University of Padova

Department of Physics and Astronomy "Galileo Galilei"

Master Thesis in Astrophysics and Cosmology

Confidence regions for the direction of all-sky neutrino alerts in IceCube

Supervisor
Prof. Elisa Bernardini
University of Padova

Co-supervisors
Sarah Louise Mancina
Diego Alberto Coloma Borja
University of Padova

Master Candidate
Alexandre Lillo

Student ID
2118977

Academic Year
2024-2025

"Experience is a lantern that only lights its own way"
— Jean-Marc Lillo

Abstract

The origin of cosmic rays is one of the principal questions at the center of multi-messenger astrophysics. Cosmic ray interactions at the site of acceleration are known to produce neutrinos and gamma-rays. Hence, a proper identification of the sources emitting these particles at high energies is instrumental in understanding where and how cosmic rays are accelerated. The field of multi-messenger astrophysics regroups information carried by different messengers (photons, neutrinos, cosmic rays, gravitational waves) and thus allows for a deeper understanding of the common source. In real-time correlation studies, simultaneous observations of several messengers, e.g. astrophysical neutrinos and gamma-rays, enhance the significance of each separate detection for a broader picture of the phenomena occurring in real-time.

The IceCube Neutrino Observatory participates in the multi-messenger effort by working to identify sources of astrophysical neutrinos. The Gamma-ray Follow-Up (GFU) branch of IceCube monitors known gamma-ray emitters and warns in real-time partner telescopes when a cluster of neutrino events are detected coming from these sources. It also performs an all-sky unbiased search for flares of neutrinos.

This work aims to investigate the creation of confidence regions around the estimated source positions for all-sky GFU cluster alerts. Showing that there is a simple, but robust way to locate the potential source of neutrino flares in the context of an unbiased search could later enable the sharing of these alerts.

We do this by comparing the performances of the algorithm at the center of the alerting process with the predictions from Wilks' theorem. This theorem would equate the distribution of the difference in log-likelihood between the algorithm estimation and the actual position of the source to a χ^2 distribution with two degrees of freedom.

We will show that this expectation proves wrong in most cases, but by letting the number of degrees of freedom for the χ^2 distribution be a free parameter, we will be able to fit the data accurately for different hypotheses of signal strength, duration and location in the sky. Ultimately, the project leads to the creation of confidence bands to encase the 50 and 90% confidence regions. This method to estimate the source position for an all-sky alert will hopefully help determine how best to direct the attention of partner telescopes to coordinate observations.

Contents

Listing of figures

Listing of tables

Listing of acronyms

AMANDA ..... Antarctic Muon and Neutrino Detector Array

ANTARES ..... Astronomy with a Neutrino Telescope and Abyss Environmental Research

BDT ..... Boosted Decision Tree

CDF ..... Cumulative Distribution Function

CMB ..... Cosmic Microwave Background

DOM ..... Digital Optical Module

DUMAND ..... Deep Underwater Muon And Neutrino Detector

EBL ..... Extragalactic Background Light

GCN ..... Gamma-ray Coordinate Network

GFU ..... Gamma-ray Follow-Up

HEALPix ..... Hierarchical Equal Area isoLatitude Pixelization

H.E.S.S. ..... High Energy Stereoscopic System

HLC ..... Hard Local Coincidence

IACT ..... Imaging Atmospheric Cherenkov Telescope

JSON ..... JavaScript Object Notation

KM3NeT ..... Cubic Kilometre (KM^3) Neutrino Telescope

LLH ..... Log-Likelihood

LST ..... Large-Sized Telescope

MAGIC ..... Major Atmospheric Gamma-ray Imaging Cherenkov Telescope

MLE ..... Maximum Likelihood Estimate

MPE ..... Multi PhotoElectron

MuEX ..... Muon Energy Estimation

PDF ..... Probability Density Function

PMT ..... Photo-Multiplier Tube

RNO-G ..... Radio Neutrino Observatory - Greenland

ROC ..... RealTime Oversight Committee

SMT-8 ..... Simple Multiplicity Trigger

SPE ..... Single PhotoElectron

TS ..... Test Statistic

UHECR ..... Ultra High Energy Cosmic Rays

VERITAS ..... Very Energetic Radiation Imaging Telescope Array System

1 Introduction

Astroparticle physics is the study of the high-energy astrophysical phenomena through the particles they emit. Among those, cosmic rays can achieve extremely high energies by being accelerated in manners which are still unknown. Interactions involving cosmic rays can produce neutrinos and gamma-rays, which also reach very high energies. The branch of astroparticle physics consisting in studying several of these *messengers* (photons, neutrinos, cosmic rays, and more recently, gravitational waves) coming from the same source to piece together a broader picture of the physics happening is called *multi-messenger astrophysics*.

The IceCube Neutrino Observatory [1], as the largest neutrino detector, contributes to this field by alerting telescopes when a significant flare of neutrinos is detected coming from the same point in the sky. IceCube is able to monitor the whole sky almost continuously, and in doing so, can notice unbiased excesses standing out from the neutrino background with a good chance of being of astrophysical origin.

This all-sky search is one of the mode of operation for the Gamma-ray Follow-Up (GFU) branch of IceCube, which contributes to the effort of the multi-messenger astrophysics community by suggesting follow-ups from Imaging Atmospheric Čerenkov Telescopes (IACTs) when it notices a neutrino flare coming from a known gamma-ray emitter. However, alerts raised by the all-sky search are not automatically shared because the source of the signal is not constrained enough.

This works aims to facilitate the sharing of all-sky alerts by creating confidence regions for the source positions of all-sky GFU alerts. This shall be done first by trying to compare the output of the time-clustering algorithm used to determine the significance of neutrino flares with Wilks' theorem, before finding a better way to represent the data and forward the confidence regions to the partner telescopes.

This chapter will first and foremost lay the basics of multi-messenger astrophysics by introducing the relevant messengers and the challenges faced by IceCube and other neutrino-oriented astronomical instruments.

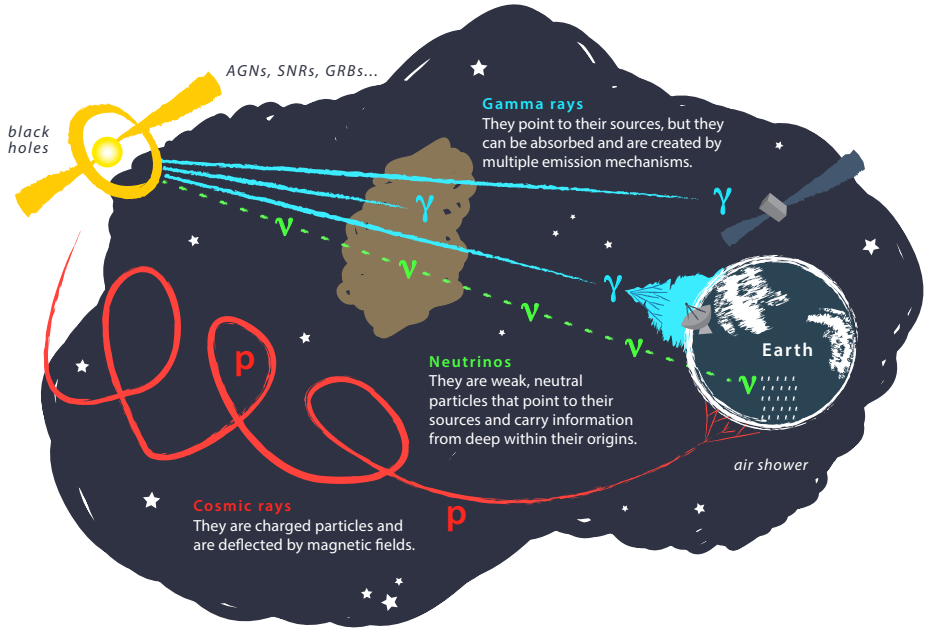


Figure 1.1: Representation of the propagation and detection techniques associated with different messengers, with an overview of what differentiate each of them. Examples of emitters are also mentioned. Credit for the image: Juan Antonio Aguilar and Jamie Yang. IceCube/WIPAC.

1.1 Photons

Light has always been mankind's first carrier of information about the universe. But beyond the narrow visible spectrum, light comes in many forms. Depending on the energy of the traveling photons (quanta of light), the information that can be extracted from a detection varies. Every photon energy is associated with a wavelength; the higher the energy, the shorter the wavelength. Since physical processes typically emit photons in a particular energy range, observations over a broad region of the electromagnetic spectrum can reveal a lot about the interactions happening.

The only type of photons relevant to this thesis are gamma-rays, which have an energy $\geq$ 511 keV, corresponding to the mass of an electron if it were converted into energy [2]. They can be produced either through leptonic interactions, when highly accelerated electrons lose energy via inverse Compton scattering or via synchrotron radiation in a magnetic field, or through hadronic interactions, when a high-energy proton or nucleus interacts in a gas or radiation field. In this case, pions are produced through a *delta resonance* [3]. For example:

$$p + \gamma \rightarrow \Delta^{+} \rightarrow \begin{cases} p + \pi^{0} \\ n + \pi^{+} \end{cases} \tag{1.1}$$

This is important, since neutral pions decay into gamma-rays, while charged pions decay into muons and neutrinos:

$$\begin{aligned} \pi^{0} &\rightarrow \gamma + \gamma \\ \pi^{\pm} &\rightarrow \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu}) \\ &\rightarrow e^{\pm} + \nu_{e}(\bar{\nu}_{e}) + \bar{\nu}_{\mu}(\nu_{\mu}) + \nu_{\mu}(\bar{\nu}_{\mu}) \end{aligned} \tag{1.2}$$

This means that if protons are accelerated to very high energies interacting with matter or radiation are expected to produce neutrinos and gamma-rays at the same time. This forms the motivation basis to coordinate observations of both messengers.

Gamma-rays are neutrally charged, they are not deflected by magnetic fields. However, the higher their energy is, the larger their cross-section becomes for the pair production resulting from the interaction of a very high-energy gamma-ray with diffuse photon fields such as

the Cosmic Microwave Background (CMB) or the Extragalactic Background Light (EBL) [3]. This opacity of the universe to gamma-rays above a certain energy means that the most energetic observed photons cannot come from outside of the Milky Way.

Throughout this thesis, we will reference Imaging Atmospheric Čerenkov Telescopes (IACTs), which are ground-based arrays of telescopes designed to detect the signatures left by Very High Energy gamma-rays (0.1 TeV $\leq E \leq$ 100 TeV). When such a photon enters the atmosphere, it will trigger an electromagnetic shower as shown in Figure 1.2. The electrons and positrons produced briefly travel faster than light in the atmosphere and emit blue Čerenkov light that is focused by mirrors onto photomultipliers. Examples of these telescopes includes the Major Atmospheric Gamma-Ray Imaging Cherenkov (MAGIC) [4] in La Palma, Spain, the High Energy Stereoscopic System (H.E.S.S.) [5] in Namibia, and the Very Energetic Radiation Imaging Telescope Array System (VERITAS) [6] in Arizona, USA. However, IACTs are only operating on clear, moonless nights, and function on a pointed mode, which means they can only monitor a narrow portion of the sky at a given time. Receiving follow-up alerts from real-time observation of other messengers is crucial to their efficiency.

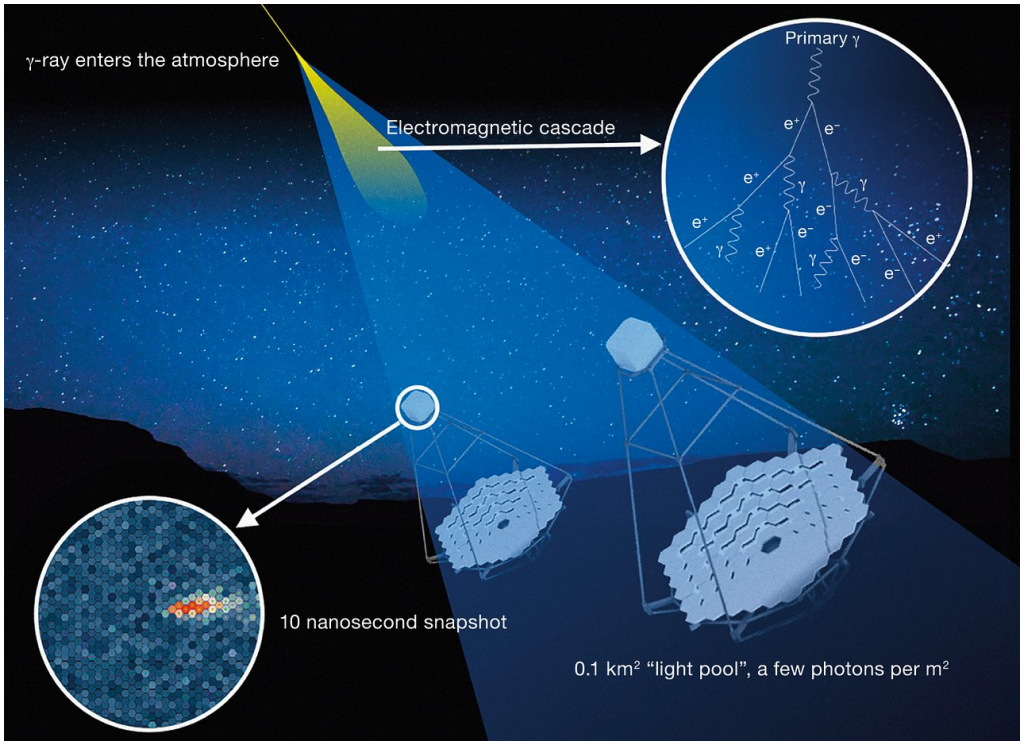


Figure 1.2: Illustration of the detection method of IACTs. Source: [7].

1.2 Cosmic rays

Discovered in 1912 by Victor Hess when he detected an increase in radiation with altitude that could not be caused by the Sun [8], cosmic rays are relativistic charged particles, mostly protons or heavier nuclei. Their charge means that they will be deflected by the magnetic fields they encounter in space, and thus their direction of origin cannot be reconstructed accurately.

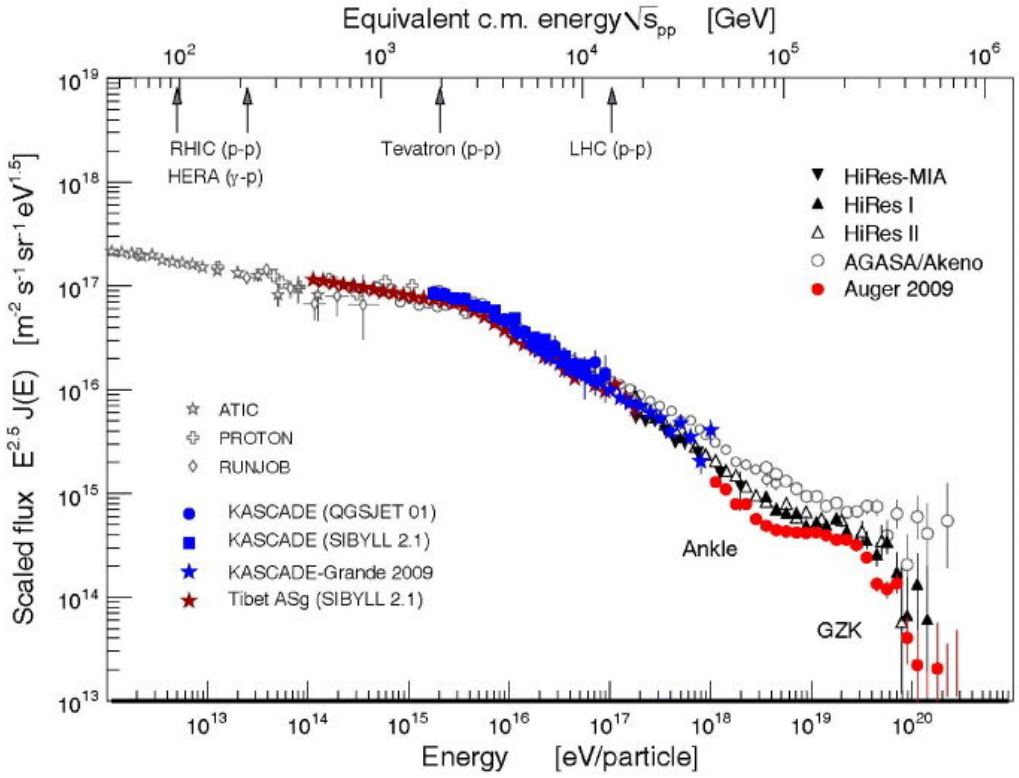


Figure 1.3: The cosmic ray energy spectrum as observed by multiple experiments. The plotted flux has been multiplied by $E^{2.5}$ to allow for the knee and ankle to be noticed more easily. Taken from [9].

The energy spectrum of cosmic rays is studied in depth but still not fully understood. It spans multiple orders of magnitude and approximately follows a power law such that [3]:

$$\frac{d\varphi}{dE} \propto E^{\gamma}, \tag{1.3}$$

where φ is the observed flux and γ is the spectral index of the power law. γ is dependent on the energy E, and so two changes are observable in Figure 1.3:

- The knee ($\approx 10^{15} - 10^{16}$ eV), where the spectrum softens from a spectral index $\gamma \approx -2.7$ to $\gamma \approx -3.1$. The commonly admitted explanation for this is that cosmic rays of higher energies are no longer confined to the galaxy by its magnetic field, but it could also mean that this is the maximum energy to which galactic sources can accelerate cosmic rays [3].

- The ankle ($\approx 10^{18}$ eV) sees a hardening of the spectrum that is not conclusively explained as of yet. It marks the lower bound for Ultra High Energy Cosmic Rays (UHECR).

The flux drops abruptly at energy above $\approx 5 \times 10^{19}$ eV. This is usually explained by the Greisen-Zatespin-Kuzmin (GZK) cut-off, which states that at these energies, UHECR interact with CMB photons to produce the same delta resonance shown in Equation 1.1 [10, 11]. However, there is no certainty that this is the reason for the drop in the spectrum, and it could be that these energies are simply the highest energies a cosmic ray can be accelerated to.

To know which sources are capable of bringing cosmic rays to such extreme energies, one can refer to the Hillas diagram, represented in Figure 1.4. It relies the Hillas criterion to evaluate what is the maximum energy to which an accelerator can bring a cosmic ray [12], namely:

$$E_{MAX} = qBR, \tag{1.4}$$

with q being the charge of the accelerated particle, B the magnitude of the magnetic field of the accelerator, and R its radius.

1.3 Neutrinos

Neutrinos are neutral leptons that interact only via the weak force. They are predicted to be massless by the Standard Model, but evidence of neutrino oscillations between the three flavors of neutrinos (electron neutrino ν_e, muon neutrino ν_μ, and tau neutrino ν_τ) prove that their mass is small but nonzero.

Neutrinos can be produced in the atmosphere as results of the interaction of a cosmic ray with a nucleus at high altitude. This produces an Extensive Air Shower with an electromagnetic component (similar to an electromagnetic shower triggered by a high-energy gamma-ray), a hadronic component, a muonic component, and finally a neutrino component. These atmospheric neutrinos are the background against which the astrophysical neutrinos need to

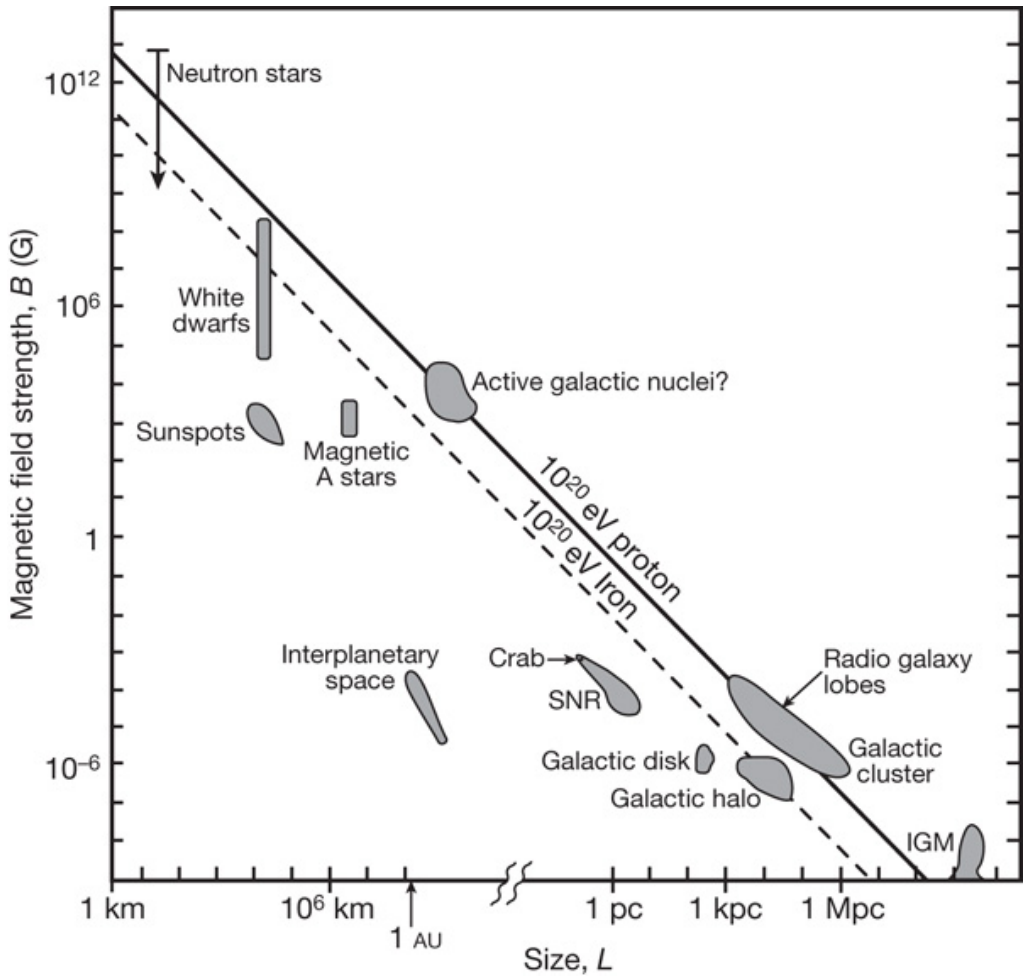


Figure 1.4: The Hillas diagram shows candidate UHECR accelerators as a function of size and magnetic field, since those properties are determining the maximum energy at which the cosmic ray can be brought to, as per Equation 1.4. Taken from [13].

compete. Astrophysical neutrinos are the high-energy products of interactions between highly-accelerated cosmic rays and matter or radiation. Neutrinos retain around 5% of their parent cosmic ray's energy [3], so the detection of a high-energy neutrino implies that its progenitor's energy was an order of magnitude greater.

1.3.1 Challenges and prospects of neutrino astronomy

Since they are electrically neutral and have a very small cross-section, neutrinos pass through matter with very little attenuation and can provide insight about the inside regions of astrophysical accelerators. The observed neutrinos have neither been deflected or scattered by mag-

netic fields or obstacles. This also means that their detection proves to be very challenging. In fact, what is detected by neutrino observatories are the results of neutrino interaction in a dense medium. It will produce a lepton moving faster than light in this medium, which in turn emits Čerenkov light detected by optical sensors. The trajectory of the particle can then be computed by taking into account that the light was emitted in a cone opened by the Čerenkov angle, defined as:

$$\cos(\theta_C) = \frac{1}{n\beta}, \tag{1.5}$$

where $\beta \approx 1$ is the velocity of the muon and n is the refractive index of the medium.

The only two suitable natural media for the installation of a neutrino detector are glaciers or the bottom of the sea. Several experiments have taken place deep underwater: DUMAND (Deep Underwater Muon And Neutrino Detector) [14] between 2000 and 4000 m in the Pacific Ocean, Baikal [15] at the bottom of lake Baikal in Russia, ANTARES (Astronomy with a Neutrino Telescope and Abyss Environmental Research) [16] in the Mediterranean sea and its current successor KM3NeT (Cubic Kilometre Neutrino Telescope) [17]. In the ice, the first instrument deployed was AMANDA (Antarctic Muon and Neutrino Detector Array) [18], at the Amundsen-Scott station at the South Pole, before being integrated in 2005 by IceCube, the current largest neutrino observatory in the world, encompassing a volume of 1 km^3 of ice. Technical aspects of IceCube are presented in Subsection 2.1.1.

IceCube and other modern neutrino observatories, such as KM3NeT, need to be this large because of the so-called Waxman-Bachall bound [19]. Based on UHECR observations and upper bound placed on the models for neutrino production, it states that the minimum volume needed to detect an astrophysical neutrino is of the order of 1 km^3. Indeed, the background is so overwhelming in these experiments that an event needs to be extremely significant to be considered as potentially signal-like. The process of filtering out the background from the data received by IceCube is detailed in Section 2.2 and Section 3.1.

An upgrade to IceCube, called IceCube-Gen2 [20], is currently under development, to bring the volume of the detector to 10 km^3, as well as integrating other means of detecting Ultra High Energy neutrinos with an array of radio antennas. This relies on the same principle than the Radio Neutrino Observatory - Greenland (RNO-G) [21] being built in Greenland. When individual particles in a shower emit Čerenkov radiation, interference between wavelengths

longer than the extent of the shower produce a front of radio waves along the Čerenkov cone. This is known as Askaryan radiation [22].

2

The IceCube Neutrino Observatory

To advance the field of multi-messenger astrophysics as a whole, it is important to observe and correctly analyse as much data as possible. It was stated in Subsection 1.3.1 that large detectors are necessary to observe high-energy neutrinos. Indeed, most of them aim to detect the Čerenkov light emitted when a charged particle travels faster than light in a dense medium after the neutrinos have interacted in said medium, then detected by optical sensors. The scattering and absorption lengths of the medium must be as long as possible, so that the observation of a photon can be associated precisely to an emission position. This is why multiple instruments have been deployed deep in water, most recently KM3NeT [17], or in ice, like IceCube [1], the primary observatory featured in this project.

This chapter will first provide an overview of the hardware used by IceCube and typical shape of the events detected by the observatory. Then it will describe the filters and reconstruction algorithms real time observations in IceCube have to go through to be qualified as relevant, and used in the Gamma-ray Follow-Up analysis such as the ones performed in this project.

2.1 Neutrino events in IceCube

Constructed between 2004 and 2011, IceCube is located at the geographic South Pole. It built upon the results of AMANDA [18], an earlier experiment in the Antarctic ice, that IceCube incorporated in 2010 before it was decommissioned. It deploys 5160 digital optical modules

(DOMs) between 1450 and 2450 m deep in the Antarctic glacier, deep enough for the pressure to rid the ice of air bubbles that would otherwise refract and deviate the Čerenkov photons [23]. Its design and performances make it the largest neutrino observatory currently, and it is working to detect astrophysical neutrinos in the TeV-PeV range.

2.1.1 Technical overview of the detector

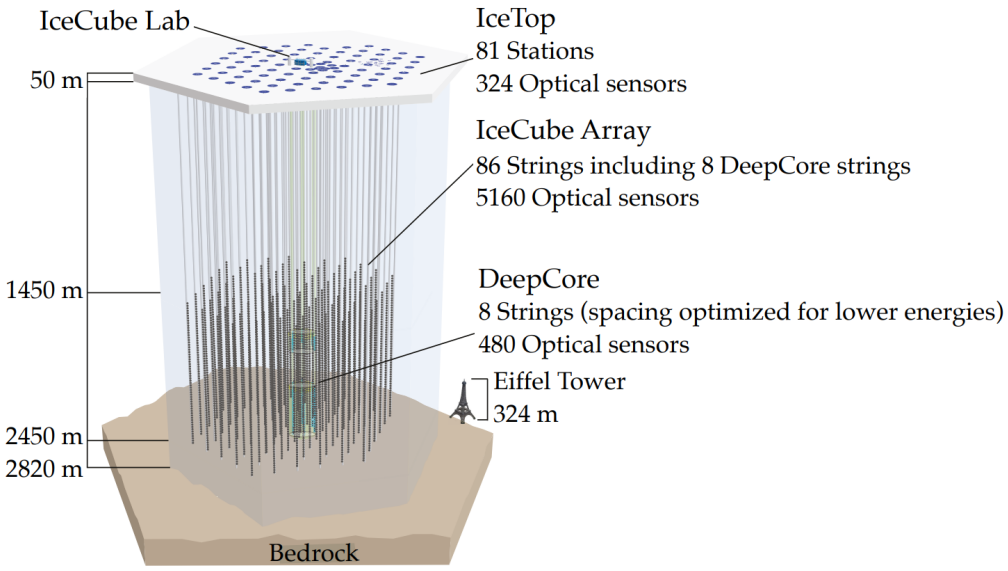


Figure 2.1: Layout of the IceCube Neutrino Observatory. Image taken from [24].

Layout Figure 2.1 shows the layout of IceCube. The DOMS are aligned in a quasi-hexagonal pattern on 86 vertical strings, each containing 60 DOMs spaced 17 m. They are linked along the strings by a cable which goes up to the *IceCube Lab*, a facility on the surface where power and communication are managed. The primary detector is comprised of 78 out of the 86 strings, aligned on an hexagonal grid, separated by 125 m each.

The rest is called *DeepCore*, it is a sub-detector where the strings and DOMs are more closely parsed (41-105 m for the strings and 7-10 m for the DOMs) [25]. The objective of DeepCore is to be sensitive to a lower energy range, with an energy threshold of 10 GeV, 10 times lower than in the rest of IceCube. It can then be used to study atmospheric neutrino oscillations and search for sterile neutrinos.

Finally, there are also 162 ice-filled tanks sitting in pairs on the ice surface, each containing two DOMs. This array, called *IceTop*, serves as a veto for air showers and aims to observe air showers induced by cosmic-rays in the PeV-EeV energy range [26]. As this is the energy range where UHECR would emerge, one might expect this array to be able to study this part of the spectrum, but the surface it covers (about 1 km^2) is too small for such observations.

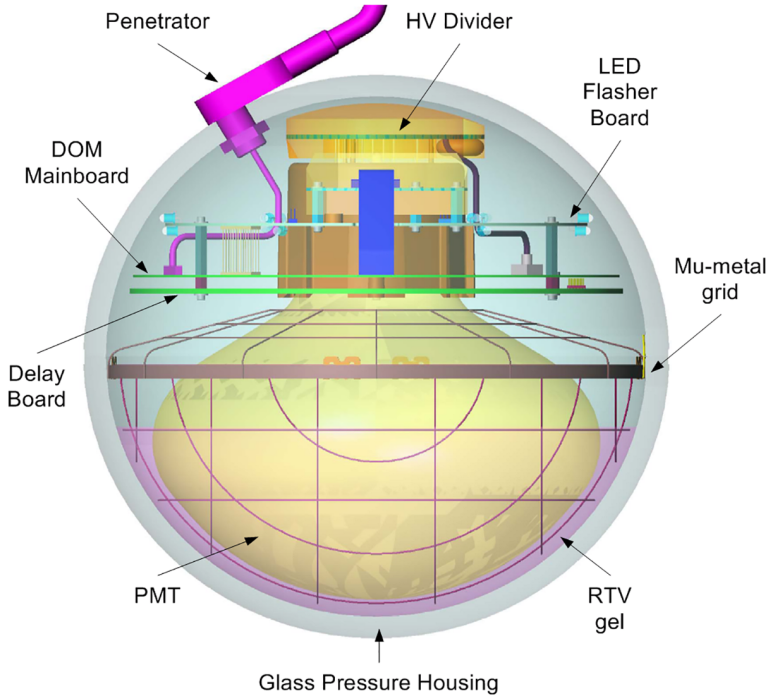


Figure 2.2: Sketch of a Digital Optical Module, as seen in [27].

The Digital Optical Modules (DOMs) Digital Optical Modules host the principal sensor components of the detector [24]. The DOMs are encased in a 13" diameter glass sphere, as schematised in Figure 2.2. This sphere is made up of two hemisphere which are joined at the equator at the end of production to englobe the circuit boards necessary to power supply, data acquisition, communication, and calibration, as well as the main component of the DOM: a downward-facing 10" photomultiplier tube (PMT). There is only one hole in the sphere to allow the 16 mm cable to exit the module.

This glass sphere protects the electronics and the PMT from the pressure from the glacier. Furthermore, the whole structure of the DOM, the cable penetrator assembly and the suspen-

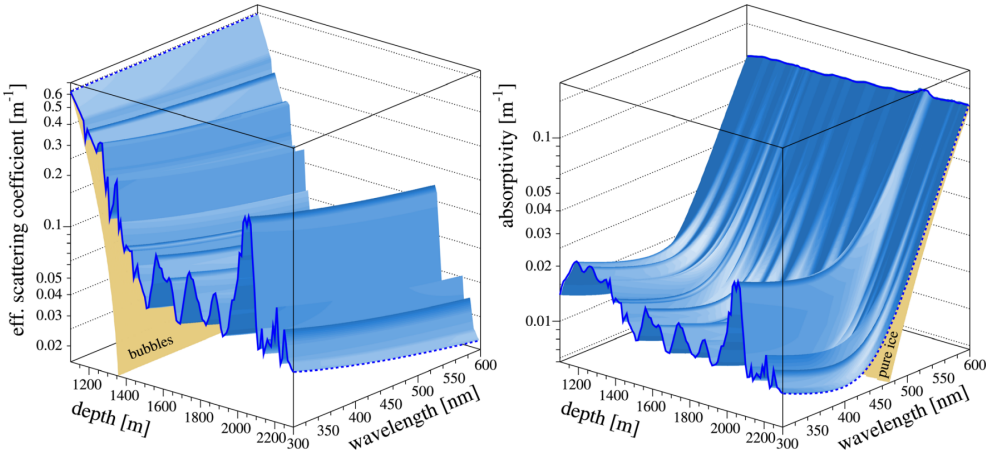


Figure 2.3: Effective scattering coefficient (left), and absorptivity (right) in the Antarctic ice, depending on the depth in the glacier and the wavelength. They are the inverse of the scattering and absorption mean free paths depending on wavelength and depth. The contribution from the air bubbles above 1400 m to the scattering coefficient is shown, as well as the expected absorptivity of pure ice for wavelengths longer than 450 nm. Taken from [23].

sion from the string have been designed to withstand the extreme conditions deep in the ice. In addition to this, a wire mesh grid surrounds the PMT and shields it from the South Pole magnetic field, which would otherwise degrade the collection efficiency and the single photoelectron resolution.

The photomultiplier, being sensitive in the 300-650 nm range, has a quantum efficiency, i.e. the probability of production of photoelectrons from photons, of 25% at 390 nm. In DeepCore, most DOMs have a higher quantum efficiency of 34%. The PMT is able to record varied waveforms, with amplitudes between 1 mV and 2 V, and widths between 12 and 1500 ns. This depends on the energy of the observed particle and its distance to the DOM. The PMT is also connected to circuit boards for various functions, two of which being the *Flasher board* and the *Delay board*. The former serves as a calibration tool and in the study of the properties of the ice [28], while the latter records and digitizes the waveform received by the PMT.

Properties of the ice It has already been mentioned that the deep Antarctic glacier was chosen because allows some of the best conditions to detect Čerenkov light. Indeed, the approximately 1 km^3 of ice delimited by IceCube is even clearer than ice produced in laboratories or in lakes [29]. However, photons are still being scattered and absorbed. To avoid losing efficiency in the detectors and accuracy in the reconstruction, a precise understanding of the ice proper-

ties deep in the glacier is fundamental. Many studies, past and ongoing, have investigated this subject, and over the years the models of the ice have been refined to account for many parameters that affect the optical properties of the ice.It is now known that as deep as IceCube, dust particles like mineral grains or sea salt crystals are responsible for scattering, and that absorption is also mainly caused by insoluble dust impurities. Figure 2.3 shows us that below 1400 m, the scattering and absorptions coefficients are very small, save for minor peaks, among which the only remarkable one occurs at 2000 m and is caused by a layer of dust approximately 65000 years old [30].

2.1.2 Event detection and signatures

Much like most high-energy observatories, in IceCube, it is not the neutrino that is being directly observed traversing the detector. Instead, it is sensitive to the deep inelastic scattering in which a neutrino will be involved with a nucleus in or near the detector. This produces hadronic showers, and most of the time a lepton of the same flavor of the neutrino, depending on whether the interaction is *neutral-current*,

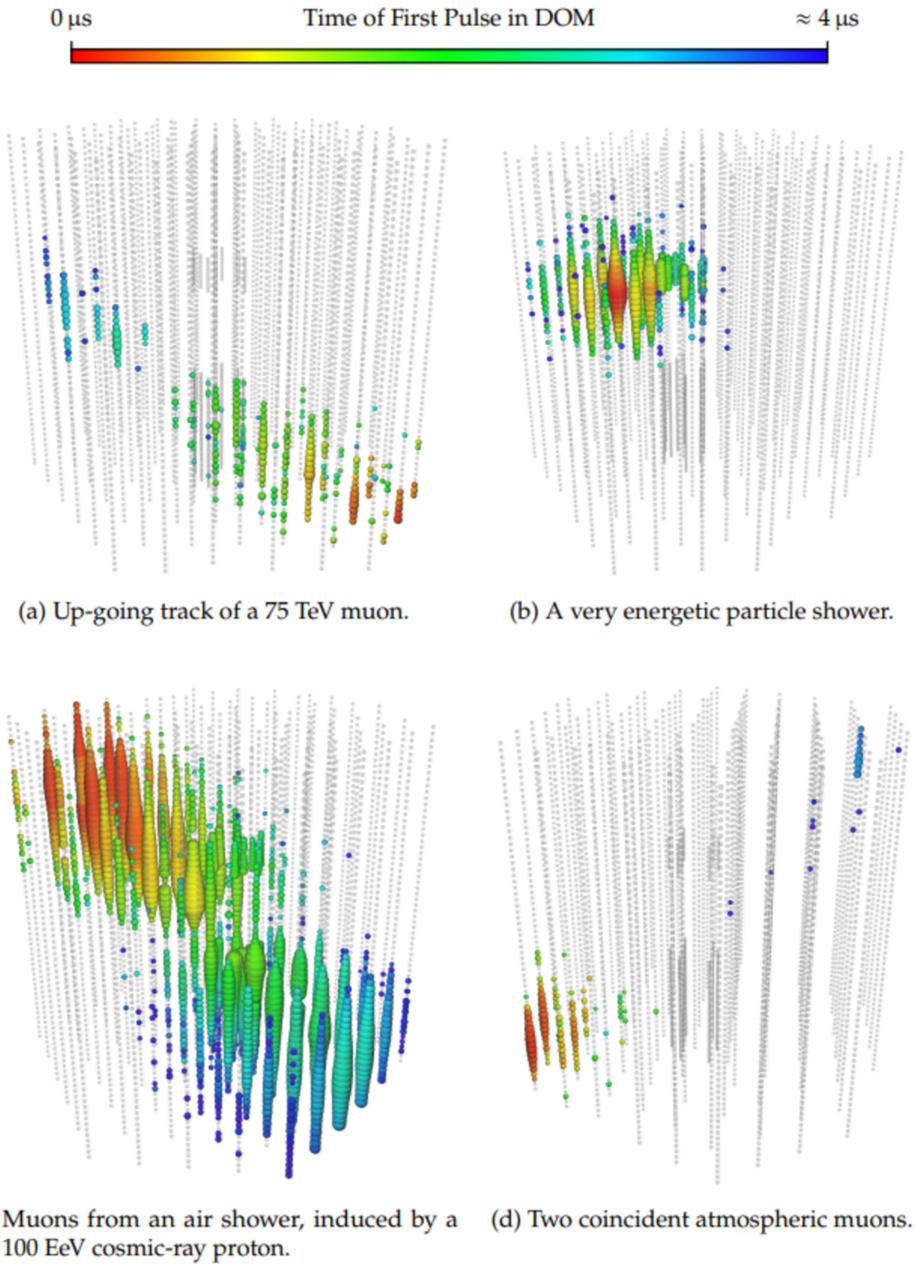
$$\nu_l + X \xrightarrow{Z} \nu_l + X' \rightarrow \nu_l + \text{hadronic shower}, \tag{2.1}$$

where ν_l is undetected but X' produces a hadronic shower ; or *charged-current*,

$$\nu_l + X \xrightarrow{W} l + X'. \tag{2.2}$$

This lepton shall travel faster than light in the ice will thus emit blue Čerenkov light and leave a distinct signature in IceCube depending on its flavor that can be recognized, pictured in Figure 2.4. These processes are also valid for anti-neutrinos, where the only difference would be the charge of the produced lepton for a charged-current interaction. However, it is impossible to distinguish between neutrino and anti-neutrino interactions, except in rare cases, such as the Glashow resonance events [31].

When the lepton created is a muon of energy higher than 1 GeV, the detector will be able to observe it travel in the detector in a linear *track-like* signature, either *through-going* (meaning they were produced outside of the effective volume of the detector), or *starting* (meaning they were produced inside the detector). Tracks are reconstructed with a good directional accuracy [33], the typical angular uncertainty being 0.3° for 1 TeV energies. However, the energy re-



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

(b) A very energetic particle shower.

(c) Muons from an air shower, induced by a 100 EeV cosmic-ray proton.

(d) Two coincident atmospheric muons.

Figure 2.4: Examples of recognizable event signatures in IceCube, where the circles are color-coded in terms of time evolution. Subfigure (a) shows a muon track, this is the kind of events relevant to this work. Subfigures (b), (c), and (d), respectively a cascade event, a muon bundle, and two coincident events, are then to be filtered out by the processes described in Section 2.2. Image taken from [32].

construction proves more difficult for this kind of event, especially when they originate from outside the detector and may have already lost energy before being first observed so the energy

losses in the detector have to be approximated.

On the other hand, *cascades* are the class of signature including the electromagnetic showers triggered by the production of an electron or a tau during the scattering of a neutrino, the hadronic showers produced by fragmented nuclei, and other decays from particles scattering in the ice, regardless of the involvement of neutrinos. Electromagnetic and hadronic showers only develop for several meters, especially for electrons. This means that on the scale of the sensors spacing, they appear point-like and thus leave spherical signatures. The advantage of this is that the signature will usually be fully contained in the detector, and this allows for an energy reconstruction accurate up to around 10% [33]. However, the angular uncertainty is a lot larger than with the muon tracks, since the best uncertainty on electron neutrino interaction directions is 4-5° [34]. Taus, on the other hand, have a median range of about 50 m / 1 PeV, still very small if the tau-neutrino did not have an energy greater than the PeV range.

2.2 Data filtering and event reconstruction

The downside to a continuous data collection with a lot of background noise is that IceCube will register an enormous amount of data, and that it is almost impossible to perform a resource-expensive analysis on all of it. A series of filter and reconstruction algorithms, pictured in Figure 2.5, will be applied successively to the PMT pulses, so here will be discussed the steps to obtain GFU neutrino candidates that will later be used to raise alerts described in Section 4.2.

2.2.1 Level1 (L1) Trigger

Since the event rate is very high, the first levels of the analysis need to be very fast to allow the RealTime framework to be efficient. The first filter applied is called *Level1 Trigger*, in which a Hard Local Coincidence (HLC) check verifies for every pulse if the nearby DOMs on the same string if they registered a pulse less than 1 µs before or after. If not, the hit is classified as noise. The data then goes through a filter called Simple Multiplicity Trigger (SMT-8), which rejects events if there have not been at least 8 HLC hits in a 5 µs window [24]. This is what marks the beginning of the trigger window. At this point, the rate of data acquisition is still very high ($\approx$ 2.3 kHz), and events happening at the same time are treated as one since at this stage, it is unclear whether they are part of the signature of a special event or not.

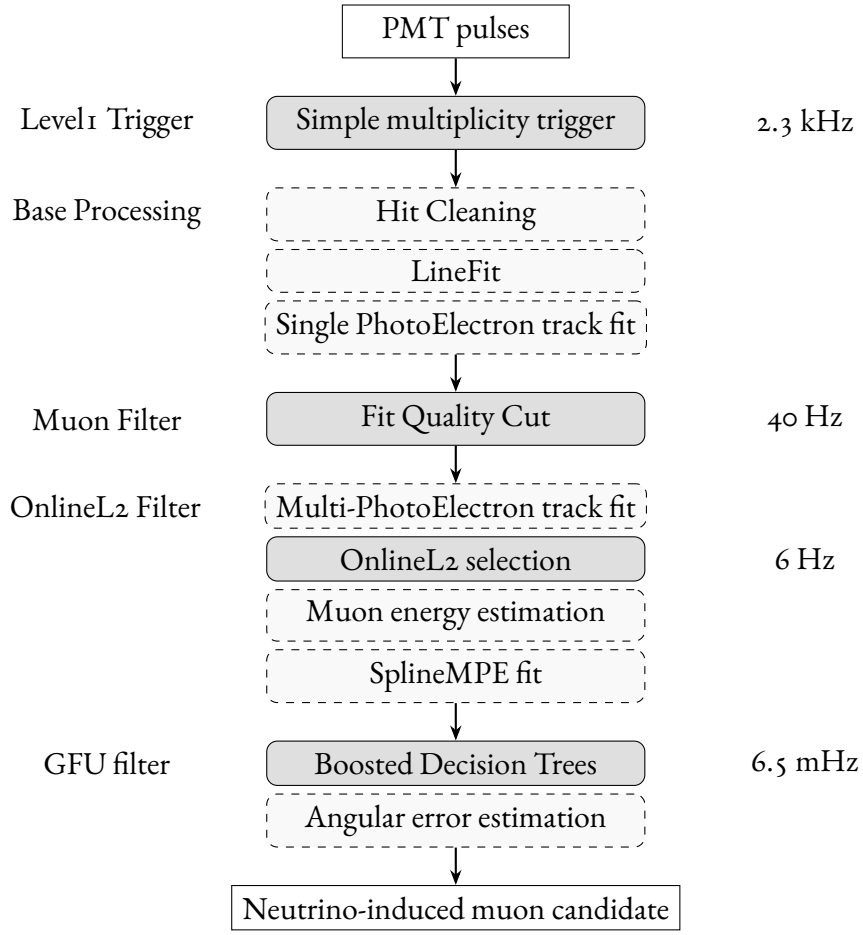


Figure 2.5: Cartoon representing the different filters and reconstruction steps the data passes through. These steps are grouped in modules of which the name are noted on the left. Light dashed boxes are the reconstruction processes, while the darker boxes represent the filters. The frequency on their right is the one for the events after the cuts performed by the filter. The low-level analysis steps (up to OnlineL2) are described in this chapter, while the GFU filter is presented in Section 3.1. Adapted from [32].

2.2.2 Base Processing

The next steps will estimate the direction of origin of the tracks, since they are supposed to come from point sources. The *Base Processing* works to do this as quickly and accurately as possible to allow for more time-consuming algorithms to be performed later, but the rate is very high at this point in the analysis. It starts with a cleaning of the pulses to remove events unlikely to originate from signal. Two reconstruction algorithms follow.

The LineFit is a first and rapid reconstruction method that consists in approximating the de-

tected photons with a plane wave of light passing through the detector. Its concept is straightforward: let $\vec{r}(t)$ be the position of a through-going muon at time t, modeled as:

$$\vec{r}(t) = \vec{r}_0(t_0) + (t - t_0)\vec{v}, \tag{2.3}$$

where $\vec{v}$ is the muon velocity and $\vec{r}_0$ is the point the muon has crossed at t_0. Then the algorithm will simply find the track that minimizes the squared distance with the positions of the DOMs that registered the hits. The LineFit has several advantages: it provides a quick first guess of the direction that can be used as a seed for the other reconstruction algorithms, it yields an estimate for $|\vec{v}|$ that will be used in the final event selection, and it will serve as an indicator of the robustness of the reconstruction [32].

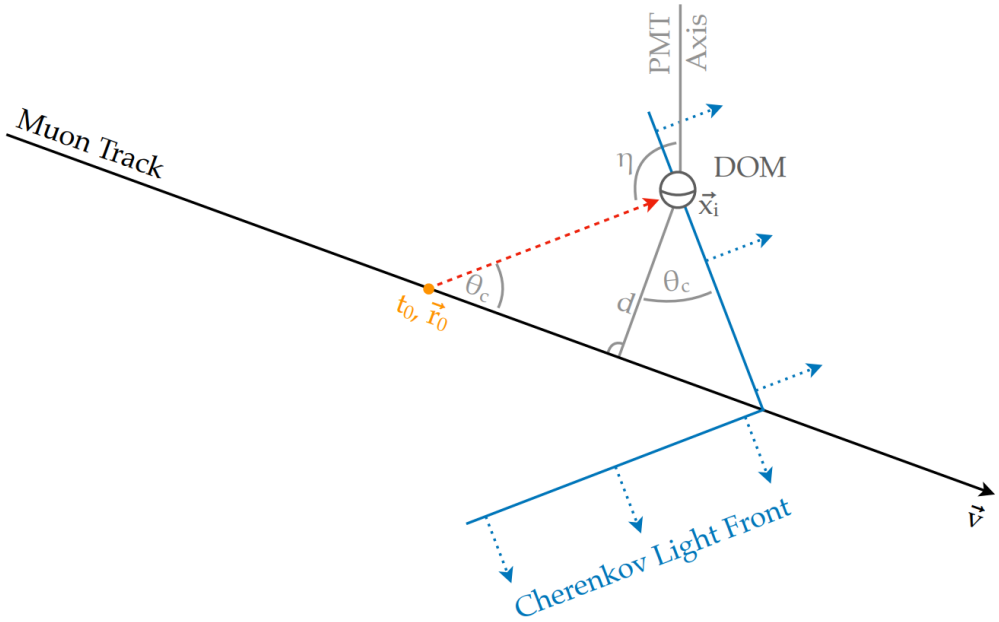


Figure 2.6: Simplified geometry for the emission of Čerenkov light by a muon traveling faster than light in the detector ice. Taken from [32].

However, it is known that the Čerenkov light front is more complex than a simple plane, but is emitted in a cone which aperture depends on the velocity of the muon. Figure 2.6 shows the geometry of the emission from a muon traveling superluminally in ice, and if no scattering happens, the photon emitted from $\vec{r}_0$ at t_0 reaches the DOM at distance d from the track at:

$$t_{geo} = t_0 + \frac{\vec{v}(\vec{r}_0 - \vec{x}_i) + d\tan\theta_C}{c_{vac}}, \tag{2.4}$$

where c_{vac} is the speed of light in a vacuum, and θ_C is the Čerenkov angle as defined in Equation 1.5.

The Single PhotoElectron (SPE) reconstruction method performs a likelihood test of the probability that the first photon reaching the DOM to trigger each pulse was emitted according to this geometry. Only the first photon is checked because it has less chances of having been scattered and thus describes the Čerenkov cone better. This is also where the *Single* in SPE comes from. To do this, it is calculating the time residual, i.e. difference, between the observation time and the one predicted in Equation 2.4:

$$t_{res} = t_{obs} - t_{geo}. \tag{2.5}$$

The probability density function used to defined the likelihood for the test is called the *Pandel* function, and it was first modelled by the Baikal [15] experiment for Čerenkov emission in water [35], then adapted by AMANDA for ice.

The likelihood will be defined as:

$$\mathcal{L} = \prod_{i=1}^{N_{ch}} p_1(t_{res,i}|\vec{r}_0, \theta, \varphi), \tag{2.6}$$

where N_{Ch} is the number of hit DOMs, and p_1 is the Pandel function. Maximizing this likelihood yields an estimate for the muon direction $(\vec{r}_0, \theta, \varphi)$, and makes it possible to judge whether those parameters describe well the distribution of pulses observed, i.e. if the event was well reconstructed and probably a through-going muon track [32].

2.2.3 Muon Filter

With the information from the SPE fits, the *Muon Filter* will eliminate all events unlikely to be muon candidates. It is not possible to perform more accurate, but also more resource-intensive reconstructions before the number of candidates is reduced. Before evaluating the probability of each event actually being a muon track, zenith-dependent conditions need to be applied on the data. For up-going tracks (zenith $\gtrsim 82°$), the SPE likelihood suffices as filter.

Down-going tracks (zenith $\lesssim 82°$), on the other hand, have to deal with atmospheric muons as background. To differentiate them, the deposited energy in the DOMs is a good indicator.

The flux of astrophysical and atmospheric neutrinos are both power laws, but with different expected spectral indices ($\gamma_{astro} \approx -2$ while $\gamma_{atm} \approx -3.7$), so the softer part of the data can be cut to eliminate the noise by requiring that the event carries at least a certain energy. To ensure this, the minimum integrated charge Q_{tot} observed by the DOMs for each event, which is proportional to the deposited energy from the emitting particle, is measured, and a cut is applied on this variable. Here are the cuts applied by the muon filter [32]:

$$\begin{aligned} -1.0 \leq \cos\theta \leq 0.2 &\Rightarrow \frac{\log \mathcal{L}}{N_{Ch} - 3} \leq 8.7 \\ 0.2 \leq \cos\theta \leq 0.5 &\Rightarrow \log Q_{tot} > 3.9 \cos\theta + 0.65 \\ 0.5 \leq \cos\theta \leq 1.0 &\Rightarrow \log Q_{tot} > 0.6 \cos\theta + 2.3, \end{aligned} \tag{2.7}$$

where $\mathcal{L}$ is the SPE likelihood, N_{Ch} is the number of hit DOMs and θ is the zenith. These cuts will reduce the rate of events from 2.3 kHz to 40 Hz, which allows IceCube to transmit daily to Madison for offline archival analysis, which was previously impractical because of the limited bandwidth due to low satellite coverage at the South Pole. More time-consuming, real time analyses are still being performed on-site.

2.2.4 Online L2 Filter

The *OnlineL2 Filter* aims at ameliorating even further the direction estimate and provide the first estimation for the energy of the muon. The Multi-PhotoElectron (MPE) Track Fit, as its name suggests, works in a similar fashion to the the SPE, except that it takes into account that, above 1 TeV, the probability of observing more than one photon per DOM becomes higher, so the likelihood used in the maximization algorithm needs to be adapted as such:

$$\mathcal{L} = \prod_{i=1}^{N_{Ch}} \left[N_i . p_1(t_{res,i}) . \left(\int_0^{\infty} p_1(t_{res,i}) \right)^{N_i - 1} \right], \tag{2.8}$$

where t_{res} is defined in Equation 2.5, p_1 is the Pandel function, and N_i is the number of photons observed in the i^{th} DOM at $t_{res,i}$. The result from the SPE is the seed used by the algorithm, and it allows for a better reconstruction, which then leads to a new filter.

The OnlineL2 selection is a series of tight cuts described in [36], where the sky is divided following declination in four regions which have different cuts. This filter brings the event rate

down to 6 Hz, which allows the application of the last reconstruction algorithm: the Spline Multi PhotoElectron (SplineMPE) Fit. What differs with the MPE is that instead of using the Pandel function for the probability density, the SplineMPE relies on a simulation of the propagation of photons in ice. The ice in which the photons travel is no longer treated as homogeneous, and this model is sensitive to depth when calculating absorption and scattering probabilities for the photons. The fit is called SplineMPE because the results of the simulation are given by interpolating splines (Segmented Polynomial Line, a function defined piecewise by polynomials but continuous up to a certain order derivative), like a lot of complex fits [37].

The last step is to compute the energy estimation of the muon. Through-going muons are created outside the detector and lose energy before they could be detected. The energy estimation is then a lower bound for that of the neutrino. The number of Čerenkov photons emitted is proportional to the energy of the muon [38], and it is modeled as a Poisson distribution. The likelihood is then:

$$\mathcal{L} = \frac{\lambda^k}{k!} e^{-k}, \tag{2.9}$$

with

$$\lambda = \Lambda.E + \rho \tag{2.10}$$

where E is the energy, k the number of observed photons, and ρ is the noise contribution to the mean expected number of photons λ [39]. This λ is dependent on the template model Λ, which depends on the positions of the track and the involved DOMs in the detector, but also their relative positions.

The main difference between the two ways of retrieving the energy estimation of the muon lies in the definition of Λ and especially the treatment of the stochastic energy losses of the muons. The only one relevant to this project is called *Muon Energy Estimation* (MuEX), in which the muon is assumed to emit light uniformly along its track, meaning the expected number of photons depends on the distance between the track and the DOMs. The template Λ contains two approximations: the scattering is neglected for short distances (photon density $\propto \frac{1}{r}$), and diffusive behavior approximated by expression found in [39] at large distances. The MuEX method provides an estimation of the energy by maximizing the Poisson likelihood using these conditions [40].

The other way to approximate the energy of the muon is called TruncatedEnergy. It will

split the track and apply Equation 2.9 to each segment separately to modelise the stochastic energy losses. The length of the segments is defined such that each segment only includes a single DOM. A cylindrical cut will also make sure not to include DOMs closer than 10 m to the track (bright DOM behavior is still not well understood) or farther than 80 m. The expectation in Equation 2.9 comes from the ice model used by the SplineMPE directional reconstruction. Half the segments with the highest energy gradient dE/dx are removed and the average of the rest of them is calibrated to give a solid estimate of the muon energy regarding the variation in stochastic losses [41].

3

The RealTime alert system

The previous chapter described the general characteristics of the IceCube Neutrino Observatory, as well as the basis of the processing and filtering of the enormous amount of data registered by IceCube. These steps are all performed in real-time at the South Pole and are common to a lot of different studies in IceCube, both on and offline. This chapter will present the next steps in the treatment of the data to raise alerts in real-time specifically in the Gamma-ray Follow-Up (GFU) branch of IceCube. It starts by describing the so-called GFU Filter in Section 3.1, which was originally designed for real-time use only, but is now also used in offline analyses. Section 3.2 will then introduce how the time-clustering algorithm raises alerts, and the modes in which GFU operates.

3.1 GFU Filter

To classify the events as neutrino-induced muon candidate to be used by GFU, this filter uses Boosted Decision Trees (BDTs) to apply a series of cuts to the data to get the rate from 6 Hz to 6.5 mHz ($\approx 200000 yr^{-1}$). BDTs are machine learning classifiers that control several variables at once and trains over several iterations to sort the data most efficiently.

In the Northern Sky, the background the BDTs are trying to eliminate from the sample is mostly constituted of mis-reconstructed cascades or down-going muons that were misclassified earlier. The BDTs will then verify that the SplineMPE fit is close to the LineFit, and that the

geometry of the track is coherent. A cut is applied on the SplineMPE likelihood, and coincident events (events less than 10 μs apart but not part of the same track) are eliminated.

In the Southern Sky, muons from atmospheric showers make up most of the background, since it is not shielded by the Earth like the Northern Sky, and atmospheric muons look like neutrino-induced muon tracks. The BDTs for down-going events identify muon bundles, sets of muons produced by a heavier nucleus producing an air shower, and remove them from the data.

The final stage is the estimation of the angular error. The first of the three methods to do this is the Cramér-Rao estimation, based on the Cramér-Rao inequality [42]. It sets a lower bound on the variance of a statistical estimator, given by the inverse of the Fisher information matrix. This matrix measures how much information the track carries about the time residual from Equation 2.5.

With $\vec{x} = (\vec{r}_0, \theta, \varphi)$ as the parameters of the track, the Cramér-Rao inequality is:

$$Cov(\vec{x}_k, \vec{x}_l) \geq \mathcal{I}_{kl}^{-1}(\vec{x}) \tag{3.1}$$

The diagonal elements of Fisher information matrix are by construction the inverse of the squared angular uncertainties along the zenith and the azimuth. This means that from the calculation of σ_θ and σ_φ, a circularized angular error can be computed [43]. This method is purely analytical [44], and fast ($\approx$ 0.01 s per event) since it does not require a minimizer, but it is also the least precise, especially at low energy.

The Paraboloid method is slower, but more precise, it approximates the shape of the negative logarithm of the likelihood around the minimum with a paraboloid (a two-dimensional Gaussian). 24 points around the minimum are evaluated then fitted to a paraboloid, from which this is known, since it is shaped like a Gaussian:

$$-\log \mathcal{L}(x \pm \sigma_x) \approx -\log \mathcal{L}(x) + 0.5, \tag{3.2}$$

Lastly, the Bootstrapping method approximates the unknown theoretical distribution of a random variable by repeatedly sampling with replacement the set of observations. Pulses

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

Table 3.1: Method of preference to calculate angular uncertainty depending on the energy and the number of hit DOMs

from the DOMs are taken at random to constitute an with the same total charge as the original detection. This is done several times to construct a larger sample in which every event is reconstructed. Then the uncertainty on the direction of the track from the original event is approximated by the median of those of the bootstrapped events.

The approach to use is chosen on a case-by-case basis, according to the properties of the event. The efficiency of the algorithms depends greatly on the energy of the event and the number of hit DOMs. As summarised in Table 3.1, low energy events are better reconstructed using the Paraboloid method, and for high energy events, the bootstrapping event is more appropriate for tracks which involve less than 300 DOMs, while the Cramér-Rao method provides the best and fastest estimation for the energetic and well-reconstructed events.

RealTime analysis demands fast and accurate performances. All the techniques, filters, and methods described above have been refined for the entire process to perform in around 1 s. The events are written and saved in JavaScript Object Notation (JSON) format, saving the direction, energy and angular uncertainty.

3.1.1 Gamma-ray Follow-up samples

The data has been reconstructed and hopefully filtered down to neutrino-induced muon tracks with a rate of 6.5 mHz. These GFU samples are sent to Madison, and are ready to be analyzed in the GFU framework, among other studies. However, it is improbable that all the events would be of astrophysical origin. Figure 3.1 confirms this assumption by showing the different data type proportions depending on the declination. In the Northern Sky, there is no contribution from atmospheric muons, sign that the GFU filter managed to reject down-going tracks. Contributions are, in order of importance: atmospheric neutrinos background, astrophysical muon neutrinos, astrophysical tau neutrinos, astrophysical electron neutrinos, and atmospheric electron neutrinos. In the Southern Sky, the data is dominated by atmospheric muons.

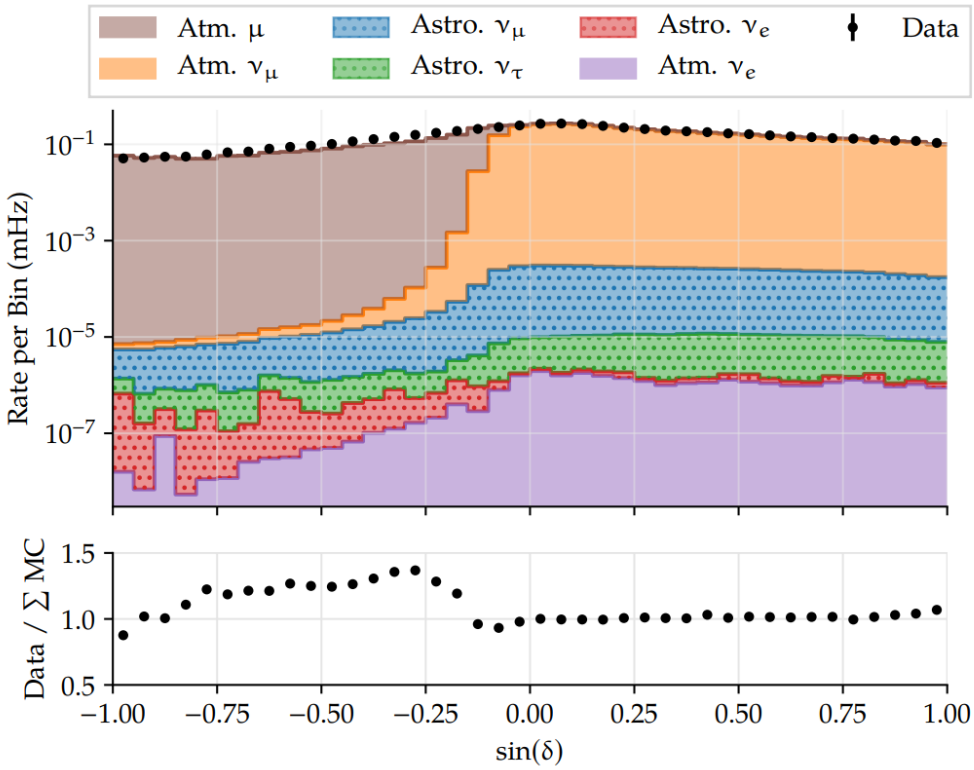


Figure 3.1: Event type proportions in the GFU data sample after the data has been filtered and reconstructed. The solid areas show the contributions of the atmospheric flux, while the dotted areas are representative of the diffuse astrophysical flux. The plot below depicts the ratio between the amount of data collected by IceCube and the simulation samples. These results are all plotted in function of the declination. Taken from [32].

The GFU sample is widely used for many analyses in the search of neutrino sources, not only for multi-messenger purposes, although this is what it was originally for. A final smoothing cut is applied to maximize discovery potential of point sources.

3.2 GFU real-time cluster alerts

IceCube is able to monitor the whole sky almost continuously. This quality makes it most suited for real time studies. The only times where IceCube is not collecting data is during software upgrades, power outages, or DOM calibration, and even when a DOM fails or a string malfunctions, the data collection does not stop. This brings IceCube's uptime to 99.8% on a typical week. There are two main approaches to alerts in the RealTime framework: either

focusing on identifying single high-energy neutrino events with a high chance of being of astrophysical origin, or noticing clusters of events that would be the result of a neutrino flare.

Even if the first method led to the coincident detection of the high energy neutrino 170922A event and the flaring blazar TXS0506+056 (still regarded as one of the most important results of multi-messenger astrophysics [45]), this work focuses on alerts produced by the Gamma-ray Follow-Up (GFU) cluster alert system, looking for neutrino flares. GFU monitors a list of known variable gamma-ray sources that are likely to also be emitters of high-energy neutrinos, since both are likely to happen at the same time. Then, when an alert is raised, IceCube forwards it privately only to its partner Imaging Atmospheric Cherenkov Telescopes (IACTs): MAGIC [4] and VERITAS [6] since 2012 [46], H.E.S.S. [5], since 2019, and LST-1 [47] since 2023 (the collaboration started in 2005 between AMANDA and MAGIC [48]) to encourage follow-up observations. GFU also performs a search over the whole sky, without the bias of looking at a predefined source list [49]. For now, these all-sky alerts are only shared publicly as Gamma-ray Coordinate Network (GCN) notices, but this work aims to show how the alerts could be shared directly to partner telescopes to ask for follow-up observations.

With the sample of muon neutrino candidates, distinguishing astrophysical sources becomes possible. The analysis is still being performed in real time, so the objective remains to implement fast and efficient ways to raise an alert when a source, whether it was already monitored (Subsection 3.2.3) or it is part of an unbiased search (Subsection 3.2.4), is flaring so that partner telescopes can be alerted and coordinate observations. This section will first introduce the probability density functions (PDFs) and test statistics (Subsection 3.2.1) relevant to the Time-Clustering algorithm (Subsection 3.2.2) which is essential to describe when an alert is triggered.

3.2.1 Point Source Likelihood Method

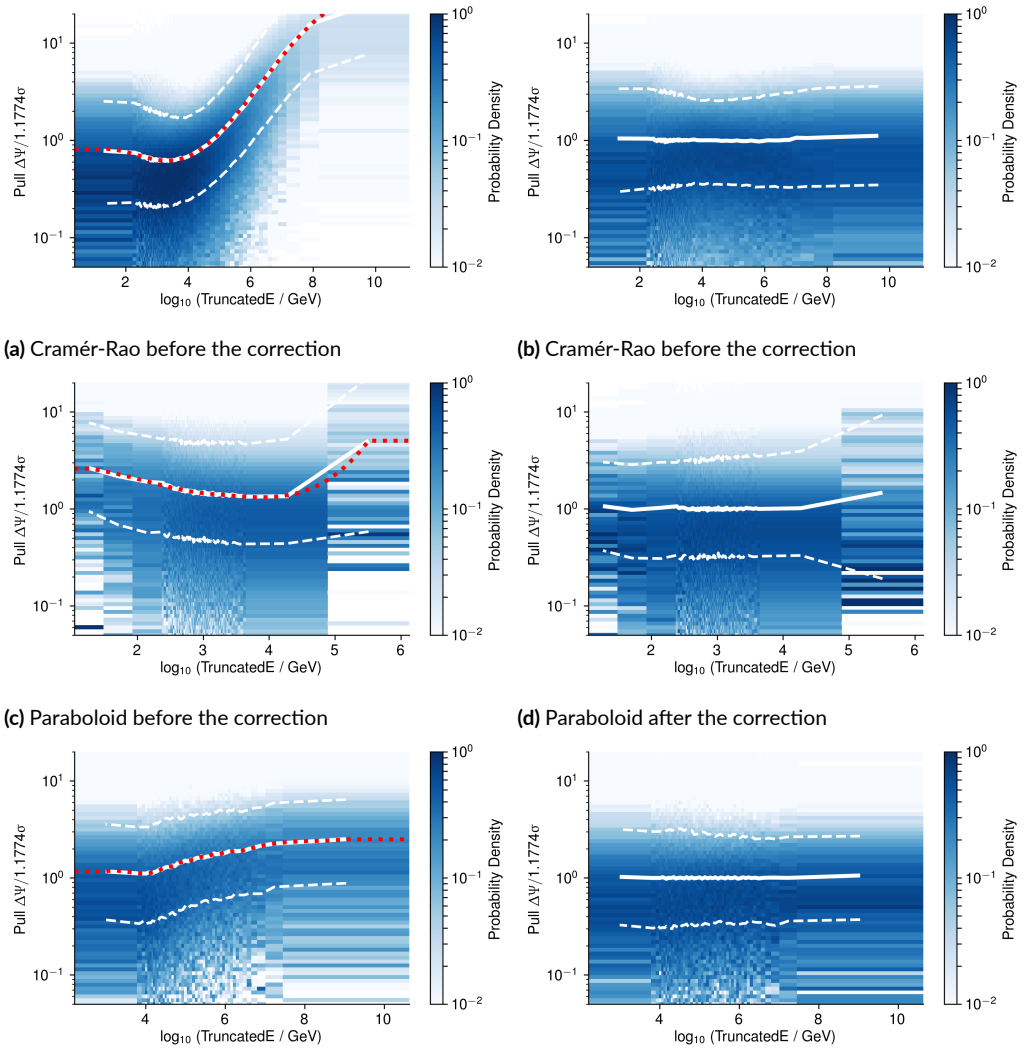
Pull correction on individual event angular error The angular uncertainty from the three angular error estimators (Cramér-Rao, Paraboloid, and Bootstrapping) applied in the GFU filter of Section 3.1 is being overestimated, depending on the energy proxy of the muon. Generally, the angular uncertainty grows with energy, as seen in Figure 3.2. This pull makes the event reconstruction less precise, so a correction needs to be applied to try and maintain the angular error as low as possible. This is accomplished in three steps [50]:

1. Each estimator's pull distribution is binned in TruncatedEnergy in such a way that the

bins contain approximately the same number of events.

2. The deviation of the median from the expectation of 1.1774σ is measured for each bin. This value is based on the assumption that events are accurately reconstructed by a Rayleigh distribution [51] of parameter σ.

3. The pull is corrected by a cubic spline for every estimator.



(a) Cramér-Rao before the correction

(b) Cramér-Rao before the correction

(c) Paraboloid before the correction

(d) Paraboloid after the correction

(e) Bootstrapping before the correction

(f) Bootstrapping after the correction

Figure 3.2: Distributions for each angular error estimator before (left) and after the pull correction (right). The solid white lines show the median, while the white dashed lines are the 10-90% quantiles. The red dashed line is an extrapolation of the median on the whole sample. Taken from [50].

SPATIAL PROBABILITY DENSITY FUNCTIONS The first PDFs examined are the background and signal spatial PDFs, i.e. the probability of an event being either background or signal emitted by a source. Building these PDFs relies on the hypothesis that while atmospheric background events are distributed almost uniformly across the sky, the signal events cluster around the point-like source they originate from. In practice however, the background event rate varies with zenith: it depends on the filtering efficiency, the volume of ice and Earth the neutrino has to travel through, etc. There is also an azimuthal dependence due to the symmetric detector layout. The strings form a hexagonal grid, hence the detector has six preferred directions where more events are observed. This is especially relevant for short timescales. This is clear in Figure 3.3, where it is shown how the background spatial PDF created from archival data deviates from a PDF depending only on declination. The axes of symmetry are visible immediately, and can modify the acceptance of signal coming from a fixed source by up to 40% [32]

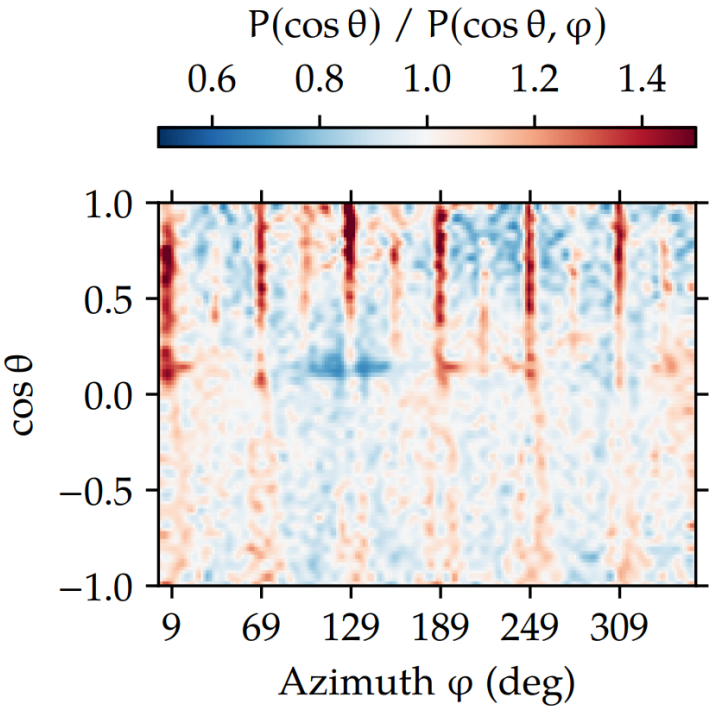


Figure 3.3: Ratio between the spatial background PDF if it depended only on declination and the PDF as it is, i.e. dependent on both declination and azimuth. This highlights the axes of symmetry of the detector as deviators from a one-dimensional PDF. Taken from [32]

The signal spatial PDF, on the other hand, relies on the assumption of a source location, and uses the pull-corrected angular uncertainty estimation to evaluate the probability of an event

having been caused by said hypothetical source:

$$\mathcal{P}^{S}_{spatial}(\vec{x}_i|\vec{x}_s,\sigma_i) = \frac{1}{2\pi\sigma_i^2}\exp\left(-\frac{\Delta\Psi_i(\vec{x}_i|\vec{x}_s)^2}{2\sigma_i^2}\right), \tag{3.3}$$

with $\Delta\Psi_i$ being the angular distance between the reconstructed direction $\vec{x}_i$ and the source location $\vec{x}_s$, and σ_i the estimated angular uncertainty. Events with a better angular uncertainty are allowed more weight in the PDF so that better localized events contribute more to the estimation of the location of the signal. The point spread function is assumed to be a two-dimensional circular Gaussian because a detailed contour in RealTime analyses would be too computationally demanding.

Energy probability density functions The energy distribution of neutrinos from astrophysical sources is expected to follow a power-law like E^{-2} [52], which is close to the observed diffuse astrophysical neutrino flux ($E^{-2.19}$). Atmospheric neutrinos exhibit a much softer spectrum, which follow a power-law of $E^{-3.7}$. So to distinguish the atmospheric from the astrophysical neutrinos, the energy proxy can be of use. The MuEX method is favored over the TruncatedEnergy, which is unusable in low-energy cases.

In simulations, this energy proxy is used to create the signal energy PDF by reweighing the simulation for spectral indices in a certain range. The chosen γ go from -1, harder than the observed diffuse astrophysical flux, to -4, which is softer than the atmospheric neutrino background. Since the signal and background are declination-dependent, the energy PDF is being built in declination bands, while the background PDF is once again created from archival data, allowing a uniform although conservative treatment of the sky.

Point source test statistic Separate PDFs have been defined for the spatial distribution and the energy because there is no correlation between the quality of the reconstruction and the estimation of the energy [53]. Below are defined a signal ($\mathcal{S}$) and background ($\mathcal{B}$) probability for each event, expressed in terms of the PDFs described above:

$$\mathcal{S}_i = \mathcal{P}^{S}_{spatial}(\vec{x}_i|\vec{x}_s,\sigma_i).\mathcal{P}^{S}_{energy}(E_i|\vec{x}_i,\gamma) \tag{3.4}$$

$$\text{and } \mathcal{B}_i = \mathcal{P}^{B}_{spatial}(\vec{x}_i) \quad .\mathcal{P}^{B}_{energy}(E_i|\vec{x}_i), \tag{3.5}$$

where E_i is the energy proxy of the event, and γ is the spectral index of the signal energy power-law. This means that for a sample of N events among which n_s are signal-like, the likelihood will be [54]:

$$\mathcal{L}(n_s, \gamma) = \prod_{i=1}^{N} \left(\frac{n_s}{N} . \mathcal{S}(\vec{\theta}_i | \gamma) + \left(1 - \frac{n_s}{N}\right) . \mathcal{B}(\vec{\theta}_i) \right), \tag{3.6}$$

where $\vec{\theta}_i$ represents all the properties of the i^{th} event (direction, angular uncertainty, energy). Here γ and n_s are free parameters that will be determined by maximizing the likelihood. There are still bounds that are imposed on these parameters: the spectral index must be between -1 and -4, like for the construction of the energy PDF, and the number of signal-like events must be between zero and the total number of events. Once the likelihood is maximized, the test statistic is defined by comparing it to the background-only hypothesis ($n_s = 0, \gamma$ irrelevant):

$$\Lambda = 2 \log \frac{\mathcal{L}(\hat{n}, \hat{\gamma})}{\mathcal{L}(0, -)}. \tag{3.7}$$

It is actually more computationally efficient to directly maximize the TS since the logarithm is monotone and the background expectation is constant. Inserting Equation 3.6 in Equation 3.7:

$$\Lambda = 2 \sum_{i=1}^{N} \log \left[1 + \frac{n_s}{N} (w_i - 1) \right], \tag{3.8}$$

where w_i is defined as the event weight, i.e. the ratio of signal and background PDFs $\mathcal{S}_i / \mathcal{B}_i$. To get rid of the events that are too far from the source hypothesis with respect to the spatial resolution, for the $N - N'$ events outside of a circular bin of 5° defined around the source, the approximation $\mathcal{S}_i \approx 0$ is made. The TS then becomes:

$$\Lambda = 2 \sum_{i=1}^{N'} \log \left[1 + \frac{n_s}{N} (w_i - 1) \right] + 2(N - N') \log \left(1 - \frac{n_s}{N}\right). \tag{3.9}$$

3.2.2 TIME-CLUSTERING ALGORITHM

The whole likelihood construction process described above is time-independent. However, neutrino emitters can be very time-variable, and observing a flare of neutrinos, i.e. a burst happening on a relatively short time scale, deals with much less background than an archival time-integrated search over several years. The goal of the algorithm presented here is to include a time component on the distributions that also allows to evaluate excesses in time, and find the

optimal time-window to represent these excesses.

An important quality of the method to find this best time window is the assumption of the time PDF: how the events are distributed in time. The most generic choices are either a box defined by a start and end time, or a Gaussian shape defined by a central time and a width. The time-clustering algorithm uses a box-profile assumption, which works well in RealTime analyses, aligning the end of the box with the end of the studied time-window.

An event with $\mathcal{S}/\mathcal{B} \geq 1$ happening at t_k, defined as the end of the box, triggers the analysis, in which the algorithm tries every other event at $t_i < t_k$ with $\mathcal{S}/\mathcal{B} \geq 1$ as possible starting times for the flare. For each possible pair, the likelihood previously presented is calculated with all the events between t_i and t_k. Figure 3.4 shows a cartoon of how the time-clustering algorithm works, where t_k is assigned the value 0 and possible clusters are being evaluated. The maximum duration accepted is fixed to a T_{max} chosen beforehand. In this work, the lookback ceiling was set at 180 days, which is the standard T_{max} for GFU cluster searches.

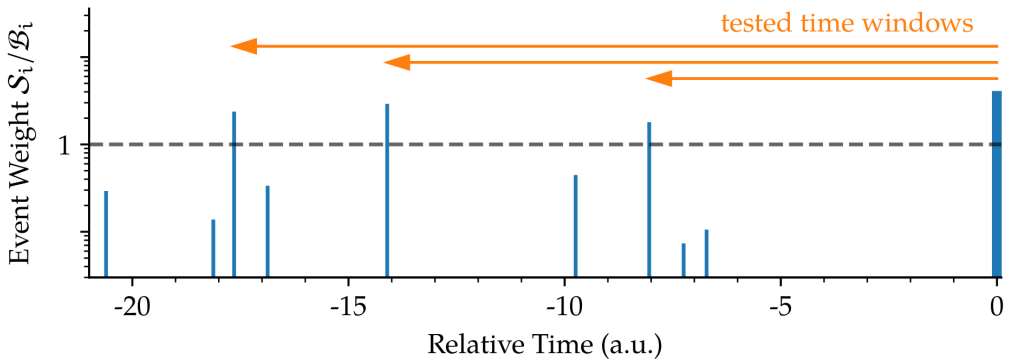


Figure 3.4: Functioning of the time-clustering algorithm. In this sketch, the vertical bars represent the weight of the event, i.e. the ratio of its signalness probability over the probability it belongs to the background. When an event triggers the algorithm by having a weight above the set threshold (1), the search for the most significant time-window is shown by the orange arrows. The algorithm only checks time-windows that start with an event above threshold and end with the triggering event. Taken from [32].

Ideally, the algorithm would be unbiased towards the possible duration of the flare, but in reality, as there is less background evaluated during shorter flares, significant combinations of events over short time windows are more numerous so the algorithm will favor them. To rem-

edy this, a penalty is introduced to the test statistic defined in Equation 3.7:

$$\Lambda = 2\log\left(\frac{\mathcal{L}(\hat{n},\hat{\gamma})}{\mathcal{L}(0,-)}\frac{T_{cluster}}{T_{max}}\right), \tag{3.10}$$

where $T_{cluster}$ is the detector uptime in the tested time window $(t_k - t_i)$. This corresponds to multiplying $\mathcal{S}$ and $\mathcal{B}$ by a time PDF such that:

$$\mathcal{P}^{S}_{time}(t) = \mathcal{P}^{B}_{time}(t) = \begin{cases} T^{-1}_{cluster} & t_i < t < t_k \\ 0 & \text{else} \end{cases}. \tag{3.11}$$

3.2.3 SEARCHES IN MONITORED SOURCES

The time-clustering algorithm serves in real time studies of GFU in two different contexts. The goals of these analyses is to identify excesses from known gamma-ray sources and warn partner telescopes in order to achieve follow-up multi-messenger observations. The alert infrastructure was originally implemented in AMANDA [48], and it has been updated several times since, but the core concept is: when a neutrino flare is identified by IceCube, it is transmitted to the community with informations about the signicance of the flare, the events it includes, the time-window and fit parameters, as well as the position of the source.

The principle of the alert production is represented in Figure 3.5, where the significance of the cluster fit for each triggering event is plotted over time. Each dot represents an event with $\mathcal{S}/\mathcal{B} > 1$ which has then triggered the time-clustering algorithm and the result of the time-clustering algorithm is showed. When this significance goes above a certain threshold, an alert is sent to the RealTime Oversight Committee (ROC) of IceCube and to partner IACTs, and then the source is muted until it goes under threshold again. On average, they receive this alert 12 minutes after the triggering event has been detected by IceCube [46]. For now, the *best fit*, as noted in Figure 3.5 is only noticed and analyzed during offline archival studies made to verify if a source has overcome the discovery threshold.

The criteria for the current list are based on visibility by the partner IACTs, variability of the high energy emissions, and energy and significance of the flux. In this mode, the time-clustering algorithm sends an alert when a trigger goes over a significance of 3σ, which yields an expected alert rates of O(10)/year.

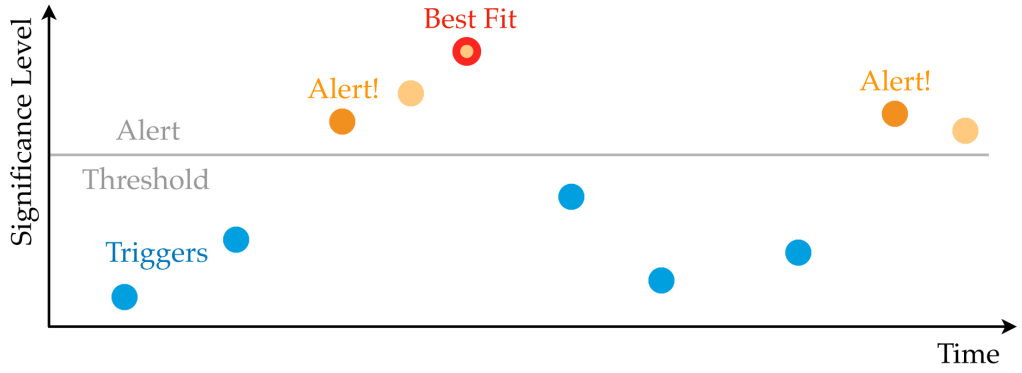


Figure 3.5: Cartoon of the alert mechanism in real time analyses. Every blue dot is a triggering event for the time-clustering algorithm, but only when the pre-trial significance of the result from the algorithm exceeds the set threshold is an alert raised (dark orange dots). The source is then muted until the significance of the time-windows constituted for subsequent triggers goes back below threshold to avoid spamming alerts. The global best fit (red dot) is then generally only studied offline during archival analyses. Taken from [32].

3.2.4 All-sky scan

This work, however, focuses on the all-sky alerts, which treat every event with $\mathcal{S}/\mathcal{B} > 1$ as a potential trigger since it treats the whole sky without biases on any source location. This increases significantly the number of triggers, and the trial factor, so the pre-trial significance threshold is currently fixed at 4.2σ for an alert rate of $1yr^{-1}$, as shown in Figure 3.6. This is higher than for the monitoring of known sources, because testing every possible direction makes it more probable to observe a random, but significant background fluctuation by chance. Also, the correction to obtain the post-trial p-value has to take the size of a cluster compared to the whole sky. An optimistic order for the area covered by a neutrino cluster would be $\approx 0.8\,\mathrm{deg}^2$, while the celestial sphere is $\approx 41253\,\mathrm{deg}^2$. This means the pre-trial p-value is to be multiplied by a trial factor that reflects the ≈ 51566 independent hypotheses for a source in the sky. Because of this, the post-trial significance level obtained for a pre-trial significance level of 5σ is only around 2.2σ, and a discovery level of 5σ demands a pre-trial significance of 6.8σ.

Since the all-sky analysis works in an unbiased manner, the time-clustering algorithm has to be triggered without an assumed source location to feed Equation 3.3. The software named Hierarchical Equal Area isoLatitude Pixelization (HEALPix) [55] and its Python wrapper `healpy`, which divides the sky in equal area bins, are used on the map, and then the time-clustering algorithm is applied to the center of each pixel. The software begins with a 0.92° grid, and iterates

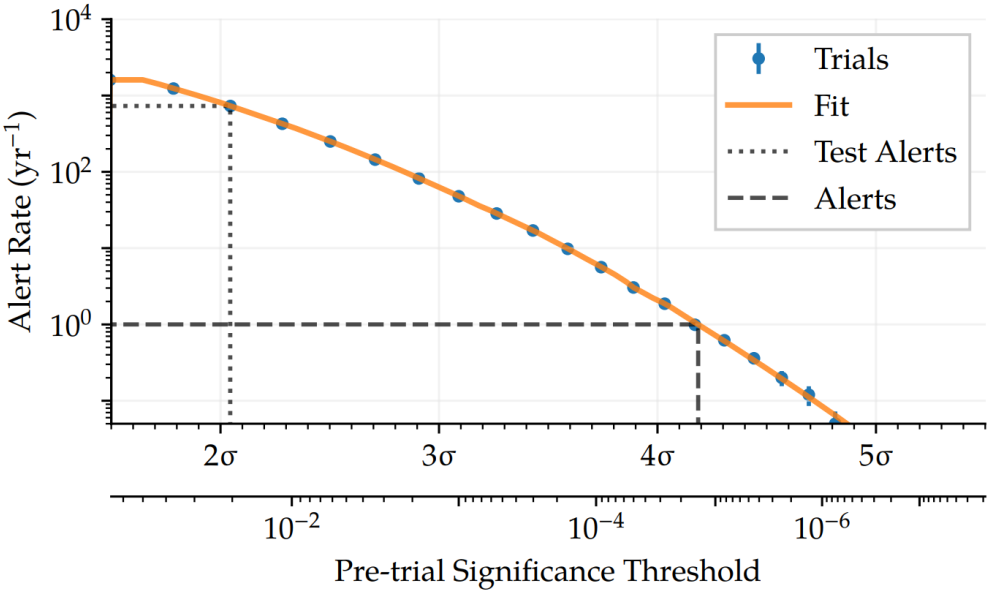


Figure 3.6: Toy simulation results to indicate the all-sky expected alert rate depending on the pre-trial significance. It is shown how a threshold of 4.2σ yields an expected alert rate of $1yr^{-1}$. Taken from [32].

with smaller and smaller grid size by refining the most significant pixel definition. Once the grid spacing has reached $\approx 0.1°$, i.e. the average angular uncertainty of the events and the definition beyond which there is no significant improvement to the fit, the best fit pixel, with the highest TS value is identified. The HEALPix method to assign a source location hypothesis to a triggering event is shown in Figure 3.7.

Until now, GFU all-sky cluster alerts have not been automatically forwarded to partner telescopes because there was no uncertainty regions around the best fit source. This work presents the inception of a procedure to generate such parameters that could eventually be forwarded to the community in the future.

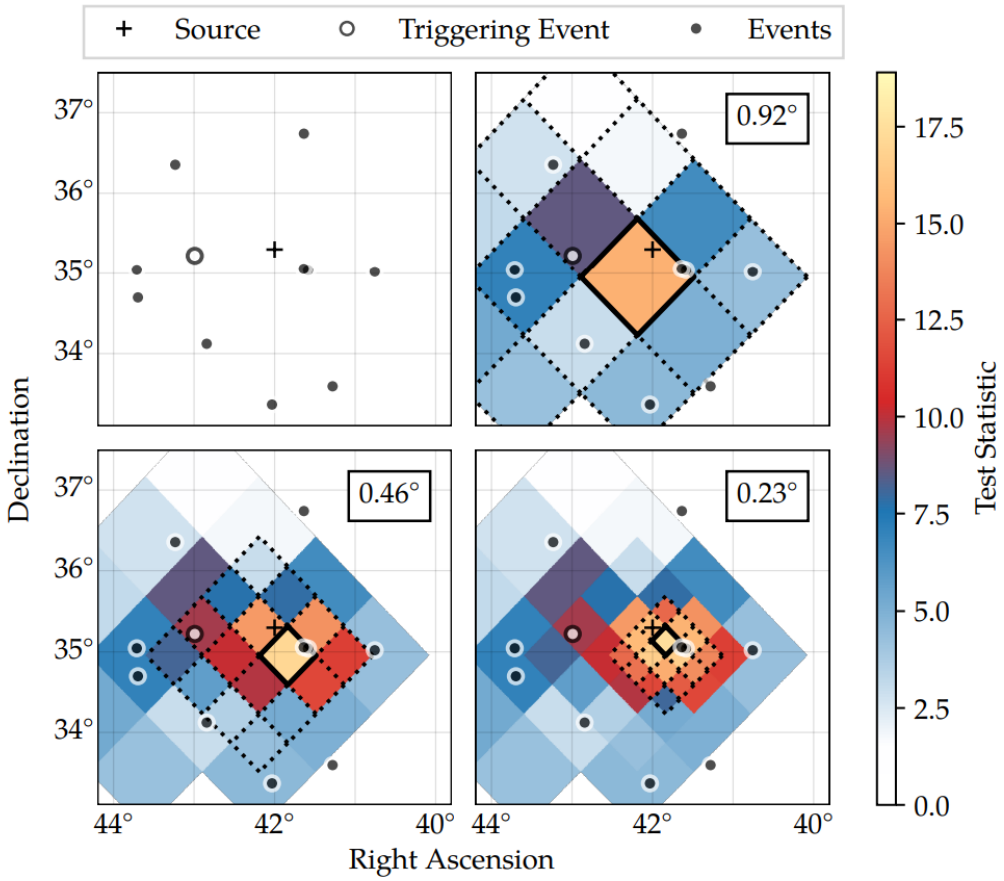


Figure 3.7: HEALPix iterative approach to assign a source hypothesis to a cluster. The dotted lines are the pixel scanned in each iteration for this simulated situation while the solid line denotes which pixel exhibits the best test statistic. The grid spacing is shown in the top right corner. Taken from [32].

4

Confidence regions in the Test statistic difference maps of alerts

Every event that makes it through the data selection processes is reconstructed with a degree of uncertainty on its energy and its direction of origin. Section 2.2 detailed how IceCube filters the background out of the data, and Section 3.2 how cluster alerts were being raised. Now, this chapter will discuss how a likelihood map of the position of its source can be constructed, and especially how regions can be parametrized in these maps to guide partner telescopes' observations. First, we will need to introduce the statistics notions important for the project, most notably Wilks' theorem (Section 4.1), before presenting the Test Statistic maps in Section 4.2 and how the simulations from this project (Section 4.3) will attempt to verify if Wilks' theorem describes accurately the situation for the cluster alerts.

4.1 Introduction to the statistics of Wilks' theorem

Formulated in 1937 by S.S. Wilks [56], this theorem gives a distribution for likelihood ratios in large samples under specific regularity conditions. In fact, it states, in its simplest form:

$$-2\log\lambda = \chi^2_{(b)}, \tag{4.1}$$

where $\lambda = \frac{\mathcal{L}_H(x_j)}{\mathcal{L}_{H_0}(x_j)}$ with $j = 1...n$ and $\chi^2_{(h)}$ is the chi-square distribution defined from h random variables as:

$$X_i \xrightarrow{d} N(0,1), \forall i = 1...h \Rightarrow \sum_{i=1}^{h} X_i^2 \xrightarrow{d} \chi^2_{(h)}. \tag{4.2}$$

This assumes that the likelihoods $\mathcal{L}_{H_0}$ and $\mathcal{L}_H$ are corrected with a factor $(1 + \varphi)$ with φ of order $1/\sqrt{n}$. Therefore, Equation 4.1 holds except for terms of order $1/\sqrt{n}$, which is why the large sample condition is fundamental in the application of Wilks' theorem. This section describes first the necessary regularity conditions, then derives Equation 4.1.

4.1.1 The Maximum Likelihood Estimates

A likelihood $\mathcal{L}(x_j; \theta_i)$ measures the probability of the combination of outcomes of n random variables x, which distributions are parametrised by h parameters noted θ. It is built by multiplying all the probability density functions of the variables x_j, while the θ_i are free parameters: their value is left free in the definition of the likelihood and it will be fixed generally when the value of the likelihood is being maximized. First of all, Wilks' theorem works on the assumption that in the h-dimensional space of the θ_i, there exists a set of $\tilde{\theta}_i$ which are the maximum likelihood estimates (MLEs) of the parameters. As their name suggests, this means that these estimated values for the parameters θ_i maximize the likelihood, and it is not trivial, as it requires the regularity conditions to be respected. These conditions are necessary to the validity of the Maximum Likelihood Estimation theorem upon which Wilks built his derivation because they impose conditions on the shape of the likelihood distribution, and are as follows [57]:

$(i = 1...h, j = 1...n)$

1. The variables X_j are independent and identically distributed with density $f(x_j; \theta_i)$.
2. The parameter space Θ is compact.
3. $\mathbb{E} \log f(X_j; \theta_i)$ exists, with $\mathbb{E}$ the mathematical expectation.
4. The true but unknown parameters values $\theta_{0i} \in Int(\Theta)$ are identified:

$$\theta_{0i} = arg \max_{\theta_i \in \Theta} \mathbb{E} \log f(X_j; \theta_i) \tag{4.3}$$

5. The log-likelihood function

$$\ell(x_j;\theta_i) = \sum_{j=1}^{n} \log f(x_j;\theta_i) \tag{4.4}$$

is continuous throughout Θ, and twice continuously differentiable in a neighborhood of θ_{0i}.

6. It is also such that $\frac{1}{n}\ell(x_j;\theta_i)$ converges almost surely in probability to $\mathbb{E}\log f(X_j;\theta_i)$ uniformly in $\theta_i \in \Theta$.

7. The information matrix

$$\mathcal{I}_{kl}(\theta_{0i}) = -\mathbb{E}\left(\frac{\partial^2 \log f(X_j;\theta_{0i})}{\partial\theta_k \partial\theta_l}\right) \tag{4.5}$$

exists, and is positive definite.

Under these conditions, the MLEs exist and are well defined. They also almost surely converge in probability to the true parameter values θ_{0i}: the events in which the MLEs *do not* converge to the true parameter values have a zero probability. This means that the estimates used to model the distribution of the parameters of the likelihood tend towards the unknown true values that maximize the likelihood. Furthermore, it is possible to prove that, given a distribution of MLEs $\tilde{\theta}_i$, its difference with that of the true parameter values, multiplied by the square root of the size of the data sample, behaves asymptotically like a normal distribution of mean zero and of variance equal to the inverse of the information matrix evaluated at θ_{0i} [57]:

$$\sqrt{n}\left(\tilde{\theta}_i - \theta_{0i}\right) \xrightarrow{d} N\left(0, \mathcal{I}(\theta_{0i})^{-1}\right) \tag{4.6}$$

Cramér-Rao inequality in Equation 3.1 tells us this distribution has the best variance possible, when n tends to infinity. However, it is impossible to evaluate the information matrix at θ_{0i} because the true values of the θ_i are unknown. Instead, the MLE asymptotic behavior allows an approximation by evaluating the information matrix at $\tilde{\theta}_i$:

$$\sqrt{n}\left(\tilde{\theta}_i - \theta_{0i}\right) \xrightarrow{d} N\left(0, \mathcal{I}(\tilde{\theta}_i)^{-1}\right) \tag{4.7}$$

This can be rewritten as:

$$\tilde{\theta}_i \xrightarrow{d} N\left(\theta_{0i}, \frac{1}{n}\mathcal{I}(\tilde{\theta}_i)^{-1}\right) \tag{4.8}$$

4.1.2 Wilks' theorem

In Wilks' theorem, all of these conditions must be respected. This means that not only the MLEs exist, but each of their density is very close to that of a normal distribution of mean θ_{0i} and of variance equal to $\frac{1}{n}\mathcal{I}(\tilde{\theta}_i)^{-1}$ [58], and then be written as [56]:

$$\frac{|\mathcal{I}_{kl}|^{1/2}}{(2\pi)^{b/2}} e^{-\frac{1}{2}\sum_{k,l=1}^{b}\mathcal{I}_{kl}z_k z_l}(1+O(1/\sqrt{n}))dz_1...dz_b \tag{4.9}$$

where $z_i = (\tilde{\theta}_i - \theta_i)\sqrt{n}$ for $i = 1...b$. The likelihood function $\mathcal{L}(x_j;\theta_i)$ will be the product of all probability density functions under some hypotheses. In the null hypotheses H_0, it is evaluated somewhere in the space of possible values for the parameters, while in the hypothesis being verified H, the values of the parameters θ_i are taken to be θ_{0i} for $i = 1...b$. The interest lies in the least upper bound of both likelihood functions at a given point x_j in the variable space. The quantity λ defined as:

$$\lambda = \frac{\mathcal{L}_H(x_j)}{\mathcal{L}_{H_0}(x_j)} \tag{4.10}$$

then represents the likelihood ratio for testing the hypothesis H that the x_j are from a population which distribution is characterized by θ_{0i} [56]. The logarithm of this quantity equals:

$$\begin{aligned} -2\log\lambda &= -2\log\frac{\mathcal{L}_H}{\mathcal{L}_{H_0}} \\ &= 2(\ell_{H_0} - \ell_H) \end{aligned} \tag{4.11}$$

since $\ell(x_j;\theta_i)$ is the log-likelihood function defined in Equation 4.4. This equation hints at how the difference of log-likelihoods in two hypotheses (ΔLLH) will be linked to the difference of Test Statistics as defined for the Time-Clustering algorithm, as shown in Equation 4.14. Replacing in Equation 4.4 the log-likelihood for the MLEs $\ell_{H_0} = \ell(\tilde{\theta}_i)$ by its Taylor expansion up to second order evaluated at $\ell_H = \ell(\theta_{0i})$, with $i = 1...b$.

$$\begin{aligned} -2\log\lambda &= 2\sum_{i=1}^{b}\left((\tilde{\theta}_i - \theta_{0i}).\ell'(\theta_{0i})\right) + \sum_{i=1}^{b}\left((\tilde{\theta}_i - \theta_{0i})^2.\ell''(\theta_{0i})\right) \\ &= 2\sum_{i=1}^{b}\left(n^{1/2}.(\tilde{\theta}_i - \theta_{0i}).n^{-1/2}.\ell'(\theta_{0i})\right) + \sum_{i=1}^{b}\left(\left(n^{1/2}.(\tilde{\theta}_i - \theta_{0i})\right)^2.n^{-1}.\ell''(\theta_{0i})\right). \end{aligned} \tag{4.12}$$

This summation is justified because the MLEs $\tilde{\theta}_i$ are independent and identically distributed,

by virtue of the regularity condition 1. Then, the regularity conditions 6 and 7 yield $n^{-1}.\ell''(\theta_{0i}) \xrightarrow{p} -\mathcal{I}(\theta_{0i})$ and $n^{-1/2}.\ell'(\theta_{0i}) = n^{1/2}.\mathcal{I}(\theta_{0i}).(\tilde{\theta}_i - \theta_{0i}) + a_n$, with $a_n \xrightarrow{p} 0$. So,

$$\begin{aligned} -2\log\lambda &\approx \sum_{i=1}^{h} \left(2\left(n^{1/2}.\mathcal{I}(\theta_{0i}).(\tilde{\theta}_i - \theta_{0i}).n^{1/2}.(\tilde{\theta}_i - \theta_{0i})\right) - \mathcal{I}(\theta_{0i}).n.(\tilde{\theta}_i - \theta_{0i})^2\right) \\ &= \sum_{i=1}^{h} \left(n.\mathcal{I}(\theta_{0i}).(\tilde{\theta}_i - \theta_{0i})^2\right) \\ &= \sum_{i=1}^{h} \left(\sqrt{n}.\left(\mathcal{I}(\theta_{0i})\right)^{1/2}.(\tilde{\theta}_i - \theta_{0i})\right)^2 \end{aligned} \tag{4.13}$$

This is very important, as it was previously shown that $\sqrt{n}.\left(\mathcal{I}(\theta_{0i})\right)^{1/2}.(\tilde{\theta}_i - \theta_{0i}) \xrightarrow{d} N(0,1)$. So there is a sum of the squares of h independent standard normal variables, and this is exactly the definition of a χ^2 distribution with h degrees of freedom (Equation 4.2). We have then found $-2\log\lambda = \chi^2_{(h)}$ again.

4.2 Test statistic space of all-sky alerts

All-sky alerts, as explained in Subsection 3.2.4, use HEALPix to apply the time-clustering algorithm on pixels of finer resolution at every iteration, each time using the center of the pixel as source position hypothesis to calculate the likelihood of Equation 3.6 and subsequently the Test Statistic (TS) of Equation 3.7. In this computation, the position of the pixel is a variable for the PDFs while the number of signal-like events in the flare and spectral index of the flux are free parameters determined by maximizing the likelihood. Since the logarithm is monotonic, it is also possible to directly maximize the TS instead.

In the case of all-sky alerts, we then have a pixelated map, for which each pixel is associated with a TS value, like shown in Figure 4.1. The pixel where the TS is the highest, referred to as *best fit* pixel, is the one which will provide the n_s and γ for the alert. However, although this pixel represents the best guess to fit the observed events at the time of the alert, it is not certain that the position of the source corresponds exactly to that of the center of the pixel. This problem is circumvented by relating the TS map to Wilks' theorem.

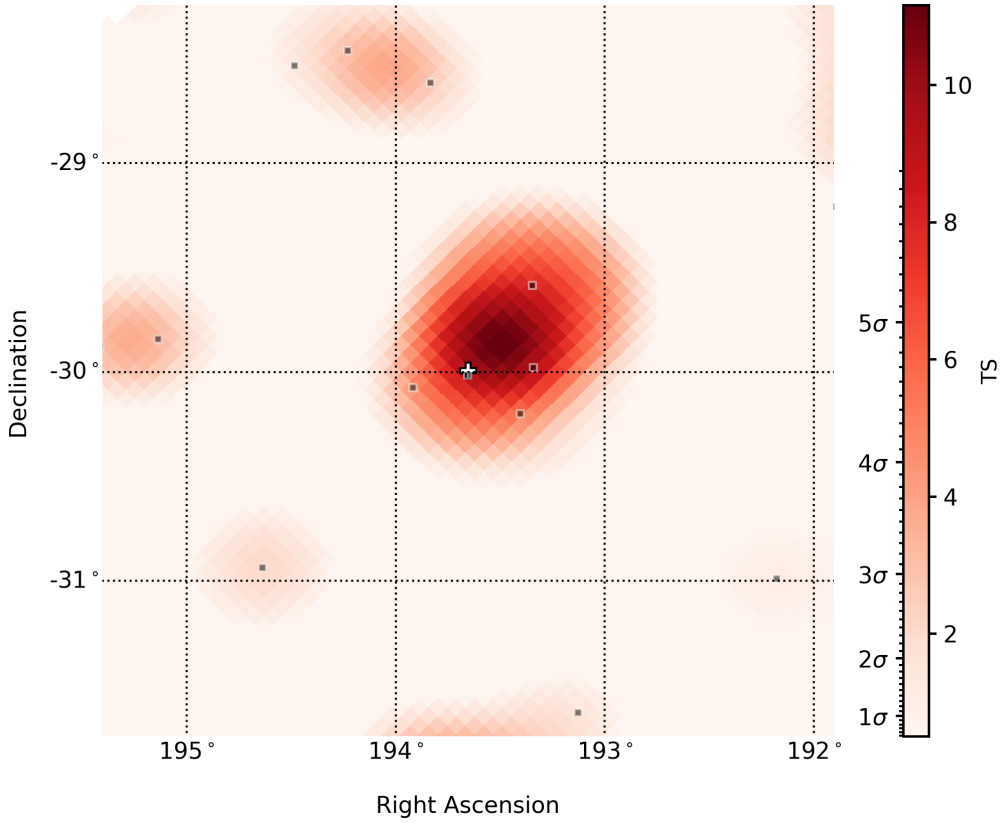


Figure 4.1: Test Statistic map of a simulated alert for a flare of 10 days at $\delta = -30°$. The color bar also shows the p-value significance in terms of σ, which serves to assess how likely the cluster is to come from a specific pixel, and so it is used to decide whether or not an alert is raised. The source injection position is showed with a white + in the middle of the map while all the events are represented by a gray square.

The regularity conditions from Subsection 4.1.1 are intuitively respected if n_s, γ are seen as the variables (they are independent and identically distributed with PDFs described in Subsection 3.2.1, as per condition 1). This means that the parameters are the right ascension (RA) and declination (δ), of which the space is indeed compact, as per condition 2. Conditions 5 & 6 require the log-likelihood function to fulfill some properties that the TS can be verified to obey, since it is the difference between log-likelihoods of two hypotheses.

To find the formulation of Wilks' theorem from the TS map, there needs to be a comparison between the TS of each pixel and the one from the best fit. Even though we will display the maps in terms of TS - TS_{best} (e.g. Figure 5.6), this yields a negative number, so to avoid sign

confusion, since the χ^2 distribution is positive by definition, it's easier to start from:

$$\begin{aligned}
\Lambda_{best} - \Lambda &= 2\log\left(\frac{\mathcal{L}(\hat{n},\hat{\gamma}|ra_{best},\delta_{best})}{\mathcal{L}(0,-|ra_{best},\delta_{best})}\frac{T_{cluster}}{T_{max}}\right) - 2\log\left(\frac{\mathcal{L}(\hat{n},\hat{\gamma}|ra,\delta)}{\mathcal{L}(0,-|ra,\delta)}\frac{T_{cluster}}{T_{max}}\right) \\
&= 2\log\left(\frac{\mathcal{L}(\hat{n},\hat{\gamma}|ra_{best},\delta_{best})}{\mathcal{L}(0,-|ra_{best},\delta_{best})}\frac{T_{cluster}}{T_{max}}\cdot\frac{\mathcal{L}(0,-|ra,\delta)}{\mathcal{L}(\hat{n},\hat{\gamma}|ra,\delta)}\frac{T_{max}}{T_{cluster}}\right) \\
&= 2\log\left(\frac{\mathcal{L}(ra_{best},\delta_{best})}{\mathcal{L}(ra,\delta)}\right) \\
&= -2\log\left(\frac{\mathcal{L}(ra,\delta)}{\mathcal{L}(ra_{best},\delta_{best})}\right) \\
&= -2\log\lambda,
\end{aligned} \tag{4.14}$$

with λ being the one defined in Equation 4.10 as the ratio between the likelihood that the fit flare properties correspond to a source located at an exact but unknown position $(ra,\delta)_0$, and the likelihood that the same flare properties correspond to a source located at the best fit pixel. Notice that since n_s,γ are the variables for the likelihood, their value is kept fixed at $\hat{n},\hat{\gamma}$ the values obtained when maximizing the likelihood at the best fit pixel.

$(1-\alpha)$ (%)	M = 1	M = 2	M = 3
68.27	1.00	2.30	3.53
90.	2.71	4.61	6.25
95.	3.84	5.99	7.82
95.45	4.00	6.18	8.03
99.	6.63	9.21	11.34
99.73	9.00	11.83	14.16

Table 4.1: Values of the $(1-\alpha)$ quantiles for a $\chi^2_{(M)}$ distribution. If the source containment satisfy Wilks' theorem expectations, the M=2 column will provide the intervals on which to build contours for the confidence regions in the TS difference map, with the confidence level shown on the left. Taken from [59].

So, if the conditions described in Subsection 4.1.1 are respected, the difference between the TS at the best fit pixel and the TS of every pixel should be distributed like a χ^2 distribution with two degrees of freedom (one for the right ascension, and one for the declination). This would, if confirmed, allow us to build confidence regions based on the known confidence intervals for χ^2 distributions, as referenced in Table 4.1, because we would know at which difference in

TS the source would be contained in the region a certain percentage of times. This also has advantages in the context of real-time analyses and alerts, because most calculations are then already performed during the scan.

4.3 Simulation of alerts

To verify this, we are going to simulate a great number of alerts by injecting simulated events in archival data, scrambled in right ascension. The randomization cannot be performed in declination without introducing a bias because the declination rate is not uniform. Every injected event has the same *true position* that corresponds to the source of the simulated flare, but different *fit positions*, which are the individual directions of origin that will be used by the algorithms to calculate the LLH for the direction of origin for the entire flare, which should be close to the true position. This way, we actually perform the analysis on simulated events and are able to evaluate the accuracy of Wilks' theorem's predictions about the source containment because unlike with a real-time alert, the true position of the signal-like events is known. Its declination is chosen manually because the performances of the analysis are expected to vary with declination, but the right ascension of every event, both signal-like and background, is randomized. The injected events are also assigned a spectral index and are spread over a defined time-window. However, the time PDF does not match a box profile like the time-clustering algorithm assumes. Instead, it has a gaussian shape, but this has but a negligible impact on the alerts: gaussian shaped flares have been shown to be reconstructed accurately by a box-profiled time-clustering algorithm, and the loss in sensitivity was only around 5% [50].

The idea is to go through the scrambled data with injected events as if the telescope was receiving the events at the referenced time and, as soon as an alert is raised, produce the ΔLLH (TS difference) map. We can retrieve from that the TS from the actual injection position, as well as the angular distance between the source and the best fit pixel. We are only interested in the analysis at the moment of the alert because this work aims to be applied to the RealTime analysis, so we are trying to reproduce the conditions in which an alert is raised. Of course, a more in-depth analysis studying the performances of the method with the best fit flare (see Subsection 3.2.3) could be interesting, and performed a posteriori, but this would be beyond the scope of this project. Once the process has been repeated a sufficient number of times, enough to satisfy the large number assumption for Wilks' theorem, and the relevant data has been ordered, we will compare the Cumulative Distribution Function (CDF) of the -2ΔLLH with

that of a χ^2 distribution with two degrees of freedom.

For the sake of robustness, this simulation has to be done for several sets of parameters, to have an idea of how suited this approach is for flares of varied properties. In this work, the impacts of declination ($\delta = -30, 0, 30°$), spectral index for the injection source ($\gamma = -2, -2.5, -3$), and the spread of the events in the flare ($T_{flare} = 1, 10, 100$ days) were evaluated. It is important to keep in mind that the properties of the injected flare will differ from those reconstructed by the time-clustering algorithm. Indeed, it will be working at the moment the alert is raised, so $T_{flare} \neq T_{cluster}$, $n_{inj} \neq \hat{n}_s$, and $\gamma_{inj} \neq \hat{\gamma}$. This is due to how the time-clustering algorithm works. Referring to Section 3.2, it is unlikely that the alert would be raised only when all the simulated signal events have been injected, so the time-window evaluated to construct the first guess for the duration of the flare has to be different than the one over which all the injections are to be spread. Similarly, the fit spectral index is expected to be different than the injected one, especially in the case of shorter flares with few accounted events.

5 Results and Interpretation

After running the simulations, we are able to plot the Cumulative Distribution Function of the results and compare them to the expectation given by Wilks' theorem. However, the goal of the project is still to be able to build robust confidence regions, even if they do not fit Wilks' predictions.

In this chapter are presented all the results from this work, but also the shortcomings of Wilks' theorem when it comes to approximate the simulations (Subsection 5.1.1) and the efforts to obtain a good fit for the data in every situation (Subsection 5.1.2). Section 5.2 presents how to display the confidence regions and the remaining steps to be done before elliptical regions based on the results of this project can eventually be included in the sharing of all-sky alerts. Finally, Section 5.3 discusses other aspects of the simulation studied in this work, as well as an alternative method of building confidence regions based on angular separation.

5.1 Comparison with Wilks' theorem

To verify if two distributions are equal or at least if their quantiles are the same, we choose to use the Cumulative Distribution Function (CDF). Not only is it a continuous representation of our discrete data, as opposed to the density function, which is plotted as a histogram, but it also immediately shows the proportion of the data under a certain level.

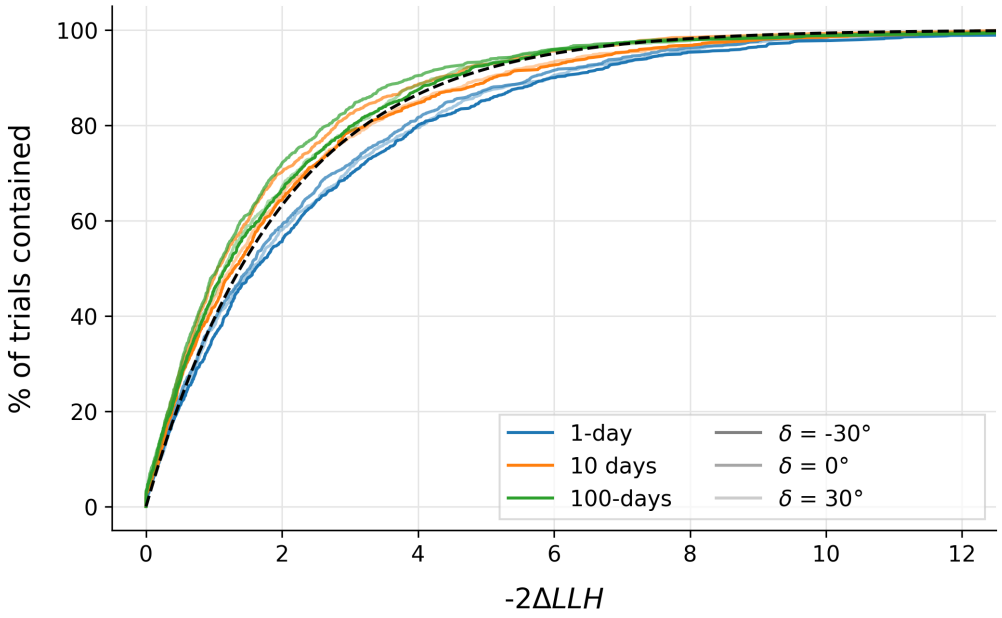


Figure 5.1: Cumulative Distribution Functions of the simulations for injected flares with a spectral index $\gamma = -2$. The duration of the flares is color-coded, and the line opacity indicates the declination. The black dashed line is the CDF for a χ^2 distribution with two degrees of freedom.

Figure 5.1 shows the CDFs for the simulated flares over different declinations and flare durations, compared to that of a χ^2 with two degrees of freedom. The injected events displayed in this plot all had the same spectral index of $\gamma = -2$, which is close to what is expected of the diffuse astrophysical flux. In these cases, the results are very promising, the χ^2_2 CDF fits the results closely. This means that even if the simulations were not distributed *exactly* like a χ^2_2, they are sufficiently close to be able to use the values in Table 4.1 to describe the confidence regions in the TS difference map.

The CDFs for the 1-day long flares, those for which the alerts were raised after very few events, however, are generally under the χ^2 curve. This is not as much a problem when the curves are above Wilks' prediction because it just means that the χ^2 values will represent a conservative interval. But the undercoverage (when the number of trials contained in the region is less than expected according to Wilks' theorem) represents an issue that will need to be addressed. For flares injected with a spectral index of -2, the effect is not too strong, but as we will see below, it becomes a real concern when the spectral index is softer.

5.1.1 Undercoverage of the simulation

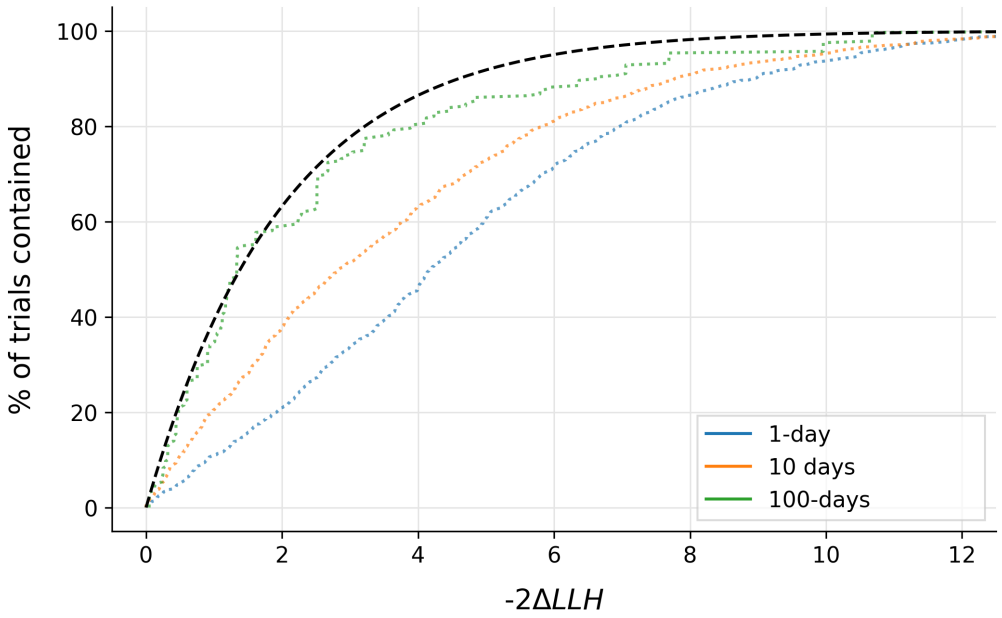


Figure 5.2: Results from the simulations of flares with spectral index $\gamma = -3$ at declination $\delta = 0°$. The black dashed line is again the CDF for a χ^2 distribution with two degrees of freedom. There is a bigger spread between the flares of different durations at this declination and soft spectrum because the signal-like background is most variable at the horizon.

Figure 5.2 serves as a good example of how the simulation strays from Wilks' theorem predictions at softer spectral indices. This first discrepancy is explained in part by the fact that at $\gamma = -2.5$ and $\gamma = -3$, the flares are more difficult to distinguish from the background in that the spectral index of the background is closer to the one of the signal. This is also a growing problem at higher declination, where the signal-like background is more present and more variable.

Also, the pull-correction for the angular error estimators described in Subsection 3.2.1 assumes a spectral index for the events of $\gamma = -2$. This means that the angular uncertainty for the individual events is slightly mis-reconstructed and overestimated for flares that deviate from $\gamma = -2$. This effect is even more present for short flares, since the alert is raised very quickly. The algorithm has to rely on few events to determine the position of the cluster source, and the individual errors contribute more to the overall precision of the prediction.

Less accuracy in the reconstruction of the parameters of the flares might affect the distribution of the variables for the likelihood used in Wilks' theorem, namely the number of accounted signal-like events and computed spectral index. If they were not respecting the regularity conditions presented in Subsection 4.1.1, e.g. if they were not identically distributed, or if their log-likelihood function was not twice continuously differentiable close to the true position, it could explain why the results of the simulation stray from the expected distribution.

The combined actions of these causes can be seen in Figure 5.3, where the residuals (difference between the simulations and the prediction) for 1-day flares are shown. The difference between the containment of the data and the one expected by Wilks' theorem reach as high as 50%, which represents a gap between expectation and simulation too important to accept these results.

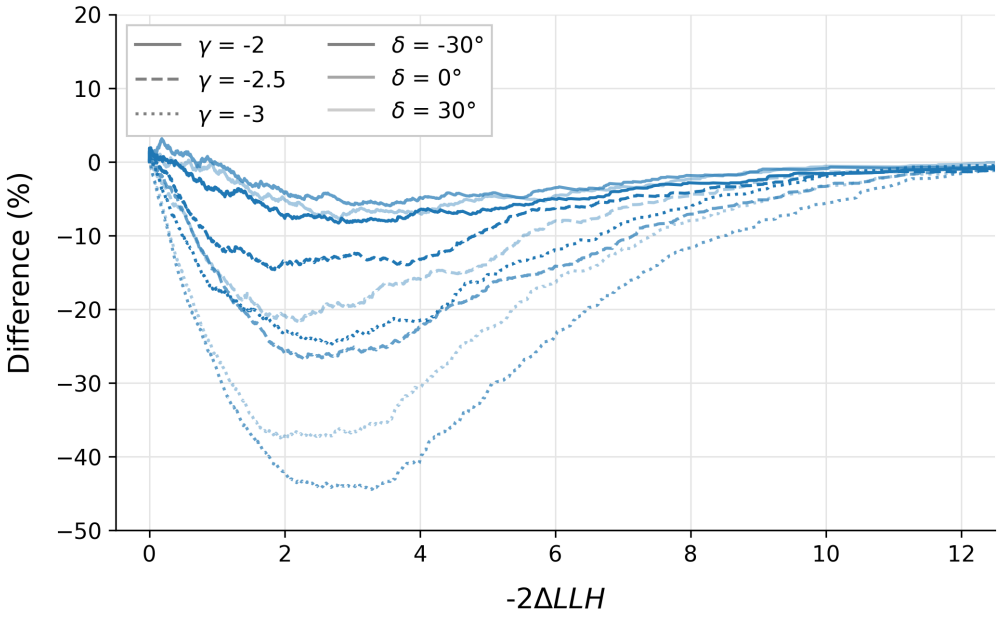


Figure 5.3: Residuals for all the 1-day flares, i.e. the difference between the CDFs of the simulation and the prediction from Wilks' theorem.

To remedy this problem, we could work towards the improvement of the individual angular error estimation. For example, the pull correction described in Subsection 3.2.1 relies on a Rayleigh function to represent the distribution of reconstructed events, but it does not take the spectral index of the events into account. An alternative that includes a dependence on the

spectral index, such as the King function [60], could replace the Rayleigh function to improve the individual angular uncertainty. In general, any improvement to the reconstruction of the tracks would mean that the uncertainty for the direction of origin is better and that the clusters are assigned more accurately to a source position.

5.1.2 Adjusting the number of degrees of freedom of the χ^2 distribution

A better localization of the source of a flare brought on by an improved individual angular uncertainty would contribute to solving the undercoverage, but there is no guarantee that the amelioration of the results would be significant, or that the prediction from Wilks' theorem would be perfectly respected. The principal objective of this work remains to establish confidence regions that match the current performance of the cluster alert algorithm. Strictly following this goal, it is not necessary to correct the undercovered cases if there is a way to accurately represent the data.

We can realize that the number of degrees of freedom of the χ^2 distribution is simply a parameter, optimizable to obtain a closer fit for the simulations in every case. Examples of this improvement to the fitting of the data are shown in Figure A.2. It becomes clear that for every set of parameters, there is a value for the number of degrees of freedom that allows for a better representation than the nominal 2 required by Wilks' theorem. The optimal number of degrees of freedom depending on the parameters of the simulations is shown in Figure 5.4. The black error bars on this plot represent the standard deviation on the optimization result, magnified 15 times for visibility. When the standard deviation associated to the value is larger, the behavior of the curve resembles more that of Figure A.2b: the fit is better than Wilks' prediction, but not perfect.

Despite the slight discrepancies present even with the best number of degrees of freedom possible, the prediction and data always coincide at one level at least: the curves cross at one point. Figure 5.5 shows the difference between the best fitting χ^2 distribution and the data at a declination of 0°, where the fits are the least accurate. It is still apparent that every fit predicts accurately the 50% containment level. The 90% containment level is also reasonably close to the simulations. As we are looking at the tails of the curves, a difference this large does not have a noticeable impact.

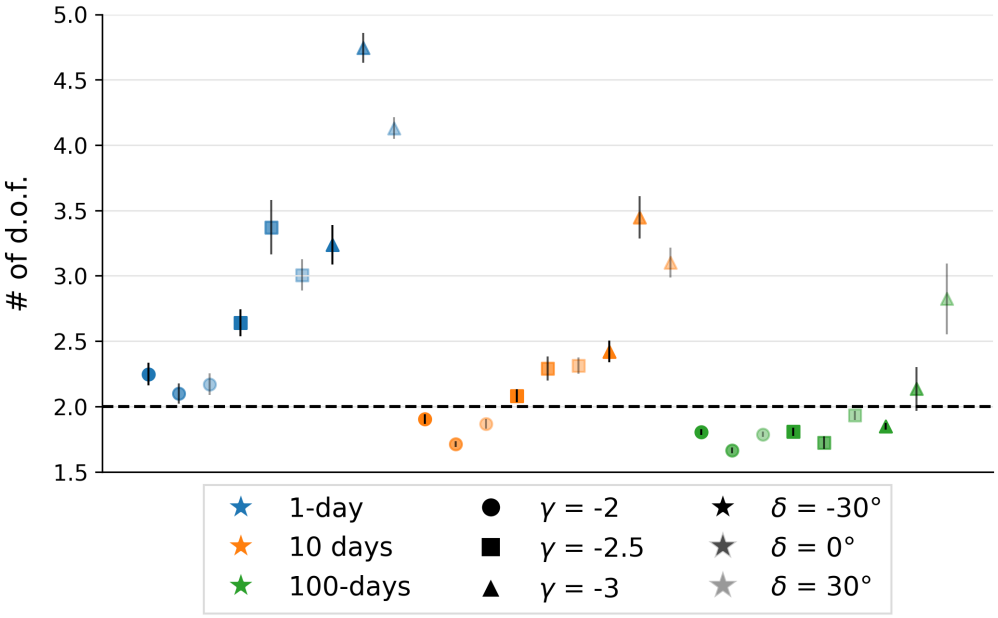


Figure 5.4: Number of degrees of freedom for all simulations. The black error bars show the standard deviation on the estimation, multiplied by 15 for visibility. We can see the tendency for flares with softer spectra or shorter durations to require more degrees of freedom.

	$\gamma = -2$			$\gamma = -2.5$			$\gamma = -3$		
δ	$-30°$	$0°$	$30°$	$-30°$	$0°$	$30°$	$-30°$	$0°$	$30°$
	1 day								
χ^2 d.o.f.	2.25	2.1	2.17	2.64	3.37	3	3.24	4.74	4.13
50%	1.63	1.48	1.55	2.01	2.73	2.37	2.6	4.09	3.48
90%	5.03	4.77	4.9	5.68	6.83	6.26	6.62	8.87	7.97
	10 days								
χ^2 d.o.f.	1.9	1.71	1.87	2.08	2.29	2.31	2.42	3.45	3.1
50%	1.29	1.11	1.26	1.46	1.67	1.69	1.8	2.81	2.47
90%	4.44	4.09	4.37	4.74	5.1	5.14	5.32	6.95	6.41
	100 days								
χ^2 d.o.f.	1.8	1.66	1.79	1.81	1.72	1.93	1.85	2.13	2.82
50%	1.2	1.06	1.18	1.2	1.12	1.32	1.24	1.52	2.19
90%	4.26	4.01	4.23	4.27	4.11	4.48	4.34	4.84	5.97

Table 5.1: Values for the number of degrees of freedom for the χ^2 distributions depending on the parameters. The 50-90% quantiles are also referenced for every χ^2.

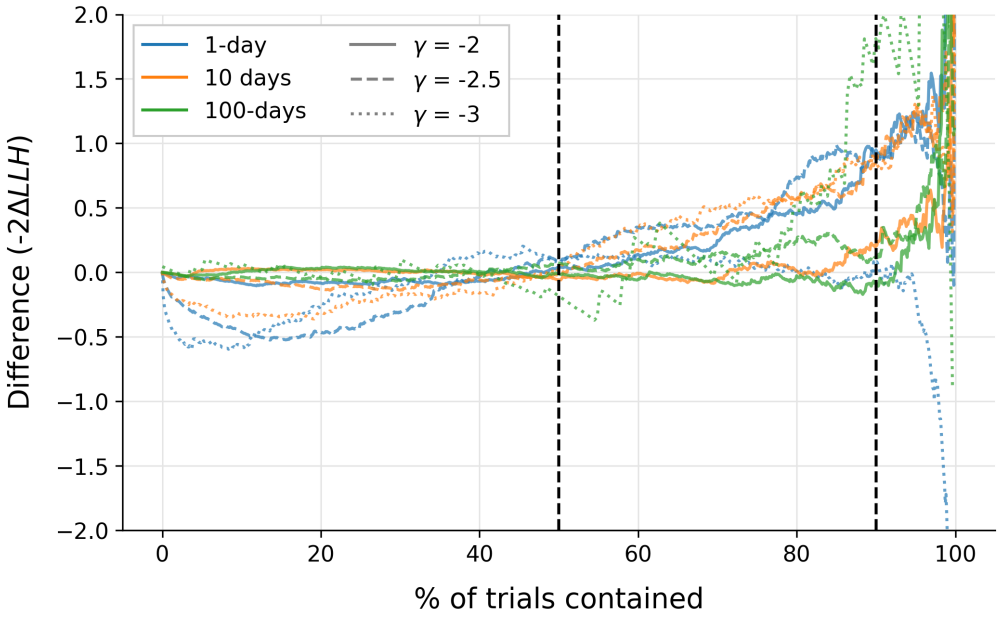


Figure 5.5: Difference between the quantiles of the data and the χ^2 distributions with the best fit number of degrees of freedom for flares injected at $\delta = 0°$. Dashed lines indicate the 50 and 90% quantiles and show that they are both close enough in every case to be used to draw the confidence regions.

At this point, we are no longer comparing to Wilks' theorem because its prediction is that the χ^2 distribution has two degrees of freedom. When the number of degrees of freedom is less than two, i.e. when the reconstruction yields more accurate results than Wilks' expectations, it is equivalent to asserting that the two free parameters (declination and right ascension) are slightly correlated. In other words, in these cases, the reconstruction is accurate enough that it seems like less parameters are required to describe the distribution of the results. On the other hand, when the number of degrees of freedom is higher than two, then the undercoverage is equivalent to introducing more variables to describe the results. However, the confidence levels for the 50 and 90% quantiles are all that we need to draw the regions that we are looking to share with partner telescopes. As such, the values presented in Table 5.1 are largely sufficient for the objective of this work.

5.2 Ellipse approximation

The final step to building the confidence regions is to draw them, then approximate them by ellipses. It will be easier to describe the regions with the ellipse parameters, along with the best

fit pixel center around which it is being drawn. This is made by recovering the positions of the pixels in which the TS difference with that of the best fit pixel is close to the level determined above. Since the TS distribution on a pixelated map is discrete, in order to have enough points to fit an elliptical curve, the level must be evaluated with a margin.

What is left to do, although it is beyond the extent of this project, is to generalize the evaluated parameter space. Simulate a larger number of flares for all declinations, and for a wider range of spectral indices. This will allow for an improved version of Table 5.1. Then, once a cluster alert with a significance above the threshold discussed in Subsection 3.2.4 is raised by the time-clustering algorithm, it will be possible to reference the reconstructed parameters of the flares in this table, and construct elliptical confidence regions.

The only issue with this method is that the spectral index of a cluster of events is usually misreconstructed, especially in the case of short flares, where a softer spectrum is estimated. The γ in the table refers to the injection source spectral index, which is then unknown, but can be assumed to lie between -2 and -3. This allows for the estimated power law used to describe the source's flux to be bounded by two limits, one harder than that of the measured astrophysical neutrino diffuse flux, and one closer to that of the atmospheric neutrino flux. That way, even if the spectral index of the source of the cluster is uncertain, the true border of the confidence region lies inside the confidence belt created. A prototype of what such confidence bands would look like on the simulated alert from Figure 4.1 is shown in Figure 5.6.

5.3 Additional considerations

This section regroups contributions that can be added to the result presented above, obtained from studying the impact of different parameters. Subsection 5.3.1 discusses how results are affected when the significance requirement for an alert to be raised by the time-clustering algorithm changes. In Subsection 5.3.2, the choice of the number of events injected in each simulated flare is justified, and in Subsection 5.3.3, an alternative confidence region based on the angular distance between best fit and injection pixels is presented.

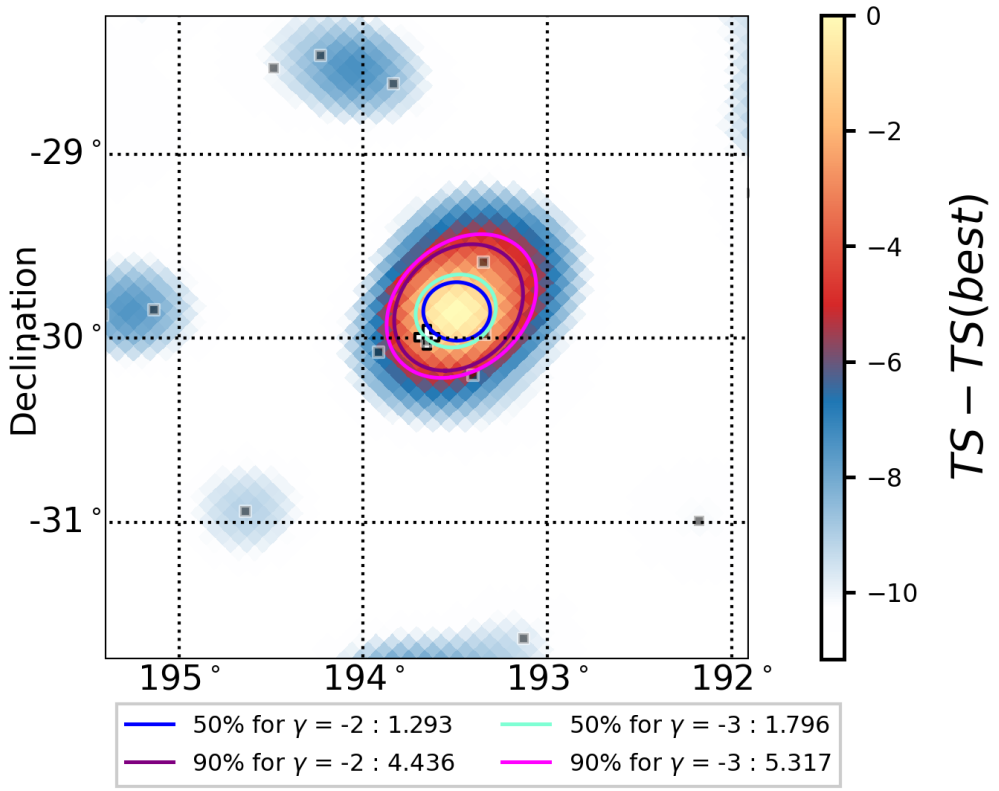


Figure 5.6: Confidence belts for a simulated alert. Ellipses are being drawn for the 50 and 90% quantiles of the flares assuming spectral indices of $\gamma = -2$ and $\gamma = -3$. The bands thus created contain the border of the actual confidence region. The numbers shown in the legend correspond to the Test Statistic difference delimited by each ellipse. They are obtained from the χ^2 distributions associated with a flare lasting 10 days at a declination of -30°.

5.3.1 Significance thresholds for alerts

As mentioned in Subsection 3.2.4, the pre-trial significance threshold that a cluster must surpass to raise an alert is set at 4.2σ for all-sky scans. This value is chosen to get an expected alert rate of one per year (see Figure 3.6). All the results presented so far were produced with this alert threshold, but all the simulations have also been run with two different thresholds.

At a threshold of 3.5σ, the expected alert rate is ten per year. Lowering the threshold allows to verify if the method and the results still hold when an alert can be raised more easily by the time-clustering algorithm. The threshold is unlikely to come down as low as this because the alerts would be too frequent, but it could be decreased to around 4σ to allow for two to three alerts per year. As such, we need to know if the results of the simulation are comparable at a

lower threshold.

Similarly, we have performed the simulation setting the threshold at 4.7σ as well, equivalent to expecting an alert per decade. This was done to verify the performances of the algorithm if we set the threshold closer to a value required for a better post-trial significance. This would mean that the raised alerts are expected to be better localized.

In the case of a lower threshold, the general trend was that the reconstruction was less precise. This was expected since a lower threshold means that fewer events are necessary to achieve a sufficient level of significance to trigger an alert. The localization of the source of the cluster is then less precise, and this effect manifests in the number of degrees of freedom of the best fitting χ^2 distribution increasing by $\approx 11\%$ on average. On the other hand, simulations with a higher significance threshold returned an overall better estimation of the source location. Again, this is expected, for the opposite reason as previously. Then, in this case, the number of degrees of freedom for the χ^2 decreases by $\approx 6\%$ on average, except for 100-days flares, for reasons that are exposed below.

5.3.2 Number of events needed to trigger an alert

The flares that were simulated in Section 4.3 all contained twenty signal-like events, even though the number of events accounted for by the time-clustering algorithm and the reconstructed number of signal-like events are both different from this injection parameter. Trials were run with only ten events in the injected flares, but in a lot of cases, this proved insufficient, i.e. the flares did not reach a significance level high enough to trigger an alert. This happened especially often for flares lasting 100 days, since the spread made it so only the clusters that were highly localized could be recognized by the algorithm. Increasing the number of injected events to twenty allowed for a more statistically significant number of alerts to be raised. When 100-days flares where injected and supposed to reach a significance threshold of 4.7σ, the number of events was increased to forty, to avoid encountering the same problem.

What can be understood from this is that dense flares will be treated like shorter flares by the algorithm. Indeed, flares of forty events were found to be reconstructed less accurately than those with a smaller threshold, but less injected events, i.e. the number of degrees of freedom of the best fitting χ^2 distribution is higher for a threshold of 4.7σ (Table A.1) than for the normal

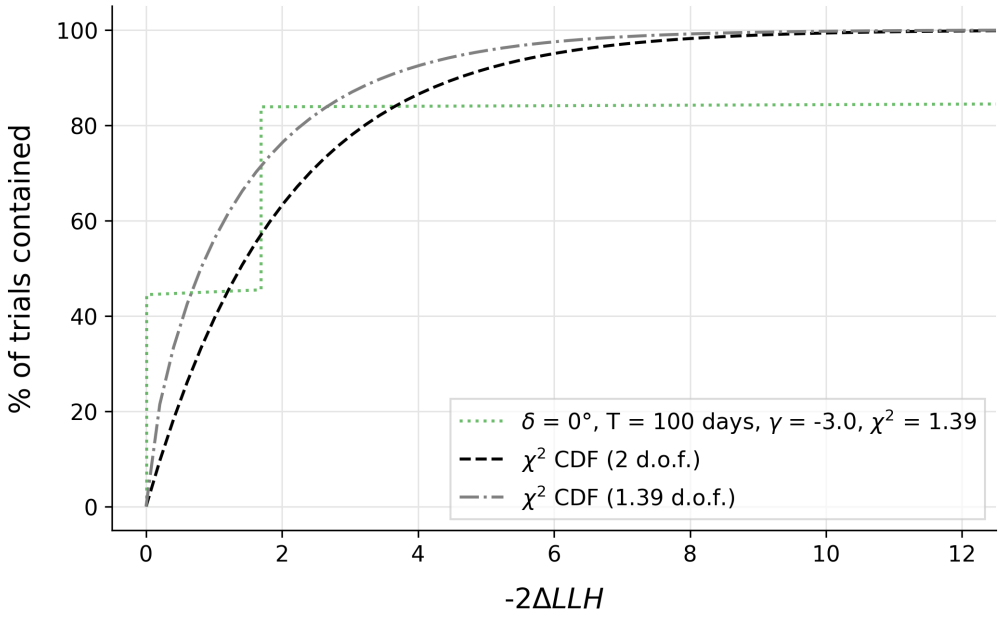


Figure 5.7: Performance of the algorithm at 30° for 100-days flares of 10 events injected with spectral index $\gamma = -3$. The injection parameters only allowed for a few very localized flares to cross the significance threshold. Comparison between the flares of 20 and 40 events with the same parameters can be found at Figure A.7.

4.2σ threshold (Table 5.1). This is because a denser injection of events raises alerts faster, and thus with less signal-like events.

5.3.3 Circular errors

In Section 4.3, we mention that while our main objective in comparing the performance of the algorithm with Wilks' theorem through the difference between the best fit pixel and the injection position, we also retrieve the angular distance between the two.

Figure 5.8 shows the CDF of the angular distance for simulations at -30°. The first thing we notice are the regular jumps in the CDF. These declination-dependent jumps come from peaks in the Probability Density Function of the results. The regularity of these artifacts indicates that they are caused by the algorithm, which favors certain distances due to how HEALPix divides the sky in pixels. Indeed, the distribution of the center of the pixels is not continuous, but discrete and declination has an effect on the orientation of these pixels.

We want to evaluate if a fit as robust as the one done on the ΔLLH is possible, we verify if a χ

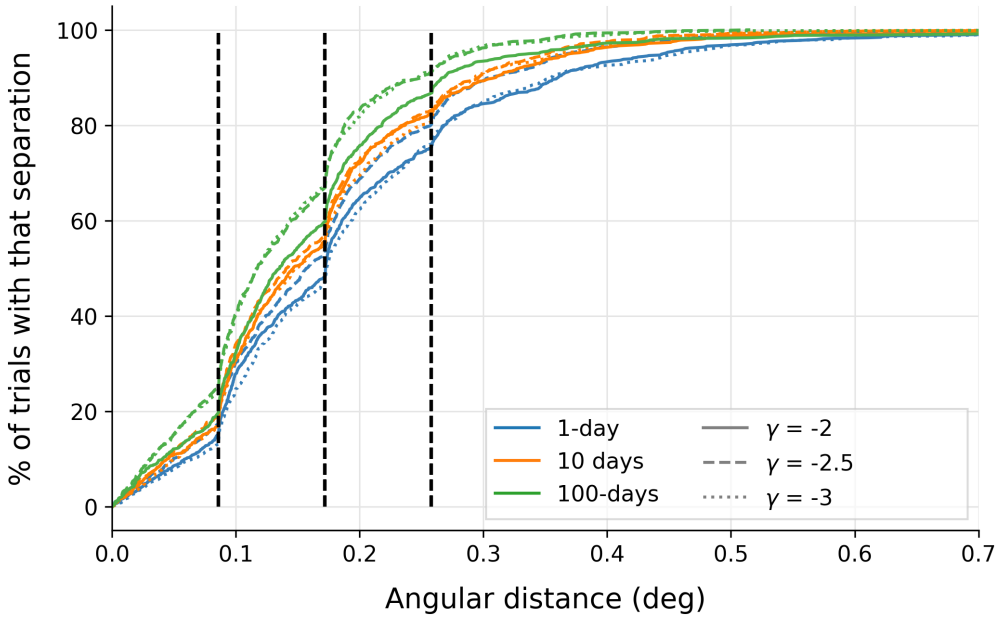
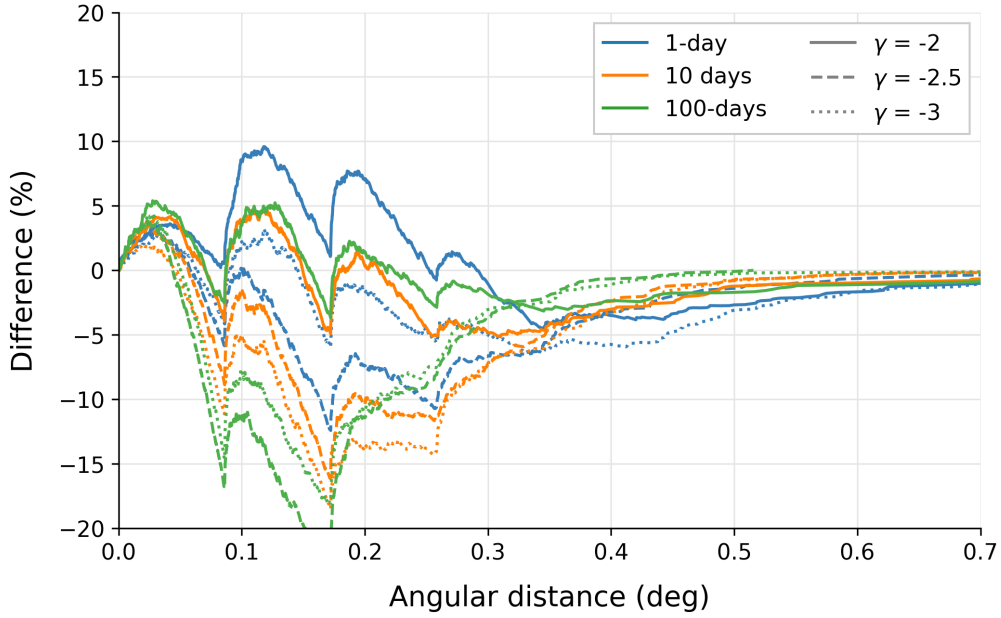
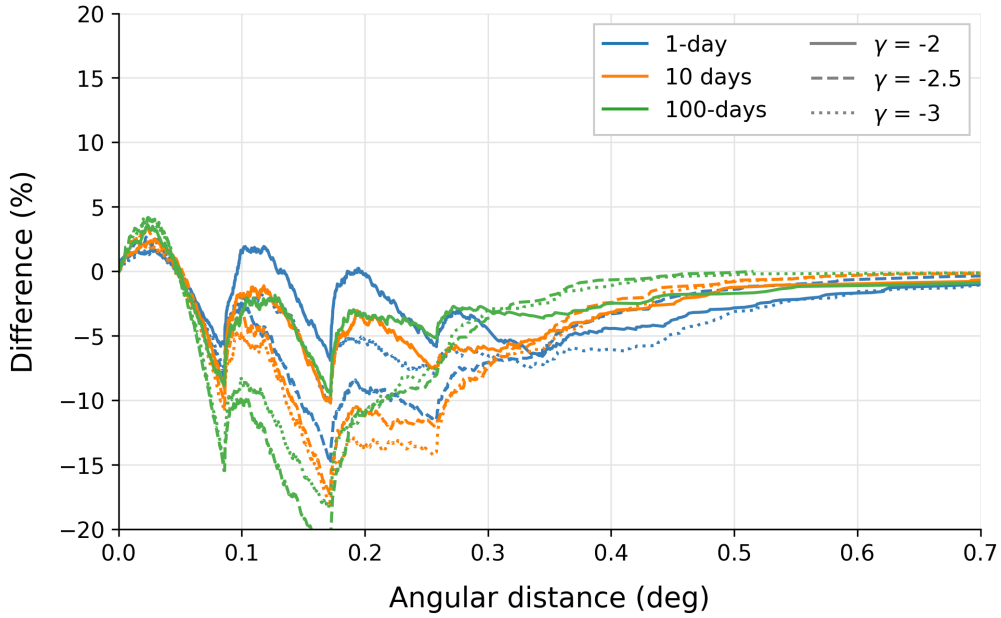


Figure 5.8: Angular distance distribution for simulations at a declination of -30°. Dashed lines show the regular jumps in the CDF.

distribution (the square root of the χ^2 used previously in this work and defined in Equation 4.2) can fit the data. The intuition behind this choice is that a χ distribution with two degrees of freedom, i.e. if the right ascension and declination are both normally distributed, is equal to the Rayleigh distribution used in the pull correction for the angular uncertainty of individual events. However, as Figure 5.9 shows, there is almost no difference between fitting the data with a Rayleigh distribution or a χ distribution with an optimized number of degrees of freedom, and none of those fits were satisfyingly good. The distribution of angular distance between the best fit and source pixels cannot be approximated in a simple and robust way like the difference in their Test Statistic can.



(a) Comparison with the χ_2 distribution



(b) Comparison with the best fitting χ distribution

Figure 5.9: Residuals for the angular distribution with respect to χ distributions of different numbers of degrees of freedom. The difference is not flagrant, meaning changing the scale parameter of the distribution does not have an effect on the efficiency of the fit.

6

Conclusion and Future Prospects

The objective of this work is to provide a reliable estimator for the angular uncertainty associated with the position of the source of an all-sky alert in GFU. To enable this, robust confidence regions must be determined for the position of the source of the all-sky alerts. We have simulated flares with different hypotheses of signal strength, duration and location in the sky and run them through the time-clustering algorithm to mimic the conditions in which an alert would be raised in real time. Then we compared the position of the injected flare with that returned as a best guess by the algorithm.

The results from these simulations show that it is possible to create confidence regions. The initial hypothesis relied on the assumption that the time-clustering algorithm would return alert positions according to the expectation from Wilks' theorem. This would equate the distribution of the difference in Test Statistic for the source of the alert and for the best fit pixel returned by the algorithm to a χ^2 distribution with two degrees of freedom.

However, this project has shown that this hypothesis is valid only for a select number of cases. When the flares could be reconstructed according to the underlying assumptions used by the time-clustering algorithm, the distribution of results fit closely the χ^2 distribution from Wilks' theorem. On the other hand, if the injected flare did not adhere to these assumptions, the precision of the reconstruction got worse and the results no longer matched Wilks' expectation. In order to provide a reliable fit of the data nonetheless, we changed the number of degrees of

freedom of the χ^2 distribution.

Each set of parameters corresponds to a new number of degrees of freedom, and the adjusted χ^2 offers satisfying values for the Test Statistic difference describing regions of 50 and 90% confidence levels in every case. The regions are approximated as ellipses and returned for a spectral index of $\gamma = -2$ and for $\gamma = -3$, creating bands to border the confidence regions. This is done to bypass the difficulty that the reconstruction of the spectral index of a flare can present: as long as it is between -2 and -3, the border it would produce lies in the created confidence belt.

In the future, this method can be implemented once the simulations have been run with greater statistics and for the whole range of declinations and flare durations. The method will still be relevant if the significance threshold for raising the alerts is changed, and if the individual angular uncertainty is improved by adding a spectral index dependence in its computation.

In the end, this project managed to fulfill its objective and create robust, yet simple confidence regions to quantify the uncertainty on the position of the source of a GFU all-sky cluster alert. Even if there is still work to do before this method can be implemented and the unbiased search alerts shared automatically with partner IACTs, it still contributes to the effort of enabling the simultaneous observations of gamma-rays and neutrinos, which could help further the field of multi-messenger astrophysics.

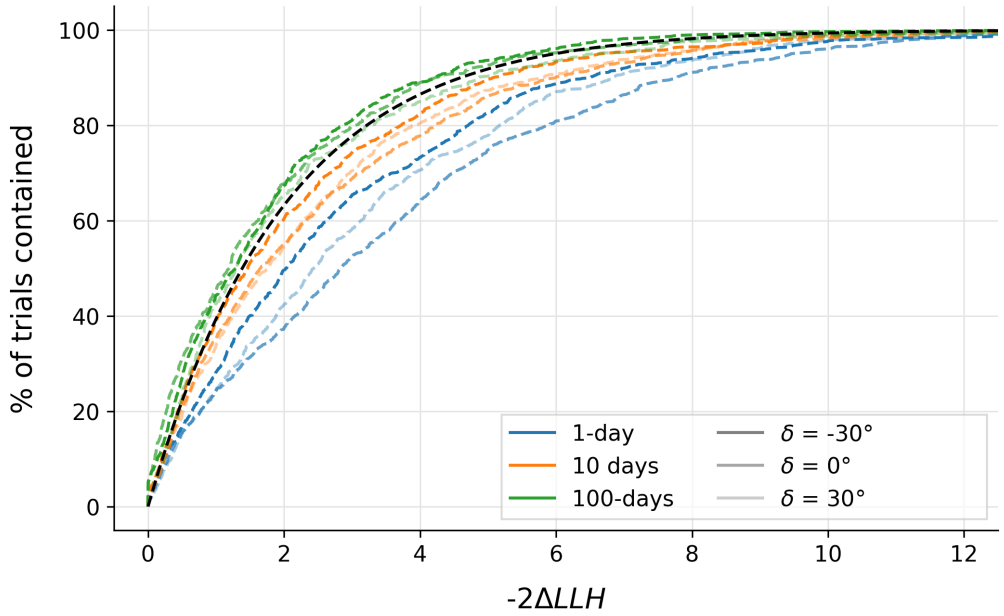
Appendices

A

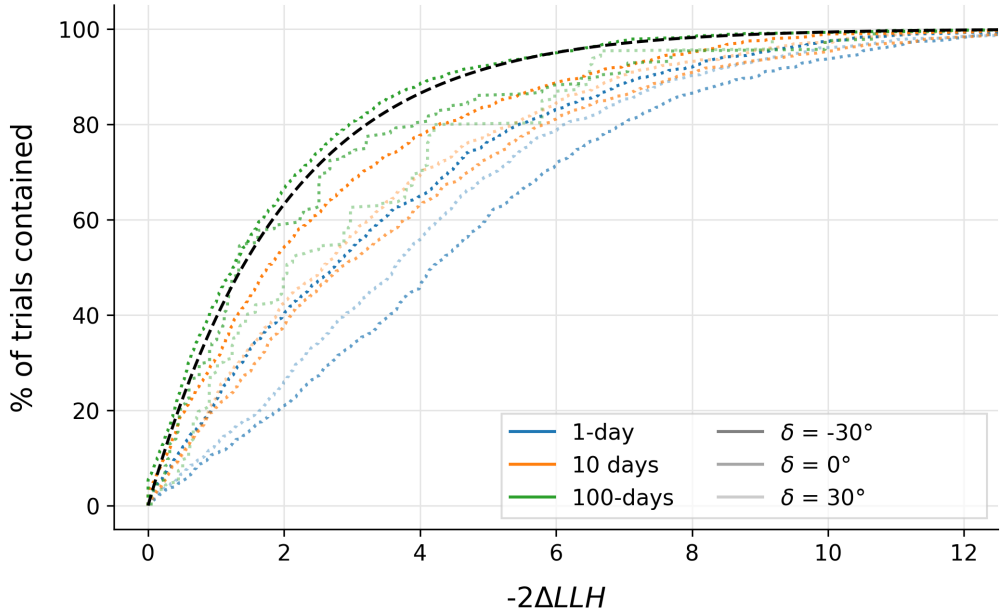
Supplementary plots and tables

A.1 Comparison of results with Wilks' theorem at softer spectra

In Figure 5.1, we show the CDFs of the results plotted against the expected χ^2_2 for flares injected with a spectral index of $\gamma = -2$. Here are the results that were less in agreement with Wilks' theory, as explained in Subsection 5.1.1.



(a) $\gamma = -2.5$

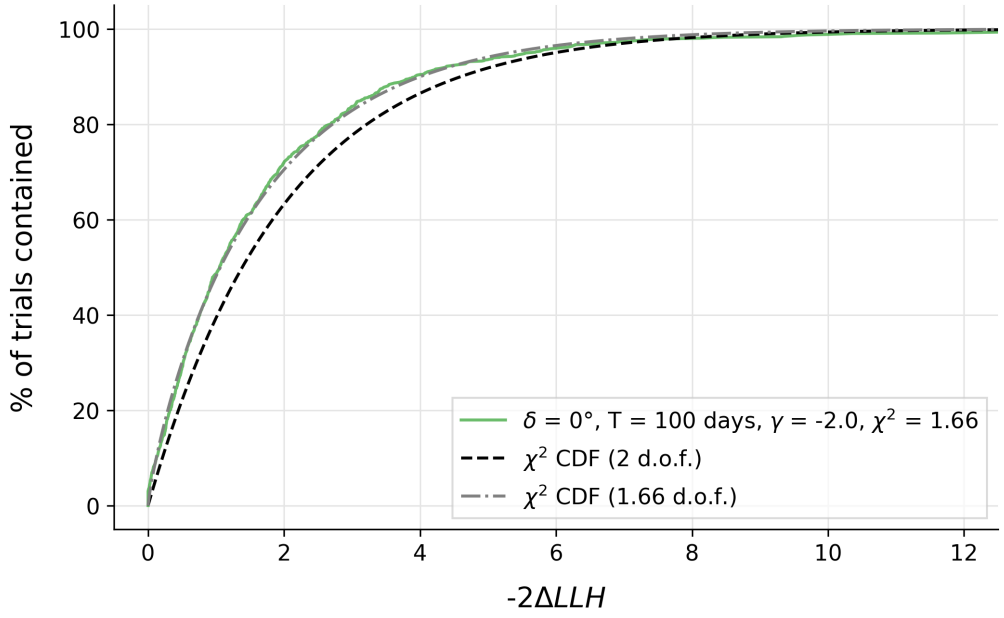


(b) $\gamma = -3$

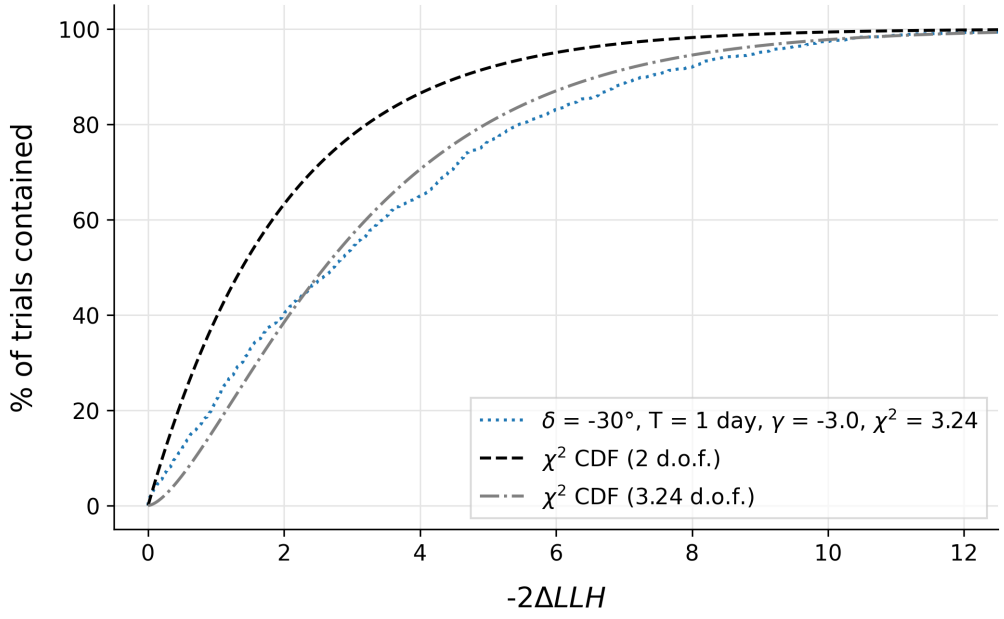
Figure A.1: Results of the simulation for softer spectra ($\gamma = 2.5, -3$, plotted against Wilks' theorem's expected χ^2_2.

A.2 Adjusted numbers of degrees of freedom for a better fit by the χ^2 distribution

Below are shown examples of how adjusting the numbers of degrees of freedom of the χ^2 distribution led to a better fit of the data in every case. Even when the new distribution didn't follow exactly the curve of the results, there is always a coincidence around the 50 and 90%. The value used for the new degrees of freedom is shown in Figure 5.4.



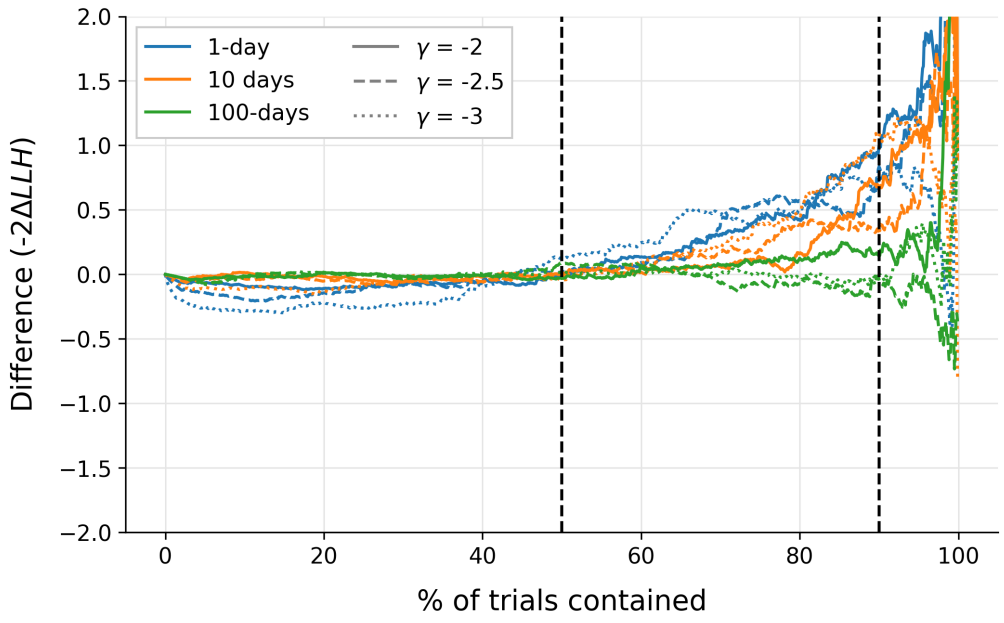
(a) # of degrees of freedom slightly lower than 2.



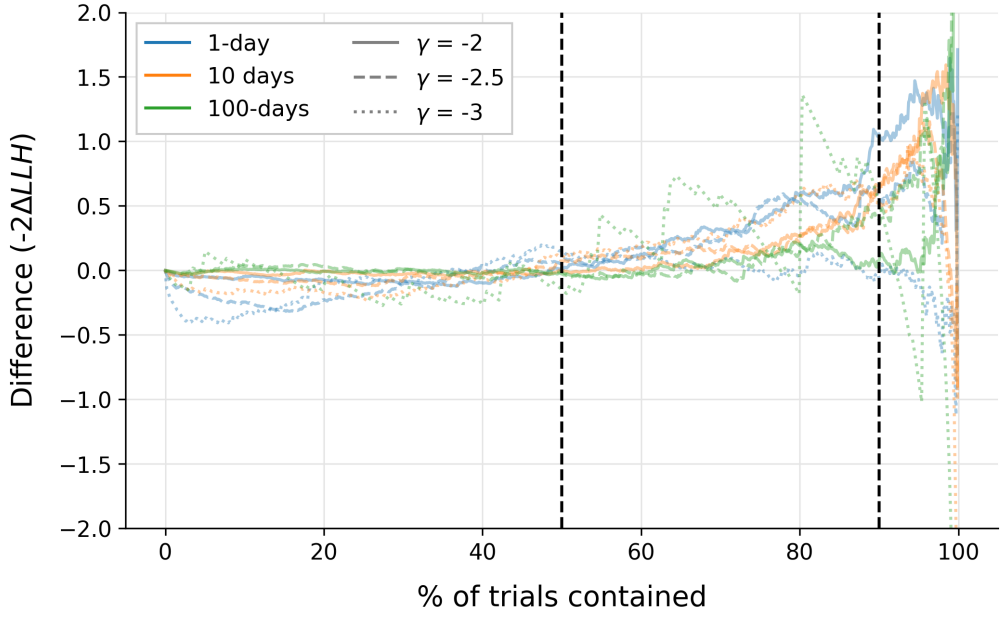
(b) # of degrees of freedom much higher than 2.

Figure A.2: Examples of the improvement for the fit by optimizing the number of degrees of freedom of the χ^2 to which the CDF is compared. The new fit is shown in gray. On the top, the new fit corresponds perfectly to the data. On the bottom however, the new fit is not perfect: the estimation was less certain, as described in Figure 5.4.

Here are shown the differences between quantiles for the data and best fitting χ^2, like Figure 5.5, but at different declinations. As mentioned in Subsection 5.1.2, these results are even better than the ones displayed in Figure 5.5, i.e. the differences at the 50 and 90% levels are smaller, which means that the χ^2 provides a more accurate estimation of the value for the border of the confidence region.



(a) $\delta = -30°$



(b) $\delta = 30°$

Figure A.3: Difference between the quantiles of the results and of their corresponding χ^2 distributions at declinations other than 0°.

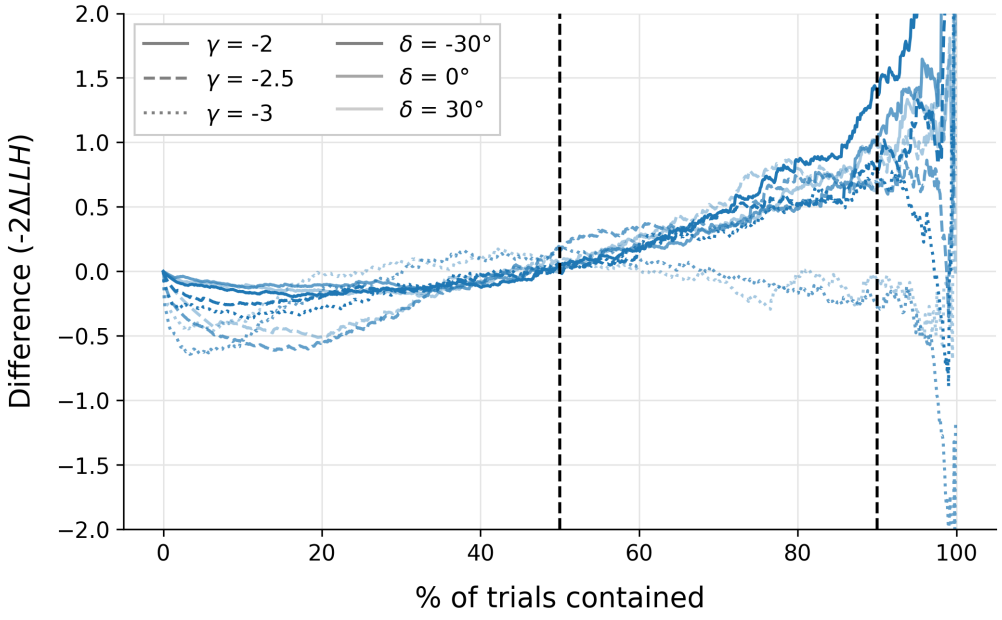
A.3 Results with different significance thresholds for alerts

Table A.1 shows the degrees of freedom and quantiles for the best fitting distributions of the data from simulations with the other thresholds, described in Subsection 5.3.1.

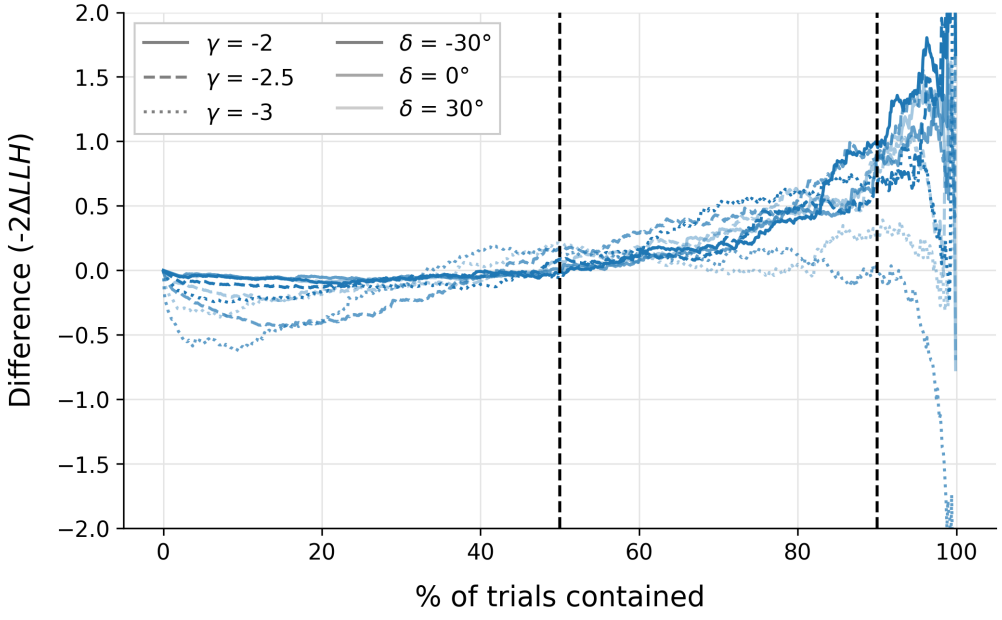
	$\gamma = -2$			$\gamma = -2.5$			$\gamma = -3$		
δ	$-30°$	$0°$	$30°$	$-30°$	$0°$	$30°$	$-30°$	$0°$	$30°$
Significance threshold $= 3.5\sigma$									
	1 day								
χ^2 d.o.f.	2.48	2.25	2.38	3.01	3.7	3.03	3.74	4.97	4.36
50%	1.86	1.63	1.75	2.37	3.06	2.4	3.1	4.32	3.71
90%	5.42	5.03	5.25	6.26	7.33	6.31	7.39	9.2	8.31
	10 days								
χ^2 d.o.f.	2.02	1.81	1.97	2.29	2.66	2.21	2.86	3.9 3	3.51
50%	1.41	1.2	1.36	1.67	2.03	1.59	2.23	3.29	2.87
90%	4.65	4.27	4.56	5.11	5.71	4.97	6.03	7.67	7.04
	100 days								
χ^2 d.o.f.	1.91	1.66	1.85	1.82	1.88	2.13	1.92	2.56	2.64
50%	1.3	1.06	1.24	1.21	1.27	1.51	1.31	1.93	2.01
90%	4.44	4	4.35	4.29	4.4	4.83	4.47	5.55	5.68
Significance threshold $= 4.7\sigma$									
	1 day								
χ^2 d.o.f.	2.1	2.01	2.05	2.4	3.13	2.82	2.91	4.63	3.94
50%	1.48	1.4	1.43	1.77	2.5	2.19	2.27	3.98	3.29
90%	4.77	4.63	4.69	5.28	6.46	5.97	6.1	8.7	7.69
	10 days								
χ^2 d.o.f.	1.82	1.64	1.78	1.97	2.12	2.23	2.24	3.19	2.8
50%	1.21	1.04	1.17	1.36	1.51	1.61	1.62	2.55	2.16
90%	4.29	3.97	4.21	4.55	4.82	5	5.02	6.55	5.93
	100 days								
χ^2 d.o.f.	1.75	1.66	1.8	1.8	1.87	1.94	1.9	2.27	2.28
50%	1.15	1.06	1.19	1.19	1.26	1.33	1.29	1.65	1.65
90%	4.16	4.01	4.26	4.26	4.37	4.5	4.44	5.07	5.08

Table A.1: Values for the number of degrees of freedom and 50-90% quantiles for the χ^2 distributions depending on the parameters for significance thresholds 3.5σ and 4.7σ.

Below are showed how trustworthy the values for the quantiles in these simulations, separated by durations of injected flares.

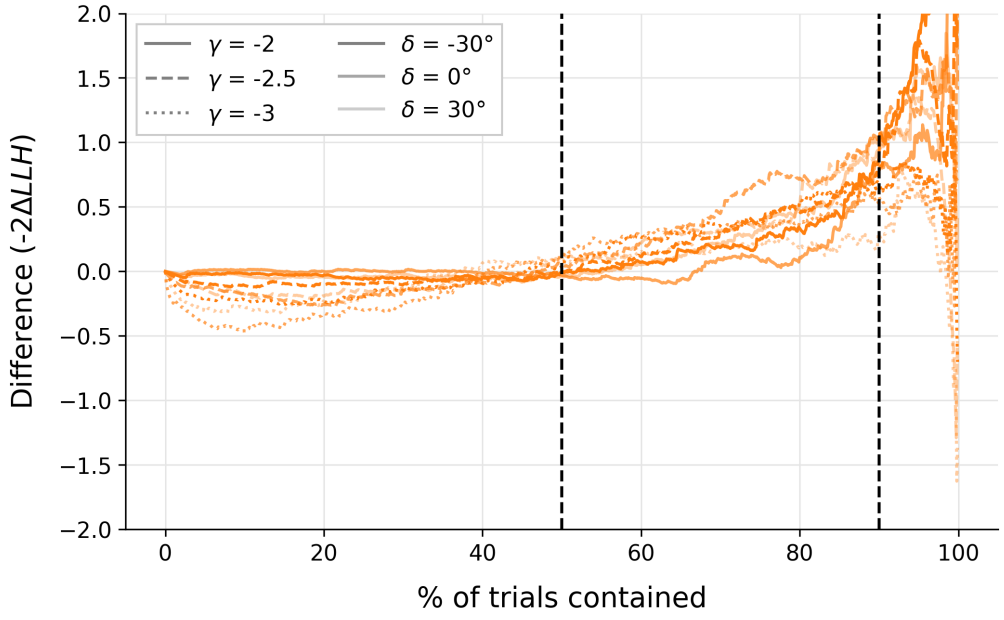


(a) Significance threshold: 3.5σ

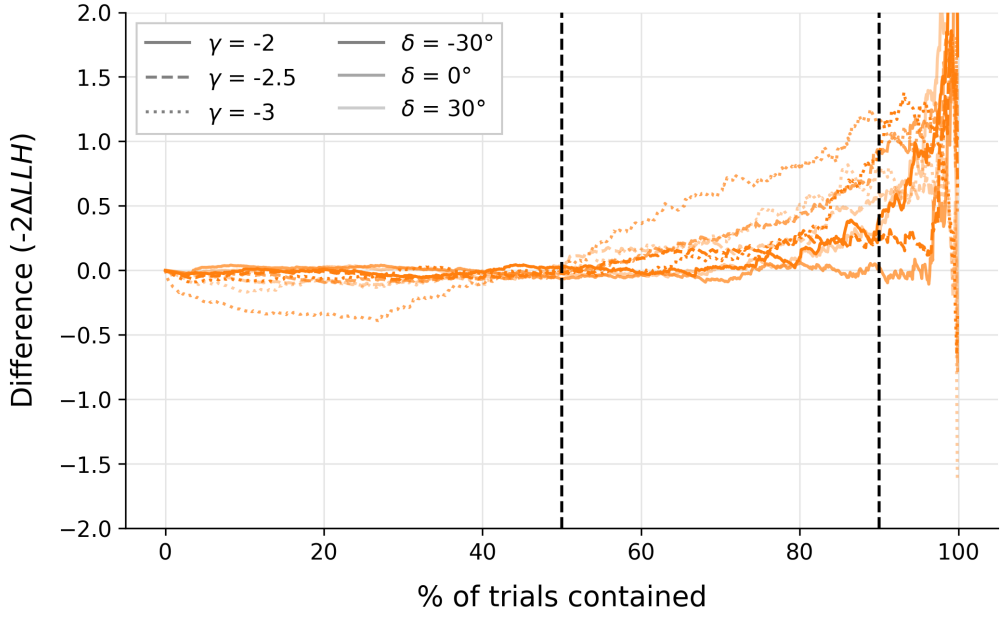


(b) Significance threshold: 4.7σ

Figure A.4: Difference between the quantiles of the results and of their corresponding χ^2 distributions for 1-day flares at different significance thresholds for the alerts.

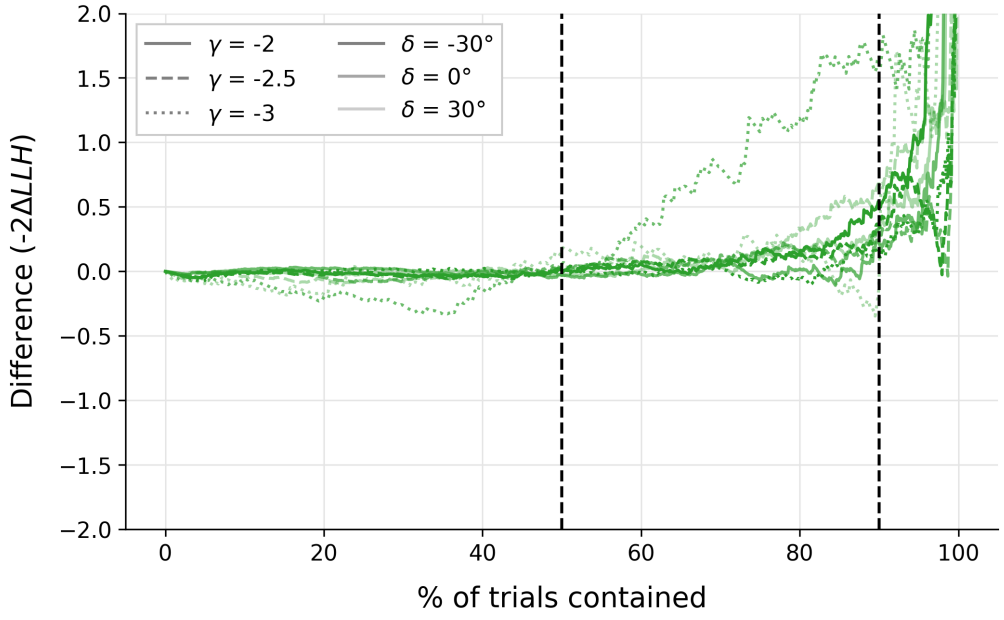


(a) Significance threshold: 3.5σ

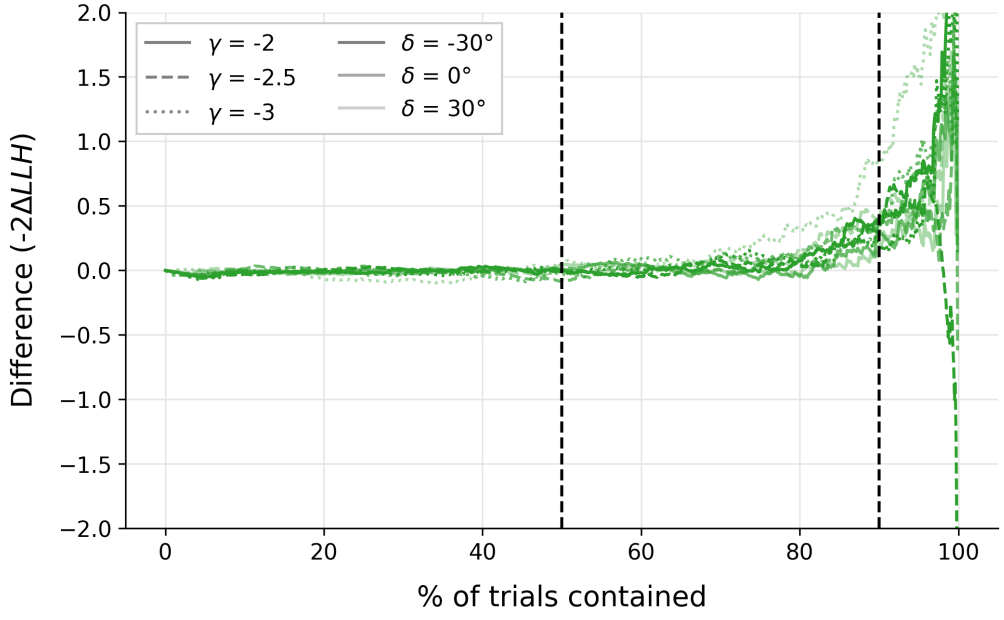


(b) Significance threshold: 4.7σ

Figure A.5: Difference between the quantiles of the results and of their corresponding χ^2 distributions for 10-days flares at different significance thresholds for the alerts.



(a) Significance threshold: 3.5σ

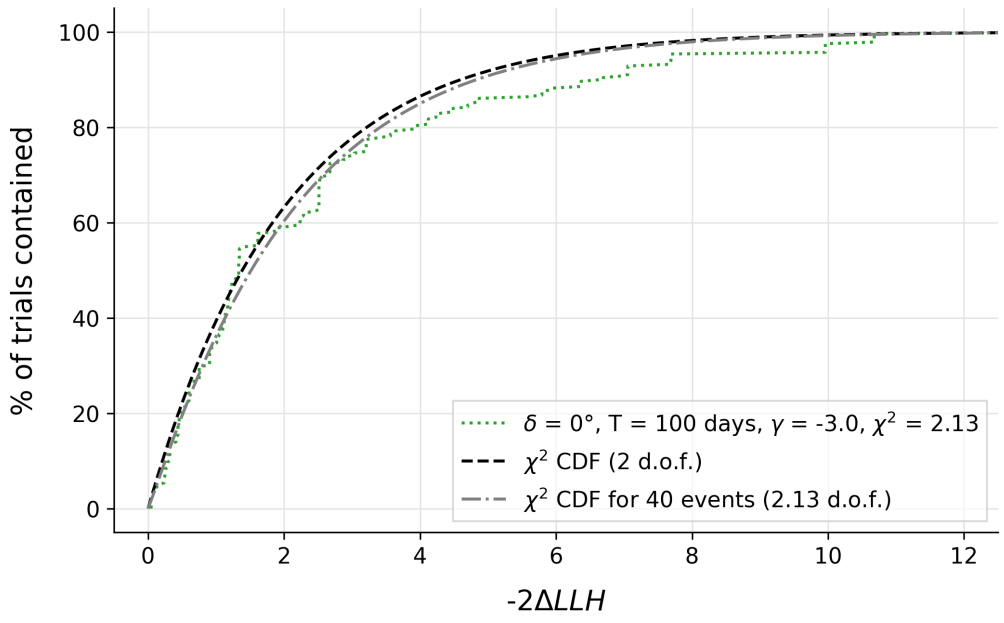


(b) Significance threshold: 4.7σ (40 injected events)

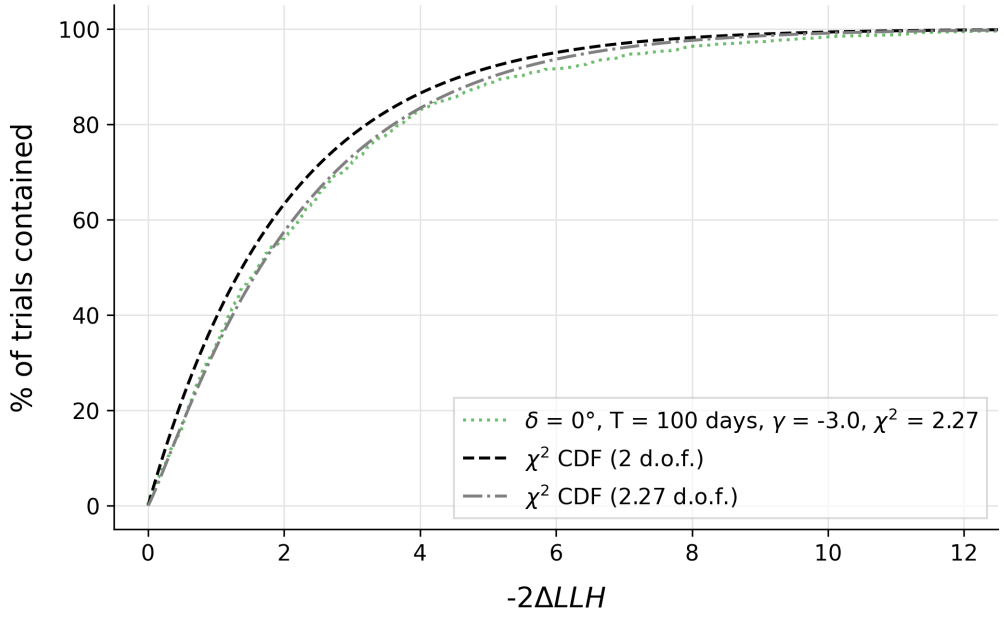
Figure A.6: Difference between the quantiles of the results and of their corresponding χ^2 distributions for 100-days flares at different significance thresholds for the alerts. For the higher threshold, the number of injected events was higher to guarantee that alerts would be raised.

A.4 Number of injected events

Below are the simulations equivalent to Figure 5.7, but with more injected events. These plots show how the algorithm only takes into account the number of events needed to surpass the significance threshold, and the time duration over which these events are spread. The higher threshold should only allow for highly localized flares, and thus be fitted by a χ^2 distribution with few degrees of freedom, but since more events were injected, less of them were necessary to trigger an alert, and they didn't need to be as localized, appearing as a shorter flare to the algorithm.



(a) 20 injected events (threshold of 4.2σ

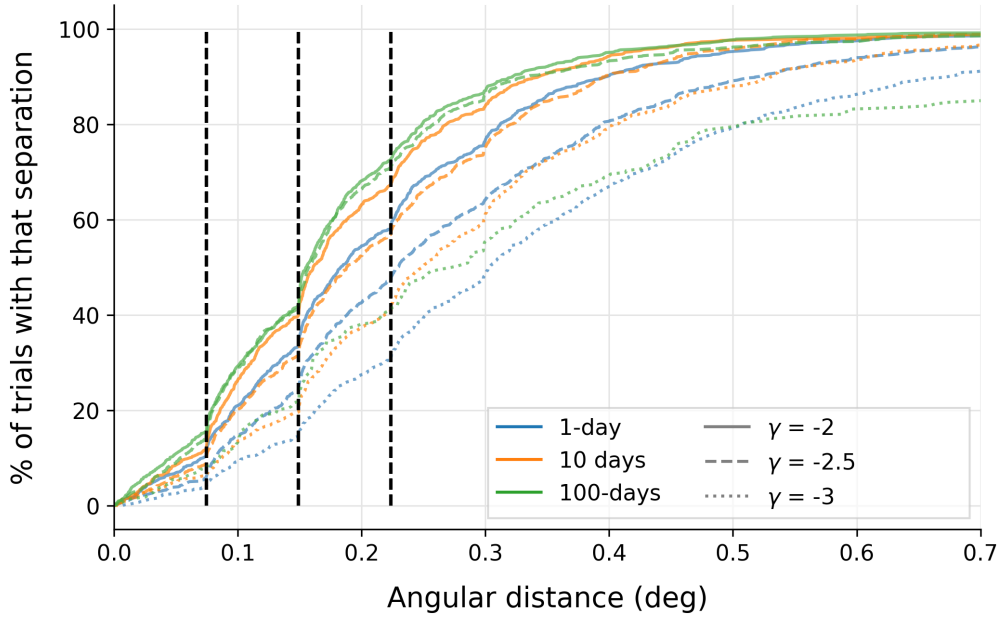


(b) 40 injected events (threshold of 4.7σ)

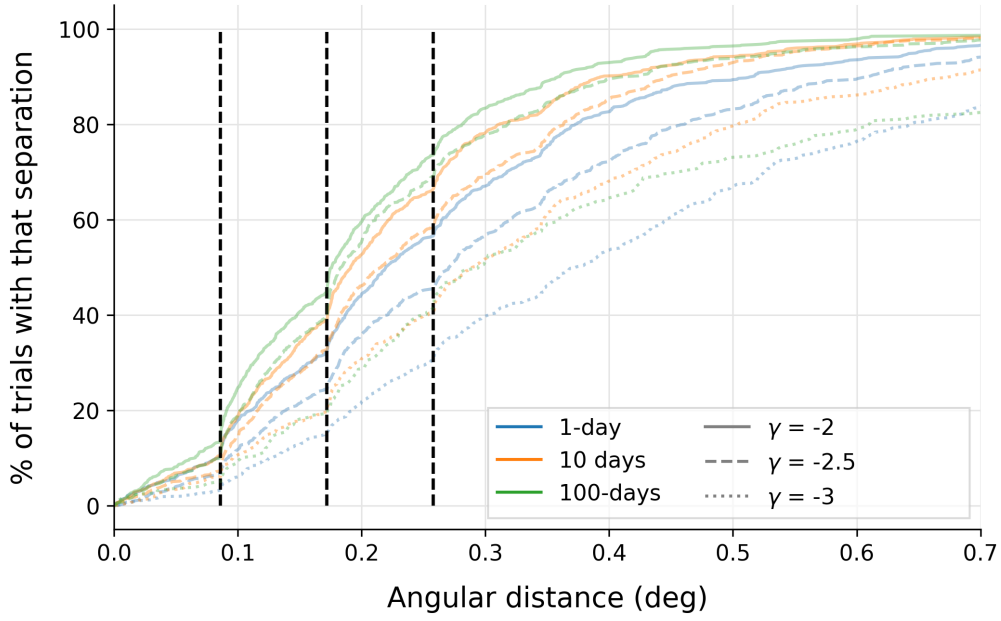
Figure A.7: Comparison between the performances of the algorithm at 30° for 100-days flares of 20, and 40 events injected with spectral index $\gamma = -3$. In the latter case, the significance threshold for the alerts was higher. The flares with 40 events are less accurately reconstructed despite their higher significance threshold, because they are denser, and thus appear as shorter to the algorithm.

A.5 Angular distance between best fit pixel and injection position

Finally, we show here, like in Figure 5.8, the CDFs for the angular separation between the injection source and the reconstructed source position at higher injection declinations.



(a) $\delta = 0°$



(b) $\delta = 30°$

Figure A.8: Cumulative Distribution Functions for the angular distance between the injection position and the best fit pixel for declinations higher than $-30°$.

References

[1] M. Ahlers and F. Halzen, "Opening a new window onto the universe with icecube," *Progress in Particle and Nuclear Physics*, vol. 102, p. 73–88, Sep. 2018. [Online]. Available: http://dx.doi.org/10.1016/j.ppnp.2018.05.001

[2] M. S. Longair, *High Energy Astrophysics*, 3rd ed. Cambridge University Press, 2011.

[3] M. Spurio, *Particles and Astrophysics. A Multi-Messenger Approach*, ser. Astronomy and Astrophysics Library. Springer, 2015.

[4] J. Aleksić *et al.*, "The major upgrade of the magic telescopes, part i: The hardware improvements and the commissioning of the system," *Astroparticle Physics*, vol. 72, pp. 61–75, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0927650515000663

[5] de Naurois, Mathieu, "H.e.s.s.-ii - gamma ray astronomy from 20 gev to hundreds of tev's," *EPJ Web Conf.*, vol. 136, p. 03001, 2017. [Online]. Available: https://doi.org/10.1051/epjconf/201713603001

[6] T. Weekes *et al.*, "Veritas: the very energetic radiation imaging telescope array system," *Astroparticle Physics*, vol. 17, no. 2, p. 221–243, May 2002. [Online]. Available: http://dx.doi.org/10.1016/S0927-6505(01)00152-9

[7] [Online]. Available: https://www.eso.org/public/images/eso1841x/

[8] V. F. Hess, "Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten," *Phys. Z.*, vol. 13, pp. 1084–1091, 1912.

[9] A. Castellina and W. Chamani, "The hecrs: Experiments and recent results," 05 2011, p. 045.

[10] K. Greisen, "End to the cosmic ray spectrum?" *Phys. Rev. Lett.*, vol. 16, pp. 748–750, 1966.

[11] G. T. Zatsepin and V. A. Kuzmin, "Upper limit of the spectrum of cosmic rays," *JETP Lett.*, vol. 4, pp. 78–80, 1966.

[12] M. Bustamante *et al.*, "High-energy cosmic-ray acceleration," 2010. [Online]. Available: https://cds.cern.ch/record/1249755

[13] P. M. Bauleo and J. R. Martino, "The dawn of the particle astronomy era in ultra-high-energy cosmic rays," *Nature*, vol. 458, no. 7240, pp. 847–851, Apr 2009. [Online]. Available: https://doi.org/10.1038/nature07948

[14] J. Babson *et al.*, "Cosmic-ray muons in the deep ocean," *Phys. Rev. D*, vol. 42, pp. 3613–3620, Dec 1990. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.42.3613

[15] I. Belolaptikov *et al.*, "The baikal underwater neutrino telescope: Design, performance, and first results," *Astroparticle Physics*, vol. 7, no. 3, pp. 263–282, 1997. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0927650597000224

[16] M. Ageron *et al.*, "Antares: The first undersea neutrino telescope," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 656, no. 1, pp. 11–38, 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900211013994

[17] A. Margiotta, "The km3net deep-sea neutrino telescope," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 766, p. 83–87, Dec. 2014. [Online]. Available: http://dx.doi.org/10.1016/j.nima.2014.05.090

[18] E. Andres *et al.*, "The amanda neutrino telescope: Principle of operation and first results," *Astroparticle Physics*, vol. 13, pp. 1–20, 03 2000.

[19] E. Waxman and J. Bahcall, "High energy neutrinos from astrophysical sources: An upper bound," *Phys. Rev. D*, vol. 59, p. 023002, Dec 1998. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.59.023002

[20] M. G. Aartsen *et al.*, "Icecube-gen2: the window to the extreme universe," *Journal of Physics G: Nuclear and Particle Physics*, vol. 48, no. 6, p. 060501, Apr. 2021. [Online]. Available: http://dx.doi.org/10.1088/1361-6471/abbd48

[21] S. Agarwal *et al.*, "Instrument design and performance of the first seven stations of rno-g," 2025. [Online]. Available: https://arxiv.org/abs/2411.12922

[22] G. A. Askar'yan, "Excess negative charge of an electron-photon shower and its coherent radio emission," *Zh. Eksp. Teor. Fiz.*, vol. 41, pp. 616–618, 1961.

[23] M. Ackermann *et al.*, "Optical properties of deep glacial ice at the south pole," *Journal of Geophysical Research: Atmospheres*, vol. 111, no. D13, 2006. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006687

[24] M. Aartsen *et al.*, "The icecube neutrino observatory: instrumentation and online systems," *Journal of Instrumentation*, vol. 12, no. 03, Mar. 2017. [Online]. Available: http://dx.doi.org/10.1088/1748-0221/12/03/P03012

[25] R. Abbasi *et al.*, "The design and performance of icecube deepcore," *Astroparticle Physics*, vol. 35, no. 10, p. 615–624, May 2012. [Online]. Available: http://dx.doi.org/10.1016/j.astropartphys.2012.01.004

[26] M. G. Aartsen *et al.*, "Measurement of the cosmic ray energy spectrum with icetop-73," *Phys. Rev. D*, vol. 88, p. 042004, Aug 2013. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.88.042004

[27] [Online]. Available: https://icecube.wisc.edu/gallery/diagrams/

[28] R. Abbasi *et al.*, "Calibration and characterization of the icecube photomultiplier tube," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 618, no. 1, pp. 139–152, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900210006662

[29] P. Askebjer *et al.*, "Optical properties of deep ice at the south pole: absorption," *Applied Optics*, vol. 36, pp. 4168–4180, 06 1997.

[30] J. R. Petit *et al.*, "Climate and atmospheric history of the past 420,000 years from the vostok ice core, antarctica," *Nature*, vol. 399, no. 6735, pp. 429–436, Jun 1999. [Online]. Available: https://doi.org/10.1038/20859

[31] M. G. Aartsen *et al.*, "Detection of a particle shower at the glashow resonance with icecube," *Nature*, vol. 591, no. 7849, p. 220–224, Mar. 2021. [Online]. Available: http://dx.doi.org/10.1038/s41586-021-03256-1

[32] T. Kintscher, "Rapid response to extraordinary events: Transient neutrino sources with the icecube experiment," Ph.D. dissertation, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2020.

[33] M. G. Aartsen *et al.*, "Search for sources of astrophysical neutrinos using seven years of icecube cascade events," *The Astrophysical Journal*, vol. 886, no. 1, p. 12, Nov. 2019. [Online]. Available: http://dx.doi.org/10.3847/1538-4357/ab4ae2

[34] R. Abbasi *et al.*, "Observation of high-energy neutrinos from the galactic plane," *Science*, vol. 380, no. 6652, p. 1338–1343, Jun. 2023. [Online]. Available: http://dx.doi.org/10.1126/science.adc9818

[35] D. Pandel, "Bestimmung von wasser- und detektorparametern und rekonstruktion von myonen bis 100 tev mit dem baikal- neutrinoteleskop nt-72 / von dirk pandel," Ph.D. dissertation, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 1996.

[36] M. Voge, "Searches for Neutrinos from Supernovae Using Cherenkov In-Ice Detectors," Ph.D. dissertation, Bonn U., 9 2016.

[37] N. Whitehorn, J. van Santen, and S. Lafebre, "Penalized splines for smooth representation of high-dimensional monte carlo datasets," *Computer Physics Communications*, vol. 184, no. 9, p. 2214–2220, Sep. 2013. [Online]. Available: http://dx.doi.org/10.1016/j.cpc.2013.04.008

[38] L. Rädel and C. Wiebusch, "Calculation of the cherenkov light yield from electromagnetic cascades in ice with geant4," *Astroparticle Physics*, vol. 44, p. 102–113, Apr. 2013. [Online]. Available: http://dx.doi.org/10.1016/j.astropartphys.2013.01.015

[39] M. G. Aartsen *et al.*, "Energy reconstruction methods in the icecube neutrino telescope," *Journal of Instrumentation*, vol. 9, no. 03, p. P03009, mar 2014. [Online]. Available: https://dx.doi.org/10.1088/1748-0221/9/03/P03009

[40] J.-D.-D. Zornoza and D. Chirkin, "Muon energy reconstruction and atmospheric neutrino spectrum unfolding with the IceCube detector," in *International Cosmic Ray Conference*, ser. International Cosmic Ray Conference, vol. 5, Jan. 2008, pp. 1275–1278.

[41] R. Abbasi *et al.*, "An improved method for measuring muon energy using the truncated mean of de/dx," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 703, pp. 190–198, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900212014234

[42] H. Cramér, *Mathematical Methods of Statistics (PMS-9)*. Princeton University Press, 1999. [Online]. Available: http://www.jstor.org/stable/j.ctt1bpm9r4

[43] R. Abbasi *et al.*, "Time-integrated searches for point-like sources of neutrinos with the 40-string icecube detector," *The Astrophysical Journal*, vol. 732, no. 1, p. 18, Apr. 2011. [Online]. Available: http://dx.doi.org/10.1088/0004-637X/732/1/18

[44] J. Lünemann, "Suche nach dunkler materie in galaxien und galaxienhaufen mit dem neutrinoteleskop icecube," Ph.D. dissertation, Johannes Gutenberg-Universität Mainz, 2014. [Online]. Available: https://openscience.ub.uni-mainz.de/handle/20.500.12030/4221

[45] M. Aartsen *et al.*, "Multimessenger observations of a flaring blazar coincident with high-energy neutrino icecube-170922a," *Science*, vol. 361, no. 6398, Jul. 2018. [Online]. Available: http://dx.doi.org/10.1126/science.aat1378

[46] E. Bernardini and D. Gora, "Very high-energy gamma-ray follow-up program using neutrino triggers from icecube," *Journal of Instrumentation*, vol. 11, no. 11, p. P11009–P11009, Nov. 2016. [Online]. Available: http://dx.doi.org/10.1088/1748-0221/11/11/P11009

[47] D. Mazin *et al.*, "Status and results of the prototype lst of cta," in *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*, ser. ICRC2021. Sissa Medialab, Jul. 2021, p. 872. [Online]. Available: http://dx.doi.org/10.22323/1.395.0872

[48] M. Ackermann, E. Bernardini, N. Galante, F. Goebel, M. Hayashida, K. Satalecka, M. Tluczykont, and R. M. Wagner, "Neutrino Triggered Target of Opportunity (NToO) test run with AMANDA-II and MAGIC," in *30th International Cosmic Ray Conference*, vol. 3, 7 2007, pp. 1257–1260.

[49] C. Boscolo Meneguolo, E. Bernardini, and S. Mancina, "Sentinel of the extraordinary: the icecube alert system for neutrino flares," in *Proceedings of 38th International Cosmic Ray Conference — PoS(ICRC2023)*, ser. ICRC2023. Sissa Medialab, Jul. 2023, p. 1500. [Online]. Available: http://dx.doi.org/10.22323/1.444.1500

[50] [Online]. Available: https://user-web.icecube.wisc.edu/~tkintscher/internal/gfu_doc/likelihood.html

[51] L. Rayleigh, "Xii. on the resultant of a large number of vibrations of the same pitch and of arbitrary phase," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 10, no. 60, pp. 73–78, 1880. [Online]. Available: https://doi.org/10.1080/14786448008626893

[52] E. Fermi, "On the Origin of the Cosmic Radiation," *Physical Review*, vol. 75, no. 8, pp. 1169–1174, Apr. 1949.

[53] G. Punzi, "Sensitivity of searches for new signals and its optimization," in *Statistical Problems in Particle Physics, Astrophysics and Cosmology (PHYSTAT2003)*, 2003, p. 79–83.

[54] J. Braun, J. Dumm, F. De Palma, C. Finley, A. Karle, and T. Montaruli, "Methods for point source analysis in high energy neutrino telescopes," *Astroparticle Physics*, vol. 29, no. 4, p. 299–305, May 2008. [Online]. Available: http://dx.doi.org/10.1016/j.astropartphys.2008.02.007

[55] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, "Healpix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere," *The Astrophysical Journal*, vol. 622, no. 2, p. 759–771, Apr. 2005. [Online]. Available: http://dx.doi.org/10.1086/427976

[56] S. S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses," *The Annals of Mathematical Statistics*, vol. 9, no. 1, pp. 60 – 62, 1938. [Online]. Available: https://doi.org/10.1214/aoms/1177732360

[57] N. M. Kiefer, "Maximum likelihood estimation (mle)," 2008. [Online]. Available: https://courses.cit.cornell.edu/econ620/reviewm5.pdf

[58] J. Dool, "Probability and statistics," *Trans. Amer. Math. Soc. 36*, pp. 759–775, 1934.

[59] S. Navas, C. Amsler, T. Gutsche, C. Hanhart, J. J. Hernández-Rey, C. Lourenço, A. Masoni, M. Mikhasenko, R. E. Mitchell *et al.*, "Review of particle physics," *Phys. Rev. D*, vol. 110, pp. 694–711, Aug 2024. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.110.030001

[60] I. King, "The structure of star clusters. I. an empirical density law," vol. 67, p. 471, Oct. 1962.

Acknowledgments

I would like to sincerely thank Elisa Bernardini for her invaluable advices and guidance during this project, and even before. Under her tutelage, I have learned about multi-messenger astrophysics and confirmed that it was the field in which I wanted to make my career, but I have also learned a great deal about what this career would look like, and she has given me many tools to succeed on this path. I am grateful for everything she has taught me and for the support she showed in my research and future endeavors.

As my co-supervisor, Diego Alberto Coloma Borja has helped me understand, assimilate, and explain the contents of my project more than anyone, and I wish to thank him for all the time he spent correcting me and sharing his experience with me. He has always accommodated me when I needed his insight, and he helped me feel in the right place.

Next, I want to address special thanks to Sarah Louise Mancina, without whom this project would not have happened, or at least not for me, Caterina Boscolo Meneguolo, who welcomed me warmly into the working group, and Michele Doro, who both listened to me and helped move my thesis forward. Their feedback on my work was instrumental in its achievement, and they motivated me to get to the bottom of what I could understand about this subject. The end result of my master studies would have been far different without them.

I want now to take the time to appropriately thank my parents. For having never failed to believe in me and feel proud of my accomplishments. For being the solid roots on which I can always rely, even when I am far away. And for having always been as attentive and as patient as I could have needed. There is no one else to whom I owe more of who I am today, and I could have never made it this far without their support. For the wisdom impossible to learn elsewhere, for all the time given, and for everything else, from the bottom of my heart, thank you.

Finally, I couldn't forget to thank my classmates, especially Viktor, Ricardo, Micole, Samir, Savina, Joya, and Nicolo. With them, I have learned and worked, but I have also shared deep

and lasting friendships, perhaps some of the best in my life. They have been there for me and understood me in moments where I really needed it, and it meant the world to me. I will treasure the memories we have created together and hope for them the very best to come.