UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master Degree in Physics

Final Dissertation

Search for transient sources of high-energy neutrinos with IceCube: tests of the real-time analysis

Thesis supervisor

Prof. Elisa Bernardini

Thesis co-supervisor

Dr. Manuela Mallamaci

Candidate

Marco Bortolami

Academic Year 2019/2020

Contents

Abstract

The astrophysical sources of neutrinos with energy above GeV are to date unknown. A discovery of these objects would provide a smoking gun signature of the unknown cosmic rays sources, since their production is correlated. In addition, neutrinos are a unique tool to probe the most extreme environments in our Universe: their interaction cross section is low and they are electrically neutral, allowing them to travel undeflected and almost unabsorbed.

The IceCube detector, located at the South Pole, observes neutrinos in the energy range from TeV to PeV. Real-time analyses process the IceCube data in order to rapidly identify signal neutrinos, sending alerts to the astrophysical community for follow-up observations in case of interesting results. Two of them are relevant for this work, both searching for signal neutrino clusters from individual sources: one monitors known transient gamma-ray sources, the other performs an unbiased all-sky search.

This work focuses on testing the core algorithm of these real-time analyses. Neutrino clusters from point-like sources, characterised by number of signal neutrinos, flux spectral index and flare duration, are simulated for sources in different positions in the sky. The reconstruction quality of the cluster parameters is studied, highlighting the behaviour of the algorithm for different simulation conditions. The discovery potential, i.e. the flux that yields a 5σ discovery in 50% of the cases, is also evaluated for every simulated source. Finally, the possibility to identify sources emitting neutrinos during time scales shorter than few hours is discussed.

Chapter 1

Introduction

Multi-Messenger Astrophysics is a branch of Physics that studies the objects in the Universe connecting the information obtained from different messengers: cosmic rays, neutrinos, photons and gravitational waves. In Chapter 2, the general properties of the four messengers are discussed, as well as some detection techniques and source candidates.

Since the discovery of cosmic rays by V. Hess in 1912 [1], physicists have asked themselves about the origin of these extra-terrestrial particles, whose sources could open our eyes to the most extreme environments in our Universe. The accelerators built by humans can reach maximum energies far below the ones of cosmic accelerators. Cosmic rays are charged particles, therefore they are deflected by magnetic fields present in the Universe: this implies that the reconstructed arrival direction at our planet does not point back to their source. In addition, cosmic rays can be absorbed, thus they can reach us only if produced by near sources. We have not identified the sources of high-energy cosmic rays yet.

After being produced in cosmic accelerators, cosmic rays can interact with ambient hadrons present near the source yielding charged pions that decay (among other particles) to neutrinos or neutral pions that decay to photons. Identifying astrophysical sources of these particles provides a smoking gun signature of cosmic rays sources. Since they are electrically neutral, they travel undeflected from their source to our planet and by reconstructing their arrival direction we can infer their production site. The neutrino interaction cross section is very low if compared to the photon one: the photon flux at very high energies is greatly reduced with respect to the neutrino flux. Thus, neutrinos rather than photons can be used to study very high energy processes.

Up to now, by performing time-integrated searches, we have detected a diffuse neutrino flux [2, 3] but we have not discovered single high-energy ($\gtrsim$ 100 GeV) neutrino sources. Only two astrophysical neutrino sources have been identified: the Sun and the supernova SN1987A. However, neutrinos produced by these two objects are in the MeV range. Supernova SN1987A has been the first Multi-Messenger observation, since both photons and neutrinos produced by this source have been detected in 1987.

Time-dependent searches, optimised for transient sources, are now expected to provide better results and to reveal the first sources of high-energy astrophysical neutrinos. By selecting short time windows of several days, rather than using several years as for the time-integrated searches, the background of atmospheric particles is greatly reduced. In addition, observing the same source or process with more than one messenger increases the significance of the discovery. For example, particular processes as black holes mergers can produce all the four messengers, thus they are very interesting for multi-messenger studies. Therefore, it is important to develop real-time analyses able to quickly provide results in order to send alerts to the astrophysical community to inspire follow-up observations, i.e. real-time observations of the same event performed by different experiments.

The largest neutrino detector in operation is IceCube, discussed in detail in Chapter 3, that is located at the South Pole. It exploits a volume of deep polar ice of the order of 1 km^3 instrumented with photo-multiplier tubes. The optical sensors detect the Cherenkov light of charged particles produced in the neutrino interaction with the ice. Depending on the neutrino flavor and on the type of interaction (weak neutral or charged currents), the products of the neutrino interaction can leave different signatures. The best signal topology to search for point-like neutrino sources is a track-like signal, since it provides a very good angular resolution. When a muon neutrino undergoes charged weak interaction, it produces a muon that can traverse all the detector instrumented volume, seen as a track with a long lever-arm.

Astrophysical neutrinos are not the only particles that leave this type of signal in the detector. Neutrinos and muons produced by the cosmic rays interaction with the atmosphere are the main component of the background, having rates 3 and 9 orders of magnitude higher, respectively, with respect to astrophysical neutrinos. A filtering procedure selects muon track-like events, after which the data are processed to reduce the background and to estimate the direction and the energy of the primary neutrinos.

The sample of events passing the filtering and reconstruction procedure is an input for a real-time time-dependent analysis searching for neutrino sources, identifying clusters of astrophysical neutrinos. This is used, for example, to monitor known gamma-ray emitters as well as to perform an unbiased all-sky search. If a neutrino cluster is marked as an interesting event, alerts are sent to the astrophysical community and to the IceCube partner detectors, like the MAGIC Cherenkov telescope in the Canary Islands [4, 5], to perform Multi-Messenger follow-up observations. Chapter 4 discusses all the steps of the event filtering and reconstruction procedure and describes the real-time analyses.

My thesis is devoted to test the core algorithms of the real-time time-dependent analyses for the search of high-energy point-like astrophysical neutrino sources already deployed or soon to be deployed at the South Pole. They exploit a maximum likelihood algorithm which aims at providing the number of astrophysical neutrinos coming from individual sources, their spectral index and the duration of the source flare. Chapter 5 describes in detail the data sample features, the probability density functions used in the analysis, the likelihood maximisation and the results of the analysis by simulating different sources, e.g. a different flare duration or different neutrino fluxes. Studying the behaviour of the analysis in several conditions is very important to understand and interpret the features of the results. In addition, the sensitivity to transient sources is discussed. Before studying the time-dependent analysis, similar tests for an easier time-integrated version are performed for comparison and to better understand the different components of the analysis.

Chapter 6 summarises the conclusions obtained in this work and discusses the future perspectives, taking into account the improvement of the IceCube detector that will be performed in the following years.

Chapter 2

Multi-Messenger Astrophysics

In 1912, V. F. Hess discovered cosmic rays, i.e. extra-terrestrial charged particles striking our planet. He detected the ionizing radiation with an electroscope while being on a balloon, in order to change the altitude of each measurement. From a certain height and above, the ionization rate increased, suggesting an extra-terrestrial origin of the radiation [1]. This measurement can be thought as the beginning of Astroparticle Physics, a branch of Particle Physics that studies particles of astronomical origin and their relation to cosmology and astrophysics.

In particular, there are four different types of messengers that arrive from space and allow us to study the objects populating our Universe:

- Cosmic rays;
- Neutrinos (to which this work is dedicated);
- Photons;
- Gravitational waves.

Each messenger provides a complementary view of the sources and has unique properties. A sketch representing the four messengers and some examples of production mechanisms discussed in the following sections are reported in Figure 2.1.

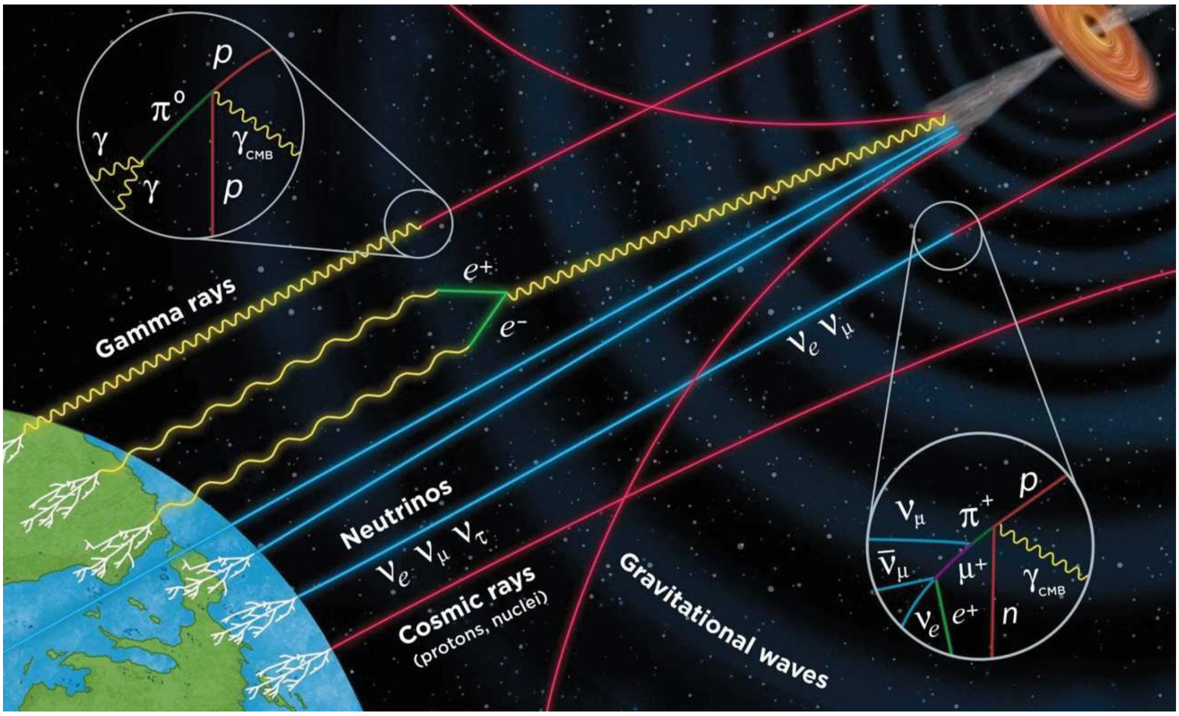


Figure 2.1: Not-to-scale sketch of observable particles striking our planet and some examples of production mechanisms. Taken from [6].

The information obtained from different messengers can be joined together, for example to increase the significance of the single observations and to provide more information on the underlying mechanism. All the messengers are studied together in the branch of Physics called Multi-Messenger Astrophysics.

In the following sections the properties and some examples of detection methods, experiments and potential sources are summarised for each messenger. Cosmic rays are described in Section 2.1, neutrinos in Section 2.2, photons in Section 2.3 and gravitational waves in Section 2.4.

2.1 Cosmic Rays

Primary Cosmic Rays (CRs) are charged particles striking our planet, like protons or heavier nuclei, whose energy varies between 10^8 eV and 10^{20} eV. Figure 2.2 shows the measurements of the cosmic rays flux provided by different experiments, as a function of the particle energy.

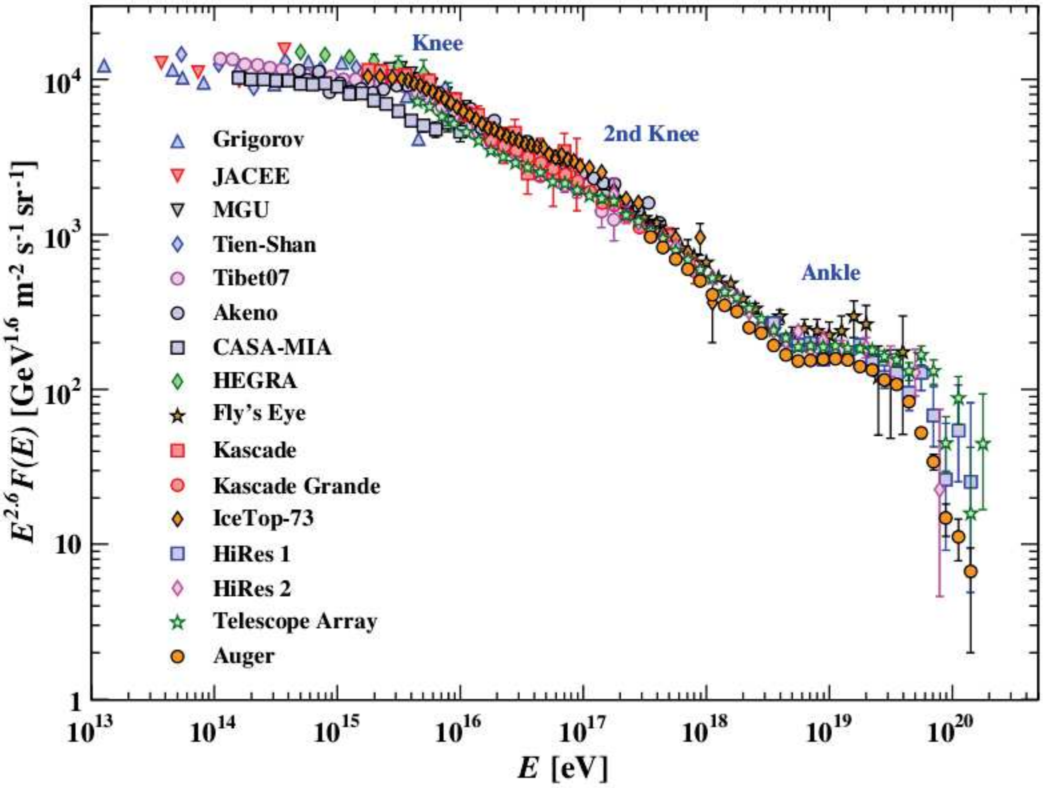


Figure 2.2: Cosmic rays flux spectrum measurements. The flux is multiplied by $E^{2.6}$ to underline its energy dependent features. The experiments names are reported on the left. The three changes of the flux spectral index are reported as blue words above the data points. Taken from [7].

The CR flux value spans several orders of magnitude. Nowadays human-made accelerators can reach maximum energies up to few TeV [8], far below the ones obtainable with astrophysical sources [9]. The flux ϕ is modeled with an unbroken power-law, defined by

$$\frac{d\phi}{dE} = \phi_0 \cdot \left(\frac{E}{E_0}\right)^{\gamma}, \tag{2.1}$$

where ϕ_0 is the flux normalisation at the reference energy E_0 (typically 1 GeV) and γ is called "spectral index". The unbroken power-law model is used for the fluxes of neutrinos in this work, too.

In Figure 2.2 the flux is multiplied by $E^{2.6}$ in order to highlight energy dependent features of the results. Before the so-called "Knee" around 3 PeV, the spectral index is approximately -2.7 and the CR are supposed to originate mainly from inside our galaxy. After the Knee but before the second Knee the spectral index changes to $\approx$ -3.1, the particle origin is thought to be extra-galactic (as, thanks to the higher energy, the charged particles can escape the galactic magnetic field) and the cosmic rays composition tends towards heavier nuclei. After the second Knee the spectral index moves to $\approx$ -3.3. At the "Ankle" around 5 EeV the spectral index is approximately -2.5 [10].

Cosmic rays with an energy above 1 EeV are called Ultra-High Energy Cosmic Rays (UHECRs). Above 100 EeV a strong suppression of the flux is evident: its origin is one of the most important open questions regarding cosmic rays. We still do not know whether is due to the fact that cosmic rays accelerators can not exceed this energy or to the Greisen-Zatespin-Kizmin (GZK) cut-off [11, 12]. The latter consists in the interaction of UHECRs with a photon from the Cosmic Microwave Background (CMB) producing a Δ^+ resonance that suddenly decays. The processes are the following:

$$\begin{aligned} p + \gamma_{CMB} \rightarrow \Delta^+ &\rightarrow p + \pi^0 \\ &\rightarrow n + \pi^+ \end{aligned} \tag{2.2}$$

The protons produced by the Δ^+ decay can repeatedly undergo this process until their energy is below the threshold of $5 \cdot 10^{19}$ eV for this process to happen.

Primary cosmic rays, i.e. the ones striking our planet, are mainly composed of hadrons (98%), while the remaining 2% is made up of leptons (mainly electrons and, to a lesser extent, positrons). The hadronic component is made up of protons (88%), α particles (11%) and heavier nuclei (Z > 2). A small fraction consists of anti-protons.

The interaction of the primary CRs with the nuclei of the atmosphere produces secondary hadronic particle showers composed primarily of protons, neutrons, pions, kaons and low Z nuclei. In particular, neutral pions decay with a branching ratio $\approx$ 100% via

$$\pi^0 \rightarrow \gamma + \gamma \tag{2.3}$$

and each photon can initiate an electromagnetic shower. Charged pions and kaons mainly decay into muons and neutrinos via these two decay channels:

$$\pi^-/K^- \rightarrow \mu^- + \bar{\nu}_\mu \tag{2.4}$$

$$\pi^+/K^+ \rightarrow \mu^+ + \nu_\mu \tag{2.5}$$

Charged kaons can also decay into pions. (Anti)muons then decay via:

$$\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu \tag{2.6}$$

$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu \tag{2.7}$$

Thus, neutrinos are produced in cosmic rays induced atmospheric showers [13]. **Atmospheric neutrinos and muons are the main component of the irreducible background for the analyses treated in this work, that aims at identifying astrophysical neutrinos**, as explained in Chapter 4. The aforementioned processes are sketched in Figure 2.3.

There exist several methods to detect cosmic rays, each one efficient in specific energy ranges. The flux around the Knee is $\approx 1 \frac{\text{particle}}{\text{m}^2\,\text{s}}$ and it can be directly observed with satellites: an example was the PAMELA detector mounted on the Resurs-DK1 satellite [15]. The flux is reduced to $\approx 1 \frac{\text{particle}}{\text{km}^2\,\text{century}}$ at the highest energies, thus large-area ground-based detectors are used to detect the secondary showers produced by the primary cosmic rays interaction with the atmosphere. The Pierre Auger observatory in Argentina, for example, exploits a surface detector array covering about 3000 km^2 with water Cherenkov tanks and a fluorescence detector to perform this kind of observation [16].

Cosmic rays are produced in the Sun and galactic SuperNova Remnants (SNRs) have been proposed in 1934 as candidate sources of cosmic rays [17]. Neutral pion decay signatures have been found for two SNRs, confirming that they contribute to galactic cosmic rays [18]. In addition, the Pierre Auger collaboration found an anisotropy in the cosmic rays arrival direction suggesting an extragalactic origin [19]. Finally, UHECRs are subject to negligible deflections due to their high energy and momentum, so the source position could in principle be estimated. Measurements of the Pierre Auger observatory suggested a correlation between UHECRs and active galactic nuclei [20]. However, the sources of the highest energy primaries are still unknown.

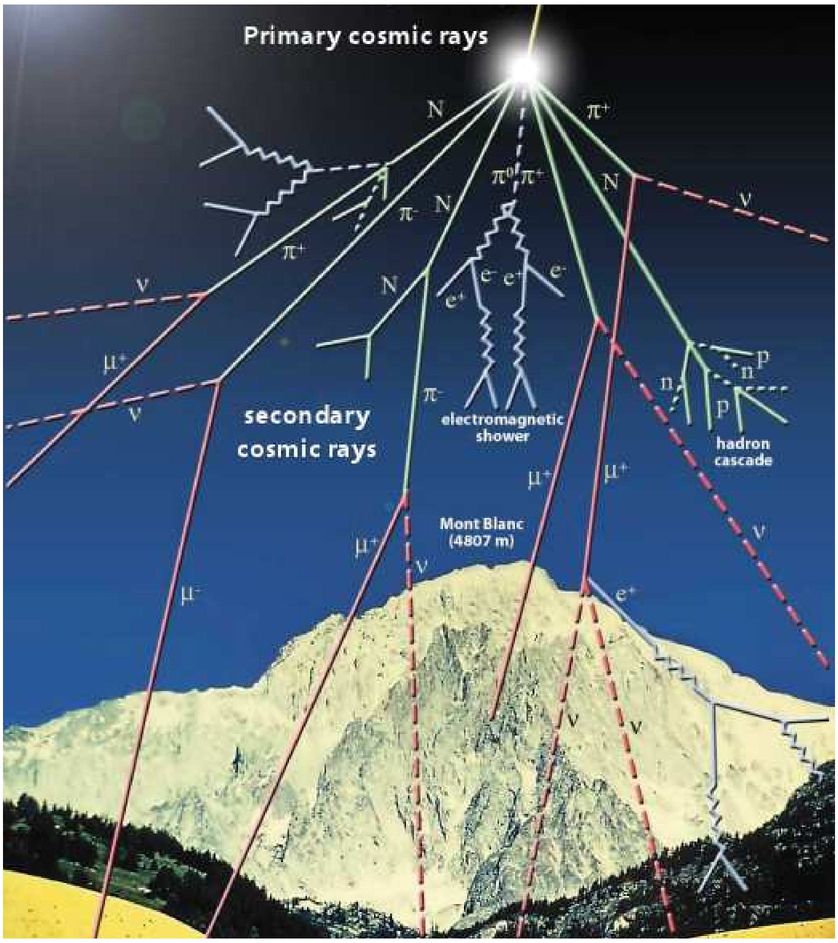


Figure 2.3: Secondary cosmic rays production for a shower initiated by a hadronic primary. Taken from [14].

Fermi developed two models for cosmic ray acceleration in supernovae [21]. The first-order Fermi acceleration requires moving magnetic shock fronts, at whose boundaries they resembles a plane wave. Charged particles with energy E_0 in front of the shock that moves at velocity β are scattered by turbulent magnetic fields and can exit the shock with a higher energy $E > E_0$. The energy spectrum is expected to follow a power-law with spectral index -2, but the interactions between the acceleration site and the Earth soften (i.e. increase in modulus) the spectral index. The second-order Fermi acceleration can take place at the boundaries of moving, magnetized interstellar clouds, for which the shock front can not be approximated with a plane wave. Particles are scattered randomly, providing a lower energy gain.

Depending on the strength B and the radius R of the accelerating magnetic field, a particle with charge z can be accelerated up to the energy

$$E_{max} = z\, B\, R. \tag{2.8}$$

The maximum attainable energy increases with larger accelerating regions and with stronger magnetic fields, and it is higher for particles with higher charge. Figure 2.4 shows the Hillas diagram, i.e. the possible accelerating sites depending on the magnetic field strength and on their size.
Usually the astrophysical sources can emit more than one type of messenger. For example, Active Galactic Nuclei (AGN) are supermassive black holes ($\mathrm{M_{BH}} = 10^6 - 10^{10}\mathrm{M}_\odot$) around which matter is accreted and by falling into the black hole it emits photons over the entire electromagnetic spectrum. Matter forms an accretion disk thanks to the angular momentum conservation and can produce two particle jets perpendicular to the disk. The accelerated protons in the jets can produce different messengers by interacting with other particles. For example, they can produce pions that decays into photons and neutrinos. Figure 2.5 shows a sketch of an AGN. If the jets are pointed to us, the AGN is called blazar.

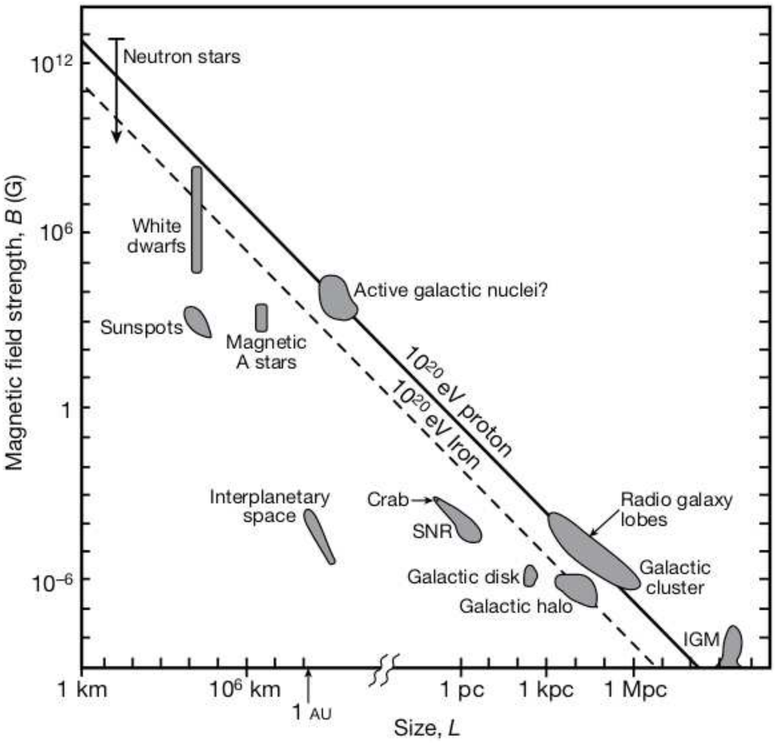


Figure 2.4: Hillas diagram showing the possible accelerating sites depending on the magnetic field strength B and on the size L. The solid (dashed) line shows the GZK cut-off for protons (Iron). (SNR = SuperNova Remnants, IGM = InterGalactic Magnetic field). Taken from [22].

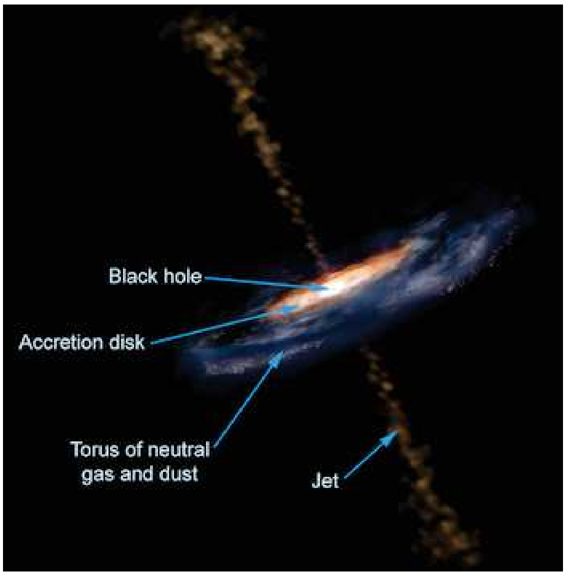


Figure 2.5: Sketch of an active galactic nucleus [23].

2.2 Neutrinos

Neutrinos are fundamental particles of Nature. They are electrically neutral leptons with spin 1/2 and they have three flavors: electron neutrino ν_e, muon neutrino ν_μ and tau neutrino ν_τ. The Standard Model of Particle Physics predicts the neutrino mass equal to zero. However, it was discovered that they can change flavor while propagating from their production site to the detector [24]. Flavor oscillations can only happen if neutrinos are massive and their flavor eigenstates (ν_e, ν_μ, ν_τ) are not equal to their mass eigenstates (ν_1, ν_2, ν_3). Therefore, studying the neutrino oscillation can unveil aspects of the so-called Beyond Standard Model Physics, that studies possible generalisations, extensions or modifications of the Standard Model to include phenomena without an explanation within the present version. The differences between the squares of the neutrino masses are $m_2^2 - m_1^2 \approx 7.59 \cdot 10^{-5}$ eV^2 and $|m_3^2 - m_2^2| \approx 2.43 \cdot 10^{-3}$ eV^2, where m_1, m_2, m_3 are the masses connected to the mass eigenstate ν_1, ν_2, ν_3, respectively [25]. However, we have only upper bounds on the absolute mass scale ($m_\nu \lesssim 0.2$ eV [26]).

Since neutrinos are leptons they do not interact via the strong interaction and because of their zero electrical charge they do not interact via the electromagnetic interaction, thus they are not deflected by magnetic fields. This is very important in Astroparticle Physics, because, as opposed to cosmic rays, the arrival direction of neutrinos at our planet points directly to their production site. Therefore, astrophysical neutrino sources can be identified by reconstructing the direction of the detected neutrinos. Neutrinos can interact via the weak interaction exchanging W$^+$, W$^-$ or Z^0 bosons with other particles, like charged leptons or quarks.

The neutrino interaction cross section in very low if compared to the strong and electromagnetic forces (typical values are 10^{-44} cm^2 [27]). Thanks to this property neutrinos are a unique tool to study the Universe, but they are also difficult to detect. In fact, on one hand, when an astrophysical source produces neutrinos they travel undeflected and almost unabsorbed, preserving the information on the flux. On the other hand they are likely not to interact even in the detectors. For this reason, special techniques have been developed to observe neutrinos.

Neutrinos in the MeV range can be detected in underground radiochemical detectors, like Super-Kamiokande in Japan [28]. However, at higher energies, the low interaction cross section connected to the reduction of the power-law flux for high-energy neutrinos results in the need of a large scale detector. It is cost effective to instrument naturally occurring detection mediums such as big deposits of water or ice like lakes, deep oceans or the polar ice. The principle of operation is the following: high-energy neutrinos, when interacting with the medium, yield charged particles that move faster than the phase velocity of light in that medium, resulting in the production of Cherenkov photons [29] detected by optical sensors. Water and ice in the aforementioned environments have large scattering and absorption lengths, which permits to collect enough photons for a precise reconstruction of the charged particles energy and direction.

This principle of operation has been exploited by different detectors. DUMAND [30] was a project for deploying a string of 7 optical detectors in the Pacific Ocean that was used to detect the Cherenkov light of atmospheric muons at ocean depths ranging from 2000 to 4000 meters. The Baikal [31] experiment, in lake Baikal (Russia), started operating in 1993 with 36 photomultipliers divided in 3 strings and it provided the first three-dimensional mapping of the deep underwater Cherenkov light. ANTARES [32] is a neutrino telescope in the Mediterranean sea off the coast of Toulon (France) at 2475 m depth composed of 12 strings holding light sensors, that provides opportunities for innovative measurements also in Earth and sea sciences. The largest neutrino telescope is the IceCube detector, located at the South Pole. It exploits a volume of ice of the order of 1 km^3 instrumented with strings of photomultipliers. Earth sciences benefited also from this experiments, since it provided the possibility to study the deep polar ice. The IceCube detector is discussed in detail in Chapter 3.

The efforts to observe astrophysical neutrinos are well justified for two main reasons. Firstly, the Universe at the highest energies is opaque to photons but not to neutrinos, as shown in Figure 2.6: neutrinos are a unique tool to observe the most extreme environments in our Universe.
Secondly, unveiling sources of cosmic neutrinos could help solving the question about the sources of cosmic rays. This is possible as astrophysical neutrinos are a smoking gun signature of cosmic rays production. In fact, high energy neutrinos can be produced by charged pions and charged kaons decays (Equations 2.4 - 2.7), where the mesons are produced by cosmic rays interacting with a photon in the photohadronic channel

$$\begin{aligned} p + \gamma \to \Delta^+ &\to p + \pi^0 \\ &\to n + \pi^+ \end{aligned} \tag{2.9}$$

or with a nucleus N in the hadronuclear channel

$$p + N \to \pi^{0,\pm} + K^{0,\pm} + ... \tag{2.10}$$

Thus, a source of protons surrounded by photon fields or matter can produce neutrinos. At the

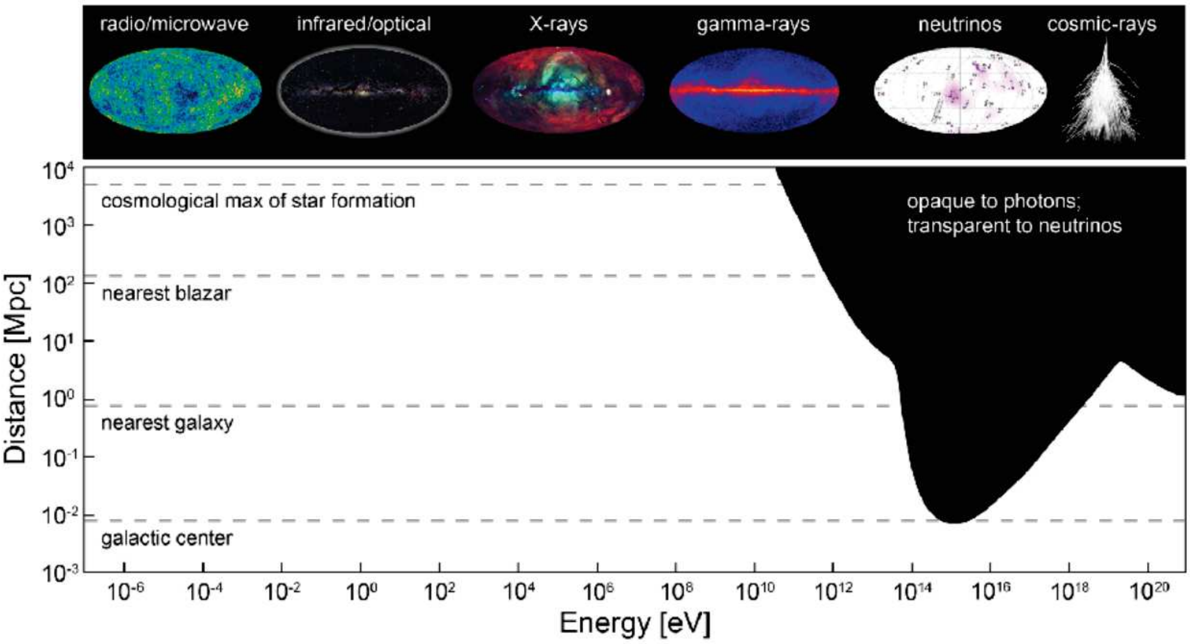


Figure 2.6: Energy and wavelength spectra versus distance of the visible Universe. The region where the Universe is opaque to photons is shown in black: about a fifth of the universe cannot be explored using photon-based telescopes. Taken from [33].

production site the flavor ratio is generally predicted to be $(\nu_e : \nu_\mu : \nu_\tau) = (1 : 2 : 0)$, but oscillations yield a ratio at the Earth of $(1 : 1 : 1)$.

Neutrinos in the MeV range coming from the Sun [27] and from the supernova SN1987A [34], that is the first multi-messenger observed source as it was also detected with photons, have been detected. These are the only two known sources of extra-terrestrial neutrinos. Several hundreds astrophysical neutrinos are expected per year and some could be emitted by the same source as a cluster of neutrinos during a flare. Searching neutrino clusters permits to disentangle the signal (astrophysical neutrinos) from the irreducible background (muons and neutrinos from atmospheric cosmic rays induced showers), and to yield a lower error on the direction of the observed source.

Up to now, after 10 years of operation of the IceCube detector, a diffuse flux of astrophysical neutrinos in the energy range from TeV to PeV has been observed [2, 3]. Sources of high-energy astrophysical neutrinos have not been identified with time-integrated searches yet. A correlation was found between a high energetic neutrino ($E_\nu = 290$ TeV) and the blazar TXS 0506+056, showing a contemporaneous increase of the electromagnetic activity, with a significance of $\approx 3\sigma$ [35]. Time-dependent approaches are now expected to provide enough significance to claim a discovery. On short time scales the atmospheric background is highly reduced, but the significance increases also thanks to possible correlated emission of gamma-rays from the same source. A real-time infrastructure (Chapter 4) able to perform a fast online analysis to find clusters of astrophysical neutrinos and to send alerts to the astrophysical community has been developed within the IceCube collaboration [10]. This work focuses on the tests of its core algorithm for different source features, as discussed in Chapter 5.

2.3 Photons

Photons are massless bosons, carriers of the electromagnetic force. They have been the first messenger to be detected, since the very beginning of humanity, when we started looking at the sky. They are electrically neutral, therefore the arrival direction at our planet points directly to the source position.

The photon interaction cross section is much higher with respect to neutrinos. The Universe can not be studied exploiting photons of the highest energies, as already mentioned in Section 2.2. In particular, high-energy photons interact with the Extragalactic Background Light (EBL), that is the product of all the stellar emission in the Universe from infrared to ultraviolet, producing an e^-e^+ pair. This process significantly reduces their flux impeding the observation of distant high-energy photons sources [36]. Neutrinos can be used to explore the region where the Universe is opaque to photons.

Different experimental techniques are used to detect photons depending on the energy range, like telescopes on board of satellites or ground-based telescopes. For example, in the high-energy regime (50 MeV - 300 GeV), where the photons would be absorbed by the atmosphere, satellites like the Fermi-LAT are used to directly detect photons via pair production inside the detector [37]. Instead, due to the low flux in the Very High-Energy (VHE) regime ($E_\gamma > 50$ GeV) that requires higher effective areas, ground-based Imaging Air Cherenkov Telescopes (IACTs) exploit the Cherenkov radiation produced by electrons and positrons of the photon-induced electromagnetic showers to estimated energy and direction of the primary photons interacting with the atmosphere. The principles of operation of the IACTs are sketched in Figure 2.7. Some examples are the MAGIC telescope [4, 5] and FACT [38] in the Canary islands, VERITAS in Arizona [39] and HESS in Namibia [40].

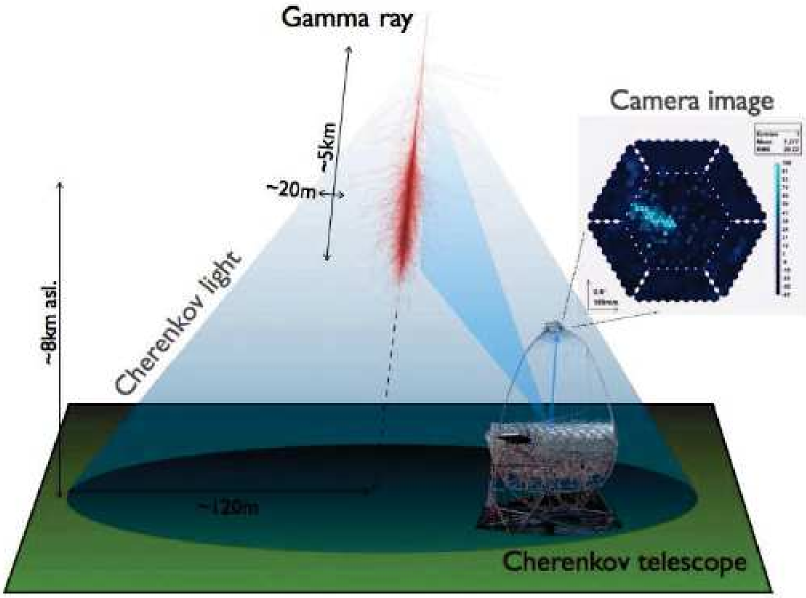


Figure 2.7: Principle of operation of the Imaging Air Cherenkov Telescopes. A gamma-ray primary interacts with the atmosphere producing an electromagnetic particle shower. The charged particles in the shower move faster than light in air, causing the emission of Cherenkov photons. Mirrors reflect the Cherenkov light onto a camera composed by photomultiplier tubes for the actual detection. Taken from [41].

Fermi-LAT can observe the entire sky in 3 hours, while IACTs have a small field of view of 4 $\deg^2$ and require dark conditions to observe. Multi-messenger observation with IACTs benefit from external alerts that can inspire follow-up observations. For example, real-time neutrino analyses (like the ones considered in this thesis) can alert partner IACTs so they can repoint the telescope to the neutrino arrival direction. This happened for the aforementioned multi-messenger observation of the blazar TXS 0506+056 [35]. Cosmic rays can interact with matter surrounding the AGN, producing charged pions decaying to neutrinos and neutral pions decaying in gamma rays: joining the information of these two messengers is very important to test models of cosmic rays production.

Blazars (AGN with a jet directed towards our planet) have shown a highly variable behaviour in gamma-ray emission. Other variable processes are Gamma Ray Bursts (GRBs), connected to supernovae [42] or neutron star mergers [43], that release up to 10^{54} erg in gamma rays on time scales from seconds to minutes. Again, to observe variable and transient sources, especially for short time scales, an external alert is of great help.

2.4 Gravitational Waves

Gravitational waves are ripples in space-time that travel at the speed of light, predicted by Albert Einstein in 1916 as a consequence of the General Relativity theory [44]. Every asymmetric motion of masses can produce gravitational waves, but nowadays we can detect the ones produced by huge masses in acceleration, like black holes.

The discovery of the binary pulsar system PSR B1913+16 in 1975 [45] and the subsequent observations of its energy loss in 1982 [46] demonstrated the existence of gravitational waves. They were discovered in 2015 together with the first observation of a two black holes merger [47].

The most sensitive gravitational wave detectors are Ligo in America [48] and Virgo in Italy [49]. They exploit laser interferometers, whose arms are 4 km long for Ligo and 3 km long for Virgo. When a gravitational wave passes through the detector, it stretches the two perpendicular arms in a different way and the interference of the light can be detected, revealing the passage of the wave.

When two astrophysical objects like two neutron stars or two black holes rapidly rotate one around the other gravitational waves are emitted and the orbit decreases. This stage is shown in Figure 2.8. The two objects merge once they are close enough. The single object then settles down to a stable form, where any distortion in the shape is dissipated as more gravitational waves. After producing gravitational waves, these events are expected to emit other messengers, like neutrinos and photons [50]. However, while neutrinos are expected to be produced in the interactions between relativistic protons and the external radiation fields of the source, gravitational waves carry information on the multi-dimensional dynamics in the source's central regions. Neutrinos and gravitational waves are thus complementary messengers [50].

For this motivation, the gravitational waves observatories can alert other collaborations to inspire follow-up observations (in the case of IACTs) or to trigger the analyses with a signal hypothesis from the arrival direction of the gravitational wave (in the case of IceCube).

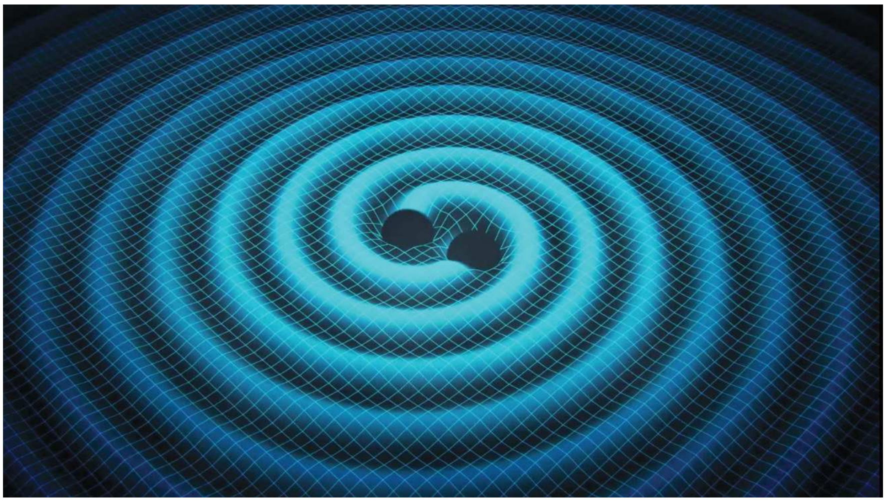


Figure 2.8: Artist's impression of gravitational waves produced by a binary system of two black holes rotating around each other. The grid represents the space-time, curved by the waves. Taken from [51].

Chapter 3

Observing neutrinos in the ice: the IceCube Detector

As anticipated in Section 2.2, in order to observe high-energy neutrinos a large scale detector is needed. The largest is the IceCube detector, that aims at observing neutrinos in the energy range from TeV to PeV exploiting approximately 1 km^3 of instrumented ice. Charged particles produced in the interaction of neutrinos with the polar ice can be faster than light in this medium: if this is the case, they yield Cherenkov photons that are detected by optical sensors. This is the principle of operation of the IceCube detector.

The deep Antarctic glacier is one of the best environments for Cherenkov light detection, both natural and human-made. The pressure of the glacier removes the air bubbles from the deep ice, reducing the scattering and absorption processes of the photons. In situ measurements demonstrated that the ice present in the South Pole is even clearer than the ice produced in laboratory [52]. However, the Cherenkov photons still undergo scattering and absorption processes worsening the detection efficiency and the reconstruction accuracy. The scattering is mainly due to dust particles present in the ice, like minerals grains and sea salt crystals, while the absorption is mainly caused by impurities in the form of insoluble dust [53]. A detailed study of the deep ice is fundamental to develop proper models of this medium, used for example to reconstruct the arrival direction of neutrinos (Section 4.4). In Figure 3.1 the scattering and absorption coefficients (the inverse of their respective mean free paths) are shown as a function of the depth from the surface and the photon wavelength.

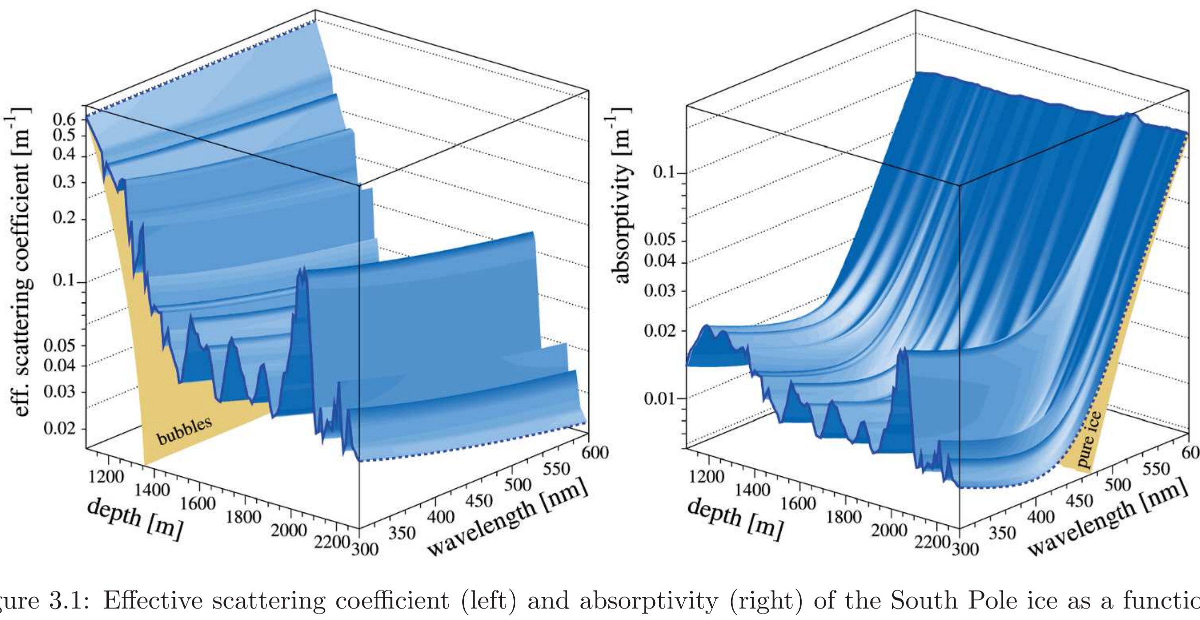


Figure 3.1: Effective scattering coefficient (left) and absorptivity (right) of the South Pole ice as a function of the depth from the surface and the photon wavelength. The yellow surfaces refers to the contribution of air bubbles to scattering and pure ice to absorption. The blue surfaces take into account all the impurities. Taken from [53].

Starting from 1400 m and deeper, where the IceCube optical sensors are located, scattering and absorption are low, a part for some peaks due to dust concentrations formed thousands of years ago. For example, the most prominent peak is due to ≈ 65 000 years old dust [54].

The knowledge about the South Pole ice and its models have seen great improvements during the years. This was possible thanks to many efforts, such as calibration with light sources of the IceCube predecessor AMANDA [55], ice measurements in different places of Antarctica [56, 57] and calibration using the LEDs present in the IceCube sensors [58]. The most updated models of the ice include many of its features, like the depth-layered structures, the effect of the refreezing of the holes created to insert the instrumentation in the deep ice and the anisotropy [59].

This chapter describes the IceCube detector. In Section 3.1 the detector hardware is described, as well as the efforts carried out to instrument the polar ice. Then, the interactions of neutrinos with the ice and the different event signatures left by the particles are discussed in Section 3.2, together with the specification of signal and background events for the analyses considered in this work and their simulation. Finally, the detector monitoring system is summarised in Section 3.3.

3.1 Construction and Hardware

The IceCube neutrino observatory is located at the geographic South Pole. The detector consists of 86 vertical strings deployed in the antarctic glacier. Each string has 60 vertically aligned Digital Optical Modules (DOMs), for a total of 5160 DOMs present between 1450 and 2450 m of depth, and a cable that connects the optical sensors to the IceCube counting house “IceCube Lab” on the surface. 86 boreholes approximately 60 cm in diameter and 2500 m in depth were created by a hot water drilling procedure. The 5 MW Enhanced Hot Water Drill produced one hole every 48 hours by pumping 760 liters/min of water at a temperature of 88°C and a pressure of 7600 kPa. The instrumentation was deployed into the water-filled boreholes before the ice was reformed. The string structure is designed to whitstand the refreezing process of the water, in order to maintain the vertical alignment of the DOMs [60].

A schematic layout of the detector is shown in Figure 3.2.

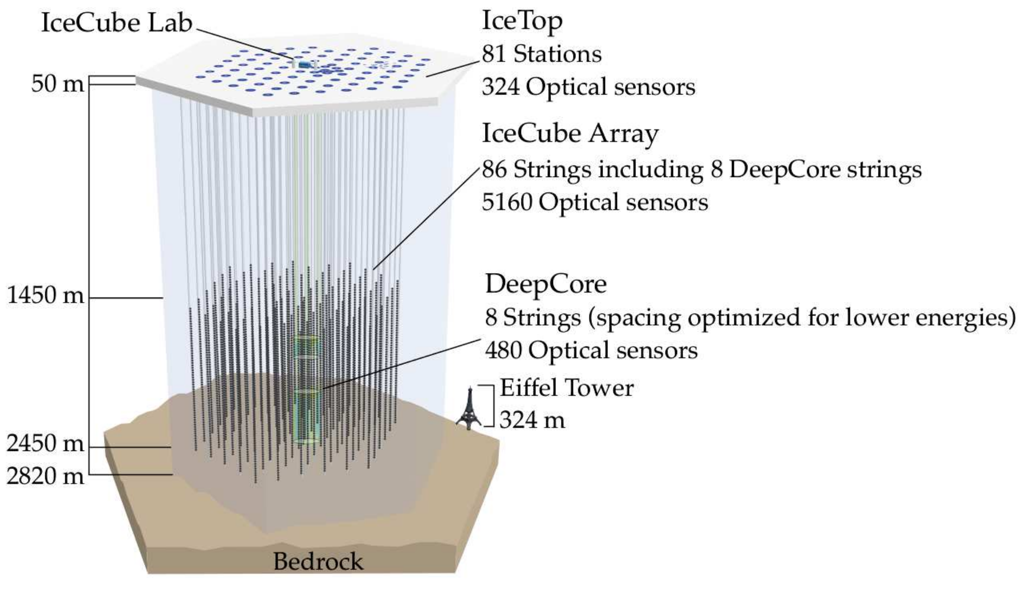
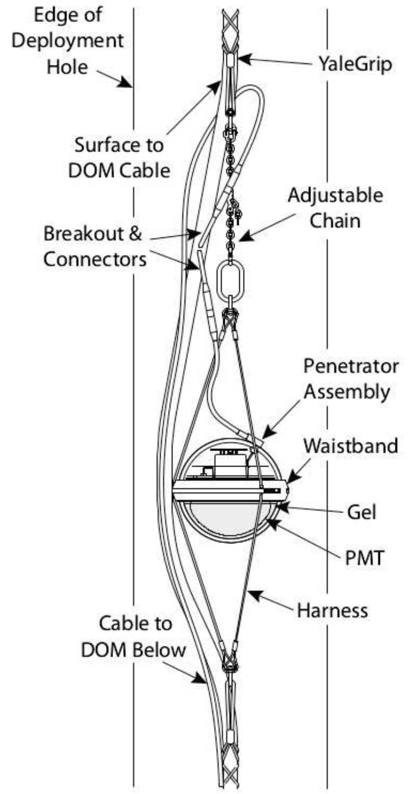


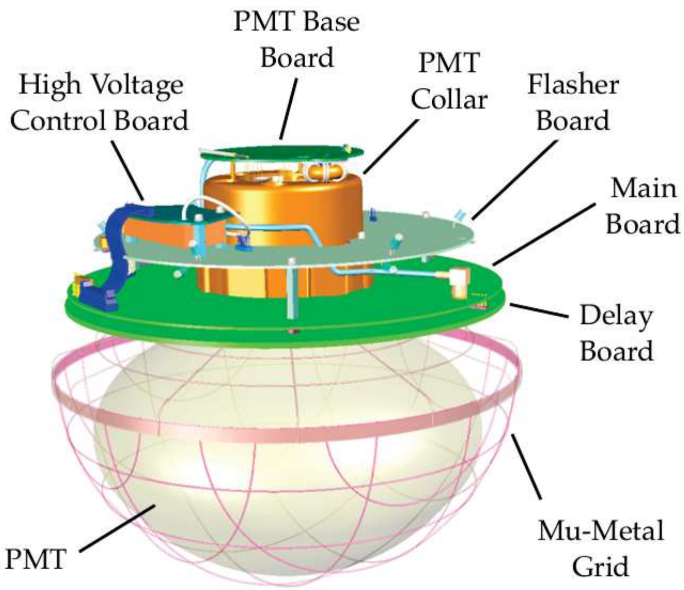
Figure 3.2: Schematic view of the IceCube detector. On the left, the depth from the ice surface is reported (the Eiffel Tower is shown on the right for comparison). Each vertical line represents a string deployed in the ice and the dots between 1450 and 2450 m show the DOMs, which instrument ≈ 1 km^3 of ice. The position of the IceCube Lab is highlighted, as well as IceTop, the IceCube main Array and DeepCore with the numbers of total stations/strings and optical sensors. Taken from [10], adapted from [60].

The primary detector is made of 78 strings arranged in a hexagonal grid with 125 m spacing between the strings and 17 m between two consequent DOMs on the same string. The energy threshold of this configuration is around 100 GeV. 8 strings have been deployed with a lower distance among themselves, between 41 m and 105 m, and the DOMs of these strings are placed between 7 and 10 m apart. This sub-detector is called DeepCore and it has a lower energy threshold of 10 GeV. On the ice surface is located IceTop, an array of 162 ice-filled tanks each one containing two DOMs, sensitive to cosmic-rays induced atmospheric showers in the energy range from PeV to EeV.

Each DOM consists of several components contained in a glass sphere with a diameter of 13" and a thickness of 0.5", sketched in Figure 3.3(b). The sphere protects the electronics and PMT from the long-term pressure of 250 bar, as well as from the temporary overpressure up to 690 bar during the refreezing of melted ice in the drill hole [60]. The optical sensor is a downward-facing 10" PhotoMultiplier Tube (PMT) secured in a high-strength silicone gel to a depth sufficient to surround the photocathode area, that provides both good optical coupling and mechanical support. The PMT is connected to circuit boards used for power supply, data acquisition, communication and calibration. The maximum supply voltage for the PMTs is 2047 V. The cable leaves the sphere through a 16 mm hole and vertically follows the string arriving at the surface. The DOM suspension and cable routing can be seen in Figure 3.3(a).



(a) DOM suspension and cable routing.



(b) Sketch of a DOM.

Figure 3.3: Left: the DOM suspension and cable routing. Right: the DOM components present inside the glass sphere (not shown). Taken from [60].

The PMT is sensitive to photons with wavelength between 300 nm and 650 nm and has a quantum efficiency of 25% at 390 nm (34% for the DeepCore modules). The collection efficiency and single photoelectron resolution of the PMT would be degraded by the South Pole magnetic field, as PMTs are sensitive to these fields. This inconvenient is solved by placing a wire mesh grid around the PMT that acts as a shield. The calibration is performed using the 12 LEDs present in the Flasher Board, that can emit light with wavelengths between 340 nm and 505 nm with a controlled light output [60]. The light produced by a DOM is detected by the others and the responses are compared to the expectations.

The PMT waveforms can be very different, depending on the energy of the particles and the distance between the particle and the DOM of interest. The amplitude of the signal has minimum values of 1 mV, but it can reach voltages up to 2 V. The width of the signals has values between 12 ns and 1500 ns (the detector time resolution is 2 ns). Not all the waveforms are recorded: noise signals are removed by a discriminator that triggers the recording only when a threshold corresponding to 0.25 photoelectrons is crossed.

For the waveforms passing this discriminator the recording starts 75 ns before the trigger. This is possible thanks to the Delay Board placed above the PMT. For a single photon the waveform is available until 6.4 μs after the discriminator trigger. The PMT signal is captured and digitised by Analog Transient Waveform Digitizers (ATWDs) and a fast Analog to Digital Converter (fADC) [61]. The ATWD collects the first 427 ns, samples at 300 Msps and has a 10-bit resolution. The fADC is continuously sampling the waveforms, thus it has a lower rate of 40 Msps, but the same 10-bit resolution.

The digitised waveform is proportional to the charge collected by the PMT. For IceCube, 1 photoelectron (the charge deposited by a single photon) corresponds to ≈ 1.6 pC. The outputs of the ATWD and fADC modules are fitted with non-negative least squares using templates of single photon pulses produced in laboratory or during in-ice studies [62]. The fit outputs a series of pulses defined by time, charge and width that is daily transmitted by satellites to the IceCube computing center in Madison, WI. In general the full digitized waveform is not included. If a goodness-of-fit test marks the least squares fitting as "unsuccessful" the full digitized waveform is sent for offline studies.

3.2 Neutrino Interactions and Event Topologies

IceCube is sensitive to neutrinos undergoing deep inelastic scattering with an atomic nucleus. The interaction does not need to happen inside the detector, as charged particles produced outside can still be detected. For example, muons produced kilometers away from the detector but passing through the instrumented volume are visible.

As discussed in Section 2.2, neutrinos can undergo charged or neutral current weak interactions. Neutral current interactions are represented by

$$\nu_l + X \xrightarrow{Z} \nu_l + X', \tag{3.1}$$

where ν_l is a neutrino related to lepton l=(e, μ, τ), X is the nucleus before the interaction and X' are the remnants of the nucleus after the interaction. The neutrino produced by the interaction is undetected because it is electrically neutral, while the remnants of the nucleus yield hadronic particle showers containing charged particles, that cause the emission of Cherenkov light.

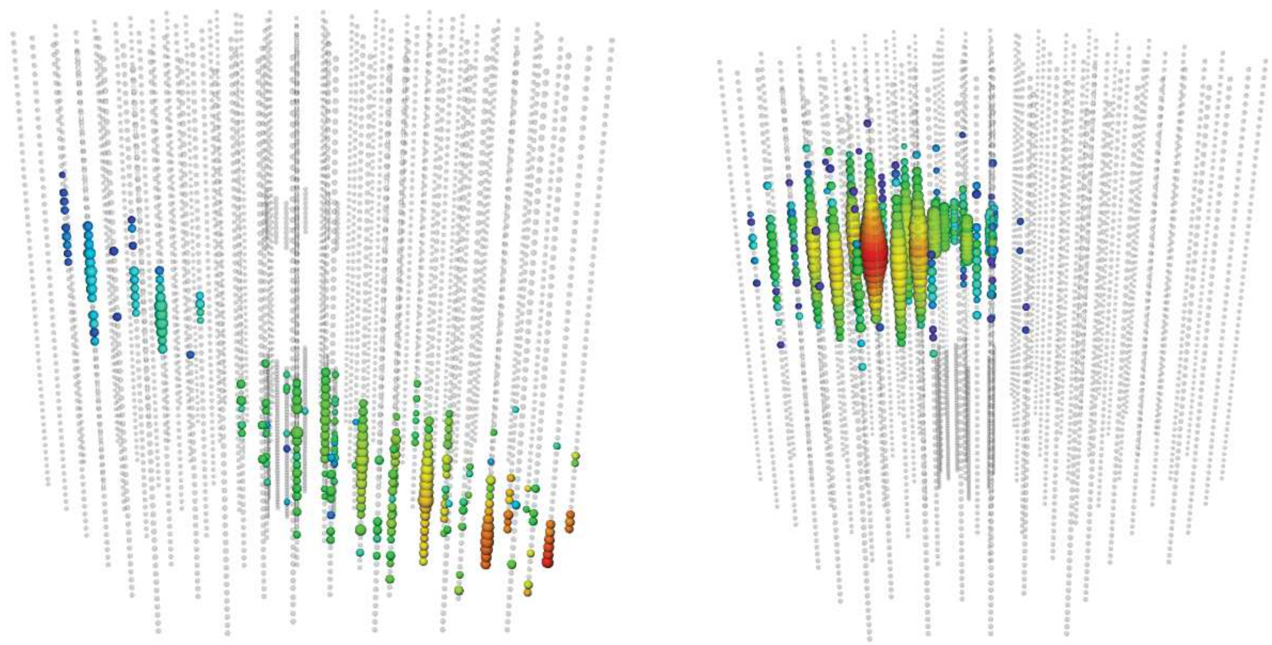
Charged current interactions can be summarised by

$$\nu_l + X \xrightarrow{W} l + X'. \tag{3.2}$$

While the remnants X' still produce a hadronic shower, the other product l can be detected since it is a charged lepton. An electron or tau neutrino produces an electromagnetic shower, whereas a muon above GeV energies travels enough to leave a track-like signature in the detector. The tau lepton has a median range of $\approx$ 50m/PeV, so it can be identified if it has an energy of at least several PeV. In addition to neutrinos, also muons from cosmic rays induced showers produced above the detector can reach IceCube and leave a track-like signal.

The electromagnetic and hadronic showers produced in the neutrino interactions with the deep ice of the South Pole have length of several meters. Compared to the distance between the DOMs, the showers are almost point-like. For this, the Cherenkov light produced by their decay products forms a spherical light pattern. This topology, called shower- or cascade-like, is optimal for energy measurements, that have an uncertainty of $\approx 10\%$ [63]. However, the directional reconstruction of the primary neutrino is difficult and the median resolution is $10°$ at 100 TeV.

In contrast, muons leave a track-like light signature in the detector. This topology is optimal for direction measurements, that have resolutions of $0.3°$ at 1 TeV. However, traversing the detector, usually a large part of the energy is released outside and it is not detected, limiting the energy resolution to a factor of 2 [10]. Examples of signatures left by a muon and by a particle shower can be seen in Figure 3.4.



(a) Up-going track of a 75 TeV muon.

(b) A very energetic particle shower.

Figure 3.4: Examples of track-like (left) and shower-like (right) event topologies. The color of the spheres is proportional to the time of the first photon detected in the DOM they represent, being red the earliest and blue the latest. The size of the spheres is proportional to the logarithm of the collected charge. Taken from [10].

In summary, the events recorded by the IceCube detector can be classified using different properties. The events can be track-like if they exhibit an elongated light pattern or shower-like if the Cherenkov light has a spherical shape. The particle producing the signal can be a muon, another charged lepton or a charged hadron. The origin of the primary particle can be astrophysical, if the event was produced by the interaction of an astrophysical neutrino, or atmospheric, if the event was produced by a particle coming from an atmospheric particle shower initiated by a cosmic ray [10].

The identification of point-like astrophysical neutrino sources needs a well reconstructed primary neutrino with a low error on the arrival direction. Thus, the track-like signatures produced by muons are the best event topology for this study. Chapter 4 describes the procedure to select this type of events and to reconstruct the properties of the primary neutrinos.

Selecting well reconstructed track-like events is not enough to discover point-like neutrino sources. In fact, signal events defined by muons produced by astrophysical neutrinos are accompanied by an irreducible atmospheric background. Muon neutrinos are not directly observed by the IceCube detector, that is able to detect only the signature of the muon. Thus, on a per-event basis, muons produced by astrophysical neutrinos can not be disentangled from atmospheric muons or muons produced by atmospheric neutrinos, where “atmospheric” means “produced in cosmic ray induced atmospheric showers”. Astrophysical neutrinos are few per year, while the rates of atmospheric neutrinos and atmospheric muons are 1 every 6 minutes and 3000 per second, respectively. Figure 3.5 shows examples of a signal neutrino and different background events.

The Northern sky is defined as the region of the sky with declination between 0° and 90°, while the Southern sky declination ranges from -90° to 0°. Then, up-going events are defined in general as coming from below (with respect to a person standing up in the detector), while down-going events leave a signal in deeper DOMs as they travel further in the detector. The background is different from the two regions of the sky.

From the Northern sky neutrinos produced by cosmic rays interaction with the atmosphere can penetrate our planet and reach the South Pole, being eventually detected by IceCube if they interact inside or near the detector and the daughter muon is directed towards the detector. From this region of the sky atmospheric muons are absorbed by the Earth. However, once being reconstructed by the algorithms explained in the next chapter, up-going tracks are not entirely neutrinos coming from the Northern sky. Atmospheric down-going muons produced directly above the detector can leave

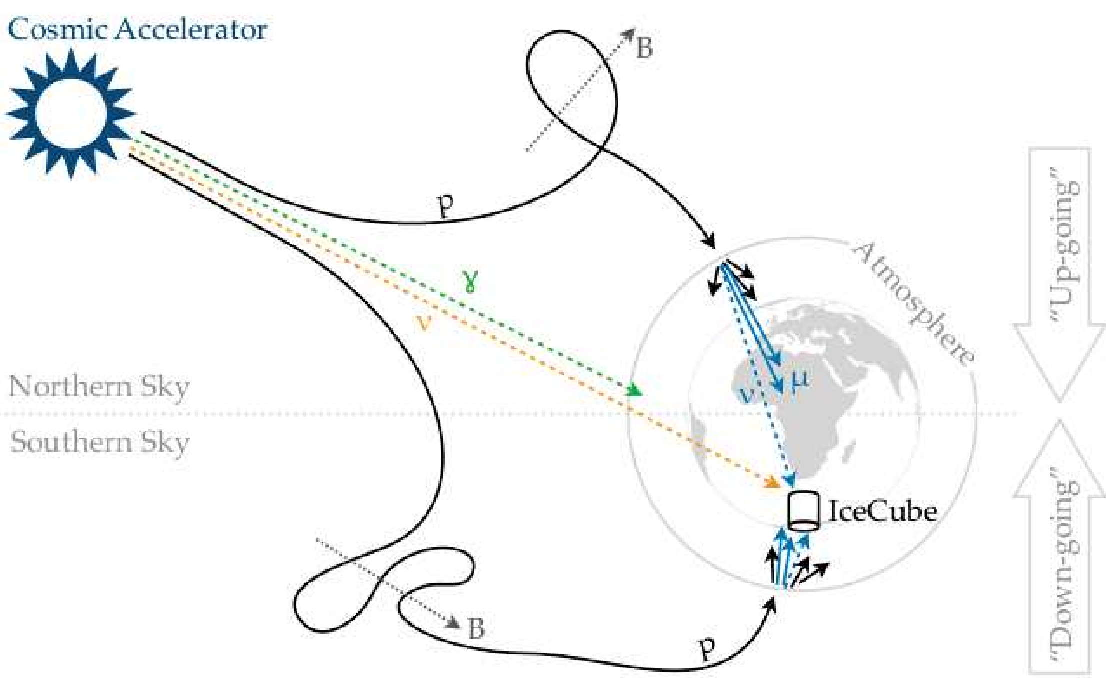


Figure 3.5: Sketch of different particles striking our planet. The dashed lines represent neutral particles, solid lines are used for charged particles, dotted grey lines represent magnetic fields. A signal neutrino, originated from an astrophysical source, is colored in orange. A photon produced by the same source is colored in green. Cosmic ray protons and secondary hadrons are colored in black. Muon neutrinos and muons forming the irreducible background for searches of point-like neutrino sources are colored in blue. Taken from [10].

an ambiguous event topology, for example because the interaction happens near the edges of the detector or a muon passes through a corner. Thus, in these cases, their arrival direction can be mis-reconstructed, falling in the wrong side of the sky. This part of the background is not negligible as the atmospheric muon flux is much higher than the atmospheric muon neutrino flux [10].

From the Southern sky both atmospheric neutrinos and atmospheric muons can reach IceCube, leaving a track-like signal in the detector at a much greater rate than signal astrophysical neutrinos. The most challenging background events are those for which several parallel muons produced by the same atmospheric shower, whose separation is much smaller than the distance between the IceCube strings, reach the detector. These close muons are usually wrongly reconstructed as a single more energetic muon [10].

The last module of the event filtering and reconstruction procedure, called GFU filter (Section 4.5), aims at selecting only properly reconstructed events, rejecting for example mis-reconstructed down-going events and muon bundles from the Southern sky.

However, even if on a per-event basis signal events can not be distinguished from the background, an analysis over all the data sample can find the astrophysical neutrinos. Different types of analyses exist, depending on the signal properties of interest. This work focuses on the search for clusters of signal events coming from individual sources. Chapter 5 illustrates the algorithm used to highlight clustered signal neutrinos among all the events present in the data sample produced by the filtering and reconstruction procedure (Chapter 4).

In order to test the point-like source analysis, Monte Carlo (MC) simulations of the events are used. A uniform flux of neutrinos following a power-law energy spectrum, simulated using the ANIS software [64], is produced at the Earth surface and the particles are propagated towards IceCube taking into account the neutrino interaction cross section and the Earth matter density [65]. Inside or near the instrumented volume of the detector neutrinos are forced to interact and a weight considering the interaction probability is stored in the file containing the simulation information [10]. This per-event weight is used when dealing with the simulated events.

In addition to neutrinos, also muons produced in air showers initiated by cosmic rays are simulated using the CORSIKA simulation program [66]. The energy of the simulated neutrinos ranges from 100 GeV to 1 EeV, while the energy of the primary cosmic rays producing the simulated atmospheric muons ranges from 600 GeV to 100 EeV. For both neutrino-induced and atmospheric muons, the Muon Monte Carlo program [67] is used to propagate the muons taking into account their continuous and stochastic energy losses. Photons produced in the muon track are propagated towards the DOMs, whose response is simulated using the IceCube-internal software DOMlauncher. Finally, the DOMs signals pass through the same event filtering and reconstruction procedure used for the real data.

3.3 Detector Monitoring

As opposed to other types of detectors like the IACTs that need particular conditions to take data (as very dark moonless nights and good weather conditions), the IceCube detector can be always operative. However, for different motivations, the data taking can be stopped [10]. Moments of downtime, i.e. when the detector is not taking data, are originated for example by software upgrades, calibration with the DOMs LEDs or power outages. In case of single DOMs or strings malfunctioning the DAQ excludes the problematic components and continue the data taking with a partial configuration of the detector.

When analysing the data provided by the IceCube detector, the downtime periods have to be handled properly. For online analyses, considered in this work as discussed in Chapters 4 and 5, it is fundamental to constantly check the detector status not to send false alerts to the Astrophysical community. The data taking is automatically stopped during planned software upgrades and calibrations and the known malfunctioning DOMs are excluded.

However, unpredictable problems can be identified by monitoring the rates of three filters of the event filtering and reconstruction procedure discussed in Chapter 4, that should not vary abruptly: the Simple Multiplicity Trigger (Section 4.1), the Muon Filter (Section 4.3) and the OnlineL2 Filter (Section 4.4). The rates of the three quantities are evaluated in 10 minute bins and the result is compared to an exponential moving average of past data, that gives more weight to more recent data [10]. If the rates and the average differ significantly for a 10 minute bin, the latter is not considered in the analysis.

This approach is used to take into account the changing conditions at the South Pole. Since the considered filters are not at the end of the event filtering and reconstruction procedure, most of the events are muons produced in the interaction of cosmic rays with the atmosphere. The temperature of the air at the South Pole varies throughout the year and in turn the air density is modified, causing a change in the expected atmospheric muons rate. From $\approx -80°$C in January, the air temperature can reach values of the order of $-40°$C in August, resulting in changes in the muon rate by $\approx 10\%$ [68].

Although the detector uptime is not 100%, the data taking interruption time is small. The detector downtime is found to be smaller than 1% [69]. Figure 5.1(a) in Section 5.1 shows the daily event rate at the final step of the event filtering and reconstruction procedure for one year of data, reported as a function of time. The effect of the exclusion of long data taking periods is seen as a sharp decrease of the event rate. However, for most of the data sample time the rate is free from abrupt changes.

Chapter 4

The Real-Time Infrastructure: Event Selection, Reconstruction and Analyses

To perform any analysis it is needed to filter and clean the pulses provided by each IceCube DOM, characterised by time, charge and width. Then, the reconstruction of the properties of the single events, like the arrival direction of the primary particle, its energy and the topology of the event (shower- or track-like), has to be performed. Events suitable for a particular analysis are subsequently selected: to search for point-like astrophysical neutrino sources, muon neutrino candidates producing a muon that left a through-going track-like signature in the detector are the best type of event to use thanks to the good angular resolution that can be achieved. At this point the analyses searching for astrophysical neutrino clusters are applied.

In order to perform real-time (online) analyses, i.e. analyses run directly when the data are collected providing results in the least amount of time, the filtering and reconstruction procedure has to be very fast, as well as the analyses algorithms. This results in the possibility to alert other observatories in order to perform follow-up observations, i.e. searches for a counterpart of any messenger correlated to the IceCube observation. This chapter illustrates the real-time infrastructure, i.e. the processing of IceCube data and the analyses for a real-time study.

Several filters and reconstruction algorithm exist within the IceCube collaboration and they are selected depending on several factors, like the analysis to be conducted, the type of events to be selected or whether an online or offline analysis is used. Here, only the algorithms used for this work, within the present real-time alert system [69], are discussed: the aim is a real-time analysis searching for point-like neutrino sources exploiting the identification of clusters of signal neutrinos.

Figure 4.1 illustrates the filtering and reconstruction steps to pass from the IceCube pulses to a set of neutrino-induced muon candidates, to which the analysis is applied. The event processing chain is executed directly at the South Pole. Since the event rate is very high at the beginning of the filtering and reconstruction procedure, only simple algorithms can be used in order to complete the whole chain very quickly to perform real-time analyses. As the filters reduce the event rate, more sophisticated but slower algorithms are used. In Section 4.1 the detection trigger conditions, composing the Level1 Trigger, are defined. Then, the first directional estimations, as well as the cleaning of the DOMs pulses, are performed within the Base Processing module described in Section 4.2. The Muon Filter, discussed in Section 4.3, aims at selecting track-like events with a high probability to be muons. The OnlineL2 Filter, explained in Section 4.4, improves the directional reconstruction, provides an estimate of the event energy and reduces again the event rate. Finally, in Section 4.5, the last (multivariate) selection and the angular error estimation making up the GFU filter are described. The name of the last module comes from the Gamma-ray Follow-Up program, a real-time program for which IceCube sent alerts to the Cherenkov telescopes MAGIC and VERITAS that performed follow-up observations [70], and the name stuck. The output of this chain is a neutrino-induced muon candidate defined by event identification numbers, event time, reconstructed direction, angular uncertainty and reconstructed energy.

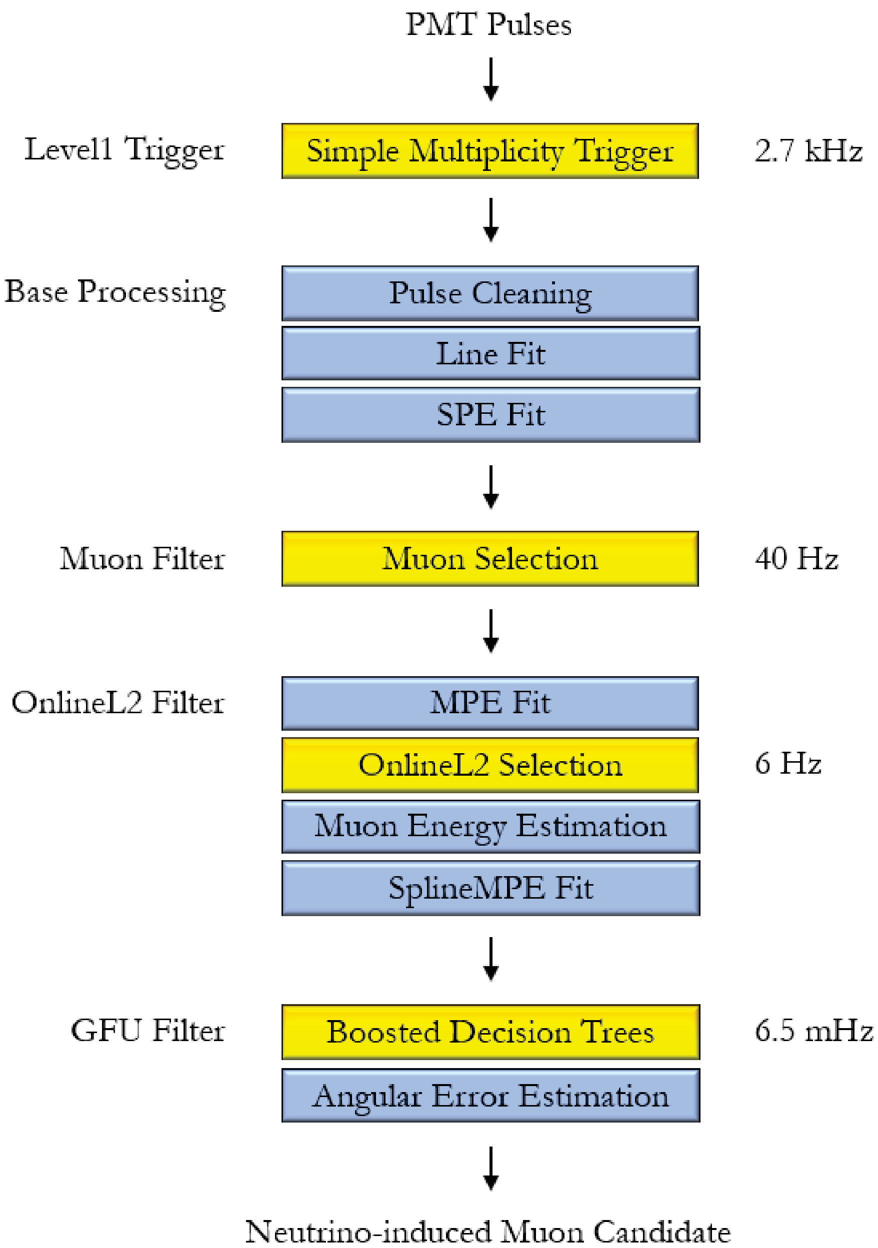


Figure 4.1: Overview of the filtering and reconstruction procedure. The names of the modules are written on the left. Each module can be composed of a single step or embed multiple operations, each one represented by a box. Light blue boxes indicate reconstruction steps, while yellow boxes refers to filters reducing the data rate by selecting the events of interest rejecting noise and background-like events. The approximate event rate *after* the selection is reported on the right. Adapted from [10].

In Figure 4.2 the most important steps of the real-time system are shown, specifying the position of the event filtering and reconstruction procedure. The signals provided by IceCube are used to monitor the correct functioning of the detector, as discussed in Section 3.3. More importantly, the waveforms recorded by the DOM photomultipliers are digitised and sent to the surface. They pass through the event filtering and reconstruction procedure at the South Pole and the output is sent to the IceCube computing center in Madison, WI. Here, the real-time analyses are run: they aim at finding and estimating the features of the astrophysical neutrinos, disentangling them from background events consisting of atmospheric neutrinos or atmospheric muons. Two examples of analyses focus on the monitoring of known gamma-ray sources and on a search of all the sky. In case of an interesting observation, alerts are sent to the astrophysical community or to IceCube partner detectors for follow-up observations.

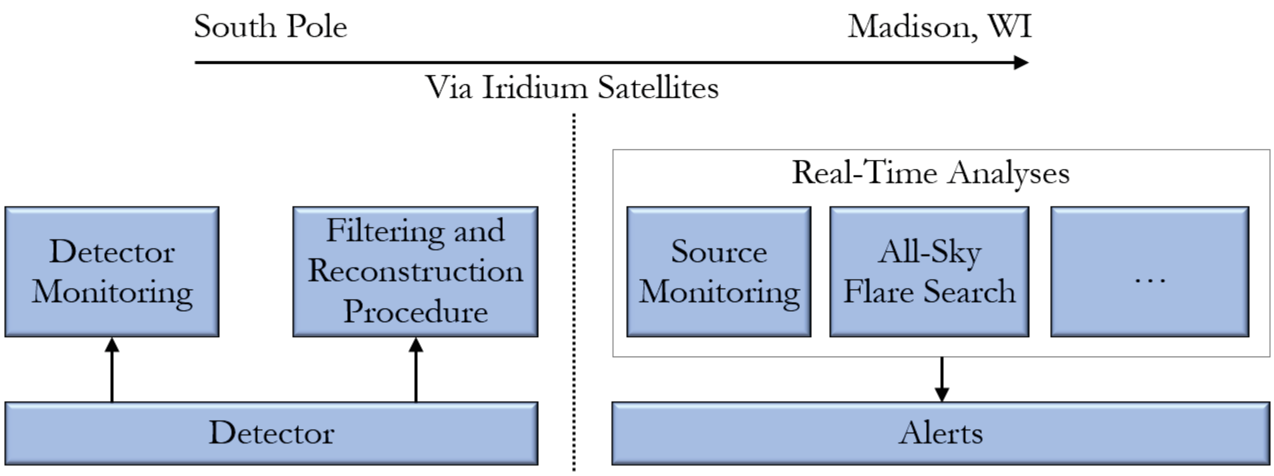


Figure 4.2: Overview of the real-time infrastructure of the IceCube experiment. The data, other that serving as checks of the correct functioning of the detector, pass through the filtering and reconstruction procedure. Then, they are sent from the South Pole to the IceCube computing center in Madison, WI, where the real-time analyses are run. Eventually, alerts are sent to the astrophysical community or to IceCube partner detectors for follow-up observations. Adapted from [10].

The analyses are run every time a new event arrives, thus each algorithm of the filtering and reconstruction chain has to be very fast. For example, the runtime of the combination of the OnlineL2 and GFU filters must not be greater than 15 seconds, or queues of events could be created [10]. In case an event processing exceeds the time limitations, the filtering and reconstruction procedure for that event is stopped and the event is handled in offline analyses. A summary of the most important analyses for this work and the alert system are discussed in Section 4.6. During this work, the core of the point-like neutrino sources search algorithm has been tested. The detail of the algorithm and the test results are discussed in Chapter 5.

4.1 Level1 Trigger

The first step encountered by the DOM pulses is called Level1 Trigger and it consists in dividing the stream of pulses in individual physics events. Since the DOMs can observe single photons, noise hits have to be removed, too.

This procedure starts already inside the DOM. Firstly, the so-called Hard Local Coincidence (HLC) is checked. After observing a hit, the DOM software checks neighboring or next-to-neighboring DOMs on the same string for coincident hits within ± 1 μs. For each HLC hit the full digitized waveform is stored. If the HLC condition is not satisfied, it is called Soft Local Coincidence (SLC) and only the amplitude is stored, since the SLC DOM hit is probably a noise hit. The DOM main board, placed above the delay board, stores the information in memory, that is requested once per second by the computers on the surface.

The information of the HLC is then used by the so-called Simple Multipliciy Trigger (SMT-8) that decides which time periods contain interesting hits. It requires a minimum of 8 HLC hits during a sliding time window of 5 μs. The number of minimum hits is chosen depending on the detector sensitivity and it is different for DeepCore and IceTop. The SMT-8 condition satisfaction defines the trigger time window start, that ends when the sliding time window does not contain any more HLC hits. Then, the different trigger windows are merged into a global trigger window. This ensures the rejection of part of the noise hits. All the HLC and SLC hits inside the global trigger window are sent to the event builder and written to disk as "DAQ events", where the processing and filtering system takes his input. The median trigger rate is 2.7 kHz, corresponding to ≈ 1 TB/day [60]. Once per year, these data, saved to hard disk at the South Pole, are shipped to Madison, WI.

Each DAQ event could contain many physics events because the triggers caused by different particles could have been overlapped and merged into the same global trigger window. Each physics event has to be separated from the others via a procedure called "trigger splitting", able to separate for example two muons arriving with a little time difference. The procedure searches for contiguous time windows during which the SMT-8 condition is satisfied and these time windows, after adding 4 μs before and 6 μs after, define physics events. Coincident events, like two muons arriving separated in space but not in time, are not divided into separate events at this stage. Other event selection steps, described in the next sections, aim to remove these coincident events since the reconstruction algorithms would otherwise treat the coincident muons as a single particle.

Time coincident but space separated particles could be divided in different events by using the spatial distribution of the hits. However, even if these events would be considered and not excluded in the analyses, other types of events would suffer from this handling as a single physics events could be wrongly separated into multiple events. For example, a dim muon track traversing a dust layer (mentioned in Chapter 3) without depositing any light in it could be divided in two parts, each one corresponding to a "fake" muon. Since the "fake" muons would have similar directions, they could be collected in an event cluster and the significance level of a point-like source analysis would increase artificially. Thus, this particular process using the spatial information is not used in this online analysis and time coincident events are removed at later steps [10].

4.2 Base Processing

After the division into separated physics events, the Base Processing steps aims, after cleaning the DOM pulses, at reconstructing the original direction of a neutrino candidate. Here and for the following algorithms the hypothesis of a neutrino-induced trough-going muon track underlies the reconstruction steps. In order to study point-like astrophysical neutrino sources, the directional reconstruction is one of the most delicate steps in the reconstruction chain. For this motivation, this step is performed four times, each one with a more sophisticated algorithm. The name of the four algorithms are Line Fit, Single Photoelectron (SPE) Fit, Multi Photoelectron (MPE) Fit and Spline Multi Photoelectron (SplineMPE) Fit. More powerful reconstruction operations require more time to be completed: to perform real-time analyses, very time consuming algorithm can not be applied to the data when the event rate is too high, but only after several filters that reduce the rate. For this motivation, Line Fit and SPE Fit are used in the Base Processing section of the filtering and reconstruction chain, where the event rate is still 2.7 kHz, and are explained in this section. As can be seen in Figure 4.1 and in the next sections of this chapter, MPE is applied when the rate decreases to 40 Hz and SplineMPE when the rate is 6 Hz, in later modules. Finally, by seeding the more sophisticated directional reconstruction algorithms with the result of the previous one, the convergence time of the single algorithms decreases.

Pulse Cleaning: each event in which the global trigger window is divided by the trigger splitting contains all the pulses recorded in that event time window, but in general not all the pulses are related to the particles creating the physics event. Pulses unrelated to the signal are removed by using a cleaning algorithm. Firstly, HLC pulses are used as a seed. Then, for three times, any SLC pulses around the considered DOMs are added if they are closer than 150 m and recorded within 1 μs. Finally, the surviving pulses are ordered in time and the 6 μs window with the largest number of pulses is selected. This process selects bright, spatially and temporally connected clusters of pulses and removes single DOMs with noise pulses or connected to very dim atmospheric muons.

Line Fit: at this step, the first directional reconstruction is performed using the Line Fit algorithm. It exploits the assumption of a plane wave of light passing through the detector and producing the observed first pulse in each DOM. The muon is assumed to be moving in a straight line with velocity $\vec{v}$. After passing through $\vec{r}_0$ at time t_0, its position $\vec{r}$ at time t is given by:

$$\vec{r}(t) = \vec{r}_0 + (t - t_0)\,\vec{v}. \tag{4.1}$$

The algorithm consists in finding $\vec{r}_0$ and $\vec{v}$ (that completely define the muon track) such that the sum of the squared distances between the (unknown) muon track and the DOMs positioned at $\vec{x}_i$ that registered a pulse at time t_i is minimum, i.e. in solving

$$\arg\min_{\vec{r}_0,\vec{v}} \sum_{i=1}^{N} ||\vec{r}(t_i) - \vec{x}_i||^2, \tag{4.2}$$

where N is the number of hit DOMs. This method is improved by excluding hits originated from scattered photons [71], since the scattering properties of the ice are not taken into account by Line Fit. Robustness to noise hits is achieved by replacing the square in Equation 4.2 with a Huber penalty function [72] which gives less weight to hits far away from the track. In addition to the first directional reconstruction, Line Fit estimates the velocity of the muon $|\vec{v}|$.

Single-photoelectron track fit: A more refined estimate of the muon direction is obtained considering the properties of the Cherenkov emission, neglected by Line Fit. The muon, moving faster than light in ice, produces a Cherenkov cone of light which wavefront is emitted at an angle θ_C with respect to the muon track [29]. Being n the ice refractive index and β the velocity of the muon in units of the light velocity in vacuum c, the cosine of the Cherenkov angle is given by

$$cos\theta_C = \frac{1}{n\beta}. \tag{4.3}$$

Assuming $\beta \approx 1$ for the (relativistic) muons and $n \approx 1.32$, the Cherenkov angle is $\theta_C \approx 41°$ [73]. Calling $\vec{v}$ the normalised direction of motion of the muon passing through $\vec{r}_0$ at time t_0, in absence of scattering the Cherenkov light front arrives at time t_{geo} at a DOM located at $\vec{x}_i$ at a distance d from the track. This time is given by

$$t_{geo} = t_0 + \frac{\vec{v}\,(\vec{x}_i - \vec{r}_0) + d\,tan\theta_C}{c}. \tag{4.4}$$

Figure 4.3 illustrates the aforementioned geometry.

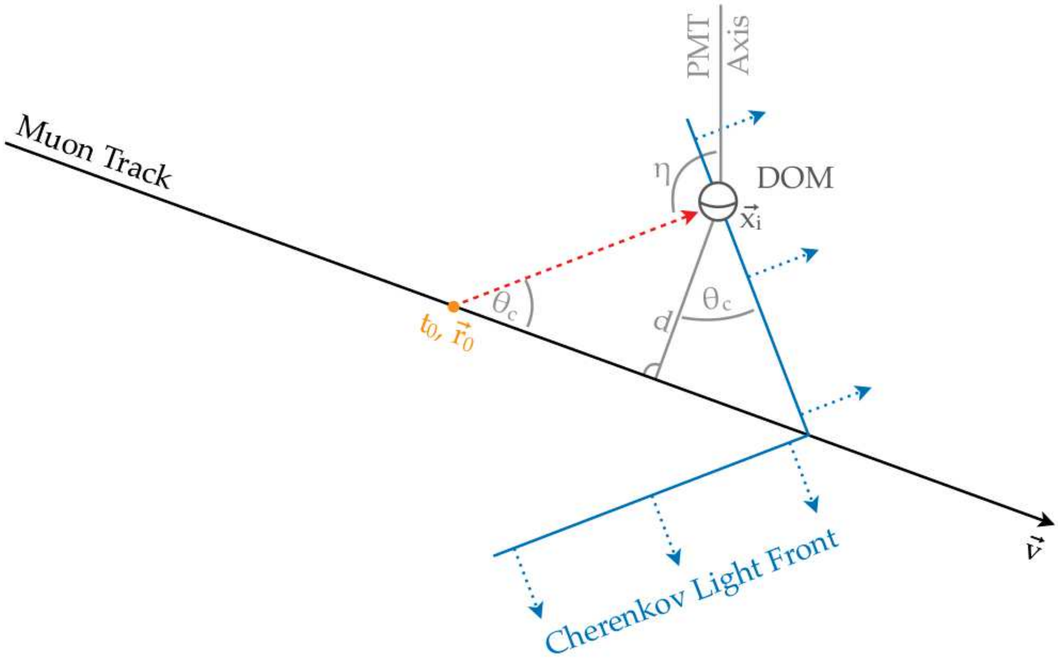


Figure 4.3: Sketch of the Cherenkov light front emitted from a muon track and detected by a DOM. The geometric photon arrival time is defined by the time it takes for an unscattered photon emitted at $\vec{r}_0$ to reach the DOM at $\vec{x}_i$. Taken from [10].

The time residual t_{res} is defined as the difference between the observed arrival time t_{obs} and t_{geo}:

$$t_{res} = t_{obs} - t_{geo}. \tag{4.5}$$

Defining the track with the arbitrary muon position $\vec{r}_0$, the zenith angle θ and the azimuth angle ϕ, the likelihood of observing all the time residuals $t_{res,i}$ is constructed exploiting the probability distribution p_1 for the time residual of a single photon and it is given by

$$\mathcal{L} = \prod_{i=1}^{N_{Ch}} p_1(t_{res,i}|\vec{r}_0, \theta, \phi), \tag{4.6}$$

where N_{Ch} is the number of hit DOMs. The likelihood maximisation provides an estimate of $\vec{r}_0$, θ and ϕ.
The single photon time residual probability density function p_1 is expressed by the so-called Panel function

$$p_1(t_{res}) = \frac{1}{N(d)} \cdot \frac{\tau^{-d/\lambda} \cdot t_{res}^{d/\lambda-1}}{\Gamma(d/\lambda)} \cdot exp\left[- t_{res}\left(\frac{1}{\tau} + \frac{c}{n\,\lambda_a}\right) - \frac{d}{\lambda_a}\right] \tag{4.7}$$

with

$$N(d) = \left(1 + \frac{c\,\tau}{n\,\lambda_a}\right)^{-d/\lambda} \cdot exp\left[-\frac{d}{\lambda_a}\right], \tag{4.8}$$

where λ_a is the absorption length and λ and τ are fit parameters obtained from a Monte Carlo simulation of the photon propagation. The Pandel function was obtained by laser calibration for the Baikal experiment [74]. The AMANDA collaboration fits this function using models of the South Pole ice, taking into account other effects like PMT jitter and the delay in case of backward illumination of the PMT, i.e. when a photon produced above a DOM scatters multiple times reaching the PMT after a certain delay [73].
To exit from possible local extrema present in the likelihood space created by detector symmetries, survived noise/scattering hits or muons hitting only the corners of the detector, after a first direction estimate $\vec{v}_0$ two random directions $\vec{v}_{1,2}$ perpendicular to $\vec{v}_0$ are chosen. $\vec{v}_{1,2}$ are used as a seed for the likelihood fit, that is repeated. This procedure is iterated and the result with the highest likelihood value is chosen.
The SPE likelihood considers only the first photon recorded by each DOM, as for Line Fit. The maximisation of the SPE likelihood requires a seed: the result of Line Fit is used to speed up the convergence. As for the Line Fit method, SPE is not used only to estimate the direction of the muon, but it also provides an estimate of the probability for an event to actually be a muon with a track-like signature, through the value of the likelihood.

At this point the information on the events, such that the SPE likelihood value, is enough to select track-like muon event candidates, as explained in the next section.

4.3 Muon Filter

The data arriving at this step are composed by candidate physics events. In order to reduce the event rate so that more complicated analyses can handle the data, an early filtering stage aims at selecting the events of interest, e.g. track-like muons or cascades. The Muon Filter is considered in this work: it excludes the events with a low probability to be track-like muons. As already specified, thanks to its long lever arm this event topology is the best for the directional reconstruction, of vital importance for point-like sources studies.

Since most of the muons are produced outside the detector in which they enter, leave a track-like signature and then exit, every muon track is considered as produced by a through-going muon. Few muon neutrinos interact inside IceCube and only part of their track is contained in the instrumented volume, but they are treated as through going muons as well. Eventually, for example due to the less number of pulses with respect to truly trough-going muons and thus a poorer directional reconstruction, track-like events starting inside the detector can be removed by a later stage of the event selection procedure. Algorithms able to identify events starting inside the detector exist, but they are quite time consuming and thus not suitable for online real-time analyses [10], like the ones considered in this work.

For up-going events, the SPE likelihood value is used as an estimate of the fit quality, providing a quantity representing the probability for an event to be track-like. For down going events, the events rate is high because of the atmospheric muon contribution. Since the astrophysical neutrino flux is expected to be much harder ($\sim \mathrm{E}^{-2}$) than the atmospheric muon contribution ($\sim \mathrm{E}^{-3.7}$), to select events with a higher energy and thus with a higher probability to have an astrophysical origin a minimum total integrated charge for the event is required. In practice, depending on the value of the zenith angle θ, three cuts are used:

$$\begin{aligned}
&\frac{log\mathcal{L}}{N_{ch}-3} \leq 8.7 && if && -1.0 \leq cos\theta \leq 0.2 \\
&logQ_{tot} > 3.9 \cdot cos\theta + 0.65 && if && 0.2 < cos\theta \leq 0.5 \\
&logQ_{tot} > 0.6 \cdot cos\theta + 2.3 && if && 0.5 < cos\theta \leq 1.0,
\end{aligned} \tag{4.9}$$

where $\mathcal{L}$ is the SPE likelihood, N_{ch} is the number of hit DOMs and Q_{tot} is the total integrated charge for the event.

To handle possible bugs in the processing, the waveforms for every event before the selection are stored on hard disk at the South Pole and shipped to Madison, WI, once a year. After the muon filter the event rate passes approximately from 2.7 kHz to 40 Hz. The latter is low enough for the daily satellite transmission to Madison of the extracted set of pulses used for offline studies and for using the more complicated reconstruction algorithms at the South Pole for online studies, explained in the next sections. The satellite bandwidth allocation is $\sim$ 100 GB/day [10].

4.4 OnlineL2 Filter

The OnlineL2 filter name derives from the word “Online”, used because this step is used for several real-time analyses, and the acronym L2, that stands for “Level 2” since it is used after the first level reconstructions of the previous steps. This section of the filtering and reconstruction procedure aims at improving the directional estimate of the Single Photoelectron likelihood fit by using firstly the Multi Photoelectron fit and, after reducing the event rate with the OnlineL2 selection, the SplineMPE fit. In addition, the muon energy is estimated after the selection: this quantity is used to estimate the energy spectrum of the sources and it helps in separating signal and background events. Since the muons from muon neutrinos are typically produced outside the detector (even several kilometers away from it), the energy observed by IceCube is only a fraction of the total. However, the estimate of the energy at the detector provides a lower bound on the neutrino energy [10].

Multi-photoelectron track fit: the result of the SPE fit is used as a seed for this more complicated algorithm, that takes into account the possibility to detect more photons in each DOM. In fact, for muon energies above 1 TeV the probability to observe more than one photon in each DOM is not negligible. Thus, taking into account the probability to observe the first photon out of N_i photons in the i-th DOM at time $t_{res,i}$, the likelihood expression is

$$\mathcal{L} = \prod_{i=1}^{N_{Ch}} \left[N_i \cdot p_1(t_{res,i}) \cdot \left(\int_{t_{res,i}}^{\infty} p_1(t_{res,i}) \right)^{N_i - 1} \right]. \tag{4.10}$$

The likelihood maximisation procedure is similar to the SPE fit. This algorithm, in addition to providing a better estimate of the event arrival direction, is able to properly identify several down-going events that were reconstructed as up-going muons.

OnlineL2 Selection: this stage reduces the event rate from 40 Hz to 6 Hz. As for the Muon Filter, the selection is based upon the directional reconstruction likelihood value and the event integrated charge and it depends on the region of the sky where the events come from. The cuts have the same aim of the Muon Filter ones and are similar to them, but they are more complicated and tighter thanks to the better directional reconstruction provided by the MPE fit. The details can be found in [75].

SplineMPE Fit: the fourth, last and most sophisticated algorithm for directional reconstruction is the SplineMPE fit. It uses the same likelihood of the MPE fit (Equation 4.10), but the time residual probability density function p_1 is modified. Instead of the Pandel function, a simulation of the photon propagation is exploited. The homogeneous ice model of MPE is substituted with a model that takes into account the depth dependence of the ice absorption and scattering. The results of the simulation in different conditions, e.g. with different values of the distance between the track and the DOM, are available as interpolating splines[1] [76, 77].
In addition, the reconstruction is slightly improved and the algorithm run time is reduced with some modifications, e.g. by taking into account some stochastic energy losses of the muons (which produce unwanted late pulses in the DOMs, that are removed) [10].

Muon Energy Estimation: the muon and the eventual secondary charged particles produced by it[2] cause the emission of Cherenkov photons as their velocity is greater than the speed of light in ice [29]. The number of Cherenkov photons is proportional to the track length, that is proportional to the particle energy [78]. The light yield scales linearly with the energy, thus the template of the expected photon count, that is usually simulated at 1 GeV, can be extended to other values of the energy. Since the number of observed photons should follow a Poisson distribution, the likelihood of the energy estimation is

$$\mathcal{L} = \frac{\lambda^k}{k!}\, e^{-\lambda} \tag{4.11}$$

with

$$\lambda = \Lambda E + \rho, \tag{4.12}$$

where E is the energy, k is the number of observed photons, λ is the mean expected number of photons, Λ is the template of the expected photon count and ρ is a parameter introduced to take into account noise and other contributions [61]. The likelihood maximisation provides an estimate of the energy E.
To parametrise Λ, MuEX approximates the Cherenkov light emission as uniform along the track [79], so the expected number of photons depends on the distance between the track and the DOM of interest. An empirical expression, that uses the tabulated absorption and scattering lengths for a layered structure of the ice, approximates the diffusion far away from the track and assumes a 1/r dependence of absorption, neglecting the scattering, near the track [61]. This model is improved by considering the stochastic energy losses of the muons. At higher energies, the reconstruction performs better thanks to the high number of photons observed for each event [10].

Figure 4.4 shows the median angular error of the four directional reconstruction algorithms exploiting simulated events, i.e. the median difference between the “true” simulated direction and the one provided by the reconstruction algorithm, as a function of the true muon energy and declination.
It is clear how the error is lower by refining the directional reconstruction with more sophisticated algorithms. For example, the median angular error for the least sophisticated Line Fit algorithm is always above 1°, while for the most sophisticated SplineMPE Fit it is almost always under 1°. By considering the SplineMPE Fit curve, the error decreases by increasing the energy, as expected. However, the stochastic light emission is not considered in the energy reconstruction likelihood, causing the error to increase at the highest energies. In the Northern sky the error increases for higher values of the declination as the high-energy muons (thus, emitting more photons, useful for a better reconstruction) are absorbed by the Earth and the detector geometry, with its different lateral and vertical DOM spacing, favors horizontal events [10].

[1] A spline (Segmented Polynomial LINE) is a function built by joining different polynomials to fit a series of points such that the function is continuous until a certain order of derivatives.

[2] Above the critical energy of ≈ 1 TeV, the energy loss happens predominantly by bremsstrahlung. Photons produced by this process can undergo pair productions, resulting in electromagnetic showers, or photonuclear interactions, resulting in hadronic showers. The showers are seen as point-like by the IceCube detector, with a spherical shape of the emitted light [61].

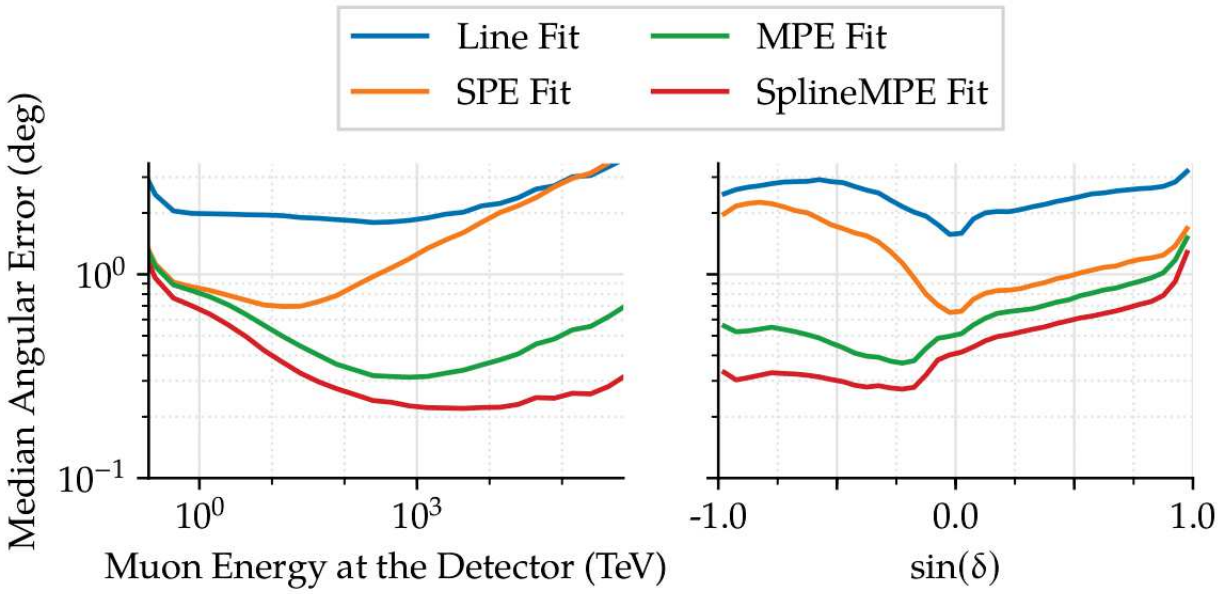


Figure 4.4: Median angular error for muons at the final event selection level for the four directional reconstruction algorithms. Left: as a function of the true muon energy, averaged over all the declinations. Right: as a function of the true declination, averaged over all the energies. Taken from [10].

4.5 Gamma-ray Follow Up Filter

This last module of the filtering and reconstruction procedure takes in input the results of the OnlineL2 filter. The aim of the GFU filter is to provide a sample of events with the best possible neutrino purity by using a machine learning classifier based on Boosted Decision Trees (BDTs). Most of the events in the sample are down-going events reconstructed as up-going that are rejected by this filter, as well as down-going muon clusters originated from the same atmospheric shower. This selection reduces the event rate from 6 Hz to 6.5 mHz (≈ 200 000 events/year), providing the final sample of events used by the analyses described in the next section. After the filter, approximately 2/3 of the events are up-going and 1/3 are down-going. Among all the events in the final data sample almost the 0.1% is estimated to have an astrophysical origin [10]. The angular uncertainty is estimated at the final step of the procedure.

Boosted Decision Trees: the results obtained up to now, for example the good directional reconstruction provided by the SplineMPE algorithm, can be used to select more strictly the events of interest, rejecting the background. Instead of splitting the data using a single variable, a multi-variate approach provides a better selection of the signal. The algorithm is based on the so-called decision trees. A decision tree is a set of binary decisions with several levels. At each node, the value of an event observable is compared to a cut. The event is passed to one of the two following nodes depending on which side of the parameter space the observable value is. After the final level, the event ends up in a signal- or background-dominated leaf and it is kept or rejected, respectively. The signal probability is defined by a parameter called purity [10].
The training process chooses the observable and its cut value for each node, building the entire decision tree (the process stops, for example, when a pre-defined maximum tree depth is reached). Then, the "boosting" procedure is performed: a completed tree is trained to choose the weights to give to the events (that are used to calculate the leaf purity) depending on whether or not the events are correctly tagged as signal or background at the end. The mis-classified events will have higher weights values. The overfit problem, i.e. the result to be stuck on the specific characteristics of the training set of events, is handled for example by randomly selecting a small number of observables for some nodes. The trees are built such that the separation gain of the nodes and the point-like source sensitivity are maximum [10]. Implementations of the algorithm can be found in the public package "scikit-learn" [80] or in the IceCube internal version "pyBDT" [81].

The dominant kind of background is not the same for Northern and Southern skies, thus they are handled differently. For the Northern sky the events to be rejected are mostly mis-reconstructed down-going muons or cascades wrongly selected by the previous filters. Atmospheric neutrinos can not be rejected at this level, with a per-event approach, but require a more sophisticated analysis (Section 4.6). The variables used in this part of the sky are summarised in the following list [10]:

- The zenith angle provided by the SplineMPE fit has to be grater than 82°. This threshold thus defines up-going events;
- The value of the SplineMPE directional reconstruction likelihood normalised using the number of hit DOMs provides an estimate of the probability for the event to be a muon track, as already mentioned for the Muon Filter and OnlineL2 Filter with the SPE and MPE likelihoods, respectively;
- The second order derivative of the SplineMPE directional reconstruction likelihood around the maximum is related to the reconstruction quality. A strong curvature is related to a narrow optimum, suggesting a good reconstruction;
- The speed of the particle estimated during the Line Fit algorithm should be close to the speed of light;
- The particle direction provided by the the Line Fit algorithm should not differ strongly with respect to the SplineMPE fit. A great change in the directional estimate can indicate unreliable fit results, typical of mis-identified cascades. Thus, the angle between the Line Fit and the SplineMPE fit directions should be small;
- Muons traversing large part of the detector are easy to reconstruct, while muons hitting only the outer surfaces of IceCube have a high change to be mis-reconstructed. The center of gravity of the hits is related to how close an event was to the borders of the detector. It is defined as

$$\overrightarrow{CoG} = \frac{1}{Q_{tot}} \sum_{i=1}^{N_{Ch}} q_i \cdot \vec{x}_i, \tag{4.13}$$

  where Q_{tot} is the total integrated charge of the event, N_{Ch} is the number of hit DOMs, q_i is the charge collected by the i-th DOM and $\vec{x}_i$ is its position. The center of gravity radial and depth components are used for some nodes of the tree to select events further away from the borders of the detector;
- A difficult class of events to identify are simultaneous events, i.e. two particles interacting in different parts of the detector very close in time, for which the event splitting fails. The directional reconstruction is performed taking into account all the pulses, even if they are related to different particles. Thus, to handle this category, the set of pulses is divided in two halves and the directional reconstruction is applied to each of the two sub-sets. The division is performed at the median pulse time and at a plane perpendicular to the SplineMPE track and passing through the center of gravity (Equation 4.13), thus creating four sub-sets of events. For single muon events the reconstructed direction of the four sub-sets should be similar to the SplineMPE one. The difference between the cosine of the SplineMPE zenith angle and the cosine of the minimum zenith angle of the four fits performed with the sub-sets of events is requested to be small. The minimum is used to prefer down-going atmospheric muons, more likely to provide coincident events;
- The MPE fit likelihood can be corrected with a zenith-dependent prior [82] considering the expectation of down-going atmospheric muons. The value of the (negative logarithmic) likelihood increases for local minima characteristic of mis-reconstructed down-going muons. This permits a better identification of the global minimum. Thus, the difference between the likelihood values with and without this correction is used;

- The time residuals defined by Equation 4.5 should have an upper limit to remove scattered photons and a lower limit to take into account the finite timing resolution of the detector. Time windows are defined with these restrictions and the DOM hits falling inside are called "direct hits". The number of direct hits is used in the trees, as few direct hits indicate mis-reconstructed events.

- The direct length, defined as the distance between the first and the last direct hits, is greater for through-going muons. Thus, this quantity is used in the trees, for example to reject starting events;

- A "smoothness" parameter, describing the distribution of direct hits along the track expected to be uniform for well-reconstructed tracks, is exploited to reject badly-reconstructed tracks;

- The separation length L_{sep}, i.e. the distance between the center of gravity (Equation 4.13) of the first and last quarter of hits, is longer for a good track-like shape. This quantity is used, for example, to reject cascades still present in the data sample;

- The empty track length L_{empty}, defined as the maximum distance along the track during which no hits are observed within a radius of 150 m around the track, is small for events not consisting of two different particles interacting in different parts of the detector. This is used to reject multiple events close in time that were seen as a single event by the event splitting procedure;

- The DOMs collecting more charge should be the ones closer to the track. The average charge-weighted track-to-DOM distance is defined by

$$d_Q = \frac{1}{Q_{tot}} \sum_{i=1}^{N_{Ch}} q_i \cdot d_i, \tag{4.14}$$

  where Q_{tot} is the total charge of the event, N_{Ch} is the number of hit DOMs, q_i is the charge collected by the i-th DOM and d_i is its distance from the track. This quantity is small for well-reconstructed tracks, thus it is used to reject badly-reconstructed events.

For the Southern sky, atmospheric muons can leave a signal in the detector because they do not encounter the Earth while they travel, while Northern sky atmospheric muons are absorbed by our planet. They mimic the neutrino-induced muon signals, as only the muons and not the neutrinos are directly visible. In addition, almost collinear clusters of muons produced in the same shower can result in a mis-reconstructed single high-energetic muon.
Veto techniques using the outer DOMs of IceCube or the IceTop DOMs are not suitable for online searches of neutrino sources. The outer DOMs of IceCube could select the events starting inside the detector, with a high probability to be of astrophysical origin, but the rate of these events from the Southern sky ($\approx$ 180/yr) is much lower than the through-going event rate ($\approx$ 50 000/yr). In addition, both the outer DOMs of IceCube and the IceTop DOMs should be always working properly (which is not guaranteed), as an absence of detected photons due to a malfunctioning can not be identified online.
The same decision tree variables of the Northern sky are used for the Southern sky, but the zenith angle has now to be lower than 82°. However, to deal with atmospheric muon clusters, some decision tree variables are added [10]:

- The time residuals defined in Equation 4.5 are different for a single muon and for a muon bundle, for example because the muons in the bundle have offsets both along the track direction and on the plane perpendicular to it. The pulses produced by the bundles are earlier or later than expected: by creating energy-dependent signal and background time residual PDFs, and by combining them in a likelihood that is maximised, separation power towards the muon bundles is gained;

- Single high-energetic muons undergo stochastic energy losses, while bundles of low-energetic muons only show a continuous energy loss due to the Cherenkov emission. The differential energy loss per unit of track length is modeled and the expected distributions for single muons and muon bundles are created and combined in a likelihood, that is maximised and used to reject the muon bundles.

Some of these variables, like the number of hit DOMs, are influenced by the higher DOM density of the DeepCore sub-detector. Thus, for some variables the DeepCore DOMs are excluded.

Angular Error Estimation: the final step of the filtering and reconstruction procedure consists in providing an estimate of the per-event angular uncertainty, i.e. the uncertainty given to the reconstructed direction in terms of an angle. Three different methods are available: the Cramér-Rao Estimation, the Paraboloid Method and the Bootstrapping.

1. The Cramér-Rao Estimation exploits the Cramér-Rao inequality [83, 84] stating that the best attainable variance of an unbiased estimator is bounded by the inverse of the Fisher information I. Defining $\vec{x} = (\vec{r}_0, \theta, \phi)$ the parameters that describe a muon track, the elements of the muon track reconstruction covariance matrix are bounded by

$$Cov(\vec{x}_i, \vec{x}_k) \geq I(\vec{x})^{-1}_{ik}. \tag{4.15}$$

The Fisher information matrix I is defined as

$$I(\vec{x})_{ik} = -\left\langle \frac{\partial^2}{\partial \vec{x}_i \partial \vec{x}_k} \, log\mathcal{L}(\vec{x}|t_{res}) \right\rangle_{t_{res}}. \tag{4.16}$$

The square root of the elements in the diagonal of the covariance matrix estimate the azimuth and zenith variances σ_ϕ^2 and σ_θ^2, respectively. They are combined to estimate the circularised error [85]:

$$\sigma = \sqrt{\frac{\sigma_\theta^2 + \sigma_\phi^2 \, sin^2\theta}{2}}. \tag{4.17}$$

Analytic expressions for the covariance matrix are found in [86]. Thanks to the latter and for the fact that this approach does not use a numerical minimisation this is the fastest and most stable algorithm. However, this is the least precise method [10].

2. The Paraboloid Method exploits the shape of the negative log likelihood profile around its minimum to estimate the angular uncertainty. The shape around the minimum is expected to resemble a paraboloid, i.e. a 2-dimensional Gaussian. The likelihood value is estimated for 24 points around the minimum (3 different zenith angles, each with 8 equally spaced azimuth angles, and the negative log likelihood is minimised with respect to $\vec{r}_0$) and a paraboloid is fitted to these points. The Gaussian shape ensures that the standard deviation with respect to a parameter x is

$$-\, log\mathcal{L}(x \pm \sigma_x) = -log\mathcal{L}(x) + 0.5 \tag{4.18}$$

and the paraboloid fit permits to estimate the semiaxes σ_x, σ_y [87] of the ellipse defined by Equation 4.18. Finally, the semiaxes are combined into [85]

$$\sigma = \sqrt{\frac{\sigma_x^2 + \sigma_y^2}{2}}\,. \tag{4.19}$$

Due to the 24 additional minimisations, this approach is slower than Cramér-Rao but the angular uncertainty estimate is better [10].

3. The Bootstrapping method is useful when the theoretical distribution producing the observation is not known and it exploits a random resampling with replacement to the observation. The sample of the event DOM pulses $X_0 = (x_1, ..., x_k)$ defines the track empirical distribution function $\hat{F}$ of the track theoretical distribution F. Bootstraped events

$X_b = (x_1^*, ..., x_k^*)$, $b = 1, ..., n$ are generated by randomly sampling with replacement n samples from the charge-weighted set of pulses in the original event until the total charge of the bootstrapped event is equal to the one of the original event. This is repeated 6 times in this online analysis. A set of tracks is created by maximising the likelihood, each one by reconstructing a track for a different bootstrapped event. From this set the average direction is estimated and the median angular difference of the bootstrapped events is the estimate of the angular uncertainty [10, 88]. The performance is similar to the paraboloid method. Again, the maximisations slow down the algorithm convergence with respect to Cramér-Rao.

Depending on the properties of each event one of the three algorithms is selected and used to estimate the angular uncertainty. This choice is supported by the fact that the completion time of the algorithms depends on the features of each event, like the number of hit DOMs N_{Ch}. A per-event approach is used in order to choose the most accurate (but more time-consuming) method that provides results within the time constraints of this part of the reconstruction chain [10]. The Paraboloid and Bootstrapping methods are preferable for the better estimate of the angular uncertainty, but the Cramér-Rao method is much faster, as can be seen in Table 4.2. From the results of a study involving the run time of the algorithms and the time limitations of the online analyses, the following selection is made [10]:

	MuEX $<$ 4 TeV	MuEX $\geq$ 4 TeV
$N_{Ch} < 300$	Paraboloid	Bootstrapping
$N_{Ch} \geq 300$	Paraboloid	Cramér-Rao

Table 4.1: Selection of the angular uncertainty estimator, based on the reconstructed energy (provided by the MuEX method) and the number of hit DOMs N_{Ch}. Taken from [10].

Finally, other corrections are performed, for example to take into account the kinematic angle between the neutrino and the muon, that is $\approx 1°$ at 1 TeV [10], since the IceCube detector is sensitive only to the muon.

The GFU filter is the last step of the filtering and reconstruction procedure producing the data samples sent to Madison via the Iridium satellites and analysed by the analyses discussed in the next section. A huge effort was made to speed up the algorithms or to choose their best combination in order to produce the best reconstructed neutrino-induced muon candidates in a reasonable time to be able to perform online analyses and eventual follow-up observations, sending alerts to other observatories as quickly as possible. In Table 4.2 the median and maximum per-event execution times of some algorithms of the filtering and reconstruction procedure are reported, as well as a comparison with the offline procedure. Even if the execution time of the online processing is much lower than the offline one, differences in the sensitivity to point-like neutrino sources are found at the percent level [10].

Name	Median (s)	Maximum (s)
SPE	0.04	1.25
MPE	0.03	2.71
SplineMPE	0.04	2.68
MuEX	0.06	0.53
Cramér-Rao	0.01	0.03
Paraboloid	0.24	10.6
Bootstrapping	0.17	4.7
Online Total	0.99	14.83
Offline Total	7.59	451.79

Table 4.2: Per-event execution times of some algorithms of the filtering and reconstruction procedure. In the first column, the name of the algorithm is reported. In the second and third columns, the median and maximum execution times are listed, respectively. Adapted from [10].

A part for a set of numbers useful for identification, each event in the final sample is defined by the arrival time, the reconstructed direction, the angular uncertainty and the reconstructed energy. They are serialised using the JavaScript Object Notation (JSON) format, compact and human readable [89]. Event summaries occupy around 1.3 kB of memory per event, while the full event containing the raw waveforms can reach 250 kB per event. The median latency between the time t_1 of the trigger at the beginning of the filtering and reconstruction procedure and the time t_2 defined as the moment when the event is received at the computing center in Madison is 29 seconds for the event summary and 39 seconds for the full event [10]. The events, as well as detector status and control messages like the event rates used to understand whether the data were taken during with a correct functioning of the detector (Section 3.3), are exchanged between IceCube and the computing center in Madison, WI, via the Iridium satellite network [90], with a bandwidth of 2.4 kbit/s per modem.

Each event surviving the filtering procedure is a neutrino-induced muon candidate, but it is not guaranteed that the event filtering procedure selects only this type of events. Figure 4.5 shows the contribution to the data sample of different event types, at the final selection level, obtained with simulations. The data are shown for comparison.

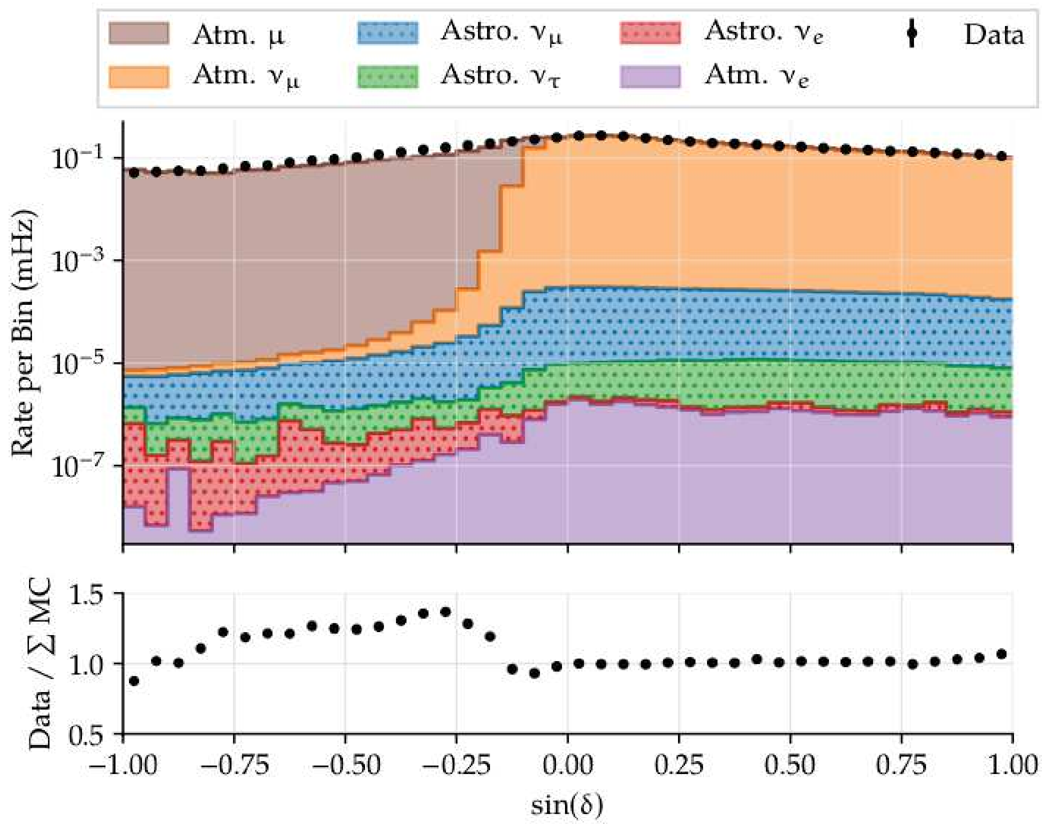


Figure 4.5: Distribution of the different components present in the data sample at the final event selection level, reported as a rate per declination bin. The colored stacked areas are obtained from simulation, while the black dots represent the data. The solid areas indicate the atmospheric contribution, while the dotted areas show the astrophysical contribution. Below, the ratio between data and the sum of the simulations is reported. Taken from [10].

In the Northern sky ($\delta > 0$) the contribution of atmospheric muons is negligible, proving the ability of the filtering procedure to remove mis-reconstructed down-going muons. The dominant component is given by atmospheric muon neutrinos. Then, in order of importance, the sample in this part of the sky contains astrophysical muon, tau and electron neutrinos. The contribution of atmospheric electron neutrinos is the smallest. The simulation agrees well with the data, as the atmospheric neutrinos simulation in this part of the sky has been fitted to the data [91].

In the Southern sky ($\delta < 0$) the sample is almost completely filled with atmospheric muons. The contribution of neutrinos, in order of importance, is atmospheric muon neutrinos, astrophysical muon neutrinos, astrophysical tau neutrinos, astrophysical electron neutrinos and atmospheric electron neutrinos. The event selection chooses more energetic events: for this motivation, the neutrino component is small in this region of the sky and decreases for lower declinations. The agreement between data and simulation is worse than in the Northern sky since a detailed simulation is not available in the Southern sky [10].

Finally, it is interesting to see the effective area for the different neutrino flavors, obtained with simulations. The effective area is the cross section that an ideal detector would exhibit to a neutrino flux and it is used to convert the rate of neutrinos to the neutrino flux (Appendix B). In addition to the flavor, it depends on the neutrino energy and declination of the arrival direction. Figure 4.6 shows the effective area for neutrinos and anti-neutrinos of different flavors, in three declination bands, as a function of the neutrino energy. The effective area value refers to the final event selection level and only events for which the simulated and reconstructed neutrino directions differ by less than 3° are considered.

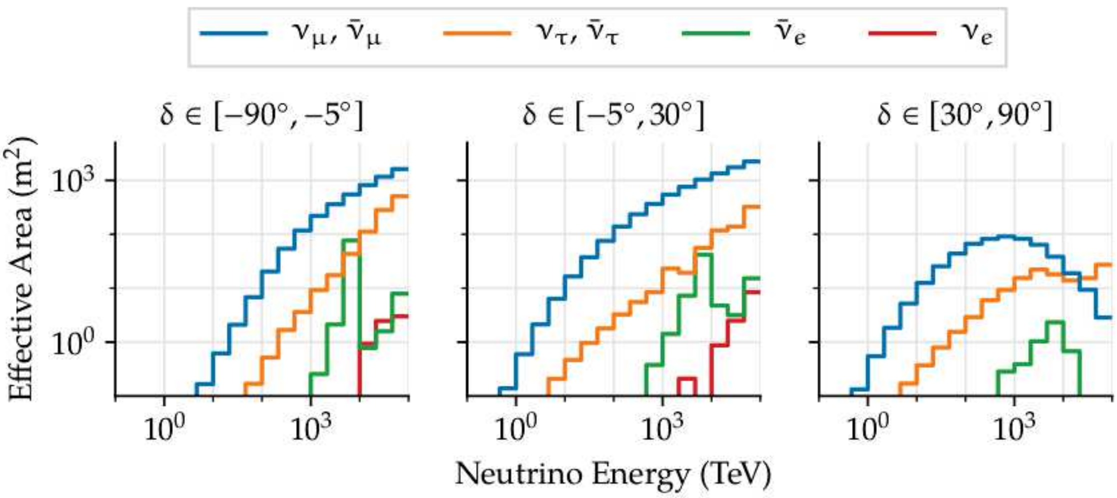


Figure 4.6: Effective area for different (anti-)neutrino flavors at the final event selection level, reported as a function of the neutrino energy for three declination bands. Taken from [10].

The highest value is almost everywhere the one of muon neutrinos and anti-neutrinos, as the event filtering procedure favors this type of events. In the declination band $30° < \delta < 90°$, at high energies the effective area is reduced with respect to the other declination bands because the neutrino interaction cross section increases and more neutrinos are absorbed by the Earth. The effective area for tau neutrinos overcomes the one of muon neutrinos at the highest energies for the so-called "regeneration effect", that renders tau neutrinos more penetrating: tau neutrinos interact with the Earth matter producing a tau lepton, that decays producing another tau neutrino that can reach the IceCube detector. Here, they can interact inside the detector producing a track-like signature, as the decay length of tau neutrinos is $\approx$ 50m/PeV. Even at lower energies, the tau decay can produce a $>$ 200 GeV muon, that is observed. Electron neutrinos have the least effective area in almost all the energy bins of the three different declinations. For electron anti-neutrinos at $\approx$ 6.3 PeV the curve peaks thanks to the Glashow resonance, a process for which the electron anti-neutrinos interact with electrons producing a W boson that can decay into a detected muon [92]. Electron neutrinos and anti-neutrinos have a non-vanishing effective area at high energies since, when producing a hadronic cascade, a daughter muon can escape and leave a track-like signature in the detector.

4.6 Real-Time Analyses and Alerts

In Madison, the real-time analyses are applied to the data sample produced by the event filtering and reconstruction procedure, sent via the Iridium satellites from the South Pole. Two analyses are the most important for this work. The core of the two is a time-dependent algorithm that aims at finding clusters of neutrinos coming from a point-like source in the sky. This work concerns a detailed testing of this algorithm with different simulated source features: the testing procedure and the results are discussed in Chapter 5.

The first analysis concerns the monitoring of a list of known gamma-ray emitters to search for potential neutrino flares. The sources in the list are chosen following some criteria, for example they must have shown time variability in the past. On one hand, gamma-ray flare observations performed online or in the past can provide well constrained time windows to apply the IceCube analysis to search for the neutrino counterpart. However, with this approach, contemporaneous gamma-ray and neutrino data are rare, for example because the Imaging Air Cherenkov Telescopes duty cycle is $\approx$ 20% and due to their few square degrees field of view they can observe essentially only one source at a time. On the other hand, instead, IceCube can observe the entire sky at once with an almost 100% uptime. All the gamma-ray sources present in the list can be observed simultaneously. Thus, the probability to obtain contemporaneous neutrino and gamma-ray data is much higher if the online alerts come from IceCube and not from the IACTs.

Once a neutrino candidate arrives, the probability for that neutrino to be a signal from one of the sources in the list is evaluated and, if this probability is high enough for one of the sources, the analysis starts[3]. In case that a pre-defined significance threshold is exceeded, alerts are sent to IceCube partner telescopes (like MAGIC [4, 5]) that can repoint to the reconstructed direction of the neutrinos to perform follow-up observations. Examples of source candidates are AGN, whose flares last from minutes to weeks. The 339 monitored sources are shown in Figure 4.7 with different symbols, depending on which IceCube partner telescopes can observe that source. Details regarding the alerts and the sources choice, as well as the complete list of the 339 sources, can be found in [10].

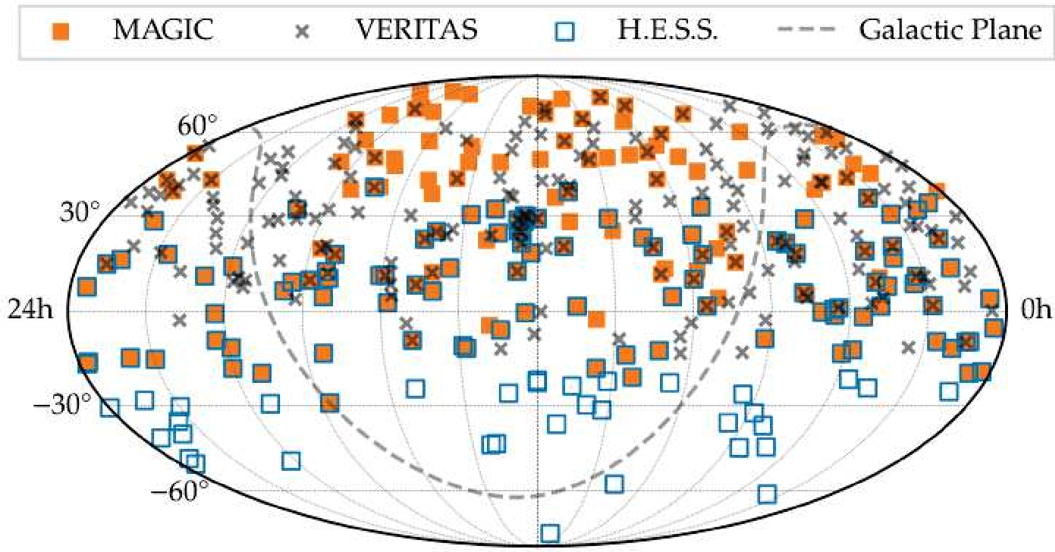


Figure 4.7: Skymap of the 339 monitored gamma-ray sources in equatorial coordinates. Different symbols are used for three IceCube partner Imaging Air Cherenkov Telescopes. Taken from [10].

For the second analysis, the entire sky is scanned to search for neutrino flares providing a tool to reveal sources not considered for time-dependent neutrino emission or even not observed in gamma-rays yet. The large trial factor due to the high number of hypothetical sources lowers the sensitivity with respect to the source monitoring, therefore the two analyses can be run in parallel. Each event is considered as a potential trigger and the area around the event direction is divided in pixels. The analysis is applied to each pixel assuming a source in its center and the pixel with the highest significance is chosen. The latter is divided in sub-pixels and the procedure is iterated until a pre-defined grid size is reached. In case the highest significance on a pixel of the last grid exceeds a

[3] In practice, the event weight w_i discussed in Section 5.3.1, that is the ratio between the per-event signal and background PDFs shown in Section 5.2, is calculated with respect to every source in the list. If $w_i \geq 1$ for a source, the analysis starts with the source hypothesis equal to that gamma-ray source location. The end of the time windows where the likelihood is maximised (see Section 5.4) is kept fixed to the triggering event. The weight is evaluated with an assumed spectral index of -3 to improve the sensitivity to low-energetic neutrino clusters.

threshold, alerts are made public for all the astrophysical community, for example using the Gamma-ray Coordinates Network (GCN) [93].

The alert format can be human-readable (like emails) or machine-readable (like VOEvent [94]), depending on the necessity. The latter are used for automated follow-up observations cutting the human-in-the-loop [10].

In addition to send alerts, IceCube can receive alerts from other observatories that trigger the analysis on the most recent events to search for a correlated neutrino emission. The observatories can be other neutrino detectors, like ANTARES [32], but also gravitational waves detector, like LIGO [48] and VIRGO [49].

Finally, the online filtering and reconstruction procedure is used as a starting point also for other IceCube analyses. For example, the most-energetic events in the data sample have a high probability to have an astrophysical origin. They can be selected and each one is treated as a valid single signal neutrino [10].

Chapter 5

Analysis Testing

The IceCube real-time analyses considered in this work aim at finding clusters of signal neutrinos coming from a source in the sky. Both the monitoring of known gamma-ray sources and the all-sky search (Section 4.6) exploit a time-dependent algorithm, that is run on the output of the event filtering and reconstruction procedure applied on the IceCube DOMs signals.

The details of the algorithm are explained in this chapter. In Section 5.1 the properties of the data sample and the Monte Carlo simulation of the events are discussed. Then, the spatial and energy Probability Density Functions (PDFs) for signal and background events used to build the analysis likelihood are shown in Section 5.2. Before the discussion of the time-dependent likelihood analysis and the results of its testing presented in Section 5.4, a time-integrated version (not used in the real-time infrastructure but useful to understand the more complicated time-dependent version) is examined in Section 5.3.

5.1 Dataset and Monte Carlo Simulations

The data sample used for the analysis tests is the output of the event filtering and reconstruction procedure presented in Chapter 4, that selects a set of N muon neutrino candidates producing a muon leaving a through-going track-like signal in Icecube. Unless otherwise specified, the data used in the following were taken during the 2014 season. Nominally the season goes from the 10^{th} of April 2014 to the 18^{th} of May 2015, but the first runs are used as test runs and are removed for the data analysis, thus the good events start from the 6^{th} of May 2014. The number of events present in the data sample is 183 314.

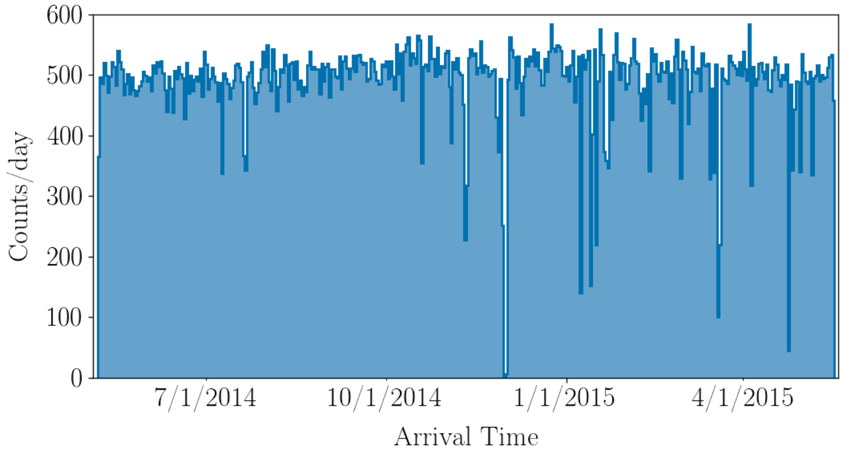
Besides the identification numbers, such as event and run ID, each event in the dataset is characterised by the following reconstructed physical quantities:

Event time: the arrival time of the event reported in Modified Julian Days (MJD, see Appendix A for the definition). The distribution of the event times, converted in month/day/year, is reported in Figure 5.1(a). The histogram resembles a uniform distribution, a part for some days for which the number of events is significantly reduced. This is due to the stop of the data taking, for example because of problems with the power supply, as explained in Section 3.3.

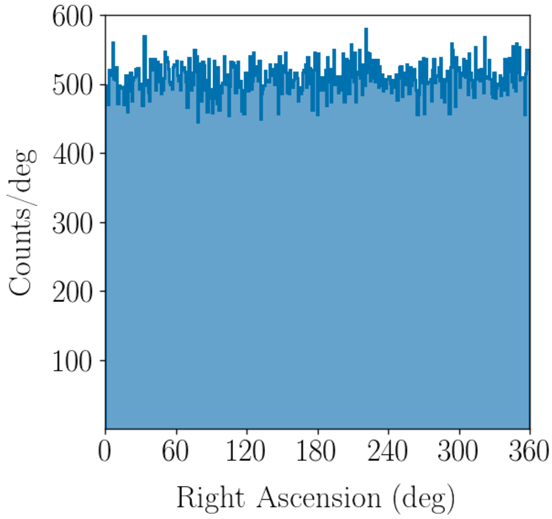
Reconstructed direction: the reconstructed right ascension and declination celestial coordinates, as well as the reconstructed zenith and azimuth angles of the local detector coordinates, reported in radians (see Appendix A for the definitions). The distributions of right ascension and declination, converted in degrees, are reported in Figure 5.1(b) and (c), respectively. The right ascension is approximately uniform. The declination increases while approaching to the Earth equator, peaking at $\approx 6°$. This is due, for example, to the cuts of the filtering and reconstruction procedure.

Estimated angular uncertainty: the angular uncertainty associated to each event, reported in radians. The distribution of the uncertainty, converted in degrees, is reported in Figure 5.1(d). Most of the events have an angular uncertainty below 1° and the distribution of this physical quantity quickly falls for higher values.

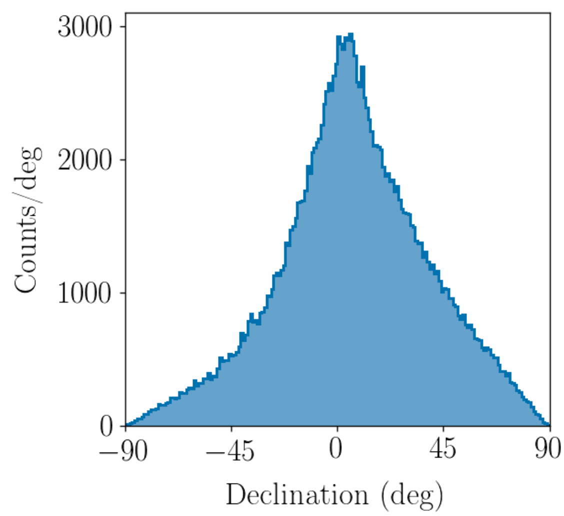
Reconstructed energy: the energy estimation, reported in log(E/GeV). The distribution of this quantity is reported in Figure 5.1(e). Most of the events have an energy of $\approx$ 3 TeV, but the distribution is not negligible up to $\approx$ 6 PeV.



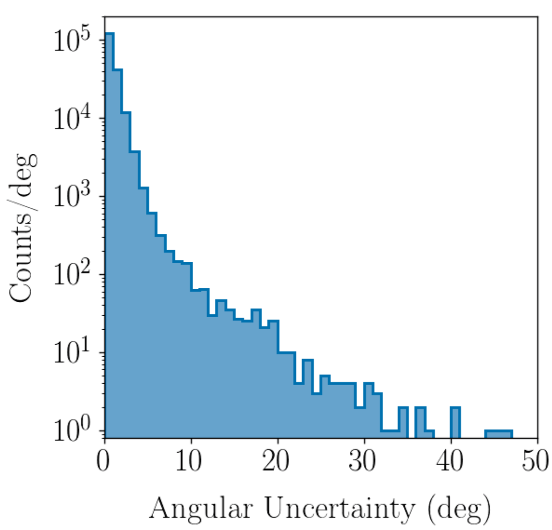
(a) Arrival time (month/day/year).



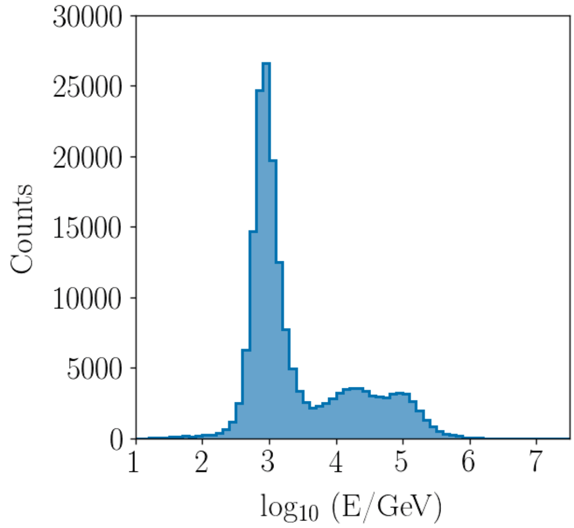
(b) Right ascension.



(c) Declination.



(d) Angular uncertainty.



(e) Energy.

Figure 5.1: Distribution of the event physical quantities for the season 2014 data sample. From (a) to (e): arrival time, right ascension, declination, angular uncertainty and energy.

The MC simulation of the events has been discussed in Section 3.2. The simulation file contains the same information of the real data file plus the true values of the physical quantities chosen at the beginning of the simulation (like true direction and true energy of neutrinos) and the per-event weights.

In Figure 5.2 the median angular error of the reconstruction algorithms, i.e. the angle between the reconstructed and the true neutrino arrival directions, is shown as a function of the neutrino energy.

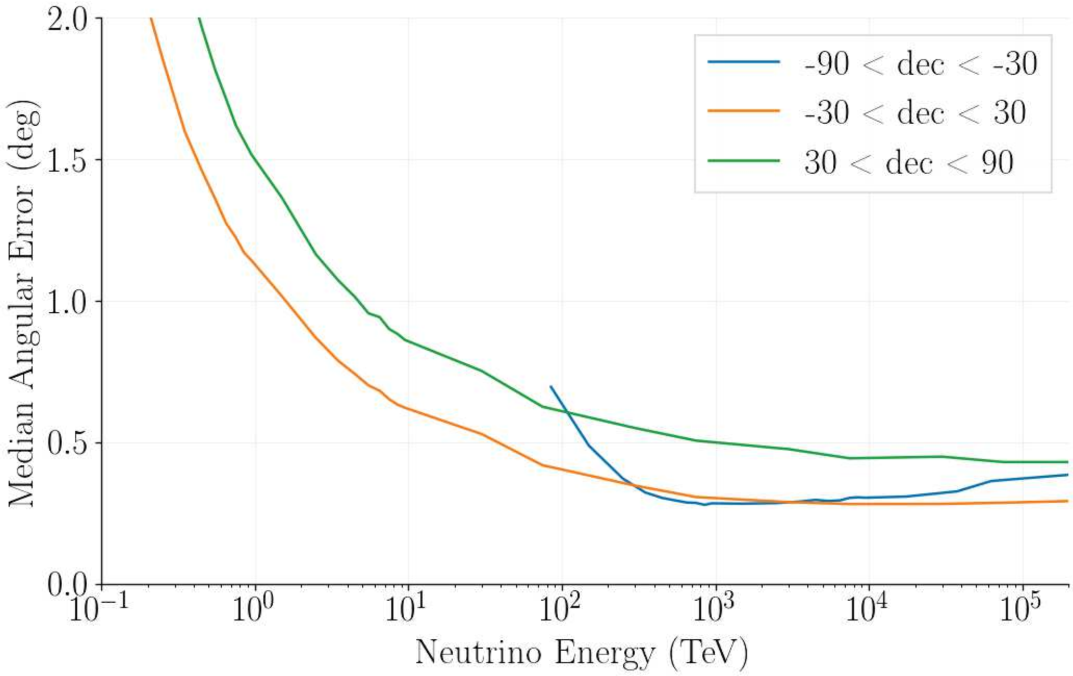


Figure 5.2: Median angular error of the reconstruction algorithm in degrees, as a function of the simulated neutrino energy in TeV, reported for three declination bands.

Figure 5.2 shows that at low energies the median angular error is higher. This happens because of the low number of available pulses and the higher kinematic angle between the neutrino and muon directions. By increasing the neutrino energy, the photon statistics increases and the muon is almost collinear with the neutrino. At the highest energies the error reaches a plateau or increases since the more frequent stochastic energy losses of the muons are not entirely taken into account by the reconstruction algorithms. Finally, the curve for the declination band $-90° < \delta < -30°$ starts around 100 TeV because in the southern sky only the most energetic events are selected by the event selection and filtering algorithms, since in this part of the sky the atmospheric muon contribution is the most important [10].

5.2 Probability Density Functions

The likelihood used by both the time-integrated and the time-dependent algorithms is built by using two per-event Probability Density Functions (PDFs): the signal probability $\mathcal{S}_i$ and the background probability $\mathcal{B}_i$, where the index i ranges from 1 to the number of events N present in the data sample.

The expression of the two PDFs is separated into a spatial PDF and an energy PDF:

$$\mathcal{S}_i = P^S_{spatial}(\vec{x}_i|\vec{x}_S, \sigma_i) \cdot P^S_{energy}(E_i|\vec{x}_i, \gamma) \tag{5.1}$$

$$\mathcal{B}_i = P^B_{spatial}(\vec{x}_i) \cdot P^B_{energy}(E_i|\vec{x}_i), \tag{5.2}$$

where $\vec{x}_S$ is the hypothetical source position, γ is the spectral index of the signal flux spectrum and $\vec{x}_i$, σ_i and E_i the direction, angular uncertainty and energy of the i-th event, respectively.

The signal spatial PDF represents the probability for an event coming from $\vec{x}_i$ with an angular uncertainty σ_i to be contributed by a source located at $\vec{x}_S$ and it is calculated by using a two-dimensional circular Gaussian function

$$P^S_{spatial}(\vec{x}_i|\vec{x}_S, \sigma_i) = \frac{1}{2\pi\sigma_i^2}\, exp\left(-\frac{\Delta\Psi(\vec{x}_i|\vec{x}_S)^2}{2\sigma_i^2}\right), \tag{5.3}$$

where $\Delta\Psi$ is the angle between the event and the source directions. This expression is well justified by the fact that the atmospheric background events are almost uniformly distributed over the entire sky while signal events should cluster around the source direction. In addition, more weight is given to well reconstructed events having a low angular uncertainty [10].

The background spatial PDF represents the probability for an event to be part of the background, based on its direction. It is constructed using archival data. This PDF is shown in Figure 5.3.

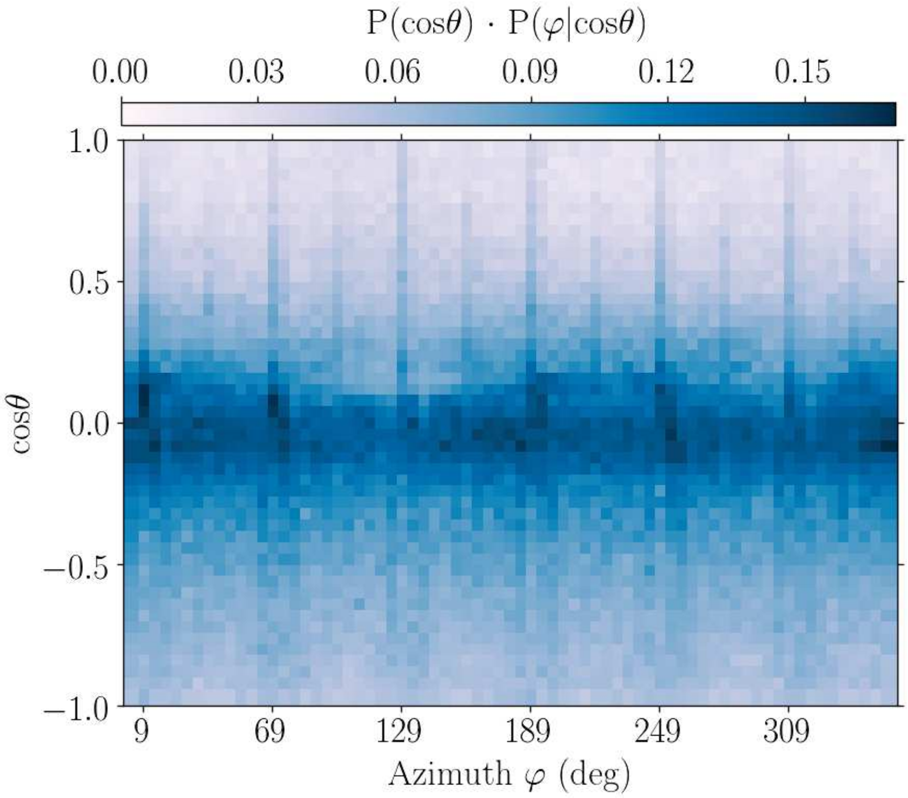
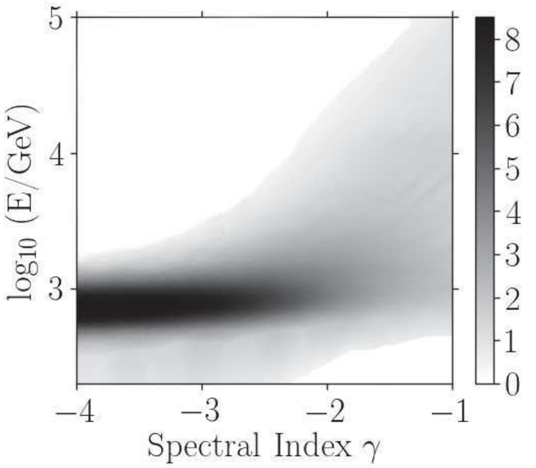


Figure 5.3: The spatial background PDF shown as a function of the cosine of the zenith angle θ and of the azimuth angle ϕ. The azimuth values reported on the abscissa highlight the preferred directions containing more strings.

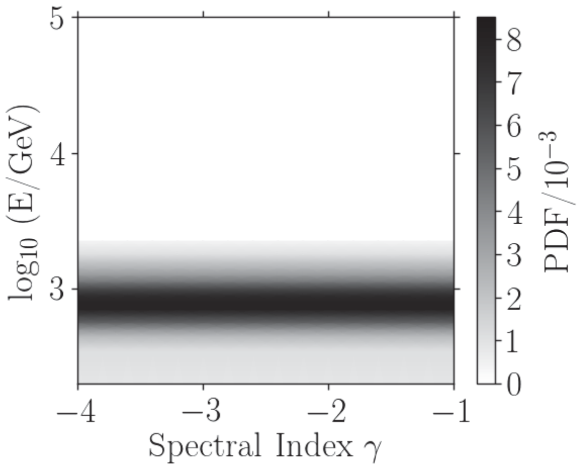
It is clear how the background spatial PDF depends both on the zenith and the azimuth angles. The zenith dependence is due to the different path length in Earth traversed by neutrinos and the zenith-dependent effective area. Then, the secondary cosmic rays distribution in zenith affects the background PDF. In addition, the zenith dependence is partly due to the event filtering procedure, that is more stringent for down-going events coming from the Southern Sky since the atmospheric background contribution is higher. After the filters, $\approx \frac{1}{3}$ of the selected events are down-going events and $\approx \frac{2}{3}$ are up-going events: this can be seen in Figure 5.3 from the more blueish color of the negative $cos\theta$ region with respect to the positive one.

For what concerns the azimuth angle, it is important to note that the instrumented volume of the IceCube detector resembles a prism with hexagonal bases, thus it is not symmetric in the 2π azimuth range. Along six directions a higher number of strings is encountered, resulting in an increased event rate seen as six equally-spaced vertical lines in Figure 5.3, recognisable by the angles reported in the abscissa. This effect is more clear for down-going events as the DOMs are positioned with the PMT facing downwards and require the photons to be scattered: the event selection procedure favors the six directions with a higher number of strings and thus the least amount of scattering [10].

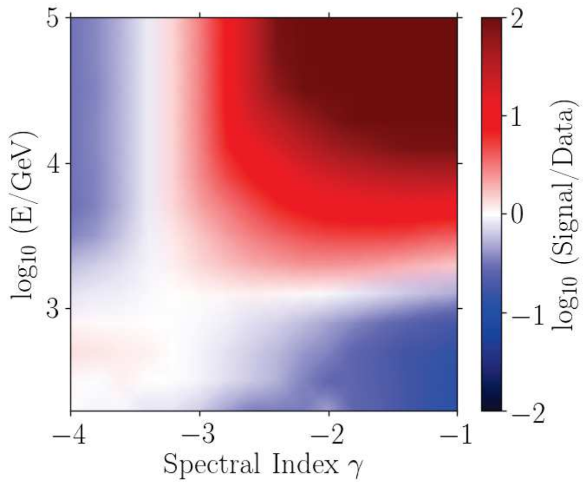
The per-event calculation of the signal and background energy PDFs in motivated by the fact that signal neutrinos are expected to follow a power-law energy spectrum $\sim E^{-2}$ (Fermi acceleration), while atmospheric neutrinos have a power-law energy spectrum $\sim E^{-3.7}$. The two PDFs are created in declination bands, since the distributions depend on this parameter.



(a) Signal energy PDF.



(b) Background energy PDF.



(c) Ratio.

Figure 5.4: The energy PDF construction for the declination band $41.9° < \delta < 45.3°$. The assumed spectral index is reported on the abscissa, the logarithm of the energy in GeV is reported on the ordinate. Up left: the signal energy PDF. Up right: the background energy PDF. Down: the logarithm of the ratio between the signal and background energy PDFs, reported in the color bar as "Signal" and "Data", respectively.

The signal energy PDF is created using signal events simulated as explained in Section 5.1. The spectral indices range from -1, harder than the diffuse astrophysical neutrino flux (with $\gamma \approx -2.19$), to -4, softer than the atmospheric neutrino flux. The result for the declination band $41.9° < \delta < 45.3°$ can be seen in Figure 5.4(a) as a function of the neutrino energy and the assumed spectral index. As expected, for harder spectral indices the high-energy region is populated, while for softer ones the flux vanishes. The background energy PDF is created using archival data and can be seen in Figure 5.4(b). It does not depend on the source spectral index, as expected.

Figure 5.4(c) shows the ratio between signal and background energy PDFs, fitted with a cubic spline. As it will be shown in Section 5.3.1, only the ratio between the signal and background PDFs is used in the analysis, not the single PDFs. As expected, for harder spectral indices and for higher energies there is a higher probability to have signal events.

Figure 5.5 shows the ratio between the signal and background energy PDFs fitted with a cubic spline, as a function of energy and declination, for an assumed spectral index for the signal of $\gamma = -2$. Signal-dominated regions are found, in general, at higher energies or at higher values of the declination.

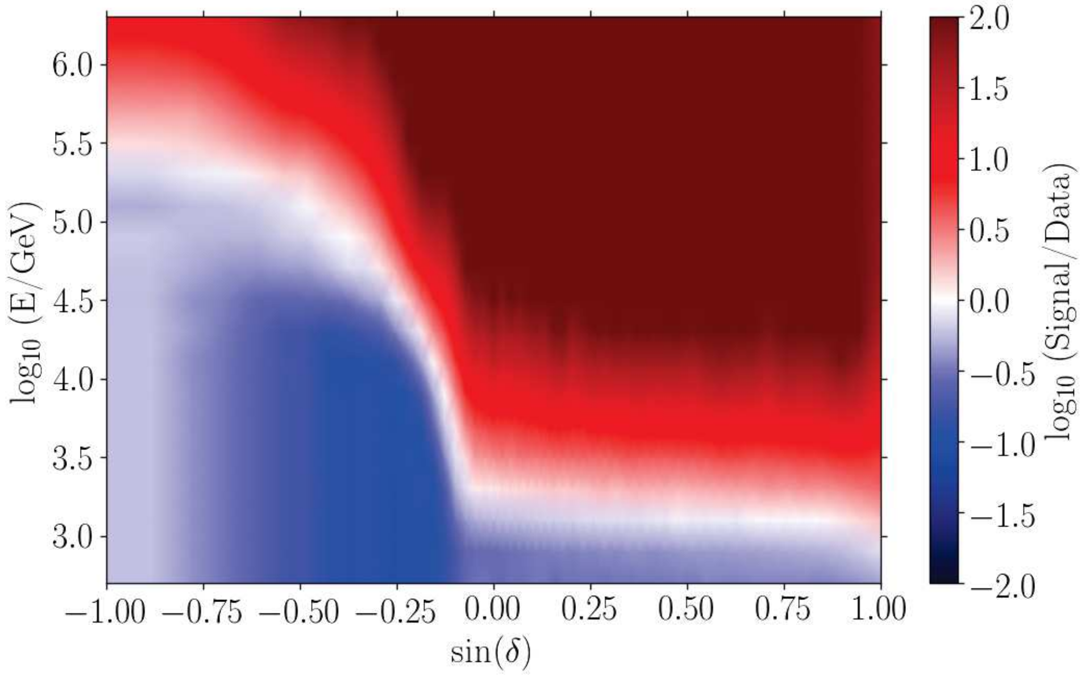


Figure 5.5: The logarithm of the ratio between the signal and background energy PDFs, reported in the color bar as “Signal” and “Data”, respectively. It is shown as a function of the logarithm of the energy in GeV and the sine of the declination and it is calculated for an assumed spectral index for the signal of $\gamma = -2$.

5.3 The Time-Integrated Analysis

The analysis used by the real-time infrastructure for the search of point-like transient neutrino sources is a time-dependent analysis. However, a time-integrated version was explored first in this work. This is done because this less sophisticated analysis provides a good example from which the time dependence is naturally implemented and the improvements of the time-dependent approach will be evident. In addition, as it is shown in Section 5.4, the time-dependent analysis exploits the time-integrated one.

Section 5.3.1 will describe the expression of the likelihood used by the time-integrated analysis and the outputs provided by its maximisation. Then, Section 5.3.2 will define the analysis testing procedure and show its results.

5.3.1 Likelihood Formulation

Both the time-integrated and the time-dependent analyses are based on the likelihood maximisation technique. The expressions of the likelihood contain the per-event signal and background PDFs explained in Section 5.2. The likelihood used by the time-integrated analysis is given by:

$$\mathcal{L}(n_S,\gamma) = \prod_{i=1}^{N} \left(\frac{n_S}{N} \cdot \mathcal{S}(\vec{x}_i, \sigma_i, E_i|\gamma) + \left(1 - \frac{n_S}{N}\right) \cdot \mathcal{B}(\vec{x}_i, \sigma_i, E_i) \right), \tag{5.4}$$

where N is the number of events present in the data sample and $\vec{x}_i, \sigma_i, E_i$ are their estimated direction, angular uncertainty and energy of the i-th event, respectively. n_S is the number of signal events in the data sample and γ is their spectral index: they are free parameters estimated by the likelihood maximisation and their fit ranges are $0 \leq n_S \leq N$ and $-4 \leq \gamma \leq -1$ (in agreement with the bounds placed on the energy PDFs) [10]. Equation 5.4 is thus the product over all the events in the data sample of the fraction of signal events $\frac{n_S}{N}$ multiplied by the probability for the i-th event to be a signal event summed to the fraction of background events $1 - \frac{n_S}{N}$ multiplied by the probability for the i-th event to be part of the background.

Rather than directly maximising the likelihood $\mathcal{L}$, defined as a product of several terms, it is computationally less expensive to find the maximum of its logarithm. For this motivation, a quantity called test statistic Λ is calculated exploiting the logarithm of $\mathcal{L}$. Since the logarithm is monotonically increasing, the number of events n_S and the spectral index γ that maximise the likelihood also maximise Λ. The test statistic is given by

$$\Lambda = 2\, log \frac{\mathcal{L}(n_S, \gamma)}{\mathcal{L}(n_S = 0)}. \tag{5.5}$$

Inside the logarithm, the likelihood is divided by the (constant) value of itself under the background assumption, thus imposing $n_S = 0$, and test statistic takes the form of a signal-over-background ratio that is always non-negative. The test statistic can be rewritten as:

$$\begin{aligned} \Lambda &= 2\, log \left[\frac{\prod_{i=1}^{N} (\frac{n_S}{N} \mathcal{S}_i + \left(1 - \frac{n_S}{N}\right) \mathcal{B}_i}{\prod_{i=1}^{N} \mathcal{B}_i} \right] \\ &= 2 \sum_{i=1}^{N} \left[log \left(\frac{n_S}{N} \mathcal{S}_i + \left(1 - \frac{n_S}{N}\right) \mathcal{B}_i \right) - log \mathcal{B}_i \right] \\ &= 2 \sum_{i=1}^{N} log \left[1 + \frac{n_S}{N} \left(\frac{\mathcal{S}_i}{\mathcal{B}_i} - 1 \right) \right] \end{aligned} \tag{5.6}$$

and the ratio between the signal and background PDFs is defined as the weight

$$w_i = \frac{\mathcal{S}_i}{\mathcal{B}_i}. \tag{5.7}$$

The spatial signal PDF shown in Equation 5.3 can be used to simplify the last expression of Λ, as it vanishes for events whose direction is few degrees away from the direction of the hypothetical source. The value of the spatial signal PDF is fixed to zero for $N - N'$ events outside a circular bin of 5 degrees around the source direction, where N' is the number of events inside the bin. The final expression of the test statistic is thus

$$\Lambda = 2 \sum_{i=1}^{N'} log \left[1 + \frac{n_S}{N} (w_i - 1) \right] + 2 \left(N - N' \right) log \left(1 - \frac{n_S}{N} \right) \tag{5.8}$$

This final form of the test statistic is maximised to estimate the number of events n_S and their spectral index γ. The value of Λ is finally calculated by plugging these estimates into Equation 5.5.

Figure 5.6 summarises the time-integrated analysis. After selecting the data sample and the hypothetical source direction, and after producing the PDFs exploiting data and Monte Carlo simulations, the time-integrated analysis provides the three outputs Λ, n_S and γ. The ability of the analysis to retrieve the correct cluster parameters has to be tested by using toy simulations. These tests, showing the behaviour of the analyses in different conditions, are the core of this work. The next Section describes the testing procedure of the time-integrated analysis.

5.3.2 Test procedure and Results

The analysis tests exploit the so-called toy simulations, i.e. simulations of different signal and background scenarios.

Firstly, the hypothetical source right ascension α_{source} and declination δ_{source} are chosen. Then, a background only scenario is created using a data sample, like season 2014. The three most energetic events in a declination band of $\pm 5°$ around δ_{source} are excluded. The event times are randomly shuffled ("scrambling") and the local detector coordinates zenith and azimuth are converted to equatorial

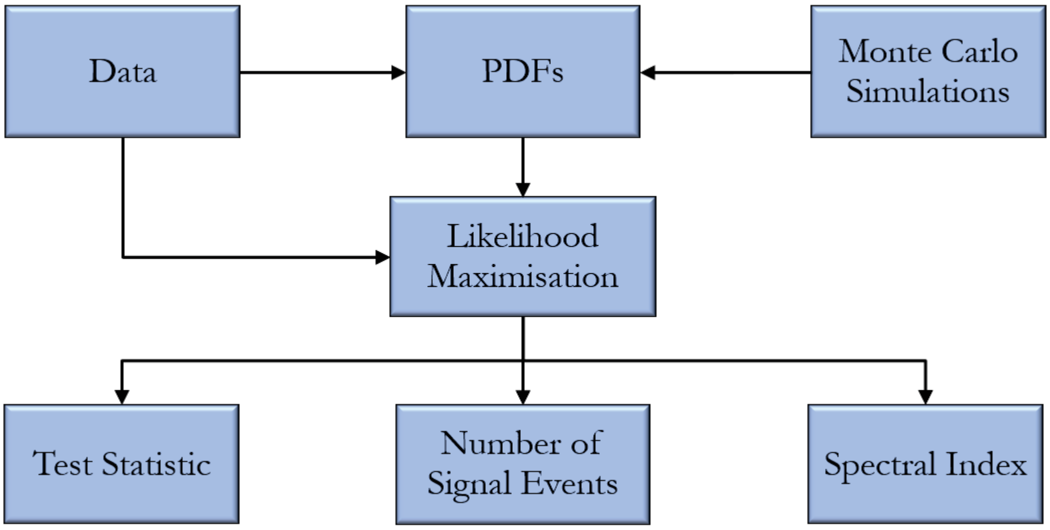


Figure 5.6: A summary of the time-integrated analysis. Data and Monte Carlo simulations (Section 5.1) are exploited to produce the spatial and energy PDFs (Section 5.2). Then, data and PDFs are used in the likelihood maximisation procedure. The algorithm produces three outputs: the test statistic, related to the value of the likelihood, the number of signal events and their spectral index.

coordinates, function of time. Being the detector at the South Pole, the declination is the same before and after the time scrambling and the right ascension is approximated with

$$\alpha_{\text{new}} = \left(\alpha_{\text{old}} + (\text{t}_{\text{new}} - \text{t}_{\text{old}}) \cdot \frac{2\pi}{\text{T}_{\text{sid}}}\right) \bmod 2\pi, \tag{5.9}$$

where $\alpha_{\text{new(old)}}$ is the new (old) right ascension, $\text{t}_{\text{new(old)}}$ is the new (old) event time and $\text{T}_{\text{sid}} = 0.99726957$ days is the length of a sidereal day [95]. This procedure is used to mask any possible signal event cluster in the dataset. This approach yields a set of background-like events embedding detector effects without using detector simulations.

Next, simulated signal events are injected in the background set. The number of injected events is drawn from a Poisson distribution. In addition, the events are injected such to resembles a power-law flux with spectral index γ_{inj}. In IceCube, the Monte Carlo simulations are produced for a diffuse flux of neutrinos with a fixed spectral index, but can be eventually weighted to resemble a different spectrum. The events are drawn from the Monte Carlo simulation explained in Section 5.1 from a declination band of $\pm 5°$ around δ_{source}. Also, the events are uniformly injected in a time window centered at time t_0, where t_0 is uniformly drawn between the minimum and maximum arrival time of the events in the data sample. Finally, the injected events direction is moved towards the source location. This is equal to simulate a point-like source [10].

Then, the test statistic is maximised, providing an estimate of the number of signal events n_{fit}, their spectral index γ_{fit} and the value of the test statistic. All this procedure is repeated for a very large number of times, each time with different scrambling and injection, resulting in a set of results that should resemble the injected quantities. The tests consist in estimating the power of the analysis to retrieve the injected physical quantities. By changing the injection parameters, like γ_{inj}, a source with different features or emitting a different flare is simulated and the behaviour of the algorithm is studied for various cases.

In the following, the behavior of the time-integrated analysis is shown. A point-like neutrino source is simulated at $\delta_{source} = 45°$ and $\alpha_{source} = 300°$, unless otherwise specified, and the dataset used is season 2014.

The test statistic maximisation does not converge every time. Figure 5.7 shows the convergence percentage of the time-integrated analysis versus the number of injected events, for three different injected spectral indices. By increasing the number of injected events, as well as by hardening the spectrum, the number of successful fits increases. This is due to the fact that for a signal cluster containing more events or for those events having a spectral index further away from the background

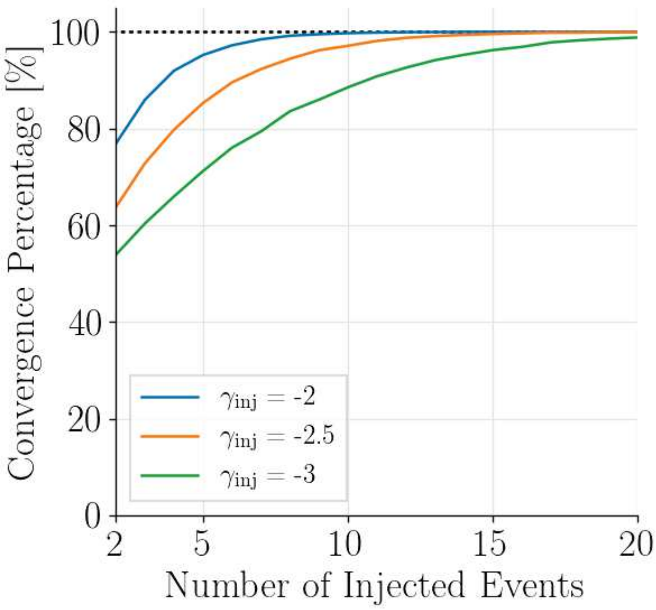


Figure 5.7: Convergence percentage for the time-integrated analysis as a function of the number of injected events, reported for three different spectral indices. The blue (orange, green) curve is related to γ_{inj} = -2 (-2.5, -3).

one (≈ -3.7) it is easier for the algorithm to converge. In general, the convergence percentage increases for lower values of source declination. When producing any results, the cases of no-convergence are excluded.

Before discussing the ability of the analysis to retrieve the injected parameters, it is useful to have a look at the distributions of the output parameters. In the following figures, the distributions are shown for two injected spectral indices (-2 and -3) and for three different number of injected events (2, 10 and 20).

Figures 5.8 and 5.9 show the distribution of the test statistic compared to the background distribution. The results relative to the background are obtained injecting events with $n_{inj} = 0$. The background test statistic distribution is peaked at zero because for a number of injected events $n_{inj} = 0$ the signal over background likelihood ratio tends to 1. By increasing the number of injected events the signal likelihood $\mathcal{L}(n_S, \gamma)$ increases and consequently the test statistic distribution moves to higher values. For the same number of injected events the distribution connected to $\gamma_{inj} = -2$ reaches higher values of the test statistic with respect to $\gamma_{inj} = -3$: this behavior is expected as a harder spectrum is more signal-like and the signal likelihood is higher.

Figures 5.10 and 5.11 show the distribution of the fitted number of signal events. When trying to estimate the number of signal events for a low n_{inj} the distribution of n_{fit} is largely asymmetric because of the fit range imposing the fitted number of signal events to be greater than or equal to zero. By injecting more signal events, the distribution is more symmetric and moves towards higher values of n_{fit}. For a high number of injected events and $\gamma_{inj} = -2$ the distribution peaks near the expected value, while for $\gamma_{inj} = -3$ the peak underestimates the number of injected events. This happens because the injected events, having a softer spectrum, are more similar to the background. Finally, the distributions are wide and reach values of the fit parameter far away from the injected one. This suggests that a deeper study and improvement of the test statistic maximisation (like a change of the minimisation algorithm) should be carried out in order to narrow the distributions of the fit parameters. This behaviour is present in all the tests carried out for this work.

Figures 5.12 and 5.13 show the distribution of the fitted spectral index. For $\gamma_{inj} = -2$ and $n_{inj} = 2$ the curve is broad and does not show a clear peak. By increasing the number of injected events, the distribution slowly peaks to the expected value, even if it spans almost the entire fit range and the peak is broad. For $\gamma_{inj} = -3$ there is little improvement by increasing the number of injected events: the distribution remains very broad, while having its maximum values near $\gamma_{fit} = -3$.

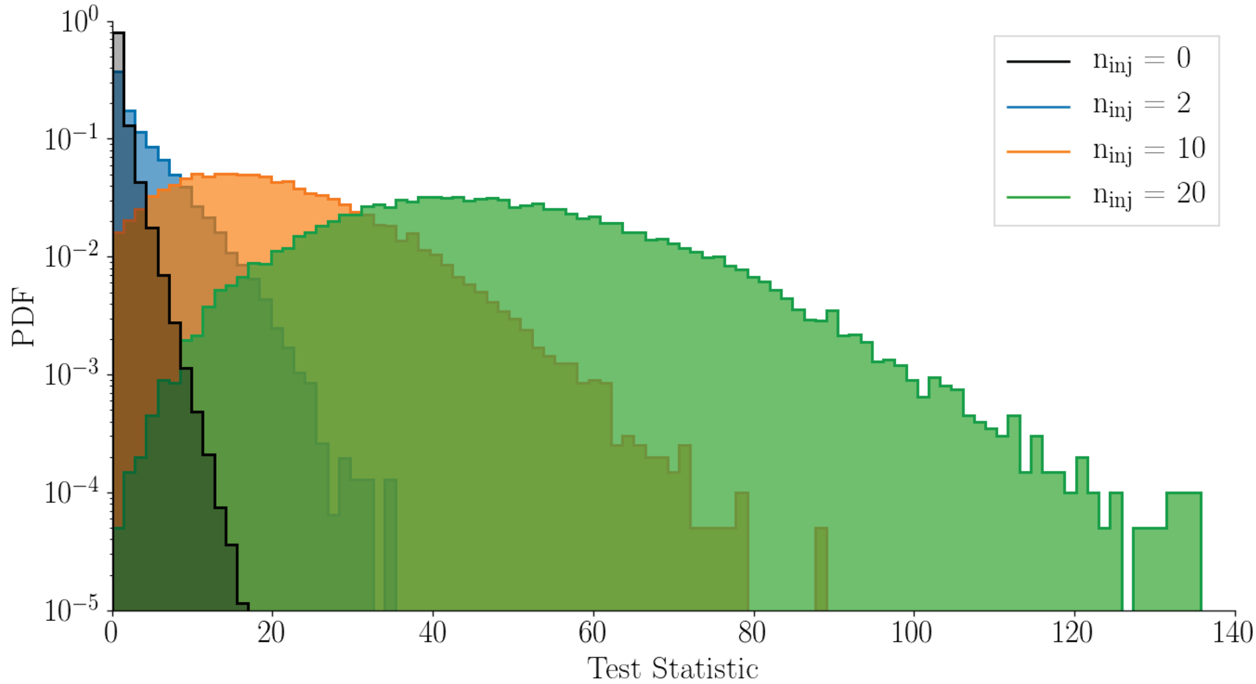


Figure 5.8: Distribution of the test statistic for the time-integrated analysis for an injected spectral index of -2. The black histogram represents the background distribution, while the blue, orange and green histograms represent the distribution for a number of injected events of 2, 10 and 20, respectively.

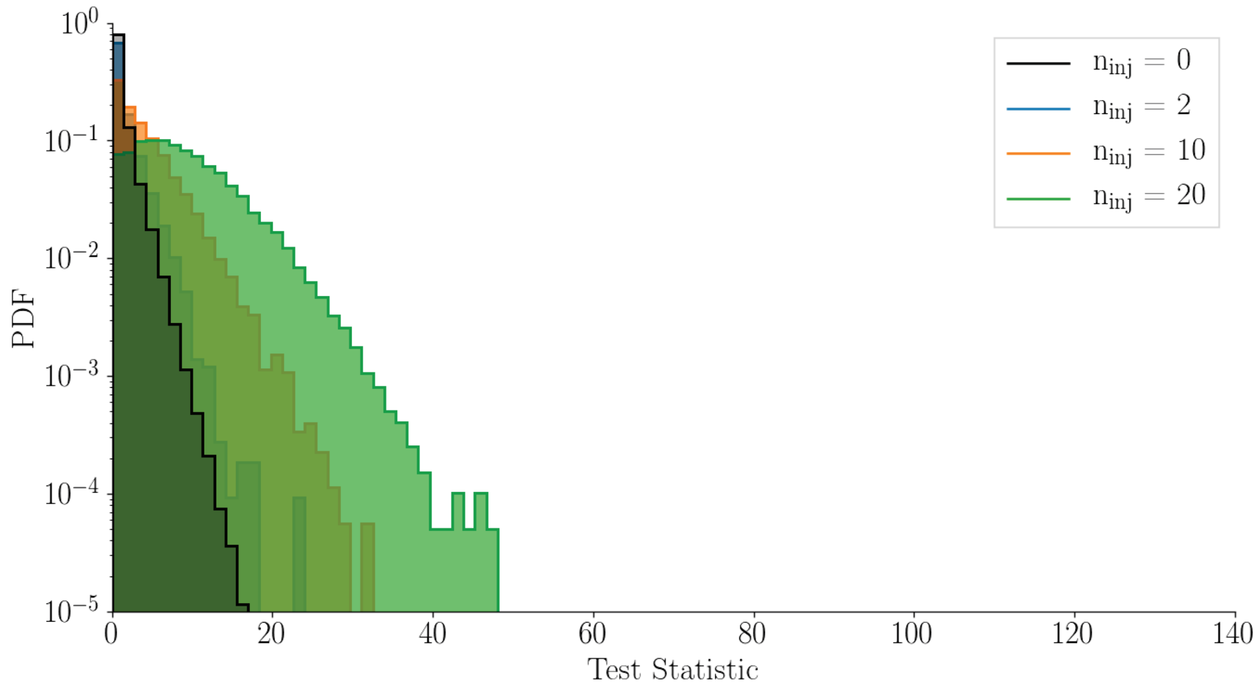


Figure 5.9: Distribution of the test statistic for the time-integrated analysis for an injected spectral index of -3. The black histogram represents the background distribution, while the blue, orange and green histograms represent the distribution for a number of injected events of 2, 10 and 20, respectively.

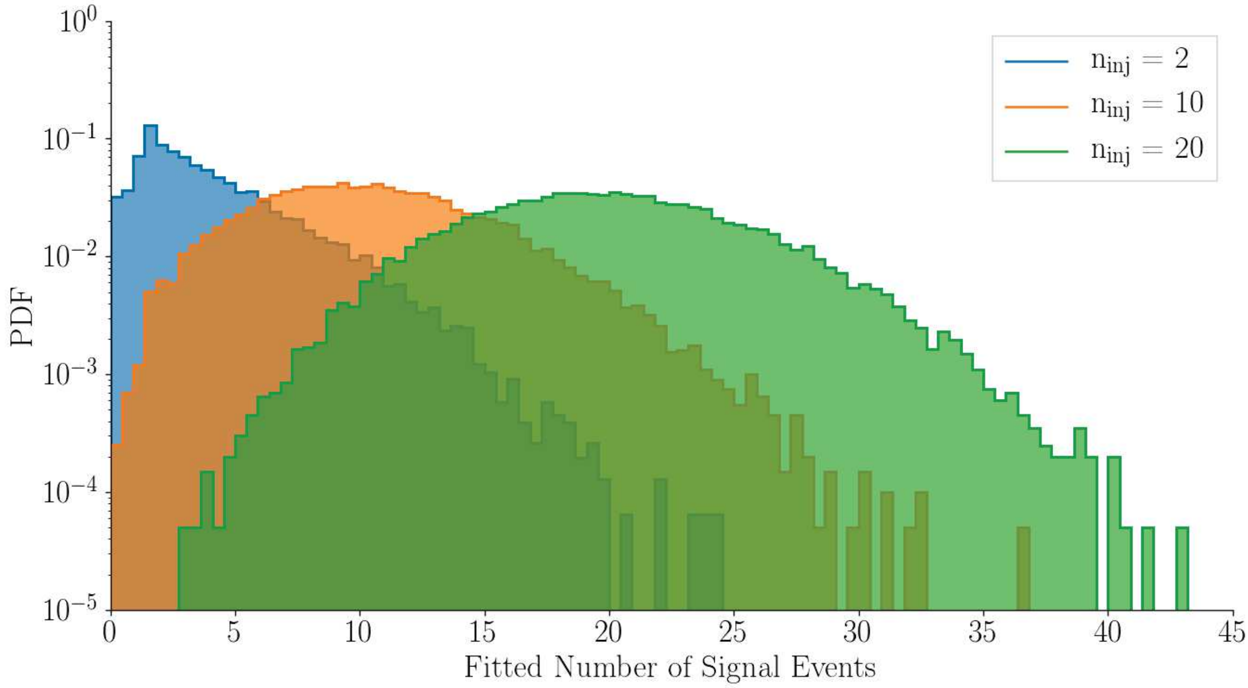


Figure 5.10: Distribution of the fitted number of signal events for the time-integrated analysis for an injected spectral index of -2. The blue, orange and green histograms represent the distribution for a number of injected events of 2, 10 and 20, respectively.

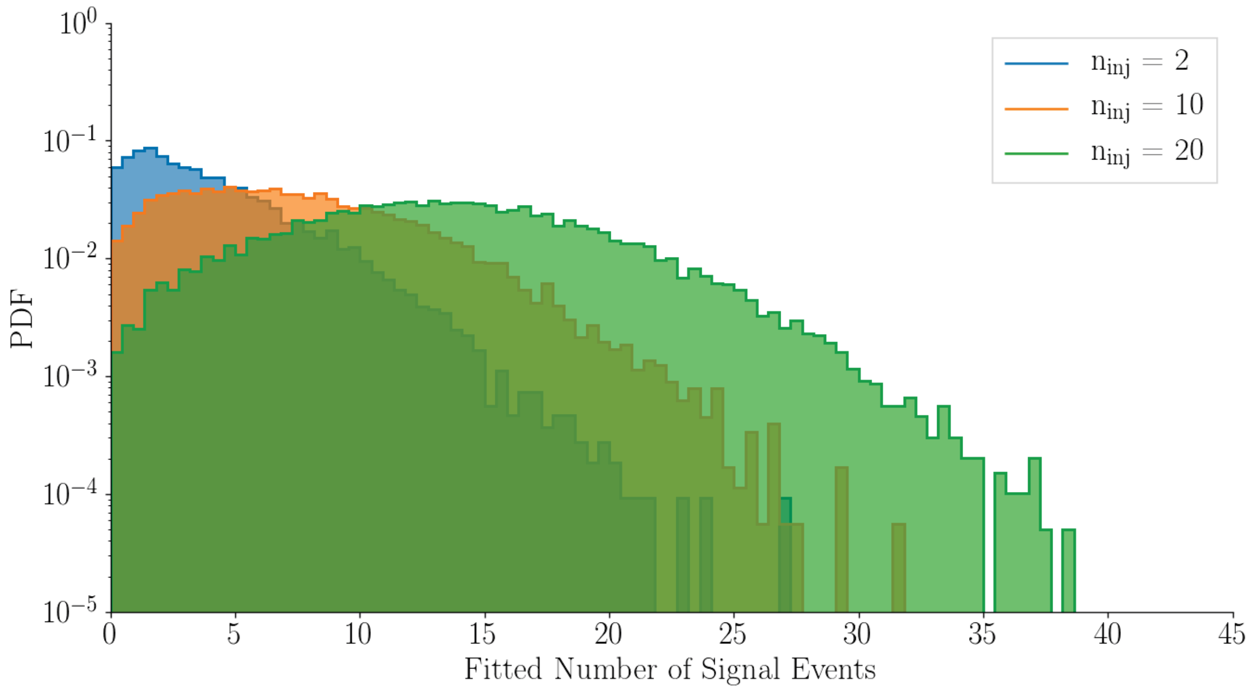


Figure 5.11: Distribution of the fitted number of signal events for the time-integrated analysis for an injected spectral index of -3. The blue, orange and green histograms represent the distribution for a number of injected events of 2, 10 and 20, respectively.

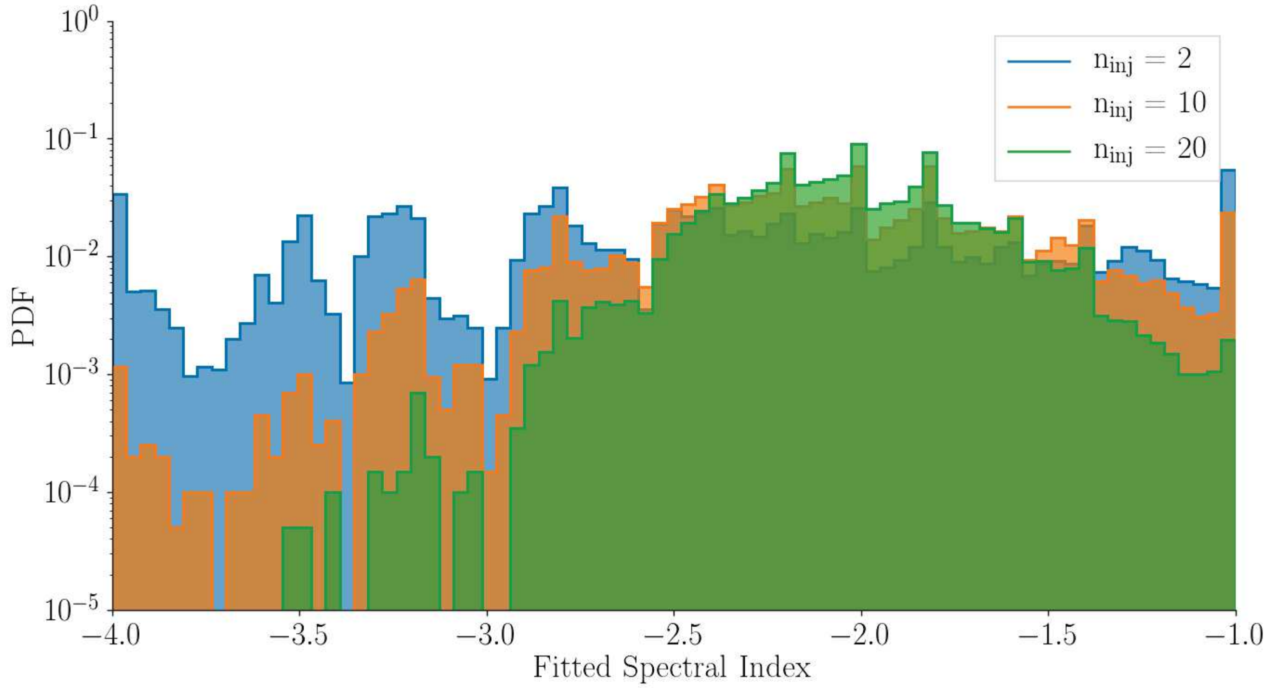


Figure 5.12: Distribution of the fitted spectral index for the time-integrated analysis for an injected spectral index of -2. The blue, orange and green histograms represent the distribution for a number of injected events of 2, 10 and 20, respectively.

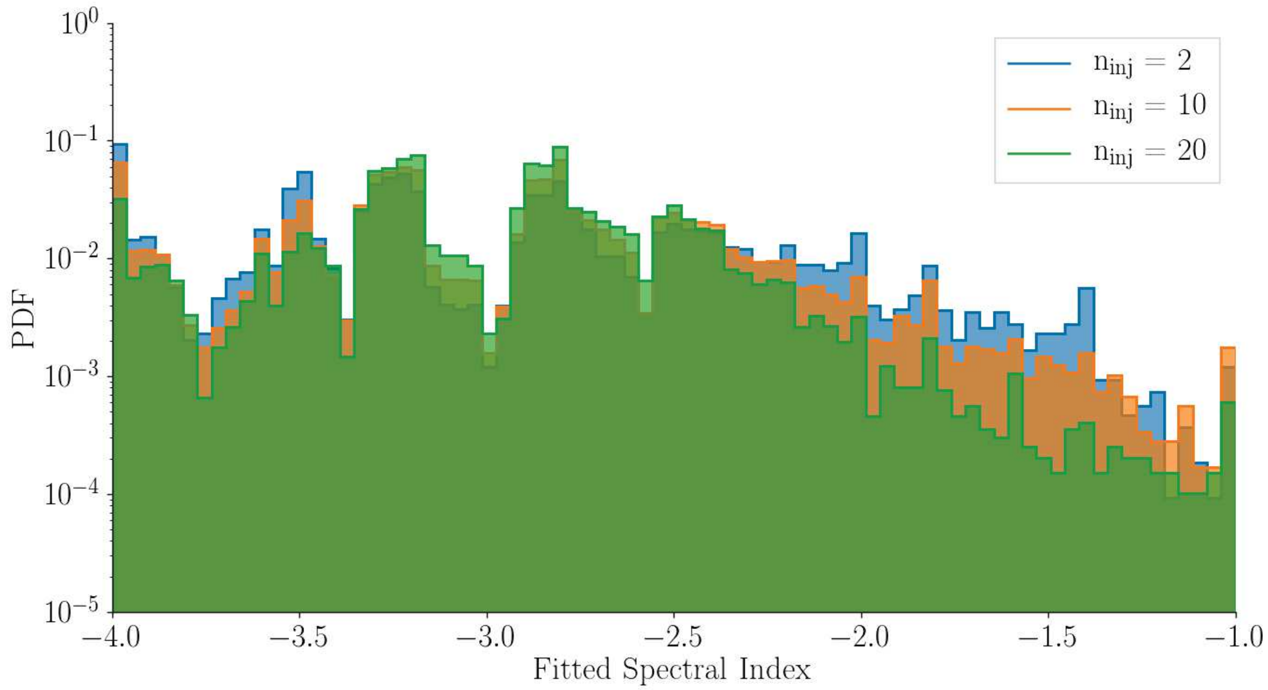
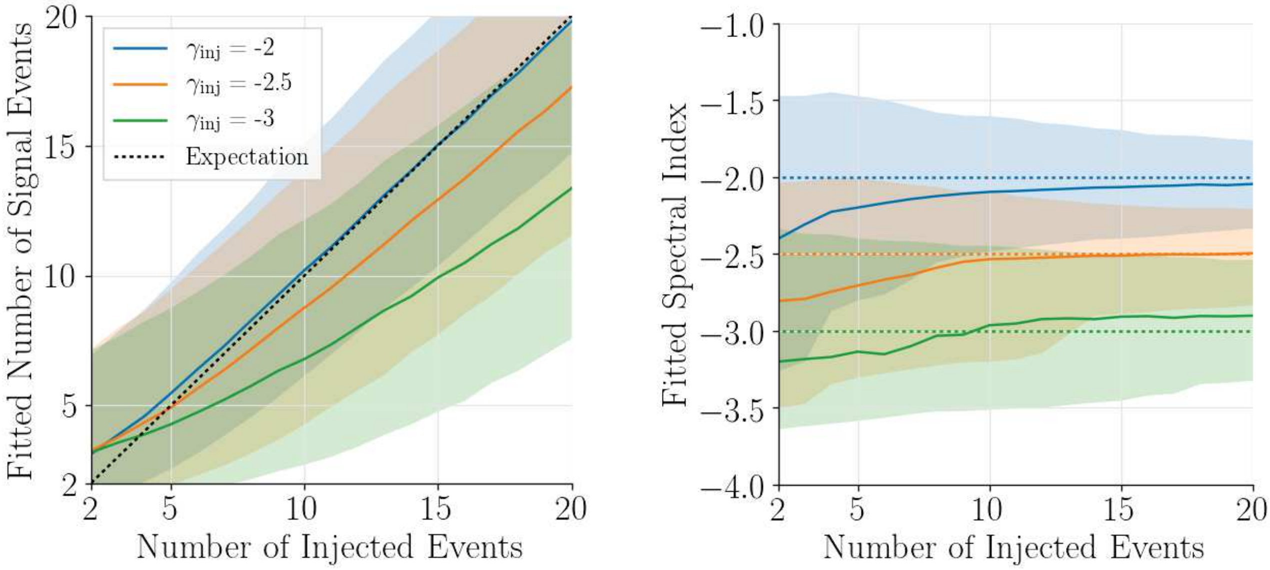


Figure 5.13: Distribution of the fitted spectral index for the time-integrated analysis for an injected spectral index of -3. The blue, orange and green histograms represent the distribution for a number of injected events of 2, 10 and 20, respectively.

Figure 5.14 shows the reconstruction quality of the injected parameters versus the number of injected events. The number of signal events is overestimated for a low number of injected events.



(a) Fitted number of signal events.

(b) Fitted spectral index.

Figure 5.14: Reconstruction quality of the fit parameters. The continuous lines represent the median of the parameter distribution, the colored area is the 68-percentile centered in the median and the dotted lines show the expectation. On the abscissa there is the number of injected signal events. The blue (orange, green) curve is related to γ_{inj} = -2 (-2.5, -3). Left: the fitted number of signal events. Right: the fitted spectral index.

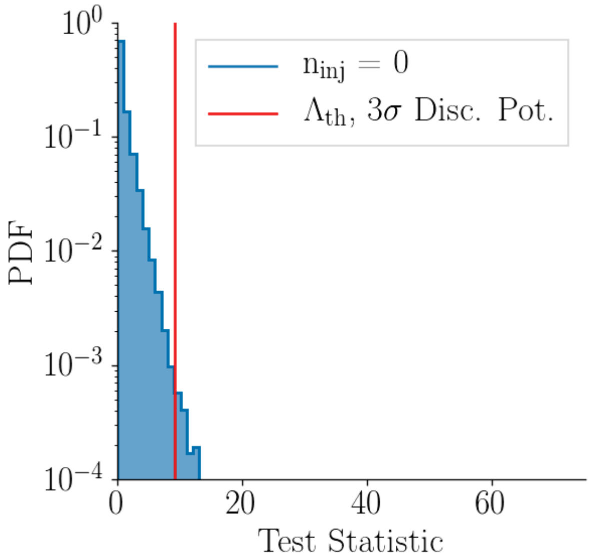
For higher n_{inj} and for $\gamma_{inj} = -2$ the median reconstructed signal events number is close to the expectation. For softer spectral indices the number of signal events is underestimated: this is due to the poor signal-like features of the injection and some signal events are considered as background. Other test results, not reported here, showed that for lower declinations the curves are in general closer to the expectation and the error areas are more narrow. The spectral index is underestimated for a small number of injected events and increases approaching to the expectation for a higher n_{inj}, stabilising near γ_{inj}. Again, the reconstruction quality improves with a higher number of injected events. In general, for δ_{source} = -45° (not shown) there is an improvement because of the more closeness of the curves to the expectation or a more narrow error area. The error areas is always very wide.

Besides the reconstruction quality of the fit parameters, another very important quantity connected to the searches for point-like astrophysical neutrino sources is the discovery potential. It is defined as the flux of neutrinos which yields a discovery with a certain significance value in 50% of the cases. For example, a 5σ discovery potential is the flux that yields a 5σ discovery in 50% of the cases. Different steps needs to be performed in order to calculate this quantity.

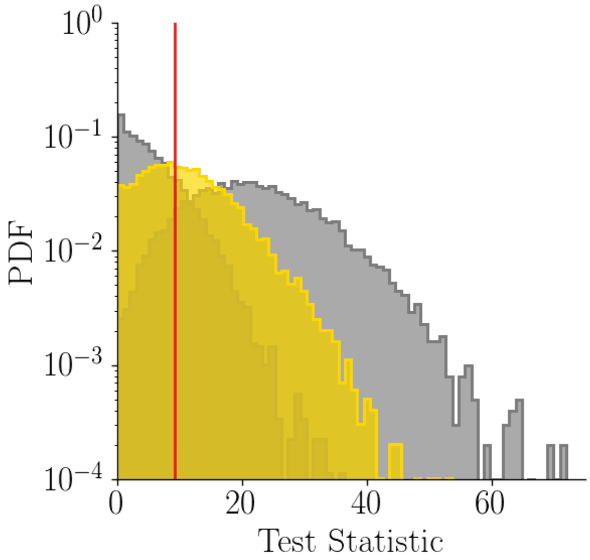
Firstly, a toy simulation without injection (i.e. with a number of injected events $n_{inj} = 0$) is used to define a threshold on the background test statistic distribution connected to the desired significance value for the discovery. This threshold Λ_{th} is such that the p-value, i.e. the probability of observing an equal or better result by chance with respect to one having a test statistic of Λ_{th}

$$p(\Lambda_{th}) = \int_{\Lambda_{th}}^{\infty} P^B(\Lambda)d\Lambda \tag{5.10}$$

is $2.86 \cdot 10^{-7}$ or $1.35 \cdot 10^{-3}$ for a 5σ or 3σ discovery potential, respectively. The quantity $P^B(\Lambda)$ is the distribution of the test statistic for a background only scenario. Since the test statistic is formed from a likelihood ratio, the test statistic distribution follows a χ^2 distribution (Wilk theorem [96]): for this, the background test statistic distribution is fitted with a χ^2 function and the value of Λ_{th} is calculated using the fit function. Secondly, different toy simulations with different number of injected events $n_{inj} > 0$ are used to extrapolate the number of events n^* such that the power of a test β, i.e. the chance of rejecting the null hypothesis when the alternative hypothesis is true, is equal to 50%. Figure 5.15 summarises the two steps exploiting an example of calculation for a 3σ discovery potential. Finally, n^* is converted to flux (see Appendix B for details).



(a) Threshold fixing.



(b) Estimation of n*.

Figure 5.15: The calculation of the number of events needed to perform a 3σ discovery in 50% of the cases. Left: the background test statistic distribution ($n_{inj} = 0$) is used to fix the threshold, represented by the red line, such that the fraction of the distribution above its value is $1.35 \cdot 10^{-3}$. Right: n^* is the number of injected events such that the test statistic distribution obtained from a test with $n_{inj} = n^*$, represented in gold, is half below and half above threshold. The two grey distributions, obtained from tests with $n_{inj} \lessgtr n^*$, are not cut in half by the red line.

By repeating the procedure with different values of the source declination we can estimate the dependence of the discovery potential on the latter. In addition, the injected spectral index can be changed to study the dependence of the discovery potential on the flux provided by different astrophysical sources. Figure 5.16 shows the 3σ and 5σ discovery potential as a function of the declination for three different spectral indices.
The discovery potential is not constant with the declination of the hypothetical source. In general, going far away from $\delta_{source} = 0$ results in an increase of the flux needed for a 3σ or 5σ discovery in 50% of the cases and the increase is more pronounced in the Southern sky due to the large background from atmospheric muons. In addition, with a slightly more background-like injection, i.e. with a slightly softer γ_{inj}, the discovery potential increases by several orders of magnitude, that results in the need of a stronger signal to discover a source. By relaxing the 5σ request to a 3σ discovery potential, the curve is lower. However, the change in the curve is very low if compared to the one connected to injecting a different spectral index. Finally, the results obtained in this work well superimpose to the points obtained for the same analysis conditions in [10].

The procedure explained in this section is valid both for the time-integrated and the time-dependent analyses.

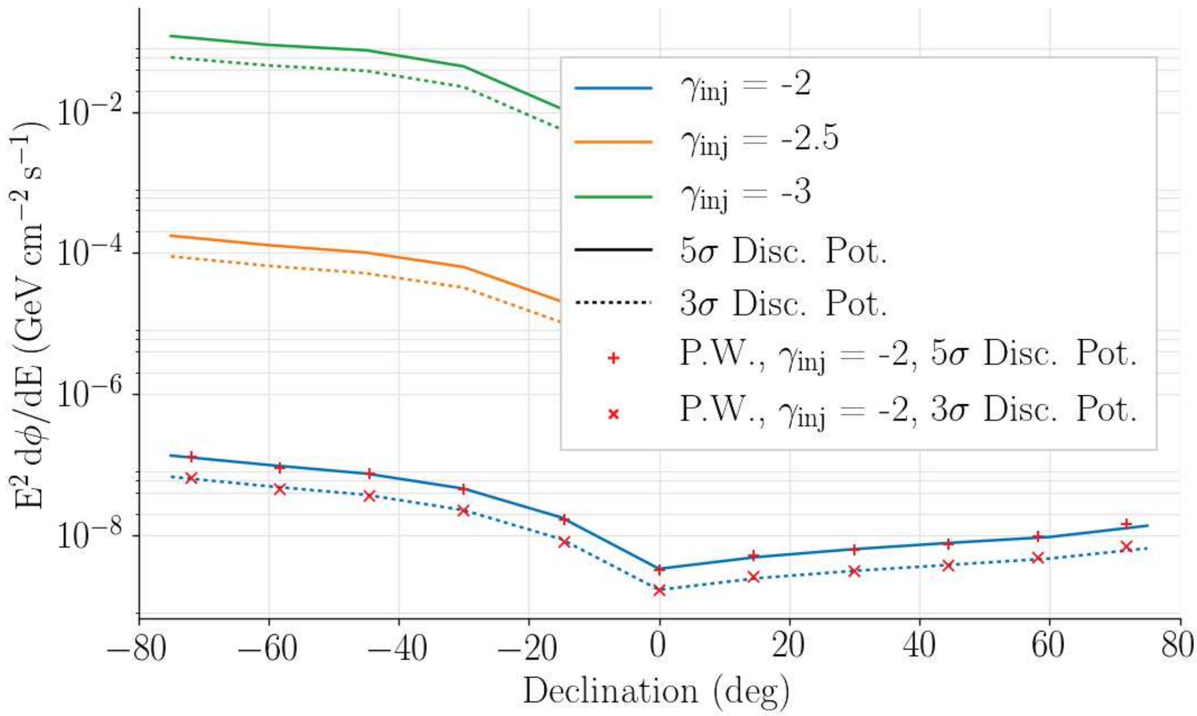


Figure 5.16: Discovery potential as a function of the source declination. The solid (dotted) lines represent the 5σ (3σ) discovery potential. The blue (orange, green) line is obtained for an injected spectral index of -2 (-2.5, -3). P.W. stands for "Previous Work", as the red pluses (crosses) were obtained in [10] for $\gamma_{inj} = -2$ and represent a 5σ (3σ) discovery potential.

5.4 The Time-Dependent Analysis

The time-integrated analysis can be improved by exploiting the time information of the events in the algorithm and thus by using a time-dependent approach. This enables the possibility to search for transient sources of astrophysical neutrinos. This search is supported from the fact that cosmic neutrinos observable with IceCube mainly originate outside the Milky Way [97]. Multi-wavelength data show that extra-galactic astrophysical sources, that are possible neutrino emitters, vary with time and assuming a connection between photon and neutrino emission we suppose that also the neutrino component varies with time showing the largest deviation from background [10].

The aim of the time-dependent analysis is to find a time window, that defines the flare, shorter than the entire dataset time (that can span one or several years). The time window is defined as a box beginning at time t_{start} and ending at time t_{end}. The quantity $T_{flare} = t_{end} - t_{start}$, i.e. the box width, represents the flare duration. The latter is a free parameter.

The core of the time-dependent analysis is the so-called time clustering algorithm, that is substituted to the simple likelihood maximisation of the time-integrated analysis. All the event arrival times in a dataset are used as window ends and, for each end, every previous event time is used as window start. The likelihood maximisation is performed in each time window, providing a set of estimates of the number of events, their spectral index and the value of the test statistic. The time window with the highest test statistic is chosen and the width of the latter is the estimated duration of the flare.

Testing all the possible time windows for a dataset is impractical. For example, the dataset of season 2014 contains a number of events N = 183 314 implying a very high number of likelihood maximisations.The number of tested time windows is reduced in two ways. Firstly, an upper limit T_{max} on the time window width is applied. Secondly, only events for which the ratio between the per-event weight $w_i = \frac{\mathcal{S}_i}{\mathcal{B}_i}$ (discussed in Section 5.2) is greater or equal than 1 are chosen as window edges to select more signal-like events. In this way the analysis is able to provide results in a reasonable amount of time.

An example of the tested time windows after the reduction of their number is sketched in Figure 5.17. For each window end (each one represented by a different color), only the events within a time of 10 arbitrary units before the end serves as window starts. In this simple example the number of tested time windows is reduced from 45 to 6.

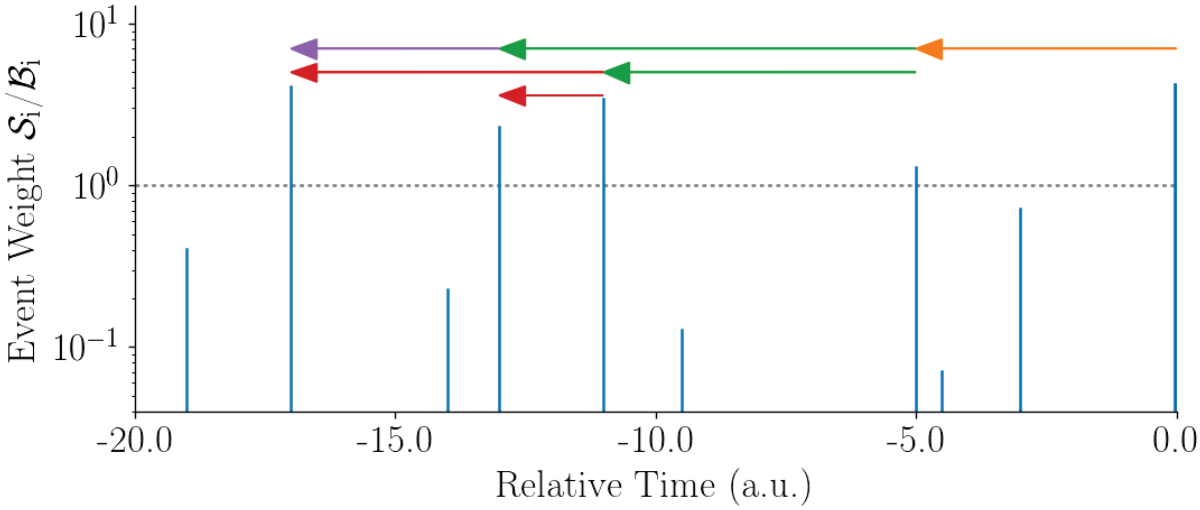


Figure 5.17: A sketch of some tested time windows when running the time clustering algorithm with T_{max} = 10 arbitrary units. On the ordinate there are the event weights, on the abscissa there is the time relative to the last event. The dotted grey line is the threshold that the events, represented by blue vertical lines, have to exceed in weight to be considered as time window edges and it is fixed to 1. The arrows represent the tested time windows and a different color represents a different window end.

Figure 5.18 summarises the time-dependent analysis. After selecting the data sample and the hypothetical source direction, and after producing the PDFs exploiting data and Monte Carlo simulations, the time-dependent analysis (via the time clustering algorithm) provides the four outputs Λ, n_S, γ and T_{flare}. The ability of the analysis to retrieve the correct cluster parameters has to be tested by using toy simulations, as for the time-integrated one. Section 5.4.2 describes the testing procedure of the time-dependent analysis and its results.

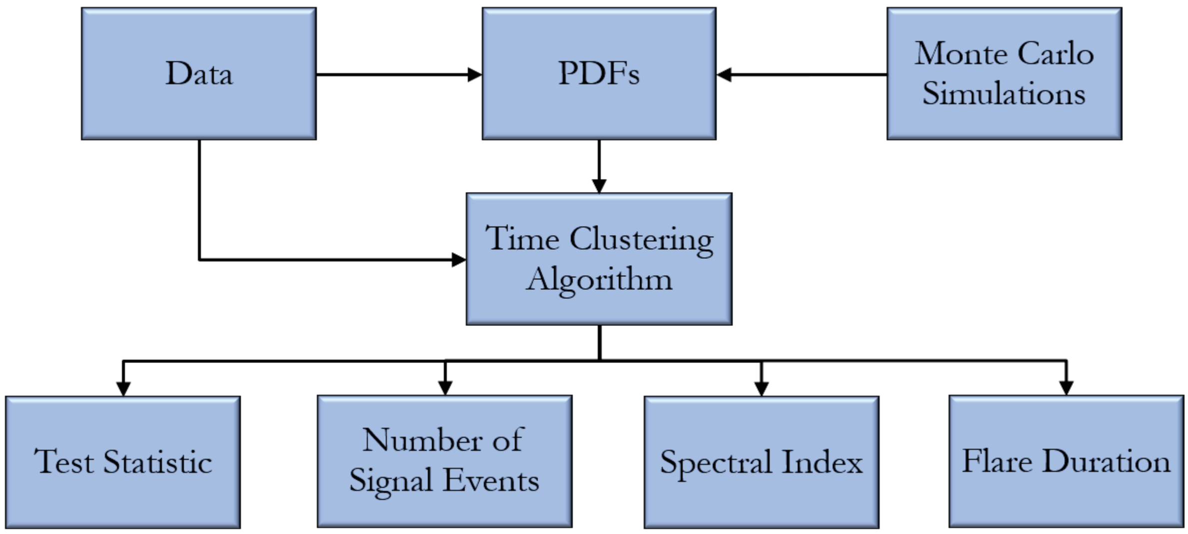


Figure 5.18: A summary of the time-dependent analysis. Data and Monte Carlo simulations (Section 5.1) are exploited to produce the spatial and energy PDFs (Section 5.2). Then, data and PDFs are used in the time clustering algorithm to maximise the likelihood in every tested time window, choosing the one with the highest test statistic. The algorithm produces four outputs: the test statistic, related to the value of the likelihood, the number of signal events, their spectral index and the flare duration.

Even after the reduction of the number of tested time windows this algorithm is still computationally more demanding with respect to the time-integrated one, that can be locally run on a PC.

The time-dependent analysis needs a more powerful environment. For this work, a cluster of CPUs was used.

5.4.1 Likelihood Formulation

The likelihood used for the time-dependent analysis is the same used for the time-integrated one, shown in Equation 5.4. However, the maximisation is performed using a modified version of test statistic. Since during time clustering algorithm many shorter time windows are tested than long time windows, there is a bias towards short flares that are selected with more preference. To reduce this effect the expression of the test statistic used in the maximisation is

$$\Lambda = 2\, log \left[\frac{\mathcal{L}(n_S, \gamma)}{\mathcal{L}(n_S = 0)} \frac{\mathcal{U}(t_i, t_k)}{T_{max}} \right], \tag{5.11}$$

where $\mathcal{U}(t_i, t_k)$ is the detector uptime during the time window starting from t_i and ending to t_k and T_{max} is the maximum width of the time windows. The new term reduces the test statistic for short time windows. In practice, in each time window the time-integrated test statistic is maximised and then the factor $2\, log \frac{\mathcal{U}(t_i, t_k)}{T_{max}}$ is added after the maximisation.

5.4.2 Test procedure and Results

The test procedure is similar to the time-integrated one. However, the injected signal events are inserted such that to resemble a flare of duration T_{inj} and the reconstruction quality of the fit parameters is studied for the flare duration, too.

The results refer to a source simulated at $\delta_{source} = 45°$ and $\alpha_{source} = 300°$, the dataset used is season 2014, the injected spectral index is $\gamma_{inj} = -2$, the injected flare duration is $T_{inj} = 10$ days and the maximum width of the time window is $T_{max} = 180$ days, unless otherwise specified.

As opposed to the time-integrated analysis, the convergence percentage of the time-dependent one is always $\approx 100\%$. This happens because, to fail the convergence, the likelihood maximisation should fail for all the tested time windows. For the time-integrated analysis only one time window is tested (the entire dataset). In addition, the decrease of the convergence percentage by decreasing the injected spectral index is negligible. These results are similar for other values of source declination. The few cases of no convergence are not used when producing any results. The presence of convergence failures suggests the need of exploring a different minimiser. Such study goes however beyond the scopes of this work.

Figure 5.19 shows the distribution of the test statistic for three different number of injected events (2, 10 and 20) compared to the background distribution. The behavior is similar to the time-integrated one (Figure 5.8), but the distributions reach higher values of the test statistic. This happens because of the lower background due to the smaller time windows (i.e. exposures). The distributions are less populated with respect to Figure 5.8 because, due to the very high run times of the algorithm, less maps were used for these tests with respect to the time-integrated ones.

Figure 5.20 shows the distribution of the fitted number of signal events. By comparing this figure to Figure 5.10, it is clear how the distributions for the time-dependent analysis are less broad. The other features of the plot are similar to the time-integrated case.

Figure 5.21 shows the distribution of the fitted spectral index. By comparing this figure to Figure 5.12, it is clear how the distributions for the time-dependent analysis for a high number of injected events are less broad. The other features of the plot are similar to the time-integrated case.

Figure 5.22 shows the distribution of the fitted flare duration. For a low number of signal events the distribution has not a clear behaviour, it only suggests an increase of the probability for a low flare duration. By increasing n_{inj} it peaks at the correct value of T_{inj}.

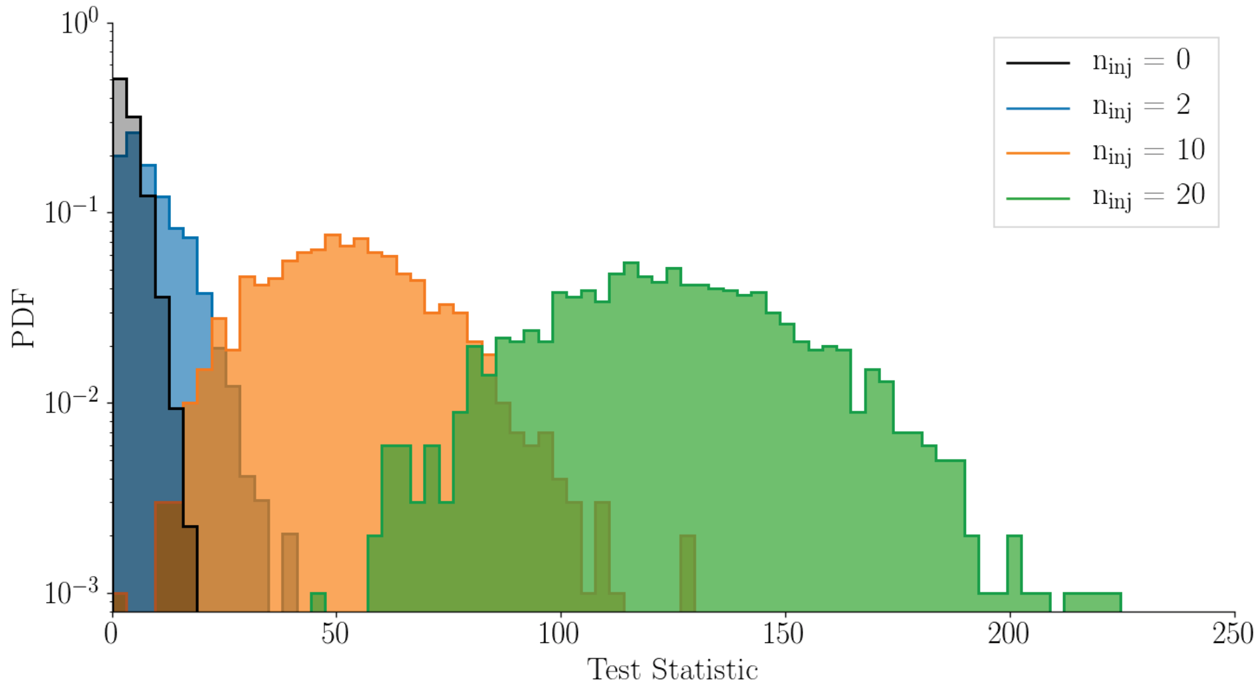


Figure 5.19: Distribution of the test statistic for the time-dependent analysis for an injected spectral index of -2, an injected flare duration of 10 days and the maximum time window size of 180 days. The black histogram represents the background distribution, while the blue, orange and green histograms represent the distribution for a number of injected events of 2, 10 and 20, respectively.

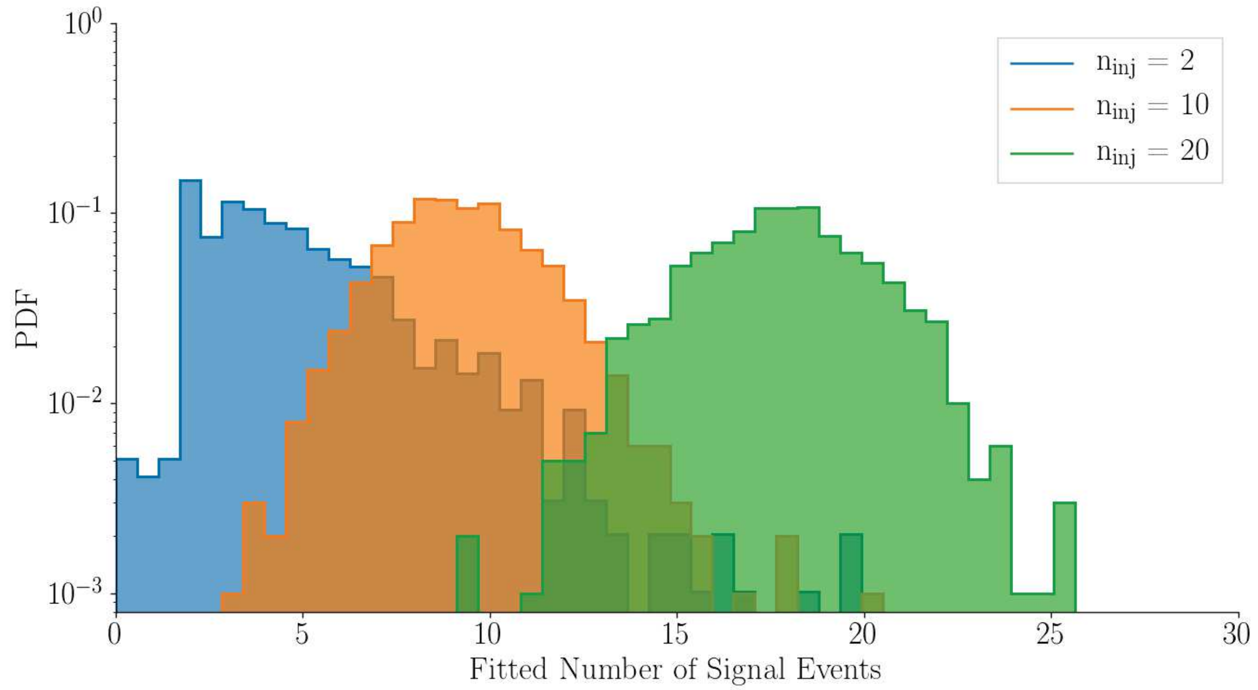


Figure 5.20: Distribution of the fitted number of signal events for the time-dependent analysis for an injected spectral index of -2, an injected flare duration of 10 days and the maximum time window size of 180 days. The blue, orange and green histograms represent the distribution for a number of injected events of 2, 10 and 20, respectively.

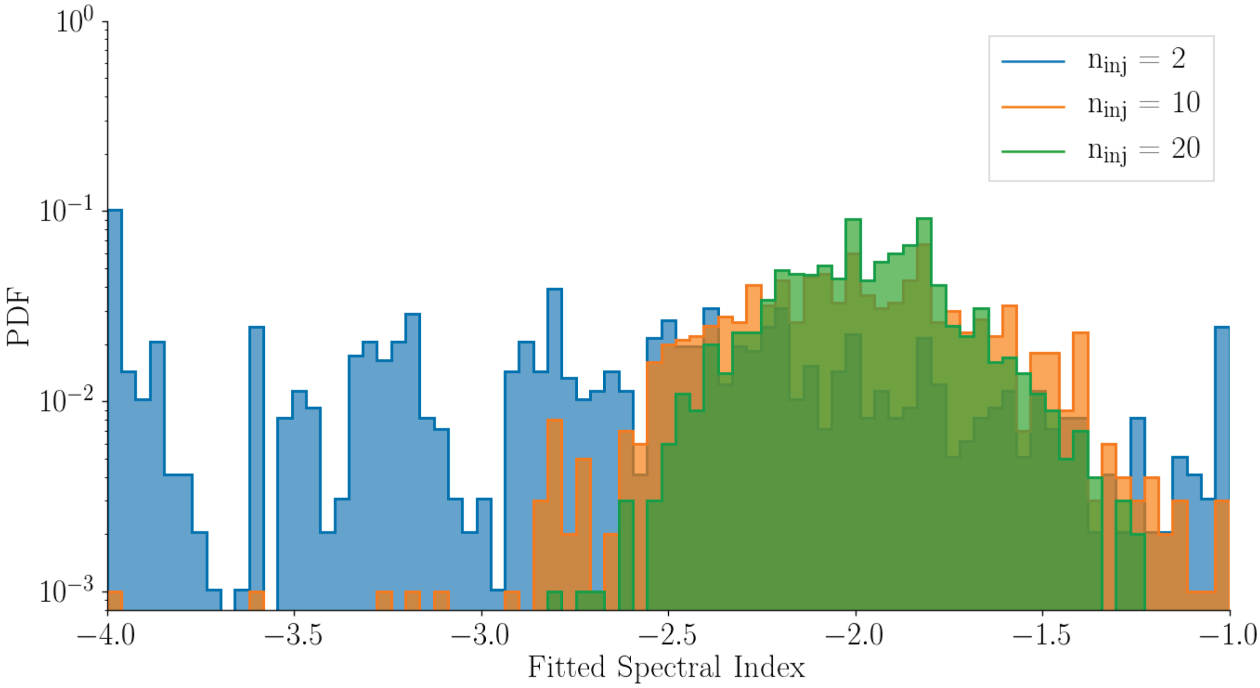


Figure 5.21: Distribution of the fitted spectral index for the time-dependent analysis for an injected spectral index of -2, an injected flare duration of 10 days and the maximum time window size of 180 days. The blue, orange and green histograms represent the distribution for a number of injected events of 2, 10 and 20, respectively.

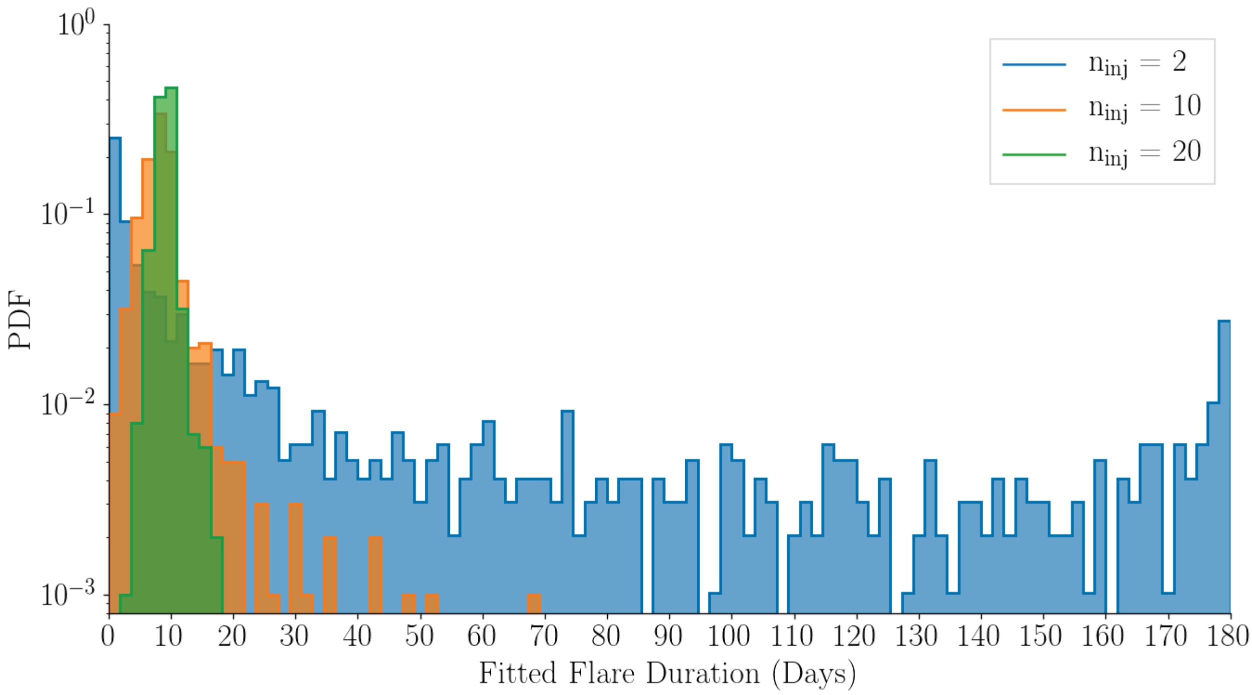
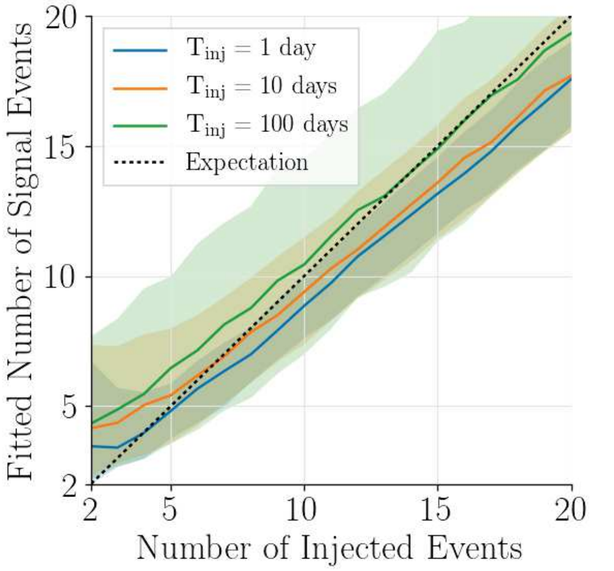
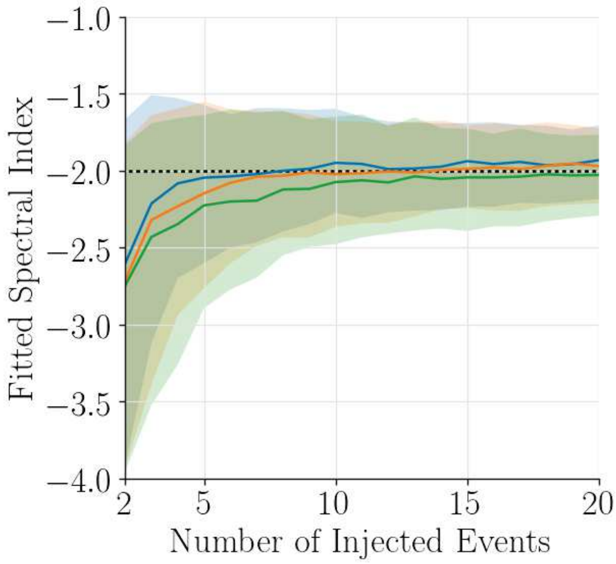


Figure 5.22: Distribution of the fitted flare duration for the time-dependent analysis for an injected spectral index of -2, an injected flare duration of 10 days and the maximum time window size of 180 days. The blue, orange and green histograms represent the distribution for a number of injected events of 2, 10 and 20, respectively.

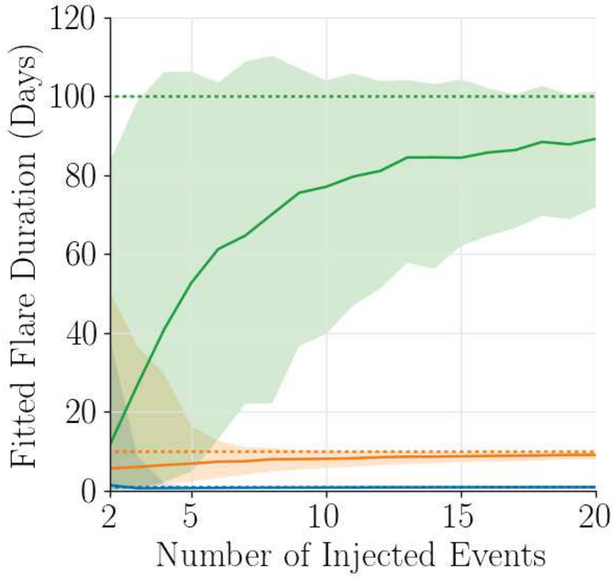
Figure 5.23 shows the reconstruction quality of the injected parameters versus the number of injected events for three different injected flare duration values.



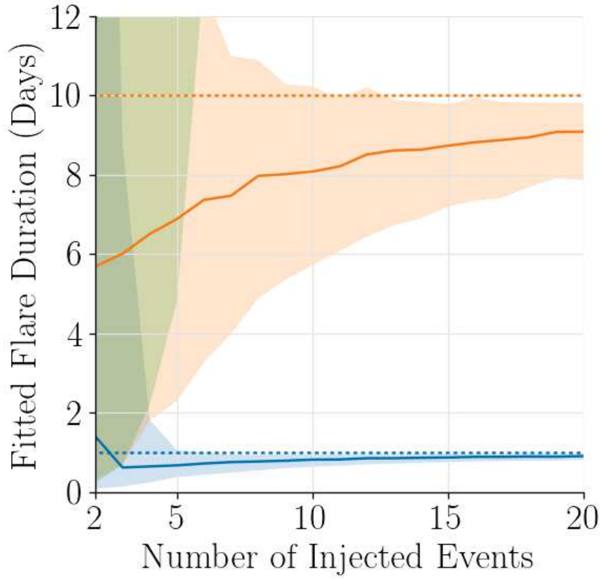
(a) Fitted number of signal events.



(b) Fitted spectral index.



(c) Fitted flare duration.



(d) Fitted flare duration (detail).

Figure 5.23: Reconstruction quality of the fit parameters for T_{max} = 180 days and γ_{inj} = -2. The continuous lines represent the median of the parameter distribution, the colored area is the 68-percentile centered in the median and the dotted lines show the expectation. On the abscissa there is the number of injected signal events. The blue (orange, green) curve is related to T_{inj} = 1 day (10 days, 100 days). Up left: the fitted number of signal events. Up right: the fitted spectral index. Down left: the fitted flare duration. Down right: detail of the fitted flare duration.

The curves almost always agree with the injection within the error. The fitted number of signal neutrinos increases for longer flares. It is overestimated for small clusters and underestimated for stronger signals. The error areas are more narrow with respect to the time-integrated case. The fitted spectral index decreases for longer flares. The plot is similar to the time-integrated case. The flare duration is almost always underestimated. For more populated clusters of neutrinos, the estimates of the spectral index and the flare duration approach the injection and the error is reduced. Other test results, not reported here, showed that for lower δ_{source} the behaviour is either similar or better.

Figure 5.24 shows the discovery potential represented as a fluence (see Appendix B for the definition) versus the declination of the hypothetical source, for three different values of the injected flare duration. Results from another study are also reported for comparison.

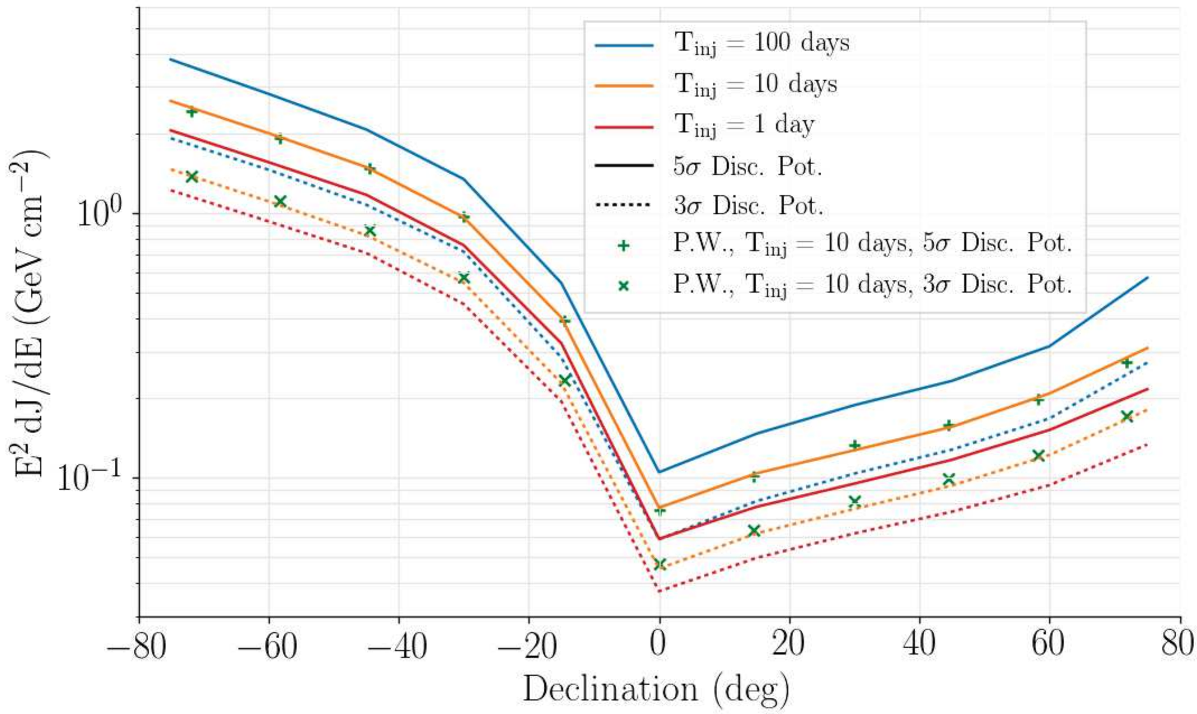
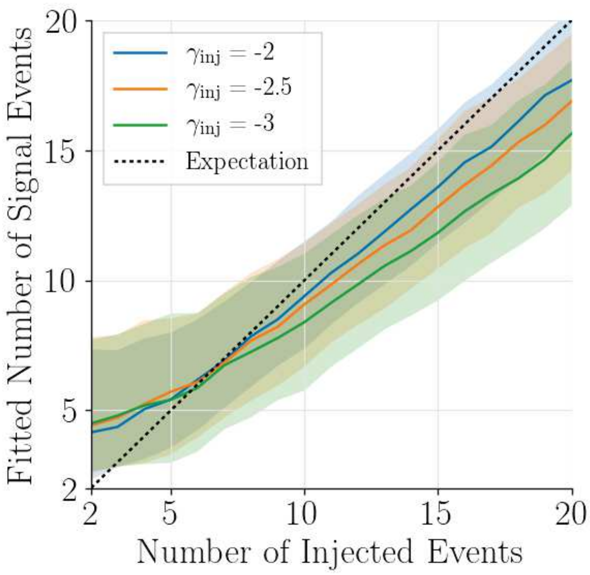


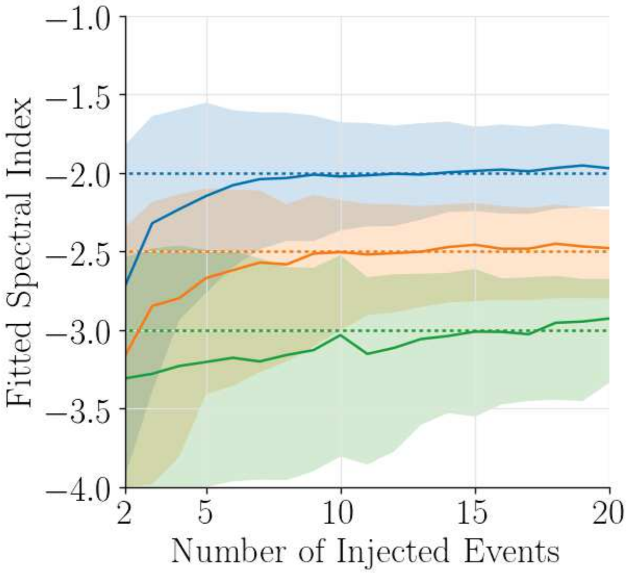
Figure 5.24: Discovery potential, reported as fluence, as a function of the source declination. The solid (dotted) lines represent the 5σ (3σ) discovery potential. The blue (orange, red) line is obtained for an injected flare duration of 100 days (10 days, 1 day). P.W. stands for "Previous Work", as the green pluses (crosses) were obtained in [10] for T_{inj} = 10 days and represent a 5σ (3σ) discovery potential.

The discovery potential depends on the injected flare duration. A transient source emitting neutrinos for a longer time needs a stronger signal to be discovered. As usual, the 3σ discovery potential is lower than the 5σ one and in the northern sky the 5σ line for T_{inj} = 1 day is lower the the 3σ line for T_{inj} = 100 days: it is easier to perform a 5σ discovery of a very fast transient source with respect to a 3σ discovery of a transient source emitting for hundreds of days. This is a very important aspect and in the next section tests tailored for very fast transient sources will be discussed. The shape of the curves is similar to the time integrated one of Figure 5.16. Finally, the analysis carried out for this work for T_{inj} = 10 days agrees with the results obtained in a previous study for the same analysis conditions [10].

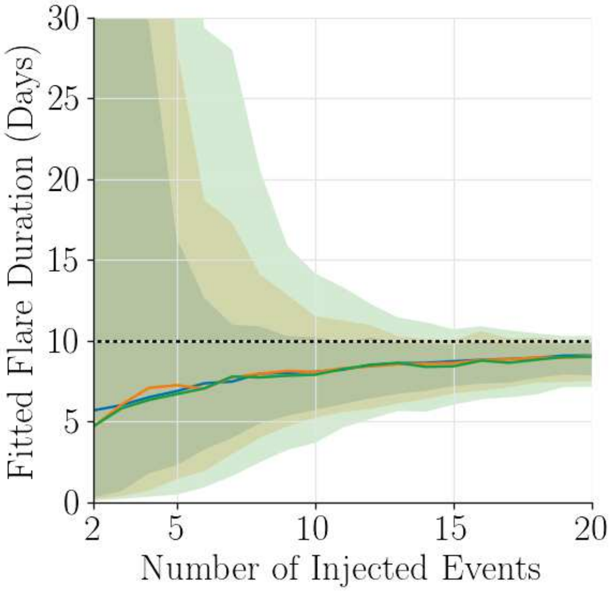
As for the time-integrated analysis, the reconstruction quality and the discovery potential depend on the injected spectral index. Figure 5.25 shows the reconstruction quality of the injected parameters versus the number of injected events for three different spectral indices and for the same injected flare duration T_{inj} = 10 days. The fitted number of signal events is overestimated for n_{inj} lower than ≈ 6 and underestimated for more populated clusters of neutrinos. Less events are recognised as signal when softening the spectral index, as expected. By comparing the plot in the upper left panel of this figure with the one in Figure 5.14, the results for γ_{inj} = -2 (-2.5, -3) is worse (similar, better) in the time-dependent case. The spectral index is underestimated for a low number of injected events and, with more signal events, it approaches the expectation. By comparing the plot in the upper right panel of this figure with the one in Figure 5.14, the results are similar. The flare duration is always underestimated, but its value gets closer to the expectation and the error significantly reduce for more injected signal neutrinos. Finally, the error is higher for softer injected spectral indices.



(a) Fitted number of signal events.



(b) Fitted spectral index.



(c) Fitted flare duration.

Figure 5.25: Reconstruction quality of the fit parameters for T_{max} = 180 days and T_{inj} = 10 days. The continuous lines represent the median of the parameter distribution, the colored area is the 68-percentile centered in the median. The dotted lines show the expectation. On the abscissa there is the number of injected signal events. The blue (orange, green) curve is related to γ_{inj} = -2 (-2.5, -3). Up left: the fitted number of signal events. Up right: the fitted spectral index. Down: the fitted flare duration.

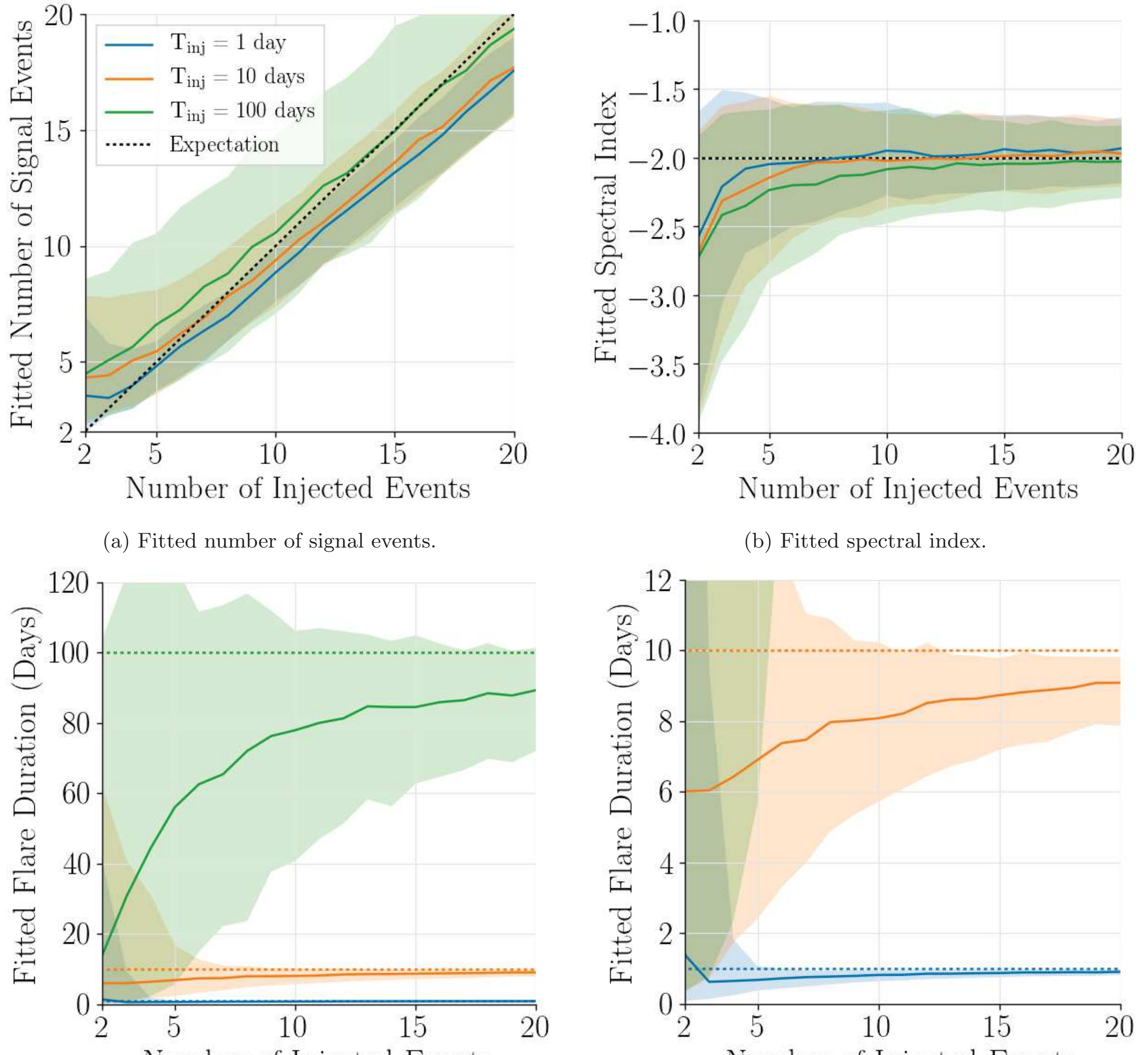
Table 5.1 shows the number of events and the fluence needed for a 3σ and 5σ discovery in 50% of the cases for the three tested spectral indices.

	3σ Disc Pot.		5σ Disc Pot.	
γ_{inj}	n^*	$E^2dJ/dE(GeV cm^{-2})$	n^*	$E^2dJ/dE(GeV cm^{-2})$
-2	4.69	$9.75 \cdot 10^{-2}$	7.58	$1.58 \cdot 10^{-2}$
-2.5	7.01	13.29	11.05	20.96
-3	9.75	825.78	16.04	1358.30

Table 5.1: Discovery potential and associated number of events n^* for three different injected spectral indices, for T_{max} = 180 days and T_{inj} = 10 days.

For softer injected spectral indices n^* increases. Consequently, the discovery potential fluence increases by several orders of magnitude. This behaviour is similar for the time-integrated case (Figure 5.16).

It is interesting and important to study the dependence of the results also on the maximum time window width, as this parameter has to be chosen beforehand when performing real-time analyses. Figure 5.26 shows the reconstruction quality of the injected parameters versus the number of injected events for the maximum time window width fixed at 365 days, higher than the previous cases.



(a) Fitted number of signal events.

(b) Fitted spectral index.

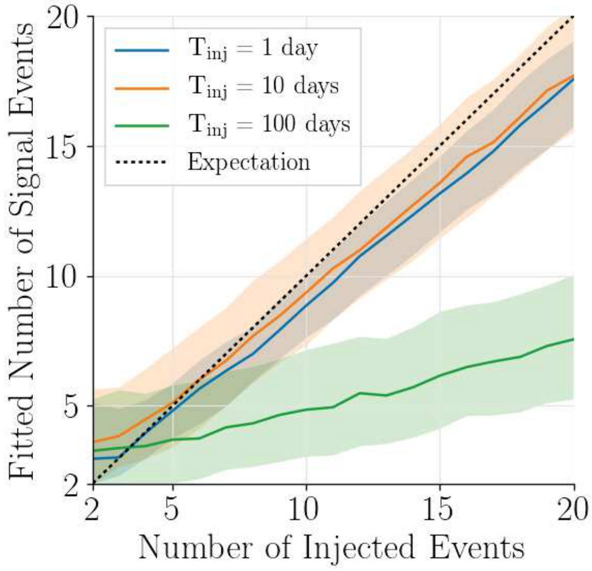
(c) Fitted flare duration.

(d) Fitted flare duration (detail).

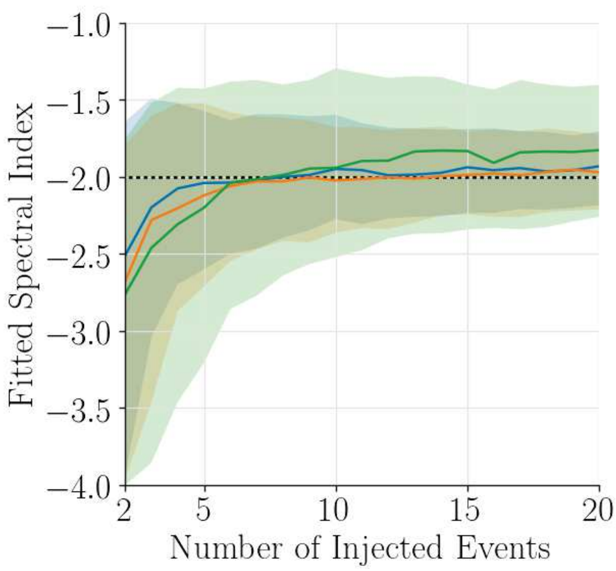
Figure 5.26: Reconstruction quality of the fit parameters for T_{max} = 365 days and γ_{inj} = -2. The continuous lines represent the median of the parameter distribution, the colored area is the 68-percentile centered in the median and the dotted lines show the expectation. On the abscissa there is the number of injected signal events. The blue (orange, green) curve is related to T_{inj} = 1 day (10 days, 100 days). Up left: the fitted number of signal events. Up right: the fitted spectral index. Down left: the fitted flare duration. Down right: detail of the fitted flare duration.

The features of the curves are the same with respect to Figure 5.23, thus there is no significant improvement. Instead, the number of tested time windows is much higher. As a consequence, the algorithm takes much more time to be executed and the trial factors worsens the discovery potential, as shown in the following. Thus, from a certain maximum time window onwards, it is worse to increase T_{max}. A choice of a high T_{max} should be done only to discover flares distributed on hundreds of days.

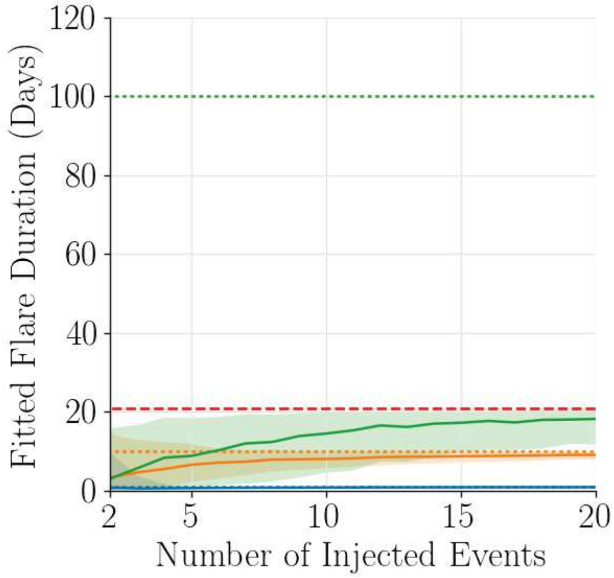
Figure 5.27 shows the reconstruction quality of the injected parameters versus the number of injected events for the maximum time window width fixed at 21 days, lower than the previous cases.



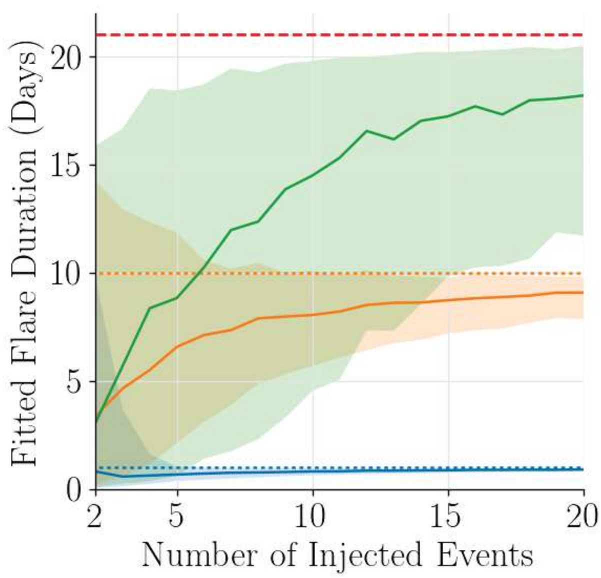
(a) Fitted number of signal events.



(b) Fitted spectral index.



(c) Fitted flare duration.



(d) Fitted flare duration (detail).

Figure 5.27: Reconstruction quality of the fit parameters for T_{max} = 21 days and γ_{inj} = -2. The continuous lines represent the median of the parameter distribution, the colored area is the 68-percentile centered in the median and the dotted lines show the expectation. On the abscissa there is the number of injected signal events. The blue (orange, green) curve is related to T_{inj} = 1 day (10 days, 100 days). The red dashed line in the flare duration reconstruction plots represents T_{max} = 21 days. Up left: the fitted number of signal events. Up right: the fitted spectral index. Down left: the fitted flare duration. Down right: detail of the fitted flare duration.

The reconstruction quality for n_S and γ with a flare shorter than T_{max} is similar to the 180 days case. However, when analysing a flare longer than the maximum time window size, the number of signal events is significantly underestimated because the time windows only cover part of the flare. The curve of the number of signal events reconstruction plot relative to T_{inj} = 100 days increases because the density of injected events in the sample increases, too, and more events fit inside the tested windows. Also the spectral index reconstruction is worse, as the median is further away from the expectation for strong signals and the error area is wider. The flare duration reconstruction for $T_{inj} < T_{max}$ is similar to the previous cases, a part for the lower error. Instead, when injecting a long flare, the curve does not approach the expectation but the maximum time window size, because longer time windows are not tested.

Thus, T_{max} should be tuned accordingly to the types of sources under search by using theoretical information on the expected flare duration. This parameter should be high enough to completely contain the flare but low enough to avoid challenging computing time.

We have seen in Table 5.1 that the discovery potential depends significantly on the flux spectral index and in Figure 5.24 that it depends on the declination and, to a lesser extent, on the flare duration. In addition, it slightly depends on the maximum window size, as shown in Figure 5.28.

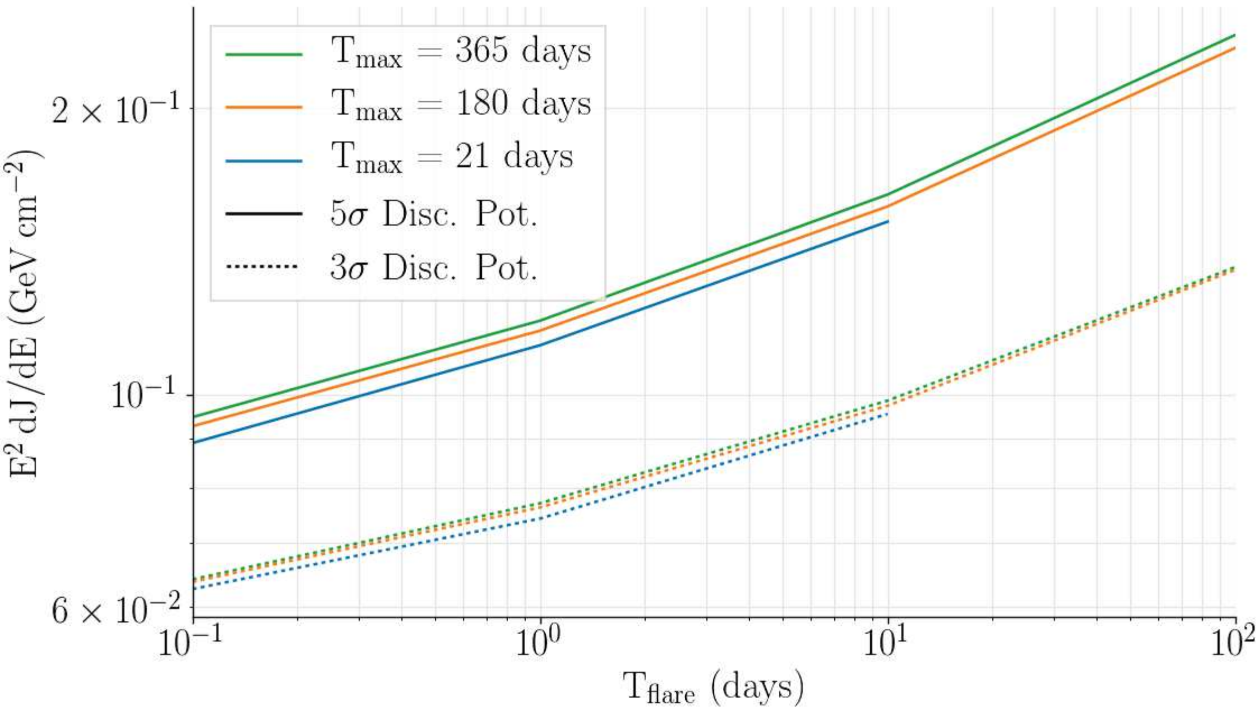


Figure 5.28: Discovery potential (reported as fluence) as a function of the flare duration. The solid (dashed) lines represent the 5σ (3σ) discovery potential. The blue (orange, green) line is obtained for a maximum time window size of 21 days (180 days, 365 day). For $T_{max} = 21$ days, the maximum tested flare duration is 10 days.

By increasing the maximum time window size the discovery potential slightly increases. This happens because, with a larger T_{max}, more time windows are tested and a penalty comes from the large number of trials (look-elsewhere effect [98]). This effect can be also seen in the $\sim log(T_{max})$ dependence of the time-dependent test statistic. The dependence of the discovery potential on the maximum size of the time windows is very weak with respect to the other dependencies.

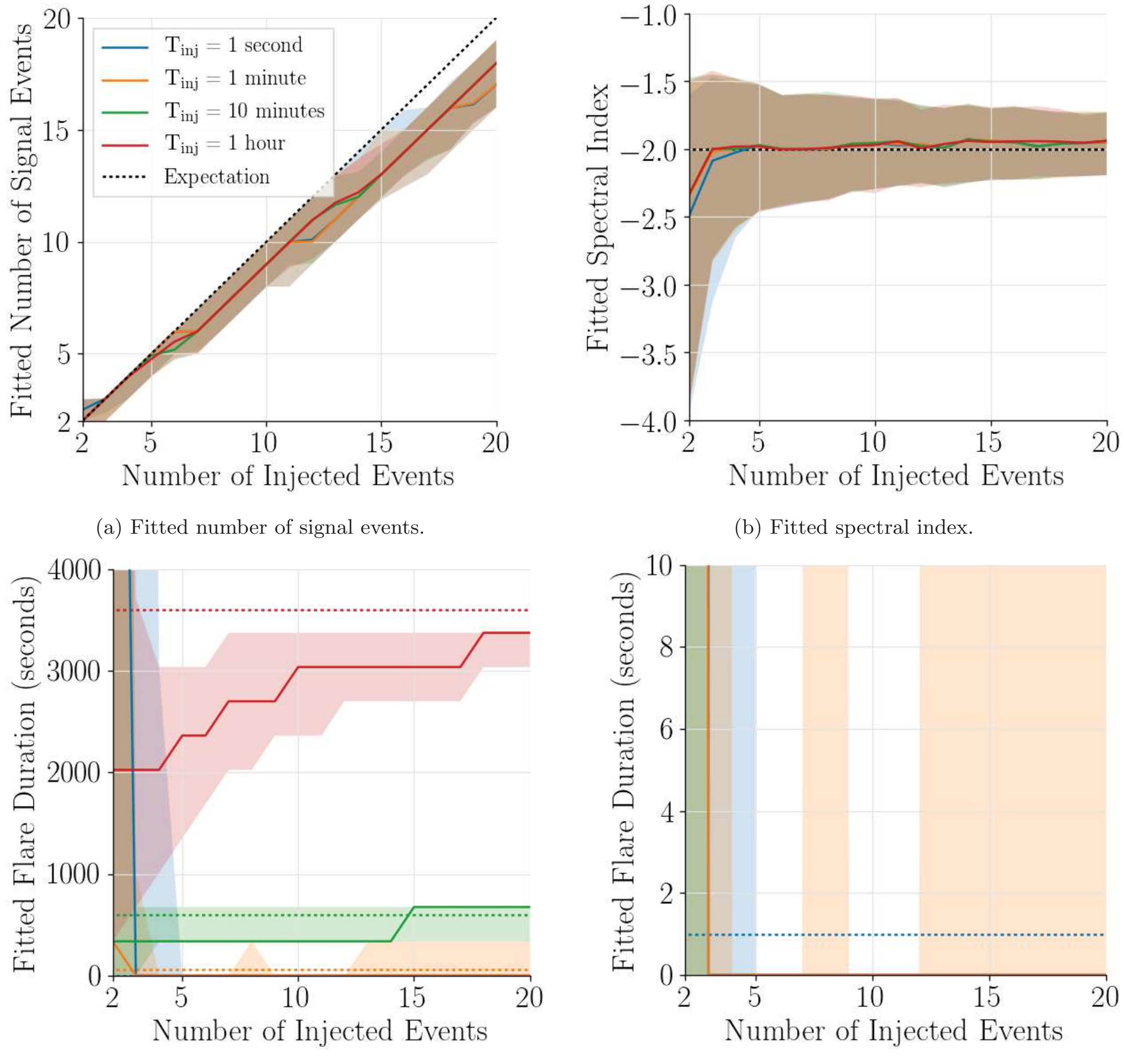
Up to know, the time-dependent analysis has been tested for flares lasting down to $\approx$ 3 hours. The next section extends the tests to very fast transient sources.

5.4.3 Tests for very fast transient sources

The rate of the GFU filter is $\approx$ 6.5 mHz. This implies that, on average, the temporal distance between two events is $\Delta t = \frac{1}{average\ rate} \approx 3$ minutes. Thus, when injecting several events to resemble a flare shorter than few hours, the background component during the flare is very low. In addition, if the flare is shorter than Δt, there will be time windows containing only injected events.

The tests for very fast transient sources are carried out to study the behaviour of the analysis in these extreme conditions. The convergence percentage is always 100% as expected, because very few background events, if any, are contained in the tested time windows.

Figure 5.29 shows the reconstruction quality of the injected parameters versus the number of injected events for the maximum time window size $T_{max} = 1$ day and for flares of 1 second, 1 minute, 10 minutes and 1 hour. The number of signal events is close to the expectation until $n_{inj} \approx 6$, then it is underestimated, and the error is low. The reconstruction of the spectral index is almost equal for all the four tested flares and shows the same behaviour of the previous tests. The flare duration for



(a) Fitted number of signal events.

(b) Fitted spectral index.

(c) Fitted flare duration.

(d) Fitted flare duration (detail).

Figure 5.29: Reconstruction quality of the fit parameters for T_{max} = 1 day and γ_{inj} = -2. The continuous lines represent the median of the parameter distribution, the colored area is the 68-percentile centered in the median and the dotted lines show the expectation. On the abscissa there is the number of injected signal events. The blue (orange, green, red) curve is related to T_{inj} = 1 second (1 minute, 10 minutes, 1 hour). Up left: the fitted number of signal events. Up right: the fitted spectral index. Down left: the fitted flare duration. Down right: detail of the fitted flare duration.

the two longer flares is underestimated for small clusters and improves for stronger signals. The flare duration for the two shorter flares is overestimated for 2 signal events and equal to zero onwards. This happens because the time information is stored with the last digit on 1/1000 of day, that corresponds to 86.4 seconds. For shorter flares the reconstructed flare duration is zero. Finally, tests with T_{max} = 180 days showed the same results.

Very fast transient sources would provide very sensitive observations due to the low number of background events in the time windows, but require a higher data rate at the final event selection level. Firstly, the detector instrumented volume and string density directly influence the number of events that are detected by IceCube and the sensitivity to high energy neutrinos. By increasing the deep polar ice instrumented volume, more events (and thus more signal-like events) are collected and the data rate at the final event selection level is higher, but a more powerful analysis may be needed to reject the increased number of background events. By placing the strings further away, the detected events have higher energies as the energy threshold for the detection would be higher, increasing the probability to have signal-like events in the data sample. An upgrade of the IceCube detector, called IceCube Gen2 [99], has been proposed. It is discussed in the next chapter.

Secondly, the event filtering and reconstruction procedure can be improved. For example, by using more detailed ice models, a more refined calibration or more powerful reconstruction algorithms results in better estimated event properties that allow more events to be selected. In addition, relaxing the event selection cuts increases the event rate at the final event selection level. For example, signal events whose reconstruction is not adequate and are rejected by the present filters could be selected. However, to handle the events with a poorer reconstruction and to reject the higher number of background events more powerful analyses are needed. In general, improved analysis methods can increase the sensitivity to point-like transient neutrino sources.

By increasing the event rate and the probability for the events to be signal neutrinos, as well as by improving the algorithms, shorter flares can be searched for and the possibility to discover transient and fast transient sources with a higher significance increases.

Chapter 6

Conclusions and Future Perspectives

Successful tests of the time-dependent algorithm, core of the real-time analyses searching for astrophysical neutrino sources with the IceCube data, have been carried out for different source features. The reconstruction quality of the signal parameters (number of signal neutrinos, flux spectral index and flare duration) and the discovery potential, i.e. the flux that yields a 5σ discovery in 50% of the cases, have been studied. Similar tests have been performed for the time-integrated version of the algorithm.

In general, the time-dependent algorithm performs better than the time-integrated one. For more populated clusters of neutrinos, for harder spectral indices or for longer flares the signal parameters are better reconstructed and the algorithms converge more often. The reconstructed signal parameters agree with the expectations within the error for most of the cases. In addition, the time windows where the time-dependent likelihood is maximised have to be larger than the flare duration to obtain reliable results.

The discovery potential mainly depends on three parameters. It improves by several orders of magnitude with a little hardening of the spectral index. Its dependency on the source declination is weaker and horizontal events have the highest probability to yield a discovery. Finally, it slightly improves for a shorter duration of the flare.

Finally, the tests have been extended to very fast transient sources, with flares shorter than few hours. By increasing the event rate and the probability for the events to be signal neutrinos, the chance to discover transient and fast transient sources increases. This can be obtained by improving the filtering and reconstruction procedure. In addition, the instrumented volume of the detector can be increased to collect more events. An improvement of the IceCube detector, called IceCube Gen2, is already planned and under design. IceCube Gen2 consists of several improvements with respect to IceCube [60, 100]:

- A high-energy in-ice array featuring an increased string spacing (rising the threshold on the minimum detectable energy) and a total effective target volume 10 times larger with respect to IceCube to study astrophysical neutrino sources at the PeV energy scale. Figure 6.1 shows the position of the IceCube Gen2 strings compared to the IceCube ones.

- An extended surface air shower array, potentially several times larger in diameter than the high-energy array, to be used both for cosmic ray studies and to veto down-going atmospheric muons and neutrinos. Figure 6.2 shows the position of the IceCube elements in IceCube Gen2, as well as the IceCube Gen2 strings and the large area covered by the extended surface air shower array.

- A low-energy array (the Precision IceCube Next Generation Upgrade, or PINGU [103]) consisting in the instrumentation of 6 MTon of ice in the center of DeepCore, enabling precision neutrino oscillation measurements down to the GeV energy range, determination of the neutrino mass ordering and dark matter searches. Updated calibration devices will be deployed in the new boreholes in order to measure the optical properties of the ice more precisely, improving the ice models and consequently the event reconstruction procedure. Figure 6.3 shows the position on the surface of the PINGU strings within the IceCube detector. The position of the 26 closer strings of PINGU is reported as red crosses.

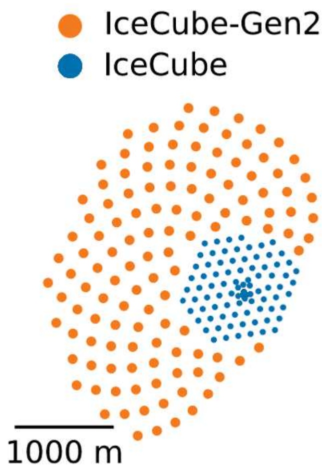


Figure 6.1: Position of the IceCube Gen2 strings (orange circles) compared to the IceCube strings (blue circles). In the initial design studies, 120 new strings are spaced 240 m apart and nominally instrumented with 80 IceCube-type optical modules over a vertical length of 1.25 km. The total instrumented volume in this design is 7.9 km^3, approaching an order of magnitude larger than IceCube alone. Adapted from [101].

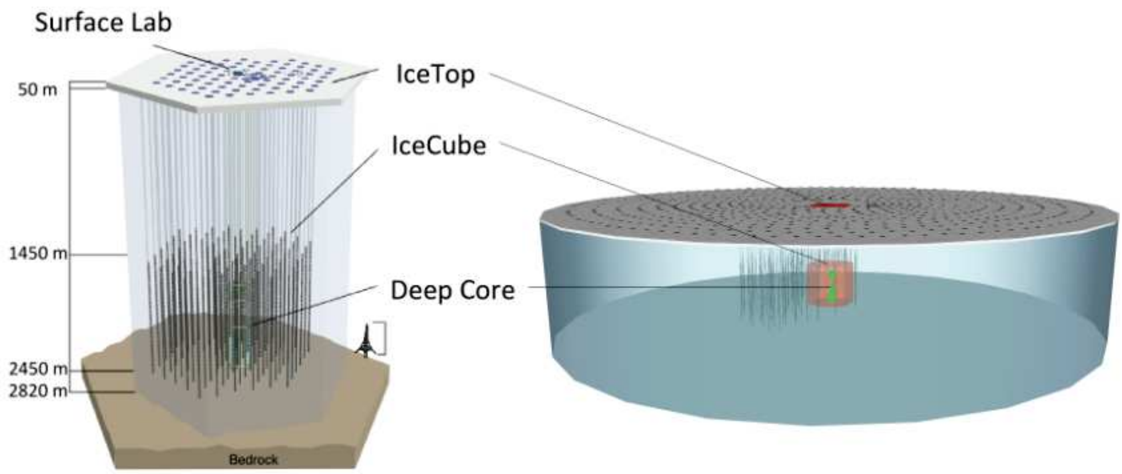


Figure 6.2: Not-to-scale comparison between IceCube (left) and IceCube Gen2 (right). On the right, IceCube and IceTop are colored in light red and red, respectively. DeepCore is colored in green. The IceCube Gen2 high-energy array strings are reported as vertical lines. The extended surface air shower array, on the top base of the cylindrical volume, covers a very wide area with respect to both IceTop and the projection of the IceCube Gen2 high-energy array on the surface. Taken from [102].

- A shallow sub-surface array of radio antennas for the detection of ultra-high-energy neutrino-induced showers in the ice, promising to achieve effective target volumes of about 100 times IceCube at 10^{18} eV, where neutrinos originating from scattering of ultra-high-energy cosmic rays on the cosmic microwave background are expected [105].

In addition, a modernized version of the IceCube DOM will be deployed [103]. Mechanical components such as the glass sphere, as well as the high-quantum-efficiency PMT, remain unchanged, while the triggering and digitization electronics are being redesigned. Alternative sensor designs that increase photocathode area, photon collection, angular coverage, and/or directional resolution are also under study [60].

The search for point-like astrophysical neutrino sources will benefit from the aforementioned improvements, like the increased instrumented volume, the better ice models and the atmospheric veto techniques.

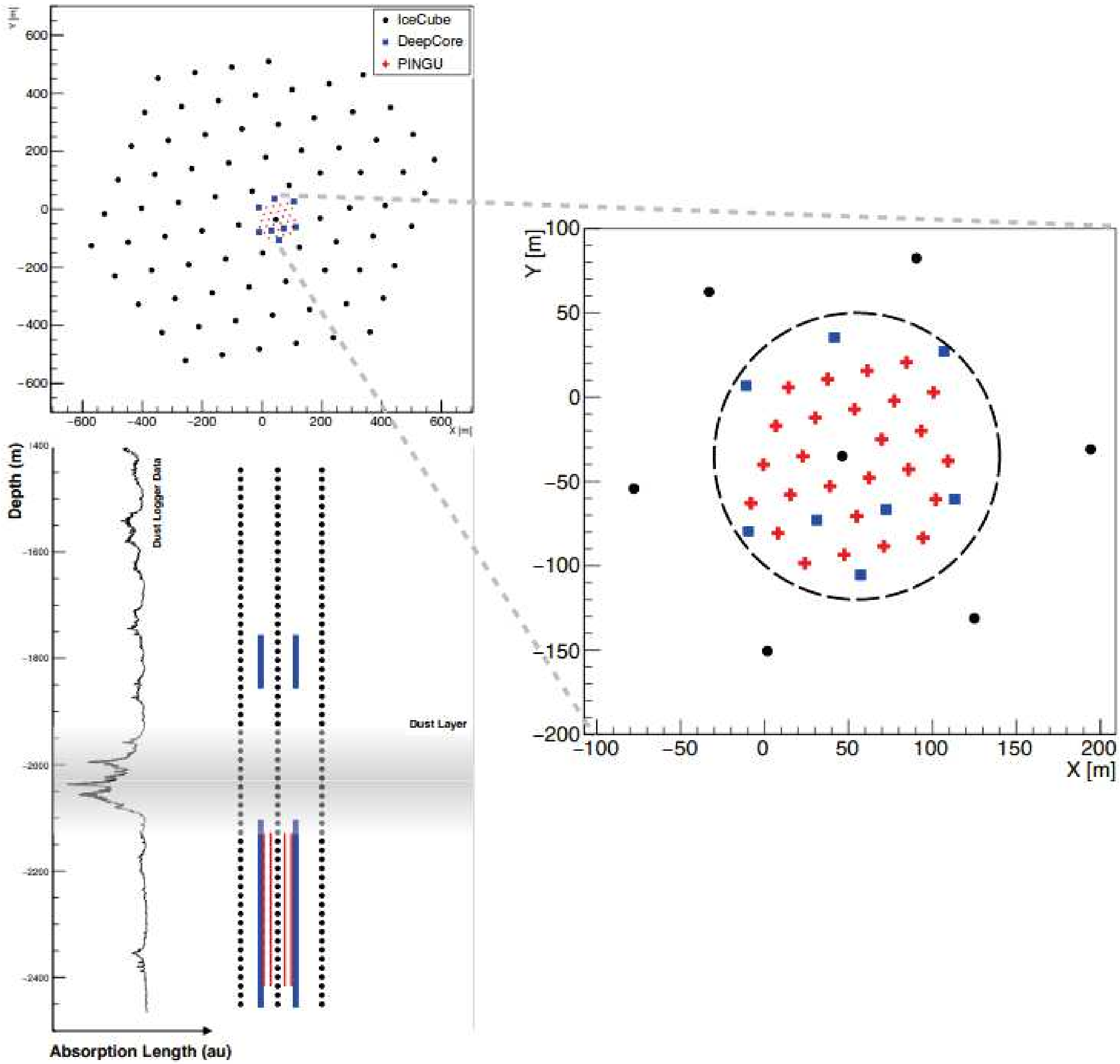


Figure 6.3: Schematic layout of PINGU within the IceCube detector. The IceCube strings are reported as black circles, the DeepCore strings as blue squares, the PINGU strings as red crosses. Top left: the strings position on the surface. Right: detail of the strings position on the surface. The PINGU optical modules will be deployed in the clearest ice at the bottom of the detector, as shown in the vertical profile on the bottom left panel, with a vertical spacing several times denser than DeepCore. Taken from [104].

Chapter 7

Acknowledgments

Several people helped and supported me during my university career and the master thesis. It is important for me to mention some of them.

I want to thank prof. Elisa Bernardini for introducing me to multi-messenger astrophysics. After inspiring me by explaining this challenging and interesting branch of Physics, she suggested me to apply for a Summer School at DESY, Zeuthen, that has been one of the most crucial experiences for my scientific career. For example, I understood that multi-messenger astrophysics is indeed the most exciting branch of Physics for me. In particular, I am grateful to Konstancja Satalecka and Cosimo Nigro for supervising my work there, giving me advice and contributing to the positiveness of that Summer. I want to thank also the other summer students because without them that experience would not have been so funny, entertaining and unforgettable. I want to express my gratitude to Manuela Mallamaci and, again, Elisa Bernardini for guiding me during the past months and for providing me useful insights and explanations, necessary for this work. I want to thank Thomas Kintscher for developing the version of the real-time infrastructure that I used and for providing me inspiration for this work with his PhD thesis. Also, thanks to Michael Larson for helping me with the job submission. I want to thank Andrés Illana Sisón and José Javier Valiente Dobón for their help and support during and after the bachelor thesis.

I want to thank my friends because I had a very great time hanging out, laughing and sharing together several beautiful experiences. Thanks also to my colleagues at university for the interesting conversations, for sharing thoughts and emotions and for making much more pleasant the university experience. I want to thank Ilaria. You have been next to me for all the five years of my university career, and more! With you I have always felt loved and appreciated. You listened to my problems and always provided me very useful advice and interesting point of views. Above all, you make me happy. I want to express my gratitude also to ER for his important help. Finally, I want to thank my family for their support during the years. In particular, thank you dad for being so proud of me!

Appendices

Appendix A

Coordinate Systems and Useful Definitions

In this appendix the coordinate systems used in this work for the events direction are defined, as well as some useful quantities.

Two coordinate systems have been used in this work. One is the local detector coordinate system defined by the usual zenith and azimuth angles. The other is the celestial coordinates system defined by right ascension and declination, sketched in Figure A.1.

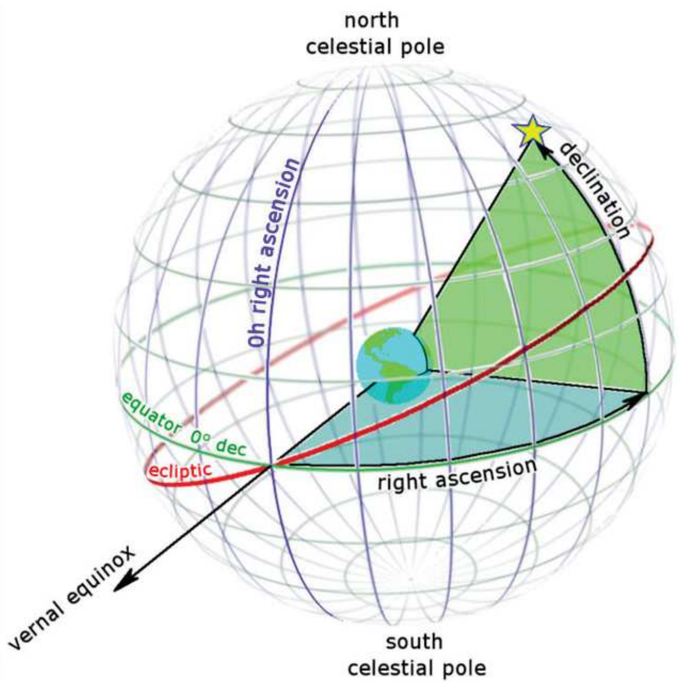


Figure A.1: Right ascension and declination example. Taken from [106].

The right ascension is the angular distance of a point measured eastward along the celestial equator, the projection of the Earth equator on the celestial sphere, from the Sun at the March (vernal) equinox. The declination is the angle between the celestial equator and the point under study, measured in a plane perpendicular to the celestial equator and containing the point. The declination is positive for the Northern sky and negative for the Southern sky. For the conversion between the two systems see [107].

Often, in astrophysics, the time is reported in Modified Julian Days (MJDs). It is a useful format to calculate the time between two events. A Julian Day (JD) is the number of days passed since noon on Monday, 1$^{\text{st}}$ of January 4713 BC. MJDs were introduced to use only 18 bits for space missions until the 7$^{\text{th}}$ of August 2576. MJDs start at midnight on the 17$^{\text{th}}$ November 1858 and are computed as

$$\text{MJD} = \text{JD} - 2400000.5\,. \tag{A.1}$$

The ".5" is used to start the count from midnight and not from noon.

In addition, the sidereal time is used to locate celestial objects as, observing from the same point on Earth at the same sidereal time, the position of a star does not change. A sidereal day, the time between two consecutive equal sidereal times, is approximately 23 hours, 56 minutes, 4.0905 seconds long.

Astronomical distances are often measured in light years, parsecs or with the redshift. A light year is the distance travelled by light in a year and corresponds to $9.461 \cdot 10^{15}$ m. A parsec is the distance at which one astronomical unit subtends an angle of one arcsecond, as seen in Figure A.2. 1 parsec corresponds to $3.086 \cdot 10^{16}$ m or to ≈ 3.26 light years.

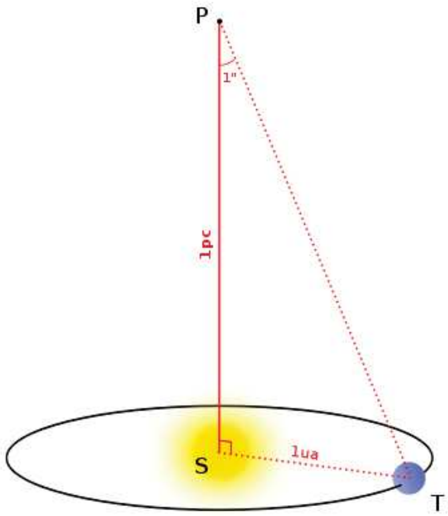


Figure A.2: Definition of parsec. Taken from [108].

The redshift z is the shift towards higher wavelengths of the electromagnetic radiation caused by the expansion of the Universe, visible for distant objects. Objects further away from us have higher values of the redshift. It is a positive number and it is calculated as

$$z = \frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}}, \tag{A.2}$$

where λ_{obs} is the observed light wavelength and λ_{em} is the emitted one.

Sometimes the energy in astroparticle physics is expressed in ergs. 1 erg corresponds to 100 nJ, hence the conversion to eV:

$$1\,\text{erg} = 6.242 \cdot 10^{11}\,\text{eV} \tag{A.3}$$

Appendix B

Flux and Fluence

The flux of neutrinos ϕ is defined as the number of neutrinos per unit area per unit time. The flux is related to the event rate by the relation

$$\frac{dN}{dt} = \int d\Omega \int_0^\infty A_{eff}(E,\delta) \frac{d^3\phi}{dt d\Omega dE} dE, \tag{B.1}$$

where dN/dt is the neutrino rate, Ω is the solid angle and A_{eff} is the effective area (discussed in Section 4.5), that depends on the energy E and on the declination δ. In this work, the effective area is averaged over muon neutrinos and muon anti-neutrinos and the fluxes are the sum of the muon neutrino and muon anti-neutrino fluxes, as done in [10].

In this work, the flux is modeled with an unbroken power-law defined by

$$\frac{d\phi}{dE} = \phi_0 \cdot \left(\frac{E}{E_0}\right)^\gamma, \tag{B.2}$$

where ϕ_0 is the flux normalisation at the reference energy of $E_0 = 1$ GeV and γ is called "spectral index". With this definition all the spectral indices in this work are negative.

When dealing with time-dependent searches the results are constrained in a time window. To compare different time scales the flux ϕ is converted to a fluence J by integrating in time:

$$\frac{dJ}{dE} = \int dt \frac{d\phi}{dE} = J_0 \cdot \left(\frac{E}{E_0}\right)^\gamma. \tag{B.3}$$

In practice, to convert a number of neutrinos n^* into the corresponding flux or fluence the following calculations are performed. Firstly, the fluence corresponding to 1 neutrino is evaluated using

$$fluence\,for\,one\,neutrino = \frac{band\,area}{\sum_j weight_j \left(\frac{E_j}{E_0}\right)^\gamma}, \tag{B.4}$$

where $band\,area$ is a quantity connected to the size of band around δ_{source}, the sum is made over the Monte Carlo events, $weight_j$ are the per-event weights taking into account the neutrino interaction probability in the simulation (Section 5.1), E_j are the Monte Carlo events energies and $E_0 = 1$ GeV. Then, the normalisation J_0 is calculated by multiplying the fluence corresponding to 1 neutrino for the number of neutrinos n^*:

$$J_0 = n^* \cdot fluence\,for\,one\,neutrino. \tag{B.5}$$

This is the end of the calculation for the fluence, as the discovery potential is reported in the plots for the time-dependent algorithm as

$$E^{-\gamma} \frac{dJ}{dE} = J_0 \cdot E_0^{-\gamma}, \tag{B.6}$$

that contains only J_0 and $E_0 = 1$ GeV. This quantity should be constant, thus this form highlights the dependencies on physical quantities like the declination or the flare duration.

For the time-integrated algorithm, the quantity ϕ_0 is calculated as

$$\phi_0 = \frac{J_0}{uptime}, \tag{B.7}$$

where *uptime* is the detector uptime for the considered dataset expressed in seconds. The uptime for season 2014, expressed in days, is 359.97. Also the discovery potential is reported in the time-integrated plots as

$$E^{-\gamma}\frac{d\phi}{dE} = \phi_0 \cdot E_0^{-\gamma} \tag{B.8}$$

to highlight the dependencies on the physical quantities different than the energy.

Bibliography

[1] V. F. Hess. Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten. *Physikalische Zeitschrift*, 13, 1084 (1912).

[2] J. Stettner (IceCube). “Measurement of the Diffuse Astrophysical Muon-Neutrino Spectrum with Ten Years of IceCube Data”, in 36th International Cosmic Ray Conference (ICRC): Madison, WI (2019), p. 1017.

[3] A. Schneider. “Characterization of the Astrophysical Diffuse Neutrino Flux with IceCube High-Energy Starting Events”, in 36th International Cosmic Ray Conference (ICRC): Madison, WI (2019).

[4] J. Aleksić et al. The Major Upgrade of the MAGIC Telescopes, Part I: The Hardware Improvements and the Commissioning of the System. *Astropart. Phys.* 72, 61 (2016).

[5] J. Aleksić et al. The Major Upgrade of the MAGIC Telescopes, Part II: A Performance Study Using Observations of the Crab Nebula. *Astropart. Phys.* 72, 76 (2016).

[6] J. Álvarez-Muñiz et al. “The Giant Radio Array for Neutrino Detection (GRAND): Science and Design.” Science China Physics, Mechanics & Astronomy 63.1 (2019).

[7] M. Tanabashi et al. (Particle Data Group). Review of Particle Physics. *Phys.Rev. D*, 98, 030001 (2018).

[8] P. Azzi et al. Standard Model Physics at the HL-LHC and HE-LHC, arXiv:1902.04070 (2019).

[9] The Pierre Auger Collaboration. The Pierre Auger Observatory I: The Cosmic Ray Energy Spectrum and Related Measurements. arXiv:1107.4809 (2011).

[10] T. Kintscher. Rapid Response to Extraordinary Events: Transient Neutrino Sources with the Ice Cube Experiment. PhD thesis, Humboldt-Universität zu Berlin (2020).

[11] K. Greisen. End to the Cosmic Ray Spectrum?, *Phys. Rev. Lett.* 16, 748 (1966).

[12] G. T. Zatsepin and V. A. Kuzmin. Upper Limit of the Spectrum of Cosmic Rays. *JETP Lett.* 4, [Pisma Zh. Eksp. Teor. Fiz.4,114(1966)], 78 (1966).

[13] M. Honda et al. Calculation of Atmospheric Neutrino Flux using the Interaction Model Calibrated with Atmospheric Muon Data. *Phys. Rev.* D75, 043006 (2007).

[14] *https://www.engineerit.co.za/engineerit/space-science-cosmic-rays-are-increasing-at-aviation-altitudes/*

[15] Peng Xiao-yan and Yuan Qiang. Overview of Direct Measurements of Cosmic Rays. *Chinese Astronomy and Astrophysics*, 43 (2019), pp. 327-341.

[16] J. Abraham. Measurement of the energy spectrum of cosmic rays above 10^{18} eV using the Pierre Auger Observatory, *Physics Letters B*, 685 (2010), pp. 239-246.

[17] W. Baade and F. Zwicky. Cosmic Rays from Super-novae. *Contributions from the Mount Wilson Observatory*, 3, 79 (1934).

[18] M. Ackermann et al. (Fermi-LAT). Detection of the Characteristic Pion-Decay Signature in Supernova Remnants. *Science*, 339, 807 (2013).

[19] A. Aab et al. (Pierre Auger). Observation of a Large-scale Anisotropy in the Arrival Directions of Cosmic Rays above 8×10^{18} eV. *Science*, 357, 1266 (2017).

[20] A. Aab et al. (Pierre Auger). Measurement of the Cosmic Ray Spectrum above 4×10^{18} eV Using Inclined Events Detected with the Pierre Auger Observatory. *JCAP*, 1508, 049 (2015).

[21] E. Fermi. On the Origin of the Cosmic Radiation. *Phys. Rev.* 75, 1169 (1949).

[22] P. M. Bauleo and J. R. Martino. The Dawn of the Particle Astronomy Era in Ultra-high-energy Cosmic Rays. *Nature*, 458, 847 (2009).

[23] Credit: Aurore Simonnet, Sonoma State University.

[24] T. Schwetz at al. Three-flavour neutrino oscillation update. *New J. Phys.*, 10, 113011 (2008).

[25] K. Nakamura and S.T Petcov. Neutrino mass, mixing, and oscillations. *J. Phys.*, G 37, 164-183 (2010).

[26] C. Kraus at al. Final results from phase II of the Mainz neutrino mass searchin tritium β decay. *Eur. Phys. J.*, C 40.4 (2005), pp. 447-468.

[27] M. Wurm et al. Spectroscopy of Solar Neutrinos. *Astron. Nachr.* 331, 512 (2010).

[28] S. Fukuda et al. The Super-Kamiokande detector. Nuclear Instruments and Methods in Physics Research A, 501 (2003), pp. 418-462.

[29] P. A. Cherenkov. Visible Emission of Clean Liquids by Action of γ Radiation. *Doklady Akademii nauk SSSR*, 451 (1934).

[30] J. Babson et al. (DUMAND). Cosmic Ray Muons in the Deep Ocean. *Phys. Rev.* D42, 3613 (1990).

[31] I. A. Belolaptikov et al. (BAIKAL). The Baikal Underwater Neutrino Telescope: Design, Performance and First Results. *Astropart. Phys.* 7, 263 (1997).

[32] M. Ageron et al. (ANTARES). ANTARES: The First Undersea Neutrino Telescope. *Nucl. Instrum. Meth.* A656, 11 (2011).

[33] *https://icecube.wisc.edu/science/highlights/neutrino_astronomy*

[34] K. Hirata et al. Observation of a neutrino burst from the supernova SN1987A. *Phys. Rev. Lett.*, 58 (1987), pp. 490-1493.

[35] M. G. Aartsen et al. (IceCube, Fermi-LAT, MAGIC, AGILE, ASAS-SN, HAWC, H.E.S.S., INTEGRAL, Kanata, Kiso, Kapteyn, Liverpool Telescope, Subaru, Swift NuSTAR, VERITAS, VLA/17B-403). Multimessenger Observations of a Flaring Blazar Coincident with High-energy Neutrino IceCube-170922A. *Science*, 361, eaat1378 (2018).

[36] A. Dominguez et al. Extragalactic Background Light Inferred from AEGIS Galaxy SED-type Fractions. *Mon. Not. Roy. Astron. Soc.* 410, 2556 (2011).

[37] M. Ackermann et al. (Fermi-LAT). The Fermi Large Area Telescope On Orbit: Event Classification, Instrument Response Functions, and Calibration. *Astrophys. J. Suppl.* 203, 4 (2012).

[38] A. Biland et al. Calibration and performance of the photon sensor response of FACT - The First G-APD Cherenkov telescope. *JINST*, 9, P10012 (2014).

[39] J. Holder et al. "Status of the VERITAS Observatory", in 4th Heidelberg International Symposium on High-Energy Gamma Ray Astronomy: Heidelberg, Germany, Vol. 1085, 1 (2008), pp. 657-660.

[40] M. de Naurois (H.E.S.S.). "H.E.S.S.-II - Gamma ray astronomy from 20 GeV to hundreds of TeV's", in 6th Roma International Workshop on Astroparticle Physics (RICAP): Rome, Italy, Vol. 136 (2016), p. 03001.

[41] A. López-Oramas. Multi-year Campaign of the Gamma-Ray Binary LS I $+61^{\circ}$ 303 and Search for VHE Emission from Gamma-Ray Binary Candidates with the MAGIC Telescopes. PhD thesis, Barcelona, Autonoma U. (2014).

[42] S. E. Woosley and J. S. Bloom. The Supernova Gamma-Ray Burst Connection. *Ann. Rev. Astron. Astrophys.* 44, 507 (2006).

[43] B. P. Abbott et al. (LIGO Scientific, Virgo). GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *Phys. Rev. Lett.* 119, 161101 (2017).

[44] A. Einstein. Die Grundlage der allgemeinen Relativitätstheorie. *Ann. Phys.*, 354 (1916), pp. 769-822.

[45] R. A. Hulse and J. H. Taylor. Discovery of a pulsar in a binary system. *Astrophys. J.* 195 (1975), pp. L51-L53.

[46] J. H. Taylor and J. M. Weisberg. A new test of general relativity - Gravitational radiation and the binary pulsar PSR 1913+16. *Astrophys. J.* 253 (1982), pp. 908-920.

[47] LIGO Scientific Collaboration and Virgo Collaboration. Observation of Gravitational Waves from a Binary Black Hole Merger, *Phys. Rev. Lett.*, 116 (2016), p. 061102.

[48] B. Abbott et al. Detector description and performance for the first coincidence observations between LIGO and GEO. Nuclear Instruments and Methods in Physics Research A, 517 (2004), pp. 154-179.

[49] F. Acernese et al. Societa Italiana di Relativita Generale e Fisica della Gravitazione XVII Congresso SIGRAV "General Relativity and Gravitational Physics" Torino, 4-7 Settembre 2006 (2020).

[50] S. Adrián-Martínez et al. A first search for coincident gravitational waves and high energy neutrinos using LIGO, Virgo and ANTARES data from 2007, *Journal of Cosmology and Astroparticle Physics* 2013.06 (2013), p. 008.

[51] *http://public.virgo-gw.eu/index.php?gmedia=zvg0K&t=g#PhotoSwipe1598176569786.* Credits: Swinburne Astronomy Productions

[52] P. Askebjer et al. Optical properties of deep ice at the South Pole: Absorption. *Appl. Opt.* 36, 4168 (1997).

[53] M. Ackermann et al. Optical Properties of Deep Glacial Ice at the South Pole. *J. Geophys. Res.* 111, D13203 (2006).

[54] P. B. Price and K. Woschnagg. Role of group and phase velocity in high-energy neutrino observatories. *Astropart. Phys.* 15, 97 (2001).

[55] M. Ackermann et al. Optical Properties of Deep Glacial Ice at the South Pole. *J. Geophys. Res.* 111, D13203 (2006).

[56] J. Petit, J. Jouzel, D. Raynaud et al. Climate and Atmospheric History of the Past 420,000 Years from the Vostok Ice Core, Antarctica. *Nature*, 429 (1999).

[57] O. Watanabe et al. The Paleoclimate Record in the Ice Core from Dome Fuji Station, Antarctica. *Ann. Glaciol.* 176 (1999).

[58] M. G. Aartsen et al. (IceCube). Measurement of South Pole ice transparency with the IceCube LED calibration system. *Nucl. Instrum. Meth.* A711, 73 (2013).

[59] D. Chirkin (IceCube). "Evidence of Optical Anisotropy of the South Pole Ice", in 33rd International Cosmic Ray Conference (ICRC): Rio de Janeiro, Brasil (2013), p. 0580.

[60] M. G. Aartsen et al. (IceCube). The IceCube Neutrino Observatory: Instrumentation and Online Systems. *JINST*, 12, P03012 (2017).

[61] R. Abbasi et al. (IceCube). The IceCube Data Acquisition System: Signal Capture, Digitization, and Timestamping. *Nucl. Instrum. Meth.* A601, 294 (2009).

[62] M. G. Aartsen et al. (IceCube). Energy Reconstruction Methods in the IceCube Neutrino Telescope. *JINST*, 9, P03009 (2014).

[63] M. G. Aartsen et al. (IceCube). Search for Astrophysical Sources of Neutrinos Using Cascade Events in IceCube. *Astrophys. J.* 846, 136 (2017).

[64] A. Gazizov and M. P. Kowalski. ANIS: High Energy Neutrino Generator for Neutrino Telescopes. *Comput. Phys. Commun.* 172, 203 (2005).

[65] A. M. Dziewonski and D. L. Anderson. Preliminary Reference Earth Model. *Physics of the Earth and Planetary Interiors*, 297 (1981).

[66] D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz and T. Thouw. CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers (1998).

[67] D. Chirkin and W. Rhode. Muon Monte Carlo: A High-precision Tool for Muon Propagation through Matter (2004).

[68] S. Tilav, P. Desiati, T. Kuwabara, D. Rocco, F. Rothmaier, M. Simmons, and H. Wissing (IceCube). "Atmospheric Variations as Observed by IceCube", in 31st International Cosmic Ray Conference (ICRC): Łódź, Poland (2009), p. 894.

[69] M. G. Aartsen et al. (IceCube). The IceCube Realtime Alert System. *Astropart. Phys.* 92, 30 (2017).

[70] D. Gora. "High-energy Gamma-ray Follow-up Program Using Neutrino Triggers from IceCube", in 34th International Cosmic Ray Conference (ICRC): The Hague, The Netherlands (2015), pp. 16-19.

[71] M. G. Aartsen et al. Improvement in Fast Particle Track Reconstruction with Robust Statistics. *Nucl. Instrum. Meth.* A736, 143 (2014).

[72] P. J. Huber. Robust Estimation of a Location Parameter. *Ann. Math. Statist.* 35 (1964), pp. 73-101.

[73] J. Ahrens et al. (AMANDA). Muon Track Reconstruction and Data Selection Techniques in AMANDA. *Nucl. Instrum. Meth.* A524, 169 (2004).

[74] D. Pandel. Bestimmung von Wasser- und Detektorparametern und Rekonstruktion von Myonen bis 100 TeV mit dem Baikal-Neutrinoteleskop NT-72. Diploma thesis, Humboldt-Universität zu Berlin (1996).

[75] M. Voge. Searches for Neutrinos from Supernovae Using Cherenkov In-Ice Detectors. PhD thesis, Humboldt-Universität zu Berlin, (2016).

[76] P. H. C. Eilers and B. D. Marx. Flexible Smoothing with B-Splines and Penalties. *Statistical Science*, 11, 89 (1996).

[77] N. Whitehorn , J. van Santen , and S. Lafebre. Penalized Splines for Smooth Representation of High-dimensional Monte Carlo Datasets. *Comput. Phys. Commun.* 184, 2214 (2013).

[78] L. Rädel and C. Wiebusch. Calculation of the Cherenkov Light Yield from Electromagnetic Cascades in Ice with Geant4. *Astropart. Phys.* 44 (2013), pp. 102-113.

[79] J.-D. Zornoza and D. Chirkin (IceCube). "Muon Energy Reconstruction and Atmospheric Neutrino Spectrum Unfolding with the IceCube Detector", in 30th International Cosmic Ray Conference (ICRC): Mérida, Mexico, Vol. 5, pp. 1275-1278 (2007).

[80] F. Pedregosa et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825 (2011).

[81] *http://software.icecube.wisc.edu/documentation/projects/pybdt/index.html*

[82] G. C. Hill. "Bayesian Event Reconstruction and Background Rejection in Neutrino Detectors", in 27th International Cosmic Ray Conference (ICRC): Hamburg, Germany, Vol. 3 (2001), p. 1279.

[83] H. Cramér. *Mathematical Methods of Statistics*, Princeton Mathematical Series (Princeton Univ. Press, 1999).

[84] C. Radhakrishna Rao. Information and the Accuracy Attainable in the Estimation of Statistical Parameters. *Bull. Calcutta Math. Soc.* 37, 81 (1945).

[85] R. Abbasi et al. (IceCube). Time-integrated Searches for Point-like Sources of Neutrinos with the 40-String IceCube Detector. *Astrophys. J.* 732, 18 (2011).

[86] J. Lünemann. Suche nach Dunkler Materie in Galaxien und Galaxienhaufen mit dem Neutrinoteleskop IceCube. PhD thesis, Johannes Gutenberg-Universität Mainz, (2013).

[87] T. Neunhöffer. Estimating the Angular Resolution of Tracks in Neutrino Telescopes Based on a Likelihood Analysis. *Astropart. Phys.* 25, 220 (2006).

[88] B. Efron. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7, 1 (1979).

[89] T. Bray. *The JavaScript Object Notation (JSON) Data Interchange Format*, RFC 8259 (2017).

[90] K. Maine, C. Devieux, and P. Swan. "Overview of IRIDIUM Satellite Network", in Proceedings of WESCON'95 (1995), p. 483.

[91] 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.* D75, 043006 (2007).

[92] S. L. Glashow. Resonant Scattering of Antineutrinos. *Phys. Rev.* 118, 316 (1960).

[93] *https://gcn.gsfc.nasa.gov/gcn_main.html*

[94] R. Seaman et al. IVOA Recommendation: Sky Event Reporting Metadata Version 2.0. (2011).

[95] R. Franke. Design, Implementation and First Results of the Neutrino Triggered Target of Opportunity Program with the IceCube Neutrino Telescope. PhD thesis, Humboldt-Universität zu Berlin (2014).

[96] S.S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *Annals Math. Statist.* 9, 60 (1938).

[97] A. Albert et al. (ANTARES, IceCube). Joint Constraints on Galactic Diffuse Neutrino Emission from the ANTARES and IceCube Neutrino Telescopes. *Astrophys. J.* 868, L20 (2018).

[98] C. E. Bonferroni. Teoria Statistica delle Classi e Calcolo delle Probabilità. *Publ. R. Istit. Super. Sci. Econ. Commerc. Firenze*, 8, 1 (1936).

[99] J. van Santen (IceCube Gen2). IceCube-Gen2: the next-generation neutrino observatory for the South Pole. *PoS*, ICRC2017 (2018), p. 991.

[100] M. G. Aartsen et al. (IceCube-Gen2). IceCube-Gen2: a vision for the future of neutrino astronomy in Antarctica. arXiv:1412.5106 (2014).

[101] M. G. Aartsen et al. Neutrino astronomy with the next generation IceCube Neutrino Observatory. arXiv:1911.02561 (2019).

[102] M. G. Aartsen et al. IceCube-Gen2 - The Next Generation Neutrino Observatory at the South Pole: Contributions to ICRC 2015. arXiv:1510.05228 (2015).

[103] M. G. Aartsen et al. (IceCube-PINGU). Letter of Intent: the Precision Icecube Next Generation Upgrade (PINGU). arXiv:1401.2046 (2014).

[104] M. G. Aartsen et al. (IceCube-Gen2). PINGU: A Vision for Neutrino and Particle Physics at the South Pole. *Journal of Physics G: Nuclear and Particle Physics* 44.5 (2017), p. 054006.

[105] P. Allison et al. (ARA). Performance of two Askaryan Radio Array stations and first results in the search for ultra-high energy neutrinos. *Phys. Rev. D* 93.8 (2016).

[106] *https://skyandtelescope.org/astronomy-resources/right-ascension-declination-celestial-coordinates/*

[107] J. Meeus. Astronomical Algorithms. *Willmann-Bell, Inc.*, Richmond, VA, second edition, chap. 13 (1998).

[108] *https://it.wikipedia.org/wiki/Parsec*