VRIJE UNIVERSITEIT BRUSSEL

Dissertation for the Degree of Master in Physics

INICE VETO STUDIES FOR THE FUTURE ICECUBE-GEN2 NEUTRINO OBSERVATORY

PAUL COPPIN
2016-2017

Promotor: Prof. Dr. Nick van Eijndhoven
Supervisors: Drs. Simona Toscano & Jan Lünemann
Science & Bio-engineering Sciences

"The first principle is that you must not fool yourself and you are the easiest person to fool. So you have to be very careful about that. After you've not fooled yourself, it's easy not to fool other scientists. You just have to be honest in a conventional way after that."

— Richard P. Feynman[1]

"*Muons are sneaky ...*"

— Kevin Meagher[2]

1 R. P. Feynman. Surely You're Joking Mr Feynman: Adventures of a Curious Character as Told to Ralph Leighton. Random House, 2014.

2 Kevin Meagher, personal communication, March 2017.

SUMMARY

In recent years, the discovery of cosmic neutrinos and gravitational waves has opened the road towards multi-messenger astronomy. However, no coincident events between photons, gravitational waves or neutrinos have so far been observed. Combined with the diffuse nature of the observed neutrino flux, this entails that the sources of the cosmic neutrinos are still unknown. Likewise, it remains a mystery which sources produce cosmic rays with energies exceeding 10^{18} eV. While we know that these sources must be extragalactic, the deflection of cosmic rays by (inter)galactic magnetic fields prevents their direction from being reconstructed. Neutrinos, in contrast, travel to Earth in a straight line. As the hadronic acceleration that takes place at cosmic ray sources should also produce copious amounts of high-energy neutrinos, these neutrinos can be used to identify the cosmic ray sources. Apart from being an interesting discipline by itself, neutrino astronomy thus also plays a vital role in the field of multi-messenger astronomy.

IceCube has undeniably been a pioneer in the field of neutrino astronomy. Currently, it still is the only detector that has been capable of observing the cosmic neutrino flux, though various experiments (e.g. KM3NET) are expected to reach the required sensitivity in the near future. One of the main difficulties of measuring the cosmic neutrino flux, is that the expected rate of events is extremely low. With its 1 km^3 volume, IceCube observes only on the order of 15 cosmic neutrinos per year. A detector upgrade is therefore planned, which will increase the rate at which cosmic neutrinos are observed by extending the detector volume. Multiple lay-outs describing this extension have been proposed and are currently still under consideration.

As a result of the low rate at which cosmic neutrinos are observed, filters with small background passing fractions $\mathcal{O}(10^{-5})$ are required to distinguish them from background events. These background events are mostly caused by atmospheric muons, created in cosmic ray air showers. An effective way to distinguish them from signal events, is to use InIce self-veto techniques. In this method, events related to entering particles are rejected by means of the light they emit upon crossing the detector border. Algorithms encoding this technique have been implemented for IceCube, but are less suited for the larger volume of the proposed IceCube-GEN2 extension.

In this thesis, we therefore developed alternative self-veto algorithms, designed to be scalable to larger detector volumes. Various self-veto concepts were first tested by applying them to one year of IceCube data and comparing their performance to that of the algorithm that is currently in use. The best performing candidate was then selected and applied on simulated events in the proposed IceCube-GEN2 geometries. This allowed optimising the filter parameters for each of the geometries and comparing their relative self-veto capabilities.

SAMENVATTING

De afgelopen vijf jaar hebben de ontdekking van kosmische neutrinos en zwaartekrachtgolven geleid tot het ontstaan van twee nieuwe takken binnen de astronomie. Voorlopig is het echter nog wachten op de eerste gelijktijdige waarneming van een extragalactisch object met fotonen, neutrinos of zwaartekrachtgolven. Mede hierdoor zijn de bronnen van de kosmische neutrinos tot op heden nog onbekend. Tevens blijft het een mysterie welke bronnen de kosmische straling met energieën groter dan 10^{18} eV produceren. Hoewel geweten is dat deze deeltjes van buiten de Melkweg komen, kan de positie van hun bronnen niet bepaald worden doordat de geladen deeltjes afbuigen in (inter)galactische magneetvelden. Neutrinos reizen daarentegen in een rechte lijn naar de Aarde. Gezien verwacht wordt dat de hadronische versnelling die plaatsvindt in de bronnen van kosmische straling ook tot de productie van neutrinos leidt, zijn neutrinos ideale kandidaten om de bronnen van de kosmische straling te achterhalen. Hoewel neutrino astronomie op zich al een interessant vakgebied is, speelt het dus ook een belangrijke rol binnen andere domeinen van de astrodeeltjesfysica.

Het IceCube experiment is onmiskenbaar een pionier geweest in het vakgebied van de neutrino astronomie. Momenteel is IceCube nog steeds het enige experiment dat de kosmische flux van neutrinos heeft kunnen waarnemen. Verscheidene nieuwe detectoren (o.a. KM3NET) zullen binnenkort hiervoor echter de vereiste sensitiviteit bereiken. Één van de grote moeilijkheden om de kosmisch neutrino flux te detecteren is de lage waarnemingsfrequentie van events. Met een volume van één kubieke kilometer neemt IceCube gemiddeld ongeveer 15 kosmische neutrinos per jaar waar. Momenteel is een uitbreiding van IceCube in ontwerp, die deze waarnemingsfrequentie zal verhogen door het volume van de detector te vergroten. Verscheidene voorstellen om deze uitbreiding te verwezenlijken worden momenteel nog in beschouwing genomen.

Ten gevolge van de lage waarnemingsfrequentie van kosmische neutrinos, zijn filteralgoritmes vereist die zeer weinig achtergrond $\mathcal{O}(10^{-5})$ doorlaten. Achtergrond events worden voornamelijk veroorzaakt door atmosferische muonen, die gecreëerd worden wanneer kosmische straling op de atmosfeer inslaat. Zelf-veto technieken zijn een effectieve methode om deze events van kosmische neutrinos te onderscheiden. Dit gebeurt door alle events waarin een deeltje de detector binnendringt te verwerpen, op basis van het licht dat aan de rand van de detector waargenomen wordt. Implementaties van deze techniek werden reeds gecodeerd voor IceCube, maar zijn niet geschikt voor het grotere volume van de IceCube-

GEN2 uitbreiding.

In deze thesis ontwikkelden we daarom alternatieve zelf-veto algoritmes, die schaalbaar zijn naar grotere volumes. Verscheidene zelf-veto concepten werden eerst getest door ze toe te passen op één jaar van IceCube data en de resultaten te vergelijken met die van het algoritme dat momenteel gebruikt wordt. De best presterende methode werd dan toegepast op gesimuleerde events in de voorgestelde IceCube-GEN2 geometrieën. Dit liet toe om de filter parameters te optimaliseren voor elk van de geometrieën en hun relatieve zelf-veto capaciteiten te vergelijken.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my promoter, Prof. Dr. Nick van Eijndhoven, and supervisors, Drs. Simona Toscano and Jan Lünemann. Unlike my fellow classmates, I already choose my master thesis topic during the first year of my master program. Thank you Nick, for immediately providing me with an interesting topic when I came knocking on your door. Whenever I got stuck on a problem during the year, I could always count on your expertise and problem solving skills to set me back on track. Simona, you been my main source of guidance throughout my master thesis. Thank you for your unwavering support. I was truly sad to see you leave the IIHE and am very grateful that you continued to help and counsel me even after you changed jobs. Nonetheless, with Jan as my new supervisor, I feel that I was left in very capable hands. Thank you Jan, for helping me finish my master thesis and becoming my new go-to person.

Apart from Nick, Simona and Jan, I would like to thank everyone from the IIHE and IceCube collaboration who has helped me during the course of my master thesis. Chris, Gwen, Kevin, Elisa, Giuliano and Krijn, thank you for letting me barge into your offices unannounced to ask you questions. Isabelle, Nadège and (for a short time) Kevin, thank you for sharing your office space with me. Your companionship always made me enjoy my stays here. I also want to thank the IIHE IT team, for helping me set up and run my filter algorithms on the IIHE cluster. From the IceCube collaboration, I specifically want to thank Jakob van Santem. I know you're a busy person and therefore appreciate you taking the time to set up and generate a new simulation dataset for my analysis all the more. In addition, I am also grateful to my fellow students Jonas and Pablo, for following up on my progress and for their companionship during tea-time.

Finally, I want to express gratitude to my family. In particular, I thank my parents for all the support that they have given me and for taking an interest in the topic of my thesis. I also specifically want to thank my sister and godmother, for all the encouragement that they have given me.

CONTENTS

ACRONYMS

AGN	Active Galactic Nuclei
AMANDA	Antarctic Muon And Neutrino Detector Array
ANTARES	Astronomy with a Neutrino Telescope and Abyss environmental RESearch project
ATWD	Analogue Transient Waveform Digitiser
CC	Charged Current
COG	Centre Of Gravity
DOM	Digital Optical Module
ESTES	Enhanced Starting Track Event Selection
fADC	Fast Analogue to Digital Converter
GRB	Gamma Ray Burst
GZK	Greisen, Zatsepin and Kuzmin
HESE	High-Energy Starting Event
HLC	Hard Local Coincidence
ICL	IceCube Laboratory
MSW	Mikheyev, Smirnov and Wolfenstein
NC	Neutral Current
pe	PhotoElectron
PINGU	Precision IceCube Next Generation Upgrade
PMT	PhotoMultiplier Tube
PnF	Processing and Filtering
PREM	Preliminary Earth Reference Model
SNO	Sudbury Neutrino Observatory

1 INTRODUCTION

1.1 ASTROPHYSICAL NEUTRINOS

In June 2013, the IceCube collaboration reported the observation of two PeV neutrinos, recorded in a dataset of 615.9 days livetime. With energies of (1.04 ± 0.16) PeV and (1.14 ± 0.17) PeV, these were at that time the most energetic neutrinos ever observed. Since the hypothesis that both neutrinos have an atmospheric origin could be excluded at 2.8σ, this observation provided the first evidence for the existence of an astrophysical neutrino flux [29].

Spurred on by this initial success, new selection criteria were devised to search for astrophysical neutrinos at lower energies. By looking for events which have their vertex inside the IceCube detector, neutrino induced events could be selected down to an energy threshold of 30 TeV[1]. Applying this technique to the same dataset resulted in the discovery of 26 additional events, with energies ranging from (30.5 ± 3.2) TeV to (253 ± 26) TeV. This is significantly larger than the total estimated background of $10^{+5.0}_{-3.6}$ events, comprised of penetrating muons and atmospheric neutrinos. As such, the hypothesis that all 28 events have an atmospheric origin could be excluded at 4.1σ [28].

As time passes, the number of observed events steadily increases. As such, it becomes possible not only to prove the existence of astrophysical neutrinos, but also to study their properties[2]. At this time, a large number of possible sources have been proposed to produce the observed neutrino spectrum. These range from cosmic accelerators, such as active galactic nuclei (AGN) [45] and gamma ray bursts (GRBs) [26], to the annihilation of dark matter in the centre of our galaxy [34]. Determining the significance with which these sources contribute to the observed neutrino spectrum, could give invaluable insight into the nature of cosmic ray accelerators, dark matter and many other phenomena. Unfortunately, the current dataset of high-energy neutrinos has only allowed placing upper limits on the contribution of the hypothesised sources [34, 61].

To locate the sources of the discovered cosmic neutrinos, a point source search can be applied to their reconstructed arrival direction. Since neu-

1 In reality, the threshold lies at 6000 photoelectrons (pe), corresponding to an effective energy cut at ~30 TeV.

2 For instance, the distribution of the energy and flavour of the observed neutrinos.

trinos travel from their creation point to detectors on Earth in a straight line, the clustering of events in galactic coordinates would indicate the location of an astrophysical neutrino source. No statistically significant clustering of events has however been observed [39]. Similarly, searches for the coincident detection of neutrinos with GRBs [26] or gravitational waves [4] have yet to yield the first confirmation of a coincident event.

A common feature of the above mentioned analyses, is that they are inconclusive in determining a clear origin for the astrophysical neutrinos due to a lack of data. Fortunately, increasing the rate at which data is recorded is a problem that can easily be solved, namely by building a bigger detector.

1.2 ICECUBE-GEN2

In its current configuration, the IceCube detector, as outlined later on, consists of 5160 Digital Optical Modules (DOMs) that monitor the emission of light inside a 1 km^3 volume of ice. These DOMs can detect the light emitted by the highly relativistic charged particles that are produced in neutrino interactions. At a given energy, the rate at which neutrino interactions are observed is proportional to the detector volume. More events could thus be detected if the detector volume was increased by adding additional DOMs. The size of the added detector volume would naturally depend on the spacing between the DOMs. That is, by adding a fixed number of DOMs with spacings larger than the one used in IceCube, significantly larger detector volumes can be obtained.

A downside of increasing the spacing is that it becomes more difficult to reconstruct low-energy events. Thus, the energy threshold of the detector effectively increases. Additionally, using a larger spacing also lowers the efficiency with which self-veto algorithms can identify background events, as will be explained in detail hereafter. These background events are mainly atmospheric muons, produced in cosmic ray air showers. Because they enter the detector, the first hits in background events typically occur at the edge of the detector. In contrast, such hits will be absent when a neutrino interacts inside the detector volume. Astrophysical neutrinos can thus be filtered from background events, by selecting those events in which no hits occurred at the edge of the detector at the start of the event. Naturally, it follows that self-veto techniques become less efficient for increased spacings. As the spacing increases, so does the probability that a muon can pass the edge of the detector without causing any hits. As such, when requiring the same signal-to-noise ratio as the current IceCube detector, the energy threshold imposed on the filter algorithms increases along with the spacing.

One way to circumvent this issue, is to use a different filtering technique altogether. For instance, neutrino events can be filtered by only selecting track-like events coming from the Northern hemisphere. Using this approach, the Earth is effectively used as a filter that blocks all atmospheric muons. Downsides to this method are that requiring a track signature restricts it to (anti-)muon neutrinos and that at energies $E_\nu > 100$ TeV, the sensitivity is reduced due to neutrino absorption in the Earth[3]. As a result, Earth absorption based filters are ill-suited for analyses that determine the flavour and energy distribution of the high-energy cosmic neutrino flux.

Improving the characterisation of the cosmic neutrino flux is one of the main goals of the upgrade of the current IceCube observatory, dubbed IceCube-GEN2. In achieving this objective, self-veto techniques will play a vital role. To lower the energy threshold from which events can reliably be filtered, new self-veto algorithms are explored, as reflected in this thesis. Various filter concepts are first investigated by applying them to IceCube data. From this analysis, the best performing candidate is selected and applied to simulated IceCube-GEN2 data. This then allows not only to determine the filter efficiency in the IceCube-GEN2 array, but also to compare the filter performance for the different proposed geometries.

1.3 OVERVIEW

To clarify the structure of this thesis, a short overview describing the content of each chapter is given below.

- Chapter 2: Developing a filter algorithm requires a thorough knowledge of the detector itself. Therefore, this chapter starts by introducing the IceCube detector. Apart from presenting its general characteristics, such as the detector geometry, elements relevant to the development of self-veto algorithms are also discussed in detail. These include the removal of noise hits, the variation of the dust concentration with depth and the occurrence of afterpulses. With the detector related effects specified, the identification of particles based on their emitted light distribution is ensuingly described. At the end of the chapter, an overview is then given on neutrino telescopes other than the IceCube detector.

- Chapter 3: Having obtained a detailed understanding of the IceCube detector, we move on to discuss its discoveries and their importance in the field of neutrino astronomy. This then brings us to a number of unanswered questions, such as: “What are the sources of the discovered cosmic neutrinos?”. To answer these questions, an upgrade to the IceCube configuration is currently under development. The second part of this chapter therefore discusses the

3 The neutrino-nucleon cross section increases with the neutrino energy (see Figure 54).

four main components of this upgrade, with special attention going to the high-energy InIce extension.

- Chapter 4: In the fourth chapter, the filter algorithms that are currently used to separate astrophysical neutrinos from background events are examined. Specifically, both the HESE and Earth absorption based filter are described in detail. As both filters had their own analysis to estimate the characterisation of the cosmic neutrino flux, they each have their corresponding parametrisation. These parametrisations and possible causes for discrepancies between them are then discussed. Returning back to IceCube-GEN2, a separate analysis is ensuingly examined, which was designed to estimate the passing rate of background events if the IceCube geometry had a DOM spacing similar to the one foreseen for IceCube-GEN2.

- Chapter 5: Chapter five introduces the concept of a new self-veto algorithm, which is based on imposing strong causal constraints on hits used to identify incoming tracks. Multiple variations of this algorithm are examined, by applying them to one year of IceCube data. From these candidates, the best performing algorithm is selected and analysed in more detail. To evaluate its efficiency, a comparison is then made between this causal filter and the HESE algorithm.

- Chapter 6: As a next step, this newly developed causal filter is used to compare the self-veto capabilities of the proposed IceCube-GEN2 geometries. Because this requires the use of simulated data, a short introduction is first given on the simulation process itself and the accompanying uncertainties. To allow determining the performance of the causal filter by comparing signal to background rates, we ensuingly discuss the procedure via which the simulated events are weighted. In addition, we determine the range of filter parameters required to optimise its performance. With all this specified, the self-veto capability of the baseline IceCube-GEN2 extension is then analysed in detail for a range of filter configurations. This analysis is then repeated for the non-baseline extensions, allowing to determine which of the proposed geometries offers the best performance.

- Chapter 7: In the seventh and final chapter, we reiterate the most important results obtained in the previous chapters. Specifically, we discuss how the developed causal filter compares with filtering techniques that are currently in use and revisit its performance for the proposed IceCube-GEN2 extension. An outlook on the future of self-veto techniques is then given, reviewing their importance in the field of neutrino astronomy.

2 THE ICECUBE DETECTOR

IceCube is a detector designed to search for neutrinos with energies larger than 100 GeV. This is done by detecting the Cherenkov photons that are emitted when charged particles travel faster than the phase velocity of light in a dielectric medium. These particles are created when a high-energy neutrino interacts with matter. As neutrinos only interact weakly, the main requirement for any astrophysical neutrino telescope is that it encompasses a large interaction volume which is transparent to Cherenkov radiation. For this reason, the IceCube detector is located in the deep ice, next to the Amundsen–Scott station at the geographical South Pole [41].

Before IceCube was build, a small scale version of the experiment was realised. AMANDA , short for the Antarctic Muon And Neutrino Detector Array, was effectively IceCube's predecessor. During the first iteration of the experiment, the concept of detecting neutrino interactions by means of Cherenkov radiation in ice was investigated. The second iteration, AMANDA-II, was composed out of 677 DOMs distributed over 19 vertical strings. It recorded events from the year 2000 onwards. In 2005, AMANDA-II became part of the IceCube detector, until it was decommissioned in 2009 [46]. *AMANDA*

2.1 DETECTOR GEOMETRY

The IceCube detector consists of 5160 InIce DOMs, which together encompass a 1 km^3 volume of ice. DOMs are distributed over 86 vertical strings in a hexagonal geometry, with a horizontal string spacing of 125 meter. Each string is equipped with 60 DOMs, vertically spaced by 17 meter[1] and located between depths of 1450-2450 meter[2] [24]. The central eight strings have a decreased vertical spacing of about 7 meter between DOMs and an average horizontal spacing of 70 meter between strings, allowing the observation of events down to 10 GeV. These central strings constitute the DeepCore detector. In its current configuration, the IceCube detector is often referred to as IC86. More generally, ICXX refers to the IceCube configuration in which only the first XX strings are included. IC40 for instance, was used to take data when about half of the detector was completed [41]. In addition, we will use *IC86* *DeepCore*

1 The horizontal and vertical spacing between DOMs differs by almost an order of magnitude due to the cost of cables and the cost of and time required to drill holes in the ice.

2 As the Amundsen–Scott station is located 2.830 meter above sea level [36], the IceCube detector has an average elevation of 880 meter.

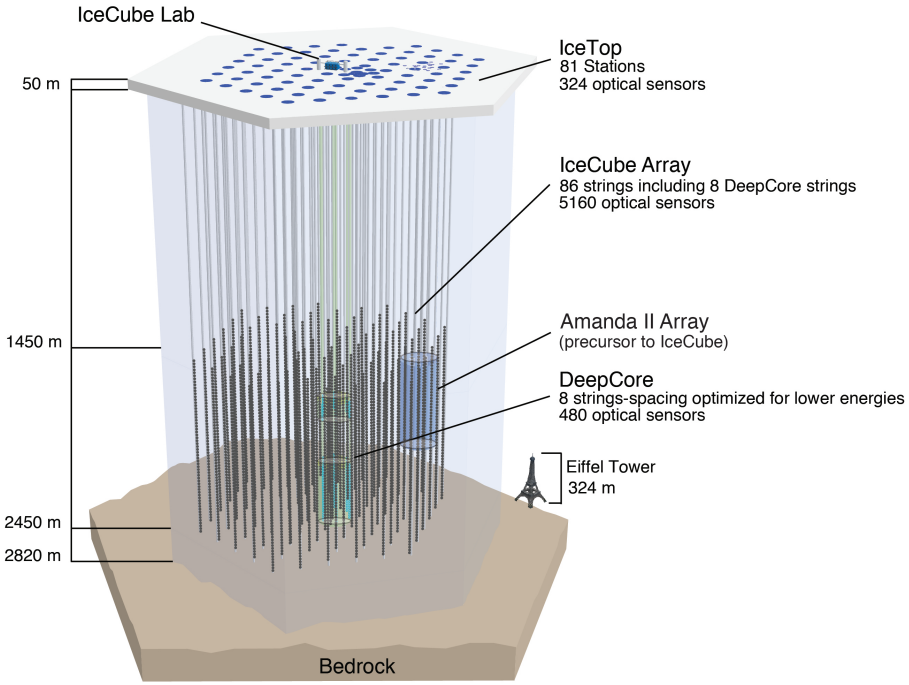

Figure 1: Graphical representation of the IceCube detector [41].

IC78 to refer to the complete IceCube detector without DeepCore, as the DeepCore strings are numbered from n° 79 to n° 86. A graphical representation of the geometry of the IceCube detector is shown in Figure 1.

IceTop

At the surface of nearly each string are two Cherenkov tanks, spaced 10 meter apart. In total, there are 162 tanks, making up 81 pairs. Each tank has a diameter of 1.86 m, is filled with ice and contains two DOMs [46]. These DOMs operate at different gains and can detect the Cherenkov radiation of charged particles created in cosmic ray air showers. The combination of all Cherenkov tanks constitute the IceTop detector, which is part of the IceCube detector. IceTop can detect air showers with energies from 1 PeV to 1 EeV [41]. As a result, this surface array can be used to filter atmospheric neutrinos from astrophysical neutrinos, as the former are accompanied by energetic muons detectable by IceTop.

Air bubbles

The depth range of the InIce DOMs, 1450 to 2450 meter, was chosen based on results of the AMANDA experiment, which showed that bubbles of air are present above depths of 1400 meter. These bubbles cause the average scattering length of photons to be less than 2 meter. Below 1400 meter, the bubbles collapse due to the pressure, increasing the average scattering length of photons by more than an order of magnitude [46]. Air bubbles are however still present in the 'hole ice'. When deploying DOMs, a hole is drilled in which a string of DOMs is let down. After the deployment, the water in this hole freezes again, locking the DOMs

in the ice sheet. This procedure causes air bubbles to remain present at all depths in the ice surrounding the DOMs [30].

Muon energy-loss

In addition to the increased scattering length, the depth range at which DOMs are located ensures that they are shielded from low-energy atmospheric muons. For muons travelling through ice, the Bethe-Bloch formula can be approximated as

$$\frac{dE}{dx} = a + bE \quad , \tag{1}$$

where $a = -0.24$ GeV m^{-1} is due to ionisation and $b = -3.3 \cdot 10^{-4}$ m^{-1} is due to bremsstrahlung, pair production and nuclear interactions [13]. By integrating equation (1), we find

$$E(x) = -\frac{a}{b} + \left(\frac{a}{b} + E_0\right) e^{bx} \quad , \tag{2}$$

where E_0 is the initial energy of the muon. Solving equation (2) for the value of E_0 for which $E(x = 1450 \text{ m}) = 0$, we find that muons travelling vertically down on average do not reach the detector if $E_0 \leqslant 0.45$ TeV.

Dust layer

Both the absorption and scattering properties of light in ice are strongly dependent on the presence of impurities. As impurities in ice are a marker for the condition of the atmosphere when the snow was deposited, the properties of the Antarctic ice sheet are strongly depth dependent. At depths between 2000 and 2100 meter, the concentration of dust is significantly higher than that of the surrounding ice [46]. This region is known as the dust layer and is characterised by both a decreased absorption and scattering distance, as shown in Figure 2.

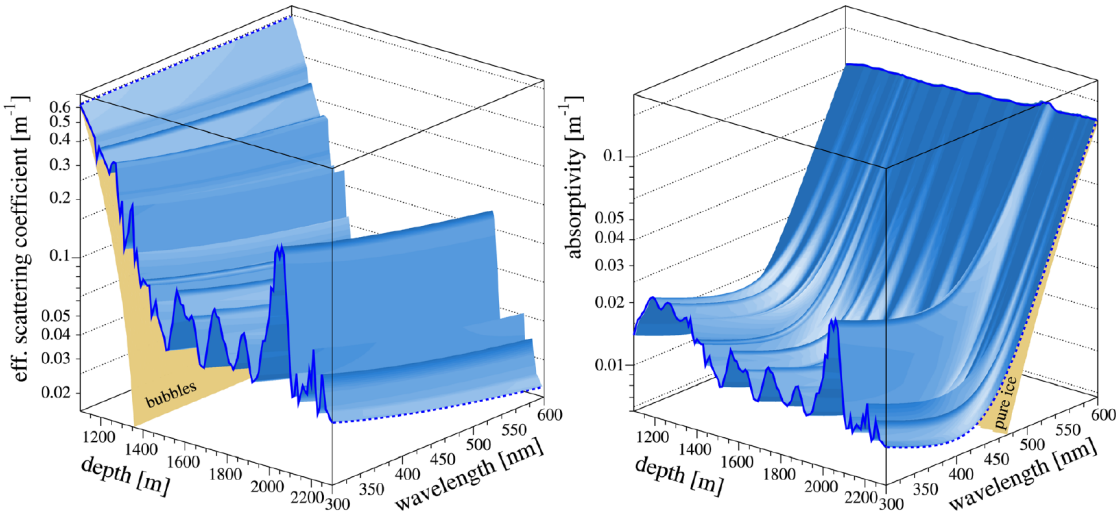

Figure 2: Average scattering and absorption distance of photons in ice as a function of depth and wavelength [1].

2.2 DATA PROCESSING AND FILTERING

ICL

Data from both the InIce DOMs and surface tanks is sent via cables to the IceCube Laboratory (ICL), which is located at the centre of the IceTop detector. Triggers are then applied to the data from the individual DOMs, grouping them into an event if one or more trigger criteria are satisfied. Examples of online triggers [41] include the

- simple multiplicity trigger, which requires N hard local coincidence[3] hits to occur within a fixed time window.
- string trigger, which requires N out of M consecutive DOMs on a single string to be hit.
- volume trigger, which requires N hits to occur in a cylinder centred on a hit DOM.

PnF

All triggered events, which occur at a rate of about 3 kHz, ensuingly go to the Processing and Filtering (PnF) system. During this step, the waveform measured by the DOMs is deconvolved based on the individual response function of the DOMs, to identify the individual Cherenkov photon pulses. Various reconstruction algorithms are then applied to the event. Based on the reconstructed properties and goodness-of-fit, 15% of all triggered events are selected[4]. This data, in combination with the detector calibration, is then send via satellite to data centres in the Northern hemisphere [41].

2.3 THE ICECUBE DOM

Glass sphere

Time resolution

The IceCube digital optical module (DOM), shown in Figure 3, is a module designed to detect and process light emitted by the Cherenkov effect. Its outer shell, capable of continuously withstanding 250 bar of pressure, is a 1.3 cm thick glass sphere, which has a diameter of 35 cm [41]. Each DOM contains a Hamamatsu R7081-02 photomultiplier tube (PMT), connected to two analogue transient waveform digitisers (ATWDs) and a fast analogue to digital converter (fADC). When triggered, one of the ATWDs digitises 427 ns of the PMT's waveform into 128 samples. To ensure capturing the entire event, the fADC also digitises 6.4 μs of the waveform into 256 samples[5]. As digitising the waveforms takes 29 μs, the other ATWD meanwhile stands by to respond to any additional triggers during this time interval [25, 24]. After digitisation, the mainboard of the DOM sends the waveform, trigger information and time stamp to the ICL via a twisted pair of copper cables. These copper

3 Hits receive a hard local coincidence (HLC) tag if one of the two upper or lower DOMs also observed a hit within 1 μs.

4 The trigger efficiency is determined by the 100 GB/day bandwidth of the satellite connection [41].

5 6.4 μs is longer than the period during which light from the most energetic events reaches any of the DOMs [46].

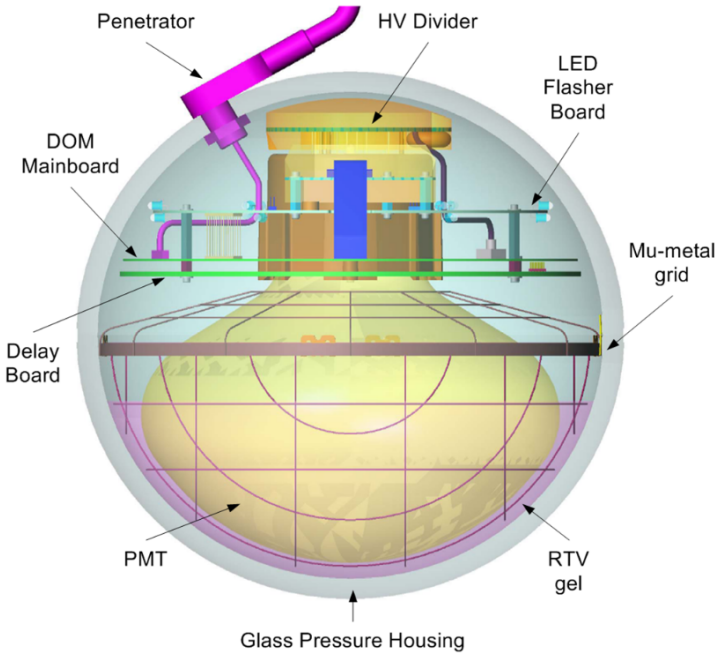

Figure 3: Graphical representation of the IceCube DOM [24].

cables also connect adjacent DOMs to each other. Via recursion, DOMs can thus exchange trigger information with other DOMs that are on the same string [41].

DOMs are triggered when they measure a pulse of which the maximum exceeds 0.25 times that of a single photoelectron (pe)[6]. In case any of the neighbouring two upper or two lower DOMs also register a hit within 1 μs, the hits are tagged as hard local coincidence (HLC) hits. Waveform information is only send to the ICL if hits received an HLC tag. Otherwise, only twelve bytes containing coarse information about the charge, timing and triggers of the hit are transmitted[7] [24].

HLC

For calibration purposes, such as measuring the timing response and position of the DOMs, each DOM is equipped with a set of 12 LEDs. Together with the DOMs mainboard, they are contained in the upper half of the DOM. The lower half is occupied by a downward-facing PMT, operating at a gain of 10^7 [25]. These PMTs have a peak quantum efficiency of 25% at 415 nm, for photons travelling upwards, parallel to the PMT's axis[8]. Absorption in the glass shell in combination with the PMT's characteristics limit the sensitivity of DOMs to photons in the wavelength range of 300 to 650 nm [30]. To eliminate effects of the Earth's mag-

LEDs

Quantum efficiency

6 The unit pe corresponds to the charge a single photoelectron is most likely to generate in a PMT. On average, the deposited charge per photoelectron is 14% lower than one pe [31].

7 The timing resolution of HLC hits is < 3.3 ns, in contrast to 7 ns for non-HLC hits [24].

8 The PMTs used in DeepCore have a quantum efficiency that is 35% higher than that of the standard IceCube PMTs [27].

netic field on the PMTs, a mu-metal wire mesh is embedded in the outer shell, reducing the magnetic field by a factor 2.8 inside the DOMs [41].

Noise hits

Dark noise hits registered by the PMTs are mostly due to radioactive decay and scintillation in the glass sphere and occur at an average rate of about 300 Hz [25]. Afterpulses are another type of noise hits, related mostly to energetic events. They occur typically 300 ns to 11 μs after photoelectron induced hits, due to ions incident on the photocathode. These ions are created when the accelerating electrons ionise residual gasses in the PMT [25].

2.4 CLASSIFICATION OF EVENTS

Cherenkov mechanism

When a charged particle travels faster than the phase velocity $\frac{c}{n}$ of light in a dielectric material with (phase) refractive index n, it creates a shock front of coherent electromagnetic radiation. The mechanism by which this radiation is produced is described by the Cherenkov effect. Photons are emitted at an angle θ with respect to the direction of the charged particle, perpendicular to the shock front, where

$$\cos(\theta) = \frac{1}{\beta n} \quad . \tag{3}$$

For highly relativistic $(\beta \to 1)$ particles in ice, θ is approximately[9] 41°. The spectrum of the emitted radiation goes as λ^{-2}, as described by the Frank-Tamm formula

$$\frac{d^2N}{d\lambda dx} = \frac{2\pi\alpha}{\lambda^2}\sin^2(\theta) \quad , \tag{4}$$

where N is the number of emitted photons, x the distance covered by the charged particle, λ the wavelength of the photons and α the fine-structure constant [30]. Figure 4 illustrates the propagation of a muon (red tube) through the IceCube detector. Emerging from the track are numerous photons, which have a colour representative of their wavelength[10]. DOMs are drawn as grey spheres if they received no hits and drawn as green sphere otherwise. In the latter case, their volume is proportional to the number of hits they received.

Comparison of particles

Apart from Cherenkov radiation, other mechanisms [31] can significantly contribute to the light yield and energy loss of high-energy particles in ice.

- Electrons quickly lose all their energy in an electromagnetic cascade by emitting bremsstrahlung.

9 This approximation only applies in the wavelength range to which the IceCube DOMs are sensitive (300 nm< λ <650 nm), as the phase refractive index $n = n(\lambda)$ is a function of wavelength.

10 UV photons are drawn in purple.

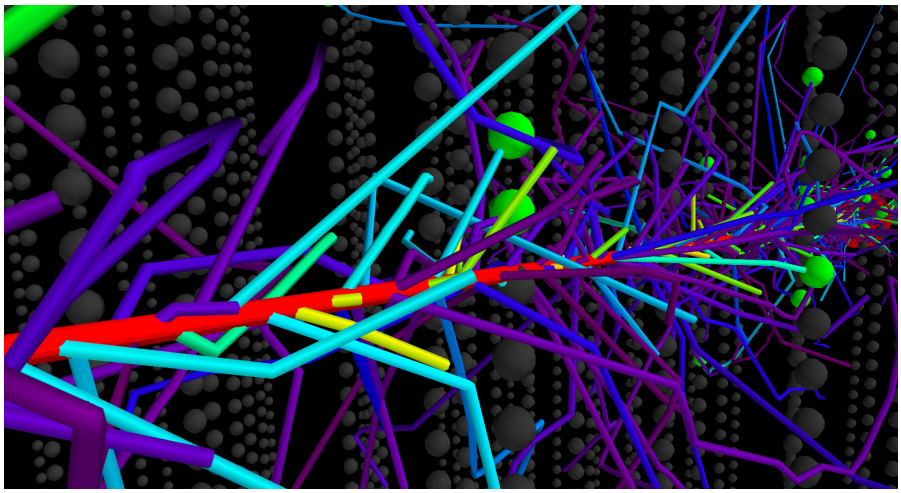

Figure 4: Visualisation of the Cherenkov photons generated by a 500 GeV muon (red) propagating through the IceCube detector.

- Muons with energies below 1 TeV lose most of their energy by ionisation, but mostly emit light via the Cherenkov effect. Muons with energies above 1 TeV lose most of their energy in stochastic energy losses related to bremsstrahlung, pair production and nuclear interactions. These stochastic energy losses then also dominate the emitted light distribution.
- Tauons, in contrast to muons, will decay before losing all their energy. This decay will produce either a hadronic or electromagnetic cascade, possibly in combination with an energetic muon.
- Hadrons quickly lose all their energy in an hadronic cascade, which on average produces a 15% lower light yield than electromagnetic cascades. As IceCube cannot distinguish the two types of cascades, all estimated energies from cascades are lower limits.

NC & CC interactions

The two possible types of neutrino interactions, neutral current (NC) and charged current (CC), are shown in Figure 5. In NC interactions, the neutrino only loses part of its energy in a hadronic cascade. As such, only lower limits on the neutrino energy can be computed. CC interactions in contrast allow estimating the full neutrino energy. Distinguishing CC from NC interactions in IceCube is however only possible for muon and tau neutrinos. In the latter case, the produced tau lepton must have an energy larger than several PeV. Otherwise, the cascade caused by the decay of the tauon will be indistinguishable from the primary hadronic cascade caused by the neutrino interaction[11] [31].

Topology

Events in IceCube can be identified as tracks, cascades or a combination of both. Tracks are associated to muons and appear as light de-

11 On average, tau particles will travel 50 $\frac{\mathrm{m}}{\mathrm{PeV}}$ before decaying.

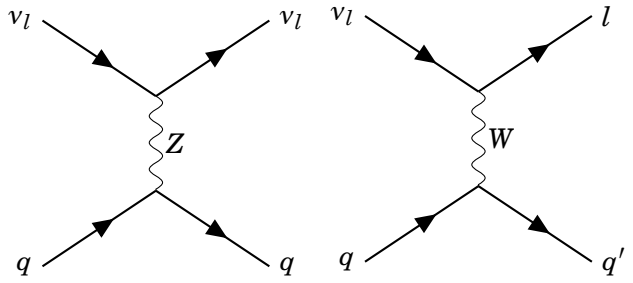

Figure 5: Leading order interactions between neutrinos and matter. NC and CC interactions are respectively displayed in the left and right panel. Corresponding diagrams can similarly be drawn for anti-neutrinos and for interactions with anti-quarks.

posited in an elongated cylinder[12]. Cascades correspond to hadronic or electromagnetic showers and have a ellipsoidal structure. As such, the angular resolution with which the neutrino direction of cascades can be reconstructed is ~15°, in contrast to 0.4° for tracks[13] [43]. For both tracks and cascades, the resolution on the deposited energy is ~15%, though this value is dependent on the analysis and energy of the event itself [31].

2.5 OTHER NEUTRINO TELESCOPES

Water based neutrino telescopes

Like IceCube, most neutrino telescopes are designed to detect the Cherenkov emission from charged particles created in neutrino interactions. As this requires a large interaction volume which is transparent to optical and/or UV photons, they are either located in large amounts of water or the Antarctic ice sheet. Currently operational water based telescopes, which are however much smaller than IceCube, include Super-Kamiokande [59], the Baikal detector [7] and ANTARES[14] [3]. An IceCube sized detector (KM3NeT [50]) is at this time under construction in the Mediterranean Sea. As it is located in the opposite hemisphere of IceCube, it will complement the detection of neutrinos, as both detectors effectively use the Earth as a muon filter. Operational neutrino telescopes in ice are currently limited to those related to the IceCube collaboration.

An alternative method to observe neutrinos is based on detecting the coherent Cherenkov radio emission coming from neutrino interactions [10]. This radio emission, described by the Askaryan effect, originates from the charge distribution of all particles due to the Cherenkov ef-

12 While tracks could also correspond to a tau lepton, the event would then have a 'double bang' topology. So far, no such events have been observed.

13 Assuming the track is 1 km long and that the cascade is fully contained within the detector.

14 ANTARES is an acronym for "Astronomy with a Neutrino Telescope and Abyss environmental RESearch project".

ARA & ARIANNA

fect. Concepts for ice based neutrino detectors designed to detect the Askaryan radio signal are under development by the ARA [5] and ARIANNA [6] collaboration. Both collaborations are currently deploying radio antennas in the Antarctic ice sheet. Once completed, they will be sensitive to neutrino events of several EeV. As such, these detectors will allow the study of neutrinos resulting from the interaction of ultra-high-energy cosmic rays with the microwave background, which is known as the GZK effect[15] [18]. Both ARA and ARIANNA will thus measure a different part of the neutrino spectrum than IceCube, which runs out of statistics around several PeV. Nevertheless, the potential to detect the same neutrino events with both the IceCube and ARA detector is being investigated. Using interferometric techniques, information on high-energy neutrino events detected by IceCube can significantly reduce the energy threshold of ARA to detect those events. If successful, the detection of the same neutrino events with both detectors would allow for a cross-calibration of the measured energy scale [62].

15 GZK is an acronym of three names: Greisen, Zatsepin and Kuzmin.

3

ICECUBE-GEN2

3.1 SCIENTIFIC MOTIVATION

3.1.1 *Neutrino astronomy*

With the discovery of a cosmic neutrino flux, IceCube opened a new window on our Universe. One of the main advantages of neutrino astronomy, is that neutrinos can pass through large amounts of matter while suffering little to no attenuation. This property is a consequence of the fact that neutrinos have neither a strong nor an electromagnetic charge. As such, neutrinos are an excellent candidate to probe dense regions from which photons and charged particles cannot escape without interacting with the surrounding medium. One example in which this scenario applies is Solar nuclear fusion. Photons take on average $\sim 2 \cdot 10^7$ year to exit the solar surface after their creation [52] and thus provide little information on the current state of the Solar core. This information is therefore instead extracted from solar neutrinos detected by e.g. Super-Kamiokande and SNO[1] [60, 58]. Other examples at higher energy scales include Gamma Ray Bursts (GRBs) and Active Galactic Nuclei (AGN).

Neutrinos as cosmic messengers

GRBs is the collective term for cataclysmic cosmic events in which two neutron stars collide or in which super-massive stars undergo an extreme supernova explosion. They are the most energetic events observed to date, detected by means of the gamma rays that are produced in the relativistic jets which emerge from the central engine [18]. AGN on the other hand, are super-massive ($10^6 - 10^9$ solar masses) black holes at the centre of galaxies which produce jets of relativistic particles by accreting the surrounding matter. Like GRBs, AGN are observed at cosmological distances, often larger than redshift $z = 1$ [49, 56].

GRBs

AGN

Both AGN and GRBs are hypothesised to be hadronic accelerators. This claim is made to explain the high-energy tail of the cosmic ray spectrum (10^{18} to 10^{21} eV), which has been observed by detectors such as the Pierre Auger observatory. No galactic sources are known which can accelerate particles up to these energies. Furthermore, galactic magnetic fields cannot contain these particles [56]. As such, they most likely have an extragalactic origin. Neutrinos are an excellent candidate to identify the sources of the ultra-high-energy cosmic rays. Unlike charged particles, their trajectory is not bend by (inter)galactic magnetic fields, mean-

Cosmic ray spectrum

1 SNO is an acronym for “Sudbury Neutrino Observatory experiment”.

ing that neutrinos point back to their source. They are also unaffected by interstellar matter, the cosmic background radiation and extragalactic background photons, which attenuate both the flux of hadrons and photons at high energies [56]. Because observing neutrinos from a known source would unambiguously identify it as a hadronic accelerator, neutrino astronomy is the most promising channel to identify the currently unknown sources of the ultra-high-energy cosmic rays.

3.1.2 *Results from IceCube*

Arrival direction of cosmic neutrinos

One of the main achievements of the IceCube experiment, has been discovering and characterising the cosmic neutrino flux. Figure 6 shows the reconstructed direction of the discovered cosmic neutrinos in galactic coordinates. The colour scale represents a test statistic indicating the uncertainty on their arrival direction. Analyses designed to look for a clustering of events along the galactic plane found no statistically significant correlation [43]. Moreover, a series of pseudo-random experiments in which points were distributed uniformly on the sky showed that an anisotropy at least as high as that of Figure 6 is obtained 58% of the time [19]. The observed distribution is thus consistent with an isotropic neutrino flux, showing no statistically significant clustering of events [61]. In combination with the neutrino flavour being consistent with $e : \mu : \tau = 1 : 1 : 1$ [19], this points to the observed neutrinos having an extragalactic origin. However, the possibility that part of the neutrino sample has a galactic origin should still be considered.

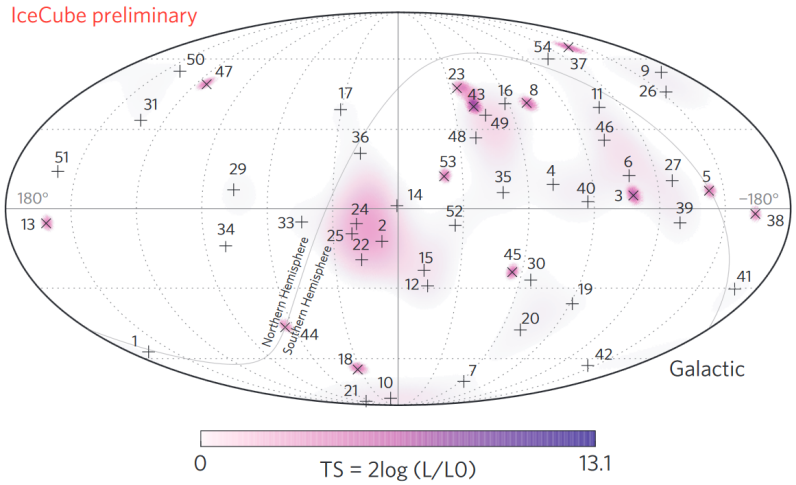

Figure 6: Reconstructed direction in galactic coordinates of the cosmic neutrinos found in four years of IceCube data. A test statistic indicating the uncertainty on their arrival direction is displayed by the colour scale. From this, it can be seen that events with a shower topology (+) have significantly larger errors on the reconstructed direction than those containing a muon track (x) [19].

Search for neutrinos from known sources

The above analyses relied on locating astrophysical sources based on the observed neutrino flux. An alternative approach is to search for neutrinos coming from known (extra)galactic sources. GRBs for instance, are excellent candidates to produce high-energy neutrinos. Both the time and position at which they occur can be measured by gamma ray space telescopes, as is currently done by e.g. FERMI [12] and SWIFT [49]. This information can then be used to search for an increased neutrino flux, coming from the GRB's direction during a time window centred around the observed photon flash. The same method can be applied to AGN. As the active lifetime of AGN, typically millions of year [63], is however much larger than that of GRBs (0.1-100 s), neutrinos coming from their direction are monitored continuously [45]. Separate analyses also exists to look for an increased neutrino flux when AGN flare [53].

Upper limits

In seven years of IceCube data, no significant correlation between neutrinos and AGN or GRBs has been found. This has restricted the prompt contribution from GRBs to the observed diffuse neutrino flux to less than 1%[2] [35]. In the energy range from 10 TeV to 2 PeV, less than 27% of the astrophysical neutrinos can be related to the observed AGN objects [40]. A combined analysis using both IceCube and Pierre Auger data also found no correlation between the observed neutrino flux and the highest-energy cosmic rays[3] [23]. To conclude, no extragalactic object has so far been identified as a source of the observed cosmic neutrinos or cosmic rays.

Other IceCube achievements

Apart from being a pioneer in the search for a cosmic neutrino flux, IceCube has achieved many other results. By searching for dark matter annihilation in the Sun, IceCube produced the best upper limits on the cross section for spin-dependent WIMP-proton scattering in the mass range 50 GeV< M <5 TeV [43]. The DeepCore part of the detector was used to measure the neutrino oscillation parameters by looking for a deficit in the flux of atmospheric muon neutrinos [22]. IceTop on the other hand provided the first confirmation of the cosmic ray anisotropy in the Southern hemisphere [57].

3.1.3 *IceCube-GEN2 goals*

The numerous results obtained using the IceCube detector have proven the scientific value of a Cherenkov neutrino telescope in the Antarctic ice sheet. IceCube's current configuration is however limited by the energies which can be observed and the rate at which events are detected. An upgrade to the InIce detector, designed to reduce these two restrictions,

2 This claim only applies to the prompt flux coming from observed GRBs and is no longer valid when considering precursor/delayed emission or hidden sources [54].

3 The trajectory of cosmic rays of the highest energies ($E \sim 10^{20}$ eV) is expected be to deflected only $\sim 5^{\circ}$ by (inter)galactic magnetic fields [23].

is therefore planned. This upgrade is mainly based on adding strings to the current InIce geometry, but will also feature an extended surface array and an array of radio antennas.

High-energy extension

With its high-energy extension, IceCube-GEN2 will increase the volume of the InIce detector by roughly a factor of ten. Correspondingly, the rate at which high-energy neutrino events are recorded will increase by up to an order of magnitude. Using this larger data sample, IceCube-GEN2 aims to address the following three currently unanswered questions.

1. Which objects are the sources of the high-energy neutrinos observed by the IceCube detector and by which interactions were those neutrinos produced?
2. Which objects are the sources of the ultra-high-energy cosmic rays and what is the acceleration mechanism that takes place at those sources?
3. Are there signatures of beyond the standard model physics in neutrino interactions at energy scales exceeding 1 PeV?

Neutrino cross section at $E_\nu > 1$ PeV

IceCube analyses related to question three have already measured the neutrino-nucleon cross section up to neutrino energies of 1 PeV. IceCube-GEN2 has the potential to significantly decrease the relative uncertainty on those measurements, which currently are 30% [51]. It will also allow to examine neutrino cross sections up to energy scales exceeding 10 PeV, far beyond what is currently reachable by man made accelerators.

Identifying the flavour of the neutrinos, in combination with their energy and direction, will give information on the processes in which they were produced. For instance, high-energy anti-electron neutrinos $\overline{\nu}_e$ are more likely to be produced in proton-proton collisions than in proton-photon collisions. Which of these two processes dominates can be determined by measuring the Glashow resonance, which greatly increases the interaction probability for $\overline{\nu}_e$ around

Glashow resonance

$$E_{\overline{\nu}_e} = \frac{m_W^2}{2m_e} = 6.3 \text{ PeV} \quad [17]. \tag{5}$$

Low-energy extension

With its low-energy extension, IceCube-GEN2 will decrease the energy threshold for neutrino detection to below 10 GeV. This will allow improving the current limits on the cross section for dark matter annihilation. More importantly, measuring the influence of the MSW effect[4] [66] on the neutrino oscillation parameters will determine whether the neutrino mass hierarchy is normal ($m_{\nu_1} < m_{\nu_2} < m_{\nu_3}$) or inverted ($m_{\nu_3} < m_{\nu_1} < m_{\nu_2}$) [9]. In addition, low-energy neutrino observations also open the door to perform neutrino tomography of the Earth [65].

4 MSW is an acronym of three names: Mikheyev, Smirnov and Wolfenstein.

3.2 PROPOSED PROJECTS

To achieve the above mentioned goals, multiple upgrades will be made to the IceCube detector. The four main components of the detector upgrade consist of

1. adding strings in the DeepCore region to lower the energy threshold below 10 GeV,
2. adding strings around the current IC86 configuration,
3. deploying a new, more extended surface array,
4. deploying an array of radio antennas.

Figure 7 shows the projected time line of the upgrade. Currently, the project is nearing the end of its 'research and development' phase, which is based on a total budget of 400 million US dollar.

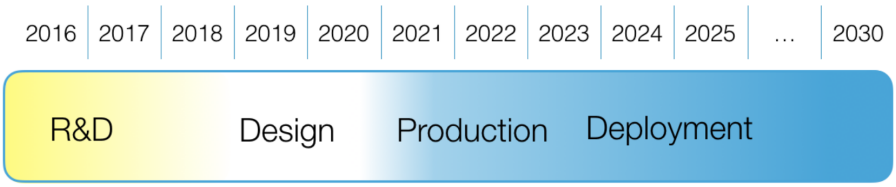

Figure 7: Projected timeline for IceCube-GEN2 [42].

3.2.1 *PINGU*

26 strings

The Precision IceCube Next Generation Upgrade (PINGU) will add 26 strings to the DeepCore region of the detector. Each string will contain 192 DOMs at depths below the dust layer. This will allow to observe events down to 3 GeV with an efficiency of 50% [38]. Based on this configuration, PINGU should determine whether the neutrino mass hierarchy is normal or inverted up to 3σ after five years of data. The deployment of PINGU strings is planned to start in 2020 and should take two polar seasons [38]. As such, PINGU will also serve as a testing ground for new DOM designs (see Section 3.3), the improved hot-water drill [43] and new calibration techniques.

Neutrino mass hierarchy

3.2.2 *High-energy InIce extension*

~120 strings

To increase the rate at which high-energy neutrino events are detected, about 120 new strings will be added around the IC86 configuration. Currently, three scenarios are still under consideration. These scenarios, shown in Figure 8, are the Sunflower, Edge-weighted and Banana geometry. In case of the Sunflower geometry, three different string spacings are being examined: 200, 240 and 300 meter. For the Edge-weighted geometry, the string spacing is 240 meter for the inner strings and 125

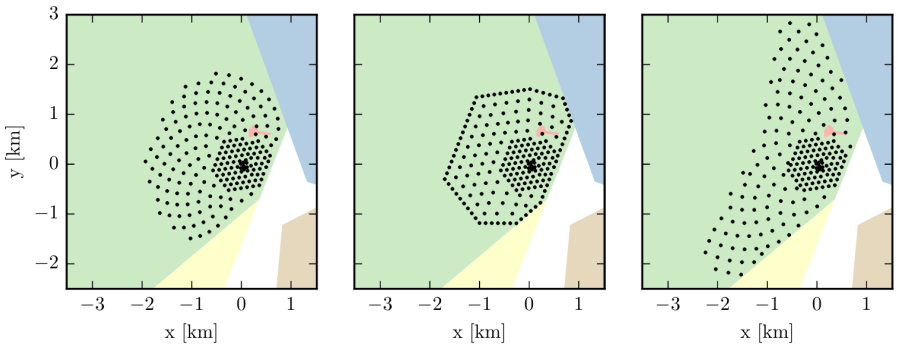

Figure 8: Proposed lay-out for the high-energy InIce extension. The left, middle and right image respectively show the Sunflower, Edge-weighted and Banana geometry. Added strings are restricted to the dark sector (green). The other regions correspond to the downwind sector (yellow), ski-way & pole station (white), old pole station (red), clean air sector (blue) and quiet sector (brown) [61].

meter for those at the edge. In the Banana geometry, strings are spaced 240 meter apart. The volume and top surface area of each of these geometries is listed in Table 1.

Table 1: Comparison of the proposed high-energy lay-out motifs.

Geometry	Volume (km^3)	Surface (km^2)	# added strings
Sunflower 200 m	5.94	4.7	122
Sunflower 240 m	7.90	6.1	122
Sunflower 300 m	11.8	9.0	122
Edge-weighted	6.18	4.8	125
Banana	10.5	8.1	120

spiral geometry

A clear difference between IC86 and the proposed geometries, is that the strings are no longer distributed in a hexagonal grid. A spiral geometry containing a random component was chosen instead, such that the average muon track reconstruction would not be dependent on the azimuthal angle of the track [43]. Another significant difference is that the string spacing, ranging from 200 to 300 meter, is much larger than in IC86. Choosing larger string spacings decreases the amount of light detected per event, thus degrading the resolution on reconstructed quantities such as energy and direction. In addition, larger string spacings decrease the efficiency with which filter algorithms can distinguish events which start inside the detector. As such, the Edge-weighted geometry will be better at identifying incoming tracks than the Sunflower geometry. This however comes at the cost of a decreased rate of signal events, as the Sunflower geometry has a larger volume than the Edge-weighted motif.

Banana geometry

The Banana geometry has both the largest surface and volume of all geometries with string spacing $\leqslant$240 meter. A downside is that a larger fraction of its strings are at the edge of the detector than in the Sunflower geometries. This effectively reduces the fiducial volume of the detector, in which events need to emit their first light to not be categorised as incoming tracks.

Dark sector

As indicated in Figure 8, the deployment of new strings is restricted to the dark sector (green). This prevents the high-energy extension from being centred on IC86, as would optimally be the case. The yellow, blue and brown areas respectively represent the down wind, clean air and quiet sector. These regions were designated by the Antarctic treaty and are reserved for balloon launches, atmospheric research and seismological research [2].

80 DOMs per string

Whilst having an increased horizontal spacing between strings, all geometries feature the same 17 meter vertical spacing between DOMs. This is the same distance as in IC86, as larger values would further degrade the efficiency of analysis algorithms looking for astrophysical neutrinos. All geometries however feature 80 DOMs per string, 20 more than in IC86. Six DOMs are added above and fourteen DOMs are added below the region covered by IC86. At low depths, the poor scattering and absorption properties of the ice are the limiting factors regarding the deployment of DOMs. At large depths, limitations are imposed by the total depth of the ice sheet and the possibility that shears in the ice sheet might break the cable.

3.2.3 *Surface array*

Plastic scintillators

IceCube's surface array, IceTop, consists of a set of tanks filled with clear ice. Each tanks contains a pair of DOMs, which can detect the Cherenkov radiation from particles generated in cosmic ray air showers. The surface array of IceCube-GEN2 will instead use 1 cm thick plastic scintillators, as they are much easier to deploy than Cherenkov tanks. Using 5000 scintillators, each 15 m^2 in size, the array would cover 0.1% of a 75 km^2 area [42]. Unlike IceTop and IceCube, this new surface array would thus have a much larger footprint than that of the InIce strings. A comparison between the size of the two components is shown in Figure 9.

Surface veto

A major advantage of the more extended surface array is that due to the larger size ratio with respect to the InIce strings, the acceptance to simultaneously detect events with the InIce DOMs is increased by a factor forty [42]. This noteworthy augmentation can be understood by considering a down-going muon going through IceTop and IC86. The requirement that it passes both detectors severely constrains the zenith

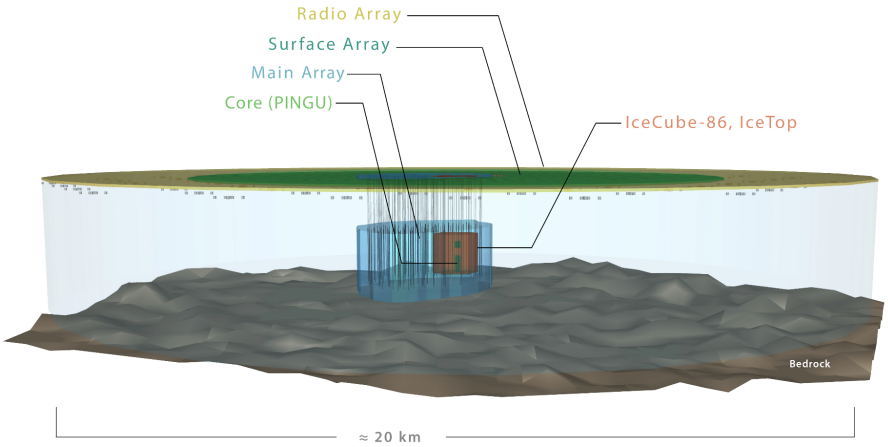

Figure 9: The four main components of the proposed IceCube-GEN2 upgrade [42].

angle of the muon track. For IceCube-GEN2, this constraint is relaxed. As a larger fraction of the InIce events will have a corresponding counterpart in the surface array, more events can be identified as incoming tracks. To supplement the capacity of the surface area as a cosmic ray detector, the addition of a grid of air Cherenkov telescopes is currently also under consideration [61].

3.2.4 *Radio array*

The final component of IceCube-GEN2 will be an array of radio antennas, covering an area of about 100 km^2. Current data suggests that the InIce detector will observe ~10 events per year for which E_ν >1 PeV. By deploying radio antennas on a grid surrounding the InIce detector, IceCube-GEN2 will be able to detect the Askaryan radiation from high-energy neutrino interactions in the ice. As such, it will be sensitive to neutrino interactions in the energy range $E_\nu = 10^{16} - 10^{20}$ eV, allowing the study of neutrinos resulting from for instance the GZK effect.

Askaryan radiation

3.3 ICECUBE-GEN2 DOMS

Deploying new strings in the ice naturally raises the question of which type of optical sensor should be attached to them. Figure 10 displays the four new modules which are at the moment under consideration. Studies for each of the modules are currently ongoing, to compare the quality with which events can be reconstructed and their collective capability to identify incoming muon tracks. Combined with the respective cost differences of the modules, these studies will determine which module will be used in the high-energy InIce extension.

P-DOM

IceCube DOMs have proven to be extremely reliable, with only 80

Figure 10: Proposed DOMs which are to be used for the high-energy InIce extension [44].

out of 5484 DOMs being permanently disabled after 7-10 years of operation[5]. A natural successor would thus be an upgraded version of the current model. The PINGU DOM (P-DOM) is the baseline version of the next-generation optical modules, featuring the same 25.4 cm diameter PMT that is used in IceCube DOMs. It has renewed electronics and is designed to be faster while consuming less power. In addition, a camera system is added to the module, which will allow to better study the properties of hole ice [44].

M-DOM

The second proposed module (M-DOM) is based on the multi-PMT design used in KM3NET. It houses 24 identical PMTs, each with a diameter of 7.6 cm. Accordingly, it has the largest total photocathode area of all the designs. In addition, the multi-PMT set-up provides directional information on the arrival direction of photons. With it's symmetric design, the M-DOM has an angular sensitivity that is close to uniform over the full 4π solid angle. A disadvantage of the large number of PMTs is that the M-DOM requires a larger diameter housing sphere than P-DOM, increasing the cost of drilling holes [61].

D-EGG

D-EGG can be seen as a compromise between P-DOM and M-DOM. Its name is an acronym for "Dual optical sensors in an Ellipsoid Glass for Gen2". By using two 20.3 cm PMTs, D-EGG has an angular sensitivity of almost 4π, resulting in an improved identification of down-going muon tracks. While providing less directional information than M-DOM, its elongated ellipsoidal shape requires smaller drilling holes. Another

5 Most permanently disabled DOMs died during their deployment or commissioning.

advantage of the design is that the glass shell is thinner at the top and bottom, as can be seen in Figure 10. This reduces the probability that a photon is absorbed by the glass shell. Combined with more sensitive PMTs, a higher photon acceptance is thus obtained [44].

WOM The wavelength shifting optical module (WOM) is the fourth and final proposed module. Unlike P-DOM, M-DOM or D-EGG, the area over which photons are collected does not correspond to the size of the photocathode. The main component of the design is a tube which confines incident photons by total internal reflection. Typical proposed dimensions are 90 cm long tubes which have a diameter of 9 cm. By coating the tube with wavelength shifting paint, the WOM would be sensitive to a larger part of the UV spectrum ($\lambda = 250 - 400$ nm) than the other designs. Considering the Cherenkov spectrum $\propto \frac{1}{\lambda^2}$ in combination with the large collection area, a WOM will observe significantly more photons than the other DOMs. These photons are collected by two small PMTs, one at each end of the tube [20]. While the design could save 10 million US dollar on drilling costs [42], a downside is that it has a reduced time resolution compared to the other DOMs[6] [20].

6 The full width at half maximum is ~ 10 ns for a homogeneous illumination [20].

4 THE HESE FILTER

Background vs signal rate

IceCube is triggered more than two hundred million times per day. Only on the order of three hundred of those events are caused by atmospheric neutrinos. The rate at which events induced by cosmic neutrinos occur is even lower, at about one event per month [44]. Observing a cosmic neutrino flux thus requires filters capable of removing these large backgrounds, while retaining nearly all signal events. IceCube's main channel via which it achieves this objective is the High-Energy Starting Event (HESE) analysis. As indicated by its name, the HESE filter was designed to look for high-energy events which start emitting light inside the detector volume. Only neutrinos can enter the detector without emitting light while doing so. By rejecting events in which the first light was emitted at the edge of the detector, the HESE analysis aims at producing a pure sample of cosmic neutrinos.

4.1 VETO-LAYER

Definition veto-layer

Implementing the HESE filter requires specifying which DOMs are part of the 'edge of the detector', correspondingly dubbed the veto-layer. Figure 11 shows a graphical representation of IC86, in which the veto-layer is indicated by a grey band. As IceCube's main background consists of down-going atmospheric muons, a 90 meter layer at the top of the detector is included to tag these events. To identify muons entering the side of the detector at an inclined angle, all strings which are at the edge of the detector are also included. This single layer of strings might however fail to catch down-going muons entering the detector at the height of the dust layer. For this reason, DOMs in an 80 meter region below the dust layer are also included in the veto-layer. Finally, DOMs in the bottom 10 meter of the detector are included to reject events induced by large stochastic energy losses of muon tracks passing below the detector[1].

Charge threshold

The choice of the veto-layer represents a trade-off between the energy at which the filter becomes effective and the fiducial volume of the detector. As low-energy events could sneak past the veto-layer unnoticed, a charge threshold is imposed above which the filter becomes effective at removing background. This charge threshold can be decreased by enlarging the size of the veto-layer. While this would increase the amount

1 Due to small differences in the deployed depth of each string, 17 strings have no DOMs which contribute to the bottom veto-layer. As multiple of these strings are adjacent, the bottom veto-layer effectively contains multiple holes. A more robust selection mechanism would thus be to include the DOM at the bottom end of each string.

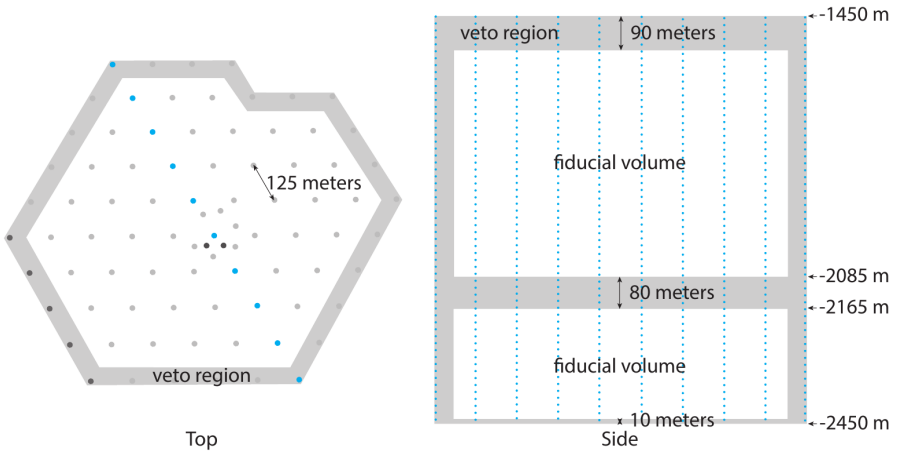

Figure 11: Arrangement of DOMs in IC86. The veto-layer in this configuration is indicated by a grey band. In the right image, a side view is shown of the strings that are coloured blue in the left image [28].

of low-energy events that can be observed, more neutrino induced starting events would then wrongfully be rejected as a result of having their interaction vertex inside the veto-layer.

The configuration of the veto-layer thus depends on the analysis that is applied. HESE uses the configuration shown in Figure 11. Analyses looking for medium and low-energy starting events, accordingly named MESE (10 TeV< E <1 PeV) and LESE (100 GeV< E <10 TeV), use a more extended veto-region [33]. Alternatively, a variable veto-layer can be used based on the reconstructed properties of the event. A new filter for track like events, called Enhanced Starting Track Event Selection (ESTES), is currently under development. ESTES estimates the probability that the event corresponds to an incoming muon track based on the number of deposited hits along various proposed track elements [28]. Preliminary studies show that ESTES will increase IceCube's sensitivity to CC interactions of ν_μ coming from the southern hemisphere.

MESE & LESE

ESTES

4.2 HESE FILTER ALGORITHM

To tag starting events, the HESE filter uses the following procedure. First, the vertex time t_v of the event is defined as the minimal value of t for which the total charge deposited in the detector during the time window $[t-3\mu\text{s}, t]$ exceeds 250 photoelectrons (pe). In the second step, the total charge Q of the N HLC hits that occurred in the VETO-layer during the time interval $[t_v-3\mu\text{s}, t_v]$ is counted. If either $N \geqslant 3$ or $Q \geqslant 3$ pe, the event is tagged as an incoming track. To remove low-energy background muons, which can pass the veto-layer while depositing less than

Event rejected if $N \geqslant 3$ or $Q \geqslant 3$ pe

three hits, an overall charge threshold $Q_{tot} > 6000$ pe is imposed[2]. Low-energy background events can however surpass this charge threshold if their muon track happens to induce a large stochastic energy loss close to one of the DOMs. To prevent these 'balloon' events from sneaking in the sample, the calculation of Q_{tot} only sums the charge of DOMs for which Q_{DOM} is smaller than 50% of the total charge of the event.

54 events in 4 years

Applying this HESE algorithm to four years of IceCube data has produced a dataset of 54 events. Figure 12 shows the energy distribution of those events[3], combined with the best signal plus background fit. The remaining background mainly consists of entering muon tracks (12.5 ± 5.1 events) and atmospheric neutrinos ($9.0^{+8.0}_{-2.2}$ events), which together dominate the distribution below 60 TeV [64]. At higher energies, a clear excess over the background becomes visible. This excess is attributed to a cosmic neutrino flux, for which the best-fit power law spectrum[4] was found to be

$$E^2\phi(E) = 1.5 \cdot 10^{-8} \left(\frac{E}{100\ \text{TeV}}\right)^{-0.3} \text{GeV cm}^{-2}\ \text{s}^{-1}\ \text{sr}^{-1} \quad [32]. \tag{6}$$

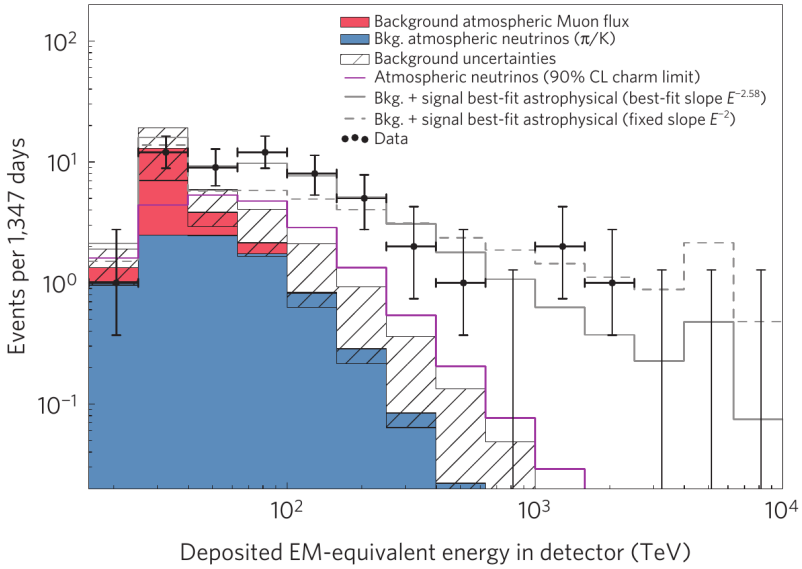

Figure 12: Energy distribution of events satisfying the HESE filter in four years of IceCube data. Filled histograms indicate the estimated number of background events. The best-fit power law spectra are shown as grey lines [19].

2 For all types of neutrino induced starting events, the imposed charge threshold of 6000 pe effectively corresponds to an energy threshold $E_\nu \gtrsim 30$ TeV.

3 The arrival direction of the 54 HESE events correspond to that of n° 1 to 54 in Figure 6.

4 The normalisation of $\phi(E)$ corresponds to the flux of $\nu_l + \bar{\nu}_l$ with flavour l, assuming equal flavour ratio's $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$ [19].

4.3 USING THE EARTH AS A FILTER

An alternative method to identify neutrino induced events is to select events containing an up-going track. Only neutrinos can pass through the Earth without being absorbed. This ensures that the selected sample will be free from charged particles created in cosmic ray air shower, which are the main background in the HESE analysis. In addition, muon neutrinos need not interact inside the detector volume for the muon to be detected. An interaction volume significantly larger than that of IC86 can thus be used.

Restricted by absorption of ν in the Earth

Using the Earth as a cosmic ray filter however comes at a price. Requiring the event to contain an up-going track effectively restricts the analysis to CC interactions from ν_μ and $\overline{\nu}_\mu$. Another important restriction is that at high energies ($E_\nu > 100$ TeV), Earth neutrino absorption can no longer be neglected. A detailed calculation describing this effect is given in Section 6.3.2. Finally, there is no way to determine if a single up-going muon originated from an atmospheric or cosmic neutrino. In contrast, the HESE filter is able to identify atmospheric neutrinos coming from the Southern hemisphere. This is done by observing muons that were created in the same air shower and are travelling parallel to the neutrino. Identifying a cosmic neutrino flux coming from the Northern hemisphere is however only possible by looking for an excess of events above the atmospheric background.

A search for up-going muons using six years of IceCube data resulted in a sample of 350.000 events. Figure 13 shows the energy distribution of those events, along with the expected atmospheric background component and best-fit astrophysical neutrino power law spectrum

$$E^2\phi(E) = 0.90^{+0.30}_{-0.27}\cdot 10^{-8}\left(\frac{E}{100\ \text{TeV}}\right)^{-0.13\pm0.13}\ \text{GeV cm}^{-2}\ \text{s}^{-1}\ \text{sr}^{-1}\quad [37]. \tag{7}$$

Results differ from HESE analysis

Naturally, this flux $\phi(E)$ only accounts for muon and anti-muon neutrinos. When comparing it to the HESE flux in equation (6), a discrepancy can be seen between the two results. Currently, the disagreement is not statistically significant, though the difference in analysis techniques might indicate that there is a physical origin. Whereas the Earth based filter is only sensitive to neutrinos from the Northern hemisphere, the HESE analysis averages over the full 4π solid angle. As such, the higher flux in the HESE analysis might be due to neutrinos which have a galactic origin[5]. Another noteworthy difference is that the filters are sensitive to different energy ranges, respectively 30 TeV $< E <$ 2 PeV for the HESE filter and 200 TeV $< E <$ 9 PeV for the Earth based filter. The discrepancy could thus be an indication for a spectral break in the cosmic

5 The galactic centre is located in the Southern hemisphere.

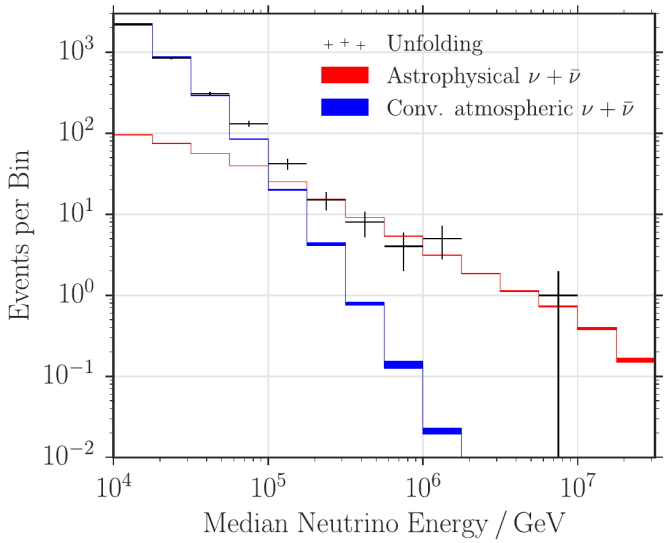

Figure 13: Energy distribution of up-going muons in six years of IceCube data. Solid lines indicate the estimated background from atmospheric neutrinos (blue) and the best-fit power law spectrum to the astrophysical component (red), which starts dominating at $E > 100$ TeV [37].

neutrino spectrum.

Apart from analysing the energy spectrum, the propagation direction of muons in the data sample was also analysed to search for astrophysical neutrino sources. As in the case of the HESE analysis, no statistically significant clustering of events was found. The data did however contain the highest energy (astrophysical) neutrino observed to date. A muon with a reconstructed energy $E_\mu = (4.5 \pm 1.2)$ PeV was observed, leading to an expected muon neutrino energy of 8.7 PeV. Given its high energy, the p-value for the event to have an atmospheric origin was found to be less than 0.005% [37].

Observation most energetic neutrino

4.4 HESE WITH INCREASED STRING SPACING

The HESE analysis has been IceCube's main channel to discover and parametrise the cosmic neutrino flux. When applying the same algorithm to events in the proposed IceCube-GEN2 high-energy extension, performance will suffer due to the increased spacing in between strings. A data-driven analysis was therefore designed by N. van Eijndhoven and S. Toscano, which studies the effect of enlarging the string spacing in the IC86 geometry. While deployed strings cannot be physically moved, IC86 data can be reduced to correspond to that of a detector with a string spacing of 250 meter. This is done by removing data recorded in over 75% of all strings, as illustrated in Figure 14.

21-string geometry

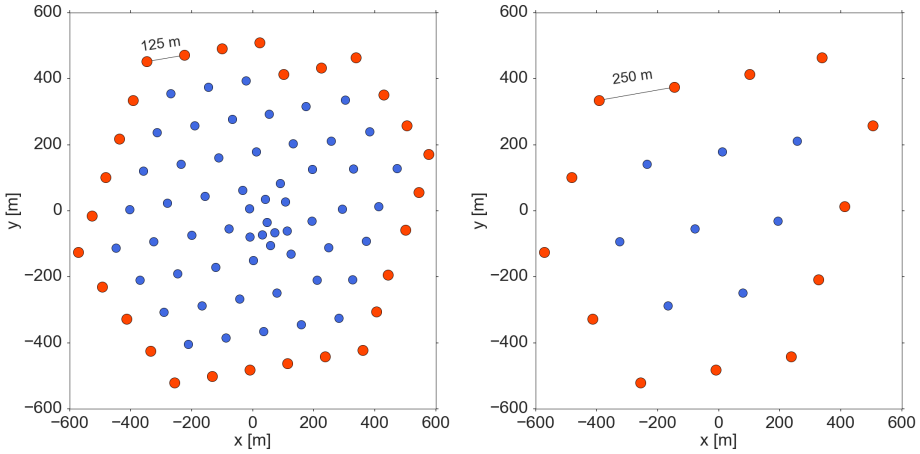

Figure 14: Reduction from the IC86 geometry (left) to a geometry with a 250 meter string spacing (right), consisting out of 21 strings.

The efficiency of the HESE filter for this reduced geometry can now be determined using IC86 data. Events[6] are first analysed using the full 86-string geometry. This allows determining the charge of the event and whether or not it passes the HESE filter. If an event was rejected by the filter, meaning it was identified as an incoming track, it is then re-analysed in the reduced geometry. By applying the HESE filter again on data from the 21 remaining strings, the number of background events which pass the veto-layer unnoticed can be determined. Figure 15 shows the number of background events passing the filter (green squares) and the total number of background events (blue dots).

To compare this passing fraction to that of the full IC86 geometry (purple triangles), the latter must first be determined. This was done by defining a second veto-layer, located at the edge of the fiducial volume. Events which were identified as incoming tracks by the outer veto-layer (blue dots), should likewise trigger the second inner veto-layer. By counting the number of events for which this is not the case (purple triangles), the passing fraction of the 125 meter string geometry was determined. Comparing the two results (green-purple) shows that the number of events wrongfully passing the filter increases by ~2 orders of magnitude when doubling the string spacing to 250 meter. This increase can be reduced by using a more extended veto-region. When including 12 DOMs in the top veto-layer (instead of 6) and re-including the 13 strings in between the outer-layer strings (similar to the Edge-weighted geometry), the passing fraction (red diamonds) decreases by ~1 order of magnitude.

6 Only events with a total charge $Q > 1500$ pe are considered, as the charge threshold of HESE in IC86 is already 6000 pe.

Increased charge/energy thresholds

From Figure 15, it follows that the charge threshold at which the number of background events passing the filter goes to zero increases from 6.000 to 20.000 pe when doubling the string spacing. This implies that the minimum energy from which the filter becomes sensitive increases from 30 TeV to 100 TeV. In case a denser, more extended veto-region is used, the charge threshold lies at 13.000 pe, corresponding to 60 TeV. It can thus be concluded that increasing the string spacing to 250 meter considerably reduces the self-vetoing capability of the detector. Improvements can be obtained by extending the veto-region, though this comes at the cost of reducing the fiducial volume, which again leads to a loss of signal. Lowering the charge threshold without losing signal thus requires considering different filter algorithms altogether. This search for new InIce self-veto techniques is the focus of the next chapter.

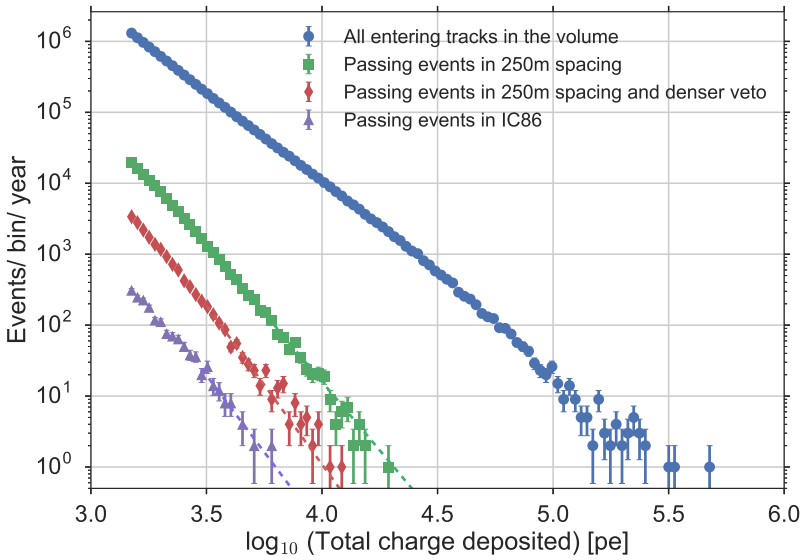

Figure 15: Number of events wrongfully accepted by the HESE filter when using the full IC86 geometry (purple triangles), the reduced 21-string geometry (green squares) or the reduced 34-string geometry with an extended veto-layer (red diamonds). The total charge distribution of all background events on which the filter was applied is indicated by blue circles.

5 DEVELOPING A NEW CAUSAL FILTER

5.1 GOAL

The HESE algorithm has proven the value of self-veto techniques in identifying neutrino induced events. When moving to a new detector geometry, applying the same algorithm might however no longer be feasible. For instance, the time window in which the HESE filter searches for hits in veto-layer (Δt =3 μs), was chosen to correspond to the average time it takes for a muon to traverse the detector. In this way, all hits which occurred at the start of the event are included, whereas most noise related hits and those occurring at the end of the event are excluded. Because the IceCube-GEN2 volume will be roughly a tenfold that of IC86, a more extended time window will be required. Combined with the increased number of DOMs belonging to the veto-layer, this will result in an increased contribution from noise hits, making it harder to identify incoming tracks. To paraphrase, a new selection mechanism is required to identify those veto-layer hits which are due to an entering track.

HESE in extended volumes

Another distinction between IC86 and IceCube-GEN2 is that to reduce costs, the latter will not include cables that allow adjacent DOMs to communicate with each other. GEN2 DOMs will therefore be unable to identify Hard Local Coincidence (HLC) hits. In IC86, a waveform is only read-out by one of the ATWDs, if the hit received an HLC tag. This option will thus no longer be available in GEN2, requiring all hits to be fully read-out. Alternatively, the principle of self-coincidence can be used, which is said to occur if one DOM observes multiple hits during a time window Δt. The feasibility of this technique will however largely depend on the type of DOM that will be used for the GEN2 high-energy extension. It is therefore more natural to move to software-based hit cleaning algorithms, abandoning hit cleaning based on real-time coincidence tags altogether. These algorithms were already developed for IC86 and can easily be extended to other geometries.

Lack of HLC

In addition to the larger volume and lack of HLC, the performance of the HESE filter will also suffer due to the larger string spacing (see Section 4.4). Self-veto techniques should however still be preferred over analyses restricted to the Northern hemisphere for cosmic neutrino searches. Earth absorption based filters are only sensitive to events containing track signatures, suffer a major loss of signal events at high-energies and are unable to distinguish between atmospheric and astrophysical neutrinos. For this reason, we examine new innovative self-veto

A new InIce filter

techniques as an alternative to the HESE filter. Different filter algorithms are first explored by applying them to IceCube data. This ensures that their resilience is tested using all types of background events. From these candidates, the best performing algorithm is ensuingly selected. Chapter 6 then discusses the use of simulated data to test the performance of this algorithm in the IceCube-GEN2 geometry. This will allow not only to optimise the filter parameters, but also to compare the performance for the different proposed geometries.

5.2 FILTER ALGORITHM

Parameters:

Distinguishing incoming tracks from starting events largely comes down to determining whether or not the first hits of the event occurred at the edge of the detector. The notion "edge of the detector" is explicated by defining a veto-layer of DOMs, which surrounds the inner/fiducial volume. In case the event consists of an entering track, hits in the veto-layer are either related to

- the muon passing through the detector,
- dark noise.

In an ideal scenario, a hit h occurring at $(\vec{x}_h, t_h)$ would indicate the position $\vec{x}_m$ of the muon m at time t_h. In reality, DOMs are not located exactly along the muon track, meaning that the propagation time of photons in ice has to be accounted for. When also accounting for the stochasticity of the light emission processes combined with the detector timing resolution, the relation softens to

$$\frac{\|\vec{x}_h - \vec{x}_m\|}{c} - |t_h - t_m| = (\Delta t)_c \quad , \tag{8}$$

where $(\Delta t)_c$ is the variation on the ideal causal connection due to the aforementioned effects. Equation (8) now can be used to select those veto-layer hits that are related to the muon track. This is done by imposing the relation $|(\Delta t)_c| < (\Delta t)_{max}$, where $(\Delta t)_{max}$ is a fixed upper limit imposed on the causal connection. As the true coordinates of the muon are unknown when using data, an estimator $(\hat{\vec{x}}_m, t_m)$ must be used instead. Once the veto-layer hits related to the event are selected, they are divided in two groups.

$(\Delta t)_{max}$

$(\hat{\vec{x}}_m, t_m)$

- Hits occurring when the muon enters the detector.
- Hits occurring when the muon exits the detector.

This subdivision is made based on whether the time of the hit t_{hit} occurred before or after a reference time t_{ref}. In Section 5.3, various candidates for t_{ref} and $(\hat{\vec{x}}_m, t_m)$ are examined. For now, we suppose that these variables have been defined and that the number of veto-hits at the start of the event N_{veto} can thus be calculated. An event is then labelled as background if N_{veto} equals or exceeds a predefined value N_{min}.

t_{ref}

N_{min}

Alternatively, a cut on the fraction of hits[1] $\frac{N_{veto}}{N_{total}} \geqslant x_{min}$ could be used instead, as high-energy events are expected to deposit more charge in the veto-layer. This prediction however does not always hold for high-energy muons $\left(E_\mu > 10\ TeV\right)$, which lose most of their energy via stochastic processes. Entering muons can thus pass through the veto-layer without having large stochastic energy losses. These events will therefore have a ratio $\frac{N_{veto}}{N_{total}}$ which is disproportionally small compared to the nominal value. For this reason, the absolute value of N_{veto} is used instead. In addition to the criterion $N_{veto} \geqslant N_{min}$, an event is also rejected if $\hat{\vec{x}}_m$ lies inside the veto-layer. This choice is made because $\hat{\vec{x}}_m$ will be chosen to lie close to either the start or centre of the muon track. The complete filter procedure can now be summarised as follows.

1. Clean events by removing balloon DOMs[2].
2. Find an estimate for the muon's position $\hat{\vec{x}}_m$ at time t_m.
3. Count the number of veto-hits that are causally connected to $\left(\hat{\vec{x}}_m, t_m\right)$, i.e. $|(\Delta t)_c| < (\Delta t)_{max}$, and happened before t_{ref}.
4. Reject the event if $N_{veto} \geqslant N_{min}$ or if $\hat{\vec{x}}_m$ lies in the veto-layer.

Filtering signal events

Apart from rejecting background, the filter should also focus on retaining as much signal as possible. This can be achieved by choosing a sufficiently small value for $(\Delta t)_{max}$ and placing t_{ref} close to the start of the event. The first choice is made to reduce the number of noise hits N_{noise} contributing to N_{veto}, as theoretically one can expect $N_{noise} \propto (\Delta t)_{max}$. The second choice is made to ensure that the contribution to N_{veto} from particles produced inside the detector is minimal. Naturally, the inverse goes for the efficiency with which background events can be rejected. In that case, larger values of $(\Delta t)_{max}$ and later times for t_{ref} are preferred, to ensure that all veto-hits related to the entering track are accounted for.

1 Similarly, the fraction of charge $\frac{Q_{veto}}{Q_{total}}$ could be used as well.

2 This step is analogous to the first step of the HESE algorithm (see Section 4.2). In IC78, i.e. neglecting DeepCore DOMs, balloon DOMs are present in 1% of all events with $Q_{total} > 1500$ pe. Not removing them results in a significant increase in the number of passing background events.

5.3 SELECTING A REFERENCE TIME AND POSITION

5.3.1 *COG-based hit selection*

Consider first the selection of a reference position $\hat{\vec{x}}_m$. A natural choice would be to use the centre-of-gravity (COG) of the event, defined as the average charge weighted position of all hits

$$\vec{x}_{\text{COG}} \equiv \frac{1}{Q_{total}} \sum_{i=1}^{N_{hits}} Q_i \cdot \vec{x}_i \quad . \tag{9}$$

For this estimator to be unbiased, DOMs should be distributed uniformly along each of the three spatial axes. In IceCube, this is not the case due to the higher density of DOMs in the DeepCore region. The COG is therefore calculated based on hits which occurred in the IC78 geometry, thus effectively neglecting DeepCore hits. Choosing $\vec{x}_{\text{COG}}$ as estimator for the reference position $\hat{\vec{x}}_m$, the time t_m at which the muon track passed closest to $\hat{\vec{x}}_m$ subsequently needs to be estimated. Unlike for the position, the average weighted value of all hit times $\overline{t_i}$ is not a proper estimator for t_m. This is due to the effect of afterpulses, which cause $\overline{t_i}$ to be shifted towards the end of the event.

$\hat{\vec{x}}_m = \vec{x}_{\text{COG}}$

A single hit as estimator

Consider therefore instead a single hit $(\vec{x}_h, t_h)$ as an estimator for $(\vec{x}_m, t_m)$. If h is chosen to be the first hit in a given DOM, it is certain that t_h is unaffected by afterpulses. An added advantage is that, of all muon-induced hits in a DOM, the first one provides the best estimate of t_m. This is due to the fact that no muon-induced hits can occur before the muon has its closest approach to the reference DOM. In case the first hit in the selected DOM was due to dark noise, t_h will however be a bad estimator for t_m. To prevent this from happening, the reference DOM needs to be selected with care and events need to be cleaned from dark noise hits before the filter is applied.

Two COG-based reference hits

Setting h to be the first hit of a given DOM, the question now remains which DOM to select. While choosing $\hat{\vec{x}}_m = \vec{x}_h$ prevents setting $\hat{\vec{x}}_m = \vec{x}_{\text{COG}}$, h can be chosen such that it is related to the COG. More specifically, we consider the following two COG-based reference DOMs.

- $d_{\text{COG},1}$: The DOM closest to the COG.
- $d_{\text{COG},2}$: The DOM with the highest charge within 150 meter of the COG.

When using either $d_{\text{COG},1}$ or $d_{\text{COG},2}$, $(\hat{\vec{x}}_m, t_m)$ is thus set equal to the coordinates of the first hit in that DOM $(\vec{x}_h, t_h)$. Apart from serving as an estimator for t_m, t_h can at the same time be used to set the value of t_{ref}. Because t_h approximates the time at which the muon was halfway through the detector, all hits related to the muon entering or exiting the

detector should respectively occur before or after t_h. Since this is exactly the behaviour desired from t_{ref}, we set

$$t_{ref} = t_m = t_h \quad . \tag{10}$$

With the choice for $(\hat{\vec{x}}_m, t_m)$ and t_m specified, the only remaining free parameters of the filter are $(\Delta t)_{max}$ and N_{min}. To allow identifying incoming tracks which deposit only a few hits when entering the detector, N_{min} should preferably be chosen as small as possible. The causal constraint $(\Delta t)_{max}$ on the other hand should offer an improved performance compared to the HESE filter, restricting its value to be smaller than 3 μs.

While the above discussion has focused on correctly identifying incoming muons, the chosen parameters are also favourable for accepting starting events. In starting events, the COG will lie close to the neutrino interaction vertex[3]. As a result, the first photon induced hits of the event will occur in DOMs located around the COG. This ensures that, if the centre of the cascade is located inside the fiducial volume, all photon induced veto-layer hits should occur after t_h. Starting events are thus only rejected if their cascade centre is not located inside the fiducial volume, or if a least N_{min} noise hits causally connected to $(\vec{x}_h, t_h)$ happen to occur in the veto-layer.

5.3.2 *T90-based hit selection*

Alternative hit selection method

In the previous section, the average weighted position of all hits was used to select a hit that reflects the coordinates of a point on the muon track. An alternative way to go about selecting a muon induced hit, is to look for a large number of hits in a small time window. We therefore develop a new hit selection technique, which is based on the T90 procedure[4] used in GRB searches [47].

X_{T90} & $(\Delta t)_{\mathrm{T90}}$

The T90 reference hit is defined as the earliest hit h for which at least X_{T90} percent of the total charge was deposited during the time interval $[t_h - (\Delta t)_{\mathrm{T90}}, t_h]$. Obviously this also determines the corresponding DOM d_{T90} and consequently the hit position. Both X_{T90} and $(\Delta t)_{\mathrm{T90}}$ are tunable parameters, which together determine how far along the muon track the reference hit occurs. As in the case of the COG-based selection, we can again set $t_{ref} = t_h$. This restricts X_{T90} to small values, to prevent photon induced veto-layer hits from occurring before t_{ref} in case of signal events. For background events, small values of X_{T90} at the same

3 In case the light emitted by a cascade is not fully contained by the detector, the COG will be biased towards the centre of the detector.

4 In GRB analyses, the T90 parameter denotes the time in between which the observed fluence increases from 5% to 95% of the total. As such, the T90 parameter provides a measure for the duration of a GRB.

time allow rejecting incoming tracks based on whether or not the reference hit is located in the veto-layer.

Standard values: $X_{\mathrm{T90}} = 10\%$ and $(\Delta t)_{\mathrm{T90}} = 1\mu s$

The choice of X_{T90} and $(\Delta t)_{\mathrm{T90}}$ represents a trade-off between minimising the contribution from noise hits and selecting values for which a reference hit exists for all events. In IC86 events, all photon induced hits will typically occur within a time period smaller than 5 μs. This implies that during the event, at least one 1 μs time window will contain more than 20% of the total number of hits. To account for non-photon induced hits and to ensure that the selected hit lies at the start of the event, $X_{\mathrm{T90}} = 10\%$ and $(\Delta t)_{\mathrm{T90}} = 1\mu$s will be used as standard parameters for IC86 events.

5.4 APPLICATION ON ICECUBE DATA

5.4.1 *Data versus simulation*

With the filter algorithm fully specified, we can now move on to testing its performance for different parameter choices. Usually, this is done by applying the filter to two simulated datasets, one signal and one background, with the goal of retaining all events in the former, while rejecting all events in the latter. Using real data should however be preferred, as the properties of simulated events are always dependent on the underlying physics models used to generate them. However, this requires a second filter algorithm, which can already separate signal from background events. In our case, this is the role played by the HESE algorithm. Because the output of the HESE filter will contain false positives and true negatives, requiring that both filters produce the same output is too strong a constraint. We therefore instead accept the output from the causal filter if it agrees with the HESE veto-value and look on an event-by-event basis if the filter values are different and $Q_{tot} > 6000$ pe.

HESE veto-value as reference

As dataset, a subset of all the IceCube level-2 events recorded during 2011 is used. The generation of this subset was based on the following two criteria. First, low-energy events were removed by requiring that the total charge of the event is larger than 1500 pe. The second criterion is that events are only included if they were recorded when the detector was running stably. Applying this selection resulted in 9.345.205 events, 266.009 of which have a total charge larger than 6000 pe.

Dataset of 9.345.205 events

5.4.2 *Analysis of a single event*

Time distribution

To illustrate how events can be classified by eye, consider the time distribution of a typical background (left) and signal (right) event shown in Figure 16. Indicated in blue is the set of all hits, which is compared to the subset of hits that occurred in the veto-layer (green). In case of

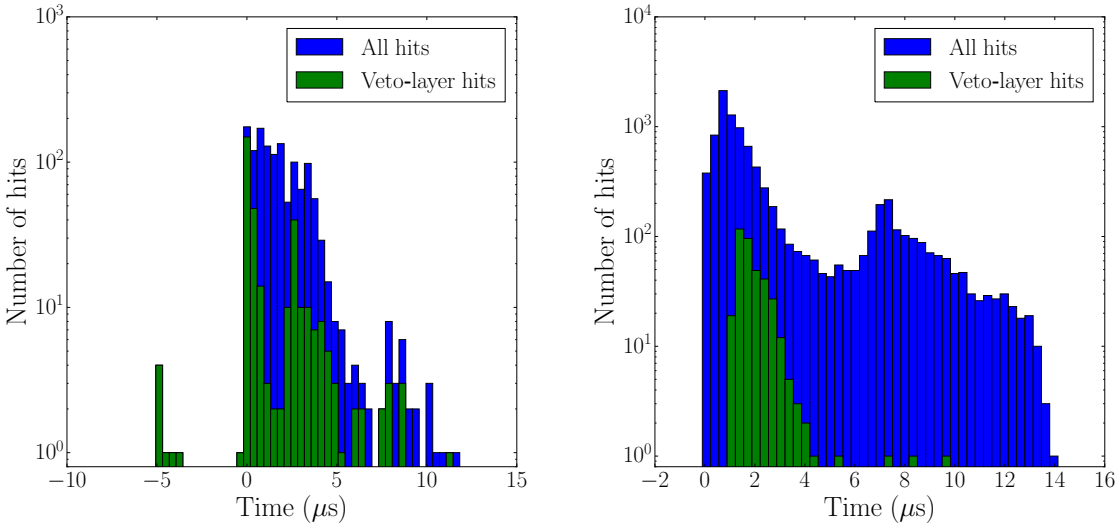

Figure 16: Comparison between the time distribution of a typical background (left) and signal (right) event. $t = 0$ corresponds to the time at which h_{T90} occurs.

the background event, the first hits (around -5 μs) are unrelated to the event. This indicates the necessity of imposing a causal constraint on veto-layer hits, to prevent rejecting events based on noise. The distribution of veto-layer hits then has two clear peaks, corresponding to the muon entering and exiting the detector. In case of the signal event, the first veto-layer hits occur 1 μs after the start of the event, indicating that the event is not due to an entering particle. An interesting feature present in both distributions, is the large number of hits occurring after $t = 5$ μs. These are non-photon induced hits, resulting from afterpulses.

Event viewer

An alternative technique that can be used to classify events, is to look at the event in an event viewer. Figure 18 shows the x-y (top) and x-z (bottom) distribution of hits from the background event shown in the left panel of Figure 16. Hits are displayed as blue spheres if they did not occur in the veto-layer and purple spheres otherwise. Inspecting Figure 18 immediately reveals that the event corresponds to a down-going[5] muon. The COG and three reference hits are also indicated, to visualise the procedure via which the reference hit is selected. It can be seen that, in this case, all three reference hits have a position close to the muon track. Spatially separated from the muon track, are a number of hits occurring at $(x, y, z) \approx (-400\ \mathrm{m}, -400\ \mathrm{m}, 400\ \mathrm{m})$. These hits correspond to those at $t \approx -5$ μs in Figure 16, indicating clearly that they are unrelated to the through going muon. Visualising events thus provides a powerful tool to distinguish starting events from incoming tracks. The disadvantage of

5 Because d_{T90} is selected in such a way that it lies at the start of the event, the muon must be down-going.

this method is that it can only be applied to a small number of events, due to the time required to analyse a single event.

$(\Delta t)_c$ distribution

Apart from plotting the time distribution and position of all hits, the distribution of the causal relation $(\Delta t)_c$ of all hits to the reference hit can also be used to classify events. Figure 17 shows the cumulative distribution of $(\Delta t)_c$ for one thousand background events and the T90-based hit selection. Indicated in green is the subset of hits which occurred inside the veto-layer before t_h. For an event to be rejected[6], the sum of the green bins inside the range $[-(\Delta t)_{max},(\Delta t)_{max}]$ should equal or exceed N_{min}. As expected, the distribution peaks around $(\Delta t)_c = 0$, indicating that the T90-based reference hit is a good estimator for the muon's position. To the right, the distribution extends to the maximal distance attainable between two DOMs. To the left, a significant number of hits can again be seen to occur due to the effect of afterpulses.

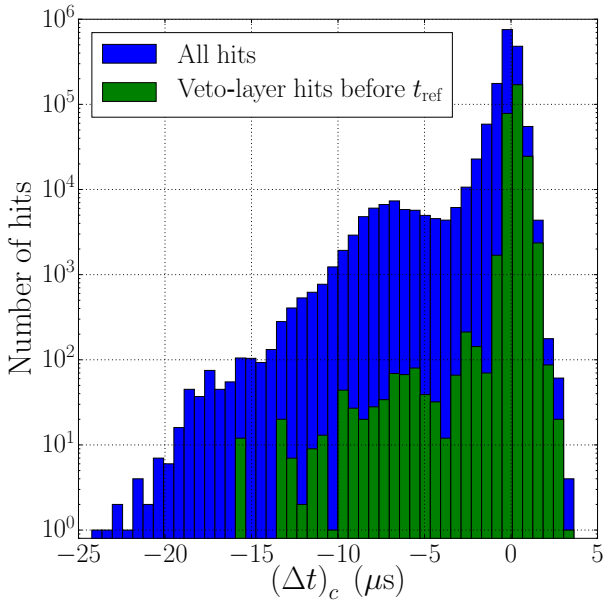

Figure 17: Combined $(\Delta t)_c$ distribution of one thousand background events using the T90-based reference hit.

6 Assuming $\vec{x}_h$ is not located inside the veto-layer.

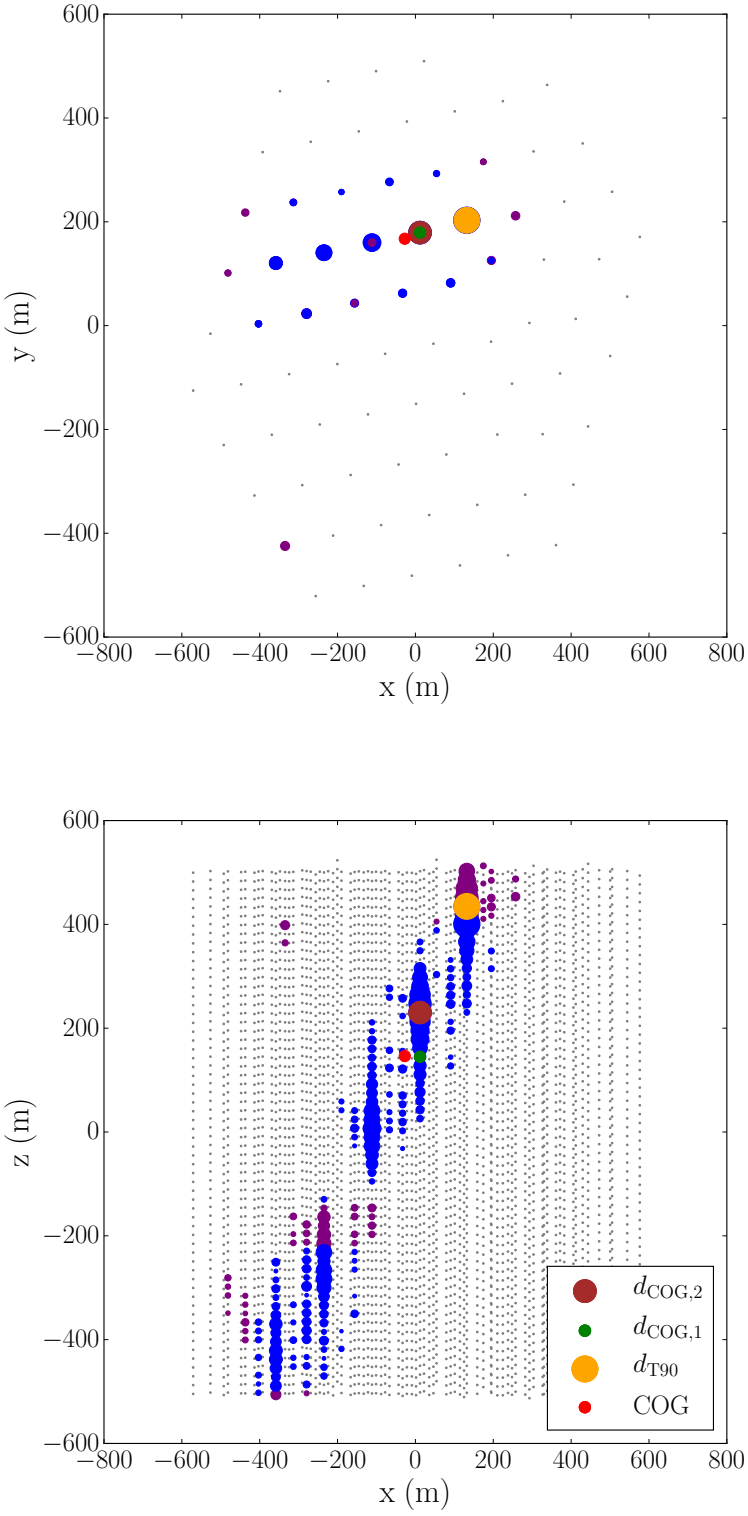

Figure 18: Visualisation of the background event shown in the left panel of Figure 16. DOMs are drawn as grey spheres. Those containing hits are coloured blue (purple if they lie in the veto-layer) and have a volume that is proportional to the number of hits they received.

5.4.3 *Comparison between COG_1, COG_2 and T90*

20 HESE events

The 2011 data sample contains twenty HESE events. One of these, when observed in the event viewer, can readily be classified as background. In this event, a low-energy muon first traverses the detector, depositing two veto-layer hits upon entering. Ensuingly, a higher energy muon passes through the detector, pushing the combined charge of both muons (Q_{tot} = 6036 pe) just above the HESE charge threshold. For this reason, the 2011 dataset only contains 19 'true' HESE events.

19 'true' HESE events

To analyse if all 19 'true' HESE events can be accepted by the causal filter, consider their combined $(\Delta t)_c$ distribution, shown in Figure 19. As none of the 19 events have a reference hit which lies in the veto-layer[7], they can only be rejected due to causally related veto-layer hits occurring before t_{ref}. In only 2 out of the 19 events, such hits are present. These can be seen as the two groups of two hits, visible in Figure 19. While Figure 19 was made using the T90-method to select a reference hit, the same scenario applies when using either of the COG-based reference hits. That is, in the same two events, two veto-layer hits are present before t_{ref}. The exact values of $(\Delta t)_c$ for these events are given in Table 2. As can be seen, $d_{\mathrm{COG},1} = d_{\mathrm{COG},2}$ in both events.

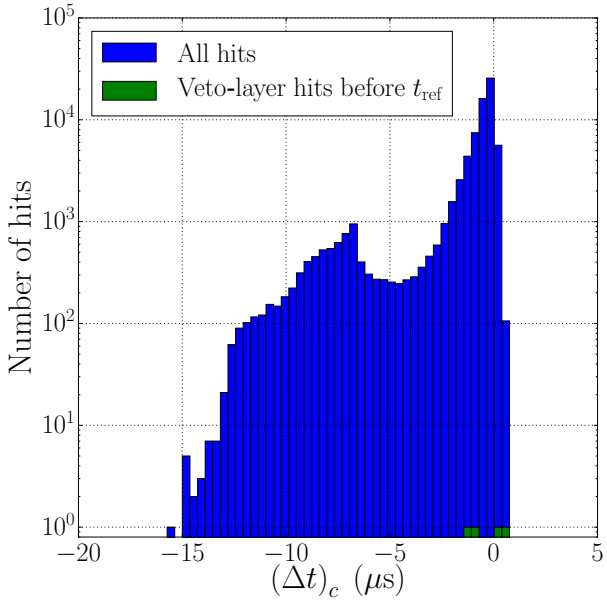

Figure 19: Combined $(\Delta t)_c$ distribution of the 19 'true' HESE events using the T90-based reference hit.

7 This applies to $d_{\mathrm{COG},1}$, $d_{\mathrm{COG},2}$ and d_{T90}.

Table 2: $(\Delta t)_c$ for the two HESE events with early veto-layer hits.

Reference hit	time (μs)			
	Event 1		Event 2	
	hit A	hit B	hit A	hit B
T90	-1.318	-1.077	0.373	0.488
COG_1	-1.279	-1.041	0.376	0.491
COG_2	-1.279	-1.041	0.376	0.491

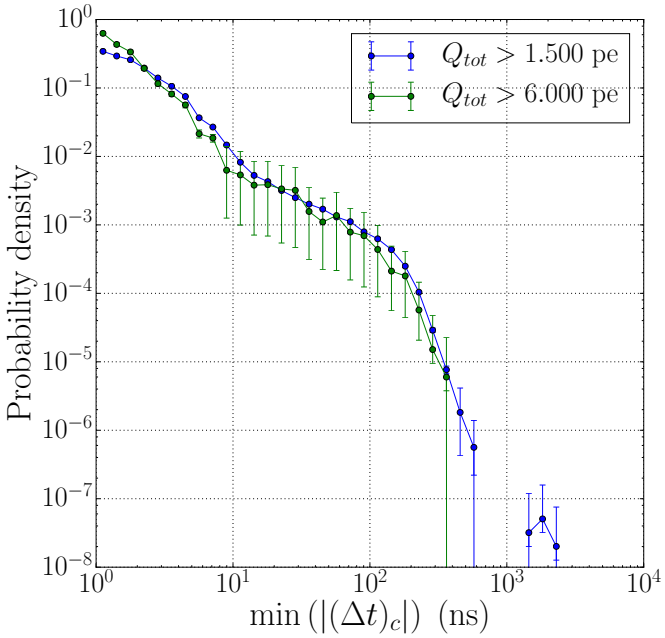

Figure 20: Probability density of $\min[(\Delta t)_{max}]$ for 136.902 background events.

Observing the first event in an event viewer[8] shows that the two veto-layer hits occur at the other side of the detector and are unrelated to the event. As such, the event should be accepted. In case of the second event, the hits are more strongly causally related to the event and occurred close to the reference DOMs. In addition, hits related to the event were detected in the IceTop detector [32]. This second event therefore most likely has an atmospheric origin and should thus be rejected as incoming background. To do so and keep the first event, we need to set $N_{min} \leqslant 2$ and $0.492\ \mu s \leqslant (\Delta t)_{max} \leqslant 1.040\ \mu s$.

18 signal events

Apart from retaining signal, the filter should also reject as much background as possible. For this purpose, large values of $(\Delta t)_{max}$ and small values of N_{min} are preferred. To place a lower-limit on the value of $(\Delta t)_{max}$, consider the probability density shown in Figure 20. On the

$\min[|(\Delta t)_c|]$ *distribution of background events*

8 Event displays of the two events in Table 2 are included as Figures 48 and 49 in Appendix A.

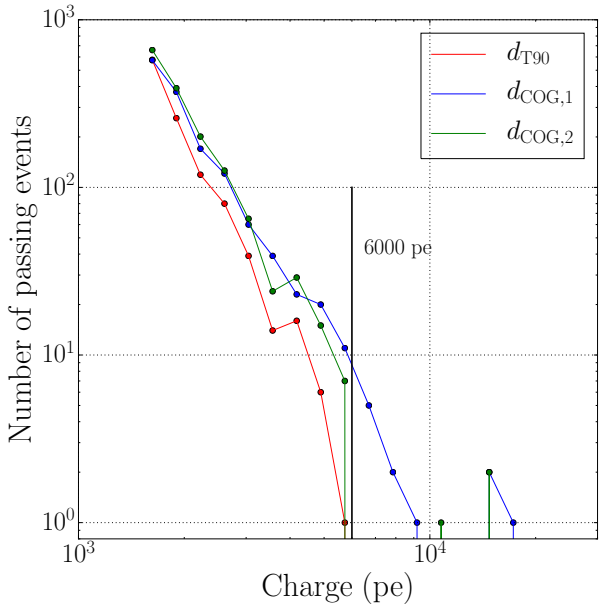

Figure 21: Charge distribution of background events that pass the causal filter for the three choices of reference hits.

x-axis, the minimal value of $|(\Delta t)_c|$ for veto-layer hits occurring before t_{ref} is displayed for 136.902 background events[9]. The y-axis displays the probability density to fall into a given bin, taking into account the width of the bin itself. For comparison, the same distribution is also shown for the 3.932 events which have a total charge larger than 6000 pe. These high-charge events have a lower average value of $|(\Delta t)_c|$, as their probability density systematically lies below that of all events at high values of $|(\Delta t)_c|$. By integrating the probability distribution of Figure 20, we find that

$$P\left(\min[\{|(\Delta t)_c|\,\text{background event}\}] > 1\ \mu\text{s}\right) \leqslant 10^{-5} \quad . \tag{11}$$

Passing background fraction $\mathcal{O}(10^{-5})$

Setting $(\Delta t)_{max} = 1\ \mu$s, we thus get a passing fraction[10] of $\sim 10^{-5}$, which is similar to that of the HESE algorithm [32]. If in addition $N_{min} \leqslant 2$, all of the 18 signal events are filtered[11] and the HESE event for which IceTop hits occurred is rejected. We will therefore use $(\Delta t)_{max} = 1\ \mu$s and $N_{min} = 2$ as the nominal values to compare the filter performance between the T90- and COG-based reference DOMs.

Figure 21 shows the number of passing events as a function of charge, obtained by applying the causal filter on all events which were rejected by the HESE filter. To ensure that none of the passing events above

9 Here, background refers to events tagged by the HESE filter.
10 This is assuming that N_{min} is equal to one.
11 Using the COG $\vec{x}_{\mathrm{COG}}$ and COG-time $\overline{t_i}$ would result in the loss of 2 out of the 18 signal events.

6000 pe were misidentified as background by the HESE filter, an event-by-event analysis was applied to each of them. No additional starting events were found via this procedure, neither using $(\Delta t)_{max} = 1$ μs and $N_{min} = 2$, nor using other filter parameters. All passing events above 6000 pe tagged as background by the HESE filter are therefore true background events that were wrongfully accepted by the causal filter. Taking this into consideration, we see that using the T90-based reference hit provides the best performance. Its curve systematically lies below that of $d_{\mathrm{COG},1}$ and $d_{\mathrm{COG},2}$, and there are no passing background events above 6000 pe. When comparing the two COG-based curves, $d_{\mathrm{COG},2}$ provides the best performance above 6000 pe with only 3 passing background events compared to 12 in the case of $d_{\mathrm{COG},1}$.

T90 rejects most background

To understand the cause for this difference in performance, the passing events were examined in more detail. This revealed that most events were filtered due to noise hits occurring before the start of the event in the COG DOM. A second scenario in which the COG-based filter failed, is for coincident muons with similar energy. In that case, the COG will be located somewhere between the two tracks, resulting in a poorly selected reference DOM. Using the T90 based reference hit solves this issue, as a hit lying along the track of the earliest muon will then be selected. Instances in which the T90 reference hit was due to noise were found, but only for events with a total charge less than 6000 pe.

Noise induced reference hits

An additional reason why the T90-based method outperforms the results obtained using the COG, is that the filter rejects events in which the reference DOM lies in the veto-layer. Like the COG, the COG-based reference hit will lie close to the centre of the charge distribution of the event. Therefore, only muon tracks which do not reach the fiducial volume will be rejected by this criterion. However, in case of the T90-based reference hit, any muon depositing more than 10% of its charge in the veto-layer upon entering the detector will be rejected. Considering this in combination with the performance difference shown in Figure 21, we will therefore from now on use the T90 procedure as the standard.

5.4.4 *Influence of filter parameters*

Fixing the reference DOM to h_{T90}, we now examine the effect of varying $(\Delta t)_{max}$ and N_{min}. First, we test the filter performance for time windows $(\Delta t)_{max} = 0.5$ μs, $(\Delta t)_{max} = 1$ μs and $(\Delta t)_{max} = 2$ μs, keeping N_{min} fixed at 2 hits. Table 2 shows that regarding the 18 true signal events, nothing changes when using the smaller time window. However, when switching to $(\Delta t)_{max} = 2$ μs, one of the 18 signal events is wrongly rejected. An upside to this larger time window is that it reduces the number of passing background events by 12%, as can be seen in the left panel of Figure 22.

Variation of $(\Delta t)_{max}$

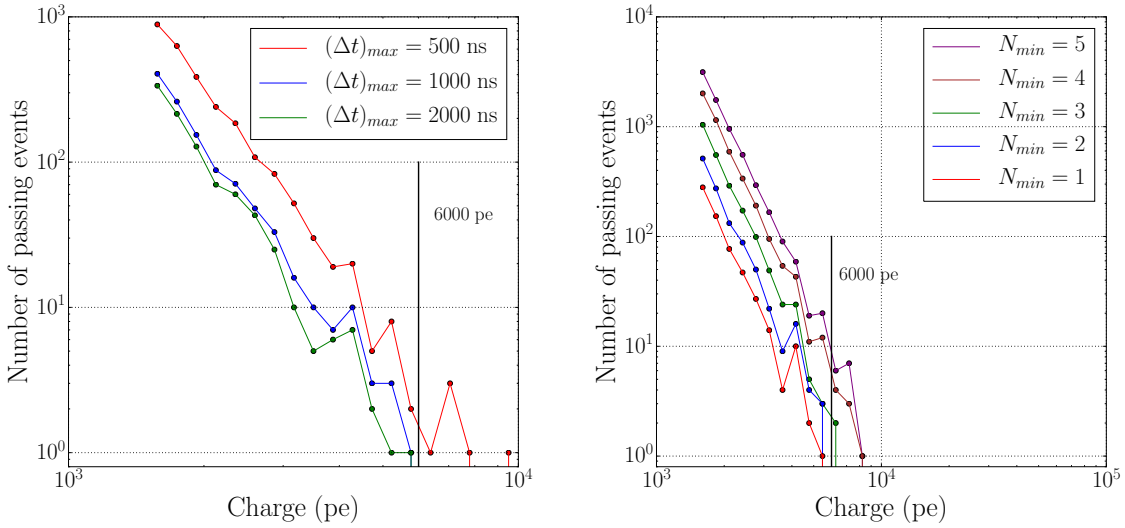

Figure 22: Comparison between the number of passing background events when varying filter parameters. In the left panel, N_{min} is fixed to 2 hits. In the right panel, $(\Delta t)_{max}$ is fixed to 1 μs.

Variation of N_{min}

Next, we fix the value of $(\Delta t)_{max}$ to 1 μs and vary N_{min} from 1 to 5 hits. For each of these five configurations, all signal events are accepted by the filter. The difference in their output thus solely lies in the number of background events that are rejected. Since any event rejected for a given value $N_{min} = A$ is also automatically rejected when $N_{min} \leqslant A$, the largest background rejection will occur for $N_{min} = 1$. For $N_{min} \geqslant 3$, some background events above the HESE charge threshold start to be accepted by the causal filter, as can be seen from the right panel of Figure 22.

The same analysis, varying the value of $(\Delta t)_{max}$ and N_{min}, was repeated with the COG-based reference hits. Appendix A contains the corresponding distributions, equivalent to those shown in Figure 22. While all 18 signal events can again be accepted (see Table 2), none of the filter parameters allows reducing the background to events below 10^4 pe. Therefore, this again demonstrates that using the T90-based hit selection is more favourable than using a reference hit based on the COG.

5.4.5 *Comparison to the HESE filter*

As a final step, we compare the filter-value produced by the HESE algorithm to that from the causal filter. For the latter, $(\Delta t)_{max} = 1$ μs and $N_{min} = 2$ are used as standard parameters, in combination with h_{T90}. Figure 23 shows the charge distribution of all events, compared to that of events passing one of the filters. Above the HESE charge-threshold, the causal filter accepts the 18 true signal events and rejects

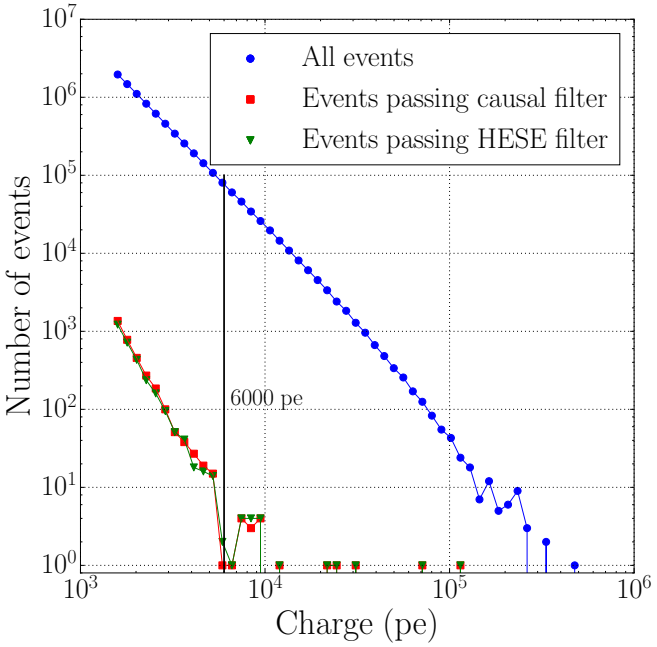

Figure 23: Comparison between the charge distribution of all events in the 2011 data sample and those accepted by the HESE or causal filter.

the two backgrounds events that were wrongly accepted by the HESE filter. A large similarity is also present for $Q < 6000$ pe, where the HESE and causal filter respectively accept 4.329 and 4.808 events. These are mostly the same events, as 3.184 out of them pass both filters. In case $N_{min} = 1$ is used, only 3.376 events pass the causal filter. Of those, 73% is also accepted by the HESE filter.

We thus conclude that the results obtained using the causal filter are consistent with those of the HESE algorithm. In fact, the causal filter can be tuned to reject the two background events passing the HESE filter. Additionally, it can reliably identify background events based on the occurrence of a single veto-layer hit and was designed to be scalable to a larger detector. The causal filter therefore has the potential to attain a significantly lower energy threshold than the HESE algorithm. Further testing of the potential of this filtering technique however requires using simulation instead of real data. This is the focus of the next chapter, where the causal filter is at the same time used to test the vetoing capabilities of the proposed IceCube-GEN2 geometries.

6

FILTER PERFORMANCE IN ICECUBE-GEN2

Switching from IceCube data to simulated data, our analysis now implicitly becomes dependent on the physics models used to generate the simulated events. For this reason, it is of the highest importance that the simulated events reflect measured data as closely as possible. Because IceCube-GEN2 plans to measure the neutrino flux over an energy region in which only a few events have so far been observed, the normalisation of the signal events contains large uncertainties. Background events, consisting of muons and atmospheric neutrinos, have however been measured in sufficient abundance to leave only small uncertainties on their parametrisation. Therefore, the filter performance will be quantified in two ways. In the first method, the flux parametrisation of cosmic neutrinos obtained by the HESE filter will be used to compare the number of passing signal events to passing background events. This then allows to determine the optimal filter parameters when requiring a given purity. In the second method, the filter performance will be quantified solely based on the number of passing background events. As such, this will allow to asses the filter performance independent of the large uncertainty on the cosmic neutrino flux parametrisation.

Uncertainty on signal parametrisation

Following a standard protocol, two sets of simulated events will be used, one signal and one background. Background events will in real life be due to entering muons and atmospheric neutrinos. The latter can only be tagged as background, if they are accompanied by one or multiple muons travelling parallel to the neutrino. As a result, background rejection in both types of events relies on correctly identifying entering muons. We therefore use a simulated background sample consisting purely out of atmospheric muons. Moving to signal, we note that most starting events identified by the HESE filter have a cascade topology. While CC interactions of muon or high-energy tau (anti-)neutrinos can produce events containing both a cascade and track, these represent a minority of all starting events. For this reason, we use a set of simulated electromagnetic cascades occurring inside the detector volume as signal.

Atmospheric muons as background

EM cascades as signal

Apart from the above mentioned features, a few final remarks about the simulated events are in order. First, the simulated signal and background dataset are both based on a high-energy InIce extension composed of P-DOMs. Results regarding the absolute performance of the filter thus only apply to this DOM choice. It is however reasonable to assume that when comparing geometries, the relative performance of the filter will be similar for the other DOM choices. In addition, P-DOM offers the smallest photo-collection area of all DOMs and no directional

information. Using a different DOM-type would thus most likely result in an improved filter performance. Finally, we note that the simulation assumes that the proposed IceCube-GEN2 geometry can be perfectly realised. In reality, there will be variations on the locations at which DOMs are deployed and there might be a small fraction of DOMs which break during deployment or are temporarily malfunctioning.

6.1 SIMULATION DATASETS

Simulation chain

The simulation of events in the IceCube-GEN2 detector proceeds via three separate stages. Only the first stage, in which a particle is sampled from a given probability distribution, is different for the signal and background simulation. In the second stage, the interactions induced by this primary particle are simulated. This includes the generation of secondary particles and the propagation of photons through the ice. The third stage, comprises the simulation of the signals in the detector itself. Photons reaching the DOMs are converted into hits, taking into account the time resolution and efficiency of the DOMs. In addition, other detector effects such as dark noise hits and afterpulses are also simulated. Once all hits corresponding to an event are generated, the last step of the simulation is to apply a hit cleaning algorithm. By using the same code to generate signal and background events in the second and third stage of the simulation chain, we thus ensure that no bias regarding the performance of the causal filter is introduced between the two datasets.

6.1.1 *MuonGun*

MuonGun is a program which generates a muon particle by sampling its properties from predefined probability distributions. These properties include its energy E_μ, propagation direction (θ, ϕ) and position (x, y, z).

Two datasets

In this thesis, two datasets of MuonGun simulation are used. Both contain muons generated on a cylindrical surface surrounding the detector, with azimuthal angles ϕ generated uniformly in $[0, 2\pi]$. The aspects in which they differ are the zenithal distribution, energy distribution and detector geometries for which they were generated. Table 3 shows these properties for both datasets.

While dataset n° 1 contains $1.6{\cdot}10^5$ events, only half of those are down-going and therefore usable in our analysis. Dataset n° 2 was generated per our request, as imposing an energy cut $E > 100$ TeV on the first dataset revealed that little to no events were passing the filter. The second dataset was therefore generated with properties much closer to our requirements, containing only down-going muons with a zenithal distribution following that of the atmospheric muon background. Computational constraints however restricted us to only generate this dataset for the Sunflower 240 geometry.

Table 3: Properties of the MuonGun datasets.

Property	n° 1	n° 2
# events	$1.6 \cdot 10^5$	10^6
$\frac{dN}{d\theta}$	uniform in $\cos(\theta) \in [-1, 1]$	atmospheric
$\frac{dN}{dE}$	$E^{-1.4}$	E^{-3}
Energy range (TeV)	$[3, 10^5]$	$[10^2, 10^4]$
Sunflower 200	✓	
Sunflower 240	✓	✓
Sunflower 300	✓	
Edge-weighted	✓	

6.1.2 *VolumeInjector*

Electrons as primary particles

Like MuonGun, VolumeInjector is program which generates an elementary particle that serves as the start of the simulation process. Unlike MuonGun, this particle is an electron with a position distributed uniformly throughout the detector volume. As such, VolumeInjector provides a good approximation for neutrino induced interactions. These interactions can be expected to occur uniformly throughout the detector volume, consistent with the position of the observed HESE events [28]. Most neutrino interactions will induce a hadronic instead of an electromagnetic cascade. However, only electrons are used here as primary particle. Since the detector cannot distinguish between the two types of cascades[1], this does not influence our results.

To test the performance of the causal filter, a single dataset of events generated with VolumeInjector was used. For each of the Sunflower geometries and the Edge-weighted geometry, a set of $1.7 \cdot 10^5$ events was generated. The propagation direction of the primary electron was sampled uniformly on the unit sphere. Its energy was generated following an $E^{-1.4}$ power-law spectrum in the energy range $E \in [3\ \text{TeV}, 10^5\ \text{TeV}]$. Because the passing fraction is much higher for signal events than for background events, no additional dataset was needed to estimate the signal efficiency for large energy cuts.

6.1.3 *SRT-cleaning*

The final step of the simulation chain is to remove those hits which are most likely due to noise. Because IceCube-GEN2 hits do not receive an HLC tag in real-time, software-based hit cleaning is instead applied on the simulated events. For the above mentioned datasets, we used the

1 Accounting for the fact that hadronic showers have a 15% lower light yield than electromagnetic showers.

RT-cleaning

Seeded RT (SRT) cleaning algorithm to obtain a set of cleaned hits before applying the filter. SRT-cleaning is an extension of the RT-cleaning algorithm, which was designed to remove noise hits that are separated in space and/or time. Specifically, a hit h_1 is only kept if another hit h_2 is found which satisfies the following two criteria.

1. The spatial distance between h_1 and h_2 is smaller than R.
2. The two hits both occur within a time T from each other.

For IC78, $R = 150$ m and $T = 1$ μs are used as standard parameters[2]. When applying the same algorithm to events in the IceCube-GEN2 geometry, R and T are rescaled such that they are linearly proportional to the string spacing. These larger values are then used to link hits in GEN2-GEN2 or GEN2-IC86 strings, whereas the standard values are still used if both hits occurred in the IC78 or DeepCore region of the detector.

SRT-algorithm

The SRT-algorithm, unlike the RT-method, starts from a small subset A of the total set of hits. In general, A will contain hits for which there is a strong indication that they belong to the event. For example, A might contain only hits within a given distance of the COG[3]. Other hits are then added to A, if there exists a hit in A with which they satisfy the two RT-criteria. This method can then be applied recursively until no more hits can be added.

6.2 VETO-LAYER & FIDUCIAL VOLUME

When the results of the causal filter were compared to those of the HESE filter in Section 5.4, both filters were applied on the exact same dataset. While this required using the same veto-layer as that of the HESE analysis, it ensured that differences in the performance of the filters were purely caused by using a different filter algorithm. For the IceCube-GEN2 geometries, there is no longer a predefined veto-layer to which our analysis needs to be restricted. As a result, we will define a general veto-layer geometry that will be used to compare the filter performance for the different geometries. In addition, multiple veto-layer choices will be examined for the baseline Sunflower 240 geometry.

Definition veto-layer

Apart from influencing the filter performance, the extent of the veto-layer also determines the rate at which signal events are observed. As any cascade starting in the veto-layer is tagged as background, the rate at which starting events occur is linearly proportional to the fiducial volume. Signal events are therefore considered to be only those cascades who's reference hit is not located in the veto-layer[4]. Before applying the

Signal events: cascades in fiducial volume

2 In DeepCore, $R = 90$ m and $T = 0.7$ μs are used as standard parameters.

3 In IC86, HLC hits are often used to define A at the start of SRT-cleaning.

4 This does not result in the signal rate being underestimated, as events in which the reference DOM lies in the veto-layer are automatically rejected by the filter.

filter to simulation, we can therefore already compare the relative total rates of signal events based on the size of the fiducial volumes. We now define the veto-layer of the combined GEN2-IC86 geometry to include

- all outer strings, as shown in the top panel of Figure 25.
- the top six and bottom one DOM of strings in the IC78 geometry.
- the top 12 and bottom one DOM of strings in the GEN2 geometry.
- all DOMs in the depth range $-215\ \text{m} \leqslant z \leqslant -135\ \text{m}$, below the dust layer.

For the Sunflower 240 geometry, we also examine veto-layers in which the top six or 18 DOMs are included. In case of the latter, the outer veto-region can also be extended to include a double row of strings, as shown in the bottom panel of Figure 25. Care should however be taken when including only the top six DOMs of the GEN2 strings. Because GEN2 strings carry 80 DOMs, six above and 14 below the IC78 DOMs, a hole in the veto-layer would be present at the transition from the GEN2 to IC86 strings. To prevent this from happening, the top 12 DOMs are always included in GEN2 strings that are located at the edge of the IC86 geometry[5], as illustrated in Figure 24.

Comparison of fiducial volumes

A comparison between the fiducial volumes of the different geometries, assuming a single outer-layer and 12 DOMs on top, is given in Table 4. Similarly, Table 5 shows the fiducial volume of the Sunflower 240 geometry corresponding to various choices of the veto-layer. For the values quoted in Table 4 and 5, the fiducial volume corresponds to the area enclosed by the convex hull of all strings that are not included in the veto-layer. As a reference, we note that applying the same formulae to the IC86 geometry results in a total volume $V = 0.85\ \text{km}^3$ and fiducial volume $V_f = 0.42\ \text{km}^3$, the ratio of which is 49.0%.

5 Formally, we include the top 12 DOMs of a GEN2 string if the distance to the closest IC78 string is smaller than 1.3 times the GEN2 string spacing.

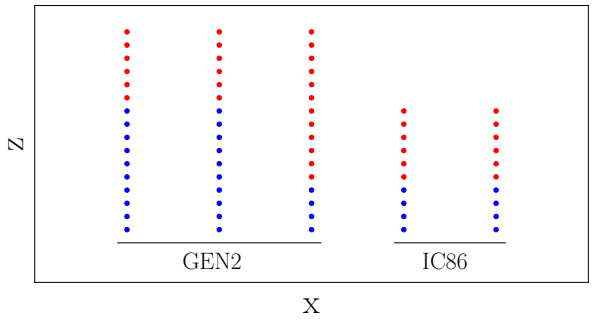

Figure 24: Illustration showing the transition from GEN2 to IC86 strings when including the top 6 DOMs of GEN2 strings in the veto-layer. DOMs are coloured red if they are included in the veto-layer and blue otherwise.

Table 4: Fiducial volumes of the different geometries, assuming a veto-layer which includes a single layer of outer strings and the top 12 DOMs of GEN2 strings.

Geometry	Total volume (km^3)	Fiducial volume (km^3)	ratio (%)
Sunflower 200	5.94	3.29	55.3
Sunflower 240	7.90	4.49	56.9
Sunflower 300	11.8	6.04	51.2
Edge-weighted	6.19	3.46	55.9

Table 5: Fiducial volumes of the Sunflower 240 geometry for various configurations of the veto-layer.

N_{top}	double edge	Fiducial volume (km^3)	ratio (%)
6		4.86	61.5
12		4.49	56.9
18		4.12	52.2
18	✓	2.67	33.8

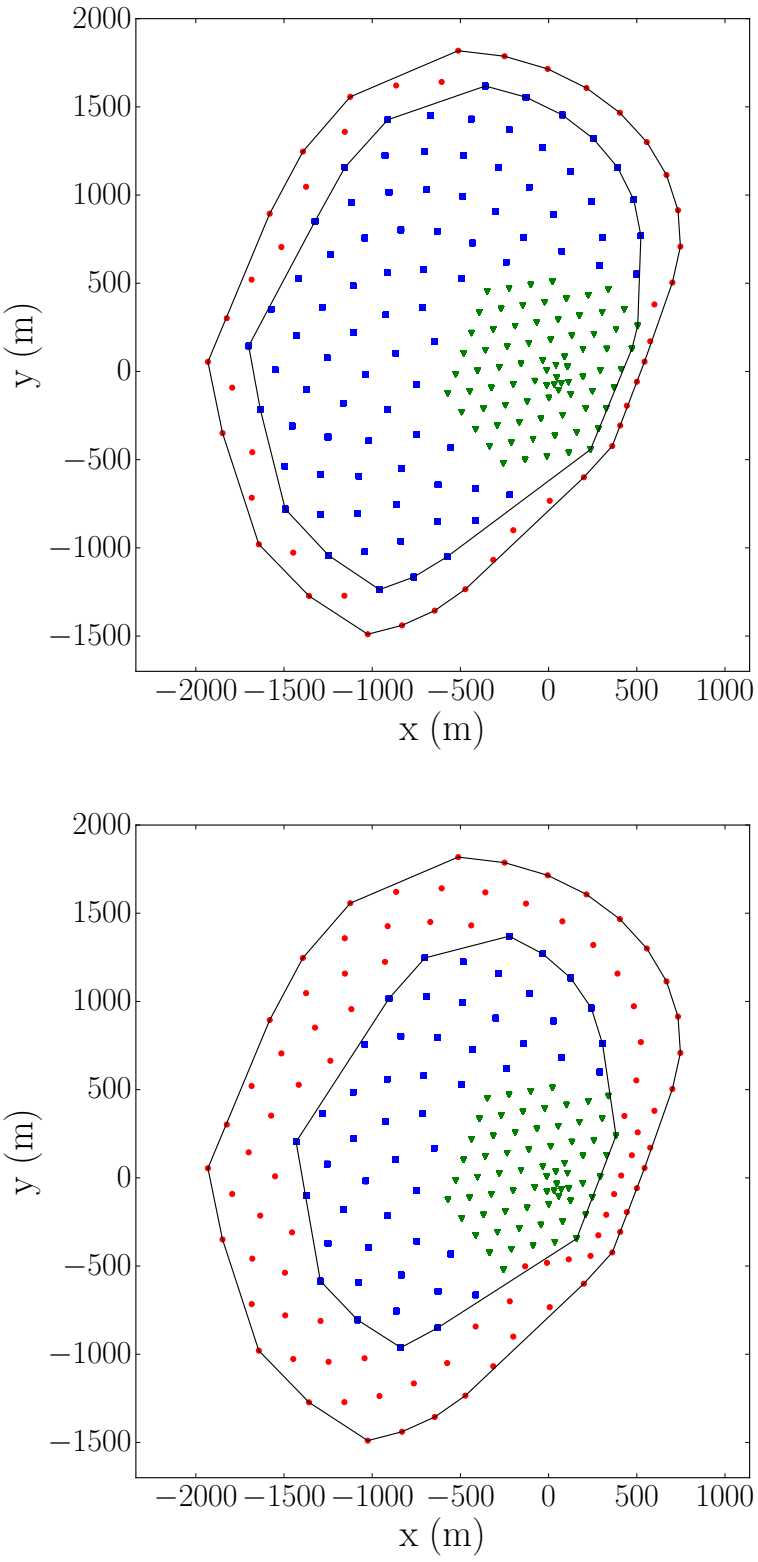

Figure 25: Comparison between the strings included in the single (top) and double (bottom) veto-layer of the Sunflower 240 geometry. The outer and inner black line respectively mark of the convex hull of the total and fiducial volume. Veto-layer strings are indicated as red circles. Strings belonging to the fiducial volume are indicated as green triangles if they are IC86 strings, and blue squares otherwise.

6.3 EXPECTED RATE OF EVENTS

Testing the filter performance requires comparing the rates at which signal and background events pass the filter. To obtain enough statistics at large energies, the simulated datasets were generated using an energy spectrum that is harder than the true power law spectrum. Since every generated primary particle is forced to interact, events need to be weighted to their respective physical fluxes in order to obtain realistic event rates. In this section, we describe this weighting procedure, along with the associated uncertainties.

6.3.1 *Weighting background events*

Based on CR model

Over the years, the flux of atmospheric muons has been extensively characterised by the IceCube collaboration (see e.g. [36]). As these muons originate from high-energy cosmic rays impinging on the Earth's atmosphere, their parametrisation at the detector can be determined starting from a cosmic ray flux model. By relating this parametrisation to that used in the simulation process, weights can be generated for the individual events. Contained in the IceCube simulation framework[6] is a module which can apply this weighting scheme to MuonGun simulation, for a set of predefined cosmic ray flux models.

Simulation chain

Parametrising the atmospheric muon flux at the detector level is done in several stages. First, cosmic ray air showers are generated using Monte Carlo techniques, starting from a given cosmic ray flux model. The muons produced in these showers are then propagated to sea-level, where their distribution is parametrised. A separate Monte Carlo simulation is then invoked, which draws muons from this distribution and propagates them through the ice until the detector depth is reached. At the detector, this muon flux is then parametrised as a function of energy, zenith angle and multiplicity.[7] [8].

Gaisser model

The question now remains which cosmic ray flux to use when weighting the events. For the MuonGun simulation, two flux models can be used. These are respectively called Gaisser and Hoerandel, after the authors of [15] and [8, 21]. While a detailed comparison of these two models is outside the scope of this thesis, we note that the flux corresponding to the Gaisser model is slightly larger than that of the Hoerandel model (see Figure 29). Therefore, weights generated with the Gaisser model are used as standard, to avoid underestimating the atmospheric muon background.

6 http://software.icecube.wisc.edu/

7 A single air shower can produce multiple muons travelling parallel to each other in a muon bundle. Here, multiplicity denotes the number of muons in such a bundle.

6.3.2 *Weighting signal events*

In contrast to the MuonGun simulation, no IceCube weighting module is available for events generated with VolumeInjector. To cope with this, we implement our own weighting algorithm. The main input of the algorithm is the differential flux of cosmic neutrinos

Self-developed weighting procedure

$$\phi_\nu = \frac{dN}{dt\, dA\, d\Omega\, dE} \quad . \tag{12}$$

As an estimator for ϕ_ν, we use the flux found by the HESE analysis [32] given in equation (6). To be consistent with the MuonGun simulation, weights are generated such that they represent the ratio of the number of expected events per second over the number of simulated events. More formally, each event is given a weight w corresponding to

$$w = \frac{dN_{\text{expected}}}{dt\, dA\, d\Omega\, dE} \cdot \left(\frac{dN_{\text{simulated}}}{dA\, d\Omega\, dE}\right)^{-1} \quad . \tag{13}$$

Integrating over the detector volume and all solid angles, equation (13) simplifies to

$$w = \frac{dN_{\text{expected}}}{dt\, dE} \cdot \left(\frac{dN_{\text{simulated}}}{dE}\right)^{-1} \quad . \tag{14}$$

As the energy distribution of the simulated events is fully known, it only remains to find the number of expected events per unit time and energy. This quantity can be written as the sum of two separate contributions, one from CC and one from NC interactions

CC and NC interactions

$$\begin{aligned}\frac{dN_{\text{expected}}(E)}{dt\, dE} = &\int d\Omega\, \phi_d(E,\theta)\, N_{target}\, \sigma^{CC}(E) \; + \\ &\int d\Omega \int_0^1 dy\, N_{target}\, f\left(\frac{E}{y},\theta\right) \quad .\end{aligned} \tag{15}$$

In this expression, ϕ_d is the differential flux reaching the detector, N_{target} is the number of nucleons in the detector and y is the fraction of energy lost by the neutrino if it undergoes a NC interaction. More details regarding the function $f\left(\frac{E}{y},\theta\right)$ and the energy loss in NC interactions are given in Appendix B. To implement equation (15), explicit expressions are now required for the differential flux at the detector ϕ_d, the number of nucleons N_{target} and the neutrino-nucleon cross section σ. As such, we will now discuss each of these components in more detail.

Cross section

For both NC and CC interactions, the cross section is at all energies dominated by interactions with nucleons [14]. An important exception however occurs for anti-electron neutrinos with energy $E_{\bar{\nu}_e} \approx 6.3$ PeV. At this energy, interactions with electrons dominate the cross section due to the Glashow resonance [17]. Currently, nor IceCube, nor any other neutrino telescope has observed a Glashow neutrino interaction [64]. Parametrisations of the cosmic neutrino flux take this absence of

events into account by either, predicting a softer power law spectrum, or by predicting a flux with non-uniform flavour ratio's for neutrinos and anti-neutrinos [64]. In the latter case, interactions induced by the Glashow resonance would also be largely absent in bigger neutrino detectors. To prevent overestimating the signal, we therefore only account for the interaction cross section with nucleons, effectively assuming a parametrisation in which anti-electron neutrinos are absent in the cosmic neutrino flux.

While the neutrino-nucleon cross section at PeV energies has not yet been measured, multiple parametrisations describing its value exist. Here, we use the one described in section VII of [14] to obtain the cross sections of CC and NC neutrino and anti-neutrino interactions. An added complication is that in NC interactions, neutrinos deposit only a fraction y of their energy. We therefore consider the differential cross section $\frac{d\sigma^{NC}}{dy}$, which is defined such that

$$\sigma^{NC}(E) = \int_0^1 dy \, \frac{d\sigma^{NC}(E,y)}{dy} \quad . \tag{16}$$

For (anti-)neutrinos interacting with nucleons, the differential cross section takes the form

$$\frac{d\sigma}{dy}(E,y) = a(E) + b(E) \cdot (1-y)^2 \quad , \tag{17}$$

where a and b are energy dependent constants which differ for neutrino and anti-neutrino interactions [16]. From equation (17), it immediately follows that the y-distribution of the differential cross sections only depends on the ratio $\frac{a}{b}$. This ratio was therefore calculated using the values $\langle y(E) \rangle$ specified in Tables 1 and 2 of [16]. Combined with the absolute normalisation taken from [14], the neutrino-nucleon cross section is thus fully specified[8]. Figures showing the used parametrisations for the various cross sections and mean values of y are included in Appendix B.

Number of nucleons

The second component required in equation (15) is the number of nucleons in the detector volume N_{target}. It can easily be seen that this quantity is equal to

$$N_{target} = V \cdot \rho \cdot \frac{1}{\mathscr{M}} \cdot N_A \cdot 18 \quad , \tag{18}$$

where V is the volume of the detector, ρ the density of the ice, $\mathscr{M}$ the molar mass of ice, N_A Avogadro's constant and 18 the number of nucleons per water molecule. For each of the proposed IceCube-GEN2 detector configurations, we set V equal to the corresponding value specified in Table 4. The density of ice is assumed to be constant at $\rho = 921.6$ kg m^{-3},

8 The alternative would be to directly calculate $a(E)$ and $b(E)$ starting from the nucleon's parton distribution functions.

as its value only varies on the order of 0.1% over the entire depth range of the detector [1].

Earth absorption

The final component required to weight the flux of signal is the conversion from the cosmic neutrino flux ϕ_ν to the flux reaching the detector ϕ_d. While ϕ_ν is isotropic as a function of both the azimuthal angle and cosine of the zenith angle $\cos(\theta)$, this is not the case for ϕ_d. At the detector, the flux of neutrinos coming from the Northern hemisphere is attenuated due to absorption in the Earth. Fully modelling this attenuation requires not only accounting for NC and CC interactions, but also for the fact that neutrinos oscillate and that the oscillation probabilities in matter differ from those in vacuum. Therefore, Monte Carlo techniques are typically used to determine the attenuation of the neutrino flux due to Earth absorption (see e.g. [55]).

Here, we use a simplified model, which distinguishes only between neutrinos and anti-neutrinos. In addition, only CC interactions are accounted for[9], eliminating the need to use Monte Carlo techniques. While this might seem a crude approximation, doing the full calculation would only influence the total rate of events in the detector by a few percent, as shown in [48]. Therefore, we can approximate the attenuation A of a neutrino beam coming from a zenith[10] angle θ as

$$A(E,\theta) = \exp\left(-\sigma^{CC}(E)\cdot N_A \cdot 1\frac{\text{mole}}{\text{g}}\int_0^L dt\ \rho(t)\right) \quad , \tag{19}$$

where we integrate over the path along which the neutrino traverses the Earth. Note that we implicitly assumed that the number of nucleons contained in one gram of matter is N_A. This is generally a good approximation and makes the calculation independent of the chemical composition of the matter inside the Earth. To perform the integration, a model describing the Earth density ρ is required. For this purpose, we use the Preliminary Earth Reference Model 500 (PREM500) [11], which describes a spherically symmetric model of the Earth. Figure 26 shows the density of the Earth $\rho(r)$ in the PREM500 model as a function of the distance from the Earth's core. Consider now the coordinate system shown in Figure 27, in which a neutrino enters the Earth at (a,b) from a zenith angle $\theta = \pi - \alpha$. In this coordinate system, the path of the neutrino (x,y) through the Earth is parametrised by

$$\begin{cases} x(t) = 2R\cos(\alpha)\sin(\alpha) - t\sin(\alpha) \\ y(t) = 2R\cos^2(\alpha) - R - t\cos(\alpha) \end{cases} , \tag{20}$$

where $t \in [0, 2R\cos(\alpha)]$. Knowing the radius $r = \sqrt{x^2+y^2}$ at any given distance t along the neutrino's path, the total attenuation can be found

9 In the energy range 10 TeV→100 PeV, the cross section of CC interactions is a factor ~3 higher than that of NC interactions [14].

10 Naturally, the attenuation factor is independent of the azimuthal angle.

by numerically integrating equation (19) from $t = 0$ to $t = 2R\cos(\alpha)$. The differential flux reaching the detector from zenith angle θ is thus $\phi_d(E,\theta) = A(\theta)\cdot\phi_\nu(E)$.

Figure 28 shows the resulting probability that a neutrino passes through the Earth unabsorbed. At energies exceeding 1 PeV, almost all vertically up-going neutrinos are absorbed by the Earth. To obtain the total number of events per unit energy and time, we now average the attenuation factor $A(\theta)$ by numerically integrating over all solid angles

$$A_{\text{tot}}(E) = \frac{1}{4\pi}\int_{-\pi}^{\pi} d\phi \int_{-1}^{1} d\cos(\theta)\, A(E,\theta) \quad . \tag{21}$$

Because neutrinos coming from the Southern hemisphere are unattenuated, the average attenuation factor is always limited between $0.5 \leqslant A_{\text{tot}} \leqslant 1$.

Total background and signal rates

With the weighting procedures fully specified, we can now compare the total rate of signal and background events. Figure 29 displays these rates for the baseline Sunflower 240 geometry as a function of energy. At all energies, the Gaisser (G) cosmic ray model predicts a larger muonic background flux than the Hoerandel (H) model. Therefore, all future analyses will use background events with weights generated based on the Gaisser model. By taking the ratio of the signal rate over the background rate, we find the maximal background passing fraction allowed to reduce the background rate to the same level as the signal rate. At an energy of 30 TeV and 100 TeV, the passing fraction should respectively be smaller than 0.007% and 0.3% to achieve this. Once the energy exceeds ~700 TeV, more signal than background events are expected to be observed. However, we stress again that due to the small number of cosmic neutrinos that have so far been observed, the expected rate of starting events comprises a large systematic uncertainty. This is especially the case in the $E > 1$ PeV region, where only 3 HESE events have so far been observed [64]. Because this uncertainty is to a much smaller extent present in the parametrisation of the atmospheric muon background, the absolute number of passing background events should closely resemble the true value. In contrast, not the absolute, but mostly the relative values of the number of passing signal events will be examined when comparing different detector geometries and veto-layer configurations.

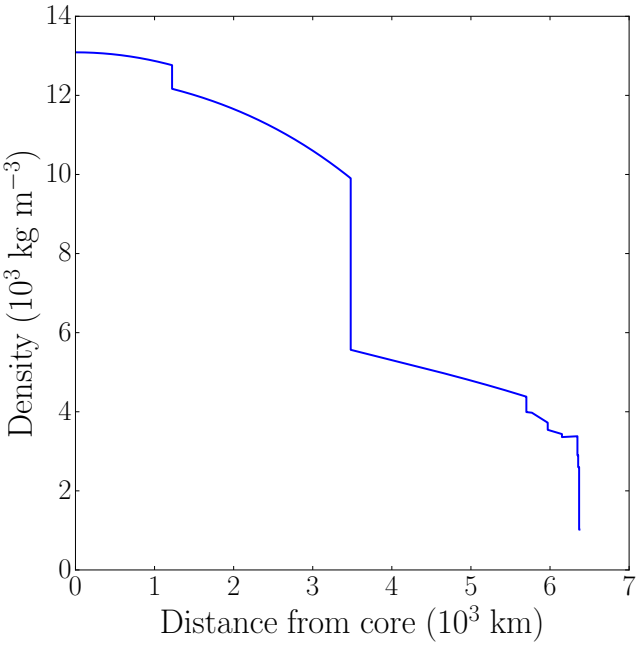

Figure 26: Density of the Earth as parametrised by the PREM500 model.

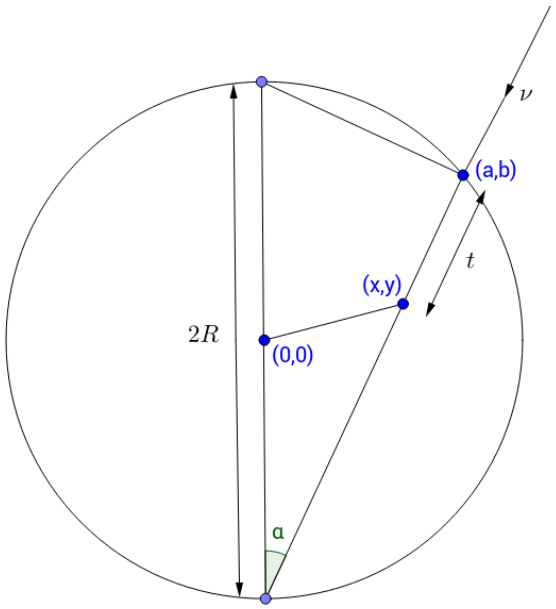

Figure 27: Geometrical illustration of the propagation path of a neutrino through the Earth for a given zenith angle $\theta = \pi - \alpha$.

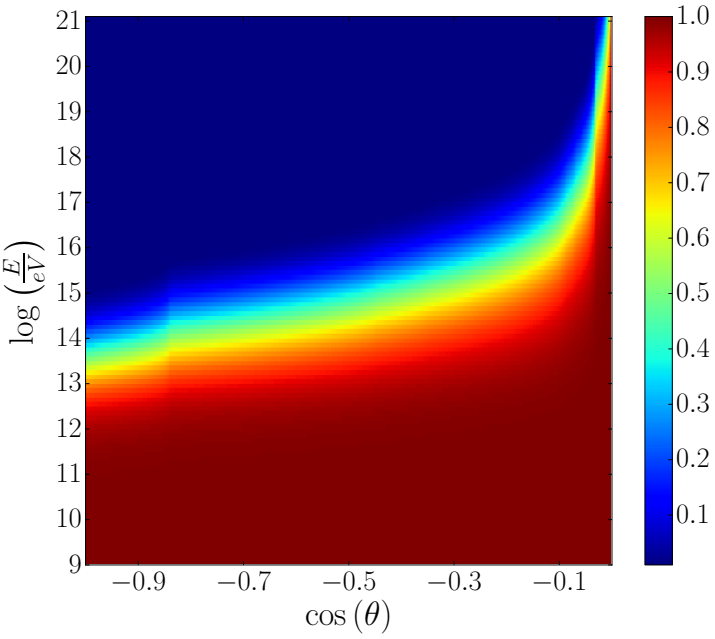

Figure 28: Colour scale showing the probability that a neutrino with energy E coming from a zenith angle θ can pass through the Earth without being absorbed.

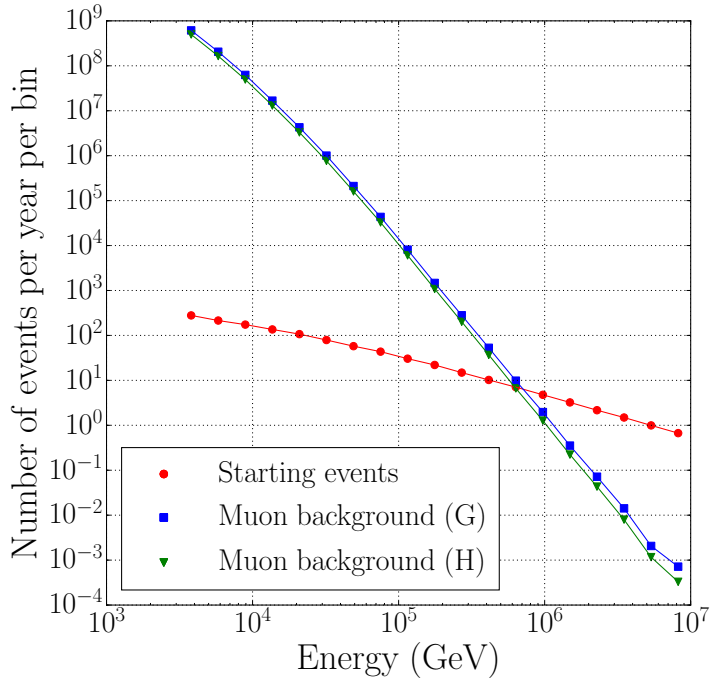

Figure 29: Predicted rate of signal and background events in the Sunflower 240 geometry. Two possible rates of background events are shown, corresponding to the weights generated based on the Gaisser (G) and Hoerandel (H) cosmic ray flux model.

6.4 T90 PROCEDURE

New filter parameters

Before applying the causal filter developed in Chapter 5 to the simulated datasets, we first need to specify the filter parameters that will be used. The parameters which provided the best performance for IC86 are not automatically the optimal ones for the IceCube-GEN2 geometries. Therefore, we will redetermine what the optimal filter parameters are and do so for each of the examined geometries and veto-layer configurations. To narrow our search of phase space, we examine in this section which range of values should be considered for each of the filter parameters.

Apart from specifying the filter parameters, another requirement is the imposition of a cut to remove the low-energy events. Examining Figure 29 shows that the ratio of signal over background events increases roughly as $\sim E^3$. As such, the ratio of signal to background events passing the filter will greatly improve with increasing energy. In IceCube, a threshold of 6000 pe was therefore imposed on the total charge of the event. This choice was made because the total charge provides a fairly robust and easily calculated measure for the energy of the event. However, in the IceCube-GEN2 geometries, this is no longer the case due to the non-uniform spacing of DOMs. Combined with a different DOM design, it follows that the charge measured in the IC86 and GEN2 part of the detector will be disproportionate, making the total charge an ill-suited candidate. We therefore opt to instead use the true energy of the particle that induces the event. While reproducing the absolute filter performance with real data would thus require an excellent energy estimator to be at hand, our goal in this chapter is mainly to examine the relative filter performance for the different geometries and veto-layer configurations. Accordingly, low energy events will be removed by placing a cut on the total energy of the primary particle. To also explore a method which is directly applicable to real data, the filter performance will be examined in Section 6.5.4 for the baseline geometry and veto-layer configuration using a cut performed on the total charge of the event.

Energy cut

6.4.1 *Selecting a reference hit*

The first two parameters which we examine, X_{T90} and $(\Delta t)_{\mathrm{T90}}$, are those related to the selection of the reference hit. Together, they determine the fraction of the total charge and time window in which it needs to occur before we select a reference hit. Following the same reasoning of Section 5.3.2, X_{T90} is again chosen to be fixed at 10%. This value stems from the expectation that if a muon deposits hits in the veto-layer upon entering the detector, at least part of these hits should be contained in the first 10% of the total charge. In contrast, the same charge interval should not contain any veto-layer hits in case of cascades occurring in-

X_{T90}=10%

side the fiducial volume.

Unlike in the case of IC86, 1 μs is no longer a suitable choice for $(\Delta t)_{\mathrm{T90}}$. This is due to the fact that tracks travelling diagonally through the detector can take longer than $\frac{1\ \mu\mathrm{s}}{10\%} = 10\ \mu$s to do so. Because of the roughly doubled string spacing, it turned out that for the IceCube-GEN2 geometries, $(\Delta t)_{\mathrm{T90}} = 2\ \mu$s is a sufficiently large time window for all types of events. Applying the algorithm to the simulated datasets confirmed this, as in each event, a 2 μs time window containing more than 10% of the total charge could be identified. Given that the function of $(\Delta t)_{\mathrm{T90}}$ is to reduce the contribution of noise hits to X_{T90}, using larger values of $(\Delta t)_{\mathrm{T90}}$ would result in a minor reduction of the filter performance. Therefore, we use X_{T90}=10% and $(\Delta t)_{\mathrm{T90}} = 2\ \mu$s in all the analyses described in this chapter.

$(\Delta t)_{\mathrm{T90}} = 2\ \mu$s

An important remark is that we do not expect the reference hit selection procedure to work correctly if the total charge Q_{tot} of the event is too low. When analysing IceCube data, this was never a problem as only events for which $Q_{tot} > 1500$ pe were examined. For the IceCube-GEN2 analysis, an energy cut is however used instead of a charge cut. While the deposited charge generally increases with energy, interactions at the edge of the detector can lead to high-energy events with a small total charge. To ensure that at least ten hits have occurred before the reference hit is selected, we impose a small charge threshold of 100 pe. As will be shown in Section 6.5.4, this leads to a negligible loss of signal events[11], given that an energy cut of at least 50 TeV will already have been imposed.

$Q_{tot} > 100$ pe

Having specified the parameters required to select a reference hit, we now examine the performance of this selection procedure in the Sunflower 240 geometry[12]. In case of background events, the coordinates of the reference hit $(\vec{x}_h, t_h)$ should ideally approximate those of the muon at time t_h. To determine whether this is the case, Figure 30 displays the fraction of background events for which the distance

Performance for background events

$$\Delta x = \|\vec{x}_h - \vec{x}_{\mathrm{muon}}(t_h)\| \quad , \tag{22}$$

is larger than a certain distance d (blue curve). Additionally, we also examine the same relation for the distance between the reference hit and the muon track (red curve)

$$\Delta x' = \|\vec{x}_h - \vec{x}_{\mathrm{muon\ track}}\| \quad . \tag{23}$$

Because $\Delta x'$ corresponds to the distance at closest approach, the relation $\Delta x' \leqslant \Delta x$ applies to each of the events. Therefore, the red curve always

11 Imposing a charge cut of 100 pe removes 3.3% of all signal events who's energy is larger than 100 TeV in the Sunflower 240 geometry.

12 An energy threshold $E > 100$ TeV is imposed before applying this analysis.

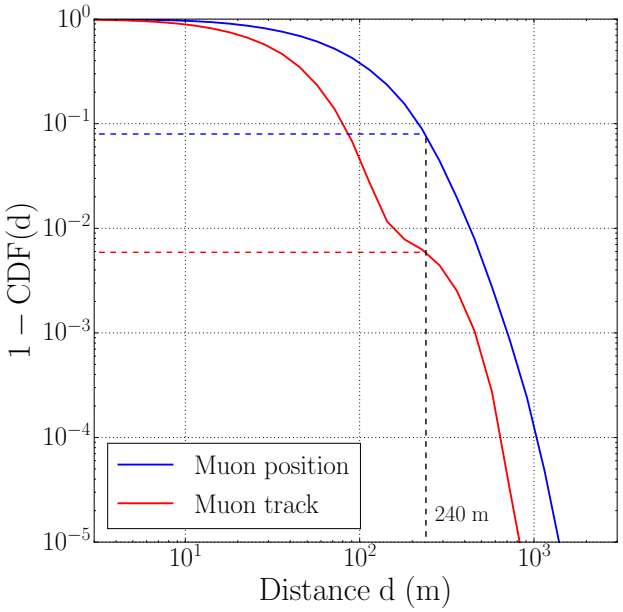

Figure 30: Fraction of background events for which the distance between the reference hit and the position of the muon (track) at its occurrence is larger than d.

lies below the blue curve. Using the string spacing as a reference distance, we note that for respectively 8% and 0.6% of all events, Δx and $\Delta x'$ are larger than 240 meter. At $d = 1000$ m, these values respectively reduce to 0.01% and <0.001%. It is fair to assume that in a certain fraction of all events, the reference hit will be noise induced and therefore unrelated to the event. The mean value of Δx for those events should then roughly correspond to the average distance between any two given DOMs, which is 1430 meter. Since this distance is reached in less than 0.001% of all events, we conclude that the reference hit in general offers a reasonable estimator for the muon's coordinates.

Performance for signal events

The same analysis can of course be performed for signal events. Figure 31 displays the fraction of events for which the distance Δx between the reference hit and the point at which the neutrino interaction occurred is larger than a certain distance d. Compared to the background events, we find that the distribution is more strongly constrained to small values, as less than 0.04% of the events have a value $\Delta x > 240$ m. Therefore, we find that the reference hit almost always occurs close to the cascade centre. Additionally, because Δx exceeds 400 meter in less than 0.001% of all events, we conclude that the probability for the reference hit to be noise induced is negligible.

Apart from examining the distance in space between the reference hit and the neutrino interaction, we also examine their time difference Δt. Figure 32 shows the probability distribution of Δt, where the normali-

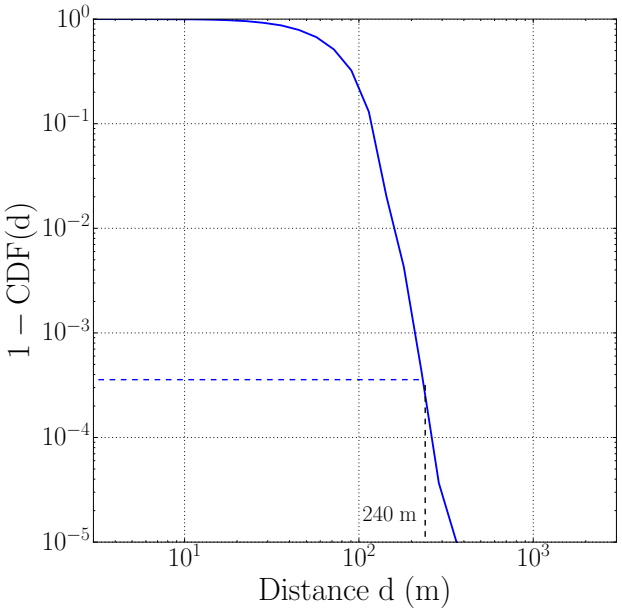

Figure 31: Fraction of signal events for which the distance between the reference hit and the position at which the neutrino interaction occurred is larger than d.

sation of the bins is such that the sum of their values equals one. First, we note that the absence of negative values indicates that as expected, all reference hits occurred after the start of the cascade. In addition, no events occur before $\Delta t = 50$ ns. This is a consequence of the time required by the DOMs to accumulate 10% of the total charge. The plateau up to $\Delta t \sim 700$ ns is then due to the difference in distance between the cascade centre and the closest DOM[13] to it. Finally, we find by integrating the distribution, that in over 98% of all events, the reference hit occurs within 1 μs of the start of the cascade.

6.4.2 *Filter parameters*

N_{min} and $(\Delta t)_{max}$

While X_{T90} and $(\Delta t)_{\mathrm{T90}}$ could be chosen independently of the filter algorithm, this is not the case for the two remaining filter parameters, N_{min} and $(\Delta t)_{max}$. These two parameters thus have to be optimised for each of the considered geometries and veto-layers configurations. Accordingly, we here investigate which values of N_{min} and $(\Delta t)_{max}$ should be considered for this purpose.

To identify low-energy atmospheric muons, the minimal number of veto-layer hits N_{min} required to tag an incoming track should preferably be as small as possible. For IC86, optimal results were obtained with N_{min} set to 1 or 2 hits. As the IceCube-GEN2 geometry will con-

13 It takes light 400 ns to travel a distance corresponding to half the string spacing, i.e. 120 meter.

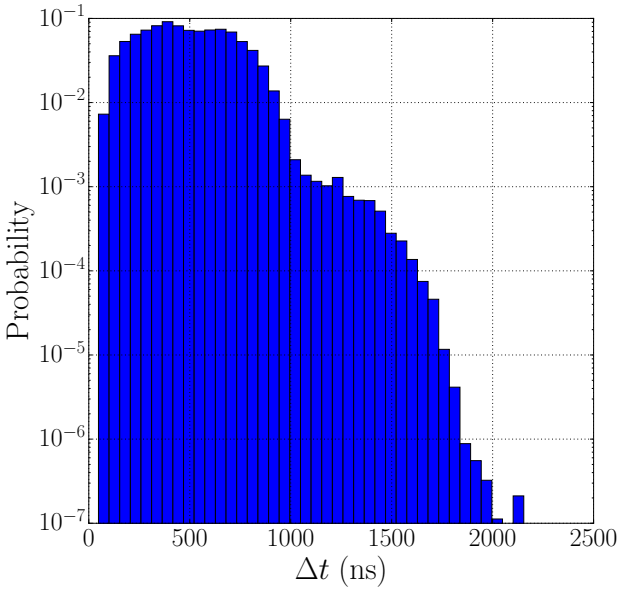

Figure 32: Probability distribution of the time difference between the occurrence of the reference hit and the neutrino interaction inducing the cascade.

tain in total about 3 times as much DOMs as IC86, higher values of N_{min} may be required due to the larger contribution from noise hits. Therefore, we here consider $N_{min} \in \{1,2,3,5,10\}$.

$N_{min} \in \{1,2,3,5,10\}$

To determine the causal time window $(\Delta t)_{max}$ in which those hits need to occur to be considered related to the event, consider the distribution shown in Figure 33. For each event, the minimal value of $|(\Delta t)_c|$ was calculated to allow determining the cumulative distribution of all events. The y-axis thus displays the fraction of events that are not rejected if $(\Delta t)_{max}$ is smaller than the corresponding x-value[14]. For background (red) and signal (blue) events, this fraction respectively converges to 0.3% and 63%. These are events for which no veto-layer hits occurred before t_{ref}, meaning that the causal filter will always accept them regardless the value of N_{min} and $(\Delta t)_{max}$. As such, only values smaller than 5 μs should be considered for $(\Delta t)_{max}$. At the same time, values larger than 0.5 μs are required to reject at least 99% of the background. Considering this in combination with the fact that in IC86, $(\Delta t)_{max}$ = 1 μs provided the best performance, we here examine $(\Delta t)_{max}$ (μs) $\in \{0.5,1.0,1.5,2.5,5\}$.

$|(\Delta t)_c|$ *distribution*

$(\Delta t)_{max}$ (μs) $\in \{0.5,1.0,1.5,2.5,5\}$

6.5 SUNFLOWER 240

To start, we first determine the filter performance for the baseline Sunflower 240 geometry. An advantage of restricting ourselves to this ge-

14 This is assuming $N_{min} = 1$ and only applies to events simulated in the Sunflower 240 geometry with an energy cut $E > 100$ TeV.

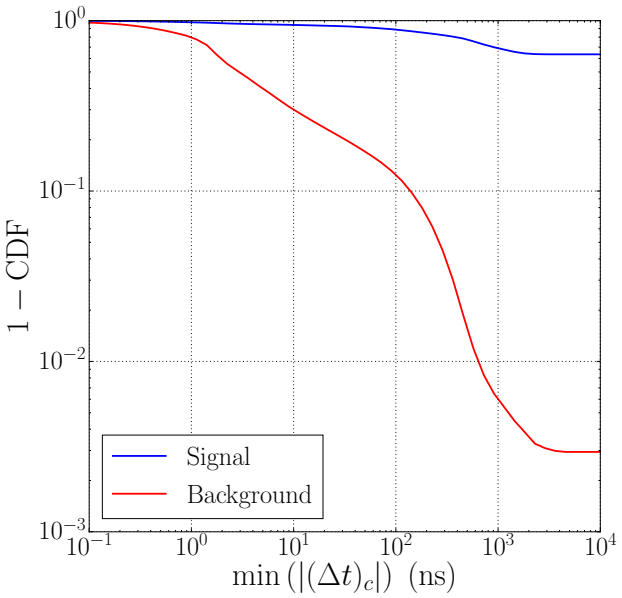

Figure 33: Fraction of events for which no veto-layer hits occurred before the reference hit and within a causal time window $|(\Delta t)_c|$.

ometry is that the larger MuonGun simulation dataset can be used, allowing an accurate estimation of the passing background at energies $E_\mu > 100$ TeV. In this section, we first discuss the parameters based on which the filter will be optimised. Then, a detailed analysis is given on the filter performance for the standard veto-layer configuration. We ensuingly consider three other configurations of the veto-layer and compare the filter performance obtained for each of them. Finally, we repeat the analysis for the baseline veto-layer configuration, using a charge instead of an energy based cut.

6.5.1 *Quantifying filter performance*

Optimising the filter performance for a given configuration will be done based on two quantities, the purity p and signal efficiency ϵ_s. Denoting the number of signal and background events passing the filter as s and b respectively, the purity is defined as

purity

$$p \equiv \frac{s}{s+b} \quad . \tag{24}$$

In other words, the purity determines the fraction of events passing the filter that correspond to signal events. The signal efficiency is defined as

Signal efficiency

$$\epsilon_s \equiv \frac{s}{s_{tot}} \quad , \tag{25}$$

where s_{tot} is the total number of signal events. It thus provides a measure for the fraction of signal events that is accepted by the filter. Optimally, both the purity and signal efficiency should equal one. In reality,

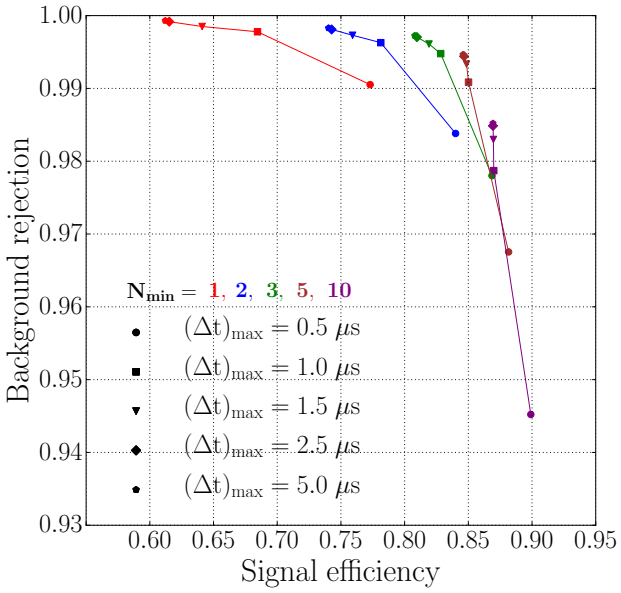

Figure 34: Background rejection as a function of signal efficiency when imposing an energy cut at 200 TeV.

we will search for the filter parameters which maximise the signal efficiency for a given minimal required value of the purity. As the purity required in future cosmic neutrino searches might vary per analysis, we consider values ranging from $p = 0.1$ to $p = 0.99$. When comparing different geometries or veto-layer configurations, the total rate of signal events will change along with the signal efficiency. Therefore, the total weight of all signal events passing the filter will be compared, instead of the signal efficiency. This will allow to determine which geometry or veto-layer results in the maximal number of filtered signal events that can be expected per year, given a required purity.

6.5.2 *Baseline veto-layer configuration*

To start, we first examine the filter performance for the baseline veto-layer configuration when imposing a high energy cut $E > 200$ TeV. Figure 34 shows the background rejection as a function of the signal efficiency for each of the 25 examined filter parameters. Naturally, the background rejection always increases for larger values of $(\Delta t)_{max}$ and smaller values of N_{min}. The inverse holds for the signal efficiency, which is maximised when using the smallest value of $(\Delta t)_{max}$ and the largest value of N_{min}. Finally, we note that the points corresponding to $(\Delta t)_{max} = 2.5$ μs and $(\Delta t)_{max} = 5$ μs almost overlap. This indicates that the plateau in Figure 33 is reached, meaning that using even larger values for $(\Delta t)_{max}$ would produce similar results.

$E > 200$ TeV

Figure 35 shows the same results, but with the purity displayed on the y-axis. An interesting observation is that the maximal purity corre-

Maximising the purity

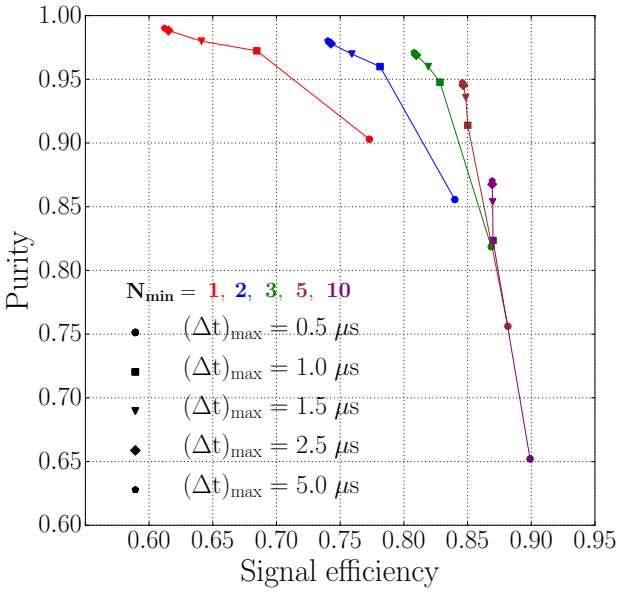

Figure 35: Purity as a function of signal efficiency when imposing an energy cut at 200 TeV.

sponds to the point at which the signal efficiency is minimised. This result can be understood by considering that the total rate of background events significantly exceeds that of signal events. Therefore, the purity $p = \frac{s}{s+b} \approx \frac{s}{b}$ is approximately inversely proportional to the number of accepted background events b. Noting that the maximal and minimal values of $\frac{s}{s_{tot}}$ differ by ~50%, whereas $\frac{b}{b_{tot}}$ varies ~2 orders of magnitude, it follows that the increase of the purity is mostly driven by the improved background rejection. Therefore, the maximal purity will typically occur for the filter parameters which minimise the number of passing background events.

Lower energy cuts

While the purities in Figure 35 are all larger than 60%, these results are based on a fairly high energy cut $E > 200$ TeV. At lower energies, the total background to signal ratio rapidly increases while at the same time, the background rejection decreases. As a result, the maximal attainable purity reduces from 99% at $E > 200$ TeV to 76% at $E > 100$ TeV and 32% at $E > 60$ TeV. For those latter two energy cuts, distributions analogous to those of Figures 34 and 35 are included in Appendix B. Comparing these distributions, we find only minor variations in the obtained signal efficiency, but a significant decrease of the background rejection.

To examine to what extent the filter depends on the energy cut-off, we examine its performance when fixing the filter parameters $(\Delta t)_{max}$ and N_{min} to 1 μs and 2 hits respectively. Figure 36 displays the signal

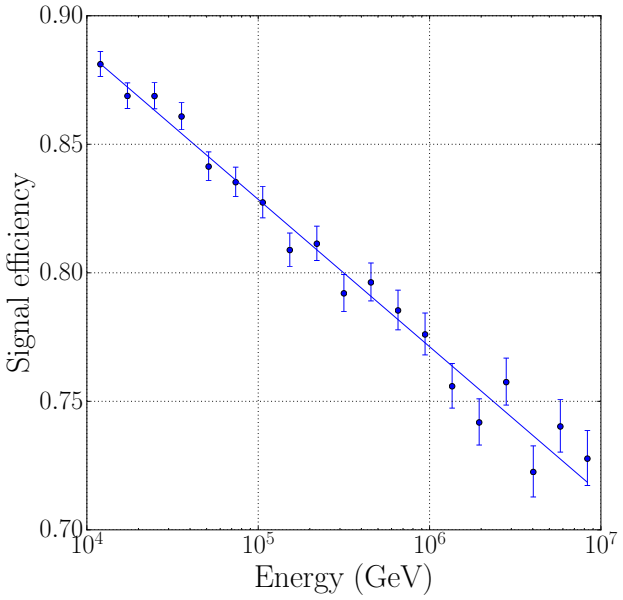

Figure 36: Energy dependence of the signal efficiency for $(\Delta t)_{max} = 1\ \mu$s and $N_{min} = 2$.

Energy dependence of signal efficiency

efficiency as a function of energy for these parameters[15]. While there is no reason to assume so a priori, we note that the points $(\epsilon_s, \log(E))$ are well described by a straight line

$$\epsilon_s = -0.057 \cdot \log\left(\frac{E}{\text{GeV}}\right) + 1.11 \quad . \tag{26}$$

This decrease of the filter performance is due to the rejection of cascades which occur close to the edge of the fiducial volume. As the energy of the event increases, so does the total charge. The probability that one of the hits lies in the veto-layer thus increases along with the total number of hits. However, because this decrease scales logarithmically with energy, it only has a minor effect on the filter performance.

Energy dependence of background rejection

Similarly, we examine the energy dependence of the background passing fraction for the same filter parameters. The corresponding distribution is shown in Figure 37. This time, we note that the points $\left(\log\left(\frac{b}{b_{tot}}\right), \log(E)\right)$ can be approximated by a straight line

$$\log\left(\frac{b}{b_{tot}}\right) = -1.12 \cdot \log\left(\frac{E}{\text{GeV}}\right) + 3.64 \quad . \tag{27}$$

Therefore, the background passing fraction decreases by more than an order of magnitude when the energy of the event is increased by a factor of ten. As the background passing fraction thus changes much more rapidly with energy than the signal efficiency, the filter performance

15 Figures 36 and 37 show the filter performance per energy bin. In contrast, the values given in Figures 34 and 35 are obtained by integrating over all energies above a given cut-off.

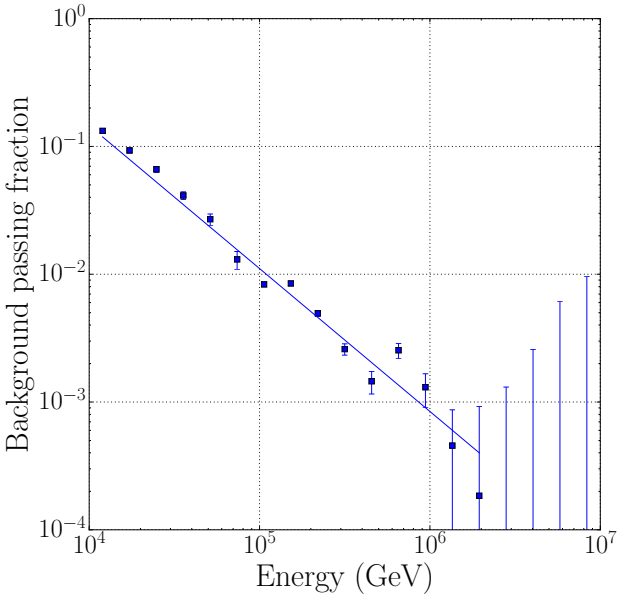

Figure 37: Energy dependence of the background passing fraction for $(\Delta t)_{max} = 1\ \mu$s and $N_{min} = 2$.

would increase at higher energies regardless of the change in the total rate of background and signal events.

Zenithal dependence

Finally, we also examine the zenithal dependence of the background passing fraction for the same filter parameters. Figure 38 shows this distribution for events with energies larger than 100 TeV. The background passing fraction is maximal for muons that are going straight down, decreasing along with the cosine of the zenith angle. Considering that the flux of muons increases with cos(zenith) [36], most passing background events will have a zenith angle close to zero. As such, the filter performance can be improved by extending the top veto-layer. In the next section, we therefore examine the filter performance for various configurations of the veto-layer.

6.5.3 *Comparison of veto-layer configurations*

Maximising the purity

To compare the filter performance for the different veto-layer configurations, we start by examining the maximal purity that can be obtained for an energy cut at 60, 100 and 200 TeV. Figure 39 displays these purities together with the corresponding number of signal events passing the filter per year. As expected, the larger the extend of the veto-layer, the larger the maximal purity that can be obtained. Extending the veto-layer however reduces the fiducial volume and thus also the number of signal events that can be observed. From Figure 39, we already observe that the improved purity when using a double edge does not outweigh the loss of signal events. In contrast, extending the number of DOMs in

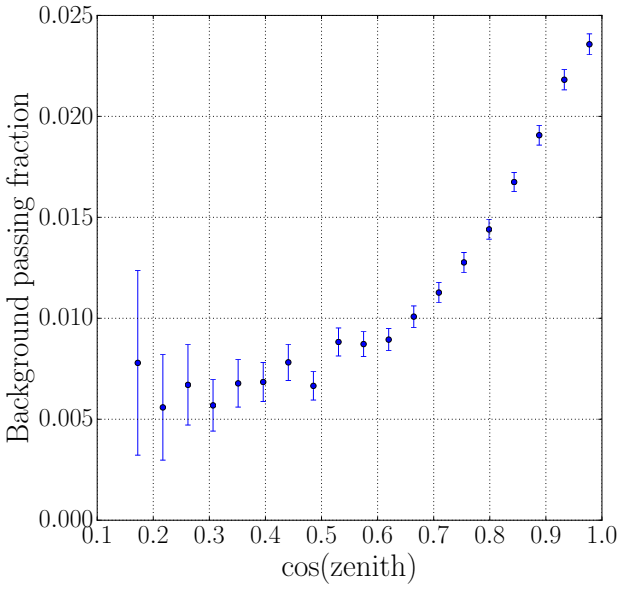

Figure 38: Zenithal dependence of the background passing fraction for $(\Delta t)_{max} = 1\ \mu$s, $N_{min} = 2$ and $E > 100$ TeV.

the top veto-region can more than double the obtained purity, while the loss of signal events is only minor. This is consistent with the results of Figure 38, which indicated that most passing background events entered via the top of the detector.

Maximising the number of signal events

For a more quantitative comparison, Tables 6, 7 and 8 show the maximal number of signal events than can be obtained when requiring the purity to be larger than a given value. Each table corresponds to a single energy cut, which is again taken at 60, 100 and 200 TeV. Missing values indicate that for that configuration, none of the filter parameters $((\Delta t)_{max}, N_{min})$ produced a purity larger than the required value. Examining the values, we confirm that for a given energy and purity, the double edge veto-layer never produces the best results if any of the other configurations reach the required purity. Using a top layer with six DOMs only provides the highest number of signal events for a high energy cut-off, provided that the required purity can be reached. The best results are thus obtained when including either 12 or 18 DOMs in the top veto-layer and a single layer of strings at the edge of the detector. Comparing these two configurations, we find only two instances in which using 18 DOMs results in a higher number of signal events[16]. By allowing the use of filter parameters that allow a higher signal efficiency, the number of signal events passing the filter can thus potentially be increased if the veto-layer is extended.

In addition to this analysis, we also compare the filter performance independent of the parametrisation of the cosmic neutrino flux. Figure

16 This is again provided that both configurations reach the required purity.

Table 6: Maximal number of signal events for an energy cut at 60 TeV.

Minimal purity (%)	20.0	50.0	70.0	90.0	99.0
6 top, single layer					
12 top, single layer	106.5				
18 top, single layer	107.3	79.1			
18 top, double layer	70.7	60.0			

Table 7: Maximal number of signal events for an energy cut at 100 TeV.

Minimal purity (%)	20.0	50.0	70.0	90.0	99.0
6 top, single layer	75.8				
12 top, single layer	79.1	74.5	56.8		
18 top, single layer	71.9	69.3	65.6		
18 top, double layer	48.8	47.1	45.0	38.0	

Table 8: Maximal number of signal events for an energy cut at 200 TeV.

Minimal purity (%)	20.0	50.0	70.0	90.0	99.0
6 top, single layer	47.5	46.6	43.8		
12 top, single layer	43.8	43.8	43.0	41.5	29.8
18 top, single layer	39.5	39.5	39.5	38.1	26.5
18 top, double layer	26.1	26.1	26.1	25.4	24.0

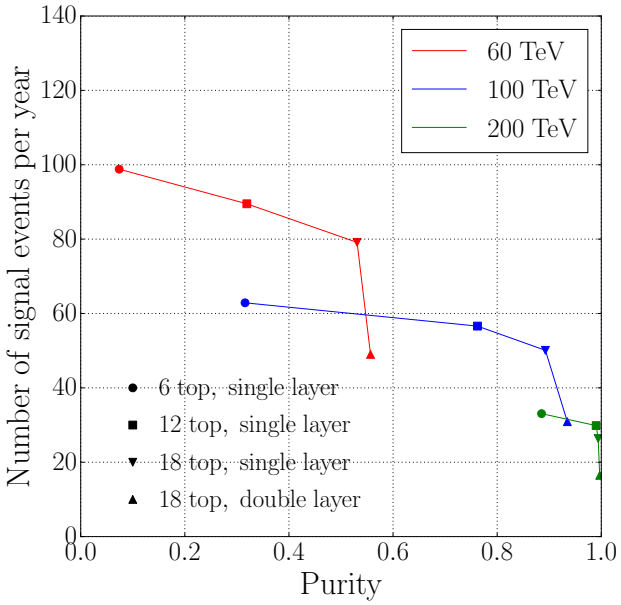

Figure 39: Number of signal events that are expected to be filtered per year for the filter parameters which maximise the purity.

Minimising the background

40 displays the minimal background passing fraction that can be obtained for the four veto-layer configurations and three energy cuts. Unlike Figure 37, the background passing fraction is not that of events with a certain energy E, but obtained by integrating over all energies larger than E. Again, we find poor performance when including only six DOMs in the top veto-layer, as the background passing fraction stays above 1% at energies E <200 TeV. By using 18 DOMs instead, ~95% of this background gets removed. At the same time, the fiducial volume and therefore the number of expected signal events is reduced by only 15% (see Table 5). Using a double edge reduces the passing background by another factor of ~3, though the fiducial volume decreases by 35%. Therefore, using a double edge is only beneficial at low energies, where the background largely dominates the signal rate. In addition, these results only apply for the filter parameters $(\Delta t)_{max} = 5$ μs and $N_{min} = 1$ hit, as they minimise the number of passing background events.

12 to 18 DOMs with single layer provides best performance

Based on these results, we conclude that including only six DOMs in the top veto-layer is no longer a viable option[17]. As most passing background events enter the top of the detector, the filter performance can greatly be improved by including 12 to 18 DOMs in the top veto-region. A further reduction of the background can be obtained by including a double layer of strings, though this considerably reduces the fiducial volume. Which veto-layer configuration is to be used will thus depend on the considered energy range of the selected events. For the energy cuts examined in this section, we find that using 12 DOMs in the top

17 Unless a very high-energy (~200 TeV) cut is imposed on the event sample.

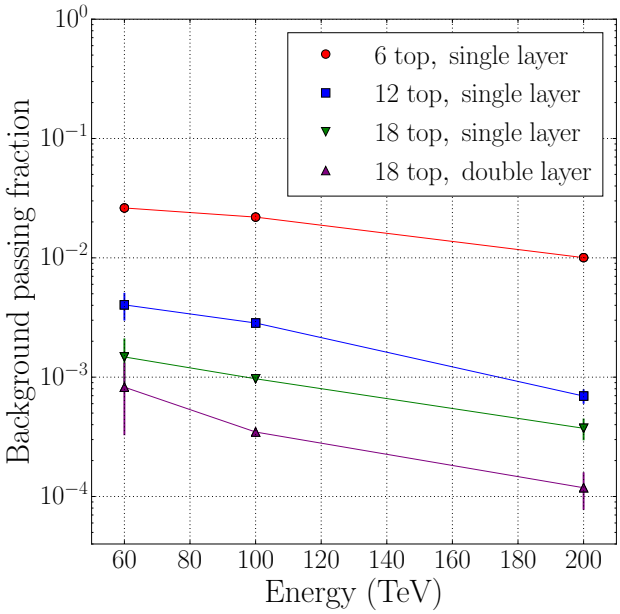

Figure 40: Minimal background passing fraction that can be attained for each of the energy cuts and veto-layer configurations.

veto-layer provides the best overall performance. In case lower cuts on the total energy are used, more extended veto regions will be needed to sufficiently reduce the background.

6.5.4 *Charge cut*

When using real data, imposing a cut based on the true energy of the event inducing particle is no longer an option. While advanced energy estimators can be used to reconstruct this true energy, a more straightforward approach is to instead base the cut on the total charge of the event. This is the method employed by the HESE analysis. As it allows a direct comparison with experimental data, we will now test its use for the causal filter in the Sunflower 240 geometry.

Correlation between charge and energy

Before blindly applying a charge cut, we first examine the relation between the energy and total charge of an event. In IC78, the uniform distribution of DOMs made the total charge a natural estimator for the energy of the event inducing particle. IceCube-GEN2 in contrast is composed out of the IC86 geometry, combined with new, more sparsely distributed strings. Additionally, the use of a newly designed DOM will also cause an asymmetry between the deposited charge in the GEN2 and IC86 strings. This latter remark however does not apply to the simulated datasets used here, as they are based on P-DOM[18]. Despite this asymmetry, Figure 41 indicates that the total charge is still a reasonable estimator for the energy of the event. In the case of background events

18 P-DOM houses the same PMT as DOMs in the IC86 geometry.

(top panel), we find that for a given energy bin, the charge is distributed over a wide range of values. This extended distribution is mostly caused by variations in the propagation distance of muons through the detector. For instance, high-energy muons can produce low-charge events if they only traverse one of the detector's corners. In contrast, signal events (bottom panel) are cascades which lie inside the fiducial volume. As a result, their charge increases to a better approximation linearly with the particle's energy.

Weighted vs unweighted charge cut

To obtain a better energy estimator, the charge can be weighted to account for the changing density of the DOMs. This then allows selecting events based on a variable which more closely resembles the true energy. However, the IC86 part of the detector will be sensitive to lower energy events than the GEN2 extension. As such, a lower energy cut could be used for events occurring in the IC86 part of the detector. Imposing a variable energy cut is automatically achieved when basing the selection on the total unweighted charge of the event. For events that are entirely contained in either the IC86 or GEN2 part of the detector, the total unweighted charge is thus a more natural parameter to cut on than the reconstructed energy. A downside to this approach is the background imposed by muons entering via GEN2 and passing through IC86. These muons can deposit little to no hits when entering the detector and still have a large total charge nonetheless. For this subclass of events, it would thus be more natural to base the cut on the weighted charge of the event.

Loss of high-energy events

Based on the results from Chapter 5, we know that a charge cut at 6000 pe results in a satisfactory filter performance for events in IC86. Therefore, we now examine the filter performance when imposing the same charge cut in IceCube-GEN2. Figure 42 shows the fraction of signal events above a given energy that are lost due to a charge cut. For $E > 100$ TeV, 47% of all signal events have a charge below 6000 pe. This loss of signal events is however countered by events passing the charge threshold with energies down to 10 TeV. To determine what the optimal filter parameters $(\Delta t)_{max}$ and N_{min} are, we again consider the background rejection and purity as a function of the signal efficiency. From these distributions, shown in Figures 43 and 44, we find that the maximal purity which can be obtained is 73%[19]. Therefore, a charge cut at 6000 pe results in purities that are comparable to those of an energy cut at 100 TeV.

Comparison with Section 4.4

Using the filter parameters $(\Delta t)_{max} = 1$ μs and $N_{min} = 2$, we now examine the charge dependence of the filter performance. Figure 45 shows the rate at which background events pass the filter compared to the total rate of background events. These rates can be directly compared to those

19 For these filter settings, 59 signal events would be observed per year.

of Figure 15 (p. 31), where the passing background was estimated using IceCube data for a geometry with a 250 meter string spacing. While the two distributions overall seem to be consistent, a few critical remarks are in order.

- More than 95% of all background events whose charge is larger than 6000 pe have an energy below 100 TeV. In addition, low-energy background events passing the charge threshold correspond to muons who lose most their energy in large stochastic energy losses. Since these events can have features similar to starting events, they are more likely to pass the causal filter. Therefore, we note that the simulated datasets used here are insufficiently large to reliable estimate the background passing fraction for a charge cut at 6000 pe. This can already be seen from the erratic behaviour of the number of filtered events in Figure 45 and is one of the reasons why our main analysis is based on an energy cut instead.
- The rates determined based on IceCube data correspond to a detector that is about a factor of ten smaller than IceCube-GEN2. Determining the charge threshold for which the passing rate of events goes to zero in IceCube-GEN2 thus requires shifting the lines in Figure 15 upwards by a factor of ten. This will result in a higher estimated charge threshold.
- In Figure 15, the x-axis displays the total charge as measured in the IC86 detector. Figure 45 in contrast displays the total charge measured in the IC86 and GEN2 region of the detector. As such, events of the same energy will on average have a lower charge in Figure 45 than in Figure 15.
- To determine the background passing fraction in Section 4.4, background events were first identified as events which are rejected by the HESE filter when using the full IC86 geometry. Any true background events which produced less than three hits when entering the detector are thus neglected. This in turn can lead to an underestimation of the charge threshold. Above 6000 pe, this effect will however be small, as only a limited fraction of the HESE events corresponds to atmospheric background muons.

We thus conclude that more simulated data is needed to reliably estimate the performance of the causal filter when imposing a charge cut. Specifically, events with a high charge but low energy impose a background that is hard to remove. While only a small fraction of all low energy muons produces a charge that exceeds the threshold, the steep power law spectrum of atmospheric muons implies that these low energy events account for a significant fraction of the total background. As such, neglecting these events could lead to a substantial underestimation of the background imposed by atmospheric muons.

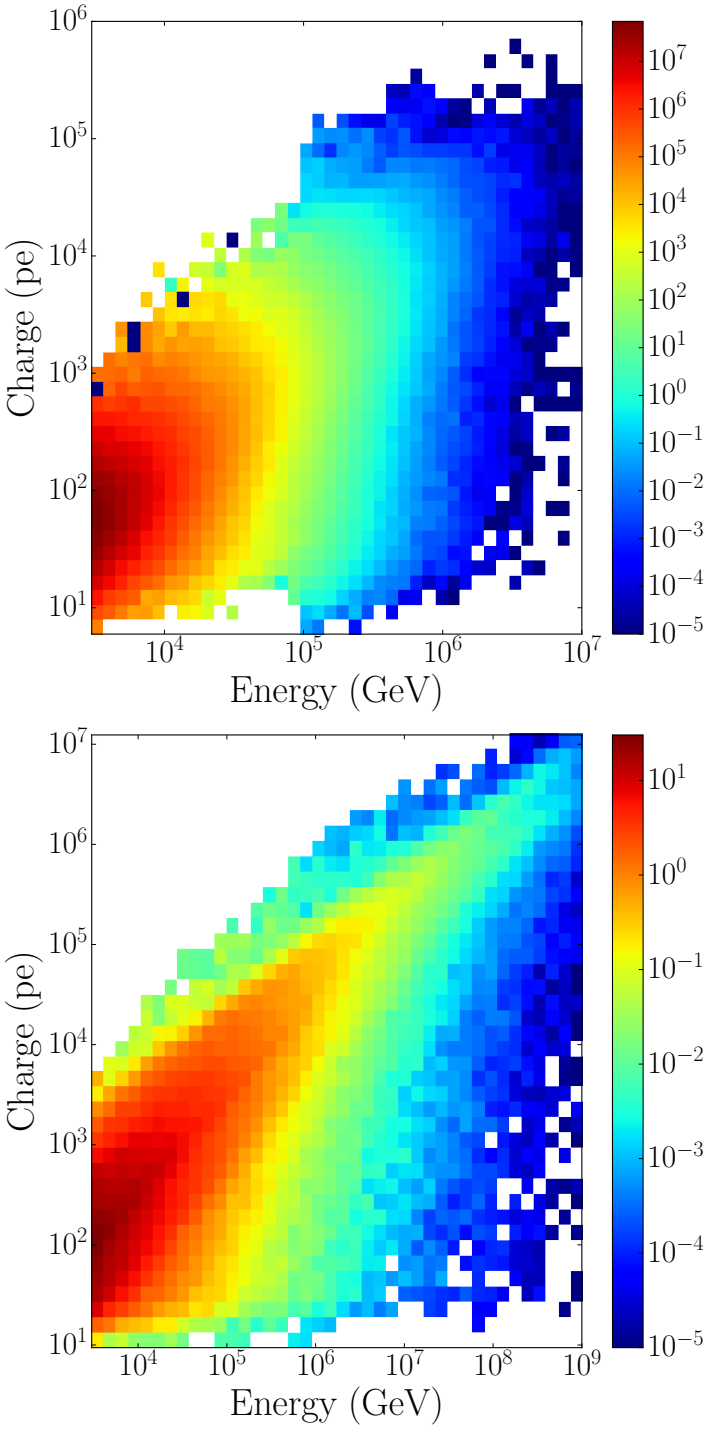

Figure 41: Charge as a function of energy for atmospheric muons (top) and cascades whose vertexes lie inside the fiducial volume (bottom). The colour-scale indicates the number of expected events per year per bin.

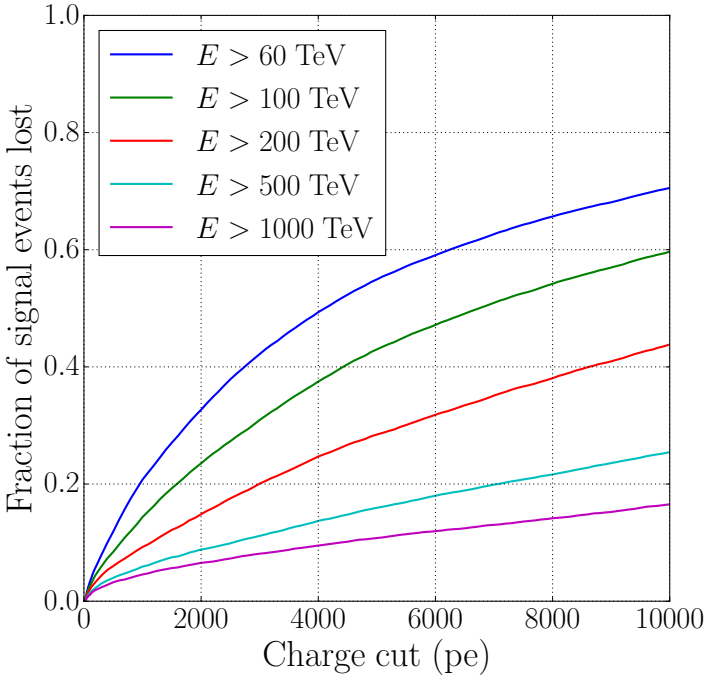

Figure 42: Fraction of signal events above a given energy that are lost by imposing a cut-off based on the total charge.

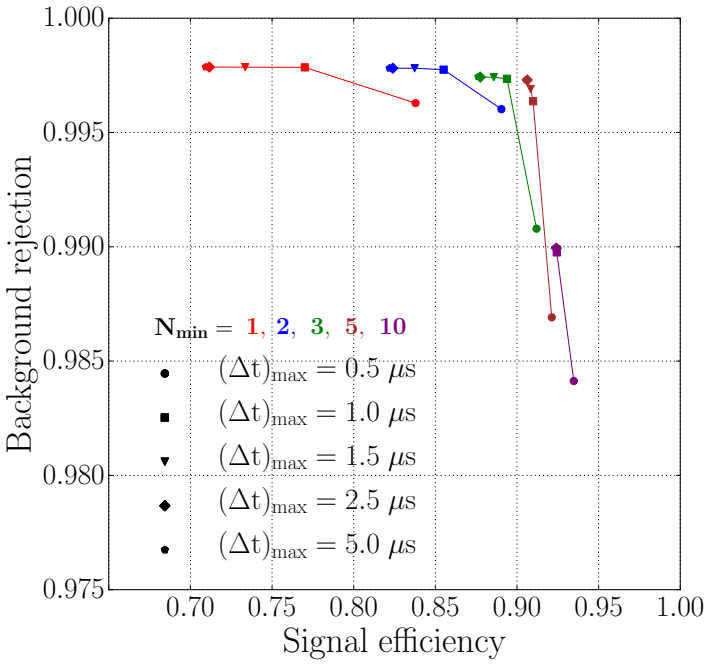

Figure 43: Background rejection as a function of signal efficiency when imposing a charge cut at 6000 pe.

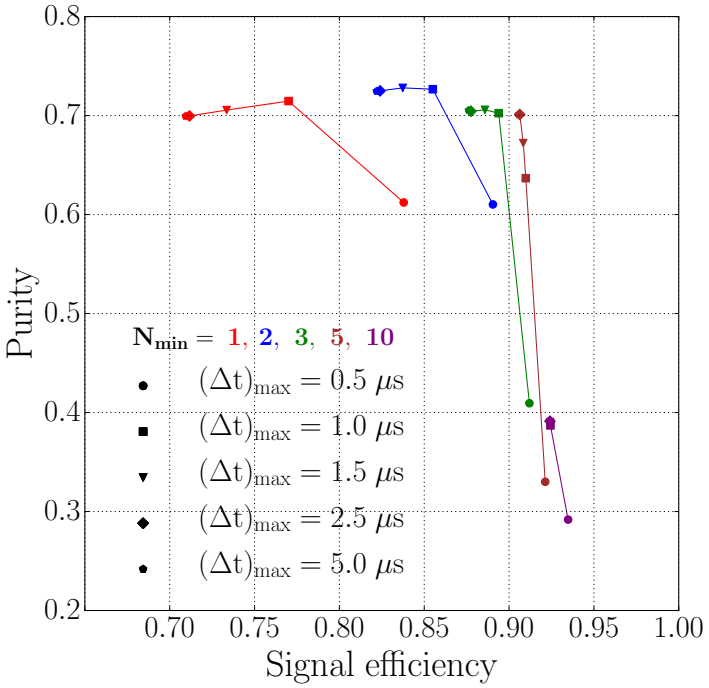

Figure 44: Purity as a function of signal efficiency when imposing a charge cut at 6000 pe.

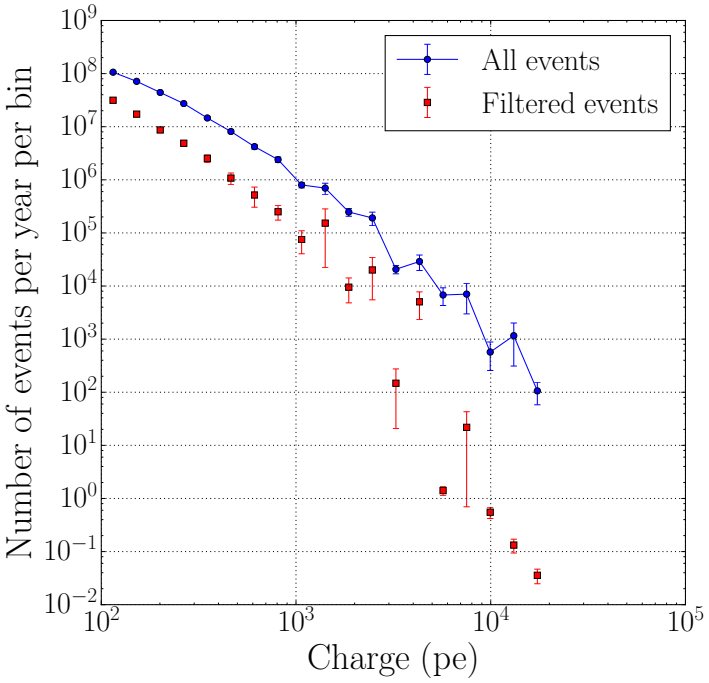

Figure 45: Total charge of all background events and those accepted by the causal filter for $(\Delta t)_{max} = 1$ μs and $N_{min} = 2$.

6.6 COMPARISON OF GEOMETRIES

In Section 6.5.3, the filter performance was tested for various veto-layer configurations in the Sunflower 240 geometry. This same analysis is now repeated for the baseline veto-layer configuration in the Sunflower and Edge-weighted geometries. Since dataset n° 2 in Table 3 is only available for the Sunflower 240 geometry, imposing a too large energy cut for the other geometries could lead to all background events being rejected by the filter. To reliable estimate the background rejection, the maximal energy cut used here is set to 100 TeV. This ensures that for each of the geometries and filter parameters, at least ten simulated background events are accepted by the filter. The performance will also be examined for two lower energy cuts, respectively placed at 50 and 70 TeV.

Maximising the purity

For each of the geometries and energy cuts, Figure 46 shows the number of expected signal events per year for the filter parameters which maximise the purity. Comparing the three Sunflower geometries at a given energy cut, we find that the purity varies more rapidly than the number of signal events. That is, if the purity is increased by a factor two, the number of signal events reduces by less than half. Another striking feature is the poor performance of the Edge-weighted geometry. As discussed in Section 6.5.2, most background events which pass the filter enter the detector via the top veto-layer. As a result, the background rejection of the Edge-weighted geometry will only slightly improve upon that of the Sunflower 240 geometry, as both have the same generic string spacing. The larger density of strings at the detector edge however causes the Edge-weighted geometry to have a fiducial volume which is 23% smaller than that of the Sunflower 240 geometry. The resulting loss of signal events is clearly visible in Figure 46 and can cause the purity to be lower than in the Sunflower 240 geometry. While the Edge-weighted geometry might thus have been designed specifically to improve the performance of InIce self-veto algorithms, we find that the improved background rejection does not outweigh the reduction of the fiducial volume.

Maximising the number of signal events

As a second means of comparison, we determine the maximal number of signal events that can be obtained when requiring that the purity exceeds a given value. Tables 9, 10 and 11 respectively show these signal rates for energy cuts at 50, 70 and 100 TeV. Again, we find that the Edge-weighted geometry has a poor performance, as its values never exceed those of the Sunflower 240 geometry. Comparing the Sunflower geometries, we find that in general, if two geometries both reach a given purity at a certain energy cut, the largest number of signal events will occur for the largest geometry. Large differences are however present in the attainable purities, as could already be seen in Figure 46. The largest geometry is therefore not automatically the optimal choice, as

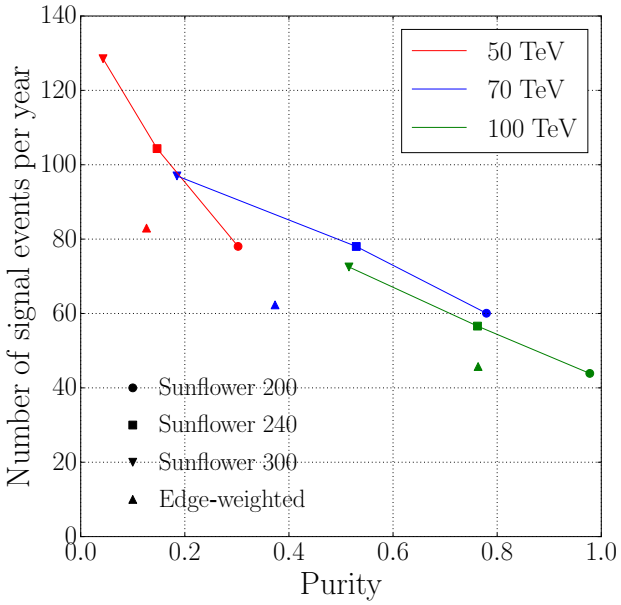

Figure 46: Number of signal events that are expected to be filtered per year for the filter parameters which maximise the purity.

obtaining the required purity would then necessitate a large energy cut. For instance, more filtered signal events occur at $p = 50\%$ for the Sunflower 240 geometry at $E > 70$ TeV than for the Sunflower 300 geometry at $E > 100$ TeV. Which of the Sunflower geometries gives the optimal results thus depends on the required purity and the energy range of the signal events which are considered.

Table 9: Maximal number of signal events for an energy cut at 50 TeV.

Minimal purity (%)	10.0	30.0	50.0	70.0	90.0
Sunflower 200	96.8	78.0			
Sunflower 240	114.4				
Sunflower 300					
Edge-weighted	88.8				

Table 10: Maximal number of signal events for an energy cut at 70 TeV.

Minimal purity (%)	10.0	30.0	50.0	70.0	90.0
Sunflower 200	77.1	74.9	72.4	60.1	
Sunflower 240	105.6	94.9	78.0		
Sunflower 300	127.9				
Edge-weighted	82.2	67.0			

Table 11: Maximal number of signal events for an energy cut at 100 TeV.

Minimal purity (%)	10.0	30.0	50.0	70.0	90.0
Sunflower 200	58.0	56.7	55.2	55.1	46.0
Sunflower 240	81.4	77.3	74.5	56.8	
Sunflower 300	111.8	94.8	72.5		
Edge-weighted	63.0	61.8	59.8	55.4	

Minimising the background

To parametrise the filter performance independently of the number of expected signal events, we examine the background passing fraction as a function of energy. Figure 47 shows the minimal background passing fraction[20] that can be obtained in each of the four geometries. For the Sunflower 240 and Edge-weighted geometry, the values overlap within their statistical uncertainty. As such, the denser string configuration at the boundary of the Edge-weighted geometry only has a small effect on the self-vetoing capabilities of the detector. The Sunflower 240 geometry is therefore the preferred choice of the two, as it has a fiducial volume which is 30% larger. In case of the Sunflower geometries, we find that decreasing the string spacing from 300 to 200 meter improves the background rejection by more than an order of magnitude. At the same time, the fiducial volume of the detector decreases by 43%. Therefore, the improved background rejection of the Sunflower 200 geometry will outweigh its smaller size if high purities are required. Apart from switching to a smaller string spacing, we remark that the passing rates in Figure 47 can be reduced by a factor ~3 if 18 instead of 12 DOMs are included in the top veto-layer. In addition, we find that both with respect to the background rejection and size of the fiducial volume, the Sunflower 200 geometry outperforms the Sunflower 240 geometry with a double edge and 18 DOMs in the top veto-layer. While extending the veto-layer can thus significantly improve the background rejection, it does not offer an alternative to the results that can be obtained with a smaller string spacing.

We thus conclude that

- the Edge-weighted geometry should not be considered, as it offers inferior performance compared to the Sunflower 240 geometry.
- the Sunflower 300 geometry will result in the largest number of signal events at a given energy and purity, provided that the required purity can be reached.

20 The background passing fraction at energy E in Figure 47 corresponds to an energy cut at E. That is, the displayed value of the background passing fraction is obtained by integrating over all energies larger than E.

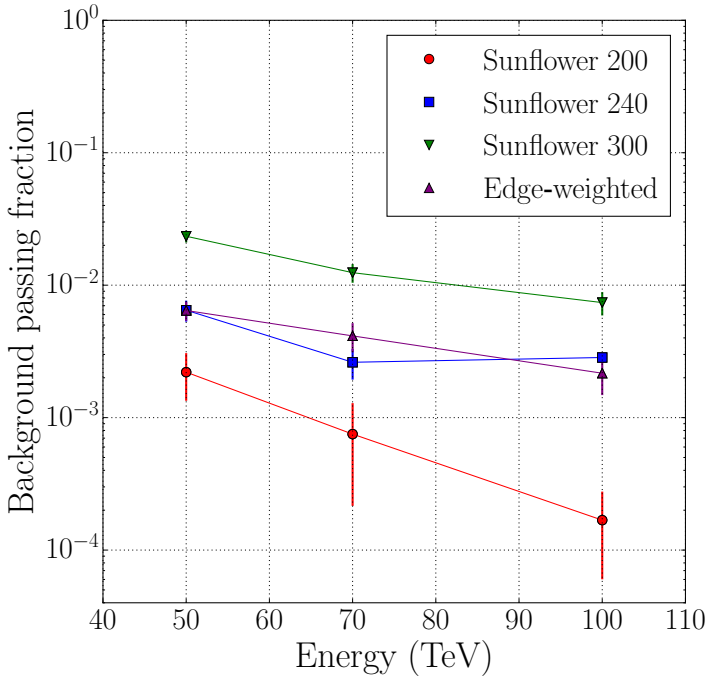

Figure 47: Minimal background passing fraction that can be attained for each of the energy cuts and proposed geometries. The 'kink' in the Sunflower 240 curve results from the use of dataset n° 2 in Table 3, which allows a more accurate determination of the background passing fraction for high energy cuts.

- the Sunflower 200 geometry results in the best background rejection, even when comparing it to the Sunflower 240 geometry with a more extended veto-layer.
- the Sunflower 240 geometry can be be seen as a comprise between minimising the background rejection and maximising the fiducial volume.

Considering that the Sunflower 300 geometry only reaches a maximal purity of 52% at an energy cut of 100 TeV, the Sunflower 200 and 240 geometries are the preferred choices. Which of these two offers the best performance then depends on the considered energy range and the purities required by the analyses.

7

CONCLUSION AND OUTLOOK

InIce self-veto techniques have played a vital role in the discovery and parametrisation of the cosmic neutrino flux. In the IceCube experiment, they are implemented by means of the HESE algorithm. Analyses using HESE events are however restricted to a small set of events, due to the limited rate at which cosmic neutrinos are detected. To increase this rate, an upgrade of the detector is currently under development that will extend its volume by roughly a factor of ten. As the HESE algorithm was specifically designed for IceCube, we in this thesis examined alternative self-veto algorithms that are better suited for larger scale detectors. IceCube data was therefore first used to test various filter concepts, from which the best performing candidate was selected. This filter algorithm was then used to compare the self-veto capabilities of the proposed IceCube-GEN2 geometries.

The first part of our research consisted of the development of new self-veto algorithms. These algorithms were designed to reject any event in which the number of hits at the edge of the detector and start of the event exceeds a given threshold. A hit selection procedure was therefore implemented, to select the subset of veto-layer hits which are due to entering particles. This selection was performed based on the time difference and causal connection between the veto-layer hits and a reference hit related to the actual event. Three different types of reference hits were examined for this purpose. By applying the filter to one year of IceCube data, the filter parameters and reference hit selection were optimised. From this analysis, we found the output of the newly developed causal filter to be overall consistent with that of the HESE-algorithm. Furthermore, the filter parameters could be tuned to reject events that were wrongly accepted by the HESE filter, while at the same time retaining all true signal events. In addition, we found that events could reliably be rejected based on the occurrence of a single hit in the veto-layer. This presents a notable improvement over the HESE filter, which requires at least three veto-layer hits to reject an event. Therefore, the causal filter has the potential to lower the energy threshold from which starting events can be distinguished from background events.

Since the causal filter was also designed to be scalable to larger detectors, it was ensuingly used to examine the self-veto capability of the proposed IceCube-GEN2 geometries. This was done based on two simulated datasets. Background events were simulated as atmospheric muons, weighted to the Gaisser cosmic ray flux model. Signal events were simulated as electromagnetic cascades occurring inside the fiducial volume

and were weighted to the HESE flux. To remove low-energy events, a cut was placed on the initial energy of the event inducing particle. Alternatively, we also examined a charge cut for the Sunflower 240 geometry, though our results were in that case limited by the size of the simulated background dataset. Optimising the filter performance was then done by varying the filter parameters and the configuration of the veto-layer.

For the Sunflower 240 geometry, four configurations of the veto-layer were examined. By comparing the obtained purity, number of passing signal events and background rejection obtained for each of them, we found that using only six DOMs in the top veto-layer no longer produces satisfactory results. If 12 to 18 DOMs are included instead, the background passing fraction decreases by over an order of magnitude. At the same time, only ~15% of otherwise accepted signal events are lost. Therefore, sufficiently reducing the atmospheric background requires a top veto-region that is more extended than that of the current IceCube detector. In contrast, we found that the improved background rejection that resulted from adding an additional layer of strings to the side of the veto-layer does not outweigh the corresponding loss of signal events.

For a veto-layer that includes 12 DOMs in the top region, we ensuingly compared the filter performance in the Sunflower 200, 240, 300 and Edge-weighted geometry. As most of the passing background events enter the top of the detector, no significant differences were found between the background rejection in the Sunflower 240 and Edge-weighted geometry. However, because the volume of the latter is 22% times smaller, the obtained results always favoured the Sunflower 240 geometry over the Edge-weighted geometry. Comparing the three Sunflower motifs, we found that very high energy cuts ($E \sim 200$ TeV) are required to obtain acceptable purities in the Sunflower 300 geometry. Overall, the Sunflower 200 and 240 geometries are thus the preferred choices. Which of these two configurations produces the optimal results then depends on the considered energy range and purity required by the analyses.

An important remark is that the absolute performance figures given in Chapter 6 most likely will differ from those of the geometry that will be realised in practise. The most relevant factors that will attribute to this difference are listed below.

- Our results implicitly rely on the various models that were used to simulate and weight the events.
- During the optimisation of the filter, the background imposed by atmospheric neutrinos was neglected, as it would not influence the filter parameters.
- The simulated datasets assumed that the IceCube-GEN2 extension can be realised without imperfections (e.g. differences in the deployed depth of strings, faulty DOMs, etc.).

- Using a DOM design other than P-DOM would result in more photons being observed. Therefore, this would positively influence the performance of the causal filter. In addition, the directional information from DOMs with a multi-PMT design could be used to improve the filter performance.
- Implementing the filter for real data requires placing a cut on the charge or estimated energy of the event. Here, we instead used a cut based on the true energy of the event inducing particle.

Nonetheless, the analysis given in the previous chapter gives a good first indication of the self-veto capability that can be expected of IceCube-GEN2. More importantly, it allowed a comparative study of the proposed geometries and veto-layer configurations.

Regarding the veto-layer, we found that examining passing events on an event-by-event basis indicated weak spots in its configuration. The filter performance could thus further be improved by fine-tuning the veto-layer to better cover certain areas. In addition, we remark that the inclusion of DOMs below the dust layer is only relevant in strings sufficiently close to the edge of the detector. As we included all DOMs within 80 meter below the dust layer, the number of signal events passing the filter could thus be further increased by excluding these DOMs from the veto-layer if they belong to the central strings. Finally, we remark that the configuration of the veto-layer largely determines the energy at which the filter becomes effective. Therefore, the configuration of the veto-layer could be based on the properties of the event. That is, to automatically use a more extended veto-layer in case of events with a low energy or charge.

To conclude, we have found that the InIce self-veto techniques currently used in IceCube can be extended to and remain highly relevant for the proposed IceCube-GEN2 extension. In contrast to filters based on the Earth's shielding, they are sensitive to high-energy ($E > 1$ PeV) neutrinos coming from the Southern sky and all types of neutrino flavours. As such, InIce self-veto techniques are an ideal candidate to further study the recently discovered cosmic neutrino flux with IceCube-GEN2. An alternative approach is to rely on a surface veto to tag all down-going events related to cosmic ray air showers. Whether a surface array can sufficiently reduce the background however remains to be seen, in contrast to the self-veto techniques, which have already proven their value. Even if the surface array proves to be too inefficient by itself, its information can be combined with the output of self-veto algorithms to improve upon the performance of the latter.

A

PERFORMANCE OF THE CAUSAL FILTER FOR ICECUBE DATA

A.1 EVENT VIEW OF SELECTED HESE EVENTS

During the discussion of the 19 'true' HESE events in Section 5.4.3, we noted that only two of those events have veto-layer hits occurring before the reference hit (see Table 2). To illustrate that in case of the first event these hits are noise induced, a visualisation of the event is given in Figure 48. Two veto-layer hits can be noted at $(x, y, z) \approx (150\text{ m}, 400\text{ m}, 100\text{ m})$, which are clearly unrelated to the event. In contrast, Figure 49 shows that for the second event, the two veto-layer hits occur at the very start of the track-like element. As the causal distance of both these hits to the reference hit is less than 0.5 μs, this second event was therefore most likely induced by a down-going muon.

A.2 INFLUENCE OF THE FILTER PARAMETERS FOR A COG-BASED REFERENCE HIT

In Section 5.4.4, we examined the influence of the parameters N_{min} and $(\Delta t)_{max}$ on the performance of the causal filter when using the T90-based reference hit. Here, we repeat the same analysis for both COG-based reference hits. Figures 50 and 51 respectively show the number of passing background events when using $h_{\text{COG},1}$ and varying N_{min} and $(\Delta t)_{max}$. Likewise, Figures 52 and 53 show the same distributions for $h_{\text{COG},2}$. Unlike in the case of the T90 analysis, none of the COG-based filter configurations reduce the background to events below 10^4 pe. Combined with the reasons listed in Section 5.4.3, we therefore conclude that using the T90 procedure results in a better filter performance than when using either of the COG-based reference hits.

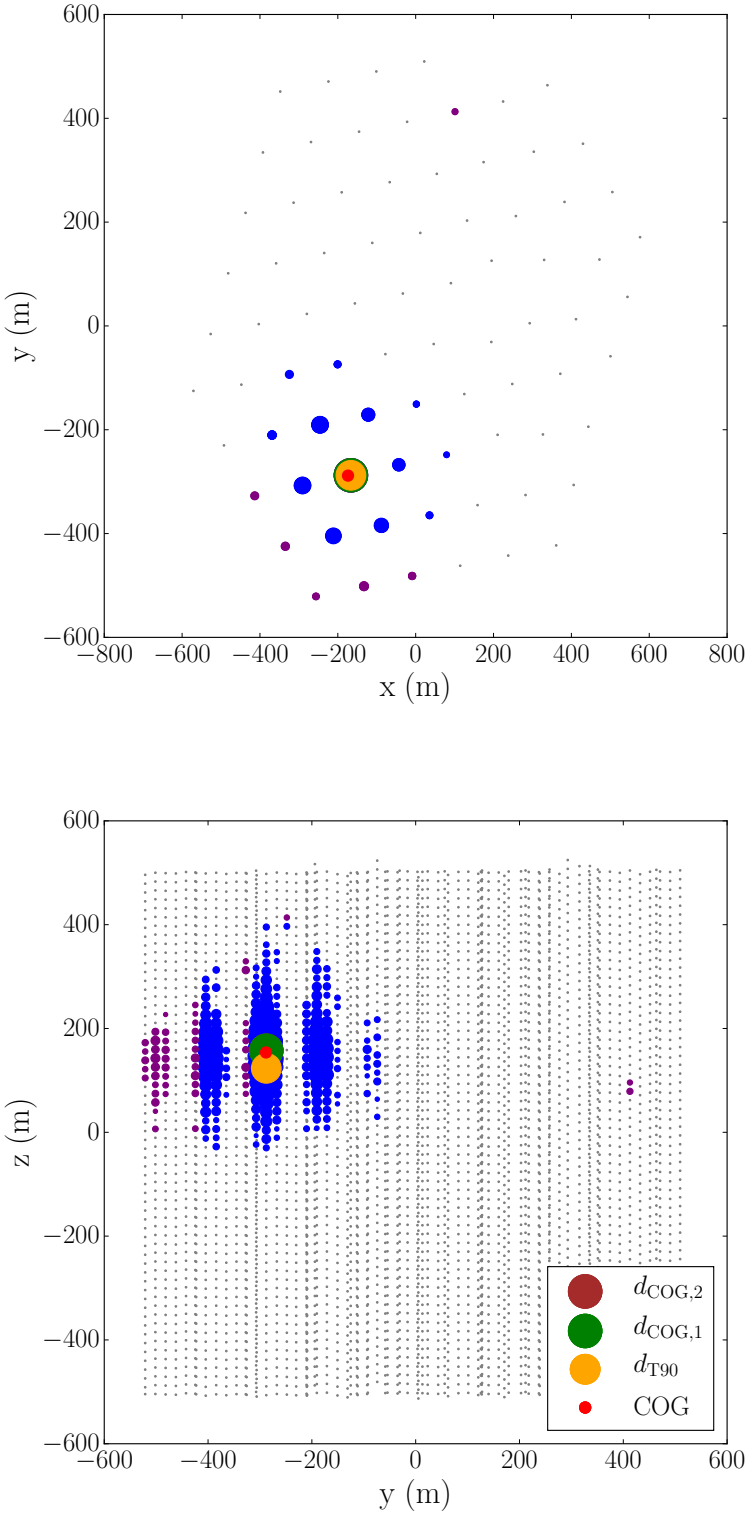

Figure 48: Visualisation of event n° 1 from Table 2. Colours are defined as per Figure 18.

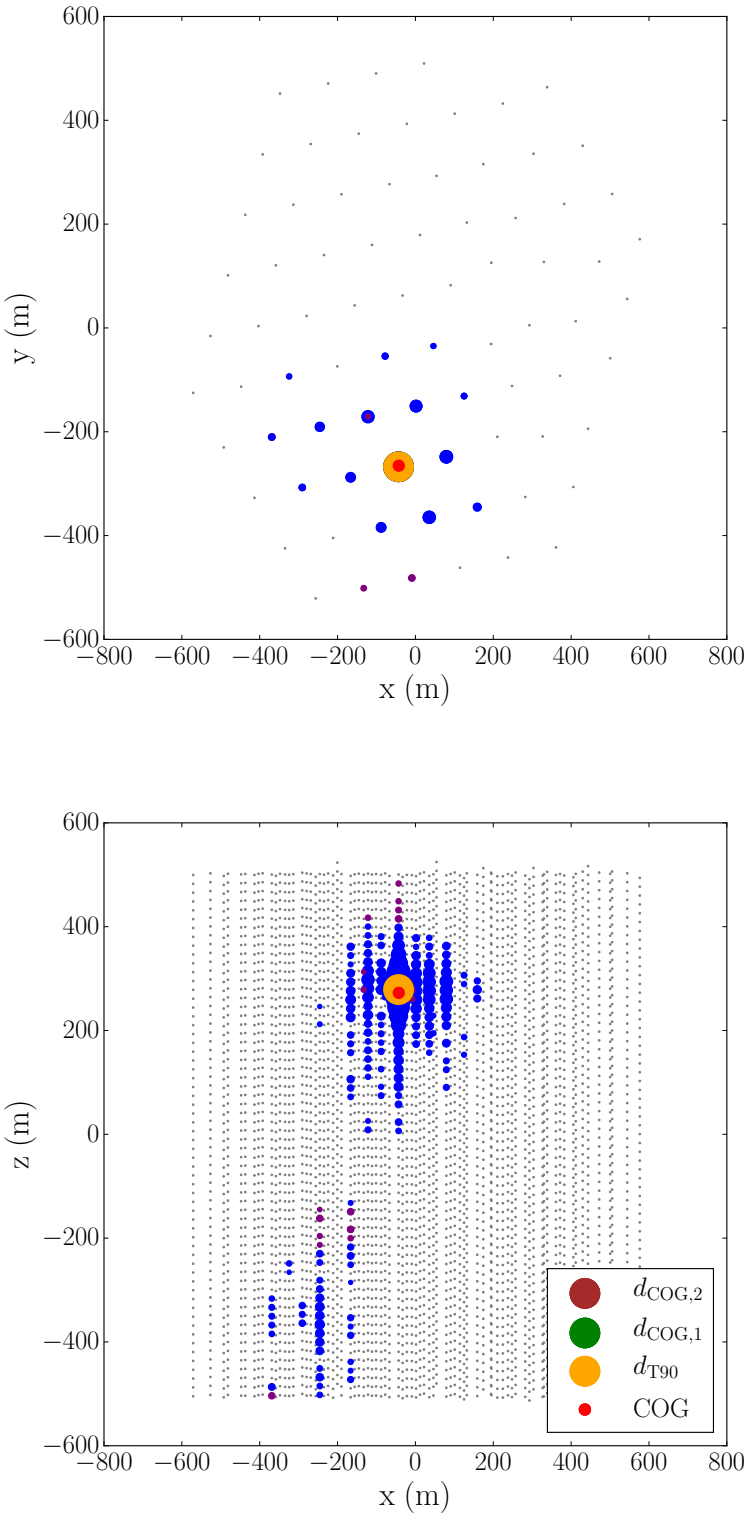

Figure 49: Visualisation of event n° 2 from Table 2. Colours are defined as per Figure 18.

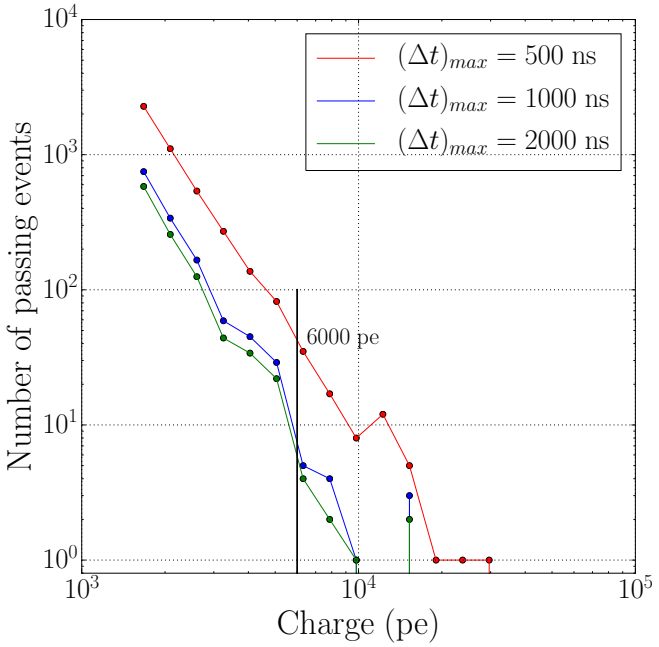

Figure 50: Number of background events passing the causal filter for $h_{\mathrm{COG},1}$, $N_{min} = 2$ and varying $(\Delta t)_{max}$.

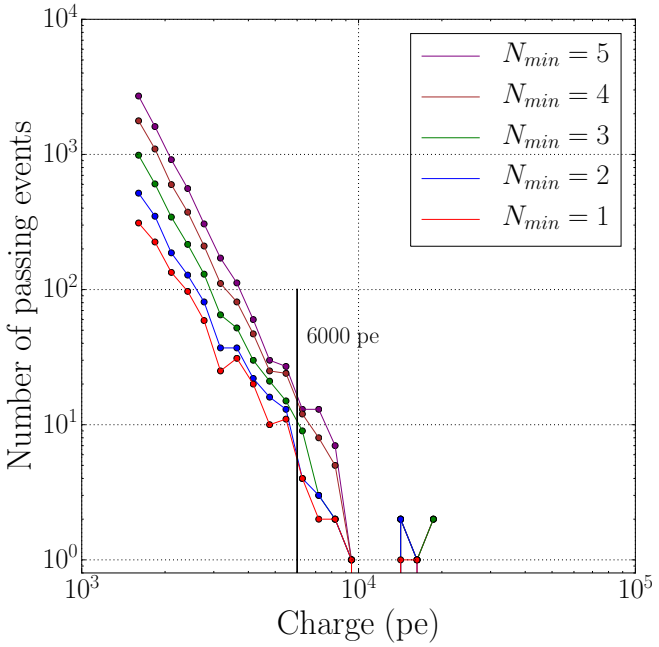

Figure 51: Number of background events passing the causal filter for $h_{\mathrm{COG},1}$, $(\Delta t)_{max} = 1$ μs and varying N_{min}.

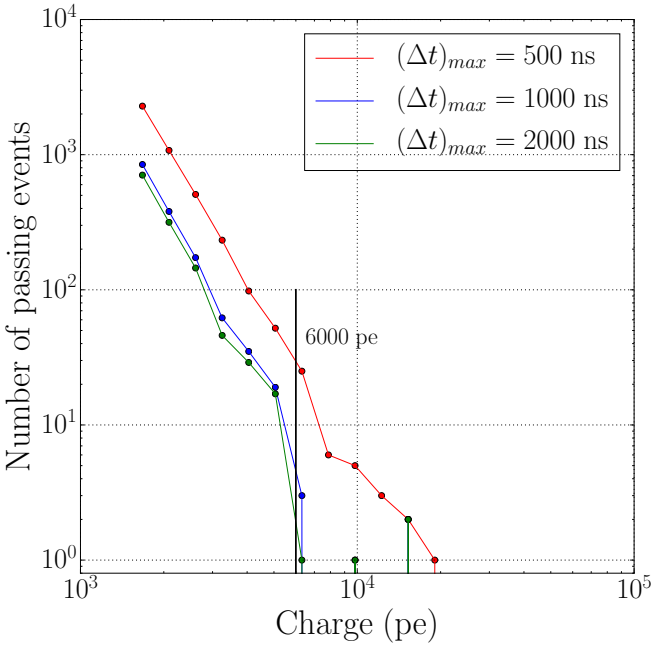

Figure 52: Number of background events passing the causal filter for $h_{\mathrm{COG},2}$, $N_{min} = 2$ and varying $(\Delta t)_{max}$.

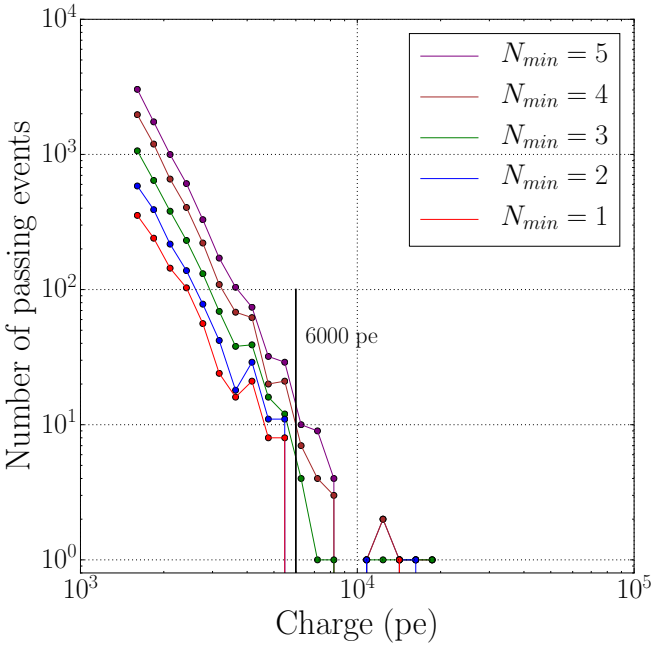

Figure 53: Number of background events passing the causal filter for $h_{\mathrm{COG},2}$, $(\Delta t)_{max} = 1$ μs and varying N_{min}.

B

RATE OF NEUTRINO INTERACTIONS

B.1 EVENT RATES

In this section, we derive the rate at which neutrino interactions depositing an energy E in the detector occur. Given a flux

$$\Phi = \frac{dN}{dt\, dA} \quad , \tag{28}$$

the rate of events $R = \frac{dN}{dt}$ can be obtained by integrating over the detector volume, which is represented by the so-called effective area A. For a detector consisting of N_{target} target particles, the rate R will be

$$R = \sigma \cdot N_{target} \cdot \Phi \quad , \tag{29}$$

where σ is the cross section between the incident and target particles. Analogously, when using a differential flux

$$\frac{d\Phi(E,\theta)}{dE\, d\Omega} = \frac{dN}{dt\, dA\, d\Omega\, dE} \quad , \tag{30}$$

the rate of events per unit energy is obtained by integrating over the detector volume and all solid angles

$$\frac{dR(E)}{dE} = \sigma \cdot N_{target} \cdot \int d\Omega\, \frac{d\Phi(E,\theta)}{dE\, d\Omega} \quad . \tag{31}$$

This recipe can now directly be applied on CC neutrino interactions to obtain the rate at which events with energy E are observed. In the case of NC interactions, a further modification to equation (29) is however needed to account for the fact that the initial neutrino energy E_i does not equal the deposit energy E. As is standard, we define $y \equiv \frac{E}{E_i}$ to be the fraction of energy lost by the neutrino. The rate of NC interactions which deposit an energy E in the detector can then be obtained by integrating over all initial neutrino energies $E_i > E$. This is implemented numerically in three steps. First, the number of neutral current interactions ΔN that neutrinos with initial energy $E_i \in [E_1, E_2]$ undergo is calculated

$$\Delta N = \int d\Omega \int_{E_1}^{E_2} dE_i\, \phi_d(E_i,\theta)\, N_{target}\, \sigma^{NC}(E_i) \quad . \tag{32}$$

The second step is to distribute these ΔN events over all energies $E = yE_i < E_i$, based on the differential cross section for NC interactions

$$\Delta N' = \Delta N \cdot \frac{1}{\sigma^{NC}(E_i)} \int_a^b dy\, \frac{d\sigma^{NC}(E_i,y)}{dy} \quad , \tag{33}$$

where $0 \leqslant a < b \leqslant 1$. These first two steps are then repeated until the entire range of initial neutrino energies E_i has been covered. In this way, we thus obtain the number of cascades that deposit an energy E due to neutral current interactions.

B.2 NEUTRINO-NUCLEON CROSS SECTION

Figure 54 displays the cross section for CC and NC neutrino and anti-neutrino interactions with nucleons, based on the parametrisation in section VII of [14]. In case of NC interactions, the average value $\langle y \rangle$ is also required to allow solving for a an b in equation (17). This distribution was therefore taken from [16], and is shown in Figure 55.

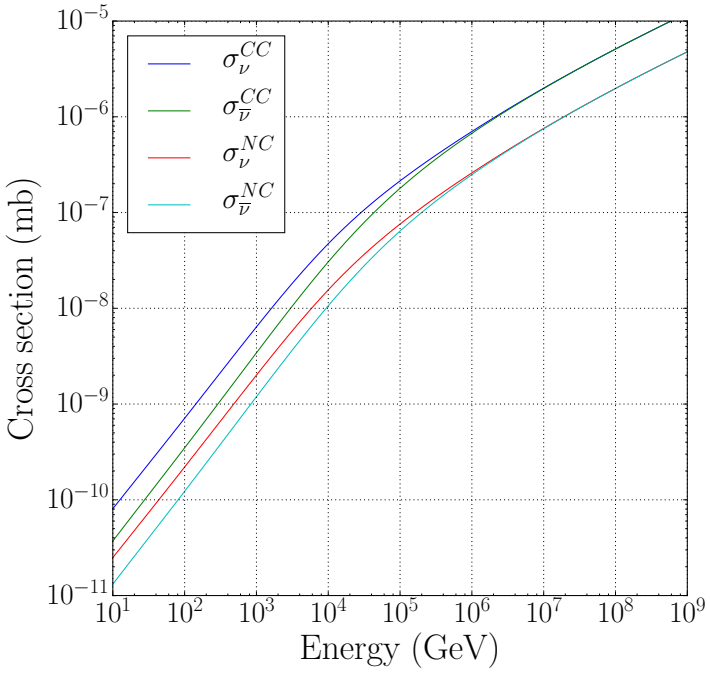

Figure 54: Energy dependent cross section of CC and NC (anti-)neutrino interactions with nucleons.

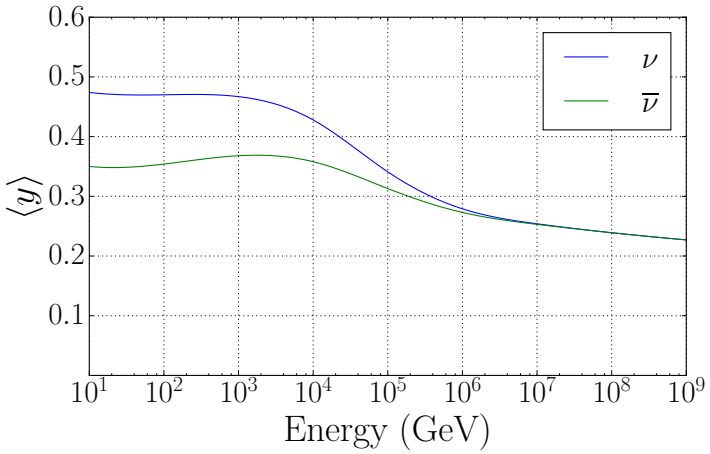

Figure 55: Distribution of the mean fraction of energy $\langle y \rangle$ deposited in NC interactions as a function of the initial energy of the (anti-)neutrino.

B.3 FILTER PERFORMANCE FOR THE BASELINE CONFIGURATION

For the Sunflower 240 geometry and baseline veto-layer configuration, we examined the performance of the causal filter as a function of the parameters $(\Delta t)_{max}$ and N_{min} in Section 6.5.2. Here, we include the distribution of the background rejection, signal efficiency and purity for energy cuts $E > 60$ TeV and $E > 100$ TeV. Due to the larger background at lower energies, the purity rapidly decreases when lowering the energy cut. At the same time, we see that using a higher energy threshold increases the background rejection and slightly decreases the signal efficiency.

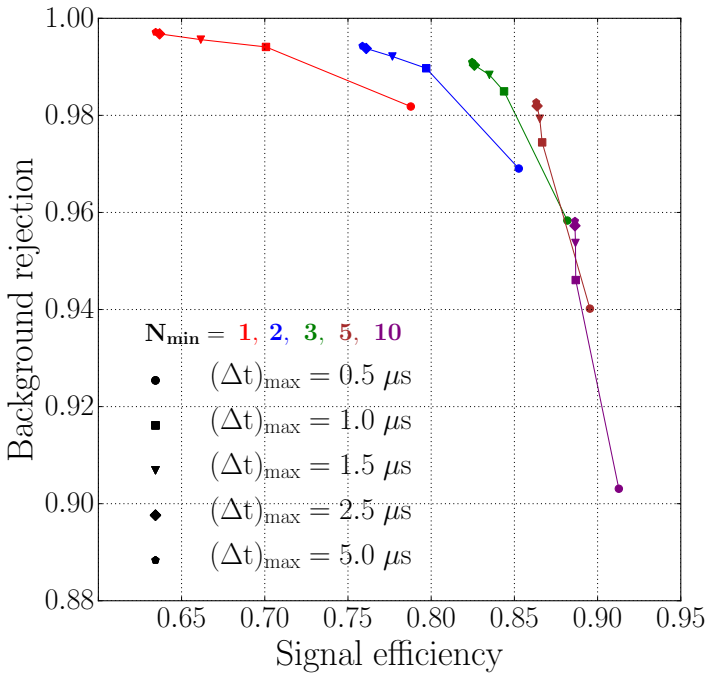

Figure 56: Background rejection as a function of signal efficiency when imposing an energy cut at 100 TeV.

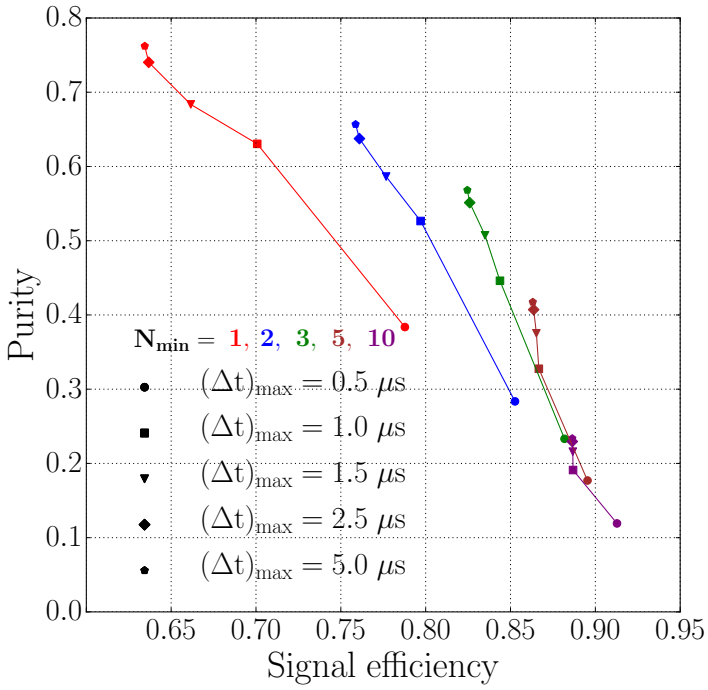

Figure 57: Purity as a function of signal efficiency when imposing an energy cut at 100 TeV.

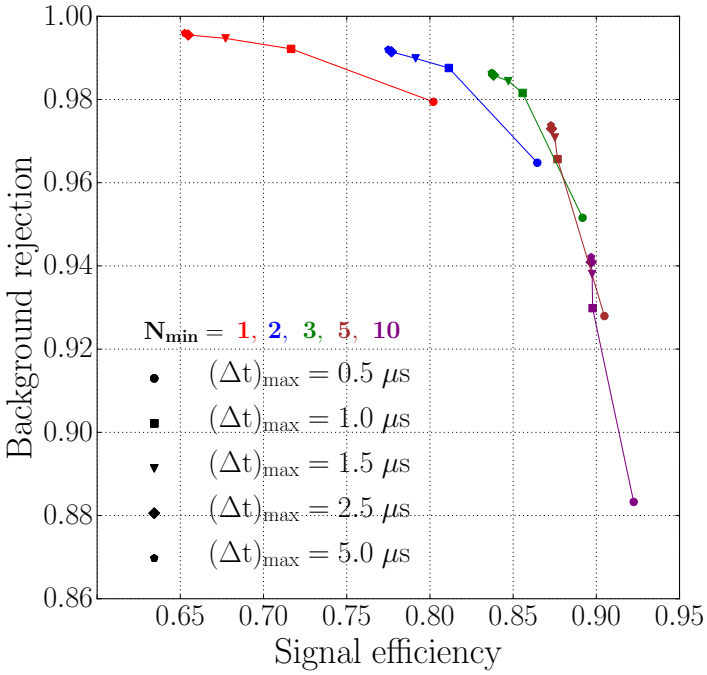

Figure 58: Background rejection as a function of signal efficiency when imposing an energy cut at 60 TeV.

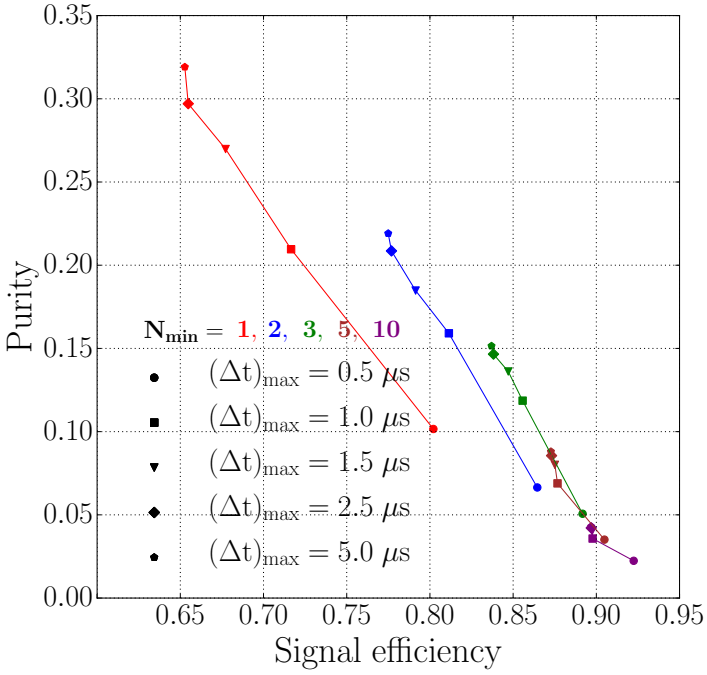

Figure 59: Purity as a function of signal efficiency when imposing an energy cut at 60 TeV.

BIBLIOGRAPHY

[1] M. Ackermann et al. Optical properties of deep glacial ice at the South Pole. *Journal of Geophysical Research: Atmospheres*, 111(D13):D13203, 2006.

[2] Antarctic Treaty Secretariat. Management Plan for Antarctic Specially Managed Area No 5. 2007. http://www.ats.aq/documents/recatt/Att357_e.pdf.

[3] ANTARES Collaboration. Antares: The first undersea neutrino telescope. *Nuclear Instruments and Methods in Physics Research A*, 656(1):11 – 38, 2011.

[4] ANTARES Collaboration, IceCube Collaboration, LIGO Scientific Collaboration, and Virgo Collaboration. High-energy neutrino follow-up search of gravitational wave event GW150914 with ANTARES and IceCube. *Physical Review D*, 93(12):122010, June 2016.

[5] ARA Collaboration. Design and initial performance of the Askaryan Radio Array prototype EeV neutrino detector at the South Pole. *Astroparticle Physics*, 35:457–477, February 2012.

[6] ARIANNA Collaboration. Time Domain Response of the ARIANNA Detector. *Astroparticle Physics*, 62:139–151, 2015.

[7] V. Aynutdinov. The BAIKAL-GVD project of km3-scale neutrino telescope in Lake Baikal. *International Cosmic Ray Conference*, 5:248, 2011.

[8] Y. Becherini, A. Margiotta, M. Sioli, and M. Spurio. A parameterisation of single and multiple muons in the deep water or ice. *Astroparticle Physics*, 25(1):1 – 13, 2006.

[9] F. Capozzi, E. Lisi, and A. Marrone. PINGU and the neutrino mass hierarchy: Statistical and systematic aspects. *Physical Review D*, 91:073011, 2015.

[10] K. de Vries, S. Buitink, N. van Eijndhoven, T. Meures, A. Ó Murchadha, and O. Scholten. The cosmic-ray air-shower signal in Askaryan radio detectors. *Astroparticle Physics*, 74:96–104, February 2016.

[11] J. J. Durek and G. Ekström. A radial model of anelasticity consistent with long-period surface-wave attenuation. *Bulletin of the Seismological Society of America*, 86:144–158, February 1996.

[12] FERMI Collaboration. The First Fermi-LAT Gamma-Ray Burst Catalog. *The Astrophysical Journal Supplement Series*, 209:11, November 2013.

[13] T. Feusels, J. Eisch, C. Xu, and for the IceCube Collaboration. Reconstruction of IceCube coincident events and study of composition-sensitive observables using both the surface and deep detector. *ArXiv:0912.4668*, December 2009.

[14] J. A. Formaggio and G. P. Zeller. From eV to EeV: Neutrino cross sections across energy scales. *Reviews of Modern Physics*, 84:1307–1341, July 2012.

[15] T. K. Gaisser, T. Stanev, and S. Tilav. Cosmic ray energy spectrum from measurements of air showers. *Frontiers of Physics*, 8:748–758, December 2013.

[16] R. Gandhi, C. Quigg, M. Hall Reno, and I. Sarcevic. Ultrahigh-energy neutrino interactions. *Astroparticle Physics*, 5:81–110, August 1996.

[17] S. L. Glashow. Resonant scattering of antineutrinos. *Physical Review*, 118:316–317, 1960.

[18] C. Grupen, G. Cowan, S. Eidelman, and T. Stroh. *Astroparticle Physics*. Springer Berlin Heidelberg, 2010.

[19] F. Halzen. High-energy neutrino astrophysics. *Nature Physics*, 2016.

[20] D. Hebecker, M. G. Archinger, S. Böser, J. Brostean-Kaiser, E. D. P. Rosendo, V. D. Lorenzo, M. DuVernois, P. J. Falke, C.-C. Fösig, T. Karg, L. Köpke, M. Kowalski, A. Looft, K. Sand, and D. Tosi. A wavelength-shifting optical module (wom) for in-ice neutrino detectors. *EPJ Web of Conferences*, 116:01006, 2016.

[21] J. R. Hörandel. On the knee in the energy spectrum of cosmic rays. *Astroparticle Physics*, 19:193–220, May 2003.

[22] IceCube and Pierre Auger Collaboration. Determining neutrino oscillation parameters from atmospheric muon neutrino disappearance with three years of IceCube DeepCore data. *Physical Review D*, 91(7):072004, 2015.

[23] IceCube and Pierre Auger Collaboration. 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. *Journal of Cosmology and Astroparticle Physics*, 1601(01):037, 2016.

[24] IceCube Collaboration. The IceCube Data Acquisition System: Signal Capture, Digitization, and Timestamping. *Nuclear Instruments and Methods in Physics*, A601:294–316, 2009.

[25] IceCube Collaboration. Calibration and characterization of the IceCube photomultiplier tube. *Nuclear Instruments and Methods in Physics Research A*, 618:139–152, June 2010.

[26] IceCube Collaboration. Search for Muon Neutrinos from Gamma-ray Bursts with the IceCube Neutrino Telescope. *The Astrophysical Journal*, 710:346–359, February 2010.

[27] IceCube Collaboration. The design and performance of IceCube DeepCore. *Astroparticle Physics*, 35:615–624, May 2012.

[28] IceCube Collaboration. Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector. *Science*, 342:1242856, 2013.

[29] IceCube Collaboration. First observation of PeV-energy neutrinos with IceCube. *Physical Review Letters*, 111:021103, 2013.

[30] IceCube Collaboration. Measurement of South Pole ice transparency with the IceCube LED calibration system. *Nuclear Instruments and Methods in Physics*, A711:73–89, 2013.

[31] IceCube Collaboration. Energy Reconstruction Methods in the IceCube Neutrino Telescope. *Journal of Instrumentation*, 9:P03009, 2014.

[32] IceCube Collaboration. Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data. *Physical Review Letters*, 113(10):101101, September 2014.

[33] IceCube Collaboration. Atmospheric and astrophysical neutrinos above 1 TeV interacting in IceCube. *Physical Review D*, 91(2):022001, 2015.

[34] IceCube Collaboration. Multipole analysis of IceCube data to search for dark matter accumulated in the Galactic halo. *European Physical Journal C*, 75:20, January 2015.

[35] IceCube Collaboration. Search for Prompt Neutrino Emission from Gamma-Ray Bursts with IceCube. *The Astrophysical Journal Letters*, 805:L5, May 2015.

[36] IceCube Collaboration. Characterization of the atmospheric muon flux in IceCube. *Astroparticle Physics*, 78:1–27, May 2016.

[37] IceCube Collaboration. Observation and Characterization of a Cosmic Muon Neutrino Flux from the Northern Hemisphere Using Six Years of IceCube Data. *The Astrophysical Journal*, 833:3, December 2016.

[38] IceCube Collaboration. PINGU: A Vision for Neutrino and Particle Physics at the South Pole. arXiv:1607.02671, 2016.

[39] IceCube Collaboration. All-sky Search for Time-integrated Neutrino Emission from Astrophysical Sources with 7 yr of IceCube Data. *The Astrophysical Journal*, 835(2):151, 2017.

[40] IceCube Collaboration. The Contribution of Fermi-2LAC Blazars to Diffuse TeV-PeV Neutrino Flux. *The Astrophysical Journal*, 835:45, January 2017.

[41] IceCube Collaboration. The IceCube Neutrino Observatory: Instrumentation and Online Systems. *Journal of Instrumentation*, 12(03):P03012, 2017.

[42] IceCube-Gen2 Collaboration. IceCube-Gen2: The Window to the PeV Universe. *In preparation*.

[43] IceCube-Gen2 Collaboration. IceCube-Gen2: A Vision for the Future of Neutrino Astronomy in Antarctica. arXiv:1412.5106, December 2014.

[44] IceCube-Gen2 Collaboration. IceCube-Gen2 - The Next Generation Neutrino Observatory at the South Pole. *Contributions to ICRC*, 2015.

[45] O. Kalashev, D. Semikoz, and I. Tkachev. Neutrinos in IceCube from active galactic nuclei. *Soviet Journal of Experimental and Theoretical Physics*, 120:541–548, March 2015.

[46] S. R. Klein. IceCube: A Cubic Kilometer Radiation Detector. *IEEE Transactions on Nuclear Science*, 56:1141–1147, 2009.

[47] T. M. Koshut, W. S. Paciesas, C. Kouveliotou, J. van Paradijs, G. N. Pendleton, G. J. Fishman, and C. A. Meegan. T_{90} as a Measurment of the Duration of GRBs. In *American Astronomical Society Meeting Abstracts #186*, volume 27 of *Bulletin of the American Astronomical Society*, page 886, May 1995.

[48] A. L'Abbate, T. Montaruli, and I. Sokalski. Effect of neutral current interactions on high energy muon and electron neutrino propagation through the earth. *Astroparticle Physics*, 23:57–63, February 2005.

[49] A. Lien et al. The Third Swift Burst Alert Telescope Gamma-Ray Burst Catalog. *The Astrophysical Journal*, 829(1):7, 2016.

[50] A. Margiotta. The KM3NeT deep-sea neutrino telescope. *Nuclear Instruments and Methods in Physics Research A*, 766:83–87, December 2014.

[51] S. C. Miarecki. Measurement of the neutrino-nucleon cross-section at multi-TeV energies with IceCube. Presented at the 38th ICHEP, August 2016.

[52] R. Mitalas and K. R. Sills. On the photon diffusion time scale for the sun. *The Astrophysical Journal*, 401:759, December 1992.

[53] K. Murase. Active Galactic Nuclei as High-Energy Neutrino Sources. arXiv:1511.01590, 2015.

[54] K. Murase, D. Guetta, and M. Ahlers. Hidden Cosmic-Ray Accelerators as an Origin of TeV-PeV Cosmic Neutrinos. *Physical Review Letters*, 116(7):071101, 2016.

[55] S. Palomares-Ruiz, A. C. Vincent, and O. Mena. Spectral analysis of the high-energy IceCube neutrinos. *Physical Review D*, 91(10):103008, May 2015.

[56] D. H. Perkins. *Particle Astrophysics, Second Edition*. Oxford Master Series in Physics. OUP Oxford, 2008.

[57] M. Santander. Anisotropy of TeV and PeV cosmic rays with IceCube and IceTop. *Nuclear Instruments and Methods in Physics Research A*, 725:85 – 88, 2013.

[58] SNO Collaboration. Measurement of the ν_e and total ^{8}B solar neutrino fluxes with the Sudbury Neutrino Observatory phase-III data set. *Physical Review C*, 87:015502, January 2013.

[59] Super-Kamiokande Collaboration. The super-kamiokande detector. *Nuclear Instruments and Methods in Physics Research A*, 501(2–3):418 – 462, 2003.

[60] Super-Kamiokande Collaboration. Solar Neutrino Measurements in Super-Kamiokande-IV. *Physical Review D*, 94(5):052010, 2016.

[61] S. Toscano. Neutrino astronomy at the South Pole: Latest results from the IceCube neutrino observatory and its future development. *Nuclear Instruments and Methods in Physics Research A*, 2017.

[62] S. Toscano, K. de Vries, and N. van Eijndhoven. HYBRID detection of PeV neutrinos at the South Pole: Simulation Plan. October 2016.

[63] R. J. Turner and S. S. Shabala. Energetics and Lifetimes of Local Radio Active Galactic Nuclei. *The Astrophysical Journal*, 806:59, June 2015.

[64] A. C. Vincent, S. Palomares-Ruiz, and O. Mena. Analysis of the 4-year IceCube high-energy starting events. *Physical Review D*, 94(2):023009, July 2016.

[65] W. Winter. Atmospheric neutrino oscillations for earth tomography. *Nuclear Physics B*, 908:250 – 267, 2016.

[66] L. Wolfenstein. Neutrino Oscillations in Matter. *Physical Review D*, 17:2369–2374, 1978.