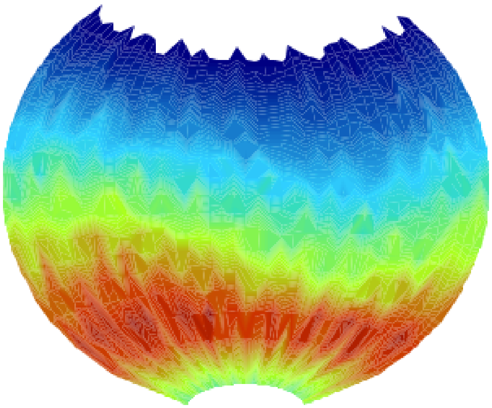


Figure 111: The effect of unrealistically strong variations of the efficiency of the data taking over time on the expected number of events in equatorial coordinates. In this color code, red indicates most, blue least events expected at a location. Image credit: Dr. Thomas Eberl, ECAP, FAU Erlangen.

2.85σ). This way of producing pseudo-experiments slightly underestimates the significance, since pairs and groups of exactly the same events are more likely to reoccur due to the limited number of possible times to assign. This underestimation is well below 0.1σ, since using pseudo-experiments with a scrambling of the exact times, but with a randomized minute value, does not suffer from such a degeneracy and yields 2.61σ.

There is another aspect that could theoretically influence the probability to observe the measured cluster in the pseudo-experiments. When the scrambling of exact event times is considered to be exact, one silently assumes that the probability to measure an event at a time t_1 with properties p_1 (zenith, azimuth, energy, noise level, ...) is exactly the same as to measure an event with p_1 at a different time t_2. If for example the zenith distribution were to be very different for day and night, the random scrambling would have to take that into account as well, for instance by scrambling only between time periods with similar distributions. The method described in chapter 8 evaluates only declination and right ascension. Hence only the parameters time, zenith and azimuth, from which the equatorial coordinates are derived, can potentially influence the result.

Earlier in chapter 9.7 studies have shown the variabilities in the time distributions, which on their own are compensated by the scrambling of exact event times. Neglecting even higher order effects, a systematic discrepancy between day and night would be expected to be necessary in one of the remaining relevant properties, zenith or azimuth, to foster an artificial cluster beyond the already discussed effects. Figure 119 shows the zenith distribution of the events separated into one day (all events between 8 - 20) and one night (20 - 8) dataset. A comparison of the azimuth distributions of these datasets can be seen in Figure 120.

Since both distributions do not show systematic deviations, it is assumed

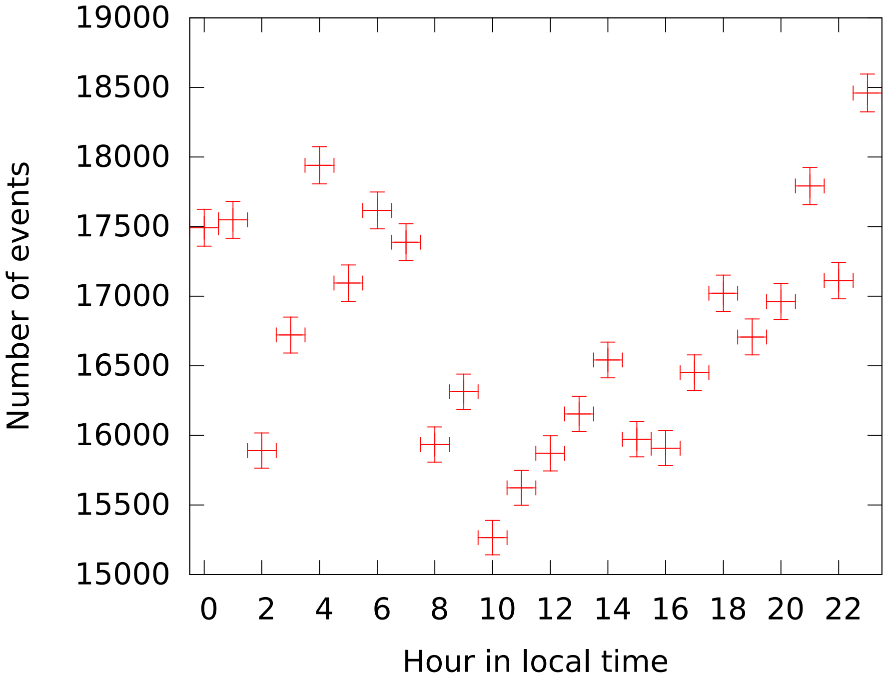


Figure 112: The distribution of the number of simulated events against the hour of the day.

here, that no additional systematic effect relevant for this analysis is introduced by variations of the probabilities to measure event properties.

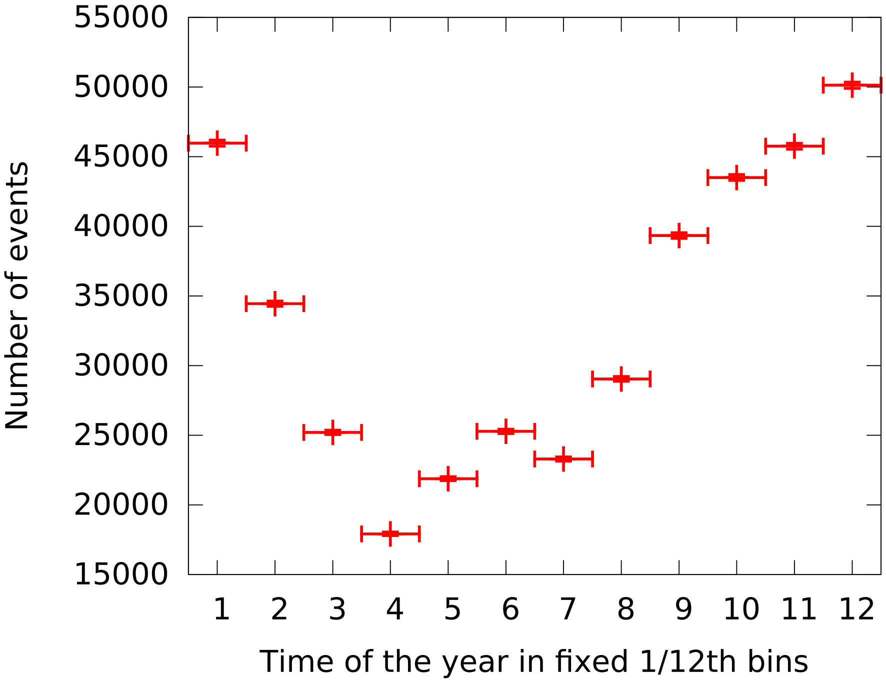
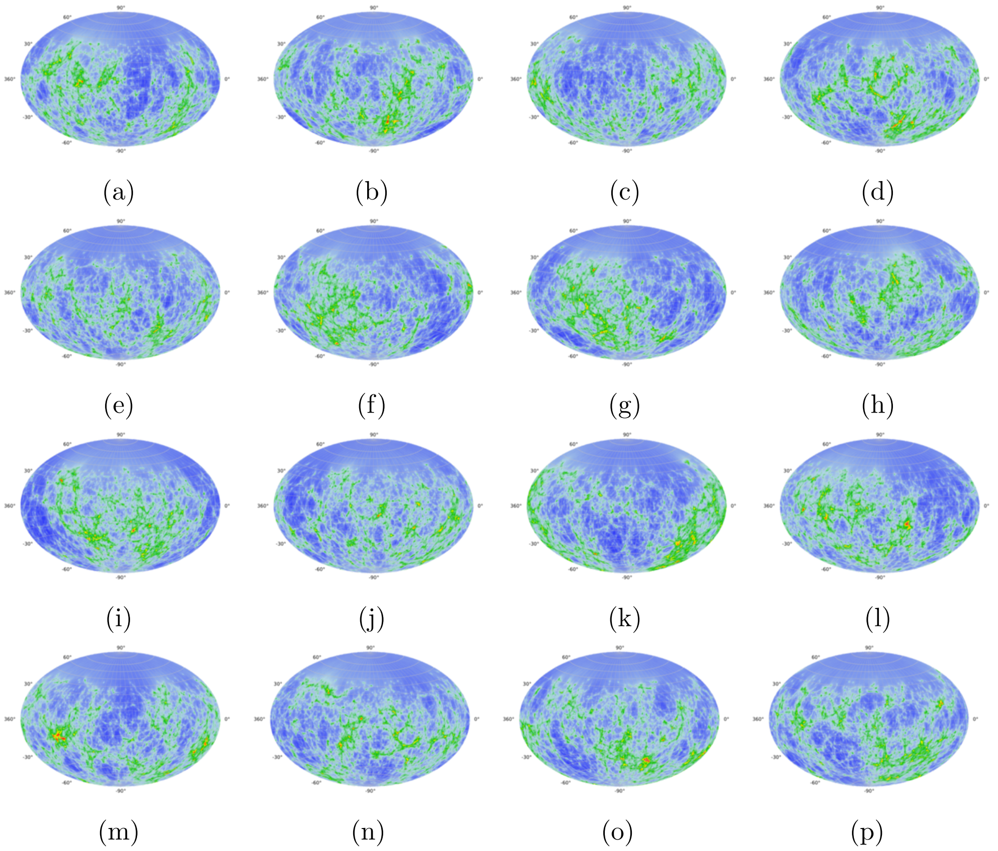


Figure 113: The distribution of the number of simulated events against $\frac{1}{12}$th of the year.



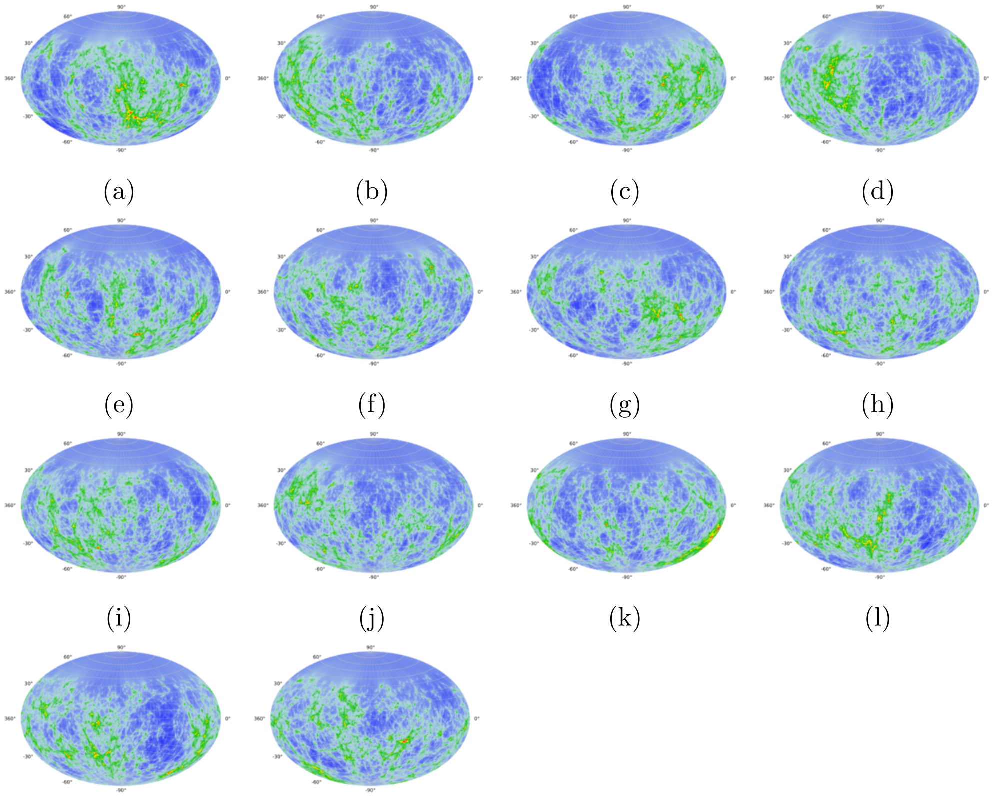
(a) (b) (c) (d)

(e) (f) (g) (h)

(i) (j) (k) (l)

(m) (n) (o) (p)

Figure 114: First 16 resulting skymaps from 30 pseudo-experiments with realistic data taking efficiency variations. (Zoomable)



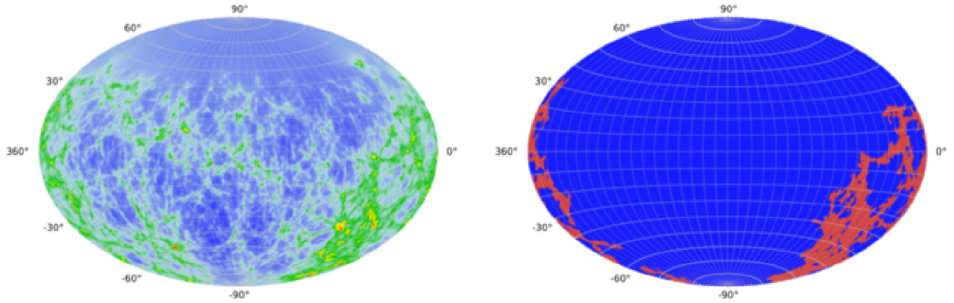
(a) (b) (c) (d)

(e) (f) (g) (h)

(i) (j) (k) (l)

(m) (n)

Figure 115: Remaining 14 resulting skymaps from 30 pseudo-experiments with realistic data taking efficiency variations. (Zoomable)



(a) (b)

Figure 116: Most significant cluster (2.96σ) of 30 pseudo-experiments with realistic data taking efficiency variations. It isn't notably similar to the results obtained for the ANTARES data.

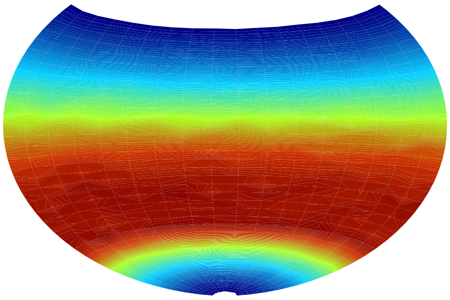
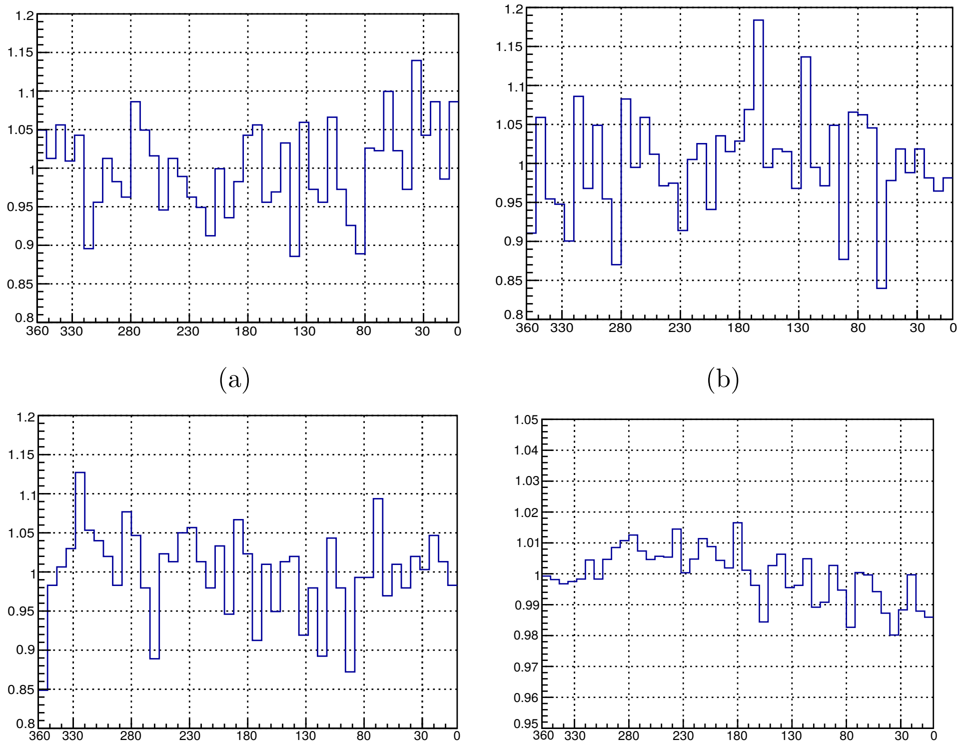


Figure 117: The effect of realistic time variations in data taking efficiency on the expected number of events in equatorial coordinates. In this color code, red indicates most, blue least events expected at a location. Image credit: Dr. Thomas Eberl, ECAP, FAU Erlangen.



(a)

(b)

(c)

(d)

Figure 118: Variation in the normalized number of events projected to right ascension for pseudo-experiments with realistic time variations. a) b) and c) The distribution for one pseudo-experiment each. The statistical fluctuations dominate. a) Shows the distribution for the most significant pseudo-experiment as shown in Figure 116. d) The distribution for 100 times the statistic of one pseudo-experiment. The statistical fluctuations are reduced, the smaller systematic effect becomes visible. Image credit: Dr. Thomas Eberl, ECAP, FAU Erlangen.

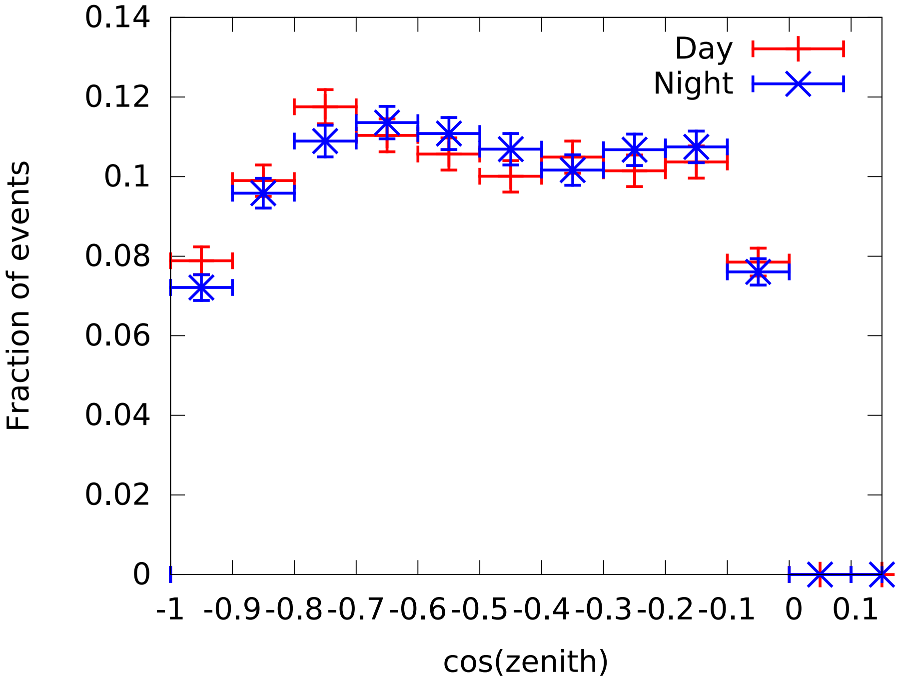


Figure 119: The normalized distribution of events for different zenith bins and separated into day and night datasets.

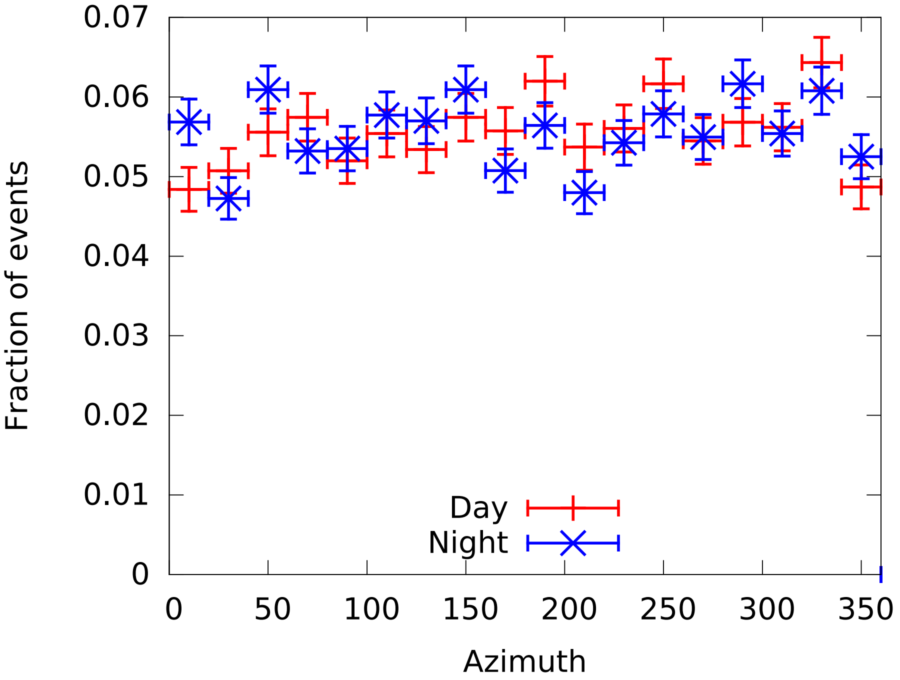


Figure 120: The normalized distribution of events for different azimuth bins and separated day and night datasets.

9.8 Results for public IceCube data

A crosscheck with the publicly available IC40 dataset [106], published by the IceCube collaboration, is performed to see if the cluster found in the ANTARES data is also present in an independent dataset. To use the found cluster as a signal hypothesis, the search had to be modified slightly as described in chapter 8.8. The template used as hypothesis in this evaluation is exactly the result as shown in Figure 84.

The two datasets are well suited for a comparison because they contain the same physical observable, muon-neutrinos. Also the achieved direction resolution of both datasets is not very far apart, with the reconstruction of ANTARES being slightly more accurate (here $0.44°$ instead of $0.7°$, see [108]). Although it is not necessary, the two datasets also happen to contain a similar number of events (13283 versus 12877 events). The only aspect that differs in a way that could influence the results notably is the different energy composition. Features that are observed in lower energies should be more pronounced in the ANTARES data, while aspects that are more dominant for higher energies should be better visible in the IC40 data sample. Still, overall this dataset offers an excellent way to crosscheck the ANTARES results.

The obtained resulting cluster from this evaluation can be seen in Figure 121. It has a significance of 2.14σ. Since the evaluation requires an overlap between the template from the ANTARES data and a cluster found in the IC 40 data, only clusters at the correct location are considered in this evaluation and are visible in the skymap. The detailed structures before any segmentation is applied can be seen in Figure 122 in equatorial, in Figure 123 in galactic coordinates. While some features observed in this skymap agree nicely with the ANTARES result, see Figure 87 for comparison, others don't. Besides the slightly different sensitivities in energy for the two datasets, it is already well known, e.g. from chapters 9.4, 9.5 and 9.7, that statistical fluctuations play a relevant role for some of the features found in these small datasets. Therefore the agreements and discrepancies between the results neither unambiguously confirm nor reject the result from the ANTARES data set.

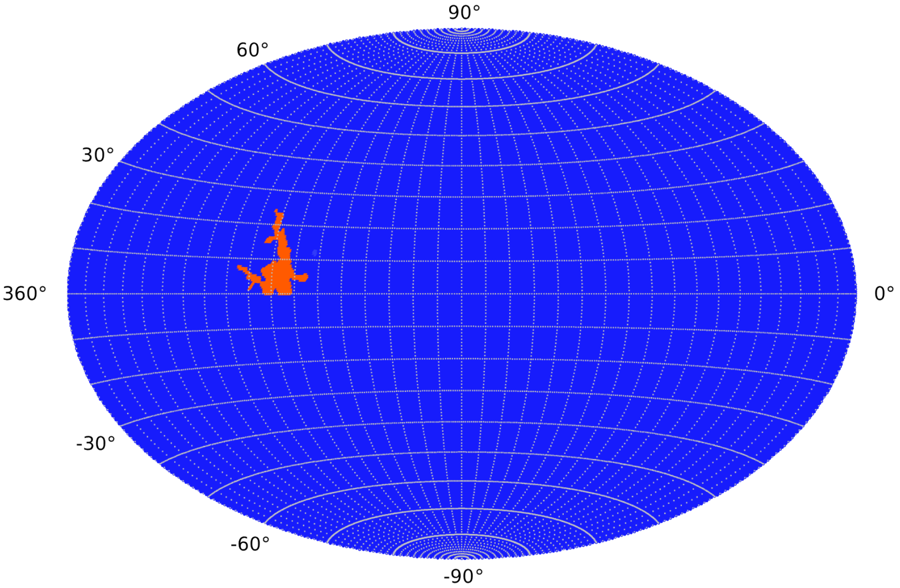


Figure 121: Result of the evaluation of the IC40 data searching with a segmentation threshold of $\alpha = 0.25$ for a cluster that overlaps with the hypothesis from ANTARES data as explained in chapter 8.8 in equatorial coordinates.

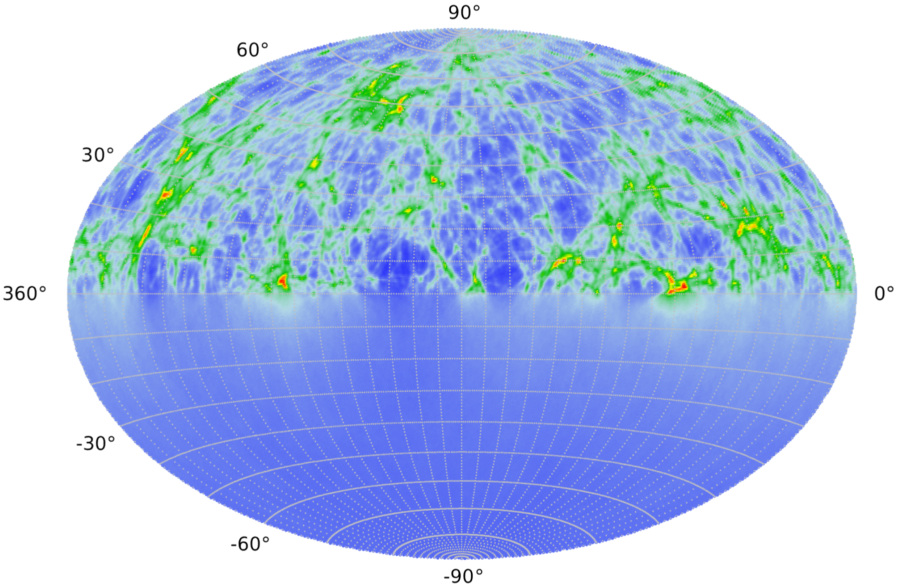


Figure 122: Result of the evaluation of the IC40 data before segmentation in equatorial coordinates.

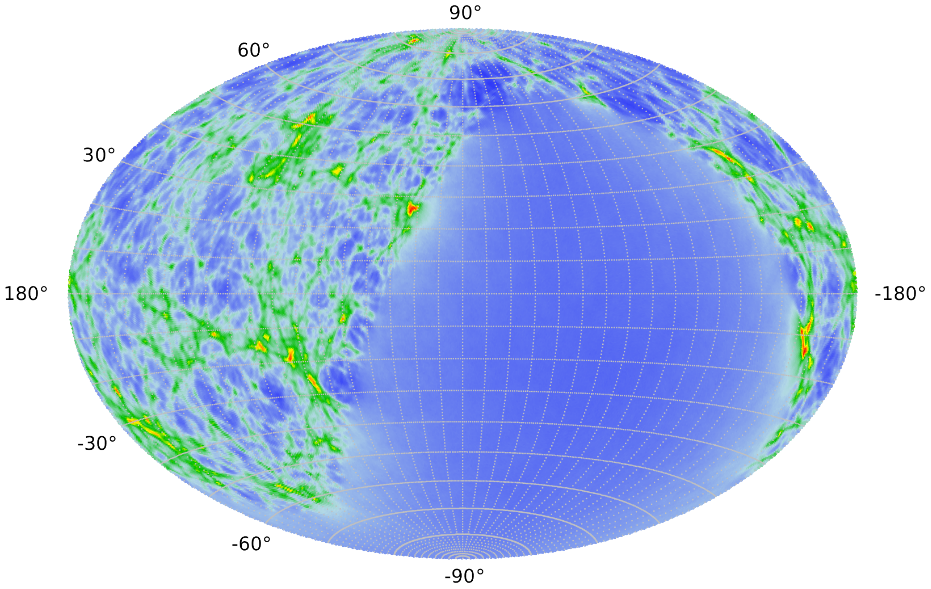


Figure 123: Result of the evaluation of the IC40 data before segmentation in galactic coordinates.

10 Interpretation of the results

The number of events obtained in this thesis is considerably larger than in most ANTARES analyses that apply the default methods and cuts. This gain is partially achieved by the improved event selection and reconstruction, but also due to looser cuts. With looser cuts, there are also more misreconstructed atmospheric muons in the obtained data set. Since misreconstructed events are distributed throughout the sky, their moderate increase constitutes no issue for an analysis aiming to detect spatially resolved sources.

The increased number of events is beneficial for this analysis not just by the availability of more statistics and therefore more signal, but it also improves the process of estimating the various expectations. Since in this search all expectations are derived from data, their estimation becomes more accurate with more data too.

The structures found with the segmentation cut of α = -0.11 as seen in Figure 81 are not particularly uncommon, as they are neither significant by their size, nor does the detailed structure at their particular locations, seen in Figure 87, look special in any way, e.g. like unexpectedly strong overfluctuations. On the other hand the obvious fact that the clusters themselves are spatially clustered is special and is also the reason why the very large structure forms in the evaluation using $\alpha = +0.25$ as seen in Figure 84.

The details of these results are not completely trivial to interpret correctly. The statement that can be made for the large cluster based on the significance of 2.52σ is "An anisotropy in the variable evaluated in the multiscale search between the region exactly defined by the cluster shape and the rest of the sky with a size as large or larger as the one observed in the ANTARES data is generated with a probability of not more than p = 1.2% by a random distribution of events that includes the data taking systematics of ANTARES." The corresponding statements are of course also true for the less significant clusters found with the segmentation threshold of α = -0.11.

The multitude of checks for systematic differences between the cluster and the rest of the sky, and also between the obtained data sample and the expectations provides additional insight into the result. The checks presented in chapters between 9.3 and 9.6 neither show any problems with the search method nor give clear hints on how to interpret the results. While these comparisons of the cluster and the rest of the sky show no obvious deviations from the expected behavior, these detailed checks led to the awareness of the relevant systematic effect investigated in chapter 9.7, which had been unaccounted for. Although it adds only a minor contribution to the observed result, investigations like this still can add to the understanding of the ANTARES neutrino telescope.

As already explained in chapter 8, the goal for this search, like for most model-independent analyses in other experiments, has been to identify the most significant excess beyond the expectations. This has been fully achieved by this search.

Since, on their own, all obtained results are well below 3.0σ, the common convention for an indication in Astroparticle physics, they are best explained by a random fluctuation of the background.

Nevertheless a few theoretically possible speculations which involve an actual neutrino signal shall be mentioned briefly here. First of all, no single object is likely to generate a cluster of the observed size. Although this is not disfavored by the obtained results, a more plausible speculation than one very large extended source region would be that multiple faint sources of unknown positions and extensions are distributed within the cluster together with random fluctuations. In simulations of similar scenarios the usage of a less strict cut for the segmentation value α allows neighboring simulated sources to be detected as one large cluster of higher significance. In this case the most likely location for sources would be at the positions of highest local overfluctuations in Figure 87. But even if this actually were to be true for some of these regions, others are *guaranteed* to be fostered or even caused by random background fluctuations as seen for example when comparing to the results in Figure 93. Therefore the highest overfluctuations cannot directly be considered "possible source locations" based on the results of this search alone. Furthermore, in simulations with remotely realistic artificial sources and positive α values used for segmentation, the obtained shape is usually not the exact shape of the artificial source, seen for instance in Figures 52b or 65. Therefore it is highly unplausible to assume that the detected shape of the large cluster should be exact. Random fluctuations at the edge are certain to influence its extension.

Concerning objects that could be expected to have similarly large extensions in a similar direction than the obtained result, one interesting aspect is that the observed cluster roughly extends in the direction of the center of our own galaxy (and beyond), but by eye its detailed structure does not coincide well with "usual galactic suspects" for neutrino sources like the Fermi Bubbles, see Figures 124 or the Galactic Center itself. Figure 125 presents the magnetic

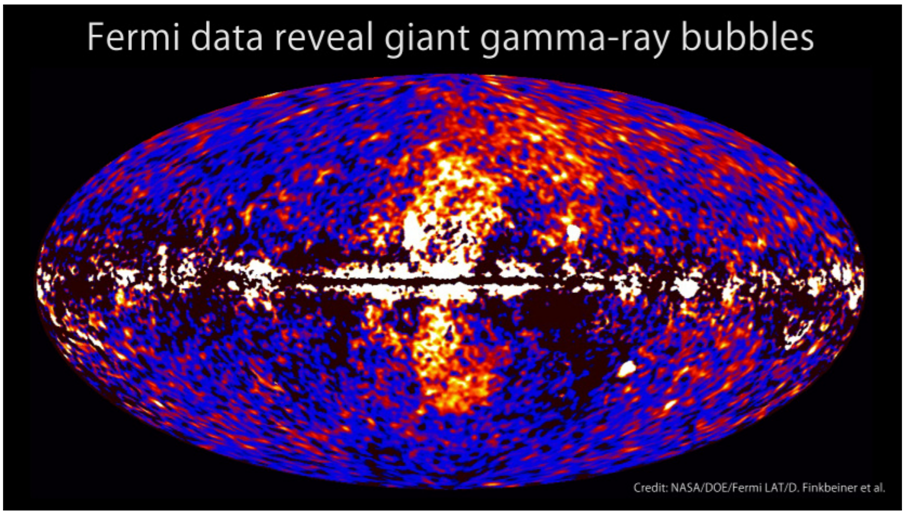


Figure 124: The Fermi Bubbles are one of the typical candidates for extended galactic sources. This skymap is in galactic coordinates. Image Credit: NASA/-DOE/Fermi LAT/D. Finkbeiner et al. Taken from [109].

substructure of the Fermi Bubbles, which shows more similar structures like the obtained ANTARES results, but it's still no good match. Detailed searches for enhanced neutrino emission from these extended regions have been performed e.g. in [110] and [111] without finding a significant excess.

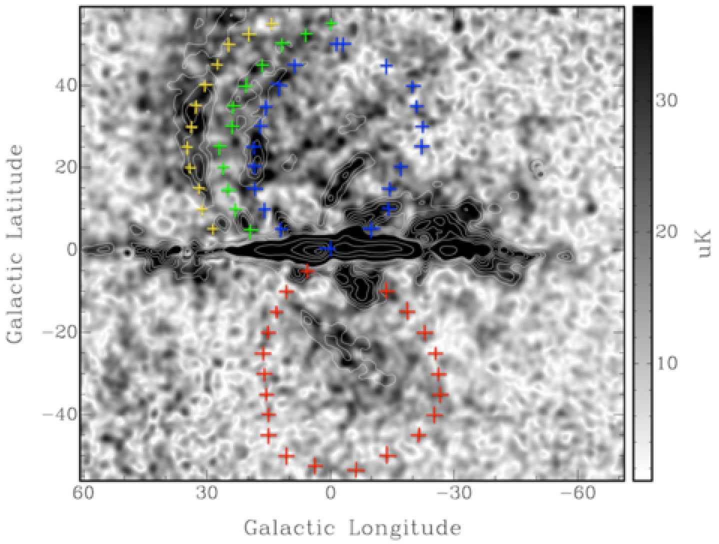


Figure 125: The magnetic substructure of the inner Galaxy in galactic coordinates. Shown is an overlay of the 33 GHz and 23 GHz total polarized intensity from 5 years data of WMAP. Blue and red are the shapes of the Northern and Southern Fermi Bubbles, yellow and green are the so called "Northern arcs". Taken from [112].

Another structure that comes to mind for these extensions is the Galactic Halo, described for instance in [113]. This diffuse, approximately spherical cloud of gas and stars around our galactic disk could house a structure with the observed extensions. In this context it can be noticed that in the detailed structure of the results in Figure 88 there is also an overfluctuation visible roughly in the direction of the Small and Large Magellanic Cloud, some of the closest smaller galaxies, see Figure 24, but not in the direction of the Andromeda Galaxy, the nearest galaxy with similar extensions than the Milky Way. Apart from a direct correlation with matter in galaxies (e.g. many faint galactic sources), dark matter could also be a candidate for a neutrino production mechanism in such a scenario, see for example [114]. Various other structures with similarly large extensions and approximately in the same direction in the sky are known, see for instance [115]. But just as for the other already mentioned objects, by the results alone there is no clear evidence to support any of these speculations, which even emphasizes the need for a dedicated follow-up analysis.

An overlay of the ANTARES skymap with the events found in the high energy starting events analysis by IceCube, described in [26], with four years of data, see [116], is show in Figure 127. This overlay is meant as an illustration

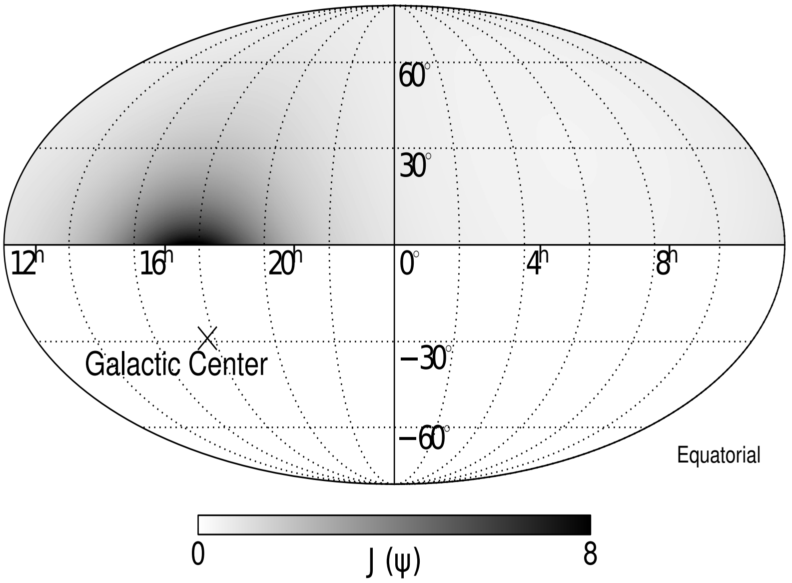


Figure 126: The anisotropy in the neutrino flux from dark matter self-annihilation is expected to follow the line-of-sight integral of the Navarro-Frenk-White halo profile. The profile is shown here (dimensionless) for the Northern hemisphere in equatorial coordinates, but with a different way of plotting right ascension. Image taken from [114].

only, no correlation analysis between the two results has been performed. The energies of the neutrinos differ, as the IceCube sample contains neutrinos with much higher energies than the ANTARES analysis. The crosses mark shower events, the x-es mark the position of tracks. It is important to note that the error on the position of the shower events is large, typically above 10°, see [27]. While the locations of some accumulations seem to agree by eye, e.g. left of the Galactic Center or at the top right end of the cluster, others do not agree. Therefore no unambiguous lesson can be learned from this comparison either.

It is noteworthy that the evaluation of the completely independent IC40 dataset found a cluster at the location predicted by the hypothesis derived from ANTARES data. Since the observed cluster at this position is not extraordinarily pronounced, it still does not provide an unambiguous confirmation of the hypothesis. It would be a mistake to straightforwardly combine the two obtained significances, since the they do not have exactly the same shape. But the evaluation cannot require them to have exactly the same shape, since the contribution of random fluctuations to the result is too large. Looking at the detailed structure of the result in Figure 122, the adaption of the search strategy to the visibility of IceCube as described in chapter 8.8 seems to slightly overestimate fluctuations at the edge of the visibility near the horizon. Nevertheless this does not invalidate the obtained results in any way, as the exact same ef-

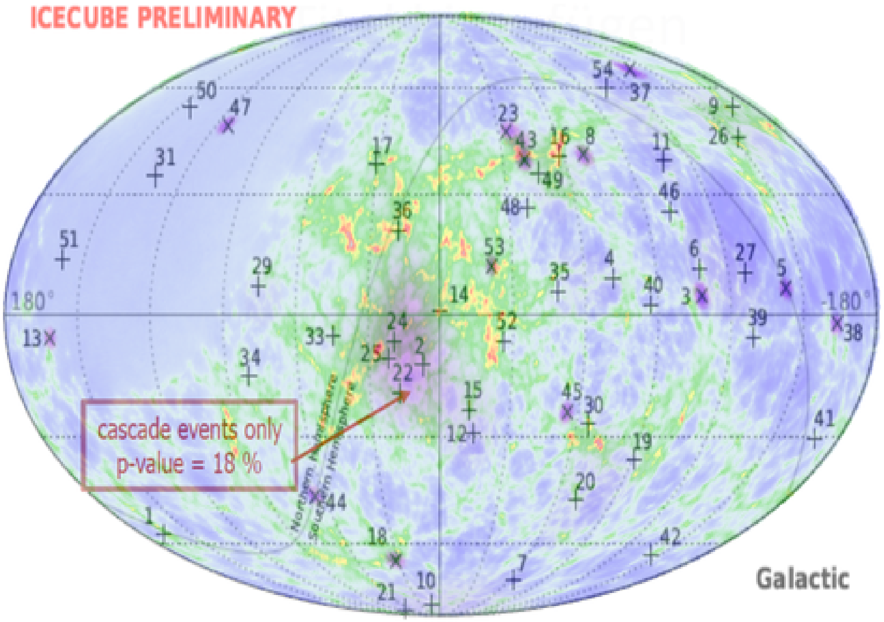


Figure 127: Overlay of preliminary 4 year IceCube HESE events with the skymap derived from ANTARES data.

fect is also contained in every pseudo-experiment, from which the significance is derived in the end. What contributes to the ambiguity of the results is that not all features in the two skymaps in Figures 87 and 122 agree well. While this is not a flaw in the method, but simply the influence of random fluctuations on small samples, it demonstrates that both of these results obtained from small statistics are not accurate enough to trust every detail.

A dedicated follow-up analysis based on the obtained results would be of great value to improve the discrimination between a statistical fluctuation and a potential indication of a faint signal. The main advantage of such an analysis would be that the location, size and shape of the source hypothesis are known instead of having to search for all possible locations and sizes at once. This decisively reduces the trial factors, which have been incorporated into this analysis implicitly by the free search in the pseudo-experiments. Using a large, independent dataset would be ideal to confirm or exclude the presence of an increased flux of neutrinos within the identified region.

Putting it all in a nutshell, the conducted analysis has fully fullfilled the given task, to identify the region with the most significant excess of recorded neutrinos, and the presented observations are worth further investigation.

11 Summary

This thesis describes three new methods designed for the data analysis of the ANTARES neutrino telescope, but versatile enough for many other applications. The first algorithm is a highly optimized scheme to deal with the signal selection required to identify the desired neutrino signal within a vast amount of data. It uses adapted cascades of Random Decision Forest classifiers to achieve this goal. This approach has been incorporated into the data selection to benefit the data analysis performed in this thesis. The second algorithm is an efficient combination of multiple independent, already available algorithms for the direction reconstruction of the observed muon-neutrinos. Internally this selection is again based on a Random Decision Forest classification. It clearly outperforms any individual reconstruction algorithm and the method in general is flexible enough to be used for different reconstruction tasks or even to combine different types of signal in one analysis. This approach has also been employed to enhance the data selection. The third explained algorithm constitutes the main focus of this work. It analyzes the spatial distribution of recorded neutrinos to detect anisotropies of arbitrary extension, shape and internal distribution at any position in the sky. Since the evaluation is based entirely on observed data, no assumptions derived from theories or Monte-Carlo simulations are necessary. Therefore it is more robust against deviations from expectations and can easily be applied to similar tasks.

The evaluation of the recorded ANTARES data using all three described approaches results in a large structure with a post-trial significance of 2.52σ. Multiple checks of the result neither revealed any flaws in the method, nor did they help much to reveal the origin of the observed cluster. On the other hand a systematic asymmetry in equatorial coordinates caused by the time variations of the data taking efficiency of ANTARES was identified and studied. But this systematic effect is already included in the quoted significance of the result. A crosscheck on an independent dataset from IceCube with this method also found a cluster at the location expected from ANTARES data. This cluster results in a significance of 2.14σ. A more detailed comparison shows that not all features seen in the ANTARES result match the ones in the IceCube result and vice versa, but this certainly doesn't invalidate them considering the influence of random fluctuations on data samples of their size.

The goal of this search, to detect the most significant excess of neutrinos in the obtained ANTARES data sample, has been achieved fully. To either identify the results as statistical fluctuations or to detect them significantly a follow-up analysis optimized to investigate the derived hypothesis is highly recommended.

Danksagung

Ich möchte mich bei Gisela Anton für das Bereitstellen meiner Doktorandenstelle und die Möglichkeit zur freien Entfaltung bedanken, bei Julia Schmid, Kerstin Fehn, Jannik Hofestädt, Tamas Gal und meinen anderen Kollegen für die angenehme Atmosphäre und faszinierende Diskussionen und bei meiner Familie und Simone Radke für den Rückhalt.
Ganz besonders bedanken möchte ich mich bei Thomas Eberl für seine außerordentlich gute Betreuung meiner Dissertation und dass er während der gesamten Dauer immer Zeit, Motivation und einen hilfreichen Rat für mich hatte.
Vielen Dank.

Part IV

APPENDIX

A Usage of Selectfit

The following examples are python code that demonstrates how to use Selectfit. They do not show the full scripts, but just the additional parts for Selectfit. A full implementation of the training examples working with ANTARES run-by-run simulations, version 2.2, can be found in Seatray under
antares-rdfclassify/resources/scripts/SelectFit/ExampleStepX.py.

A.1 Application of existing Selectfit

Already trained RDFs are available at the computing center in Lyon at /afs/in2p3.fr/throng/km3net/i3-tools/params/antares-rdfclassify/2014-12-11. The files are called bestReco_AA_BB_BBME_Grid_fit_r14956.rdf for the selection
which algorithm to use and errorEstimate_AA_BB_BBME_Grid_fit_r14956.rdf for the error estimation. Both are trained with the tracksOnly flag, meaning that only Aafit, Bbfit, Bbfit MEstimator and Gridfit are considered.
Here is how to use Selectfit with pretrained RDFs:

```
# Import the classes needed for Selectfit
from icecube.rdfclassify import selectfit
# Load the library used for classification
load(”librdfclassify”)

# This extracts all features from the reconstructions results
# So the results must be already in the frame at this point
tray.AddModule(selectfit.SelectFitFeatureExtractor,”extractFeat”)(
        (”trackOnly”, False)
        # True means Aafit, Bbfit, BBfit MEst and Gridfit
        # False also adds Dusj and Q Strategy
        )

# Selecting which direction reconstruction to use
tray.AddModule(”I3RDFClassify”,”selectFit”)(
        (”RDFToLoad”, ”bestReco.rdf”)
        # The path and filename of an already trained RDF
        # The RDF must be trained with the same number of features
        # and the best reco output classes
        )

# Estimating the error class of the used reconstruction
tray.AddModule(”I3RDFClassify”,”errorEstimate”)(
        (”RDFToLoad”, ”errorEstimate.rdf”)
        # Same as above, but trained with
        # the error estimate output classes
        )

# Collect the information stored by the classifications
# and write an AntaresRecoParticle to the frame named ”SelectFit”.
```

```
# Zenith, Azimuth and QualityParam are set accordingly.
tray.AddModule(selectfit.SelectFitWriter, "writeSelectFitToFrame")
```

A.2 Training anew

Again the following pieces of code are just the additional parts for Selectfit. Have a look at antares-rdfclassify/resources/scripts/SelectFit/ExampleStep1.py to ExampleStep5.py for a chain of scripts that perform the training.

The first step for a training is to extract features from Monte-Carlo simulations. The class numbers are by default automatically set to the best reconstruction algorithm by the module SelectFitFeatureExtractor:

```
from icecube.rdfclassify import selectfit

tray.AddModule(selectfit.SelectFitFeatureExtractor,"extractFeat1")(
        ("fileOutputName", "./featureFilesBestReko/SelectFit.dat"),
        # The location to write the feature files to.
        # All files for the same classification
        # should end up in the same folder
        ("trackOnly", True)
        # True means Aafit, Bbfit, BBfit MEst and Gridfit
        # False also adds Dusj and Q Strategy
        )
```

The second step uses the written .dat files containing the features to train a RDF which will then be able to select the best reconstruction result for unseen events. Potential complaints about missing .meta files can be safely ignored for this step.

```
load("librdfclassify")

tray.AddModule("I3RDFClassify","rdfclassify")(
        ("NumberOfTrees", 101),
        # The RDF will consist of 101 individual decision trees
        ("RDFToSave", "bestReko.rdf"),
        # The filename to save the RDF to
        ("DataFolder", "./featureFilesBestReko/")
        # The folder where all .dat files
        # for this task have been stored
        )
```

The third step is to extract the features for the error estimation. To be able to estimate what the error of the selected reconstruction for an event is, it must perform the selection first. Therefore this step extracts the features, uses them and the trained bestReco.rdf to select a reconstruction, and then evaluates the angular error of the selected reconstruction according to section 6.5. It is **very important to use different events** than the ones used for the first feature extraction! Using the same events again will result in an unrealistic, perfect selection of the direction reconstruction and therefore change the error estimation.

```
from icecube.rdfclassify import selectfit
load("librdfclassify")
```

```
tray.AddModule(selectfit.SelectFitFeatureExtractor,"extractFeat2")(
        # No need for a feature file here
        # the features are used in the Seatray frame directly
        ("trackOnly", True)
        )

tray.AddModule("I3RDFClassify","selectFit")(
        ("DataName", "SelectFitFeatures"),
        # (optional): The name of the features in the frame
        ("RDFToLoad", "bestReko.rdf"),
        # The filename of the RDF trained in step2
        )

# The class number for the features is changed here
# from best reconstruction to angular error class
tray.AddModule(selectfit.SelectFitErrorEstimator,"resetClassErr")(
        ("fileOutputName","./featureFilesErrorEstimate/errEst.dat")
        )
```

The fourth step again uses the written .dat files for training. This time the class numbers are set such, that the resulting RDF will be able to estimate the angular error of the selected reconstruction for unseen events. Again, potential complaints about missing .meta files can be safely ignored for this step.

```
load("librdfclassify")

tray.AddModule("I3RDFClassify","rdfclassify")(
        ("NumberOfTrees", 101),
        ("RDFToSave", "errorEstimate.rdf"),
        ("DataFolder", "./featureFilesErrorEstimate/")
        )
```

After these four steps the training is done. The two resulting RDFs for selection and error estimation can be used just as explained in section A.1.

B Feature list for Selectfit

The features are stored in files ending in ".dat". The format of these files is one line per event, each line starting with the class number followed by the feature numbers. The ordering is the one presented in this list. All numbers within a line are separated by one whitespace. All variables are taken from the Seatray physics frame. The features for Selectfit are:

1. Number of pulses from "CalibratedPulses"
2. Zenith of Aafit from "AafitFinalFit"
3. Azimuth of Aafit from "AafitFinalFit"
4. Lambda of Aafit from "AafitLambdaFinalFit"
5. Beta of Aafit from "AafitErrorEstimateFinalFit"
6. Zenith of Bbfit from "BBFitTrack"

7. Azimuth of Bbfit from "BBFitTrack"
8. BBFitBChi2 of Bbfit from "BBFitBChi2"
9. BBFitTChi2 of Bbfit from "BBFitTChi2"
10. Zenith of Bbfit MEstimator from "BBFitMEstTrack"
11. Azimuth of Bbfit MEstimator from "BBFitMEstTrack"
12. QualityParam of Bbfit MEstimator from "BBFitMEstTrack"
13. Chi2 of Bbfit MEstimator from "BBFitMEstTrack"
14. NumberOfHits of Bbfit MEstimator from "BBFitMEstTrack"
15. NumberOfUsedLines of Bbfit MEstimator from "BBFitMEstTrack"
16. Zenith of Gridfit from "GridFit_FinalFitResult"
17. Azimuth of Gridfit from "GridFit_FinalFitResult"
18. QualityParam of Gridfit from "GridFit_FinalFitResult"
19. Ratio of Gridfit from "GridFit_Ratio"
20. Ratio of Gridfit from "GridFit_Ratio_Precut"
21. XParameter of Gridfit from "GridFit_X_parameter"
22. Zenith of Dusj from "DusjShowerRecoFinalFit_FitResult"
23. Azimuth of Dusj from "DusjShowerRecoFinalFit_FitResult"
24. QualityParam of Dusj from "DusjShowerRecoFinalFit_FitResult"
25. CutValue of Dusj from "DusjShowerRecoCutValues"
26. ReducedLogLikelihood of Dusj from "DusjShowerRecoFinalFitReducedLogLikelihood"
27. ReducedLogLikelihood of Dusj from "DusjShowerRecoVertexFitReducedLogLikelihood"
28. ReducedChi2 of Dusj from "ShowerIdentifierReducedChiSquare"
29. ReducedChi2 of Dusj from "FitTimeResidualChiSquare"
30. Zenith of QStrategy from "QStrategyReconstructionResult"
31. Azimuth of QStrategy from "QStrategyReconstructionResult"
32. QualityParam of QStrategy from "QStrategyReconstructionResult"
33. Chi2 of QStrategy from "QStrategyReconstructionResult"

C Sequential execution of linear filters

Multiple application of linear filters gives the same results as the one time application of a larger linear filter. Applying a linear filter twice that operates on three gridpoints is analog to a filter that operates on five gridpoints.

```
twice          [       1          2          1      ] =

1*[      1          2          1          0          0     ] +
2*[      0          1          2          1          0     ] +
1*[      0          0          1          2          1     ] =

  [1*1+2*0+1*0  1*2+2*1+1*0  1*1+2*2+1*1  1*0+2*1+1*2  1*0+2*0+1*1] =
  [      1          4          6          4          1     ]
```

The weights in front of the filters multiply.

D Segmentation example

Figure 128 shows that different segmentation thresholds allow to observe completely different results.

E List of relevance metrics

This is an incomplete list of the metrics tested to assess the relevance of a found cluster [15] :

Direct metrics:

0. The size N (number of gridpoints) of the cluster
1. The mean value of the highest $\sqrt{N}$ values in the cluster, $topSqrt$
2. The L2 norm of all values in the cluster
3. The maximum value $maxVal$ of any gridpoint of the cluster
4. The $average$ value of all values in the cluster
5. The L1 norm of all values
6. The mean value of the highest up to 9 values in the cluster, $top9$
7. The mean value of the highest up to 5 values in the cluster, $top5$
8. The number of surface gridpoints of the cluster
9. The mean value of the highest 9 values in the cluster, adding 0 for smaller clusters
10. The mean value of the highest 5 values in the cluster, adding 0 for smaller clusters

---

[15] The numbering presented here matches the numbering in the source code.

11. The ratio of N to the number of surface gridpoints
12. The average distance between all gridpoints of the cluster
13. The maximum distance between two gridpoints of the cluster
14. The ratio of the maximum to the average distance

Linear combination meta-metrics:

21. $\sqrt{maxVal} \cdot top5 \cdot \sqrt{L2}$
22. $average \cdot \sqrt{L2}$
23. $topSqrt \cdot \sqrt{L2}$
24. $topSqrt \cdot \sqrt{L2} \cdot \sqrt{maxVal}$
25. $average \cdot \sqrt{L2} \cdot \sqrt{maxVal}$
26. $average \cdot \sqrt{L2} \cdot \sqrt{top5}$

More general meta metrics:

28. The sum of the two metrics that overfluctuated most
29. The sum of the three metrics that overfluctuated most
30. The sum of the half of the direct metrics that overfluctuated most
33. The sum of the overfluctuation of all direct metrics
34. The L2 norm of the overfluctuation of all direct metrics

F Conducting the multiscale search

This section explains how to actually perform the multiscale search. It assumes that the required prerequisites (pretrained RDFs, various scripts, ...) have been obtained from the svn. An example setup can be found in the file “multiscale-ExampleScripts.tar.gz”.

F.1 Filtering the data

The first step is extract the desired events. If a list of the events in equatorial coordinates is already available in the correct file format, this step may be skipped. The input for this step is required to be one large .i3 (or .i3.gz) file containing the whole data. Hint: .i3 files can be concatenated simply by “cat file1.i3 file2.i3 ... > filesAll.i3”. This doesn’t work for .i3.gz files without uncompressing them first.

For now we assume that all data runs to analyze are contained in a file called “allData.i3”. The script “applyAllCuts.py“ can then be used to filter the desired events by executing: ”./applyAllCuts.py -i allData.i3“. Depending on the size of the input file this step may take days. The script performs the following tasks:

- Remove all runs that have been used as burn sample (the run numbers are hardcoded).
- Remove known sparking runs (hardcoded run numbers).
- (For technical reasons) Remove events that occur multiple times.
- Keep only events from runs that fullfill all desired quality criteria (hardcoded run numbers).
- Use the pretrained up/down RDFs from the folder ./RDFs/ to get rid of events classified as downgoing.
- Use the pretrained Selectfit RDFs from the folder ./RDFs/ to apply Selectfit to each remaining event.
- Keeps only events that have an upgoing direction reconstruction as result of Selectfit.
- Keeps only events that have a Selectfit error estimate class less than 5.

The output is a file called allData_processed.i3.

F.2 Obtaining expectations

The next step is to extract the events from the .i3 file to ASCII files and to estimate the expected number of events as described in chapter 8.3.3. This is done by executing the command: "./doSignalFirstAnalysis.py –benchmark -1 -i allData_processed.i3“. This writes the files "candidateDump.txt“ and "candidateDumpRaw.txt“, and the file "normalizationDataStep1/normalizationResults_Size165016_bins181_nr1.nres2". The folder “normalizationDataStep1“ is assumed to be present, it is not created by the script. The file candidateDump.txt is a plain ASCII text file containing all events. Each line contains one event in the file, with the information stored in the following ordering, each entry separated by one space: Right ascension, declination, time of the event, energy of the event (currently not implemented), and two zero values (the same functions that read and write this information are used for other computations where these values have a meaning). The file candidateDumpRaw.txt contains the same events, but with azimuth and zenith instead of right ascension and declination, so in local instead of equatorial coordinates. The file normalizationResults_Size165016_bins181_nr1.nres2 contains the estimation how many events are expected at every location of the search sphere for all search scales, derived from the event sample and then smoothed by lowpass filters.

If an already available event sample is to be used where only equatorial coordinates are available, this step is a bit different. The file candidateDump.txt is required to be present in the same format as described above. The only relevant values are right ascension and declination, all other values are not required in the current version and can be zero values. The file candidateDumpRaw.txt also needs to be present and must follow the described format since it is read in and parsed, but it may contain arbitrary information, e.g. it may be a copy of candidateDump.txt. The command to invoke the expectation estimation in this case is: "./doSignalFirstAnalysis.py –benchmark -10 -i allData_processed.i3". The scrambling in the pseudo-experiments to obtain all estimates is then not performed with random times when converting from local to equatorial coordinates, but random right ascensions are used instead. The input of a valid .i3 file is still required because the script expects this input, not to obtain any events. The result is stored in the file "normalizationDataStep1/normalizationResults_Size165016_bins181_nr1.nres2" just as before. Inputs in galactic coordinates are not supported.

The difference in the number of lowpass filters required for IceCube data, explained in chapter 8.8.1, is to be changed in the source code before this step is executed. This adaption can be performed in the file "src/signal-first/private/signal-first/signalFirstInternal.cpp". To make this piece of code easier to find, the location for this task within the file is marked with the tag #CODETODETERMINETHENUMBEROFLOWPASS.

F.3 Conducting pseudo-experiments

To compute pseudo-experiments that are required for the determination of the renormalized relevances and the significances of clusters, the call is "./doSignalFirstAnalysis.py –benchmark -2 -i allData_processed.i3". The option -i allData_processed.i3 is only used for the first step, but again the current implementation complains if no valid i3 file is specified. By default the script is processing the whole i3 file now before doing the actual work. To save a large fraction of this time, the line "tray.Execute()" near the end of the script doSignalFirstAnalysis.py can now be replaced by "tray.Execute(10)". Of course this is optional. If you do this, don't forget to revert it to the original line before step 1 is executed again.

If no local coordinates are available (or they aren't to be used for whatever reason), the call changes to "./doSignalFirstAnalysis.py –benchmark -20 -i allData_processed.i3".

Both calls by default perform 20 pseudo-experiments. They store their results in the folder normalizationDataStep2 (the output fails if it doesn't exist). The files are called clusters_segment(ALPHA_FOR_SEGMENTATION)_(UNIXTIME).txt, so for example clusters_segment0.25019999999999964_1433713240.txt. Their format is one line per cluster, each line: Number of metrics, one value for each metric, number of gridpoints the cluster has, a list of the indices of all gridpoints of this cluster. Each value is separated by one space.

The segmentation thresholds to use can be changed near the beginning of file "src/signal-first/private/signal-first/signalFirstInternal.cpp" at the location marked with the tag #ALPHASEGMENTATIONSTOUSE. If a change of the used metrics is desired, this is also possible in file signalFirstInternal.cpp at the location tagged with #SETTHEMETRICSMASKHERE. Since all metrics are evaluated for every cluster in every evaluation, a mask is used to only keep the metrics that are desired. If the metric with index n (see list in Appendix E) is to be included in the evaluation, a line “metricsMask[n] = 1.0;” is to be added there.

Computing pseudo-experiments takes some time (about 15 minutes per pseudo-experiment with two segmentations on a Intel(R) Xeon(R) CPU X5650), but it can be parallelized easily in independent processes. To be able to perform various fits later on, a minimal number of pseudo-experiments is required. This has not been investigated in detail, but 800 experiments have been sufficient (to perform the analysis without error, not for statistical purposes).

The unixtime in the filename is used to distribute the output of many such processes computing simultaneously to independent result files in a simple way. With the computing power used for this thesis it only occurred in very rare cases that two processes wrote a result for the same computation during the same second and hence wrote uncoordinately to the same file. Such files can be identified easily, as they produce errors in the next step. If this has occurred, ALL file with this unixtime have to be deleted. If not ALL files with this unixtime in their filename are deleted, the next step will fail again, since it assumes, that the same number of pseudo-experiments are available for every segmentation and metric.

F.4 Renormalizing the relevances

As described in chapter 8.6, the renormalized relevances are not derived from even more pseudo-experiments, but instead are derived from the already available ones. This computation is done by calling “./doSignalFirstAnalysis.py –benchmark -4 -i allData_processed.i3”. There is no difference anymore between evaluations starting with local or equatorial coordinates. The resulting output is also stored in folder “normalizationDataStep2” in files called hist_Size(ALPHA_FOR_SEGMENTATION)_Metric(NUMBER_METRIC).txt, e.g. hist_Size0.25019999999999964_Metric4.txt. They contain the histograms of the relevance values for one segmentation and one metric. Additionally, the files hist_Size(ALPHA_FOR_SEGMENTATION)_Metric(NUMBER_METRIC)fit.txt are stored, containing the histograms with the tail substituted by a fit. With this information this step can also compute the distribution of renormalized relevance values used later to determine the significance. This distribution is stored in folder “normalizationDataStep3” in file “normalizationStep3_Maxima.txt”. For every pseudo-experiment it contains the maximum of the found renormalized relevance values. The folder has to exists already.

The savvy reader may have noticed that –benchmark -3 is missing. This step is deprecated. It can be used to greatly extend the number of evaluation metrics by the computation of various meta-metrics (Metrics based on other metrics). Since this didn’t lead to an improvement of the final result, it has been dropped.

F.5 Evaluating the data

When all prerequisites are computed, finally the actual evaluation of the data sample can be performed. This is done by calling “./doSignalFirstAnalysis.py

–benchmark -5 -i allData_processed.i3". It evaluates the events from the file "candidatesDump.txt", uses the estimated expectation of event numbers from folder "normalizationDataStep1" to compute Poisson probabilities, uses the distributions of relevance values from pseudo-experiments from folder "normalizationDataStep2" to renormalize the relevances of found clusters, and finally it uses the distribution of the renormalized relevances from folder "normalizationDataStep3" to determine a significance for a found cluster. The results are various outputs in the vtk file format for some of the intermediate steps of the algorithm and the results of the evaluation for visualization with Paraview. The main results are also output as skymaps in Hammer-Aitoff projection labeled hammerProj*.txt in equatorial and galactic coordinates. They can be visualized e.g. by the following commands in gnuplot:

- set view map
- set palette defined
- splot "hammerProjEquatorial.txt" w p pt 7 ps 1 palette notitle

Note: If the window size of Gunplot and the point size are too small, nothing is displayed. In that case, enlarge the window.

Most steps also output some .txt files containing debug information on various computations which can also be plotted with Gnuplot.

F.6 Additional steps for specific tasks

Other values for benchmark are possible when calling doSignalFirstAnalysis.py. –benchmark 0 computes an evaluation for a simulated source scenario (Note: the - in front of e.g. -5 has been to indicate a negative number). For (positive) values from 1 up to 25, one value of a benchmark is computed as shown in chapter 8.7. The shown scenarios in the sensitivity chapter are examples of these 25. A value of -6 performs post-processing on previously computed sensitivity result files. The computations performed for -7, -80 and -9 assume that a file "./stencilGrid.grid" is present, which indicates a template where to look for a cluster. -7 reads this template and evaluates the sky normally, but sets all clusters that do not overlap with the template by at least 51% to 0.0σ. -80 computes pseudo-experiments that also perform this check for overlap with the template, hence incorporating the effect of the template into the results of the pseudo-experiments. These functions were necessary for the evaluation of the IC40 dataset. -9 allows to distribute additional events within the template. The number has to be hard coded in the source code. This is how the checks for Figure 103 have been done.

G Sources tested during development

The most common test setup used during the development of the multiscale search contained three sources of 3.0° x 3.0° at a declination of -57° with 50° between them in right ascension with 6, 10 and 15 events, one extended source of 10° x 10° with 60 events at a declination of 0° and one point source, 0.5° x 0.5°, containing 6 events at -29° with a shift of 90° in right ascension with respect to the extended source.

Table 1 shows a list of the sources that have also been used to to test the method at various times during the development and therefore influenced the heuristic optimization that led to the choice of the metrics and segmentation thresholds. Apart from these, several variations of these setups have been tested, but they cannot all be listed here.

Source description	Extension	Declination	# Events
Point source	0.5° x 0.5°	-70°	0 - 20
Point source	0.5° x 0.5°	-29°	0 - 20
Point source	0.5° x 0.5°	+10°	0 - 20
Point source	1.0° x 1.0°	-70°	0 - 20
Point source	1.0° x 1.0°	-29°	0 - 20
Point source	1.0° x 1.0°	+10°	0 - 20
Extended source	5° x 5°	-70°	0 - 80
Extended source	5° x 5°	-29°	0 - 80
Extended source	5° x 5°	+10°	0 - 80
Extended source	10° x 10°	-70°	0 - 150
Extended source	10° x 10°	-29°	0 - 150
Extended source	10° x 10°	+10°	0 - 150
Extended source	30° x 30°	-29°	0 - 300
Extended source	50° x 50°	-29°	0 - 500
Diagonal line	20° x 0.5°	-70°	0 - 80
Diagonal line	20° x 0.5°	-29°	0 - 80
Diagonal line w. interruption	20° x 0.5°	-29°	0 - 80
Curved line	20° x 0.5°	-29°	0 - 80
Curved line w. interruption	20° x 0.5°	-29°	0 - 80
Three point sources	1.0° x 1.0°	-25°, -29°, -33°	0 - 20
Three point sources	1.0° x 1.0°	-66°, -70°, -74°	0 - 20
Three point sources	1.0° x 1.0°	-60°, -70°, -80°	0 - 20
Two extended sources	5.0° x 5.0°	-29° (20° apart in RA)	0 - 50
Two extended sources	5.0° x 5.0°	-70° (20° apart in RA)	0 - 50
Two touching diagonal lines	20° x 0.5°	-39°	0 - 50
Three extended sources	3° x 3°	-57° (50° apart in RA)	4 - 15

Table 1: Sources tested during development

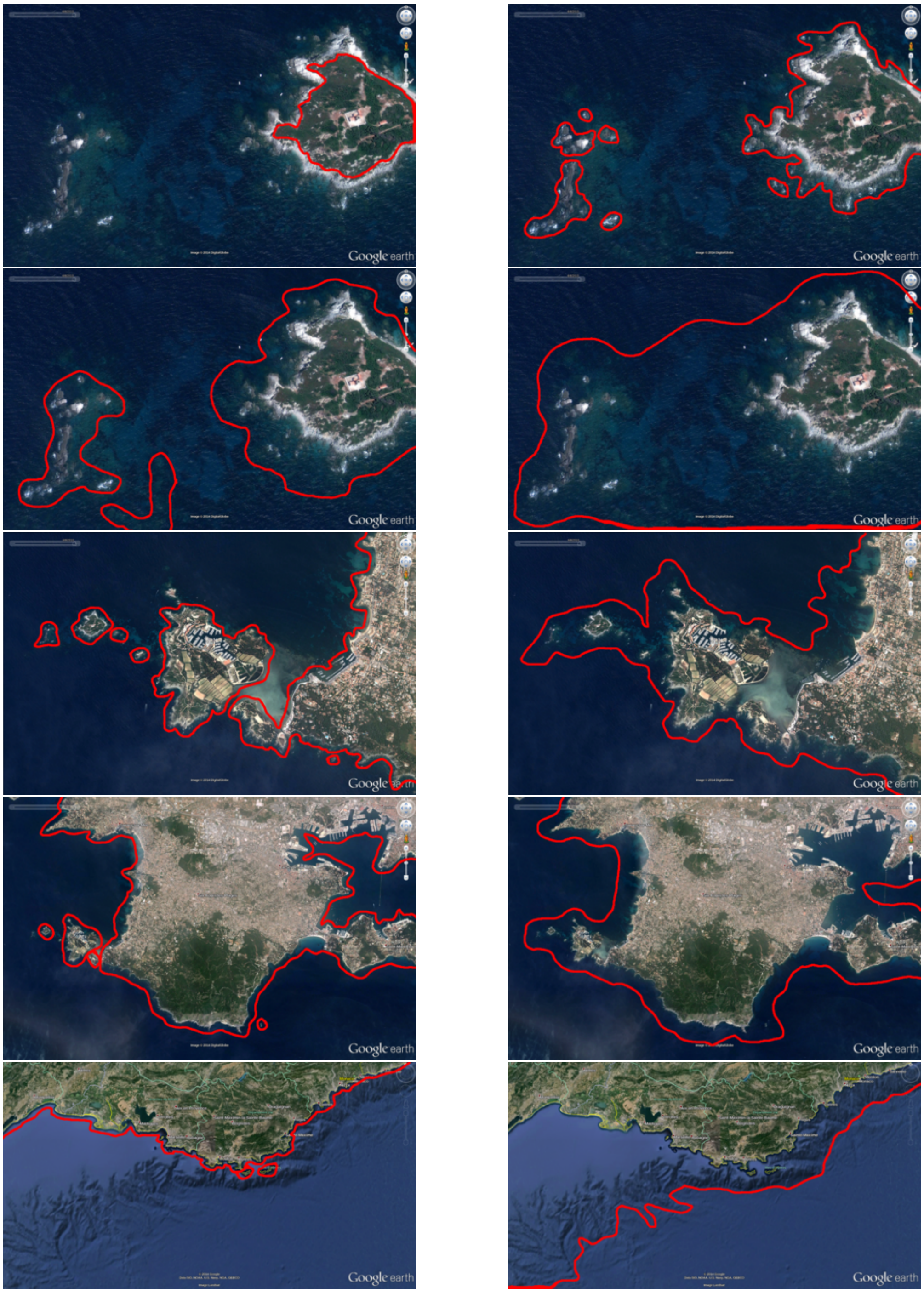


Figure 128: Real world example of different segmentation thresholds on an almost randomly selected region north-east of Spain and west of Italy. Different thresholds (here: water depths) reveal different underlying structures, which are all valid. It is therefore not absolutely defined what parts should be regarded to be connected, multiple thresholds can be used to find the most interesting structures. Image Credit: Google Earth, Google.

H Used runs

Here are the run numbers of all data runs used for the analysis:

25691 25700 25705 25706 25707 25713 25714 25715 25716 25717 25720 25721 25722 25723 25724 25725 25728 25730 25731
25800 25863 25864 25872 25880 25892 25893 25895 25905 25920 25921 25922 25923 25926 25927 25928 25930 25931 25932
25933 25934 25935 25937 25947 25953 25954 25955 25958 25964 25965 25972 25986 25988 25990 25991 26010 26030 26062
26087 26110 26131 26132 26133 26134 26135 26144 26160 26161 26165 26190 26227 26228 26230 26238 26265 26267 26270
26271 26272 26273 26274 26275 26276 26277 26281 26283 26331 26333 26334 26335 26340 26345 26355 26361 26363 26364
26395 26396 26397 26398 26405 26406 26407 26462 26463 26465 26469 26470 26471 26480 26483 26484 26485 26487 26528
26534 26535 26536 26537 26538 26540 26542 26543 26544 26545 26546 26592 26593 26594 26606 26633 26637 26638 26640
26647 26660 26676 26677 26684 26702 26703 26704 26710 26713 26714 26715 26716 26718 26720 26721 26723 26726 26770
26795 26797 26800 26801 26802 26803 26804 26805 26806 26807 26808 26810 26811 26812 26813 26814 26815 26816 26817
26818 26820 26821 26824 26825 26826 26827 26828 26829 26830 26831 26832 26833 26834 26864 26865 26866 26867 26868
26870 26871 26872 26873 26874 26875 26876 26877 26878 26891 26893 26894 26895 26896 26897 26898 26902 26903 26904
26905 26917 26918 26920 26936 26937 26938 26940 26956 26957 26958 26959 26963 26964 26965 26967 26968 26970 26971
26974 26978 26982 26983 26984 26987 26994 26995 27002 27003 27008 27009 27013 27015 27080 27081 27090 27091 27092
27094 27095 27100 27130 27143 27153 27162 27163 27164 27165 27170 27172 27173 27174 27175 27176 27177 27180 27182
27185 27216 27217 27218 27226 27227 27228 27230 27231 27232 27233 27234 27235 27244 27246 27248 27270 27271 27272
27290 27291 27292 27293 27294 27296 27297 27351 27352 27353 27379 27403 27404 27405 27406 27407 27408 27413 27424
27425 27426 27427 27440 27458 27460 27461 27462 27463 27464 27465 27467 27468 27470 27471 27501 27502 27504 27505
27550 27551 27558 27565 27566 27567 27569 27571 27572 27573 27574 27575 27576 27577 27578 27580 27581 27582 27616
27617 27618 27620 27622 27623 27624 27625 27627 27632 27633 27634 27635 27638 27640 27641 27642 27644 27646 27647
27648 27657 27658 27659 27660 27667 27668 27669 27721 27722 27723 27724 27726 27727 27728 27742 27743 27744 27745
27746 27748 27750 27751 27752 27753 27754 27755 27756 27757 27758 27760 27761 27762 27763 27764 27808 27812 27813
27814 27815 27822 27824 27825 27826 27828 27830 27831 27832 27853 27871 27872 27874 27875 27887 27890 27891 27892
27893 27894 27895 27896 27897 27898 27900 27941 27946 27952 27953 27954 27963 27979 27980 27981 27982 27984 27988
27990 27991 28007 28034 28035 28037 28044 28058 28061 28062 28063 28064 28065 28066 28067 28068 28070 28071 28106
28110 28114 28131 28132 28133 28143 28148 28161 28162 28163 28182 28211 28224 28225 28226 28233 28235 28236 28237
28238 28243 28244 28246 28247 28278 28280 28281 28285 28303 28304 28305 28306 28312 28315 28316 28317 28327 28334
28335 28336 28337 28340 28341 28342 28343 28344 28345 28346 28347 28348 28350 28351 28352 28354 28386 28387 28388
28395 28401 28402 28404 28421 28422 28432 28433 28434 28441 28442 28443 28444 28445 28446 28448 28450 28451 28452
28453 28454 28455 28456 28457 28458 28459 28460 28487 28494 28495 28496 28497 28503 28507 28508 28510 28512 28542
28552 28553 28554 28563 28566 28567 28568 28570 28583 28584 28585 28586 28613 28615 28616 28617 28618 28620 28621
28622 28623 28624 28625 28628 28630 28631 28632 28642 28643 28644 28645 28647 28648 28650 28651 28652 28653 28654
28655 28681 28686 28687 28688 28702 28711 28712 28713 28714 28715 28716 28717 28718 28720 28721 28722 28723 28724
28726 28727 28728 28730 28731 28732 28733 28734 28735 28736 28737 28738 28761 28762 28764 28765 28766 28767 28768
28770 28771 28772 28774 28775 28776 28777 28780 28782 28784 28785 28786 28787 28790 28791 28792 28793 28794 28795
28796 28797 28798 28801 28802 28803 28804 28832 28833 28834 28835 28836 28839 28841 28842 28845 28846 28847 28848
28850 28851 28852 28853 28854 28856 28857 28858 28860 28861 28862 28863 28864 28866 28867 28868 28870 28871 28872
28873 28874 28913 28948 28950 28952 28954 28955 28956 28957 28958 28966 28967 28968 28970 28971 28972 28973 28974
28975 28980 28981 28982 28983 28984 28985 28986 28987 28988 28990 28991 28992 28993 28994 28995 29022 29023 29024
29025 29026 29027 29028 29035 29036 29037 29038 29040 29041 29042 29043 29044 29045 29047 29048 29050 29051 29054
29055 29056 29057 29058 29060 29061 29062 29063 29064 29065 29066 29067 29068 29091 29092 29093 29094 29096 29097
29098 29100 29101 29103 29104 29105 29106 29107 29108 29110 29111 29112 29113 29114 29115 29116 29117 29118 29120
29121 29122 29123 29124 29125 29126 29127 29128 29152 29153 29154 29155 29156 29157 29158 29160 29161 29162 29165
29172 29173 29174 29175 29176 29177 29178 29180 29182 29183 29186 29187 29188 29190 29191 29192 29193 29194 29196
29197 29213 29214 29215 29216 29217 29218 29220 29221 29222 29223 29224 29246 29247 29250 29251 29252 29253 29254
29255 29256 29257 29258 29259 29261 29262 29263 29264 29265 29266 29267 29270 29271 29272 29273 29274 29275 29276
29277 29278 29279 29280 29281 29282 29283 29284 29285 29286 29287 29288 29292 29317 29318 29320 29321 29322 29323
29324 29325 29327 29328 29330 29335 29338 29341 29343 29350 29360 29361 29362 29363 29364 29365 29366 29367 29368
29370 29411 29413 29414 29417 29420 29421 29422 29423 29424 29426 29433 29434 29435 29436 29437 29438 29440 29441
29448 29472 29473 29474 29476 29477 29478 29480 29481 29483 29484 29485 29487 29488 29490 29500 29538 29540 29541
29542 29543 29547 29551 29552 29553 29554 29557 29558 29559 29561 29562 29563 29564 29566 29567 29580 29581 29582
29594 29595 29596 29599 29600 29604 29605 29606 29623 29624 29625 29626 29643 29646 29657 29658 29664 29678 29688
29693 29695 29697 29698 29716 29722 29728 29733 29740 29751 29752 29753 29754 29755 29756 29757 29758 29760 29761
29762 29763 29764 29788 29790 29791 29802 29803 29804 29805 29806 29807 29808 29810 29812 29813 29814 29815 29816
29817 29818 29820 29821 29822 29823 29825 29826 29827 29830 29831 29832 29833 29835 29836 29837 29838 29840 29867
29871 29872 29873 29874 29877 29891 29892 29895 29903 29904 29905 29906 29907 29908 29910 29911 29912 29913 29914
29915 29916 29917 29918 29920 29921 29922 29923 29924 29925 29926 29963 29968 29971 29978 29980 29981 29982 29993
30017 30018 30020 30021 30023 30024 30025 30026 30027 30028 30030 30031 30032 30033 30034 30035 30036 30060 30061
30062 30063 30064 30065 30066 30067 30068 30070 30071 30072 30073 30074 30075 30076 30077 30078 30080 30081 30082
30083 30084 30085 30086 30087 30088 30090 30091 30092 30093 30094 30095 30096 30097 30098 30100 30101 30102 30103
30104 30105 30106 30107 30108 30109 30110 30111 30112 30113 30114 30115 30116 30117 30118 30120 30140 30141 30142
30143 30144 30145 30146 30147 30148 30149 30150 30151 30152 30153 30154 30155 30156 30157 30158 30160 30161 30162
30163 30164 30165 30166 30167 30169 30170 30171 30172 30173 30174 30175 30176 30177 30203 30205 30206 30207 30208
30210 30211 30212 30213 30215 30216 30217 30218 30220 30221 30222 30223 30224 30225 30232 30233 30234 30235 30236
30237 30238 30240 30241 30242 30243 30244 30307 30308 30310 30311 30312 30313 30314 30315 30316 30317 30318 30319
30320 30321 30322 30324 30326 30327 30328 30332 30333 30335 30336 30337 30340 30341 30343 30344 30345 30346 30370
30371 30372 30373 30374 30375 30376 30377 30378 30380 30381 30384 30385 30386 30387 30388 30390 30391 30392 30393
30394 30395 30396 30397 30398 30401 30402 30407 30408 30416 30417 30422 30423 30424 30427 30428 30430 30431 30433
30434 30437 30438 30440 30441 30442 30450 30452 30453 30454 30455 30457 30459 30460 30508 30514 30521 30526 30533
30538 30541 30545 30551 30555 30557 30561 30563 30568 30571 30575 30577 30580 30583 30586 30590 30593 30598 30599
30601 30604 30606 30610 30614 30618 30623 30624 30626 30629 30633 30642 30654 30662 30670 30678 30706 30710 30711
30712 30713 30714 30715 30716 30717 30718 30719 30720 30721 30722 30723 30724 30725 30726 30727 30728 30729 30730
30784 30785 30787 30792 30793 30795 30799 30801 30802 30804 30806 30815 30816 30818 30821 30823 30825 30826 30829
30832 30841 30842 30855 30857 30861 30863 30864 30866 30867 30870 30872 30873 30876 30880 30881 30883 30885 30887
30889 30891 30892 30894 30896 30897 30942 30944 30946 30947 30950 30952 30953 30955 30956 30959 30961 30963 30964
30967 30969 30970 30972 30974 30977 30979 30981 30983 30986 30987 30989 30992 30993 30995 30997 30998 31000 31003
31005 31006 31008 31010 31011 31013 31016 31017 31020 31022 31024 31025 31027 31029 31030 31033 31035 31036 31038
31040 31042 31044 31046 31048 31049 31051 31053 31055 31057 31061 31062 31065 31067 31069 31070 31072 31074 31075
31078 31080 31082 31083 31085 31087 31091 31094 31096 31098 31099 31101 31104 31105 31107 31109 31111 31112 31115
31117 31118 31120 31122 31123 31125 31128 31130 31132 31134 31136 31137 31139 31141 31144 31145 31147 31149 31152
31154 31160 31164 31165 31167 31169 31170 31172 31176 31207 31209 31211 31212 31219 31221 31226 31228 31230 31232
31234 31235 31237 31247 31256 31261 31263 31264 31266 31269 31270 31275 31277 31280 31282 31283 31285 31287 31294
31312 31314 31316 31318 31355 31357 31359 31360 31362 31365 31368 31374 31381 31383 31385 31387 31388 31394 31399
31401 31404 31405 31407 31409 31415 31417 31424 31426 31427 31430 31432 31434 31435 31436 31438 31440 31443 31444
31446 31448 31449 31451 31454 31478 31479 31481 31483 31485 31490 31493 31505 31508 31509 31511 31513 31515 31516
31519 31520 31522 31523 31524 31527 31528 31530 31532 31534 31536 31540 31542 31543 31550 31552 31554 31581 31584
31586 31587 31589 31591 31592 31595 31612 31614 31615 31617 31619 31644 31646 31648 31650 31651 31654 31657 31662
31664 31675 31678 31680 31682 31683 31685 31687 31690 31695 31697 31698 31700 31702 31704 31706 31708 31709 31712
31714 31716 31718 31719 31721 31723 31724 31728 31729 31731 31732 31734 31736 31783 31785 31786 31788 31794 31797
31799 31803 31807 31811 31813 31823 31834 31835 31837 31839 31841 31849 31863 31871 31873 31875 31877 31879 31888
31893 31896 31897 31899 31901 31903 31905 31907 31911 31913 31915 31916 31918 31920 31958 31962 31963 31965 31966
31968 31970 31975 31979 31987 31988 31992 31999 32008 32009 32011 32013 32015 32021 32023 32028 32029 32031 32033
32034 32039 32049 32064 32077 32078 32080 32082 32083 32086 32088 32090 32091 32093 32095 32096 32098 32100 32102
32103 32105 32107 32109 32110 32150 32160 32165 32167 32168 32170 32173 32176 32177 32179 32181 32183 32185 32187
32189 32196 32233 32239 32241 32243 32244 32246 32248 32252 32258 32260 32262 32263 32265 32267 32269 32271 32273
32275 32276 32278 32281 32283 32284 32286 32288 32290 32292 32294 32338 32340 32343 32344 32346 32348 32350 32354
32356 32359 32360 32381 32387 32397 32399 32400 32402 32404 32411 32423 32430 32436 32437 32439 32441 32442 32445
32447 32450 32452 32454 32455 32458 32460 32462 32463 32468 32470 32472 32474 32475 32478 32479 32482 32485 32486
32488 32490 32491 32525 32529 32531 32533 32535 32570 32571 32576 32578 32580 32582 32584 32588 32592 32594 32597
32599 32600 32602 32604 32606 32617 32621 32622 32624 32626 32628 32629 32632 32634 32636 32637 32639 32641 32644
32645 32647 32649 32651 32652 32655 32657 32659 32700 32704 32706 32708 32710 32712 32717 32727 32740 32742 32743
32745 32747 32749 32751 32753 32755 32757 32758 32760 32762 32767 32769 32770 32771 32772 32773 32775 32778 32781
32784 32786 32788 32815 32817 32821 32823 32826 32828 32830 32833 32835 32848 32850 32852 32855 32858 32875 32880
32889 32895 32897 32900 32901 32904 32942 32954 32958 32961 32970 32978 32984 32986 32989 32990 32993 32995 32998
33002 33003 33005 33008 33010 33036 33038 33039 33041 33046 33047 33049 33053 33056 33058 33060 33062 33064 33067
33069 33071 33075 33076 33079 33081 33083 33094 33102 33110 33111 33113 33115 33117 33126 33131 33134 33161 33172
33173 33176 33178 33186 33194 33210 33215 33217 33219 33222 33223 33225 33227 33230 33231 33234 33236 33238 33239
33241 33243 33246 33247 33249 33251 33281 33284 33285 33288 33290 33292 33296 33300 33307 33308 33310 33312 33314
33316 33320 33321 33323 33325 33327 33328 33331 33338 33341 33343 33345 33346 33349 33351 33353 33355 33357 33359
33362 33364 33365 33367 33369 33371 33373 33375 33377 33379 33380 33382 33385 33387 33389 33390 33429 33431 33433
33435 33436 33439 33444 33447 33449 33451 33452 33454 33459 33461 33463 33464 33466 33468 33471 33472 33474 33476
33478 33479 33482 33484 33486 33487 33491 33494 33496 33499 33500 33502 33505 33507 33509 33511 33513 33515 33516
33518 33521 33523 33525 33527 33529 33530 33559 33563 33565 33567 33568 33570 33573 33584 33585 33587 33588 33590
33592 33595 33596 33598 33600 33611 33613 33615 33617 33618 33621 33623 33624 33626 33628 33630 33632 33634 33636
33638 33639 33641 33644 33647 33651 33653 33654 33656 33658 33660 33668 33695 33697 33698 33700 33702 33705 33706
33709 33711 33713 33714 33716 33718 33721 33722 33724 33726 33728 33729 33738 33746 33748 33749 33751 33754 33756
33757 33758 33763 33764 33766 33768 33770 33773 33776 33785 33798 33805 33806 33809 33810 33812 33814 33815 33817
33821 33823 33824 33826 33828 33832 33834 33868 33869 33871 33875 33878 33880 33883 33885 33886 33888 33889 33891
33896 33897 33899 33902 33903 33921 33924 33926 33927 33929 33932 33934 33935 33938 33940 33941 33943 33944 33946
33948 33950 33952 33954 33956 33958 33960 33961 33964 33965 33967 33969 33970 34007 34011 34012 34014 34016 34017
34019 34021 34026 34031 34033 34041 34045 34046 34048 34075 34080 34082 34083 34085 34087 34088 34091 34092 34094
34096 34097 34099 34100 34102 34106 34108 34109 34111 34113 34117 34118 34122 34124 34125 34127 34130 34131 34133
34140 34142 34146 34149 34151 34152 34212 34217 34230 34242 34243 34245 34246 34248 34250 34252 34254 34255 34268
34270 34366 34368 34370 34371 34374 34375 34378 34383 34385 34386 34388 34389 34394 34395 34396 34397 34398 34399
34400 34401 34403 34405 34407 34414 34416 34417 34419 34420 34422 34423 34425 34426 34427 34432 34433 34439 34443
34444 34445 34448 34451 34452 34454 34455 34457 34458 34459 34462 34465 34466 34467 34469 34470 34473 34474 34481
34486 34487 34489 34490 34491 34493 34494 34496 34497 34498 34502 34521 34522 34549 34550 34551 34552 34553 34554
34560 34561 34562 34563 34564 34565 34566 34571 34572 34573 34574 34575 34576 34577 34579 34580 34583 34588 34598
34600 34601 34603 34604 34605 34608 34609 34610 34616 34621 34622 34625 34626 34628 34630 34632 34633 34635 34636
34639 34640 34642 34643 34644 34646 34647 34649 34650 34656 34658 34660 34661 34669 34670 34672 34673 34675 34678
34679 34685 34687 34688 34692 34695 34703 34708 34711 34712 34718 34719 34721 34723 34724 34726 34734 34740 34742
34744 34745 34747 34748 34754 34756 34759 34761 34771 34773 34775 34780 34781 34783 34785 34786 34789 34790 34792
34793 34796 34798 34800 34801 34806 34808 34809 34813 34819 34821 34832 34834 34835 34841 34843 34845 34846 34848
34851 34853 34856 34882 34883 34893 34896 34898 34899 34906 34907 34909 34912 34913 34915 34917 34919 34921 34927
34928 34930 34931 34933 34935 34937 34938 34941 34942 34944 34945 34947 34953 34955 34957 34959 34960 34961 34963
34964 34968 34969 34971 34974 34976 34978 34980 34981 34983 34984 34986 34988 34990 34996 34998 34999 35000 35002
35003 35006 35007 35008 35011 35016 35019 35021 35022 35026 35028 35030 35031 35033 35037 35038 35040 35041 35044
35049 35052 35053 35054 35056 35057 35058 35060 35061 35063 35064 35068 35069 35071 35073 35074 35151 35153 35155
35158 35159 35160 35161 35162 35163 35164 35166 35168 35170 35172 35173 35175 35176 35178 35179 35180 35190 35193
35195 35196 35198 35199 35200 35201 35202 35203 35236 35238 35240 35242 35243 35244 35246 35247 35248 35250 35253
35261 35264 35277 35285 35288 35289 35291 35292 35293 35294 35296 35300 35304 35306 35308 35309 35311 35312 35314
35316 35318 35319 35321 35322 35325 35327 35328 35330 35332 35333 35335 35336 35337 35339 35340 35342 35343 35346
35348 35349 35351 35353 35354 35355 35357 35359 35360 35362 35368 35370 35372 35374 35375 35377 35413 35417 35419
35420 35421 35422 35423 35424 35425 35444 35463 35465 35470 35472 35473 35475 35476 35477 35479 35480 35483 35484
35485 35488 35490 35491 35493 35495 35496 35498 35501 35503 35504 35506 35511 35512 35513 35515 35517 35519 35520
35521 35523 35529 35530 35532 35534 35535 35539 35541 35543 35545 35548 35550 35551 35553 35555 35561 35562 35563
35565 35571 35573 35575 35577 35579 35580 35627 35635 35636 35639 35640 35651 35661 35662 35663 35665 35666 35672
35674 35678 35681 35683 35685 35686 35689 35691 35696 35698 35700 35702 35704 35706 35707 35709 35710 35712 35716

35717 35719 35721 35723 35724 35726 35728 35730 35732 35734 35735 35746 35748 35750 35751 35753 35754 35781 35783
35785 35788 35790 35793 35795 35797 35798 35799 35800 35802 35838 35843 35846 35849 35851 35852 35853 35855 35856
35858 35863 35864 35872 35873 35875 35876 35878 35880 35881 35884 35885 35886 35887 35889 35890 35891 35892 35894
35896 35897 35899 35900 35901 35902 35904 35905 35906 35907 35908 35910 35911 35912 35914 35916 35917 35919 35921
35923 35924 35926 35927 35929 35930 35932 35933 35935 35936 35937 35939 35941 35942 35944 35948 35953 35954 35956
35958 35959 35962 35963 35964 35966 35968 35969 35971 35972 35975 35976 35977 35979 35981 35983 35984 35985 35987
35989 35991 35992 35993 35996 35997 35999 36000 36002 36029 36070 36071 36072 36075 36076 36079 36088 36091 36093
36095 36097 36099 36101 36103 36104 36107 36108 36110 36112 36113 36119 36123 36124 36128 36139 36149 36151 36153
36155 36156 36157 36159 36160 36161 36163 36165 36215 36218 36220 36221 36234 36238 36244 36254 36255 36256 36258
36259 36260 36262 36263 36264 36267 36268 36269 36271 36272 36273 36275 36276 36277 36279 36280 36281 36283 36284
36285 36287 36288 36289 36291 36294 36297 36298 36300 36302 36303 36305 36306 36309 36315 36352 36359 36360 36362
36364 36366 36367 36369 36371 36372 36373 36375 36380 36385 36386 36389 36392 36394 36396 36399 36401 36403 36406
36409 36411 36413 36417 36418 36420 36422 36424 36426 36428 36430 36431 36434 36435 36436 36438 36439 36442 36443
36444 36446 36448 36450 36452 36453 36456 36458 36459 36461 36466 36467 36469 36471 36473 36474 36475 36477 36478
36480 36481 36483 36484 36485 36488 36489 36491 36492 36494 36495 36497 36498 36500 36501 36502 36505 36506 36507
36509 36511 36512 36514 36516 36517 36519 36520 36532 36534 36543 36548 36555 36556 36557 36561 36570 36572 36576
36577 36582 36587 36590 36592 36593 36594 36596 36598 36599 36623 36734 36736 36738 36745 36747 36749 36750 36752
36755 36759 36761 36763 36772 36775 36776 36783 36785 36786 36787 36789 36791 36809 36813 36814 36815 36816 36817
36818 36825 36826 36832 36835 36836 36837 36838 36840 36850 36853 36858 36860 36862 36863 36865 36867 36869 36870
36872 36874 36877 36879 36881 36883 36885 36887 36889 36891 36892 36893 36895 36897 36898 36900 36901 36903 36904
36906 36907 36908 36911 36914 36915 36918 36922 36923 36925 36926 36927 36929 36930 36932 36934 36935 36938 36939
36940 36942 36943 36945 36947 36948 36950 36951 36953 36956 36957 36958 36960 36961 36963 36964 36966 36968 36970
36973 36975 36978 37015 37016 37018 37019 37021 37024 37025 37026 37028 37029 37031 37033 37035 37037 37044 37045
37047 37049 37051 37053 37055 37057 37061 37064 37065 37068 37070 37072 37073 37075 37076 37078 37081 37082 37084
37085 37090 37091 37093 37095 37096 37098 37099 37103 37104 37105 37107 37108 37109 37111 37112 37114 37115 37118
37119 37121 37123 37125 37128 37130 37131 37132 37134 37137 37139 37141 37142 37144 37146 37147 37149 37150 37151
37153 37155 37158 37159 37161 37163 37165 37167 37169 37171 37176 37209 37211 37213 37215 37217 37218 37220 37222
37224 37226 37228 37230 37231 37234 37236 37238 37239 37241 37242 37246 37248 37249 37251 37252 37253 37255 37257
37259 37261 37263 37266 37268 37271 37272 37276 37280 37281 37283 37284 37285 37287 37289 37291 37292 37293 37296
37297 37298 37300 37301 37302 37304 37305 37307 37308 37309 37312 37313 37314 37316 37317 37319 37320 37321 37323
37324 37327 37328 37330 37331 37332 37334 37335 37337 37339 37340 37342 37344 37346 37348 37349 37350 37352 37353
37389 37391 37393 37395 37397 37399 37402 37404 37407 37409 37412 37415 37417 37419 37420 37423 37424 37426 37428
37430 37432 37434 37436 37437 37439 37440 37442 37443 37445 37447 37448 37450 37452 37454 37455 37456 37458 37459
37461 37462 37463 37465 37466 37468 37470 37471 37473 37601 37603 37605 37607 37608 37616 37618 37619 37620 37622
37623 37625 37631 37632 37634 37635 37637 37638 37644 37650 37651 37653 37655 37656 37658 37659 37662 37664 37665
37668 37670 37672 37673 37675 37678 37679 37682 37683 37685 37687 37688 37689 37701 37712 37714 37716 37717 37719
37720 37722 37723 37731 37732 37733 37735 37736 37738 37739 37741 37744 37746 37751 37752 37753 37755 37757 37759
37761 37763 37764 37765 37767 37768 37770 37771 37772 37774 37776 37780 37782 37783 37785 37788 37789 37871 37873
37874 37876 37990 37992 37993 37994 37996 37997 37999 38002 38004 38007 38009 38011 38045 38052 38054 38056 38058
38060 38063 38065 38067 38070 38072 38074 38075 38076 38078 38080 38082 38084 38086 38088 38090 38093 38095 38097
38099 38100 38102 38104 38106 38108 38110 38112 38113 38115 38117 38118 38119 38121 38122 38123 38125 38126 38128
38129 38153 38155 38156 38157 38158 38159 38161 38168 38170 38172 38175 38176 38177 38179 38180 38182 38183 38185
38187 38188 38190 38193 38195 38196 38197 38199 38200 38202 38204 38205 38207 38208 38209 38211 38212 38213 38215
38216 38218 38223 38225 38227 38228 38230 38233 38234 38241 38242 38244 38245 38249 38252 38264 38285 38290 38292
38293 38294 38391 38397 38398 38400 38402 38403 38404 38406 38407 38410 38421 38423 38424 38426 38427 38429 38431
38432 38435 38438 38440 38442 38443 38445 38447 38448 38450 38451 38461 38463 38465 38466 38467 38470 38472 38474
38475 38477 38478 38480 38481 38485 38486 38488 38489 38491 38492 38493 38495 38496 38500 38502 38503 38504 38506
38511 38512 38515 38517 38518 38519 38520 38521 38523 38524 38526 38527 38529 38533 38534 38536 38537 38539 38540
38542 38543 38546 38548 38549 38551 38552 38554 38555 38557 38558 38560 38562 38569 38571 38572 38574 38575 38577
38578 38581 38582 38583 38585 38586 38588 38589 38590 38592 38593 38596 38597 38599 38600 38601 38603 38604 38606
38607 38610 38611 38612 38614 38615 38617 38618 38620 38621 38623 38625 38626 38628 38629 38631 38632 38634 38635
38636 38639 38640 38641 38643 38645 38646 38647 38649 38650 38652 38654 38656 38658 38659 38664 38685 38687 38688
38690 38691 38692 38695 38696 38697 38699 38700 38702 38703 38711 38712 38716 38717 38719 38720 38722 38723 38724
38726 38727 38730 38731 38733 38734 38736 38738 38739 38741 38742 38744 38745 38747 38748 38750 38752 38753 38755
38756 38757 38760 38762 38765 38766 38770 38771 38773 38774 38780 38783 38784 38787 38788 38790 38791 38793 38795
38796 38799 38800 38801 38803 38805 38806 38807 38809 38810 38813 38814 38823 38824 38826 38827 38829 38830 38833
38834 38835 38837 38838 38840 38841 38843 38844 38847 38849 38850 38852 38853 38855 38856 38869 38899 38900 38902
38903 38905 38906 38908 38909 38913 38917 38918 38925 38926 38928 38929 38931 38932 38934 38936 38938 38940 38941
38943 38944 38946 38947 38949 38951 38953 38954 38955 38957 38958 38960 38961 38962 38965 38966 38968 38969 38971
38972 38974 38975 38977 38978 38980 38982 38983 38985 38986 38988 38989 38990 38992 38994 38996 38997 38999 39000
39001 39003 39004 39006 39008 39015 39021 39022 39025 39028 39030 39031 39033 39034 39036 39037 39038 39040 39042
39044 39045 39047 39048 39049 39051 39052 39054 39056 39057 39059 39060 39062 39063 39065 39066 39067 39069 39071
39072 39074 39076 39078 39079 39080 39082 39083 39085 39086 39087 39090 39091 39095 39099 39101 39102 39104 39105
39106 39108 39109 39111 39112 39114 39116 39117 39119 39120 39122 39124 39125 39127 39128 39131 39133 39134 39144
39147 39150 39151 39153 39154 39155 39157 39158 39159 39161 39168 39169 39180 39181 39182 39184 39196 39197 39199
39204 39205 39207 39209 39211 39212 39214 39215 39217 39218 39220 39221 39223 39225 39227 39228 39229 39231 39232
39234 39235 39237 39238 39240 39242 39244 39246 39247 39249 39250 39254 39255 39257 39259 39260 39262 39263 39265
39266 39273 39307 39313 39319 39321 39324 39327 39331 39339 39341 39342 39345 39349 39351 39357 39360 39362 39364
39365 39369 39372 39373 39377 39379 39381 39382 39384 39387 39389 39391 39393 39395 39398 39400 39402 39410 39420
39425 39427 39430 39432 39438 39442 39453 39456 39458 39460 39462 39564 39565 39567 39569 39570 39573 39575 39580
39581 39583 39586 39589 39590 39603 39609 39617 39620 39623 39625 39627 39629 39631 39633 39636 39638 39640 39642
39644 39647 39649 39652 39662 39669 39671 39673 39675 39677 39680 39682 39683 39685 39689 39700 39702 39705 39707
39711 39712 39713 39717 39718 39723 39725 39726 39727 39728 39729 39757 39760 39762 39763 39765 39772 39774 39775
40098 40111 40119 40150 40152 40154 40155 40157 40160 40161 40167 40169 40171 40177 40179 40189 40191 40195 40197
40199 40203 40205 40206 40213 40216 40218 40453 40473 40475 40478 40480 40483 40507 40509 40516 40530 40545 40553
40555 40591 40594 40603 40606 40608 40616 40618 40623 40647 40649 40653 40655 40657 40659 40660 40663 40664 40666
40668 40669 40672 40675 40676 40677 40678 40681 40682 40684 40686 40688 40690 40691 40695 40697 40698 40705 40706
40707 40710 40712 40713 40714 40716 40717 40718 40719 40721 40722 40741 40751 40753 40756 40758 40762 40763 40765
40766 40768 40770 40772 40777 40780 40783 40784 40786 40788 40790 40867 40870 40872 40879 40880 40884 40908 40910
40917 40957 40959 40962 41005 41089 41092 41094 41096 41192 41194 41203 41204 41206 41208 41209 41212 41214 41215
41217 41222 41224 41226 41228 41231 41233 41234 41247 41249 41251 41252 41254 41257 41259 41261 41262 41265 41268
41270 41279 41281 41283 41285 41287 41289 41292 41295 41298 41301 41303 41305 41308 41310 41313 41314 41316 41318
41320 41321 41350 41352 41353 41355 41361 41364 41365 41367 41369 41370 41373 41374 41376 41377 41381 41384 41385
41388 41390 41391 41393 41395 41397 41400 41402 41403 41404 41406 41408 41410 41411 41416 41417 41418 41421 41423
41424 41426 41428 41429 41430 41431 41433 41434 41435 41437 41439 41440 41441 41442 41444 41445 41446 41449 41450
41452 41453 41481 41482 41524 41528 41530 41533 41534 41536 41538 41540 41543 41546 41548 41550 41552 41555 41557
41560 41563 41567 41569 41572 41578 41580 41582 41584 41586 41591 41597 41605 41608 41610 41612 41614 41626 41628
41630 41631 41632 41633 41634 41635 41661 41662 41665 41666 41673 41675 41677 41679 41684 41687 41690 41691 41692
41694 41696 41697 41699 41702 41703 41706 41711 41712 41714 41715 41717 41719 41722 41724 41727 41731 41737 41740
41742 41744 41769 41780 41781 41784 41787 41808 41810 41812 41815 41830 41835 41836 41839 41842 41844 41845 41847
41850 41852 41854 41913 41937 41939 41940 41971 41973 41975 41977 41980 41982 41984 41986 41988 41991 41993 41995
41996 41998 42001 42003 42005 42020 42032 42035 42038 42042 42045 42047 42050 42053 42056 42059 42062 42064 42066
42069 42072 42074 42076 42082 42086 42088 42090 42091 42096 42099 42102 42109 42111 42114 42116 42117 42119 42121
42129 42132 42174 42178 42181 42185 42187 42190 42208 42213 42215 42218 42222 42224 42226 42230 42233 42242 42244
42248 42250 42252 42254 42259 42260 42261 42264 42266 42267 42271 42272 42274 42277 42278 42279 42280 42281 42283
42285 42286 42288 42290 42292 42296 42298 42300 42303 42305 42309 42311 42314 42316 42318 42320 42323 42326 42329
42334 42337 42341 42344 42348 42349 42350 42353 42355 42358 42359 42362 42365 42369 42373 42376 42379 42381 42407
42408 42412 42415 42417 42420 42423 42425 42432 42433 42438 42445 42451 42453 42455 42456 42458 42460 42462 42464
42467 42469 42473 42475 42477 42482 42494 42497 42500 42502 42504 42515 42518 42521 42523 42526 42528 42530 42533
42536 42541 42543 42545 42547 42550 42554 42577 42580 42585 42588 42592 42594 42597 42602 42606 42609 42612 42616
42623 42625 42628 42631 42634 42638 42640 42643 42646 42649 42652 42655 42658 42661 42668 42671 42676 42686 42689
42756 42758 42765 42771 42777 42781 42785 42792 42796 42799 42804 42810 42811 42813 42815 42819 42821 42825 42829
42831 42835 42838 42842 42851 42856 42860 42863 42866 42889 42893 42897 42900 42906 42909 42912 42922 42926 42932
42936 42941 42944 42947 42950 42954 42959 42962 42964 42967 42971 42973 42976 42983 42986 42990 42994 42997 43003
43007 43010 43015 43019 43022 43026 43031 43034 43040 43046 43050 43059 43062 43066 43069 43073 43078 43081 43086
43091 43098 43101 43105 43108 43112 43115 43140 43144 43148 43158 43161 43164 43169 43173 43176 43180 43183 43185
43191 43199 43219 43222 43226 43229 43234 43237 43240 43242 43245 43250 43253 43256 43258 43262 43265 43269 43271
43274 43282 43321 43325 43328 43353 43359 43362 43366 43372 43376 43379 43385 43390 43442 43450 43462 43466 43476
43480 43487 43491 43495 43516 43521 43525 43540 43546 43551 43555 43563 43568 43574 43584 43587 43591 43596 43600
43606 43610 43612 43615 43620 43624 43627 43629 43632 43634 43639 43646 43649 43653 43657 43663 43666 43669 43673
43678 43681 43689 43692 43695 43698 43701 43705 43710 43713 43716 43719 43722 43725 43728 43731 43734 43737 43741
43743 43745 43748 43751 43758 43760 43764 43769 43773 43776 43778 43782 43787 43796 43800 43806 43812 43817 43823
43826 43831 43834 43837 43842 43846 43849 43852 43855 43860 43866 43869 43872 43875 43878 43882 43885 43888 43890
43893 43899 43902 43905 43911 43922 43925 43930 43932 43936 43940 43959 43964 43969 43974 43978 43981 43984 43987
43991 43999 44003 44007 44010 44013 44016 44018 44020 44025 44029 44037 44040 44042 44045 44049 44053 44056 44060
44064 44072 44075 44078 44081 44083 44087 44089 44098 44105 44107 44108 44109 44112 44115 44149 44152 44165 44169
44174 44182 44184 44188 44195 44199 44211 44214 44218 44225 44228 44266 44271 44274 44277 44281 44285 44289 44293
44296 44298 44304 44307 44311 44315 44322 44328 44334 44338 44340 44344 44348 44353 44356 44358 44361 44364 44367
44372 44375 44378 44383 44385 44395 44400 44407 44412 44419 44421 44423 44426 44428 44431 44440 44473 44477 44485
44487 44491 44493 44497 44499 44502 44513 44529 44532 44535 44539 44544 44556 44560 44564 44566 44568 44571 44574
44577 44580 44583 44585 44588 44596 44598 44602 44605 44608 44611 44614 44617 44620 44623 44633 44634 44637 44640
44644 44648 44653 44656 44658 44662 44664 44666 44670 44782 44787 44790 44792 44824 44833 44835 44838 44840 44842
44844 44855 44860 44864 44866 44869 44871 44883 44885 44894 44897 44963 44968 44971 44978 44985 44989 44992 44994
45000 45006 45008 45012 45015 45018 45021 45024 45030 45034 45037 45040 45044 45046 45050 45054 45058 45061 45065
45069 45071 45098 45102 45105 45108 45110 45112 45116 45118 45123 45126 45128 45130 45132 45134 45137 45138 45140
45142 45145 45147 45150 45165 45168 45169 45172 45177 45179 45181 45183 45185 45188 45191 45193 45196 45197 45199
45202 45210 45212 45214 45218 45222 45225 45227 45229 45232 45234 45238 45243 45246 45248 45250 45252 45254 45256
45259 45261 45263 45265 45267 45269 45271 45274 45275 45277 45279 45281 45283 45285 45287 45290 45291 45294 45295
45326 45330 45332 45334 45338 45342 45349 45352 45354 45356 45358 45361 45363 45367 45371 45373 45374 45377 45379
45381 45383 45385 45388 45389 45391 45393 45395 45397 45399 45400 45403 45405 45408 45410 45412 45414 45416 45418
45419 45422 45424 45426 45428 45430 45431 45433 45435 45438 45440 45442 45444 45446 45448 45450 45466 45468 45470
45472 45474 45477 45479 45481 45484 45485 45487 45489 45491 45493 45495 45497 45500 45503 45504 45506 45509 45511
45513 45516 45518 45521 45523 45525 45527 45529 45531 45533 45536 45538 45540 45542 45544 45546 45548 45551 45553
45555 45557 45559 45561 45563 45565 45567 45569 45571 45573 45575 45577 45579 45581 45583 45586 45589 45592 45595
45598 45601 45607 45616 45617 45620 45626 45628 45630 45632 45639 45641 45648 45650 45663 45665 45668 45670 45681
45683 45685 45687 45689 45692 45695 45697 45700 45701 45703 45705 45707 45710 45714 45717 45719 45721 45724 45726
45728 45730 45746 45749 45750 45753 45755 45757 45759 45761 45766 45768 45771 45773 45775 45776 45778 45780 45784
45788 45790 45795 45797 45800 45802 45804 45807 45812 45814 45816 45818 45820 45822 45825 45827 45829 45831 45833
45835 45838 45841 45843 45845 45848 45851 45853 45855 45858 45860 45862 45864 45867 45869 45871 45873 45875 45877
45879 45881 45884 45905 45909 45912 45915 45918 45920 45924 45931 45933 45936 45938 45940 45942 45944 45946 45952
45956 45958 45960 45962 45964 45966 45971 45976 45978 45981 45983 45985 45987 45991 45994 45996 45998 46001 46003
46006 46009 46011 46014 46016 46018 46020 46022 46026 46027 46030 46033 46036 46039 46040 46043 46045 46048 46050
46052 46054 46092 46094 46097 46099 46101 46102 46104 46106 46109 46112 46118 46120 46122 46124 46126 46127 46129
46130 46133 46135 46137 46139 46140 46141 46143 46145 46147 46150 46153 46159 46193 46195 46200 46202 46207 46211
46225 46228 46310 46324 46341 46353 46356 46360 46362 46367 46369 46383 46388 46389 46392 46394 46396 46399 46401
46404 46406 46407 46409 46411 46413 46415 46418 46420 46422 46423 46425 46428 46430 46432 46434 46438 46442 46449
46451 46453 46455 46457 46460 46462 46464 46468 46469 46472 46474 46475 46478 46480 46482 46485 46487 46489 46492
46494 46498 46500 46502 46505 46507 46510 46512 46514 46516 46518 46521 46523 46525 46527 46529 46531 46533 46536
46538 46540 46542 46544 46546 46549 46551 46554 46556 46558 46559 46561 46563 46564 46567 46570 46572 46574 46576
46584 46586 46591 46593 46595 46597 46599 46601 46602 46604 46605 46607 46610 46611 46613 46615 46616 46618 46624
46626 46633 46635 46638 46640 46650 46653 46665 46667 46670 46673 46676 46684 46687 46689 46692 46694 46696 46698
46700 46702 46704 46707 46710 46712 46714 46716 46718 46721 46724 46726 46729 46731 46733 46736 46745 46747 46750
46752 46754 46756 46763 46765 46768 46770 46772 46774 46777 46807 46820 46823 46825 46827 46831 46834 46838 46852
46855 46858 46861 46863 46865 46867 46869 46872 46874 46876 46878 46879 46881 46883 46885 46887 46890 46892 46893
46896 46898 46899 46901 46903 46905 46908 46909 46911 46913 46916 46918 46921 46924 46949 46953 46955 46957 46959
46962 46964 46966 46968 46970 46972 46974 47012 47016 47025 47039 47042 47044 47053 47058 47061 47076 47082 47120
47210 47214 47218 47223 47228 47235 47236 47238 47241 47244 47246 47247 47534 47536 47574 47575 47577 47591 47594

47596 47599 47601 47617 47620 47623 47629 47634 47662 47665 47667 47670 47672 47674 47676 47679 47682 47685 47686
47688 47692 47694 47706 47708 47710 47712 47713 47716 47718 47719 47730 47733 47735 47784 47785 47788 47792 47794
47795 47797 47799 47801 47830 47833 47835 47836 47838 47840 47843 47846 47859 47861 47862 47865 47866 47868 47870
47873 47874 47876 47878 47879 47882 47883 47885 47891 47893 47895 47897 47898 47933 47943 47947 47949 47951 47952
47969 47973 47987 47990 47992 47994 48092 48094 48095 48096 48097 48098 48100 48112 48115 48117 48118 48122 48124
48125 48127 48128 48129 48131 48133 48135 48138 48140 48141 48143 48144 48145 48146 48148 48149 48151 48152 48154
48156 48157 48158 48160 48161 48163 48164 48165 48167 48168 48169 48172 48173 48175 48176 48178 48179 48181 48182
48184 48186 48188 48190 48198 48199 48200 48202 48203 48209 48211 48213 48216 48228 48230 48231 48276 48280 48366
48368 48370 48417 48420 48422 48425 48426 48429 48454 48461 48463 48465 48466 48468 48471 48472 48474 48476 48478
48479 48480 48482 48484 48487 48488 48490 48492 48494 48495 48512 48514 48515 48517 48518 48520 48521 48522 48524
48525 48527 48532 48534 48536 48756 48904 48906 48908 48909 48910 48915 48916 48928 48929 48931 48932 48934 48937
48939 48940 48941 48943 48945 48947 48950 48953 48956 48958 48959 48961 48964 48966 48968 48970 49223 49233 49236
49249 49251 49272 49273 49275 49278 49281 49288 49290 49333 49336 49339 49341 49344 49346 49350 49358 49360 49362
49366 49369 49372 49375 49388 49393 49395 49396 49398 49401 49403 49407 49409 49411 49413 49415 49417 49420 49422
49425 49427 49429 49431 49441 49447 49450 49452 49455 49457 49464 49474 49478 49558 49586 49588 49591 49593 49598
49601 49604 49606 49608 49609 49619 49621 49624 49627 49631 49633 49634 49636 49640 49643 49658 49660 49661 49663
49667 49669 49670 49672 49674 49677 49683 49685 49689 49690 49693 49696 49699 49701 49703 49706 49708 49710 49712
49715 49716 49722 49725 49728 49731 49733 49737 49740 49741 49743 49745 49749 49755 49757 49759 49764 49770 49773
49775 49777 49780 49783 49787 49790 49792 49795 49803 49805 49826 49831 49833 49850 49853 49858 49860 49863 49866
49870 49872 49897 49900 49901 49902 49906 49908 49911 49914 49932 49942 49945 49948 49951 49954 49958 49959 49980
49983 49986 49989 49991 49995 50002 50012 50016 50018 50021 50023 50026 50028 50030 50032 50035 50037 50038 50041
50043 50046 50049 50053 50055 50057 50061 50063 50064 50068 50069 50071 50072 50073 50075 50078 50080 50082 50086
50088 50115 50159 50162 50164 50166 50175 50198 50204 50207 50211 50216 50218 50221 50225 50227 50230 50232 50236
50241 50244 50339 50340 50344 50346 50348 50351 50355 50359 50361 50365 50377 50379 50383 50386 50388 50390 50391
50392 50394 50396 50397 50398 50400 50402 50405 50407 50411 50416 50419 50422 50426 50429 50434 50692 50695 50698
50700 50709 50724 50738 50740 50753 50756 50767 50771 50776 50779 50782 50785 50789 50792 50796 50800 50802 50821
50824 50826 50829 50832 50834 50837 50839 50843 50845 50848 50850 50854 50859 50861 50863 50867 50869 50872 50874
50877 50878 50880 50882 50884 50887 50889 50891 50897 50901 50914 50937 50941 50943 50947 50949 50951 50955 50958
50960 50964 50967 50969 50972 50976 50983 50986 50990 50993 50995 50999 51002 51003 51005 51007 51009 51011 51013
51014 51017 51022 51024 51025 51033 51037 51040 51050 51057 51061 51063 51067 51070 51072 51077 51079 51083 51085
51088 51092 51094 51097 51101 51104 51107 51110 51113 51116 51118 51122 51124 51127 51130 51132 51133 51136 51138
51140 51144 51146 51149 51152 51154 51156 51158 51160 51162 51163 51165 51166 51168 51169 51172 51174 51175 51177
51179 51199 51207 51216 51222 51223 51225 51226 51228 51230 51232 51234 51235 51237 51238 51241 51242 51245 51246
51251 51252 51254 51256 51257 51260 51261 51263 51264 51266 51268 51271 51273 51274 51276 51278 51285 51290 51292
51295 51299 51301 51305 51312 51315 51320 51323 51326 51327 51330 51332 51334 51336 51337 51339 51341 51343 51345
51348 51354 51358 51359 51362 51367 51370 51373 51375 51379 51382 51384 51387 51390 51393 51396 51399 51403 51406
51409 51411 51413 51415 51423 51435 51441 51443 51445 51447 51449 51451 51452 51455 51456 51460 51463 51483 51485
51488 51493 51497 51499 51503 51506 51511 51514 51529 51531 51536 51538 51541 51544 51547 51550 51552 51553 51555
51558 51559 51561 51564 51566 51568 51569 51572 51574 51576 51578 51580 51582 51584 51586 51588 51589 51591 51594
51596 51598 51600 51603 51605 51607 51609 51614 51616 51630 51632 51634 51636 51638 51641 51645 51648 51657 51669
51672 51677 51681 51684 51686 51690 51692 51695 51697 51700 51702 51704 51705 51708 51710 51715 51717 51720 51722
51724 51727 51729 51731 51733 51736 51740 51742 51745 51747 51751 51792 51794 51800 51803 51806 51810 51813 51815
51820 51825 51828 51834 51835 51838 51840 51842 51844 51846 51848 51850 51853 51855 51857 51859 51861 51863 51865
51868 51870 51872 51876 51879 51881 51883 51885 51886 51892 51897 51899 51901 51903 51905 51906 51908 51909 51912
51914 51916 51917 51919 51921 51923 51924 51926 51929 51931 51933 51934 51936 51938 51941 51943 51964 51965 51967
51969 51971 51975 51977 51979 51983 51984 51986 51988 51989 51991 51993 51994 51997 51999 52001 52002 52004 52006
52007 52009 52011 52013 52015 52017 52018 52021 52023 52025 52026 52028 52031 52032 52034 52035 52038 52041 52043
52044 52046 52048 52051 52053 52055 52057 52059 52062 52064 52066 52068 52071 52073 52075 52077 52079 52081 52083
52085 52087 52089 52092 52096 52098 52101 52103 52105 52108 52112 52114 52116 52119 52123 52125 52128 52131 52133
52136 52138 52142 52145 52147 52149 52151 52155 52157 52159 52162 52164 52166 52169 52171 52174 52178 52179 52182
52184 52186 52187 52189 52191 52193 52195 52197 52199 52201 52202 52204 52206 52208 52209 52216 52220 52240 52242
52244 52246 52247 52250 52252 52254 52258 52260 52261 52263 52265 52268 52269 52271 52273 52274 52276 52278 52280
52282 52284 52285 52287 52288 52290 52292 52293 52295 52297 52302 52305 52307 52309 52312 52314 52316 52319 52321
52323 52326 52328 52330 52332 52335 52337 52338 52340 52341 52343 52345 52347 52350 52353 52356 52359 52361 52363
52365 52366 52368 52370 52371 52376 52377 52393 52395 52397 52399 52401 52403 52404 52406 52408 52409 52411 52413
52414 52417 52419 52421 52423 52425 52426 52428 52430 52433 52435 52437 52439 52441 52442 52444 52447 52449 52451
52453 52455 52456 52458 52460 52463 52465 52467 52469 52471 52473 52475 52478 52479 52481 52483 52485 52487 52489
52491 52493 52495 52497 52499 52502 52504 52506 52509 52511 52513 52515 52517 52518 52521 52523 52526 52528 52530
52532 52533 52535 52537 52539 52542 52545 52547 52550 52553 52555 52557 52563 52565 52568 52570 52573 52575 52578
52580 52583 52585 52587 52590 52592 52596 52599 52601 52603 52605 52608 52610 52612 52615 52617 52621 52623 52626
52629 52631 52634 52638 52640 52642 52644 52646 52649 52656 52659 52662 52664 52666 52668 52671 52673 52675 52676
52678 52679 52681 52685 52687 52689 52705 52707 52710 52712 52713 52715 52717 52719 52721 52725 52732 52733 52735
52737 52739 52741 52742 52750 52752 52754 52756 52758 52759 52761 52764 52767 52769 52771 52772 52774 52776 52779
52781 52783 52784 52786 52788 52790 52799 52824 52826 52827 52829 52831 52832 52835 52837 52838 52840 52843 52845
52846 52848 52850 52853 52871 52872 52873 52874 52878 52879 52880 52882 52884 52885 52886 52896 52897 52905 52907
52910 52922 52924 52928 52932 52941 52943 52945 52947 52950 52952 52954 52956 52965 52967 52969 52971 52972 52975
52977 52979 52981 52983 52985 52987 52989 52992 52994 52996 52998 53000 53003 53005 53007 53008 53009 53014 53016
53019 53021 53023 53025 53027 53028 53031 53032 53034 53037 53039 53041 53042 53044 53045 53047 53049 53050 53052
53054 53055 53057 53059 53060 53062 53064 53066 53068 53070 53072 53075 53077 53079 53080 53082 53084 53085 53087
53089 53091 53093 53094 53097 53099 53100 53102 53103 53105 53107 53109 53110 53112 53113 53116 53118 53119 53121
53122 53124 53126 53127 53129 53131 53134 53135 53136 53139 53140 53144 53148 53151 53153 53154 53156 53158 53159
53161 53162 53164 53166 53169 53171 53173 53175 53177 53179 53181 53212 53214 53216 53218 53220 53223 53224 53226
53229 53230 53232 53234 53235 53237 53238 53240 53241 53243 53245 53246 53248 53249 53251 53253 53254 53256 53257
53259 53261 53262 53264 53265 53267 53269 53270 53272 53274 53275 53277 53279 53280 53282 53284 53285 53287 53288
53290 53292 53293 53295 53297 53298 53300 53302 53303 53305 53306 53308 53350 53352 53353 53355 53356 53358 53359
53361 53363 53364 53365 53367 53368 53371 53373 53374 53376 53378 53379 53381 53383 53384 53408 53410 53412 53413
53415 53417 53418 53420 53422 53423 53425 53427 53428 53431 53433 53434 53436 53438 53439 53441 53443 53445 53446
53448 53450 53451 53453 53455 53456 53458 53460 53461 53463 53465 53466 53467 53469 53471 53472 53474 53476 53477
53479 53481 53482 53484 53486 53489 53490 53492 53493 53495 53497 53498 53500 53502 53503 53505 53507 53510 53512
53514 53516 53517 53519 53520 53522 53523 53525 53527 53528 53531 53533 53534 53536 53537 53539 53541 53543 53545
53547 53548 53550 53552 53553 53555 53557 53559 53561 53563 53565 53568 53570 53571 53573 53575 53576 53578 53580
53581 53583 53584 53586 53588 53589 53592 53594 53595 53597 53600 53601 53603 53605 53607 53608 53610 53612 53613
53615 53617 53618 53620 53622 53623 53625 53627 53628 53631 53633 53635 53637 53639 53641 53642 53644 53646 53647
53649 53651 53654 53656 53657 53660 53662 53664 53665 53667 53669 53670 53672 53674 53675 53677 53679 53680 53682
53683 53685 53687 53688 53690 53692 53693 53695 53696 53698 53700 53701 53703 53704 53706 53707 53708 53710 53712
53714 53715 53718 53719 53721 53722 53724 53726 53728 53730 53732 53734 53736 53739 53741 53743 53745 53747 53749
53751 53753 53755 53756 53758 53760 53762 53764 53766 53768 53770 53772 53773 53777 53779 53781 53783 53785 53787
53804 53806 53808 53809 53811 53813 53814 53818 53819 53820 53822 53823 53825 53826 53828 53830 53832 53834 53836
53854 53856 53857 53859 53860 53862 53863 53865 53866 53868 53869 53900 53901 53903 53904 53905 53907 53909 53910
53912 53914 53916 53917 53919 53924 53926 53927 53928 53930 53932 53933 53935 53937 53938 53940 53942 53943 53945
53948 53950 53952 53953 53955 53957 53958 53960 53962 53963 53964 53966 53968 53969 53971 53973 53974 53980 53983
53986 53988 54001 54003 54005 54007 54027 54029 54031 54033 54034 54036 54037 54039 54041 54042 54043 54045 54047
54049 54050 54052 54053 54054 54056 54057 54059 54060 54062 54063 54065 54066 54088 54089 54091 54093 54094 54096
54097 54099 54100 54102 54103 54105 54107 54108 54110 54111 54113 54115 54116 54118 54119 54121 54123 54124 54125
54127 54129 54131 54133 54135 54144 54146 54148 54156 54158 54161 54163 54165 54167 54169 54171 54173 54175 54177
54179 54181 54182 54184 54186 54188 54189 54191 54193 54194 54196 54198 54199 54201 54202 54204 54205 54207 54209
54210 54212 54213 54215 54217 54218 54220 54222 54224 54226 54228 54229 54231 54233 54234 54236 54237 54239 54241
54242 54244 54246 54247 54249 54250 54252 54253 54255 54257 54258 54260 54261 54263 54265 54266 54268 54269 54271
54272 54275 54277 54279 54281 54283 54284 54286 54287 54289 54290 54292 54294 54295 54297 54298 54300 54301 54303
54305 54306 54309 54311 54312 54314 54315 54317 54319 54320 54322 54323 54325 54326 54328 54330 54331 54333 54334
54336 54337 54339 54341 54342 54344 54345 54347 54348 54350 54351 54365 54366 54368 54370 54371 54373 54374 54376
54378 54379 54381 54382 54384 54385 54387 54388 54390 54391 54393 54394 54396 54397 54399 54401 54403 54405 54406
54408 54410 54412 54414 54415 54417 54418 54420 54421 54423 54425 54426 54428 54429 54431 54434 54436 54437 54439
54441 54443 54445 54447 54449 54451 54453 54454 54456 54457 54459 54461 54462 54464 54465 54473 54475 54476 54478
54479 54481 54483 54485 54486 54488 54491 54493 54495 54497 54498 54504 54506 54523 54529 54536 54546 54549 54573
54584 54586 54588 54606 54608 54610 54612 54614 54616 54618 54619 54621 54623 54624 54626 54627 54629 54630 54632
54634 54635 54637 54638 54640 54642 54644 54646 54648 54649 54651 54653 54654 54656 54658 54671 54673 54675 54676
54678 54680 54681 54683 54685 54686 54690 54692 54693 54695 54697 54698 54700 54702 54703 54705 54707 54708 54710
54712 54714 54715 54717 54719 54720 54722 54724 54725 54727 54730 54732 54733 54738 54740 54742 54744 54745 54747
54749 54751 54752 54754 54756 54758 54760 54761 54763 54765 54767 54769 54771 54772 54774 54776 54778 54779 54781
54783 54785 54786 54788 54789 54791 54793 54795 54796 54798 54799 54801 54802 54804 54805 54807 54809 54811 54812
54814 54818 54820 54822 54824 54826 54827 54829 54831 54832 54834 54837 54838 54840 54842 54844 54846 54849 54850
54852 54856 54857 54859 54860 54862 54864 54866 54868 54869 54871 54872 54874 54876 54877 54879 54881 54885 54887
54889 54890 54893 54895 54897 54898 54900 54902 54903 54905 54906 54908 54910 54914 54916 54917 54919 54921 54922
54924 54925 54927 54928 54929 54931 54933 54935 54937 54938 54940 54942 54943 54945 54947 54948 54950 54952 54953
54967 54969 54971 54972 54974 54975 54977 54978 54980 54982 55001 55003 55004 55006 55008 55010 55012 55014 55016
55018 55020 55021 55024 55025 55028 55030 55032 55033 55035 55037 55038 55040 55042 55043 55045 55047 55048 55050
55052 55054 55055 55057 55059 55061 55062 55069 55071 55073 55074 55076 55078 55080 55081 55083 55085 55086 55088
55090 55092 55093 55095 55097 55099 55101 55103 55105 55106 55108 55110 55112 55114 55115 55117 55119 55121 55124
55126 55142 55144 55146 55147 55149 55151 55152 55154 55167 55173 55176 55177 55178 55180 55182 55183 55185 55188
55190 55192 55193 55195 55207 55208 55210 55212 55230 55232 55233 55235 55309 55311 55313 55316 55317 55593 55621
55622 55627 55633 55634 55636 55638 55641 55657 55659 55661 55663 55672 55674 55676 55681 55683 55685 55687 55689
55691 55693 55695 55697 55699 55701 55703 55705 55707 55709 55712 55714 55716 55718 55720 55722 55726 55728 55729
55731 55733 55734 55736 55737 55739 55741 55743 55744 55746 55748 55749 55751 55753 55754 55757 55759 55760 55762
55763 55765 55810 55812 55813 55815 55817 55818 55820 55821 55823 55827 55828 55830 55832 55836 55837 55839 55841
55842 55843 55863 55868 55870 55871 55873 55875 55877 55879 55881 55883 55886 55888 55890 55892 55894 55896 55898
55903 55904 55906 55910 55912 55914 55916 55918 55920 55922 55923 55925 55938 55940 55943 55945 55947 55966 55969
55986 55987 55988 55990 55993 56001 56011 56012 56062 56063 56067 56069 56071 56073 56075 56077 56079 56084 56086
56093 56095 56099 56101 56103 56105 56107 56109 56111 56113 56116 56118 56123 56125 56127 56129 56131 56132 56134
56135 56157 56165 56171 56179 56186 56192 56200 56207 56213 56220 56228 56235 56242 56250 56256 56264 56271 56277
56284 56291 56297 56304 56312 56318 56326 56347 56376 56381 56386 56390 56393 56395 56396 56397 56399 56402 56424
56426 56427 56429 56431 56433 56436 56437 56438 56440 56444 56446 56447 56514 56535 56536 56539 56552 56554 56562
56564 56565 56567 56568 56570 56571 56573 56574 56575 56577 56578 56579 56581 56582 56588 56589 56591 56592 56594
56596 56597 56599 56600 56602 56604 56605 56606 56608 56609 56610 56618 56619 56620 56627 56628 56631 56633 56639
56641 56643 56645 56646 56648 56650 56651 56654 56655 56656 56658 56659 56661 56662 56664 56665 56667 56669 56670
56672 56674 56675 56677 56690 56692 56694 56697 56699 56700 56703 56706 56708 56710 56711 56713 56715 56716 56718
56719 56721 56722 56725 56728 56730 56732 56734 56736 56737 56739 56740 56742 56743 56745 56746 56749 56751 56754
56755 56756 56758 56760 56761 56763 56794 56796 56797 56798 56800 56802 56804 56805 56807 56828 56829 56832 56834
56836 56839 56841 56842 56844 56845 56857 56858 56860 56861 56862 57042 57044 57046 57047 57049 57051 57052 57053
57056 57057 57059 57060 57062 57063 57065 57067 57070 57072 57081 57088 57090 57091 57093 57095 57096 57098 57099
57102 57105 57108 57114 57116 57118 57119 57123 57132 57134 57136 57138 57141 57162 57170 57172 57214 57216 57222
57224 57228 57244 57246 57249 57250 57252 57254 57256 57258 57260 57266 57268 57269 57271 57272 57274 57275 57276
57278 57279 57280 57282 57284 57285 57287 57289 57291 57293 57295 57297 57299 57302 57304 57306 57308 57310 57311
57313 57314 57316 57317 57319 57320 57321 57323 57324 57326 57327 57329 57331 57332 57334 57335 57336 57338 57339
57341 57342 57344 57346 57348 57350 57352 57354 57356 57358 57360 57362 57364 57370 57372 57374 57376 57378 57380
57382 57384 57386 57392 57394 57396 57398 57405 57410 57411 57413 57415 57417 57419 57420 57422 57424 57426 57427
57429 57431 57433 57435 57436 57438 57440 57442 57444 57446 57452 57462 57464 57466 57468 57470 57472 57474 57476
57484 57486 57495 57498 57500 57502 57504 57506 57509 57511 57513 57515 57517 57519 57520 57522 57524 57526 57528
57542 57544 57546 57547 57549 57551 57553 57555 57557 57558 57560 57564 57566 57567 57569 57571 57572 57574 57576
57578 57579 57580 57582 57584 57587 57588 57590 57592 57594 57595 57597 57598 57599 57601 57604 57606 57608 57610
57614 57615 57617 57619 57621 57623 57629 57631 57633 57635 57637 57639 57641 57642 57644 57646 57648 57650 57652
57654 57656 57658 57660 57662 57668 57671 57673 57676 57678 57679 57681 57682 57684 57686 57687 57689 57690 57692
57694 57695 57697 57698 57700 57703 57705 57706 57708 57710 57711 57713 57715 57717 57718 57720 57722 57724 57726
57728 57729 57731 57733 57734 57766 57767 57769 57771 57772 57774 57775 57777 57779 57784 57786 57788 57789 57792
57795 57797 57799 57801 57803 57805 57806 57809 57811 57813 57815 57817 57818 57820 57822 57824 57826 57829 57831
57833 57835 57837 57839 57841 57843 57846 57855 57857 57859 57866 57881 57883 57885 57887 57889 57891 57893 57899
57901 57903 57905 57907 57909 57911 57913 57914 57916 57918 57920 57922 57924 57926 57928 57930 57932 57933 57935

```
57937 57957 57962 57964 57966 57976 57978 57980 57982 57989 57991 57993 57994 57996 57998 58000 58002 58006 58008
58010 58013 58015 58017 58019 58022 58024 58030 58032 58034 58035 58036 58038 58039 58041 58043 58045 58047 58049
58050 58052 58053 58054 58057 58058 58059 58061 58063 58065 58067 58069 58071 58072 58074 58077 58078 58080 58081
58083 58084 58086 58087 58088 58090 58091 58093 58094 58095 58097 58099 58101 58102 58104 58106 58108 58110 58112
58114 58115 58117 58118 58121 58123 58124 58130 58131 58132 58134 58136 58138 58140 58142 58143 58145 58147 58149
58150 58152 58154 58156 58163 58165 58166 58168 58170 58172 58174 58176 58178 58180 58182 58184 58185 58187 58189
58191 58193 58195 58197 58199 58200 58202 58204 58206 58208 58210 58212 58214 58216 58218 58220 58221 58223 58225
58227 58229 58231 58233 58240 58265 58334 58361 58363 58365 58366 58368 58370 58372 58374 58376 58378 58379 58381
58383 58385 58387 58389 58391 58393 58395 58397 58399 58401 58403 58405 58407 58411 58413 58415 58417 58418 58420
58422 58424 58426 58428 58430 58432 58434 58436 58438 58440 58441 58449 58451 58453 58455 58457 58458 58460 58462
58464 58466 58467 58469 58472 58474 58476 58478 58480 58482 58483 58485 58487 58488 58490 58491 58493 58495 58496
58498 58500 58502 58534 58536 58538 58540 58542 58545 58547 58549 58550 58552 58554 58555 58556 58558 58560 58562
58563 58565 58566 58568 58570 58572 58573 58575 58577 58583 58585 58586 58588 58590 58595 58596 58598 58599 58601
58602 58604 58606 58612 58616 58617 58619 58620 58622 58624 58625 58627 58629 58631 58633 58636 58638 58640 58642
58644 58646 58648 58650 58652 58654 58656 58658 58660 58662 58663 58664 58666 58667 58669 58670 58672 58674 58675
58677 58679 58681 58683 58684 58686 58688 58690 58692 58693 58695 58697 58699 58701 58702 58704 58706 58708 58710
58711 58713 58715 58717 58719 58721 58727 58729 58730 58732 58734 58736 58753 58759 58762 58764 58766 58772 58775
58777 58778 58802 58804 58806 58808 58810 58812 58814 58815 58817 58819 58821 58844 58865 58868 58881 58882 58883
58884 58885 58886 58887 58888 58890 58892 58894 58896 58898 58900 58902 58904 58906 58908 58909 58911 58913 58915
58917 58919 58921 58923 58925 58927 58929 58931 58933 58935 58937 58939 58940 58942 58944 58946 58948 58950 58952
58958 58960 58962 58964 58966 58967 58971 58973 58975 58977 58980 58982 58984 58986 58988 58990 58992 58998 59007
59012 59014 59016 59022 59028 59029 59032 59034 59044 59046 59055 59057 59069 59071 59080 59155 59157 59159 59163
59165 59166 59168 59170 59172 59174 59176 59179 59180 59182 59184 59188 59190 59192 59194 59196 59198 59200 59202
59203 59205 59207 59209 59211 59213 59279 59329 59341 59342 59344 59346 59348 59361 59362 59364 59366 59368 59370
59372 59374 59376 59378 59380 59382 59384 59386 59388 59390 59417 59419 59421 59423 59425 59428 59430 59432 59434
59435 59437 59439 59441 59443 59445 59447 59448 59450 59451 59453 59456 59458 59460 59462 59464 59466 59467 59469
59471 59472 59474 59476 59477 59479 59481 59483 59485 59486 59487 59489 59491 59492 59494 59496 59498 59500 59501
59503 59509 59511 59512 59515 59517 59519 59521 59523 59525 59527 59529 59532 59535 59536 59538 59540 59542 59544
59546 59550 59552 59555 59557 59559 59561 59563 59567 59570 59572 59574 59576 59578 59580 59582 59583 59585 59587
59589 59590 59594 59596 59598 59600 59601 59603 59605 59607 59609 59611 59613 59615 59617 59619 59620 59622 59623
59625 59627 59629 59631 59633 59640 59642 59644 59646 59647 59649 59651 59653 59654 59656 59658 59660 59661 59663
59664 59666 59668 59670 59672 59674 59677 59679 59681 59682 59684 59686 59688 59690 59692 59694 59696 59698 59699
59701 59702 59704 59705 59707 59709 59710 59712 59714 59716 59736 59738 59739 59741 59743 59745 59747 59749 59751
59753 59770 59778 59782 59785 59787 59789 59793 59795 59797 59799 59802 59803 59810 59813 59816 59829 59831 59832
59834 59836 59838 59840 59842 59843 59845 59847 59849 59851 59853 59855 59857 59860 59862 59864 59866 59868 59870
59872 59874 59876 59878 59880 59882 59884 59886 59888 59890 59892 59894 59896 59898 59900 59904 59906 59908 59910
59912 59914 59916 59918 59919 59921 59923 59924 59926 59928 59930 59931 59933 59939 59941 59942 59944 59946 59947
59949 59951 59953 59954 59955 59957 59959 59960 59962 59964 59965 59967 59969 59971 59972 59974 59976 59978 59979
59981 59983 59985 59987 59988 59990 59992 59994 59995 59996 59998 59999 60001 60003 60005 60007 60009 60011 60013
60014 60016 60018 60020 60022 60024 60026 60028 60030 60031 60033 60035 60037 60039 60041 60043 60045 60047 60049
60050 60052 60054 60056 60058 60060 60062 60068 60070 60071 60073 60074 60076 60077 60079 60082 60084 60086 60087
60089 60091 60093 60095 60097 60098 60100 60101 60103 60104 60106 60108 60110 60112 60114 60116 60118 60120 60122
60124 60125 60127 60129 60130 60132 60133 60135 60136 60138 60139 60141 60142 60144 60146 60147 60149 60150 60152
60153 60155 60156 60158 60159 60161 60163 60164 60166 60167 60168 60170 60172 60175 60177 60179 60181 60183 60185
60187 60189 60191 60193 60195 60215 60216 60218 60220 60222 60224 60226 60228 60231 60233 60235 60237 60238 60240
60241 60242 60244 60245 60247 60248 60249 60251 60253 60254 60256 60257 60260 60262 60263 60265 60267 60269 60272
60274 60275 60281 60283 60284 60286 60288 60290 60292 60294 60296 60298 60300 60302 60303 60305 60306 60308 60310
60311 60313 60315 60316 60319 60320 60322 60324 60325 60327 60329 60331 60332 60334 60336 60337 60339 60341 60342
60344 60346 60347 60349 60351 60352 60354 60356 60357 60359 60361 60362 60364 60366 60367 60369 60371 60373 60374
60376 60378 60380 60382 60383 60385 60387 60388 60390 60392 60393 60396 60398 60399 60401 60403 60405 60407 60409
60410 60412 60414 60416 60418 60420 60421 60423 60425 60426 60428 60429 60431 60433 60435 60436 60438 60440 60441
60443 60444 60446 60448 60449 60451 60453 60455 60456 60458 60460 60461 60463 60465 60466 60468 60469 60471 60473
60474 60478 60479 60481 60483 60484 60489 60490 60492 60494 60495 60497 60499 60500 60502 60503 60505 60508 60511
60512 60514 60516 60517 60519 60520 60522 60524 60525 60527 60533 60535 60536 60538 60539 60541 60543 60544 60546
60548 60549 60551 60553 60554 60556 60557 60559 60561 60562 60564 60565 60567 60568 60570 60571 60573 60574 60576
60578 60579 60581 60582 60583 60585 60586 60587 60589 60590 60592 60593 60594 60596 60597 60599 60600 60601 60603
60604 60606 60608 60610 60614 60615 60617 60619 60621 60622 60624 60625 60627 60629 60631 60633 60635 60637 60639
60671 60673 60675 60677 60679 60685 60686 60688 60690 60691 60693 60695 60696 60698 60699 60701 60703 60704 60706
60708 60709 60711 60712 60714 60715 60717 60719 60720 60722 60723 60725 60727 60728 60730 60731 60733 60735 60736
60738 60740 60741 60743 60746 60747 60749 60751 60752 60754 60756 60758 60759 60761 60763 60766 60767 60771 60775
60776 60779 60780 60782 60783 60785 60786 60788 60789 60791 60792 60794 60796 60798 60800 60801 60803 60805 60806
60808 60814 60815 60818 60820 60821 60823 60824 60826 60827 60829 60831 60832 60834 60835 60837 60838 60840 60842
60843 60845 60847 60848 60850 60852 60853 60855 60856 60858 60859 60861 60862 60864 60865 60867 60869 60870 60872
60873 60875 60877 60878 60880 60881 60883 60884 60886 60888 60889 60891 60894 60896 60897 60899 60901 60903 60905
60907 60909 60910 60912 60913 60915 60916 60918 60919 60921 60922 60924 60925 60927 60928 60930 60932 60934 60936
60938 60940 60941 60943 60946 60948 60949 60951 60953 60954 60956 60958 60959 60965 60967 60969 60970 60972 60974
60976 60977 60979 60981 60982 60984 60986 60987 60989 60990 60992 60994 60995 60997 60999 61000 61003 61005 61006
61008 61010 61011 61013 61015 61017 61018 61020 61022 61023 61025 61027 61028 61030 61032 61033 61035 61037 61038
61040 61042 61043 61045 61047 61048 61050 61052 61053 61055 61057 61058 61060 61062 61063 61065 61066 61069 61070
61072 61074 61075 61077 61079 61080 61082 61083 61085 61087 61088 61090 61092 61093 61095 61097 61098 61100 61103
61104 61110 61113 61114 61116 61118 61119 61121 61123 61124 61126 61127 61129 61131 61133 61134 61137 61139 61140
61142 61143 61145 61147 61148 61150 61151 61154 61155 61157 61158 61160 61161 61163 61165 61167 61169 61171 61173
61176 61178 61180 61182 61184 61186 61187 61189 61191 61193 61194 61196 61198 61201 61203 61206 61208 61210 61212
61214 61216 61218 61220 61222 61223 61225 61226 61228 61229 61231 61233 61235 61237 61239 61241 61251 61254 61255
61257 61259 61261 61263 61265 61267 61269 61271 61273 61275 61277 61279 61281 61284 61286 61292 61294 61296 61298
61301 61345 61347 61349 61351 61353 61355 61357 61359 61361 61363 61365 61367 61373 61375 61377 61379 61381 61383
61385 61387 61389 61390 61392 61393 61395 61396 61398 61399 61401 61402 61404 61406 61408 61410 61411 61413 61415
61416 61419 61421 61423 61424 61426 61428 61429 61431 61433 61434 61436 61438 61440 61441 61443 61444 61447 61449
61450 61452 61454 61455 61457 61459 61460 61462 61464 61466 61468 61470 61471 61473 61474 61476 61478 61480 61481
61483 61485 61486 61488 61489 61491 61493 61496 61497 61499 61501 61502 61504 61507 61522 61524 61525 61527 61529
61530 61532 61534 61536 61537 61539 61541 61542 61544 61546 61548 61549 61551 61553 61555 61556 61558 61560 61562
61563 61565 61567 61569 61571 61572 61574 61576 61578 61579 61581 61583 61585 61586 61588 61590 61592 61594 61595
61598 61606 61608 61609 61611 61613 61615 61617 61618 61620 61622 61624 61625 61627 61629 61631 61632 61634 61636
61638 61641 61642 61644 61646 61648 61649 61651 61653 61654 61656 61662 61664 61665 61667 61669 61671 61672 61674
61675 61677 61679 61681 61682 61684 61686 61687 61689 61691 61693 61694 61696 61698 61699 61701 61703 61704 61706
61707 61709 61710 61712 61713 61715 61716 61718 61719 61720 61721 61723 61725 61727 61728 61729 61730 61732 61734
61736 61742 61744 61746 61748 61750 61752 61754 61756 61758 61760 61762 61764 61766 61768 61770 61772 61775 61777
61779 61781 61783 61785 61787 61789 61796 61798 61800 61801 61803 61805 61806 61808 61809 61811 61813 61814 61816
61817 61819 61821 61822 61824 61825 61827 61828 61830 61832 61833 61835 61837 61838 61840 61842 61843 61845 61847
61848 61850 61852 61853 61855 61856 61858 61859 61861 61863 61864 61866 61868 61869 61871 61872 61874 61875 61877
61879 61880 61882 61884 61886 61887 61889 61891 61892 61894 61896 61898 61900 61902 61904 61908 61909 61911 61913
61915 61917 61919 61920 61922 61924 61926 61928 61930 61931 61933 61935 61941 61943 61945 61947 61949 61951 61952
61954 61956 61958 61960 61962 61964 61966 61969 61971 61973 61975 61976 61978 61980 61982 61984 61986 61988 61989
61990 61992 61993 61996 61997 61999 62001 62002 62004 62006 62008 62009 62011 62013 62015 62017 62019 62021 62022
62024 62026 62028 62030 62032 62033 62035 62037 62039 62041 62043 62044 62046 62048 62050 62052 62054 62056 62057
62059 62061 62072 62074 62076 62078 62080 62081 62083 62085 62087 62089 62091 62093 62095 62096 62098 62100 62102
62104 62106 62108 62110 62112 62113 62115 62117 62119 62121 62123 62124 62127 62128 62130 62132 62134 62135 62137
62139 62140 62143 62145 62150 62152 62153 62155 62157 62159 62161 62163 62165 62166 62168 62170 62172 62174 62176
62178 62180 62182 62184 62187 62189 62191 62192 62228 62230 62232 62234 62236 62238 62240 62241 62243 62245 62247
62249 62250 62252 62254 62256 62258 62260 62261 62263 62265 62267 62269 62270 62272 62274 62276 62277 62279 62281
62283 62285 62286 62288 62290 62291 62293 62295 62297 62298 62300 62302 62303 62305 62307 62309 62310 62312 62314
62315 62317 62319 62321 62322 62324 62326 62328 62329 62331 62334 62336 62338 62340 62342 62344 62345 62347 62349
62351 62352 62354 62355 62357 62359 62369 62370 62384 62386 62388 62390 62392 62394 62396 62398 62400 62402 62404
62406 62408 62416 62418 62420 62422 62424 62426 62442 62444 62446 62448 62450 62452 62454 62456 62458 62460 62462
62464 62466 62468 62475 62477 62479 62484 62486 62489 62491 62498 62500 62502 62505 62506 62508 62509 62511 62513
62515 62517 62523 62525 62527 62529 62531 62533 62535 62537 62539 62541 62543 62545 62547 62549 62551 62553 62554
62556 62559 62560 62565 62566 62568 62570 62577 62579 62581 62583 62596 62599 62601 62602 62610 62612 62614 62616
62623 62625 62634 62636 62638 62642 62647 62649 62651 62672 62673 62675 62677 62679 62681 62683 62687 62689 62693
62695 62697 62699 62701 62703 62705 62707 62714 62716 62722 62724 62726 62728 62730 62745 62747 62749 62751 62753
62755 62757 62759 62766 62773 62780 62782 62784 62786 62788 62789 62791 62794 62796 62798 62800 62802 62804 62805
62807 62811 62812 62815 62817 62819 62821 62822 62824 62826 62827 62829 62831 62832 62834 62836 62838 62839 62841
62843 62844 62846 62855 62857 62859 62862 62863 62865 62867 62869 62872 62875 62877 62879 62881 62883 62885 62887
62891 62895 62897 62899 62901 62902 62906 62911 62913 62915 62917 62935 62937 62943 62945 62947 62949 62951 62953
62955 62957 62969 62977 62999 63001 63003 63007 63009 63011 63013 63017 63019 63021 63024 63058 63060 63062 63064
63066 63068 63070 63072 63107 63109 63111 63113 63115 63150 63204 63206 63218 63220 63222 63243 63245 63259 63261
63263 63265 63267 63269 63271 63273 63275 63277 63279 63281 63283 63285 63287 63289 63291 63293 63296 63326 63328
63330 63332 63334 63336 63338 63340 63342 63343 63344 63345 63346 63348 63350 63352 63353 63355 63361 63363 63365
63367 63369 63371 63373 63375 63377 63379 63380 63382 63384 63387 63388 63390 63392 63393 63395 63396 63398 63400
63402 63404 63407 63409 63411 63413 63415 63417 63519 63521 63523 63525 63527 63529 63531 63532 63534 63536 63538
63539 63541 63542 63543 63544 63545 63547 63548 63550 63552 63735 63736 64074 64076 64077 64083 64087 64090 64092
64094 64096 64098 64100 64101 64110 64113 64114 64116 64117 64119 64121 64123 64124 64126 64128 64130 64132 64133
64135 64136 64138 64140 64142 64144 64146 64148 64150 64151 64153 64155 64159 64162 64166 64168 64196 64272 64277
64278 64400 64403 64404 64406 64408 64409 64411 64413 64417 64419 64423 64424 64426 64428 64430 64432 64434 64436
64437 64439 64440 64442 64444 64446 64448 64450 64454 64455 64592 64600 64602 64604 64606 64608 64609 64611 64658
64659 64661 64662 64665 64666 64668 64675 64688 64690 64692 64725 64735 64736 64737 64738 64739 64741 64743 64744
64746 64748 64750 64751 64753 64755 64756 64758 64760 64772 64773 64776 64780 64782 64783 64785 64787 64789 64791
64793 64795 64797 64799 64801 64803 64805 64807 64887 64889 64905 64907 64908 64923 64981 64983 64984 64986 65002
65004 65005 65011 65013 65015 65016 65018 65019 65066 65068 65069 65071 65072 65088 65091 65094 65095 65100 65109
65116 65117 65118 65119 65120 65122 65123 65125 65126 65133 65135 65136 65138 65139 65140 65143 65145 65146 65148
65149 65150 65152 65153 65154 65160 65176 65177 65178 65180 65181 65183 65184 65185 65187 65188 65189 65191 65192
65193 65195 65196 65198 65199 65201 65203 65205 65211 65212 65214 65215 65217 65220 65221 65223 65224 65226 65228
65229 65231 65233 65234 65236 65238 65239 65241 65243 65244 65246 65247 65249 65251 65252 65254 65256 65258 65259
65261 65263 65265 65266 65268 65269 65271 65273 65274 65276 65278 65280 65281 65283 65284 65286 65288 65289 65291
65292 65294 65295 65296 65298 65310 65311 65313 65315 65317 65319 65321 65323 65325 65326 65328 65329 65331 65332
65333 65335 65336 65338 65339 65341 65342 65344 65345 65346 65347 65349 65350 65351 65352 65353 65355 65356 65357
65358 65360 65365 65366 65368 65369 65371 65372 65374 65376 65378 65379 65381 65383 65384 65386 65388 65390 65393
65396 65398 65400 65402 65404 65415 65417 65419 65421 65423 65425 65427 65428 65434 65436 65438 65440 65441 65443
65445 65447 65449 65453 65457 65459 65461 65462 65464 65465 65467 65469 65471 65472 65474 65475 65477 65478 65480
65482 65484 65487 65489 65491 65493 65494 65495 65497 65499 65500 65502 65503 65505 65506 65508 65509 65511 65512
65518 65520 65521 65523 65525 65527 65528 65530 65531 65532 65533 65535 65537 65539 65540 65542 65544 65545 65547
65548 65550 65551 65552 65554 65555 65556 65558 65559 65561 65563 65565 65566 65568 65570 65572 65573 65575 65577
65579 65581 65583 65585 65587 65588 65590 65592 65593 65595 65597 65598 65600 65602 65603 65605 65607 65609 65610
65612 65613 65615 65617 65619 65620 65622 65624 65625 65627 65628 65630 65632 65633 65635 65637 65638 65640 65642
65643 65645 65647 65649 65651 65652 65654 65656 65658 65660 65661 65667 65669 65671 65673 65675 65676 65678 65680
65682 65684 65685 65688 65694 65696 65698 65699 65701 65710 65712 65713 65715 65716 65718 65719 65721 65722 65724
65726 65727 65729 65730 65732 65737 65740 65742 65743 65745 65747 65748 65751 65752 65754 65755 65757 65758 65760
65762 65764 65765 65767 65768 65779 65780 65782 65784 65790 65792 65793 65795 65796 65798 65799 65801 65809 65854
65856 65858 65860 65870 65872 65874 65875 65882 65884 65885 65887 65888 65890 65892 65894 65896 65898 65915 65917
65919 65921 65922 65924 65926 65928 65930 65932 65933 65935 65937 65939 65940 65942 65943 65945 65947 65949 65951
65954 65955 65958 65961 65963 65965 65967 65969 65971 65973 65975 65981 65982 65984 65986 65987 65989 65990 65992
65993 65995 65997 65999 66001 66002 66004 66006 66007 66009 66011 66013 66014 66016 66018 66020 66022 66024 66025
66027 66029 66031 66033 66035 66041 66044 66048 66051 66055 66058 66060 66062 66064 66065 66067 66069 66071 66073
66075 66077 66079 66081 66083 66085 66087 66089 66090 66092 66093 66095 66097 66098 66100 66101 66103 66105 66107
66114 66115 66117 66119 66120 66122 66124 66125 66127 66129 66131 66132 66134 66135 66137 66139 66140 66142 66144
66146 66153 66154 66155 66159 66161 66163 66164 66166 66168 66170 66171 66173 66176 66177 66183 66185 66187 66189
66191 66193 66195 66197 66199 66200 66202 66204 66206 66215 66216 66221 66222 66224 66226 66232 66233 66235 66237
66244 66246 66248 66250 66251 66253 66255 66257 66258 66260 66262 66264 66266 66267 66269 66271 66273 66277 66278
```

66280 66282 66284 66286 66287 66289 66291 66293 66295 66297 66299 66300 66304 66306 66308 66310 66312 66314 66316
66318 66320 66322 66324 66326 66328 66330 66332 66334 66336 66338 66339 66340 66341 66343 66345 66347 66349 66351
66353 66359 66361 66362 66364 66365 66367 66369 66371 66373 66375 66378 66380 66382 66384 66386 66388 66390 66392
66394 66396 66398 66399 66400 66402 66404 66407 66454 66461 66462 66464 66466 66468 66470 66472 66474 66476 66478
66480 66482 66484 66491 66493 66495 66497 66499 66501 66506 66508 66510 66512 66514 66515 66517 66519 66522 66524
66526 66528 66530 66532 66534 66536 66538 66540 66541 66543 66544 66546 66548 66550 66552 66554 66555 66557 66558
66560 66562 66563 66565 66568 66569 66571 66573 66575 66577 66578 66580 66581 66583 66585 66587 66589 66590 66592
66594 66596 66598 66599 66602 66604 66606 66608 66614 66616 66617 66619 66621 66622 66624 66626 66628 66630 66632
66634 66636 66638 66640 66642 66643 66645 66647 66649 66651 66653 66654 66656 66658 66660 66662 66664 66665 66667
66669 66671 66673 66675 66677 66678 66680 66682 66684 66686 66687 66689 66691 66693 66694 66696 66698 66700 66701
66703 66705 66707 66709 66711 66713 66714 66716 66718 66720 66722 66724 66726 66728 66730 66732 66733 66735 66737
66743 66745 66747 66749 66751 66752 66754 66756 66758 66760 66762 66763 66765 66767 66768 66770 66772 66774 66776
66778 66779 66781 66783 66785 66786 66788 66790 66792 66793 66795 66797 66798 66800 66804 66805 66807 66809 66811
66812 66814 66816 66818 66821 66823 66825 66827 66829 66831 66832 66834 66836 66838 66839 66841 66843 66844 66846
66847 66849 66851 66853 66855 66857 66859 66861 66862 66864 66866 66867 66869 66870 66871 66873 66875 66877 66878
66884 66886 66888 66892 66894 66896 66898 66899 66901 66903 66909 66911 66913 66915 66917 66919 66923 66925 66934
66936 66938 66939 66941 66943 66944 66946 66947 66949 66950 66952 66953 66955 66957 66958 66960 66961 66963 66965
66967 66969 66970 66975 66978 66998 67021 67023 67025 67027 67029 67035 67037 67039 67041 67043 67045 67047 67049
67051 67053 67055 67057 67059 67062 67063 67065 67067 67069 67071 67073 67075 67077 67079 67081 67082 67084 67086
67088 67089 67091 67093 67098 67100 67102 67104 67106 67108 67109 67111 67113 67115 67117 67119 67121 67122 67124
67126 67128 67130 67132 67133 67135 67136 67138 67140 67141 67143 67145 67146 67148 67150 67156 67185 67187 67189
67190 67192 67194 67196 67198 67200 67202 67204 67206 67208 67210 67212 67213 67215 67217 67219 67220 67222 67224
67226 67228 67229 67231 67234 67236 67237 67239 67242 67244 67248 67250 67252 67253 67255 67257 67259 67261 67263
67266 67269 67270 67272 67274 67276 67277 67279 67280 67282 67284 67286 67288 67289 67291 67292 67294 67295 67297
67298 67300 67301 67306 67307 67310 67312 67313 67315 67317 67323 67325 67327 67328 67330 67331 67333 67335 67336
67338 67339 67341 67342 67343 67345 67347 67348 67350 67352 67353 67355 67357 67359 67360 67362 67364 67366 67367
67369 67371 67372 67374 67375 67377 67379 67380 67382 67384 67385 67387 67389 67391 67392 67394 67396 67397 67399
67401 67403 67406 67409 67414 67418 67420 67421 67423 67425 67428 67429 67431 67433 67434 67436 67438 67440 67442
67444 67446 67448 67450 67452 67461 67463 67465 67473 67475 67477 67478 67480 67482 67484 67488 67490 67492 67493
67495 67497 67499 67501 67502 67507 67509 67510 67512 67514 67516 67518 67520 67522 67523 67525 67527 67529 67531
67533 67535 67537 67539 67540 67542 67544 67545 67547 67549 67551 67553 67554 67556 67561 67562 67564 67566 67568
67569 67571 67573 67575 67576 67578 67580 67582 67583 67585 67586 67588 67590 67592 67594 67596 67597 67605 67607
67609 67611 67613 67615 67617 67619 67622 67627 67628 67630 67632 67634 67637 67638 67639 67640 67641 67645 67647
67649 67651 67653 67656 67657 67659 67661 67663 67665 67667 67669 67671 67673 67675 67677 67679 67681 67683 67685
67687 67689 67691 67692 67694 67696 67698 67701 67703 67705 67706 67715 67717 67719 67721 67723 67725 67727 67733
67735 67737 67739 67741 67743 67745 67747 67749 67750 67752 67754 67756 67757 67759 67761 67765 67767 67768 67770
67772 67774 67775 67777 67779 67783 67785 67787 67790 67792 67793 67795 67797 67799 67804 67805 67807 67809 67811
67812 67814 67816 67817 67819 67821 67822 67824 67826 67828 67830 67832 67833 67835 67837 67838 67844 67846 67848
67850 67852 67854 67856 67858 67863 67865 67867 67869 67871 67873 67875 67877 67879 67881 67883 67885 67887 67889
67891 67893 67895 67897 67899 67901 67903 68138 68140 68142 68143 68145 68147 68174 68175 68176 68178 68179 68181
68183 68185 68187 68189 68191 68193 68196 68198 68200 68202 68204 68206 68207 68210 68212 68214 68216 68218 68222
68224 68225 68227 68229 68231 68233 68235 68237 68239 68241 68243 68245 68246 68249 68251 68253 68255 68257 68260
68262 68264 68266 68269 68271 68273 68275 68277 68279 68281 68283 68286 68287 68289 68291 68293 68295 68297 68299
68301 68303 68321 68323 68325 68327 68329 68346 68348 68350 68352 68354 68378 68380 68382 68384 68386 68388 68390
68392 68397 68399 68401 68403 68412 68414 68416 68418 68420 68422 68424 68426 68428 68432 68434 68436 68438 68450
68452 68454 68456 68458 68460 68461 68463 68466 68467 68471 68473 68475 68476 68507 68509 68511 68513 68516 68518
68520 68525 68530 68532 68534 68540 68544 68546 68549 68552 68554 68556 68558 68560 68562 68564 68566 68568 68570
68573 68575 68577 68579 68581 68583 68585 68587 68590 68592 68595 68597 68599 68601 68603 68605 68608 68610 68638
68640 68642 68644 68646 68649 68651 68653 68655 68657 68659 68661 68663 68666 68668 68670 68672 68674 68676 68678
68681 68683 68685 68687

I Bibliography

References

[1] K. G. Jansky. Radio Waves from Outside the Solar System. *Nature*, 132:66, 1933.

[2] National Aeronautics and Goddard Space Flight Center Space Administration. Multiwavelength Milky Way. `http://mwmw.gsfc.nasa.gov/mmw_product.html`, 2015. 2015-03-19.

[3] Viktor Franz Hess. Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten. *Physikalische Zeitschrift*, 13:1084—-1091, 1912.

[4] J. Abraham et al. [Pierre Auger Collaboration]. Properties and performance of the prototype instrument for the Pierre Auger Observatory. *Nuclear Instruments and Methods A*, 523:50–95, 2004.

[5] T. K. Gaisser, T. Stanev, and S. Tilav. Cosmic Ray Energy Spectrum from Measurements of Air Showers. `http://arxiv.org/abs/1303.3565v1`, 2013.

[6] Pavel A. Cherenkov. Visible emission of clean liquids by action of γ radiation. *Doklady Akademii Nauk SSSR*, 2:451, 1934.

[7] M. Ageron et al. ANTARES: The first undersea neutrino telescope. *Nucl. Inst. and Meth. A*, 656(1):11–38, 2011.

[8] I. A. Belolaptikov et al. The BAIKAL underwater neutrino telescope: Design, performance, and first results. *Astropart. Phys.*, 7(3):263–282, 1997.

[9] T. DeYoung. Neutrino astronomy with IceCube. *Mod. Phys. Lett. A*, 24(20):1543–1557, 2009.

[10] Wolfgang E. Pauli. Offener Brief an die Gruppe der Radioaktiven bei der Gauvereins-Tagung in Tübingen. 1930. Letter.

[11] C. L. Cowan et al. Detection of the free neutrino: a confirmation. *Science*, 124(3212):103–104, 1956.

[12] D. Fargion. Deflection of massive neutrinos by gravitational fields. *Lettere al Nuovo Cimento 9*, 31(2):49–52, 1981.

[13] V. Gribov and B. Pontecorvo. Neutrino astronomy and lepton charge. *Physics Letters B*, 28:493–496, 1969.

[14] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys.Rev.Lett.*, 81:1562–1567, 1998.

[15] M.C. Gonzalez-Garcia and Y. Nir. Neutrino Masses and Mixing: Evidence and Implications. *Rev.Mod.Phys.*, 75:345–402, 2003.

[16] M.C. Gonzalez-Garcia and Michele Maltoni. Phenomenology with Massive Neutrinos. *Phys.Rept.*, 460:1–129, 2008.

[17] K.A. Olive at al. (Particle Data Group). REVIEW OF PARTICLE PHYSICS. *Chin. Phys. C*, 38:090001, 2014.

[18] Julien Lesgourgues and Sergio Pastor. Neutrino mass from Cosmology. *Adv. High Energy Phys.*, 2012:608515, 2012.

[19] Julien Lesgourgues and Sergio Pastor. Neutrino cosmology and PLANCK. *New Journal of Physics*, 16:065002, 2014.

[20] PBS NOVA, Fermilab, Office of Science, United States Department of Energy, and Particle Data Group. The standard model of elementary particles. http://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg, 2015. 2015-03-28.

[21] R. Davis et al. Search for neutrinos from the Sun. *Phys. Rev. Lett.*, 20:1205–1209, 1968.

[22] C. Grupen. *Astroparticle physics.* Springer-Verlag Berlin Heidelberg New-York, 2005.

[23] K. Hirata et al. Observation of a neutrino burst from the supernova SN1987A. *Phys. Rev. Lett.*, 58:1490–1493, 1987.

[24] R. M. Bionta et al. Observation of a neutrino burst in coincidence with supernova 1987A in the Large Magellanic Cloud. *Phys. Rev. Lett.*, 58:1494–1496, 1987.

[25] E. N. Alexeyev et al. Detection of the neutrino signal from SN 1987A in the LMC using the INR Baksan underground scintillation telescope. *Phys. Lett. B*, 205(23).

[26] M. G. Aartsen et al. [IceCube Collaboration]. Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector. *Science*, 342:1242856, 2013.

[27] M. G. Aartsen et al. [IceCube Collaboration]. Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data. *Phys. Rev. Lett.*, 113:101101, 2014.

[28] E. Fermi. On the Origin of the Cosmic Radiation. *Physical Review*, 75:1169–1174, 1949.

[29] J. Hester, A. Loll, (NASA, and ESA). The crab nebula. http://upload.wikimedia.org/wikipedia/commons/0/00/Crab_Nebula.jpg, 2015. 2015-04-05.

[30] M. Su, T. R. Slatyer, and D. P. Finkbeiner. Giant gamma-ray bubbles from Fermi-LAT: Active galactic nucleus activity or bipolar galactic wind? *Astrophys. J.*, 724(2):1044, 2010.

[31] U.F. Katz and Ch. Spiering. High-Energy Neutrino Astrophysics: Status and Perspectives. *Progress in Particle and Nuclear Physics*, 67:651–704, 2012.

[32] L. Nellen, K. Mannheim, and P.L. Biermann. Neutrino production through hadronic cascades in AGN accretion disks. *Phys. Rev. D*, 47:5270–5274, 1993.

[33] Angel R. Lopez-Sanchez, Australian Astronomical Observatory, and Macquarie University. Centaurus A multiwavelength. `http://oldweb.aao.gov.au/local/www/alopez/images/Cen_A_multiwavelength_7.jpg`, 2015. 2015-04-12.

[34] C.A. Meegan, G.J. Fishman, and R.B. Wilson et al. Spatial distribution of γ-ray bursts observed by BATSE. *Nature*, 355:143, 1992.

[35] N. Gehrels and P. Mészáros. Gamma-Ray Bursts. *Nature*, 337:932, 2012.

[36] Julia Schmid. *Searches for High-Energy Neutrinos from Gamma-Ray Bursts with the Antares Neutrino Telescope.* PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg.

[37] J. F. Beacom et al. Measuring Flavor Ratios of High-Energy Astrophysical Neutrinos. *Phys. Rev. D*, 68:093005, 2003.

[38] Adriaan Heijboer. *Track Reconstruction and Point Source Searches with ANTARES.* PhD thesis, Universiteit van Amsterdam.

[39] Florian Folger. *Search for a diffuse cosmic neutrino flux using shower events in the ANTARES neutrino telescope.* PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg.

[40] P. A. Cherenkov. Visible radiation produced by electrons moving in a medium with velocities exceeding that of light. *Phys. Rev.*, 52:378—379, 1937.

[41] Camille Moirenc and CNRS. Optical module. `http://antares.in2p3.fr/Gallery/CNRS/LineConstruction/116_Antares_Vincenttriplet.jpg`, 2015. 2015-04-09.

[42] Ronald Bruijn. *The Antares Neutrino Telescope: Performance Studies and Analysis of First Data.* PhD thesis, Universiteit van Amsterdam.

[43] Francois Montanet. Depiction of ANTARES overview. `http://antares.in2p3.fr/Gallery/3D/antares_pub.jpeg`, 2015. 2015-04-09.

[44] S. Adrian-Martinez et al. The positioning system of the ANTARES Neutrino Telescope. *JINST*, 7:T08002, 2012.

[45] B.A.P. van Rens. The software trigger in ANTARES. Antares Internal Note ANTARES-SOFT-2004-001, 2004.

[46] M. De Jong. The ANTARES Trigger Software. Antares Internal Note ANTARES-SOFT-2005-005, 2005.

[47] John Carr, Stephanie Escoffier, and Dmitry Zaborov. Proposition for an alternative trigger based on the T3 cluster trigger. Antares Internal Note ANTARES-SOFT-2007-016, 2007.

[48] S. Fukuda et al. (SuperKamiokande). The Super-Kamiokande detector. *Nucl. Instrum. Meth. A*, 501:418, 2003.

[49] Ulrich F. Katz (for the KM3NeT Collaboration). The ORCA Option for KM3NeT, 2013. PoS(Neutel 2013)057.

[50] Bettina Hartmann. *Reconstruction of neutrino-induced hadronic and electromagnetic showers with the ANTARES experiment.* PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg.

[51] C. Tamburini, M. Canals, and X.D. de Madron et al. Deep-Sea Bioluminescence Blooms after Dense Water Formation at the Ocean Surface. *PloS one*, 8:e67523, 2013.

[52] Nick Hobgood. Comb jellyfish. `http://en.wikipedia.org/wiki/File:Combjelly.jpg`, 2015. 2015-04-12.

[53] Kevin Raskoff and oceanexplorer.noaa.gov. Arctic comb jellyfish. `http://en.wikipedia.org/wiki/File:LightRefractsOf_comb-rows_of_ctenophore_Mertensia_ovum.jpg`, 2015. 2015-04-12.

[54] The ANTARES Collaboration. A deep sea telescope for high energy neutrinos. `http://arxiv.org/abs/astro-ph/9907432`, 1999. 2015-03-19.

[55] M. Honda, T. Kajita, K. Kasahara, S. Midorikawa, and T. Sanuki. Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data. *Phys.Rev.D*, 75:043006, 2007.

[56] Morihiro Honda, Takaaki Kajita, Katsuaki Kasahara, and Shoichi Midorikawa. Improvement of low energy atmospheric neutrino flux calculation using the JAM nuclear interaction model. *Phys.Rev.D*, 83:123001, 2011.

[57] Thomas Eberl and Claudio Kopper. The SeaTray software framework. Antares Internal Note ANTARES-SOFT-2009-013, 2009.

[58] Claudio Kopper et al. A software framework for KM3NeT. In *Proceedings of the ICRC 2009 conference*, 2009.

[59] D. J. L. Bailey. Genhen v5r1: Software Documentation. Antares Internal Note ANTARES-SOFT-2002-004, 2002.

[60] A. L'Abbate et al. GENHEN v6: ANTARES neutrino generator extension to all neutrino flavors and inclusion of propagation through the earth. Antares Internal Note ANTARES-SOFT-2004-010, 2004.

[61] Alexander Kappes and Uli Katz. On the Use of Event Weights for the Simulation of Atmospheric Muons in ANTARES. Antares Internal Note ANTARES-SOFT-2005-003, 2005.

[62] A. Carminati, G. Margiotta and M. Spurio. MUPAGE user guide. Antares Internal Note ANTARES-SOFT-2007-004, 2007.

[63] P. Antonioli, C. Ghetti, E. V. Korolkova, V. A. Kudryavtsev, and G. Sartorelli. A three-dimensional code for muon propagation through the rock: Music. *Astropart. Phys.*, 7:357–368, 1997.

[64] M. De Jong. The TriggerEfficiency program. Antares Internal Note ANTARES-SOFT-2009-001, 2009.

[65] thinkastronomy.com. The Galactic Coordinate System. `http://www.thinkastronomy.com/M13/Manual/common/galactic_coords.html`, 2015. 2015-05-16.

[66] Infrared Processing and Analysis Center/Caltech & University of Massachusetts. The galactic plane. `http://upload.wikimedia.org/wikipedia/commons/e/e4/Milky_Way_infrared.jpg`, 2015. 2015-05-16.

[67] Stefan Geißelsöder. Classification of events for the ANTARES neutrino detector. Master's thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg.

[68] Thomas Wittenberg. Algorithms for Pattern Recognition and their Realisation 2. Lecture Slides, Selective Undergraduate Course, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2008.

[69] Lutz Priese and Patrick Sturm. Introduction to the Color Structure Code and its Implementation. In *International Conference on Image Processing, ICIP '04, (Volume:2)*, 2004. `http://www.uni-koblenz.de/~lb/lb_downloads/download/csc.pdf`.

[70] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*. John Wiley & Sons, 1998.

[71] Joachim Hornegger and Dietrich W. R. Paulus. *Pattern Recognition of Images and Speech in C++*. Verlag Vieweg, 1997.

[72] J. Han, M. Kamber, and J. Pei. *Data Mining*. Elsevier & Morgan Kaufmann, Third edition, 2012.

[73] Tin Kam Ho. The Random Subspace Method for Constructing Decision Forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:832–844, 1998.

[74] Vladimir Bystritsky and Sergey / The ALGLIB Project Bochkanov. ALGLIB. `http://www.alglib.net`, 2011. 2011-11.

[75] L. Breiman, J.H. Friedman, R.A. Olshen, and C.J. Stone. *Classification and Regression Trees.* Wadsworth, Belmont, CA, 1984.

[76] J. R. Quinlan. Induction of Decision Trees. *Mach. Learn. 1*, 1 (Mar. 1986):81–106, 1986.

[77] Laurent Hyafil and RL. Rivest. Constructing Optimal Binary Decision Trees is NP-complete. *Information Processing Letters*, 5:15–17, 1976.

[78] Kaggle. `http://www.kaggle.com`, 2015. 2015-03-28.

[79] Thomas Kittler. Erweiterung der Ereignisklassifizierung des ANTARES Neutrinoteleskops mittels Random Decision Forests. Bachelors thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg.

[80] J.A. Aguilar et al. A fast algorithm for muon track reconstruction and its application to the ANTARES neutrino Telescope. *Astroparticle Physics*, 34:652–662, 2011.

[81] Hans Joachim Bremermann. Optimization through evolution and recombination. 1962. in Self-Organizing Systems, M. C. Yovits, G. T. Jacobi, and G. D. Goldstein (eds), Spartan Books, Washington D. C.

[82] Raymond Chiong, editor. *Nature-Inspired Algorithms for Optimisation.* Springer, 2009.

[83] Thomas Kittler. Abschlussbericht zur Optimierung der Kennzahlen für den Random Decision Forest. Project report, Friedrich-Alexander Universität Erlangen-Nürnberg, 2014.

[84] Geoffrey E. Hinton, Simon Osindero, and Yee-Whye Teh. A fast learning algorithm for deep belief nets. *Neural Computation*, 18:1527–1554, 2006.

[85] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 25:1097–1105, 2012.

[86] Erwin Visser and Stefanie Wagner. Studies of the performance of different track reconstruction strategies using the FilteringFit track reconstruction. Antares Internal Note ANTARES-SOFT-2013-002, 2013.

[87] Florian Folger. Diffuse flux analysis using the Dusj Shower Reconstruction. Antares Internal Note ANTARES-PHYS-2013-012, 2013.

[88] Florian Folger. Updates on the Dusj diffuse flux analysis. Antares Internal Note ANTARES-PHYS-2013-014, 2013.

[89] Q. Dorosti Hasankiadeh and H. Lohner. Search for Showers Induced by Diffuse Flux of Cosmic Neutrinos. Antares Internal Note ANTARES-PHYS-2012-011, 2012.

[90] The ANTARES data quality working group. Data quality flags. http://antares.in2p3.fr/internal/dokuwiki/doku.php?id=quality_flags, 2015. 2015-05-16.

[91] S. Adrian-Martinez et al. Search for cosmic neutrino point sources with four year data of the ANTARES telescope. *Astrophysics J.*, 53:760, 2012.

[92] S. Adrian-Martinez et al. Search for muon neutrinos from gamma-ray bursts with the ANTARES neutrino telescope using 2008 to 2011 data. *A&A*, A9:559, 2013.

[93] S. Adrian-Martinez et al. Searches for Point-like and Extended Neutrino Sources Close to the Galactic Center Using the ANTARES Neutrino Telescope. *The Astrophysical Journal Letters*, L5:786, 2014.

[94] S. Adrian-Martinez et al. A search for neutrino emission from the Fermi bubbles with the ANTARES telescope. *Eur. Phys. J. C*, 74:2701, 2014.

[95] S. Adrian-Martinez et al. First results on dark matter annihilation in the Sun using the ANTARES neutrino telescope. *JCAP11*, 032, 2013.

[96] P. Bagley et al (KM3NeT Coll.). Technical Design Report. ISBN: 978-90-6488-033-9, Available from www.km3net.org, 2010.

[97] Kitware Inc. *VTK User's Guide*. Kiteware Inc., 2010.

[98] Kitware Inc. *The ParaView Guide (Full Color Version): A Parallel Visualization Application*. Kiteware Inc., 2015.

[99] Thomas Williams and Colin Kelley. gnuplot 4.6 An Interactive Plotting Program. http://www.gnuplot.info/docs_4.6/gnuplot.pdf, 2014. 2015-06-16.

[100] Bernd Jähne. *Digital Image Processing*. Springer, 2005.

[101] William K. Pratt. *Digital Image Processing: PIKS Scientific Inside*. John Wiley & Sons, 2007.

[102] Hans W. Schüßler, G. Dehner, R. Rabenstein, and P. Steffen. *Digitale Signalverarbeitung 2: Entwurf diskreter Systeme*. Springer, 2009.

[103] John G. Proakis and Dimitris K. Manolakis. *Digital Signal Processing*. Prentice Hall, 2006.

[104] The boost community. Boost C++ libraries. http://www.boost.org, 2015. 2015-03-28.

[105] M. G. Aartsen et al. (IceCube Collaboration). Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data. *Phys. Rev. Lett.*, 113:101101, 2014.

[106] R. Abbasi et al. (IceCube Collaboration). A Search for a Diffuse Flux of Astrophysical Muon Neutrinos with the IceCube 40-String Detector. *Phys. Rev. D*, 84:082001, 2011. arXiv:1104.5187, `http://icecube.wisc.edu/science/data/ic40`.

[107] S. Cecchini. Is there a seasonal temperature wave in the neutrino counting rate of ANTARES? ANTARES internal note ANTARES-PHYS-2011-007, 2011.

[108] Albrecht Karle (for the IceCube Collaboration). IceCube. In *Proceedings of the ICRC 2009 conference*, 2009.

[109] Trent Perrotto (NASA Headquarters) and Md.) Lynn Chandler (Goddard Space Flight Center, Greenbelt. NASA's Fermi Telescope Finds Giant Structure in our Galaxy. `http://www.nasa.gov/mission_pages/GLAST/news/new-structure.html`, 2015. 2015-08-18.

[110] Steffen Hallmann. Search for neutrinos from the Fermi Bubbles with the ANTARES telescope. In *Proceedings of the ICRC 2015 conference*, 2015.

[111] Erwin Lourens Visser. *Neutrinos from the Milky Way*. PhD thesis, Universiteit Leiden.

[112] David I. Jones, Roland M. Crocker, Wolfgang Reich, Jürgen Ott, and Felix A. Aharonian. Magnetic substructure in the Northern Fermi Bubble revealed by polarized microwave emission. *The Astrophysical Journal*, 747:L12 (4pp), 2012.

[113] Paul Murdin. *Encyclopedia of Astronomy and Astrophysics*. Nature Publishing Group, 2001.

[114] M. G. Aartsen et al. (IceCube Collaboration). Multipole analysis of IceCube data to search for dark matter accumulated in the Galactic halo. *European Physical Journal C*, 75:20, 2015.

[115] R. J. Egger and B. Aschenbach. Interaction of the Loop I supershell with the Local Hot Bubble. *Astronomy and Astrophysics*, 294(2):L25–L28, 1995.

[116] Olga Botner. IceCube Present and Future. In *The IceCube Particle Astrophysics Symposium*, 2015. https://events.icecube.wisc.edu/conferenceTimeTable.py?confId=68#20150504.