Thesis

Time Integrated searches for Astrophysical Neutrino Sources using the IceCube Detector and Gender in Physics studies for the Genera Project

CARVER, Tessa

Abstract

This thesis presents point-like neutrino source searches using ten years of IceCube data collected between Apr.~6, 2008 to Jul.~10, 2018. These searches evaluate the significance of an astrophysical signal by looking for an excess of clustered neutrino events with energies typically above 1 TeV among a background of atmospheric muons and neutrinos. This involved optimizing the event selection for the final six years of data and applying it to: a full sky scan, a selected source catalog, a catalog population study, and stacked galactic catalog searches. The improved sensitivity of these analyses did not find new evidence for a steady state neutrino source. Upper-limits were placed on the expected flux from sources in the catalogs and an excess was identified as inconsistent with background at the level of 2.9σ at the coordinates of NGC 1068. These results motivate interest in point-like sources given future updates to the detector or a longer lifetime of data.

Reference

CARVER, Tessa. *Time Integrated searches for Astrophysical Neutrino Sources using the IceCube Detector and Gender in Physics studies for the Genera Project*. Thèse de doctorat : Univ. Genève, 2019, no. Sc. 5329

DOI : 10.13097/archive-ouverte/unige:120924
URN : urn:nbn:ch:unige-1209245

UNIVERSITÉ DE GENÈVE
Section de Physique
Department de Physique Nucleaire et Corpusculaire

FACULTÉ DES SCIENCES
Professeur Teresa Montaruli

Time Integrated searches for Astrophysical Neutrino Sources using the IceCube Detector and Gender in Physics studies for the Genera Project

THÈSE

présentée à la Faculté des sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences, mention physique

par

Tessa Carver
de
Royaume-Uni & États-Unis d'Amérique

Thèse N°5329

GENÈVE
2019

UNIVERSITÉ
DE GENÈVE

FACULTÉ DES SCIENCES

DOCTORAT ÈS SCIENCES, MENTION PHYSIQUE

Thèse de Madame Tessa CARVER

intitulée :

«Time Integrated Searches for Astrophysical Neutrino Sources Using the IceCube Detector and Gender in Physics Studies for the Genera Project»

La Faculté des sciences, sur le préavis de Madame T. MONTARULI, professeure ordinaire et directrice de thèse (Département de physique nucléaire et corpusculaire), Monsieur X. WU, professeur associé (Département de physique nucléaire et corpusculaire), Monsieur S. PALTANI, professeur ordinaire (Département d'astronomie) et Monsieur J. A. AGUILAR, professeur (Inter University Institute for High Energies, Université Libre de Bruxelles, Belgique), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 15 avril 2019

Thèse - 5329 -

Le Doyen

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

Résumé

La lumière a depuis longtemps été utilisée comme manière d'explorer l'Univers. Cependant, en 2013, une nouvelle étape a été franchie [1] lorsque l'expérience IceCube a permis d'explorer l'Univers non thermique grâce à la première observation d'un flux astrophysique de neutrinos à haute énergie [2]. Depuis, l'existence des ondes gravitationnelles a été confirmée au moyen d'observations directes [3] et il est devenu d'intérêt majeur de localiser les sources multimessagères responsables de ces émissions non thermiques. Depuis maintenant plus de six ans, le détecteur IceCube, situé au pôle Sud, collecte des données, ce qui a permis d'établir de façon solide la présence d'un flux de neutrinos astrophysiques [4]. Cependant, la majeure partie de ce flux n'a pas de source connue associée, avec pour seule piste, une source rayonnante, le blazar TXS 0506+056 [5], qui serait à l'origine d'environ 1% du flux astrophysique, laissant le reste du flux inexpliqué.

Le travail entrepris dans cette thèse consiste à mettre en évidence des sources pouvant être resonsables d'un flux astrophysique de neutrinos. Dans ce but, une amélioration de la sélection d'évènements a été effectuée afin de gagner en sensibilité; cette sélection mise à jour contient les six années de données les plus récentes, dont trois sont complètement nouvelles par rapport à la dernière analyse de sources de neutrinos intégrée en temps et couvrant la totalité du ciel [6]. La durée totale de la sélection ainsi obtenue est de dix ans, avec des données recueillies entre le 6 avril 2008 et le 10 juillet 2018. Additonnellement à ces techniques existentes, une nouvelle méthode de sélection d'évènements à variables mutliples (multivariée) a été introduite et la reconstruction de la direction des traces détectées a été améliorée. La sélection finale a été utilisée afin de chercher des sources émettant un flux constant de neutrinos. Ces recherches ont toutes pour cibles des émetteurs ponctuels de neutrinos astrophysiques dans le ciel et les méthodes utilisées mettent en oeuvre différents niveaux de détails connus à-priori à propos des sources candidates.

La première analyse discutée est une carte de l'entièreté du ciel en fonction de la probabilité que l'émission observée ne provienne que du bruit diffus. À cette fin, aucune information sur les coordonnées ou le spectre de la source ne sont requis. Ensuite, un nouveau catalogue de sources potentielles de neutrinos prometteuses, basé sur des observations de rayons gammas associés à ces sources, est proposé. La probabilité d'une émission de neutrino provenant de chaqune des sources présentes dans ce catalogue est testée individuellement. De plus, de nouveaux catalogues de candidats de sources galactiques, organisés par types d'objets et mis à jour grâce aux dernières observations de TeVCat, ont été créés. Les signaux détectés dans la direction de chaqun de ces objets ont été superposés, permettant ainsi la détection d'un signal global même dans le cas où l'émission par source

individuelle est faible. Finalement, une analyse populationnelle du catalogue basé sur les observations de rayons gammas est réalisée. Le principe de cette analyse est de considérer les résultats individuels des objets du catalogue comme un ensemble et de déterminer si, statistiquement, une telle combinaison de résultats pourrait être dûe au bruit uniquement ou non.

Aucune de ces recherches n'a permis la découverte d'une nouvelle source ponctuelle de neutrinos. L'excès le plus significatif dans le catalogue des objets individuels de l'hémisphère nord se trouve dans la direction de la galaxie Seyfert NGC 1068; le flux mesuré à ce point n'est pas consistant avec l'hypothèse du bruit seul et l'hypothèse d'un bruit accompagné d'un signal prévaut avec un intervalle de confiance final qui s'élève à de 2.9 σ. De plus, la direction de cette source est contenue dans la région autour du point le plus significatif globalement de la carte du ciel. Ceci, combiné avec la présence de trois autres excès en direction de sources candidates (incluant TXS 0506+056), résulte en une déviation globale à un niveau de 3.3 σ par rapport à l'hypothèse d'un flux complètement diffus uniquement, pour la liste des sources dans le catalogue nord.

Au vu du nombre de collaborateurs présents au sein de l'expérience IceCube, qui est menée par de grandes équipes basées dans de nombreuses universités, il est utile d'évaluer l'efficacité de cet environnement de recherche. C'est pourquoi la recherche purement astrophysique de cette thèse est accompagnée d'une étude pour le consortium Genera [7] à propos de l'environnement et des mesures prises pour les questions de genre en physique à l'Université de Genève (UNIGE). Il y a une longue histoire de sous-représentation des femmes dans les sciences dures, et cela en particulier dans les positions académiques élevées. Le but de ce travail est d'évaluer l'état actuel de la situation à l'UNIGE et d'identifier d'éventuels problèmes qui pourraient être ciblés par un plan d'action contre l'inégalité liée au genre. Les effets de telles mesures pourront ensuite être observés à travers le temps, afin d'en déterminer le succès. De plus, une journée a été organisée grâce au consortium Genera, le Gender in Physics Day (Jour du genre en physique), afin de sensibiliser les participants aux problèmes dans ce domaine. Additionnellement, le projet Gender in Physics (le genre en physique) a été alimenté par la collecte de données à propos du corps académique, obtenues grâces aux services de ressources humaines de l'université. Afin de recueillir les détails manquants, un sondage a été envoyé à l'entièreté de la faculté des sciences de l'UNIGE, afin d'obtenir des informations sur diverses variables liées à la carrière. Les données issues des services de ressources humaines et les données obtenues grâce au sondage sont analysées et comparées dans cette thèse.

Acknowledgments

I would firstly like to thank my supervisor Professor Teresa Montaruli for giving me the opportunity to undertake exciting new physics research in the form of this PhD. During this time I have had my first taste of astrophysics and independent work and Teresa has always had great ambitions for my career and pushed me to try new challenges I would have never thought I was capable of. I have learnt a tremendous amount from these experiences which I hope keep with me forever.

Secondly I am very grateful to my colleagues at UNIGE with whom I had many physics discussions. In particular, Stephanie Bron, who never lost sight of the bigger picture and her determination, hard work, and curiosity was inspiring. It was my privilege to work alongside her and also a great pleasure. Although Anastasia Barbano was a late addition to the IceCube group in Geneva, her presence has had an excellent impact already. I am immensely grateful for her generosity in sharing enlightening insights into our work in addition to kind words of support.

I would also like to thank the whole IceCube collaboration which has supported the work I undertook and spent their precious time understanding it, reviewing it, and suggesting excellent improvements. In particular I have to thank those in the collaboration who were always helpful, generous, and kind with their expertise, without whom this PhD would have taken decades longer. In particular; René Reimann who gave expert advice on my analyses, Josh Wood who ensured that the process of implementing the software was a smooth and reliable as possible, Chris weaver who was seemingly always available for a technical problem, and Thomas Kintscher who somehow managed to do it all in real time!

None of this would be possible without the unconditional support from my close friends and family. Their belief in my capacity to succeed and patience in listening to my struggles has been a light guiding me through. I could not finish the acknowledgement without mentioning Ana Barbara and Lara Jost who have always stepped in when I needed a helping hand, a hug, or a good meal. Finally, I am eternally in debt to the rock of my life, Catalina Garcia. The fact that she has even read and corrected most of the pages of this thesis is just the tip of the iceberg.

List of Publications and Proceedings

- M. G. Aartsen et al. Constraints on Galactic Neutrino Emission with Seven Years of IceCube Data. *Astrophys. J.*, 849(1):67, 2017. doi: 10.3847/1538-4357/aa8dfb
- *A Summary of Recent Updates in the Search for Cosmic Ray Sources using the IceCube Detector*. SISSA, Proceedings, 2017 European Physical Society Conference on High Energy Physics (EPS-HEP 2017). doi: https://doi.org/10.22323/1.314.0004

Contents

Introduction

There is a long history of exploring the visible universe using light. In 2013 IceCube marked a breakthrough [1] in exploring the non-thermal universe with the first observation of a high energy astrophysical neutrino flux [2]. Since then the first gravitational waves have been observed directly [3] and the interest to locate non-thermal multi-messenger sources has reached its peak. IceCube now has a well observed astrophysical neutrino flux based on over 7 years of data taken at the South Pole [4]. It has been found that sources cannot be distributed only in the Galactic plane, but they should also exist outside of the Galaxy since high energy events have been observed coming from all directions. Recently, IceCube found evidence for a flaring blazar, TXS 0506+056 [5], which can account for only 1% of the observed astrophysical flux leaving the rest of this flux currently unexplained.

The work in this thesis searches for specific sources of an astrophysical neutrino flux with a new and improved sensitivity achieved with an updated event selection. The new selection was applied to the most recent six years of data using the full detector, three of which are additional years compared to the latest published all-sky time-integrated neutrino source search [6]. The lifetime of this search is ten years from Apr. 6, 2008 to Jul. 10, 2018. The new data selection is closer in its principles to the seven year analysis by Aartsen et al. [6] for the southern hemisphere, and the selection applied to find diffuse muon tracks in the northern hemisphere by Aartsen et al. [10]. This work uses a completely new multi-variate event selection and an improved direction reconstruction. This work targets point-like steady-state astrophysical neutrino emitters in the sky. Each analysis is optimized to take advantage of different levels of prior source information in the signal hypothesis.

The first analysis discussed in this thesis is an all-sky point-source search for a single bright neutrino source determined by scanning the whole sky, which requires no prior source information. Secondly, a new catalog of motivated neutrino source candidates is described, where each source is tested individually for neutrino emission. In addition, new catalogs of Galactic neutrino source candidates sorted by object type are evaluated for their neutrino cumulative or stacked emission. The sensitivity for these catalogs, which contain many southern sources, is improved by stacking the emission in the directions of all the sources in the same catalog. The flux required per source for a discovery is then reduced. Finally, a catalog population analysis is conducted by searching for an excess

in the rate of significant points from the individual source list. The population search is sensitive to statistical deviations from the background for the entire catalog, but does not identify the sources eventually producing the significance.

Research on the scale of IceCube is conducted by large teams of people based at universities and within a collaboration it is useful to evaluate the effectiveness of this research environment. Alongside the astrophysics research for this thesis, studies were conducted on behalf of the Genera Consortium [7] into the environment and policies regarding gender in physics at the University of Geneva (UNIGE). There is a long history of women being under-represented in STEM subjects, particularly at more senior research positions. The aim of this work was to evaluate the current gender equality status at UNIGE and identify any outstanding issues which can be addressed in a Gender Equality Plan (GEP). The effectiveness of policies could be monitored over time to evaluate its success through the survey I established and conducted following my acquired expertise in data analysis. My gender in physics work was also fundamental for the organisation of a Gender in Physics Day (GiP day) to raise awareness and discover current issues and concerns in this field. In addition, the Gender in Physics status was monitored by collecting data from Human Resources covering demographic data on the staff at the university. For details that were missing from this data, a survey was launched across the entire Science Faculty at UNIGE covering a wide range of career related variables. Both data from HR and the survey is analysed and compared in this thesis.

The layout of this thesis is summarized in brief below:

- *Chapter 1*: This chapter introduces the messengers currently being used to explore the Universe and compares their various advantages and disadvantages when it comes to learning about astrophysical objects. Cosmic ray particles are introduced as charged particles constantly arriving at the Earth in a broad spectrum of energies on which we would like to understand the origin. It is here that the search for an astrophysical neutrino flux is motivated through its connection to cosmic ray accelerators which produce their high-energy flux. Hypotheses for this acceleration which would allow for these energies and strong accelerator candidates are also explained in more detail.

- *Chapter 2*: This chapter provides context for IceCube functioning as a detector in the ice at the South Pole. This begins with the principles behind neutrino detection in their interactions with the detector. These interactions are then observed via the resulting Cherenkov radiation which is monitored by Digital Optical Modules and this data is then filtered and transmitted to central servers in the northern hemisphere where further processing and reconstructions are run.

- *Chapter 3*: This chapter details the most significant changes in the event selection for the last 6 years of the sample. This involves a description of the improved angular resolution and the updated multivariate selection in the northern hemisphere. In addition the performance of this selection is evaluated and compared with the previous all-sky event selection.

- *Chapter 4*: This chapter defines the likelihood method applied to point source searches. This method allows the analysis to be sensitive to a deviation from background assuming signal originates from a point-source. The sensitivity is then further improved if the astrophysical signal demonstrates a harder energy spectra, producing more events at higher energy than the background flux. It also explains how to determine the minimum fluxes at which this analysis is expected to be sensitive and the point at which a flux would result in a discovery.

- *Chapter 5*: This chapter provides information on how the likelihood method is applied for the separate point-source searches. This includes the technical details of the all-sky scan as well as criteria for a revised individual neutrino source candidate list and the Galactic catalogs. The method applied to find a deviation from background using these results in the catalog population search is also explained.

- *Chapter 6*: This chapter summarises the results seen from the proposed neutrino source searches.

- *Chapter 7*: This chapter outlines the work studying gender in physics at UNIGE and summarizes the results from the monitoring of HR data and a faculty wide Career progress survey. The disproportionate loss of women at senior academic levels is confirmed and the results of additional variables are shown and discussed.

- *Chapter 8*: A summary of the work done in this thesis and the conclusions of the results are presented in this chapter.

Chapter 1

High Energy Astrophysics

1.1 Astrophysical Messengers

Humans have long been curious about the universe surrounding them, as documented since written history began. We began our exploration by simply looking at the visible light emitted by stars in the sky. For hundreds of years there have been many contrasting ideas for the physics outside of the Earth's atmosphere compared to what we are more easily able to measure within it. During the ptolemaic era the Earth was assumed to be flat and at the centre of the universe. These theories were followed by the theories of Copernicus proposing a heliocentric universe where objects instead orbit the Sun, marking the beginning modern astronomy. We now know that most of the universe lies beyond this solar system and other galaxies far away from our own have been observed. The great challenge in this field is figuring out how to understand objects which are so far away that we will, for the most part, never reach them in person. Even the closest planet in our solar system to the Earth is 56 million km away at the minimum distance in its orbit around the sun. These distances are so vast that they are often referred to in the time it takes light, the fastest messenger, to travel them. The diameter of our own galaxy, the Milky Way, takes over 100,000 years for light to traverse. For thousands of years the only messenger known to mankind able to carry information across such large distances was light. Studies of light have undergone radical changes along with technological discoveries, from observations with our eyes alone to sketch out maps of the sky, to improvements in the precision of these maps using lenses in telescopes and creating new detectors sensitive to wavelengths outside the visible part of the spectrum or even able to measure single photons. Light from the lowest to the highest known parts of the energy spectrum is still observed today as a powerful tool to understand the universe.

However, we have now discovered a range of other fundamental particles such as electrons, positrons, and protons, which are also able to traverse vast distances across the

universe to be detected by a variety of experiments on Earth. In fact the high energy particles arriving at the Earth's atmosphere were the best way to study the growing particle zoo until the construction of ground based particle accelerators. This is demonstrated by the fact that up to 20 particles were discovered up to 1959 using these high energy particles from space. Just as mapping the sky at different energies of the photoelectric spectrum is able to provide important information on the electromagnetic behaviour of observed astronomical objects, remapping the sky using a different messenger altogether could provide a totally new perspective in the universe and the interactions taking place within it.

The work in this thesis concerns the physics of high-energy messengers far above the visible part of the electromagnetic spectrum ($> 10^9$ eV). At this energy range the objects observed are found outside of, not only our atmosphere, but also the solar system. At the highest energies we expect the astrophysical sources to lie even outside of our Galaxy. To learn about objects so far away we must utilize all the astrophysical messengers available as each one is optimal at providing different source information.

1.1.1 Cosmic Rays

To begin our journey in understanding the theories for production of an astrophysical neutrino flux we must first discuss the presence of charged particles known as Cosmic Rays (CRs) penetrating the Earth's atmosphere from outside our own solar system. In this section we will go over what CRs are, how they have been observed over time, what processes could allow for these observations, and therefore which objects we expect to be strong candidates for CR sources.

Cosmic ray physics began in the early 1900s when electroscopes were observed to discharge even when kept entirely in the dark and away from known sources of radioactivity. The reasons for such discharge were not understood until the 1920s, when manned balloons were used to collect information on atmospheric ionization as a function of height [11]. Contrary to the hypothesis of an ionization source on Earth, it was found that the average ionization rates increased at high altitudes. This was the first piece of evidence for the source to be located outside of the Earth's atmosphere. It was not until 1929 that cloud chambers were used to demonstrate how CRs initiated showers of charged particles [12]. In 1939, a collection of separated detectors was used to discover the full extent of such showers with millions of ionized particles [13] when implementing a separation distance more than 100 m. At these dimensions the original particles must have more than 10^{15} eV when entering the atmosphere. The challenge became to determine how astrophysical sources could accelerate particles to such high energies.

Energy Spectra

For close to a hundred years the composition and spectra of cosmic rays has been studied. The spectra has now been characterised up to energies of 10^{20} eV as seen in Figure 1.1 by multiple different experiments. At the lower energy end of the spectra the CRs can be measured directly from satellites or balloon experiments which detect the particles before they can interact with the atmosphere. The flux of cosmic rays can be estimated as a power-law with spectrum described by:

$$N(E)dE \propto E^{-x}dE \tag{1.1}$$

Where N is the number of CR particles, E is the CR energy and x is the spectral index which changes with energy. The equation 1.1 demonstrates how the CR flux normalization drops rapidly at higher energies.This spectrum then steepens at what is called the *knee* at about 10^{15} eV, which is referred to as a softening of the spectrum. The energy associated with the *knee* in the spectrum is dependent on the charge z of the CR particle as this will dictate the magnetic deflection and acceleration from a magnetic field of strength B within a region with a radius R. In order to make observations at energies above the knee a large effective area and long observation times are required to measure enough events. For such large detector areas, up to hundreds of metres, the detectors must be ground based and the CRs are measured indirectly through the particles produced upon interacting with the atmosphere. This is done by measuring the cascade particles or the radiation they emit as they ionize the molecules around them.

The CR energy spectrum is observed up to 10^{15} eV with a flux $\propto E^{-2.7}$. Here the spectral index shifts from 2.5-2.7 to 3.1. At even higher energies there is another prominent spectral feature where the spectrum *hardens* flattening out at energies above $10^{18.5}$ eV. This region is referred to as the *ankle*. These differences in the spectrum are often attributed to different particle origins and physical processes. It is thought that the knee demonstrates the energies at which the galactic magnetic field is no longer strong enough to contain CRs. At lower energies the CR gyro-radius is smaller than the width of the Galaxy such that the CRs are highly likely to be bent or scattered back into the Galactic plane without escaping. It is also supposed that galactic CR sources begin to reach a limit in terms of energy they can accelerate particles to. This would explain the steeper fall in the flux as a function of energy above the *knee*. The hardening of the spectra at the *ankle* would then signify where the flux from extra-galactic sources begins to dominate. Although the normalization of the total extra-galactic flux observed at Earth is lower than the galactic flux, the harder nature of the spectrum allows for the flux above 10^{19} eV to make a

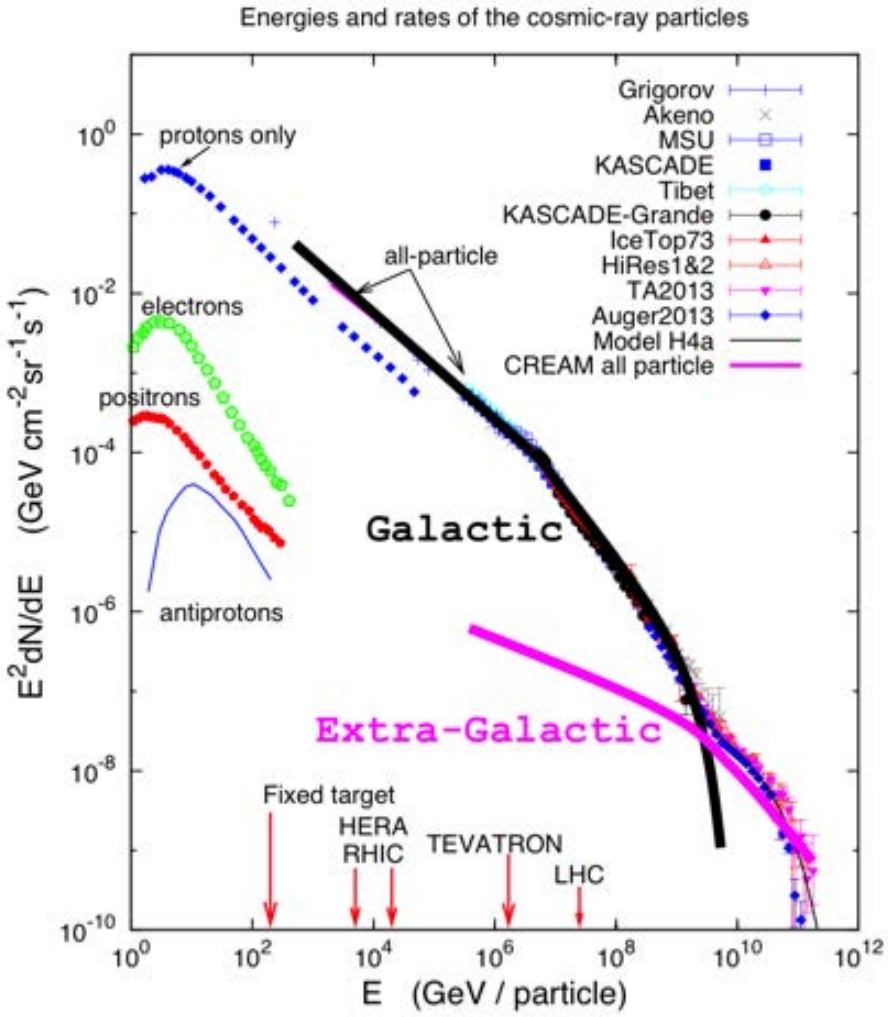


FIGURE 1.1: Cosmic ray (CR) spectrum as observed by various experiments (see legend). Each experiment targets different types of CR particles—primarily atomic nuclei—and they range from satellites to ground based ones, for lower to higher energies respectively.

significant detectable contribution to the total observed CR flux. This is likely due to the more powerful acceleration processes occurring in certain objects outside of the Galaxy. After the *ankle* in the spectrum the rate of cosmic rays has a fairly sharp cut-off at energies of approximately 10^{20} eV which is largely measured by air showers arriving at the Pierre Auger Telescope [14].

Greisen-Kuzmin-Zatsepin (GZK) cut-off

The GZK cut-off describes a cut off-energy for CR particles around $10^{20}\ eV$ depending on the particle's mass number—heavier particles have a higher cut-off, up until iron nulei. The GZK cut-off is due to interactions between the CR hadrons and the ambient photons from the Cosmic Microwave Background (CMB)[15, 16]. At these energies, even in the rest frame of the cosmic ray, the CMB photons have relativistic energies. For photons above a threshold energy $\epsilon_t = 200\ MeV$, they begin to interact with the CR protons producing pions in the following reactions:

$$\gamma + p \rightarrow \Delta^+ \rightarrow n + \pi^+$$
$$\gamma + p \rightarrow \Delta^+ \rightarrow p + \pi^0 \rightarrow p + \gamma + \gamma$$
$$\gamma + p \rightarrow \Delta^+ \rightarrow p + N\pi$$

Charged pions produced in this interaction, decay into high energy muons and muon neutrinos. Given a high enough rate or a large enough effective area, the resulting neutrinos are stable particles and therefore should be detectable on Earth. In order for photons to reach energies above the threshold in the rest frame for the CRs, the CRs require energies above 5×10^{19} eV. This energy is below the highest energy showers measured which peak at around 3×10^{20} eV. Both the Auger Observatory[17] and the Telescope Array[18] have observed a suppression in CRs above these energies which could be attributed to the GZK cut-off.

An alternative explanation for the observed suppression is provided where high energy CRs are composed of heavier nuclei (closer to iron)[19], where the GZK cut-off would not be responsible. Although photonuclear interactions still occur, the photons in the rest frame of the CR start to reach the binding energies per nucleon of the CR nuclei; the high energy photons are absorbed, leading to the excitation and ejection of nucleons, and beginning the process of nuclear disintegration. The expected path-length from this process is shorter than that from pion production, and it can be estimated that the disintegration cross section would be greater than the cross section in photo-pion production. These estimates would result in a cut-off lower than $3\times10^{20}eV$, even for iron nuclei. Therefore, at the highest

energies, the over-all cut-off effect is dependent on the CR composition; given that these events have been measured by shower arrays on the ground, the composition is difficult to determine and is not well known at the highest energies [20, 21]. As a consequence, the cause for the CR cut-off is unknown although current measurements by Piere Auger and TA are compatible with the GZK cut-off. Only the photo-pion production would lead to a corresponding ultra-high energy neutrino flux which—if measured—could be evidence of which suppression process is taking place. Until now there has been no confirmed signal from IceCube of a cosmogenic neutrino flux resulting from the GZK cut-off [22] and the question of the origins of the flux suppression remains unanswered.

1.1.2 Acceleration Mechanisms

Although still unknown, there must be a mechanism for accelerating CR particles to ultra-high energies. The general principle of acceleration for these charged particles relies on electric and magnetic fields. The general expression for acceleration of a charged particle is

$$\frac{d}{dt}(\gamma m\mathbf{v}) = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \tag{1.2}$$

for a particle of mass m and charge e with a velocity $\mathbf{v}$, in the presence of electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$, respectively; γ is the Lorentz factor. This method of acceleration is usually limited since powerful electric or magnetic fields are rapidly diminished by the free movement of charged particles. High electric or magnetic fields usually are not stable as a static field but instead vary in time.

Second Order Fermi Acceleration

In 1949 Fermi proposed a mechanism by which charged particles could be gradually accelerated stochastically up to the extreme energies observed in CRs[23]. The mechanism could take place in colliding magnetized gas clouds, which act as reflection surfaces (magnetic mirrors) for the charged particles. It is assumed that both of the walls from the colliding clouds of gas move randomly with an overall typical velocity V and contains particles randomly scattered isotropically.

The CR particles are assumed to strike the *wall* at an angle θ, and the wall itself is sufficiently massive for its velocity V to remain unchanged by the collision. In the rest frame of the wall, the CR particle energy E_{cr} is conserved and the CR momentum p_{cr} is reversed. Back in the observer's frame of reference, the CR energy becomes:

$$E''_{cr} = \gamma_V(E'_{cr} + Vp'_{cr}), \text{ where } \gamma_V = \left(1 - \frac{V^2}{c^2}\right)^{-1/2} \tag{1.3}$$

By expanding Equation 1.3 as a function of the collision angle θ and averaging the change in energy ΔE as a fraction of the original energy, it can be shown that:

$$\langle \frac{\Delta E}{E} \rangle = \frac{8}{3} (\frac{V}{c})^2. \tag{1.4}$$

This model results in the CR energy increasing with $\left(\frac{V}{c}\right)^2$ per collision which is why it is referred to as Fermi second order acceleration. Since the energy gain then increases with particle velocity for each collision, the CR energy increases exponentially. Measurements then suggest that the overall CR spectrum remains relatively constant with time and so it can be assumed that there is a steady state solution to resolve the produced CR flux as a function of energy. The conclusion on this model is a final number of CR events N with energy E is given by:

$$N(E) \propto E^{-x}, \tag{1.5}$$

a power-law spectrum where x depends on the mean free path of the CR particles in the acceleration region between the 2 walls, the typical escape time and the speed of the walls.

This solution has been shown to have major discrepancies when applied to our current knowledge of interstellar cloud collisions. Firstly, for this model to work efficiently and accelerate particles from low energies, the cloud collisions would have to occur almost relativistically; in fact, these collisions are significantly slower than the speed of light. Secondly, the energy gain—which comes from CR collisions—should be maximized in order to achieve high energies; this is done by maximizing the number of collisions. However, in the interstellar medium, the mean free path is on the order of 0.1 pc, which means there are only a few collisions per year. This slow collision rate that would require long escape times to amass the energy observed in the CR spectrum. Thirdly, there is the problem of energy losses from ionisation: if particles are not injected with initial energies above the peak in ionisation losses, then they must have an extremely rapid acceleration in order to exceed that of the ionisation losses; this is commonly known as the *injection problem*. Finally, there is nothing in the theory with motivates the specific spectral index of the power-law spectrum to correspond to the measured spectrum; as such, there is no reason why the acceleration processes in many different types of objects would result in the same spectral index. In this sense it is difficult to prove that this model applies to astrophysical CR sources.

First Order Fermi Acceleration

As discussed above, second-order Fermi acceleration poses several problems when used to explain the CR spectrum at high energies. A more compatible model—first order Fermi

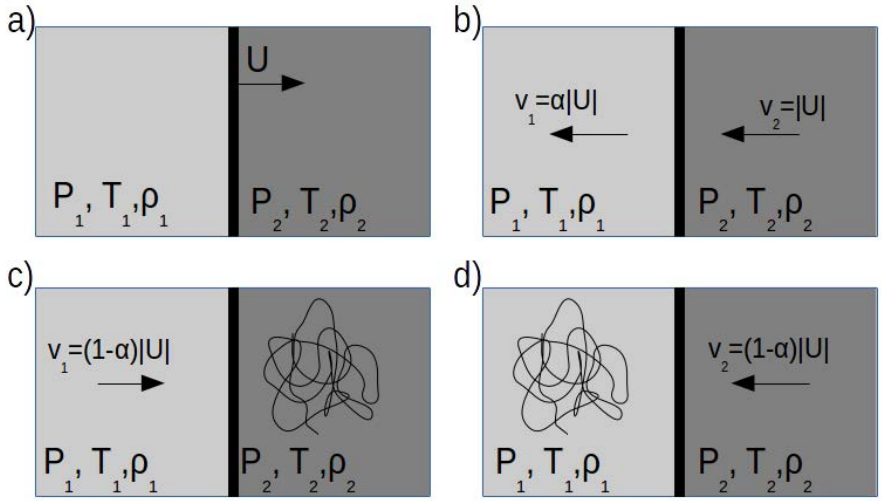


FIGURE 1.2: Diagram of diffusive shock acceleration where a shock wave with speed U travels through a relatively stationary gas with pressure P_1, temperature T_2 and density ρ_2 *upstream* is on the right and *downstream* is shown on the left. b) demonstrates the same scenario in the rest frame of the shock wave where the preceding gas is moving in the direction of the shock with a fraction of its velocity $1 - \alpha$. c) is the same scenario again in the rest frame of the surrounding gas. d) is the same scenario in the rest frame of the gas following the shock wave.

acceleration—based on diffuse shock acceleration has been applied to the problem[24].

This alternative model requires an ultra-sonic shock wave to propagate through the surrounding interstellar medium. The acceleration process occurs as particles pass from one side of the shock to the other, gaining energy each time as depicted in Fig. 1.2. Assuming conservation of mass across the shock, the product of the gas density ρ and speed v should be equal on both sides of the shock wave. If we assume the mean velocity of the gas surrounding the shock wave is 0 relative to the velocity of the shock U then matter is arriving at the shock front from upstream with a velocity v_2 where $|v_2| = |U|$. Since the shock is moving faster than the local speed of sound, the matter *downstream* behind the shock front is moving at a fraction α of this speed. In the case of a strong shock and assuming fully ionised gas on both sides, the ratio of the densities as depicted in Fig. 1.2 becomes $\frac{\rho_1}{\rho_2} = 4$ which leads to $v_1 = \alpha v_2 = \frac{1}{4} v_2$.

For this model it is assumed that, although the shock wave is not travelling close to the speed of light, the individual charged particles being accelerated have sufficient energy to be considered relativistic. It is also assumed that once a particle passes the shock to the gas on the other side it is quickly scattered such that in the rest frame of that gas where the particle directions are isotropic. In the scenario where a particle with initial energy E

crosses the shock into the downstream side which has a speed $V = \frac{3}{4}U$ relative to the gas, the downstream energy becomes:

$$E' = \gamma_V(E + p_x V), \text{ where } E = pc \text{ and } p_x = \frac{E}{c}cos(\theta) \tag{1.6}$$

Where θ is the angle at which the particle approaches the shock. By making the assumption that the shock wave is travelling much less than the speed of light then $\gamma_V \approx 1$ and therefore:

$$\frac{\Delta E}{E} = \frac{V}{c}cos(\theta), \tag{1.7}$$

Considering θ, the average energy gain from crossing the shock can be shown to be:

$$\langle \frac{\Delta E}{E} \rangle = \frac{2}{3}\frac{V}{c}. \tag{1.8}$$

From Figure 1.2 it can be seen that, whether the CR particle is in the rest frame of the upstream or downstream side of the shock, the other side is always approaching at a speed of $1 - \alpha V = \frac{3}{4}V$. High energy particles crossing the shock from one side are quickly and isotropically scattered without energy loss before eventually returning to the original side; the resulting average gain in energy becomes $\frac{4}{3}\frac{V}{c}$.

The total flux of of particles crossing the shock f_s given an average particle speed $\frac{c}{4}$ and a particle number density n would be $nc/4$. However the gas downstream of the shock front is moving away from the shock front in the frame of the shock at a speed of $v_1 = \frac{1}{4}U$. Thus the flux advected away f_a is $nv_1 = \frac{n}{4}v_2$. The fraction of the flux which is lost due to advection is then :

$$\frac{f_a}{f_s} = \frac{nv_1}{nc/4} = \frac{nU/4}{nc/4} = \frac{U}{c} \tag{1.9}$$

Which can be thought of as the probability P of a single particle of N particles escaping the shock after a collision. In the event that a particle doesn't escape, the increase in energy each time it returns back to its original frame can be written as $\beta = E''/E_0$, such that after k collisions:

$$E = E_0\beta^k \text{ and} \tag{1.10}$$

$$N = N_0P^k, \text{ hence} \tag{1.11}$$

$$N = N_0\Big(\frac{E}{E_0}\Big)^{ln(P)/ln(\beta)} \tag{1.12}$$

Using the expressions for P and β found above and Taylor expanding assuming $U << c$ the ratio of the logs can be found as :

$$ln(P) = ln(1 - \frac{U}{c}) = -\frac{U}{c} \tag{1.13}$$

$$ln(\beta) = ln(1 + \frac{4V}{3c}) = \frac{4V}{3c} = \frac{U}{c} \tag{1.14}$$

$$\implies \frac{ln(P)}{ln(\beta)} = -1 \tag{1.15}$$

The final differential particle energy spectrum from Equation 1.10 then becomes:

$$N(E)dE \propto E^{\frac{ln(P)}{ln(\beta)}-1}dE. \tag{1.16}$$

$$\implies N(E)dE \propto E^{-2}dE. \tag{1.17}$$

Although there is still an undeniable discrepancy, this begins to approach the spectrum of $E^{-2.5}$ observed by cosmic ray experiments in Fig. 1.1. It is a feasible explanation as it does not require extreme wave front velocities, only a strong shock containing scattered relativistic charged particles; these conditions are already fulfilled by various sources, including Supernova Remnants (SNR), Active Galactic Nuclei (AGN) and components of extragalatic radio sources. This mechanism, although not exactly matching the observations, goes a way towards explaining why a reasonably fixed CR spectrum across large ranges in energy has been observed, despite likely being produced by a range of different sources in a variety of environments.

The limiting factor for the maximum particle energies this process can produce in these sources is the lifetime during which the shock wave can accelerate the charged particles before either they escape or it the shock has considerably decelerated. The origin behind this maximum lifetime before an energetic CR particle escapes the shock region comes from the maximum energy $\mathcal{E}_{max}$ for an accelerator with a magnetic field of strength B and size L. This maximum energy can be estimated assuming L as the Larmor radius for a particle of charge q in a magnetic field B such that $\mathcal{E}_{max} = qBL$. This limitation is known as the *Hillas' Condition* (plot in Section 1.4) and can be used to approximate the maximum CR energy for different objects.

If certain well observed objects of known type were tied to CR emission, this could help clarifying the circumstances under which ultra-high energy cosmic rays (UHECR) are produced, and hence which model could be favoured. Since cosmic ray particles themselves are charged they are easily deflected by ambient or emergent magnetic fields in the universe; this deflection is particularly strong at lower energies and renders the overall CR emission as almost isotropic, destroying any information about the direction of its origins.

The magnetic deflection becomes less effective at very high energies and the deflection is expected to be dominated by a regular galactic magnetic field. There are currently many possible models for the galactic magnetic fields which would have a significant effect on the overall CR deflection event at EeV energies, but it is estimated that the deflection could be on the order of 3° [25]. Although an anisotropy has been observed in the CR flux at these energies [26], with such a large deflection and so few sources it is difficult to correlate these high energy events with specific astrophysical sources. For these reasons, there continues to be a lot of research using secondary particles—expected to be produced near the acceleration site—as CR messengers to provide more information on CR sources.

1.2 Photons

Historically, photons are the oldest messengers used to observe the universe in the visible spectrum and they are still a key ingredient in experimental astrophysics. There is a large spectrum of energies observable in the electromagnetic spectrum from 10^{-6} eV in radio waves up to 10^{12} eV gamma rays. Different energy regions can correspond to different physical processes and can also provide different spatial resolution in terms of locating astrophysical sources. Although other astrophysical messengers, such as CRs and gravitational waves, have been observed they have very large angular errors in direction. This makes it difficult to associate these signals with specific known objects in the sky and so until now most known astrophysical objects are monitored using photons. Therefore the electromagnetic observations are used to motivate source coordinates and other source properties. When choosing source candidate coordinates, if the object is expected to be point-like compared to the resolution of an analysis, then observations from the most precise electromagnetic wavelength are assumed as the true coordinates. If the object is extended then the origins for different particle energies could vary. In this case it is often advisable to use coordinates from observations measuring photons expected from the same physical processes but with optimal angular resolution. In the case of searches for CR sources and or neutrino source candidates this requires information from the highest energy photons which, in effect, corresponds to gamma-rays starting at about 100 keV. In practice before trying to observe astrophysical objects at certain energies it is important to understand possible sources of background. The Sun itself has a magnetic field that can act as a gamma-ray source, therefore it is helpful to optimize searches for radiation above energies which local sources can accelerate particles to when searching for astrophysical objects; in practice, the threshold is 1 GeV—the energy of particles accelerated by solar flares— and above it there are various radiative processes that can result in a gamma ray flux, including leptonic and hadronic processes.

1.2.1 Leptonic Radiative Processes

Leptonic radiative processes are those which originate from interactions of high energy leptons, typically electrons. There are many well-observed mechanisms for electrons to produce high energy photons, and the predominant one depends on the energy range in question. Typically up to GeV energies, synchrotron radiation dominates the observed electromagnetic spectrum depending on the maximum energy achieved by the particle accelerator in question, which corresponds to a maximum electron energy of 1-10 PeV. Above these energies observed gamma-rays typically originate from energy losses from Compton-Scattering, pair-production, or pion decay. An example of such a Spectral Energy Distribution (SED) for the Crab nebula can be seen in Figure 1.3 formed from experimental measurements at different wavelengths. The first and second peaks are associated with synchrotron emission and Compton scattering, respectively.

Understanding the signal in a detector of high-energy gamma-rays requires insights into the processes in which they lose energy. These insights are used to extrapolate information about their origins. One of the many gamma-ray telescopes whose measurements are used to drive searches for potential neutrino candidates is the Fermi Large Telescope Array (Fermi-LAT). Fermi-LAT is one of the two instruments on board of the Fermi satellite. It is sensitive to gamma rays of energies ranging from $\sim$ 10 MeV to $\sim$ 1 TeV through electron positron pair production. A γ-ray with energy larger than $2m_e$ can produce an electron and a positron upon interacting with an electron or nuclei in matter. Observations of photons in this energy range is an effective tool for measuring the high-energy emission of extra-galactic and galactic objects. Fermi-LAT is able to provide measurements of different astrophysical sources in the γ-ray energy range with angular resolution of $< 3.5^\circ$ at 100 MeV and $< 0.15^\circ$ above 10 GeV, allowing the observations to be well correlated with objects already observed at lower energies. At 1 GeV Fermi-LAT has a high enough event rate that it is even able to measure the variability of the observed sources on the time scale of a day and therefore to form light curves.

However due to the limited size of the detector (from the constraints on space and weight hosted on the satellite) and the flux decreases according to a power law, the detected number of photons becomes too low above $\sim$ 300 GeV to be observable. Meanwhile, the scintillators, which detect pair-production, become saturated at such high energies. At this point high energy photons can be observed individually and detected indirectly through electromagnetic showers upon interacting with the atmosphere. The Cherenkov light produced by electromagnetic showers can be detected by arrays of Imaging Air Cherenkov Telescopes. These are sensitive to photons from a few hundreds of GeV up to hundreds of TeV.

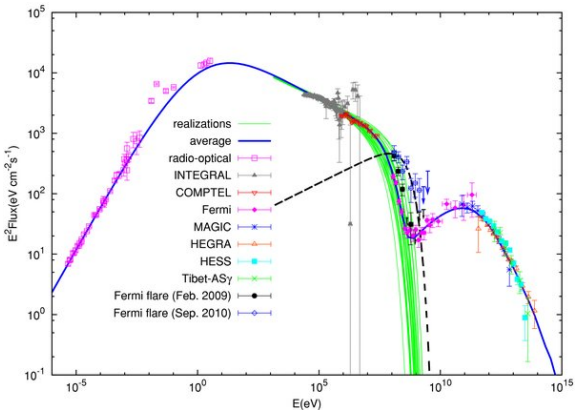


FIGURE 1.3: Multi-wavelength gamma ray spectrum of the Crab Nebula taken from Yuan et al. [27].

1.2.2 Hadronic Radiative Processes

Gamma rays at very-high energies can originate from both leptonic and hadronic interactions. Photons from hadronic interactions should be directly related to neutrinos from the same sources as they are produced together. In sources where beams of protons are accelerated to high energies, the protons can interact with hydrogen atoms, for example in the galactic disk or a supernova shock; the result is the following interaction:

$$p + p \rightarrow [\pi^0 + \pi^+ + \pi^-] + X, \tag{1.18}$$

where X is the associated nuclei or hadron depending on the pions produced. On average there is equal production of charged versus neutral pions. The magnetic fields expected in a CR accelerator would trap hydrogen and charged particles as well as radiation, such as Ultra-Violet photons, in some active galaxies. Accelerated charged particles may then interact with these photons:

$$p + \gamma \rightarrow \Delta^+ \rightarrow \pi^0 + p, \tag{1.19}$$

$$p + \gamma \rightarrow \Delta^+ \rightarrow \pi^+ + n. \tag{1.20}$$

The high-energy charged hadrons are expected to interact with the local matter and radiation to produce charged and neutral pions. In turn, these pions—being unstable particles—decay into messenger particles which depend on the pion type: neutral pions decay into gamma rays, that is, $\pi^0 \rightarrow \gamma + \gamma$ as there is no charge to conserve, whereas charged pions decay as $\pi^+ \rightarrow \nu_\mu + \mu^+ \rightarrow \nu_\mu + (e^+ + \nu_e + \bar{\nu}_\mu)$. The production of π^0 from hadronic interactions around a CR accelerator would therefore produce a gamma-ray flux. Gamma-rays,

unlike CRs, are uncharged and so are expected to travel straight from their source to the Earth allowing information to be extracted on their source direction; the γ-ray energy also provides information on the primary particle energy for a known production mechanism.

The disadvantage of using γ-rays as astrophysical messengers stems from their ability to interact electromagnetically. Firstly, gamma-rays, just like CRs, can interact with the ambient radiation in the universe undergoing pair-production in the form:

$$\gamma + \gamma \rightarrow e^+ + e^-. \tag{1.21}$$

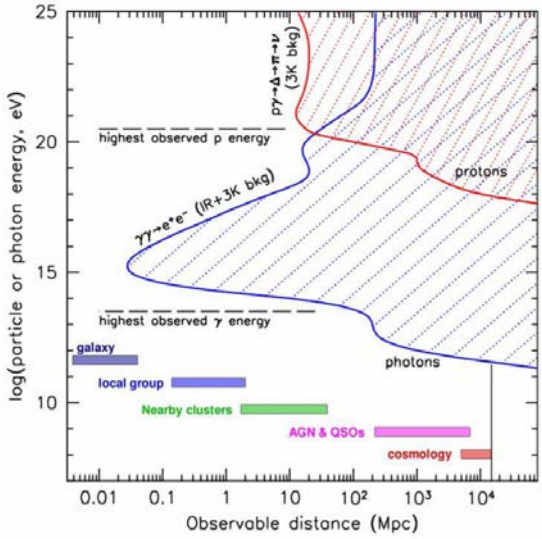


FIGURE 1.4: Event horizon for different messenger particles as a function of particle energy. This is compared to typical distances of astrophysical objects to demonstrate which objects are best seen with which messenger. Taken from Chen and Hoffman [28].

When γ-rays start interacting with the CMB, an event horizon is established where γ-rays above 300 TeV undergo pair-production with a mean free path $\lambda_{\gamma\gamma}$ given by:

$$\lambda_{\gamma\gamma} = \frac{1}{\sigma_{\gamma\gamma} n_{CMB}}, \tag{1.22}$$

where $\sigma_{\gamma\gamma}$ represents the cross section for this interaction and n_{CMB} the density of soft photons, i.e the CMB. Figure 1.4 shows this horizon for γ-rays and protons as compared to the typical distances of astronomical objects. The curve which represents protons (red) provides information on the distances which would not be observable by CRs due to the reaction producing GZK cut-off (see Eg. 1.20). The blue curve for γ-rays shows the equivalent horizon for gamma-rays; the effect is that at energies above approximately 300 TeV

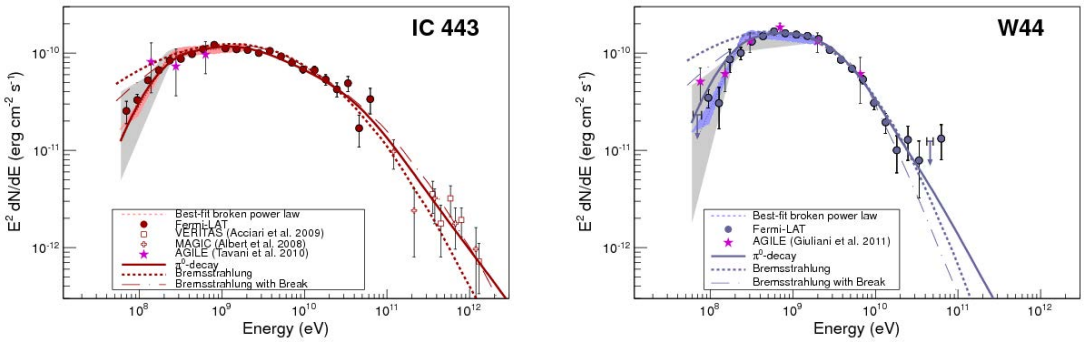
the γ-rays would have a mean free path so short that we would not expect to observe those produced far outside our Galaxy. At TeV energies the mean-free path is of the order of some hundreds of Mpc, corresponding to the closest extra-galactic active galactic nuclei distances. The photon horizon from pair-production with the CMB means that photons cannot be used as suitable messenger to provide information about the high-energy tail of the SED of extra-galactic sources beyond few hundreds of Mpc. For Galactic sources, while absorption of photons can still happen inside sources, photons are much less attenuated during their propagation to Earth and the only limitation comes from the exposure of detectors.

Astrophysical accelerators can accelerate both electrons and charged nuclei in their magnetic fields, rendering the observed γ-rays as potential products of both leptonic and hadronic interactions. Characterizing a source by placing limits on how leptonic or hadronic it is can be done by fitting the observed flux with the predicted spectra resulting from the different mechanisms. The spectra were measured in γ-rays for two galactic supernova remnants (SNR) (discussed in section 1.4.1) as shown in Figure 1.5; both of these objects show a better fit with hadronic models than leptonic models using bremsstrahlung radiation, although a reasonable fit can be made by fitting a lower energy cut-off in the electron spectrum at 300 MeV. This is evidence supporting SNR as hadronic sources, making them also sources for an astrophysical neutrino flux.

In summary, electromagnetic measurements of astrophysical objects are still a very powerful tool to understand them. They currently demonstrate the most precise angular resolution and cover a range of energies spanning almost 19 orders in magnitude from 10^{-6} eV to 10^{13} eV. However, due to the various production paths for photons, it is difficult to use the measurements to differentiate between the physical processes, and hence identify which objects accelerate high energy hadrons versus electrons. Since photons also interact with matter and other radiation with a fairly high cross-section, gamma-rays are limited by an event horizon at high energies; they are also likely to be absorbed and attenuated by clouds of matter or objects surrounding the CR accelerator.

1.3 Neutrinos

Astrophysical objects with powerful magnetic fields can accelerate charged particles to high energies. At ultra-high energies there is a well observed spectrum of CRs made of atomic nuclei, dominated by protons with a much smaller fraction of He and other nuclei. In order to produce such a flux, objects must accelerate hadrons to high energies; the hadrons interact with each other locally and with radiation producing charged and neutral pions as described in section 1.2.2. The charged pions then decay into muons, and these decay



(A) IC443 (B) W44

FIGURE 1.5: Gamma-ray spectra for Galactic Supernova Remnants IC443 and W44 taken from Fermi [29]. Color-shaded areas bound by dashed lines denote the best-fit broad-band smooth broken power law (60 MeV to 2 GeV), gray-shaded bands show systematic errors below 2 GeV due mainly to imperfect modeling of the galactic diffuse emission. Solid lines denote the best-fit pion-decay γ-ray spectra, dashed lines denote the best-fit bremsstrahlung spectra, and dash-dotted lines denote the best-fit bremsstrahlung spectra when including an arbitrary low-energy break at 300 MeV in the electron spectrum. These fits were done to the *Fermi*-LAT data alone (not taking the TeV data points into account).

further into neutrinos creating an extra-galactic neutrino flux. The charge of the pion determines whether more neutrinos or anti-neutrinos are produced according to:

$$\begin{aligned} \pi^+ &\rightarrow \nu_\mu + \bar{\mu} \rightarrow \nu_\mu + \bar{e} + \nu_e + \bar{\nu_\mu} \\ \pi^- &\rightarrow \bar{\nu_\mu} + \mu \rightarrow \bar{\nu_\mu} + e + \bar{\nu_e} + \nu_\mu. \end{aligned} \tag{1.23}$$

Assuming that the final decay products share the pion energy equally and that the average energy transferred from the proton to the pion is $\langle x_{p\rightarrow\pi} \rangle = \langle \frac{E_\pi}{E_p} \rangle \simeq 0.2$, the relation between the energy of the origin CR and the photon and neutrino messenger particles can be approximated respectively as:

$$\begin{aligned} \frac{E_\nu}{E_p} &= \frac{1}{4}\langle x_{p\rightarrow\pi} \rangle \simeq \frac{1}{20} \\ &\text{and} \\ \frac{E_\gamma}{E_p} &= \frac{1}{4}\langle x_{p\rightarrow\pi} \rangle \simeq \frac{1}{10}. \end{aligned} \tag{1.24}$$

The expected neutrino flavor ratio at production should be $\nu_e : \nu_\mu : \nu_\tau = 1 : 2 : 0$, no differentiation between neutrinos and anti-neutrinos can be made on the scale of neutrino detectors. This differentiation is typically done by comparing the charge of the interaction products of the neutrinos using a magnet, but they cannot be constructed on the km

scale typical for interaction products associated with the neutrino energies probed in this thesis. Since all charged pions eventually decay into neutrinos and leptons regardless of their energy, assuming a CR flux dominated by hadrons with a fixed charged-to-neutral secondary pion ratio, the neutrino spectrum should follow directly. For a power law CR spectrum the neutrino flux would be expected to have the same form and the same spectral index with just a fraction of the normalisation. It should be noted that in dense sources at high energies the neutrino flavor ratio may vary as the produced muons have a higher chance of interacting with other matter before decaying to further neutrinos; in an extreme case, the flavor ratio would be $\nu_e : \nu_\mu : \nu_\tau = 0 : 1 : 0$.

It is important to understand the expected flavor ratios in order to extrapolate the resulting ν flux associated with a certain CR flux from a given astrophysical source. Given the high energies and long travel distances from the neutrino origins to the Earth as well as the broad spectrum of neutrino energies, it is expected that given a $\nu_e : \nu_\mu : \nu_\tau = 1 : 2 : 0$ ratio at production, the neutrino oscillations combined across all energies will result in an even flavor ratio of $1 : 1 : 1$ [30, 31]. This prediction arises from measurements of the neutrino flavor and mixing properties conducted from known neutrino beams, for example at T2K [32], as well as observed neutrino flux from solar and atmospheric neutrinos [33]. For these neutrino beams the source distance (baseline) is known as well as an estimated origin flux and flavor ratio from the reactions at the source which produces the neutrinos. By observing high-energy astrophysical neutrinos, important contributions to oscillation measurements can also be made [34].

At the high energy end, the GZK cut-off discussed in Section 1.1.1 CRs interact with the CMB in the same interaction described by Equation 1.20. For heavy CRs, such as iron nuclei, the nuclei could absorb the CMB photon only to excite hadrons in the nuclei, prompting the ejection of nucleons from the nucleus with no high energy neutrino production. This process is known as photo-disintegration and would have a higher cross-section and shorter mean free path than the pion production process at energies above 10^{18}A eV, where A is the CR atomic mass number. In the case where the highest energy CRs are light elements that undergo proton-proton collisions or interact with the CMB, a flux of ultra-high energy cosmic neutrinos would be expected.

Neutrinos are a advantageous messenger to observe high energy extra-galactic CR accelerators. They are predicted to have a large flux from these sources—close in magnitude to that of the γ-ray and CR spectrum—which is unimpeded on its journey to the Earth. This is because neutrinos only interact via the weak force with low cross-sections, so the majority of neutrinos are unaffected when passing through gas clouds or even solid objects; they are also neutral, which means their direction is unchanged from the site of production. These advantages become considerably more relevant when searching for ultra-high

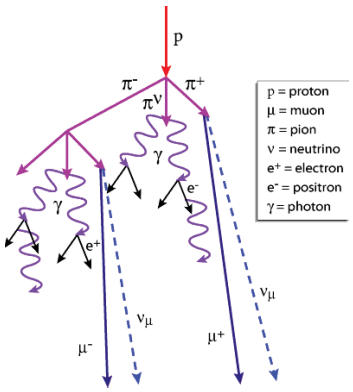


FIGURE 1.6: Diagram of an air-shower produced by a high energy CR interacting with the atmosphere. These showers are detected to measure the highest energy CRs and are the dominant source of background for astrophysical ν measurements. Taken from [35].

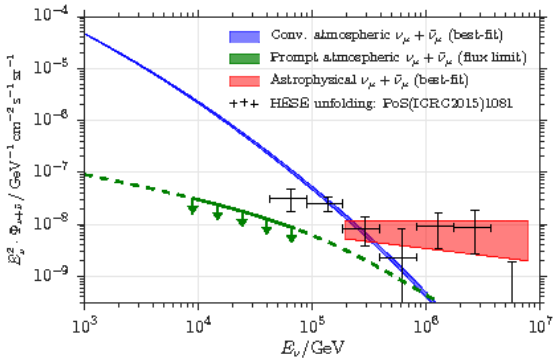


FIGURE 1.7: Best fit of the astrophysical, prompt, and atmospheric neutrino flux seen by IceCube as high quality muon tracks in the northern hemisphere starting at TeV energies. Taken from Aartsen et al. [10].

energy CR sources; at such high energies photons cannot travel the long distances from extra-galactic sources and the CRs are, in general, too deviated in their direction to point precisely to their sources. Since neutrinos—unlike photons—can only come from hadronic interactions, an astrophysical neutrino flux associated with a possible CR source candidate would be a smoking gun for CR production.

1.4 Sources of High-Energy Messengers

As already explained in Sec. 1.2.2, hadronic processes in CR accelerating sources produce radiation and neutrinos from interactions of CRs with matter or radiation. The question then remains, what kind of objects in the universe could be the site of such processes and what evidence exists to support these objects as CR accelerators.

As discussed already in section 1.1.1, it is supposed that the flux contribution at highest energies is dominated by extra-galactic sources, whereas the flux below 10^{18} eV is dominated by galactic emission. It is also the case that electromagnetic observations have limited horizon due to photon absorption during propagation from the source to the Earth. Hence, more information is known about Galactic objects. Nonetheless, the Galaxy cannot host the most powerful accelerators able to produce CRs in the region well beyond the knee.

1.4.1 Galactic Objects

As previously mentioned in section 1.1.1, it is thought that the part of the CR spectrum (until about 10^{15} eV) is dominated by Galactic sources. Their emission is expected to have a diffuse component since CRs remain trapped and gyrate in the Galactic magnetic fields for million of years. There might be also a component which arrives from extra-galactic sources and becomes trapped in the Galactic magnetic field. These CRs would then interact with radiation and matter in the Galaxy to produce pions which can decay to create a high-energy neutrino flux. There are models for diffuse galactic emission [36] assuming that the γ-ray emission measured by the Fermi-LAT detector is from π^0 decay. IceCube recently set 90% confidence level upper limits on the possible neutrino emission assuming π^0 decay from the Fermi-LAT model as $6.13 \times 10^{-18}\left(\frac{E}{100\ TeV}\right)^{-2.5}$ ($\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$)[8], or 14% of the observed diffuse astrophysical neutrino flux reported in Aartsen et al. [37].

There are several promising candidates for point-like CR sources within the Galaxy, including objects of currently unknown type which have been observed emitting high-energy γ-rays which can be associated with CR acceleration. The physics of these objects can be closely related to the highest energies attainable by the accelerated particles. Charged particles at high energies in a magnetic are impossible to contain in a volume considerably smaller than their Larmor radius as they are not sufficiently deflected. Using the magnetic field intensity B and the size of the accelerating region R where particles are contained, a condition can be applied to deduce the maximum energy particles can be accelerated to:

$$\mathcal{E}_{max} = eBR \simeq 10^{21}\left[\frac{R}{1\ pc}\right]\left[\frac{B}{1\ G}\right]\ eV, \tag{1.25}$$

This is known as the Hillas' condition. Equation 1.25 is not necessarily fulfilled in the case on an inefficient accelerator, because it does not account for energy losses inside the accelerating region, but it marks an upper limit in acceleration capacity under these conditions. One can apply this condition to the highest energy observed CRs and determine the possible astrophysical objects that can accelerate CRs up to a certain energy. This is

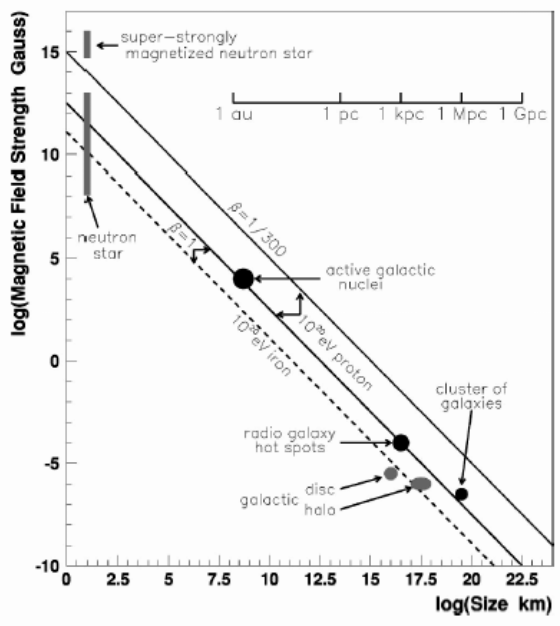


FIGURE 1.8: Hillas' plot demonstrating the Hillas' criterium for the maximum energy of a charged particle from a source with a magnetic field B and a size R.

typically shown by the "Hillas' plot", as seen in Figure 1.8. Sources able to accelerate protons above 10^{20} eV would have to lie above either black line depending on the velocity of the accelerating region $\beta = \frac{U}{c}$ which functions as an efficiency. This is achieved by either having a stronger magnetic field or larger accelerator region.

Supernova Remnants (SNRs)

SNRs have long been considered one of the most probable candidates for CR acceleration within the Galaxy. The death of a massive star can cause a powerful explosion known as a supernova; the resulting structure is the supernovae remnant. The energy from a supernovae is transferred to the stellar envelope as kinetic energy, causing it to quickly heat up and expand. At the energies in question this shell begins to expand supersonically creating an abrupt discontinuity between the interstellar matter and the hot expanding shock front.

Since the shock front is moving fast relative to the surrounding matter, the interstellar medium (ISM) begins to accumulate in the outer shell. This effect gradually decelerates the shock front compared to the interior expansion of the SNR which in turn creates an internal shock front traveling towards the center of the SNR. This interior shock front converts the kinetic energy from the expansion into heating of the ejected gas inside the

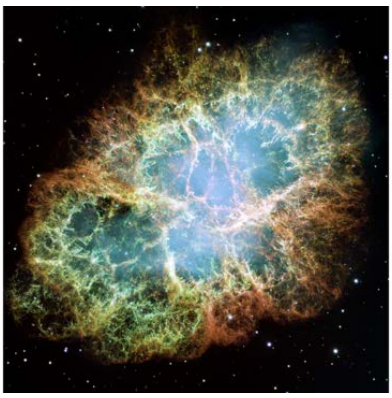


FIGURE 1.9: Hubble Space Telescope image of the Crab Nebula which is associated with an inner pulsar acting as a Pulsar Wind Nebula (PWN). This was the first PWN to be observed at TeV energies and is typically used to calibrate astrophysical results as one of the brightest objects in the sky.

shell. This shell serves as a prime candidate for first order Fermi acceleration presented in section 1.1.2 for as long as the shell acts as a strong shock.

Eventually, as the shell expands, the temperature inside the shell decreases; initially it causes compression at the shock front to preserve pressure balance until, finally, the expansion becomes subsonic. At this point the matter is able to disperse from random motion and is no longer considered a SNR. Although the acceleration mechanism continues until this stage of the SNR, it is most efficient at earlier stages (the first 1000 years) when the blast wave has not substantially decelerated. This is why often young SNRs are of more interest as CR emitters. However, their contribution to the CR energy spectra is typically limited to below 10^{15} eV due to their relatively small size and magnetic field, as well as the limited lifetime as an efficient accelerator. High-energy gamma-ray observations of such SNRs act as evidence supporting the hypothesis of particle acceleration, however they do not clarify the leptonic or hadronic origin of the γ-ray flux. For some SNRs the hadronic models reproduce well the spectral emission distribution (SED), as seen in results from Fermi studies of the γ-ray emission from two known SNRs that are better fit by an hadronic model than leptonic models [29].

Pulsar Wind Nebula (PWN)

In the case of a supernova often a neutron star is formed at the centre after the core collapse of a massive star. These neutron stars rotate rapidly. Their rotation can be observed through precessing beams of radiation originating from their poles. When such a pulsar is surrounded by a nebula of matter, charged particles close to the neutron star are accelerated to relativistic energies powered by its strong magnetic fields into a powerful

wind. Such pulsars are typically found at the center of a SNR where the wind meets the relatively slow moving particles in the surrounding matter creating a shock front where particles can be further accelerated. These objects typically are strong X-ray emitters and have been observed in photons at optical wavelengths to have an extended shock front with a small central nebulous region corresponding to the pulsar powering the wind. In 1989, the first observation of PWN at TeV energies was conducted by Whipple Observatory by detecting γ-rays associated with the Crab Nebula [38]. Nowadays the crab is well observed as seen by Hubble Space Telescope in the optical energy region shown in Figure 1.9. About 40 [39, 40] of these objects have been identified at TeV energies by the gamma Telescopes such as HESS [41].

The Crab Nebula has an associated PWN and is currently one of the brightest objects in the sky and is often used to calibrate telescopes acting as a standard candle in the X-ray energy range.

1.4.2 Extra-galactic Objects

Once we begin to consider objects outside of the Galaxy the distances begin to become immense (greater than 10 Mpc). At these distances even large galaxies become point-like when observed from earth. When considering point-like neutrino sources produced by CRs above the *knee* at $\sim 10^{15}$ eV, the usual suspects are other galaxies. It can be difficult to understand exactly which processes are occurring in these objects due to their large distances, but they can often be sorted into several general categories.

Active Galactic Nuclei (AGN)

Just as the Milky-way contains a super-massive black-hole at the centre, many other galaxies are thought to have a central black hole known as the nucleus. When the luminosity from the black whole nucleus outshines the remainder of the galaxy, it is classified as an Active Galactic Nuclei. The luminosity from the nuclei is expected not from the black-hole itself but from matter which forms an accretion disk around the nucleus. This matter is gradually pulled in towards the centre, at which point the pressure on the cold matter becomes enough to heat it up to millions of degrees kelvin. Due to the strong magnetic field lines around the black-hole, jets can become collimated around the magnetic poles forming a jet of radiation [42]. High rates of acceleration can be reached within the jet and—when it is pointing in the direction of the Earth—it is seen as luminous and referred to as a *blazar*. The luminosity expected from such a jet where radiation and matter can escape is much brighter than radiation from the rest of the galaxy which is limited to the value of the *Eddington Luminosity*. Above this luminosity L_{edd}, the radiation would provide a

greater expansive pressure than the gravitational pressure which is attracting all of the matter to the black-hole; it does not depend on the distance from the centre but just on the mass M of the central object and can be written as:

$$L_{edd} = \frac{4\pi G M m_p c}{\sigma_T} = 1.3 \times 10^{31}\ W, \tag{1.26}$$

where σ_T is the Thomson cross-section and m_p is the proton mass. This calculation assumes the accreting matter is dominated by protons and radiation produced from Thompson scattering of high-energy photons upon interacting with charged particles in the source. The cross-section for this interaction is then defined as :

$$\sigma_T = \frac{d\mathcal{E}/dt}{|\vec{S}|} \tag{1.27}$$

Where $\vec{S}$ is the Poynting vector which is taken to be the incident energy flux, and $\frac{d\mathcal{E}}{dt}$ is the radiated energy. By integrating for the average value assuming collisions between 0 and π and assuming the Larmor formula for radiated energy such that:

$$\left\langle \frac{d\mathcal{E}}{dt} \right\rangle = \frac{e^4}{3m_e^2} E_0^2, \tag{1.28}$$

$$\text{and } \langle \vec{S} \rangle = \frac{\langle \vec{E} \times \vec{B} \rangle}{4\pi} = \frac{E_0^2}{8\pi} \tag{1.29}$$

$$\Rightarrow \sigma_T = \frac{8\pi e^4}{3m_e^2} \approx 6.65 \times 10^{-25} cm^2 \tag{1.30}$$

where E_0 is the maximum amplitude of an linearly polarized electromagnetic plane wave.

The orientation of the AGN with respect to the Earth determines which part of the AGN can be observed. Hence, they are primarily categorized based on this criterium, as seen in Figure 1.10. Within different angles there can be even more subcategories based on the presence of a jet or the nuclear activity levels and hence their emission lines [43, 44]. The subcategories of AGN are defined by the relative flux seen at different wavelengths in the electromagnetic spectrum. Blazars are observed with a broad spectrum of emission in wavelengths, including bright radio cores. They can be further subdivided based on their broad spectrum emission. For instance, the presence of surrounding gas which absorbs and emits photons in the optical range should exhibit clear optical spectral lines. Objects with luminous broad emission lines and continuous thermal emission from optical to UV wavelengths are called Flat Spectrum Radio Quasars (FSRQ). Those with weak or absent broad emission lines are BL Lacs—so named after the galaxy BL Lacertae.

Traditional blazars are bright and easy to detect, unless the jet is obscured. Where there

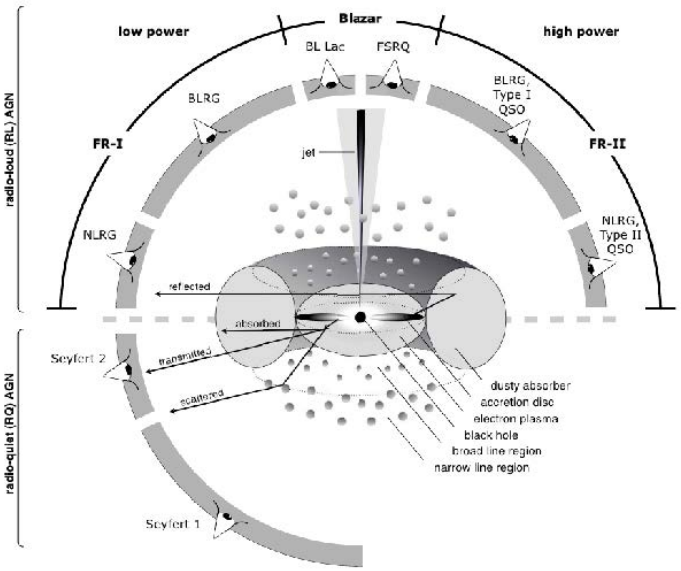


FIGURE 1.10: AGN classification as a function of their observation angle with respect to a possible jet from the black-hole magnetic poles taken from Singh [45]

is a lot of matter in the galaxy surrounding the black-hole, higher energy photons can be absorbed and scattered making their emission severely attenuated. However, lower energy photons may travel through the gas clouds without scattering and a high flux could indicate an AGN jet pointing in the direction of the Earth; these can be referred to as choked jets. In addition to extreme luminosity, blazars often show high rates of variability emitting flares lasting from minutes to months. There is considerable effort to correlate signals from different wavelengths during these flares. The first evidence for a high energy astrophysical neutrino source comes from such a flaring blazar TXS 0506+506, which was triggered by an alert which occurred during a period of observed high emission in γ-rays [46].

Even galaxies without jets, or where the jet is not oriented towards the Earth, could be strong CR emitters. For example, there are galaxies with particularly dense clouds of gas and matter with high rates of star formation and supernovae. The high density in these galaxies does not only lead to high numbers of possible cosmic ray accelerators, but also to a large quantity of matter with which the accelerated CRs can interact and produce pions—which decay into high energy γ-rays—and neutrinos—which escape the galaxy. These galaxies, known as Star-Burst Galaxies (SBG), could also act as sites of scattering for ambient CRs originating from other extra-galactic sources, making them another strong candidate for CR point-like sources detectable by neutrinos [47, 48].

Gamma Ray Bursts (GRBs)

Gamma Ray Bursts (GRBs) were discovered during the Vela military mission as emissions of high energy γ-rays over a short duration. They are brighter than any other source in the γ-ray sky with a typical luminosity of around 10^{52} erg s^{-1}. These are commonly divided into short GRBs (sGRB)—lasting from 10^{-3} s to 2 s—and long GRBs—lasting from 2 s to 10^3 s. These GRBs show peak luminosity between 100 keV and 1 MeV which is largely non-thermal.

It was predicted that such powerful γ-ray emission originated from relativistic shocks in extra-galactic sources. Shocked electrons attaining high energies would then emit synchrotron radiation corresponding to the observed γ-ray bursts, lasting only a short period of time with lower energy emission expected to last longer. This emission is referred to as the *afterglow* and is expected at X-ray, optical and radio wavelengths. The first evidence of such an afterglow was observed by X-ray emission 8 hrs after a detected GRB [49]. Observed afterglows not only provide evidence for the mechanisms behind the observed emission, but they also drastically improve the direction resolution of the event. Such short duration flares of γ-ray flux allow a minimum resolution of a few degrees around the burst direction, whereas longer duration observations in X-ray and optical can pinpoint the flare within 1° to a single associated object.

The time-scales of the two burst categories are usually associated with different production mechanisms and display distinctive properties. Typically, long GRBs are more energetic, demonstrate a longer rise-time and a softer γ-ray spectra. The duration and variability of these bursts places a limit on the size of the burst source to roughly 3×10^5 Km, which corresponds to stellar objects. These Long GRBs are expected to be created during a core-collapse supernovae from massive stars in star-forming galaxies whilst the short GRBs correspond to even more compact objects such as neutron stars or black-holes. High luminosities are expected from the merger of these objects and they are now being observed by advanced gravitational detectors [3]. The correlation of short GRBs with a gravitational wave signal occurred for the first time providing strong evidence of neutron star mergers as the origins of sGRBs [50]. In this case these mergers would be very interesting candidates for an associated high energy neutrino flux[51]. The first observation of a GRB close to TeV energies was made on the 15^{th} of January 2019 by MAGIC[52] for energies above 300 GeV. These recent observations begin to illuminate the physics behind these extremely luminous events.

Chapter 2

IceCube Neutrino Detector

2.1 The Detector

The motivation behind the IceCube experiment was to identify and characterize cosmic neutrinos and locate their sources. When motivating the construction of such an experiment it is important to understand how this signal can be observed. Neutrinos arriving to the Earth's surface at energies below 1 TeV are overwhelming atmospheric. In order to achieve a reasonable fraction of neutrino events from outside the solar system, the particle energies must be extremely high, to the point where the number of background neutrinos start to dwindle. Before the IceCube detector, Baikal [53], AMANDA [54] and ANTARES [55] were constructed to observe high energy neutrinos. Although they were key in characterizing high-energy atmospheric neutrino properties, a significant astrophysical signal was not detectable given the low rate of events at the highest energies. In order to probe higher energies a detector must be constructed that is capable not only of observing a considerable number of events for neutrinos above 1 TeV, but also a way of estimating and constraining the energy of these events. Such requirements are only achieved by using a detector in the km^3 scale, which was motivated by theoretical predictions, using a medium in which the neutrinos can interact and be observed.

2.2 Neutrino Detection

The IceCube experiment takes advantage of an already tried and tested method for measuring neutrino interactions. Neutrinos are uncharged and interact only via the weak force. Even in their weak interactions, neutrinos have a very small cross section meaning that they rarely interact with the matter that they pass. Several factors come into play when observing astrophysical neutrinos at the Earth. A high rate of neutrino production is expected from astrophysical sources, their small cross section prevents them from being attenuated from their original source, and neutrino detectors employ large detector volumes. Hence,

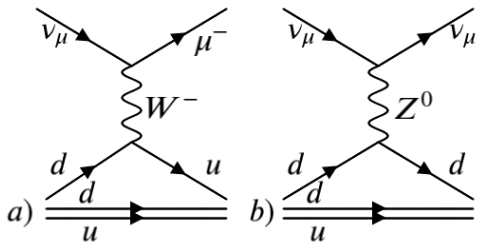


FIGURE 2.1: Fermi diagrams showing charged a) and neutral b) current interactions of neutrinos with atomic nuclei via the weak force.

enough neutrinos will pass through the detector that some of them will be observed though their interactions with the detector medium. This occurs when neutrinos interact with the ice molecules through either charged current (CC) or neutral current (NC) interactions. At very high energies probed by IceCube, neutrinos enter the deep inelastic scattering (DIS) regime with the atomic nucleons in the ice. Charged current interactions are mediated by charged W bosons in the following interactions:

$$\nu_\mu + n \rightarrow \mu^- + p, \tag{2.1}$$

$$\bar{\nu_\mu} + p \rightarrow \mu^+ + n \tag{2.2}$$

which are depicted respectively in Fig.2.1. To conserve lepton number, in a charged current interaction, each different flavour of neutrino will produce its respective lepton. Neutral current DIS is mediated via the neutral Z boson and the high energy neutrino is transformed into a lower energy one. Similarly to CC interactions, a significant fraction of the neutrino energy is transferred to the atomic nucleus depending on the ν energy, ranging from 50% at low energies to 20% at higher ν energies. This energy transfer triggers nuclear disintegration and hence hadronic cascades in the ice whilst the neutrino escapes undetected.

Both CC and NC interactions occur with the cross section increasing at higher energies according to Fig.2.2 with $\sigma \propto E$ until roughly $E_\nu \equiv 10^3\ GeV$. Below that energy the interaction is limited by the large mass of the W and Z bosons. At energies above this the momentum transfer k^2 begins to play a significant role in the cross sections so that:

$$\sigma_{cc/nc} \propto \left(\frac{1}{k^2 + m^2_{W/Z}} \right)^2 E. \tag{2.3}$$

It can been seen that at lower energies the ν cross-section is greater than that of $\bar{\nu}$. This

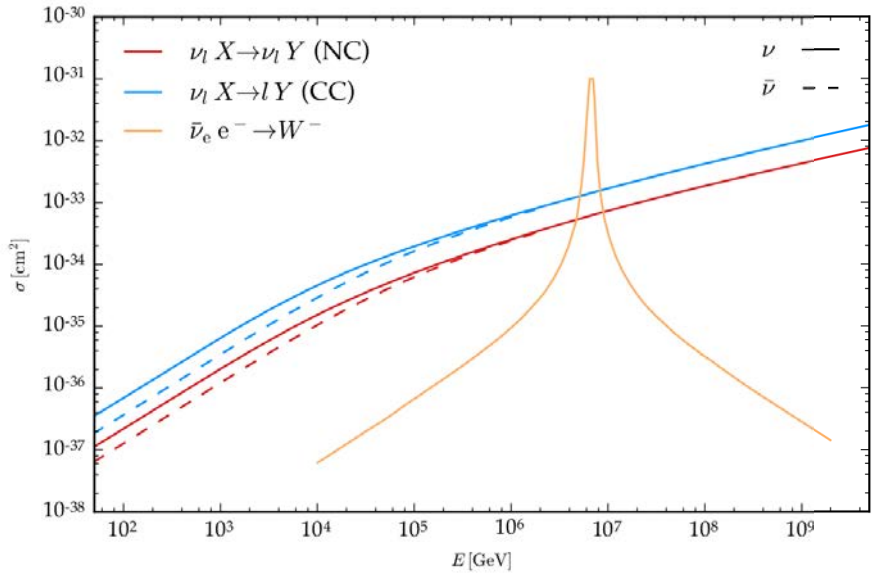


FIGURE 2.2: ν cross sections for charged and neutral current interactions with atomic nuclei in the ice as well as $\bar{\nu_e}$ cross section with electrons which peaks at high energies known as *Glashow resonance* [56]. Data taken from Gandhi et al. [57]

can be explained by the target hadrons for DIS. At lower energies the neutrinos primarily interact with the valence quarks in the hadrons composed of matter not antimatter (uud or udd for protons and neutrons respectively). At higher energies scattering begins to occur increasingly within the sea of quarks in the nucleons which are composed equally of matter and antimatter thereby evening out the relative ν and $\bar{\nu}$ cross-sections.

In addition to DIS with the atomic nuclei, energetic ν can also interact directly with the electrons bound to the atoms in the ice. In the case of ν_e this is effectively just a scattering, for other flavours of neutrino this can result in producing a muon or tau lepton accompanied by a ν_e. Due to the small electron mass, these interactions typically have such a small cross-section that they are negligible compared to the nuclear reactions. However at 6.3 PeV [56] the center of mass energy for the $\bar{\nu_e}$ and e interaction equals that of the W boson resulting in a resonant cross section. The cross-section at this peak exceeds that of nuclear ν interactions up to extremely high energies as seen in Fig.2.2 and is referred to as *Glashow* resonance. Such an interaction would produce two showers of particles from both the W boson production and decay verticies. This is the only interaction IceCube could be sensitive to which differentiates between ν_e and $\bar{\nu}_e$ since atoms in the ice contain electrons not positrons allowing no equivalent interaction of $\nu_e + e^+$. Such an interaction would be observed as a double cascade of particles in the ice.

The mean free path of neutrinos is found by calculating a total cross section for all

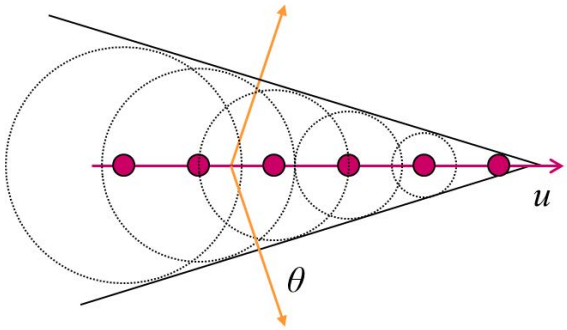


FIGURE 2.3: Diagram of Cherenkov radiation emitted at an angle θ by a particle shown in pink moving at a velocity u faster than the local speed of light in the medium.

possible interactions of CC and NC nuclear interactions and interactions with the shell electrons for each flavour. When traversing a medium of density ρ the neutrino mean free path λ_ν becomes:

$$\lambda_\nu = \frac{1}{\sigma_{tot}(E)\rho}. \tag{2.4}$$

When considering the density of the Earth this means that the interaction length for neutrinos above PeV energies actually becomes shorter than the diameter of the Earth. Since the cross section increases with energy, at higher energies the neutrinos have a smaller chance of crossing the Earth without interacting making the Earth effectively opaque. Some of the high energy flux can emerge at lower energies in the case of NC or ν_τ CC interactions where another high energy neutrino escapes after the interaction. This effect is referred to as *regeneration* [58] and also applies for CC- ν_τ interactions since τ leptons are unstable particles which decay into ν_τ with a lifetime of only 2.9×10^{-13} s.

2.2.1 Cherenkov Radiation

Given the event of a neutrino interaction in the ice a series of very energetic charged particles are produced or scattered. When the velocity of such particles exceeds that of the local speed of light in the medium then the resulting radiation is referred to as Cherenkov radiation [59]. The local speed of light is estimated as $\frac{c}{n}$, with c being the speed of light in a vacuum and n being the index of refraction of the medium, which is 1.31 for ice. Since a particle is ionising surrounding matter faster than the resulting radiation can travel in ice, the radiation is emitted at an angle θ_C depicted in Fig.2.3. This angle is a function of the particle speed v and the refractive index of the medium according the following equation:

$$\theta_C = \cos^{-1}(\frac{c}{vn}). \tag{2.5}$$

In water or ice the resulting emission angle is found to be roughly 41°. The number of expected photons per track length x and radiation wavelength λ is estimated using the Frank-Tamm formula [60]:

$$\frac{dN}{dxd\lambda} = \frac{2\pi z\alpha}{\lambda^2} sin^2(\theta_C), \tag{2.6}$$

where z is the charge of the ionizing particle and α is the fine structure constant taken to be $\frac{1}{137}$. Since the photon emission is dependent on $\frac{1}{\lambda^2}$, the peak emission is found at short wavelengths. For deep polar ice this lies principally between 350 and 600 nm which appears blue in the optical region.

2.3 IceCube Detector

The deep Antarctic ice acts as a high quality medium in which to observe the neutrino interactions described in the sections above. It is a transparent to local radiation although most light from above the surface is blocked. In order to observe charged particles from neutrino interactions the ice is instrumented with an array of Photo-Multiplier Tubes (PMTs) which detects the emitted Cherenkov radiation. The layout of this array was designed to optimize multiple features. Firstly the set up must be compatible with the energy range of the neutrino events that IceCube aims to detect. The IceCube detector was constructed to differentiate the atmospheric and astrophysical neutrino flux. Given at lower energies no astrophysical signal was yet observed, IceCube targets energies $\mathcal{O}(TeV) - \mathcal{O}(PeV)$ for a detectable contribution from the astrophysical flux to emerge. The targeted neutrino energy and the cost dictates how large the detector should be and the maximum spacing which can be used between strings with optical modules. Higher energy neutrinos are expected to arrive at a considerably lower rate and therefore require a much larger detector volume in order to ensure the observation of some of these events within the detector lifetime.

Neutrino events at TeV energies and above are also expected to light up a much larger volume of the detector. This event volume means fewer strings of optical modules are needed to achieve a sub degree precision in the direction reconstruction. The distance between the strings is chosen as a balance between the higher and lower energy events the experiment targets as well as experimental set-up costs.

2.3.1 Set-up

The final IceCube array [61] consists of 86 strings each with 60 Digital Optical Modules (DOMs), described in section 2.6.1, which detect the cherenkov radiation in the ice. The strings of DOMs are placed in a hexagonal prism formation with a spacing of 125 m between

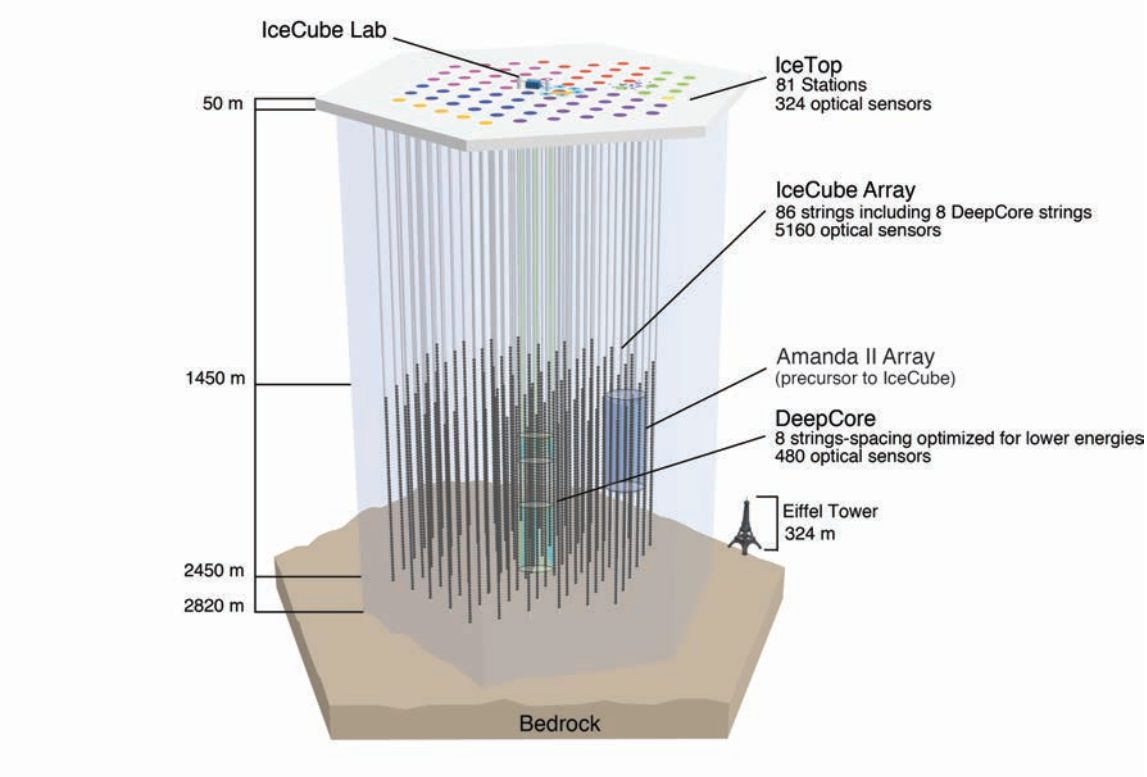


FIGURE 2.4: Diagram of IceCube detector layout below and above the surface of the ice.

each string and 17 m between each DOM as seen in Fig. 2.4. This array begins 1450 m below the surface of the ice in order to ensure an environment with the minimum amount of noise possible. Ambient light and low energy atmospheric neutrinos cannot penetrate through this distance of ice. The final detector volume is 1 km^3, making IceCube the largest particle detector in the world.

DeepCore

As discussed before, the IceCube detector as a whole has been optimized to discover astrophysical neutrinos above TeV energies. The large separation between the detector strings makes reliable reconstructions of lower energy events difficult. In order to probe $\mathcal{O}(GeV)$ neutrino events relevant for different analyses, such as neutrino oscillation and solar flare analyses, the centre of the detector has subset of additional DOMs called DeepCore. DeepCore is optimized for lower energy events by employing a more dense string spacing. 8 extra strings are added resulting in an inter-string spacing which is on average half that of the rest of the detector and which varies between 41 and 105 m. The DOM spacing within the additional strings varies as a function of depth. The shallowest DOMs between 1750 and 2000 m have a wider spacing of 10 m and are used as a veto for those which are below. Between 2000 and 2100 m no additional DOMs was installed because this region, named the *"dust layer"*, shows a significantly increased optical absorption and scattering. This is thought to be due to dust particles contaminating the ice which reduce the event detection efficiency. Below this layer until 2450 m, the DOMs have the most dense spacing of only 5 m. Most of the DOMs in this region are an upgraded version with higher quantum efficiency of 35%, compared to about 25% for the rest of the DOMs. These upgraded DOMs are installed on 6 complete strings and are combined with the standard DOMs for the remaining 2 strings. The DeepCore DOMS are placed at the center of the lower half of the IceCube array, beginning deeper than 1750 m below the surface. This allows for the maximum amount of noise to be already removed by the surface ice and by using the surrounding IceCube strings as a veto region.

The improved efficiency, dense spacing, and use of veto regions, extends the lower range of detectable event energies from 100 GeV to 10 GeV. These lower energy events are critical for neutrino oscillation analyses, dark matter searches and the ability to detect Galactic supernovae.

IceTop

Just as the DeepCore region of the detector is able to use the rest of the DOM array as a veto, the Location of the IceTop surface array directly above IceCube allows it to

function also as a veto for IceCube events. IceTop is a cosmic air shower array on the surface of the ice directly above the detector. This array is made of 162 ice-filled tanks arranged in pairs separated by 10 m referred to as stations. The stations align roughly with the in-ice strings. Each tank contains PMTs that can observe if a high energy CR, producing many background muon and neutrino events, arrives above the detector. As in the IceCube array, IceTop estimates the energy deposited based on the light measured by the 2 standard IceCube DOMs within each IceTop ice tank. For a single cosmic ray event, it is typical that the shower particles will cover several IceTop stations. The deposited energy across all the affected tanks is used to derive an estimate for the original CR energy and direction. Given a typical tank spacing, in a hexagonal array mirroring that of IceCube DOMS, IceTop is sensitive to cosmic rays of $\mathcal{O}$(PeV) up to EeV in energy for the larger station spacing. Above DeepCore where the stations are arranged close together, the lower energy threshold can be as low as 100 TeV.

IceTop measures the electromagnetic component of the showers induced cosmic rays, while the in-ice detector measures the more penetrating muons. This allows one to infer the composition of cosmic rays. Hence, IceTop measures the spectrum and composition of cosmic rays in the energy region of the knee. Another use for this detector is to make studies of high energy muons, gamma rays, and transient incidents including solar flares. A critical use for this surface detector concerning this thesis is it's implementation as a veto for events observed in the IceCube array in coincidence with a cosmic ray observed directly above the detector by IceTop. This is an effective tool but has a limited veto capabilities as it only covers a small fraction of the surface from which background atmospheric events arrive. The spacing between the stations also results in a high energy threshold meaning lower energy background events could not be effectively vetoed.

2.4 Antarctic Ice Properties

The properties of the antarctic ice dictate the optical properties in the detector and therefore the light yield expected for tracks of different energies and orientations. Overall the antarctic ice is exceptionally pure and clear, making it an excellent medium for optical transmission. There are however position dependant imperfections in the ice, especially as a function of depth in the detector. These imperfections need to be well understood in order to precisely reconstruct the interactions occurring in the ice.

The first discernible issue is the air bubbles in the ice, which significantly increase the light scattering in the ice due to the different density between the air in the bubbles and the surrounding ice. This provides the primary motivation in positioning the DOMs starting from 1450 m below the ice surface where the pressure compresses the air bubbles into

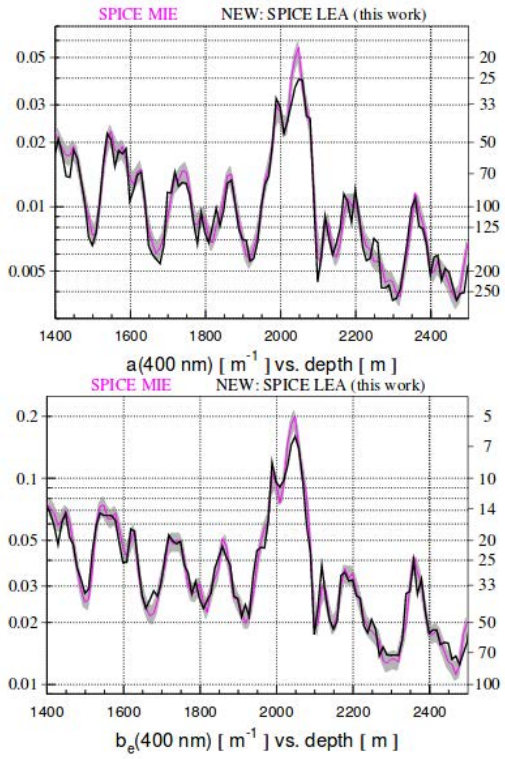


FIGURE 2.5: Results of depth dependent scattering and absorption frequencies in the ice comparing Spice-Mie and Spice LEA which accounts for horizontal anisotropies as well.

clathrates of a similar density to the surrounding ice. There can be further imperfections which need to be well measured to calibrate the simulated light yield as accurately as possible.

This calibration is conducted by *flasher runs*. On each DOM there are 12 LEDs which emit light at frequencies expected by Cherenkov radiation. During the flasher runs the light yield across the array of DOMs is recorded when the LEDs from each DOM emit photons. In this way the properties of the ice which are estimated in the simulation, such as the absorption and scattering coefficients, can be fit to give the same as the observed results as a function of DOM position. The final result is a complete ice model known as SpiceMie [62], named for South Pole Ice and the theory of electromagnetic radiation scattering off small particles documented by Gustav Mie [63]. In this model scattering and absorption coefficients a_{400} and b_{400} respectively at a frequency of 400 nm are fit, as described by Ackermann et al. [64] and Aartsen et al. [62], as well as the depth dependent temperature and six global ice parameters.

Further investigation into the ice properties resulted in an observed anisotropy [65].

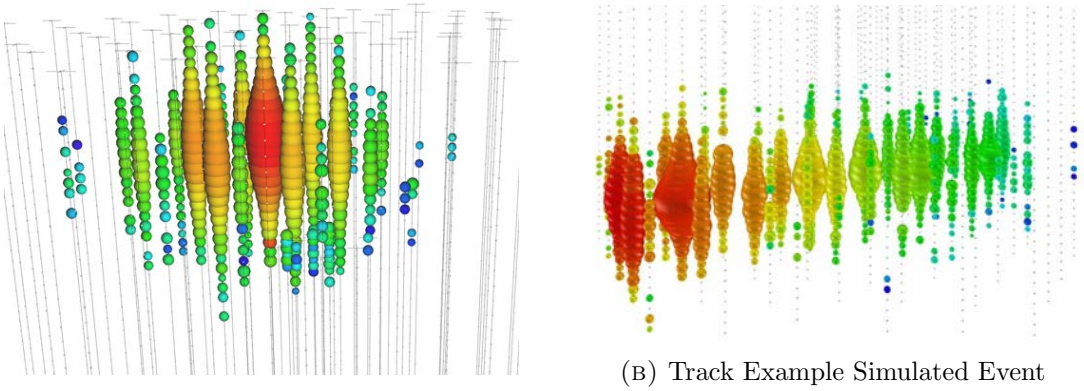
A depth dependant anisotropy was already well documented and can be seen in Fig. 2.5 for Spice-Mie. The absorption and scattering coefficients fluctuate with the depth in the detector which demonstrates the different mean free path of photons before absorption or scattering respectively across the x-y plane of each layer in the ice. These fluctuations represent the varying purity of the ice for different layers in depth and a general tendency can be seen for the mean free path to increase with depth implying purer ice. This difference arises due to the air bubbles reaching a maximum pressure and being condensed into hydrates [66]. These hydrates are air-ice clathrates with a close enough index of refraction to the surrounding ice to allow minimal refraction of passing photons [64].

Another noticeable feature of Fig. 2.5 is the peak in the absorption and scattering coefficients at around 2000 m depth. This feature is due to much higher rates of absorption and scattering for these layers of ice which implies many impurities. These impurities are thought to be originated from dust from a large volcano explosion during the years in which these ice layers were forming [67]. The antarctic snow becomes compacted into ice at a rate of about 2 cm a year. The work by Aartsen et al. [65] also details an azimuthal anisotropy. Previously in the Spice-Mie model the ice was assumed to be uniform in azimuth with the properties only possessing a depth dependency. It was found that the light scattering has a preference in azimuthal direction possibly due to orientation of the dust particles in the ice. This effect has now been taken into account in a newer model referred to as Spice-LEA and is compared to the old model in Fig. 2.5.

The final notable impurity to the ice comes from the installation of the detector itself. In order to install the instruments in the ice, it was essential to drill large holes to insert the instruments on long cables down to 2450 m below the surface. During this process the ice was first drilled and then melted in a diameter of approximately 60 cm. Once the instrumented cables were installed the melt water was left to refreeze around the instruments. The refrozen ice in the drill hole is referred to as *hole-ice*. In order to best understand the qualities of the *hole-ice*, a camera system was installed on string 80 to monitor the refreezing of the ice. It was found that the ice on the edge of the hole freezes in the first few days. The hole becomes completely frozen after 15 days and the inner 16 cm layer contains air bubbles pushed to the center during the refreezing process. This results in much shorter scattering lengths in this region compared to the bulk ice [68].

2.5 Event types

Events in the detector fall into two main categories according to their topological differences, tracks and cascades, depicted in Fig.2.6. Tracks arise from muons traversing the detector originating from CC ν_μ interactions or from the atmospheric showers. Cascades



(A) Cascade Example event

(B) Track Example Simulated Event

FIGURE 2.6: Example of a cascade event observed by the detector (A) and a simulated track event from a CC ν_μ interaction in the ice (B).

are more spherical originating from NC interactions of all flavor neutrinos or from ν_e and ν_τ CC interactions.

Track Events

High-energy muons are able to travel distances of the order of kilometers before interacting or decaying, which allows them illuminating long sections of the detector. These muons are produced either by CC ν_μ interactions in the ice or by CR interactions in the atmosphere, which is a large source of noise for neutrino detection. Atmospheric muons at TeV energies traverse the deep polar ice losing energy primarily through continuous ionisation and stochastic losses. In practice the stochastic losses take place in the form of brehmsstrahlung, photo-nuclear scattering and pair production. The energy losses are then approximated as :

$$-\left\langle \frac{dE_\mu}{dx} \right\rangle = a + bE_\mu, \tag{2.7}$$

where a accounts for ionisation losses and b accounts for the sum of stochastic radiative processes. As such the energy can be estimated based on the observed energy losses across the detector. The long lever arm of the muon track is also critical for the muon direction reconstruction. Track-like events typically can be reconstructed to a precision of $< 1^\circ$ depending on the topology and energy.

Another advantage of track events is that the effective detector volume is increased by the ability to observe muons from neutrinos which interact outside of the detector. However, the muons producing such tracks are not entirely contained within the detector. In the cases where the muon is produced outside of the instrumented space, the distance

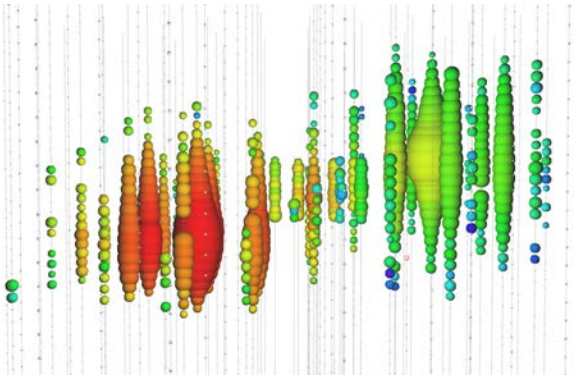


FIGURE 2.7: Simulated *double-bang* event from a charged current ν_τ interaction.

that the muon has traveled already is unknown. This has the effect that energy reconstructed from 'visible energy' in the detector is a lower limit to the real muon energy and so the estimated neutrino energy is affected by the missing energy of muons outside of the instrumented region.

Cascade Events

When electrons or tau particles are produced by corresponding neutrinos in the detector they have an interaction length typically shorter than the distance between 2 detector strings. This means that only the electromagnetic shower, as opposed to the individual lepton, is observed which produces Cherenkov light that expands spherically from the original interaction vertex. Similarly for NC deep inelastic scattering, a hadronic shower is also produced and is centered on the interaction point of the ν. The number of light is proportional to the produced number of shower particles which reflects the original ν energy for CC ν_e and ν_τ events. NC events behave similarly but usually with only a fraction of the original ν energy since the scattered neutrino escapes with the remainder. The size of the event scales logarithmically with the energy such that even ultra-high-energy events can be contained in the detector. In cases where the showers are contained, the energy can be reconstructed with a precision of the order of 10% [69]. However the spherical symmetry means that the direction reconstruction, based on the timing information, is relatively poor. The direction precision for cascade events is often as large at 15° or 20°. In the event of a ν_τ interaction in the ice, a high-energy τ lepton is produced with a very short lifetime. Typically these τ particles decay immediately to produce a quark-antiquark pair or a lepton pair, with a probability 0.65 and 0.35 respectively. The τ lifetime is about 100 m above PeV energy. Below this energy, the shower produced in the DIS interaction by the ν_τ and the shower produced in the decay of the τ cannot be separated in IceCube,

because the photomultipliers in the detector are not close enough. As a matter of fact the τ decay length needs to be comparable to the distance between the detector strings in order to distinguish the two cascades and distinguish ν_τ from a typical cascade event from NC or ν_e interactions. However, above PeV energies τ leptons are capable of traveling a detectable distance in the detector before decaying. This is done by separating the showers stemming from the interaction point of the DIS and the τ decay. In the case of a CC ν_τ interaction above PeV energies, if the decay vertex lies outside of the instrumented region this event becomes indistinguishable from a muon track in the detector but if both vertices occur within the instrumented region of ice this event is referred to as a *double bang event* simulated in Fig. 2.7. A double bang event is expected to be dominated by astrophysical signal since the expected background of atmospheric ν_τ above PeV energy is negligible. However a double bang event is also possible from other neutrinos. For Glashow events from a $\bar{\nu}_e$, a $W^{\pm}$ could decay into a τ lepton leaving a similar double bang in the detector. Observing ν_τ interactions would also allow more precise measurements of the flavour composition of the astrophysical neutrino flux.

2.5.1 Atmospheric Background Flux

IceCube point source analyses are optimized for well reconstructed muons passing through the detector. For this reason they are also sensitive to the background of high-energy muons which arrives at the detector (see Figure 1.6). High-energy atmospheric muons and neutrinos are produced in particle showers originating from high-energy CRs interacting with atmospheric atoms and molecules. The reactions are similar to the expected interactions surrounding a CR source which would result in an astrophysical ν flux: protons or ionized nuclei interact with other protons and resulting mesons decay into neutrinos and muons. At the astrophysical source the energetic charged protons or nuclei can also interact with the ambient radiation around the source.

IceCube is between 1.5 and 2.5 km below the surface of the ice at the South Pole and it requires a correlated signal across multiple light sensors to trigger an event. This means that the dominant background events are atmospheric neutrinos and muons which are able to penetrate the detector. Such neutrinos and muons are from high-energy CRs arriving at the atmosphere and they can be very difficult to differentiate from an astrophysical neutrino arriving directly from a source and interacting in the ice inside and surrounding the detector. Methods for filtering out such atmospheric events will be described in more detail in Sec. 3. A single high-energy CR will produce many muons and neutrinos. Atmospheric muons can often arrive in bundles. Since these are high-energy charged particles, they will be easily observed from radiation losses upon traversing the detector while most of the

neutrinos do not interact. The background for IceCube is then significantly effected by the rates of muons able to reach the detector. At trigger level IceCube detects about 2.7 kHz of atmospheric muons.

In the northern hemisphere the atmospheric muon flux is almost completely eliminated because muons cannot pass through large distances of matter without decaying or interacting. The distance that must be traversed through the Earth for particles to arrive at the detector is maximum for particles originating from the north pole where the zenith is 180° and minimum at the south pole, where zenith is 0°. Although many atmospheric muons are able to penetrate the 1.4 km of ice from the south pole to reach the detector, none of them are able to traverse the entire diameter of the Earth from the north pole. The horizon region at zenith $\sim 90°$, is the critical region where the distance between the surface of the earth and the detector becomes too large for the atmospheric muons to traverse. Hence, it is this region where the hemispheres are divided, indicating a transition from a background dominated by atmospheric muons to a background dominated by atmospheric neutrinos. Truly atmospheric muons may remain in an event selection for the northern hemisphere if the event direction is poorly reconstructed. Hence, by selecting for high quality tracks and high energy events, the background in the northern hemisphere becomes dominated by atmospheric neutrinos which are not impeded by the Earth.

The only method of differentiating atmospheric neutrinos from astrophysical neutrinos is to examine the event energy where the typical energy fluxes are shown in Fig. 1.7, and to search for a coincident muon which could have resulted from the same atmospheric shower. Consequently, at TeV energies and in the absence of a coincident muon, an individual event cannot not be identified as signal or background. The *signalness* is greater at higher energies, however the Earth becomes effectively opaque to neutrinos at energies of about 1 PeV since the neutrino-nucleon cross section increases with neutrino energy. Although the down-going tracks from the southern hemisphere are dominated by atmospheric muon and neutrino background, these events have a softer flux than the astrophysical neutrinos, that is, the atmospheric muon rate rapidly decreases as the energy increases. Fortunately the thickness through which the particle passes in the southern hemisphere is not sufficient to render it opaque to neutrinos, and highly energetic neutrino events can interact in the detector. As a result, selecting energies beyond an energy cut value can eliminate more background than signal in the event selection.

The different energy spectra between atmospheric muons and astrophysical neutrinos also applies to atmospheric neutrinos, known as the conventional flux. The atmospheric neutrinos come from the decay of high energy pions and kaons from a CR air-shower. At energies below around 100 TeV this follows closely the CR spectrum as a power law with an index of -2.7. At higher energies the mesons are able to travel further which gives them

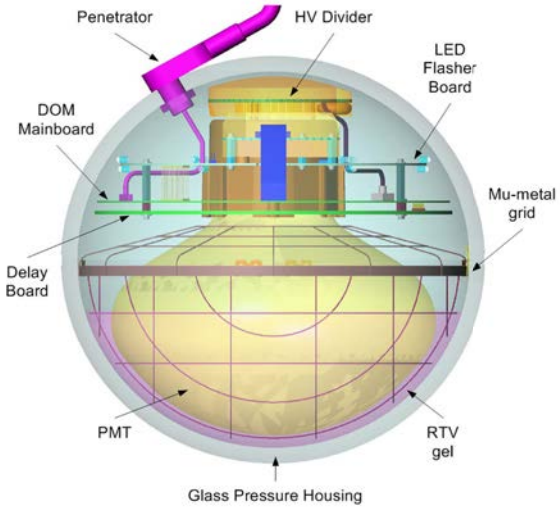


FIGURE 2.8: Diagram of a Digital Optical Module (DOM) installed in the ice. Each DOM is able to make light deposit measurements, run calibration tests, and send and receive information from the detector base.

a greater chance of interacting before decaying to produce neutrinos; this has the effect of steepening the spectrum to a spectral index $\gamma \simeq 3.7$ [70]. Given a diffuse astrophysical fit power law flux with a spectral index $\gamma = 2.13 \pm 0.13$ [10], a greater fraction of astrophysical events is expected at higher energies.

2.6 Data Acquisition

As discussed in Section. 2.2, the key ingredient to indirect neutrino detection is the observation of Cherenkov radiation in the ice. This light is measured at a series of positions in a km^3 of ice through the use of DOMs that were installed there.

2.6.1 Digital Optical Module (DOM)

The DOM, pictured in Fig. 2.8, is a piece of almost autonomous equipment installed in the ice capable of making light deposit measurements, running calibration tests and sending and receiving information from the detector base in a short time scale. In order to perform the essential role of data acquisition the DOM contains a 10" Hamamastu PhotoMultiplier Tube [71] (PMT), which converts the Cherenkov radiation on the level of individual photons into an analog waveform representing the PMT electrical current. The PMT is arranged facing downwards in the DOM with a 13 mm layer of protective gel, which serves as an optical coupling medium as well as mechanical support between the PMT and the DOM's

spherical glass encasing. The DOM casing is built to survive the high pressure deep in the ice withstanding 250 bar of long term pressure and 690 bar temporary over pressure. There is also a μ-metal cage of wires around the PMT bulb to reduce effects of the earth's magnetic field at the South Pole by a factor of 2.8 whilst having a minor impact on the effective light yield. The PMT orientation was chosen to be optimal for signal arriving from below the detector through the centre of the earth. This is the direction from which IceCube is expected to have optimal background to signal differentiation.

The PMTs are designed to detect photons with energies of 10 Gev to 10 PeV up to 500 m away, a larger distance than the distance between the detector strings. The PMTs in place are powered by high voltage. The PMTs feature a box-and-line 10-stage dynode-chain, which is operated at a gain of 10^7. This device demonstrates a peak quantum efficiency of around 25% for the standard DOMs and 34% for the high-quantum-efficient (HQE) DOMs for wavelengths near 390 nm. Once a photon enters the DOM, the PMT must register a signal greater than 0.25 photo-electrons (PE) which triggers an initial IceCube "hit". The PMT hit time and PMT waveform is saved to allow the collected photon data to be compared to this reference time. After collecting data for approximately 1 s, the data is sent up to the IceCube Lab (ICL) directly above the detector. The PMT hit rate is also monitored independently by the DOMs periodically every 1.6384 ms, which is intended to be sensitive to a sudden increase in low energy neutrino events generated by the electron neutrinos emitted by Galactic supernova, which produce positrons above the Cherenkov threshold emission. The PMTs constantly observe noise dominated by dark noise, where no external photon has stimulated electron emission. This noise rate occurs at a rate of about $\sim$ 300 Hz for temperatures between -40°C and -20°C, which has been minimized by using low potassium glass which has minimal radioactivity.

The DOMs are powered by and transfer the data through twisted wires from pairs of DOMs which run up the cables to the ICL. There is also a separate wiring between neighbouring DOMs that allows Local Coincidences (LC) to be determined between neighbour or next to neighbour DOMs along a string within $\pm 1\mu$s. As well as a PMT for signal detection, and LEDs for detector calibration, there must be a local computer capable of processing and digitizing this information before transmitting it to the surface computers. These steps are implemented on a DOM main board which contains digitizers, a computer, communications circuits and a high-voltage supply necessary for the PMT. The analogue PMT signal is sent through 4 different paths. The first path is passed through a variable discriminator and is used to evaluate if the PMT signal amplitude passes the threshold required for a *hit* determining if the signal is digitized.

If an analog signal passes the 0.25 photoelectron (PE) threshold, the PMT signal continues along 4 different paths. Three of these paths run after an enforced delay of 75 ns

and they undergo digitization by Analogue Transient Waveform Digitizers (ATWDs). The ATWD samples the PMT amplitude at rate of 300 MHz for 426 ns, limiting the longest waveform to this duration. During the digitization process, the ATWD is unavailable to process future events, creating dead time in the DOM. To compensate for this 2 ATWDs are in place reading the signal alternately removing any possible dead time. Three different paths are used here because the ATWD has a very limited resolution. The waveform is amplified by 0.25, 2, and 16 on the 3 different paths, which allows it to cover the entire dynamic range of the PMT. The smaller amplification allows sensitivity to larger amplitudes and vice-versa. Finally the signal is also processed by a Fast Analog to Digital Converter (FADC) which covers longer waveforms than 426 ns. To allow for this the FADC samples at a lower rate of 40 Mhz over 6.4 μs. The combined digitized waveform is referred to as a *DOMLaunch*.

2.6.2 IceCube Trigger

IceCube DOMs detect hits at a rate of $\sim$ 600 Hz per DOM mostly due to background effects. The largest background effect on the individual DOM level is that of dark noise in the detector when the PMT emits electrons from the cathode even when there is no external photon stimulating the emission. This noise has many possible sources including electronic noise, field emission in the PMT, thermionic emission or radiation from radioactive decays in the DOM materials. Since this noise does not originate from any signal external to the DOM, it is expected that the noise between DOMs is entirely independent. Using this assumption, the rate of dark noise is radically reduced by requiring hits within $\pm 1\mu$s to be neighbours or next-to-neighbours to be classified as a Hard Local Coincidence (HLC). Other hits are considered Soft Local Coincidences (SLC) and for these hits only the time-stamp and minimal charge information is transmitted to the DOM-Hub. For HLC events the PMT waveform data is stored in full detail and sent to the surface computers at ICL.

Online Filtering

HLC events are further filtered and analyzed before being compiled into possible neutrino events. This filtering process involves testing many trigger algorithms looking for temporal, or less commonly, spatial coincidences between the HLC hits. All hits passing the triggers within a certain time window are compiled into an event. The result is a Level 1 median event rate of 2.7 kHz which varies seasonally. This yields a DAQ data rate of approximately 1 TB per day. Due to the remote location of the IceCube Lab, the data collected at the South Pole must eventually be transmitted to the Wisconsin IceCube Particle Astrophysics Center (WIPAC) for final processing and analyses. The data is sent via satellite from the

South Pole and this transmission is limited to roughly 100 Gb per day. That means the data rate at ICL must be reduced through the use of multiple filters, each optimized for different target event selections. Due to the high rate of events requiring processing at ICL, the reconstructions applied are limited in terms of time and computing resources and so they are fairly basic. This filtering is referred to as online filtering as it is conducted in real-time at the South Pole and the filtered and reconstructed events are sent directly to the servers in the northern hemisphere. This includes multiple alert triggers for events likely to be of astrophysical interest, in practice these very high energy events with either track topology or with a starting vertex inside the detector. Sent along with the events is also information on the configuration of the detector for each 8h-period, known as a run, when the data were taken. This is an essential ingredient during the event reconstruction that allows the data to be properly calibrated using the detector monitoring information at the time the data were taken.

Further offline filtering with fewer constraints on time and resources is then conducted on the data after is stored at the servers in the northern hemisphere for different data selections defined for different analyses. This will be discussed further as the Event Selection in Chapter 3.

2.7 Reconstructions

At the final level of a point source analysis, only the energy and direction variables are used to characterize the signal and background in the likelihood probabilty density functions (pdfs) as will be discussed in Chapter 4. As such they are the only information used to establish to what degree an event can be associated with a particular source location and how much it resembles astrophysical signal. The only data available from the detector is the energy deposited and PMT waveforms across the matrix of DOMs with the relative timing of these deposits. It is therefore essential to use this information to reconstruct the best estimate of the original neutrino energy and direction. The capabilities of these reconstructions depend heavily on the kind of event which has occurred in the detector.

2.7.1 LineFit

An initial direction reconstruction can be quickly established by making rudimentary assumptions about the event in the detector which do not account for the geometry of Cherenkov radiation, but instead treat the light emission as a plane wave perpendicular to the produced lepton, ignoring detailed local ice properties. The best fit is then

determined by minimizing χ^2 [72], defined as :

$$\chi^2 \equiv \sum_{i=1}^{N_{hit}} (\mathbf{r}_i - \mathbf{r} - \mathbf{v} \cdot t_i)^2 \tag{2.8}$$

where the light travels at speed $\mathbf{v}$ through the ice to reach a DOM at the position $\mathbf{r_i}$ at time t_i for the the sum of the hits included in that event. The free parameters for this minimization are the vertex point $\mathbf{r}$ and the Cherenkov light velocity $\mathbf{v}$ which is assumed to be the same magnitude as the incoming muon. This minimization can be solved quickly analytically to give a resolution within a few degrees of the true neutrino direction. Improvements have also been put in place which penalize late pulses as well as hits from DOMs too far away to contribute to the event [73].

2.7.2 SplineMPE

Improved reconstruction techniques are able to converge more precisely on a simulated true neutrino direction when taking into account the geometry of Cherenkov radiation and the ice properties [72]. For these methods an initial seed direction is used, such as the LineFit reconstruction described above, and then the track parameters $\vec{a}$ are estimated by maximizing the likelihood $\mathcal{L}(x|\mathbf{a})$ with x. Where:

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

where $p(x_i|\mathbf{a})$ is the pdf for the observations $\mathbf{x}$ given true event parameters $\mathbf{a}$.

In this case the track parameters are:

$$\mathbf{a} = (\mathbf{r}_0, \vec{p}, t_0,) \text{ with } \mathbf{r_0} = (x_0, y_0, z_0) \text{ and } \vec{p} = (\theta, \phi) \tag{2.10}$$

Where $\vec{r_0}$ and t_0 represent a position and respective time for an arbitrary point along the track and $\vec{p}$ the track's direction along the detector. This parameterization is simplified by assuming that the Cherenkov radiation is emitted as a cone uniformly along the path of a muon track.

For the simplest likelihood, only the first measured photo-electron arriving at each DOM is used, under the assumption that the first PE has undergone the least scattering. This is known as single-photo-electron (SPE). Given the timing of these photons along the array of DOMs an analytical function describes the timing distributions more accurately than a simple LineFit. By convolving this function with a Gaussian noise factor, jitter and dark noise are also accounted for [74].

For high energy events, where each DOM can detect multiple photons, it can be more useful to take into account the scattered photons which arrive with a slight delay to gain more information about the overall event, such as the distribution of deposited energy. This is taken into account using a modified likelihood known as multi-photo-electron model (MPE). This model uses the photon arrival time pdf to incorporate the information on the number of photons arriving at each DOM. For this model it is important to take into account the ice features which would influence light scattering. This scattering has a direct effect in the observed delay of photons arriving from the same event. These ice features are taken into account by using a complete depth dependant ice-model with a photon transport simulation [62]. The photon simulation is implemented by tabulating the timing distributions and light yield from a range of different configurations. Originally the binned results were used directly leaving the results sensitive to binning effects. To handle this issue without having increasingly fine binning that would be impractical to store, a multi-dimensional spline is fit to the simulated photon tables using a photo-spline package [75]. This step analytically parameterizes the photo-spline tables and can be used directly as pdfs in the MPE likelihood, currently known as SplineMPE.

2.7.3 Paraboloid

Given a neutrino source, it is intuitive that most of the observed neutrinos originating from this source will not be perfectly reconstructed with the exact source location. In fact it is expected that they will be reconstructed in a distribution around the true source position with an associated error. For a point source analysis the error for each event is vital to determine, within a certain confidence level, the region in which one would expect to find the source of that event.

The methods for determining the angular error in direction [76], known as paraboloid, require a likelihood map and a best fit location from this map for each event to be already calculated. Using these outcomes, an ellipse is fit to the likelihood map that best describes the 1σ contour around the best fit position. The axes σ_1 and σ_2 are then combined to give and average error estimated circular error:

$$\sigma_{err} = \sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}}. \tag{2.11}$$

Before implementing this error in the likelihood calculations, it was compared to the true angular difference between the true and reconstructed direction from simulations and found to be systematically underestimated. One such cause for these under estimations is that the error estimate on the likelihood map only accounts for the error in the muon track

reconstruction and ignores any additional potential offset between the muon track and the original neutrino direction. At energies below 1 TeV, this effect can become a dominant factor compared to the track reconstruction resulting in a significant underestimated direction error. To account for such effect the ratio between the track error and the true error is fit as a function of energy and used to rescale the estimated error as a function of energy to take into account kinematic errors.

2.7.4 Energy Reconstruction

When analysing a selection dominated by muon tracks traversing the instrumented ice, it is very difficult to have an accurate estimate of the true neutrino energy. As previously discussed, when tracks traverse the detector the distance travelled since a possible neutrino interaction is not known. It is possible for a neutrino to interact producing a muon very far from the instrumented ice therefore radiating most of its energy before being detected. Hence, the energy estimate for through-going tracks acts as a lower limit on the neutrino energy. Cascade events which are mostly contained in the detector are an exception to this and the energy can be determined much more precisely than for muon tracks up to an accuracy of about 15%.

For the events in this selection the energy estimate is made assuming that the light yield from the detected photons is directly proportional to the event energy [69] and can be estimated by using a template of an expected energy for a given light yield. The estimated energy scaling factor can be determined as:

$$\lambda = \Lambda E + \rho. \tag{2.12}$$

where Λ is the number of photons in the template, λ is the mean number of photons for an event with an energy E times the template energy and ρ is the contribution from noise hits.

Given the relatively low number of predicted photon events per DOM, the likelihood can be constructed assuming Poisson statistics such that the likelihood $\mathcal{L}$ of observing k_i photons at the i^{th} DOM is:

$$\begin{aligned} \mathcal{L} &= \prod_i \frac{\lambda_i^{k_i}}{k_i!} \cdot e^{-\lambda_i} \\ &= \prod_i \frac{(\Lambda_i E + \rho)^{k_i}}{k_i!} \cdot e^{-(\Lambda_i E + \rho)} \end{aligned} \tag{2.13}$$

This likelihood is then maximized to give the most probable energy scaling E given the observed number of photons at each DOM. This is referred to as the MuE algorithm. This algorithm has since been improved by applying a kernel function $G(\lambda_i, \lambda_j)$ to accommodate more thoroughly the statistical uncertainties on λ_i. More information on the kernel is provided by Aartsen et al. [69]. The result is a modified likelihood:

$$\mathcal{L} = \prod_j \int_0^\infty G(\lambda_i, \lambda_j) \frac{\lambda_i^{k_i}}{k_i!} \cdot e^{-\lambda_i} d\lambda_i \tag{2.14}$$

which demonstrates a 35% precision in $log_{10}E$ for muon detector entry energy of 10^4 GeV which improves to about 30% in $\log_{10} E$ at 10^6 GeV.

2.8 Event Simulation

Once a basic detection method has been put in place it is vital to be able to optimize the event processing and filtering techniques for the signal one intends to observe. Without real data examples of such a signal, one is forced to simulate these examples. These Monte Carlo simulations have been built over years to take into account many factors, primarily concerning:

- Atmospheric neutrino and muon production from CR showers striking the atmosphere.
- Neutrino and Muon propagation through the Earth and its atmosphere.
- Neutrino and Muon propagation through the polar ice to the detector.
- Cherenkov photon generation and propagation.
- Detector response to the radiated light.

The simulation is conducted by *neutrino generator* (NuGen), a modification on the code "All Neutrino Interaction Simulation" (ANIS) [77]. This simulation starts from the Earth's surface injecting neutrinos with a fixed hard spectrum and following step by step the propagation and interactions through the profile of the Earth using the Preliminary Earth Model [78]. These neutrinos are capable of passing unaffected, oscillating in flavour, or having a few neutral and charged current interactions within the matter they pass. The rates of these processes are dependent on the neutrino energy and the matter and relative distances they pass through. At the high neutrino energies of interest for this analysis, charged current neutrino-nucleon interactions can result in considerable losses in the flux

making the simulation process inefficient at creating measurable event rates at the detector. To compensate for this, charged current interactions in the Earth far from the detector are forbidden but the neutrinos are assigned a *Propagation weight* to correct for this. Similarly it is known that the cross section for neutrino interactions is very small, so generating a reasonable number of events in the detector would mean generating enormous rates of initial simulated neutrinos thus, within the detector region, the neutrinos are forced to interact. These forced interactions are then associated with an *Interaction weight* to correct for the fact that this simulated interaction actually has a small probability of occurring. In order to estimate the initial neutrino flux from an observations it is key to take into account both the *Interaction weight*, the *Propagation weight*, as well as the detector volume in which the events can be observed and finally the generated simulated neutrino spectrum. All these factors are combined by Nugen into a final weighting called *OneWeight*. The OneWeight is applied to the simulated events along with the desired signal or atmospheric spectra hypothesis to estimate the final observable flux. This simulation generates equal rates of neutrinos and antineutrinos of the desired flavor.

When training and estimating the best event filtering techniques it is also critical to be able to simulate the expected background. In IceCube at high energies this is principally atmospheric neutrinos and muons. Atmospheric neutrinos are simulated using NuGen and weighted according to the best known atmospheric spectrum. Atmospheric muons are then generated by Cosmic Ray Simulation for Kascade (CORSIKA). This simulation starts with a CR nuclei from Hydrogen, Helium, Nitrogen, Aluminum, and Iron, and generates the full particle shower, including the essential muon component IceCube would be sensitive to.

Given a muon flux from the atmosphere or from charged current neutrino interactions, it must be estimated how these muons would be observed in the detector. This is done by tracking the propagation through the ice with Muon Monte Carlo (MMC [79]). Models are used for the Cherenkov light yield from radiative losses of the muons traversing the instrumented ice. The light from the charged particles must then also be propagated until it reaches a PMT in a DOM. This propagation is done with CLSim [80] or Photon Propagation Code [81]. The light yield arriving at the PMTs depends on the properties of the ice for which the best available ice model is used, or for high energy cascades which are more sensitive to the ice properties, spline light propagation tables are used [75].

The final separate step is to simulate the response of the detector to such an incoming light yield and apply the real-time trigger filters which are applied immediate to the real data. This also includes noise from the PMT, the electronic readouts and the coincidence requirements. Once this has been applied all the remaining processing is applied equally to data and simulation. Only after all these steps the analysis can be replicated with simulation of the true signal and background.

Chapter 3

IceCube Event Selections

As discussed in section 2.6, online filters at the South Pole during the data acquisition filter the event selection reducing the data rate to $\sim$ 100 Gb per day in order to be transmitted to the northern hemisphere. However, the data sent to the northern hemisphere are still vastly dominated by background events, in particular atmospheric muons. Even for events expected to have passed through the Earth, a large muon contamination comes from mis-reconstructed event directions. The event selection, which is applied to identify possible astrophysical neutrino sources, should comply with the following criteria: a) maximization of the number of astrophysical neutrino events in the analysis, b) minimization of the contamination of background events, and c) high quality direction reconstructions.

The neutrino flux from astrophysical sources is expected to have a harder spectrum than the atmospheric background flux as shown in Fig. 1.7. The sources that make up this flux are still unknown. The final analysis for this selection is implemented to discriminate between astrophysical events around a source and the atmospheric background sensitive to a clustering of many events around a precise location. Thus a precise direction reconstruction is a critical parameter for this analysis.

The selection will be applied to the full lifespan of the detector to date, which is composed of over 10 years of data from April 2008 until July 2018. The first three years each used a partial configuration of the detector (40, 59, 79 strings) and the first year of the full detetor (86 years) used a different filtering applied to the events than the remaining 6 years of the full detector data. These samples are kept exactly as they were used for the previous all sky point source searches and are summarized by Aartsen et al. [6] and in the table 3.1. From 2012 on-wards, the detector configuration and initial software processing is mostly unchanged. An optimized event selection and reconstructions were applied to this six year sample and will be described in the following chapter.

In section 2.6 the basic filtering steps were described, from light triggering the detector to the data transmission to the northern hemisphere. The the data is stored in servers in the northern hemisphere where further processing reduces the sample using a filter for

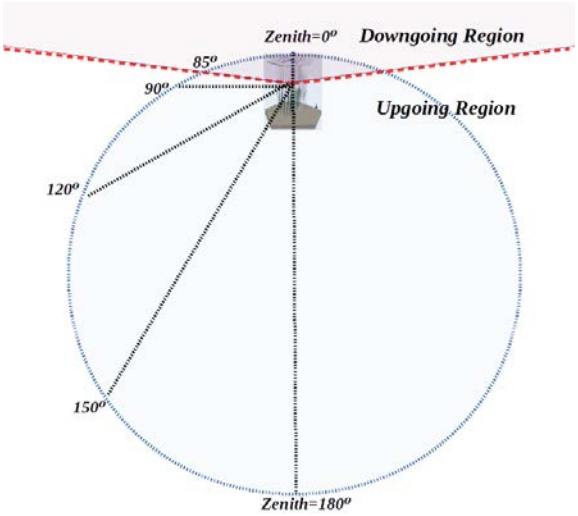


FIGURE 3.1: Cross section of the Earth showing the South Pole with the IceCube detector (at the top of the diagram). The dotted line separates the down/up-going regions: the down-going region corresponds to the southern hemisphere with zenith below 85°; the up-going region corresponds to the northern hemisphere above 85°, these events must pass through a much larger fraction of the Earth.

muon track-like events. Such an event selection decreases the data rate to a few Hz, and can be optimized further by various analyses with a range of criteria. This data level is referred to as Muon Level 3 and has been used in multiple muon track analyses looking for both a diffuse and a point-source astrophysical neutrino flux [6, 10]. Once the data rate has been reduced to Muon Level 3, it becomes feasible to study the improvements of more intensive cuts and reconstructions. This thesis presents the work done to optimize and update the previous final filtering applied after Muon Level 3. This led to a $\sim$ 35% improvement in the analysis sensitivity when combined with the 3 extra years of data. A different filtering is applied to the northern and southern hemisphere as the composition of event types is dramatically different, hence the expected background rate and sensitivity changes comparably. Events which arrive from the northern hemisphere, i.e. with a zenith angle greater than 85°, must first pass through a large fraction of the Earth (depending on their original direction). These events are seen to arrive from the bottom of the IceCube detector and are called *up-going events*, while events which arrive from the southern hemisphere are called *down-going*. For example events from the north pole must pass through almost the entire Earth, as can be seen in Fig.3.1. The earth acts as a shield removing muons, hence the only atmospheric muons able to reach the detector are those from the southern hemisphere that can penetrate the distances on the order of only a few km through the earth. Atmospheric muon events observed from the up-going region are only those which

have been mis-reconstructed. Often, misreconstruction is due to two cosmic rays producing events in the same data acquisition time frame. These are called coincident events.

Most of the criteria in both hemispheres are designed to select high quality single tracks traversing large sections of the detector. This reduces the confusion with cascade events which have minimal pointing capabilities, and atmospheric events that often arrive at the detector in bundles. However, the cuts in the northern hemisphere are generally looser as the Earth already acts as a powerful filter on the atmospheric muon background.

3.1 Updated Direction Reconstruction

The final data processing implements an updated direction reconstruction to both hemispheres inspired by the sensitivity improvements observed using a northern tracks only selection [82]. In the previous point source event selection, the direction reconstruction from Muon Level 3 is taken as the final event direction used in the analysis. This direction reconstruction uses the Spline MPE direction reconstruction described in section 2.7.2 using the results of LineFit as a seed. The major difference for the new selection with respect to the previous selection is that the Spline MPE reconstruction is applied again using the previous direction likelihood map from Spline MPE as a seed and this time also including the energy estimation. This has the effect that the reconstruction has not only more precise seed information but also more event information available to make a new direction fit. This is useful as it is known that the angular resolution can depend on the event energy. Once a new event direction is reconstructed, the estimated uncertainty for this direction must also be recalculated using the updated SplineMPE direction likelihood map.

The improvement in the angular resolution is shown in Fig. 3.2 compared to the most recent all-sky event selection. The angular difference between the direction of the reconstructed muon track and the true original neutrino direction is calculated as $\Delta\Psi$. This reconstruction is highly energy dependent, so instead of calculating a single average $\Delta\Psi$ for the weighted simulated neutrino events, the median $\Delta\Psi$ is calculated for different bins of the true ν_μ energy. The average improvement in angular resolution is $\sim 15\%$ which increases at higher energies up to almost 30% reaching a median resolution of 0.2° as can be seen in Fig. 3.2. In addition a minimum floor to the angular uncertainty has also been applied at 0.2°. This has been applied to prevent some tendency of the reconstruction to produce angular uncertainties that are too small.

The improvement in the angular resolution compared to the past for this selction seen in Fig. 3.2, should cause events to be more closely clustered around the true source direction. The expected effect is to improve the sensitivity for finding point-like neutrino sources

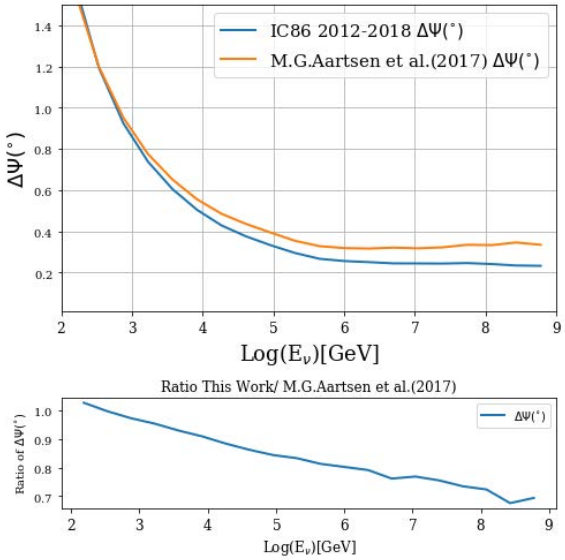


FIGURE 3.2: The median angular difference between the true neutrino and reconstructed muon track direction as a function of energy for the new data selection compared to the previous 7 year analysis [6].

among background dominated data, and the estimated direction of such a source would be more precise.

3.2 Up-going/Northern Selection

After the Muon filtering selection at level 3 a selection is applied consistent with that described in Aartsen et al. [82].

An initial criterion which is common for both the southern and northern selection is a maximum length of the reconstructed track without observed hits at the DOMs. For a high energy track emitting radiation almost uniformly along its path, one would expect to observe radiation at the DOMs closest to the track, with a maximum distance given by the separation between DOM strings (of the order of 125 m). In reality a less energetic track may not be sufficiently luminous to trigger hits on every single DOM, especially if it arrives at a less efficient orientation with respect to the DOM strings and PMTs. It should also be recalled that there is a *dust layer* in the ice, discussed in Section 2.4, which can inhibit the detection of radiation. Therefore, there is a cut on the maximum length of empty track ($l_{empty} \leq 400$ m) for both hemispheres, chosen to filter out false tracks but which still allows for dimmer tracks passing through the detector at a non-optimal orientation.

Another track length quality cut which is fairly consistent in both hemispheres is a cut on the minimum total track length l_{dir}. In the northern hemisphere a track is removed if it has a length reconstruction less than 200 m within a certain time window. This filter is even stricter and raised to 250 m in the southern hemisphere.

It was already mentioned before that atmospheric events can often arrive in the same readout window as coincident events. An optimised procedure to remove such events which are often mis-reconstructed involves splitting the initial track in half, then reapplying the track reconstruction algorithms using only the DOMs from the respective half of the detector and comparing the track reconstructions. For a high-quality track, both reconstructions should lie almost parallel, whereas coincident events can cause the reconstructions in each half to differ drastically. An event passes the selection if $\cos(\theta^2_{geo}) < 0.2$, where θ_{geo} is the angle between the two reconstructed halves of the track.

The final straight cuts are aimed at selecting events sufficiently energetic to be well reconstructed. Firstly, at least 12 DOMs in the detector must register hits for an event, where the number of hit DOMs (n_{DOMs}) is highly correlated to the event energy. DOM hits often occur from light that has scattered in the ice before reaching a PMT. The inclusions of these hits is useful to determine the total energy of the event, but can deteriorate the reconstruction of the muon direction. For this reason, a time residual is defined relative to the most likely start time of an event, and the number of direct hit DOMs (n_{dir}) and *direct* track length l_{dir} are defined as those occurring within a fixed time window such that they are relatively unscattered before reaching the DOMs. In the northern hemisphere there must be at least 6 direct DOMs registered for an event to pass the selection.

These filters are considered *pre-cuts* applied to the events before the final multivariate selection which goes on to remove most of the atmospheric muon background. The summary of these pre-cuts is written as:

- $l_{empty} \leq 400$ m
- $l_{dir} > 200$ m
- $n_{DOMs} \geq 12$
- $n_{dir} \geq 6$
- $cos(\theta^2_{geo}) < 0.2$

.

3.3 Down-going/Southern Selection

The event selection in the southern hemisphere, excluding the updated direction reconstruction, is described in Aartsen et al. [6].

The selection applies a minimum direct track length and maximum empty track length as described above. In addition, an initial estimated uncertainty on the direction reconstruction $\sigma_{paraboloid}$ is calculated as described in section 2.7.3. Those with a $\sigma_{paraboloid}$ above 5° are removed from the selection as they are expected to be lower quality tracks.

The event direction is reconstructed by maximising a likelihood map of possible event directions as discussed in section 2.7. The reduced log likelihood ($Rlog(\mathcal{L})$) is derived from likelihood reconstructions where the maximum likelihood value, $\mathcal{L}$ is divided by the number of degrees of freedom n_{dof}. In this case n_{dof} is $N_{DOMs} - 5$ due to the 5 free parameters in the likelihood fit for a through-going track. This parameter can be used to evaluate the precision of the direction for this event; where $Rlog(\mathcal{L}) > 9$, the events are removed.

Further filters are applied as in the southern hemisphere to ensure enough light has been detected for a reliable event reconstruction. This results in removing events where fewer than 5 different strings detected hits, and where the number of direct DOM hits was fewer than 12.

3.4 Multivariate selection

A multivariate selection was introduced separately as a Boosted Decision Tree (BDT) for the up-going and down-going selections to the Muon level 3 selection after the initial filters described above. The BDT for the down-going selection is described in Aartsen et al. [6], whereas a completely new BDT was trained and applied for the up-going events in the northern hemisphere.

This BDT is trained using simulated ν_μ events after re-weighting their spectrum with an E^{-2} power-law function for the signal class and simulated atmospheric muons and cascade events as background. The principle of the BDT is to apply more optimized selection by exploiting the multi-dimensional phase space of the different event properties for signal and background. Including many variables allows the selection to be sensitive to any unexpected correlations, however given the same sample sizes, increasing the number of BDT parameters will increase the risk of over-training the BDT, for this reason the number of parameters is kept low at 11. Almost all of the variables available refer to the energy or direction reconstruction, either using high level reconstructions or simple fits and basic DOM information.

The following list presents the variables adopted as BDT parameters:

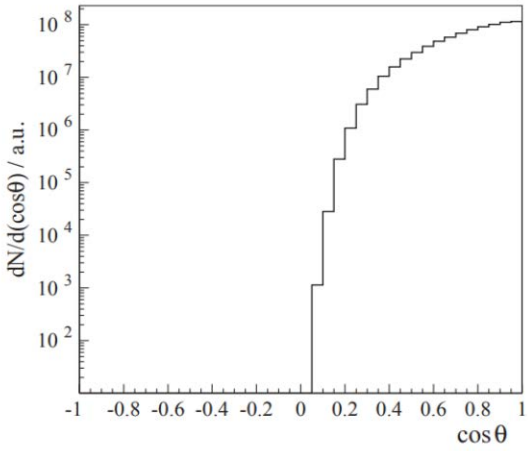


FIGURE 3.3: Bayesian prior for downgoing atmospheric muons from Ahrens et al. [83]

- *COG(ρ)*: The *Center of Gravity* for an event is defined as the energy-weighted central position associated to each event. For a single track or cascade with uniform energy losses, the center of gravity should lie towards the centre of the track or the vertex of the interaction respectively. This COG is defined in cylindrical coordinates with respect to the detector, where ρ is the radial distance in the horizontal plane from the central vertical axis of the detector and z is the depth from the surface of the ice. Each hit DOM making up an event then has a deposited charge q_i. The COG is then defined for general coordinates $\vec{x}$ as :

$$COG(\vec{x}) = \frac{\sum_{i=0}^{N_{DOMs}} q_i \vec{x}_{i=0}}{\sum_{i}^{N_{DOMs}} q_i} \tag{3.1}$$

  This center of gravity can be broken down into its radial and vertical components by applying this definition separately for $\vec{x} = \rho$ and $\vec{x} = z$ to find COG_ρ and COG_z. COG_ρ describes how close the center of the event lies to the central vertical axis of the detector. Events that lie close to the edge or the corners of the detector are

called *corner-clippers* and are often poorly reconstructed due to the limited lever arm available. This information can be useful in distinguishing possible misreconstructed atmospheric muons.

- *COG(z)*: This parameter describes how close the energy-weighted COG for an event lies to the central horizontal plane in the detector using the definition above.

- *Separation Length* l_{sep}: is the distance along the axis of the track between the center of gravity for the first and last quartile of time-ordered hits within a cylinder with a radius 150 m around the track reconstruction. This value is expected to be higher for energetic through-going tracks and smaller for *corner-clippers*.

- *Number of DOMs* n_{DOMs}: This variable describes the total number of DOMs in the detector that register hits associated with this event. This number is highly correlated with the event energy.

- *Bayes Likelihood* $\Delta\mathcal{L}_{bayes}$: This variable is calculated by taking the difference between a maximum likelihood for the final direction reconstruction ($\mathcal{L}_{up}$) and the maximum likelihood fit of the same likelihood direction map but using a Bayesian prior ($\phi(\tilde{\theta})$) on the fit direction $\tilde{\theta}$ taken from Ahrens et al. [83]. The prior, which can be seen in Fig. 3.3 effectively forces a down-going fit, hence the resulting maximum likelihood is referred to as $\mathcal{L}_{down}$. Thus a significantly larger maximum likelihood from $\mathcal{L}_{up}$ than $\mathcal{L}_{down}$ implies the event is much more likely to originate from the northern hemisphere.

$$\Delta log(\mathcal{L}) = log(\mathcal{L}_{up}) - log(\mathcal{L}_{down} * \phi(\tilde{\theta})) \tag{3.2}$$

  The smaller the difference between these likelihoods, the larger the possible error in the direction reconstruction. Thus the higher is the probability that the event could actually originate from the southern hemisphere.

- $cos(\tilde{\theta}_{zenith})$: cosine of the reconstructed zenith angle.

- *Reduced Log Likelihood* $Rlog(\mathcal{L})$: see Section 3.3.

- *Direct Track Length* l_{dir}: see section 3.2.

- *Number of Direct Hit DOMs* n_{dir}: see Section 3.2.

- *Estimated Uncertainty in Direction* $\sigma_{paraboloid}$: see Section 2.7.3.

- *Track Smoothness* s_{dir}: Well reconstructed tracks should deposit energy smoothly across the track path through the detector. One way of quantifying this smoothness

is approximated using the following formula taken from Ahrens et al. [72]:

$$s_{dir} = maximum_j\Big(\frac{j-1}{N-1} - \frac{l_j}{l_N}\Big) \tag{3.3}$$

Where l_j is the length of the track segment from the beginning of the track (of length l_N) to the position on the track where the distance to the j^{th} DOM (out of a total of N doms) is minimized. This variable was motivated by a Kolmogorov-Smirnov test to examine the consistency of any two random samples, in this case the fraction of time ordered hit DOMs and the fraction of track length.

- *Empty Track Length* l_{empty}: see Section 3.2.
- *Average DOM distance* d_{Qavg}: the minimum distance between each hit DOM and the reconstructed track is weighted by the observed deposited energy ($\frac{q_i}{Q_{tot}}$ for the i^{th} DOM) to determine an average DOM distance:

$$d_{Qavg} = \frac{1}{Q_{tot}} \sum_{i=1}^{n_{DOMs}} q_i ||\vec{x}_i - \vec{p}||_{min} \tag{3.4}$$

where $\vec{x}_i$ is the position of the ith DOM whilst $\vec{p}$ is reconstructed track. It is expected for a clear track that the energy will be roughly uniformly distributed around the track path making this weighted average distance fairly small. A large deviation could indicate a poor track reconstruction.

The BDT algorithm is trained on a Montecarlo generated sample of signal and background events. The BDT uses the list of variables above to differentiate signal and background. The sample that is used to train the BDT is referred to as a *burn sample* and is removed from the selection for further analysis. A folding method was implemented by training the BDT on two thirds of the urn sample to test it on the remaining third. This can be repeated rotating which third is used to test the BDT. Consistent results across all the tests demonstrate a lack of bias in the selection known as *over-training*. Over training in the selection procedure occurs when the BDT has to handle too many parameters with a limited statistics. In this case the algorithm learns to separate perfectly the supplied data but then demonstrates a deteriorated performance when differentiating new statistically independent data. One way to avoid this is to limit the number of BDT parameters and to assign parts of the burn sample to training and testing respectively.

The BDT was trained using the ADA-Boost algorithm as part of scikit-learn toolkit [84]. The method used in training is consistent with that applied for the selection used by Aartsen et al. [10] with one major update. Instead of training two separate BDTs designed to

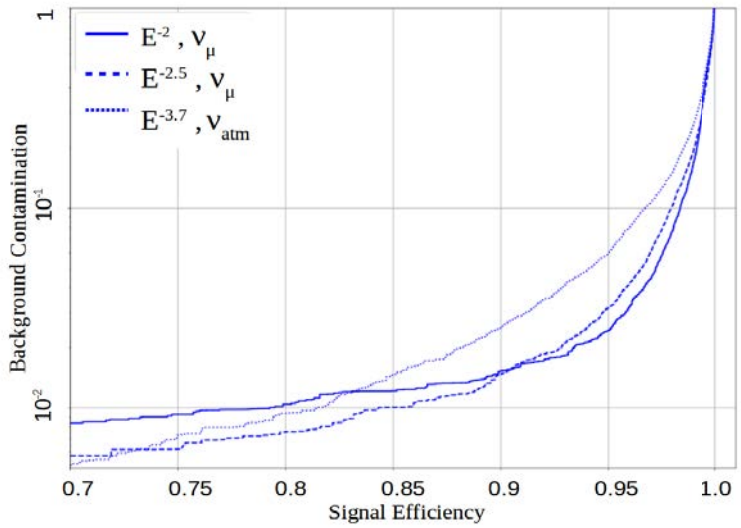


FIGURE 3.4: BDT ROC curve when trained with the signal class at different spectra, E^{-2}, $E^{-2.5}$, and an atmospheric flux.

remove atmospheric muons and cascade events respectively, a single multi-class BDT was implemented with the aim of differentiating between neutrinos, atmospheric muons, and cascades. This is motivated by the fact that all three event types are expected to be present in the sample at all times, although with different abundances.

The classes of events the BDT is trained for are then:

- astrophysical muon tracks: simulated muons from ν_μ charged current interactions weighted for a flux with an E^{-2} power-law spectra.
- atmospheric muon tracks: simulated muons from CORSIKA particle showers generated by CR interactions in the atmosphere.
- cascade events: neutral current ν interactions and ν_e and ν_τ charged-current interactions were used to generate cascade events.

The performance of such a BDT can be estimated using Receiver Operating Characteristic (ROC) curves. For a given trained BDT, a ROC curve shows the background contamination as a function of the signal efficiency of a certain BDT score cut. The efficiency in this case is the fraction of ν_μ tracks that remains after the BDT cut and contamination shows the fraction of total atmospheric muons and cascades that also remain in the selection. The best BDT should maximise efficiency while minimizing contamination. The final BDT version and cut value will depend on the acceptable levels of background contamination and required signal efficiency.

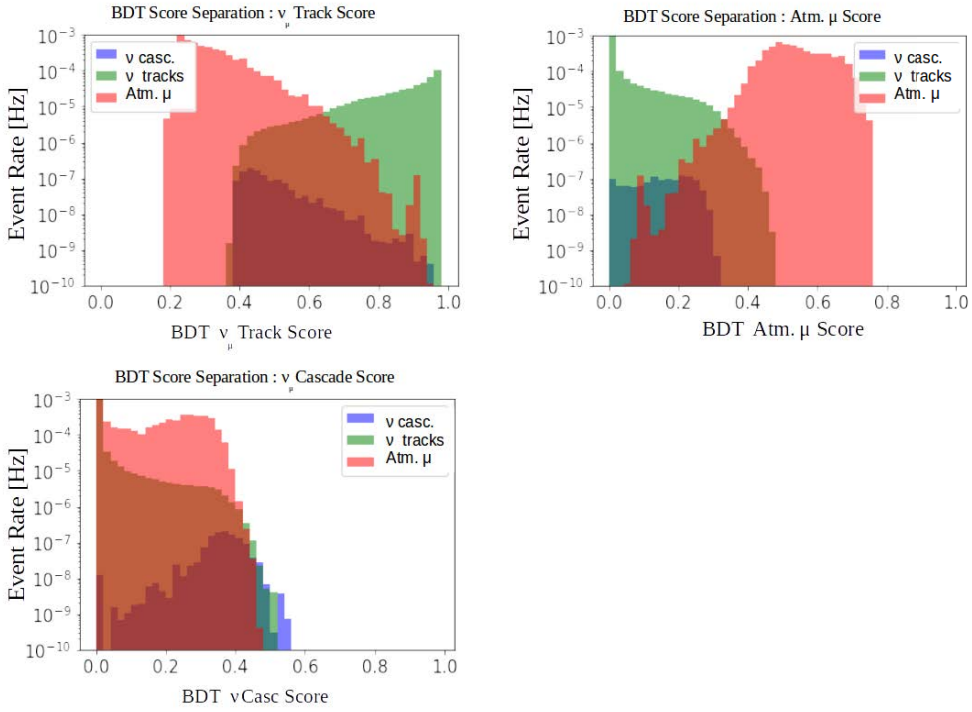


FIGURE 3.5: BDT category separation for the 3 different event Classes where only neutrino charged current tracks are considered signal.

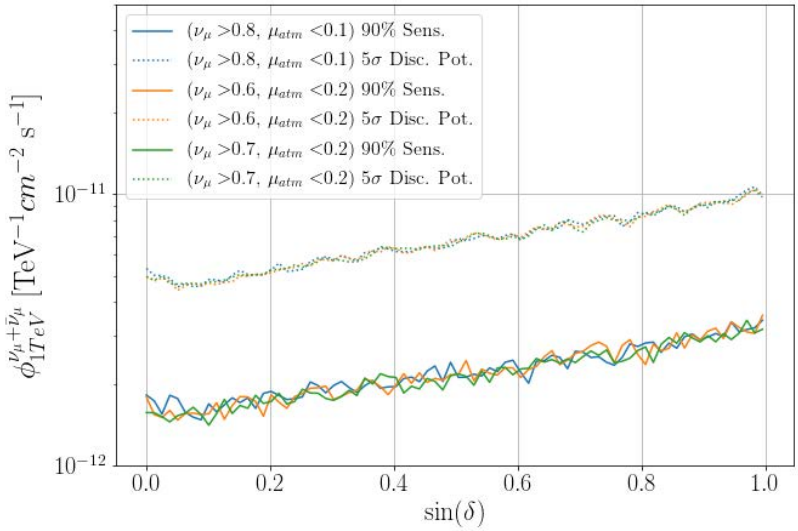
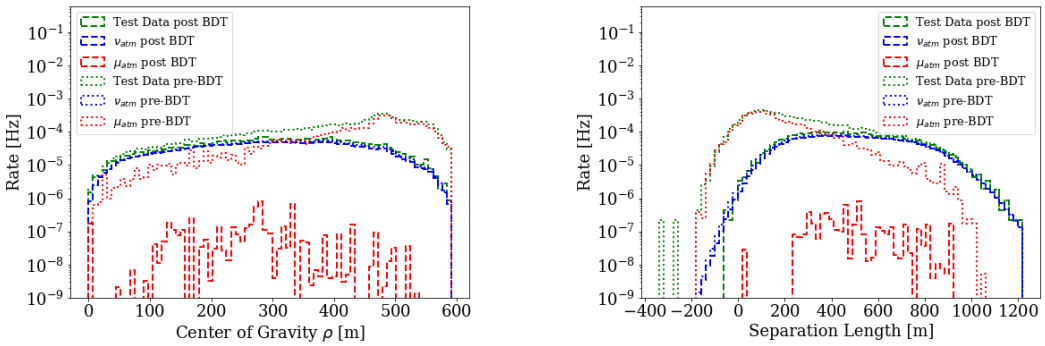


FIGURE 3.6: The 90% sensitivity and 5σ discovery potential flux using a single year of data of the updated selection with a variety of BDT cuts. The results can be seen to be within errors for the different scores. ν_μ >0.7 and μ_{atm} <0.2 were chosen as the final filters.



(A) COG_ρ variable (B) Separation Length

FIGURE 3.7: Distributions of two of the eleven BDT variables before and after applying the final BDT cut. These are shown for data, simulated atmospheric muons and neutrinos.

The BDT in the northern hemisphere is capable of high atmospheric muon rejection such that the contamination is less than 1% whilst keeping $\sim$ 80% of the atmospheric and astrophysical neutrino events as seen in Fig. 3.4. In this case the BDT efficiency was compared for different signal weighting schemes: i) for the traditional hard astrophysical flux of E^{-2}, ii) a softer flux of $\mathrm{E}^{-2.5}$ and finally iii) the atmospheric ν flux as a softest possible ν flux. The relative efficiency of each weighting depended on the BDT cut value. Although the softer flux weighting schemes demonstrate a higher ν_μ efficiency for the same levels of background contamination below a signal efficiency of 0.9, a large fraction of these ν_μ will have atmospheric and not astrophysical origins. The weighting scheme which maximizes the astrophysical signal efficiency is the E^{-2} power-law flux. Thus this spectrum was used to weight the astrophysical muon class during the training of the BDT.

Each individual event in the selection will have three separate scores associated with the probability of that event originating from the 3 different respective classes. These scores will then sum to 1, and so the cuts on these scores are dependent on each other. In order to evaluate where to apply a cut on up to three of the available BDT class scores, it is good to compare the distributions of the different true event categories as a function of the different BDT class scores which can bee seen in Fig. 3.5. The intention of the cut would be to create a selection where the sample becomes dominated by atmospheric and astrophysical neutrino events. This was done by requiring a ν-track score above 0.7 and an atmospheric μ score below 0.2. The final score combination was tested using the point source analysis sensitivity which will be described in more detail in section 4.2.1. Some slight variations on this score were tested for their impact on the final analysis sensitivity although the differences in sensitivity were all within the fluctuations as seen in Fig. 3.6.

TABLE 3.1: Details of data selection for the different years of data.

Data Samples					
Year	lifetime (Days)	Number of Events	Start Day	End Day	Ref.
IC40	376.4	36900	2008/04/06	2009/05/20	[85]
IC59	352.6	10711	2009/05/20	2010/05/31	[86]
IC79b	316.0	93133	2010/06/01	2011/05/13	[87]
IC86-I	332.9	136244	2011/05/13	2012/05/15	[88]
IC86-II-VII	2198.2	760923	2012/04/26	2018/07/10	this work

The resulting effect of these cuts can be seen in Fig. 3.7 which shows the distributions of the COG_ρ and separation length for the different simulated categories and data before and after the BDT cuts are applied. It is clear that the atmospheric muons make up a considerable part of the data before the cut and only around 0.1% of the simulated atmospheric muons remain after the cut. The data then corresponds closely with the distribution of simulated atmospheric neutrinos alone implying they dominate the final event selection. The figure also shows that almost 90% of these atmospheric neutrinos survive the BDT selection. It is expected for the atmospheric neutrinos to dominate the event selection and it is part of the likelihood analysis to differentiate these neutrinos from the astrophysical neutrinos.

3.5 Performance of the optimized selection

The event selection for each sample has changed over time with the updates to the configuration of the detector and new and improved data processing and reconstruction techniques. The work in the remainder of this thesis is conducted using the optimised event selection described above applied to six years of data from April 2012 until July 2018, in addition to the previous four years of data from 2008 as summarised in table 3.1. The optimised selection replaces that used in the previous all-sky analysis [6]. The performance for each of these events selections can be compared using three main quantities:

- Effective Area: the effective area (A_{eff}) plays a role in defining the expected number of neutrino events N_ν which would pass through an event selection as a function of

the neutrino declination δ and energy E_ν such that:

$$N_\nu = T * \int_{E_\nu} A_{eff}(E_\nu, \delta) \frac{d\Phi_\nu(E_\nu)}{dE_\nu} dE_\nu \tag{3.5}$$

Where T is the lifetime of the detector used for the selection, $\frac{d\Phi_\nu(E_\nu)}{dE_\nu}$ is the given differential neutrino flux for which the effective area is calculated.

- Angular Resolution, see section 3.1.
- Analysis Sensitivity: The analysis sensitivity can be highly dependant on the analysis implemented, more details for this calculation are included in section 4.2.1. The comparisons below compare the minimum astrophysical flux required for the point-source analysis to be sensitive to a source when the the true source position is being tested.

The effective area can be thought of as the equivalent surface area of a detector with 100% efficiency. It can be used to determine the rate of events for a certain flux (see eq. 3.5). It varies considerably for different declinations due the declination dependant cuts as well as the ratio of neutrinos that survive arriving to the detector depends on the fraction of the Earth they must pass. Fig. 3.8 compares the effective areas in different declination bands as a function of energy for the different event selections. The biggest difference in the effective areas can be seen between IC40 and the later years of data where the effective area in the southern hemisphere increases by almost a factor of 10. This is because when increasing the number of DOM strings from 40 to 86, the volume of the detector was more than doubled thus amplifying the rates of neutrinos which can be detected by a factor of 10 depending on the energy and declination.

The neutrinos must pass through different fractions of the Earth before arriving at the detector; this places a physical limit on the events arriving at the detector. This effect is observed as a diminishing effective area for events above 100 TeV with declination greater than 30° and so passing through most of the Earth, at energies where the ν interaction cross-section with matter is the highest. The effective area for down-going events arriving from the southern hemisphere is also significantly smaller than that of the up-going events below energies of a few hundred TeV. This difference is explained by the very strict cuts applied in the down-going event selection to remove atmospheric muon background. This effect no longer dominates at high energies since the muon flux has a softer spectral index and so should naturally diminish compared to an astrophysical neutrino flux, allowing looser cuts on high-energy events.

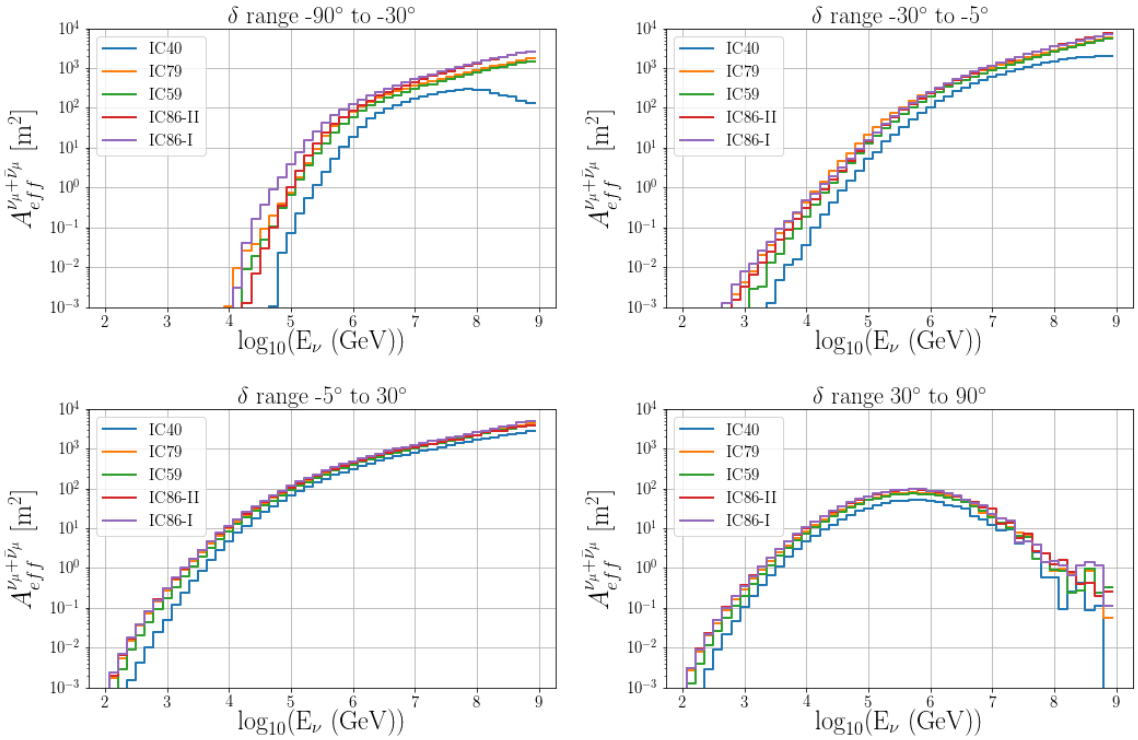


FIGURE 3.8: The average effective areas for $\nu_\mu + \bar{\nu}_\mu$ for the different years of data selection as a function of the neutrino energy for different declination regions in the sky

The improvements in the angular resolution compared the the previous event selection were discussed in section 3.1. Further comparisons with the older samples can be seen in Fig. 3.9. The updated six year selection demonstrates an up to 50% improvement in resolution compared to IC40, as the events in this configuration have the smallest lever arm to reconstruct the direction. The average angular resolution improvement of the new selection is $\sim 15\%$ compared to previous years although it is equivalent to IC79 depending on the event energy.

The selection effective area, angular resolution and the lifetime of the sample all play a key role in the final analysis sensitivity. To estimate the sensitivity of this sample an unbinned likelihood method is applied to search for an injected point source at the same location as the source injection. The minimum normalization of the flux that the analysis is sensitive to is compared in Fig.3.10. An average $\sim 35\%$ improvement with respect to the seven year all-sky analysis [6] comes from the 15% improvement in the angular resolution for the final years of the event selection along with the additional three years of data. The analysis used in Aartsen et al. [82] uses a different eight year selection with comparable angular resolution, however the analysis is optimised for a neutrino flux with an $E^{-2.19}$

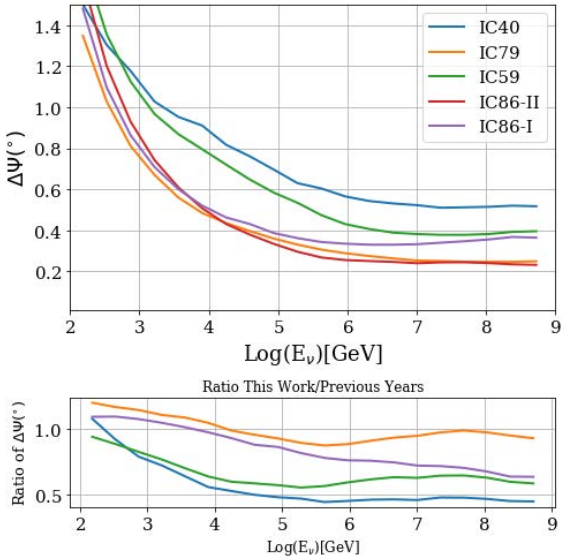


FIGURE 3.9: The median angular difference between the true neutrino and reconstructed muon track direction as a function of energy for the new data selection divided by that of previous data selections.

power-law spectrum with the use of a prior on the neutrino spectral index γ. This improves the sensitivity when injecting an E^{-2} spectrum so that, despite the shorter lifetime, the sensitivity is comparable. However the sensitivity of the updated selection in this thesis demonstrates an average $\sim 30\%$ improvement when testing an E^{-3} spectrum.

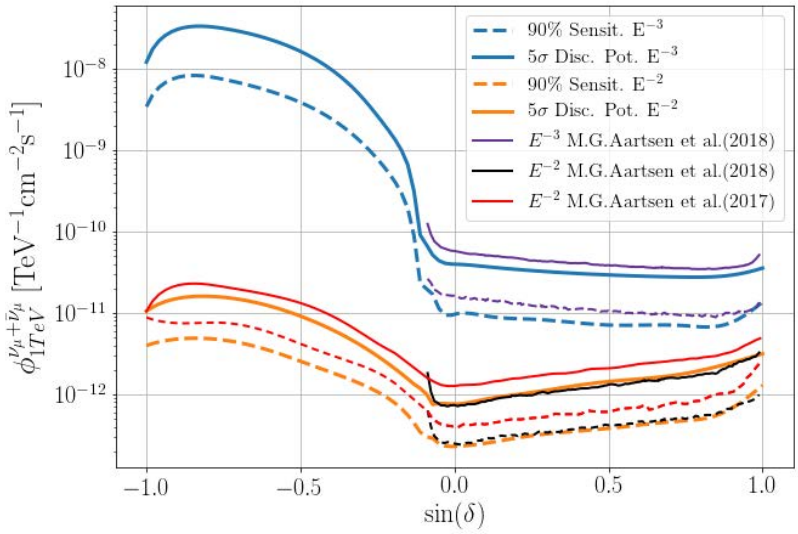


FIGURE 3.10: The minimum 90% sensitivity flux as a function of declination estimated for the likelihood analysis when injecting, and testing for a point source at the same location comparing the 10 year selection from this thesis to the 7 year analysis (red) and the 8 year northern sky analysis (black and purple) for an E^{-2} and an E^{-3} spectrum respectively.

Chapter 4

Analysis Method

A huge flux of particles generated from high energy CR showers in the atmosphere continuously penetrates the polar ice. It is a big challenge for the IceCube detector to distinguish rare astrophysical events among a large atmospheric background. To assess the spectrum of astrophysical neutrinos precisely, large statistics of events above TeV energies need to be collected. It is difficult to establish on an event by event basis the astrophysical or atmospheric nature. On the other hand, atmospheric neutrinos and muons are expected to be distributed uniformly across the sky, whereas astrophysical neutrinos should be observed coming from a point-like source or a small region due to the finite experimental angular resolution. In addition to spatial clustering the events originating from astrophysical sources are expected to obey a harder spectra than atmospheric events. An unbinned likelihood method consistent with that described by Braun et al. [89] is used to combine the spatial and energy event information and to build the probability distribution functions (PDFs) for the signal and background in order to estimate the likelihood that a signal of astrophysical origin emerges on top of the background.

The most reliable method for constructing background PDFs is to use the true event selection data. However, since astrophysical events cannot be determined on an event-by-event basis, it is not possible to construct an equivalent data selection of only signal events, and so simulated ν_μ events are used to construct the signal PDFs. Although it is assumed that any true astrophysical signal would make a component of the data, the background and signal PDFs should still be distinct. This is firstly possible because it is assumed that the event selection is dominated by background events uniform in right ascension (RA) due to the rotation of the Earth around its axis and integrating events over many years. Secondly, the spatial and energy PDFs are both constructed for bins of declination but not RA. This means, within a declination bin of the PDF, any point-like clustering should be smeared by the other events in the full right ascension range between 0 and 360°.

The disadvantage of this technique is that the analysis is limited by the amount of processed data and thus by the lifetime of the detection. At least for atmospheric neutrinos

we can use very high statistics of events through simulation while simulated atmospheric muons can be generated for more limited times. The very important advantage of using data to build up background PDFs is that the analysis is not sensitive to possible data-simulation disagreements.

4.1 Unbinned Likelihood

The IceCube point-source searches utilize an unbinned maximum likelihood method to determine a significant excess of events clustered around a point. This clustering is described by a signal PDF on top of a high rate of events modelled by a background PDF. An unbinned analysis considers the properties of each event separately which takes into account the individual event angular resolution. This can result in more precise discrimination between the signal and background spatial PDF. As mentioned, the energy distributions for astrophysical signal are expected to differ from that of atmospheric background, so that including the event energy information in the unbinned likelihood can further improve the analysis sensitivity. In order to include this information some assumptions must be made about the astrophysical spectra. Depending on the astrophysical source, there are multiple models for possible astrophysical neutrino energy spectra, often based around a power-law function. Although the true spectrum could be more complicated, for example with a break or a cut-off in the spectra, there are too few statistics in a point source analysis currently to be able to fit the spectra precisely. In this case the simplest most general case is implemented as a single unbroken power-law spectrum, $\frac{dN}{dE_\nu} \propto E^{-\gamma}$. However, since different sources are already observed in by gamma ray telescopes to have a range of spectral indices (γ) often as soft as 3.0 [90] or as hard as 1.75 [91], the spectral index is left as a free parameter in the analysis to maximise the sensitivity to a range of astrophysical source spectra.

The signalness of the collection of events around a specified direction can be evaluated using the likelihood defined as:

$$\mathcal{L}(n_s, \gamma_s, \overrightarrow{x_s}) = \prod_{i=1}^{N_\nu} \Big(\frac{n_s}{N_\nu} \mathcal{S}_i(\gamma_s, \overrightarrow{x_s}) + \Big(1 - \frac{n_s}{N_\nu}\Big) \mathcal{B}_i \Big) \tag{4.1}$$

Where $\mathcal{S}_i$ and $\mathcal{B}_i$ are the signal and background PDFs respectively, n_s is the number of signal events out of the total N_ν detected events and γ_s is the signal spectral index. The number of signal events n_s and the spectral index γ_s are free parameters in the likelihood function which are determined by maximizing the likelihood when using the entire data sample in the maximisation procedure. Although mathematically it is possible to fit an n_s

less than zero, physically one cannot have negative astrophysical neutrinos and so for these searches, where $n_s < 0$, the likelihood is assumed to maximally compatible with background and n_s is set to 0. γ_s is limited to between 1 and 4, because typical source accelerators are not observed with a harder spectrum hard than E^{-1}, and the analysis would still be sensitive to this flux by fitting the minimum spectral index, also any spectrum softer than E^{-4} would be difficult to differentiate from the background in terms of energy as this would become softer than the atmospheric flux spectrum of $E^{-3.7}$ which dominates the event selection.

The signal and background PDFs for the unbinned likelihood method depend on the tested source hypothesis. IceCube neutrino source analyses typically assume a diffuse, continuous background flux that is uniform in RA. A signal hypothesis can be applied for different neutrino source morphologies, energy spectra or time correlated emission by modifying the signal PDFs. For example, a time component in the PDF can be added for sources that are expected to only emit bursts or act as flaring objects. The analyses in this thesis are optimised for steady state neutrino emitters and thus do not have a time component so the events are integrated over the entire detector lifetime. The signal and background PDFs then consist of an energy ($\mathcal{E}$) and spatial ($\mathcal{S}$) component as described below:

$$\mathcal{S}(E_i, \overrightarrow{x_i}) = p(\vec{x_i}|\vec{x_s}) \times \mathcal{E}_{\mathcal{S}}(E_i, \sin\delta_i, \gamma) \tag{4.2}$$

$$\mathcal{B}(E_i, \overrightarrow{x_i}) = p(\vec{x_i}|bkg) \times \mathcal{E}_{\mathcal{B}}(E_i, \sin\delta_i) \tag{4.3}$$

In the following sections the spatial and energy terms for signal and background will be discussed in detail.

4.1.1 Spatial Likelihood Component

The spatial part of the signal PDF is based on the assumption that the angular uncertainty of an event (σ_i), which is calculated during the reconstruction, corresponds to a 1σ (68%) probability region for the event direction.

Hence, the spatial component of the likelihood assumes signal events will be clustered according to a 2D Gaussian around a given source position with σ_i angular uncertainty. The probability for an event originating from a source direction $\vec{x}_s$ can then be written as:

$$p(\vec{x_i}|\vec{x_s}) = \frac{1}{2\pi\sigma_i^2} e^{-\frac{|\vec{x}_s - \vec{x}_i|^2}{2\sigma_i^2}} \quad , \tag{4.4}$$

where the neutrino direction reconstruction and its associated error are $\vec{x_i}$ and σ_i respectively. This distribution can be assumed for small direction uncertainties below $\sim 5^\circ$ as is the case for the vast majority of events in the final selection. For events with a much larger angular error the 2D Gaussian assumption does not correspond well enough to the true event distribution to be applicable, and in the southern hemisphere these events are removed from the selection. This is why we cut them during selection. The spatial component for background however is dependent on the tested source declination. It has already been discussed that the arrival rate and selection of neutrinos and atmospheric muons is very dependent on how much of the Earth the event must propagate through. As a result, when filtering the data selection for the highest quality tracks, the final data rate varies with declination making it not entirely uniform in space. This is the reason why the expected rate of background events is calculated as a function of declination (δ_i of an event) and normalized to 1 by dividing by 2π to construct the background spatial component as written below:

$$p(\vec{x_i}|Background) = \frac{\mathcal{P}_B(\sin(\delta_i)}{2\pi} \tag{4.5}$$

The background spatial PDF only depends on the declination component, δ_i, of the event direction. For sufficiently long time scales, the spatial PDF for background is uniform as the Earth's rotation smears any RA dependency.

Pull Correction

It can been seen in the expression 4.4 that the signal spatial PDF is dependent on the estimated angular uncertainty σ of each event. The angular uncertainty associated to the track reconstruction ($\sigma_{paraboloid}$) is calculated based on the uncertainty in the muon track direction as described in section 2.7.3. This uncertainty does not take into account the angular difference between the incoming neutrino and the produced muon. In the weak interaction between the neutrino and the nuclei in the ice, the neutrino's momentum is transferred to the nuclei and the produced muon. The atomic nuclei is bonded in a lattice in the ice and therefore does not gain much velocity, however it is orders of magnitude more massive than the produced muon. For this reason the nuclei can take a non negligible part of the neutrino's momentum particularly for lower neutrino energies. This has the effect that the muon track direction is more deviated from the neutrino direction according to the following equation [92]:

$$\Delta\Psi = 0.7^\circ \times \left(\frac{E_\nu}{\text{TeV}}\right)^{-0.7} \tag{4.6}$$

This deviation is referred to as the kinematic angle between the muon and the neutrino. Given only information on the muon track, it represents a physical limit on the angular

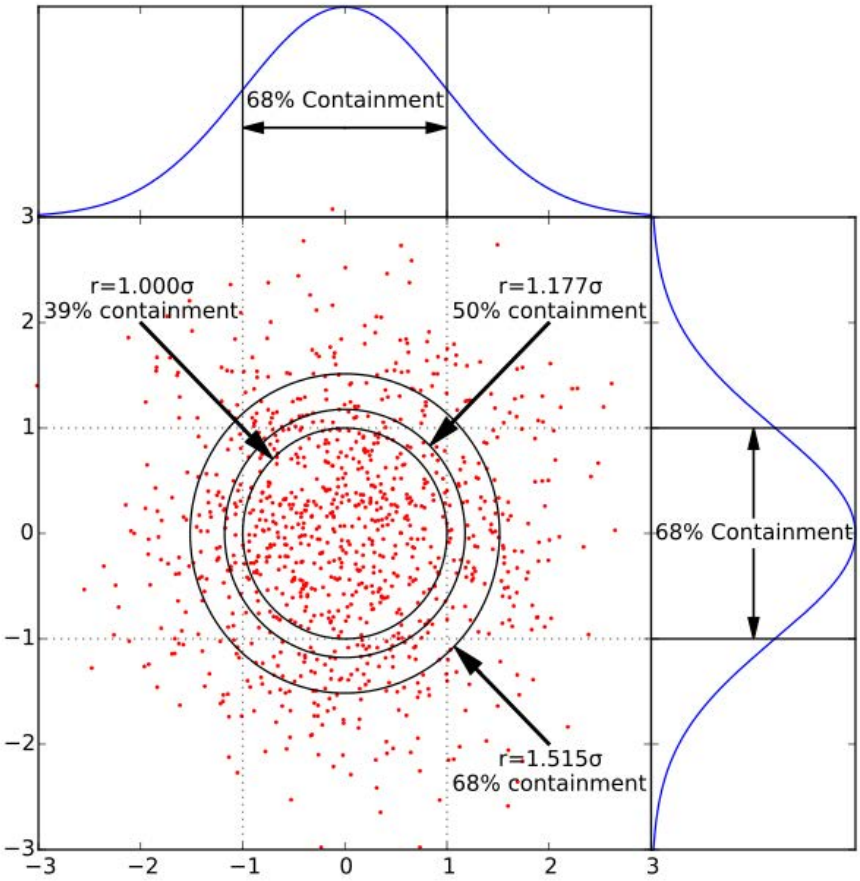
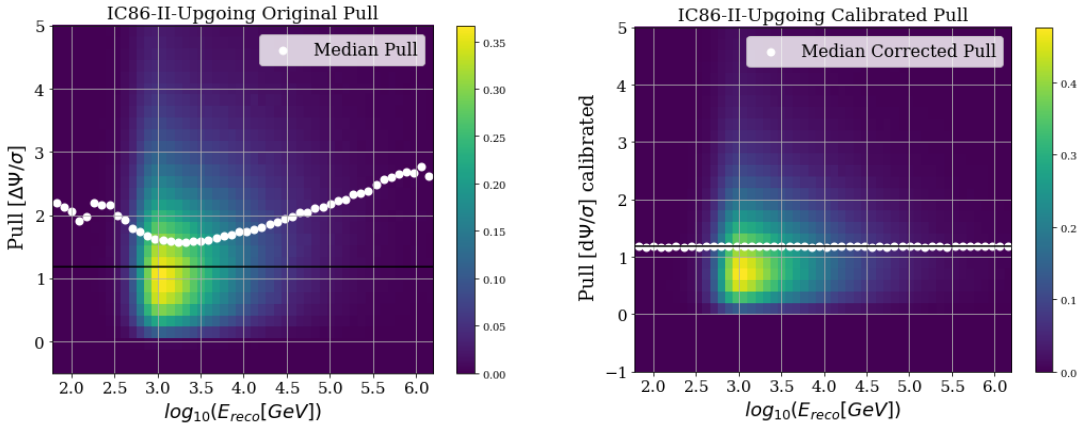


FIGURE 4.1: Events distributed according to a 2D Gaussian around a central source position for a certain angular error estimation σ taken from [93]

resolution as a function of neutrino energy. It should be reiterated however than the observed energy deposited in the detector from the muon does not directly correspond to the original neutrino energy. The energy losses from the muon are stochastic and the number of losses along the track through the detector are variable. The best indicator for neutrino energy from a charged current interaction is the length of the muon track, but for through-going tracks the length already traversed by the muon is unknown. In this scenario, the muon energy estimation cannot be used directly to determine the lower limit in angular resolution.

To account for the kinematic angle together with the uncertainty on the track reconstruction ($\sigma_{paraboloid}$), a method of calibrating the angular error estimation, σ, is applied using simulated events. This calibration is called the pull correction. The *pull* is defined as the angular difference between the true neutrino and the muon reconstruction ($\Delta\Psi$)



(A) Original Median Pull (B) Corrected Pull

FIGURE 4.2: Original pull as a function of the energy proxy (left) the corrected pull distribution as a function of energy proxy (right) for an E^2 spectrum.

over the event angular uncertainty estimation. Under the assumption that the distribution of *pull* within a fixed energy bin can be described by a 2D Gaussian distribution, 50% of events with an angular uncertainty estimation σ should fall within 1.17σ from the true source position as seen in Fig. 4.1. Although this assumption is not totally valid for the simulated neutrino events, it is the best approximation applicable so far. As indicated in Fig. 4.2a there is a deviation between the median of the pull distribution as a function of energy and the black 1.17σ line. The pull calibration involves calculating a scaling factor to shift the distribution of events such that the median of the pull is at 1.17. When the median pull is greater than this value it means that the angular difference from the reconstructed to the true direction is on average much larger than the estimated angular uncertainty, hence the uncertainty is underestimated. Since the median pull curve shows a dependency on the reconstructed event energy, the scaling factor is calculated in bins of energy proxy using simulated astrophysical neutrino events with an E^{-2} spectrum. A spline function is used to interpolate the scaling factors as a function of event energy to produce a continuous function. The reconstructed energy is used, despite being less accurate, because this calibration needs to be applied also to the data where the true neutrino energy is not available. The pull correction is calculated separately for each data sample and also each hemisphere since the event filtering is different in each hemisphere. The corrected event angular uncertainty (σ_{corr}) is then re-calculated according to:

$$\sigma_{corr} = \frac{\sigma_{paraboloid} \times pull_{50\%}(E_{reco})}{1.17} \tag{4.7}$$

Where $pull_{50\%}(E_{reco})$ is the median pull correction factor as a function of the reconstructed event energy.

4.1.2 Energy Likelihood Component

The energy component of the likelihood is constructed in a similar way. As discussed in section 3.5, the effective area of the detector varies with energy and declination of the events. The background energy spectra is known and expected to remain constant meaning that the background energy PDF for a given data sample is simply taken to be the probability density of the data as a function of $\sin(\delta_i)$ and $\log_{10}(E_{reco})$ for an event. The background energy PDF for the final data selection can be seen in Fig. 4.3 along with the curves showing the upper and lower limits containing 90% of events for a year of data and for simulated astrophysical signal for an E^{-2} and E^{-3} spectrum. The 90% energy range is a critical parameter that describes the relevant energy range probed by IceCube. Although IceCube is capable of detecting events outside that energy range, especially higher energy events, the 90% energy range describes the energy range where IceCube can provide reliable constraints or flux measurements.

Since the astrophysical spectral index is a free parameter in the analysis, the signal energy PDF ($\mathcal{E}_S$) is more complicated to construct. A given astrophysical signal spectrum, $\frac{dN}{dE} \propto E^{-\gamma}$, with a fixed γ is used to weight the astrophysical flux populating the 2D PDF as a function of declination and reconstructed energy, this will apply more or less weight to the high-energy end of the distribution. Once the 2D signal PDF for a given spectral index has been constructed, it is divided by the equivalent fixed background PDF to create a 2D map showing the relative signalness as a function of reconstructed energy and declination. These maps are created for a range of γ values between 1 and 4. The likelihood maximization procedure then interpolates between these maps when fitting the γ for a point-source. When evaluating the likelihood for a point-source, a single spectral index is used in the energy PDF for all samples used in the analysis. The γ which is fit by maximizing the likelihood will be value that maximises the energy and spatial likelihood for the sum all of the events around the source position.

4.1.3 Fit Spectra

When the likelihood is maximized, it is done at a fixed tested source position for all of the available data. The principle is to integrate all of the possible signal from that source location over the lifetime of the detector. However, the lifetime of the used sample is composed of different event selections (see table 3.1) each with different signal and background PDFs. For this reason the likelihood is calculated separately for each data

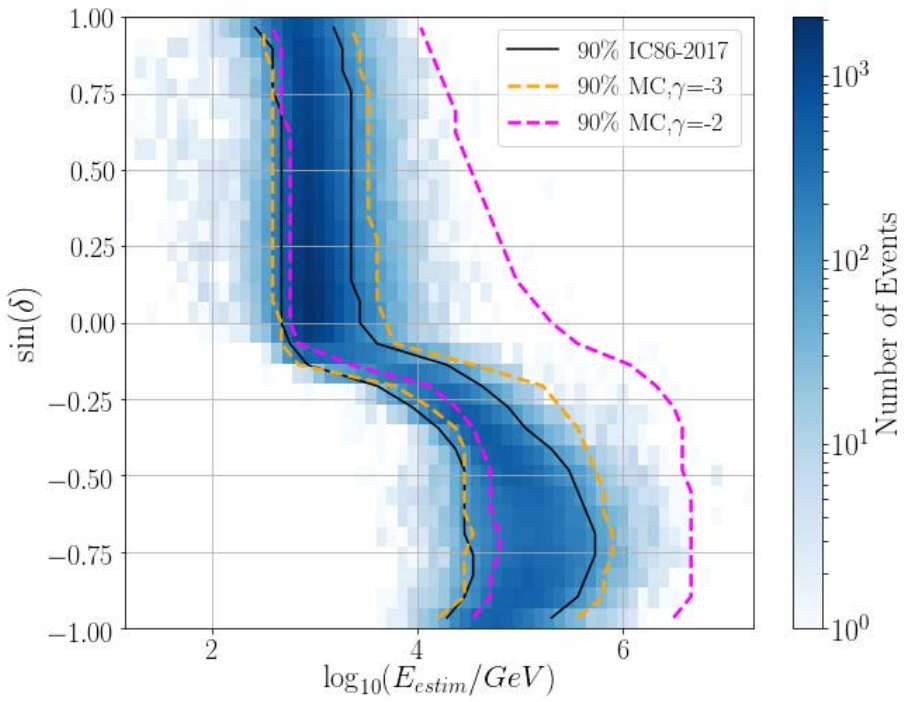


FIGURE 4.3: 2D distribution of events in IC86-2017 as an example of a background energy PDF. The 90% energy range for the data, as well as an astrophysical spectrum of E^{-2} and E^{-3} are shown as a guide of the relevant energy range of IceCube.

sub-sample and the final likelihood for a given signal hypothesis at a specific source location is the product of them all:

$$\mathcal{L}(n_s, \gamma) = \prod_j \mathcal{L}^j(n_s^j, \gamma) \ . \tag{4.8}$$

In this calculation the spectrum is assumed to be consistent across all the seasons and so the fit γ is forced to be the same for all the sub-samples. The number of events for each sample n_s^j is defined such that their sum is the best fit value $\hat{n}_s$. Assuming a steady flux for the lifetime of the detector, the contribution from each sample is taken to be as a function of the effective area and the lifetime of the j^{th} sample out of K samples according the formula below:

$$n_s^j(\delta_{src}, \hat{n}_s, \hat{\gamma}) = \hat{n}_s \times \frac{\int_0^{E_{max}} A_{eff}^j(\delta_{src}, E) E^{-\hat{\gamma}} dE}{\sum_{k=1}^{K} \int_0^{E_{max}} A_{eff}^k(\delta_{src}, E) E^{-\hat{\gamma}} dE} \tag{4.9}$$

The final likelihood is maximized keeping the γ the same in all seasons and with each season contributing n_s^j towards the final fit $\hat{n}_s$. This fit is then performed in the IceCube software *Skylab* by minimizing the negative TS using the L-BFGS algorithm implemented

in *Python* [94] .

4.2 Test Statistic

In Section 4.1 it was discussed how the likelihood for a best fit spectra can be found by maximizing a signal plus background likelihood function for the given sample. The final test-statistic for the point-source analyses is built to compare this likelihood to that of a background only hypothesis. The aim is to compare the *null hypothesis* ($\mathcal{H}_0$) with that of a signal hypothesis ($\mathcal{H}_1$) to determine if the signal hypothesis is significantly more likely than the null hypothesis, otherwise the null hypothesis is taken to be consistent with the data. This comparison is quantified as the Test Statistic by taking the ratio of the probabilities corresponding to each hypothesis as:

$$\mathrm{TS} = 2\log\left[\frac{p(Data|\mathcal{H}_1)}{p(Data|\mathcal{H}_0)}\right] = 2\log\left[\frac{\mathcal{L}(\hat{n_s},\hat{\gamma})}{\mathcal{L}(n_s=0)}\right], \tag{4.10}$$

Where $\mathcal{H}_1$ is the hypothesis that the event selection is dominated by background with a signal component with an unbroken power-law spectrum and $\mathcal{H}_0$ is the *null hypothesis* that the event selection is background only. Thus the ratio of the probabilities corresponding to $\mathcal{H}_1$ and $\mathcal{H}_0$ is equivalent to the ratio of the maximum likelihood functions with the best fit parameters n_s and γ over the likelihood function of background only where $n_s = 0$.

The null hypothesis is tested by evaluating the likelihood for a point-source using a sample where it is certain that $n_s = 0$, i.e a background only sample. Such a sample is achieved by using the data sample with the true event information except for the RA. Assuming that the event selection is background dominated and uniform in RA, a random RA value between 0° and 360° is assigned to each event. Any events clustered around point-like neutrino source become randomly scrambled and are assumed to be few enough in number that they will have negligible impact to the total number of events within a certain declination bin. Once all of the events have a randomly assigned RA value, the data is considered scrambled and *blind* to any astrophysical neutrino signal. One background trial consists of one copy of the data sample with scrambled RA values. To produce many background trials a new random RA value is assigned to each of the events, re-scrambling the events for every trial. In order to find an astrophysical signal in the sample the data is *unblinded* to astrophysical signal by using the true RA event information.

Before the data is unblinded, the analysis efficiency can be tested by injecting a known astrophysical signal corresponding to the signal hypothesis $\mathcal{H}_1$. This is done by generating a background trial and injecting simulated ν_μ events according to the spectrum of the

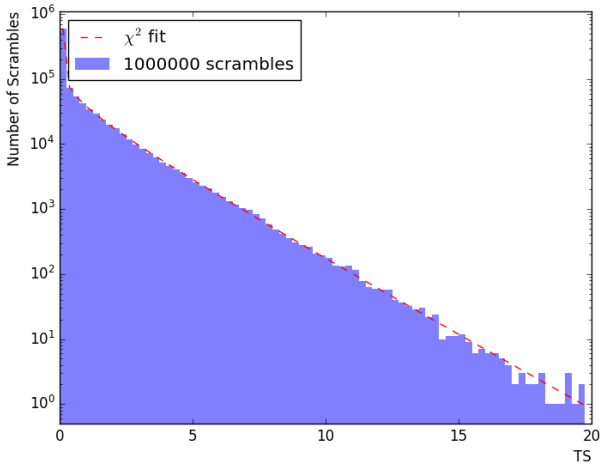


FIGURE 4.4: Example Background TS distribution using many trials of scrambled data testing the source location at a declination of $-0.52°$.

signal hypothesis. The likelihood method remains the same for blind background trials, background trials with injected signal, or the unblinded data.

Once the likelihood is maximized to calculate a TS, larger TS values should correspond to cases where the signal hypothesis is much more likely than the background only hypothesis. The case where TS= 0 from the best fit can be due to $n_s = 0$. Statistically one expects both over-fluctuations around and under-fluctuations around TS=0 roughly 50% of the time each way. The former can be due to random spatial clustering or high-energy events in the tested source direction and, according to Wilk's theorem [95], should follow a χ^2 distribution with n degrees of freedom. Since this analysis includes n_s and γ as free parameters, although they are partially correlated in that a higher n_s usually requires a softer spectra to fit the same data, this results in a number of degree of freedom between 1 and 2 degrees of freedom. In this analysis the TS under-fluctuations are set to 0 because the likelihood does not allow fits of $n_s < 0$. This is the reason why there is a large peak in the distribution where TS= 0.

The resulting TS is used to determine a *p-value*. The p-value estimates the probability of finding an equal or larger TS value than that of the observed data, given background only. The analyses in this thesis involve calculating TS when testing a point source hypothesis at a specific location. The final obtained result is then compared with a TS distribution generated entirely by background trials at the same declination. An example is shown in Fig. 4.4 for a point-source test at a declination $\delta = -0.52°$ using 1000,000 background trials. The probability that the tested result could be due to background alone is equivalent

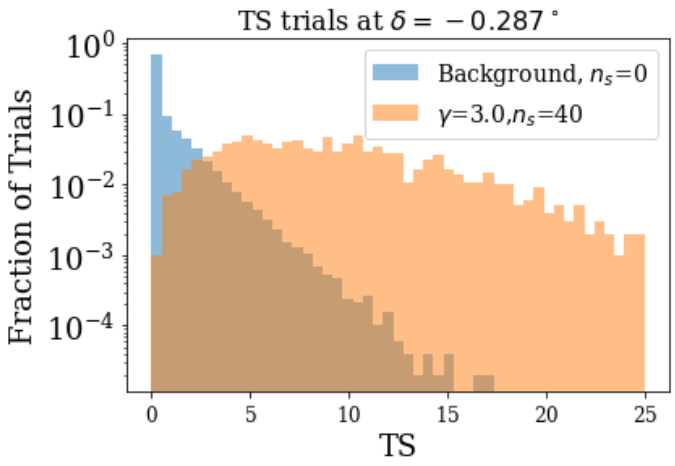


FIGURE 4.5: Example Background TS distribution using many trials of scrambled data testing the source location at a declination of $-0.28°$ compared to TS distribution of background trials with injected signal of 40 events from a spectrum of $E^{-3.0}$ at the source location. Distributions such as these are used to establish the sensitivity of the analysis.

to the fraction of the distribution above the observed TS value. This can be done by either taking directly the fraction of background trials or by integrating a fit function of the distribution. Although there can be some errors in the fit estimation, at high values of TS, it is in general more accurate to integrate a χ^2 function which is fit from the background distribution, as the true distribution will usually suffer from statistical fluctuations due to a limited number of background trials.

4.2.1 Analysis Sensitivity

Before evaluating the p-value for any source hypothesis, it is important to estimate how sensitive is the analysis to a signal from such a source. This involves estimating what spectra the analysis should be sensitive to and, in the case of a null result, which spectra can be ruled out and constrained by the given result. Just as the p-value is established using the distribution of background trials, the sensitivity is established also using trials with injected signal events.

Right ascension and time of the events are scrambled to produce background trials, then some signal is injected by sampling from simulated astrophysical neutrino events according to the tested source spectrum. Simulated events are chosen which fall within $\sim \sin(\Delta\delta) = 0.1$ of the tested source position, then the entire event is rotated until the true neutrino direction matches that of the tested source position. The energy and number of the sampled events depends on the tested spectrum. The likelihood is hence evaluated for the signal and background events together, as is done when using data alone. This

process is repeated injecting the same flux many times to create a signal TS distribution corresponding to a given flux. This signal TS distribution is then compared with that of background only to establish the sensitivity to signal. The important parameters are used to quantify this comparison are:

- α: describes the probability of finding a *false positive*, in the case that a background only trial is able to reproduce a certain result.
- β: describes the power of the search and is the probability, for a certain flux, that the result will be able to be reproduced.

The *Sensitivity* of an analysis is then defined as the flux required to have $\alpha = 0.5$ and $\beta = 0.9$ so that 90% of signal trials injecting that flux have a higher TS than 50% of the background-only trials. The flux is determined iteratively by injecting $n_s = 1$ in many trials and calculating β for the TS value where $\alpha = 0.5$. The injected n_s is then increased step by step until $\beta = 0.9$, and the respective flux is calculated from the n_s sampled from a fixed spectrum. An example of a single step can be seen in Fig. 4.5.

The *Discovery potential* of an analysis repeats the exact same procedure as the sensitivity but requires α to be 5σ and the flux is increased until $\beta = 0.5$. The discovery potential estimates the flux required so that 50% of the time this flux will produce a result inconsistent with the background hypothesis at the discovery level of 5σ. The *Sensitivity* flux at 90% c.l. can then be interpreted as the flux at which one expects the analysis to begin to be sensitive to the source by observing a higher signal likelihood than the background likelihood for 90% of the trials. Although at such a low α a source with this flux cannot be confirmed as inconsistent with background, it will begin to be of interest. The *Discovery* flux describes the flux for which, 50% of the time, can be declared inconsistent with a background hypothesis and hence a neutrino source can be discovered. This flux is usually much higher than the sensitivity flux but this difference depends on the background TS distribution.

For the 10 year analysis the discovery potential and sensitivity to a point source is shown as a function of declination given different injected neutrino spectra in Fig. 4.6.

Once the analysis sensitivity is considered satisfying the data can be *unblinded*. This means, instead of using the *blind* data scrambled in RA, thus removing any spatial and energy clustering, the exact same procedure is implemented using the true direction and energy information for all of the events.

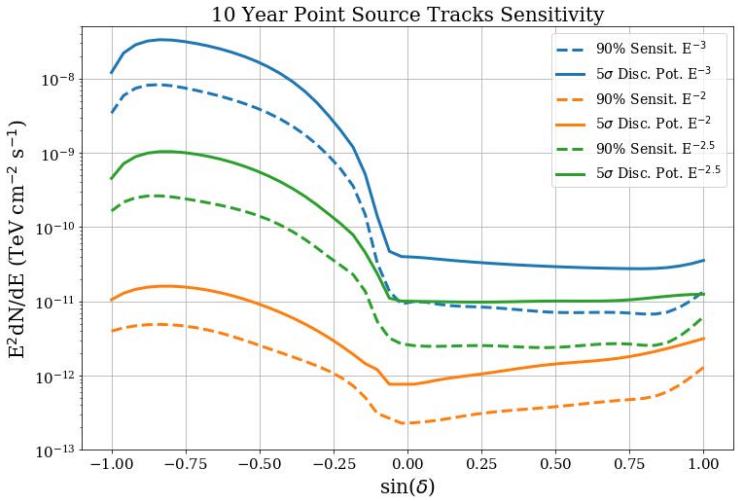


FIGURE 4.6: 90% Sensitivity and 5σ Discovery Flux to a point-like neutrino source as a function of declination shown for different γ using the final 10 year Point source Tracks selection.

Chapter 5

Point Source Searches

All the analyses in this thesis are designed to search for point-like steady state astrophysical sources of neutrinos. In this case we assume a background of atmospheric muons and neutrinos which are uniform in right ascension (RA). These background events are known to depend on energy and declination (as described in chapter 3). Background trials are generated to calculate the sensitivity and discovery potential of the analyses, as well as determining the post-trial p-value. Data themselves are used scrambling the right ascension which is equivalent to the time of the events. By using data and not simulation we ensure that there is no systemic error from differences in data and simulation. Given that IceCube measured a diffuse astrophysical signal, these events are present in the sample but are subdominant with respect to the atmospheric background. They are assigned different times (and RA) and kept in the sample at their energy and declination; this constitutes a diffuse neutrino flux. However, in a point-source search, the significance of the fit is determined by a spatial clustering of high energy events and it would not be sensitive to a single high energy event. It is also possible to observe a signal-like result where the weighting is dominated only by the spatial clustering of low energy events. This requires a higher flux of events, with respect to clustering of high-energy events, in order to observe a signal that significantly surpasses the background.

All of the searches conducted in this analysis test a source hypothesis expecting observed signal events to appear clustered around a point with an extension comparable to the IceCube angular resolution. If the source candidates themselves have a spatial extension σ_{src} (in general $\sigma_{src} < 1^\circ$), this is modelled as a two dimensional Gaussian term in the spatial PDF that becomes:

$$p(\vec{x_i}|\vec{x_{src}}, \sigma_{src}) = \frac{1}{2\pi(\sigma_i^2 + \sigma_{src}^2)} e^{-\frac{|x_{src}-x_i|^2}{2(\sigma_i^2+\sigma_{src}^2)}} , \tag{5.1}$$

where σ_i is the angular uncertainty of an observed event in the direction $\vec{x}_i$ around a tested source position $\vec{x}_s$. Therefore, the searches in this thesis for point-like time-integrated

neutrino sources over ten years of data taking have a common theme: they are all optimised to be most sensitive to objects emitting astrophysical neutrinos with a harder spectrum than the atmospheric background originating from a region less than 1° around the object position in the sky. Since we fit for a single spectrum, this analysis is most suitable for not variable sources with a fixed energy spectral index. However the time-integrated analysis does not exclude sensitivity to variable neutrino sources, since it can still detect variable neutrino sources with a bright flare emission or a long flare duration as an excess of events. Time-dependent analyses typically overtake the sensitivity of a time-integrated analysis for flare duration below $\sim$ 100 days. Within this common steady-state point-source theme, several different neutrino source searches have been conducted testing different signal hypotheses that will be described in the following sections.

5.1 All-sky Scan

The motivation for an all-sky scan is to provide a map of the sky in terms of the probability of an astrophysical neutrino source at every location. This can be a key tool used to find a source of astrophysical neutrinos that has not been observed or strongly motivated by other astrophysical messengers. It is possible that an object could be a strong source of high energy neutrinos but matter around the source would obscure light from arriving to Earth. For example the first gravitational wave observation [3] came from two black holes merging which was indeed never observed by electromagnetic observations. By a similar argument, since neutrinos have an extremely small interaction cross section compared to photons, an astrophysical object surrounded by a dense gas of matter could be blocked from view in high energy photons but act as a bright neutrino source.

To locate such objects, the pre-trial probability of the observed emission being totally compatible with background is evaluated at a grid of directions across the entire sky, excluding the regions close to the poles. The polar regions are not examined because it is important to compare the final result with scrambles of the events in background trials. The regions around the poles are so small that the same few events would be scrambled and would not create a reasonable background distribution for comparison with the results. Between $\pm 82°$ in declination, a local p-value is determined for each tested location in the sky to create an entire sky map such as that of Fig. 5.1, as the probability the observed signal around the tested location is consistent with background only at that same location.

The all-sky scan is conducted iteratively in order to improve efficiency and to avoid testing over a million different grid positions at once.

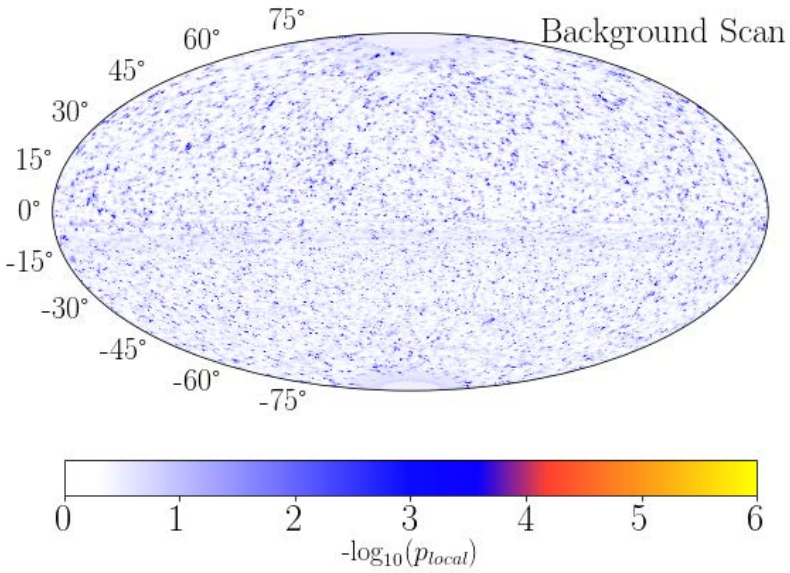


FIGURE 5.1: Pre-trial p-value map of the sky from a background trial.

- At first the sky is scanned with a course binning of 0.9° in order to save processing time and establish which regions of the sky are the so called "warmer" regions with a more signal-like result. This way time is not wasted evaluating points in regions where there are few interesting events.
- A scaling factor is applied effectively increasing the resolution by increasing the number of grid points across the sky.
- the course grid map is interpolated to estimate TS values across the new finer grid of points.
- A test statistic threshold (TS_{th}) is calculated such that the only grid points with an interpolated TS$> TS_{th}$ are evaluated directly by maximising the likelihood function, such that the number of new grid points evaluated directly is roughly equal to the number of points in the course grid.
- This process is repeated until the final resolution of the grid map is $\sim 0.1°$. Assuming the uncertainty of the events is limited to 0.2°, this should find even the most compact hotspots in the sky.

Once a final grid of test statistic values has been constructed to a resolution of $\sim 0.1°$ for the entire sky, the local pre-trial p-value from the TS value is estimated for every grid point in the sky. The *hotspots* from this map are defined as the locations of the most significant p-values in the northern and southern hemisphere respectively.

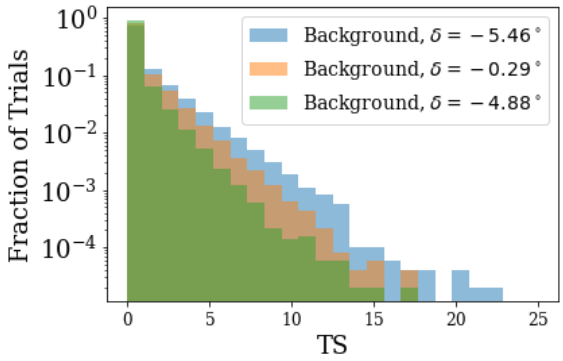


FIGURE 5.2: Background TS distributions for a point source at three different declinations. The region close to the horizon, with a declination close to $-5°$ shows the largest differences in the TS distributions

5.1.1 P-value Calculation

In order to calculate a local pre-trial p-value, the resulting TS value at a certain location needs to be compared with a background distribution generated around that same declination, as described in 4.2. When testing a single specific source hypothesis the background distribution can be constructed by fitting the many scrambled data samples to be compared to a injected flux or an unblinded value. In the all-sky scan hundreds of thousands of different directions are each tested for a local p-value at a fine grid of points. It is not feasible to run many background trials for every single grid point in the sky to calculate this p-value and so instead a function is constructed which can estimate a local pre-trial p-value as a function of TS and declination. It has previously been mentioned how a TS distribution constructed from many trials can be fit in order to avoid p-value fluctuations due to a lack of statistics at the tails of the distribution. A similar procedure is implemented to tackle the p-value estimation when an exact background TS distribution is unavailable.

The importance for the declination dependency in the p-value function is because the event selection, and the effective area are both declination dependent. This declination dependence has the effect that the TS distributions are not exactly the same at every declination in the sky. This effect is particularly strong at the boundary between the northern and southern hemisphere. The reason for this is that the background cannot be considered uniform in the region a few degrees across the hemisphere boundary. At the horizon the data selection rapidly changes and stricter criteria result in a much lower data rate in the southern hemisphere. An example of different TS distributions in different declination bands can be seen in Fig. 5.2.

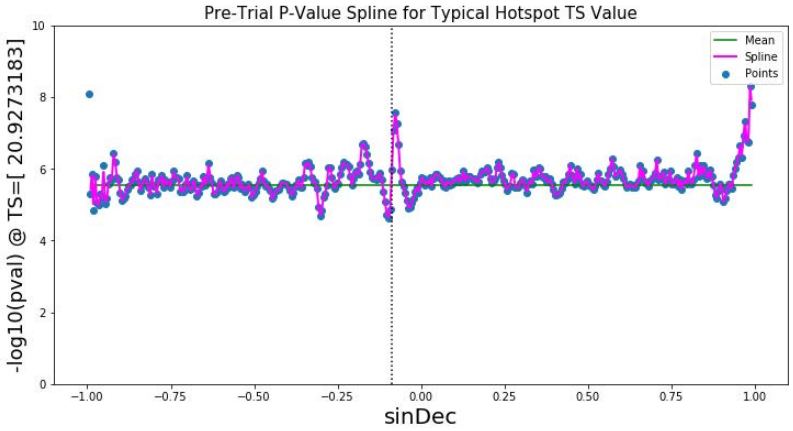


FIGURE 5.3: Pre-trial local p-values as a function of declination for a fixed TS value, roughly that of a typical hotspot. This is shown at points calculated from the best fit from background trials at that direction as compared with a spline of these points and the mean fit result.

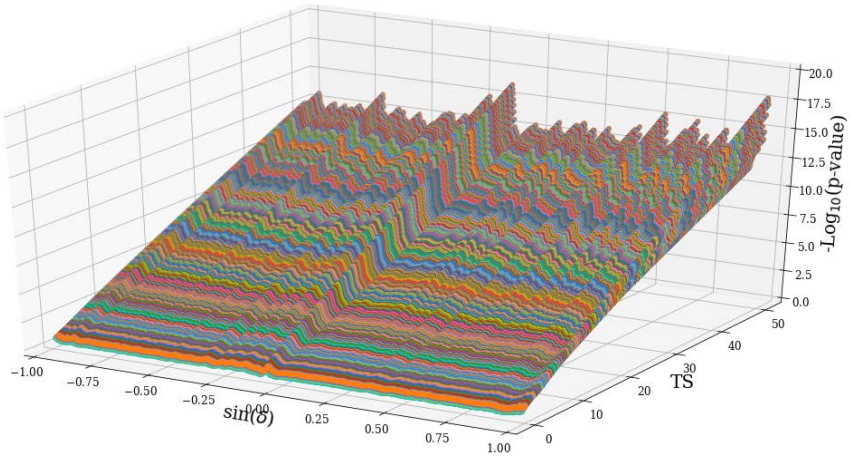


FIGURE 5.4: Pre-trial local p-values as a function of declination and TS value. This map is created using splines of the best fit background TS distributions in 400 different declination bins.

Instead of using a single TS fit for the entire sky, the TS fit is made as a function of declination. This process involves:

- constructing a background TS distribution from 1 million scrambled trials each at a fine grid of declinations in $sin(\delta)$ such that the region around the horizon has the finest grid distribution. In practice a fit for the background distribution was made for 400 different declinations in the sky so that the difference in $sin(\delta)$ between each declination bin is 0.0025.
- A χ^2 fit was made for the background TS at each declination bin. Either fitting the whole function at once or using only the part of the distribution where TS$>$ 5.
- A Kolmogorov-Smirnov (KS) test [96] was conducted to select which of the two fits corresponded more closely to the true distribution and this fit was selected.
- The final list of fits for each declination was used to estimate a p-value for a range of TS values in steps of 0.125 between 0 and 50.
- At each TS value a spline function is fit to estimate the p-value as a function of declination and then saved. An example of one of these splines for the TS value of a typical hotspot if shown in Fig. 5.3.

The result of these steps is the two dimensional function seen in Fig. 5.4, that is able to estimate the pre-trial p-value as a function of declination and the observed TS value at the tested source location. Once TS values are calculated for the final grid of points in the sky, this function is applied to estimate their corresponding p-values. The hottest spots in the northern and southern hemisphere are selected by the most significant p-values and not TS values to avoid a bias to picking hotspots in declinations bins where the background TS distribution is shifted to higher values.

5.1.2 Post-trial p-value

A scan is first conducted on many background trials where the right ascension information is scrambled. In this way the level of background fluctuations from random event clustering can be observed and compared with final hotspot information. The parameters of the hotspots in each hemisphere are stored for each background trial. These trials are used to create expected distributions of the typical parameters of the background hotspots. This can be used to check that these hotspots are distributed roughly uniformly in space. In addition the distribution of background hotspot p-values permits a post-trial p-value estimate for the probability that the hottest spot in the unblinded data is due to a typical background fluctuation. The all-sky scan requires the signal to surpass the highest

background fluctuations from many trials in order to seem significant. This means that, although this analysis is sensitive to a source from any direction, it requires a strong flux to be seen as significant.

5.2 Source List

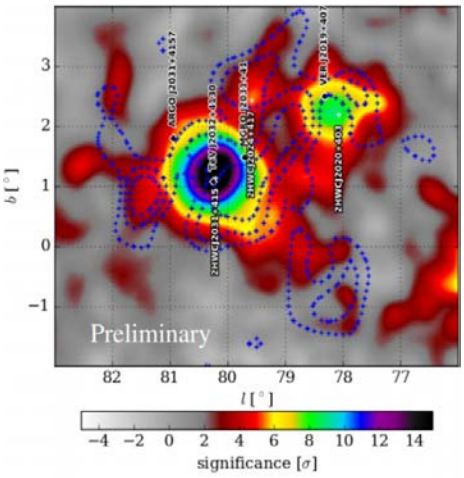
As mentioned above, the neutrino signal must be very strong in order to significantly surpass the highest fluctuations. By incorporating more source information into the signal hypothesis a different search can be conducted, in principle more sensitive to a lower neutrino flux provided that the source information is accurate. A possible way is to evaluate the TS only at specific source locations corresponding to different multi-messenger evidence. This approach could lead to the discovery of multi-messenger astrophysical sources. Objects which have been observed in high energy gamma rays have already demonstrated an ability to accelerate particles to high energies. The question about which processes are taking place remains open; if there are hadronic interactions at the accelerator sites which would result in a high energy neutrino flux.

5.2.1 Individual Source Catalog

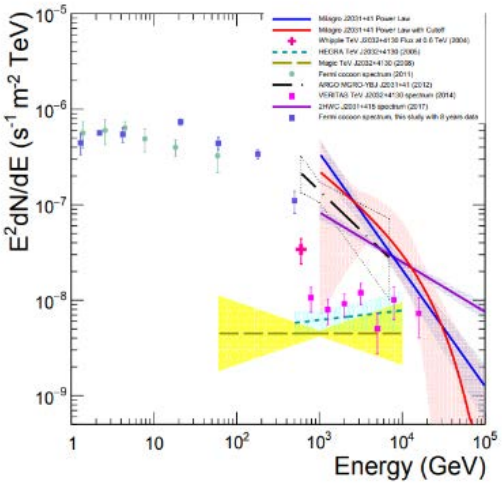
A list of 110 objects most likely to be steady state neutrino emitters was composed based on gamma-ray observations. This number of sources was chosen because it is consistent with the order of magnitude of sources monitored by IceCube in the past [6] and given a 5σ pre-trial p-value for a certain source the post-trial p-value is expected to remain above 4σ after penalizing for the trial factor due to the size of the catalog.

The composition of these 110 sources was broken down into 2 sections, galactic and extra-galactic objects. This was motivated by the fact that galactic sources are much closer to the Earth and therefore the gamma-rays from these objects are less attenuated on their journey from source to detector. High energy photons from far distances otherwise encounter en event horizon due to interactions with the Cosmic Microwave Background (CMB). This means there is more precise information available concerning galactic objects and the expected possible associated neutrino flux.

For the galactic sources, the gamma source properties were assembled from the most recent observations at TeV energies found in the online catalog TeVCat [98]. Known associated sources were considered as a single source. The most luminous observation was used for example in the Cygnus region with lots of closely located sources. The objects HAWC J2031+415 and TeV J2032+4130 were selected with the locations and spectra shown in Fig. 5.5a. To select the galactic objects of highest interest to this analysis a



(A) Map of Cygnus Region in Gamma-rays as seen by HAWC



(B) Spectra for Sources in Cygnus Region

FIGURE 5.5: Summary of information on the Cygnus region taken from Hona et al. [97] and used to estimate a flux for HWC J2031+415.

power-law flux of the form below was considered :

$$\frac{dN_{\nu_\mu+\bar{\nu}_\mu}}{dE_\nu} = \phi_0\Big(\frac{E_\nu}{E_0}\Big)^{-\gamma} e^{-E_\nu/E_{cut-off}}\ TeV^{-1}cm^{-2}s^{-1}\ , \tag{5.2}$$

where E_ν represents the neutrino energy, ϕ_0 is the flux normalization at the reference energy E_0 and $E_{cut-off}$ represents the cut-off energy for a break in the spectrum. The exponential cut-off factor is ignored if no cut-off is yet measured in gamma-rays. The measured normalizations and spectral index for gamma at 10 TeV, as well as any observed cut-off, were used to predict a model spectrum for each source as shown in table 5.1. This spectrum was injected directly into the point source analysis to determine the normalization of the spectra that would allow 90% of the signal trials to have a significance greater than 50% of the background trials. Assuming up to 50% possible improvement of the analysis sensitivity over the next 10 years, the Galactic sources with an estimated emitted flux greater than 50% of the required sensitivity flux were selected. This process determined 12 galactic objects described in Table 5.1 to be of interest to this analysis over the next few years.

In addition to the 12 galactic sources, other 98 extra galactic objects are chosen. Similarly to the galactic source selection, extra-galactic sources were chosen which were of most interest to this analysis. However, due to the more limited knowledge about the source flux at high energies, it is not possible to directly estimate the possible neutrino spectrum.

TABLE 5.1: Spectra used for the 12 Galactic Neutrino Source Candidates

Galactic Sources							
Name	$\alpha(^\circ)$	$\delta(^\circ)$	E_0[TeV]	Φ_0	γ	$E_{cut-off}$	Ref.
Crab-nebula	83.63	22.01	1.0	3.76×10^{-11}	2.39	14.3	[99]
HESS J1841-055	280.23	-5.55	1.0	1.28×10^{-11}	2.41	-	[100]
HESS J1843-033	280.75	-3.30	1.0	1.28×10^{-11}	2.41	-	[101]
MGRO J1908+06	287.17	6.18	1.0	4.23×10^{-12}	2.20	-	[100]
HESS J1837-069	279.43	-6.93	1.0	5.00×10^{-12}	2.27	-	[100]
HESS J1857+026	284.30	2.67	1.0	6.10×10^{-12}	2.39	-	[100]
PSR B0656+14	104.95	14.24	7.0	$2.30\text{e}\times10^{-14}$	2.0	-	[102]
MGRO J2019+37	304.01	37.20	1.0	1.35×10^{-12}	2.0	-	[91]
HESS J1849-000	282.26	-0.02	2.74	$7.66\text{e}\times10^{-14}$	1.97	-	[101]
HESS J1852-000	283.0	0.0	2.06	2.96×10^{-13}	2.17	-	[101]
Gamma Cygni	305.56	40.26	1.0	6×10^{-12}	2.37	-	[103]
2HWC J2031+415	307.93	41.51	7.0	6.15×10^{-14}	2.52	-	[102]

Hence, the sensitivity for each source is estimated as the flux integral above 1 TeV as measured by Fermi LAT [104] weighted by the point source sensitivity for a spectrum of E^{-2} at the source declination. In this way the brightest objects at declinations of optimal sensitivity are prioritized. Next the extra-galactic objects are sorted into different categories of BL Lac objects, Flat Spectrum Radio Quasar (FSRQ), Active Galactic Nuclei (AGN), Unidentified Blazars (UNIDB) and Star Burst Galaxies (SBG). The top sources from each category were selected to give a total of 98 sources. This was used instead of a straight cut in order to ensure that sources were selected from each category even if they had a lower flux on average. The result is the top 5% of BL Lac, FSRQ, whilst including blazars of unidentified type (AGN and UNID objects) that pass the average lower energy threshold set by the 5% limit which was found to be $3.6 \times 10^{-9}\text{cm}^{-2}\text{s}^{-1}$. An exception was made for the SBG because there were only 8 objects and they are well motivated to be possible sources of hadronic emission [105]. In this case all 8 galaxies were included. For this catalog there were 97 sources in the northern hemisphere and 13 sources in the southern hemisphere. This large difference comes from weighting the source flux by the IceCube sensitivity which is considerably more powerful in the northern hemisphere for this analysis.

The aim of this analysis is to evaluate the probability to observe an astrophysical neutrino flux at each individual source location. The number of estimated astrophysical neutrino events and the estimated spectral index for an unbroken spectrum is fit separately at each location to maximize the signal-over-background likelihood function. Different criteria are applied to select track-like events from the Northern and Southern hemisphere (with a boundary between them at declination $\delta = -5^\circ$), because the background differs in these two regions. As a result the properties of the event selection and the resulting analysis sensitivity differs considerably between the two hemisphere described further in chapter 3. As such, the sources in the northern and southern hemisphere are treated separately and the total source list is divided into a northern-sky catalog containing 97 objects (87 extragalactic and 10 Galactic) and a southern-sky catalog containing 13 sources (11 extragalactic and 2 Galactic). The large north-south difference is due to the difference in the sensitivity of IceCube in the northern and southern hemispheres; recall that the sensitivity was used to construct the catalog. The post-trial p-value for each catalog is calculated from the fraction of background trials in which the pre-trial p-value of the most significant source is smaller than the pre-trial p-value found in data.

5.2.2 Source List Population Search

Taking into account the fact that multiple sources were examined, the post-trial result for each catalog in the individual source-list search is the probability that the most significant source from the catalog is due to background alone. Similarly to the all-sky scan analysis, the individual source list search requires a single source to show a fairly substantial significance to surpass possible background fluctuations. Alternatively, weaker sources could be discovered by searching for a population of sources within the catalog as opposed to a single bright source. Although, given a significant result from the population search, the specific sources responsible for an observed signal are not identified.

A population of sources of unknown size and origins is determined by finding an excess in the rate of pre-trial p-values below a certain threshold for a given catalog. The sources in each catalog are re-ordered from most to least significant and the threshold is iteratively set as the pre-trial p-value (p_k) of the k^{th} source. This threshold is set as the background probability, p_k, of the k^{th} source in the source list. The number of sources which pass this threshold are counted. Assuming a binomial background distribution for sources which pass the threshold, the final probability, p_{bkg}, for having k or more sources with a pre-trial p-value smaller than p_k for background is:

$$p_{bkg} = \sum_{i=k}^{N} P_{binom}(i|p_k, N) = \sum_{i=k}^{N} \binom{N}{i} p_k^i (1-p_k)^{N-i} \tag{5.3}$$

where N is the total number of objects, i.e, 97 for the Northern catalog and 13 for the southern catalog. The final result of this search is the most significant p_{bkg} from all the different tested population sizes between 1 and N. The associated post-trial p-value from this search takes into account the fact that multiple population sizes were tested.

5.2.3 Stacking Searches

Given the very large contamination of background events in the point-source event selection, a source must emit a very high flux to surpass the atmospheric background events in the same region, or have a very hard spectrum such that signal events can be differentiated from the background due their to higher energies. Ideally both cases would be true for the best chances of being observed. Given the low rate of high-energy high-quality track events reconstructed in the detector, sensitivity increases as a function of the accumulated statistics. A possible way to increase the analysis sensitivity is to stack the events from multiple source candidate directions and to fit the likelihood once for all the source candidates together. Although the overall flux required to be observed may be slightly higher

with respect to a single source case due to a larger background rate, the flux per source can be dramatically decreased. This effect can be particularly powerful when searching for source candidates in the southern hemisphere where the point-source sensitivity is orders of magnitude lower than in the northern hemisphere. This fact is key when investigating galactic objects, since a large portion of the Galactic plane, and hence known nearby objects, are found in this region. To implement this method a relative flux contribution must be estimated for each source. This is applied as a weight in the analysis during the fit so that events around the object where the largest observable flux is predicted, will have a higher weight in the likelihood fit. By default this weighting assumes equal contributions from all the sources in the catalog.

7 Year Stacked Catalogs

Before the optimized 10 year data selection was completed, multiple stacking analyses were performed with the same data selection applied in Aartsen et al. [6] over only 7 years of IceCube data. Since the stacked analysis assumes the same spectra and steady emission from all of the stacked source candidates, instead of using one large catalog containing any motivated candidate, objects are sorted by type. Each catalog should be expected to have similar particle acceleration techniques in order to have the same spectral index, and also to choose a relative weighting scheme for the stacking.

Previous work published in Aartsen et al. [106] used 4 years of IceCube data to perform stacking searches targeting SNR, the Milagro 6 catalog, and starburst galaxies. This analysis benefits from 3 more years of data as well as updated SNR catalogs and a list of 10 new HAWC hotspots observed in the inner-Galactic plane [107]. The final catalogs implemented in this analysis were:

- Ten new HAWC hotspots in the inner-Galactic Plane [108]. Although some of these hotspots were associated with or close to known Galactic objects when the 7 year IceCube sample became available there was not strong enough evidence to motivate a weighting scheme other than uniform. The flux normalisation observed by HAWC for each hotspot was not implemented because this could depend on the relative exposure of HAWC which would differ to that of IceCube. Each of these 10 hotspots were tested assuming an average source extension of 0.5° as many of the objects in the inner Galactic plane are seen as extended. This catalog was inspired by the large overlap in sensitive energy range between the HAWC and IceCube detectors, permitting a multi-messenger view of the same candidate of astrophysical particle accelerators.

- The Milagro 6 catalog motivated by a single high energy gamma ray telescope Milagro. The Milagro catalog contains six Milagro sources in the Cygnus region originally reported by Abdo et al. [109] and modeled in Halzen et al. [110] and Gonzalez-Garcia et al. [111] as possible PeVatron candidates. This catalog has been used for stacking analysis previously using four years of IceCube data where a 2% p-value was found [106].

- Three new SNR catalogs were used to update the previous SNR results. The final three catalogs using the 7 year selection are sub-catalogs of a group of SNRs taken from SNR Cat [112] which have been observed in the TeV region with an age less than 3000 years. The selection of young SNRs was inspired by results demonstrating that SNRs less than 3000 years old are more efficient accelerators in the TeV region [113]. This group was then divided into three subgroups of sources based on their observed environment: those with known molecular clouds, those with associated PWN, and those with neither. A weighting of $\frac{1}{age*distance^2}$ was used for the sources corresponding to the stronger signal expected from young nearby sources.

- The extra-galactic catalog of 127 Starburst Galaxies. The presence of high-mass stars and high rates of star formation in a galaxy leads to a high rate of supernovae and associated SNR. SNR have already been discussed in Section. 1.4.1 as a possible accelerator of charged hadrons below the *knee* in the CR spectrum. This is the motivation behind stacking groups of known SNR within the Milky Way. A dense collection of potential CR accelerators from the Earth's perspective could lead to an interesting neutrino source. The associated dust clouds in an around these galaxies could also provide a dense target for proton-nuclear interactions which, despite blocking Gamma rays, would produce more neutrino emission. The Far Infrared (FIR) emission from these galaxies has been associated with the hot dust around these galaxies whilst radio observations are associated with synchrotron losses of accelerated electrons [105]. Assuming that both hadrons and electrons are accelerated together, then the high rate of star formation would explain the strong correlation between the FIR and radio observations. In this case a neutrino flux should also follow [48]. For this reasoning a list of known Starburst Galaxies compiled in Table A.1 in Becker et al. [105] is used as a stacking catalog in search of a neutrino signal. The relative source weighting is taken to be the FIR flux at 60 μm.

10 Year Stacked Catalogs

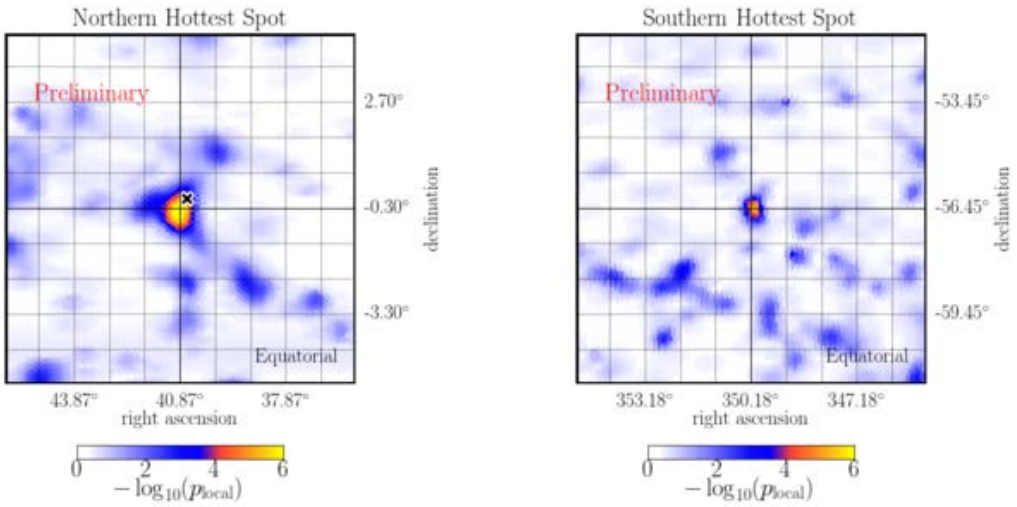
Using the 10 year IceCube point-source sample, 3 catalogs were stacked targeting galactic sources intending to update the Galactic Stacking results using the 7-year IceCube point-source event selection and using the most recent galactic source information. Galactic objects observed by gamma telescopes at TeV energies were assembled from the online catalog TeVCat [98]. These objects were then categorized into Pulsar Wind Nebula (PWN), Supernova Remnants (SNR) and Unidentified objects (UNID) in order to group objects likely to be neutrino emitters with similar spectra since a common likelihood fit is performed per category. The final lists were composed of 33 PWN, 23 SNR, and 58 UNID. Within each catalog the source candidates were then attributed a weight according to their estimated integral flux above 10 TeV. This weighting scheme was implemented as it was assumed to already take into account factors such as age, distance and luminosity without relying on each of these individual measurements and their associated errors. Since the stacking search benefits from using larger catalogs by lowering the per-source sensitivity flux, no further cuts were applied to the Galactic source candidates. One pre-trial p-value was found for each catalog determining the probability of the collective emission to be from background only. A single post-trial p-value is then evaluated for the most significant catalog from these all three Galactic catalogs.

Chapter 6

Search Results

6.1 All Sky Scan

In order to be sensitive to a neutrino source anywhere in the sky, including locations where other messenger particles have not yet provided strong motivation for the presence of a neutrino source, the entire sky is scanned by maximizing the test statistic values at a fine grid of points in the sky, as described in section 5.1. The grid of TS values is used to construct a map of pre-trial p-values describing the probability at each point that the observed neutrino emission originated from background only. The most significant p-value in the northern hemisphere was found to be $10^{-6.45}$ (at the RA 40.87° and a declination of −0.30°) as shown in Fig. 6.1a alongside the most significant point in the southern hemisphere in Fig. 6.1b. Fig 6.2 compares the significance of the unblinded northern and southern hotspots with the distributions of the most significant hotspots in each hemisphere generated from background trials. These background distributions are used to find the post-trial p-value for the unblinded hotspots in each hemisphere by calculating the fraction of background hotspots with a more significant p-value. Due to the extremely large number of trials when examining the entire hemisphere, the northern hotspot p-value significance is reduced to 9.9% post-trial. As can be seen in Fig. 6.1a, this hotspot is seen as slightly extended and it is in coincidence with the coordinates of the galaxy NGC 1068 in the Source List such that the hottest northern point falls only 0.35° away. The most significant hotspot in the southern hemisphere is less significant with a pretrial p-value of $10^{-5.37}$ which reduces to 75% when compared to the largest over-fluctuations in the hemisphere from many background trials. Both hotspots from the unblinded all-sky scan are therefore consistent with background fluctuations.



(A) Northern Hotspot (B) Southern Hotspot

FIGURE 6.1: Unblinded significance map of local p-values around the hottest points in the northern and southern hemisphere respectively. The black cross marks the coordinates used for the galaxy NGC 1068 in the source catalog.

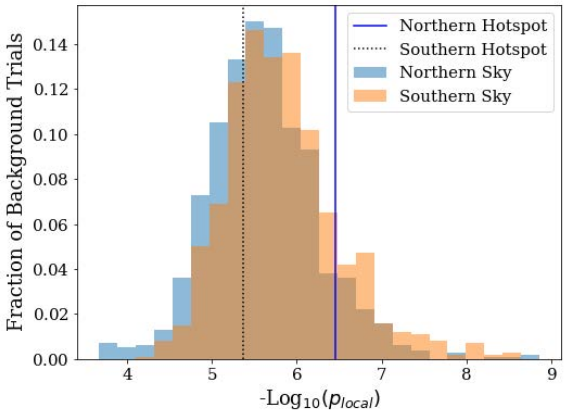


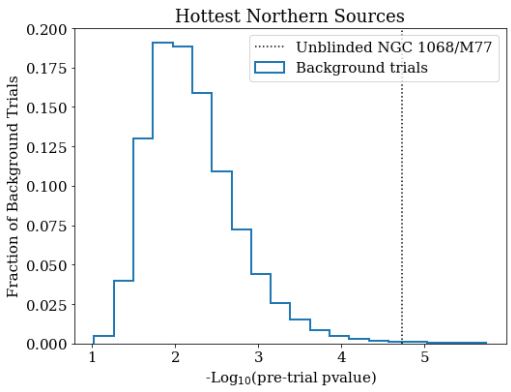
FIGURE 6.2: Distributions of the most significant p-values in the northern and southern hemisphere from all-sky scans of background scrambles of the sky compared with the most significant unblinded values. (dashed and solid lines for the south and north respectively)

6.2 Source List Results

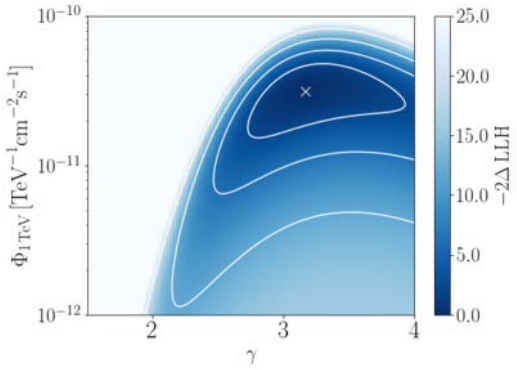
The coordinates of each potential astrophysical neutrino source in the individual source list are evaluated for the probability that the observed emission is due to background alone at the tested position. This analysis can reduce the penalty in significance from testing many different source hypotheses by limiting the source positions to only those motivated by recent gamma ray observations. The results of this search are shown in Table 6.3. Out of the 97 northern source candidates tested, excesses with pre-trial p-values less than 1% were found in the directions of 4 of these objects. These objects include 3 BL Lac objects: the BL Lac TXS 0506+056 with a pre-trial p-value of 1.89×10^{-4} for which evidence has recently been presented by IceCube for it to be a flaring neutrino source [5], as well as the BL Lac objects PKS 1424+240 and GB6 J1542+6129 with pre-trial p-values 1.59×10^{-3} and 1.82×10^{-3} respectively. The largest excess was found in the direction of the Seyfert Galaxy with known Starburst activity NGC 1068, with a pre-trial p-value of 1.84×10^{-5} which corresponds to 4.1σ pre-trial. In the southern hemisphere no excess was found with a p-value less than 1%, with the most significant excess found in the direction of the BL Lac PKS 2233-148 with a p-value of 5.4%. The post-trial significance for the most significant excess in the direction of the source candidates in each hemisphere is calculated by creating distributions of pre-trial p-values of the most significant excess in the catalog for each hemisphere using 10,000 background scrambles and calculating the fraction of the 10,000 scrambles with a more significant pre-trial p-value than that observed from the unblinded result. The probability of observing a more significant excess from an over-fluctuation in background than the unblinded result in the southern catalog is 0.53. This result is therefore completely consistent with background.

The most significant excess in the unblinded northern catalog, (4.1σ pre-trial seen in the direction of NGC 1068), resulted in a probability of 0.002 of occurring due to an over-fluctuation in a background only trial. The distribution of pre-trial p-values of the most significant source in the northern source list generated from 10,000 background trials is compared to the most significant unblinded p-value in the northern source list in Fig. 6.3a. This translates to a 2.9 σ level inconsistency with the background only hypothesis. This is below evidence level of 3 σ so it cannot be concluded that this is necessarily due to a neutrino source. However, the relatively high significance, not seen before by other individual point source searches by IceCube, demonstrates that this source location is of interest. In this case it is worth looking into the possible astrophysical flux that could cause such an excess. The fit power-law spectra from this excess is :

$$\frac{dN}{dE} = 3.14 \times 10^{-14} \left(\frac{\mathrm{E}}{\mathrm{TeV}}\right)^{-3.16} \mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1} \tag{6.1}$$



(A) Northern Hot Sources



(B) Spectral fit of Hottest Northern Source NGC 1068

FIGURE 6.3: The Observed pretrial p-value of the most significant Northern source, NGC 1068, as compared with the values from the most significant source from over 10,000 background trials. Results in 2.9σ post-trial p-value. The spectral fit for this source at 1 TeV shown in 6.3b

The flux normalisation and spectral index for the spectrum of this source are fit by maximizing the delta log likelihood map shown in Fig. 6.3b assuming an unbroken power-law spectrum at 1 TeV, ϕ_{TeV}. The fit spectra for this excess is seen as fairly soft and very close to the background spectrum from atmospheric events. This means that the significance at this location comes almost entirely from the spatial correlation of the IceCube events and only a minor part from the energy component. Fig. 6.4 shows from left to right the distribution of events in RA and declination weighted by the spatial PDF, energy PDF, and their global product assuming a point source hypothesis in the direction of NGC 1068. It can be seen that the energy weights are uniform in space around the direction of NGC 1068. This result must then be due to an excess in events with overlapping direction reconstructions within their estimated angular uncertainties but indicates there are no particularly high energy events.

Figure 6.5b compares the distribution of the angular distance between the events and NGC 1068 and the distribution of the angular uncertainty of these events both weighted by the signal over background likelihood assuming a neutrino source at the coordinates of NGC 1068. This plot shows that the biggest contribution to the likelihood comes from events reconstructed within 1° of the position of NGC 1068 and from events which have an estimated angular uncertainty less than 1°. The energy distributions of events within different radial distances of NGC 1068 in Fig. 6.5a show that in the considered region there is only 1 event greater than 10 TeV. At these energies the energy distribution is still dominated by atmospheric background events. This means that, although there is an excess of clustered events, it is difficult for the fit to establish precisely how many events

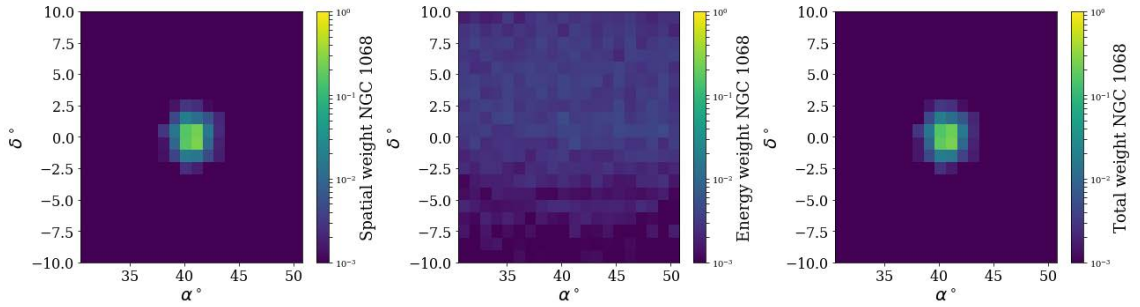
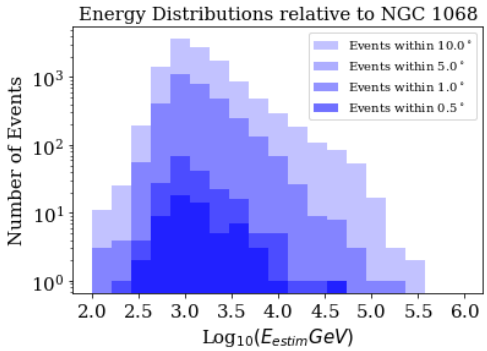


FIGURE 6.4: 2D event distribution around the coordinates of NGC 1068 where are weighting by the Spatial PDF (left), the energy PDF (middle) and the final Signal over Background likelihood (right). It can be seen that the final likelihood is dominated by the spatial clustering and the event weight plays a negligible role.

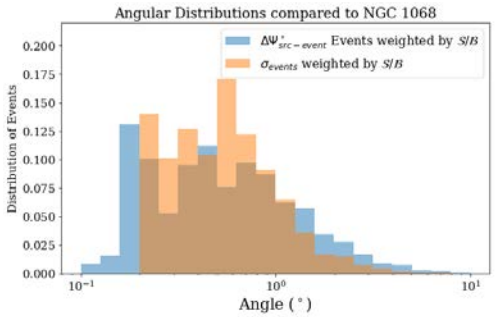
may be due to an astrophysical source signal, and how many are from the atmospheric flux until such clustering reaches above 5σ significance. This is why the best fit of the spectrum in Fig. 6.3b shows large errors, in particular the spectral index γ displays errors of $\sim \pm 0.5$. Given the low numbers of high-energy events observed around NGC 1068, the likelihood drops significantly for harder spectral indices and a high flux normalisation, whereas there are larger uncertainties for lower values of flux normalisation.

Although the excess in the direction of NGC 1068 was not significant enough to be considered as evidence for a neutrino source, tests were done to evaluate how feasible it would be for a neutrino source to give the same results. In this case, the significance in the direction of NGC 1068 comes from an extended hotspot in the sky seen in Fig. 6.1a where the hottest point is actually 0.35° degrees away. However, NGC 1068 is an extragalactic object and so any associated neutrino flux would be expected to appear point-like with respect to IceCube's angular resolution. To study the possibility that the 0.35° offset between the all-sky hotspot and NGC 1068 is a statistical fluctuation, we inject a soft-spectrum source similar to the best-fit $E^{-3.2}$ flux at the position of NGC 1068 in our background samples. Scanning in a 5° window around the injection point, we find the most significant hotspot is identified within 1° of the injection point 70% of the time, and within 0.35° of the injection point 51% of the time. Thus, it is plausible that the offset between the all-sky hotspot and NGC 1068 is a statistical fluctuation.

Regarding the extension of such offset hotspots, one can examine the difference in TS value between the shifted hotspot and the TS at the true source position for the same background scrambles with the best fit spectrum injected. Fig. 6.6 demonstrates that when the NGC fit spectrum is injected the hottest spot 1° around the source is often found 0.2° away, and in this case the TS is rarely less than 30% lower. This implies that such a hotspot has a diameter of at least this distance. By examining the red curve which



(A) Energy distribution of events within different angular distances of NGC 1068



(B) The distribution of angular distances from NGC 1068 and the estimated angular errors weighted by their signal over background likelihood.

FIGURE 6.5: The distributions of events from the 10 year selection which contributed towards the 2.9σ excess seen when testing the NGC 1068 direction

represents values from the unblinded result it can be also seen that, when the hotspot is shifted 0.35°, a 30% drop the TS value between the hotspot and the source could be fairly typical, although it is still not the peak in the distribution. Thus the excess in the direction of NGC 1068 is currently consistent with both a background only or an injected soft spectrum hypothesis.

6.3 Stacked Source Searches

The stacking analysis combines the emission observed from the directions of multiple potential neutrino sources and performs a single fit for the estimated astrophysical neutrino flux. This analysis assumes a catalog of sources with a similar spectra and a weighting scheme which describes their estimated relative contribution to a total astrophysical neutrino flux. If the source assumptions are correct this analysis can reduce the required flux per source to discover the astrophysical neutrino sources. Each of the three Galactic catalogs was evaluated using the point-source likelihood to fit a single spectrum per catalog. All the results from the 10 year stacking searches are shown in Table. 6.1. All of the pre-trial p-values from these three searches are consistent with background fluctuations. The PWN catalog returned a test statistic of 0 (corresponding to a p-value of 1.0) which indicates there was no over fluctuation at all when the positions of these sources were stacked. The UNID catalog returned a low test statistic indicating a slight excess in the source directions and the most significant p-value of 11% comes from the SNR catalog. This inconsistency with background is not significant and so all three catalogs are considered consistent with the background hypothesis.

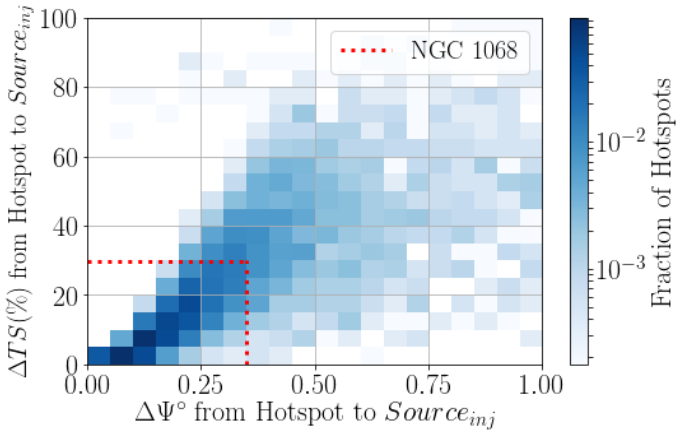


FIGURE 6.6: 2D distribution of background trials of a 1° scan around the NGC 1068 injecting the best fit spectrum of the unblinded results. Showing the percentage difference in TS between the hottest spot and the source position is shown as a function of the angular distance between them.

10 year Stacking Results

Source Catalog	Number of sources	p-value	n_s	γ	$\phi_{90\%}$
Supernovae Remnants	23	0.11	23.9	3.55	4.96×10^{-15}
Unidentified Objects	58	0.4	3.28	2.39	1.56×10^{-15}
Pulsar Wind Nebula	33	1.0	-	-	2.64×10^{-15}

TABLE 6.1: Note: ns and γ are the best-fit number of signal events and spectral index, respectively. Fluxes are given as the sum over the catalog and parameterized as $d\phi_{\nu_\mu+\bar{\nu}_\mu}/dE = \phi_{90\%}\cdot(E/TeV)^{-2.}\ GeV^{-1}cm^{-2}s^{-1}$ with 90% confidence level upper limits quoted for $\phi_{90\%}$.

7 year Stacking Results

Source Catalog	Number of sources	p-value	n_s	γ	$\phi_{90\%}$
Milagro 6	6	0.3	31.8	3.95	4.96×10^{-20}
HAWC hotspot	10	0.31	17.3	2.38	9.48×10^{-21}
SNR with mol. clouds	10	0.25	16.5	3.95	2.23×10^{-19}
SNR with PWN	9	0.34	9.36	3.95	1.17×10^{-18}
SNR alone	4	0.42	3.82	2.25	2.06×10^{-19}

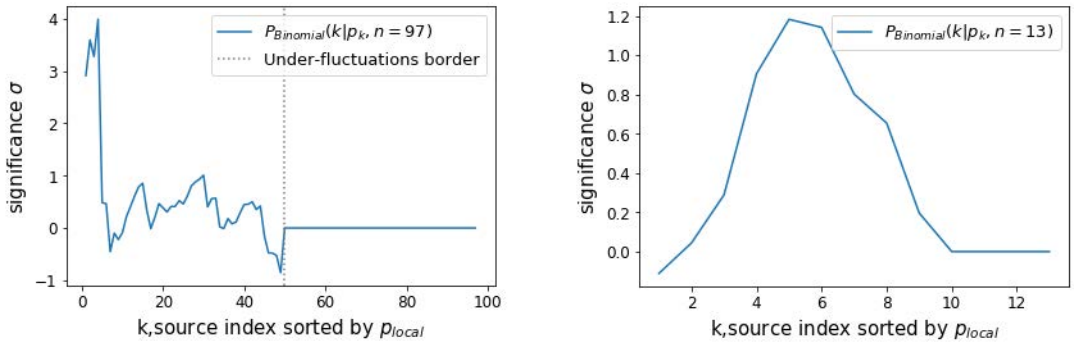
TABLE 6.2: Note: ns and γ are the best-fit number of signal events and spectral index, respectively. Fluxes are given as the sum over the catalog and parameterized as $d\phi_{\nu_\mu+\bar{\nu}_\mu}/dE = \phi_{90\%} \cdot (E/100TeV)^{-2.5}\ GeV^{-1}cm^{-2}s^{-1}$ with 90% confidence level upper limits quoted for $\phi_{90\%}$.

The results of the 7-year Galactic searches are shown in Table 6.2 [8]. All of the catalogs found statistically insignificant excesses. The most significant result with a p-value of 25% results from SNR with molecular clouds, which fit 16.5 excess events and a soft spectral index approaching the limit of $\gamma < 4.0$. The most promising Galactic catalog for stacked analysis was that of the six Milagro sources. The previous iteration of this search found a p-value of 2% [106]. However, the updated results of this search, using 3 more years of IceCube data, showed a increase in p-value from 2% to 29%. This result is able to exclude the model of Halzen et al. [110] at a 90% confidence level. The extra-galactic Starburst catalog was also found to be totally consistent with background with an unblinded p-value of 0.473. This results in a 90% upper-limit of 4.01×10^{-15} $\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ for a spectrum $d\phi_{\nu_\mu+\bar{\nu}_\mu}/dE = \phi_{90\%} \cdot (E/\mathrm{TeV})^{-2}$ $\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

6.4 Source List Population Search

The source list population analysis tests how consistent and entire catalog is with the background only hypothesis by looking for an excess of significant sources from a catalog of N sources. This is implemented by calculating the probability of observing k or more out of N sources pass a given significance threshold and iterating the threshold to allow k to vary between 1 and N.

The unblinded binomial probabilities associated with each combination of sources, ordered by their significance are shown in Figure 6.7. It can be seen that the most significant combination in the northern hemisphere comes from the probability that background could reproduce 4 or more sources with a pre-trial p-value of the 4^{th} most significant unblinded source or less. The associated pre-trial p-value is found as $p_{pre-trial} \leq 10^{-2.7}$. This significance becomes 3.3 σ post-trial when the same process of selecting the most significant



(A) Northern Sources (B) Southern Sources

FIGURE 6.7: Unblinded significance for binomial p-values from the individual source list, calculated as a function of the number of sources passing the threshold for the least significant event.

combination of source in the catalog is applied to 10,000 background trials. This significance does not directly correspond to the probability of the individual objects being neutrino sources. This result demonstrates an inconsistency with the hypothesis of only background fluctuations in the source list to a significance of 3.3 σ. The most probable reason for the inconsistency would be the presence of at least one neutrino source amongst the directions tested by the Northern Catalog. The result from the population analysis cannot determine which candidates are the most likely neutrino sources. To have an idea of which are the most likely candidates one must examine the individual p-values for the source candidates from Section 6.2. The sources making up the excess in this result are NGC 1068, TXS 0506+056, PKS 1424+240, and GB6 J1542+6129. Since evidence has already been presented for TXS 0506+056 to be a flaring neutrino source, this search was checked for the contribution without this source. The pre-trial significance went from 4 σ to 3.17 σ for the same 3 remaining sources, which results in 2.25 σ post-trial inconsistency with a background only hypothesis.

6.5 Outlook

The source list catalog with pre-trial p-values and TS values for each object in the list can be used to establish an upper-limit on the possible astrophysical neutrino flux in each location. This upper-limit is calculated in a similar way to the sensitivity described in Section 4.2.1. The difference is that instead of injecting a mean signal until 90% of trials result in a TS which surpasses the median of background TS distribution, the TS threshold is raised to be the unblinded TS value in the given source.

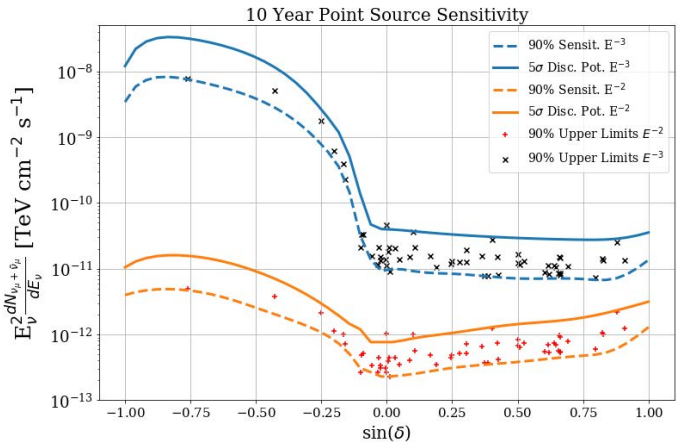


FIGURE 6.8: 90% Sensitivity and 5σ Discovery flux for the 10 year event selection given a hard E^{-2} and a soft E^{-3} spectrum as compared to the 90% upper limits for the sources in the source list with over-fluctuations for the hard and soft spectra respectively.

In Fig. 6.8 the 90% upper-limit for each of the over-fluctuations found in the source catalog are plotted for a hard spectrum of E^{-2} and a soft spectrum of E^{-3} as compared to the expected 90% sensitivity flux and the 5σ discovery potential flux for the 10 year data selection. The 90% upper limit for four of the sources in the source list can be seen to lie very close to the 5σ discovery flux. These are the same as the four most significant source in the source list highlighted by the catalog population analysis (NGC 1068, TXS 0506+056, PKS 1424+240, and GB6 J1542+6129). This comparison means that we can place a 90% C.L. on having seen a flux greater than a $\sim$ 5σ discovery level flux in the direction of these objects with 90% confidence. This is consistent with the source results where the excesses in the directions of these objects were found at roughly 3-4 σ, but not 5 σ. However the limit is well above the sensitivity flux which demonstrates the region where IceCube expects to be sensitive to.

The same method can be applied to the stacking analysis, although the results correspond to the entire catalog and applied weighting scheme and so cannot be attributed to a single direction. The 90% upper-limits for this analysis are included in the Table 6.1. Although for the PWN catalog where the TS= 0, the 90% upper-limit is the same as the sensitivity flux.

Given the 2.9 σ excess seen for NGC 1068, it can be interesting to examine what other telescopes have observed in the direction of this source, especially at energies approaching those of IceCube events. NGC 1068 is the closest and brightest Seyfert 2 galaxy to Earth. This galaxy has an AGN with a mis-aligned jet and known star-burst activity [116][117]. Both the AGN jet and a diffuse shock generated between the winds from starburst activity

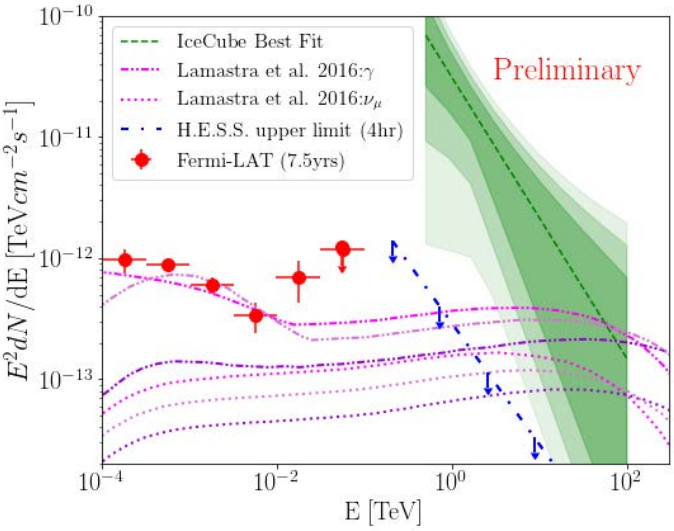


FIGURE 6.9: The best fit steady state astrophysical power law flux using the 10 year IceCube event selection in the direction of NGC 1028 including the 1, 2 & 3σ error regions on the spectrum as seen in Fig. 6.3b. This fit is compared to the γ and ν models and the Fermi Pass8 (P8) results found in Lamastra et al. [114]. The Upper-limit from HESS is taken from Aharonian et al. [115].

and the surrounding Inter-Stellar Medium (ISM) could explain the high fluxes in the electromagnetic spectrum observed from NGC 1068 [118]. Models for the resulting neutrino flux given CR acceleration from shocked winds around the AGN are described in Lamastra et al. [114], although these models currently under-estimate the observed gamma ray flux. A summary of gamma ray observations and models can be seen in Fig. 6.9. The current best fit seen by IceCube in the direction of NGC 1068 appears to have a much larger normalisation and a softer spectrum than already observed by gamma ray telescopes such as Fermi [104][114]. It is possible for the neutrino spectrum to differ slightly from the gamma ray spectrum if gamma rays are absorbed or scattered by dense clouds of gas surrounding the acceleration region, as is the case for this source. Although depending on the surrounding matter one could expect the high energy TeV gamma rays to cascade down in energy and appear as a higher flux at lower energies. Without a well observed astrophysical neutrino flux for any similar object it is difficult to assess how feasible such a difference between the neutrino and gamma flux could be.

The likelihood of this excess originating from an astrophysical source could be established with either a longer lifetime of data or a more precise angular resolution. Given a more precise angular resolution, especially for lower energy events, an astrophysical neutrino source with a softer spectrum close to background could be spatially resolved within a less extended hotspot. In this case a background fluctuation would be expected to become

more distinctly off-set such that the hotspot is no-longer in coincidence with the coordinates of NGC 1068. In the case of a true neutrino source the excess would be expected to converge more closely to the coordinates of the neutrino source. Similarly having a longer lifetime would allow the total number of potential signal events to increase assuming a steady neutrino state source. This would result in a higher n_s which should also be easier to separate from the background rate. Assuming roughly an order of magnitude improvement in the sensitivity after 15 years of IceCube Gen-2[119], the discovery potential of the point source analysis could fall to a smaller flux normalisation than the best fit neutrino flux of NGC 1068 from this analysis which lies very close to the flux upper-limit for a 90% confidence level (C.L.). In this case using the reconstruction improvements and increased statistics from Gen-2, NGC 1068 could either be observed as an astrophysical neutrino source or the current best-fit spectrum would be ruled out at a 90% C.L.

TABLE 6.3: Results of the *a priori* defined source list search. Coordinates are given in equatorial coordinates (J2000). The neutrino Flux upper-limit ($\phi_{90\%}$) is parametrized as: $\frac{dN_{\nu_\mu+\bar{\nu}_\mu}}{dE_\nu} = \phi_{90\%} \cdot \left(\frac{E_\nu}{TeV}\right)^{-2} \times 10^{-13}\text{TeV}^{-1}\text{cm}^{-2}\text{s}^{-1}$

Source List Results

Name	α [deg]	δ [deg]	$\hat{n}_s$	$\hat{\gamma}$	-Log$_{10}(p_{local})$	$\phi_{90\%}$
Mkn 421	166.12	38.21	2.10	1.92	0.38	5.29
PG 1553+113	238.93	11.19	0.00	2.82	0.32	3.17
BL Lac	330.69	42.28	0.00	2.70	0.31	4.90
PKS 0235+164	39.67	16.62	0.00	3.03	0.28	3.07
TXS 0518+211	80.44	21.21	15.75	3.85	0.92	6.64
PKS 1424+240	216.76	23.80	41.47	3.94	2.80	12.32
S5 0716+71	110.49	71.34	0.00	2.51	0.38	7.41
3C 66A	35.67	43.04	0.00	2.80	0.30	3.86
4C +01.28	164.61	1.56	0.00	2.91	0.26	2.40
TXS 0506+056	77.35	5.70	12.32	2.08	3.72	10.10
Mkn 501	253.47	39.76	10.27	4.00	0.61	7.32
S2 0109+22	18.03	22.75	2.02	3.13	0.30	3.70
B2 1215+30	184.48	30.12	18.60	3.39	1.09	8.45
PKS 0735+17	114.54	17.71	0.00	2.84	0.30	3.54
MG1 J021114+1051	32.81	10.86	1.55	1.75	0.43	3.48
OT 081	267.87	9.65	12.21	3.22	0.73	4.80
OJ 287	133.71	20.12	0.00	2.63	0.32	3.49
1H 1013+498	153.77	49.43	0.00	2.58	0.29	4.47
OJ 014	122.87	1.78	16.07	4.00	0.99	4.43

Continuation of Source List Results: Table 6.3

Name	α [deg]	δ [deg]	$\hat{n}_s$	$\hat{\gamma}$	-Log$_{10}(p_{local})$	$\phi_{90\%}$
TXS 0141+268	26.15	27.09	0.00	2.53	0.31	3.45
NVSS J141826-023	214.61	-2.56	0.00	2.96	0.25	2.05
PKS 0829+046	127.97	4.49	0.00	2.93	0.28	2.13
1ES 1959+650	300.01	65.15	12.60	3.27	0.77	12.33
1H 1720+117	261.27	11.88	0.00	2.66	0.30	3.18
RGB J2243+203	340.99	20.36	0.00	3.01	0.33	3.14
S4 0814+42	124.56	42.38	0.00	2.30	0.30	4.90
RX J1931.1+0937	292.78	9.63	0.00	2.88	0.29	2.76
PKS B1130+008	173.20	0.58	15.79	4.00	0.96	4.41
MG4 J200112+4352	300.30	43.89	6.15	2.26	0.67	7.75
ON 246	187.56	25.30	0.94	1.71	0.37	4.20
B3 0609+413	93.22	41.37	1.81	1.68	0.42	5.27
1RXS J194246.3+1	295.70	10.56	0.00	2.68	0.33	2.60
B3 0133+388	24.14	39.10	0.00	2.61	0.28	4.08
GB6 J1542+6129	235.75	61.50	29.72	3.02	2.74	22.01
W Comae	185.38	28.24	0.00	2.98	0.32	3.66
PG 1246+586	192.08	58.34	0.00	2.83	0.35	6.35
PKS 0422+00	66.19	0.60	0.00	2.92	0.27	2.34
1ES 0647+250	102.70	25.06	0.00	2.90	0.27	3.01
PKS 0048-09	12.68	-9.49	3.92	3.27	0.87	10.03
PKS 1717+177	259.81	17.75	19.82	3.65	1.32	7.28
1ES 0806+524	122.46	52.31	0.00	2.75	0.31	4.71
TXS 1902+556	285.80	55.68	11.70	4.00	0.85	9.95
PKS 0019+058	5.64	6.14	0.00	2.87	0.29	2.44
PKS 2233-148	339.14	-14.56	5.32	2.80	1.26	21.41
PG 1218+304	185.34	30.17	11.13	3.94	0.70	6.68
MG2 J043337+2905	68.41	29.10	0.00	2.66	0.28	4.55
B2 2114+33	319.06	33.66	0.00	3.00	0.30	3.89
S4 1250+53	193.31	53.02	2.22	2.49	0.39	5.94
S4 1749+70	267.15	70.10	0.00	2.48	0.37	7.99
PKS 1216-010	184.64	-1.33	6.88	4.00	0.45	3.13
3C 454.3	343.50	16.15	5.41	2.24	0.62	5.10
PKS 1510-089	228.21	-9.10	0.09	1.70	0.41	7.11
3C 279	194.04	-5.79	0.26	2.38	0.20	2.67

Continuation of Source List Results: Table 6.3

Name	α [deg]	δ [deg]	$\hat{n}_s$	$\hat{\gamma}$	-Log$_{10}(p_{local})$	$\phi_{90\%}$
PKS 1502+106	226.10	10.50	0.00	2.99	0.33	2.60
CTA 102	338.15	11.73	0.00	2.70	0.30	2.77
4C +21.35	186.23	21.38	0.00	2.59	0.32	3.49
4C +01.02	17.16	1.59	0.00	3.00	0.26	2.40
B2 1520+31	230.55	31.74	7.09	2.40	0.83	7.31
4C +38.41	248.82	38.14	4.21	2.28	0.66	7.03
3C 273	187.27	2.04	0.00	2.99	0.28	1.89
PKS 0502+049	76.34	5.00	11.19	3.02	0.66	4.05
MG1 J123931+0443	189.89	4.73	0.00	2.61	0.28	2.43
PKS 0336-01	54.88	-1.77	15.50	4.00	0.99	4.42
PKS 0736+01	114.82	1.62	0.00	2.79	0.26	2.40
OG +050	83.18	7.55	0.00	3.24	0.28	2.87
MG2 J201534+3710	303.92	37.19	4.43	4.00	0.40	5.55
PKS 2320-035	350.88	-3.29	4.79	3.63	0.45	3.34
PKS 0727-11	112.58	-11.69	1.94	3.51	0.59	11.36
PKS 2032+107	308.85	10.94	0.00	2.45	0.33	3.19
PKS 0805-07	122.07	-7.86	0.00	2.73	0.31	4.68
4C +28.07	39.48	28.80	0.00	2.79	0.30	3.55
PKS 0420-01	65.83	-1.33	9.31	4.00	0.52	3.36
Ton 599	179.88	29.24	0.00	2.21	0.29	4.55
PKS 0440-00	70.66	-0.29	7.59	3.94	0.46	3.11
S3 0458-02	75.30	-1.97	5.45	4.00	0.33	2.67
4C +55.17	149.42	55.38	11.88	3.28	1.02	10.64
B3 1343+451	206.40	44.88	0.00	2.83	0.32	4.96
B2 0218+357	35.28	35.94	0.00	3.08	0.33	4.35
PKS 1441+25	220.99	25.03	7.49	2.38	0.94	7.34
4C +14.23	111.33	14.42	8.54	2.89	0.60	4.81
OX 169	325.89	17.73	2.02	1.73	0.69	5.09
PKS 0215+015	34.46	1.74	0.00	3.16	0.27	2.32
TXS 2241+406	341.06	40.96	3.76	3.79	0.42	5.65
PKS 1329-049	203.02	-5.16	6.06	2.68	0.77	5.08
NVSS J190836-012	287.20	-1.53	0.00	2.86	0.22	2.31
GRS 1285.0	283.15	0.69	1.67	3.82	0.27	2.27
NGC 1275	49.96	41.51	3.63	3.09	0.41	5.53

Continuation of Source List Results: Table 6.3

Name	α [deg]	δ [deg]	$\hat{n}_s$	$\hat{\gamma}$	-Log$_{10}(p_{local})$	$\phi_{90\%}$
PMN J0948+0022	147.24	0.37	9.32	4.00	0.76	3.92
M 87	187.71	12.39	0.00	2.76	0.29	3.14
PKS 1502+036	226.26	3.44	0.00	2.75	0.28	2.86
LMC	80.00	-68.75	0.00	3.10	0.36	41.10
NGC 1068	40.67	-0.01	50.39	3.16	4.74	10.47
M 82	148.95	69.67	0.00	2.61	0.36	8.76
SMC	14.50	-72.75	0.00	2.37	0.37	44.14
M 31	10.82	41.24	10.99	4.00	1.09	9.59
NGC 253	11.90	-25.29	3.01	4.00	0.75	37.70
NGC 598	23.52	30.62	11.40	4.00	0.63	6.32
NGC 4945	196.36	-49.47	0.31	2.65	0.31	50.25
Crab nebula	83.63	22.01	1.08	2.23	0.31	3.67
HESS J1841-055	280.23	-5.55	3.55	4.00	0.55	4.84
HESS J1843-033	280.75	-3.30	0.00	2.78	0.31	2.50
MGRO J1908+06	287.17	6.18	4.22	1.96	1.42	5.70
HESS J1837-069	279.43	-6.93	0.00	2.84	0.30	4.03
HESS J1857+026	284.30	2.67	7.44	3.12	0.53	3.53
PSR B0656+14	104.95	14.24	8.36	4.00	0.51	4.45
MGRO J2019+37	304.85	36.80	0.00	3.08	0.33	3.98
HESS J1849-000	282.26	-0.02	0.00	2.99	0.28	2.22
HESS J1852-000	283.00	0.00	3.28	3.75	0.38	2.63
Gamma Cygni	305.56	40.26	7.43	3.75	0.59	6.92
2HWC J2031+415	307.93	41.51	13.36	3.77	0.97	9.16

End of Results

Chapter 7

Gender in Physics at the University of Geneva

In order to probe the boundaries of scientific knowledge and discover new phenomena, we are past the days of people being able to derive and test hypotheses alone or using few resources. In practice, for hundreds of years interested parties have gathered to work together sharing their knowledge and resources at universities and research institutes. The ability to pool together resources like this has allowed dramatic improvements in research, a greater rate of new ideas and the construction of elaborate equipment like telescopes able to resolve the sky to a brand new precision.

For most of recorded history, documented scientific breakthroughs have been attributed to men, with women often being assigned more menial tasks or discouraged from participating at all. The first female professor in Europe was Laura Bassi in Italy in 1732 [120], although it was not until the twentieth century that scientific research started to become accessible to many women in Europe. Despite increasing participation of women in science and their contributions, there is still a well-documented discrepancy between the incidence of women compared with men in the senior positions in science and technology. Some subjects, such as medicine and biology, have grown to see equal to slightly higher ratios of women to men at early stages in their career and studies in some countries; however this ratio falls dramatically at professor level. The physical sciences and engineering typically start with a low proportion of girls at the student level. It is thought that by tackling this gender bias in science, the scientific field could achieve a fairer and healthier working environment. It should also be the case that by harnessing the potential from the entire population by not impeding able researchers from minorities from accessing senior positions, the field would benefit with more high quality research output.

Gender Equality Plans (GEPs) have been proposed in order to evaluate how fair the scientific environment is to all genders and to help improve any noticeable injustices. These plans mark out how an institution aims to monitor and act on known gender issues. For

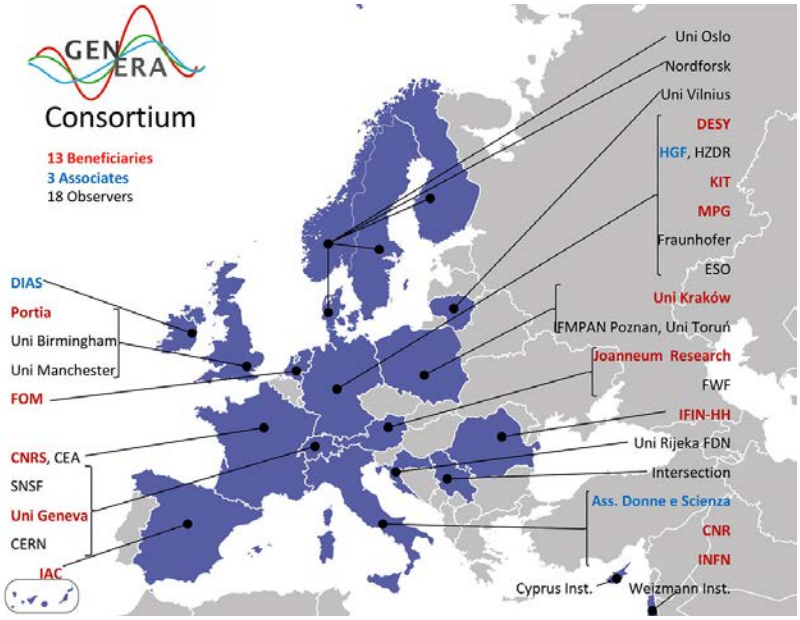


FIGURE 7.1: Map of different institutes participating in the GENERA Project [7]

these reasons the GENERA Consortium [7], a Horizon 2020 [121] project, seeks to aid in creating or continuing and improving Gender Equality Plans specialising in Physics, for its participating institutions. In total, thirteen different institutes across Europe are partnered with GENERA as seen in Fig. 7.1. Representatives from each of these institutions met together to derive a uniform strategy that could be tailored to each institution by engaging with local scientists, senior administrative staff, and local leadership.

This project aimed to establish long term practices and monitoring techniques able to pinpoint issues contributing towards gender equality in Physics. This included benchmarking already active or past measures and sharing these ideas with other members of the GENERA consortium. The minimum required actions to help initiate the construction or improvement of local GEPs were to supply statistics from a minimum data set and to organise a Gender in Physics Day (GiP Day) where Gender equality measures could be discussed and awareness in the community could be raised.

7.1 Gender in Physics Day (GiP) at Geneva

In January 2017 the PI of GENERA in Geneva, Prof. Montaruli, and myself, who acted as Implementation Manager at the Université de Genève (UNIGE) throughover the project, with the help of Maya Widmer, formerly in charge of the Equality program of the Swiss funding agency SNSF, worked to organise a GiP day at the University. The primary topic

of this meeting was the analysis of the existing actions towards gender equality practices in higher education institutes for Physics and Astronomy in Switzerland. Since UNIGE was the only GENERA Partner in Switzerland this involved inviting Research Performing Organizations (RPOs) and Research Funding Organizations (RFOs), Cantonal and Federal Institutes in Switzerland to discuss current issues and future prospects for gender equality.

On behalf of UNIGE, the statistics already taken for the minimum data set and plans already in motion for the Physics Section were presented. Members from other local Swiss universities and Federal institutes, such as the Swiss National Science Foundation (SNSF) were equally invited to demonstrate the work in their sectors. Following these presentations a panel discussion on the proposals for GEPs was chaired by Maya Widmer, as well as the head of the Equality Office at UNIGE, Brigitte Mantilleri. Leading the discussion were: Prof. Thomas Brage, University of Lund (Sweden), Prof. Camille Bonvin (UNIGE), Dr. Laura Cattaneo (ETHZ), Prof. Michel Oris, Vice-rector (UNIGE), and Dr. Susanne Wampfler (University of Bern). This discussion aimed to determine which proposals were believed to be the most inspiring and sustainable in terms of increasing the numbers of women in Physics.

The GENERA project is a collaboration of people joining forces to share ideas and monitor each other's progress. It was thus beneficial for other GENERA members to attend the UNIGE GiP day, especially since Geneva was one of the first partners to plan a GiP day. This way lessons from the day could be passed on when planning future similar events. Given the aim of raising female participation at all levels, the topic of science and gender equality in education was also included, as well the connection between University and School. The target group was people working in science or education or anyone else who could benefit from understanding in more detail the issues of today in gender equality in STEM. There were presentations by A. Gasperini, a Physics School professor very engaged in Gender issues, and about two initiatives, Athena by Prof. Maggiore and Physiscope by Prof. Renner. Physiscope, now extended to all Sciences (Sciencescope) [122], hosts thousands of school students to illustrate scientific phenomena at the UNIGE premises, and it is often present in the media. Athena is a program born in the Section of Physics at UNIGE [123] during which selected college students in their last year can prepare a science university exam being followed by PhD or Postdocs in the University. The Athena Program saw success in rates of female participation reaching over 50% [123].

It was from this GiP day that the priority of having more access to child-care emerged as a consistent issue for many women aiming at more senior positions in research.

After the GiP the GENERA project was invited by the Commission d'égalité of the Science Faculty and proposed a series of possible actions in the field domain identified by GENERA.

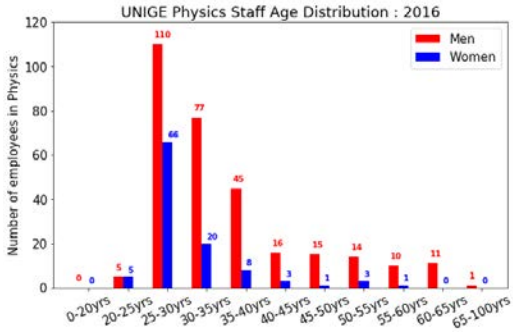
7.2 Monitoring Career Status

It has been already mentioned that GENERA required a minimum data set to monitor the status of gender equality at each institute. Initially the aim was to have access to the full statistics from Human Resources (HR) to compare the possible differences between men and women under the following categories:

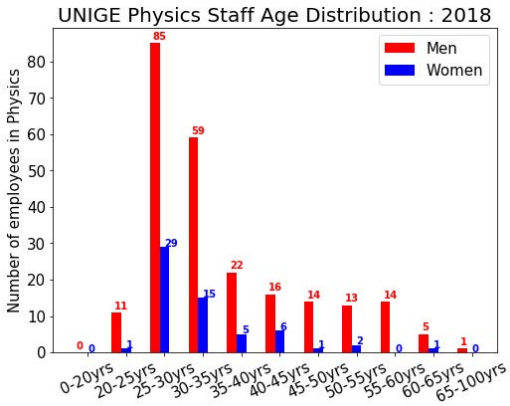
- Demographic Data:
  - Gender, Age, Civil Status, Country of Birth, Citizenship, Number of Children.
- Qualifications
  - Level of education, Field of Degree.
- Career Path
  - Contract Start date, Type of Contract (Full-time/Part-time), Long-term or Short Term contract, Level of contract, Level of first contract.
- Responsibilities
  - Start date, end date, Type of responsibility (Project/Laboratory/Teaching).
- Leaving Periods
  - Start date, end date, Reason for break (Parental leave/Career Break/Other).
- Research Output
  - Journal articles, Conference contributions, Chapter in Book, Patent/Trademark.

Most of this information was unavailable via HR, or we could not obtain permission to have access to it. This was a common problem among the different GENERA partners and so the Minimum data-set (MDS) was established. This MDS was designed to be used to monitor and compare all the partner institutions; as such the variables had to be kept uniform and not institute or country specific and also limited to the information to which every partner institution was able to supply. The result was:

- Gender: (Man/Woman)
- Age: In 5 years bins
- Job Level: 4 different career levels from PhD student to full permanent professor were defined.



(A) Distribution of Staff ages in the Physics Section at UNIGE in November 2016.



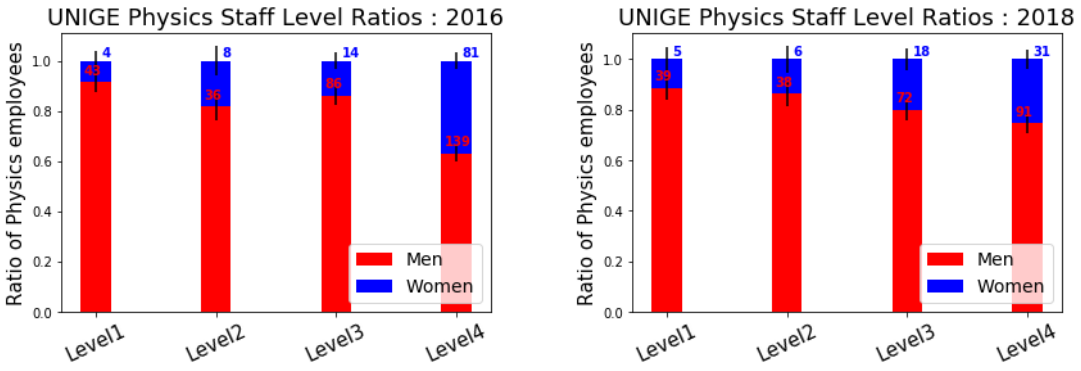
(B) Distribution of Staff ages in the Physics Section at UNIGE in November 2018.

FIGURE 7.2: The distributions of ages of Staff in the Physics section in 2016 and 2018.

The Job Levels are defined as:

- **Level 1: Researcher/Research Director/Professor**: Top researchers, full professors, research directors. This level includes the top management and organization directors. Activities primarily focused on high level research and specialist teaching, including research project leading/ research and institutional policy orientation.
- **Level 2: Established Researcher/Senior scientist/Assistant professor**: Senior researchers/scientists/lecturers including research managers and research group leaders. Activities could include research management, research group coordination and teaching.
- **Level 3: Recognised Researcher/Post-doc/Junior Academic**: Researchers/ scientists/ lecturers and post-doctoral scientists. Activities primarily focused on research, teaching, and program management. This is considered the basic level of the career development in the organizations.
- **Level 4: First stage researcher/Research assistant/Doctoral candidates** Research and project assistants/doctoral students/research grant winners including any researcher without doctoral degree. Activities focused on supervised research and training. Frequently participating in a structured program or working collaboratively in a larger team. They may have student status and/or be employed by the organizations. This is considered the entry level in the organization.

At UNIGE, data was granted for this work for the entire Faculty of Sciences covering: Age, Gender, Civil Status, Nationality, Activity Rate, Job Title, Department, and Contract



(A) Distribution of Staff Career Levels in the Physics Section at UNIGE in November 2016.

(B) Distribution of Staff Career Levels in the Physics Section at UNIGE in November 2018.

FIGURE 7.3: The distributions of Staff career levels in the Physics section in 2016 and 2018.

Start Date. This information was gathered in 2016 and again in 2018 and was more than enough to fulfill the MDS. Plots showing the results of the MDS can be seen in Fig. 7.2 for the age distributions and Fig. 7.3a, 7.3b for the Career levels. Fig 7.2 shows that at all ages there are fewer women than men. The fraction of women at each age group starts to fluctuate due to low statistics at the higher ages but it is clear that the fraction of women is significantly lower at older age ranges. For 2016 the majority of the staff (69%) are less than 35, and 32% of these staff are women. However for the more senior staff who are over 40 years old, only 10.7% are women. The staff become skewed to an even younger cohort in 2018 with 72.2% less than 35 years of ages and amongst this group only 25% are women whilst above 40 years the picture is largely unchanged with about 12% women.

This difference is reflected in the ratios of women at different hierarchical levels in Fig. 7.3. Since the ratio is calculated using a trial with n elements with a binary result of staff being *Men* or *Women*, a binomial error can be estimated as :

$$\sigma_{ratio} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{p(1-p)}{n}} \tag{7.1}$$

where p is the probability of staff being a woman, and q the probability of them being a man. Using this information to compare the changes between 2016 and 2018 it can be seen that the senior positions Level 1 and 2 are well within errors of about 4-5%. Even at Level 3 where the errors are roughly 3.5% these ratios could be consistent where the error regions cross. However at Level 4 on the level of young researchers the drop in women from 2016 to 2018 is not within the error bars of either year. The *leaky pipeline* here is also evident in the consistent and significantly higher ratios of women in Level 4 than all of the senior

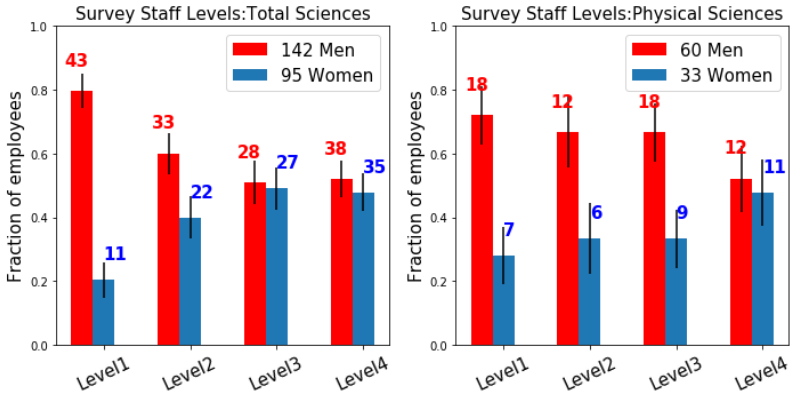


FIGURE 7.4: Results from 2018 survey for the fraction of women at different Career levels for all responses in the Faculty of Sciences and those only in Physical Sciences.

levels even when taking statistical errors into account. This indicates that the perceived drop in women in more senior positions is real and long term since it is outside of the error regions and remains relatively unchanged in the space of two years.

This *leaky-pipeline* is a well known phenomena in Science Technology Engineering & Maths (STEM). In order to try and probe into the origins of such a loss of women in research, especially in a way that can be dealt with at UNIGE, one can examine other variables included in the initial GENERA minimum data set. Since this data was not universally available from HR, it was sought instead by launching a survey across the Physics Section. This would allow any desired variable to be demanded from the surveyed group, although typically surveys suffer from much lower response rates; and a biased data selection because of a bias the in the group of people who choose to respond.

7.2.1 Career Progress Survey

In 2016 a survey was approved to be sent to all the UNIGE staff in the Physics Section to harvest more detailed information pertaining to the careers of male and female researchers. This survey, under the request of the Faculty Commission d'égalité, had to be combined with another targeted at staff who may otherwise have missed an opportunity at promotion and as such a large variety of questions were included. The result was that there were very limited responses, with only a small fraction of total staff responding and many submitting only partially filled responses. For this reason, when presenting the results of the survey the total of number of responses used for that variable are indicated in the Legend. In some cases it is simple to eliminate all those who did not respond to the question. In other cases it is not clear whether someone simply did not reach that question or if they left it blank because they did not have anything to include. For example they may leave

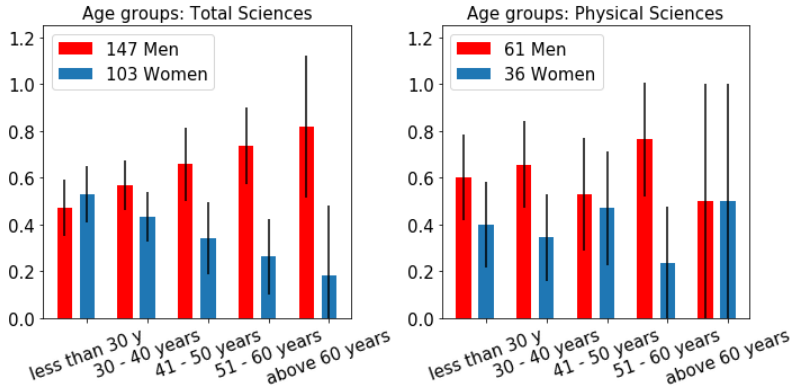


FIGURE 7.5: Results from 2018 survey for the fraction of women in different age groups for all responses in the Faculty of Sciences and those only in Physical Sciences.

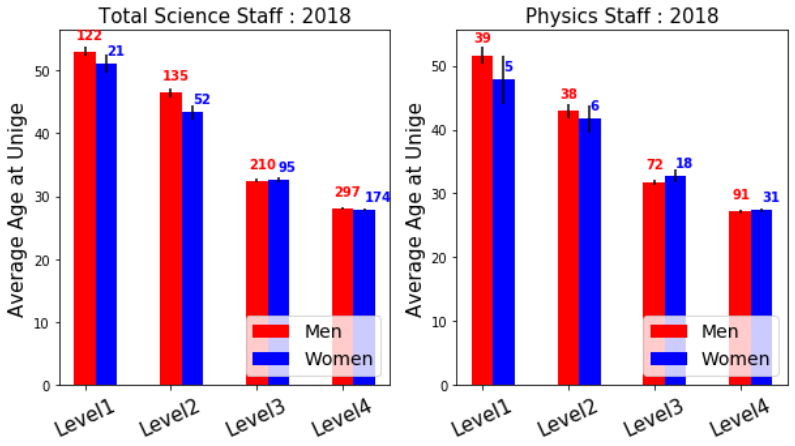


FIGURE 7.6: Results from 2018 HR data for the average age of Men and Women at each Career level.

blank a question on the number of conferences if they have not attended any conferences. When analyzing the HR data the total numbers of men and women are fixed; however the numbers within each category vary depending on the category. As such the reliability of each result can vary depending on how many responses are used. Where there are few responses one can expect the response rate to be more prone to fluctuations. For this reason the numbers of men and women above each bar are shown in the following bar charts as already observed in Fig. 7.2 and Fig. 7.3.

This survey was analyzed and updated to remove some irrelevant questions and improve others before sending again to the entire science faculty in 2018. The results from 2018 are seen as updates to those of 2016 and so are shown in this thesis. As has been discussed, the survey is likely to have a number of weaknesses based on the low response rates and the bias

in the responses, with a selection of people not answering some questions or not filling out the survey at all. People can also make mistakes and fill out their information incorrectly. Although the individual results cannot be verified, it is possible to make a rough estimate of the bias in the results by comparing results for variables which are included in the HR data, assuming the HR data as true. The first plots shown for the 2018 survey compare the age and career level distribution using the same criteria as above. From Fig. 7.4 we can already see that there are slightly higher fractions of women from the survey at different levels compared to those of the HR data. For example, the survey indicates over 20% women at Level 1 and almost 50% women at Level4 compared to 12% at Level 1 and 28% at Level 4 from HR data. This bias is likely due to the fact that more women are motivated to volunteer their time and information to monitor career progress and gender equality at the faculty, although the true reasons for such a bias cannot be confirmed by this result. In Fig. 7.4 the ratios for Physical Sciences are shown alongside the same results using all the responses from the entire Faculty of Sciences. This has been done because within each Faculty Section there were only ~ 100 responses, and for many of the last or more complicated questions there were even fewer. These low statistics are prone to fluctuations and make it difficult to discern any significant deviations in the results. By incorporating all the responses, any noticeable difference become more meaningful. However it should be noted that the behavior of different variables changes from department to department where teaching requirements, typical rates of conferences and awards, among other details, can vary significantly. Fig. 7.4 shows that the leaky pipeline of lower fractions of women at more senior levels still holds in the survey data despite the bias in responses. However one can also see that the significant drop for all sciences begins after Level 3 as opposed to Level 4. This could be the effect of adding Life Sciences where typically there are higher ratios of women in junior research positions.

Demographic Information

The effect of the leaky pipeline is again visible in Fig. 7.5 where the fraction of women steadily declines in higher age brackets for all sciences, although the statistics are too limited confirm this for Physics alone. However, there may also be other factors at work. Fig. 7.6 demonstrates the average age of men and women at each career level for the total science faculty and only physical sciences. From this we can see that for all sciences the average age of Women in Level 2 is slightly younger than that of men.

It is often posited that having a family can pose problems with advancing in a competitive career like academia, particularly for the primary care givers. Although it is difficult to parametrize qualities such as family life, one indicator which can be monitored is the civil

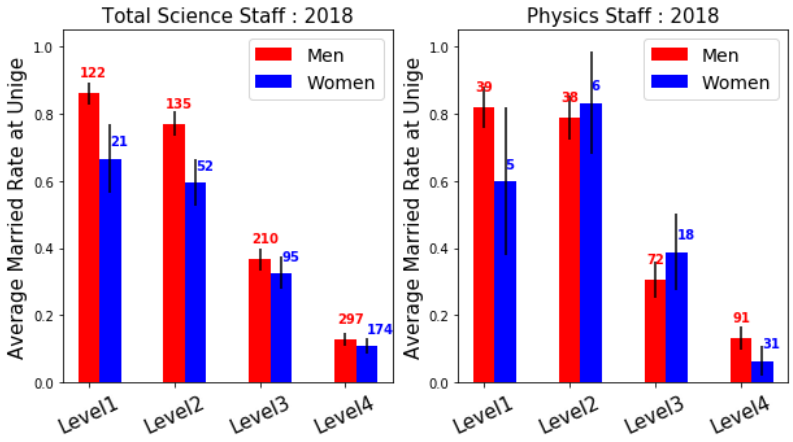


FIGURE 7.7: Results from 2018 HR data for Fraction of Male and Female researcher which are Single, Married or Other in the Faculty of Sciences and those only in Physical Sciences Broken down into separate Career levels.

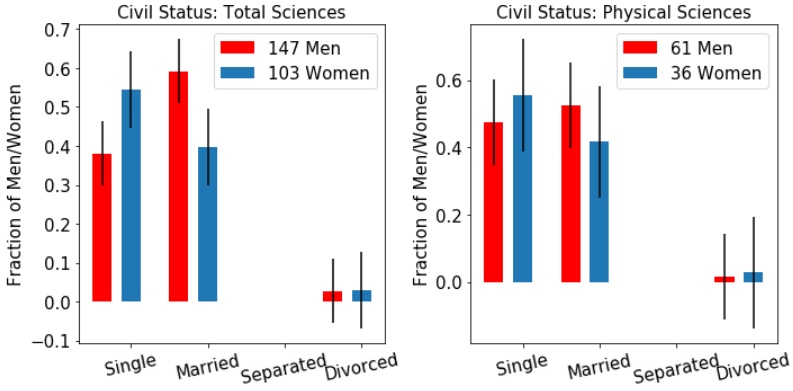


FIGURE 7.8: Results from 2018 survey for Fraction of Male and Female researchers which are Single, Married or Other in the Faculty of Sciences and those only in Physical Sciences.

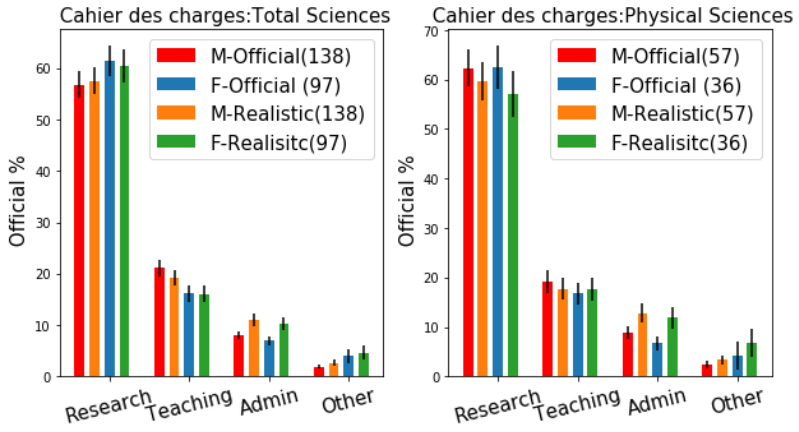


FIGURE 7.9: The average official and the realistic break down of responsibilities for men and women from the 2018 Survey out of the total science faculty and only the Physical Science departments.

status. Researchers who are married are more likely to have a family and or greater home life demands outside of work. This is one of the few details which was available from the HR data where there are more available statistics. Using the survey data seen in Fig. 7.8 for Physics alone, although there is a slightly lower fraction of married researchers amongst the female responses, the large errors make this not significant. The same behavior, however, is observed for the total Science faculty from the survey where the larger statistics make this trend more significant. This result is held up in the HR data and the larger statistics allow for the ratios of married staff to be broken down by career level. Although the error bars are large due to the low numbers of women, the difference in fractions of married staff becomes significant at Level 1 and Level 2 for the entire science faculty. For physical sciences alone there is a lower rate of women married at Level 4 just outside of the error bars. The ratio of women who are married at each Level never seems to significantly surpass that of men. This is a strong indicator that home life could be a factor in the differences in the fraction of women in senior research positions.

Responsibilities

A common hypothesis for the differences between male and female researchers is that women are often burdened with more tasks outside of their research or that they dedicate more time to such tasks in order to ensure they are completed. This was examined by asking respondents to include the official break down of their responsibilities known as their *Cahier des Charges* in their contract, and to compare that to what they actually found themselves doing on a regular basis. The results can be seen in Fig. 7.9 and the differences between

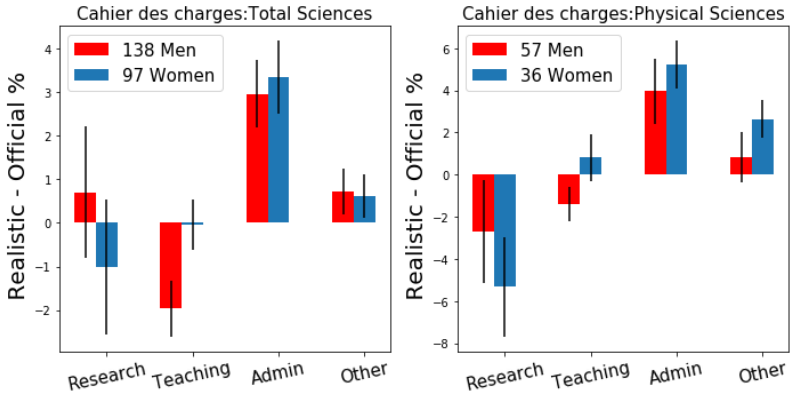


FIGURE 7.10: The average differences between the official and the realistic break down of responsibilities for men and women from the 2018 Survey out of the total science faculty and only the Physical Science departments.

the official and realistic *Cahier des Charges* are shown directly in Fig. 7.10. The trends in behavior between official and realistic responsibilities seems to be similar for both men and women which becomes more clear in Fig. 7.10. One can see that in both all sciences and only physical sciences, everyone considers that they spend more time performing administrative tasks and "Other" tasks than is in their contract. The differences between teaching and research time however can vary depending on the department. Across all departments it appears that men spend slightly more time on research than their official contract, although there is a lot of variation. In physical sciences, though, men tend on average to feel that they spend slightly less time on their research. Both of these differences are within the margins of errors for each other and so not incompatible. Concerning teaching, both in total sciences and only physical sciences, men feel that they teach slightly less than is dictated by their contract. This contrasts with women who feel that they spend slightly more time teaching than their contract. Although the average deviation from the official contract is very slight compared to the large statistical errors for women, there is a significant difference between women and men where women feel that the extra time required for teaching is greater than the extra time the men put in. These statistics are subjective according to who is filling them out which explains the large errors in these statistics; Accordingly, the reasons behind such a trend are difficult to disentangle. Nevertheless, there is an emerging trend for women to perceive a greater positive difference in teaching time compared to their contract than men.

In an attempt to quantify further possible discrepancies in the teaching portion of the contract, the average time spent teaching was also included in the survey. This was broken down into average hours per week, weeks per year, total number of years, as well as the

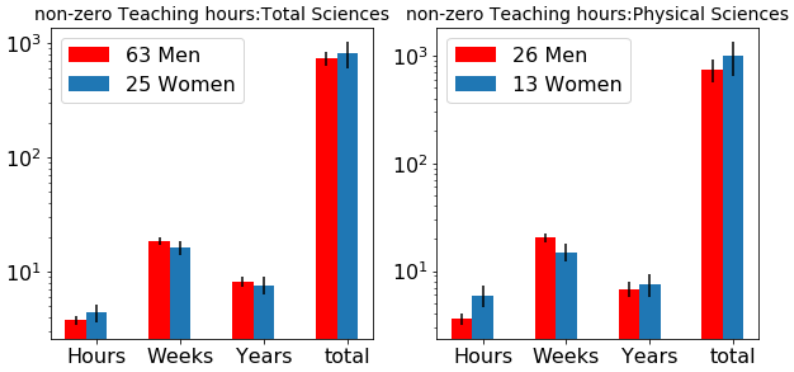


FIGURE 7.11: The average time spent teaching for those who have taught as the Primary Teacher for senior research staff (not including PhD students and Post-docs) also broken down into hours per week, weeks per year, total number of years.

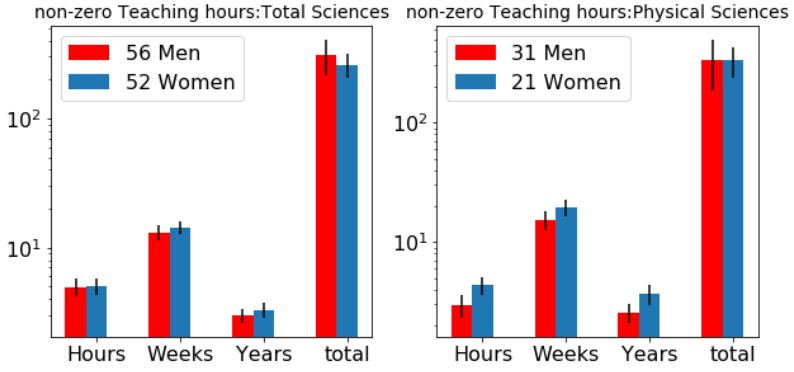


FIGURE 7.12: The average time spent teaching for those who have taught as an assistant for senior research staff (not including PhD students and Post-docs) also broken down into hours per week, weeks per year, total number of years.

Fraction of Responses filling in Teaching Hours

Category	Teaching Role	Total	Male	Female
All Sciences	Principle Teacher	0.68	0.72	0.61
All Sciences	Assistant	0.45	0.39	0.53
Physical Sciences	Principle Teacher	0.72	0.72	0.72
Physical Sciences	Assistant	0.55	0.53	0.6

TABLE 7.1: .

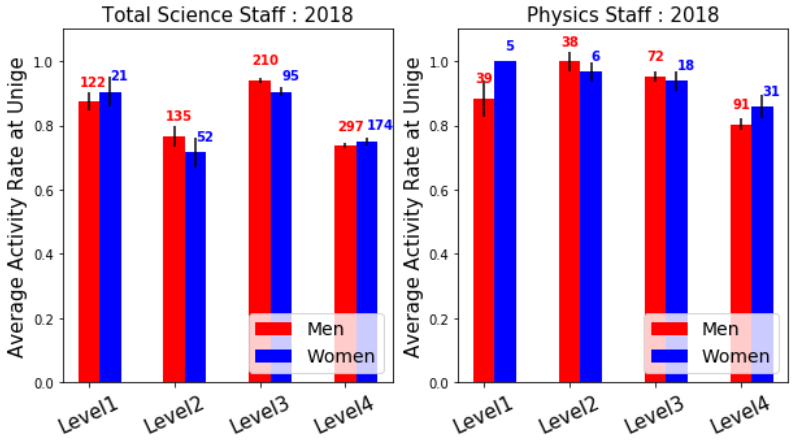


FIGURE 7.13: The average Activity Rate as a fraction of Full-time for Men and Women in the entire faculty of Science and only the Physical Sciences using the full HR 2018 data.

product of all of these, to give the total hours as a principal teacher and as an assistant for each staff member. The average of each of these factors was then plotted in Fig. 7.11 for principal teacher, not counting PhD students or Post-docs and in Fig. 7.12 for assistants. These plots show that women in Physics tend to indicate that they spend more hours per week teaching than men and at the assistant level they teach more years than men. However all other differences are within error bars and also do not hold when applied to all the sciences. For these plots only those who have indicated some teaching hours are included. It cannot be determined if respondents who filled in other parts of the survey did not reach this question or if they did not teach in this role at UNIGE. The fraction of responses to this question compared with the total numbers that filled in the initial demographic data are shown in Table. 7.1. It can be seen that the ratios of people replying to this question varies depending on the gender for the assistant level where men tend to respond less to this question.

Since many arguments around the progression of women in competitive careers often revolve around family responsibilities, it was interesting to consider the activity rate of UNIGE staff. This information was also available from HR and so the full faculty statistics

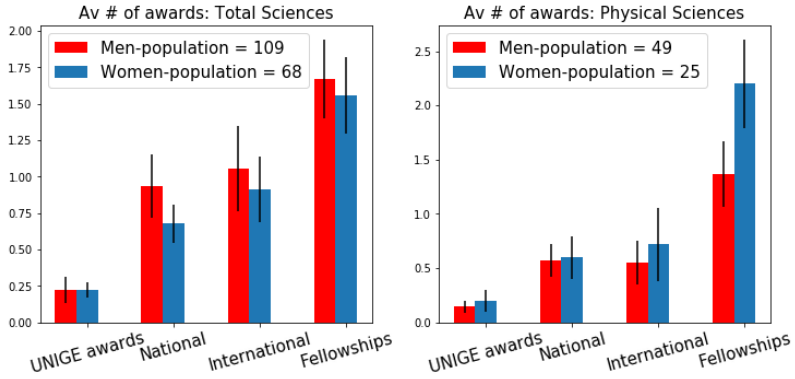


FIGURE 7.14: Results from 2018 survey for the average number of different awards per Male or Female researcher in the Faculty of Sciences and those only in Physical Sciences.

can be inspected. A full-time contract in this case would have an activity rate= 1. Many staff such as PhD students are typically employed with only part-time contracts. Fig. 7.13 shows that the differences between levels are much more apparent than any difference between the genders. At Level 1 there is a slightly higher Activity Rate for female researchers than male researchers. All other gender differences are within error bars and are likely just fluctuations.

Research Output

An example of the differences between the different departments can be seen in Fig. 7.14 which shows the average number of awards per male or female researcher. When examining only Physical sciences one can see a a slightly higher rate of awards per female researcher than male researcher, the difference becoming significant for fellowships; By contrast, for total sciences this trend is reversed with women having fewer awards. These sorts of differences would be sensitive to a certain department having access to a wider range of awards or awards being more common in some departments. It is possible that in some departments there are more awards available only for women thus increasing their possibility of gaining one. On the other hand it is also possible that, due to unconscious bias, in general women must be more qualified in order to be perceived to be at the same level and thus employed in the same roles. The reasons for such a difference cannot be derived from these results but there is evidently some variance in behavior between departments.

An important parameters which can be used in the decision of whom to hire for an academic position can be attendance and presentations at conferences. Showing work at conferences not only provides experience in presenting research but also demonstrates that

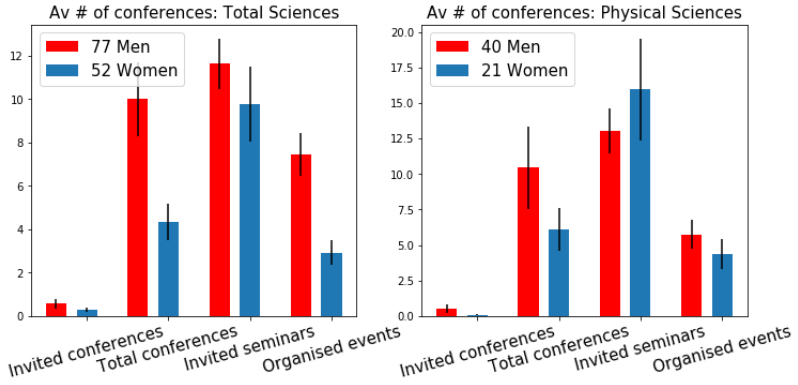


FIGURE 7.15: The average number of conferences per male or female researcher in the last 5 years for people who have participated in one of the following events.

the work conducted is of sufficient caliber that it was accepted by the conference, especially if this was an invited talk. By examining this variable one can try and determine if this has a strong part to play in determining why women may not reach senior academic levels. The survey thus asked for the total number of: invited conferences, total conferences, invited seminars, and other organised events in the last 5 years. The results shown in Fig. 7.15 show that within the physical sciences the rate of invited seminars and organised events are within errors of each other but women seem to attend fewer invited and total conferences just outside of the error bars. This differences in conference attendance is even more apparent when considering the entire Faculty of Science where there is in addition a considerable drop in female organised events. These results were only calculated using people who marked attendance for at least one of the events to exclude those who did not reach this part of the survey or who were not eligible to participate in conferences or events.

7.2.2 Conclusion

The issue of the *leaky-pipeline* and the low numbers of women in physics is a well known phenomena for many years appearing consistently but to varying degrees depending on the country or the institution as seen by the European study *She-figures* [124] and the ETAN report [125]. Although the acknowledgement and participation of women in scientific research has increased, especially in higher education, these changes are very slow to reach the top levels of academia. As such there is a concern that this implies that there are more subtle factors preventing women from fully participating in scientific research and being acknowledged at a high level, which could even impact effectiveness of science to further progress by not harnessing the full potential of the population.

The purpose of GEPs is to address whatever factors may be hindering female participation rates at any level in their careers in science research, particularly at the higher career levels where participation is at its lowest. The very marked difference in participation rates by gender suggests that there are such factors at play. Some factors may be beyond the scope of institutional policy; others may be potential levers for actions by employing institutions or academic bodies such as conference organizers. A better understanding of what factors affect participation rates, and to what degree, is key to formulating more effective GEPs. The work done under this thesis has been aimed at both improving the data and analysis relevant to female participation in scientific research, in particular at UNIGE, and also updating it in time so that, over time, one can build up a picture of how past events affect future outcomes. Limited data is available at a reasonable quality and this has affected the significance of results found. In general, the working hypothesis is that if there is a significant difference between men and women then that is a potential cause for differences in participation rates, and thus a potential lever for action. Some data was obtained from HR, other data from survey(s). The table below summarises which factors were examined and the headline outcomes:

Data Source	Variable	Differences outside of error bars
HR	Career Levels	Fewer women in Physical sciences at all levels
HR	Age	25% of physical sciences staff below the age of 35 are women goes to 12% above the age of 40. Average age of women is younger at Level 2 for all sciences
HR	Civil Status	Smaller fraction of the women at Level 1 and Level 2 are married compared to men in all sciences. In physical sciences there are also fewer women married at Level 4.
HR	Activity Rate	For all sciences the average activity rate for women at Level 4 is higher than for men.
Survey	Career Levels	Fewer women than men in Level 1 and 2 (and 3) for all sciences (physical sciences)
Survey	Age	Fewer women than men above the age of 45 for all sciences at for 51-60 years in physical sciences.
Survey	Civil Status	Smaller fraction of women are married than men in all sciences.
Survey	Awards	Higher average rate of fellowships among women than men in physical sciences
Survey	Conferences	In both all sciences and physical sciences alone women have a lower average number of conferences in the last 5 years. In all sciences the average number of organised events is lower for women.
Survey	Cahier des Charges	The extra time women perceive they dedicate to teaching is greater than the extra time men perceive that they make for teaching.
Survey	Teaching	In physical sciences the average teaching hours per week is greater for women but the average weeks of this teaching per year is greater for men.

Although these results are still unclear regarding how to improve the situation for women in physics, the aim of this work was to provide a current status for UNIGE and monitor how this changes with time. This is an important task as it can allow the impact of changing policy to be monitored and evaluated independently. For example, further study can go into why the fraction of women at Level 4 showed a drop from 2016 to 2018. These long term studies, especially with access to the full faculty statistics from HR, could open up a

new understanding of what causes these vast major discrepancies in careers in physics and could be used to update policies to create a fairer research environment. The work done under this thesis should thus be seen as part of a larger picture which needs to be built up over time but which offers the prospect of formulating significantly more effective GEPs in the future.

Chapter 8

Conclusion

For over a hundred years humans have been observing high energy particles arriving from outside the earth's atmosphere [11] and for even longer we have been recording observations of light produced by distant stars and galaxies. Since then new fundamental particles and messengers have been discovered and with them a whole new way of creating a map of the universe. The IceCube detector targets high energy neutrinos as astrophysical messengers due to their undeviated path through space even at high energies and their intrinsic link to hadronic interactions of cosmic rays. Determining high energy astrophysical neutrino sources would not only be a new messenger to observe objects but would also be a smoking gun to hadronic interactions at CR accelerators.

The first evidence for a high energy astrophysical source was presented in Aartsen et al. [5] in the form of the flaring blazar TXS 0506+056. This thesis performed all-sky searches for any point-like steady-state neutrino source with ten years of IceCube data. The most recent equivalent searches include: the last all-sky search with 7 years of IceCube data [6], and the northern sky search optimised to target the neutrino sources responsible for the $E^{-2.19}$ best fit astrophysical neutrino flux [4] from a selection of through going muon tracks in the northern hemisphere [82]. Both of these searches found no new sources of astrophysical neutrinos. The additional years of data combined with an improved data selection for the final 6 years demonstrate a 35% improvement in sensitivity compared to the 7 year general all-sky search. This ten year analysis also shows a comparable sensitivity to the Northern sky analysis for a hard E^{-2} spectrum but a roughly 30% average improved sensitivity given a softer E^{-3} spectrum.

This analysis included an updated event selection based on the latest samples in Aartsen et al. [6] and Aartsen et al. [4] with an improved angular resolution and a new multi-class multivariate cut in the northern hemisphere. This selection was applied to the final six years of data, taking the sample to ten years of through-going muon tracks starting at TeV energies. The selection was used to conduct searches for point-like steady-state astrophysical neutrino sources using a maximum likelihood method and assuming a background

dominated event selection. Different searches were optimised to be sensitive to different neutrino source signals. Firstly, the all-sky search allows the detection of any point-like steady-state neutrino source given no prior information on the source spectrum or position. Since the all-sky method is also sensitive to every background fluctuation in the sky, there is a large penalty factor in the significance requiring a very bright source to overcome it. To compensate for this, prior information on possible source locations is implemented in the form of a source catalog of motivated source directions using gamma ray observations and the spectrum and likelihood is evaluated at each individual source. This method requires only a limited number of sources to be tested to avoid another large trial factor. This method, along with the all-sky scan is also limited by the sensitivity at the individual source location. This limitation is particularly relevant in the southern hemisphere where there are many Galactic objects. To combat this problem the signal from a catalog of objects is stacked to allow a smaller sensitivity flux per source.

The result from the all-sky scan was a 9.9% and 75% post-trial p-value for the most significant hotspot in the northern and southern hemisphere respectively. Both these hotspots are consistent with background alone. However, the northern hotspot is 0.35° from the galaxy NGC 1068 in the northern source list; this leads to an pre-trial p-value of 4.13σ in the direction of NGC 1068 which becomes 2.9σ when compared to the most significant excesses in the source list when using background trials. Simulation tests demonstrate that is it possible to observe a hotspot with such a shift when injecting a soft signal spectrum.

Finally, in the absence of a single discovery-level excess from one of the sources in the source list, a population search was conducted for an excess in the rate of hotspots in the unblinded catalog. This search can be sensitive to the presence of any signal in the catalog but does not identify which or how many of said objects may produce astrophysical signal. Using the 10 year analysis, an excess was found for the 4 most significant sources resulting in a 3.3σ post-trial inconsistency with a background only hypothesis to find at least 4 sources of the significance great than or equal to the 4^{th} source. These sources include NGC 1068, TXS 0506+056, PKS 1424+240, and GB6 J1542+6129. Evidence has already been presented for TXS 0506+056 to be a flaring neutrino source [5]. Redoing this search excluding TXS 0506+056 from the catalog instead results in 2.25σ post-trial p value for the remaining top 3 most significant sources.

NGC 1068 is a very nearby Seyfert galaxy with an AGN, as such there have already been multiple studies in its direction using photons. The best-fit neutrino spectrum from this analysis finds a much larger flux normalisation at 1 TeV than predicted by current models or than current gamma ray observations at lower energies. With no known similar neutrino source with which to compare these observations it is difficult to determine how feasible such a difference could be. The results in the direction of NGC 1068 and from the

catalog as a whole demonstrate a strong motivation to follow up these sources using future updates to the analysis.

The final chapter of this thesis treated gender studies conducted at UNIGE on behalf of the GENERA consortium. This work was to raise awareness and provide local information supporting the implementation of a university Gender Equality Policy (GEP). Much of this work went into monitoring and establishing the status of gender equality in the Physics section at UNIGE currently. The typical *leaky pipeline*, where the ratio of women decreases with more senior positions in STEM was observed using data from both a survey and full statistics from human resources. There were also slight differences in other variables such as marital status or number of awards, which were less significant. It could prove vital to continue to monitor these variables in order to gauge any changes in this status with time, in particular after any considerable policy changes. The aim would be to support changes that result in a fair selection and diverse academic staff at the university in the Physics section.

Appendix A

Appendix

Catalog	Associated Names	R.A (◦)	Declination (◦)	σ_{src} (◦)	Age (yrs)
Milagro Six	MGRO J1852+01	283.12	0.51	0.0	-
	MGRO J1908+06	286.68	6.03	1.3	-
	MGRO J2019+37	304.68	36.70	0.64	-
	MGRO J2032+37	307.75	36.52	0.0	-
	MGRO J2031+41	307.93	40.67	1.5	-
	MGRO J2043+36	310.98	36.3	1.0	-
HAWC	HWC J1825−133	276.3	-13.3	0.5	-
	HWC J1836−090c	278.9	-9.0	0.5	-
	HWC J1836−074c	279.1	-7.4	0.5	-
	HWC J1838−060	279.6	6.0	0.5	-
	HWC J1842−046c	280.5	-4.6	0.5	-
	HWC J1844−031c	281.0	-3.1	0.5	-
	HWC J1849−017c	282.3	-1.7	0.5	-
	HWC J1857+023	284.3	2.3	0.5	-
	HWC J1904+080c	286.1	4.44	0.5	-
	HWC J1907+062c	286.8	6.2	0.5	-
SNR with mol. cloud	Tycho	6.33	64.15	0	443
	IC443	94.3	22.6	0.16	3000
	SN 1006 SW	225.7	-41.9	1.06	1009
	HESS J1708−410	258.4	-39.8	1.36	1000
	HESS J1718−385	259.5	-37.4	0.15	1800
	Galactic Centre Ridge	266.4	-29.0	0.2	1200
	HESS J1813−178	274.5	-15.5	0.77	2500
	HESS J1843−033	281.6	-3.0	0	900
	SNR G054.1+00.3	292.6	18.9	0	2500
	Cassiopeia A	350.9	58.8	0	316
SNR with PWN	Crab	83.6	22.01	0	961
	RX J0852.0−4622	133.0	-46.3	0.7	2400
	MSH 15−52	228.6	-59.1	0.11	1900
	HESS J1634−472	249.0	-47.3	0.63	1500
	HESS J1640−465	250.3	-46.6	0.87	1000
	SNR G000.9+00.1	266.8	-28.2	0	1900
	HESS J1808−204	272.9	-19.4	0.14	960
	HESS J1809−193	273.4	-17.8	0.92	1200
	HESS J1825−137	278.4	-10.6	1.63	720
SNR alone	RCW 86	220.8	-62.5	0.98	2000
	HESS J1641−463	250.3	-46.3	0.62	1000
	RX J1713.7−3946	258.5	-38.2	0.65	350
	HESS J1858+020	284.5	2.2	0.08	2300

TABLE A.1: Details of Galactic sources used in the 7 year stacking analysis

Bibliography

[1] Hamish Johnston. Cosmic neutrinos named Physics World 2013 Breakthrough of the Year. https://physicsworld.com/a/cosmic-neutrinos-named-physics-world-2013-breakthrough-of-the-year/.

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

[3] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016. doi: 10.1103/PhysRevLett.116.061102.

[4] M. G. Aartsen et al. The IceCube Neutrino Observatory - Contributions to ICRC 2017 Part II: Properties of the Atmospheric and Astrophysical Neutrino Flux. 2017.

[5] M. G. Aartsen et al. Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert. *Science*, 361(6398):147–151, 2018. doi: 10.1126/science.aat2890.

[6] M. G. Aartsen et al. All-sky Search for Time-integrated Neutrino Emission from Astrophysical Sources with 7 yr of IceCube Data. *Astrophys. J.*, 835(2):151, 2017. doi: 10.3847/1538-4357/835/2/151.

[7] GENERA. https://genera-project.com/index.php.

[8] M. G. Aartsen et al. Constraints on Galactic Neutrino Emission with Seven Years of IceCube Data. *Astrophys. J.*, 849(1):67, 2017. doi: 10.3847/1538-4357/aa8dfb.

[9] *A Summary of Recent Updates in the Search for Cosmic Ray Sources using the IceCube Detector*. SISSA, Proceedings, 2017 European Physical Society Conference on High Energy Physics (EPS-HEP 2017). doi: https://doi.org/10.22323/1.314.0004.

[10] M. G. Aartsen et al. Observation and Characterization of a Cosmic Muon Neutrino Flux from the Northern Hemisphere using six years of IceCube data. *Astrophys. J.*, 833(1):3, 2016. doi: 10.3847/0004-637X/833/1/3.

[11] Victor F. Hess. Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten. *Phys. Z.*, 13:1084–1091, 1912.

[12] D Skobelzyn. Über eine neue art sehr schneller -strahlen. *Zeitschrift für Physik A Hadrons and Nuclei*, 54:686–702, 09 1929. doi: 10.1007/BF01341600.

[13] Pierre Auger, P. Ehrenfest, R. Maze, J. Daudin, and A. Freon Robley. Extensive cosmic ray showers. *Rev. Mod. Phys.*, 11:288–291, 1939. doi: 10.1103/RevModPhys.11.288.

[14] Alexander Aab et al. Combined fit of spectrum and composition data as measured by the Pierre Auger Observatory. *JCAP*, 1704(04):038, 2017. doi: 10.1088/1475-7516/2018/03/E02,10.1088/1475-7516/2017/04/038. [Erratum: JCAP1803,no.03,E02(2018)].

[15] Kenneth Greisen. End to the cosmic ray spectrum? *Phys. Rev. Lett.*, 16:748–750, 1966. doi: 10.1103/PhysRevLett.16.748.

[16] G. T. Zatsepin and V. A. Kuzmin. Upper limit of the spectrum of cosmic rays. *JETP Lett.*, 4:78–80, 1966. [Pisma Zh. Eksp. Teor. Fiz.4,114(1966)].

[17] Ines Valino. The flux of ultra-high energy cosmic rays after ten years of operation of the Pierre Auger Observatory. *PoS*, ICRC2015:271, 2016. doi: 10.22323/1.236.0271.

[18] T. Abu-Zayyad et al. The Cosmic Ray Energy Spectrum Observed with the Surface Detector of the Telescope Array Experiment. *Astrophys. J.*, 768:L1, 2013. doi: 10.1088/2041-8205/768/1/L1.

[19] R. Aloisio, V. Berezinsky, and A. Gazizov. Ultra High Energy Cosmic Rays: The disappointing model. *Astropart. Phys.*, 34:620–626, 2011. doi: 10.1016/j.astropartphys.2010.12.008.

[20] Alexander Aab et al. Depth of maximum of air-shower profiles at the Pierre Auger Observatory. I. Measurements at energies above $10^{17.8}$ eV. *Phys. Rev.*, D90(12):122005, 2014. doi: 10.1103/PhysRevD.90.122005.

[21] R. U. Abbasi et al. Study of Ultra-High Energy Cosmic Ray composition using Telescope Array's Middle Drum detector and surface array in hybrid mode. *Astropart. Phys.*, 64:49–62, 2015. doi: 10.1016/j.astropartphys.2014.11.004.

[22] M. G. Aartsen et al. Constraints on Ultrahigh-Energy Cosmic-Ray Sources from a Search for Neutrinos above 10 PeV with IceCube. *Phys. Rev. Lett.*, 117(24):241101,

2016. doi: 10.1103/PhysRevLett.117.241101,10.1103/PhysRevLett.119.259902. [Erratum: Phys. Rev. Lett.119,no.25,259902(2017)].

[23] Enrico Fermi. On the Origin of the Cosmic Radiation. *Phys. Rev.*, 75:1169–1174, 1949. doi: 10.1103/PhysRev.75.1169.

[24] A. R. Bell. The Acceleration of cosmic rays in shock fronts. I. *Mon. Not. Roy. Astron. Soc.*, 182:147–156, 1978.

[25] M. G. Aartsen et al. Search for correlations between the arrival directions of IceCube neutrino events and ultrahigh-energy cosmic rays detected by the Pierre Auger Observatory and the Telescope Array. *JCAP*, 1601(01):037, 2016. doi: 10.1088/1475-7516/2016/01/037.

[26] A. Aab et al. Large-scale cosmic-ray anisotropies above 4 EeV measured by the Pierre Auger Observatory. *Astrophys. J.*, 868(1):4, 2018. doi: 10.3847/1538-4357/aae689.

[27] Qiang Yuan, Peng-Fei Yin, Xue-Feng Wu, Xiao-Jun Bi, Siming Liu, and Bing Zhang. A statistical model for the γ-ray variability of the crab nebula. *The Astrophysical Journal Letters*, 730(2):L15, 2011. URL http://stacks.iop.org/2041-8205/730/i=2/a=L15.

[28] Pisin Chen and K. D. Hoffman. Origin and evolution of cosmic accelerators - the unique discovery potential of an UHE neutrino telescope: Astronomy Decadal Survey (2010-2020) Science White Paper. *arXiv e-prints*, art. arXiv:0902.3288, February 2009.

[29] M. Ackermann et al. Detection of the Characteristic Pion-Decay Signature in Supernova Remnants. *Science*, 339:807, 2013. doi: 10.1126/science.1231160.

[30] Werner Rodejohann. Neutrino Mixing and Neutrino Telescopes. *JCAP*, 0701:029, 2007. doi: 10.1088/1475-7516/2007/01/029.

[31] Zhi-Zhong Xing. Flavor distribution of UHE cosmic neutrino oscillations at neutrino telescopes. *Nucl. Instrum. Meth.*, A602:58–62, 2009. doi: 10.1016/j.nima.2008.12.018.

[32] K. Abe et al. Combined Analysis of Neutrino and Antineutrino Oscillations at T2K. *Phys. Rev. Lett.*, 118(15):151801, 2017. doi: 10.1103/PhysRevLett.118.151801.

[33] Roger Wendell. Atmospheric Results from Super-Kamiokande. *AIP Conf. Proc.*, 1666:100001, 2015. doi: 10.1063/1.4915569.

[34] M. G. Aartsen et al. Measurement of Atmospheric Neutrino Oscillations at 6–56 GeV with IceCube DeepCore. *Phys. Rev. Lett.*, 120(7):071801, 2018. doi: 10.1103/PhysRevLett.120.071801.

[35] HAWC. *HAWC- Cosmic rays*, 2018. URL https://www.hawc-observatory.org/science/cosmicrays.php.

[36] M. Ackermann et al. FERMI-LAT OBSERVATIONS OF THE DIFFUSE γ-RAY EMISSION: IMPLICATIONS FOR COSMIC RAYS AND THE INTERSTELLAR MEDIUM. *The Astrophysical Journal*, 750(1):3, apr 2012. doi: 10.1088/0004-637x/750/1/3. URL https://doi.org/10.1088%2F0004-637x%2F750%2F1%2F3.

[37] M. G. Aartsen et al. A combined maximum-likelihood analysis of the high-energy astrophysical neutrino flux measured with IceCube. *Astrophys. J.*, 809(1):98, 2015. doi: 10.1088/0004-637X/809/1/98.

[38] T. C. Weekes et al. Observation of TeV gamma rays from the Crab nebula using the atmospheric Cerenkov imaging technique. *Astrophys. J.*, 342:379–395, 1989. doi: 10.1086/167599.

[39] H. Abdalla et al. The population of TeV pulsar wind nebulae in the H.E.S.S. Galactic Plane Survey. *Astron. Astrophys.*, 612:A2, 2018. doi: 10.1051/0004-6361/201629377.

[40] TeVCat: online catalogue of TeV sources. http://tevcat.uchicago.edu/.

[41] J. A. Hinton. The Status of the H.E.S.S. project. *New Astron. Rev.*, 48:331–337, 2004. doi: 10.1016/j.newar.2003.12.004.

[42] Roger Blandford, David Meier, and Anthony Readhead. Relativistic Jets in Active Galactic Nuclei. 2018.

[43] Robert Antonucci. Unified models for active galactic nuclei and quasars. *Ann. Rev. Astron. Astrophys.*, 31:473–521, 1993. doi: 10.1146/annurev.aa.31.090193.002353.

[44] C. Megan Urry and Paolo Padovani. Unified schemes for radio-loud active galactic nuclei. *Publ. Astron. Soc. Pac.*, 107:803, 1995. doi: 10.1086/133630.

[45] K. P. Singh. An X-ray view of quasars. *Bull. Astron. Soc. India*, 41:137, 2013.

[46] M. G. Aartsen et al. Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A. *Science*, 361(6398):eaat1378, 2018. doi: 10.1126/science.aat1378.

[47] Julia K. Becker, Peter L. Biermann, Jens Dreyer, and Tanja M. Kneiske. Cosmic Rays VI - Starburst galaxies at multiwavelengths. *arXiv e-prints*, art. arXiv:0901.1775, January 2009.

[48] Abraham Loeb and Eli Waxman. The Cumulative background of high energy neutrinos from starburst galaxies. *JCAP*, 0605:003, 2006. doi: 10.1088/1475-7516/2006/05/003.

[49] E. Costa et al. Discovery of an X-ray afterglow associated with the gamma-ray burst of 28 February 1997. *Nature*, 387:783–785, 1997. doi: 10.1038/42885.

[50] B. P. Abbott et al. Multi-messenger Observations of a Binary Neutron Star Merger. *Astrophys. J.*, 848(2):L12, 2017. doi: 10.3847/2041-8213/aa91c9.

[51] Shigeo S. Kimura, Kohta Murase, Peter Mészáros, and Kenta Kiuchi. High-Energy Neutrino Emission from Short Gamma-Ray Bursts: Prospects for Coincident Detection with Gravitational Waves. *Astrophys. J.*, 848(1):L4, 2017. doi: 10.3847/2041-8213/aa8d14.

[52] R. Mirzoyan. First time detection of a GRB at sub-TeV energies; MAGIC detects the GRB 190114C. *The Astronomer's Telegram*, 12390, January 2019.

[53] I. A. Belolaptikov et al. The Baikal underwater neutrino telescope: Design, performance and first results. *Astropart. Phys.*, 7:263–282, 1997. doi: 10.1016/S0927-6505(97)00022-4.

[54] E. Andres et al. The AMANDA neutrino telescope: Principle of operation and first results. *Astropart. Phys.*, 13:1–20, 2000. doi: 10.1016/S0927-6505(99)00092-4.

[55] M. Ageron et al. ANTARES: the first undersea neutrino telescope. *Nucl. Instrum. Meth.*, A656:11–38, 2011. doi: 10.1016/j.nima.2011.06.103.

[56] Sheldon L. Glashow. Resonant Scattering of Antineutrinos. *Phys. Rev.*, 118:316–317, 1960. doi: 10.1103/PhysRev.118.316.

[57] Raj Gandhi, Chris Quigg, Mary Hall Reno, and Ina Sarcevic. Ultrahigh-energy neutrino interactions. *Astropart. Phys.*, 5:81–110, 1996. doi: 10.1016/0927-6505(96)00008-4.

[58] Edgar Bugaev, Teresa Montaruli, Yuri Shlepin, and Igor A. Sokalski. Propagation of tau neutrinos and tau leptons through the earth and their detection in underwater / ice neutrino telescopes. *Astropart. Phys.*, 21:491–509, 2004. doi: 10.1016/j.astropartphys.2004.03.002.

[59] P. A. Cerenkov. Visible radiation produced by electrons moving in a medium with velocities exceeding that of light. *Phys. Rev.*, 52:378–379, 1937. doi: 10.1103/PhysRev.52.378.

[60] I. M. Frank and I. E. Tamm. Coherent visible radiation of fast electrons passing through matter. *Compt. Rend. Acad. Sci. URSS*, 14(3):109–114, 1937. doi: 10.1007/978-3-642-74626-0_2,10.3367/UFNr.0093.196710o.0388. [Usp. Fiz. Nauk93,no.2,388(1967)].

[61] M. G. Aartsen et al. The IceCube Neutrino Observatory: Instrumentation and Online Systems. *JINST*, 12(03):P03012, 2017. doi: 10.1088/1748-0221/12/03/P03012.

[62] M. G. Aartsen et al. Measurement of South Pole ice transparency with the IceCube LED calibration system. *Nucl. Instrum. Meth.*, A711:73–89, 2013. doi: 10.1016/j.nima.2013.01.054.

[63] G. Mie. Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen. *Annalen der Physik*, 330:377–445, 1908. doi: 10.1002/andp.19083300302.

[64] M. Ackermann et al. Optical properties of deep glacial ice at the South Pole. *J. Geophys. Res. Atmos.*, 111(D13):D13203, 2006. doi: 10.1029/2005JD006687.

[65] M. G. Aartsen et al. The IceCube Neutrino Observatory Part VI: Ice Properties, Reconstruction and Future Developments. In *Proceedings, 33rd International Cosmic Ray Conference (ICRC2013): Rio de Janeiro, Brazil, July 2-9, 2013*, 2013.

[66] P. B. Price. Kinetics of conversion of air bubbles to air hydrate crystals in Antarctic ice. *Science*, 267:1802, 1995. doi: 10.1126/science.267.5205.1802.

[67] Icecube. South Pole glacial climate reconstruction from multi-borehole laser particulate stratigraphy. *J. Glaciol.*, 59(218):1117–1128, 2013. doi: 10.3189/2013JoG13J068.

[68] Martin Rongen. Measuring the optical properties of IceCube drill holes. *EPJ Web Conf.*, 116:06011, 2016. doi: 10.1051/epjconf/201611606011.

[69] M. G. Aartsen et al. Energy Reconstruction Methods in the IceCube Neutrino Telescope. *JINST*, 9:P03009, 2014. doi: 10.1088/1748-0221/9/03/P03009.

[70] John G. Learned and Karl Mannheim. High-energy neutrino astrophysics. *Annual Review of Nuclear and Particle Science*, 50(1):679–749, 2000. doi: 10.1146/annurev.nucl.50.1.679. URL https://doi.org/10.1146/annurev.nucl.50.1.679.

[71] R. Abbasi et al. Calibration and Characterization of the IceCube Photomultiplier Tube. *Nucl. Instrum. Meth.*, A618:139–152, 2010. doi: 10.1016/j.nima.2010.03.102.

[72] J. Ahrens et al. Muon track reconstruction and data selection techniques in AMANDA. *Nucl. Instrum. Meth.*, A524:169–194, 2004. doi: 10.1016/j.nima.2004.01.065.

[73] M. G. Aartsen et al. Improvement in Fast Particle Track Reconstruction with Robust Statistics. *Nucl. Instrum. Meth.*, A736:143–149, 2014. doi: 10.1016/j.nima.2013.10.074.

[74] N. van Eijndhoven, O. Fadiran, and G. Japaridze. Implementation of a Gauss convoluted Pandel PDF for track reconstruction in Neutrino Telescopes. *Astropart. Phys.*, 28:456–462, 2007. doi: 10.1016/j.astropartphys.2007.09.001.

[75] Nathan Whitehorn, Jakob van Santen, and Sven Lafebre. Penalized Splines for Smooth Representation of High-dimensional Monte Carlo Datasets. *Comput. Phys. Commun.*, 184:2214–2220, 2013. doi: 10.1016/j.cpc.2013.04.008.

[76] Till Neunhoffer. Estimating the angular resolution of tracks in neutrino telescopes based on a likelihood analysis. *Astropart. Phys.*, 25:220–225, 2006. doi: 10.1016/j.astropartphys.2006.01.002.

[77] Askhat Gazizov and Marek P. Kowalski. ANIS: High energy neutrino generator for neutrino telescopes. *Comput. Phys. Commun.*, 172:203–213, 2005. doi: 10.1016/j.cpc.2005.03.113.

[78] A. M. Dziewonski and D. L. Anderson. Preliminary reference earth model. *Phys. Earth Planet. Interiors*, 25:297–356, 1981. doi: 10.1016/0031-9201(81)90046-7.

[79] Dmitry Chirkin and Wolfgang Rhode. Muon Monte Carlo: A High-precision tool for muon propagation through matter. 2004.

[80] Claudio Kopper. An OpenCL-based photon-tracking simulation using a (source-based) ray tracing algorithm modeling scattering and absorption of light in the deep glacial ice at the South Pole or Mediterranean sea water. https://github.com/claudiok/clsim, 2011.

[81] Dmitry Chirkin. Photon tracking with GPUs in IceCube. *Nucl. Instrum. Meth.*, A725:141–143, 2013. doi: 10.1016/j.nima.2012.11.170.

[82] M. G. Aartsen et al. Search for steady point-like sources in the astrophysical muon neutrino flux with 8 years of IceCube data. 2018.

[83] J. Ahrens et al. Observation of high-energy atmospheric neutrinos with the Antarctic Muon and Neutrino Detector Array. *Phys. Rev.*, D66:012005, 2002. doi: 10.1103/PhysRevD.66.012005.

[84] Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Edouard Duchesnay, and Gilles Louppe. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12, 01 2012.

[85] R. Abbasi et al. Time-Integrated Searches for Point-like Sources of Neutrinos with the 40-String IceCube Detector. *Astrophys. J.*, 732:18, 2011. doi: 10.1088/0004-637X/732/1/18.

[86] M. G. Aartsen et al. Search for Time-independent Neutrino Emission from Astrophysical Sources with 3 yr of IceCube Data. *Astrophys. J.*, 779:132, 2013. doi: 10.1088/0004-637X/779/2/132.

[87] K Schatto. *Stacked searches for high-energy neutrinos from blazars with IceCube*. PhD thesis, Johannes Gutenber Universit at Mainz, 2014.

[88] J Feintzeig. *Searches for Point-like Sources of Astrophysical Neutrinos with the IceCube Neutrino Observatory*. PhD thesis, University of Wisconsin, Madison, 2014.

[89] Jim Braun, Jon Dumm, Francesco De Palma, Chad Finley, Albrecht Karle, and Teresa Montaruli. Methods for point source analysis in high energy neutrino telescopes. *Astropart. Phys.*, 29:299–305, 2008. doi: 10.1016/j.astropartphys.2008.02.007.

[90] Abdo Abdo, B T. Allen, T Aune, D Berley, Emanuele Bonamente, Grant Christopher, T DeYoung, Brenda Dingus, R W. Ellsworth, Jessica Galbraith-Frew, M M. Gonzalez, Jordan Goodman, C M. Hoffman, P H. Huentemeyer, Brian Kolterman, James Linnemann, J E. McEnery, A I. Mincer, T Morgan, and Gaurang Yodh. Spectrum and morphology of the two brightest milagro sources in the cygnus region: Mgro j2019+37 and mgro j2031+41. *The Astrophysical Journal*, 753, 02 2012. doi: 10.1088/0004-637X/753/2/159.

[91] E. Aliu et al. Spatially Resolving the Very High Energy emission from MGRO J2019+37 with VERITAS. *Astrophys. J.*, 788:78, 2014. doi: 10.1088/0004-637X/788/1/78.

[92] U. F. Katz and Ch. Spiering. High-Energy Neutrino Astrophysics: Status and Perspectives. *Prog. Part. Nucl. Phys.*, 67:651–704, 2012. doi: 10.1016/j.ppnp.2011.12.001.

[93] IceCube. Software Documentation: Paraboloid. http://software.icecube.wisc.edu/documentation/projects/paraboloid/index.html.

[94] José Luis Morales and Jorge Nocedal. Remark on “algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound constrained optimization”. *ACM Trans. Math. Softw.*, 38(1):7:1–7:4, December 2011. ISSN 0098-3500. doi: 10.1145/2049662.2049669. URL http://doi.acm.org/10.1145/2049662.2049669.

[95] S. S. Wilks. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Statist.*, 9(1):60–62, 03 1938. doi: 10.1214/aoms/1177732360. URL https://doi.org/10.1214/aoms/1177732360.

[96] N. Smirnov. Table for estimating the goodness of fit of empirical distributions. *Ann. Math. Statist.*, 19(2):279–281, 06 1948. doi: 10.1214/aoms/1177730256. URL https://doi.org/10.1214/aoms/1177730256.

[97] B. Hona, A. Robare, H. Fleischhack, and P. Huentemeyer. Correlated GeV–TeV Gamma-Ray Emission from Extended Sources in the Cygnus Region. *PoS*, ICRC2017:710, 2018. doi: 10.22323/1.301.0710. [35,710(2017)].

[98] S. P. Wakely and D. Horan. TeVCat: An online catalog for Very High Energy Gamma-Ray Astronomy. *International Cosmic Ray Conference*, 3:1341–1344, 2008.

[99] F. Aharonian et al. Observations of the Crab Nebula with H.E.S.S. *Astron. Astrophys.*, 457:899–915, 2006. doi: 10.1051/0004-6361:20065351.

[100] F. Aharonian. HESS VHE Gamma-Ray Sources Without Identified Counterparts. *Astron. Astrophys.*, 477:353–363, 2008. doi: 10.1051/0004-6361:20078516.

[101] H. Abdalla et al. The H.E.S.S. Galactic plane survey. *Astron. Astrophys.*, 612:A1, 2018. doi: 10.1051/0004-6361/201732098.

[102] A. U. Abeysekara et al. The 2HWC HAWC Observatory Gamma Ray Catalog. *Astrophys. J.*, 843(1):40, 2017. doi: 10.3847/1538-4357/aa7556.

[103] A. U. Abeysekara et al. A Very High Energy γ-Ray Survey towards the Cygnus Region of the Galaxy. *Astrophys. J.*, 861(2):134, 2018. doi: 10.3847/1538-4357/aac4a2.

[104] F. Acero et al. Fermi Large Area Telescope Third Source Catalog. *Astrophys. J. Suppl.*, 218(2):23, 2015. doi: 10.1088/0067-0049/218/2/23.

[105] Julia K. Becker, Peter L. Biermann, Jens Dreyer, and Tanja M. Kneiske. Cosmic Rays VI - Starburst galaxies at multiwavelengths. 2009.

[106] M. G. Aartsen et al. Searches for Extended and Point-like Neutrino Sources with Four Years of IceCube Data. *Astrophys. J.*, 796(2):109, 2014. doi: 10.1088/0004-637X/796/2/109.

[107] A. U. Abeysekara et al. The HAWC Gamma-Ray Observatory: Design, Calibration, and Operation. 2013.

[108] A. U. Abeysekara et al. Search for TeV Gamma-Ray Emission from Point-like Sources in the Inner Galactic Plane with a Partial Configuration of the HAWC Observatory. *Astrophys. J.*, 817(1):3, 2016. doi: 10.3847/0004-637X/817/1/3.

[109] A. A. Abdo et al. TeV Gamma-Ray Sources from a Survey of the Galactic Plane with Milagro. *Astrophys. J.*, 664:L91–L94, 2007. doi: 10.1086/520717.

[110] Francis Halzen, Alexander Kappes, and Aongus O'Murchadha. Prospects for identifying the sources of the Galactic cosmic rays with IceCube. *Phys. Rev.*, D78:063004, 2008. doi: 10.1103/PhysRevD.78.063004.

[111] M. C. Gonzalez-Garcia, Francis Halzen, and Soumya Mohapatra. Identifying Galactic PeVatrons with Neutrinos. *Astropart. Phys.*, 31:437–444, 2009. doi: 10.1016/j.astropartphys.2009.05.002.

[112] Gilles Ferrand and Samar Safi-Harb. A Census of High-Energy Observations of Galactic Supernova Remnants. *Adv. Space Res.*, 49:1313–1319, 2012. doi: 10.1016/j.asr.2012.02.004.

[113] Mathieu de Naurois. The Very High Energy Sky from 20 GeV to Hundreds of TeV - Selected Highlights. *PoS*, ICRC2015:021, 2016. doi: 10.22323/1.236.0021. [34,21(2015)].

[114] A. Lamastra, F. Fiore, D. Guetta, L. A. Antonelli, S. Colafrancesco, N. Menci, S. Puccetti, A. Stamerra, and L. Zappacosta. Galactic outflow driven by the active nucleus and the origin of the gamma-ray emission in NGC 1068. *Astron. Astrophys.*, 596:A68, 2016. doi: 10.1051/0004-6361/201628667.

[115] F. Aharonian et al. Observations of selected AGN with H.E.S.S. *Astron. Astrophys.*, 441:465–472, 2005. doi: 10.1051/0004-6361:20053478.

[116] A. S. Wilson and J. S. Ulvestad. Radio jets and high velocity gas in the Seyfert Galaxy NGC 1068. , 275:8–14, December 1983. doi: 10.1086/161507.

[117] S. G. Neff, M. N. Fanelli, L. J. Roberts, R. W. O'Connell, R. Bohlin, M. S. Roberts, A. M. Smith, and T. P. Stecher. Ultraviolet imaging of the AGN+starburst galaxy NGC 1068. , 430:545–549, August 1994. doi: 10.1086/174429.

[118] Tova M. Yoast-Hull, J. S. Gallagher III, Ellen G. Zweibel, and John E. Everett. Active Galactic Nuclei, Neutrinos, and Interacting Cosmic Rays in NGC 253 and NGC 1068. *Astrophys. J.*, 780:137, 2014. doi: 10.1088/0004-637X/780/2/137.

[119] Jakob van Santen. IceCube-Gen2: the next-generation neutrino observatory for the South Pole. *PoS*, ICRC2017:991, 2018. doi: 10.22323/1.301.0991.

[120] Gabriella Berti Logan. Women and the practice and teaching of medicine in bologna in the eighteenth and early nineteenth centuries. *Bulletin of the history of medicine*, 77:506–35, 02 2003. doi: 10.1353/bhm.2003.0124.

[121] Horizon 2020. https://ec.europa.eu/programmes/horizon2020/en/what-horizon-2020.

[122] Physiscope. https://scienscope.unige.ch/physiscope/.

[123] GiP UNIGE. Gender in Physics Day , UNIGE : Athena Program. https://indico.cern.ch/event/525539/contributions/2352169/attachments/1402271/2140757/Genera2016.pdf.

[124] European Comission. She Figures 2015 - Gender in Research and Innovation. https://ec.europa.eu/research/swafs/pdf/pub_gender_equality/she_figures_2015-final.pdf.

[125] Teresa Rees. Mainstreaming gender equality in science in the european union: The 'etan report'. *Gender and Education*, 13(3):243–260, 2001. doi: 10.1080/09540250120063544. URL https://doi.org/10.1080/09540250120063544.