Vrije Universiteit Brussel

Faculty of Science & Bio-Engineering Sciences
Department of Physics

A search for dark matter from the Sun with the 59-string IceCube neutrino detector

Proefschrift ingediend met het oog op het behalen van de graad van
"Master in de Wetenschappen"

Jan Kunnen

Promotor: Prof. Dr. Catherine De Clercq
Co-promotor: Dr. Erik Strahler

June 2011

Acknowledgements

Although this is one of the first pages in my thesis, it is the very last part that I wrote. The advantage of waiting till the very end to write this part is that I can really think about who to thank without having anything else on my (almost empty) mind.

First of all I want thank the Interuniversity Institute for High Energies (IIHE) of the Vrije Universiteit Brussel-UniversitÃĭ Libre de Bruxelles. There are many reasons to thank the institute, but I'm especially grateful that the institute gave me the opportunity to travel to Madison (WI) 2 times, where I had the chance to find my way in the IceCube collaboration. Bedankt Marleen Goeman om dat je altijd alles zo goed regelt.

I also thank the IceCube collaboration and in particular all the people of the Brussels IceCube group. This group, or should I say team, is so nice that I don't want to leave it. Therefor I'm very happy that I can stay 4 more years to do my PhD! I thank Debanjan, Mark and David for the nice badminton games, Sabrina and Mathieu to remind me how to speak French and Thomas, Kael and Stijn because they are nice :). Also a thanks to Alberto for taking good care of students (all of us). Dankjewel Nick voor alle hulp en voor het vertrouwen dat je in mij had toen je me naar Amerika stuurde. En natuurlijk ook bedankt voor de Havanna's :).

A special thanks goes to my promotor, Prof. Dr. Catherine De Clercq and my co-promotor, Dr. Erik Strahler. Erik, really thanks for all the help you gave me during the last months. It's nice to work with you, so I'm looking forward to the next year(s). Bedankt Catherine voor alle mogelijkheden die je mij gegeven hebt : een leuk thesis onderwerp, een zomer in de VS, de kans om te doctoreren, ... Ik apprecieer ook enorm hard dat je je tijd hebt genomen om mijn thesis grondig door te nemen en te verbeteren, zodat ik nu (hopelijk) een mooie versie kan indienen.

Many many many... thanks to Alfio Rizzo, for helping me A LOT with the technical computing parts of my analysis. I couldn't have done this in such a short time without your help, so THANK YOU ALFIO!

Dan zijn er natuurlijk ook de mensen die mij al jaren steunen : de vrienden en familie :). Waar zal ik beginnen...?

Bedankt aan alle klasgenoten (allemaal = 4) voor de aangename tijd die we samen doorgebracht hebben de laatste jaren.

Vermits de rest van mijn vrienden dit niet zullen lezen, kan ik die groep al overslaan, hetgeen mij bij de familie brengt (die ik zal verplichten om toch ten minste dit stuk te

lezen).

Danku mama en papa om mij altijd zo goed te verzorgen en mij te steunen door dik en dun. Het is dankzij jullie vertrouwen en steun dat ik al zover geraakt ben en ik ben jullie daar enorm dankbaar voor. Dus nogmaal, bedankt mama en papa!

Bedankt ook aan de (toekomstige) schoonfamilie. Chiara, bedankt dat ik in jou 5 sterren hotel mocht verblijven de voorbije weken. De service is uitstekend en de aangename omgeving heeft mij toegelaten om in alle rust aan deze te werken.

Ten slotte wil ik ook mijn vriendin Leyla bedanken. Liefje, ik besef heel goed dat ik vaak ondraaglijk geweest ben de laatste weken. Ik apprecieer dan ook enorm dat je mij hebt verdragen en daarenboven nog eens lief en zorgzaam met mij was. Bedankt voor alles #hvj.

Samenvatting

Uit verscheidene experimenten blijkt dat ons heelal doordrongen moet zijn van een onbekende soort materie. Deze zogenaamde *donkere materie* is één van de grote mysteries in de moderne fysica. Volgens recente waarnemingen zou de donkere materie zorgen voor maar liefst 23% van de totale energie-dichtheid in ons Universum, terwijl de zichtbare materie slechts bijdraagt voor 4%. Massieve, zwak interagerende deeltjes (WIMPs) zouden de discrepantie in de dichtheid kunnen oplossen.

De Minimale Suprsymmetrische uitbreiding van het Standaard Model (MSSM) is een hypothetische theorie waarin elk deeltje uit het standaars model een supersymmetrische partner heeft. Deze theorie postuleert dus allerlei nieuwe (tot vandaag ongeobserveerde) deeltjes, waaronder het neutralino. Dit neutralino χ is één van de beste WIMP kandidaten.

Als deze neutralino's bestaan, zullen zij gravitationeel aangetrokken worden door massieve objecten, zoals onze zon, waardoor er een ophoping aan neutralino's zal ontstaan in de kern van de zon. Eens de dichtheid van deze neutralino's groot genoeg is, zullen zij paarsgewijs annihileren en standaard model deeltjes produceren. Deze standaard model deeltjes kunnen in hun verval o.a. neutrinos produceren. De geproduceerde neutrinos kunnen ontsnappen uit de zon en, als ze in de richting van de aarde komen, kunnen ze gedetecteerd worden door neutrino detectors zoals bijvoorbeeld IceCube. Onze analyse werd geoptimaliseerd voor neutralino's met een massa van 1000GeV die annihileren in W^+W^-.

In deze thesis hebben we data gebruikt die genomen werd door de IceCube detector wanneer deze uit 59 kabels bestond, dit was in 2009-2010. Vandaag is de detector af en bestaat hij uit 86 kabels. Hoewel de IceCube detector is gebouwd om neutrinos te detecteren, wordt de data gedomineerd door atmospherische muonen. Het eerste deel van onze analyse bestond eruit al deze atmosferische muonen weg the filteren. Hiervoor gingen we in 2 stappen te werk. Eerst werd er een filter ontwikkeld door manuele snedes aan te brengen op 5 variabelen waarvoor er een duidelijk verschil waarneembaar was tussen signaal en achtergrond muonen. Deze filter reduceerde de hoeveelheid atmosferische achtergrond muonen met een factor $\sim 10^2$, terwijl slechts $\sim 25\%$ van het gesimuleerde signaal weggesneden werd. In de tweede filter stap werd gebruik gemaakt van een multivariate techniek, de zogenaamde *Boosted Decision Trees* (BDT). Het doel van deze BDT filter was om het grootste deel van de atmosferische achtergrond muo-

nen weg te filteren en toch zoveel mogelijk signaal over te houden. Na de BDT filter waren bijna geen atmosferische muonen meer over en was de data gedomineerd door atmosferische achtergrond neutrino's.

De resterende experimentele gegevens (voornamelijk atmosferische neutrino's) werden dan gebruikt om de *gevoeligheid* van de detector voor ons WIMP signaal in te schatten. De gevoeligheid van een detector wordt gegeven door de bovenlimiet op het aantal signaal eventen in een *'background-only'* scenario, d.w.z. in het geval dat er geen echt signaal aanwezig is in de geanalyseerde data. Door statistische fluctuaties is het mogelijk dat het lijkt alsof er signaal gemeten wordt, zelfs als er geen signaal is. Vooraleer we kunnen besluiten dat het gemeten signaal wel degelijk fysisch signaal is en geen statistische fluctuatie is, moeten we zeker zijn dat het gemeten signaal groter is dan de *gevoeligheid* van de detector voor ons signaal.

Nadat we deze gevoeligheid bepaald hadden, konden we gevoeligheden berekenen voor een aantal fysische grootheden zoals het neutralino annihilatietempo, de neutrino-geïnduceerde muon flux en de spin-afhankelijke en -onafhankelijke neatralino-proton werkzame doorsneden.

Globaal gezien presteert onze analyse zeer goed. De gevoeligheid die we bereiken voor neutralino's met een massa van 1000GeV die annihileren in W^+W^- is beter dan de voorgaande analyses, waarbij data van de IceCube detector gebruikt werd wanneer deze nog uit 22 kabels bestond. Onze analyse werd gedaan met 1 maand data. Een volgende stap is om de data set uit te breiden naar 6 maanden, waardoor we 6 keer meer statistiek zullen hebben, hetgeen ons zal toelaten om onze gevoeligheid te verbeteren.

Contents

Chapter 1

Introduction

Almost 80 years ago, an astronomer, Zwicky, estimated the mass of the Coma cluster of galaxies, based on the motions of galaxies near its edge and compared that estimate to one based on the number of galaxies and total brightness of the cluster. The amount of mass he estimated was much higher higher than the visually observable mass. Zwicky concluded that there should be matter that is not observable by its electromagnetic radiation, so-called *Dark Matter*. In chapter 2 we will briefly describe the current model of the universe and the role of dark matter (section 2.1). Some examples of observational evidence for the existence of Dark Matter are given in section 2.2.

Today, there are several theoretical candidates for this dark matter (DM) particle, but so far none of them have been detected. The candidates for these DM particles are divided in 3 classes : hot, warm and cold DM, depending on the speed of the particles. By observing e.g. the Cosmic Microwave Background radiation, most cosmologists favor the cold CM (CDM). Popular candidates for the CDM particles are Weakly Interacting Massive Particles (WIMPs), Axions and Massive Compact Halo Objects (MACHOs).

Recent observations show that the latter candidate (MACHOs) are not likely to account for the total amount of DM in the universe. Theories that go beyond the standard model of elementary particles, such as a supersymmetric extension of the standard model, naturally yeald DM candidates in the form of WIMPs (see section 2.3). Most Supersymmetric models predict a Lightest Supersymmetric Particle (LSP), which I will note as χ, that could account for the CDM. Since the χ hasn't been detected yet, we don't know what its mass m_χ is, but there are theoretical and experimental constraints on this mass, which narrow the possibilities. The mass of the particle is an important parameter for experimentalists.

There are different experiments that try to detect DM. These experiments can be divided in three categories : Direct detection, indirect detection and collider experiments (see section 3.2).

In this thesis, we will focus on indirect detection of solar WIMPs. If WIMPs are Majorana particles (i.e. they are their own anti-particle) then they can annihilate pair-wise when they collide. So WIMPs that are gravitationally captured in e.g. the Sun will collide and annihilate. This annihilation generates particles which can also be standard

model particles, which we already know and which we can observe, like heavy quarks, tau leptons, gauge bosons and/or Higgs bosons, which will decay into energetic neutrinos, which can escape from the bulk of the Sun. All of this is described in more detail in section 3.1.
If these neutrinos come in the direction of the Earth, they can be detected by neutrino telescopes, such as IceCube (see section 4). Depending on the mass m_χ, we expect to observe a flux of neutrinos with a certain energy, produced by $\chi\chi$ annihilations, coming from the direction of the Sun. One of the goals of the IceCube neutrino detector is to look for these neutrinos.

The author's contribution is the subject of chapters 6 and 7. In chapter 6 a filter is developed to remove as much background as possible. The experiment is background dominated, and this background mainly consists of atmospheric muons. The filtering will be done in two steps. The first step, described in section 6.4, consists of making cuts 'by eye' to reduce the atmospheric muon background by a factor ~ 100. In a second step we'll use*Boosted Decision Trees* (BDT), which will allow us to remove almost all the remaining atmospheric muons. This BDT filter is the subject of section 6.5.

Events that pass the filtering will then be used in chapter 7 to estimate the sensitivity of 59-string IceCube detector (IC59) for a solar WIMP signal, by looking for an excess of neutrinos coming from the direction of the Sun, under the assumption that they are produced in the decay of the annihilation products of χ particles. The upper-limit on a possible signal will be interpreted in the framework of the Minimal Supersymmetric extension of the Standar Model (MSSM) with R-parity conservation. In this model the LSP is the neutralino.

The calculation of the IC59 sensitivity for a solar WIMP signal is the aim of this thesis. We didn't unblind our data yet, so we will have to work *blind*. This *blindness* policy is introduced to prevent the introduction of a bias by the researcher. During an analysis it's quite possible that small errors are made and certain things are overlooked. If a researcher is allowed to look at the end result of his analysis and has a certain outcome in mind he can go back to the analysis to remove errors and recheck for possible oversights. But if he does this until his results check out with what he expects and stops removing errors then he will have introduced a bias because its very well possible that he did not remove all the errors and oversights in his analysis. Before unblinding the data one has to prove that the atmospheric background is well understood and under control.

We will calculate the IC59 sensitivity for a solar WIMP signal, obeying the *blind* procedure. The sensitivity is defined as the expected 90% confidence level upper limit on the number of signal events in a background only scenario. One expects that the number of signal events in a background only scenario is zero, but due to statistical fluctuations this number can be different from zero. In order to be sure that one measures real signal events and not just statistical fluctuations, the amount of measured signal events should exceed the sensitivity of the detector.

In section 7.2 we will calculate sensitivities of some interesting physical quantities, such as the $\chi\chi$ annihilation rate Γ_A, the neutrino induced muon flux Φ_μ and the neutralino-proton cross section $\sigma_{\chi p}$.

The next step will be an unblinding request, but before we do this we should calculate the systematic errors which is beyond the scope of this thesis. Once the data will be unblind, we can check if there are signal events in the data sample. If we don't see any signal, we will be able to set upper-limits on Γ_A, Φ_μ and $\sigma_{\chi p}$.

IceCube was built over several years, so the detector configuration changed every year. The upper-limits on Γ_A, Φ_μ and $\sigma_{\chi p}$ have already been calculated using the 22-string (2007-2008) and 40-string (2008-2009) data, and right now people are analyzing at the 78-string (2010-2011) data. But this thesis is the only analysis that is done with the 59-string (2009-2010) data, so this is the author's contribution to the collaboration.

In chapter 8 we will summarize our work and discuss the outlook.

Chapter 2

A Short Introduction To Dark Matter

2.1 Cosmology

Cosmology is the study of the large-scale structure of the Universe and its dynamics. Thanks to the progress on both theoretical and experimental side, cosmology has now become a serious branch in science.

The *standard cosmological model* [2] is based on the so-called *Big Bang* scenario. According to this model, the universe expanded from a hot and highly compressed state to the form it has today, in about 10^{10} year. This Big Bang model finds its roots in Hubble's law, which states that the observed recession velocity v is directly proportional to the distance d[3]

$$v = H_0 \cdot d, \tag{2.1}$$

where $H_0 = 73 \pm 3 kms^{-1}Mpc^{-1}$ is the present value of the Hubble parameter $H(t)$. From eq. (2.1) it follows that the longer the distance between two objects, the faster they move away from each other. This is nicely explained by an expanding universe, as is the case in the Big Bang model.

The model is built by using the *Einstein Field Equations*[5], from which we learn that the geometry of the universe is determined by its energy content. The energy content consists of 3 different components, each with its own abundance Ω_i [4, 5] (see Fig. 2.1) :

1. radiation, $\Omega_{R_0} \sim 10^{-4}$,
2. matter, $\Omega_{M_0} \sim 0.27$,
3. vacuum energy, $\Omega_{\Lambda_0} \sim 0.73$.

So about 30% of the energy density in our current universe consists of matter. It turns out[1, 5, 4] that the ordinary baryonic matter density is only a fraction of the

total matter density :

$$\Omega_{b_0} \sim 0.04. \tag{2.2}$$

This means that most of the matter density must be in the form of nonbaryonic dark matter. There is quite some evidence for the existence of dark matter, some observational evidences will be discussed in the next section (see 2.2).

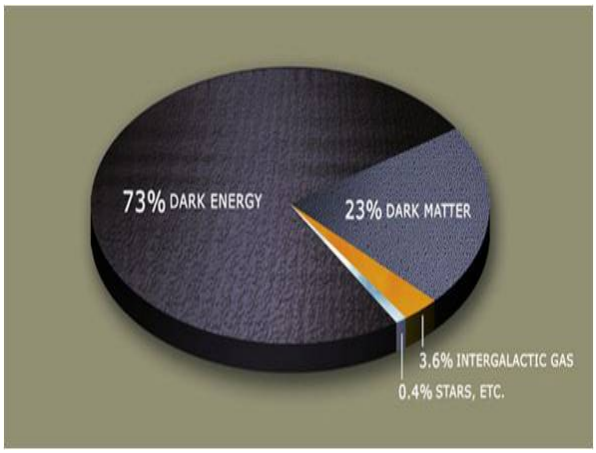

Figure 2.1: Overview of the content of the universe.

Up till today, we don't really know what the dark matter is made of. We only have some theoretical candidates that arise from yet unobserved theories like the SuperSymmetric extension of the standard model [6, 7] or model with extra space dimensions[8]. These candidates should have an abundance

$$\Omega_{DM_0} = \Omega_{M_0} - \Omega_{b_0} \sim 0.23. \tag{2.3}$$

Note that this number (Ω_{DM_0}) can be the sum of many different kind of particles. So far we already know that the standard model neutrinos have a small contribution, but the experimental upper limit on their relic density is only [1, 5]

$$\Omega_\nu h^2 \lesssim 0.07, \tag{2.4}$$

where $h = H_0/100 kms^{-1}Mpc^{-1}$ is a dimensionless parameter. From eq. (2.4), we see that neutrinos are not abundant enough to be the dominant component of the dark matter.

The relic density (like e.g. $\Omega_\nu h^2$) for a generic particle species X can be approximated by [1]

$$\Omega_X h^2 \approx \frac{3 \times 10^{-27} cm^3 s^{-1}}{\langle \sigma v \rangle}. \tag{2.5}$$

If we know the $\langle \sigma v \rangle$ for a generic particle species X, we can calculate its abundance and *vice versa*. We will use this relation later (section 2.3) to determine how the dark matter particles should interact in order to get the correct DM density (see eq. 2.11).

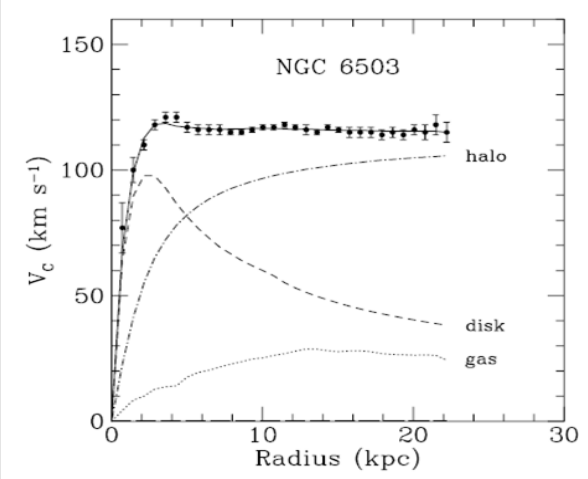

Figure 2.2: Rotation curve of NGC 6503. The dotted, dashed and dash-dotted lines are the contributions of gas, disk and dark matter, respectively. Picture taken from [1].

2.2 Observational Evidence

The astrophysicist Zwicky was the first to notice the presence of so-called dark matter. This was in 1933, when he estimated the mass of the Coma Cluster of galaxies by looking at it in two different ways[9] :

1. he looked at the motions of galaxies near the edge of the cluster and
2. he estimated the brightness in the cluster by counting the number of galaxies.

He found that the brightness was much smaller than he expected for the amount of mass that was present, so there had to be a large amount of non-luminous or dark matter.

Today, almost 80 years later, we have much more evidence for the presence of this dark matter. The most convincing and direct evidence comes from the observations of orbital velocities of stars versus their distance from the center of spiral galaxies. From observations we learn that most of the luminous matter in spiral galaxies is concentrated in the center, so we expect that the radial velocity v_{rad} goes as

$$v_{rad} = \sqrt{\frac{G_N M(r)}{r}}, \tag{2.6}$$

where G_N is Newtons constant, $M(r)$ is the mass density profile and r is the distance to the center of the galaxy. Fig. 2.2 shows the expected and the measured rotation curves. A possible explanation for the discrepancy between the observed rotation curve and the expected rotation curve from the centered luminous matter (dashed line) is provided by including a halo of dark matter (dashed-dotted line).

There are many other observational evidences for the presence of dark matter, like gravitational lensing[10], the velocity dispersions of dwarf spheroidal galaxies[11] and

much more (see [1] for an overview). If we now want to make predictions of the matter and energy content in the universe, we can study the cosmic microwave background radiation (CMBR). The WMAP experiment has measured the CMB very precisely and has found the following values [12]

$$\Omega_b h^2 = 0.02267^{+0.00058}_{-0.00059} \qquad \Omega_m h^2 = 0.1358^{+0.0037}_{-0.0036}, \tag{2.7}$$

which means that about 4.6% of the universe is made up of ordinary baryonic matter, while about 23% is made up of dark matter. Dark energy, a component we currently don't know anything about, would account for the remaining 73% of the energy.

2.3 Candidates

As was already mentioned in section 2.1, there are several theoretical candidates for the dark matter particles. We also mentioned in that section that the standard model neutrinos have a small contribution to the dark matter. In this discussion, we will focus on todays most promising candidate : *the neutralino*. There are many more candidates, for which an overview is given in [1].

The neutralino is a particle that arises from supersymmetry, which is a theory that goes beyond the standard model. According to supersymmetry, each standard model (SM) particle has a supersymmetric partner with spin different by one half. The new generators that are introduced change fermions into bosons and vise versa

$$Q|\text{fermion}\rangle = |\text{boson}\rangle; \quad Q|\text{boson}\rangle = |\text{fermion}\rangle. \tag{2.8}$$

Supersymmetry provides a nice solution to the *hierarchy problem* and unifies the gauge coupling scales at $M_U \sim 2 \cdot 10^{16}$ GeV, which has been taken as a strong hint in favor of Grand Unified Theories [6, 7].

The *minimal* extension of the standard model of particle physics is described by the Minimal Supersymmetric Standard Model (MSSM). The MSSM contains the smallest possible field content necessary to give rise to all the fields of the standard model (an overview is given in Table 2.1).

The fact that the MSSM is such an interesting theory in the search for dark matter, is because one of its basic elements is R-parity conservation. R-parity is a multiplicative quantum number, defined as

$$R \equiv (-1)^{3B+L+2s}, \tag{2.9}$$

where B is the baryon number, L the lepton number and s the spin of the particle. From eq. (2.9) it follows that

$$\begin{cases} R &= +1 \text{ for SM particles,} \\ R &= -1 \text{ for SUSY particles.} \end{cases} \tag{2.10}$$

Conservation of R-parity implies that SUSY particles can only decay into an odd number of SUSY particles (plus SM particles), from which we can conclude that there should be a Lightest Supersymmetric Particle (LSP) if the theory is correct. This LSP is an excellent dark matter candidate, since it's stable and can only be destroyed via pair annihilation. This LSP is electrically neutral and colorless[7]. The MSSM contains 4 of these electrically neutral and colorless particles, namely the *neutralinos* $\tilde{\chi}_1^0$, $\tilde{\chi}_2^0$, $\tilde{\chi}_3^0$ & $\tilde{\chi}_4^0$ (see table 2.1). In this thesis, we will focus on the lightest neutralino, namely $\tilde{\chi}_1^0$ or simply noted as χ. If we fill in the results of WMAP (eq. 2.7) in eq. (2.5) and

Standard Model particles and fields		Supersymmetric partners			
		Interaction eigenstates		Mass eigenstates	
Symbol	Name	Symbol	Name	Symbol	Name
$q = d, c, b, u, s, t$	quark	$\tilde{q}_L$, $\tilde{q}_R$	squark	$\tilde{q}_1$, $\tilde{q}_2$	squark
$l = e, \mu, \tau$	lepton	$\tilde{l}_L$, $\tilde{l}_R$	slepton	$\tilde{l}_1$, $\tilde{l}_2$	slepton
$\nu = \nu_e, \nu_\mu, \nu_\tau$	neutrino	$\tilde{\nu}$	sneutrino	$\tilde{\nu}$	sneutrino
g	gluon	$\tilde{g}$	gluino	$\tilde{g}$	gluino
$W^\pm$	W-boson	$\tilde{W}^\pm$	wino	$\tilde{\chi}^\pm_{1,2}$	chargino
H^-	Higgs boson	$\tilde{H}^-_1$	higgsino		
H^+	Higgs boson	$\tilde{H}^+_2$	higgsino		
B	B-field	$\tilde{B}$	bino	$\tilde{\chi}^0_{1,2,3,4}$	neutralino
W^3	W^3-field	$\tilde{W}^3$	wino		
H^0_1	Higgs boson	$\tilde{H}^0_1$	higgsino		
H^0_2	Higgs boson	$\tilde{H}^0_2$	higgsino		
H^0_3	Higgs boson				

Table 2.1: SM particles and their superpartners in the MSSM (adapted from [1]).

we assume that χ is the dominant component of dark matter, we see that in order to match observation, $\Omega_\chi h^2$ should be

$$\Omega_\chi h^2 \approx \frac{3 \cdot 10^{-27} c m^m s^{-1}}{\langle \sigma v \rangle} \sim \mathcal{O}(0.1), \tag{2.11}$$

The above equation (2.11) is satisfied if $\langle \sigma v \rangle$ is of the order of a picobarn to a femtobarn, which is a typical size for an weak process. So our dark matter particles should be massive and weakly interacting, that's why they are often referred to as Weakly Interacting Massive Particles (WIMPs). Neutralinos (which are a combination of gauginos and Higgsinos) satisfy these conditions and are generically called WIMPs, for the remainder of this thesis.

Chapter 3

Detecting WIMPs

The neutralino described in the previous chapter is an excellent dark matter candidate. But where and how should we look for these particles? First (see section 3.1) we'll explain where the WIMPs accumulate. In section 3.2 we'll give an overview of the different methods to detect them.

3.1 O WIMP, Where Are Thou

In section 2.2 we mentioned that the observed rotation curves can be explained by a halo of dark matter around the spiral galaxies (see Fig. 2.2). Assuming that this model is correct, we expect that there is dark matter in our immediate environment. Some experiments try to observe these particles by looking for their interactions in a detector. These *direct detection experiments* will be briefly explained in section 3.2.2.

Another way to look at these WIMPs is by looking at their annihilation products. This is what *indirect detection experiments* do, as will be described in section 3.2.3. The idea behind this is that if WIMPs are Majorana particles, then two colliding WIMPs can annihilate into SM particles, which we can detect. For WIMPs to annihilate, they should be very densely populated, so to detect their decay products, one should look at some particular places in the universe where there's a high concentration of WIMPs. It is believed that these high concentrations would be present in the center of heavy objects such as the Sun. This is because WIMPs will sometimes scatter elastically (by the weak interaction) with nuclei in the Sun, as the Solar System moves in the galactic halo. After this scattering, the WIMP has lost momentum and can become gravitationally bound. The solar capture rate of WIMPs $C^{\odot}$ can be approximated by[13]

$$\begin{aligned} C^{\odot} &\approx 1.3 \times 10^{21} s^{-1} \left(\frac{\rho_{\text{local}}}{0.3 GeV/cm^3} \right) \left(\frac{270 km/s}{v_{\text{local}}} \right) \\ &\quad \times \left(\frac{100 GeV}{m_{\chi}} \right) \sum_i \left(\frac{A_i(\sigma_{\chi i,SD}) + \sigma_{\chi i,SI} S(m\chi/m_i)}{10^{-6} pb} \right), \end{aligned} \tag{3.1}$$

where ρ_{local} is the local dark matter density, v_{local} is the local rms velocity of halo dark matter particles, $\sigma_{\chi i,SD}$ and $\sigma_{\chi i,SI}$ are respectively the spin-dependent and spin-independant elastic scattering cross sections of the WIMP (with mass m_χ) with nuclei species i (with mass m_i), and A_i is a factor denoting the relative abundance and form factor for each species. The quantity S is is given by

$$S(x) = \left[\frac{A(x)^{3/2}}{1 + A(x)^{3/2}}\right]^{3/2}, \tag{3.2}$$

where

$$A(x) = \frac{3}{2}\frac{x}{(x-1)^2}\left(\frac{v_{\text{esc}}}{v_{\text{local}}}\right)^2, \tag{3.3}$$

with $v_{\text{esc}} \approx 1156 km/s$ the escape velocity of the Sun. During its lifetime, the Sun may already have captured a huge amount of WIMPs. If the density of the WIMPs is high enough, they will start to annihilate. If we neglect evaporation, the present WIMP annihilation rate is given by

$$\Gamma = \frac{1}{2} C^\odot \tanh^2\left(\sqrt{C^\odot A^\odot}\; t_\odot\right), \tag{3.4}$$

where $t_\odot \approx 4.5 \times 10^9$ years is the age of the Solar System and $A^\odot$ is the annihilation cross section times the relative WIMP velocity per volume which can be approximated by [14, 15]

$$A^\odot \approx \frac{\langle \sigma v \rangle}{5.7 \times 10^{27} \text{cm}^3 \left(100 GeV/m_\chi\right)^{3/2}}. \tag{3.5}$$

If now

$$\sqrt{C^\odot A^\odot}\; t_\odot \gg 1, \tag{3.6}$$

equilibrium between capture and annihilation is reached, so the annihilation rate and the corresponding event rate are maximized and are entirely determined by the capture rate.

Colliding WIMPs could annihilate into a variety of SM particles, which could then be detected. In this thesis, we are looking at neutrinos that come from WIMP annihilation. In some models, WIMPs can annihilate directly into neutrino-antineutrino pairs, but the predicted rates for this process is extremely small[1], so we will focus on neutrinos that come from the decay of the annihilation products, such as heavy quarks, tau leptons, gauge bosons and/or Higgs bosons. These neutrinos are very interesting particles, since they are the only kind of particles that will be able to escape from the core of the Sun and reach the Earth, where they can (hopefully) be detected by neutrino detectors such as IceCube. Fig. 3.1 gives an overview of the described process.

The total neutrino energy spectrum is given by the sum of all the contributions from each channel X, weighted by its branching ration B_X

$$\left(\frac{dN_\nu}{dE_\nu}\right) = \sum_X B_X \left(\frac{dN_\nu}{dE_\nu}\right)_X. \tag{3.7}$$

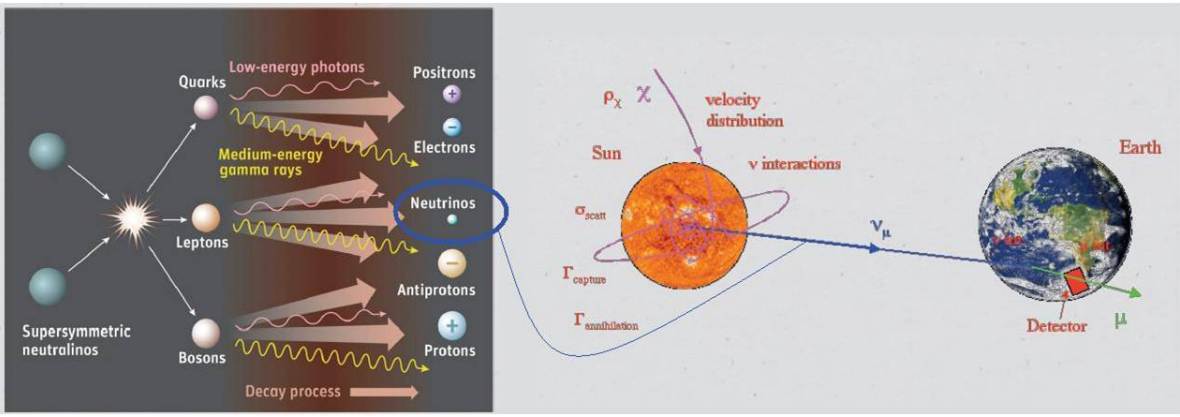

Figure 3.1: Left : two annihilating neutralinos produce SM particles, that can eventually decay into e.g. neutrinos. Right : Neutralinos are captured by the Sun as our Solar System moves in the galactic halo.Once the WIMP density in the core of the Sun is high enough, the WIMPs will self-annihilate and produce SM particles. Neutrinos can be produced as decay products of these SM particles. The produced neutrinos can escape from the core of the Sun, since they only interact very rarely with ordinary matter and if they do, its mainly via the weak interaction. The neutrinos that come in the direction of the Earth, could be detected by neutrino detectors such as IceCube.

This energy spectrum is unknown, since it depends on unknown SUSY parameters, such as mass and composition of the neutralino. For what concerns the mass, one could just scan a mass region. A lower limit on the mass of the WIMPs was determined by the LEP experiments at CERN. By combining the searches for sleptons, charginos and Higgs bosons in the mSUGRA scenario, they found[16]

$$m_\chi \geq 47GeV. \tag{3.8}$$

Since the annihilation cross section of particles goes as (see eq. (3.5))

$$A^\odot \sim m_\chi^{3/2}, \tag{3.9}$$

we know that the mass m_χ can't be too high, otherwise the WIMPs would have annihilated too much and the relic WIMP density would be too low to explain the dark matter.

Since we don't know what the annihilation channel of the WIMP is, we look at the extremes of the energy distribution, with at the lower edge

$$\chi\chi \to b\bar{b}, \tag{3.10}$$

which we call the *soft channel* and at the higher edge we have the *hard channel*

$$\chi\chi \to W^+W^-. \tag{3.11}$$

Since these are the extreme cases, any other choice of final state would lead to an intermediate energy spectrum. That's why we assume 100% branching into the soft and hard channels. The reason why we call eq. (3.11) the hard channel is because the neutrinos are produced immediately by the annihilation products ($W \to \nu_\ell \ell$, 30% of the time). In the soft channel 3.10, most of the neutrinos are produced indirectly in decays of the hadrons that were created in the quark jets.

In this work we will focus on WIMPs with a mass of 1 TeV that decay via the hard channel. All the optimizations and calculations will be done for this kind of WIMP, but we will from time to time check what the results of our analysis are for other masses ($50GeV \to 5TeV$) and channels (soft and hard). Fig. 3.2 shows the neutrino energy spectrum E_ν for a 1 TeV WIMP for both the hard channel (blue solid line) and the soft channel (black dashed line).

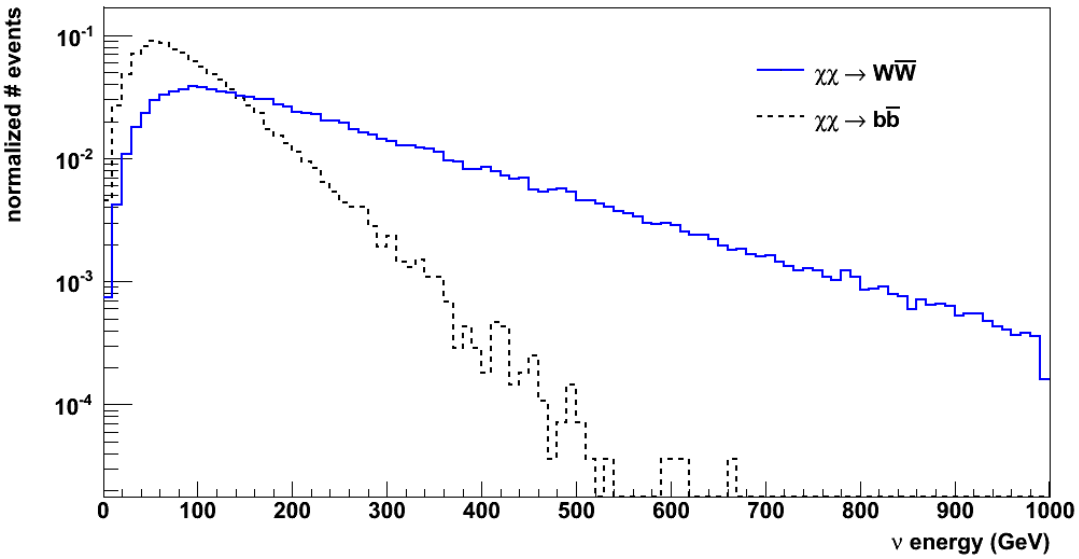

Figure 3.2: The neutrino energy spectrum E_ν for a 1 TeV WIMP for both the hard channel (blue solid line) and the soft channel (black dashed line). Note that this is a logarithmic scale.

3.2 Experiments

In the previous section 3.1 we already mentioned two type of experiments that are looking for dark matter : the *direct* and the *indirect detection experiments*. There is also a third type of experiment, namely the *collider experiments*. In this section we briefly describe these three detection methods. A more elaborate description is given in [1, 4].

3.2.1 Collider Experiments

According to supersymmetry, every SM particle has a superpartner. These superpartners could maybe be created in collider experiments such as those at the Large Hadron Collider (LHC) at CERN[17]. Neutral SUSY particles (such as the neutralino) will not be detected by the LHC experiments, but if they are produced, they have a certain energy and momentum, so one would measure this missing energy and momentum. This missing energy would not be a proof for the existence of dark matter though, since it would only tell us that the produced particle was stable enough to exit the detector, which only tells us that the particle has a lifetime $\tau \gtrsim 10^{-7}s$, a value that is still far from the required lifetime of dark matter $\tau \gtrsim 10^{17}s$[]. On the other hand, the LHC should be able to detect charged SUSY particles, if their masses are not to high. The detection of such a charged SUSY particle would be a good indication for the neutralino as a dark matter candidate, since it would confirm supersymmetry, the theory from which the neutralino arises.

3.2.2 Direct Detection

As was already mentioned in section 3.1, we expect that there is an amount of dark matter in our immediate environment. Direct detection experiments try to measure the recoil energy of a WIMP particle scattering on an atom in a lattice. By placing these detectors deep under ground, the background coming from cosmic rays is reduced, but there still remains a background from radioactive materials and photons.

One of the direct detection experiments is DAMA, which is located at the INFN laboratory under the Gran Sasso mountain in Italy. DAMA claims to have detected dark matter. They measured an annual modulation of the event rate, which would be caused by the revolution of the Earth around the Sun [19]. So far, no other experiment has detected nuclear recoils due to dark matter interactions, so the results of DAMA are controversial and many people think that the annual modulation is not coming from WIMPs, but from other factors that the DAMA collaboration didn't take into account.

The CoGeNT collaboration has recently reported a rising low energy spectrum in their ultra low noise germanium detector. This is particularly interesting as the energy range probed by CoGeNT overlaps with the energy range in which DAMA has observed their annual modulation signal [20].

The current limits on spin-independent scattering scattering, from some direct detection experiments are summarized in Fig. 3.3.

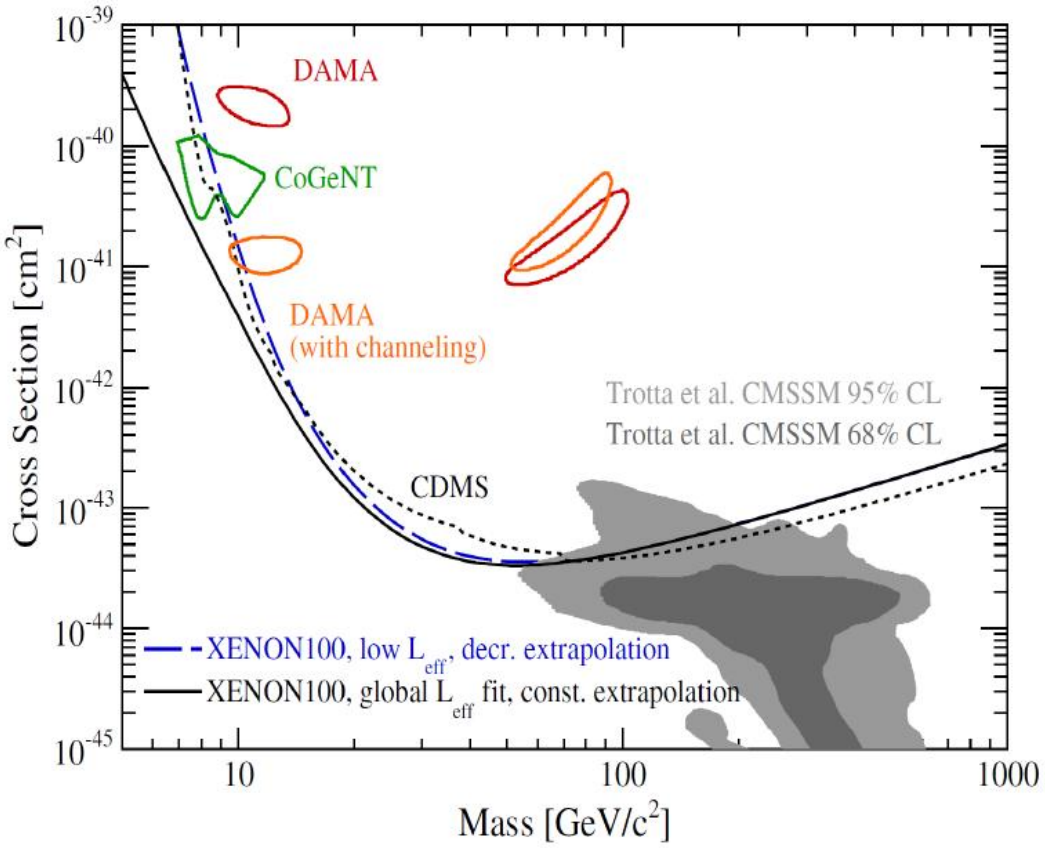

Figure 3.3: The 90% confidence upper limits on the spin-independent elastic WIMP-nucleon cross section (solid and long dashed lines) from XENON100, together with the best limit to date from CDMS (dotted line), recalculated assuming an escape velocity of 544 km/s and $v_0 = 220$km/s. Expectations from a constrained MSSM model [**?**]SSMconstr, and area (90% C.L.) favoured by CoGeNT (green) and DAMA (red/orange) are also shown. Picture taken from [26]

3.2.3 Indirect Detection

Indirect detection experiments look for the radiation that is produced in dark matter self-annihilation. In section 3.1 we described how WIMPs are accumulated in the center of heavy objects like our Sun, and how their annihilation products (e.g. neutrinos) make it to the Earth. By measuring these decay products, one could find an indirect indication for WIMPs as dark matter. In this thesis, we'll look for neutrinos that have been produced in the decay of the annihilation products of WIMPs captured

in the Sun, hereafter referred to as solar WIMPs. These neutrinos can be observed by detectors such as IceCube (this is the experiment that we will use), ANTARES[21] and SuperKamiokande [23]. In chapter 8 we will look at the current results of these experiments.

Other experiments, such as PAMELA[24] look for positrons and electrons coming from annihilation of WIMPs in the galactic halo. There are many more experiments that look for all kinds of annihilation products of WIMPs, but it is not our intention to list them all. A nice overview is given in [1].

Chapter 4

Detecting Neutrinos With The IceCube Detector

In the first part of this chapter we'll explain how neutrinos interact and why it's interesting to build a detector in ice (or water). The more technical part of IceCube is explained in section 4.2. What we explain here is only a very brief summary of what IceCube is and does. For a more complete description we refer the reader to [22].

4.1 Neutrino Interactions

Neutrinos are electrically neutral particles with an extremely small, but nonzero mass. Since they only rarely interact with ordinary matter, they are able to pass almost unaffected through it. If they interact with ordinary matter it's mainly via the weak force.

There are two ways in which a neutrino can weakly interact with a nucleon

1. via the exchange of a W boson :

$$\nu_\ell(\bar{\nu}_\ell) \ + \ N \ \rightarrow \ \ell^-(\ell^+) \ + \ X \tag{4.1}$$

   this process is called charged-current (CC) interaction. In this interaction the neutrino interacts with a nucleon and produces a lepton ℓ and a hadronic shower X. The signature of these kind of reactions depends on the flavor of the final lepton. Fig. 4.1 (A), (B) and (C) show the signatures of respectively the electron, the muon and the tau lepton.

2. via the exchange of a Z boson :

$$\nu_\ell(\bar{\nu}_\ell) \ + \ N \ \rightarrow \ \nu_\ell(\bar{\nu}_\ell) \ + \ X, \tag{4.2}$$

   this is the so-called neutral-current (NC) interaction. In this kind of interaction the neutrino scatters off the nucleon N and produces a hadronic shower X (see Fig. 4.1 (D)).

The cross section $\sigma_{\nu N}$ of neutrino nucleon interactions increases with energy and the CC interaction has a higher cross section then the NC interaction[30]. We also know that, because of helicity constraints, these cross sections are higher for neutrinos than for anti-neutrinos.[30]
From Fig. 4.1, we see that the most interesting signature for our analysis comes from CC interactions with a muon neutrino ν_μ as primary particle, since in this case the produced muon has almost the same direction as the primary neutrino. This is important if one wants to determine the position of the neutrino source, like the Sun in our case. Also, muons have a long ($> 10m$) range, while electrons and tau leptons are quickly absorbed.

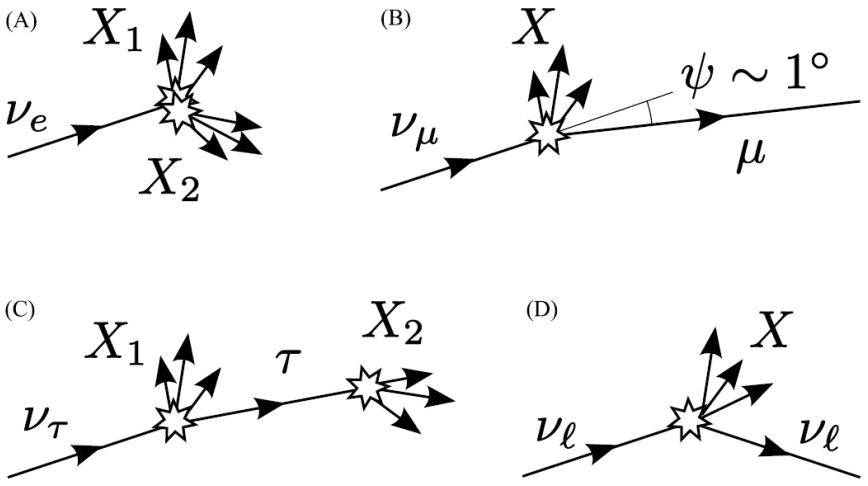

Figure 4.1: Different scenarios for neutrino-nucleon interactions. (A), (B) and (C) show the CC interactions for the three flavors, respectively ν_e, ν_{mu} and ν_τ. (D) shows the general NC interaction (similar for three flavors). Picture taken from [31].

The mean opening angle between ν and μ_{true} can be approximated by [32]

$$\langle \theta_{\nu\mu} \rangle \approx \frac{0.7°}{(E_\nu/TeV)^{0.7}}, \tag{4.3}$$

which means that the average opening angle is less than 1° above 0.6 TeV. In our analysis we will go down to $\sim 10GeV$ neutrinos, for which the opening angle isn't that small as is shown in the left plots of Figs. 4.2 to 4.5.

Figs. 4.2 to 4.5 show the distribution of the opening angle between ν and μ vs. the energy of the neutrino. The left plots show how the opening angle between the neutrino and the true muon decreases with increasing neutrino energy. In Figs. 4.2 and 4.3 we used simulated atmospheric neutrinos that passed the low level filters (level 1 and 2 : see section 6.3), and Figs. 4.4 and 4.5 only the atmospheric neutrinos that also passed our high level filter (level 3 + BDT : see sections 6.4 and 6.5). In Figs. 4.3 and 4.5, we

zoomed in on the $0 - 1$TeV range of respectively Figs. 4.2 and 4.4.
The thick black line represents the mean opening angle. The right plots in Figs. 4.2 to 4.5 show the same distribution, but this time between the LLH reconstructed muon μ_{reco} (see section 6.1 for an explanation of the reconstruction) and the primary neutrino ν. We see that in this case, the mean opening angle is much larger. This is due to many badly reconstructed tracks in the distribution, so this opening angle gets smaller after removing the badly reconstructed tracks. Fig. 4.4 shows the same distributions as Fig. 4.2, but after the filter we've developed. This filter is the subject of sections 6.4.2 to 6.5. Note that the opening angle becomes much smaller, as expected, since after the filtering, most of the poorly reconstructed tracks are removed. This small opening angle will allow us to look for neutrinos coming from the Sun direction with a pretty good resolution.

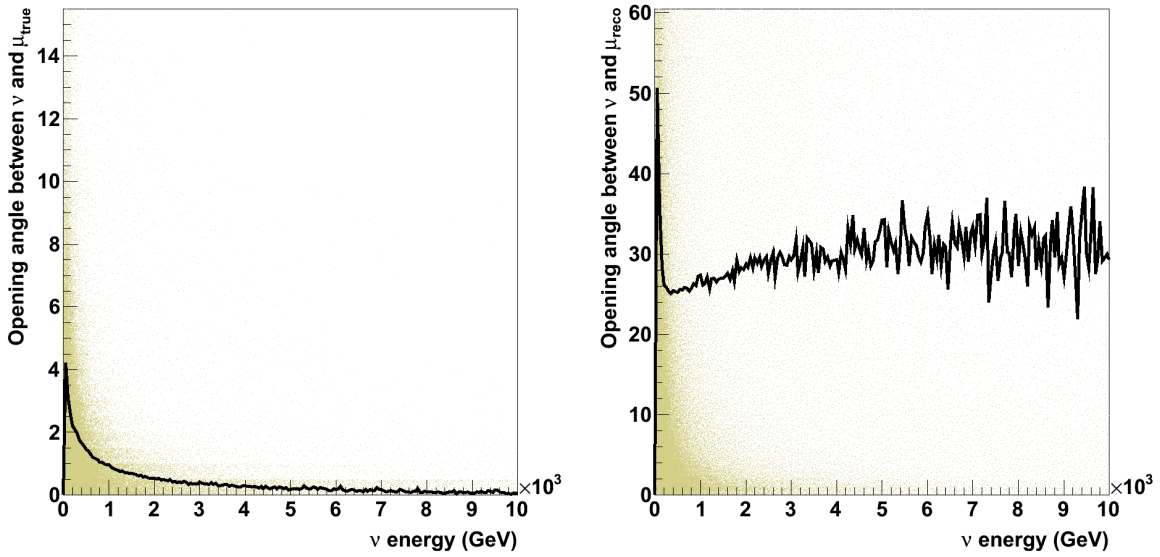

Figure 4.2: Distribution of the opening angle between the primary ν_μ and the produced μ_{true} (left) (reconstructed μ_{reco} (right)) vs. the energy of the primary neutrino E_ν. The thick black line shows the mean opening angle vs. E_ν. The fluctuations of the mean opening angle are due to small statistics.

So now we know that it is interesting to look for muons coming from the CC interaction

$$\nu_\mu(\bar{\nu_\mu}) \; + \; N \; \rightarrow \; \mu^-(\mu^+) \; + \; X. \tag{4.4}$$

The next question we should ask ourselves is how to detect these muons? The detection is based on the*Cherenkov radiation*, which is electromagnetic radiation emitted when a charged particle (such as the muon) passes through a dielectric medium (like ice in our case) at a speed that exceeds the phase velocity of light in that medium

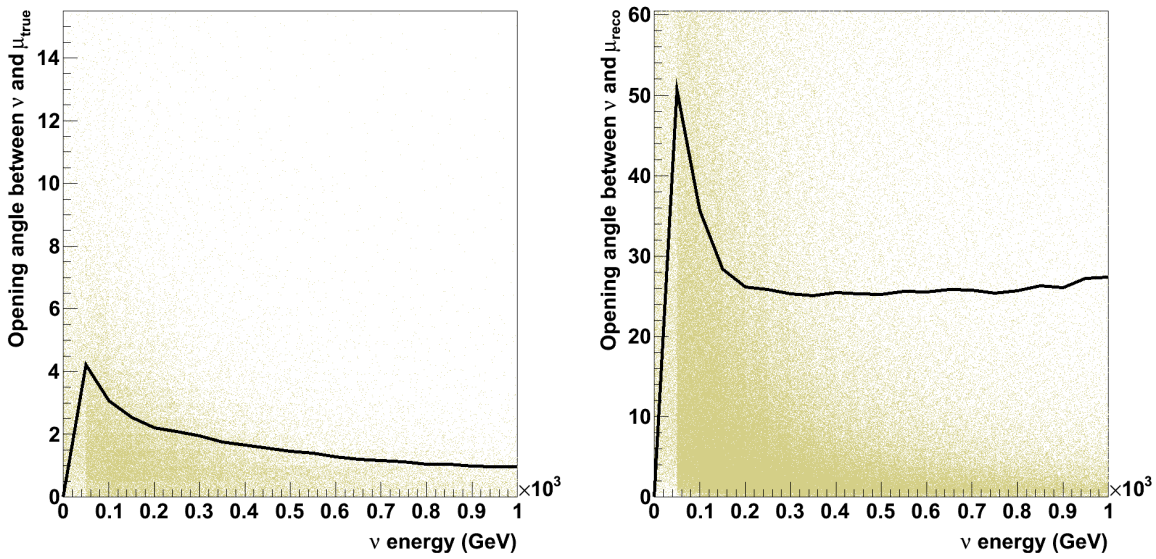

Figure 4.3: A zoom on the energy region $0 < E_\nu < 1000GeV$ of Fig. 4.2

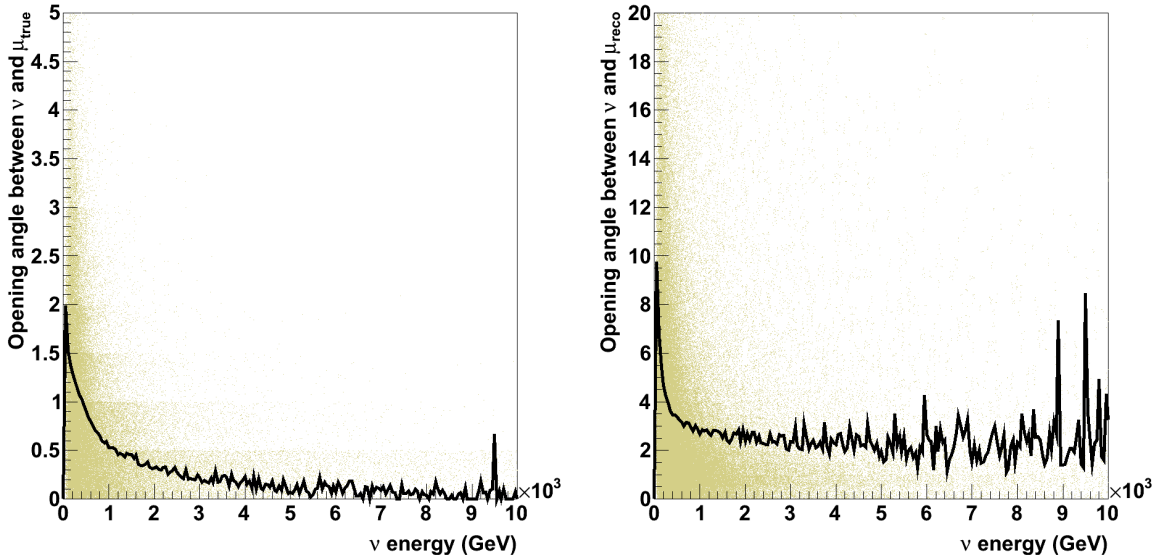

Figure 4.4: Distribution of the opening angle between the primary ν_μ and the produced μ_{true} (left) (reconstructed μ_{reco} (right)) vs. the energy of the primary neutrino E_ν after the filtering described in sections 6.4.2 to 6.5. The thick black line shows the mean opening angle vs. E_ν.

(c_{vac}/n_{ice}). What happens is that the local electromagnetic (EM) field in the medium is disrupted when the charged particle (e.g. muon) travels through it. The atoms of

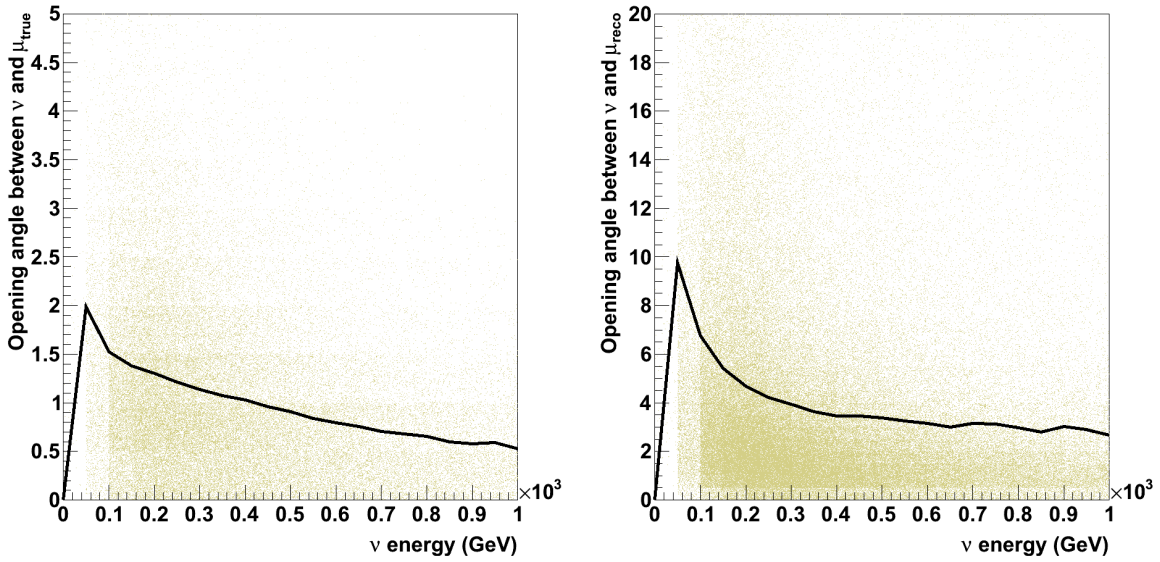

Figure 4.5: A zoom on the energy region $0 < E_\nu < 1000GeV$ of Fig. 4.4

the medium become polarized because of this passing EM field. Once the EM field has passed by, the electrons of the polarized atoms restore themselves to equilibrium and emit photons while doing this. These photons will interfere constructively and form a shock wave at a certain angle θ_c if the speed of the disruption exceeds the light speed in the medium. Whenever this happens, one observes the so-called Cherenkov radiation, shown in Fig. 4.6. For this to happen for a muon (mass m_μ) that passes through the ice (refractive index n_{ice}), a minimum muon energy E_μ is needed

$$E_\mu^{\min}(\lambda) = \frac{m_\mu}{\sqrt{1 - n_{ice}(\lambda)^{-2}}}, \tag{4.5}$$

whereλ is the wavelength of the emitted Cherenkov photons.
The number of Cherenkov photons emitted per unit path length is given by[33]

$$\frac{dN}{dx} = 2\pi\alpha \int \left(1 - \frac{1}{n^2\beta^2}\right) \frac{d\lambda}{\lambda^2} \tag{4.6}$$

where α is the fine structure constant and λ the wavelength of the emitted photon. The maximum number of Cherenkov photons per unit path are emitted in the visible spectrum $300nm < \lambda < 600nm$, so the glass sphere around the DOM is designed such that it allows these photons to pass. In this range of wavelengths $n_{ice} \approx 1.33$, so using eq. 4.5 we find that $E_\mu^{min} = 160$ MeV is the minimum energy for Cherenkov emitting muons in IceCube.

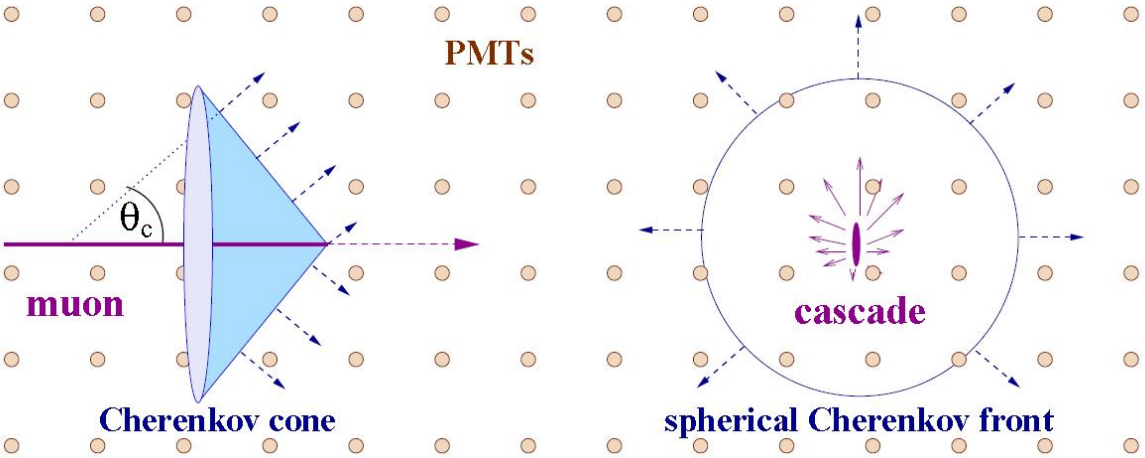

Figure 4.6: Left : Cherenkov cone produced by a muon as it passes through the ice. Right : Cherenkov light produced by an electron or tau particle during creation or decay.Picture taken from [45].

The Cherenkov angle θ_c is given by

$$\cos\theta_c = \frac{1}{n(\lambda)\beta}, \tag{4.7a}$$

$$= \frac{c_{vac}}{n_{ice} v_\mu}, \tag{4.7b}$$

$$\Rightarrow \theta_c \approx 41^\circ, \tag{4.7c}$$

where $\beta = v/c_{vac}$. Eq. (4.7b) gives the specific case for a muon traveling through ice and in eq. (4.7c), we used the fact that the muon is a relativistic particle and we approximated its speed by the light speed in vacuum c_{vac}. The Cherenkov photons will hit the Digital Optical Modules (DOMs see section 4.2) in the detector and be recorded. The direction of the muon can be reconstructed by using sophisticated algorithms that are based on the timing and the positions of the hit DOMs (see section 6.1).

4.2 IceCube

IceCube is a neutrino detector located at the geographical South pole. It consists of 78 kilometer-length strings, each instrumented with 60 Digital Optical Modules (DOMs) spaced by 17 m. The strings are lowered into the Antarctic ice at depths between 1450 meters to 2450 meters, this way a huge amount of background from cosmic ray showers is filtered away by the ~ 1.5 km ice layer above the detector. The strings are arranged at the apexes of equilateral triangles 125m on a side. This part of the detector is able to detect neutrinos with energies $E_\nu \geq 100$ GeV. In the center of the IceCube array there is the DeepCore detector which consists of an infill of 8 strings with 60 DOMs with high quantum efficiency. DeepCore allows us to measure neutrinos with energies down to $E_\nu \sim 10$ GeV, which will be useful to look for WIMPs with small masses.
The total detector volume is a cubic kilometer of sterile Antarctic ice. Fig. 4.7 gives an overview of the detector. More details can be found in [22].

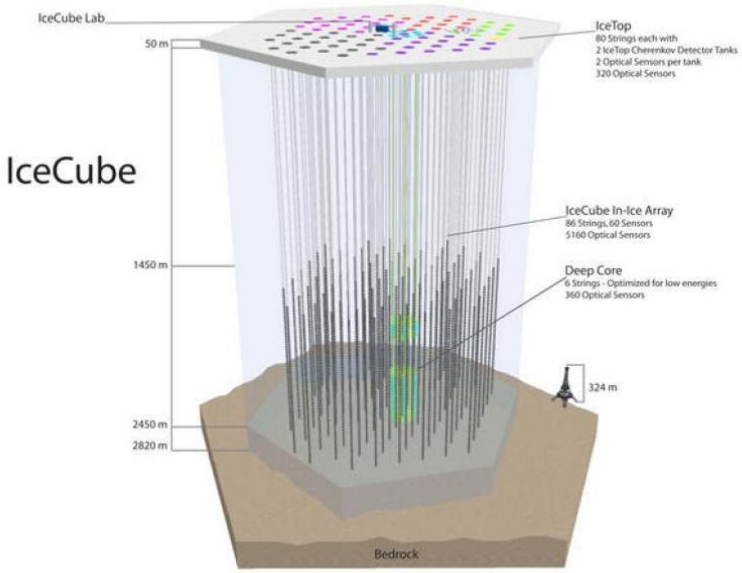

Figure 4.7: Overview of the total IceCube detector.

The detector was built over a time span of several years. Each year (during the South Pole Summer), more strings were added to the detector. A brief overview of the construction steps [29] :

- 2005 : 1 string deployed,
- 2006 : 9 strings deployed,
- 2007 : 22 strings deployed,

- 2008 : 40 strings deployed,
- 2009 : 59 strings deployed (of which 1 DeepCore string),
- 2010 : 79 strings deployed (of which 6 DeepCore strings),
- 2010 : 86 strings deployed (of which 8 DeepCore strings) $\Rightarrow$ IceCube is now complete.

In this thesis, we will analyze data taken by the 59 string IceCube detector (IC59). This detector configuration took data from the 30th of May 2009 till the 31th of May 2010. Fig. 4.8 shows the detector as seen from the top. The dots show the positions of the strings and the colors show in which year they were deployed. The green region represents the IC59 part of the detector.

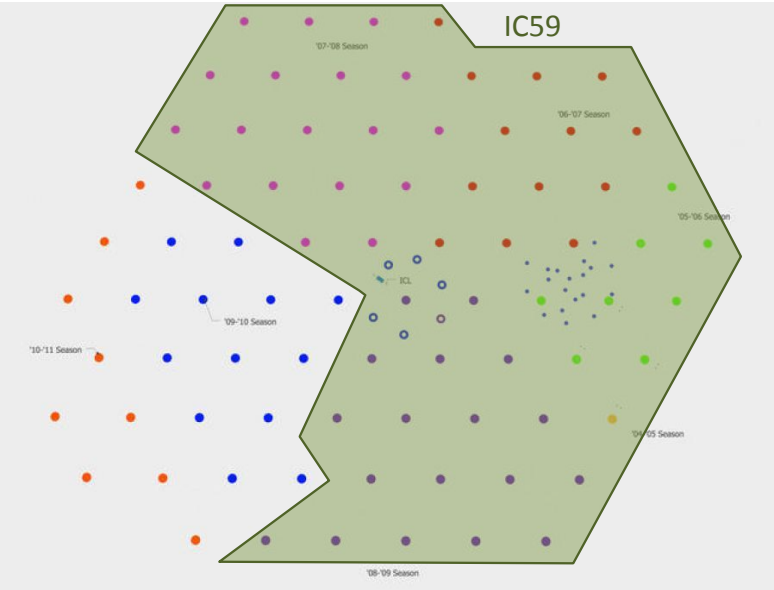

Figure 4.8: IceCube as seen from above. The green region shows the IC59 detector.

We already explained that IceCube detects muons by measuring the Cherenkov photons they emit when they travel through the ice. The Cherenkov photons are detected by the DOMs in the detector. DOMs are optical sensors that consist of a 13 mm thick glass sphere that contains a 25 cm photomultiplier tube (PMT) with 20% quantum efficiency to measure the Cherenkov photons, a 2kV high voltage power supply for the PMT, a DOM Main Board that digitizes the signals locally using an on-board computer, and many other components (see Fig. 4.9). Once the signals are digitized, they get a global time stamp with a resolution $\lesssim 3$ ns and are subsequently sent to the surface, where they are collected by processors.

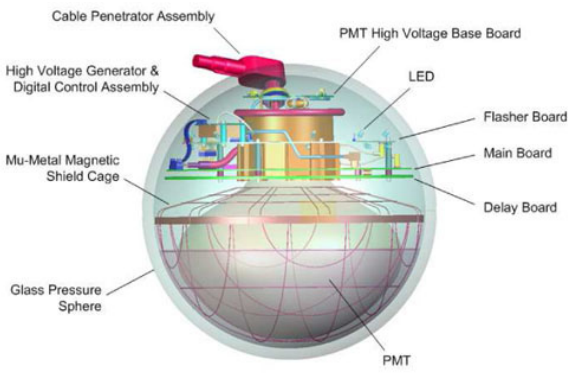

Figure 4.9: Schematic drawing of a Digital Optical Module (DOM). Picture taken from [22].

4.3 Background Events

In our analysis, we will look for muon neutrinos ν_μ coming from the direction of the Sun, where they originate from by the annihilation products of solar WIMPs. We will be able to detect these neutrinos if they interact within IceCube via the CC interaction in which a muon μ is produced, that can be detected. Unfortunately for us, IceCube also detects other muons, that either come directly into the detector after being produced in the atmosphere (the so-called atmospheric muons μ_{atm}), or from a CC interaction of an atmospheric neutrino (ν_{atm}).
Atmospheric muons and neutrinos are produced in *cosmic ray showers*[4]. The *primary cosmic rays* are charged particles that are accelerated by astrophysical sources. When these particles reach the Earth and collide with nuclei in the atmosphere, jets of hadrons like pions and kaons are produced, which in turn decay into secondary particles like muons and neutrinos. A scheme of a cosmic ray shower is shown in Fig. 4.10.

Figure 4.10: Overview of the evolution of a cosmic ray. An extraterrestrial particle (e.g. proton) comes into the atmosphere and collides with a nucleus in the atmosphere (e.g. Nitrogen). After this collision jets of hadrons, like pions and kaons, are produced, which in turn decay into secondary particles like muons and neutrinos.

The down-going atmospheric muons are the main background for the IceCube de-

tector ($\sim 2000Hz$). However, the flux of the atmospheric muons is strongly reduced once they reach the Earth's surface and start to travel through it, which means that the Earth can be used as a filter for atmospheric muons. So the higher the zenith angle, the lower the atmospheric muon flux, since more matter should be passed before the detector is reached. This means that up-going and horizontal muons can't be atmospheric muons, but have to come from CC neutrino interactions close to the detector.
The up-going and horizontal muons are mainly produced by CC interactions of atmospheric neutrinos ν_{atm}, that are also products of the cosmic ray showers (~ 100/day). Since neutrinos only interact weakly with ordinary matter, they are able to travel across a huge amount of matter. So they can travel through the Earth and interact close to the detector via a CC interaction and in this way produce up-going and horizontal muons
Fig. 4.11 shows the neutrino energy spectrum, which is determined by a mixture of observations and theoretical predictions. In our analysis we will be looking for GeV-TeV neutrinos coming from the direction of the Sun. We see in Fig. 4.11 that in this energy range the flux of the atmospheric neutrinos is still strong, so they will represent an irreducible background. We also see that there is a high solar neutrino flux, coming from the nuclear fusion reactions in the center of the Sun. Since these neutrinos have energies in the MeV domain, they can't be detected by IceCube, because this detector isn't sensitive for neutrinos in this energy region.

An overview of the different kinds of events in IceCube is given in Fig. 4.12. It shows how the down-going events are dominated by atmospheric muons, while up-going events will mainly be atmospheric neutrinos and hopefully also extraterrestrial neutrinos, like e.g. neutrinos coming from the decay of the annihilation products of solar WIMPs.

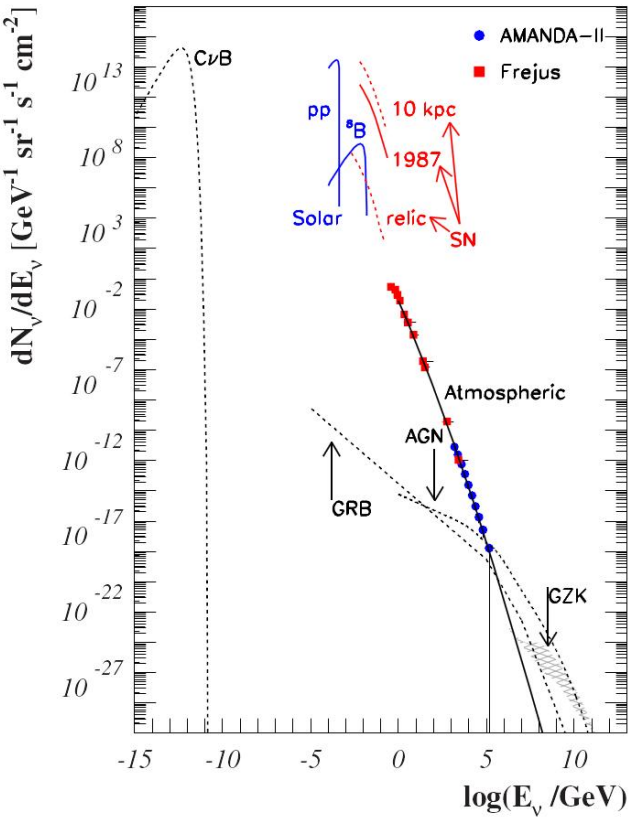

Figure 4.11: The cosmic-neutrino spectrum. Sources are the Big Bang (CνB), the Sun, supernovae (SN), atmospheric neutrinos, active galactic nuclei galaxies, and GZK neutrinos. The data points are from detectors at the Frejus underground laboratory [34] and from AMANDA [35]. Picture taken from [22]

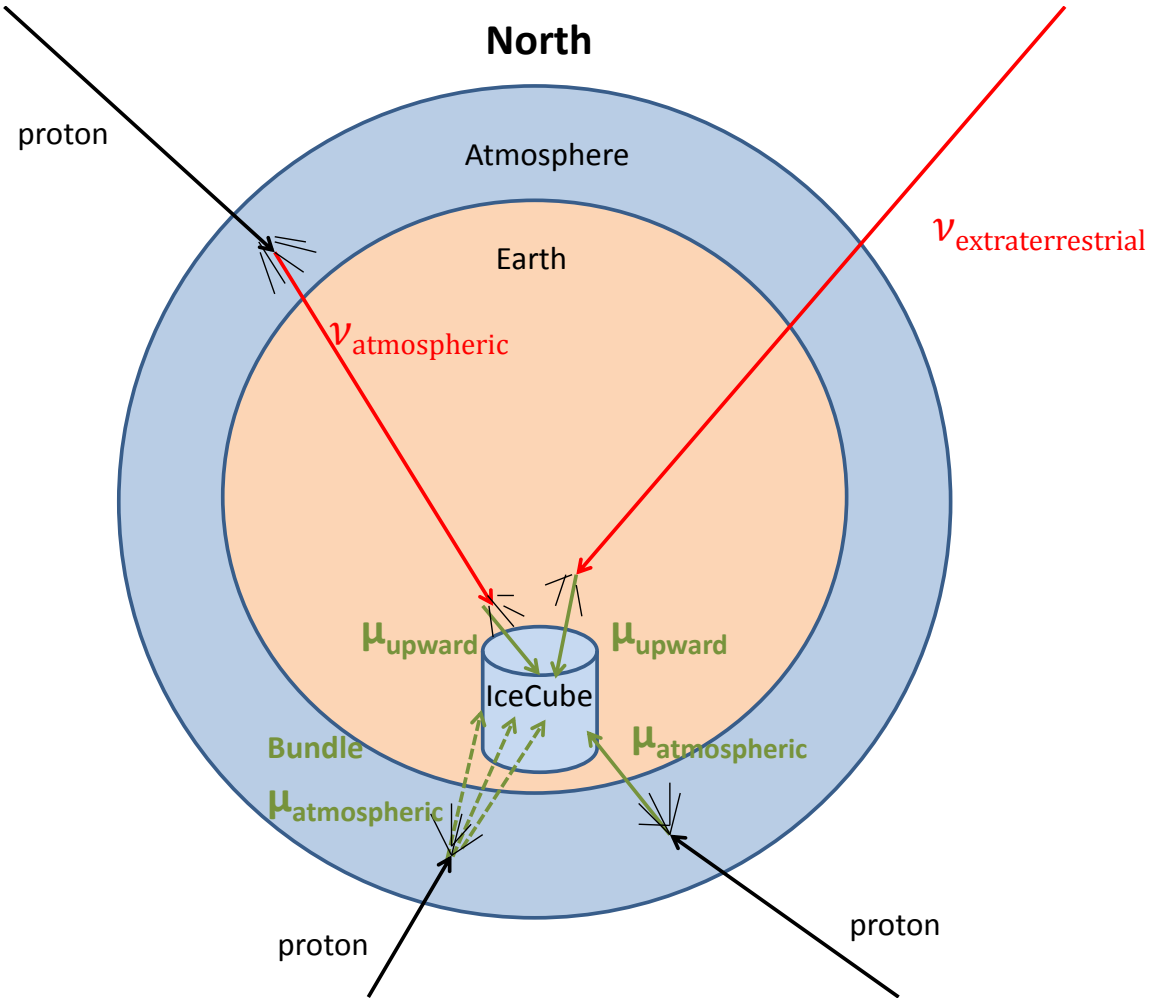

Figure 4.12: Schematic view of the different kind of events measured by IceCube. The down-going events are dominated by atmospheric muons, while up-going events will mainly be atmospheric neutrinos. For our analysis, these atmospheric muons and neutrinos are background. The extraterrestrial neutrino shown in this figure could represent e.g. a neutrino coming from the decay of the annihilation products of solar WIMPs, so this would be the kind of event we are looking for in this analysis.

Chapter 5

Simulated and Experimental Data

In the previous chapters we described what WIMPs are and how we could try to detect them by using the IceCube neutrino detector. We also talked about what kind of backgrounds we have.
In chapter 6, we will develop filters that cut away a lot of background and keep most of the signal. Although we will use off-source data to develop our filters, it is useful to have good background simulation because then we can check if there is a general agreement between the expectation and the measurement and it is also interesting to see what the contamination of various backgrounds in the final data set is.

The signal neutrinos must be simulated in order to get an idea of how the signal would look like and most importantly to calculate the effective volume, which we will need to calculate the sensitivity of the detector for our analysis.
In section 5.1, we describe how simulation works. In section 5.2 we'll say some words about the data set that was used for the analysis. The comparison between simulation and data is made in section 5.3.

5.1 Simulation

Simulation proceeds in several stages. In the first step the primary particles are *generated*, with a particular flux, energy, direction, etc.. In the second step, these particles are *propagated* through various media, (like ice, rock), until they reach the detector. This propagation takes into account energy losses and production of numerous secondaries, and it also tracks the Cherenkov photons. The last step consists of simulating the detector response.

5.1.1 Generation of events

In this thesis, we are dealing with three different kind of events : atmospheric muons, atmospheric neutrinos and signal neutrinos, each with their own event generation.

- **Atmospheric Muons** are simulated with the program COsmic Ray SImulation for KAscades (CORSIKA)[36]. This program simulates primary cosmic ray particles and tracks them through the atmosphere until they undergo reactions with nuclei or decay. In this way, extensive air showers initiated by high energy cosmic ray particles are simulated. In IceCube, we use an internally released version of CORSIKA, that includes modifications specifically created for IceCube. This version is called dCORSIKA and is developed by Dmitry Chirkin[37]. The simulation used in this thesis contains both single and coincident atmospheric muons.
- **Atmospheric Neutrinos** are simulated with the program NuGen. It simulates all neutrino events, up-going and down-going. Since only neutrinos are able to pass through the earth, only neutrinos have to be considered as primaries for up-going events in IceCube. More information about this can be found in [38].
- **WIMP Neutrinos** are simulated with the program WimpSim [39]. This program uses the DarkSUSY[40] package, which simulates several steps:

  1. the WIMP annihilation in the center of the Sun,
  2. the interactions of the annihilation products with the surrounding medium,
  3. the decays of these annihilation products,
  4. finally the propagation of the neutrinos towards the Earth surface.

  The neutrino directions are generated such that they follow the declination of the Sun throughout the year.

5.1.2 Propagators

To propagate the produced muons (atmospheric or coming from neutrino CC interactions) through the Earth (ice, rock), we use Muon Monte Carlo (MMC) [42]. MMC takes both continuous and stochastic energy losses into account.
The propagation of the produced Cherenkov photons in the ice is simulated with the software package Photonics [43]. This package takes into account the structure of the ice by varying scattering and absorption as a function of both wavelength and depth. At each point in space the photon intensity and time residual information is stored in lookup tables.

5.1.3 Detector simulation

The last step is to simulate how the detector responds on the simulated Cherenkov photons.
The photon density at each DOM is retrieved from the Photonics lookup tables. This photon density is used to generate a PMT waveform. A simulation of the DOM is then performed and random noise hits are added. If the simulated event is triggered, the information is read out and processed identically to data.

5.2 Experimental Data

As was already mentioned in section 4, IceCube is a huge detector and so there's an extremely high event rate, that is dominated by atmospheric muons. The first thing that should be checked is that the measured events are really coming from a particle and not from noise. That's why an event is only triggered if :

- there are $\geq$8 Hard Local Coincidence hits (HLC) within 2.5μs or
- there are $\geq$5 hits (HLC or SLC, see Fig. 5.1) on 1 string within 1.5μs.

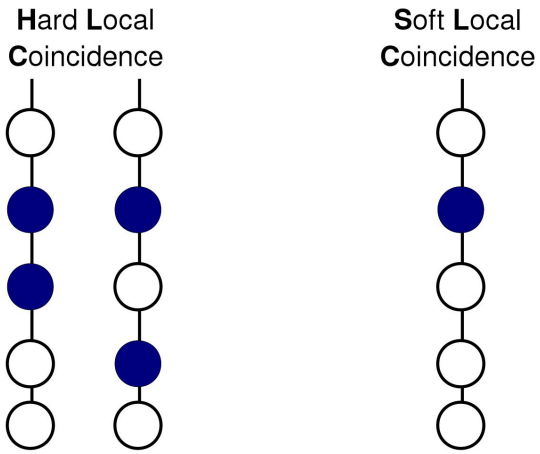

Figure 5.1: If $\geq$2 neighboring or next-to-neighboring DOMs are hit within $\sim 1\mu s$, we say there is Hard Local Coincidence (HLC) and all the hits that satisfy this condition are called HLC hits. Hits that don't satisfy this condition are called Soft Local Coincidence (SLC) hits.

Events that pass at least one of the triggers are recorded. In a second step, these events are cleaned of bad DOMs, feature extracted and fit.

Our signal comes from the Sun's direction, so it comes from the zenith range $[67°, 113°]$. We can now use the Earth as a filter against upwards-going muon background. This means that we will split our dataset into two parts :

- data taken when $\theta_{Sun} > 90°$ can be used for the analysis (on source data),
- data taken when $\theta_{Sun} < 90°$ is used as background sample (off source data).

To develop the level 3 filter we took level 2 experimental data taken during 1 run (8 hours) on the 29th of June 2009. We took this day, simply because we had that data in our hands. The level 3 filter will be developed such that the amount of atmospheric muon events is reduced to $\sim 5\%$. Since at level 2 the data is still extremely dominated

	Simulated Background				Data
	ν_{atm}	μ_{single}	$\mu_{multiple}$	total	
rate (Hz) at level 2	0.0113	29.93	10.06	39.99	46.23
# events used for L3 filter	345	861984	289728	1115712	1331424

Table 5.1: Rates and absolute numbers of the simulated background (rescaled to a 8h livetime) and data at level 2 for 1 run (8 hours) of the 29th of June 2009.

by atmospheric muons we didn't need to worry about off source and on source data yet. The data sample that we used consisted of 1331790 events, and we can see from table 5.1 that almost all these events are atmospheric muons. We really don't except that our signal will be visible at level 2, so at this level it is not a problem if we use on source data to develop the level 3 filter.

The level 4 filter will be developed by using *Boosted Decision Trees*[49] (see section 6.5), which is a multivariate method of data classification. For this we will use one month of off source data, more specifically all the events recorded in October 2009 that passed the level 3 filter. Table 5.2 shows the rates and absolute numbers of the simulated background and data events after the level 3 filtering for whole month of October 2009.

	Simulated Background				Data
	ν_{atm}	μ_{single}	$\mu_{multiple}$	total	
rate (Hz) at level 3	0.0045	1.242	0.361	1.607	2.027
# events used for BDT filter	12053	3326573	966902	4304188	5429117

Table 5.2: Rates and absolute numbers of the simulated background (rescaled to a 1 month detector livetime) and data events at level 3 for whole month of October 2009.

The events of the October data sample that pass the level 4 filter will be used to estimate the sensitivity of the IC59 detector for our analysis.

Since we didn't unblind the analysis, we will not use on source data in this thesis.

5.3 Comparing Simulation and Data

Now we compare the selected data with simulation. Comparison between data and simulation is important to understand the detector systematics.

We expect that the (low level filtered) data is dominated by atmospheric muons, so the experimental data should should be comparable with CORSIKA. Figs. 5.2 and 5.3 are 2 examples that show that there are differences between data and simulated muons. From Fig. 5.2 we see that the ratio # data/#CORSIKA is quite constant, but that the atmospheric muon flux is underestimated by the simulation. This is probably due to the fact that there is a theoretical uncertainty of $\sim 20\%$ in the estimation of the atmospheric muon flux. In Fig. 5.3 we see that for this variable not only the rates, but also the shapes of simulation and data differ. The reason that our simulation isn't perfect, is because there are many uncertainties, like e.g. the ice properties.
The aim is to use off source data to estimate the background in the selected sample. We will therefore use experimental data to develop our filters and to estimate the sensitivity. The difference between data and CORSIKA isn't really an issue for our analysis. We'll only use CORSIKA to have an idea of what the effect of our filters on the atmospheric muon rates is.

In Appendix B.1 we make this comparison between data and CORSIKA for all the variables that will be used in the level 4 (BDT) filter.

We are looking for neutrinos coming from the Sun's direction, so what matters for us is the quality of the neutrino simulation, NuGen. Before we filter the atmospheric muons away, the fraction of atmospheric neutrinos in our data set is negligible, but after applying all the filters (Level 1-4) to the data, almost all the atmospheric muons will be removed and the sample will mainly consist of atmospheric neutrinos. In Figs. 5.5 and 5.4 the data and NuGen that passed the high level filters are compared. We see that the simulation (NuGen) describes the data quite good. Note that we have much less statistics in these figures than in Figs. 5.2 and 5.3, that's why the error bars are bigger and there's more statistical fluctuation.

In Appendix B.2 we make this comparison between data and NuGen for all the variables that will be used in the level 4 (BDT) filter.

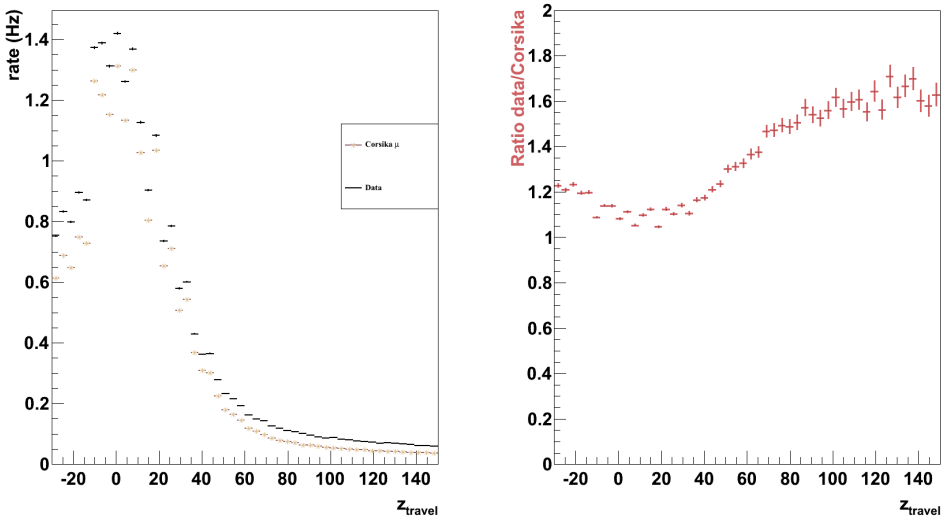

Figure 5.2: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : mean spread in z from the mean z calculated from the first quartile of hits in time (z_{travel}) of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for COGY.

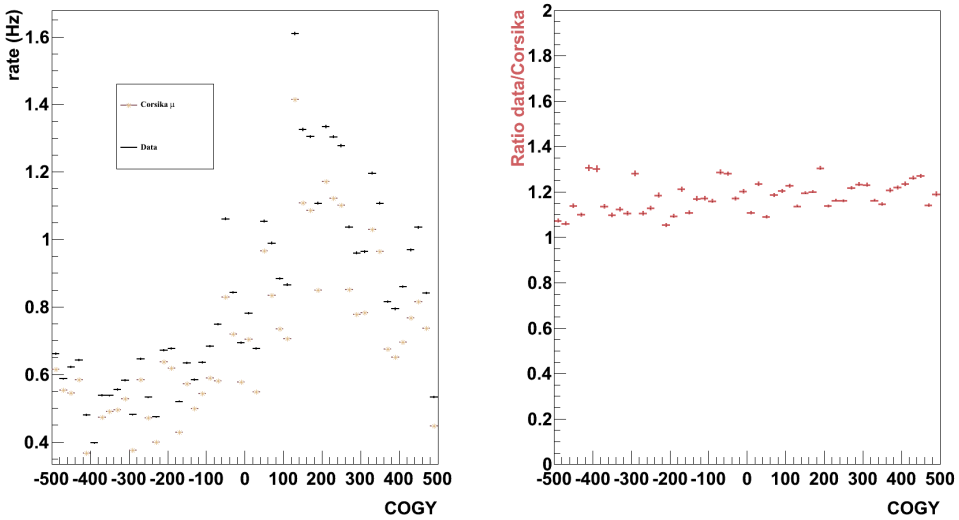

Figure 5.3: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : y component of the Center Of Gravity (COG) of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for COGY.

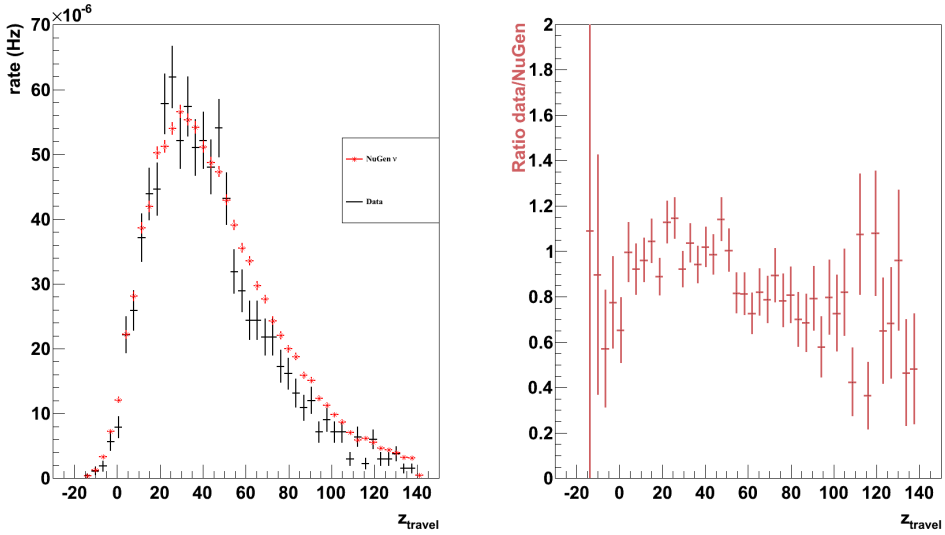

Figure 5.4: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : mean spread in z from the mean z calculated from the first quartile of hits in time (z_{travel}) of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for z_{travel}.

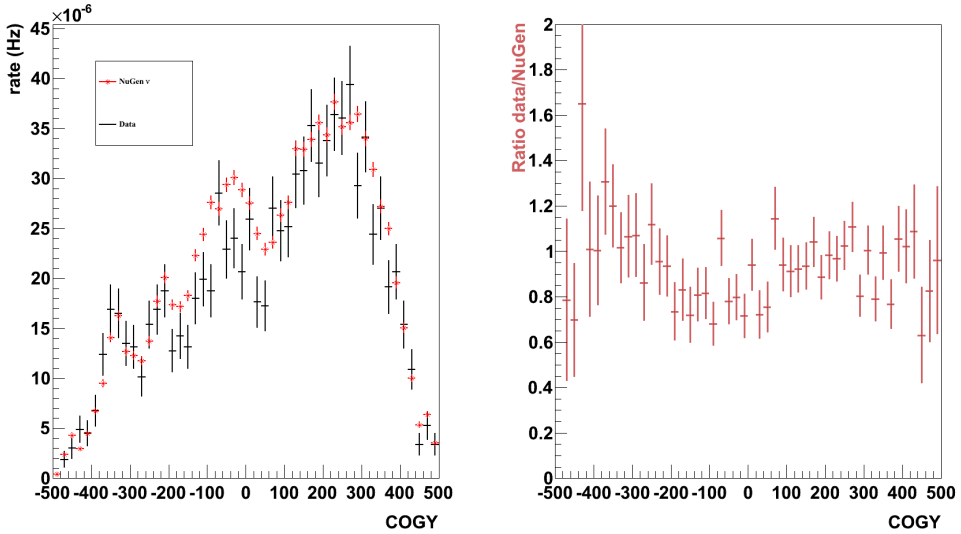

Figure 5.5: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : y component of the Center Of Gravity (COG) of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for COGY.

Chapter 6

Event selection

In this chapter we will discuss the filters we developed for this thesis. First we introduce a basic element needed for the filtering : event reconstruction. Next the concepts *effective volume* (V_{eff}) and *effective area* (A_{eff}) are explained. After this, we give an overview of the low level filtering : level 1 & 2. Next, we discuss the higher level filters that were developed by the author of this thesis.
The filtering will go in two steps : The first step will be the development of a level 3 filter that reduces the amount of background by 96%. In a second step, a level 4 filter will be developed by using *Boosted Decision Trees* (BDTs) which will allow us to remove almost all the atmospheric muons, after which we will be left with only atmospheric neutrinos as background. The events that survive all the filters will be used in the hypothesis testing in chapter 7.

6.1 Reconstruction Algorithms

The two reconstruction methods that were used in the analysis are described in this chapter. A detailed description of these and other reconstruction methods used in IceCube can be found in [45].

6.1.1 Line-Fit

The first algorithm, *line-fit*, produces an initial track on the basis of the hit times with an optional amplitude weight. Line-fit doesn't take into account the geometry of the Cherenkov cone, nor the optical properties of the medium. It assumes light travels through the detector with a velocity $\mathbf{v}$ along a 1-dimensional path. DOMs that are located at $\mathbf{r}_i$ and hit at time t_i can be connected by a line

$$\mathbf{r_i} = \mathbf{r} + \mathbf{v} \cdot \mathbf{t_i}. \tag{6.1}$$

Using eq. (6.1) we can define a χ^2 that has to be minimized

$$\chi^2 \equiv \sum_{i=1}^{N_{hit}} (\mathbf{r}_i - \mathbf{r} - \mathbf{v} \cdot t_i)^2, \tag{6.2}$$

where N_{hit} is the number of hits. Eq. (6.2) is minimized by differentiation with respect to the free parameters $\mathbf{r}$ and $\mathbf{v}$. If we solve this analytically, we find

$$\begin{cases} \mathbf{r} = \langle \mathbf{r}_i \rangle - \mathbf{v} \cdot \langle t_i \rangle, \\ \mathbf{v} = \frac{\langle \mathbf{r}_i \cdot t_i \rangle - \langle \mathbf{r}_i \rangle \cdot \langle t_i \rangle}{\langle t_i^2 \rangle - \langle t_i \rangle^2}, \end{cases} \tag{6.3}$$

with $\langle x_i \rangle \equiv \frac{1}{N_{hit}} \Sigma_{i=1}^{N_{hit}} x_i$. The line-fit thus yields

- a vertex point $\mathbf{r}$, and
- a direction $\mathbf{e} = \mathbf{v}_{\mathrm{LF}}/|\mathbf{v}_{\mathrm{LF}}|$. The zenith angle is given by

$$\theta_{\mathrm{LF}} \equiv -\arccos(v_z/|\mathbf{v}_{\mathrm{LF}}|). \tag{6.4}$$

6.1.2 Likelihood Fit

A more elaborate track reconstruction is obtained by using a Likelihood fit.

In this reconstruction method, the parameters of the track are denoted by the set of values $\mathbf{a} = \{a_i\}$ and the measured quantities by $\mathbf{x} = \{x_i\}$. The likelihood that we measure a set of quantities $\mathbf{x}$, given the track parameters $\mathbf{a}$ is given by

$$\mathcal{L}(\mathbf{x}|\mathbf{a}) = \prod_i p(x_i|\mathbf{a}), \tag{6.5}$$

where $p(x_i|\mathbf{a})$ is the probability density function (PDF) for observing the values x_i, given the hypothetical track parameters $\mathbf{a}$.

The likelihood PDF can be written in terms of time residuals t_{res} [46], which is the difference between the measured time and the expected time of the hit $t_{res} \equiv t_{\text{measured}} - t_{\text{expected}}$. The expected time of a hit represents the time it takes for a Cherenkov photon to travel that same distance without being scattered. Writing the likelihood in terms of the time residuals

$$\mathcal{L} = \prod_i \mathcal{P}(t_{res,i}|\mathbf{a}), \tag{6.6}$$

where $\mathcal{P}$ is the *Pandel function*[45], which is an analytic function that takes into account the absorption and scattering of the Cherenkov photons in the ice. The Pandel function doesn't take into account the PMT jitter and the potential for negative time residuals from random noise in the detector, so we have to convolute it with a Gaussian to get the probability of observing the measured hits at each PMT. If we

now multiply the likelihoods of all the PMTs, we get a total likelihood $\mathcal{L}_{tot}$ that can be maximized. The muon track that maximizes $\mathcal{L}_{tot}$ is then our best guess hypothesis. Actually, we don't look for a maximum of the log likelihood, but for a minimum of $-\mathcal{L}_{tot}$.

The likelihood based event reconstructions described above require a track vertex and direction as a seed. The result of the line-fit reconstruction is used as a seed in our analysis. It could be that, using this direction as a seed, the minimum found for $-\mathcal{L}_{tot}$ is just a local and not the true, global minimum. In order to mitigate this effect, we perform 8 iterations for the advanced reconstructions, with a variety of seed directions.

6.2 Generated And Effective Volume And Effective Area

6.2.1 Generated And Effective Volume

In the simulation of WIMP events, a certain volume is generated around the detector array for each event. The interaction vertices of the simulated events are spread in this *generated volume* V_{gen}, that is different for every event and depends on the energy and direction of the neutrino. V_{gen} is defined as a box around the detector with a length L along the neutrino track and the projected area of the sensitive volume as cross section. Fig. 6.1 shows V_{gen} for a certain event. The average generated volume $\langle V_{gen} \rangle$ will be shown ofr the event sample at every step of the filtering. We expect that this $\langle V_{gen} \rangle$ increases, because during the filtering we will cut away the lowest energy events, and the lower the energy of the event, the smaller V_{gen}. This is because the generated volume depends on the range of the muon, so the higher the energy of the neutrino, the higher the energy of the produced muon and the bigger the generated volume. Table 6.1 shows $\langle V_{gen} \rangle$ after the low level filtering (see section 6.3), for the 14 different WIMP models.

From the generated volumes of the events, we can derive the effective volume V_{eff} of the detector. V_{eff} is a measurement of the efficiency of the detector. It can be interpreted as the proper volume of a detector, which has an ideal efficiency ($\varepsilon \equiv 1$). An ideal efficiency means that every event within the volume is triggered and selected. So within this effective volume

1. every WIMP neutrino with muon flavor interacts and produces a muon,

2. this muon triggers the detector and passes the cuts.

The effective volume is thus a fraction of the generated volume. Within V_{eff} we measure the fraction of generated events that passed the triggers and selection criteria within V_{gen} :

$$V_{eff} = \frac{N_{sel}}{N_{gen}} V_{gen} = \frac{\sum_{i=1}^{N_{gen}} w_i(E_\nu) \delta_i V_{gen,i}(E_\nu, \theta_\nu)}{\sum_{i=1}^{N_{gen}} w_i(E_\nu)}, \tag{6.7}$$

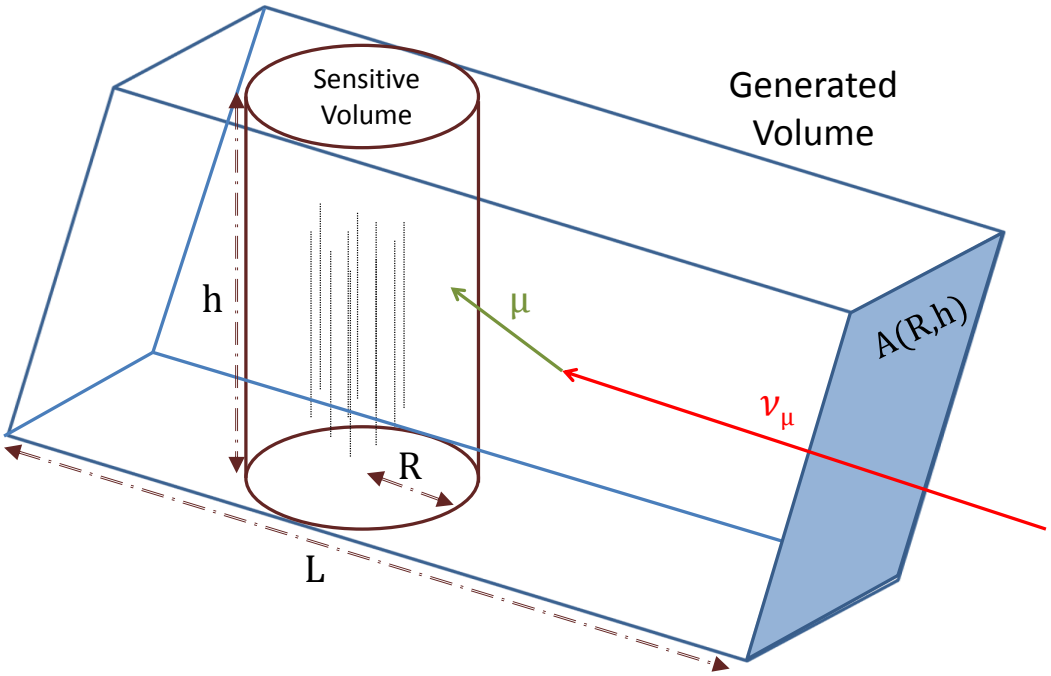

Figure 6.1: The generated volume V_{gen} for a WIMP neutrino is shown, together with the detector volume. This generated volume has a length L along the neutrino track and the projected area of the sensitive volume as cross section.

where $w_i(E_\nu)$ is the weight of the event, δ_i is either 1 or 0 if the event respectively did or didn't pass the selection criteria and $V_{gen,i}(E_\nu, \theta_\nu)$ is the generated volume of the event. The weight w_i of an event denotes the statistical importance of generated event i in the total simulated sample. So each generated event i corresponds to w_i physical events, where w_i is a combination of different weights : w_χ, w_{abs}, w_σ and w_{vertex}[44]. Let us now explain what the meaning of these different weights is

- The physical distribution of the neutralino-induced neutrino energy spectrum at the surface of the Earth is given by $\frac{dN_\nu}{dE_\nu}$. This spectrum is however sampled from $\frac{dN_\nu}{dE_\nu} \times E_\nu^\alpha$, where E_ν^α is chosen such that the low statistics regions are artificially populated[1]. In order to correct for the unphysical E_ν^α factor, each event is weighted with $E_\nu^{-\alpha} \equiv w_\chi$.

- When a neutrino travels through the Earth, it is possible that it gets absorbed

[1]The cross-section increases almost linearly with neutrino energy.[30]

WIMP Mass (GeV)	WIMP decay channel	$\langle V_{gen} \rangle (km^3)$
50	soft	3.792
	hard	3.976
100	soft	3.969
	hard	4.164
250	soft	4.114
	hard	4.695
500	soft	4.295
	hard	5.101
1000	soft	4.451
	hard	5.330
3000	soft	4.610
	hard	5.306
5000	soft	4.646
	hard	5.256

Table 6.1: Average generation volume ($\langle V_{gen} \rangle$) after the level 2 filtering, for different WIMP models.

before it reaches the detector. In this case we won't measure the neutrino. To remove the absorbed neutrinos from the file, we introduce a weight w_{abs}, that is 0 (1) for absorbed (surviving) events.

- The interaction probability depends on the cross section of the neutrino, so every event has a different weight w_σ, depending on this cross section.

- We already mentioned that a certain volume is generated around the detector for each event. Consider N_{gen} neutrino interactions is a reference volume V^0_{gen}, with interaction density $\rho = N_{gen}/V^0_{gen}$. If we want to distribute these N_{gen} vertices in another volume V_{gen}, we need to weight the interactions in order to recover the correct interaction density : $N_{gen} \rightarrow N'_{gen} = w_{vertrex} \times N_{gen}$, with

$$\begin{aligned} \rho &= \rho' \\ \frac{N_{gen}}{V^0_{gen}} &= \frac{N'_{gen}}{V_{gen}} \\ \Rightarrow w_{vetrex} &\equiv \frac{N'_{gen}}{N_{gen}} = \frac{V_{gen}}{V^0_{gen}}. \end{aligned}$$

The total weight of an event is thus given by :

$$w_i = w_\chi \times w_{abs} \times w_\sigma \times w_{vertex}. \tag{6.8}$$

The effective volume V_{eff} of the detector, when looking at 1000GeV hard annihilating WIMPs, after the low level (level 1 and 2) filtering (see section 6.3) is

$$V_{eff} = 0.420 km^3. \tag{6.9}$$

Unfortunately we don't know the values of V_{eff} and $\langle V_{gen} \rangle$ at trigger level, since the simulated WIMP-files we have for our analysis were already (low level) filtered.

6.2.2 Effective Area

The expected rate of events for a certain theoretical model is given by

$$\frac{\mathrm{d}n(\nu)}{\mathrm{dt}} = \int \frac{\mathrm{d}\phi}{\mathrm{d}E} \cdot \boxed{\sigma_{\nu\to\mu}(E,\theta) \cdot P_\mu(E) \cdot \epsilon_{\mathrm{det}}(E,\theta,...) \cdot \epsilon_{\mathrm{reco}}(E,\theta) \cdot \mathrm{d}A} \cdot \mathrm{d}E \cdot \mathrm{d}\Omega, \tag{6.10}$$

where $\frac{\mathrm{d}\phi}{\mathrm{d}E}$ is the model dependent neutrino flux, $\sigma_{\nu\to\mu}$ is the cross section for $\nu \to \mu$, is the probability that the μ is detectable (close enough to the detector), $\epsilon_{\mathrm{det}}(\mathrm{E},\theta,...)$ is the detector efficiency to detect the incoming μ, $\epsilon_{\mathrm{reco}}(\mathrm{E},\theta)$ is the efficiency that the event is reconstructed and triggered, and Ω is the solid angle :

$$\Omega = 2\pi \times \Big(\cos(\theta_{min}) - \cos(\theta_{max})\Big). \tag{6.11}$$

The boxed part of eq. (6.10) is defined as the *effective area* of the detector.

$$A_{eff}(E,\theta) = \int \sigma_{\nu\to\mu}(E,\theta) \cdot P_\mu(E) \cdot \epsilon_{\mathrm{det}}(E,\theta,...) \cdot \epsilon_{\mathrm{reco}}(E,\theta) \cdot \mathrm{d}A. \tag{6.12}$$

This effective area is the surface corresponding to 100% detection efficiency :

$$A_{eff}(E,\theta) = \frac{N_{selected}(E,\theta)}{N_{gen}(E,\theta)} A_{gen}, \tag{6.13}$$

with A_{gen} the generated detector area, $N_{gen}(E,\theta)$ the amount of generated events and $N_{selected}(E,\theta)$ the amount of selected events, i.e. the amount of events that are really detected and passed the selection criteria.

In Fig. 6.2, the effective area in m^2 is given as a function of the energy of the neutrino coming from 1000GeV hard WIMP annihilation E_ν. The histogram shows the effective area for different solid angles (see section 6.4.1):

- $[\theta_{min}, \theta_{max}] = [67°, 113°]$, this is the Sun's zenith range (thick dashed dotted line),
- $[\theta_{min}, \theta_{max}] = [67°, 90°]$, this is the zenith range in which the Sun is in the Southern hemisphere (gray small dashed dotted)

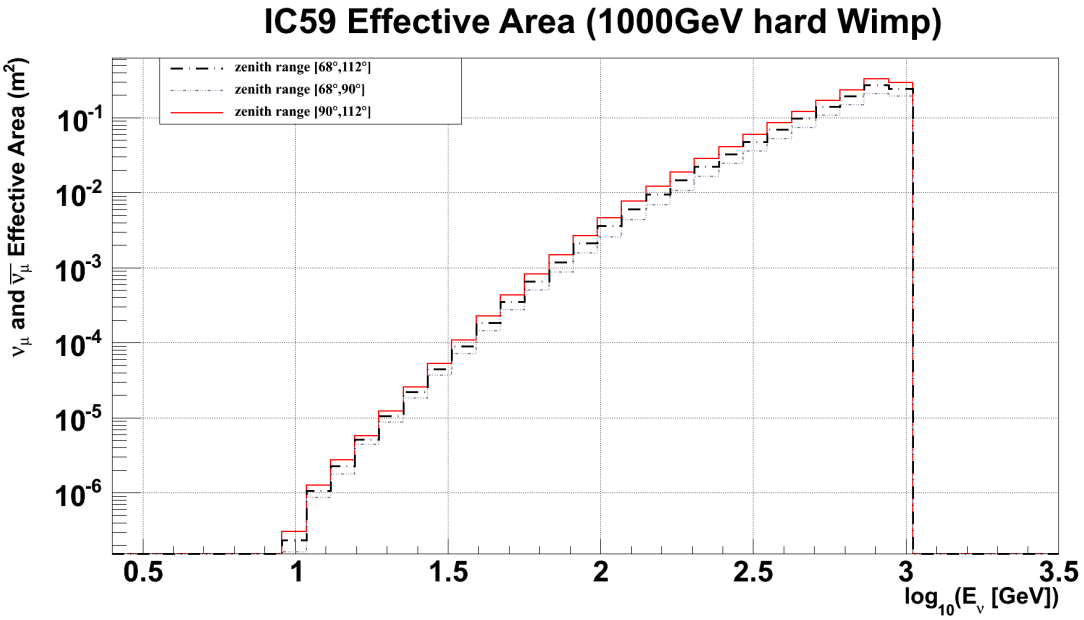

Figure 6.2: Neutrino effective area of the IC59 detector in function of $\log_{10} E_{\nu_\mu}$. This effective area is shown for neutrinos that pass the low level (1 & 2) filters (see next section 6.3. The three different lines represent different zenith ranges : $[67°, 113°]$ (thick dashed dotted line), $[67°, 90°]$ (gray small dashed dotted) and $[90°, 113°]$.

- $[\theta_{min}, \theta_{max}] = [90°, 113°]$, this is the zenith range in which the Sun is in the Northern hemisphere (solid red line)

Now that the effective area of the detector is calculated, one can calculate the expected rate of events for a certain theoretical model :

$$\frac{\mathrm{d}n(\nu)}{\mathrm{dt}} = \int \frac{\mathrm{d}\phi}{\mathrm{d}E} \cdot A_{eff}(E, \theta) \cdot \mathrm{d}E \cdot \mathrm{d}\Omega, \tag{6.14}$$

This can now be used to reject a model or put limits on parameters of the model, if the observed event rate is lower than the expected event rate.

We will show the effective area distribution after every step, but we will use a different method to analyze the final data set.

6.3 Level 1 and Level 2 Filtering

The data is dominated by atmospheric muons, which are background for us, so we want to filter these events away. Another reason why we want to filter away a lot of background, is because the data has to be sent over by satellite, which is limited in bandwidth. In a first step, a simple filter is applied on all the events, which are extremely dominated by background. Once the data is sent to the North, more elaborate filters are used.

For our analysis, we only want up-going events (see Fig. 4.12), that's why we only selected the data that passed the *MuonFilter* or the *LowUpFilter*.

The *MuonFilter* rejects most of the low energy, down-going background. In this filter, the events are divided into 5 categories and for each category there's a different energy threshold :

1. events with $\theta_{\text{track}} \in [80°, 180°]$ are all accepted $\Rightarrow$ all the up-going and horizontal events (also low energy) pass this filter.
2. events with $\theta_{\text{track}} \in [70°, 80°]$ are accepted if $^{10}\log(\text{total charge}) > 1.5$,
3. events with $\theta_{\text{track}} \in [60°, 70°]$ are accepted if $^{10}\log(\text{total charge}) > 2.0$,
4. events with $\theta_{\text{track}} \in [50°, 60°]$ are accepted if $^{10}\log(\text{total charge}) > 2.3$,
5. events with $\theta_{\text{track}} \in [0°, 50°]$ are accepted if $^{10}\log(\text{total charge}) > 2.7$.

Since the total charge is a measurement of the energy, we can deduce from (2)-(5) that the more vertical a down-going event, the higher the energy must be to pass this filter. We are looking for events that come from the direction of the Sun, which are $\sim$ horizontal events. We will also only look for signal during the South Pole winter, when the Sun is below the horizon. This means that we will only look for horizontal and up-going events, which are exactly the events that pass the filter.

The *LowUpFilter* really selects horizontal and up-going events. It has the following criteria

1. the amount of fired DOMs ≥ 5,
2. $\theta_{\text{track}} > 80°$, where θ_{track} is the reconstructed zenith of the muon track,
3. $z_{travel} \geq -10$m, where $z_{travel} \equiv$ is the mean spread in z from the mean z calculated from the first quartile of hits in time,
4. $z_{ext} \leq 600$m, where z_{ext} is the extension of the track in the z-direction (vertical),
5. at least 1 inner string should be hit,

6. and it uses the upper part of the detector as veto.

Criterion (1) removes the events for which aren't reconstructible, (2) and (3) filter away the down-going events (see section 6.4), (4) removes high energy events that cross more than 600m of the detector vertically, (5) makes sure that the event takes place in the detector and isn't just the afterglow of a muon outside the detector and (6) removes most the down-going muons.

6.4 Developing A Level 3 Filter

After the level 2 filtering, more cuts need to be made in order to

- reduce the background as much as possible,
- keep as much signal as possible.

The level 3 cuts are made by looking at the histograms of some variables and choosing those variables for which there is a clear separation between signal and background. The cuts are optimised for a WIMP mass of 1000 GeV, annihilating via the hard channel. To decide how exactly the cut values are chosen, one could look at the significance of the cut (Λ_{cut}). A possible definition of this significance can be

$$\Lambda_{cut} \;=\; \frac{n_{sig,ac}}{n_{sig,bc}} \times \left(1 \;-\; \frac{n_{bkg,ac}}{n_{bkg,bc}}\right) \;\equiv\; \varepsilon_{sig} \times (1 \;-\; \varepsilon_{bkg}) \tag{6.15}$$

where $n_{sig,ac}(n_{sig,bc})$ and $n_{bkg,ac}(n_{bkg,bc})$ are respectively the amount of signal events after (before) the cuts and the amount of background events after (before) the cuts. One could now chose the cut such that this efficiency is maximal. As will be shown, this is not the best way to choose the cut values, since it is hard to find a maximum significance for certain parameters.

The cut values will be chosen such that most of the signal is kept. The goal is to reduce the background to $\sim 5\%$. After doing the level 3 filtering, there will be a further, more fine tuned level 4 cut. This level 4 cut will be determined by the Boosted Decision Tree method (see later).

6.4.1 Season Selection

Our signal comes from the Sun's direction, so it comes from the zenith range $[67°, 113°]$. Neutrinos only interact rarely with matter, so they are able to cross the Earth and sometimes CC interact in (or close to) the IceCube detector, in which case they can be detected. Atmospheric muons on the other hand are absorbed very fast on their way through the Earth, so we use the Earth as a filter against upwards-going muon background, by splitting our data set into 2 parts :

- data taken when $\theta_{Sun} > 90°$ will be used for the analysis (on source data, green region in Fig. 6.3), so we keep signal events with $\theta_{\text{sig}} > 90°$,
- data taken when $\theta_{Sun} < 90°$ will be used as background sample (off source data, red region in Fig. 6.3), so we remove signal events with $\theta_{\text{sig}} < 90°$,. This is done to ensure that there is no possible up-going signal contamination.

Fig. 6.4 shows the simulated zenith angle θ_{WIMPMC} for neutrinos coming from 1000GeV hard WIMPs. We remove events with $\theta_{WIMPMC} < 90°$. The effect of this cut on the amount of signal is shown in Tables 6.2 and 6.3. The distribution is asymmetric, because this plot is made using level 2 data, so most of the low energy, down-going events are cut away.

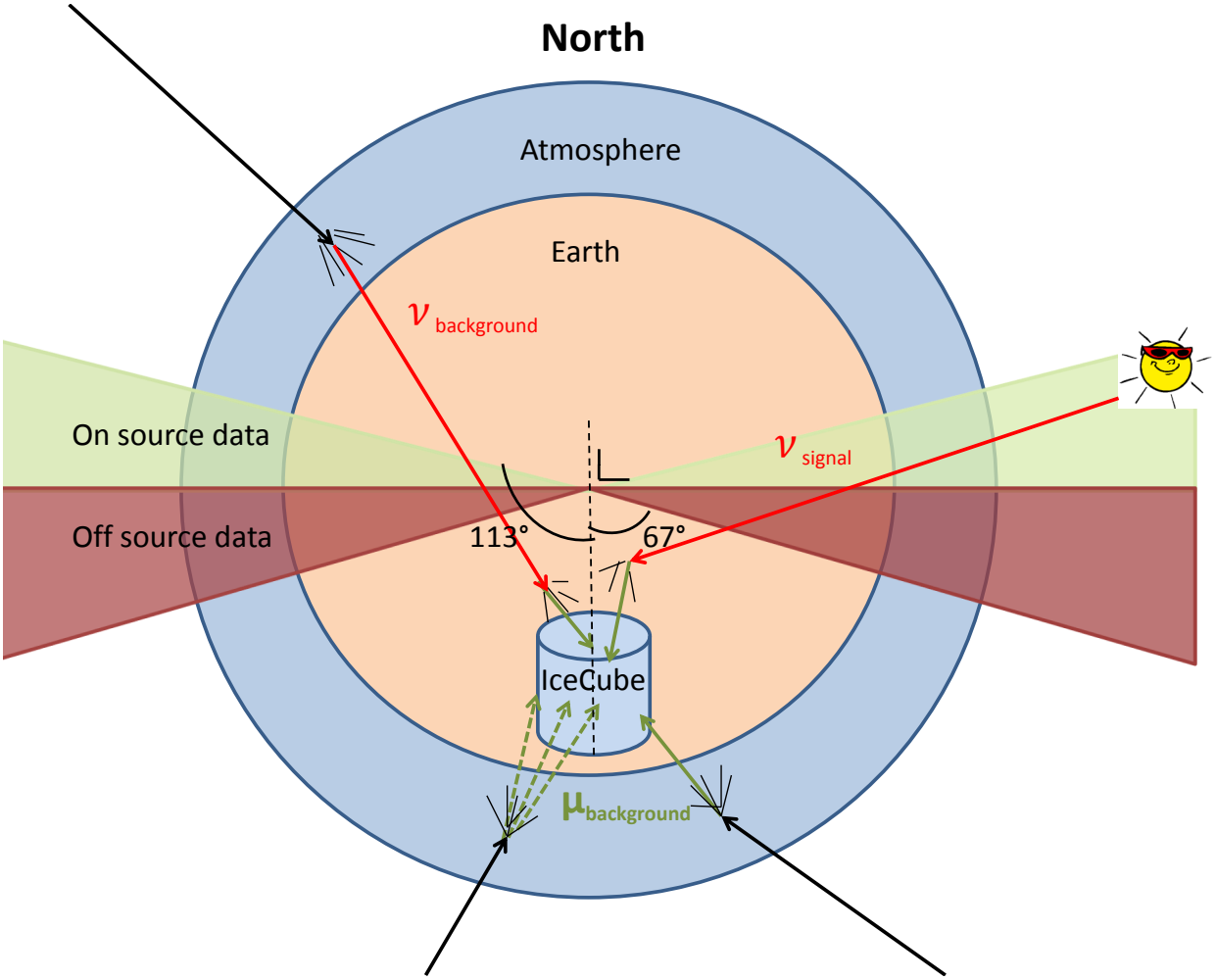

Figure 6.3: Schematic view of the directions of the on source and off source data.

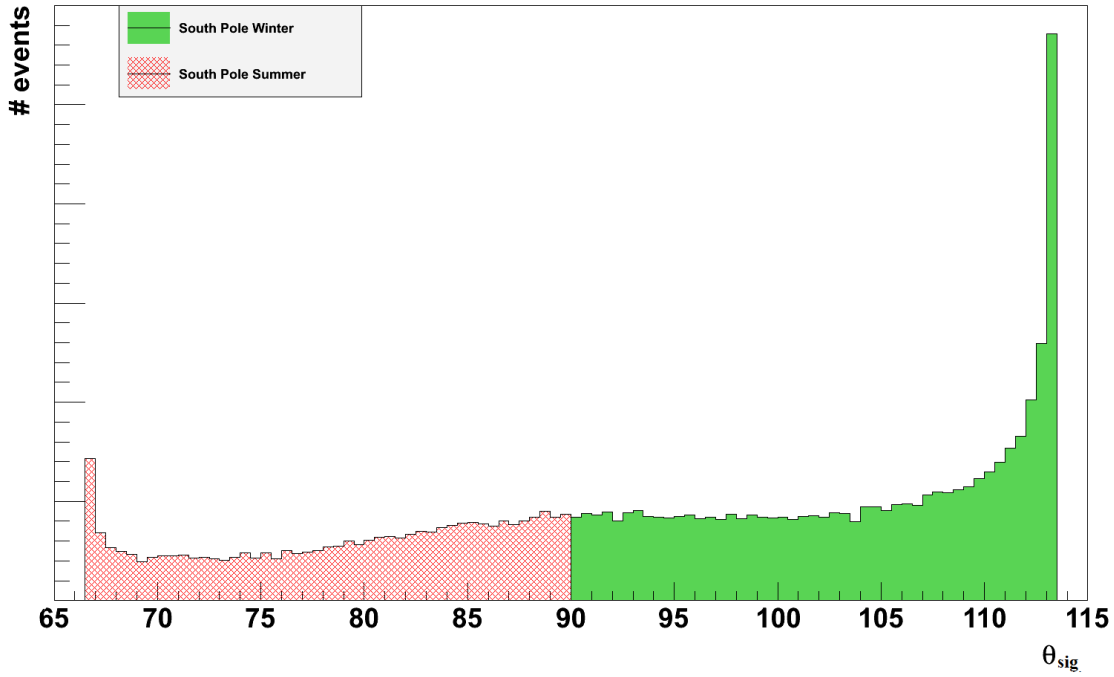

Figure 6.4: Simulated zenith angle θ_{WIMPMC} for neutrinos coming from 1000GeV hard WIMPs. The histogram is divided into 2 parts : South Pole Summer (off source data) and South Pole Winter (on source data).

6.4.2 LLH_{Fit} Filter

Now a filter is introduced which only allows events that are reconstructed by LLH_{Fit} (see section 6.1.2). This advanced reconstruction algorithm will be needed later to reconstruct e.g. the track direction.

Tables 6.2 shows the effect of this filter and the season selection of previous section on the amount of simulated signal events and on the rates of the simulated background and data. In Table 6.3, the effect of the season cut and the trigger on the amount of signal volume and the average generated volume, is shown for the different WIMP models.

	# WIMP (%)	Simulated Background (Hz)				Data (Hz)
		ν_{atm}	μ_{single}	$\mu_{multiple}$	total	
level 2	100	0.0113	29.93	10.06	39.99	46.23
θ_{WIMPMC} >90°	64.59	0.0113	29.92	10.06	39.99	46.23
$\exists LLH_{Fit}$	63.22	0.0103	25.98	9.625	35.62	41.86

Table 6.2: Percentages of the amount of simulated signal and rates of the simulated background and data events before and after the filter(s)

WIMP Mass (GeV)	WIMP decay channel	# WIMP (%) after the precuts	$\langle V_{gen}\rangle(km^3)$
50	soft	26.23	3.866
	hard	35.84	4.060
100	soft	35.06	4.050
	hard	52.52	4.298
250	soft	51.96	4.283
	hard	63.69	4.885
500	soft	58.32	4.488
	hard	63.96	5.310
1000	soft	61.47	4.656
	hard	63.22	5.540
3000	soft	62.31	4.825
	hard	63.04	5.530
5000	soft	62.40	4.858
	hard	63.14	5.467

Table 6.3: Percentages of the amount of simulated signal and the average generated volume ($\langle V_{gen}\rangle$) after the level 3 precuts (w.r.t. level 2), for different WIMP models.

The effective volume V_{eff} of the detector, when looking at 1000GeV hard annihilating WIMPs, after the level 3 precuts is

$$V_{eff} = 0.266 km^3. \tag{6.16}$$

The effective area for the 1000GeV hard signal, after these precuts is shown in Fig.6.5. We see that there's a decrease in effective area mainly at low energies, as these neutrinos create muons that do not generate a lot of light in the detector, and thus fail the reconstruction.

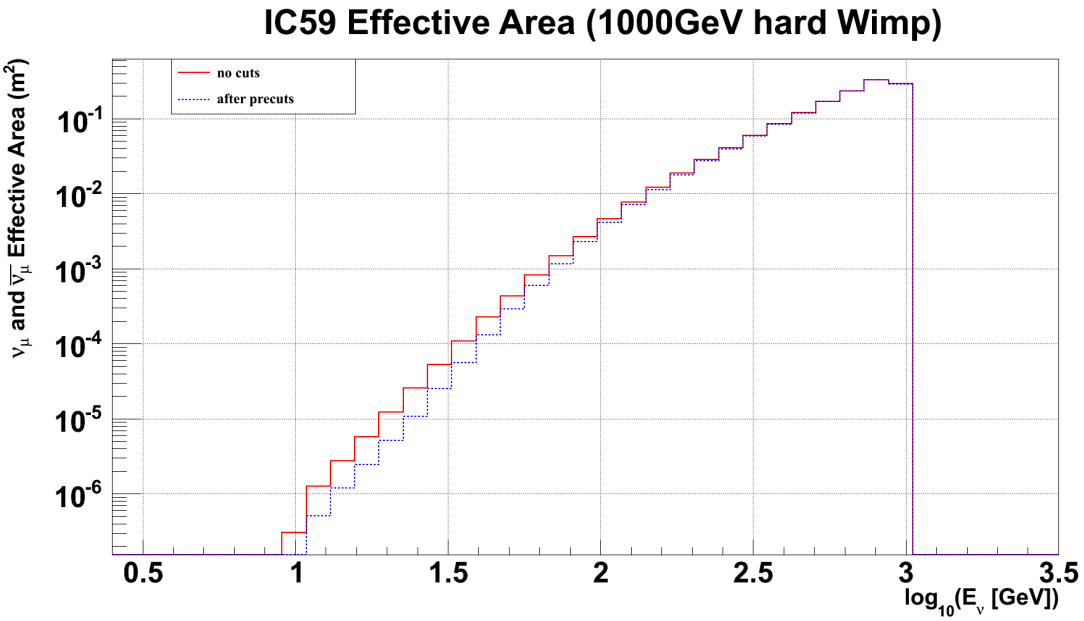

Figure 6.5: Neutrino effective area of the IC59 detector in function of $\log_{10} E_{\nu_\mu}$. This effective area is shown before level 3 filtering (solid red line) and for events that passed the level 3 precuts (blue dotted line).

6.4.3 Determining the Level 3 Cuts

Events which pass the season selection and the LLH_{Fit} trigger from previous subsection, will now be used in what follows. We looked at several (~ 40) different parameters to cut on and chose to use the following variables for our level 3 filter :

1. $\theta_{LLH} \equiv$ Reconstructed muon Zenith. Very logical parameter to cut on : signal comes from the Sun, so you expect it to come from a region between $90° < \theta_{\text{sig}} < 113°$.

2. $z_{travel} \equiv$ Mean spread in z from the mean z calculated from the first quartile of hits in time. So negative values come from down-going events (e.g. background muons) while positive come from up-going events (e.g. signal).

3. $t_{extension} \equiv$ Time between start of first and last pulse in the event; everything that is in the tail of this is pure (badly reconstructed) background. This tail is dominated by coincident muons, that are reconstructed as single muons. This is because nor the line-fit, nor the log likelihood reconstruction are able to distinguish coincident muons. Because of this merging of multiple muons, the event seems to last very long, which is wrong.

4. $\Delta_{LLh,linefit} \equiv$ Angle between linefit reconstructed track and LLH reconstructed track. This is a selection on well reconstructed tracks, in which case this angle is small, because there was a clear track that was reconstructed in the same way by both linefit and the LLH fit.

5. $D_{dom,reco} \equiv$ Average distance between the reconstructed track and the fired DOMs. For low energy events (e.g. our signal), this distance is small, the higher the energy, the bigger this distance gets.

Let's go through the histograms of each of these variables. These histograms show :

- The experimental data (with errors) : black dots with error bars. This is data of 1 run (8 hours) taken at the 29th of June 2009 (1331790 events at level 2). We explained in section 5.2 that there it is not problematic to use on source data to develop the level 3 filter.

- The simulated background (total) : brown dashed line, this simulated background consists of

  - atmospheric neutrinos (ν_{atm}) : red line,
  - single muons (μ_{single}) : green line,
  - multiple/coincident muons ($\mu_{multiple}$) : dotted green line,

- The simulated WIMP signal : blue line. As was already said, this is simulation for 1000 GeV hard annihilating WIMPs.

The single and multiple muons are both products of cosmic ray showers. It could be that only one (single) muon reaches/triggers the detector, or that (multiple) coincident muons trigger the detector. Coincident muons could come from the same shower or from different, coincident showers.

The histograms are shown both on a linear scale (left) and a logarithmic scale (right). This is because the rate of atmospheric neutrinos is really small, so they are only visible on the logarithmic plot. Although these atmospheric neutrinos only have a small contribution to the background at this level, we can certainly not neglect them, because atmospheric neutrinos are indistinguishable from the expected neutrinos coming from WIMP decays. We will see that these atmospheric neutrinos will become the dominant background after all the filtering (see Figs. 6.23 and 6.24).

Note that the signal simulation on the plots doesn't show the true signal rate, but a scaled rate. This is because the expected (unknown) rate is so small that it wouldn't be visible on the histograms. By scaling the signal rate to the experimental data rate, the shape of the histograms of signal and background can be compared and it is this comparison that is needed to make the cuts.

Let's have a look at the histograms of the 5 parameters now. The histograms show the uncut data and simulation, that passed the filter of the previous section (6.4.2). At this level we don't take the influence of the cuts on one parameter to another into account.

The first parameter we look at is the reconstructed muon zenith angle (θ_{LLH}); the histogram of this variable is shown in Fig 6.6.

This histogram clearly shows that the neutrinos coming from the solar WIMPs have a restricted zenith area between $\sim 85°$ and $\sim 130°$. The fact that this reconstructed zenith is not strictly in the region $90° < \theta_{LLH} < 113°$ has two reasons. First of all, this is because the neutrino coming from the WIMP annihilation in the Sun has not exactly the same direction as the detected muon coming from that neutrino, as shown in the left part of Fig. 4.3. The second reason is that the reconstruction algorithm isn't perfect, so the reconstructed muon direction isn't exactly the same as the true muon direction, this is what we see in the right part of Fig. 4.3. These effects are shown in figure 6.7.

Now, we have to choose the cut values. As was already mentioned before, these values can be determined by cutting at the value which gives the highest significance (see eq. 6.15). The significance of the cut for the lower bound of θ_{LLH} is shown in Fig. 6.8.

Figure 6.8 shows that the cut value with the highest significance is around 89°. As was already mentioned before, the cuts will not be based on these significances, but they will be more relaxed, to make sure that as much signal as possible is left after the cut. In the case of the lower value of the reconstructed zenith, a value of 85° is chosen.

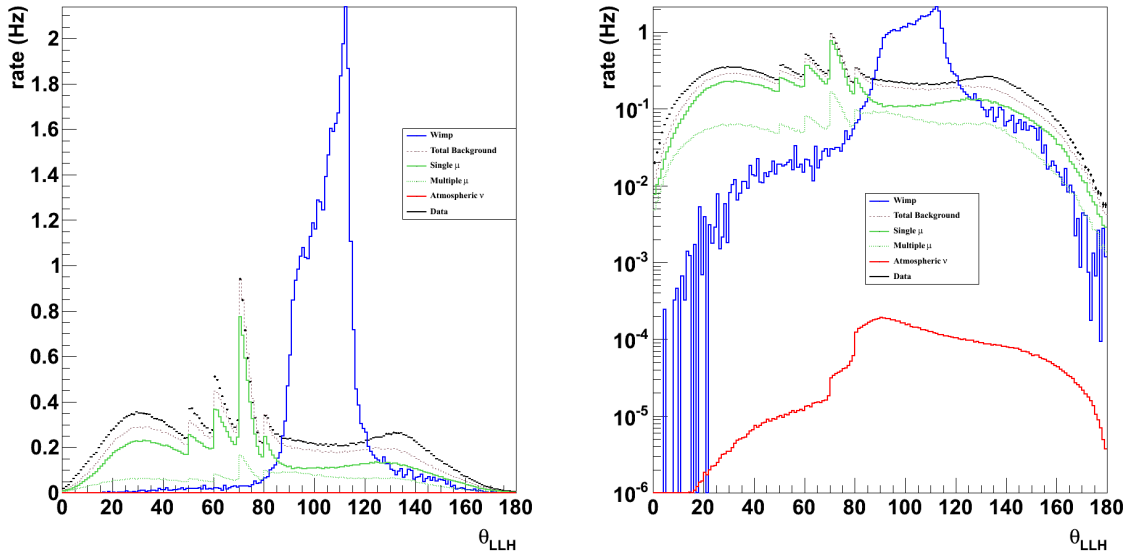

Figure 6.6: Histogram of the Reconstructed Zenith angle of the detected muon, expressed in °. Left : linear scale, right : logarithmic scale

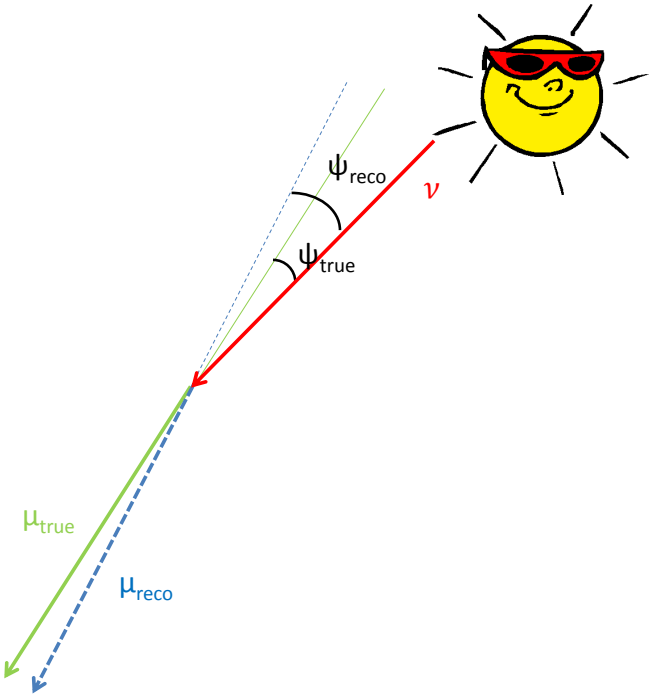

Figure 6.7: Explanation of the difference between the neutrino direction and the direction of the reconstructed muon.

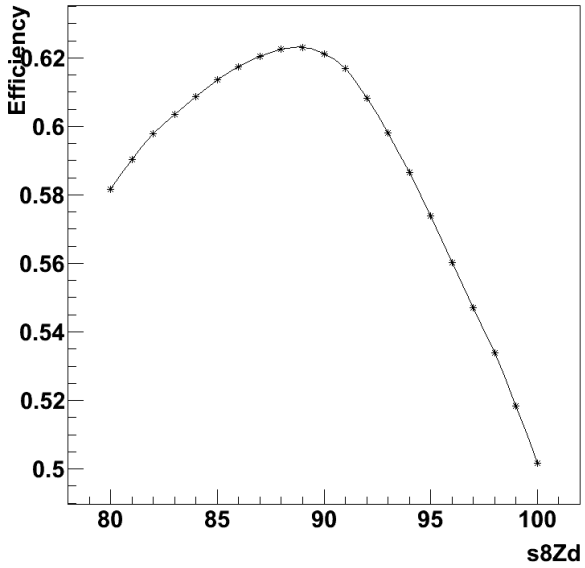

Figure 6.8: Significance of the cut for the lower bound of θ_{LLH}.

As upperbound, a value of 130° is chosen. So only events which satisfy

$$85° < \theta_{LLH} < 130°, \tag{6.17}$$

are kept.

The next parameter we look at is z_{travel}. Its histogram is shown in Fig. 6.9.

Negative values of z_{travel} mean that the event is reconstructed as down-going. So most of the muon background will have a negative z_{travel} value, while signal events should have positive values. We see that a small amount of signal has $z_{travel} < 0$. This is because neutrinos that are coming from the Sun come from a quite horizontal region $[90°, 113°]$ in the point of view of the IceCube detector. So the reconstructed tracks will also be quite horizontal, and by the effects explained in figure 6.7, these tracks could be (reconstructed as) down-going.

We choose to only keep events with

$$-20m < z_{travel} < 140m. \tag{6.18}$$

The next cut is made on $t_{extension}$. Fig. 6.10 shows the histogram of this parameter.

The histograms 6.10 clearly show that the tail of the background events is dominated by multiple muons. This is expected, because coincident muons make that the event lasts longer. To cut away these long lasting events, an upperbound on $t_{extension}$ of 4000ns is chosen. So the next cut is

$$t_{extension} < 4000ns. \tag{6.19}$$

We see in Fig. 6.10 that there's a cut off at $6\mu s$. This is because the event readout window closes automatically at $6\mu s$. Since the detector is a cube of 1 km^3 the longest

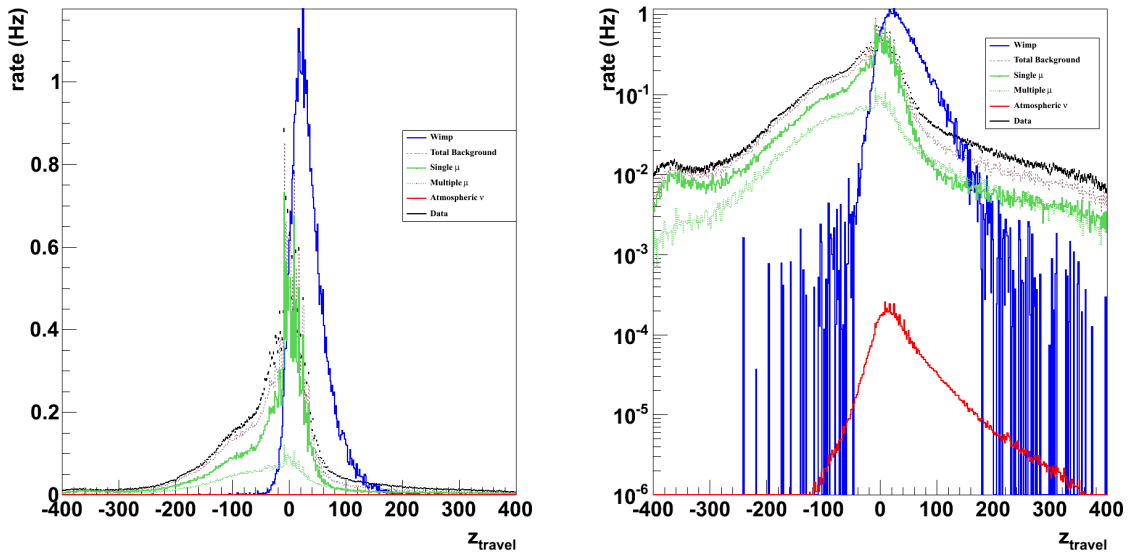

Figure 6.9: Histogram of the spread in Z coordinates from the average calculated using only the first 25%, expressed in meters. Left : linear scale, right : logarithmic scale

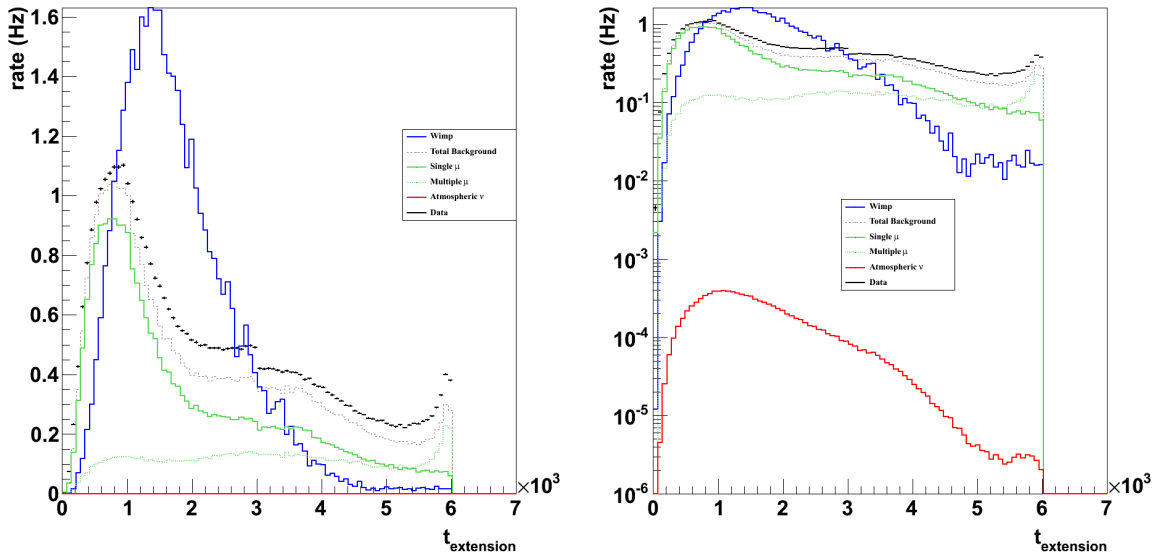

Figure 6.10: Histogram of the time between start of first and last pulse in the event, expressed in ns. Left : linear scale, right : logarithmic scale

possible track within IceCube is $\sqrt{3}$ km as shown in Fig. ??. For a particle traveling with a velocity $\sim c_{vac}$, it takes $\sim 6\mu s$ to travel across this longest possible track in the detector.

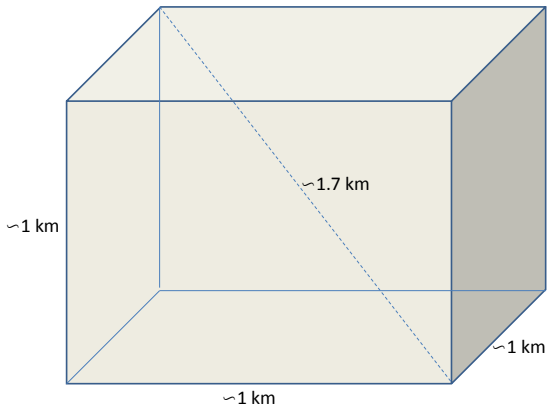

Figure 6.11: Scheme of the longest possible track within IceCube .

The next parameter we look at is $\Delta_{LLh,linefit}$. Fig. 6.12 shows the histogram of this parameter

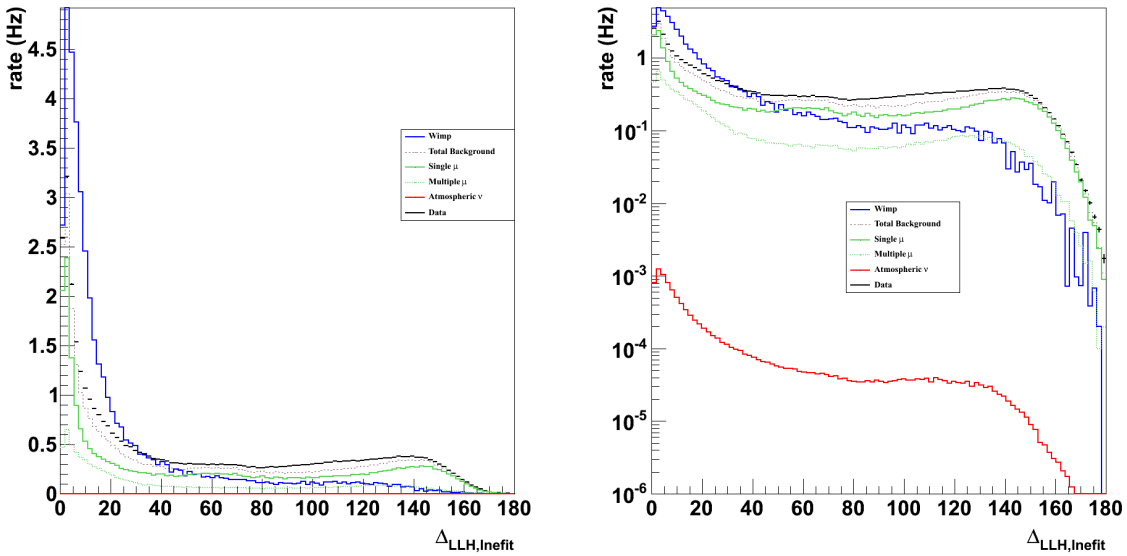

Figure 6.12: Angle between the line-fit and the LLH_{Fit} expressed in °. Left : linear scale, right : logarithmic scale

The physical meaning of this parameter is that events where this angle is big, are poorly reconstructed. So we only select events with a "small" value of $\Delta_{LLh,linefit}$, this way we reject the poorly reconstructed tracks. In order to keep enough signal and still throw away some background, we choose to keep only events with

$$\Delta_{LLh,linefit} < 40°. \tag{6.20}$$

The last parameter in the level 3 filter is $D_{dom,reco}$, which is shown in Fig. 6.13.

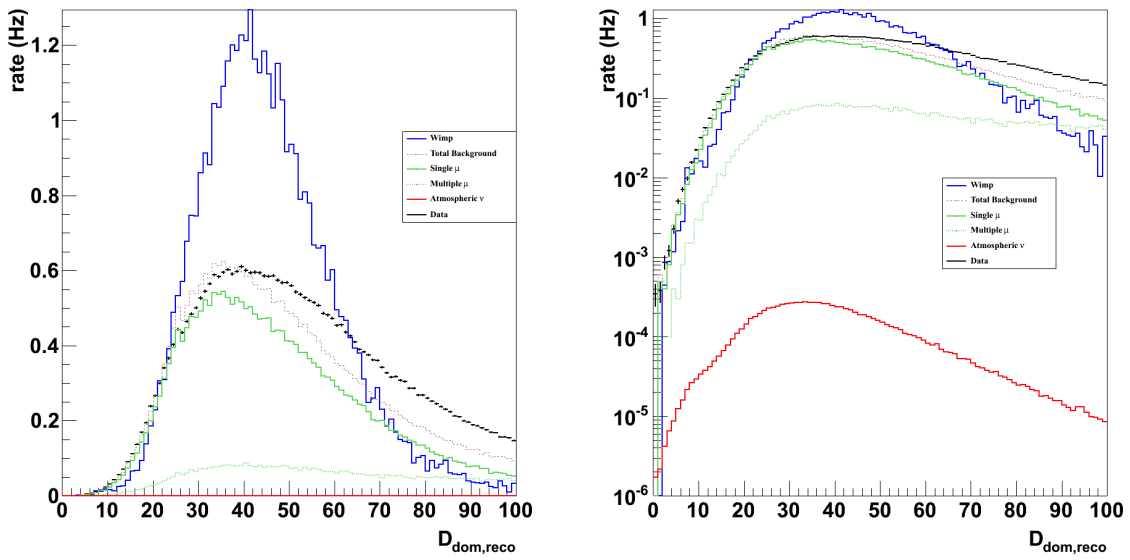

Figure 6.13: Average distance between DOMs and track, expressed in meters. Left : linear scale, right : logarithmic scale

For low energy events (e.g. our signal), this distance is small, the higher the energy, the bigger this distance gets. So we choose to only keep events with a small value for $D_{dom,reco}$, in order to select our signal. We decided to cut at

$$D_{dom,reco} < 70m \tag{6.21}$$

Table 6.4 gives an overview of the effect of the previous cuts.

	# WIMP (%)	Simulated Background (Hz)				Data (Hz)
		ν_{atm}	μ_{single}	$\mu_{multiple}$	total	
level 3 precuts	100	0.0103	25.98	9.625	35.62	41.86
$85° < \theta_{LLH} < 130°$	91.47	0.0062	5.355	3.330	8.691	10.18
$-20m < z_{travel} < 140m$	89.32	0.0058	4.203	1.649	5.856	6.621
$t_{extension} < 4000ns$	87.76	0.0057	4.116	1.284	5.406	6.058
$\Delta_{LLh,linefit} < 40°$	77.89	0.0047	1.584	0.881	2.470	3.265
$D_{dom,reco} < 70m$	76.90	0.0045	1.242	0.361	1.607	2.027

Table 6.4: Percentages of the amount of simulated signal (w.r.t. level 3 precuts) and rates of the simulated background and data events before and after the level 3 cut(s)

The effective area after the L3 cuts is shown in Fig. 6.14.

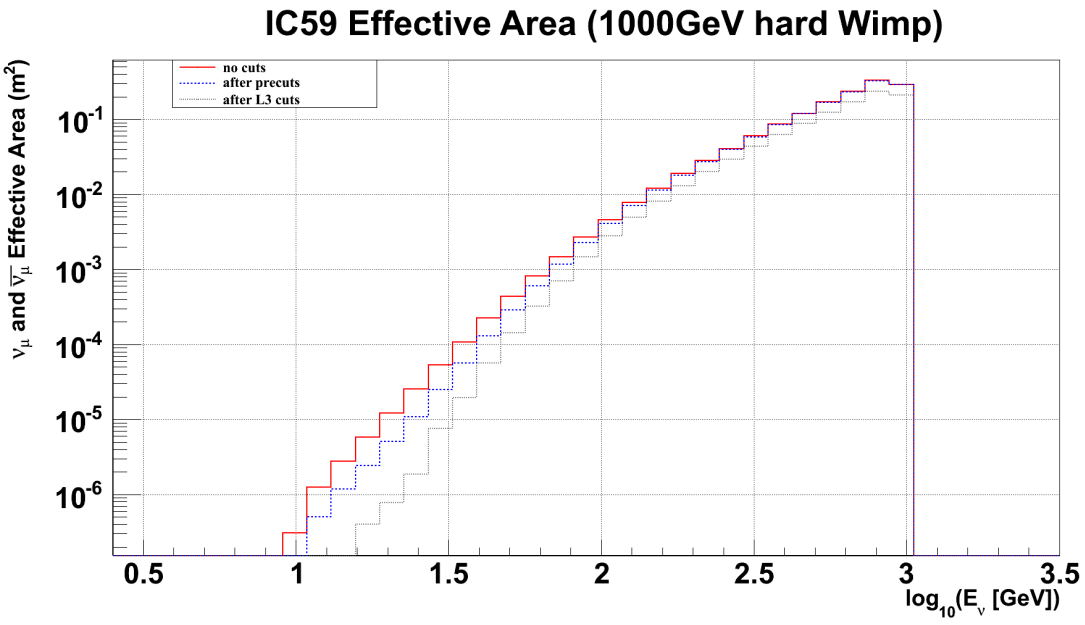

Figure 6.14: Neutrino effective area of the IC59 detector in function of $\log_{10} E_{\nu_\mu}$. This effective area is shown before level 3 filtering (solid red line), for events that passed the level 3 precuts (blue dotted line) and events that passed the level 3 cuts (gray small dots).

The effective volume V_{eff} of the detector, when looking at 1000GeV hard annihilating WIMPs, after the L3 cuts is

$$V_{eff} = 0.204 km^3. \tag{6.22}$$

As was already mentioned, these cuts are optimised for WIMPs with a mass of 1000GeV decaying via the hard channel. Table 6.5 shows the effect of the cuts on the amount of signal events and the average generated volume for different masses and annihilation channels.

WIMP Mass (GeV)	WIMP decay channel	# WIMP (%) after the L3 cuts	$\langle V_{gen} \rangle (km^3)$
50	soft	18.48	3.878
	hard	33.19	4.086
100	soft	33.53	4.109
	hard	55.72	4.361
250	soft	57.011	4.398
	hard	73.84	4.885
500	soft	65.43	4.651
	hard	76.93	5.514
1000	soft	70.23	4.816
	hard	76.90	5.771
3000	soft	72.79	5.019
	hard	77.04	5.762
5000	soft	73.35	5.072
	hard	77.08	5.710

Table 6.5: Percentages of the amount of simulated signal and the average generated volume ($\langle V_{gen} \rangle$) after the level 3 cuts (w.r.t. level 3 precuts), for different WIMP models.

6.5 Level 4 : Boosted Decision Tree.

From table 6.4,we see that $\sim 95\%$ of muonic backgound has been filtered out, while $\sim 73\%$ of the data is kept (w.r.t level 3 precuts). So we already did a good job in selecting good events. However, the experimental data will still be contaminated by e.g. misreconstructed atmospheric muons, which make it hard to extract any neutralino signal from the Sun. A more sophisticated filter is needed now, to remove most of the remaining background without removing too much signal. We will build this filter by using the Boosted Decision Tree (BDT) method, provided by the TMVA root[48] package. A very good description of this method can be found in [49].

Decision Trees (DT) are well known classifiers that allow a straightforward interpretation as they can be visualized by a simple two-dimensional tree structure (see Fig. 6.15). The DTs can be used to look for the ideal cuts to make on a set of parameters in order to separate the signal from the background.

The first step is to build or *grow* a tree. This growing is described as follows in [49]

> "The training starts with the root node, where an initial splitting criterion for the full training sample is determined. The split results in two subsets of training events that each go through the same algorithm of determining the next splitting iteration. This procedure is repeated until the whole tree is built. At each node, the split is determined by finding the variable and corresponding cut value that provides the best separation between signal and background. The node splitting stops once it has reached the minimum number of events which is specified in the BDT configuration (see Table 6.6). The leaf nodes are classified as signal or background according to the class the majority of events belongs to."

The node splitting can be done in different ways. The quality of the separation between signal and background can be checked by looking at the purity p.

$$p = \frac{\sum_s w_s}{\sum_s w_s + \sum_b w_b} \tag{6.23}$$

where $w_{s(b)}$ is the signal (background) event weight and the sums are taken over the signal (s) and background (b) events.

- $p = 0.5$ when the samples are purely mixed,
- $p > 0.5$ for "signal-leaves",
- $p < 0.5$ for "background-leaves".

The cuts should be chosen in such a way that they separate background and signal as good as possible. This means that they should create *pure* leaves :

- $p \to 1$ for "signal-leaves",
- $p \to 0$ for "background-leaves".

In order to get this good result, an impurity measure has to be defined. We chose to use the *Gini Index* (see Table 6.6)

$$\text{Gini Index} \equiv \Big(\sum_{i=1}^{n} w_i\Big) p(1-p), \tag{6.24}$$

where n is the number of events on that branch.

During the training, the cut parameters and - values are chosen in such a way that the increase in the separation index between the parent node and the sum of the indices of the two daughter nodes, weighted by their relative fraction of events is optimized. So the following expression is maximized

$$\text{Gini Index}_{\text{parent}} - (\text{Gini Index}_{\text{daughter, left}} + \text{Gini Index}_{\text{daughter,right}}) \tag{6.25}$$

The cut values are optimized by scanning over the variable range with a granularity that is set via the option nCuts. We decided to set nCut = 40 (see Table 6.6).

In principle, we could reach a perfect discrimination between signal and background, by continuing the splitting until we get pure background and pure signal leaves. This would however lead to strongly overtrained decision trees. The decision trees must thus be *pruned* in order to avoid this overtraining.

> "Pruning is the process of cutting back a tree from the bottom up after it has been built to its maximum size. Its purpose is to remove statistically insignifcant nodes and thus reduce the overtraining of the tree." [49]

We decided to use the *CostComplexity pruning* (see Table 6.6) as pruning method. To see what's the gain in adding extra nodes, a cost estimate $\mathscr{R}$ is defined as

$$\mathscr{R} = 1 - \max(p, 1-p) \tag{6.26}$$

for each node. The cost complexity for this node is now defined as

$$\rho = \frac{\mathscr{R}(\text{node}) - \mathscr{R}(\text{subtree below that node})}{\#\text{nodes(subtree below that node)} - 1}. \tag{6.27}$$

By recursively pruning away the nodes with the smallest value of ρ, the problem of overtraining is avoided. This pruning continues as long as $\rho <$ **PruneStrength** (see Table 6.6).

Then there's another problem about decision trees, namely that they are very instable with respect to statistical fluctuations in the training sample from which the tree is derived. This problem can however be overcome by constructing a forest of decision trees and classifying an event on a majority vote of the classifications done by each tree in the forest.

> "The trees are derived from the same training ensemble by reweighing events, and are finally combined into a single classifier which is given by a (weighted) average of the individual decision trees. Boosting stabilizes the response of the decision trees with respect to fluctuations in the training sample and is able to considerably enhance the performance w.r.t. a single tree." [49]

For the boosting method, we decided to use **AdaBoost**. (see Table 6.6).

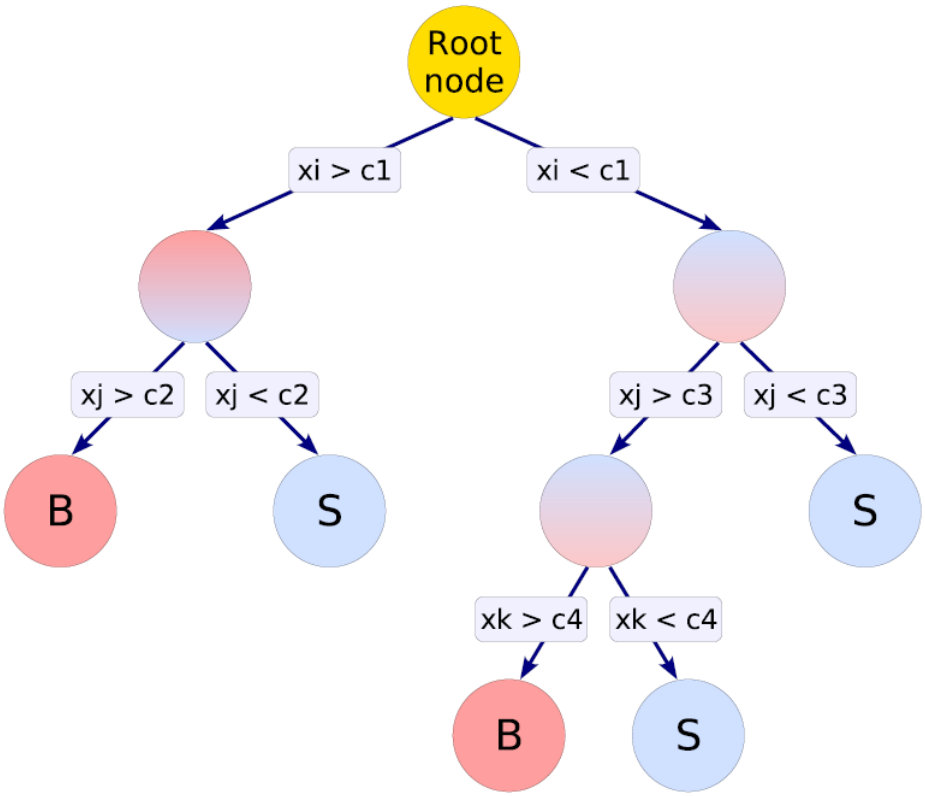

Figure 6.15: Scheme of a decision tree.

For the BDT training we used a signal (WIMP) file which contained $\sim 33 \times 10^3$ at level 3 and a background file that contained $\sim 90 \times 10^3$ events. Since we are optimizing all the cuts for 1000GeV hard neutralinos, we do this training on a 1000GeV hard file. As background file we used off source data : a subrun taken on the 10th of October in 2009. The advantage of taking data instead of MC background is that the off source

Option	Value	Description
NTrees	400	# trees in the forest
BoostType	AdaBoost	Boosting type for the trees in the forest
SeparationType	GiniIndex	Separation criterion for node splitting
nCuts	40	Number of steps during node cut optimisation
PruneMethod	CostComplexity	Method used for pruning (removal) of statistically insignificant branches
PruneStrength	2.5	Pruning strength
nEventsMin	10	Minimum number of events required in a leaf node

Table 6.6: Configuration options for the BDT classifier in TMVA package.

data really represents our background, without uncertainties, while the simulation isn't perfect and so it doesn't represent reality perfectly.

In a first step, we apply the BDT method on almost all the potentially good variables (45 to be exact) in the root files. After this training, we get a ranking of variables. If a variable gets a high value in this ranking, it means that this variable is often used during the training to decide how to split signal and background[2]. So we want to keep the variables that are the highest in this ranking. Another interesting output of the BDT training is the correlation matrix, that shows how the variables on which we trained are correlated. Since we don't want to train on highly correlated variables, we decide to only keep variables which are correlated by less then 60%. We don't train on highly correlated variables, since their effect on the training should be similar and we want to reduce the number of parameters in the training, to have a better understanding about what's happening. So if 2 variables are correlated by more then 60%, we only keep the one with the highest ranking.

After doing this, we are left with only 13 variables, which (in decreasing ranking) are :

1. $\underline{s8rlogl}$: LLH_{Fit} reduced log likelihood. This parameter is given by the ratio of the minimum of the log likelihood fit, divided by the number of hits in the track. The smaller this value, the higher the probability that the minimum is really the global minimum, so the track truely represents the best LLH track.

2. $\underline{s8NDirA}$: Number of hits in time residual (-15, +15) ns according to the LLH_{Fit}. The time residual t_{res} is the difference between the measured time and the expected time of the hit $t_{res} \equiv t_{\text{measured}} - t_{\text{expected}}$. The expected time of a hit represents the time it takes for a Cherenkov photon to travel that same distance

[2]the ranking is derived by counting how often a variable is used to split decision tree nodes and by weighting each split occurrence by the separation gain-squared it has achieved and by the number of events in the node.

without being scattered. So if this time residual is small, we know that the measured photons haven't scattered much, so they give better information about the track. The bigger the number of hits in time residual (-15, +15) ns the better the quality of the track.

3. z_{travel} : Mean spread in z from the mean z calculated from the first quartile of hits in time. This parameter was already used in the level 3 filter, where we found that this is a good parameter to distinguish signal neutrinos and down-going muons.

4. θ_{LLH} : Reconstructed Zenith. This parameter was already used in the level 3 filter, where we found that this is a good parameter to distinguish WIMPs and down-going muons. Since our signal neutrinos really come from the direction $90° < \theta_{sun} < 113°$, the reconstructed zenith can also be used to filter away the atmospheric neutrinos that don't have the same direction as our signal neutrinos.

5. *s8SallCharge*: *Smoothness* of the reconstructed LLH_{Fit} track, using the total charge in the event as a weight. The smoothness of a track represents how good or *smooth* the hits are distributed over the length of the track. In this case, the smoothness of the charge of the hits is given. If the hits are smoothly distributed over the track, the track has a smoothness ~ 0. If there are more hits at the beginning (end) of the tracks of the track, the smoothness of the track goes to $\sim$ -1 (+1).

6. z_{av} : Mean z coordinate of channels (COGZ unweighted by hits). This is again a good parameter to filter away the down-going muons, since for these events z_{av} will be mostly positive (top of the detector), while for our (horizontal/up-going) signal it will be quiet evenly distributed.

7. $D_{dom,reco}$: Average distance between the reconstructed track and the fired DOMs. This parameter was already used in the level 3 filter, where we found that this is a good parameter to select low energy events (like our signal).

8. ρ_{av} : Mean distance in the xy plane of channels from the origin. Even though this parameter is a combination of two other parameters used in the training : $\rho_{av} = \sqrt{COGX^2 + COGY^2}$, the correlation of this parameter with COGX and COGY is smaller than 60% (see Figs. 6.17 and 6.18). The reason that this is a good parameter is good to distinguish signal and background is the same as that for COGX and COGY, which will be explained lower (see (12) and (13)).

9. $\Delta_{LLh,linefit}$: Angle between linefit reconstructed track and LLH reconstructed track. This parameter was already used in the level 3 filter, were we found that this is a selection on good reconstructed tracks, in which case this angle is small, because in this case there was a clear track, that was reconstructed in the same way by both linefit and the LLH fit.

10. t_{acc} : Time to accumulate 75% of the total charge in the event. If we look at the level 3 distribution of this variable, we see that it's small for atmospheric muons, and it gets bigger for signal neutrinos. We think that this is due to the fact that the muons that are left at level 3 have a higher energy than the signal neutrinos at this level. The high energy event will have stochastic energy losses, while low energy particles are minimum ionizing. Assuming that this indeed correct, we expect that the high energy atmospheric muons loose most of their energy in the beginning of the track, while this energy loss is more evenly distributed for the low energy signal events.

11. $\underline{s8SDirC}$: Smoothness of the reconstructed LLH_{Fit} track, in the time residual (-15, +75) ns. Comparable explanation

12. $COGY$: Center of gravity in Y. To explain why this is a good parameter we need to look at the shape of the detector, shown in Fig. 6.16. We see that the IC59 detector is asymmetric and a whole part is still "missing". Most of the missing strings are in the lower left part (negative values for the x and y coordinates). High energy tracks could have their COG in that missing part of the detector and still be detected, while this is less probable for low energy tracks, that travel less far. That's why we expect that most of the low energy tracks, like our signal neutrinos, have the COG in within the IC59 detector, while the higher energy background tracks can have the COG outside of the IC59 detector (e.g. in the "missing" lower left part). This is why we expect the COGY (and COGX) distribution to have a peak at positive values in the case of our (low energy) signal, while the (higher energy) background distribution is expected to be more evenly distributed.

13. $COGX$: Center of gravity in X. Same explanation as for $COGY$.

There are two kinds of parameters :

1. parameters that are used to select good quality tracks. These parameters are underlined.

2. parameters that are used to distinguish signal from background.

The histograms of these variables, after the level 3 cuts, are shown in Appendix A.

The signal and background correlation matrices for these parameters are shown in resp. Fig. 6.17 and Fig. 6.18.

The next step was to apply the output of the training on signal and background (data and MC). For the signal we applied the output on the same 1000GeV hard WIMP file than the one we trained on.

For background, we used the whole month of data taken in October 2009 and we also applied the output on simulated background (atmospheric muons and neutrinos). This application of the training results provided to each event a certain weight between

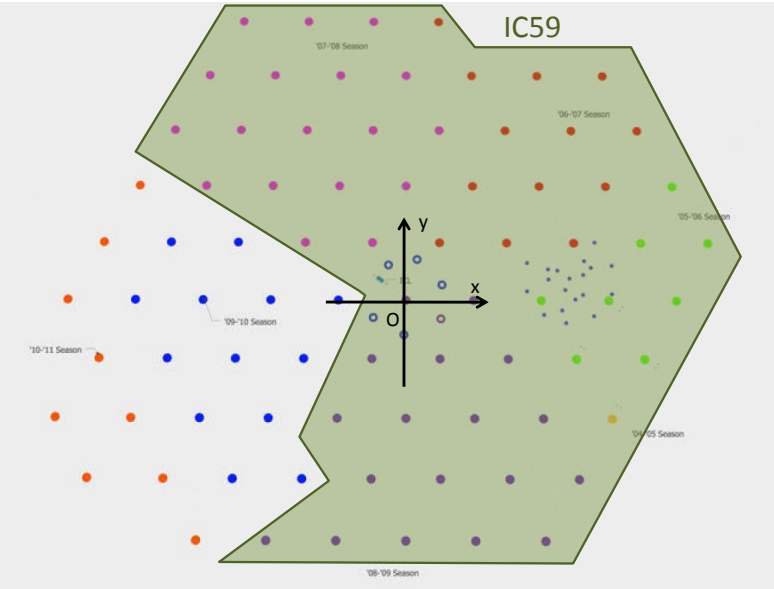

Figure 6.16: IceCube as seen from above. The green region shows the asymmetric IC59 detector and the axes show how the (x,y) coordinate system is defined.

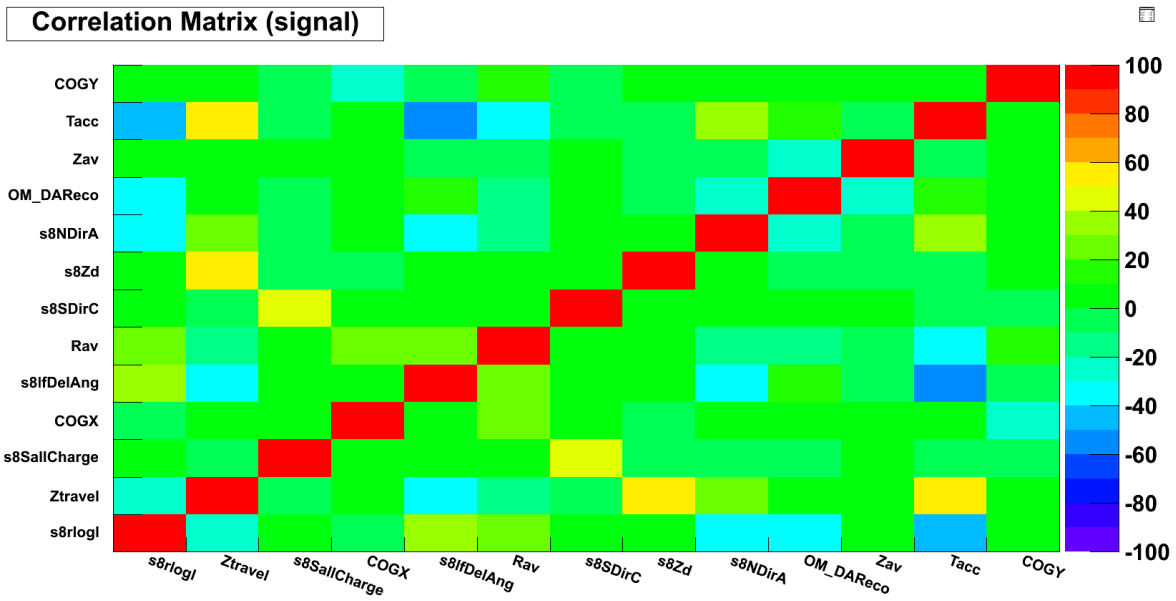

Figure 6.17: Correlation matrix of the 13 BDT training parameters for the signal sample.

$[-1, 1]$, where 1 (-1) stands for a pure signal (background) event. These BDT outputs are shown in Fig. 6.19, where the same line colors and styles were used as in the previous section :

- The experimental data (with errors) : black dots with error bars. This is data of

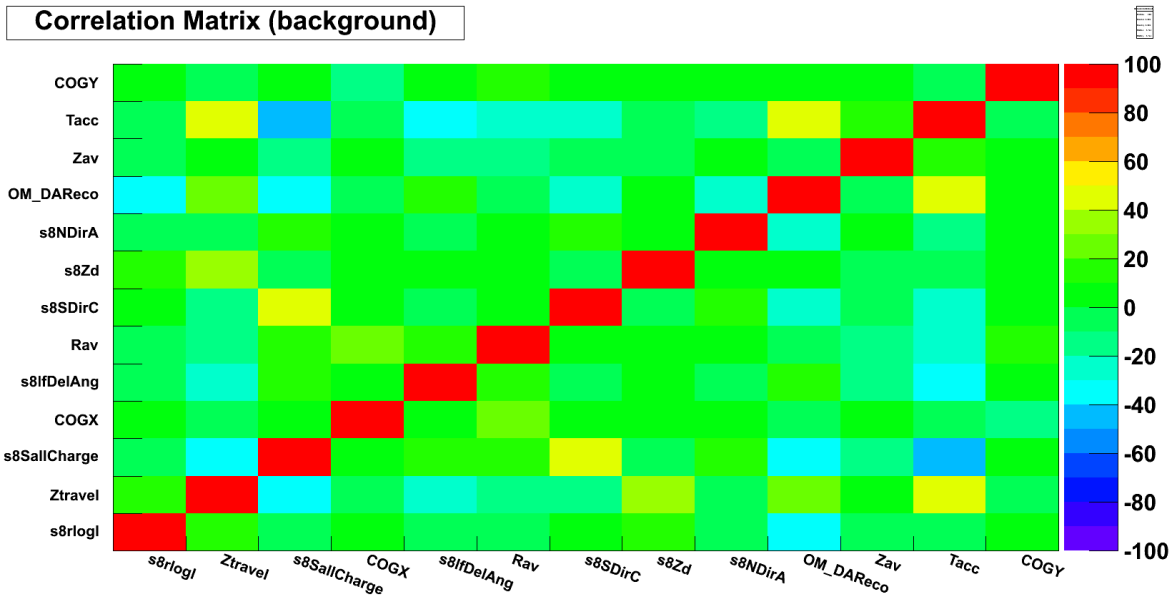

Figure 6.18: Correlation matrix of the 13 BDT training parameters for the background sample.

the whole month of October 2009.

- The simulated background (total) : brown dashed line, this simulated background consists of
  - atmospheric neutrinos (ν_{atm}) : red line
  - single muons (μ_{single}) : green line
  - multiple muons ($\mu_{multiple}$) : dotted green line
- The simulated WIMP signal : blue line. The signal rate is scaled to the experimental data rate. As was already said, this is simulation for 1000 GeV hard annihilating WIMPs.

Fig. 6.19 clearly shows that we can separate the muonic background from our signal, but we can't cut away a lot of atmospheric neutrinos, since they look like signal. Now we have to decide where we cut on the BDT output value. To do this, we looked at three different significance calculations :

1. the significance calculated as in eq. 6.15

$$\Lambda_{cut,1} = \varepsilon_{sig} \times (1 - \varepsilon_{bkg}) . \tag{6.28}$$

Fig. 6.20 shows how the significance calculated in this way changes for different cut values. We see that this is not a good way to look for the best cut value, since it doesn't return a clear maximum.

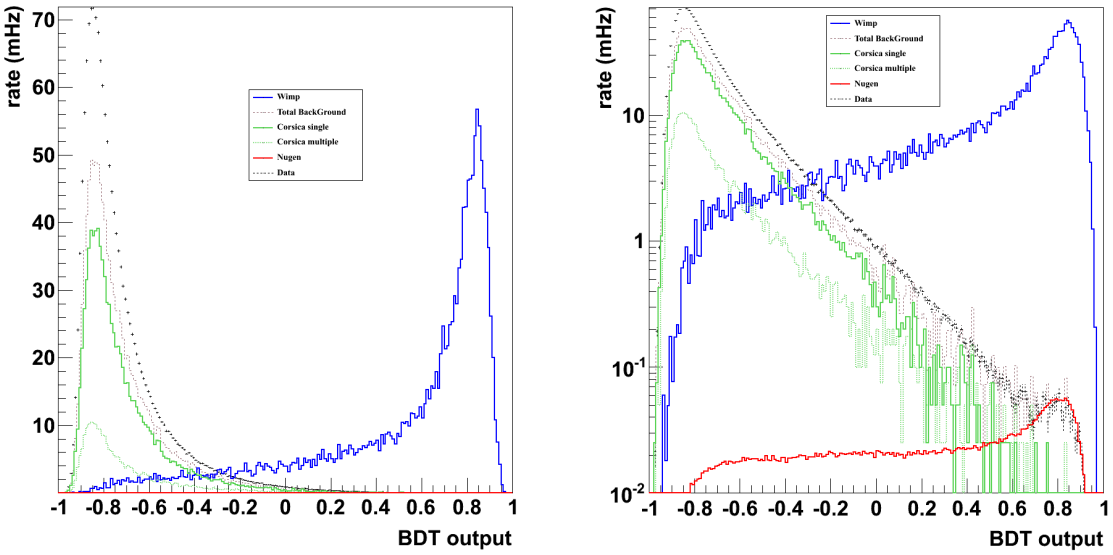

Figure 6.19: BDT output for signal and background samples after training on 13 parameters.

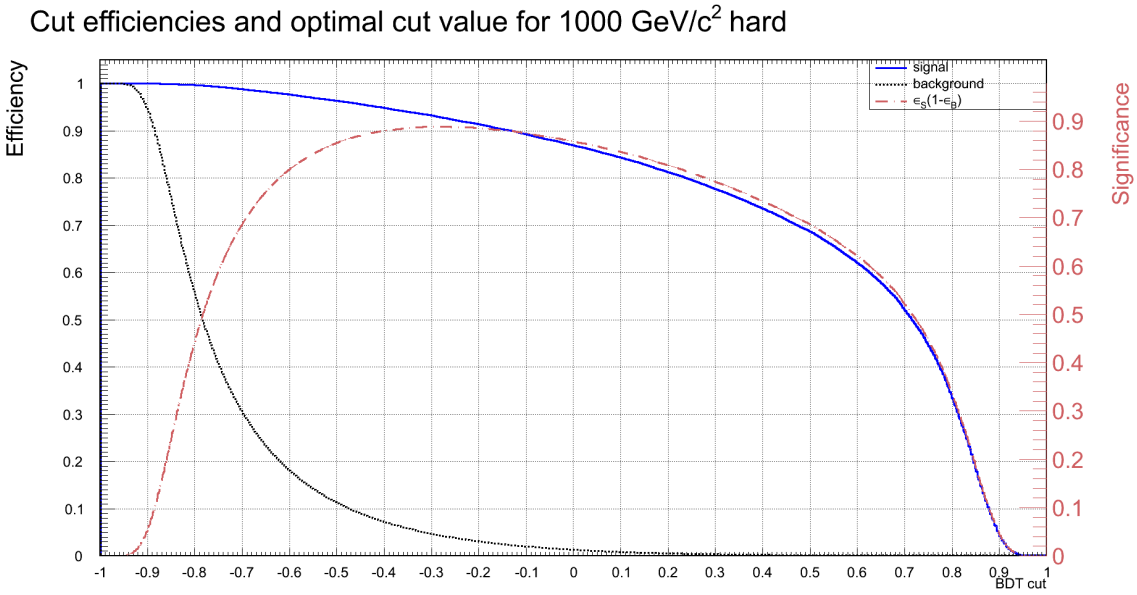

Figure 6.20: Distribution of the significance according to eq. 6.28. The black solid line shows the background efficiency ($\varepsilon_{bkg} = n_{\text{bkg aftercut}}/n_{\text{bkg before cut}}$), the blue line shows the signal efficiency ($\varepsilon_{sig} = n_{\text{sig after cut}}/n_{\text{sig before cut}}$) and the maroon line shows the distribution of the significance of the cut according to eq. 6.28

2. the "standard" way to calculate the significance of the cut is

$$\Lambda_{cut,2} = \frac{\varepsilon_{sig}}{\sqrt{\varepsilon_{bkg}}} \tag{6.29}$$

Fig. 6.21 shows how $\Lambda_{cut,2}$ depends on the BDT cut. Here we see a clear maximum at $BDToutput = 0.74$.

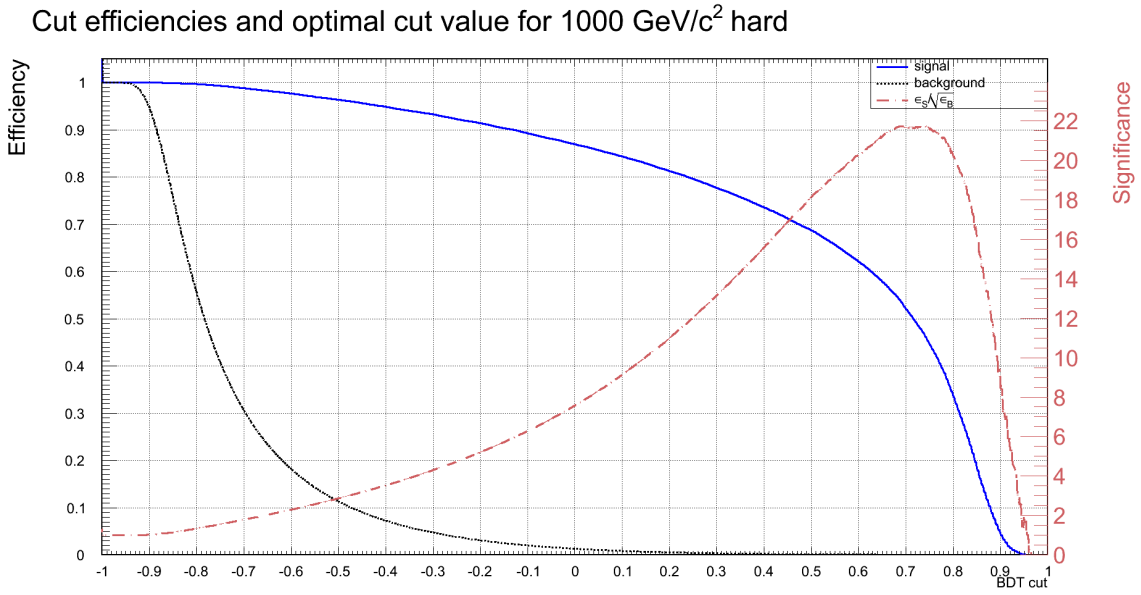

Figure 6.21: Distribution of the significance according to eq. 6.29. The black solid line shows the background efficiency, the blue line shows the signal efficiency and the maroon line shows the distribution of the significance of the cut according to eq. 6.29

3. the *Model Discovery Potential* (MDP) as defined in [50]

$$\Lambda_{cut,3} = \frac{\varepsilon_{sig}}{\frac{a^2}{8} + \frac{9b^2}{13} + a\sqrt{n_B} + \frac{b}{2}\sqrt{b^2 + 4a\sqrt{n_B} + 4n_B}} \tag{6.30}$$

This figure-of-merit was meant to minimize the strength of the signal flux needed for a discovery with 5σ significance (a=5) at 90% confidence level (b=1.28). Fig. 6.22 shows how $\Lambda_{cut,3}$ depends on the BDT cut. There is a maximum at $BDToutput = 0.69$, comparable with 6.21.

Figs. 6.21 and 6.22 show that $BDToutput = 0.69$ is a good value to cut at. By cutting at this value, we reduce the amount of level 3 background by a factor $\sim 10^4$ and we keep $\sim 40\%$ of the signal events that passed the level 3 filter. So our level 4 cut is

$$BDT\ output\ > 0.69. \tag{6.31}$$

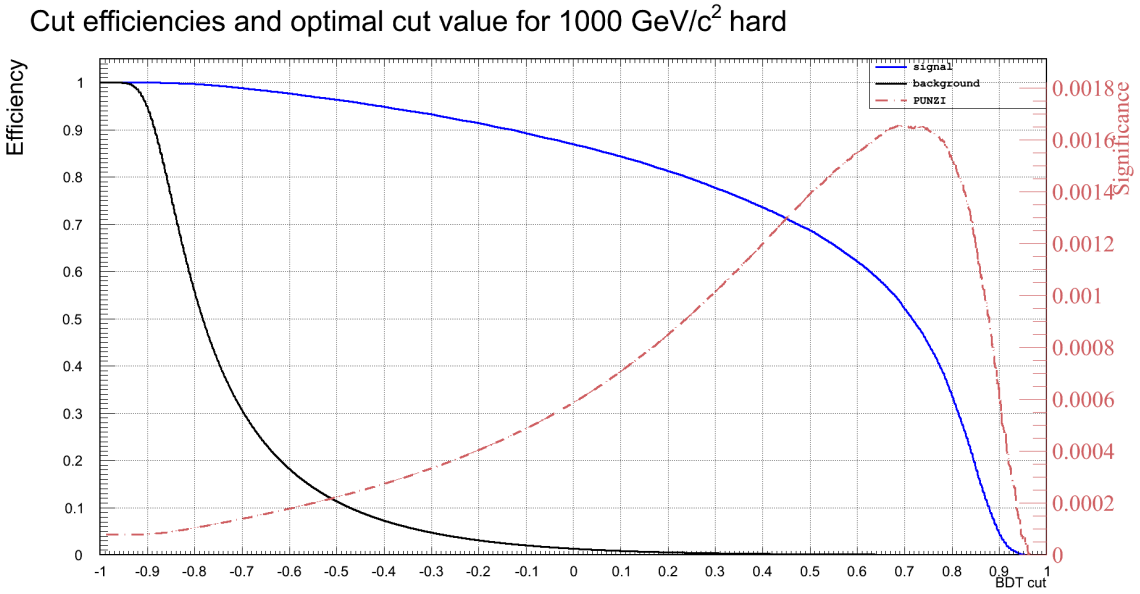

Figure 6.22: Distribution of the significance according to eq. 6.30. The black solid line shows the background efficiency, the blue line shows the signal efficiency and the maroon line shows the distribution of the significance of the cut according to eq. 6.30

We can see from Fig. 6.19 that this cut removes most of the atmospheric muons, but a lot of atmospheric neutrinos survive the cut. In chapter 5 we talked about the difference between simulation and data. At filter level 2 there was a clear difference between data (dominated by atmospheric muons) and simulated muons. We mentioned that this difference wasn't important for our analysis, since we use off source data to develop the filters (levels 3 and 4). The atmospheric neutrino simulation (NuGen) however had to be good, since our signal events are simulated neutrinos.
We now succeeded to remove almost all these muons from our dataset, and we're left with mainly atmospheric neutrinos as can be seen in Figs. 6.23 and 6.24, which respectively show the θ_{LLH} and z_{travel} distributions before and after the BDT cut (6.31)[3]. Figs. 6.25 and 6.26, and Table 6.7 also show that the simulated background event rate matches the experimental event rate very well. From this, we can conclude that the simulation of the atmospheric neutrinos is good, so we can trust the effective volume calculations.

Table 6.7 summarizes the effect of the different level cuts on the signal and the background (data and MC). From this table it is again clear that the BDT cut removed a lot of atmospheric muons, and that the atmospheric neutrinos are now the dominating background. The analysis in the next section will again be done with the data of a whole

[3]The distributions before and after the BDT cut of the other parameters used in the BDT training are shown in Appendix A

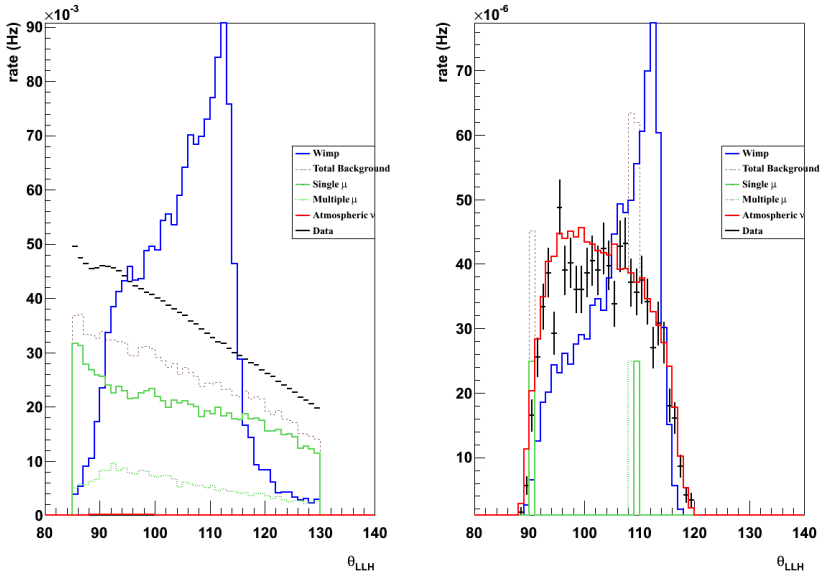

Figure 6.23: Histogram of the Reconstructed Zenith angle of the detected muon before (left) and after (right) BDT cut (6.31)

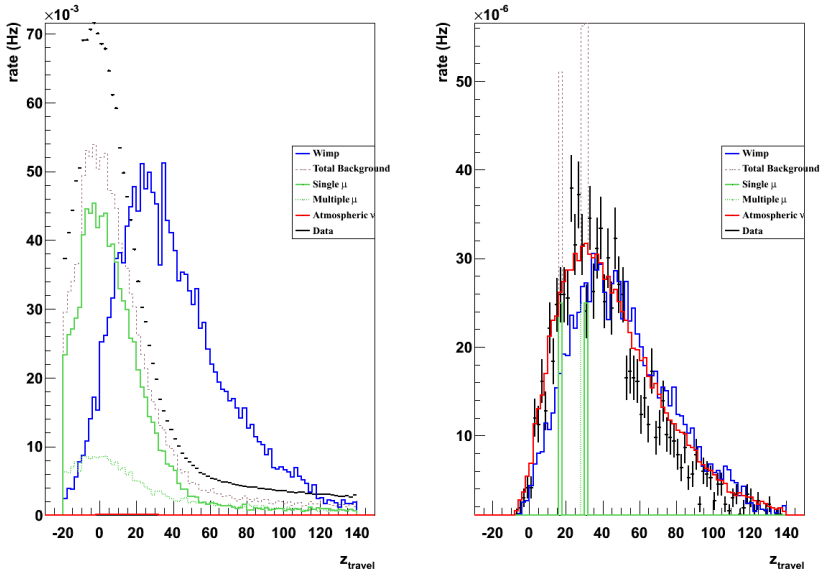

Figure 6.24: Histogram of the mean spread in z from the mean z calculated from the first quartile of hits in time, before (left) and after (right) BDT cut (6.31)

month of October 2009. The amount of events (both experimental and simulated) that are left after all the cuts, are shown in Table 6.8.

The effective area after the BDT cut is shown in Fig. 6.27.

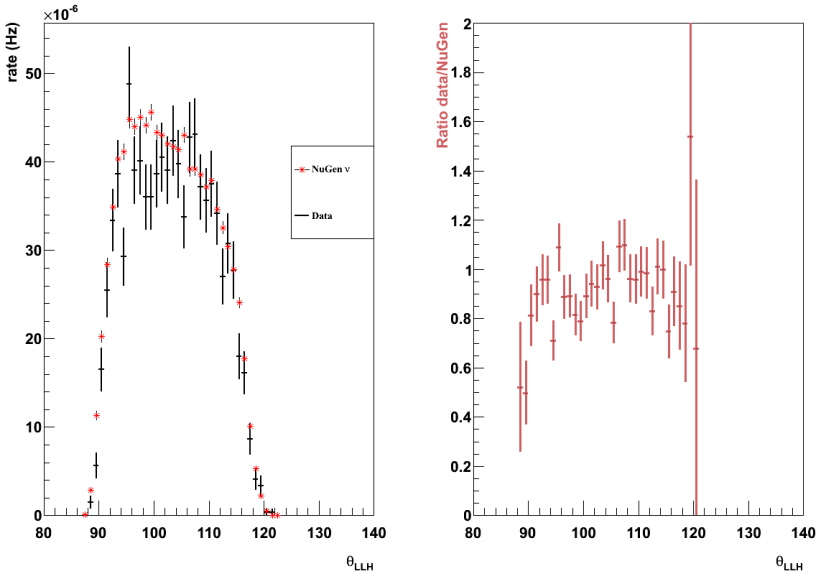

Figure 6.25: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : θ_{LLH} of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for θ_{LLH}.

	# WIMP (%)	Simulated Background (Hz)				Data (Hz)
		ν_{atm}	μ_{single}	$\mu_{multiple}$	total	
L3 precuts	100	0.0103	25.98	9.625	35.62	41.86
L3 cuts	76.90	0.0045	1.242	0.361	1.607	2.027
BDT cut	39.12	0.0010	4.98×10^{-5}	2.49×10^{-5}	0.0011	0.00095

Table 6.7: Percentages of the amount of simulated signal (w.r.t. level 3 precuts) and rates of the simulated background and data events before and after the cuts.

# Simulated Background				# Data
ν_{atm}	μ_{single}	$\mu_{multiple}$	total	
2759	133	66	2958	2536

Table 6.8: Amount of simulated background and data events left after the cuts.

The effective volume V_{eff} of the detector, when looking at 1000GeV hard annihilating WIMPs, after the BDT cut is

$$V_{eff} = 0.104 km^3. \tag{6.32}$$

The BDT cut was optimised for WIMPs with a mass of 1000GeV decaying via the hard channel. Table 6.9 shows the effect of the cuts on the amount of signal events and

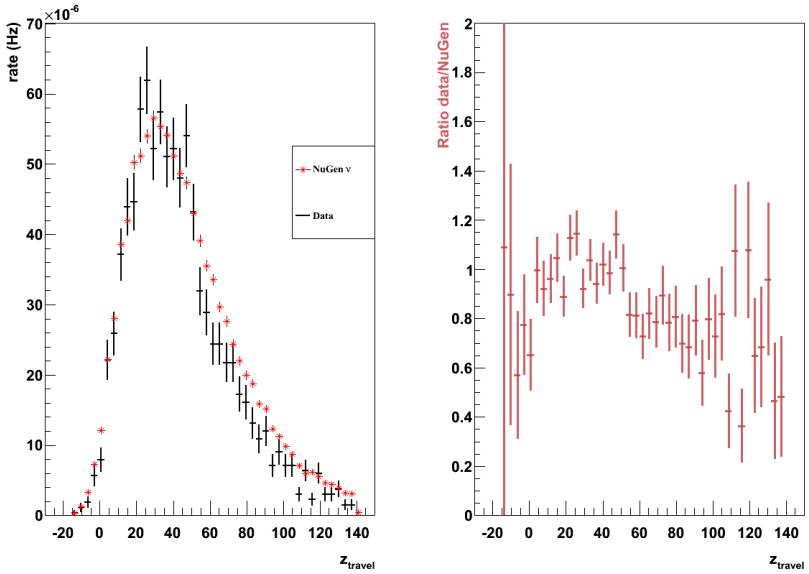

Figure 6.26: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : z_{travel} of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for z_{travel}.

the average generated volume for different masses and annihilation channels.

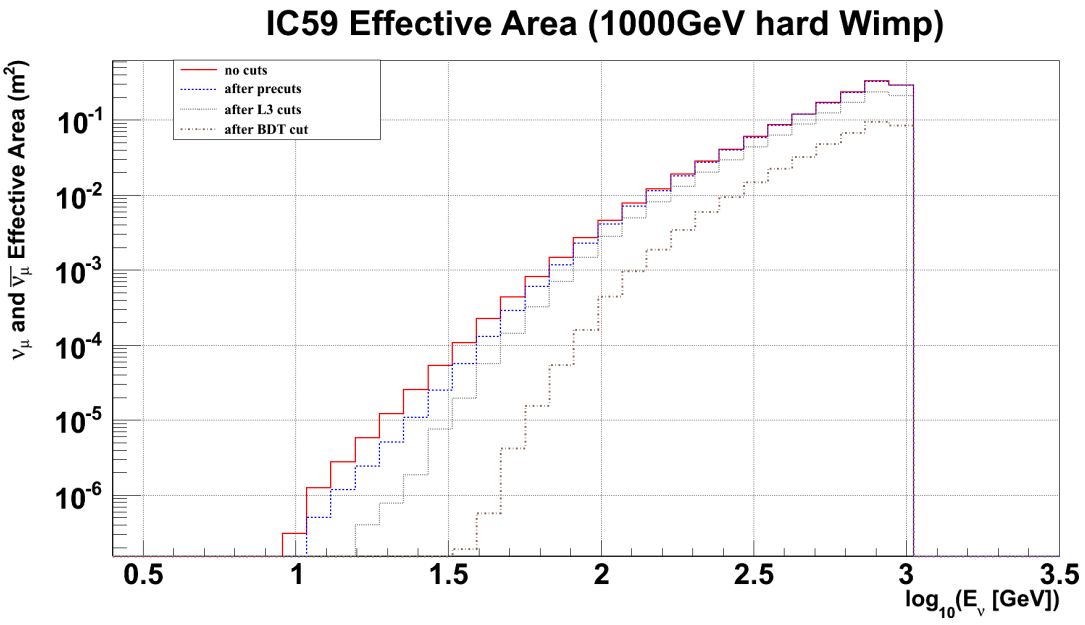

Figure 6.27: Neutrino effective area of the IC59 detector in function of $\log_{10} E_{\nu_\mu}$. This effective area is shown before level 3 filtering (solid red line), for events that passed the precuts (blue dotted line), events that passed the level 3 cuts (gray small dots) and events that passed the BDT cut (brown dashed-dotted line).

WIMP Mass (GeV)	WIMP decay channel	# WIMP (%) after the BDT cut	$\langle V_{gen} \rangle (km^3)$
50	soft	0.00	0.00
	hard	0.99	4.368
100	soft	2.51	4.463
	hard	6.76	4.697
250	soft	8.17	4.885
	hard	27.91	5.459
500	soft	16.78	5.142
	hard	35.52	6.056
1000	soft	23.35	5.337
	hard	39.12	6.393
3000	soft	28.89	5.610
	hard	38.35	6.424
5000	soft	29.71	5.658
	hard	37.47	6.357

Table 6.9: Percentages of the amount of simulated signal and the average generated volume ($\langle V_{gen} \rangle$) after the BDT cut (w.r.t. level 3 precuts), for different WIMP models.

Chapter 7

Sensitivity of The Detector and the Analysis

In the previous chapter we developed a high level filter that removed almost all the atmospheric muons. These atmospheric muons dominated the data samples at level 2, and they are pure background. We reduced the atmospheric muon rate from $\sim$ 40Hz at level 2, to $\sim 10^{-4} - 10^{-5}$Hz after the high level filtering. The experimental data sample that we are left with after the filtering, is dominated by atmospheric neutrinos ($\sim 10^{-3}$Hz).

In this chapter we will develop a method to analyze the 2536 data events that passed all filters. These events were recorded during the whole month of October 2009, so they are off source events. We will use this month of off source data to estimate the sensitivity of the IC59 detector for this analysis. The sensitivity is defined as the 90% confidence level upper limit on the number of signal events in a background only scenario. One expects that the number of signal events in a background only scenario is zero, but due to statistical fluctuations this number can be different from zero. In order to be sure that one measures real signal events and not just statistical fluctuations, the amount of measured signal events should exceed the sensitivity of the detector.

The aim of this thesis was to calculate this sensitivity. We didn't unblind our data yet, so we had to work *blind*. In order to do this we scrambled the azimuth of the Sun, such that the real position of the Sun is unknown during our analysis. This *blindness* policy is introduced to prevent the introduction of a bias by the researcher. During an analysis it's quite possible that small errors are made and certain things are overlooked. If a researcher is allowed to look at the end result of his analysis and has a certain outcome in mind he can go back to the analysis to remove errors and recheck for possible oversights. But if he does this until his results check out with what he expects and stops removing errors then he will have introduced a bias because its very well possible that he did not remove all the errors and oversights in his analysis. Before unblinding the data one has to prove that the atmospheric background is well

understood and under control.

The next step will be an unblinding request, but before we do this we should calculate the systematic errors which was beyond the scope of this thesis.

In the last part of this chapter, we will use the estimated sensitivity to calculate the resulting upper limits on some physical quantities.

7.1 Hypothesis Testing

What we want to do now, is look if there are signal events in our final data set. We already pointed out that this final data set is dominated by atmospheric neutrinos (see Figs. 6.23 and 6.24 and Table 6.7). These atmospheric neutrinos come from all over the sky [1], while our signal comes from the Sun's direction. We will use the *space angle* ψ between the source direction $(\theta_\odot, \phi_\odot)$ and that of the reconstructed track $(\theta_{\mu,reco}, \phi_{\mu,reco})$ to evaluate the number of signal events μ_s. The space angle is defined as (see Fig. 7.1):

$$\psi \equiv \arccos\left(\cos\theta_{\mu,reco}\cos\theta_\odot + \sin\theta_{\mu,reco}\sin\theta_\odot\cos\left(\phi_{\mu,reco} - \phi_\odot\right)\right) \tag{7.1}$$

For signal, we expect this space angle to be small. The spread in space angle of the signal depends on two effects that were already explained before (see section 6.4). First of all, there is an opening angle between the neutrino and the CC interaction produced muon. The second source of spread is that the reconstruction algorithm isn't perfect, so the reconstructed muon direction isn't exactly the same as the true muon direction. These effects are shown in Fig. 7.1.

A word about the space angle distributions

For the signal events, the space angle distribution $f_S(\psi)$ is directly calculated using the $(\theta_{\mu,reco}, \phi_{\mu,reco})$ and $(\theta_\odot, \phi_\odot)$ in the simulation. There are systematic uncertainties in the simulation, but the calculation of these uncertainties is beyond the scope of this thesis.

To determine the background space angle distribution $f_B(\psi)$, we used off source data (the whole month of October 2009) that passed the filters. There are 2 big advantages of using data instead of simulated background. The first advantage is that we have larger statistics and the second, major advantage is that the data is almost systematic uncertainty free. Because of the blindness policy, we aren't allowed to calculate ψ directly by using $(\theta_{\mu,reco}, \phi_{\mu,reco})$ and $(\theta_\odot, \phi_\odot)$, like we did for $f_S(\psi)$. To take the blindness policy into account, we simulated a fake Sun position $(\theta_\odot^{\text{fake}}, \phi_\odot^{\text{fake}})$ for each off source event. Actually we *scrambled* $\phi_\odot^{\text{true}}$ to get $\phi_\odot^{\text{fake}}$ and left $\theta_\odot^{\text{fake}} = \theta_\odot^{\text{true}}$, because the detector has a θ acceptance.

[1] These atmospheric neutrinos only have to pass the zenith(θ) cuts that were made in the high level filtering, so they should be downgoing. There is however no restriction in the azimuth(ϕ) for atmospheric neutrinos.

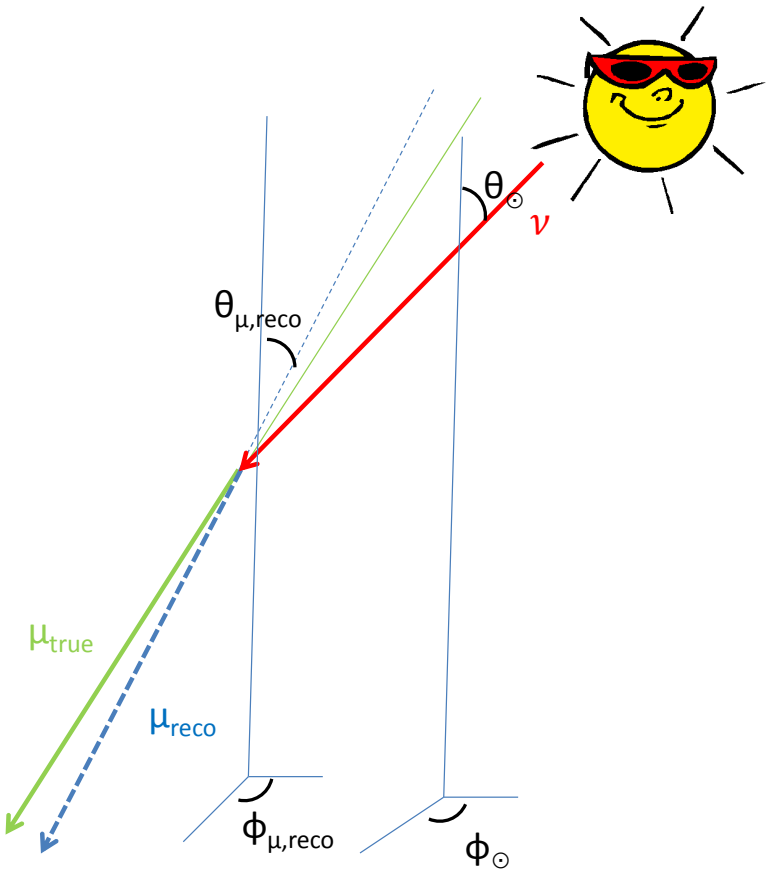

Figure 7.1: Sketch of construction of the space angle ψ between the Sun direction and the reconstructed track.

For each event, a random value for $\phi_{\odot}^{\text{fake}}$ was calculated, evenly distributed between $[0°, 360°]$. Then the corresponding space angle was calculated between the reconstructed track and the fake Sun direction. This procedure was repeated 10^4 times for each event, to get a continuous space angle distribution without intermediate empty bins. Finally, we normalized the resulting $f_S(\psi)$ and $f_B(\psi)$ distributions to one, so that they represent the p.d.f. for signal and background. Fig. 7.2 shows the output of this procedure for 1000GeV hard annihilating WIMPs.

The blue dots with error bars in Fig. 7.2 show the space angle distribution for the simulated 1000GeV hard annihilating WIMPs remaining after the BDT selection. The black line shows the data (the whole month of October 2009) that passed all the filters. We see that the signal peaks towards low values of ψ, while the data is *evenly* distributed over the sky. The fact that the data drops of at the edges is because we run out of phase space there.

The median opening angle $\tilde{\psi}$ for the 1000GeV hard annihilating WIMP sample is $\sim 1.9°$. Table 7.1 gives the $\tilde{\psi}$ for different WIMP models. This median opening value

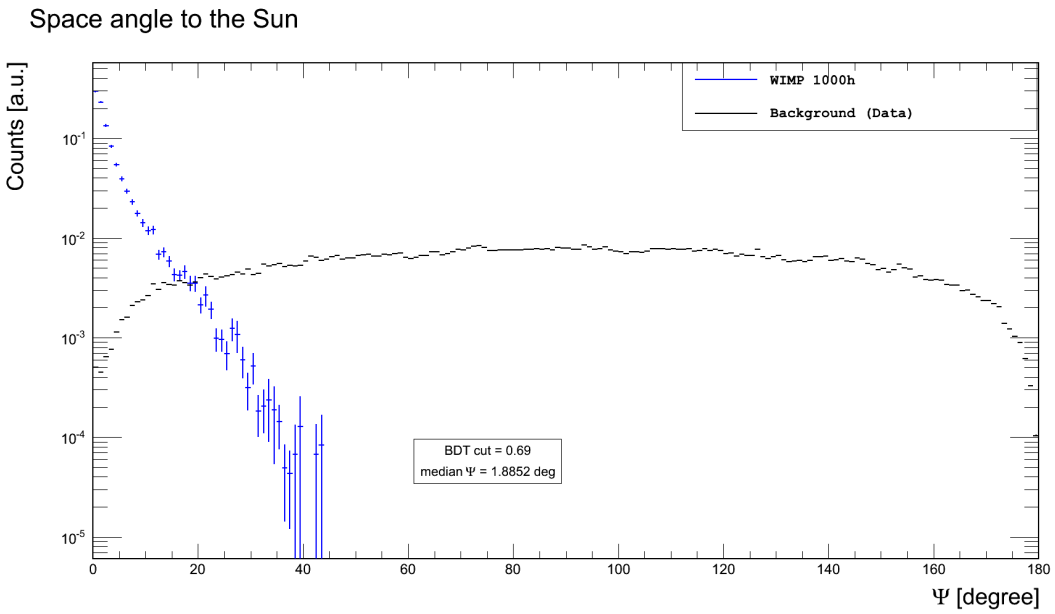

Figure 7.2: Space angle distribution for the 1000GeV hard annihilating WIMPs (blue crosses) and scrambled experimental data (black lines) that passed all the filters.

gives an estimation of the angular resolution of the IC59 detector for this analysis.

WIMP Mass (GeV)	WIMP decay channel	$\tilde{\psi}(^\circ)$
50	soft	no signal left...
	hard	6.08
100	soft	6.59
	hard	5.55
250	soft	4.47
	hard	2.80
500	soft	3.52
	hard	2.03
1000	soft	2.83
	hard	1.89
3000	soft	2.51
	hard	1.82
5000	soft	2.29
	hard	1.90

Table 7.1: Median opening angle $\tilde{\psi}$ for different WIMP models.

We could now use the *cut-and-count* approach to look for signal events in our data sample. In this approach, we would look for the optimal opening angle $\psi_{optimal}$ by using e.g. eq. (6.29) or (6.30). We would then define a cone around the Sun with opening an-

gle $\psi_{cone} = \psi_{optimal}$ and count the number of hypothetical signal events and background events in this cone. The problem with this approach is that it gives the same weight to every event inside the cone. So we can't distinguish e.g. 10 events inside this cone that are compatible with the number of background, from 10 events of which 8 are really close to the Sun's direction. This latter case should give a strong hint for the existence of signal though. It is because of this reason that we decided to use a different method in which we use the space angle distributions to do a signal content likelihood study. The method we used is based on the one that is used by A. Rizzo in his PhD thesis [47]

7.1.1 Estimation of the Sensitivity for WIMP Neutrinos

Now that we have the p.d.f. of the signal and the background (see section 7.1), we can derive the amount of signal events μ_s compatible with the observed data sample at a certain confidence level α. Up till now, we applied our cuts and calculations on 14 different WIMP models (7 masses with each two channels). This last step of the analysis will only be done for the 1000GeV hard WIMPs. This is because this step is very CPU time consuming and since we didn't do an unblinding request yet, it wouldn't be very interesting to do this last step for every model. Note that the described procedure would be the same for every model.

The probability to observe an event with space angle ψ_i, when μ_s signal events are present among a total number of observed events n_{obs}, is given by

$$f(\psi_i|\mu_s) = \frac{\mu_s}{n_{obs}} f_S(\psi_i) + \left(1 - \frac{\mu_s}{n_{obs}}\right) f_B(\psi_i), \tag{7.2}$$

where $f_S(\psi_i)$ and $f_B(\psi_i)$ are respectively the signal and the background p.d.f. . From eq. 7.2, we can now calculate the likelihood of the presence of μ_s signal events in an experiment that observed n_{obs} events with an ensemble of space angles $\{\psi_i\}$, as

$$\mathcal{L}(\mu_s) = \prod_{i=1}^{n_{obs}} f(\psi_i|\mu_s). \tag{7.3}$$

This likelihood (7.3) can be used in the likelihood-ratio test statistic, suggested by Feldmann and Cousins [51]

$$\mathcal{R}(\mu_s) = \frac{\mathcal{L}(\mu_s)}{\mathcal{L}(\hat{\mu}_s)}, \tag{7.4}$$

where $\hat{\mu}_s$ is the result of the best fit to the observed ensemble of space angles. So $\mathcal{R}(\mu_s) \leq 1.$ for each μ_s.

Feldmann and Cousins used this likelihood-ratio $\mathcal{R}(\mu_s)$ as a *rank*, since starting at high values, it ranks the experimental results in order of inclusion in the acceptance

interval. Another nice feature about the Feldmann and Cousins approach is that the parameter μ_s is bound to physical values, so $\mu_s \in [0, n_{obs}]$, while in the *standard* χ^2 frequentist approach, non-physical best fits ($\mu_s \notin [0, n_{obs}]$) could arise at the boundaries of the interval.

The acceptance intervals in this approach at the desired α CL are

$$[\mathcal{R}^{\alpha}_{\text{crit}}(\mu_s), 1]\,, \tag{7.5}$$

where $\mathcal{R}^{\alpha}_{\text{crit}}(\mu_s)$ is defined such that the interval 7.5 contains a fraction α of the $\mathcal{R}(\mu_s)$-values.

Once this $\mathcal{R}^{\alpha}_{\text{crit}}(\mu_s)$ value is known, we can calculate the confidence interval $[\mu^{\alpha}_{s,\text{low}}, \mu^{\alpha}_{s,\text{up}}]$ for a certain number of signal events μ_s as

$$[\mu^{\alpha}_{s,\text{low}}, \mu^{\alpha}_{s,\text{up}}] = \{\mu_s | \ln \mathcal{R}(\mu_s) \geq \ln \mathcal{R}^{\alpha}_{\text{crit}}(\mu_s)\}. \tag{7.6}$$

Since we didn't unblind the data yet, we will keep using the background p.d.f. described before, with scrambled $\phi_{\odot}$. This off source data set is the whole month of October 2009, of which 2536 events are left after the BDT cut. We calculate the confidence interval for μ_s (eq. (7.6)) in the background-only scenario to find what the sensitivity of the detector for this analysis is. The sensitivity of the detector is the median upper limit (90% confidence level) of the amount of signal events μ_s in a background only scenario (i.e. absence of signal), estimated with 10000 pseudo-experiments. Statistically it is possible that events seem to be signal events, even though there is no signal, but only background. The lower this median upper limit, the better the sensitivity of the detector for the analysis.

To calculate this confidence interval we first need to know the $\mathcal{R}(\mu_s)$ distribution. The procedure to get this $\mathcal{R}(\mu_s)$ distribution is based on the one used in [47] and goes as follows:
For each $\mu_s \in [0, 10]$, with step-size $\Delta\mu_s = 0.1$

1. for each pseudo-experiment k = 1, ..., 10000
   (a) given μ_s, sample a set $\{\psi_i\}_k$ with n_{obs} space angles from eq. (7.2)
   (b) calculate $\mathcal{L}_k(\mu_{s_k})$ with eq. (7.3)
   (c) find $\hat{\mu}_{s_k}$ with maximum likelihood $\mathcal{L}_k(\hat{\mu}_{s_k})$
   (d) calculate $\ln \mathcal{R}_k(\mu_s)$
2. find the critical value $\ln \mathcal{R}^{\alpha}_{\text{crit}}(\mu_s)$ following the FC ordering principle.

The output of the procedure is shown in Fig. 7.3. The blue-green dots represent the distribution of $\ln \mathcal{R}(\mu_s)$, the black thick line shows the $\mathcal{R}^{0.9}_{\text{crit}}(\mu_s)$ distribution and the red dashed-dotted line shows $-0.5 \cdot \chi^2(0.9, 1)$. There is a theorem (Wilks' Theorem [52])

that states that $-2\ln\mathcal{R}$ becomes χ^2 distributed in a Gaussian scenario ($n_{obs} \to \infty$). So in our case, with one degree of freedom, this theorem states that

$$-2\ln\mathcal{R}^{\alpha}_{\text{crit}}(\mu_s) = \chi^2(\alpha, 1) \tag{7.7}$$

We see that this approximation is valid once μ_s is far enough from the physical boundary region. In the case of Fig. 7.3, eq. (7.7) is valid for $\mu_s \gtrsim 4$, where $\ln\mathcal{R}^{0.9}_{\text{crit}}(\mu_s) \approx -0.5 \cdot \chi^2(0.9, 1) \backsimeq -1.35$.

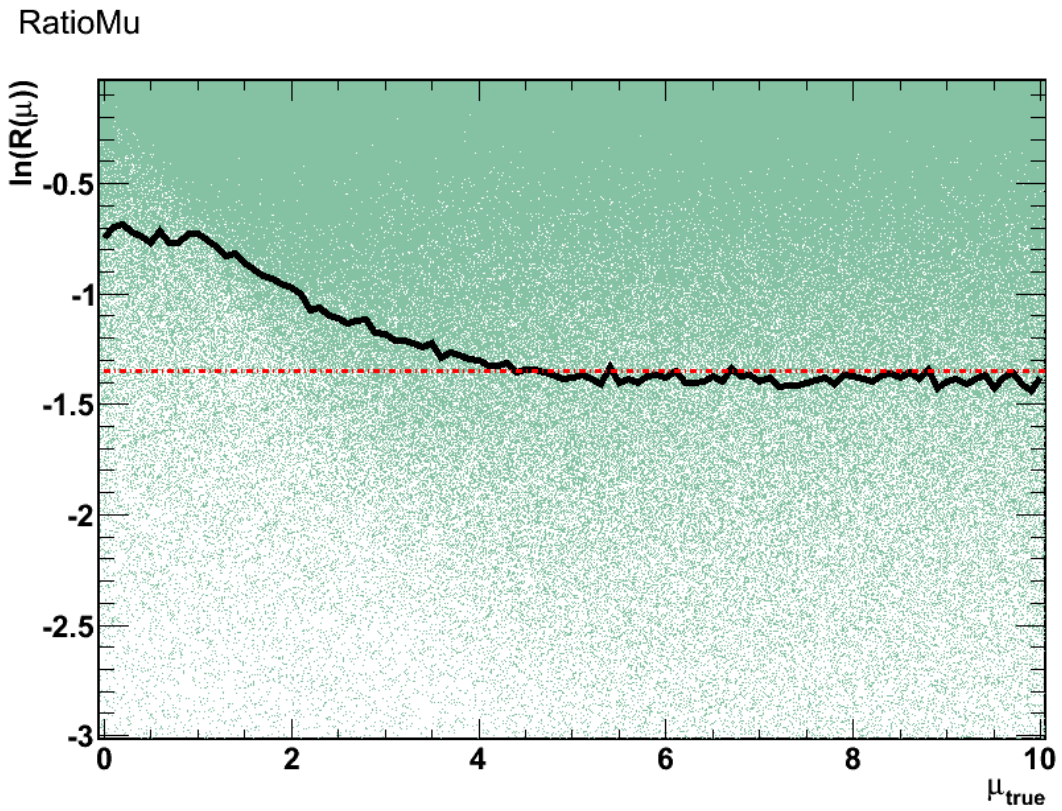

Figure 7.3: Determining the $\ln\mathcal{R}^{90}_{\text{crit}}(\mu_s)$ distribution. The blue-green dots represent the distribution of $\ln\mathcal{R}(\mu_s)$, the black thick line shows the $\mathcal{R}^{0.9}_{\text{crit}}(\mu_s)$ distribution and the red dashed-dotted line shows $-0.5 \cdot \chi^2(0.9, 1)$

Now we have all the information to calculate the confidence intervals from eq. (7.6) in the background-only scenario ($\mu_{s,true} = 0$). The lower- and upper limits of the confidence intervals at 90% confidence level, for the 10000 pseudo-experiments are shown in Fig. 7.4. Notice that the lower limit of the interval is sometimes greater than zero. We expect this to happen in a fraction 1 - 0.9 = 0.1 of the pseudo-experiments and indeed, it happened in a fraction 0.1042 of the pseudo-experiments.

From Fig. 7.4, we can calculate the median upper limit at 90% confidence level on μ_s for the 1000GeV hard WIMPs. This median upper limit is (with it's 1σ spread)

$$\tilde{\mu}^{90}_s = 5.6^{+4.4}_{-3.0} \text{ events} \tag{7.8}$$

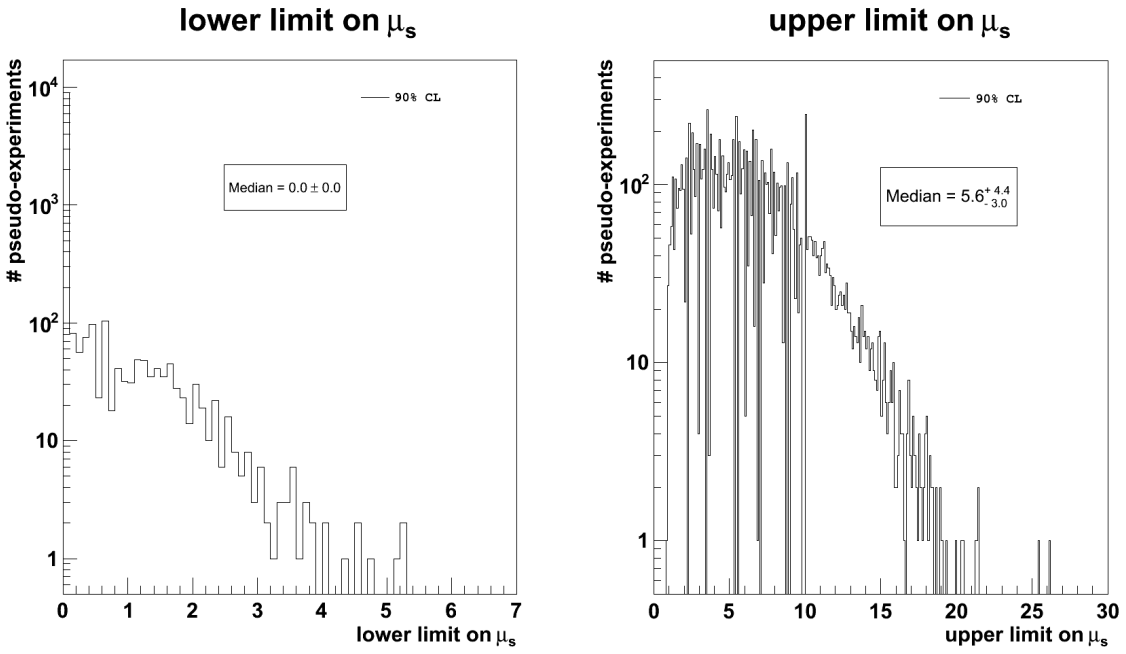

Figure 7.4: The lower (left) and upper (right) limits for μ_s in the background-only scenario, at 90% confidence level. The values shown in the plots are the median limits with their 1σ uncertainties.

If we would now look at 1 month of on source data we expect in case of no neutralino signal, the upper limits on μ_s at 90% confidence level, to be inside the 1σ band with 68% probability. So this result will be important when we will unblind the data, but that's beyond the scope of this thesis, because of two reasons :

1. getting the approval of the collaboration to unblind is an elaborate procedure and takes more time than allocated for this thesis.

2. a more detailed study of the systematic uncertainties needs to be done. This study is beyond the scope of this Masters thesis.

7.2 Physical Quantities

We can now use the median upper limit from eq. (7.8) to calculate limits on some interesting physical quantities. First of all, we can calculate the *Volumetric Flux* $\Gamma_{\nu\to\mu}$. This volumetric flux is a quantity that tells how many signal-like neutrinos would convert per second into muons in the detector if it would have a size of 1 km^3. The upper limit in eq. (7.8) was calculated with 1 month of data and only takes into account the signal-like neutrinos that would really be detected and therefor interact in the effective volume V_{eff}. To convert this number to an upper-limit on the flux of signal-like events

μ_s per km^3, we should calculate

$$\tilde{\Gamma}^{90}_{\nu\to\mu} = \frac{\tilde{\mu}^{90}_s}{V_{eff} \times t_{\text{live}}}, \tag{7.9}$$

where t_{live} is the detector livetime. If we now fill in the numbers, we find that for a 1000GeV WIMP annihilating in W^+W^- the median upper-limit of the volumetric flux at 90% C.L. is

$$\Gamma_{\nu\to\mu} \leq \frac{5.6 \text{ events}}{0.104\ km^3 \times 1\text{month}} = 2.01 \times 10^{-5}\ km^{-3} \cdot s^{-1} = 646\ km^{-3} \cdot yr^{-1} \tag{7.10}$$

From $\Gamma_{\nu\to\mu}$ we can calculate the neutralino annihilation rate in the Sun Γ_A, since [41]

$$\Gamma_{\nu\to\mu} = \frac{\Gamma_A}{4\pi D^2_\odot} \int_0^\infty dE_\nu \sigma_{\nu N}(E_\mu > E_{thr}|E_\nu)\rho_N \sum_X B_X \left(\frac{dN}{dE_\nu}\right)_X, \tag{7.11}$$

where $D_\odot$ is the distance between the detector and the Sun, $\sigma_{\nu N}$ is the neutrino-nucleon cross section, ρ_N is the nucleon density at the detector, E_{thr} is the muon energy threshold of the simulation (this is 1 GeV in the simulation we used), and B_X is the branching ratio for the annihilation channel X with associated neutrino spectrum $\left(\frac{dN}{dE_\nu}\right)_X$.
The unknown parameters in the above equation (7.11) are

- $\Gamma_{\nu\to\mu}$, but we calculated an upper limit on this quantity (eq. 7.8),
- Γ_A, for which we would like to calculate an upper-limit,
- The total neutrino spectrum

$$\left(\frac{dN_\nu}{dE_\nu}\right) = \sum_X B_X \left(\frac{dN}{dE_\nu}\right)_X. \tag{7.12}$$

  this spectrum depends on the model you use to describe the neutralino particles. We already discussed this in section 3.1 and as we already mentioned several times, we are calculating everything for 1000 GeV neutralinos annihilating into W^+W^-.

The median upper-limit at 90% C.L. for Γ_A is

$$\tilde{\Gamma}^{90}_A = 1.13 \times 10^{21} Hz \tag{7.13}$$

Another, very interesting parameter is the neutrino induced muon flux, which is given by [41]

$$\Phi(E_\mu > E_{thr}) = \frac{\Gamma_A}{4\pi D^2_\odot} \int_{E_{thr}}^\infty dE_\mu \frac{dN_\mu}{dE_\mu}, \tag{7.14}$$

where $\frac{dN_\mu}{dE_\mu}$ is the neutrino induced muon energy spectrum at the detector, and $E_{thr} = 1GeV$ is the muon energy threshold in the simulation. We find that the median upper-limit at 90% C.L. is

$$\tilde{\Phi}(E_\mu)^{90} = 341 km^{-2} yr^{-1}. \tag{7.15}$$

The last quantity we calculate is the median upper-limit at 90% C.L. of the neutralino-proton cross section [41]

$$\tilde{\sigma}^{90}_{SD} = 8.26 \times 10^{-4}\ pb, \tag{7.16a}$$
$$\tilde{\sigma}^{90}_{SI} = 6.38 \times 10^{-7}\ pb. \tag{7.16b}$$

The results of Γ_A, Φ_μ and $\sigma_{SD(SI)}$ were obtained through the procedure explained in [41].

Chapter 8

Summary and Outlook

8.1 Summary of the Presented Analysis

In this thesis, we determined what the sensitivity of the 59-string IceCube detector was to detect neutrinos from 1000 GeV neutralinos annihilating to W^+W^- in the Sun. These neutralinos are hypothetical Weakly Interacting Massive Particles, that could account for the missing mass in our Universe. If these particles really exist, they will be captured gravitationally by heavy objects, such as the Sun. Once their concentration is high enough, they will self-annihilate and produce Standard Model (SM) particles. These SM particles will decay into e.g. neutrinos, which will escape from the core of the Sun. If the neutrinos travel in the direction of the Earth, they will sometimes be detected by neutrino detectors such as IceCube.

The first step of our analysis was to filter away most of the atmospheric muons recorded by the 59-string IceCube detector. These muons make up the biggest part of the background. We developed a high level filter (levels 3 and 4) in two steps. First we made some manual cuts by looking at the histograms of 5 variables : θ_{LLH}, z_{travel}, $t_{extension}$, $\Delta_{LLH,linefit}$ and $D_{dom,reco}$. For this we used 1 run (8 hours) of data which corresponds to $\sim 10^6$ events at level 2. These cuts (level 3 filter) removed $\sim 95\%$ of the background the background and only $\sim 25\%$ of signal.

In a second step we used *Boosted Decision Trees* (BDT) to remove the remaining atmospheric muons (level 4 filter). The BDT method allowed us to make a multidimensional cut which removed almost all the atmospheric muons. In this step we used 1 month of (of source) data taken in October 2009.

Table 8.1 summarizes the effects of our high level cuts. The numbers in this table show that we developed a good filter, since we removed almost all the remaining atmospheric muons, but still kept $\sim 40\%$ of the signal w.r.t. the level3 pre-selections.

	# WIMP (%) (1000GeV hard)	Simulated Background (Hz)				Data (Hz)
		ν_{atm}	μ_{single}	$\mu_{multiple}$	total	
before L3	100	0.0103	25.98	9.625	35.62	41.86
after L3	76.90	0.0045	1.242	0.361	1.607	2.027
after BDT	39.12	0.0010	4.98×10^{-5}	2.49×10^{-5}	0.0011	0.00095

Table 8.1: Percentages of the amount of simulated signal and rates of the simulated background and data events before and after the high level cuts.

We optimized our cuts for the 1000GeV hard annihilating WIMPs, and then we looked at the effect on the other WIMP masses and channels. Table 8.2 shows the amount of signal that survived the L3 cuts and the BDT cut for the 14 WIMP models considered. The *soft* channel refers to annihilation in $b\bar{b}$

WIMP Mass (GeV)	WIMP decay channel	# WIMP (%) after the L3 cut	# WIMP (%) after the BDT cut
50	soft	18.48	0.00
	hard	33.19	0.99
100	soft	33.53	2.51
	hard	55.72	6.76
250	soft	57.01	8.17
	hard	73.84	27.91
500	soft	65.43	16.78
	hard	76.93	35.52
1000	soft	70.23	23.35
	hard	76.90	39.12
3000	soft	72.79	28.89
	hard	77.04	38.35
5000	soft	73.35	29.71
	hard	77.08	37.47

Table 8.2: Percentages of the amount of simulated signal after the different cuts (w.r.t. level 3 precuts), for different WIMP models. The *hard* channels refer to annihilation in W^+W^- ($\tau^+\tau^-$ at 50GeV); and the *soft* channels refer to annihilation in $b\bar{b}$.

We also described that the effective area is a useful tool to check theoretical models. Although we didn't use the effective area in our analysis, we still showed it at every filtering level, since it can be useful for other people. Fig. 8.1 shows the effective area at the different filtering levels.

We see from Fig. 8.1 that the filters have strongest rejection at low energies. If the WIMPs have a mass < 250GeV, it is impossible for us to detect them with this detector, since we run out of statistics (see table 8.2). DeepCore will *save* us in this low

mass region.

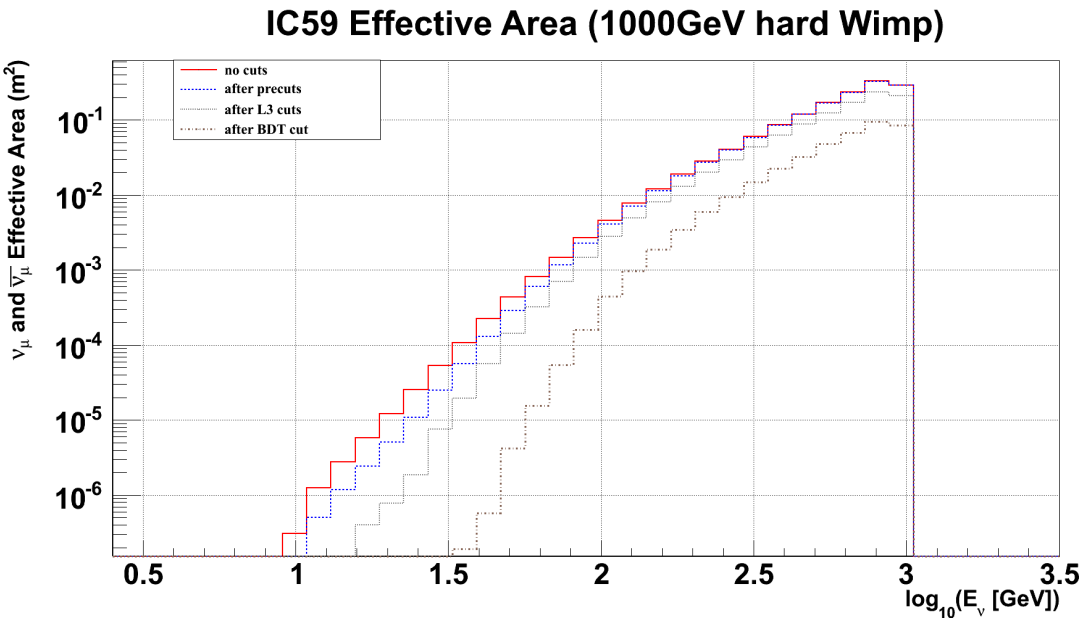

Figure 8.1: Neutrino effective area of the IC59 detector in function of $\log_{10} E_{\nu_\mu}$. This effective area is shown before level 3 filtering (solid red line), for events that passed the level 3 precuts (blue dotted line), events that passed the level 3 cuts (gray small dots) and events that passed the BDT cut (brown dashed-dotted line).

We then used the data that passed the filters to determine the sensitivity of the analysis and the IC59 detector. We determined this sensitivity for the 1000GeV hard WIMP model. In a first step we calculated the angular resolution, which is the median opening angle between the reconstructed direction of the muon track and the direction of the Sun. We found from the simulated WIMP events that the angular resolution for the 1000GeV hard annihilating WIMP model was

$$\tilde{\psi} = 1.89°. \tag{8.1}$$

The angular resolution for the 14 WIMP models is given in Table 8.3.

WIMP Mass (GeV)	WIMP decay channel	$\tilde{\psi}(^\circ)$
50	soft	no signal left...
	hard	6.08
100	soft	6.59
	hard	5.55
250	soft	4.47
	hard	2.80
500	soft	3.52
	hard	2.03
1000	soft	2.83
	hard	1.89
3000	soft	2.51
	hard	1.82
5000	soft	2.29
	hard	1.90

Table 8.3: Median opening angle $\tilde{\psi}$ for different WIMP models.

In the second step we calculated the median upper-limit at 90% confidence level (C.L.) on the amount of signal events measure in the background-only scenario. This is what we call the sensitivity of the detector and the analysis. We calculated this sensitivity for the 1000GeV hard annihilating WIMPs only. The median upper limit we found was

$$\tilde{\mu}_s^{90} = 5.6^{+4.4}_{-3.0} \text{ events} \tag{8.2}$$

This result allowed us to calculate the sensitivities on the following physical quantities:

- The volumetric flux $\Gamma_{\nu\to\mu}$

$$\tilde{\Gamma}^{90}_{\nu\to\mu} = \frac{\tilde{\mu}_s^{90}}{V_{eff} \times t_{\text{live}}}, \tag{8.3}$$

- the neutralino annihilation rate Γ_A

$$\tilde{\Gamma}^{90}_{A} = 1.13 \times 10^{21} Hz \tag{8.4}$$

- the neutrino induced muon flux $\Phi(E_\mu)$

$$\tilde{\Phi}(E_\mu)^{90} = 341 km^{-2} yr^{-1}. \tag{8.5}$$

- the spin-dependent (SD) and spin-independent (SI) neutralino cross sections

$$\tilde{\sigma}^{90}_{SD} = 8.26 \times 10^{-4}\ pb, \tag{8.6a}$$
$$\tilde{\sigma}^{90}_{SI} = 6.38 \times 10^{-7}\ pb. \tag{8.6b}$$

We can now compare these results with the results of other analyses. Fig. 8.2 shows the upper limits on the muon-flux according to 2 different analyses

1. IC22 data (blue lines),

2. the AMANDA data.

The red line is preliminary. It shows the expected sensitivity of the IC86 detector after 5 years. The shaded regions in the plot show the theoretically accepted regions that are still allowed by the spin-independent cross section limits set by direct experiments.

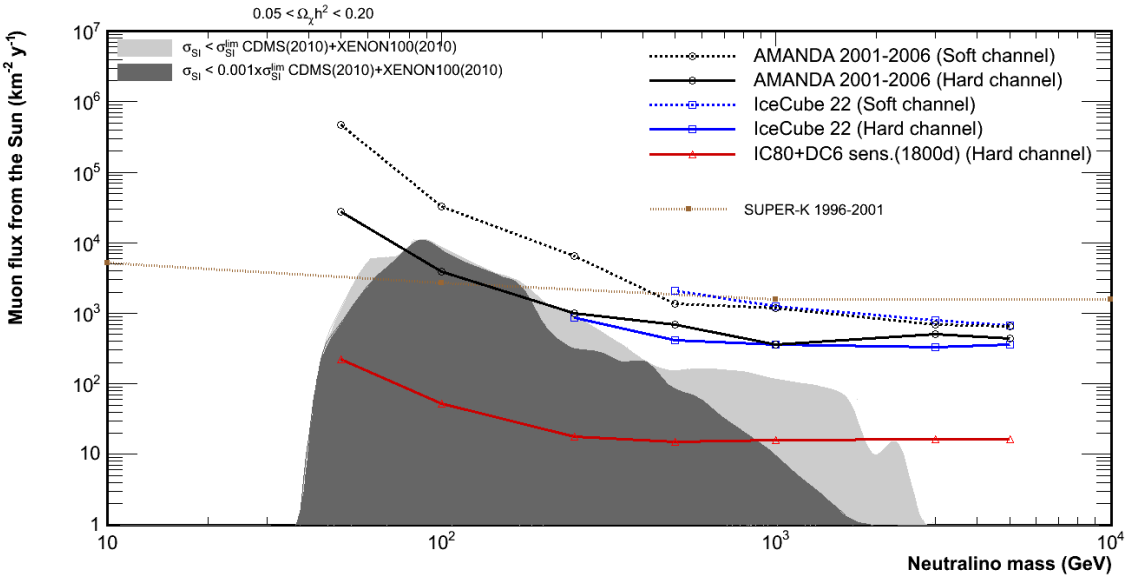

Figure 8.2: The upper limits for each neutralino model on the muon flux from the Sun at 90% confidence level, calculated by using AMANDA data (black line) and IC22 data (blue line). The red line is preliminary. It shows the expected sensitivity of the IC86 detector after 5 years. The shaded regions in the plot represent the MSSM-7[25] parameter space that are still allowed by the spin-independent cross section limits set by direct experiments [47]

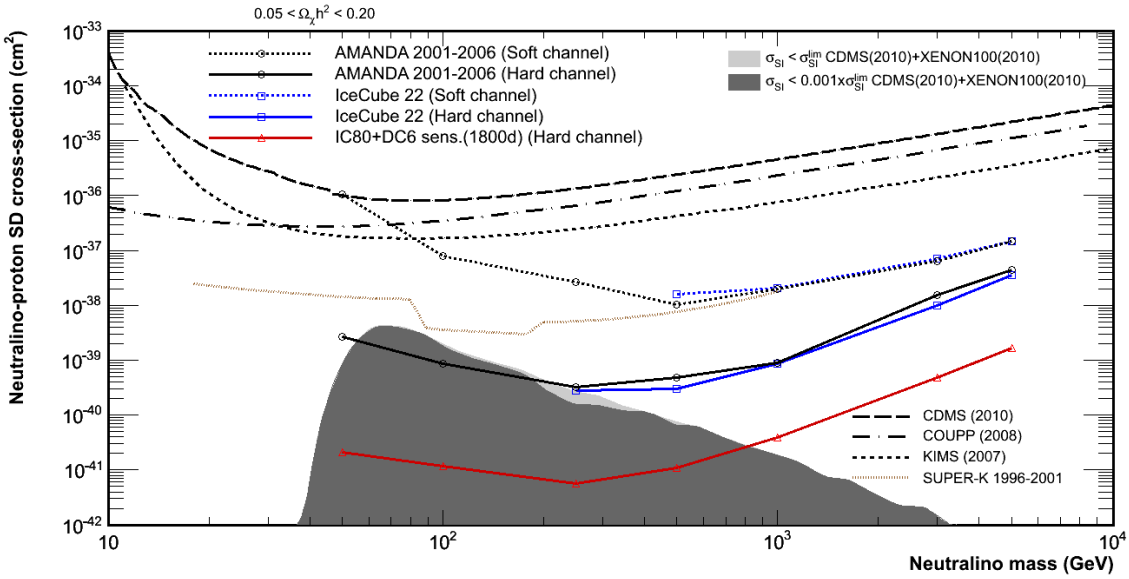

Figure 8.3: The upper limits at 90% confidence level on the spin-dependent neutralino scattering cross section on protons, including systematics. These upper-limits are calculated by using AMANDA data (black line) and IC22 data (blue line). The red line is preliminary. It shows the expected sensitivity of the IC86 detector after 5 years. The shaded regions in the plot represent the MSSM-7[25] parameter space that is not excluded by the spin-independent cross section results from CDMS[27] and XENON100[26]. [47]

If we now compare our results in equations (8.5) and (8.6a) with the results shown in resp. Figs. 8.2 and 8.3, we see that we do slightly better than the previous results. We can off course only make the comparison for the 1000GeV hard annihilating WIMPs, since we didn't calculate the sensitivities for the othe models shown in the plots.

We see that the upper limit that we found on the muon flux in case of a 1000GeV hard annihilating neutralino (see eq. (8.5)) is slightly better than the results that are shown in Fig. 8.2. The same is true for the spin-dependent cross section : compare eq. (8.6a) with the result in Fig. 8.3. Note that we only used 1 month of data in our analysis, so we could improve our result by using the 6 months of (on source) data.

8.2 Outlook

We already mentioned that our analysis wasn't unblind so far. One major part that has to be done is the calculation of the statistical and systematic uncertainties of the quantities in the analysis. Once these uncertainties are calculated, we could start thinking about asking for an unblinding. Once the analysis is unblinded, there are two possibilities

1. we find WIMP signal in the 6 months on source data or

2. we don't find WIMP signal. In this case we can set upper limits on the discussed physical quantities.

Once the analysis is unblinded and the on source data is analyzed, we can go to the next step : redo this analysis with the complete IC86 detector. Since the IC86 detector is bigger and will have a longer lifetime than the IC59 detector, it will provide us more statistics. Also, the IC86 detector contains the complete DeepCore detector, which will allow us to look at lower energy neutrinos. Thanks to this we will be able to analyze the low mass WIMP models.

Appendix A

Parameters Used In BDT

In this chapter, we show the histograms of the 13 parameters used for the BDT training. The parameters are shown before (left plot) and after (right plot) the BDT cut. The line colors and styles are plotted according to the same conventions as the ones used before :

- The experimental data (with errors) : black dots with error bars. This is data of the whole month of October 2009.
- The simulated background (total) : brown dashed line, this simulated background consists of
  - atmospheric neutrinos (ν_{atm}) : red line
  - single muons (μ_{single}) : green line
  - multiple muons ($\mu_{multiple}$) : dotted green line
- The simulated WIMP signal : blue line. The signal rate is scaled to the experimental data rate. As was already said, this is simulation for 1000 GeV hard annihilating WIMPs.

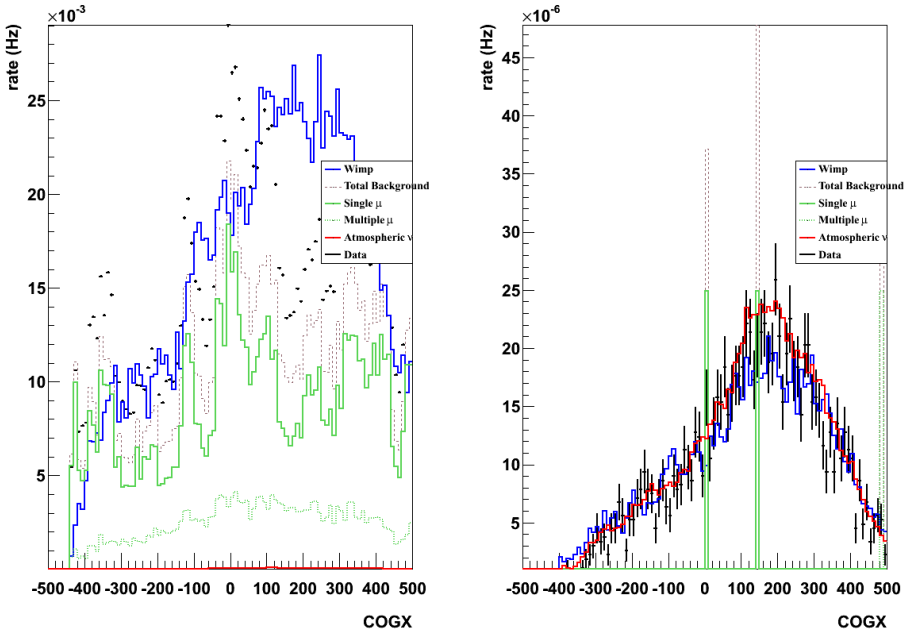

Figure A.1: Center of gravity in X distribution before (left) and after (right) the BDT cut.

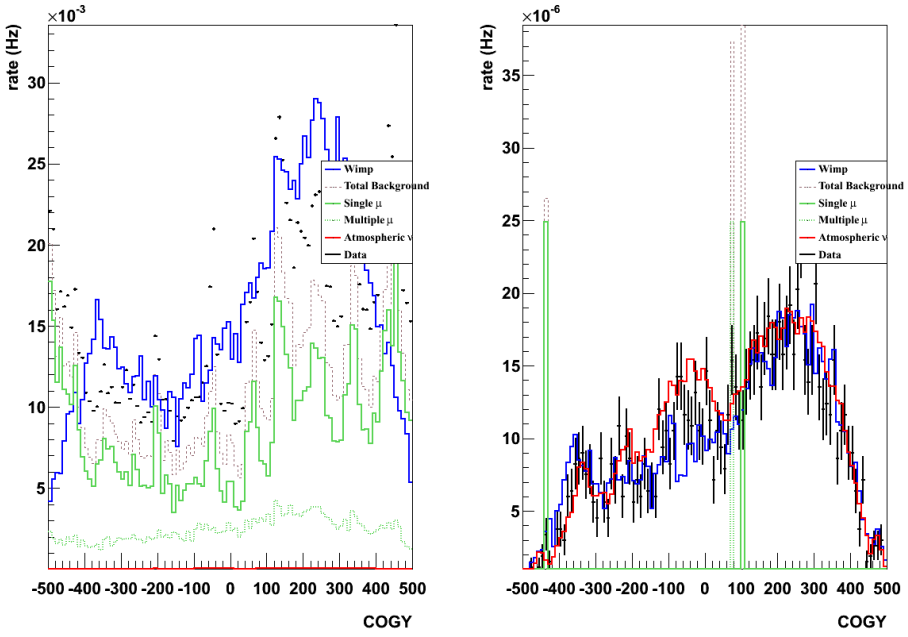

Figure A.2: Center of gravity in Y distribution before (left) and after (right) the BDT cut.

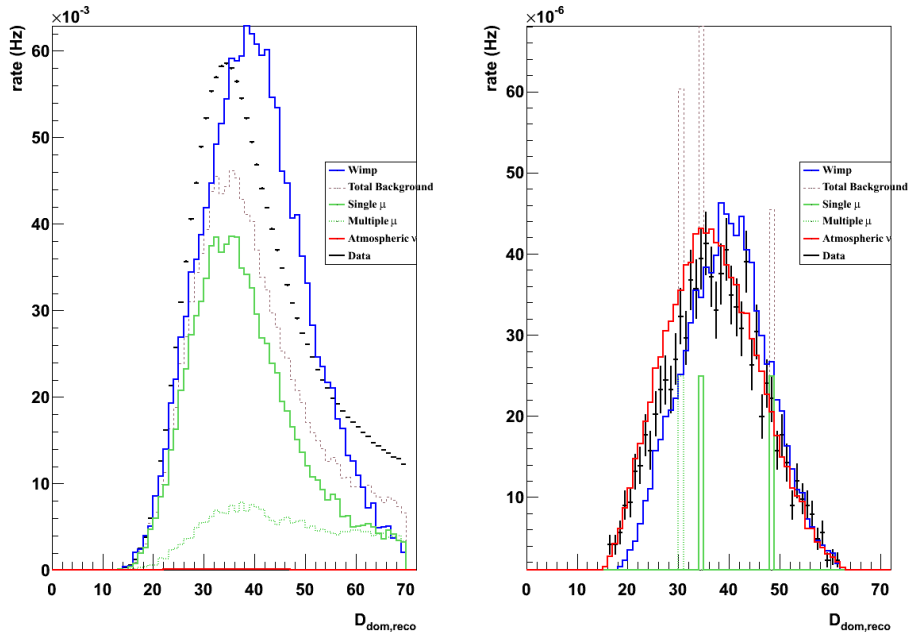

Figure A.3: Distribution of the average distance between the reconstructed track and the fired DOMs, before (left) and after (right) the BDT cut.

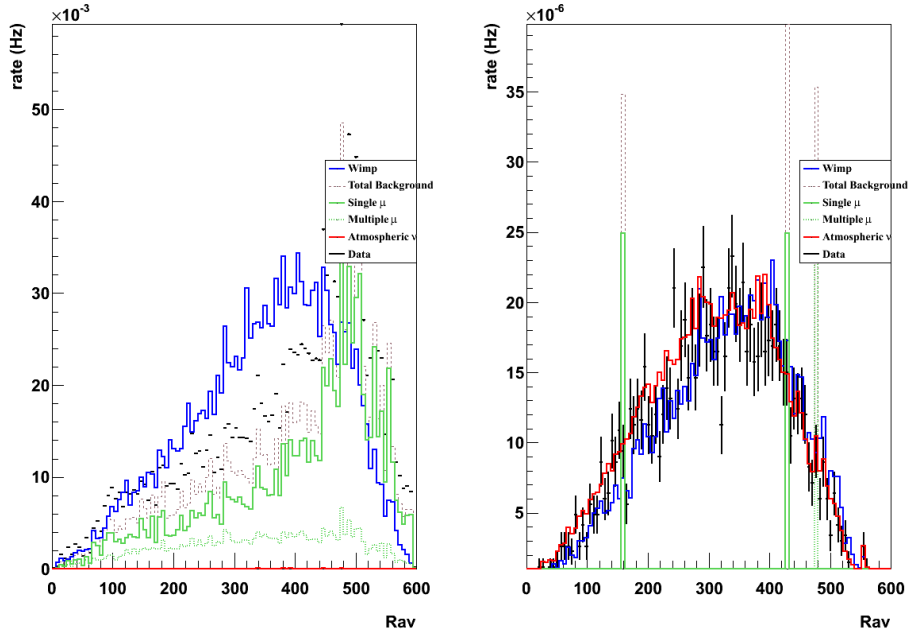

Figure A.4: Distribution of the mean distance in the xy plane of channels from the origin, before (left) and after (right) the BDT cut.

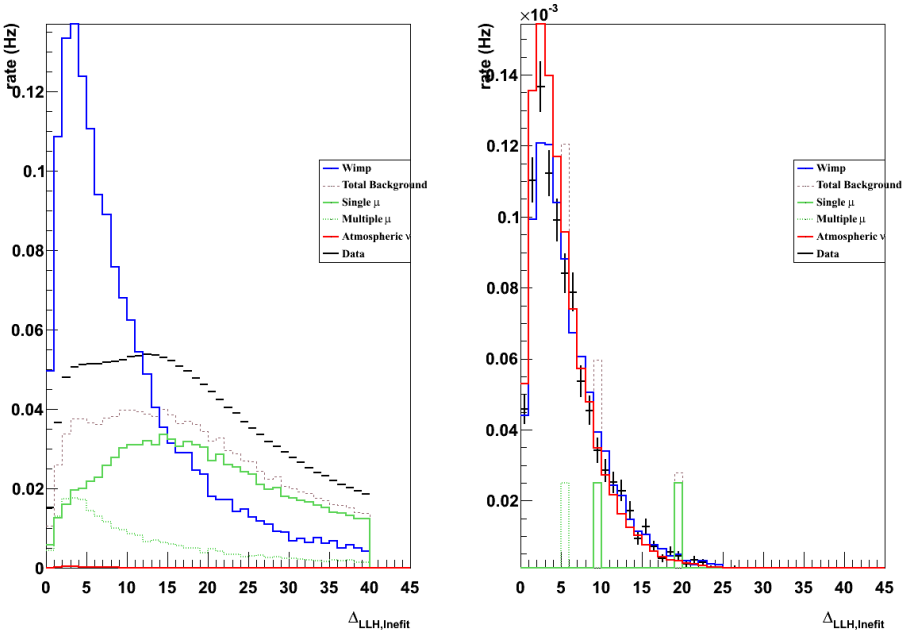

Figure A.5: Distribution of the angle between linefit reconstructed track and LLH reconstructed track, before (left) and after (right) the BDT cut.

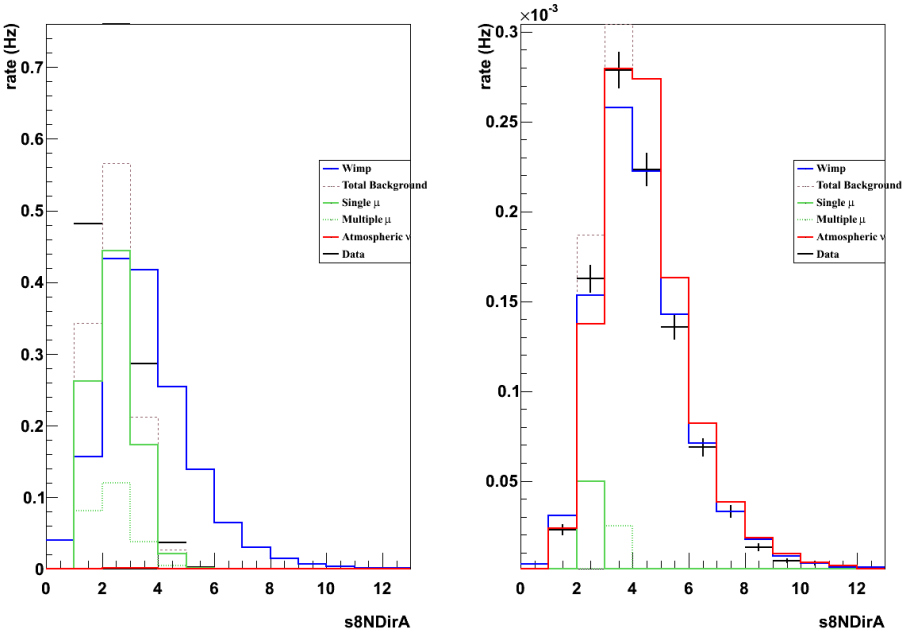

Figure A.6: LLH_{Fit} reduced log likelihood distribution before (left) and after (right) the BDT cut.

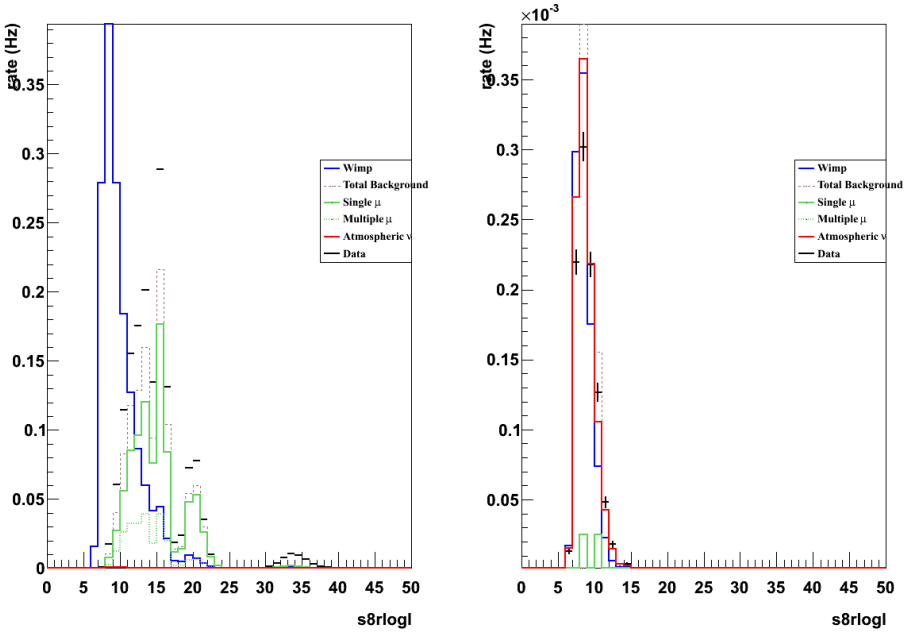

Figure A.7: distribution before (left) and after (right) the BDT cut.

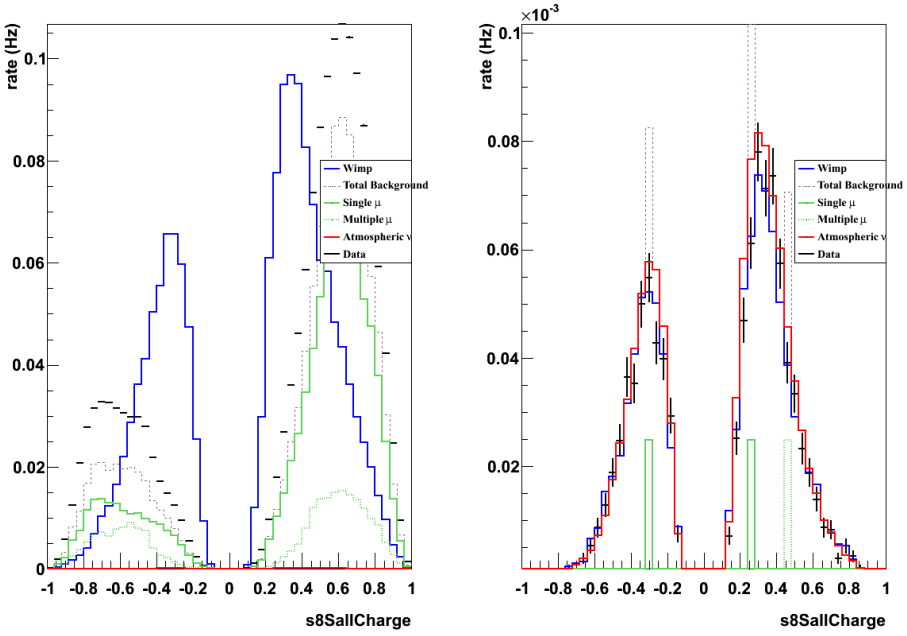

Figure A.8: Distribution of the smoothness of the reconstructed LLH_{Fit} track, using the total charge in the event as a weight, before (left) and after (right) the BDT cut.

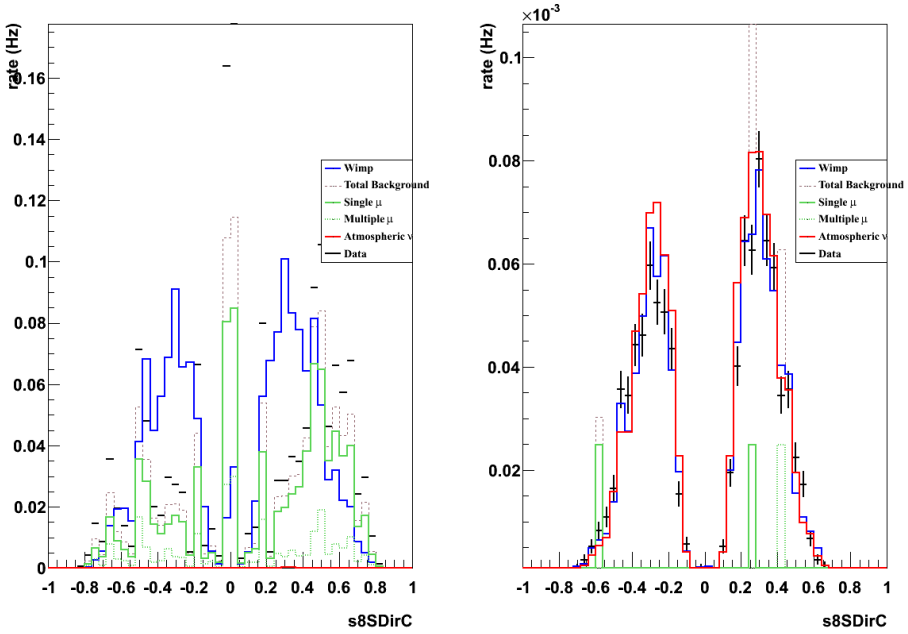

Figure A.9: Distribution of the smoothness of the reconstructed LLH_{Fit} track, in the time residual (-15, +75) ns, before (left) and after (right) the BDT cut.

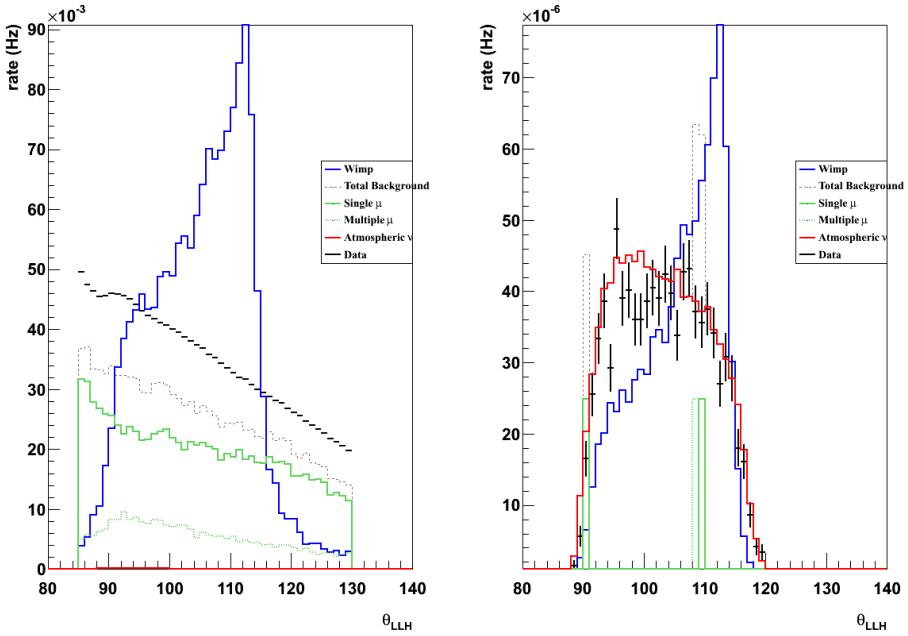

Figure A.10: Reconstructed Zenith distribution before (left) and after (right) the BDT cut.

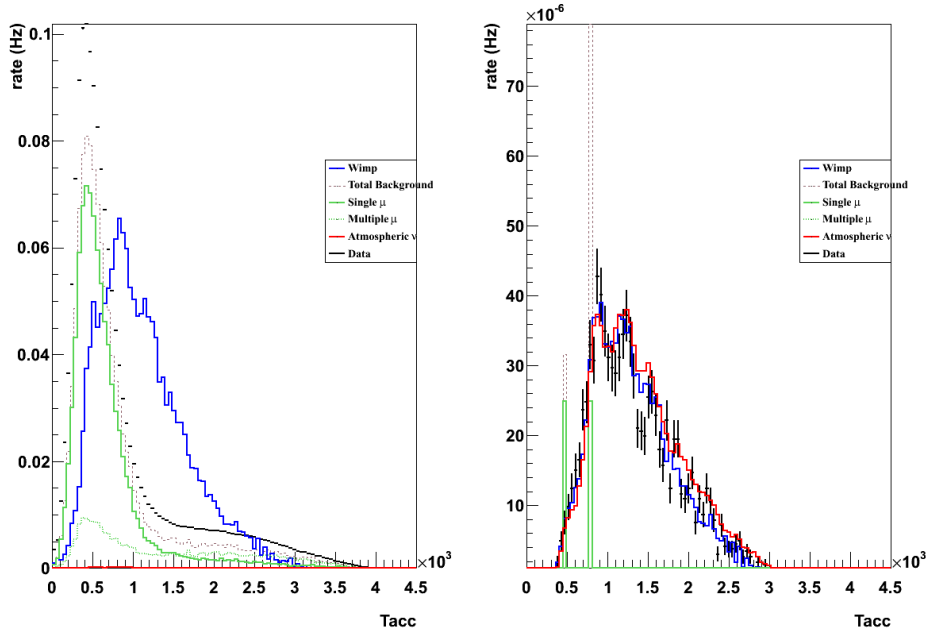

Figure A.11: Distribution of the time to accumulate 75% of the total charge in the event, before (left) and after (right) the BDT cut.

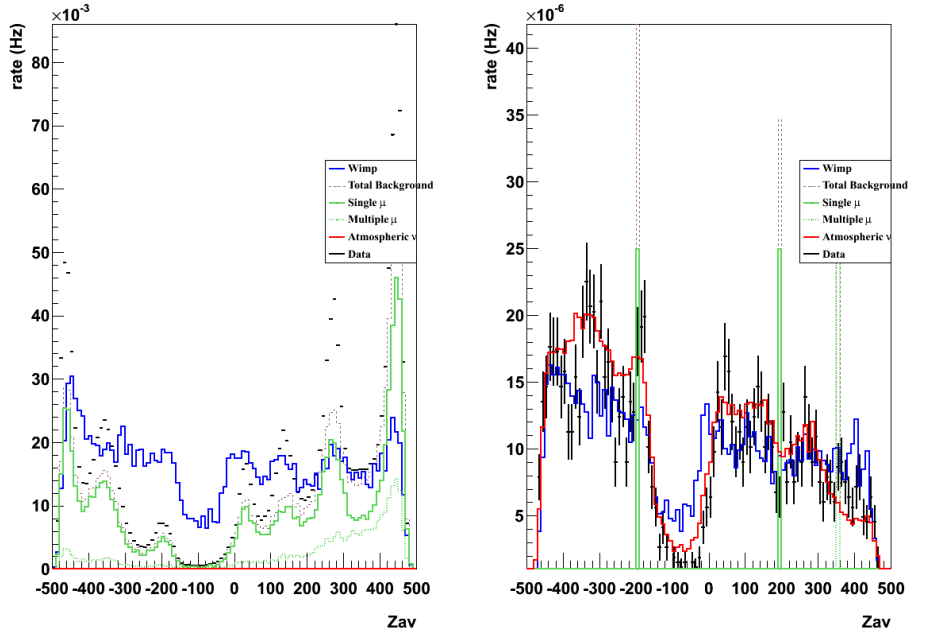

Figure A.12: Distribution of the mean distance in the xy plane of channels from the origin, before (left) and after (right) the BDT cut.

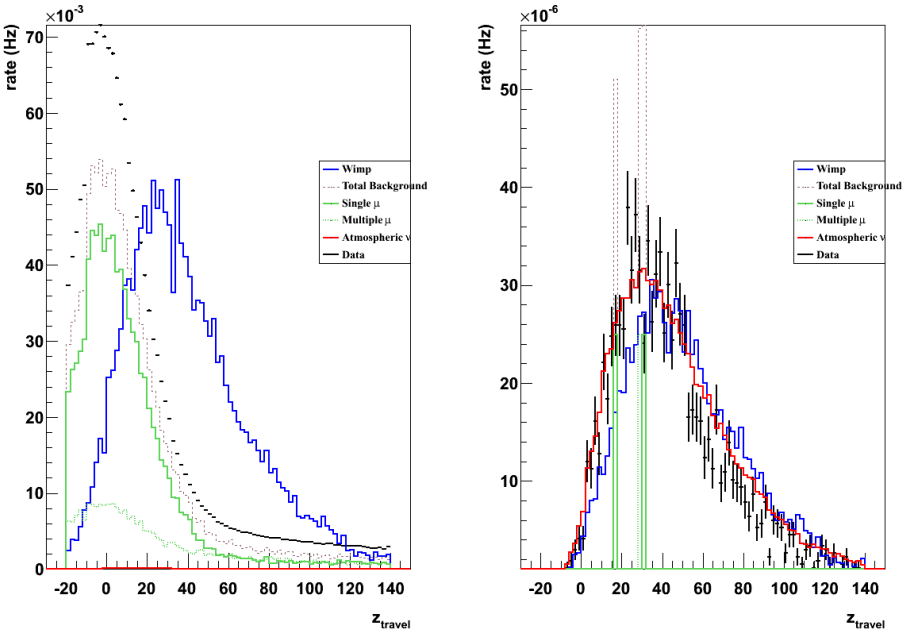

Figure A.13: Distribution of the mean spread in z from the mean z calculated from the first quartile of hits in time, before (left) and after (right) the BDT cut.

Appendix B

Comparing experimental data and simulation

In this chapter we compare the simulated distribution with the experimental distribution of the 13 parameters used for the BDT training. In the first section we compare the CORSIKA events with the experimental data. This comparison is made for events that passed the level 3 filter. In the second section we compare the NuGen events that passed the BDT cut with the experimental data events that passed the BDT cut.

B.1 CORSIKA

In this section we compare the CORSIKA events with the experimental data. The left plots show the distribution of the simulation and experimental data. The simulation is represented by brown * and the data is represented by black dots. The right plots show the ratio $(amountofdata)/(amountofCORSIKA)$. The events (both experimental and simulated) that are used in the plots are events that passed the low level filters (levels 1 and 2). These events are strongly dominated by atmospheric muons, so we expect that the ratio ~ 1 if the CORSIKA simulation is good. The figures show that the ratio isn't ~ 1, but we already explained that this is not a problem for our analysis.

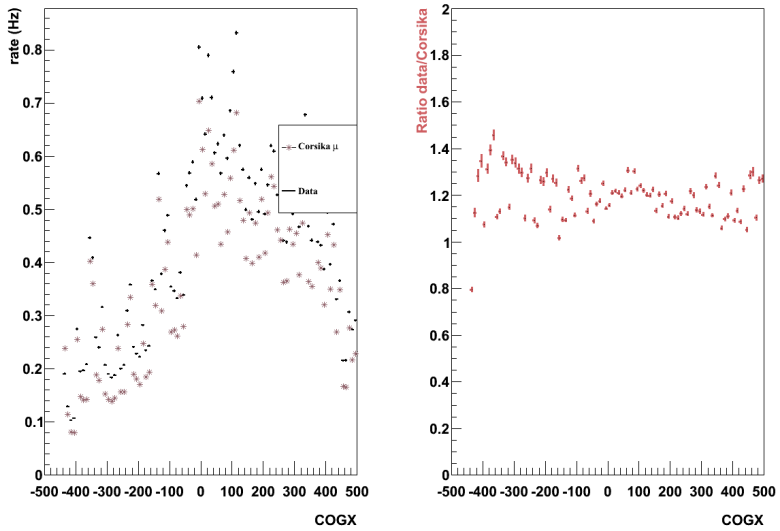

Figure B.1: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : $COGX$ of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for $COGX$.

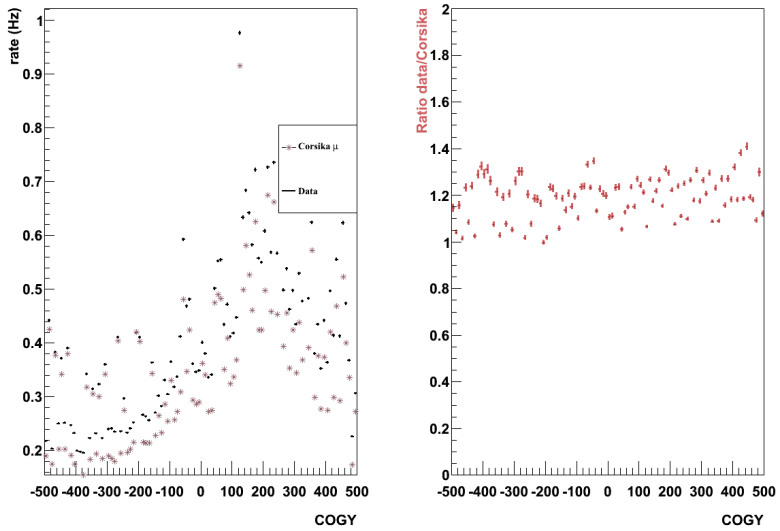

Figure B.2: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : $COGY$ of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for $COGY$.

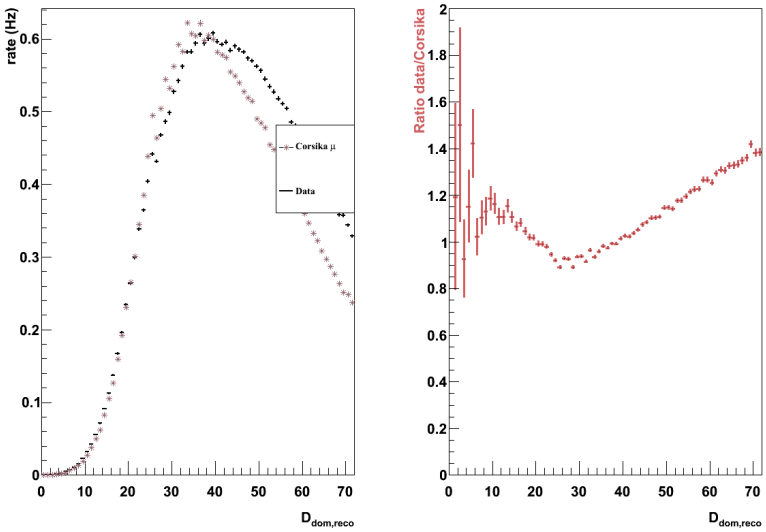

Figure B.3: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : $D_{dom,reco}$ of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for $D_{dom,reco}$.

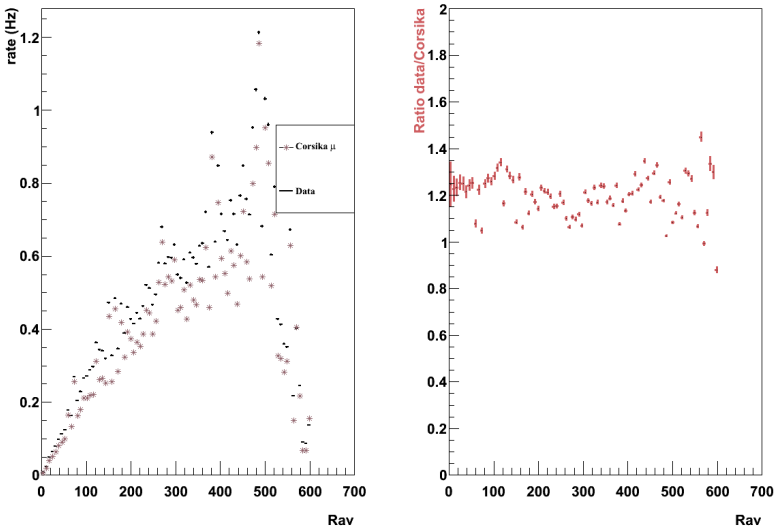

Figure B.4: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : ρ_{av} of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for ρ_{av} .

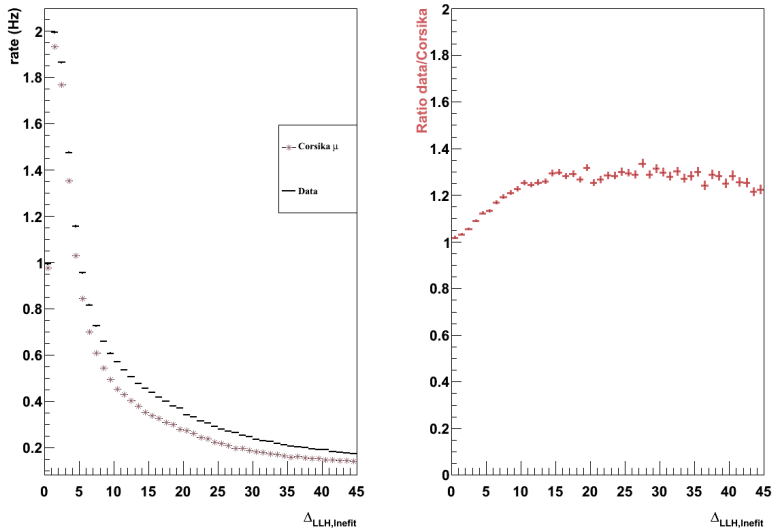

Figure B.5: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : $\Delta_{LLh,linefit}$ of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for $\Delta_{LLh,linefit}$.

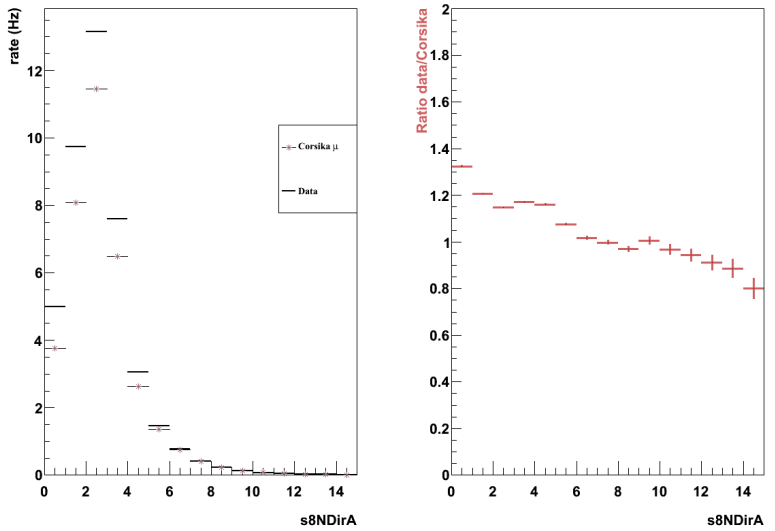

Figure B.6: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : $s8NDirA$ of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for $s8NDirA$.

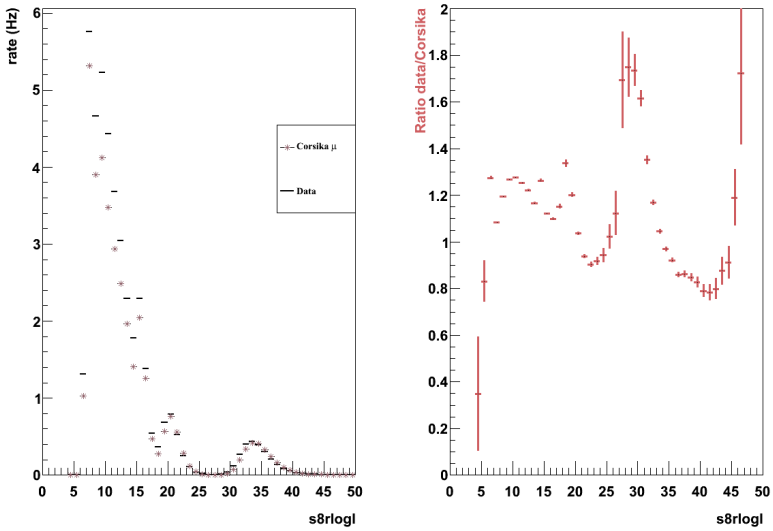

Figure B.7: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : *s8rlogl* of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for *s8rlogl* .

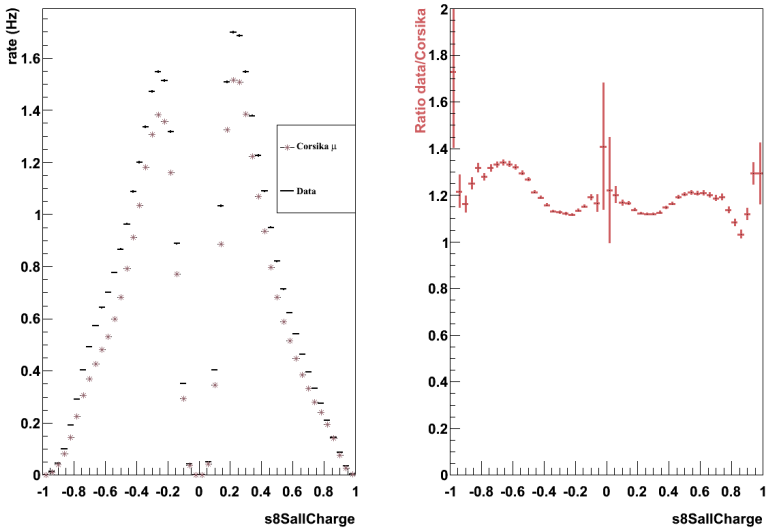

Figure B.8: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : *s8SallCharge* of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for *s8SallCharge* .

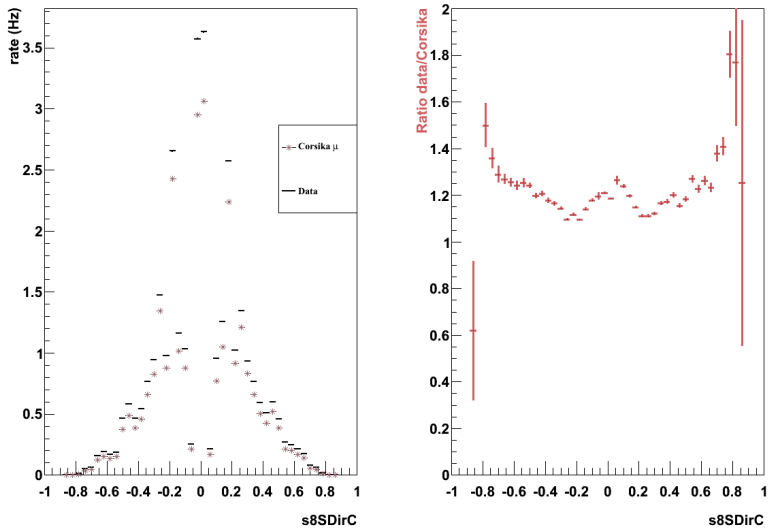

Figure B.9: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : $s8SDirC$ of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for $s8SDirC$.

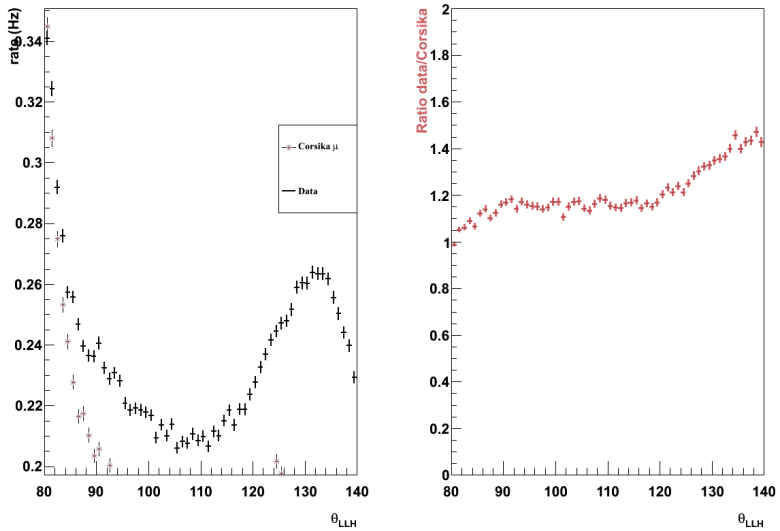

Figure B.10: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : θ_{LLH} of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for θ_{LLH} .

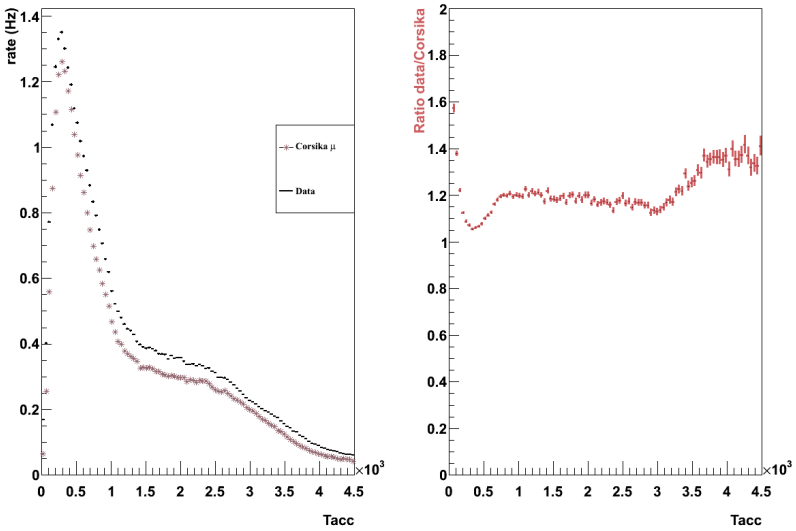

Figure B.11: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : t_{acc} of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for t_{acc} .

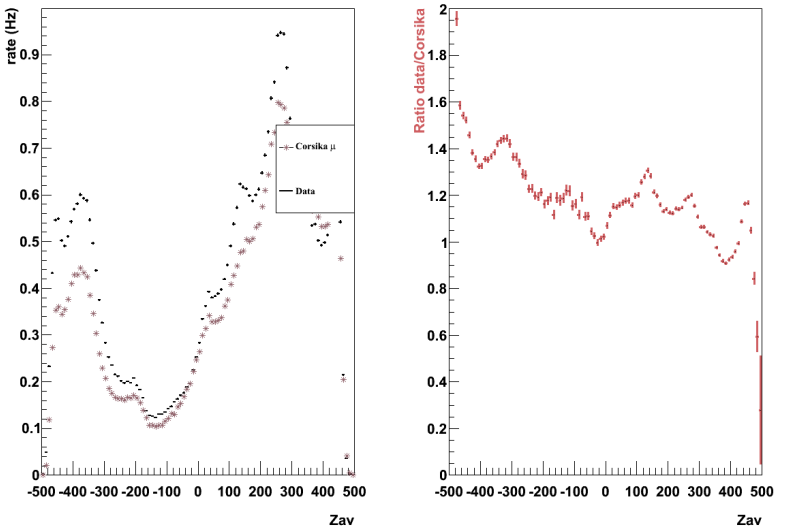

Figure B.12: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : z_{av} of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for z_{av} .

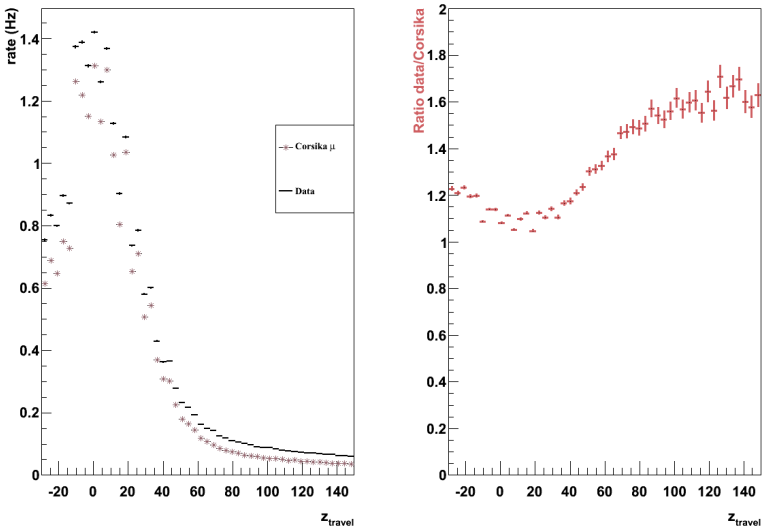

Figure B.13: Comparison between data and simulated atmospheric muon background after the low level filtering (levels 1 & 2). Left : z_{travel} of the event, for both experimental data (black dots) and simulated muon background (brown *). Right : ratio (experimental data)/(simulated muon background) for z_{travel} .

B.2 NuGen

In this section we compare the NuGen events with the experimental data. The left plots show the distribution of the simulation and experimental data. The simulation is represented by red * and the data is represented by black dots. The right plots show the ratio $(amountofdata)/(amountofNuGen)$. The events (both experimental and simulated) that are used in the plots are events that passed the BDT cut. These events are strongly dominated by atmospheric neutrinos, since most of the atmospheric muons are removed by the BDT cut. This means that we expect that the ratio ~ 1 if the NuGen simulation is good. The figures show that this ratio is indeed ~ 1, within the error bars, which is important for our analysis.

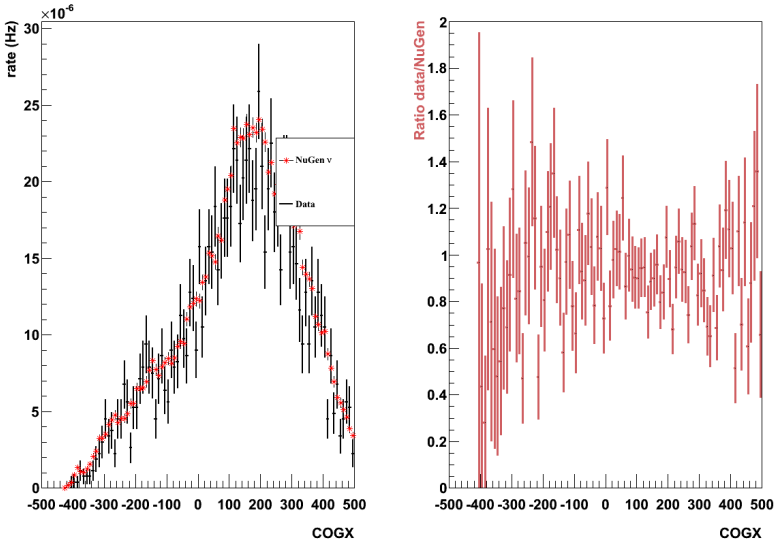

Figure B.14: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : $COGX$ of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for $COGX$

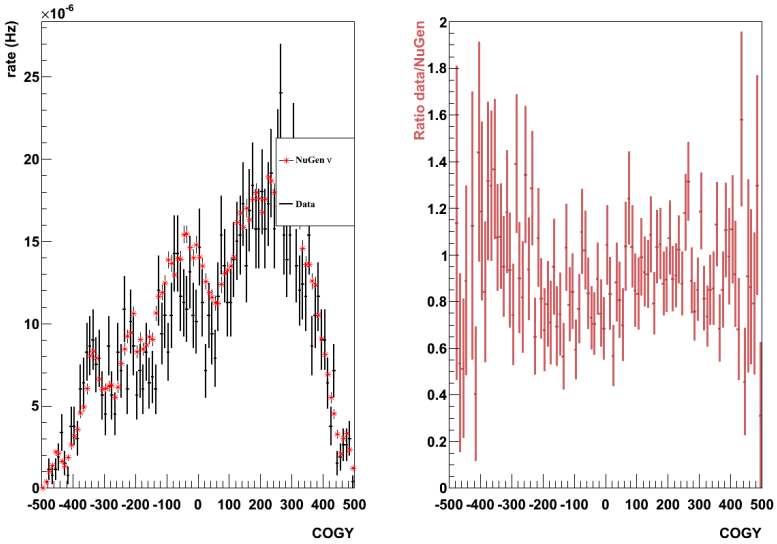

Figure B.15: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : $COGY$ of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for $COGY$

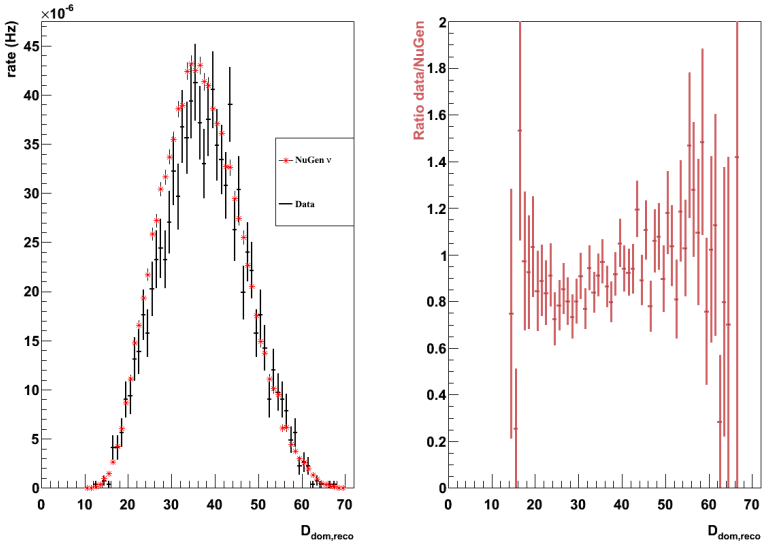

Figure B.16: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : $D_{dom,reco}$ of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for $D_{dom,reco}$

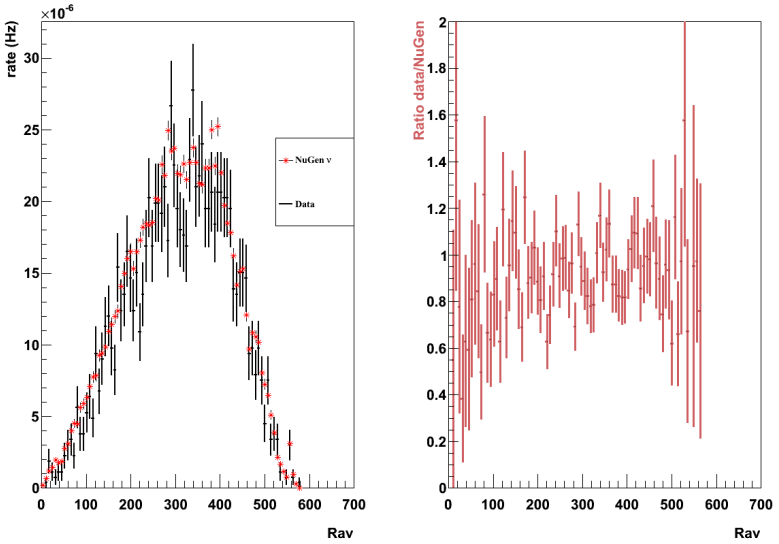

Figure B.17: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : ρ_{av} of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for ρ_{av}

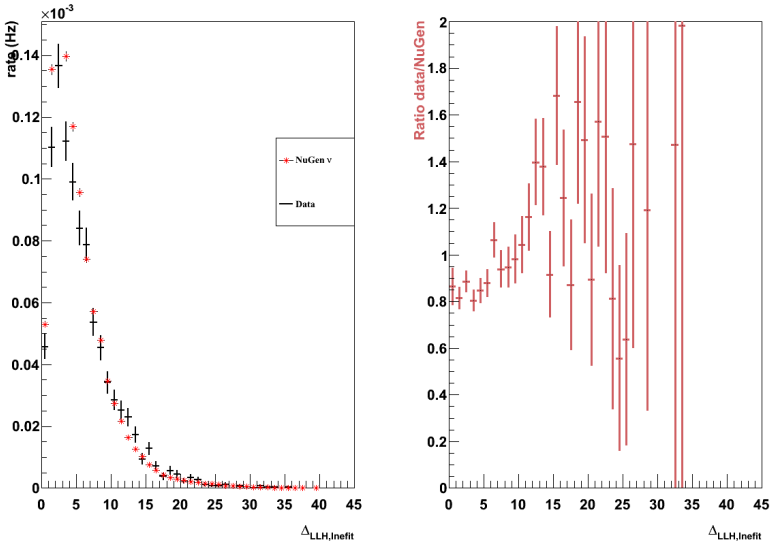

Figure B.18: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : $\Delta_{LLh,linefit}$ of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for $\Delta_{LLh,linefit}$

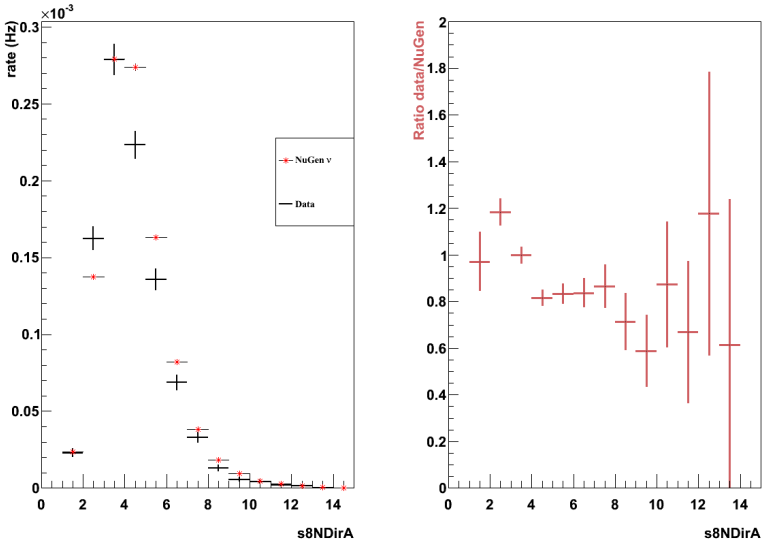

Figure B.19: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : $s8NDirA$ of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for $s8NDirA$

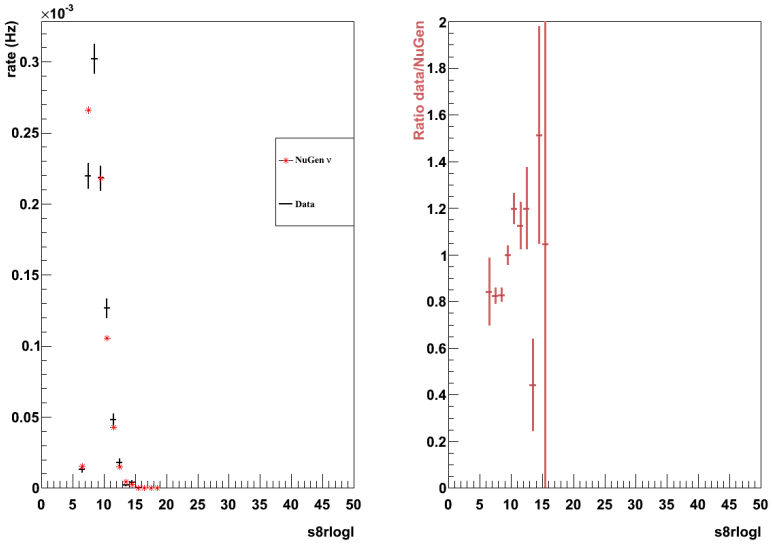

Figure B.20: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : $s8rlogl$ of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for $s8rlogl$

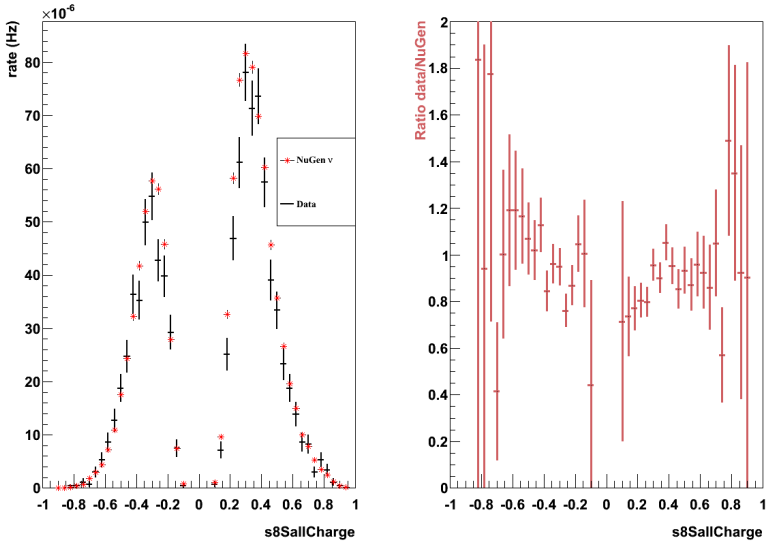

Figure B.21: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : *s8SallCharge* of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for *s8SallCharge*

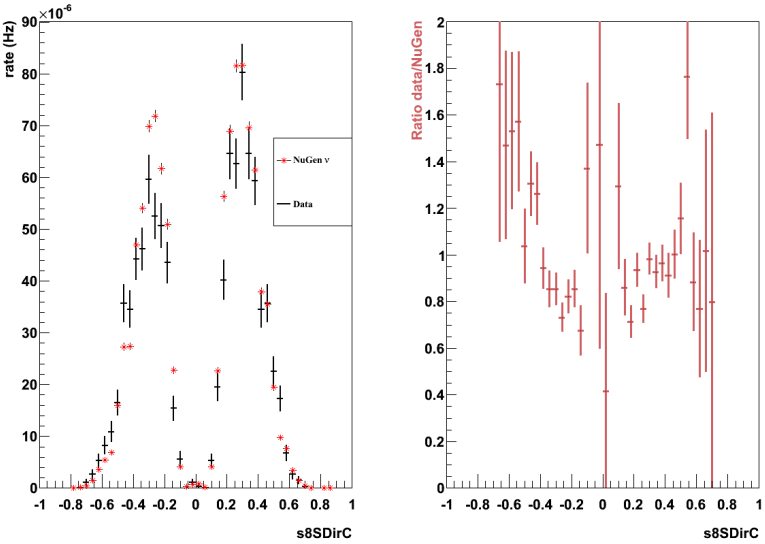

Figure B.22: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : $s8SDirC$ of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for $s8SDirC$

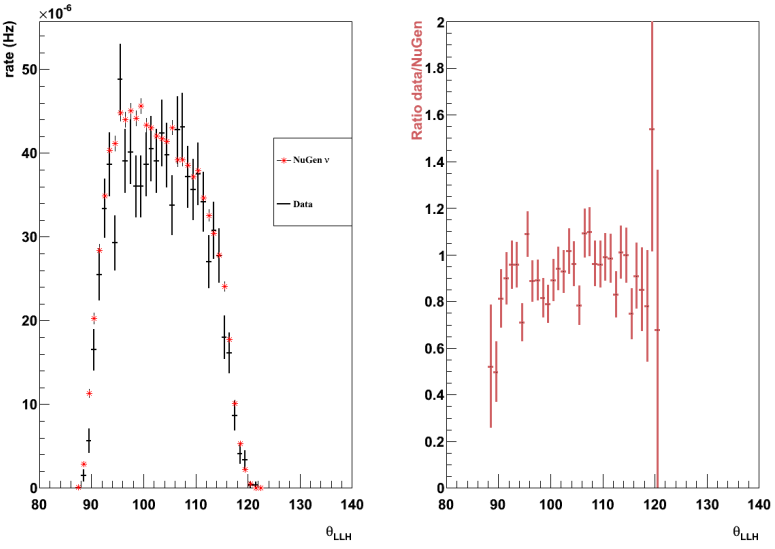

Figure B.23: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : θ_{LLH} of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for θ_{LLH}

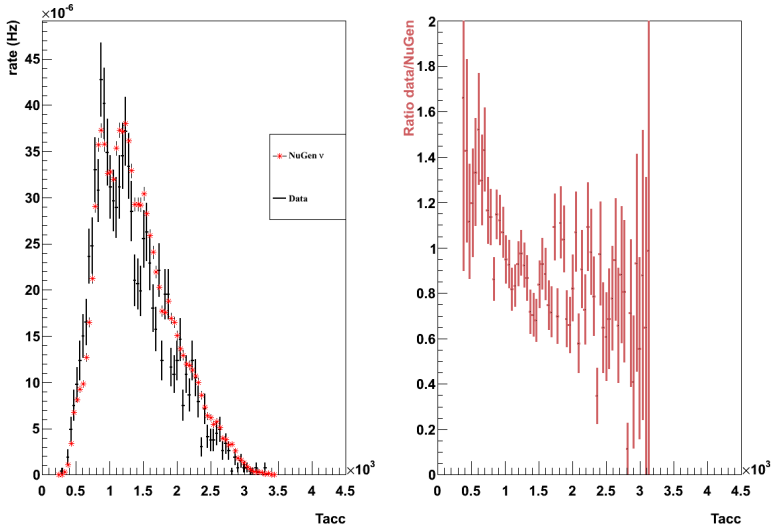

Figure B.24: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : t_{acc} of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for t_{acc}

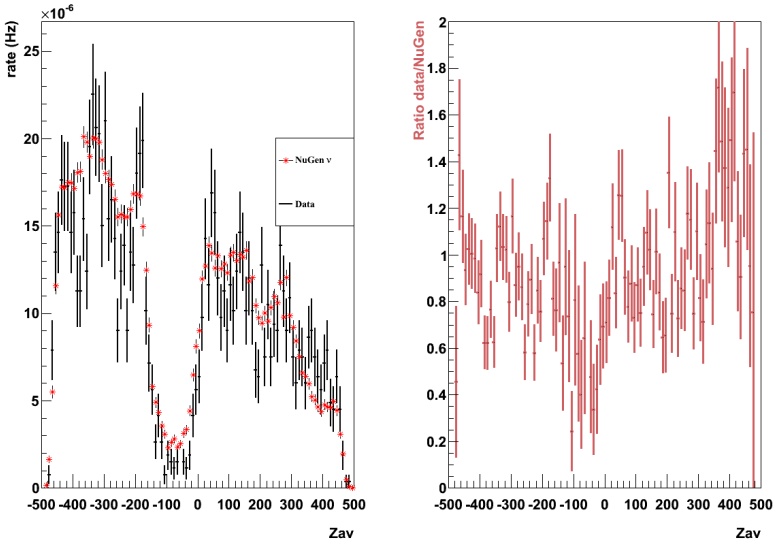

Figure B.25: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : z_{av} of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for z_{av}

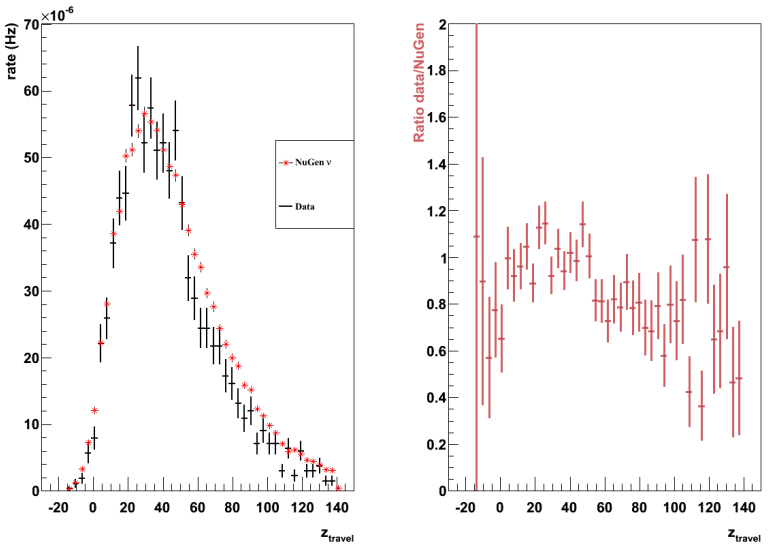

Figure B.26: Comparison between data and simulated atmospheric neutrino background after the high level filtering (levels 3 & 4). Left : z_{travel} of the event, for both experimental data (black dots) and simulated neutrino background (red *). Right : ratio (experimental data)/(simulated neutrino background) for z_{travel}

Bibliography

[1] G. Bertone, D. Hooper J. & Silk, *Particle Dark Matter : Evidence, Candidates and Constraints,* ArXiv.org:physics/0404175v2 (2008).

[2] D. Scott, *The standard cosmological model,* (2005).

[3] E. Hubble, *A relation between distance and radial velocity among extragalactic nebulae,* Proceedings of National Academy of Science, 15, 168-173 (1929).

[4] D. Perkins, *Particle Astrophysics second edition*, Oxford University Press (2009)

[5] M. Carroll, *An introduction to General Relativity spacetime and Geometry*, Benjamin Cummings 2004

[6] N. Polonsky, *Supersymmetry : Structure and Phenomena,* ArXiv.org:physics/0108236 (2001).

[7] S.P. Martin, *A Supersymmetry Primer,* ArXiv:hep-ph/9709356v2 (2008).

[8] L. O'Raifeartaigh and N. Straumann, *Early history of gauge theories and Kaluza-Klein theories, with a glance at recent developments,* (1998).

[9] F. Zwicky,(1937). *On the Masses of Nebulae and of Clusters of Nebulae.* Astrophysical Journal 86: 217.

[10] A. N. Taylor *et al.*, *Gravitational Lens Magnification and the Mass of Abell1689.* The Astrophysical Journal 501: 539. arXiv:astro-ph/9801158.

[11] E. L. Lokas, *Velocity dispersions of dwarf spheroidal galaxies: dark matter versus MOND,* MON.NOT.ROY.ASTRON.SOC. 327 (2001) L21.

[12] E. Komatsu, *et al.*, *Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation.* Astrophys. J. Suppl. Ser. **180**, 330-376 (2009).

[13] F. Halzen, D. Hooper, *The Indirect Search for Dark Matter with IceCube.* ArXiv.org:astro-ph:0910.4513v2.

[14] A. Gould, Astrophys. J. **321**, 571 (1987)

[15] Griest, K. and Seckel, D., Nucl. Phys. B **283**, 681 (1987)

[16] The Particle Data Group Collaboration, C. Ansler *et al.*, *Review of Particle Physics.* Phys. Lett. **B667**, 1 (2008).

[17] The CERN website `http://public.web.cern.ch/public/`

[18] Jonathan L. Feng, *Dark Matter Candidates from Particle Physics and Methods of Detection* ArXiv.org:astro-ph/1003.0904v1

[19] B. Bernabei *et al.*, *New results from DAMA/LIBRA* ArXiv.org:physics/1002.1028.

[20] C. Aalseth *et al.*, *Results from a Search for Light-Mass Dark Matter with a P-type Point Contact Germanium Detector.* ArXiv.org/astro-ph:1002.4703v2 (2010)

[21] F. Montanet, *Design and expected performance of the ANTARES neutrino telescope,* Nucl.Phys.Proc.Suppl. 87 (2000) 436-438, ArXiv:astro-ph/0001380.

[22] F. Halzen, S. R. Klein, *Invited Review Article: IceCube: An instrument for neutrino astronomy,* (2010).

[23] S. Desai *et al.*, *Search for Dark Matter Wimps using upward through-going Muons in Super-Kamiokande.* Phys. Rev. D **70**, 083523 (2004).

[24] PAMELA collaboration, Nature **458**, 607 (2009).

[25] R. Trotta *et al.*, *The impact of priors and observables on parameter inferences in the constrained MSSM.* JHEP **12**, 024 (2008).

[26] E. Aprile *et al.*, *First Dark Matter Results from the XENON100 Experiment.* Phys. Rev. Lett. **105**, 131302 (2010).

[27] CDMS collaboration, ArXiv.org:astro-ph/0912.3592v1 (2009).

[28] A. Achterberg *et al.*, Astropart. Phys. **26**, 155 (2000).

[29] `http://www.icecube.wisc.edu/science/timeline.php`

[30] R. Roosen , *Fenomenologie van de elektromagnetische- Sterke- en zwakke wisselwerkingen*, College nota's (2009).

[31] T. Burgess, *A search for Solar Neutralino Dark Matter with the AMANDA II Neutrino Telescope.* PhD thesis, Stokholm University (2008).

[32] J. G. Learned, K. Mannheim, *High-Energy Neutrino Astrophysics*, Annual Reviews of Nuclear and Particle Science 50 (2000) 679âĂŞ749.

[33] I. Frank and I. Tamm, *Coherent radiation of fast electrons in a medium,* C. R. Ac. Sci. U.S.S.R. 14 (1937) 107.

[34] Fréjus Collaboration, Astropart. Phys. 4, 217 (1996).

[35] IceCube Collaboration, R. Abassi *et al.*, e-print arXiv:1004.2357; IceCube Collaboration, A. Achterberg, M. Ackermann, J. Adams *et al.*, Phys. Rev. D 76, 042008 (2007).

[36] Heck, D. *et al.*, *CORSIKA: A monte carlo code to simulate extensive air showers,* Tech. Rep., FZKA, 6019,1998.

[37] `http://dima.lbl.gov/~dima/work/CORSIKA/` (2010)

[38] `http://wiki.icecube.wisc.edu/index.php/Neutrino_Generator`

[39] M. Blennow, J. Edsj and T. Ohlsson, *Neutrinos from WIMP Annihilations Obtained Using a Full Three-Flavor Monte Carlo Approach*, JCAP 0801:021,2008 , arXiv:0709.3898

[40] P.Gondolo *et al.*, JCAP 0407 (2004) 08

[41] G. Wilkström and J. Edsjö ArXiv.org:0903.2986

[42] D. Chirkin, W. Rhode, *Propagating leptons through matter with Muon Monte Carlo (MMC)*, ArXiv.org:hep-ph/0407075

[43] J. Lundberg *et al.*, *Light tracking through ice and water - Scattering and absorption in heterogeneous media with Photonics,* Nucl.Instrum.Meth.A581:619-631,2007 , astro-ph/0702108.

[44] D. Hubert, *Aspects of neutralino weights and effective volumes.* Internal IceCube report (2007) + PhD thesis, Vrije Universiteit Brussel.

[45] The AMANDA Collaboration, *Muon Track Reconstruction and Data Selection Techniques in AMANDA*, ArXiv.org:astro-ph/0407044v1 (2004).

[46] E. Strahler, *Searches for Neutrinos from Gamma Ray Bursts with the AMANDA-II and IceCube Detectors* PhD thesis, University of Wisconsin, Madison (2009).

[47] A. Rizzo, *Search for neutralino dark matter with 6 years of data of the AMANDA-II neutrino telescope.* PhD thesis, Vrije Universiteit Brussel (2010).

[48] *ROOT : An Object-Oriented Data Analysis Framework.* `http://root.cern.ch`

[49] *TMVA : Toolkit for Multivariate Data Analysis with ROOT.* `http://tmva.sourceforge.net`

[50] G. Punzi, *Sensitivity of searches for new signals and its optimization,* ArXiv.org:physics/0308063 (2003).

[51] G.J. Feldman & R.D. Cousins, *Unified approach to the classical statistical analysis of small signals,* Phys. Rev. D, **57**, 3873 (1998). 142

[52] S.S. Wilks, *Mathematical Statistics* (Wiley, New York, 1962). 2nd printing, corrected, 1963.

[53] `http://www.icecube.wisc.edu/`