Vrije Universiteit Brussel

Faculteit Wetenschappen
Departement Natuurkunde

Search for neutralino dark matter with 6 years of data of the AMANDA-II neutrino telescope

Alfio Rizzo

Proefschrift ingediend met het oog
op het behalen van de academische graad

Doctor in de Wetenschappen

Promotor: Prof. Dr. Catherine De Clercq

December 2010

Print: Silhouet, Maldegem

Cover Image: Cluster Crash Illuminates Dark Matter Conundrum
Credit: X-ray: NASA / CXC/ U. Victoria/ A. Mahdavi et al.
Optical/Lensing: CFHT/ U. Victoria/ A. Mahdavi et al.

2010 Uitgeverij VUBPRESS Brussels University Press
VUBPRESS is an imprint of ASP nv (Academic and Scientific Publishers nv)
Ravensteingalerij 28
B-1000 Brussels
Tel. +32 (0)2 289 26 50
Fax +32 (0)2 289 26 59
E-mail: info@vubpress.be
www.vubpress.be

ISBN 978 90 5487 844 5
NUR 924
Legal Deposit D/2010/11.161/165

Examencommissie voor Doctoraat

Voorzitter: Prof. Dr. J. P. De Greve (VUB)

Promotor: Prof. Dr. C. De Clercq (VUB)

Secretaris: Prof. Dr. N. Van Eijndhoven (VUB)

Prof. Dr. S. Tavernier (VUB)

Prof. Dr. C. Wiebusch (RWTH Aachen Universiteit)

Prof. Dr. M. Tytgat (ULB)

To my family and
to the memory of my father...

Samenvatting

Volgens verscheidene experimentele waarnemingen moet ons heelal doordrongen zijn van een onbekende soort materie, de zogenaamde donkere materie. Volgens de moderne kosmologie moeten deze deeltjes stabiel en zwaar zijn en enkel met materie interageren via de zwaartekracht of zwakke wisselwerking. Daarom worden zulke deeltjes ook WIMP's (*Weak Interactive Massive Particles* of zwak interagerende massieve deeltjes) genoemd.

Een van de beste WIMP kandidaten is het neutralino zoals gepostuleerd in de supersymmetrische uitbreiding van het Standaard Model van de deeltjesfysica; het kan op zichzelf de volledige donkere materie vormen, of het is misschien een van de ingrediënten. Als neutralino's door middel van de zwaartekracht door de zon werden gevangen, konden ze zich ophopen in de kern en vervolgens paarsgewijs annihileren. Een mogelijke manier om de neutralino donkere materie te onthullen is daarom het detecteren van hun Standaard Model vervalproducten, zoals het neutrino. In dat geval wordt een mogelijke overschrijding boven de atmosferische achtergrond verwacht van neutrino's die uit de richting van de zon komen.

We zijn dit werk begonnen met als doel het zoeken naar neutrino's afkomstig uit neutralino-interacties in het centrum van de zon. De gegevens verzameld gedurende de periode van 2001 tot 2006 met de AMANDA neutrino telescoop, die gelegen is op de Zuidpool, werden benut voor dit werk. Daarvan hebben we ongeveer het equivalente van 812 dagen opnametijd gebruikt die geschikt zijn voor de specifieke vereisten van deze analyse.

Het belangrijkste doel van onze analyse was de indirecte waarneming van het neutralino, maar alvorens dit punt te bereiken was voorbereidend werk nodig om de contaminatie door de atmosferische achtergrond te verwijderen uit de experimentele data. Vanwege de positie van de zon op de Zuidpool verwachtten we bijna horizontale sporen van laag-energetische gebeurtenissen, die een echte uitdaging voor het reconstructie-algoritme vormde. Een ander belangrijk as-

pect was de aanwezigheid van de *string trigger* die de drempel om deze gebeurtenissen te accepteren verlaagde.

Wij introduceerden een multivariate techniek, de zogenaamde *Boosted Decision Trees* (BDT), om het grootste deel van deze achtergrond van atmosferische muonen te verwijderen, en tegelijkertijd zo veel mogelijk signaal over te houden.

De prestaties van deze methode stonden met kop en schouders boven een eenvoudige een-dimensionale selectie methode, die in de voorgaande AMANDA analyses werd gebruikt. Na het toepassen van de BDT selectie hebben we gekozen om een nieuwe en verfijnde methode aan te nemen om de signaalsterkte te onttrekken uit de resterende experimentele gegevens.

Het uiteindelijke resultaat van onze analyse was dat in de uiteindelijke steekproef geen statistisch significante toename van gebeurtenissen uit de richting van de Zon werd gevonden. Een bovenlimiet met 90% betrouwbaarheidsniveau voor het aantal verwachte signaalgebeurtenissen werd verstrekt voor de selecties van verschillende neutralinomodellen. Deze bovengrens werd benut om een bovenlimiet af te leiden voor het neutrino-muon conversietempo in de detector, het annihilatietempo van het neutralino in de zon, de neutrino-geïnduceerde muonflux door de detector en de spinafhankelijke en -onafhankelijke neutralino-proton werkzame doorsneden.

Globaal gezien presteert onze analyse zeer goed; de resultaten geciteerd voor de harde-kanaalmodellen met lage neutralinomassa en voor alle zachte-kanaalmodellen, zijn tot dusver de meest gevoelige AMANDA/IceCube resultaten met betrekking tot donkere materie. De belangrijkste reden is het gebruik van de multidimensionale methode die een betere scheiding tussen signaal en achtergrond mogelijk maakt. Een onderzoek naar donkere materie uitgevoerd met IceCube gegevens (met slechts 22 *strings*) reikt niet tot de lage-energieregio, omdat deze een hogere energiedrempel had in vergelijking met AMANDA. Onze resultaten, vergeleken met een andere analyse van donkere materie uitgevoerd op AMANDA gegevens uit 2000 tot 2006 (geoptimaliseerd voor hoog-energetische neutrino's), presteren nog beter als we denken dat we een jaar minder gegevens hebben gebruikt.

De gefilterde dataset gebruikt in de laatste fase van deze analyse kan ook nog worden benut om een zoektocht naar andere WIMP kandidaten, zoals het lichtste Kaluza-Kleindeeltje in het kader van universele extra dimensies, na te streven.

Contents

Introduction

The twentieth-century started with a period of great discoveries for physics, both in the theoretical and experimental parts. Einstein's special and general relativity theories and quantum mechanics laid the foundations of a new way to understand Nature. In the experimental part, the discovery of the extraterrestrial nature of cosmic rays was fundamental to start to understand the phenomena of the deep Universe.

It was just by studying the orbital velocities of galaxies in clusters that Fritz Zwicky, in 1934, postulated the existence of an unknown kind of matter, the so-called *dark matter*, to account for evidence of missing mass in the orbital velocities. Over the time other observations have indicated the presence of dark matter in the Universe; these observations include the rotational speeds of galaxies, gravitational lensing of background objects by galaxy clusters such as the Bullet Cluster.

One possible scenario to explain the mismatch between the required mass, needed to supply the derived gravitational potential, and the observed mass is represented by the Cold Dark Matter (CDM) model, i.e. non-relativistic massive (GeV or heavier) particles produced in the Big Bang. This CDM makes up 23% of the energy density of the Universe, according to the WMAP measurements of the temperature anisotropies in the Cosmic Microwave Background, in combination with data on the Hubble expansion and the density fluctuations in the Universe. The thermally averaged cross section for scattering of the dark matter (DM) particles at the freeze-out temperature explains why the DM can only have weak and gravitational interactions. Therefore, the DM particles are generically called WIMPs, Weakly Interacting Massive Particles.

One of the most popular and widely studied WIMP candidates is the lightest supersymmetric neutralino $\widetilde{\chi}_1^0$ (or simply $\widetilde{\chi}$). In the Minimal Supersymmetric extension of the Standard Model (MSSM), where the multiplicative quantum number R-parity is conserved, the neutralino is the mixture of the superpartners of the B and W_3 gauge bosons and the neutral Higgs bosons, H_1^0 and H_2^0. The attractiveness of this candidate stems from the fact that it is electrically neutral, and thus neither absorbs nor emits light, is stable, and can only disappear via pair annihilation (it is a Majorana particle) or coannihilation with the next-to-lightest supersymmetric particle; therefore the neutralinos may have survived since the Big Bang. Consequently, relic neutralinos in the galactic halo will pass through massive bodies like the Sun or the Earth, where they can lose energy by scattering off nuclei. Over time, the neutralinos concentrate near the centres of these celestial bodies and annihilate producing Standard Model particles. The products of these annihilations will, in general, decay and produce neutrinos. The latter will be able to escape and would potentially be visible in a high energy

neutrino telescope at the surface of the Earth as an excess over the atmospheric neutrino flux.

We started this work with the aim of discovering indeed neutrinos from annihilations of neutralino dark matter particles in the Sun, using the data collected during six years (from 2001 to 2006) by the AMANDA (Antarctic Muon And Neutrino Detector Array) neutrino detector located at the South Pole, near the Amundsen-Scott station. The events collected during the live-time of the detector (about 7 billion in 812 effective days) were mostly muons induced by cosmic ray interactions in the atmosphere. Hence, we introduced a multivariate technique, the so-called Boosted Decision Trees (BDTs), to remove the bulk of this atmospheric muon background and in the same time preserve as much signal as possible. The performance of this method stood head and shoulders above a simple one-dimensional selection method, which was used in previous AMANDA analyses. After applying the BDT selection, we adopted a new and refined method to extract the signal strength from the remaining experimental data. This is actually the outcome of our analysis, which can be exploited to get an estimation of the neutrino-to-muon conversion rate, the neutralino annihilation rate in the Sun, the neutralino-induced muon flux at the detector and the neutralino-proton elastic scattering cross section.

This thesis is structured in the following way:

In Chapter 1 we will review the experimental evidences that give a strong hint for the existence of dark matter in the Universe, from galactic to cosmological scale. Then, we will focus our attention on the neutralino which, as we have already mentioned, is one of the best candidates for dark matter. At the end of the chapter we will discuss about the possible methods to detect neutralinos, via a direct or indirect way, and the status of the related experiments. We will also briefly discuss how collider results could constrain Supersymmetric parameter space, from which our neutralino composition depends.

In Chapter 2 we will discuss the underlying physics related to indirect detection of dark matter from the Sun using muon neutrinos as neutralino probe. We will discuss about its interaction with matter (e.g. ice), and subsequent production of charged particles like muons, which in turn will produce a Cherenkov light cone at a well-defined angle. At the end of the chapter we will present the atmospheric background concerning this search, i.e. atmospheric neutrinos and atmospheric muons.

In Chapter 3 we will talk about the AMANDA neutrino telescope, located 2000 m below the polar South Pole ice cap. We will describe its technology, calibration and operation. At the end of the chapter we will describe the ice properties, essential to understand the effect of scattering and absorption of the light in the medium, along with their measurements.

In the first part of Chapter 4 we will focus our attention on the experimental

data collected from 2001 to 2006 by AMANDA, which were used for this work. The data to be analysed were subject to a further check to verify the stability of the detector, which could be altered by some trigger issues, and in that case reject bad data. In the second part we explain how we performed our Monte Carlo simulations, from generator to detector simulation level, both for the neutralino signal and for the atmospheric background. We have chosen 14 different neutralino models to analyse; in more detail, 7 different neutralino masses[1], from 50 GeV to 5 TeV, which in turn annihilate in two extreme channels, yielding a soft and hard neutrino spectrum.

Chapter 5 consists of several sections related to the event processing and analysis. After giving some basic elements of event reconstruction, we pass to describe event filtering with the aim of removing badly reconstructed tracks, mostly due to the dominant atmospheric muon background. We have divided this event filtering in two main steps. The first, called *low level* filtering, required a condition on only one observed variable, usually the reconstructed theta angle, to select the event. The second step, the *high level* filtering, consisted in a more refined method which combined several variables to distinguish signal from background. This multivariate approach, as we have already mentioned, has been pursued through the classifier called Boosted Decision Trees (BDTs).

In Chapter 6 we describe a consolidated method, to extract the signal strength from the combined 2001-2006 final sample, which passed all the filtering steps. Next, we investigate all the possible sources of systematic uncertainties which could affect our final results. Then, through our signal estimation, we can infer some relevant physical quantities like the neutrino-to-muon conversion rate, the muon flux at the detector, the neutralino annihilation rate in the Sun, and the elastic neutralino-proton cross section.

In the last Chapter, number 7, we will summarise the work done in this search and put forward the more interesting parts of the analysis and possible lessons for future improvements.

As final remark, we have to point out that since 2009 AMANDA has finished its operation. Now the detection of extraterrestrial neutrinos is entirely in the hands of the AMANDA successor called IceCube, a kilometre-cube detector also located at the South Pole whose completion is foreseen in 2011, and of its low energy extension called DeepCore.

[1] Over this work we will use natural units adopted in particle physics, hence $c = 1$

...È dunque l'universo uno, infinito,
immobile; una è la possibilità assoluta,
uno l'atto, una la forma o anima, una la
materia o corpo, una la cosa, uno lo ente,
uno il massimo et ottimo;...

De la causa, principio et uno
Giordano Bruno (1548 - 1600)

1 The dark matter question

We start this work reviewing the experimental evidences that give a strong hint for the existence of dark matter in the Universe, from galactic to cosmological scale. Then, we will focus our attention on the neutralino which is one of the best candidates for dark matter. At the end of the chapter we will discuss about the possible methods to detect neutralinos, via a direct or indirect way, and the status of the related experiments. We will also briefly discuss how collider results could constrain Supersymmetric parameter space, from which our neutralino composition depends

1.1 Standard cosmology

In the past, cosmology was considered a pure speculative and theological doctrine. Nowadays it has become a scientific branch of knowledge which is approaching to a phase of full maturity, thanks to the progress of the observational techniques along with the development of the theoretical model, like the current *concordance model* ΛCDM (see next section).

Modern cosmology is based upon the so-called *Big Bang* scenario, in which the Universe evolved from a highly compressed state around 10^{10} years ago.
Hubble's law [1], discovered at the beginning of the past century, is the frame of the fundamental picture which describes our expanding Universe. Distant galaxies, indeed, move away from us with a recessional velocity v which is proportional to the intervening distance d:

$$v = H_0 \cdot d \tag{1.1}$$

where H_0 is the present value[1] of the Hubble parameter $H(t)$.

If we assume the isotropy and homogeneity[2] of the Universe, Einstein's field equations [3] can be solved, one of its components leading to the Friedmann

[1] A recent estimate [2] gives $H_0 = 70.5 \pm 1.3$ km s^{-1} Mpc^{-1}

[2] These hypotheses are corroborated by the observations of the Cosmic Microwave Background and by galaxy surveys (see sec. 1.2.3).

equation:

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G_N}{3}\rho_{tot} \tag{1.2}$$

where G_N is the Newton's constant, $a(t)$ is the so-called *scale factor*, the constant k describes the spatial curvature (being $k = -1,0,+1$) and ρ_{tot} is the total average energy density[1] of the Universe. Writing the Hubble parameter as $H(t) = \frac{\dot{a}(t)}{a(t)}$, we see from eq. 1.2 that the Universe is flat ($k = 0$) if the total average energy density is equal to the *critical density*

$$\rho_c \equiv \frac{3H^2}{8\pi G_N} \tag{1.3}$$

Introducing then a quantity:

$$\Omega = \sum_i \Omega_i \equiv \sum_i \frac{\rho_i}{\rho_c} \tag{1.4}$$

where Ω_i describes the abundance of a substance (matter, radiation or vacuum energy) of density ρ_i in units of ρ_c, eq. 1.2 can be written as follows:

$$\Omega - 1 = \frac{k}{H^2 a^2} \tag{1.5}$$

The values of k are consequently determined following the scheme below:

$$\begin{array}{ll} \Omega < 1 \Rightarrow k = -1 & \text{(open Universe)} \\ \Omega = 1 \Rightarrow k = 0 & \text{(flat Universe)} \\ \Omega > 1 \Rightarrow k = +1 & \text{(closed Universe)} \end{array}$$

The relic density of a generic particle species X (for instance neutralinos) can be expressed in terms of the critical density and in the dimensionless parameter $h = H_0/100$ km s^{-1}Mpc^{-1}[4]:

$$\Omega_X h^2 \approx \frac{3 \cdot 10^{-27} cm^3 s^{-1}}{\langle \sigma v \rangle} \tag{1.6}$$

where the thermal average of the annihilation cross section should be calculated taking into account, in the Boltzmann equation, the coannihilations with the next-to-lightest particle [5].

[1] $\rho_{tot} = \rho_m + \rho_{rad} + \rho_{vac}$, where ρ_m and ρ_{rad} are the energy densities in matter and radiation, and $\rho_{vac} = \frac{\Lambda}{8\pi G_N}$ is the vacuum energy density.

1.1.1 The concordance model: ΛCDM

ΛCDM or Lambda-CDM is an abbreviation for Lambda-Cold Dark Matter. It is referred to as the *concordance model* of big bang cosmology, since it endeavours to interpret cosmic microwave background observations, as well as large scale structure observations and supernovae observations of the accelerating expansion of the universe. Further, it is the simplest known model that is in general agreement with observed phenomena [6].

In the ΛCDM model, the Universe is spatially flat, homogeneous and isotropic on large scales. It is composed of atoms, dark matter, and dark energy, with nearly scale-invariant adiabatic Gaussian fluctuations.

The term Λ stands for the cosmological constant which is connected with a vacuum energy or dark energy, which accounts for the current accelerating expansion of the universe. Currently the fraction of the energy density $\Omega_\Lambda \simeq 0.74$, suggesting that 74% of the energy density of the present universe is dark energy [2, 7].

Cold dark matter is a form of matter necessary to account for gravitational effects observed in all astrophysical scale structures (see next sections) that cannot be interpreted by the quantity of observed matter. Dark matter is described as being cold (i.e. its velocity is non-relativistic at the epoch of radiation-matter equality), and possibly non-baryonic (consisting of matter other than protons and neutrons), interacting only weakly or gravitationally with each other and other particles. This component is currently estimated to constitute about 23% of the mass-energy density of the universe [2, 7].

The ΛCDM model has six primary parameters: physical baryon density Ω_b, physical dark matter density Ω_c, dark energy density Ω_Λ, scalar spectral index n_s, curvature fluctuation amplitude $\Delta^2_{\mathcal{R}}$ and reionisation optical depth τ. From these the other model values, including the Hubble constant H_0 and age of the universe t_0, can be derived. The optical depth to reionisation determines the redshift of reionisation. Information about the density fluctuations is determined by the amplitude of the primordial fluctuations (from cosmic inflation) and the spectral index, which measures how the fluctuations change with scale ($n_s = 1$ corresponds to a scale-invariant spectrum).

1.2 Evidence for dark matter

1.2.1 The Galactic scale

The observations of the rotation curves of galaxies (i.e. the circular velocities of stars and gas as a function of their distance from the galactic centre) represent the most strict evidence for the existence of dark matter on galactic scales. In

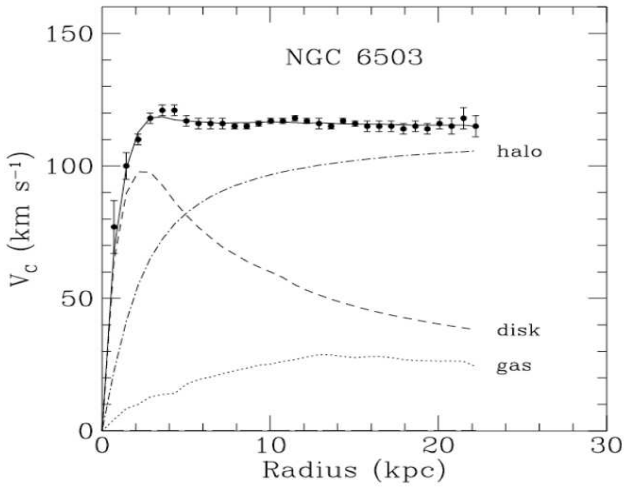


Figure 1.1: *Graph of the circular velocities of objects in NGC 6503 as a function of the distance to the centre. The different lines show the contributions of gas (dotted line), disk (dashed line) and dark matter (dash-dotted line). Figure taken from ref. [8]*

the Newtonian dynamics frame, the circular velocity is expressed by:

$$v(r) = \sqrt{\frac{G_N M(r)}{r}} \tag{1.7}$$

where $M(r) \equiv 4\pi \int \rho(r) r^2 dr$ and $\rho(r)$ is the mass density profile.
From eq. 1.7, the circular velocity should be falling as $\frac{1}{\sqrt{r}}$ beyond the edge of the visible disk; but what is observed, instead, is a flat behaviour of $v(r)$ at large distance which hints to the existence of a dark halo (see fig. 1.1) with $M(r) \propto r$ and $\rho \propto \frac{1}{r^2}$.

Some numerical simulations, the so called N-body simulations, have proposed a *universal* dark matter profile with the same shape for all masses, epoch and input power spectra [9, 10, 11]. The slope of the density profile should increase if we move from the centre of a galaxy to the external region; however, the exact value of the power-law index in the innermost galactic regions is still questionable as the various simulation groups ended up with different results for the spectral shape in those galactic regions.

The dark matter profile in the inner region of our galaxy, the Milky Way, is even more uncertain. Several observations suggest the presence of a Super Massive Black Hole (SMBH) in the centre of our galaxy [12, 13]; then the process of adiabatic accretion of dark matter on it would produce a "spike" in the dark matter density profile in this region [14].

1.2.1.1 The local density

The density of dark matter in the region of our solar system is a quantity which is substantially better known than the density near the galactic centre, and it is calculated by observing the rotation curves of our galaxy. This observation is difficult to perform from our location within the Milky Way; moreover to calculate with accuracy the dark matter profile we need to take into account the density distributions of the galactic bulge and disk.
Throughout this thesis, as canonical value for the local dark matter density we will use $\rho_0 = 0.3$ GeV cm^{-3}; for a detailed discussion about the local dark matter distribution, see refs. [4, 15].
Another quantity that is inferred from the observation of rotation curves, is the velocity distribution of the dark matter in the local region, which is typically described by its average velocity $\bar{v} = \langle v^2 \rangle^{1/2} \cong 270$ km s^{-1}.

These two quantities, ρ_0 and $\bar{v}$, are crucial to both the direct and indirect methods of dark matter searches.

1.2.2 The Galaxy Clusters scale

The mass of a galaxy cluster can be inferred via several techniques, like the application of the virial theorem to the distribution of the radial velocities. Fritz Zwicky, in 1933, by measuring the velocity dispersion of galaxies in the Coma cluster, derived that the observed mass-to-light ratio exceeded the ratio in the solar neighbourhood by two orders of magnitude [16]; this was the first strong indication of the existence of a large amount of non-luminous (or dark) matter.

Gravitational lensing is one of the confirmed predictions of Einstein's theory of general relativity; in presence of an intense gravitational field, light propagates along geodesics which deviate from straight lines. The shape of the potential well, and thus the mass of the cluster, can be then drawn from the distortion of the images of background objects due to the gravitational field produced by the cluster. The cluster of galaxies 1E0657-56 , called Bullet cluster, is one of the hottest, most X-ray luminous cluster known and it is still subject to intense ongoing studies. Chandra X-ray Observatory showed that the cluster is a supersonic merger in the plane of the sky [17]. Due to its unique geometry and physical state, the Bullet cluster is the best known system to test the dark matter hypothesis [18]. The combined weak and strong lensing mass reconstructions show two substructures that are offset with respect to the baryon distribution (hot gas) observed in X-ray by the Chandra Observatory (see fig. 1.2). In contrast, the cluster galaxy population follows the dark matter distribution.

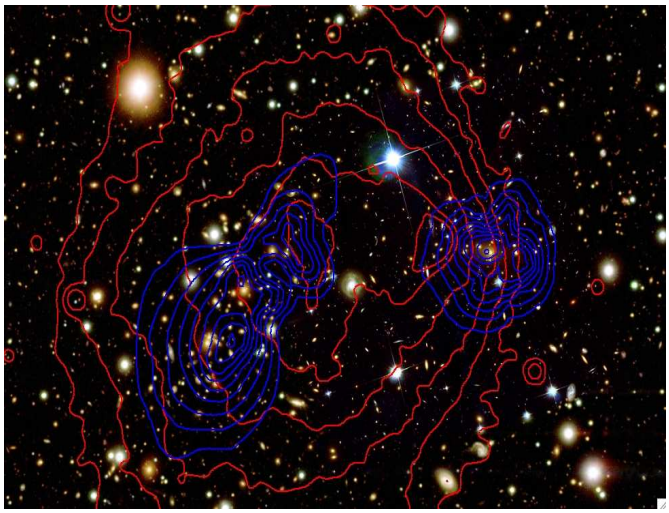


Figure 1.2: *The joined strong and weak lensing mass reconstruction (blue contours) of the cluster 1E0657-56. The X-ray emission (shown in red contours) is clearly offset from the total mass distribution. Picture taken from [19]*

1.2.3 The Cosmological scale

In the previous sections we have shown how, in galaxy and in cluster of galaxies scales, the need of dark matter appears to be compelling to explain some experimental observations; however, through these observations, we cannot determine the total amount of the dark matter in the Universe.

The existence of the Cosmic Microwave Background (CMB) [20, 21], which was predicted as relic radiation from the early Universe, was confirmed in 1965. Further studies have established that the CMB is isotropic at the 10^{-5} level, and that it follows the spectrum of a black body with a temperature $T = 2.73$ K. Through the study of CMB anisotropies, accurate tests of cosmological models can be performed, which, in turn, can put stringent constraints on cosmological parameters.

The observed temperature anisotropies can be decomposed in spherical harmonics:

$$\frac{\Delta T(\hat{\boldsymbol{n}})}{T} = \sum_{lm} a_{lm} Y_{lm}(\hat{\boldsymbol{n}}) \tag{1.8}$$

and if the temperature fluctuations are assumed to be Gaussian, the information included in the CMB maps can be compressed into the power spectrum. Given a cosmological model with a fixed number of parameters, the best-fit parameters are extracted from the peak of the N-dimensional likelihood surface (see fig. 1.3). The simple six-parameter power-law ΛCDM cosmological model fits not only the

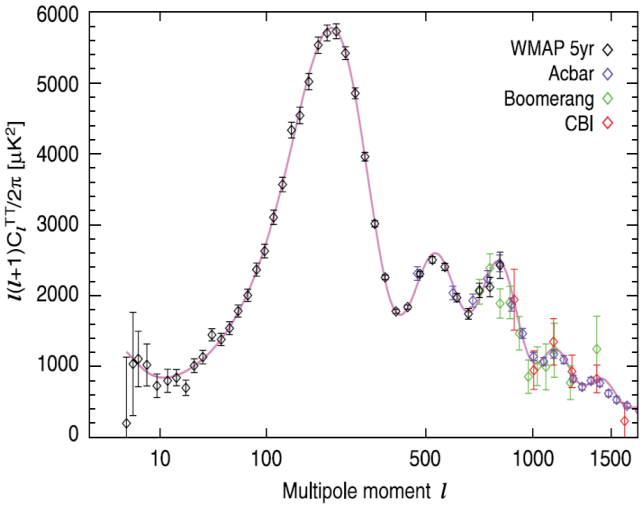


Figure 1.3: *WMAP five-year angular power spectrum along with recent results from other experiments. The red curve is the best-fit ΛCDM model to the WMAP data, which agrees well with all datasets when extrapolated to higher ℓ. Picture taken from [24].*

Wilkinson Microwave Anisotropy Probe (WMAP) data, but also a wide range of astronomical data [22, 23].

The WMAP has put the most stringent constraints on the abundances of baryons ($\Omega_b h^2$) and matter ($\Omega_m h^2$) in the Universe. Indeed, from the analysis of WMAP five years data, combined with the distance measurements from the Type Ia supernovae (SN) and the Baryon Acoustic Oscillations (BAO) in the distribution of galaxies, the following values[1] are found [2]:

$$\Omega_b h^2 = 0.02267^{+0.00058}_{-0.00059} \qquad \Omega_m h^2 = 0.1358^{+0.0037}_{-0.0036} \tag{1.9}$$

The obtained value of $\Omega_b h^2$ is consistent with predictions from Bing Bang nucleosynthesis (BBN) [25].

WMAP has measured the basic parameters of ΛCDM cosmology to high precision: with the WMAP 5-year data alone, a dark matter density of 21.4%, and an atoms density of 4.4% were found.

1.3 Dark matter candidates

In the previous sections, we have shown that dark matter is compelling at all observed astrophysical scales. Among the various dark matter candidates

[1] For other relevant cosmological parameters, see the reference [2].

(see ref. [5] for a general review), we focus our attention on those which are (in principle) detectable with present or near-future technology.

One possible excellent candidate, the neutrino, which has the "undisputed virtue of being known to exist" [26], is simply not abundant enough to be the dominant component of dark matter, since current upper bounds on its mass from particle physics and from cosmology give as relic density $\Omega_\nu \lesssim 0.012$ [7]. Besides, due to their tiny mass, neutrinos constitute what is called Hot Dark Matter (HDM), as they decouple from the radiation equilibrium in the early Universe at relativistic energies and move with high velocities. Hot dark matter cannot account for structure formation in the Universe, hence most DM should be "cold", i.e. non-relativistic at the onset of galaxy formation [27].

From the measurement of the baryon relic density (see previous section), we infer that we cannot account for all of the dark matter with cold massive objects of ordinary matter (such matter clumps in galactic halos are called MACHOs, Massive Compact Halo Objects). Hence, the particle candidates for Cold Dark Matter (CDM) that are best motivated remain supersymmetric non-baryonic Weakly Interacting Massive Particles (WIMPs), and we will focus our search on the most widely studied candidate: the supersymmetric neutralino (see next sections).

Other non-baryonic CDM noteworthy candidates are WIMPs in universal extra dimension models and the axion (see sec. 1.3.3).

1.3.1 Supersymmetry

From the Standard Model of particle physics we know that fermions are the constituents of matter while bosons are the mediators of interactions. At this point one could ask if a (super)symmetry exists which relates bosons and fermions thus providing a frame which unifies matter and interactions. Another argument in favour of a supersymmetric model is its role in understanding the so-called *hierarchy problem*, which is linked to the huge difference between the electroweak and Planck energy scales[1]. One possible solution to this problem is to postulate the existence of new particles with similar masses but with spin different by one-half. The corresponding algebra of supersymmetry (SUSY) naturally ensures the existence of new particles with the required properties: to all of the SM particles are associated superpartners with the same mass but opposite spin-type.
The new generators of SUSY change fermions into bosons and *vice versa*:

$$Q|\text{fermion}\rangle = |\text{boson}\rangle; \qquad Q|\text{boson}\rangle = |\text{fermion}\rangle \tag{1.10}$$

The operators Q, since their fermionic nature, must carry spin 1/2, which hints that SUSY must be a spacetime symmetry.

[1] The Planck scale is the energy scale at which gravitational interactions become comparable in strength to electroweak and strong interactions, roughly $M_P \sim (G_N)^{-\frac{1}{2}} \sim 10^{19}$ GeV.

The SUSY model is also interesting because by introducing supersymmetry at the TeV scale the gauge couplings unify at a scale $M_U \sim 2 \cdot 10^{16}$ GeV [28], which is a strong hint in favour of a Grand Unified Theory (GUT).

1.3.1.1 The minimal supersymmetric extension of the SM: MSSM

The minimal supersymmetric extension of the SM (MSSM) contains the smallest possible field content necessary to generate all the SM fields. In table 1.1 there is a sketch of the resulting particle content of MSSM.
One fundamental component of the MSSM is a conserved multiplicative quantum number, the R-parity[1], which is defined as:

$$R \equiv (-1)^{3B+L+2S} \tag{1.11}$$

where B is the baryon number, L the lepton number and S the spin of the particle. This implies that R = +1 for SM particles and R = -1 for SUSY particles. This means that SUSY particles can only be created or annihilated in pairs in reactions of SM particles. Consequently, a single SUSY particle can only decay into final states containing an odd number of SUSY particles (plus SM particles). In particular the lightest supersymmetric particle (LSP) is stable, since there is no kinematically allowed state with negative R-parity which it can decay to. Therefore, the LSP can be considered as an excellent dark matter candidate [30].

Several observations have put some constrains on the nature of the LSP, for instance it cannot have an electric charge different from zero or colour, otherwise it would have condensed with baryonic matter to generate heavy isotopes.

1.3.2 The lightest neutralino as dark matter candidate

In the MSSM the four Majorana fermionic mass eigenstates called neutralinos, are the result of the admixture of the superpartners of B, W_3 gauge bosons (or the photon and the Z, equivalently) and the neutral Higgs bosons, H_1^0 and H_2^0, which are called binos ($\widetilde{B}$), winos ($\widetilde{W}_3$) and higgsinos ($\widetilde{H}_1^0$ and $\widetilde{H}_2^0$) respectively. The four neutralinos are typically labelled $\widetilde{\chi}_1^0$, $\widetilde{\chi}_2^0$, $\widetilde{\chi}_3^0$ and $\widetilde{\chi}_4^0$, sorted with increasing mass. In this work we focus our attention on the lightest of the neutralinos, $\widetilde{\chi}_1^0$, which we can simply denote as χ and refer to as *the* neutralino.

The neutralino is therefore a linear combination of $\widetilde{B}$, $\widetilde{W}_3$ $\widetilde{H}_1^0$ and $\widetilde{H}_2^0$:

$$\chi = N_{11}\widetilde{B} + N_{12}\widetilde{W}_3 + N_{13}\widetilde{H}_1^0 + N_{14}\widetilde{H}_2^0 \tag{1.12}$$

We can define the *gaugino fraction*, f_G, and the *higgsino* fraction, f_H, as:

$$f_G = N_{11}^2 + N_{12}^2 \tag{1.13}$$

[1] R-parity was actually first introduced to suppress the rate of proton decay [29].

Normal particles/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 1.1: *Standard model particles and their superpartners in the MSSM (adapted from [31]).*

and

$$f_H = N_{13}^2 + N_{14}^2 \tag{1.14}$$

The neutralino is an excellent candidate for Cold Dark Matter, which is one of the ingredients of ΛCDM cosmological model, since it is a non-relativistic particle produced as a thermal relic of Big Bang. In the stage when the Universe, in addition to cooling, is also expanding and thus becoming so large, the neutralinos are so dilute that they cannot self-annihilate. Then the neutralinos "freeze out" approaching their relic density $\Omega_\chi h^2$. From equation 1.6 and from the WMAP results (see values in 1.9) we infer that the total cross section at the freeze-out temperature is typical for weak interactions, which is simply too small to have large energy losses when falling towards the center of galaxies, and thus clustering like baryons do. Therefore the neutralinos are generically called WIMPs, Weakly Interacting Massive Particles.

1.3.3 Alternative candidates

In this section we will give a brief report on two of the best motivated and studied alternative candidates: the axion and the lightest Kaluza-Klein particle.

Axions are particles introduced to solve the CP violation problem in the strong interactions [32, 33]. Several observations have constrained the mass of the axion to be very light, $m_a < 0.01$ eV [34], and since it couples very weakly to ordinary matter [35], it was never in thermal equilibrium in the early Universe, thus behav-

ing as CDM today. A large part of the available parameter space has, however, been ruled out by searches for axion conversion to photons in magnetic fields [36].

The suggestive and original idea of the possible existence of extra-dimensions at high energy scales, was proposed in 1921 by Kaluza [37] in an attempt to unify electromagnetism and gravity. More recently, extra-dimensional models were proposed, called Universal Extra Dimensions (UED) [38], where all the particles propagate in flat, compact extra dimensions. The simplest UED model predicts one extra dimension of size R compactified on a circle, and, based upon this, all the fields propagating in the bulk have their momentum quantised (in units of $p^2 \sim 1/R^2$). Thus, for each bulk field, a set of so-called Kaluza-Klein (KK) states appear as a Fourier series (called a tower) with masses $m_n = n/R$, where n is the mode number. The UED model provides a viable dark matter candidate as a consequence of the conservation of momentum in higher dimensional space which leads to the conservation of KK number. All odd-level KK particles are charged under this symmetry called KK-parity, thus ensuring that the lightest (first level) KK state, the Lightest Kaluza-Klein Particle (LKP), is stable and a possible dark matter candidate [39].
The LKP is associated with the first KK excitation of the photon (or the first KK excitation of the hypercharge gauge boson), labelled as B^1. The regions of parameter space, to account for the correct relic density of the B^1, have been investigated by several groups [40, 41, 42]; these studies give a LKP mass range of 600 GeV $\lesssim m_{B^1} \lesssim$ 1400 GeV.

1.3.4 Alternative theory, changing the law of gravity?

The ΛCDM cosmological model has achieved a remarkable success to explain and to predict different data sets at large scales. However, several observations at galactic scales [43], which show that the baryons dominate kinematically in the inner parts of rotation curves, are in conflict with the predicted "cuspy" halo by ΛCDM model. The Modified Newtonian Dynamics (MOND) is an alternate theory [44] which explains the mass discrepancy with a modification of Newtonian gravity and not requiring any dark matter halo. The observed kinematics at the low gravitational accelerations in the outer region of the galaxies, is described in MOND theory assuming that below a certain acceleration, $a_0 \sim 10^{-8}$cm s^{-2}, Newtonian gravity is no longer valid. The actual gravitational acceleration g_{MOND} is related with the Newtonian one g_N as following:

$$g_{MOND} = \frac{g_N}{\mu(g_{MOND}/a_0)} \tag{1.15}$$

where $\mu(x)$ is an interpolation function which regulates the transition between Newtonian regime and deep MOND regime (for $g_{MOND} \gg a_0$, $\mu(x) \longrightarrow 1$ and for $g_{MOND} \ll a_0$, $\mu(x) \longrightarrow g_{MOND}/a_0$).

Although the predictions of MOND are very successful at galactic scale, this phenomenological theory comes upon serious difficulties when it is applied to larger scales. A large amount of dark matter is still necessary to interpret correctly the data to fully explain the lensing and dynamics of clusters of galaxies [45].

1.4 Detection of dark matter

What we know about dark matter, currently, stems only from its macroscopic gravitational effects. The way to understand the distribution of dark matter at the galactic scale and smaller, is to try to detect dark matter particles individually. We have to point out that current experiments, which aim to detect directly (sec. 1.4.2) or indirectly (sec. 1.4.3) WIMPs as dark matter, need more information to disentangle their nature, either supersymmetric or extra-dimension, since candidates like LSP or LKP have the same signature.
Collider experiments (sec. 1.4.1) are probing significant regions of the parameter space of these hypothetical particles, but are not able to set conclusively their stability or relic density. Conversely, a positive astrophysical detection of dark matter would provide remarkable information about the physics "beyond the Standard Model".

Hence, only by combining all the contributions, that came from these different experimental approaches and cosmological observations, we hope to shed light on the dark matter mystery.

1.4.1 Collider constraints

The Large Hadron Collider (LHC) has produced its first pp collision at CERN in 2009 at $\sqrt{s} = 900$ GeV, reaching 2.36 TeV at the end of the same year [46]; then at the beginning of 2010, the energy has been increased to 7 TeV.
After a long shut-down foreseen in 2012, to prepare design-energy running, in 2013 it will reach energies up to 14 TeV, with a predicted luminosity of 10^{33} cm^{-2}s^{-1}. Given this energy, and the requirement that WIMPs have mass $\sim m_{\text{weak}}$, then particles like the weakly interacting neutralino will be almost certainly produced; but unfortunately direct production of $\chi\chi$ pairs is invisible. Therefore the indirect production is the only feasible search at the LHC, for instance SUSY pair production of squarks and gluinos followed by their decay through some cascade chain, and ending up to neutralinos. The existence of these latter, which escape the detector, is inferred then by the missing energy $\not{E}_T$ in the transverse plane.
The observation of missing particles in collider experiments, however, is not sufficient to claim evidence, since this observation will tell us only that the produced particle was stable enough to exit the detector, typically with a lifetime $\tau \gtrsim 10^{-7}$s, too far from $\tau \gtrsim 10^{17}$s required for dark matter. Hence, the WIMP's thermal

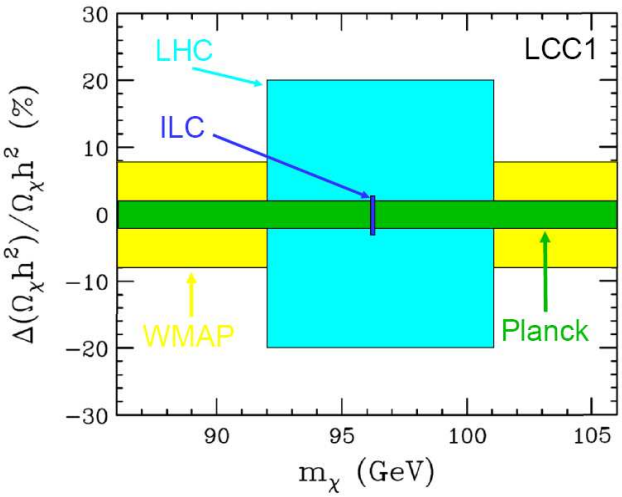


Figure 1.4: *Constraints from the LHC and the ILC in the LCC1 frame, in conjunction with WMAP and Planck experiments. WMAP and Planck measure Ω_χ but are insensitive to m_χ, while collider experiments bound both. Picture taken from [49].*

relic density, calculated through the collider's constraints, should be consistent with the cosmologically observed density to prove that the particle produced at colliders is indeed the dark matter.

The International Linear Collider (ILC) is a proposed high energy e^+e^- collider and it is designed to operate with $\sqrt{s}$ from 250 GeV to 500 GeV, which could be upgradable to $\sim$ 1 TeV [47]. The WIMP is not directly observable in lepton collider experiments (as at hadron colliders), therefore its mass has to be inferred through some indirect methods. One possible and precise technique is the scan for the threshold of the $e^+e^- \to X\bar{X}$ pair production process tuning the centre of mass energy. The mass of the SUSY particle X and its width can be extracted from a fit to the signal event yield as a function of $\sqrt{s}$ [48]. Using a threshold scan at $\sqrt{s} \sim 2M_X$ to determine the mass of X, the mass relation between the lightest neutralino and X can be settled by running at higher centre of mass energy.

Currently, the strongest lower limit on SUSY particles were set by the LEP e^+e^- collider at CERN, which ran at $\sqrt{s} \leq 208$ GeV. The LEP experiments restrict electrically charged SUSY particles to have a mass above 100 GeV [50] and a lower bound on the lightest neutralino mass, $m_\chi \geq 47$ GeV, was set by combining searches for sleptons, charginos and Higgs bosons, in the mSUGRA[1]

[1]This is a restricted subspace of the MSSM (called also constrained MSSM), where the soft-SUSY breaking parameters are unified at the grand unification scale.

scenario [51].

The possible performances of the LHC on dark matter scenario have been studied by different groups [52, 53, 54]; the results of one exemplary study in the LCC1[1] frame are shown in figure 1.4, along with the constraints from WMAP and its successor Planck [55]. The LHC can determine so many of the SUSY model parameters that the neutralino mass can be evaluated with ± 5 GeV uncertainty, and the relic density with a precision of ± 20%. The ILC could improve these constraints on mass and relic density to ± 50 MeV and ± 3%, respectively.

1.4.2 Direct searches

If the galaxy is filled with WIMPs, then plenty of them should pass through the Earth and eventually interact with matter. The basic idea underlying the dark matter direct search experiments, is to record the recoil energy of nuclei as WIMPs scatter off them [56]. The density and the velocity distributions of WIMPs in the solar neighbourhood, and the WIMP-nucleon cross section are the main ingredients to calculate the signal rate, R, which is approximately given by:

$$R \approx \sum_i N_i n_\chi \langle \sigma_{i\chi} \rangle \tag{1.16}$$

where the index i runs over nuclei species in the detector, $N_i = \frac{\text{Detector mass of species } i}{\text{Atomic mass of species } i}$ is the number of target nuclei in the detector of species i, $n_\chi \equiv \frac{\text{WIMP energy density}}{\text{WIMP mass}}$ is the local WIMP density and $\langle \sigma_{i\chi} \rangle$ is the averaged cross section[2] for the scattering of WIMPs off nuclei of species i.

Elastic scattering of WIMPs off nuclei is usually described in the context of two classes of couplings. One class is the axial-vector or spin-dependent (SD) interaction which results from coupling to the spin content of a nucleon. The cross sections for SD scattering are proportional to $J(J+1)$ rather than the number of nucleons, hence the gain is little if heavier target nuclei are used.
The other class is the scalar or spin-independent (SI) interaction whose cross section increases rapidly with the mass of the target nuclei, and normally dominates over the SD scattering in the experiments which use heavy atoms as target.

Several experiments are currently operating; among others we cite XENON100 [58], ZEPLIN III [59], which use both scintillation and ionisation techniques, EDELWEISS II [60] and CDMS II [61] which use both ionisation and photon techniques. The DAMA/LIBRA experiment [62], which uses the scintillation technique, attempts to separate WIMP signature from background by looking

[1] This is a particular mSUGRA model with parameters:
$(m_0, M_{1/2}, \tan\beta, A_0, \text{sign}(\mu))$ = (100 GeV, 250 GeV, 10, -100 GeV, +).

[2] The cross section is averaged over the relative WIMP velocity with respect to the detector.

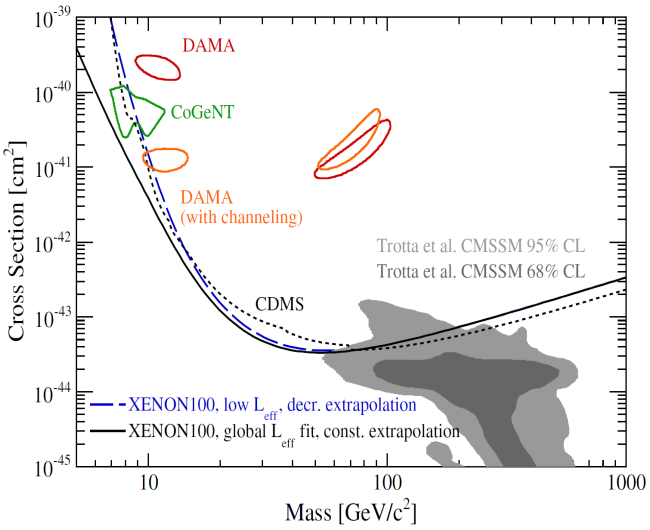


Figure 1.5: *The 90% confidence upper limits on the spin-independent elastic WIMP-nucleon cross section (solid and long dashed lines) from XENON100, together with the best limit to date from CDMS (dotted line), recalculated assuming an escape velocity of 544 km/s and $v_0 = 220$ km/s. Expectations from a constrained MSSM model [57], and the areas (90% C.L.) favoured by CoGeNT (green) and DAMA (red/orange) are also shown. Picture taken from [58].*

at the annual modulation in their rate; this effect is due to the Earth's annual motion around the Sun. The DAMA/LIBRA experiment claimed to have found a signal in annual modulation with period and maximum at the expected values[1], but other experiments have explored the parameter space favoured by DAMA without finding any evidence of dark matter. Some solutions to this diatribe have been proposed, like the possible inelastic scattering of the dark matter [63]; channelling effect [64, 65] and astrophysics arguments [66] have been put forward in pursuit of explaining the DAMA results with elastic scattering, and without violating other constraints. 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 region in which DAMA has observed their annual modulation signal [67].

The current limits on spin-independent scattering, from some direct detection experiments, are summarised in fig. 1.5; these results assume an isothermal WIMP halo with $v_0 = 220$ km/s, a local dark matter density $\rho_\chi = 0.3$ GeV/cm^3,

[1]The Earth should be crossed by a larger flux of DM particles around roughly June 2$^{\text{nd}}$ (when its orbital speed is summed to the one of the Solar System with respect to the Galaxy).

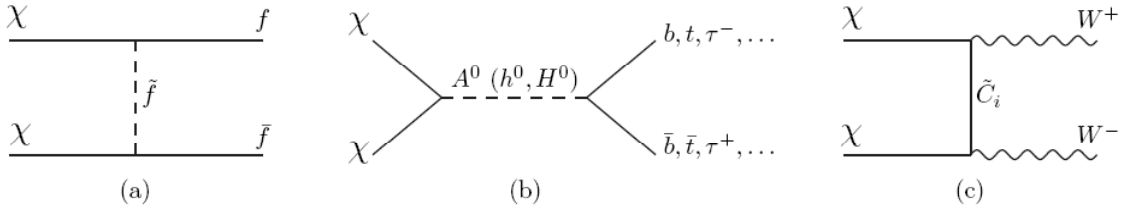


Figure 1.6: *Contributions to the annihilation cross-section for neutralino dark matter from (a) slepton and squark exchange, (b) near-resonant annihilation through a Higgs boson and (c) chargino exchange. Adapted from [70].*

and escape velocity 544 km/s [58]. The XENON100 limit, below 20 GeV, constrains the interpretation of the CoGeNT and DAMA signals as being due to light mass WIMPs.

The current direct detection experiments which are looking at the spin-dependent scattering, such as CDMS, COUPP [68] and KIMS [69], give less performances in term of probing the core of SUSY parameter space.

1.4.3 Indirect searches

Dark matter can be detected indirectly by observing the radiation produced in dark matter annihilations. Dark matter indeed continues to annihilate after freeze out in high density regions, and the flux of the radiation is proportional to the annihilation rate which in turn depends on the square root of the dark matter density: $\Gamma_A \propto \rho_\chi^2$. Hence, the regions which we should look for significant fluxes, since large dark matter density accumulate there, are the galactic centre or astrophysical objects like the Sun, the Earth or the Milky Way halo. It could be also possible to observe annihilation radiation from galaxies outside the Milky Way (even if they are far more distant than the galactic centre, the emitting region is much larger) or from dwarf galaxies within the Milky Way.

Some of the neutralino annihilation channels are:

$$\chi\chi \longrightarrow \begin{cases} q\bar{q}, l^+l^-, W^+W^-, Z^0Z^0, ... & \text{(tree level)} \\ gg, Z\gamma, \gamma\gamma, ... & \text{(one loop level)} \end{cases} \tag{1.17}$$

and three of these mechanisms are illustrated in fig. 1.6.

The fluxes of positrons and electrons are possible radiations produced by neutralino annihilation in the galactic halo, and some indirect detection experiments have reported anomalies in the measured flux which could be interpreted as evidence for dark matter. The most significant recent observations come from the PAMELA [71], ATIC [72] and Fermi LAT [73] collaborations, which have detected positrons and electrons with energies between 10 GeV and 1 TeV. These

data, displayed in fig. 1.7, clearly show an excess above the expected background, which was estimated by the code GALPROP [74].
However, as we can see from the picture, the ATIC "bump" is not confirmed by the Fermi LAT experiment, which has much higher statistics. These latter data could be explained by modifying the spectral index of the cosmic ray background [75], even if this explanation worsens the PAMELA discrepancy which in turn seems to be consistent with expectation from pulsar radiation [76, 77], and may have also other astrophysical explanations [78].

The possibility that the positron excess has an origin in dark matter annihilation has been widely reviewed as well; the energies of the excess (around m_{weak}) are as expected for the neutralino, but the positron fluxes result far larger to accommodate the neutralino thermal relic density. Thus the annihilation cross section should be enhanced by two or three orders of magnitude to explain the positron data [79]. Some exotic dark matter explanations have been put forward: the annihilation cross section may be boosted by resonances from states with mass $2m_\chi$ [80] or by the Sommerfeld enhancement factor [81], or the dark matter may be produced by decays and not by thermal freeze out [82].

Collecting more data from Fermi or from the AMS experiment (an anti-matter detector to be set on the International Space Station [83]) could maybe disentangle the scenarios for the positron excess. Searches for gamma rays by space-based experiments like Fermi and AMS, or by ground-based atmospheric Cherenkov telescopes (e.g. MAGIC [84], VERITAS [85] and H.E.S.S. [86]) are also promising. The most evident gamma ray signal would be photons from $\chi\chi \to \gamma\gamma$ which are mono-energetic, but since the neutralino is a chargeless particle, these processes are loop-induced and highly suppressed. More usually gamma rays are produced when neutralinos annihilate to other particles which radiate photons, leading to a smooth gamma ray energy spectrum; however, an advantage from photons is that they point back to their source providing a powerful signature.

The space-based telescope EGRET [87], on-board the Compton Gamma Ray Observatory (CGRO), has measured an excess of gamma rays in the region of the galactic centre which could be interpreted as the product of dark matter annihilation [88]. However, preliminary studies with Fermi LAT are consistent with the expected background and suggest that the EGRET excess may have an instrumental origin [89].

Searches for neutrinos are really unique among the indirect detection techniques and, given some assumptions, they can be competitive with direct searches. The underlying idea of neutrino searches revolves around the following: if neutralinos in the galactic halo pass through massive bodies like the Sun (or the Earth, or the galactic centre), they can lose energy by scattering off nuclei and eventually be gravitationally trapped. Once captured, they settle to the centre where their density increases greatly, thus they annihilate producing SM particles

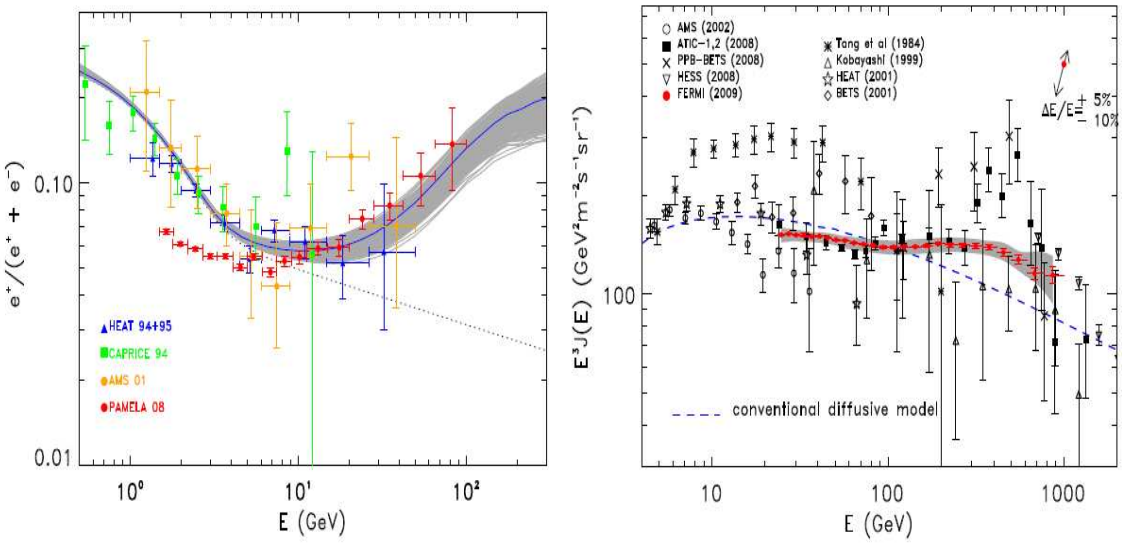


Figure 1.7: *The cosmic positron fraction measured by PAMELA and other experiments is showed on the left, along with predictions of pulsars with different parameters (grey contours); the discrepancies at energies below 10 GeV are claimed to be due to the solar modulation. On the right, the plot of the total $e^+ + e^-$ flux as measured by ATIC, Fermi and other experiments; the dashed band represent the background prediction from GALPROP.*

which, in general, decay and produce neutrinos. These latter are not immediately absorbed, unlike the other produced particles, thus escaping from the centre and travelling to the surface of the Earth, where they may be detected through the conversion to charged leptons.
We will discuss in more details about this technique in the next chapter, since the subject of this thesis is the search for neutralino dark matter using six year of data taken by the AMANDA-II neutrino telescope.

Liebe Radioaktive Damen und Herren,

...Nämlich die Möglichkeit, es könnten elektrisch neutrale Teilchen, die ich Neutronen nennen will, in den Kernen existieren,...

Letter to a group of physicists meeting in Tübingen
Wolfgang Ernst Pauli (1900-1958)

2 Neutrino as neutralino probe

In this chapter we will discuss the underlying physics related to indirect detection of dark matter from the Sun using muon neutrinos as neutralino probe. We will discuss about its interaction with matter (e.g. ice), and subsequent production of charged particles like muons, which in turn will produce a Cherenkov light cone at a well-defined angle. At the end of the chapter we will present the atmospheric background concerning this search, i.e. atmospheric neutrinos and atmospheric muons.

2.1 Neutrinos from neutralino annihilation in the Sun: capture and annihilation rate

The neutralino capture rate in the Sun can be approximated by the following expression [90]:

$$C_\odot \simeq 3.35\times10^{20}\mathrm{s}^{-1}\left(\frac{\rho_\chi}{0.3\,\mathrm{GeV/cm^3}}\right)\left(\frac{270\,\mathrm{km/s}}{\bar{v}_\chi}\right)^3 \times\left(\frac{\sigma^{SD}_{\chi H}+\sigma^{SI}_{\chi H}+0.07\sigma^{SI}_{\chi He}}{10^{-6}\,\mathrm{pb}}\right)\left(\frac{100\,\mathrm{GeV}}{m_\chi}\right)^2, \tag{2.1}$$

where ρ_χ is the local neutralino density, $\bar{v}_\chi$ is the local neutralino velocity dispersion and m_χ the neutralino mass. The neutralino loses its energy through spin-dependent (SD) and spin-independent (SI) elastic scattering with hydrogen nuclei and, in less measure, through spin-independent elastic scatterring with helium nuclei. The capture rate suppression is regulated by two factors which depend on the neutralino mass; one of these is the kinematic suppression of neutralinos much heavier than the target nuclei, and the second factor is the number density of the neutralinos ($n_\chi \propto 1/m_\chi$).
The evolution equation for N neutralinos in the Sun is given by

$$\frac{dN}{dt} = C_\odot - C_A N^2 - C_E N \tag{2.2}$$

where in addition to the capture rate $C_\odot$, we have twice the annihilation rate ($\Gamma_A = \frac{1}{2} C_A N^2$) and the last term which is the neutralino thermal evaporation. This last one can be safely neglected in our analysis (indeed it holds for neutralinos below 10 GeV [91]), and if we solve equ. 2.2 for the annihilation rate Γ_A we obtain:

$$\Gamma_A = \frac{1}{2} C_\odot \tanh^2 \frac{t}{\tau} \tag{2.3}$$

where $\tau = (C_\odot C_A)^{-\frac{1}{2}}$ is the time scale for capture and annihilation equilibrium to occur. In the Sun, equilibrium will have occurred for many neutralino models and the annihilation rate is at "full strength", $\Gamma_A \simeq \frac{1}{2} C_\odot$. Indeed, this occurs when $\sqrt{C_\odot C_A} t_\odot \gg 1$, where $t_\odot \simeq 4.5 \times 10^9$ years is the age of the solar system. Hence, at the equilibrium, the annihilation rate (and corresponding neutrino flux and event rate) does not depend on the neutralino annihilation cross section, but it is directly proportional to the capture rate.

The neutralinos can annihilate to heavy quarks, tau leptons, gauge bosons and Higgs bosons thus producing neutrinos in the subsequent decay. These neutrinos can escape from the centre of the Sun and travel to the Earth where they can be eventually detected. The neutrino differential flux at the Earth from neutralinos annihilating in the Sun is given by

$$\frac{d\Phi}{dE_\nu} = \frac{\Gamma_A}{4\pi D_\odot^2} \sum_X B_X \left(\frac{dN_\nu}{dE_\nu} \right)_X \tag{2.4}$$

where $D_\odot$ is the Earth-Sun distance, B_X is the branching ratio of annihilation channel X and $\left(\frac{dN_\nu}{dE_\nu}\right)_X$ its differential neutrino spectrum.
The annihilation channel $\chi\chi \to \nu\bar{\nu}$ is strongly suppressed by the tiny neutrino mass, therefore neutrino fluxes come from decays of primary annihilation products, with a mean energy $\langle E_\nu \rangle \simeq \frac{m_\chi}{2}$ to $\frac{m_\chi}{3}$. The most energetic spectra, referred as "hard", come from neutralino annihilations into W^+W^-, ZZ, $\tau^+\tau^-$, while the less energetic ones, referred as "soft", come from $b\bar{b}$. The reason is because in hard channels, neutrinos are produced directly from the neutralino decay products. The softest spectra come from the quark channels where most neutrinos are produced indirectly in decays of hadrons created in the quark jets. Due to $\nu_\mu - \nu_\tau$ vacuum oscillations, the muon neutrino flux observed at the Earth is the average of the ν_μ and ν_τ components. In fig. 2.1 are shown the muon neutrino spectra[1] at production in the Sun, after propagation to the Sun's surface and at the Earth's surface for two particular neutralino masses [92].

[1] The energy spectra are a combination of several annihilation channels: $W^+W^-, ZZ, Zh, t\bar{t}$.

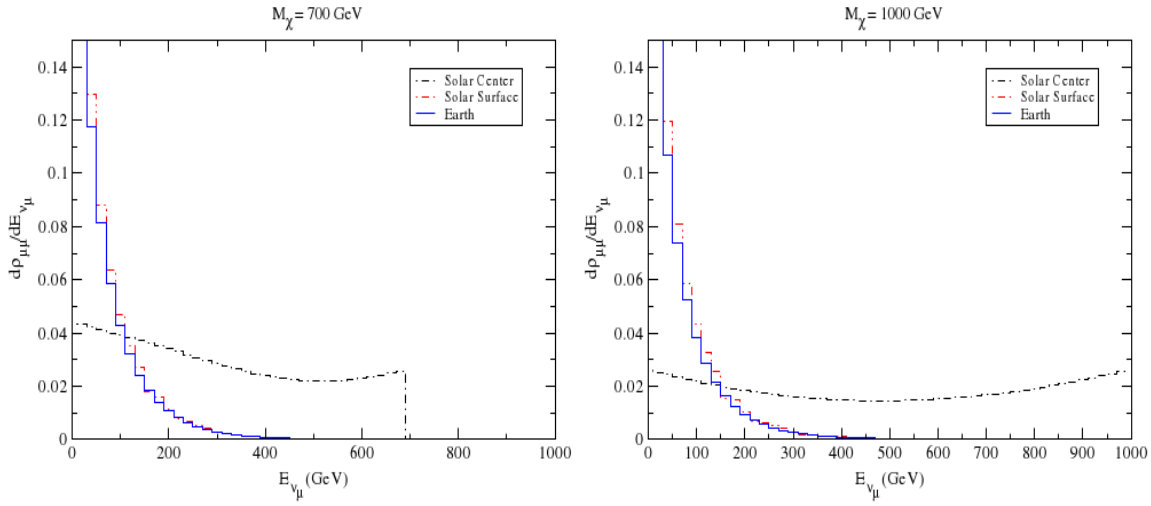


Figure 2.1: *The neutrino energy spectra at production in the Sun, after propagation to the Sun's surface, and at the Earth's surface (picture adapted from [92]).*

2.2 Neutrino detection

The first proposal, about the possibility of detecting high-energy cosmic neutrinos in underground experiments, was submitted to the international scientific community around the '60s [93]. Such equipments should permit the identification of the neutrino source and therefore they are usually named *neutrino telescopes*. In order to detect the characteristic signatures of high energy neutrino interactions, the proposal foresaw the instrumentation of large volumes of ice (or water as well) with different sensors, optical or acoustic. This is due both to the low intensity of expected neutrino fluxes and to the low neutrino interaction cross section.

The most discriminating information of a neutrino detector, when searching for neutrinos from point-like sources like the centre of the Sun, comes from the reconstructed direction of the observed events. The proposed technique to pinpoint the neutrino source consists in detecting the optical signals emitted by the muons generated in charged current neutrino interactions (see sec. 2.3). The muons, indeed, can propagate with a velocity greater than the speed of light in the medium, and thus emit Cherenkov radiation at fixed angle with respect to the muon track (see sec. 2.5). Hence, the muons provide ample directional information for the reconstruction methods, while cascade signatures from ν_e and ν_τ are too short for a reliable angular reconstruction. Further, the lifetime of taus (below 10^6 GeV) is too short to produce tracks of significant length.

This kind of neutrino-induced muon detectors have to be shielded from the intense flux of atmospheric muons, which originate in cosmic ray interactions with the atmosphere (see sec. 2.6). That's why the neutrino telescopes should be deployed deep under the ice (or underwater). However, even at great depths ($\sim$ 4000 m), the atmospheric muon flux at the detector is about 6 orders of magnitude more intense than the neutrino-induced flux. This is the reason why the so-called *downgoing* muon events are not useful in the search for astrophysical sources, like in our case where we search for neutralino annihilations in the Sun. However, a *downward looking* neutrino telescope still suffers from the atmospheric muon background because of down-going muons misreconstructed as *upgoing* ones.

In conclusion, the ice (water) surrounding the detector has a triple function:

- it screens the detector from the atmospheric muon background
- it is a target for neutrino interactions producing muons
- it acts as a transparent radiator where Cherenkov light is emitted and propagated

The AMANDA (Antarctic Muon And Neutrino Detector Array) neutrino telescope, described in chapter 3, meets the requirements of a large-volume (200 m wide, 500 m high) sparse-instrumentation (optical sensors separated by $\sim$ 30 m horizontally, $\sim$ 15 m vertically) equipment, which is deployed deep in the Antarctic ice cap with the attempt to detect astrophysical neutrinos.

2.3 Neutrino interactions

The neutrino can interact weakly with matter exchanging a charged W boson with a quark in nucleon N. This process is called charged-current (CC) interaction:

$$\nu_\ell(\bar{\nu}_\ell) + N \longrightarrow \ell^-(\ell^+) + X \tag{2.5}$$

where a charged lepton ℓ and a hadronic shower X are produced (see fig. 2.2(a)). The neutrino can exchange a neutral Z boson as well, and in this case the process is referred to as a neutral-current (NC) interaction (see fig. 2.2(b)):

$$\nu_\ell(\bar{\nu}_\ell) + N \longrightarrow \nu_\ell(\bar{\nu}_\ell) + X \tag{2.6}$$

where the neutrino scatters off the nucleon N and a hadronic shower X is produced.

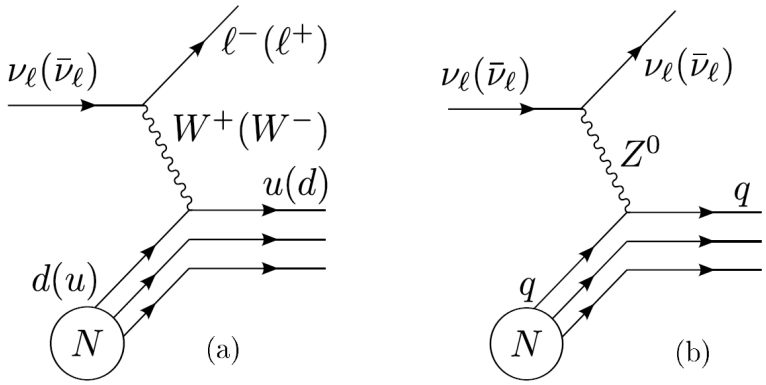


Figure 2.2: *Charged-current interaction (a). The neutrino (antineutrino) converts a down (up) quark to an up (down) quark by exchange of a W^+ (W^-) boson. A ℓ^- (ℓ^+) lepton is created and the nucleon produces a hadronic shower. Neutral-current interaction (b). The neutrino scatters off the nucleon through a Z boson exchange and a hadronic shower is produced. (Picture adapted from [94].)*

The differential cross section for the neutrino CC interaction with an isoscalar nucleon (i.e. $N \equiv \frac{n+p}{2}$), considering $\ell \equiv \mu$, can be written in terms of the Bjorken scaling variables $x = Q^2/2M(E_\nu - E_\mu)$ and $y = 1 - (E_\mu/E_\nu)$ as

$$\frac{d^2\sigma}{dxdy} = \frac{2G_F^2 M E_\nu}{\pi}\left(\frac{M_W^2}{Q^2 + M_W^2}\right)^2 \left[xq(x,Q^2) + x\bar{q}(x,Q^2)(1-y)^2\right] \tag{2.7}$$

where $-Q^2$ is the invariant squared four-momentum transfer between the incident neutrino and the outgoing muon, M and M_W are the nucleon and intermediate W boson masses, $q(x,Q^2)$ and $\bar{q}(x,Q^2)$ are the parton distributions of the nucleon, and G_F is the Fermi constant.
Muons produced in such a process are not collinear with the generating neutrinos; the mean scattering angle between the neutrinos and the muons, indeed, decrease with increasing neutrino energy and it is about 1.5° at $E_\nu = 1$ TeV and it reduces to $\sim 0.3°$ at $E_\nu = 10$ TeV, following the law $\langle\theta_{\nu\mu}\rangle \propto E_\nu^{-0.5}$ (this last relation is obtained integrating eq. 2.7 with respect to x).
A similar expression, like eq. 2.7, can be derived for the neutrino NC interaction as well, considering the Z as the intermediate (neutral) boson and a parton distribution functions.

The CC and NC cross sections are well described by a linear function of energy, in the range 1 GeV $< E_\nu <$ 10 TeV, while at high energies the increase in cross section becomes suppressed by the W boson propagator. The results of a calculation [95] with the CTEQ6 [96] parton distribution functions, for a proton (p), neutron (n) and isoscalar Nucleon (N) target, are shown in fig. 2.3.

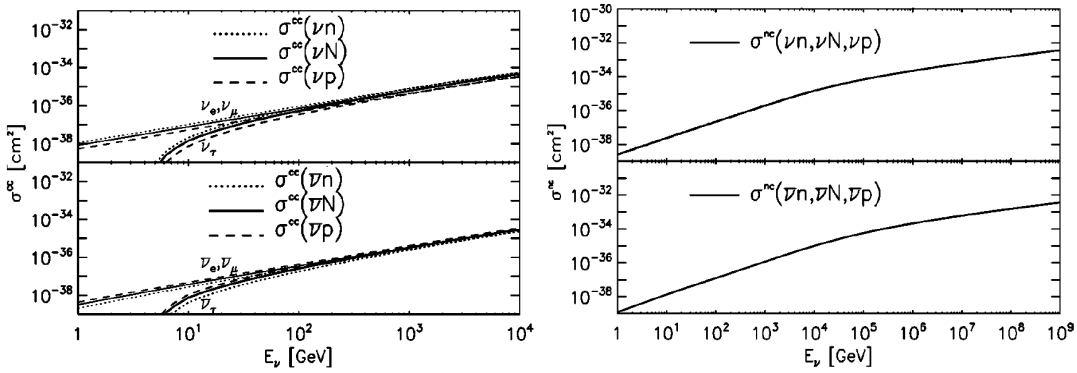


Figure 2.3: *On the left, the CC interaction cross sections for neutrinos (top) and anti-neutrinos (bottom); one should notice the kinematic suppression due to the τ mass. On the right, the NC interaction cross sections for all the three flavours. (Pictures taken from [95].)*

If the CC interaction on one hand permits the neutrino detection, on the other hand it reduces the flux of neutrinos on their way out of the Sun, an effect being particularly important for $E_\nu > 100$ GeV (see fig. 2.1). The NC interaction, does not absorb the neutrinos, but reduces their energy instead, and hence their further interaction probability.

The probability of neutrino transmission through the Earth is given by

$$P_{Earth}(E_\nu, \theta_\nu^{\rm zen}) = e^{-N_A \times \sigma(E_\nu) \times \rho_l(\theta_\nu^{\rm zen})} \tag{2.8}$$

where N_A is the Avogadro number, E_ν and $\theta_\nu^{\rm zen}$ are the energy and the zenith angle of the neutrino, σ is the neutrino interaction cross-section, and ρ_l is the integrated column density of the Earth [97]. This probability is shown in fig. 2.4 for three different neutrino energies; notice that the neutrino absorption in the Earth is negligible in the sub-TeV energy domain.

2.4 Muon propagation in ice and energy loss

High energy muons which propagate in a transparent medium like the ice, lose a small amount of their energy through Cherenkov radiation (see sec. 2.5).
The main muon energy loss, indeed, is due to ionisation and radiative processes (bremsstrahlung, e^+e^- pair productions, photo-nuclear interactions). These processes settle the muon path (the so-called *muon range*), and may generate secondary high energy particles (mostly electrons), which can emit Cherenkov radiation too.

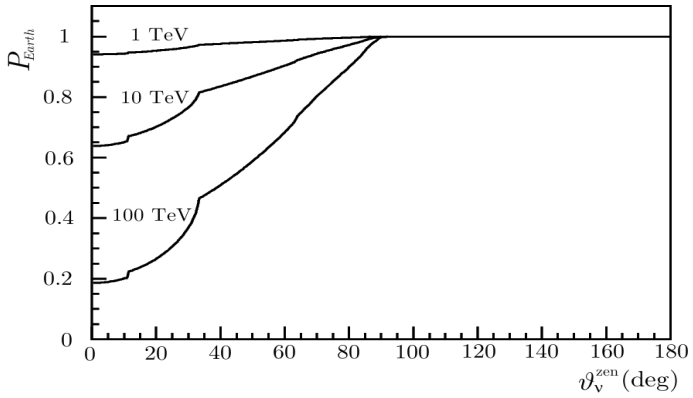


Figure 2.4: *Neutrino transmission probability through the Earth in function of the zenith angle, for three different neutrino energies.*

Muons which traverse the medium, can also undergo multiple elastic Coulomb scattering off nuclei, although the energy loss is small or negligible. At small scattering angles, the space angle distribution per solid angle is well approximated by a Gaussian, while as the angle increases a long broad tail is observed [98]. However, the effects of multiple Coulomb scattering can be neglected in this analysis, since $\theta_{\mathrm{RMS}}^{\mathrm{Coul}} \ll \langle\theta_{\nu\mu}\rangle$.

2.4.1 Ionisation energy loss

The mean rate of energy loss (or stopping power) of a moderate relativistic charged particle is given by the Bethe-Bloch formula [51]:

$$-\frac{dE}{dx} = Kz^2\frac{Z}{A}\frac{1}{\beta^2}\left[\frac{1}{2}\ln\frac{2m_e c^2\beta^2\gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2}\right] \tag{2.9}$$

where $K \simeq 0.307$ MeV cm^2 mol^{-1}, z and β are the charge (in units of e) and velocity (in units of c) of the particle, Z and A are the atomic number and mass number of the atoms of the medium, and x is the path length in the medium measured in g cm^{-2} (or kg m^{-2}). The quantity I is an effective ionisation potential, averaged over all electrons, whose value is $\simeq 10Z$ eV; T_{max} is the maximum kinetic energy which can be imparted to a free electron in a single collision, and $\delta(\beta\gamma)/2$ is a density effect correction to the ionisation energy loss.

The incident particle produces primary ionisation in atomic collisions. High energy electrons knocked out in this process (called δ-rays), can themselves produce fresh ions in traversing the medium (secondary ionisation).

2.4.2 Radiative processes

At sufficiently high energy, radiative processes, such as bremsstrahlung, pair production and photo-nuclear interactions, become more important than ionisation for all charged particles. These processes of stochastic nature, are characterised by small cross sections, hard spectra, large energy fluctuations, and the associated generation of electromagnetic and (in the case of photo-nuclear interactions) hadronic showers [51].

The average rate of muon energy loss can be written in a more convenient form [99]:

$$-\frac{dE}{dx} = a(E) + b(E)E \tag{2.10}$$

where $a(E)$ is the ionisation energy loss given by eq. 2.9 and $b(E)$ is the sum of e^+e^- pair production, bremsstrahlung and photo-nuclear contributions. Under the approximation that these slowly-varying functions are constant, the mean range R_μ of a muon with initial energy E_μ can be found solving eq. 2.10:

$$R_\mu \approx \frac{1}{b} \ln\left(1 + \frac{E_\mu}{E_{\mu_{\text{crit}}}}\right) \tag{2.11}$$

where $E_{\mu_{\text{crit}}} = a/b$ is the *muon critical energy*, defined as the energy at which radiative and ionisation losses are equal.
For stochastic losses, $a \simeq 2.68$ MeV g^{-1} cm^2 and $b \simeq 4.7 \times 10^{-6}$ g^{-1} cm^2, hence the average muon range varies between ~ 300 m w.e.[1] at $E_\mu = 100$ GeV and ~ 31 km w.e. at $E_\mu = 10^9$ GeV.
Figure 2.5 illustrates the survival probability for a muon at a certain distance, indicating the range calculated by means of eq. 2.11. The large fluctuations of the actual muon range around the mean value are clearly visible.

Bremsstrahlung

The electric field of a nucleus, or atomic electrons, decelerate the muons which traverse the medium, and the energy change appears in the form of photons; hence the term *bremsstrahlung*, "braking radiation". The calculation of the cross section of this process was first derived by Bethe and Heitler [101], and subsequently improved by the work of Petrukhin and Shestakov [102].

Direct electron pair production

A muon can radiate a virtual photon, which in the electric field of a nucleus, can convert into a real e^+e^- pair. A parametrisation of the differential cross

[1] The effect of medium densities normally rescales out by working in "water equivalent" (w.e.) units, where the density of all materials is set to one. Rescaling the resulting water equivalent distances by $1/\rho$, we can then obtain the physical distances.

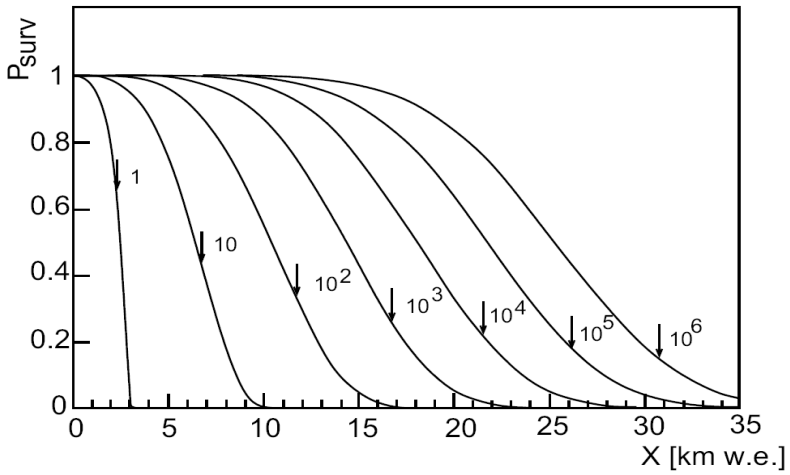


Figure 2.5: *Survival probabilities for muons of different energies (1 TeV–10^6 TeV, indicated by the numbers besides the curves) in rock. The arrows point to the average range resulting from eq. 2.11. Figure taken from [100].*

section of direct pair production was derived by Kokoulin and Petrukhin [103].

Photo-nuclear interaction

A muon can also radiate a virtual photon which directly interacts with a nucleus. The differential cross section of this process is proportional to the total cross section $\sigma_{\gamma N}$, for absorption of a real photon of energy $E_\gamma = s/2m_N = \nu_E$ by a nucleon. A parametrisation of the differential cross section was derived by Bezrukov and Bugaev [104].

Figure 2.6 shows the different contributions to the average energy loss of a muon when travelling through the ice. Decay energy loss is shown for comparison, and it is evaluated by multiplying the probability of decay by the energy of the particle. In the region below 1 GeV, bremsstrahlung energy loss has a double cut-off structure. This is due to a difference in the kinematic restrictions for muon interaction with oxygen and hydrogen atoms.

2.5 The Cherenkov effect

When high-energy charged particles traverse dielectric media, part of the light emitted by excited atoms appears in the form of a coherent wavefront at fixed

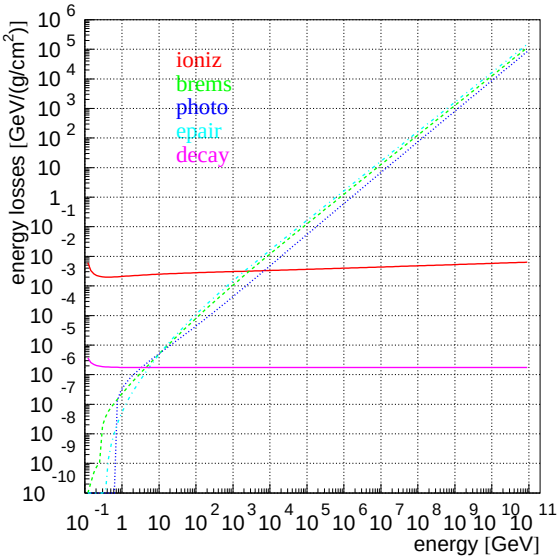


Figure 2.6: *Ionisation (upper solid curve), bremsstrahlung (dashed), photo-nuclear (dotted), e^+e^- pair production (dashed-dotted) and decay (lower solid curve) losses in ice. Figure taken from [105]*

angle with respect to the trajectory. This phenomenon is known as Cherenkov effect, after its discoverer. Such radiation is produced whenever the velocity of the particle, βc, exceeds c/n, where n is the refractive index of the medium. From the Huygens construction of fig. 2.7, we can see that the wavefront forms the surface of a cone around the trajectory as axis, such that

$$\cos\theta_c = \frac{ct/n}{\beta ct} = \frac{1}{\beta n} \tag{2.12}$$

For muons passing through the ice ($n = 1.32$ at 400 nm) the energy threshold[1] for the Cherenkov effect is approximately 160 MeV. For our concerns, we can safely assume $\beta \simeq 1$, hence $\theta_c \simeq 41°$.

Cherenkov radiation appears as a continuous spectrum; in a dispersive medium, both n and θ_c will be functions of the frequency ν. The total energy content of the radiation, per unit track length, is given by [98] (charge $z = 1$)

$$\frac{dE}{dx} = \frac{4\pi^2 e^2}{c^2} \int \left(1 - \frac{1}{\beta^2 n^2}\right) \nu d\nu \tag{2.13}$$

The spectral distribution of the emitted photons is given by the Frank-Tamm

[1] The velocity threshold is $\beta_t = 1/n$, hence $E_t = \gamma_t m_\mu = m_\mu/\sqrt{1-n^{-2}}$

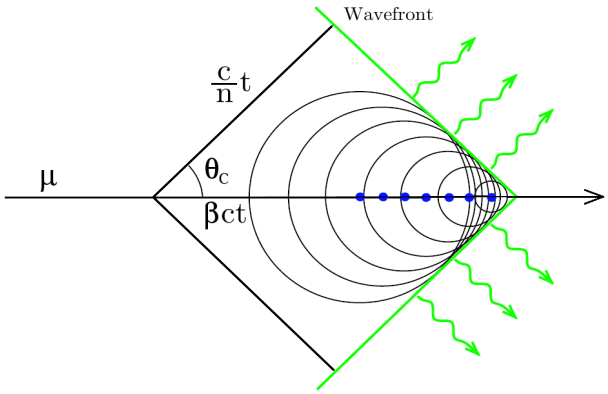


Figure 2.7: *Illustration of the formation of a coherent Cherenkov light wavefront from spherical waves emitted along the particle trajectory. The Cherenkov angle θ_c is defined as the angle between the particle trajectory and the propagation direction of the light.*

formula [51] (charge $z = 1$)

$$\frac{d^2N}{dxd\lambda} = \frac{2\pi\alpha}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2(\lambda)}\right) \tag{2.14}$$

where $\alpha = 1/137$ is the fine structure constant. The number of photons at a particular frequency or wavelength is proportional to $d\nu$ or $d\lambda/\lambda^2$, thus the blue light predominates. From integration of eq. 2.14, we can infer the total number of photons expected per unit track length. In the range between 300 nm and 500 nm, 2.6×10^4 photons are emitted by a muon in one meter. The energy loss due to the radiation of these photons is approximately 86 keV/m, negligible compared to the dominant processes of energy loss of the muon described in sec. 2.4.

2.6 Particles background

Charged particles accelerated by astrophysical sources that reach the Earth, the so-called *primary cosmic rays*, are mainly composed of protons and α-particles. These particles interact with nuclei in the atmosphere producing jets of hadrons, like pions and kaons, which in turn can decay into muons and neutrinos. Hence, according to their origin, these particles are called *atmospheric* muons and neutrinos. The latter constitute then a background to the search for astrophysical neutrinos.

The decay length of the π- and K-mesons ($d_m = \beta\gamma c\tau_m$) might be larger than their interaction length λ_m, depending on their lifetime τ_m. The turnover point

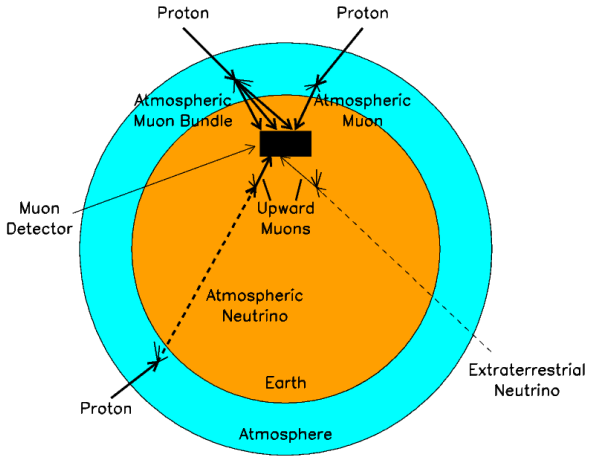


Figure 2.8: *Illustration of the possible processes which produce muons that can be detected by an underground telescope.*

where $d_m = \lambda_m$, is called the critical energy ϵ_m: above that energy interaction dominates over decay; we have $\epsilon_{\pi^\pm} = 115$ GeV, $\epsilon_{K^\pm} = 850$ GeV. Hence, below ~ 100 GeV, where decay dominates over interaction, the atmospheric muon energy spectrum follow that of the primary cosmic rays, $d\phi/dE \propto E^{-2.7}$. Above the critical energy, on the contrary, the interaction dominates, steepening the spectrum by one power in energy, $d\phi/dE \propto E^{-3.7}$. For atmospheric neutrinos the same steepening of the spectrum can be observed. Above $E \simeq 1$ PeV, the spectrum flattens again due to the contributions of the "prompt" flux, originating from the decay of charmed mesons, which have very short lifetimes and correspondingly high critical energies ($\epsilon_c > 10^7$ GeV).

The angular spectrum below ~ 10 GeV is proportional to $\cos^2\theta_z$ and flattens at higher energies approaching a $\sec\theta_z$ distribution, where the zenith angle θ_z spans the angle between the vertical and the bundle axis. At large zenith angles the π- and K-mesons traverse a lower density atmosphere, hence they decay even at larger energies; this explain the so-called *secant theta* effect.

The flux of atmospheric muons is strongly reduced once they reach the Earth's surface and start to travel through it. The Earth, hence, can be used as a muon filter if the neutrino telescopes are installed deep underground. The energy threshold for a vertical downward-going atmospheric muon, which could reach the AMANDA-II detector (located at 1730 m below the Antarctic ice cap) is about $E_{tr} = 400$ GeV. The energy threshold increases with zenith angle, since more matter should be passed to reach the detector. This suppresses the high zenith angles in the atmospheric muon flux at the detector, until they vanish at around $\theta_z \simeq 85°$. This means that upward-going and horizontal muons can only

be generated close to the detector and not in the atmosphere.

Atmospheric neutrinos follow the same angular and energy directives as the atmospheric muons. However, neutrinos have a small probability to be absorbed during their path through the Earth (see fig. 2.4). Therefore they are able to produce horizontal and upward-going muons in a detector, that is the same signature left by neutrinos from astrophysical sources. Hence, the atmospheric neutrinos represent an irreducible background, especially in the GeV-TeV energy range, where the flux is still strong.

An illustration of the possible processes, which produce muons that can be detected by an underground telescope, are shown in fig. 2.8.

The primary cosmic rays can interact with the Sun's atmosphere as well, which, completely analogous to the Earth's case, subsequently generate *solar atmosphere neutrinos*. This can be seen as a third source of background; however, the number of events expected is rather small, less than one event per year in a detector like AMANDA-II [106]. This should be compared to the trigger rate of atmospheric muons (~ 100 Hz), and atmospheric neutrinos ($\sim 3.5 \times 10^{-4}$ Hz).

Finally we should mention that nuclear fusion reactions in the centre of the Sun produce neutrinos in the MeV energy domain, which is much lower than the energy threshold of a neutrino telescope, around $\mathcal{O}(15$ GeV).

Lo 'mperador del doloroso regno
da mezzo 'l petto uscia fuor de la ghiaccia;
e più con un gigante io mi convegno,
che i giganti non fan con le sue braccia;
vedi oggimai quant' esser dee quel tutto
ch'a così fatta parte si confaccia.

Inferno - Canto XXXIV
Dante Alighieri (1265 - 1321)

3

The AMANDA-II neutrino telescope

In this chapter we will talk about the AMANDA neutrino telescope, located 2000 m below the polar South Pole ice cap. We will describe its technology, calibration and operation. At the end of the chapter we will describe the ice properties, essential to understand the effect of scattering and absorption of the light in the medium, along with their measurements.

3.1 Development and structure of the detector

As we will discuss in sec. 3.3, below 1500 m the Antarctic glacier is very pure with good optical properties for the detection and reconstruction of high energy neutrinos. The use of the existing research station and established infrastructure located at the Amundsen-Scott South Pole base, represented the appropriate way of getting access to deep glacier ice.

The construction of the AMANDA (Antarctic Muon And Neutrino Detector Array) neutrino telescope, near the Amundsen-Scott base, started in the austral summer season[1] of 1995/96. The deployment in the ice of the detector cables (referred as *strings*), involved the drilling of 60 cm wide holes using hot water (at $\sim 80°$ C), down to a depth of about 2000 m. Each string equipped with several optical modules (OMs), was then lowered into the water filled hole, which afterwards fully refroze. The final detector configuration was reached in February of 2000: an instrumented cylindric array of 19 strings arranged over three concentric rings (see fig. 3.2 top), with 677 OMs. It was located at depths between 1500 m and 2000 m, with a height of 500 m and a diameter of 200 m, and its centre was 1730 m below the surface. All the supply and readout cables in the strings were bundled at the surface and connected to the data acquisition system located in the Martin A. Pomerantz Observatory (MAPO). The layout of AMANDA in

[1]The harsh climate affected the construction schedules, and naturally divided the development into stages defined to one season. During the winter (no transports came to or from the Pole), the station was manned only by a small crew to do maintenance work and data-taking.

the ice is shown in fig. 3.1, along with the historical development of the detector array, which we will briefly discuss below.

In 1993/94 a prototype detector with four strings, called AMANDA-A, was installed at 800-1000 m depth. AMANDA-A is not part of the detector used for this analysis. Because of the presence of air bubbles at shallow depths, the resulting scattering length of light in the ice was too short to allow accurate track reconstruction; however the prototype has contributed to the understanding of the ice properties [107].

The first and innermost four strings of the detector were installed in the season 1995/96. The optical modules on these strings were separated by 20 meters and connected to the surface electronics by coaxial cables.

In the season 1996/97 strings 5 to 10 were installed; these ones were deployed with an OM separation of only 10 meters and twisted pair quad cables. We refer to the detector configuration reached in this campaign as AMANDA-B10.

The strings installed in the season 1997/98 were number 11 to 13. From string 11 on, analog fiber-optic signal transmission was used. Unfortunately about 10% of the fibers were destroyed in the refreezing, due to their frailty and those of their connections. Additional modules, located in pairs above and below the bulk of the detector (see fig. 3.2, bottom), were used for studies of the ice properties; however they were excluded from the trigger and also from the current analysis.

The season 1999/2000 marked the final configuration of the detector as we said before, with the deployment of the last six strings, the numbers 14 to 19. We refer to this 19 string configuration as AMANDA-II, and, since we only use data taken with the whole detector array, we will use its shorthand notation, AMANDA, in the following.

A close look at fig. 3.2 (bottom) indicates that the centre of string 17 is 500 m higher than the rest of the array. That is the result of a deployment incident: the string got stuck while lowering it into the molten hole and froze in above the designated position. The modules in string 17 were thus excluded from the trigger as well as from the analysis.

3.2 AMANDA detector technology

3.2.1 The optical module

The optical module is composed of an 8-inch Hamamatsu photo-multiplier tube (PMT), enclosed in a pressure resistant glass sphere together with some basic electronics. The photo-cathode is optically connected by silicon gel to the pressure sphere to reduce light reflection. Apart these common aspects, the OMs in the different strings were quite heterogeneous. The different generations of

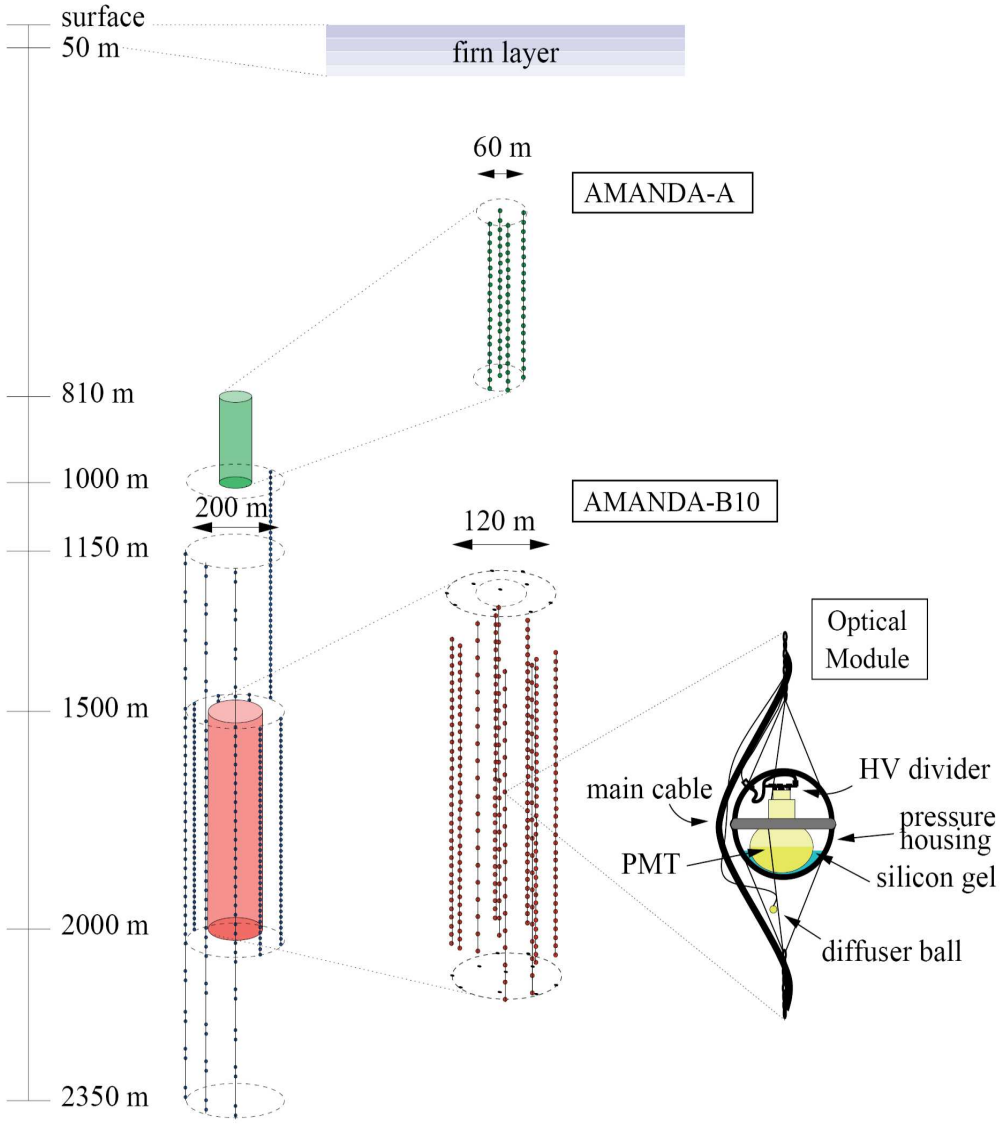


Figure 3.1: *Schematic overview of the development of the detector in different stages, from AMANDA-A to the final configuration AMANDA-II. On the bottom right, a zoom on one optical module.*

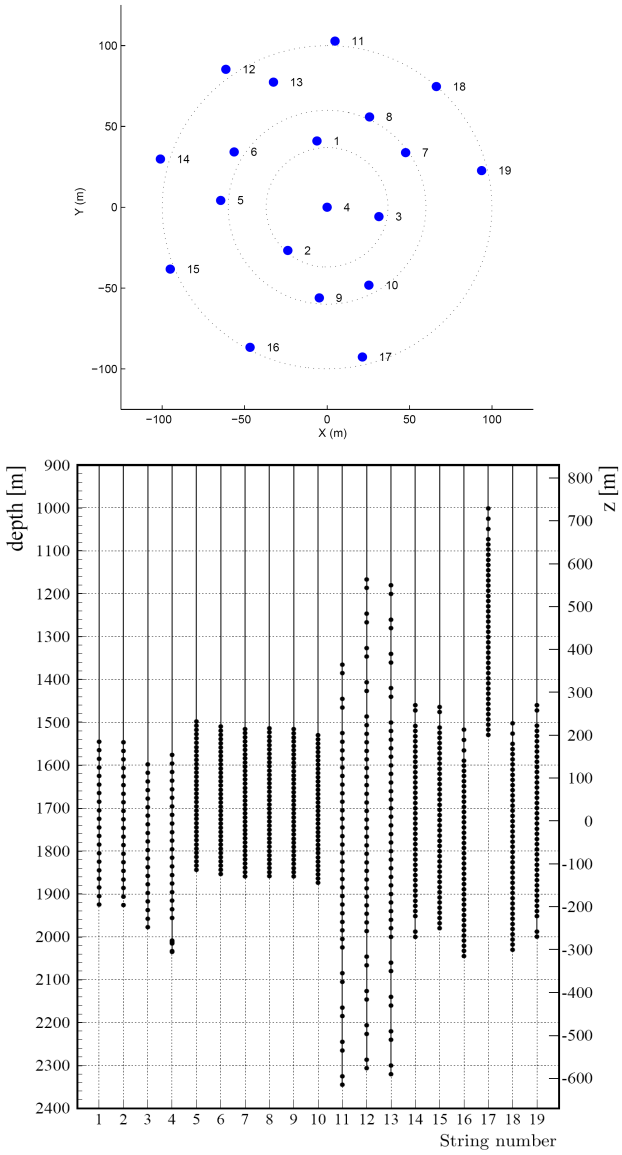


Figure 3.2: *On the top, geometry of the 19 AMANDA strings in the horizontal plane, according to the AMANDA coordinate system (see sec. 4.2). On the bottom, the depth and the Z coordinate of the OMs along each string.*

OMs deployed over the years, reflect the efforts of improvement of the technology used before. A short description of the various OM generations follows below.

In the strings 1 to 4, the high voltage power supply and transmission of the PMT pulses to the surface was fulfilled by coaxial electrical cables, which were insensitive to noise pick-up from induction, but suffered from high dispersion. Single photo-electron pulses were broadened to 200 ns - 400 ns, thus contiguous pulses could not be resolved. The dark noise rate of these OMs (about 0.5 kHz) was quite low, since the glass sphere manufactured by the Billings company was made of low radioactive material. When the PMT operated at high gain of 10^9 the OM was sensitive to single photo-electrons in the wavelength range between 300 nm and 600 nm. The glass became opaque below 350 nm, thus reducing the OM quantum efficiency by about 25% compared to the glass sphere manufactured by the Benthos company, which was used later.

The introduction of twisted pair electric cables in the strings 5 to 10, reduced dispersion to about 100 ns - 200 ns. However, the sub-optimal shielding led to *crosstalk* pulses, i.e. pulses from one OM induced fake signals in a neighbouring cable of another OM. The Benthos glass spheres were used with improved UV transparency, but on the other hand the rate of the dark noise increased to about 1 kHz, due to the presence of radioactive material.

As we have discussed in the previous section, optical fibers were deployed from the strings 11 onwards; therefore a LED was used to convert the electrical PMT pulse into an optical pulse. In this way the dispersion was eliminated and the pulse-width was around 20 ns. The optical fibers, however, had a high failure rate from mechanical stress, hence to guarantee the signal transmission, twisted pair electrical cables were used as backup solution.

String 18 was equipped with Digital Optical Modules (DOMs); they were active modules generating HV locally and, in addition, the PMT pulses were digitised at the OM and transmitted to the surface as binary data via electrical cables. Analog optical signal transmission was provided as backup. String 19 was equipped with dAOMs, digital-analog optical modules, providing analog optical and electrical readout, but digital module control.

Not all pulses transmitted to the surface originated from charged leptons; besides non-photon pulses produced by the detector hardware itself, like the crosstalk and the dark noise, we can mention afterpulses caused by ionisation of residual gas in the PMT tubes. And more, an unidentified occasional source of non-photon pulses was outside the detector, so-called flare events. Anyhow, most of these *background* pulses can be afterwards removed from raw data (more details in sec. 5.2).

3.2.2 Data acquisition system

The pulses transmitted to the surface were processed by the data acquisition system (DAQ). An illustration of the AMANDA DAQ is sketched in fig. 3.3. The numbers in the following description refer to the numbers in this figure.

The pulses that travel through the electrical cable were amplified by the SWAMP (SWedish AMPlifier) that delivered one prompt and one delayed version of 2 μs of the amplified signal (2). Optical signals instead went first to the Optical Receiver Boards (ORB) that generated corresponding prompt and delayed electrical pulses of 2 μs as well (3).

The prompt signals were then sent to the combined discriminator and trigger electronics, the *Discriminator and Multiplicity ADder* (DMAD) (4,9,10), and to the Transient Waveform Recorder (TWR) (4)[1]. Besides the MuonDaq currently described, the TWRDaq was a second more advanced system that sampled the full wave form of the pulses [108, 109]; however, since this information has not been used in the analysis discussed in this thesis, we will not add more details.

The second component of the DMAD, the multiplicity adder, added the discriminated pulses of all OMs in a sliding window of 2.5 μs and verified whether they were above a certain threshold. The timings of the discriminated signals were recorded with a *Time to Digital Converter* (TDC) (5); it measured the time of the leading (LE) and trailing edge (TE) of the pulse, corresponding to the positive and negative crossing point with the discriminator threshold (see fig. 3.4). The TDC was capable to store at most 16 edges in its 32 μs buffer, or 8 complete pulses which arrived between -22 μs and +10 μs around the time of the issued trigger (see description in the next section). A *peak sensing Analog to Digital converter* (pADC) (6) received the delayed SWAMP/ORB output and registered the maximum pulse amplitude among the pulses arriving within a 10 μs time window, [-2 μs, +8 μs], around the trigger time.

The DAQ system collected all the information from the pADCs and TDCs and combined them with the absolute time received from a GPS clock (7,8); the triggered events were thus red out and eventually stored on tape. During readout of an event, TDCs and pADCs were not able to collect data, hence the detector was unable to record information: this period was the so-called *dead-time*. The fraction of dead-time was a function of the overall trigger rate, i.e. higher trigger rate led to higher dead time. The true trigger rate was about 100 Hz which yielded about 80 recorded events per second with 20% dead-time.

[1] Note that label 4 is used in two places in fig. 3.3.

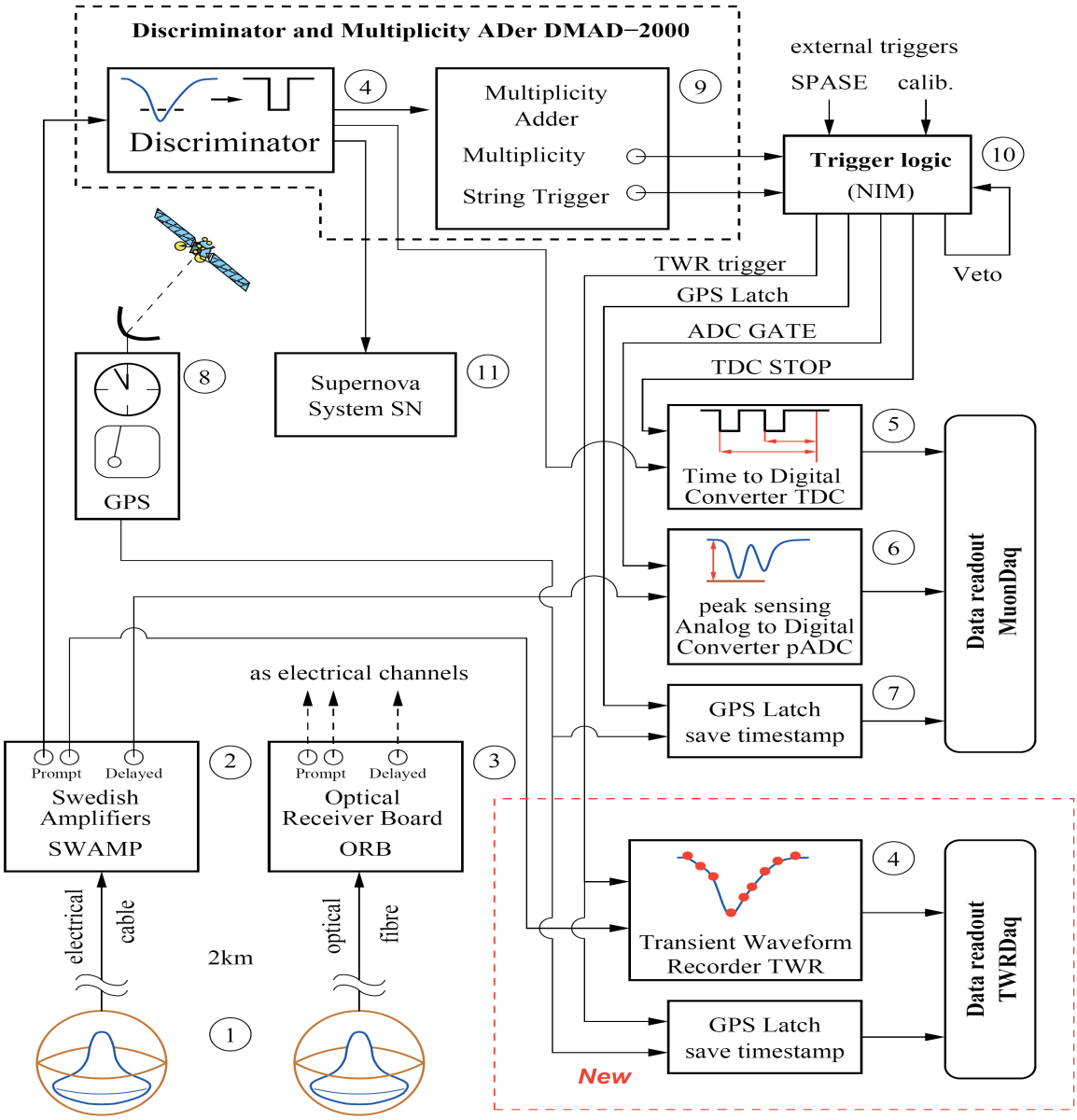


Figure 3.3: *Schematic layout of the data acquisition system. Picture taken from [108].*

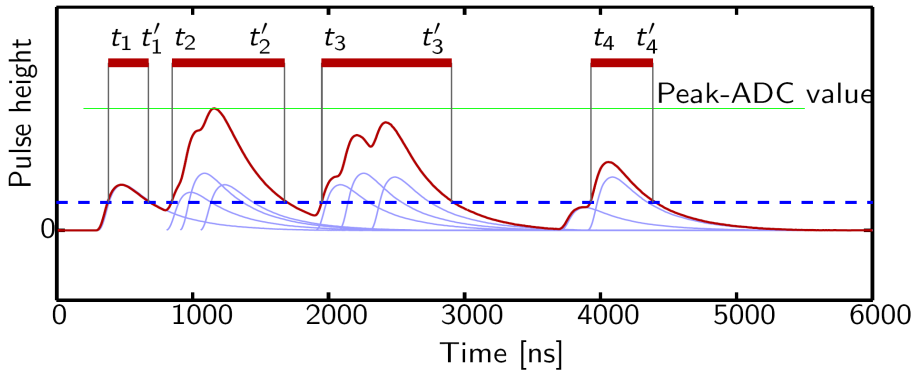


Figure 3.4: *A pulse from the PMT (thin red curve) becomes a hit when it passes the discriminator threshold (blue dashed line). The TDC records the time of the leading edge t_i and trailing edge t'_i of up to eight hits (thick red lines). The recorded pADC value is common for all hits in a given OM: that of the maximum pADC (V_{ADC}) of all pulses (horizontal green line). Picture taken from [110].*

3.2.3 Event trigger

The choice of the trigger condition for the detector should be set carefully, if we want to find a good balance between technical constraints and physics interests. For instance, a higher trigger rate would increase the number of readout cycles, hence the detector dead-time. Another fact that should be taken into account is the storage capacity, which could restrict the number of events to be processed. On the other hand, a too limiting trigger would result in the loss of many low energy events, potentially interesting for dark matter searches.

As a compromise in the period concerning our analysis, i.e. during 2001-2006, two main trigger conditions were available.

The first trigger condition, the *multiplicity trigger*, required a minimum number $M = 24$ of OMs hit within a time window of 2.5 μs (see fig. 3.5 on the left).

The second trigger condition, the *string correlation trigger*, was implemented to improve the sensitivity for low energy events, without including too many fake events due to noise in the detectors. By using a correlation criterion on neighbouring hits, the number of hits needed for a trigger can be lowered. The string correlation trigger required at least M out of N consecutive modules fired on a string within a time window of 2.5 μs. For the inner four strings $M/N = 6/9$ (see fig. 3.5 on the right), while on the remaining strings (which have a smaller vertical spacing between OMs), the condition was set to 7/11. To keep the dead-time as low as possible, it was decided in 2002 to downscale the string trigger by

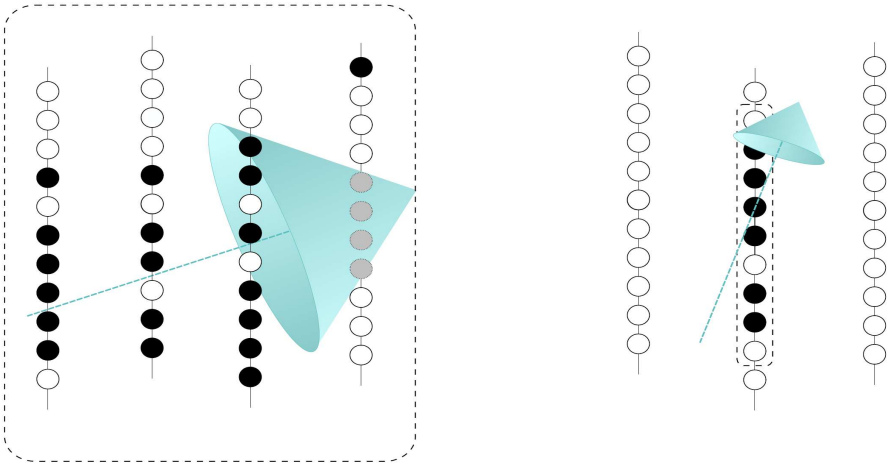


Figure 3.5: *Sketch of the AMANDA triggers: on the left the $M = 24$ multiplicity trigger condition, on the right the string trigger condition ($M/N = 6/9$). Picture taken from [111].*

a factor of two, i.e. only half of the events which exclusively satisfied the string trigger were stored to the disk.

A third trigger, more precisely, an external trigger from SPASE (South Pole Air Shower Experiment) located on the surface above AMANDA [112], was also available during 2001-2006. Other external triggers might be applied for example when performing calibration or other detector diagnostics. These triggers were not useful for our search for low energy neutrino from the Sun, and thus discarded from our analysis.

3.2.4 Calibration

The information recorded for every hit consisted of: the time of the leading (t_{LE}) and trailing edge (t_{TE}) measured by the TDC on surface, the maximum amplitude V_{ADC} recorded by the pADC, and the channel number (n_{OM}) which detected the photon.

In chapter 5 we will see how the hit information of a triggered event was used by the reconstruction algorithm for the data analysis. Basically, the information needed to reconstruct the events was the arrival time of the photon at the Optical Module (t_{OM}), and the photo-electron multiplicity n_{pe}, which produced the pulse. Moreover, the geometrical location of the Optical Modules had to be known with high precision. Hence, an accurate calibration in time, amplitude and geometry was performed to get the previous quantities from the recorded information.

Time calibration

The time of arrival of a photon at the Optical Module was calculated through the formula:

$$t_{\mathrm{OM}} = t_{\mathrm{LE}} - t_0 - \frac{\alpha}{\sqrt{V_{\mathrm{ADC}}}} \tag{3.1}$$

where the time t_0 accounted for all the delays of the pulses along the cable and in the electronics. The last term, $\alpha/\sqrt{V_{\mathrm{ADC}}}$, was a correction for the amplitude dependent time interval between pulse start and discriminator threshold crossing, so-called "amplitude time walk". The pulse width was due to the dispersion of the cable, hence larger pulses crossed the fixed discriminator threshold earlier than smaller pulses.

The two calibration constants, t_0 and α, were derived using frequency doubled Nd:YAG laser light pulses, sent out through the ice via optical fibers and emitted by diffuser balls located below the OM (see fig. 3.1). During the calibration run, sufficient statistics of the t_{LE} and V_{ADC} were collected. Hence a linear fit to equ. 3.1, in form of $y = mx + q$ ($x = 1/\sqrt{V_{\mathrm{ADC}}}$, $y = t_{\mathrm{LE}}$), was performed to extrapolate the parameters $m = \alpha$ and $q = t_0 + t_{\mathrm{OM}}$. The value of t_0 was then calculated subtracting from this last parameter the value of t_{OM}, which was the sum of the time when the laser pulse was emitted, t_{laser}, the travel time of the light pulse inside the optical fibers, t_{fiber}, and the travel time, between the diffuser ball and the OM, of the light pulse in the ice, t_{ice}. The time t_{laser} was calculated using a photodiode on the surface to trigger the data acquisition system, t_{fiber} was measured for each individual fiber from the round trip time of light reflected at the end of the fiber using an Optical Time-Domain Reflectometer (OTDM), and t_{ice} was calculated from the known speed of light in ice and the known distance OM-diffuser ball. Such a procedure was repeated for all OMs in the detector [113]; however t_0 and α changed with the years due to the maintenance, tuning or upgrade of the surface hardware and its relative setting.

A second calibration method, the so called "muon-calibration", which used down-going muons from cosmic ray induced air showers, selected well reconstructed muon tracks to iteratively fine-tune the t_0 constants, comparing the time distribution of the recorded hits to their expected time distribution. In addition to have an independent cross check of the laser calibration, this method was used for channels which could not be calibrated by the laser method [114].

The time resolution of the registered pulses after calibration was about 5 ns for the strings with coaxial or twisted pair cables, and about 3.5 ns for strings with optical readout, accurate enough to get a negligible systematic effect on the reconstruction results [115].

Amplitude calibration

The merged pulses of photons which arrived nearly simultaneously at the PMT, resulted in a pulse of amplitude V_{ADC} (see fig. 3.4). The number of photo-electrons which produced the pulse depended on the gain of the amplifier, the PMT gain and the signal attenuation in cables and electronics. Mainly the pulses in downgoing muon data were single photo-electron (*spe*) pulses from either dark-noise or distant tracks. The pADC distribution of such events was a Gaussian spe-peak superimposed on a rapidly falling exponential contribution from PMT noise. The mean amplitude for single photo-electrons V_{spe} was found from a fit to the position of the Gaussian peak in the pADC spectrum, and was different for each OM. Assuming a linear response (in the range where saturation did not play a significant role) the measured V_{ADC} value could be converted to the number of photo-electrons:

$$n_{\mathrm{pe}} = \frac{V_{\mathrm{ADC}}}{V_{\mathrm{spe}}} \tag{3.2}$$

Geometry calibration

The surface position of the holes was determined by triangulation before the deployment. Pressure sensors located at the lowest and highest part of the OMs were used to determine their Z-position in the melted hole, and the string expansion during deployment. With this method an accuracy of about 1 m in X-Y-direction and about 2 m in Z-direction was reached [116].

After deployment, intra-module light sources were used to obtain the relative positions of the OMs in the array; these laser pulses were sent from at least five locations on every string. The distances were then accurately determined by observing the propagation time of unscattered photons to neighbouring strings. The accuracy reached by both methods combined was about 50 cm [117].

3.3 South Pole ice properties

The knowledge of the optical properties of glacial ice, like the deep ice at the South Pole, is of striking importance to scientific endeavours beyond fields like optics and glaciology. One of this scientific endeavour is neutrino astrophysics, which we investigated by means of the AMANDA detector; hence, since the optical sensors in AMANDA were sensitive to light with wavelengths between 300 and 600 nm, we need to understand and take into account the effects of scattering and absorption of light in the ice at wavelengths in the visible and near ultraviolet.

Antarctica, on average, is the coldest, driest, and windiest continent, and it is almost considered a desert with very low precipitation, especially inland.

About 98% of Antarctica is covered by the Antarctic ice sheet, a sheet of ice averaging at least 1.6 km thick, being 2820 meters deep at the detector site. The icing of Antarctica began with ice-rafting[1] about 45.5 million years ago and escalated inland widely about 34 million years ago [118]. Ice enters the sheet through precipitation as snow. This snow is then compacted to form glacier ice which moves under gravity towards the coast. The bulk of AMANDA was in a depth region where the ice moves as a rigid block, and only the lowest modules at 2300 m (which were used for investigation of ice properties but not for data analysis) might lag behind the rest of the detector [119].

Antarctic ice is a quite and clean medium with very low biological activity, if any, and the only noise registered by the OMs came from decays of ^{40}K in the pressure spheres and from the dark noise of the photomultipliers. Mineral dust, sea salt, biological material, debris thrown into the atmosphere by volcanic eruptions, or meteorites, are the only impurities carried by air from other regions of the Earth, or that come from space. The amount of such impurities, which settle with the growing ice, varies with climate and geological conditions; this can be seen in form of layers with different optical properties. Figure 3.6b shows, for one wavelength, the scattering coefficient as a function of depth found by AMANDA [119]. The measurements of absorption and scattering were made using pulsed and steady light sources in the ice. Besides the Na:YAG laser, there were also two *in situ* nitrogen lasers and several LED "flashers". One of the steady light sources was a UV lamp at 313 nm with variable intensity. The other one was the *rainbow module*, a halogen lamp with variable wavelength between 340 and 560 nm, used to investigate the wavelength dependence of the ice parameters.

At depths in South Pole ice shallower than ∼1300 m, scattering by air bubbles dominates scattering by dust, including most of the dust peak corresponding to the Last Glacial Maximum[2] (LGM). Since air bubbles scattered light without absorbing it, analysis of scattering as a function of depth showed a strong peak at a depth of 1300 m corresponding to the LGM. Figure 3.6a shows the relevant dust peaks at Vostok [121], obtained using Ca^{2+} concentration as a proxy for dust. A rough age versus depth relationship for South Pole ice for the last 70 ka (kiloyears before present) was obtained by Price et al. [122], who identified the peaks at LGM, A, B, C, and D with the Vostok ages at the corresponding peaks and included the age versus depth relationship obtained from analysis of a 200-m South Pole core [123]. Figure 3.6c shows the high-resolution signal from the "dust-logger" device lowered in the first IceCube borehole at the South Pole for the age interval ∼25 to ∼70 ka [124]. For comparison, fig. 3.6d shows the

[1] Ice-rafting is the transport of various material by ice.

[2] The Last Glacial Maximum refers to the time of maximum extent of the ice sheets during the last glacial period, approximately 20,000 years ago [120].

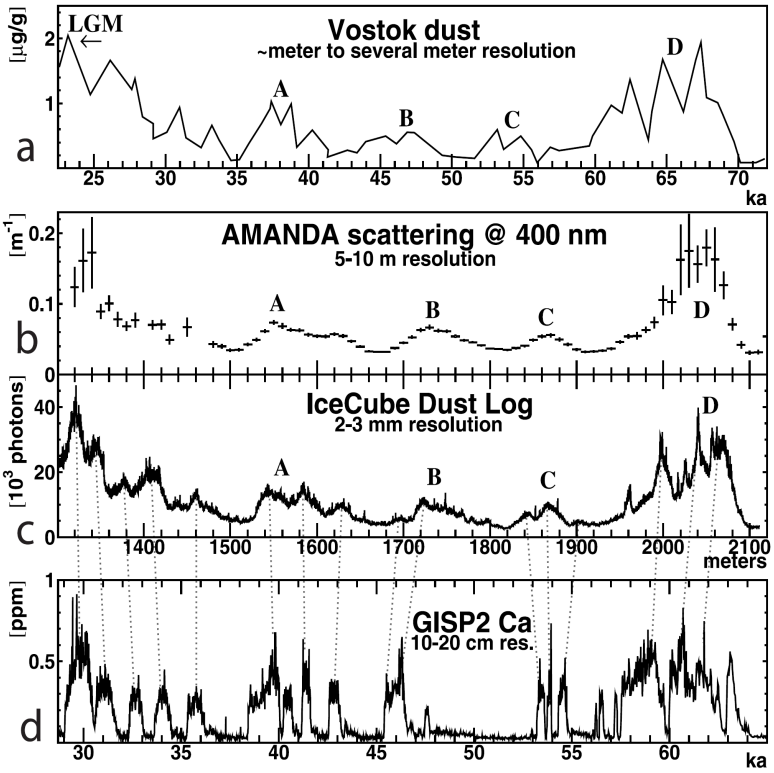


Figure 3.6: *Comparison of dust (a) measured in a Vostok ice core with a Coulter counter; (b) inferred from travel time of light from emitters and receivers in the AMANDA array in South Pole ice; (c) with a new dust logger in a water-filled IceCube borehole; and (d) inferred from analyses of calcium ions in the GISP2 ice core. Picture taken from [124].*

GISP2[1] calcium record for the same time interval [125].

As the pressure increases with depth, air bubbles compress and eventually become unstable against a transition from the gas phase to the solid air-hydrate clathrate phase [126]. Since the rate of transformation is slow, bubbles and air hydrate crystals coexist over a depth range of several hundred meters [127]. The AMANDA scattering results led to predictions [128], later confirmed [122], that all bubbles have transformed into the solid phase at 1500 m, and that at greater depths the optical properties in the visible region depend almost solely on the concentration of dust in the ice.

[1]The Greenland Ice Sheet Project (GISP) was a decade-long project to drill ice cores in Greenland, GISP2 was the follow-up project.

3.3.1 Scattering

Scattering of photons off the dust particles in the ice is, hence, the predominant complication in the reconstruction of particle track directions from the patterns of Cherenkov light. The scattering occurs predominantly on submillimeter-sized air bubbles and micron-sized dust grains, but not on air hydrate crystals which have a refractive index very similar to that of ice.

The scattering length (or geometric scattering length) λ_s is particularly short (of the order of a few meters), and calculations based on Mie scattering theory show that air bubbles and dust grains tend to scatter photons in the forward direction, respectively with $\langle\cos\theta\rangle = 0.75$ [107] and $\langle\cos\theta\rangle = 0.94$ [119]. This process is hence not isotropic, and the scattering length, λ_s, does not represent the length scale over which the photon direction becomes randomised. It is more convenient to define then an effective scattering length:

$$\lambda_e = \frac{\lambda_s}{1 - \langle\cos\theta\rangle} \tag{3.3}$$

Hence when light propagates through a dense medium, the centre of the photon cloud moves along the incident direction at a decreasing step until it comes to a halt at one λ_e from the point of injection. The effective scattering length is thus for anisotropic scattering what the geometric scattering length is for isotropic scattering.

It is often more appropriate to discuss scattering in terms of the reciprocal of λ_e, the effective scattering coefficient $b_e = 1/\lambda_e$. From Mie scattering theory we expect a simple power law relation between effective scattering coefficient and wavelength of the form $b_e \propto \lambda^{-\alpha}$, with α close to one. This last relation was confirmed by measurements; longer wavelengths suffer less scattering on dust, while scattering on air bubbles can be considered wavelength independent. The results of these measurements are shown on the left of fig. 3.7, as a map of the variation of b_e in function of the depth and the wavelength.

3.3.2 Absorption

The absorptivity of a medium a describes the fraction of light which is absorbed in the material per unit length, and it is defined as the reciprocal of the absorption length (the distance at which the survival probability drops to $1/e$), $a = 1/\lambda_a$. It is well parameterised in the ice by a three components empirical model [129]

$$a(\lambda) = A_U e^{-B_U \lambda} + C_{\text{dust}} \lambda^{-\kappa} + A_{\text{IR}} e^{-\lambda_0/\lambda} \tag{3.4}$$

where the second term is due to insoluble dust particles in the ice, and the two exponentials characterise light absorption by the ice itself and are independent

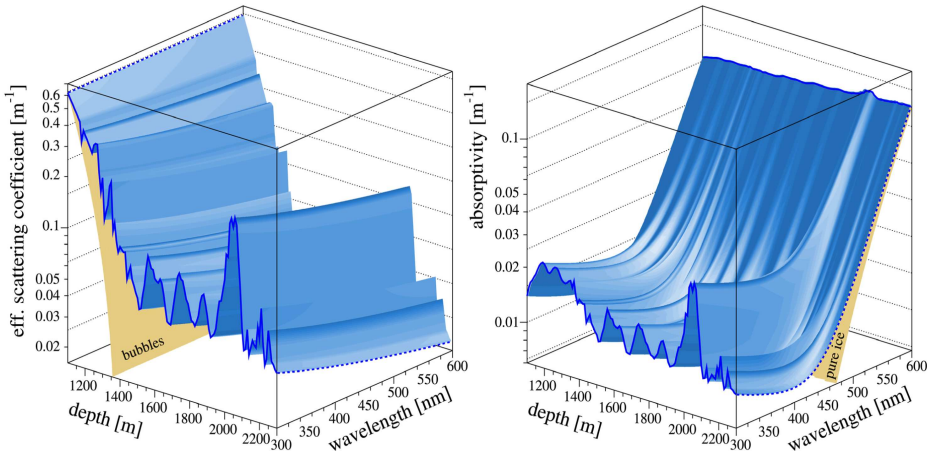


Figure 3.7: *Maps of optical scattering and absorption for deep South Pole ice. The depth dependence between 1100 and 2300 m and the wavelength dependence between 300 and 600 nm (left) for the effective scattering coefficient and (right) for absorptivity are shown as shaded surfaces, with the bubble contribution to scattering and the pure ice contribution to absorption superimposed as (partially obscured) steeply sloping surfaces. The dashed lines at 2300 m show the wavelength dependences: a power law due to dust for scattering and a sum of two components (a power law due to dust and an exponential due to ice) for absorption. The dashed line for scattering at 1100 m shows how scattering on bubbles is independent of wavelength. The slope in the solid line for absorptivity at 600 nm is caused by the temperature dependence of intrinsic ice absorption. Picture and caption text taken from [119].*

of dust content. The absorption increases exponentially at short wavelengths ($\lambda < 200$ nm) due to the electronic band structure of the ice crystal. Above 500 nm the absorption is dominated by excitation of the H_2O molecules. Between 200 nm and 500 nm pure ice is extremely transparent and the main component causing absorption is the dust. The fig. 3.7, on the right, shows the depth and wavelength profiles of the absorptivities derived from *in situ* light sources.

A global ice model, based on the measurements and fits, was derived with the absorptivity a and the effective scattering coefficient b_e dependent both on the depth z and wavelength λ. This model was used in the simulation of the photon propagation from the muon trajectory to the Optical Module (see next chapter).

3.3.3 Hole ice

In the holes (∼60 cm of diameter) in which the OMs of AMANDA have been deployed, the phase transition of the air bubbles to air-hydrate crystal happened

on much larger time scales than the detector operation. Therefore the scattering was dominated by these bubbles, leading to effective scattering lengths of less than one meter. However, since the hole diameter was small compared to the total travel distance of the photons between point of emission and OM, the main effect of the hole ice was to modify the angular photon acceptance of the modules: bubbles scattered photons from the dark side of the module back on the photocathode of the PMT. The additional arrival time delays from the scattering of the photons in the hole ice were negligible.

Measurements of the angular acceptance have been compared to simulations of different scattering lengths λ_h [130], with a best fit $\lambda_h = 50$ cm. The expected angular acceptance distribution for this value of λ_h was used in the photon propagation simulation.

- **Lieutenant Worf**:
"Computer simulation was not like this.
That delivery was very orderly."
- **Keiko O'Brien** [having a child]:
"Well, I'm sorry!"

Star Trek - The Next Generation,
episode: "Disaster"
Eugene Wesley Roddenberry
(1921-1991)

4 Experimental Data and Monte Carlo simulations

In the first part of this chapter we will focus our attention on the experimental data collected from 2001 to 2006 by AMANDA, which were used for this work. The data to be analysed were subject to a further check to verify the stability of the detector, which could be altered by some trigger issues.
In the second part we explain how we performed our Monte Carlo simulations, from generator to detector simulation level, both for the neutralino signal and for the atmospheric background.

4.1 Experimental Data

4.1.1 Data-taking period

The experimental data, used for our search, were taken between the 6th of March of 2001 and the 31st of October of 2006. During this span, only the data collected when the detector was proved to be stable were kept. The calibration and maintenance tasks on AMANDA were performed during austral summer (early November to mid February), as well as maintenance of power systems and other equipment of the South Pole station. In this period, the detector setup was continuously changing and therefore it was impossible to simulate reliably its response, hence the data collected during this time-frame were rejected.

Every day a *run* automatically started; it was a continuous period of data taking lasting up to a maximum of 24 hours, if not stopped before. During a run, the triggered events were written into a new file every 10 minutes roughly. Hence, every year more than 1 TB of raw data were stored to tapes and shipped out to the northern hemisphere for further processing.

As we explained in sec. 2.6, in a search for solar neutralino-induced events, the Earth can be used as a filter against upward-going muon background. In this way a signal source below the horizon will be easier to disentangle from the

downward-going muon background. Hence, the experimental data we used for our analysis concerned the period when the Sun was below the horizon, i.e. when its zenith $\theta_\odot > 90°$. The other data registered when the Sun was above the horizon ($\theta_\odot < 90°$) were used also in this analysis but as background sample for the optimisation of the event selection (see secs. 4.1.4 and 5.5).

Since the Earth's axis is tilted with respect to the ecliptic by approximately 23 degrees 27 minutes, the Sun does not set at high latitudes during the summer at the South Pole, where the time span with the Sun above the horizon reaches about 186 days. During these six months the Sun spends the days constantly "moving around" the horizon, reaching its highest circuit of the sky at the summer solstice. At this extreme latitude, one usually refers to these six months of daylight as *polar day*, while the remaining six months of the year, when the Sun is below the horizon, are referred to as *polar night*.

Figure 4.1 shows the evolution of the solar zenith angle from the year 2001 to 2003 (left to right) and fig. 4.2 from the year 2004 to 2006 (top to bottom) during the AMANDA data-taking period. The black markers in fig. 4.1 show the period when the Sun was below the horizon, while the grey ones show when the Sun was above[1] the horizon. Along the same line, the red markers in fig. 4.2 show the period when the Sun was below the horizon, while the black ones when the Sun was above the horizon. On both figures, only the runs considered for the present search are shown; as we will explain in sec. 4.1.3, the data which picked out unstable detector conditions were further disregarded from the analysis.

4.1.2 Detector live-time

As we outlined at the end of sec. 3.2.2, when the trigger logic accepted an event, the read out and storage processes took around 2.2 μs; this gap was the "dead" time of the detector, during which it could not record any additional events. As a consequence, the total time of data-taking as read from a wall-clock (t_{obs}) overestimated the period in which the detector was actually ready to detect events (t_{live}). Apart from the initial 2 μs, the time interval between consecutive events followed an exponential distribution; then from its slope it was possible to determine the actual event rate R_{true}. The comparison with the observed rate R_{obs} permitted then to find the dead-time $\mathcal{D} = 1 - R_{\text{obs}}/R_{\text{true}}$ of the DAQ. This procedure was repeated for every experimental data file [131], leading to the calculation of the live-time

$$t_{\text{live}} = t_{\text{obs}} \times (1 - \mathcal{D}) \tag{4.1}$$

[1] As we will explain in the next chapter, the experimental data 2001-2003 up to the filter level two, come from another analysis [111], in which the data were limited to the interval $80° < \theta_\odot < 113.5°$.

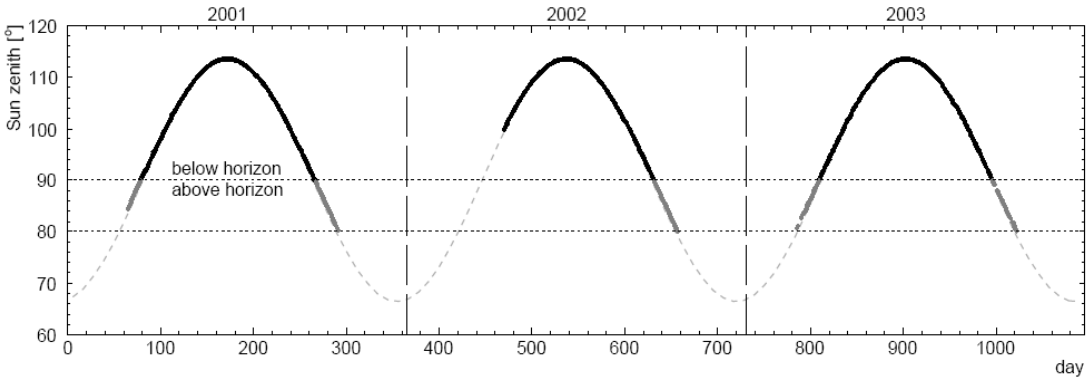


Figure 4.1: *Annual variation of the Sun's zenith at the South Pole from 2001 to 2003. The black markers represent the zenith evolution during the austral winter data-taking, the grey ones the data-taking during the austral summer. The dotted grey line completes the picture from Jan 1st 2001 until December 31st 2003. Picture taken from [111].*

The statistical error from the exponential fit to the time gap distribution translated to an error of less than 0.1% on the live-time of a data file.

4.1.3 Data stability check

Detector maintenance, test runs and calibration work were performed mainly in the austral summer, consequently less regular data-taking were made in this period. After the end of the summer, with the concurrent closing of the station, the achievement of stable data-taking conditions laid with the winter-over team and their northern collaborators. The performance of the global detector (trigger rates, dead-time, ...), as well as of the individual optical modules (dark noise rates, TDC/ADC information, ...), were monitored on-line via a web-interface which produced the relevant plots[1].

We already discussed in sec. 3.2.3 about the string trigger, which was installed in 2001; its configuration was set in the first part of the year to a 6/9 coincidence on the inner four strings and a 7/11 coincidence on the remaining strings. However, later that year, precisely on June 29th 2001, these settings were relaxed to 6/9 on all strings, causing an additional 20 Hz in trigger rate. This non-uniformity in the 2001 string trigger settings was compensated in the offline data analysis by means of a retriggering procedure [111] (see also sec. 5.2).

The events collected during the first part of 2002 were excluded from this work (see fig. 4.1), since the string trigger settings changed again in order to

[1]Part of the data were transferred to the northern hemisphere via a high bandwidth TDRS (Tracking and Data Relay Satellite) link.

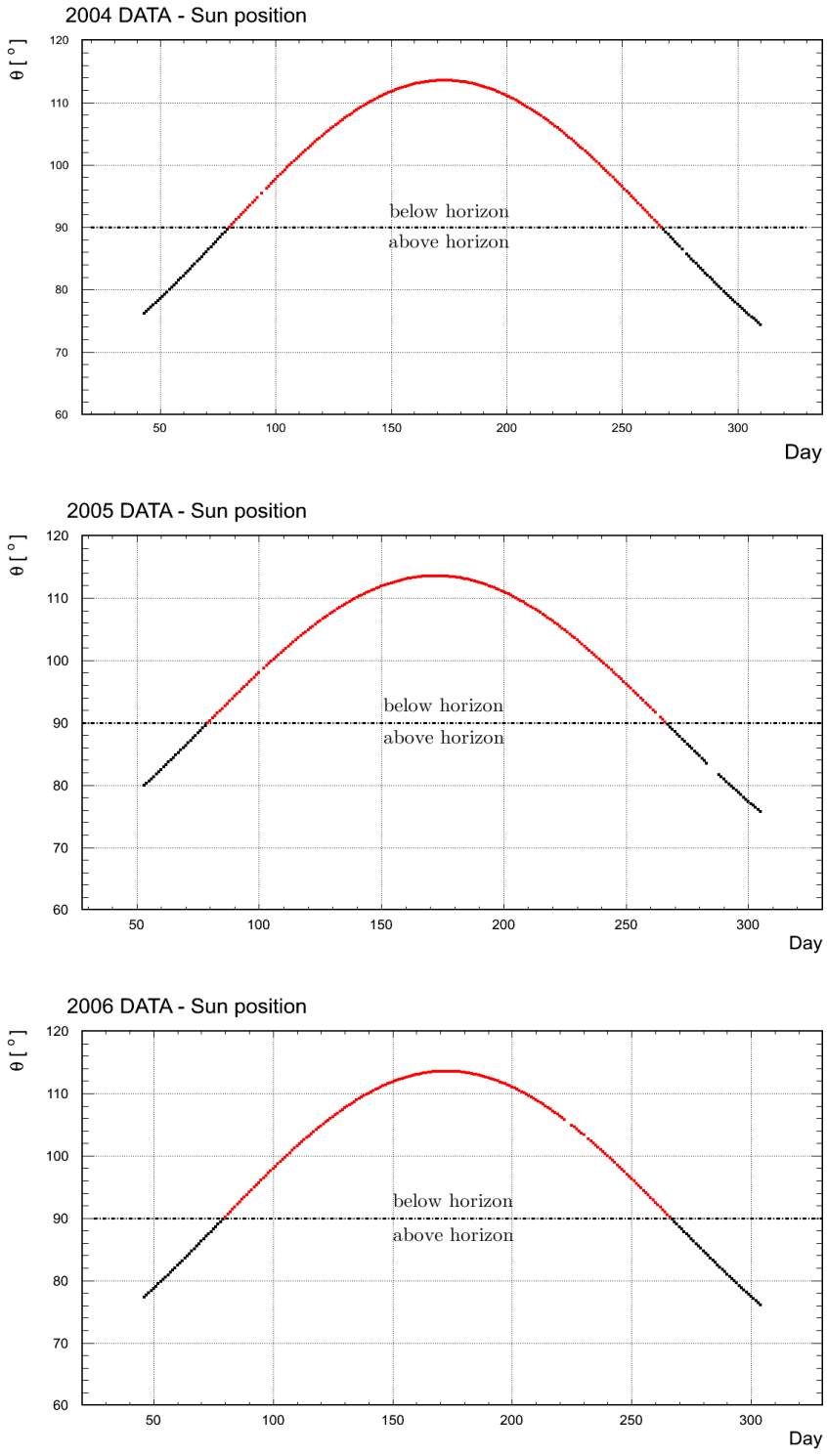


Figure 4.2: *Annual variation of the Sun's zenith at the South Pole of 2004 (top), 2005 (middle) and 2006 (bottom). The red markers represent the zenith evolution during the austral winter data-taking, the black ones the data-taking during the austral summer.*

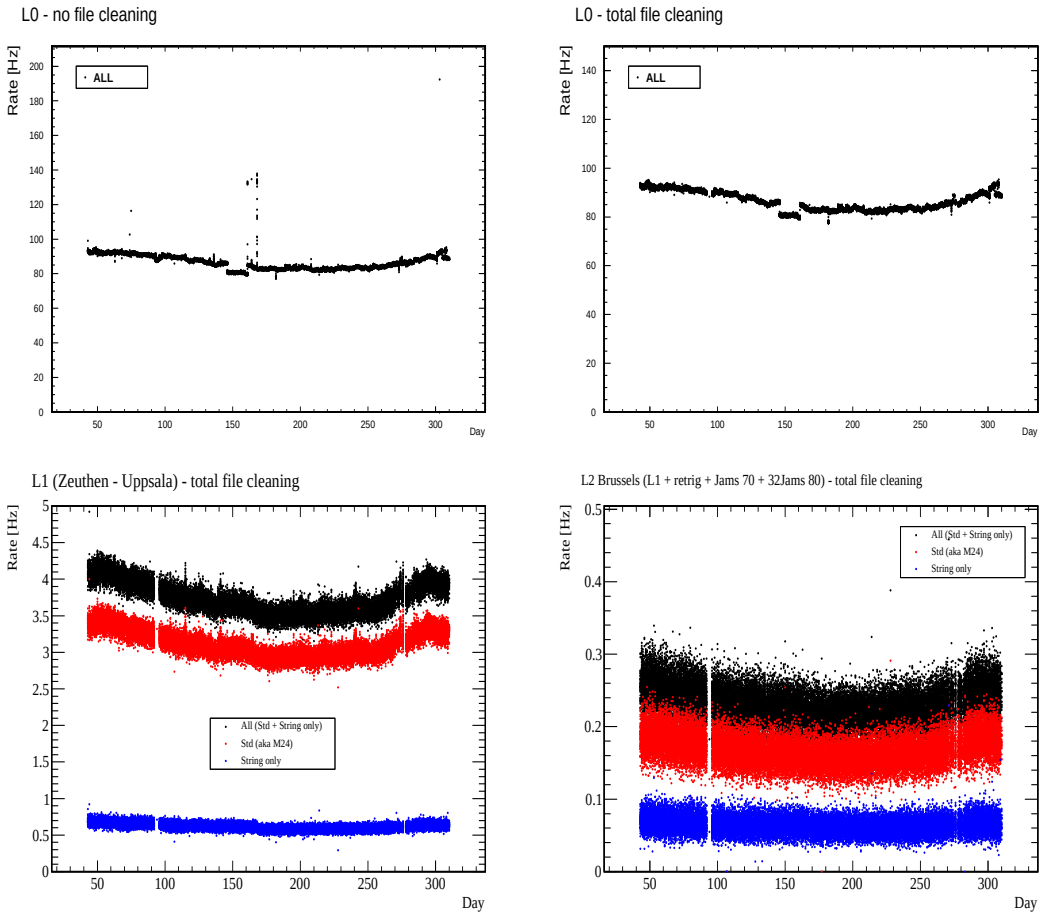


Figure 4.3: *Event rates (corrected for the dead-time) at different filter levels for the year 2004. At the top the trigger level before (left) and after (right) applying the total file cleaning, i.e. the standard plus some extra file selection. At the bottom the total file cleaning applied at Level 1 (left) and at Level 2 (right).*

keep the trigger rate and the detector dead-time fairly reasonable. From April 15th 2002 onwards, the string trigger configuration was set to 6/9 on strings 1-4 and 7/11 on strings 5-19, with the introduction of the downscaling; these settings were preserved throughout the rest of 2002 up to 2006.

As we will find out in the following of this work, the string trigger will play an important role in our low energy analysis; hence we should be sure that it did not introduce any instabilities to the data-taking. We refer to the analysis described in ref. [111], performed on AMANDA data from 2001 to 2003, for the stability check and for the lists of good runs of these first three years.

The existing standard lists of files with good runs, compiled by various analyses [100, 132, 133], for the remaining three years (2004, 2005 and 2006), were not

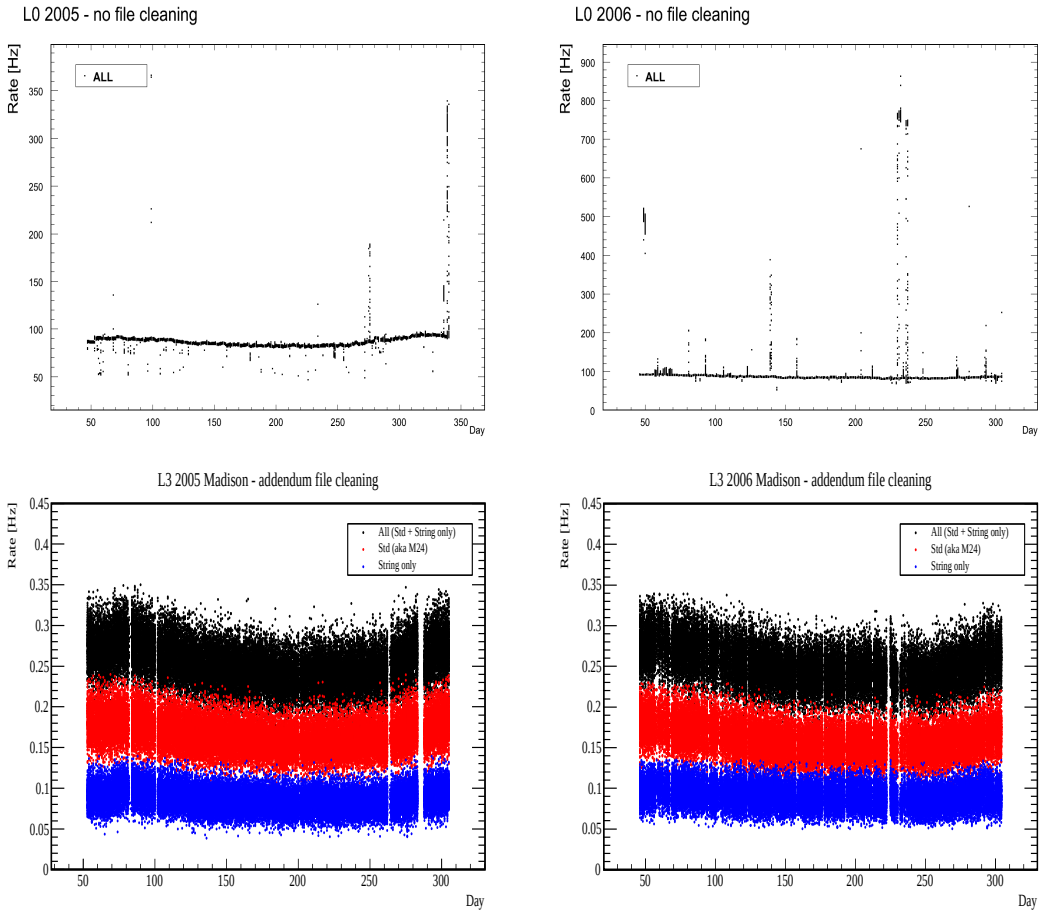


Figure 4.4: *Event rates (corrected for the dead-time) at trigger level for the year 2005 (top left) and the year 2006 (top right) before the file cleaning. At the bottom the event rates at Level 3 after the total file cleaning, i.e the standard plus some extra file selection, for the year 2005 (bottom left) and 2006 (bottom right).*

sufficient to remove all the instabilities noticed at trigger level (see fig. 4.3 top left and fig. 4.4 top half) and even at further filter level[1]. Hence we performed an extra stability check by requiring a smooth evolution of the event rate after the trigger level and after the first and the second selection level, thus to identify and remove additional unstable periods. The top half of fig. 4.3 shows the evolution of the global trigger rate (referred to as L0, see sec. 5.5.2) for 2004 before and after the total file cleaning; the bottom left of fig. 4.3 shows the rate evolution of the events passing both the first level selection criteria (also L1, see sec. 5.5.3.1) and the retriggering requirements after the total file cleaning. At L1 with retriggering (L1+retrigger, see secs. 5.2.2and 5.5.3.1) three sets of events were distinguished:

[1] For filter level definition see sec. 5.5.

all triggers (black markers), the events having only a string trigger flag (blue markers) and the events satisfying the multiplicity trigger and possibly another trigger as well (red markers). The right bottom part of fig. 4.3 shows the rate evolution of the events passing the second level selection criteria (also L2, see sec. 5.5.3.2), where the colours of the markers have the same meaning as the ones of L1+retrigger plot. On the same line, the top half of fig. 4.4 shows the evolution of the global trigger rate for 2005 (left) and 2006 (right) before the file cleaning. On the same picture, on the bottom it is showed the rate evolution of the events passing the level selection criteria called Level 3[1] for 2005 (left) and 2006 (right), after the total file cleaning. The event rates were corrected for the dead-time of the detector.

Eventually, a total of 7.7 days of live-time were further removed from the standard good runs lists of 2004-2006, and we concluded that after our total file cleaning no sudden, non-physical jumps in the event rate remained and that the selected data samples were safe for analysis purposes. Besides, the string trigger performance was at least as stable as that of the multiplicity trigger.

4.1.4 Final 2001-2006 data sets

We already mentioned, at the beginning of this chapter, that the total sets of data files taken on for further analysis, were divided in two samples (see figs. 4.1 and 4.2). The data sample, which refers to the Sun's zenith $\theta_\odot < 90°$, offered little hope for the detection of high-energy neutrinos from the Sun, due to the difficulty to disentangle downward-going neutrinos from atmospheric muons. Hence, this data sample was used instead for the optimisation of a multivariate background-rejecting selection criteria, called Boosted Decision Tree (BDT, see sec. 5.5.5), and therefore we called it *BDT optimisation sample*. The rest of the data set, which refers to the Sun's zenith $\theta_\odot > 90°$, was exploited at the end of our analysis, when the optimised event selection was applied and the final results calculated; this data sample was called *analysis sample*.

At trigger level, the experimental data sets were completely dominated by atmospheric muons, whose event rate depends on the interaction length of $\pi-$ and $K-$mesons in the Earth atmosphere (see sec. 2.6). We expected then, since this depends on the density and temperature of the atmosphere[2] above the South Pole, a seasonal variation of the trigger rates (higher in the austral summer, lower during the winter), as it is shown in figs. 4.3 and 4.4. However, fortunately, the

[1] As we said, the definition of the filter levels will be explained in sec. 5.5, however, for clarity's sake, we anticipate here that the filter L2 of 2001-2004 is similar to the filter called Level 3 of 2005-2006. The explanation of this latter is also outlined in sec. 5.5.3.2.

[2] The connection between atmospheric temperature and atmospheric muon rates gave the opportunity of study physics of the Earth atmosphere with AMANDA [134, 135].

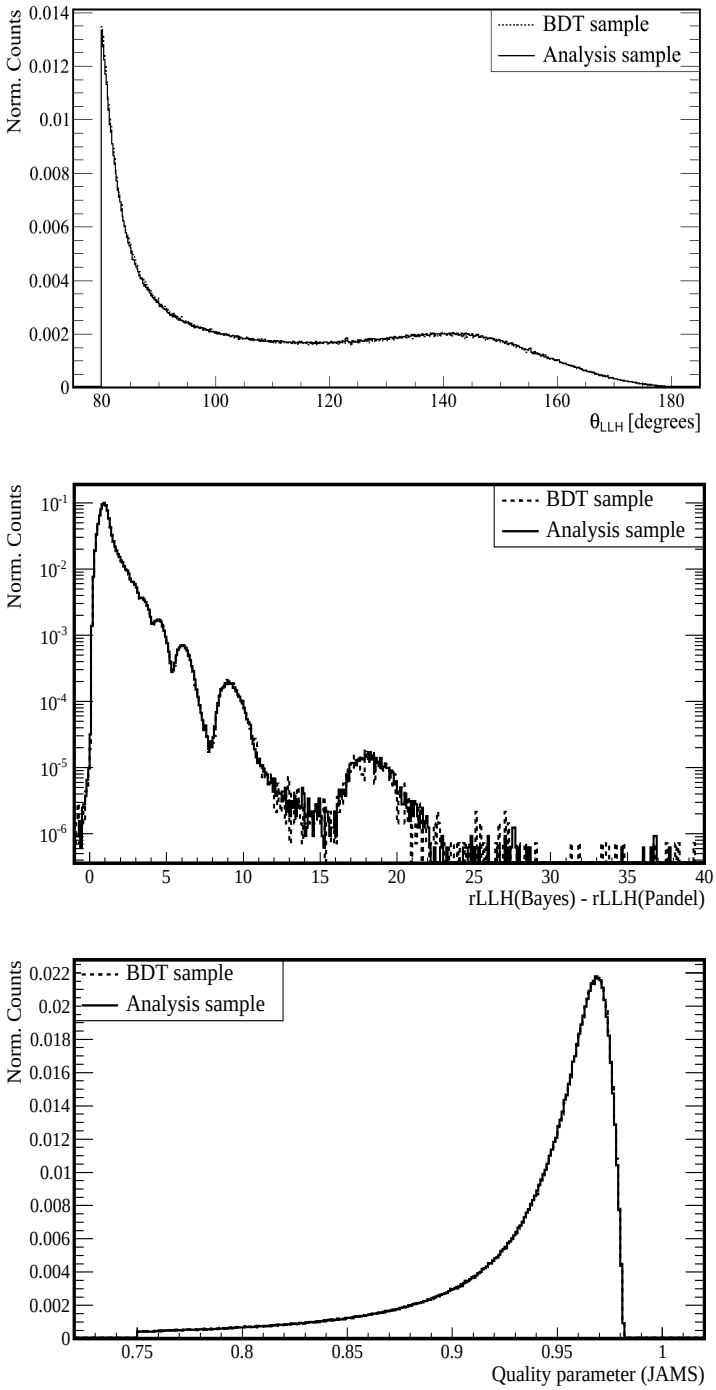


Figure 4.5: *Normalised distributions of some observables for experimental data at filter L2 (see sec. 5.5), for the BDT optimisation and for the analysis sample.*

Year	Day range	Run range	t_{obs} [d]	$\mathcal{D}$ [%]	t_{live} [d]	$N_{\text{trig}}^{(\times 10^9)}$
2001	6 Mar - 19 Mar	3145 - 3158	39.0	22.8	30.1	0.337
	24 Sep - 19 Oct	3370 - 3397				
2002	24 Sep - 20 Oct	5885 - 5914	25.6	12.5	22.4	0.204
2003	24 Feb - 19 Mar	6940 - 6984	33.0	15.8	27.8	0.252
	24 Sep - 20 Oct	7310 - 7341				
2004	12 Feb - 19 Mar	7800 - 7836	64.8	16.3	54.3	0.506
	24 Sep - 5 Nov	8042 - 8090				
2005	22 Feb - 19 Mar	9155 - 9193	57.6	15.6	48.1	0.445
	24 Sep - 1 Nov	9464 - 9520				
2006	15 Feb - 20 Mar	9785 - 9819	60.6	16.3	50.7	0.462
	25 Sep - 31 Oct	10063 - 10108				
2001-2006			280.3	16.7	233.4	2.206

Table 4.1: *For the BDT optimisation data sample ($\theta_\odot < 90°$) the day range, the run range, the observation time, the dead-time, the live-time and the number of the triggered events are shown.*

characteristics of the additional atmospheric background events in the summer were not different from those collected in the winter. This trend was verified at various filter levels in the analysis and for different observables (see e.g. fig. 4.5). We concluded that the atmospheric background in both data samples behave identically and we could safely use the BDT optimisation sample in the event selection procedure.

Year	Day range	Run range	t_{obs} [d]	$\mathcal{D}$ [%]	t_{live} [d]	$N_{\text{trig}}^{(\times 10^9)}$
2001	19 Mar - 23 Sep	3159 - 3369	181.0	21.3	142.5	1.456
2002	15 Apr - 23 Sep	5634 - 5884	124.5	13.4	107.8	0.921
2003	20 Mar - 23 Sep	6985 - 7309	156.8	15.0	133.3	1.146
2004	19 Mar - 24 Sep	7837 - 8041	166.4	15.2	141.1	1.219
2005	19 Mar - 24 Sep	9194 - 9463	176.5	15.8	148.6	1.291
2006	21 Mar - 24 Sep	9820 - 10062	164.5	15.8	138.6	1.213
2001-2006			969.7	16.3	811.9	7.246

Table 4.2: *For the analysis data sample ($\theta_\odot > 90°$) the day range, the run range, the observation time, the dead-time, the live-time and the number of the triggered events are shown.*

The total live-time of the various samples are found in tabs. 4.1 and 4.2, along with the dead-time and other useful information. At the end we got in

total 811.9 days of live-time for the 2001-2006 analysis sample and 233.4 days for the BDT optimisation sample. The downscaling of the string trigger in 2002-2006 reduced the dead-time from about 20% to about 15%, and obviously the trigger rate too. The issues with string trigger settings in the beginning of 2002 resulted in considerably lower live-times for that year.

4.2 Monte Carlo simulations

Monte Carlo[1] simulations are successfully used in data analysis in pursuit of modelling the known backgrounds and the hypothetical signal which one is looking for. The quality of the simulated data which aim to describe correctly the experimental data, even after some event selection, give us an indication how well the background and the detector response are understood.

Hence, for a start, the physics processes and the detector hardware responses (previously described in chapters 2 and 3) have to be taken into account. To that end, the AMANDA simulation software package is the frame in which several codes were developed and chained together, which coped from the neutrino generation up to its detection. This simulation chain was divided into three main parts: *event generators* (sec. 4.2.1), *muon propagation* (sec. 4.2.2) and *light and detector response simulation* (sec. 4.2.3). In the generation step, an incoming lepton interacts in the vicinity of the detector, which can be induced by neutralino signal or by cosmic ray background. Then, in the next step, the outgoing lepton (we remind that in this work we handle only muons) was propagated through the ice surrounding the detector, and secondary particles were produced with the energy loss. Finally, the Cherenkov photons radiated by leptons and their secondaries were traced, and the resulting detector response to the light was simulated.

Every geometrical information which was written in any specific *f2000* file format [137], was referred to the AMANDA coordinate system. In this frame the origin of the coordinates was located close to OM 69[2] on string 4, at a depth of 1730 m below the surface (see fig. 3.2). It was a right-handed orthogonal system, with the Y-axis pointing towards Greenwich and the Z-axis vertically up, towards

[1] The term "Monte Carlo method" was coined in the 1940s by physicists working on nuclear weapon projects in the Los Alamos National Laboratory; it is referred to the gambling house in Monaco because one of the researchers had an uncle who would borrow money from relatives because he "just had to go to Monte Carlo" [136].

[2] Historically, the origin of the system was defined as the location of OM 70 on string 4. Then some more refined geometry determinations indicated that this OM was actually a bit away from the origin. However, it was decided not to change the absolute position of the origin of the frame, thus the OM 69 turned in the one closest to the origin.

the sky. The spherical coordinates (θ, ϕ)[1] of a particle, like θ_ν and ϕ_ν, identified in this system a unit vector of the particle's track which pointed backwards from the direction of travel. Hence, vertically downward-going and upward-going tracks have $\theta = 0°$ and $\theta = 180°$ respectively.

The huge Monte Carlo mass production was carried out by the author of this work thanks to the **BEgrid** [138] and **SARA Matrix** [139] GRID facilities located respectively in Belgium and in the Netherlands. Besides, several scripts to chain and run the AMANDA software on the GRID system were developed also by the author of this work as well as other graphical scripts to make more user-friendly the interplay with the GRID system.

4.2.1 Event generators

Neutralino signal

The codes used to simulate the neutralino annihilation in the Sun, the lepton generation and propagation to the Earth up to its interactions close to the detector, were `WimpSim 2.06` [140] in combination with `WimpEventF2k` [141].

The `WimpSim` code is made up of two parts, called `WimpAnn` and `WimpEvent`. In the first step, `WimpAnn` simulated the neutralino annihilation in the core of the Sun[2] with the support of `PYTHIA` code [142] and of `DarkSUSY` code [143], where the Standard Solar Model [144] is used to extract the solar density.

The neutrino (antineutrino) interactions in the Sun were simulated with the fast Monte Carlo called `nusigma 1.15` [145], which uses for the NC and CC interactions the CTEQ6-DIS structure functions [96] for protons and neutrons; hence, having both as a target, it didn't assume any isoscalar particle.

The full three-flavour neutrino oscillations are also included in these simulations and the following parameters were used [146]:

$$\theta_{12} = 32.215° \qquad \theta_{13} = 6.875° \qquad \theta_{23} = 44.977°$$

$$\Delta m^2_{12} = 8.2 \cdot 10^{-5}\,[\mathrm{eV}^2] \qquad \Delta m^2_{13} = 2.4 \cdot 10^{-3}\,[\mathrm{eV}^2]$$

The neutrino regeneration[3] from τ decay, arising in CC interactions on the way out of the Sun, was also included. Eventually, `WimpAnn` propagated neutrino

[1] In the AMANDA system these angles are sometimes improperly called zenith and azimuth of the particle. For clarity's sake, in this work we will simply refer to these angles as "theta" and "phi" of the particle.

[2] The code can handle also other WIMP candidate like LKP, and WIMP annihilation in the centre of the Earth as well.

[3] This process reduces the probability that the ν_τ are absorbed, increasing the total neutrino flux to the detector.

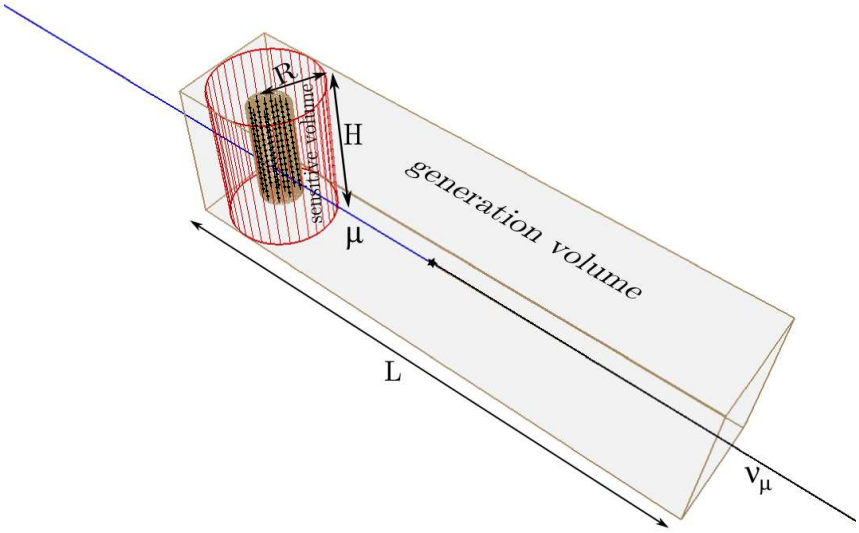


Figure 4.6: *The* `WimpEventF2k` *generation volume (in grey), which encompasses the sensitive volume (in red), which in turn includes the AMANDA detector.*

events (including all neutrino state amplitudes and phases) to a distance of 1 A.U. from the Sun.

In the next step, `WimpEvent` propagated the neutrinos, left by `WimpAnn`, through the Earth to a given detector location, without defining the neutrino interaction point position, projecting them out on flavour eigenstates there and including the neutrino oscillations along their path. The Monte Carlo used to simulate the neutrino-nucleon interaction in the Earth was again `nusigma`. `WimpEvent` attached also a time stamp (time of the year) and a weight to each event, since the interactions were forced to occur[1].

Finally, the `WimpSim` software produced the output text files which were filled with information (angle, energy, weight) about the incoming neutrino, the outgoing lepton and the hadronic shower.

The format of the output files produced by `WimpSim`, did not match the specific AMANDA *f2000* file format. Hence, the `WimpEventF2k` code was written to convert the `WimpSim` events to the *f2000* standard, so that they were suitable for particle propagation and detector simulation performed with the other AMANDA codes. Another task of `WimpEventF2k`[2] was to distribute efficiently the interaction points, randomly in a *generation volume* which encompassed the physical detector. This generation volume V_{gen} could be either a vertical *cylin-*

[1] See sec. 4.2.4 for a thorough discussion about event weight.

[2] The algorithm used here is based on a previous AMANDA code called `GenN++` [147], developed by the Author of this thesis.

der or a so-called *muon-box*. In our simulation setting we have chosen the latter option, thus our generation volume was a box around a *sensitive volume*[1], which was a cylinder with a model-dependent height H and radius R (see tab. 4.3). This box was rotated and aligned with the incoming neutrino direction, and its length L was stretched towards this direction. The length could be set as *fixed* or as *flexible*, in order to speed up the simulation forcing the interaction near the detector (this was also taken into account in the weighting scheme).
In the flexible case, which was our setting, the length was a function of the neutrino spherical coordinate θ_ν and of the maximum muon range $R_\mu^{\max}$. Hence, once fixed H and R, it was defined as follows

$$L \equiv L(E_\mu, \theta_\nu) = R_\mu^{\max}(E_\mu) + L_{\min}(\theta_\nu) \tag{4.2}$$

where $R_\mu^{\max} = \mathcal{K} \cdot R_\mu$ was obtained from the average muon range R_μ (see eq. 2.11) multiplied by $\mathcal{K}$, which was a scaling factor to make the range conservative [105]. The term $L_{\min}(\theta_\nu)$ has been added in order to take into account that neutrino interactions could take place inside or outside the sensitive volume. Summarising, the geometrical generation volume was then defined (still fixing R and H) as

$$V_{\text{gen}}(E_\mu, \theta_\nu) = A(\theta_\nu) \cdot L(E_\mu, \theta_\nu) \tag{4.3}$$

where $A(\theta_\nu)$ was the projection of the sensitive volume on a plane orthogonal to the neutrino direction. All this explanation about the geometrical construction of the generation volume, can be visualised in fig. 4.6.

In the MSSM, several parameters related to the neutralino mass are still unknown (see secs. 1.3.1 and 1.3.2). From accelerator searches and relic density constraints from WMAP data a lower bound and an upper bound on the neutralino mass was set to 47 GeV and to 10^4 GeV respectively (see sec. 1.4.1). Hence, to cover this range the following seven neutralino masses were simulated:

$$M_\chi = 50, 100, 250, 500, 1000, 3000, 5000\,[\text{GeV}]$$

We have already discussed in sec. 2.1 that neutrinos are produced from several neutralino annihilation products with unknown branching ratios. Therefore separate simulations were needed to span over the extreme cases where the channels produce the softest and hardest possible neutrino energy spectra. Hence, for each neutralino mass, we have simulated two annihilation channels: the soft $b\,\bar{b}$, and the hard $\mathrm{W}^+\,\mathrm{W}^-$ (for 50 GeV the hard channel is $\tau^+\,\tau^-$). Besides, an energy threshold of 15 GeV on the incoming neutrino and a threshold of 10 GeV on the coming outgoing muon were set; these energy cut-offs will be properly

[1] It is called sensitive volume because the light produced into it can still reach the OMs.

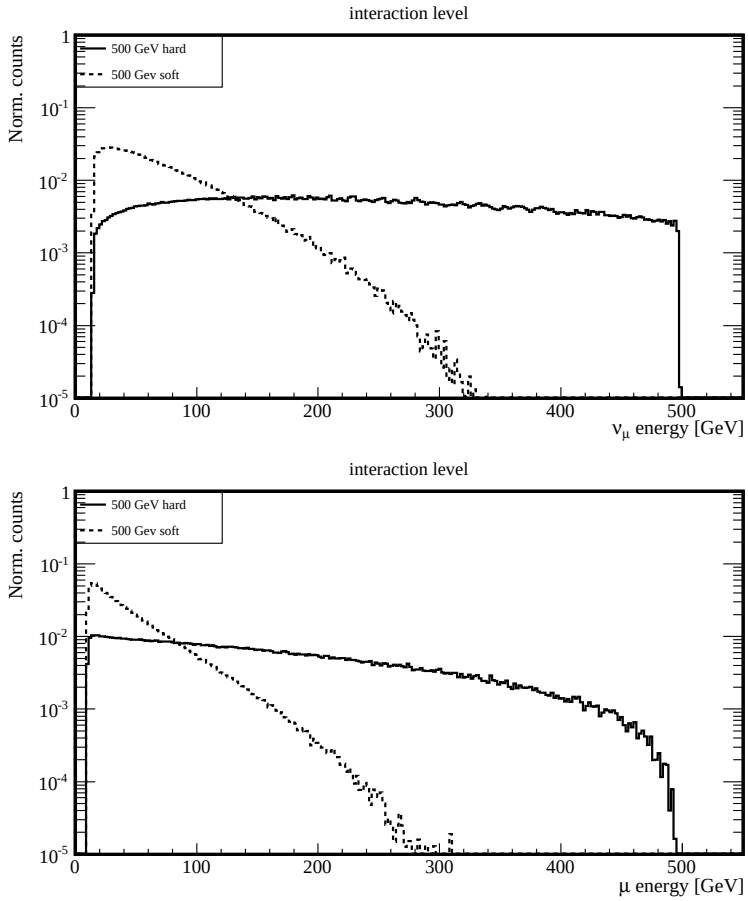


Figure 4.7: *The normalised energy spectra for 500 GeV neutralino-induced neutrinos from hard (soft) channel, are drawn in solid (dotted) line on the top. On the bottom, the normalised energy spectra of the outgoing muons are drawn in solid (dotted) line.*

considered when the muon-flux will be calculated (see sec. 6.3.3). The different hard and soft normalised energy spectra[1] at interaction level, for 500 GeV neutralino-induced neutrinos, are shown on the top of fig. 4.7, in solid and dotted line respectively. Whereas, on the bottom of the same picture, the normalised spectra of the outgoing muons produced in the neutrino interactions are shown.

We have already remarked, at the beginning of this chapter, that it is very hard to disentangle downward-going sub-TeV neutrino events from atmospheric muons; consequently, the possibility to detect a neutrino signal from the Sun during daytime is extremely reduced. Hence, we have only generated neutrino

[1] In all the calculated spectra (both energy and angular distribution) the event weight was taken into account.

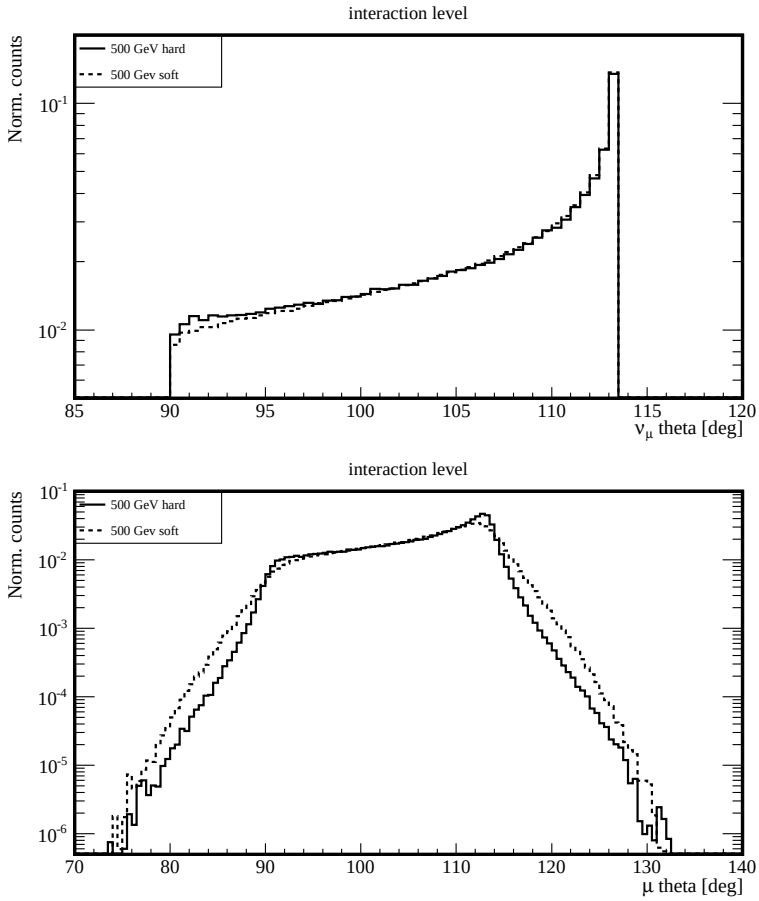


Figure 4.8: *The normalised angular spectra for 500 GeV neutralino-induced neutrinos from hard (soft) channel, are drawn in solid (dotted) line on the top. On the bottom, the resulting normalised angular spectra of the outgoing muons are drawn in solid (dotted) line.*

events when the Sun is below the horizon, where the solar zenith angle $\theta_\odot$ spans from 90° to 113.45°. The different normalised angular distributions at interaction level, for 500 GeV neutralino-induced neutrinos, are shown on the left of fig. 4.8, in solid and dotted line respectively. Whereas, on the right of the same picture, the normalised angular distribution of the outgoing muons produced in the neutrino interactions are shown.

In tab. 4.3 all the relevant numbers about our neutralino signal simulation setting, for each model (mass and channel), are shown, along with the number of generated and triggered unweighted events.

M_χ [GeV]	Channel	H_{sens} [m]	R_{sens} [m]	N_{gen}^{raw} ($\times 6 \cdot 10^3$)	N_{trig}^{raw}
50	soft	700	200	267634	935750
	hard	700	200	93085	833134
100	soft	700	200	103690	765086
	hard	700	200	21482	721922
250	soft	800	200	39786	686122
	hard	800	250	11391	662851
500	soft	800	250	34386	694771
	hard	800	250	8576	643654
1000	soft	800	250	15734	722276
	hard	800	250	8630	686228
3000	soft	800	250	29068	995868
	hard	800	250	10622	790753
5000	soft	800	250	19527	698836
	hard	800	250	9505	675694

Table 4.3: *For each model: H_{sens} is the height of the sensitive volume and R_{sens} is the radius, N_{gen}^{raw} is the number of the unweighted signal events generated, N_{trig}^{raw} is the number of unweighted signal events triggered by the detector.*

Atmospheric Muons

A code derived from `CORSIKA 6.500` [148], which was adapted and optimised to meet the AMANDA requirements, called `dCORSIKA` [149], was used to simulate the atmospheric muon background at the Earth surface. The `CORSIKA` code was first developed to perform simulations for the KASCADE experiment [150], to study the evolution and properties of extensive air showers in the atmosphere. The program allows to simulate interactions and decays of nuclei, hadrons, muons, electrons, and photons in the atmosphere up to energies of some 10^{20} eV. The simulation of the interactions of high energy cosmic rays with nuclei of the Earth atmosphere was indeed a rather challenging task, since the primary flux is not well known at the highest energies. Moreover, the development of the shower is a complicated process which involves several particles, which in turn are subjected to interactions and/or decays.

In our setting the high energy hadronic interaction model was simulated using, as option, the Monte Carlo called `SYBILL` [151], with the primary cosmic ray composition following the Hörandel parametrisation [152], from a $d\phi/dE \propto E^{-2.7}$ energy spectrum. The primary Monte Carlo samples were generated isotropically over the Southern hemisphere[1] ($\theta_p \in [0°, 90°]$) between 800 GeV and 10^{11} GeV. We

[1] We remind that atmospheric muons from the Northern hemisphere will never reach the detector.

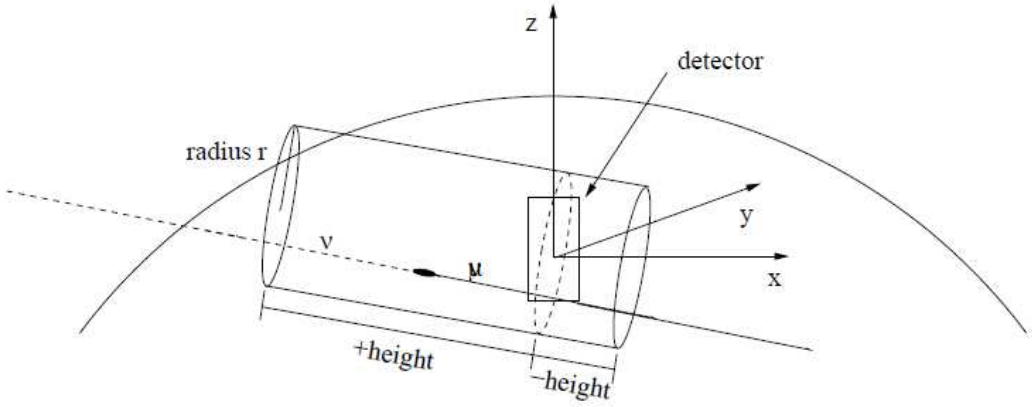


Figure 4.9: *In ANIS the coordinate system is anchored at the detector centre, with the Z-axis pointing away from the Earth centre. The cylinder represents the "final volume", in which potentially detectable neutrino interactions are simulated. Picture taken from [153].*

generated in total 10^{11} air showers which produced about 1.3×10^9 atmospheric muons; this corresponded to ~ 9.1 days of live-time, for each year of data-taking. The normalised theta distribution of the atmospheric muons at generator level is shown on the top of fig. 4.12 (dotted line), while their normalised energy distribution is shown on the bottom of the same figure (dotted line).

Atmospheric Neutrinos

The atmospheric neutrino background was simulated with the `ANIS 1.8.2` code [153], which is a general tool for generating all-flavour neutrinos in the energy range 10 GeV-10^{12} GeV. For each year, 5×10^7 events were simulated isotropically with $\theta_\nu \in [80°, 180°]$, since we assumed that downgoing atmospheric neutrinos will be completely removed at later filter stage. Those events were sampled in energy, between 10 GeV to 325 TeV, from a power law spectrum with $d\phi/dE \propto E^{-1}$. This allowed to use the same Monte Carlo sample, endowing a flux weights to each event to pass from an isotropic flux to some neutrino flux models, like atmospheric muon neutrinos, according to the Lipari parametrisation [154]. These neutrinos were then propagated through the Earth (neutrino oscillations of ν_μ to ν_τ were not taken into account) and eventually forced[1] to interact (CC or NC) near the detector. Hence, to optimise the computing time, the interaction vertices were spread in a cylindrical volume aligned with and extended to the incident direction of the neutrino, whose dimensions should not alter the shape of the triggered spectra. The chosen geometrical setting is then visualised in

[1] This is also taken into account in the weighting scheme (see sec. 4.2.4).

fig. 4.9, where a radius (r) of 375 m and a length of 15 km before the detector centre ($+height$) and 375 m behind ($-height$) were used.

The normalised theta distribution of the atmospheric neutrinos at generator level is shown on the top of fig. 4.13 (dotted line), while their normalised energy distribution is shown on the bottom of the same figure (dotted line). The normalised theta distribution of the muons from atmospheric neutrinos at generator level is shown on the top of fig. 4.14 (dotted line), while their normalised energy distribution is shown on the bottom of the same figure (dotted line).

4.2.2 Muon propagation

A three-dimensional propagation of the generated muons was simulated with the `MMC 1.4.3` code [105]. The simulation included all the processes we described in sec. 2.4, i.e. all continuous and stochastic processes concurring to the muon energy loss, which took place through four different media around the detector: air, firn, ice and rock.

The muon propagation in `MMC` was divided into three separate steps: 1) before entering, 2) inside, and 3) after leaving the sensitive detector region, being this time a vertical cylinder surrounding the physical AMANDA detector, with height and radius of 800 m and 400 m respectively.

In the first step, before entering the sensitive detector region, the muon energy was evaluated. Then, stochastic energy losses larger than 5% of the muon energy were simulated, whereas lower energy losses were handled with a continuous energy loss approximation.

In the second step, the simulation of muon propagation through the sensitive detector region took place; here the light from the muon and secondary particles had a good chance to be detected. Inside the sensitive volume, all secondary particles with energies above 500 MeV were processed separately, each giving a contribution of Cherenkov light, and kept in a file output for later processing stages.

In the final step, the stopping point of the muon was estimated from the average muon range.

4.2.3 Light and detector response simulation

Photon propagation

The next step in the simulation chain was the tracking of the Cherenkov light, produced by the muon and its associated secondaries, which was collected by the OMs. Hence, for a good performance, we needed to estimate correctly the number of photons and their arrival time at the OM, taking into account the scattering and absorption properties of the ice (see sec. 3.3).

The expected number of photons emitted by a muon per traversed meter is about $3{\cdot}10^4$ (see sec. 2.5); hence, to avoid high CPU-time processing, some devices were needed: unlike to track all photons individually, this was accomplished once for a large number of photons and track-OM configurations. The yielded probability density distributions for photon flux and arrival times, were then stored in look-up tables, also referred to as *photon tables*. This task was done by the `Photonics 1.54` photon tracking tool [155].

The photon tables were produced by generating photons from sources according to their wavelength and angular distributions; these sources were either the single Cherenkov emission along a muon track, or shower events of many short Cherenkov tracks. Then the photons were tracked through the ice, taking into account scattering and absorption modelled by the `AHA v2` ice model [156], which accommodates for the heterogeneous ice[1]. So, summing up, the photon tables contained the mean amplitude and hit time delay distributions (due to the photons scattering in the ice) as a function of the particle direction (θ, ϕ), the point of closest approach between the track and OM, ρ, the length from the track vertex to the point of closest approach, L, and the OM coordinates[2].

Detector simulation

The final step in the simulation chain was the detector response simulation (the read-out electronics and data acquisition), which was done with the `AMASIM-OPT5`[3] code [157].

Given a particle track, the expected photo-electron multiplicity $\langle n_{pe} \rangle$, for each OM in the array, was simulated by `AMASIM` using the photon tables made by `Photonics`. This $\langle n_{pe} \rangle$ value was then scaled with the PMT effective area $\mathcal{A}_{\text{eff}}$ and relative OM sensitivity $\mathcal{S}$, so that the number of hits was inferred by sampling a Poissonian distribution with mean $\lambda \equiv \langle n_{pe} \rangle \cdot \mathcal{A}_{\text{eff}} \cdot \mathcal{S}$. After the hit time delay distributions were evaluated, `AMASIM` added some random hits, due to dark noise, and afterpulse hits in the OMs; both effects were tuned for each OM.

The hit amplitudes were randomly picked from an experimentally measured single photo-electron response distribution, scaled to fit different OM types. If more than one hit was present in a module, the individual waveforms were summed up; the saturation effects of the amplifier for large pulses was also taken into account. The TDC and ADC electronics, as well as the delay of the pulses in the cables, were also simulated.

[1] This had the drawback that to reach necessary accuracy the photon tables became very large, more than 20 Gb.

[2] In this work the depth-resolution bin was set to 20 m and the θ bin to 10°.

[3] The tuned parameters developed for 2003 were used also for the other years of the analysis upon verification.

Although most features of the detector were simulated, some others were not, like the noise introduced by the crosstalk in twisted pair cables (see sec.3.2.1), or unusual behaviour of OMs like high bursts or sudden drops in dark noise, or even correlated noise due to bad weather conditions at the South Pole. Hence, pulses generated by these effects were identified and removed from the experimental and simulated data by a later procedure called "hit cleaning" (see sec. 5.2.1).

The hits in the event were then used as input to the trigger simulation, and concerning our analysis, as we explained in sec. 3.2.3, two patterns of trigger had to be simulated: the string trigger (downscaled by a factor of two from 2002 onwards) and the multiplicity trigger.

4.2.4 Event weights

The generation of neutralino signal and atmospheric neutrino events were performed optimising the required CPU-time processing, in order to enhance low statistics regions. However, this technique biased the generated angular and energy spectra; hence a proper weight for each event was needed to compensate for this effect.

Neutralino signal

The weight for the outgoing lepton w_i is constructed so that the volumetric flux Γ per annihilation results as follows:

$$\Gamma = \frac{1}{N_{Ann}} \sum_{i}^{N_{Ann}} w_i(E_\nu) \tag{4.4}$$

where N_{Ann} is the number of annihilations simulated, and w_i the neutrino cross section.

Since we simulated event-wise generation volumes $V_i(E_\nu, \theta_\nu)$, the number of the observed physical events should be calculated as follows:

$$N_{obs} = \sum_{i}^{N} \delta_i V_i(E_\nu, \theta_\nu) w_i(E_\nu) \quad \text{where} \quad \delta_i = \begin{cases} 1 & \text{event observed} \\ 0 & \text{event not observed} \end{cases} \tag{4.5}$$

and where N is the number of the generated events simulated (i.e the events at the Earth surface). If we consider a volume V_{gen} around the detector which encompasses all the volumes V_i, i.e. $V_{gen} \supseteq V_i$, we can then write:

$$N_{gen} = V_{gen} \sum_{i}^{N} w_i(E_\nu) \tag{4.6}$$

where N_{gen} represents the number of the generated physical events.

Atmospheric neutrino

The weight expression for atmospheric neutrino events results in:

$$w(E_\nu, \theta_\nu) = w^{\nu_\mu}_{atm.}(E_\nu) \times w_\sigma(E_\nu, \theta_\nu) \times w_{norm} \tag{4.7}$$

where $w^{\nu_\mu}_{atm.}$ is the normalised spectral shape for atmospheric muon neutrinos at the surface of the Earth, w_σ is the cross section, and $w_{norm} \equiv N \times t_{\text{live}}$ is a normalisation factor that takes into account the correct amount of events per year in the chosen volume (N) scaled to the experimental live-time (t_{live}).

The term $w^{\nu_\mu}_{atm.}$ can be replaced, thus to obtain other neutrino source spectra from a sample of atmospheric neutrino events, generated from a $E^{-\gamma}$ distribution; for instance a spectral shape $\Phi(E_\nu)$ can be recovered using $E^\gamma \Phi(E_\nu)$ as event weight. Non-isotropic sources can be recover as well adding a direction-dependent factor in the weighting order.

4.2.5 Summary plots

In this section some plots drawn from the Monte Carlo simulation are shown. These plots are normalised to one just to show the shape of theta (on the top) and energy (on the bottom) distribution of the particles (neutrinos and muons) at generator level, or better, at the interaction level (dotted line), and at trigger level (solid line). The fig. 4.10 shows the distribution of neutrinos from the 500 GeV hard neutralino model, while the fig. 4.11 shows the corresponding outgoing muon distributions. The fig. 4.12 shows the atmospheric muon distributions, while the fig. 4.13 shows the atmospheric neutrino distributions and fig. 4.14 the corresponding outgoing muon distributions.

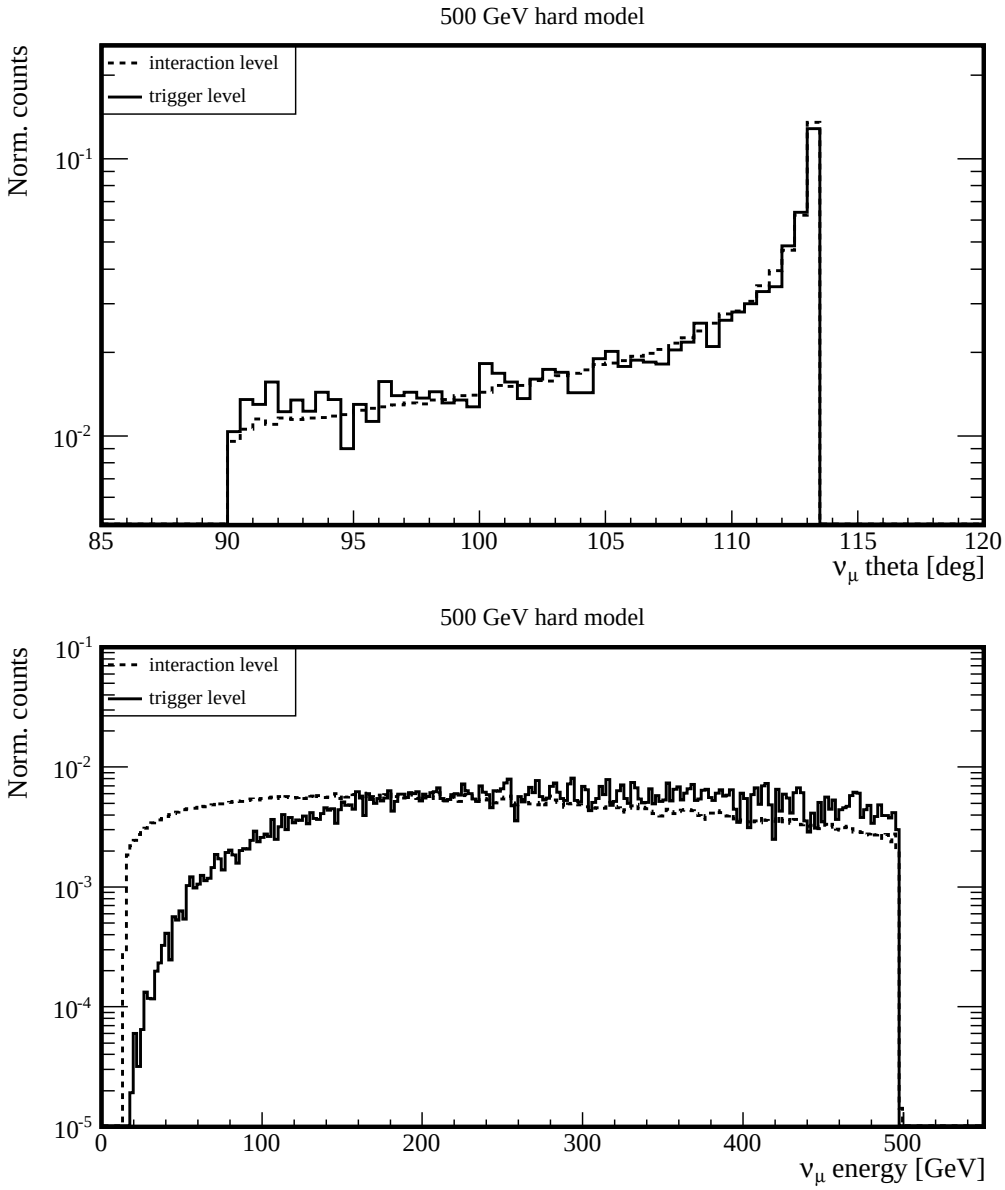


Figure 4.10: *The normalised angular (top) and energy (bottom) spectra of neutrinos from 500 GeV hard neutralino model, at interaction level (dotted line) and at trigger level (solid line).*

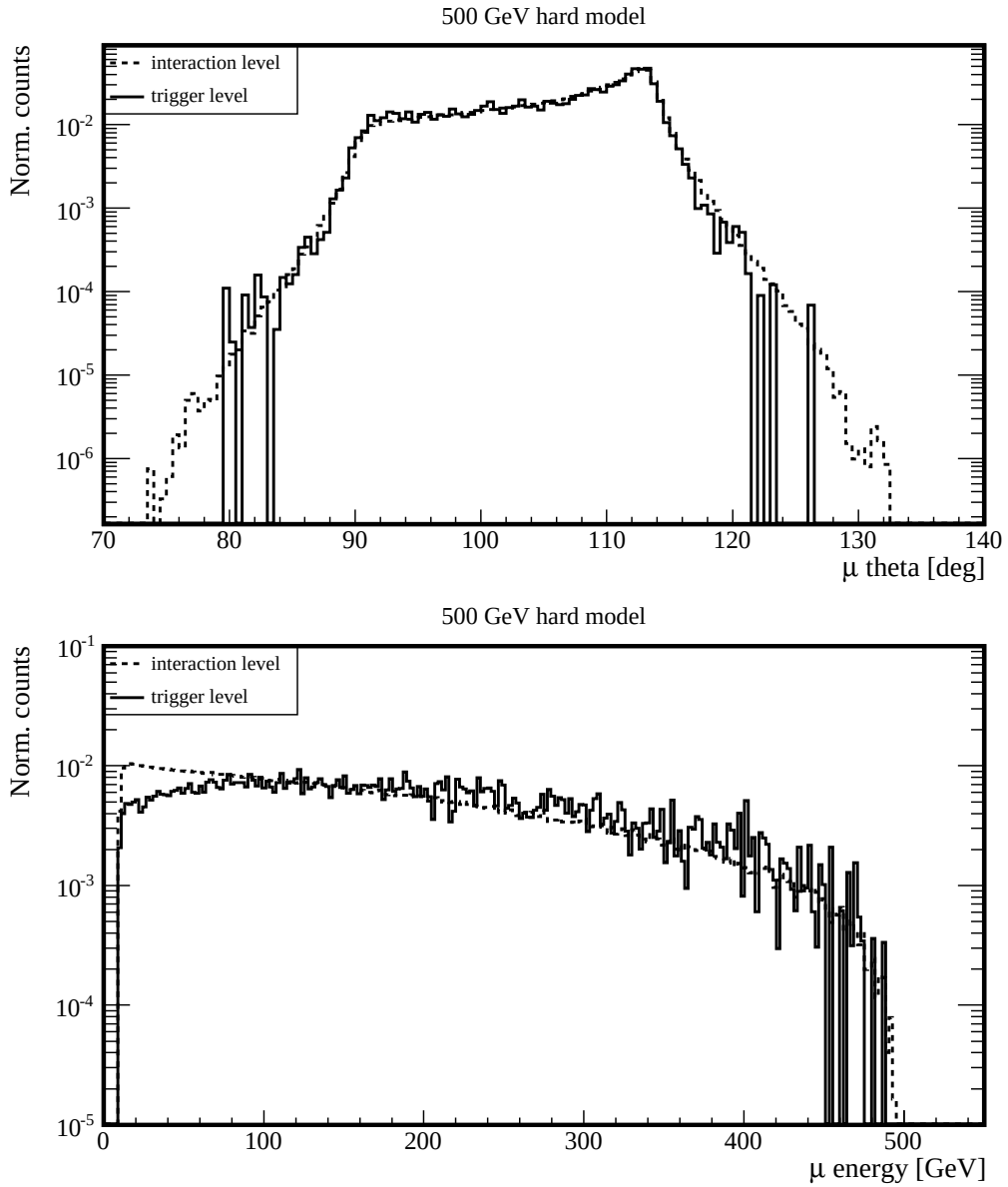


Figure 4.11: *The normalised angular (top) and energy (bottom) spectra of outgoing muons from 500 GeV hard neutralino model, at interaction level (dotted line) and at trigger level (solid line).*

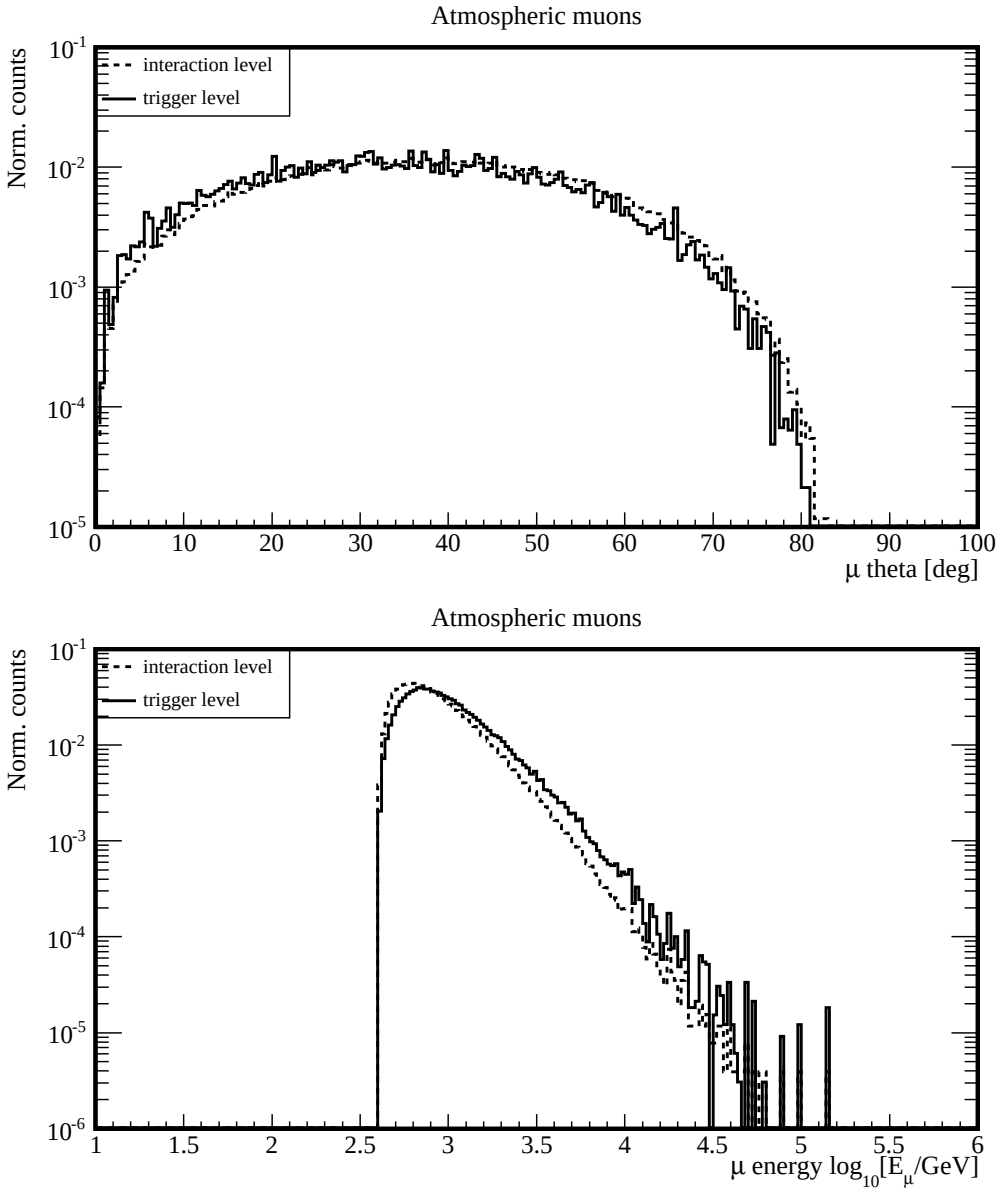


Figure 4.12: *The normalised angular (top) and energy (bottom) spectra of atmospheric muons, at interaction level (dotted line) and at trigger level (solid line).*

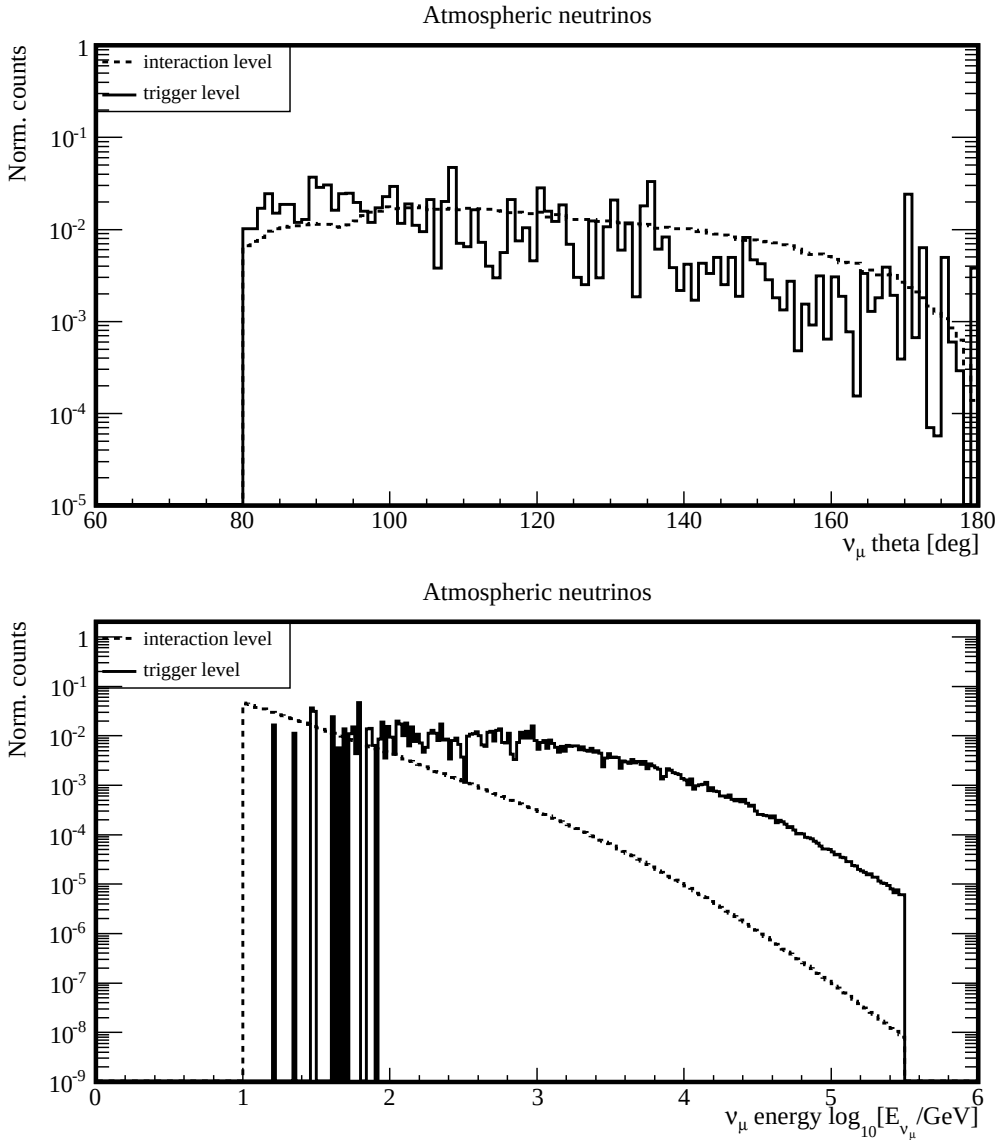


Figure 4.13: *The normalised angular (top) and energy (bottom) spectra of atmospheric neutrinos, at interaction level (dotted line) and at trigger level (solid line).*

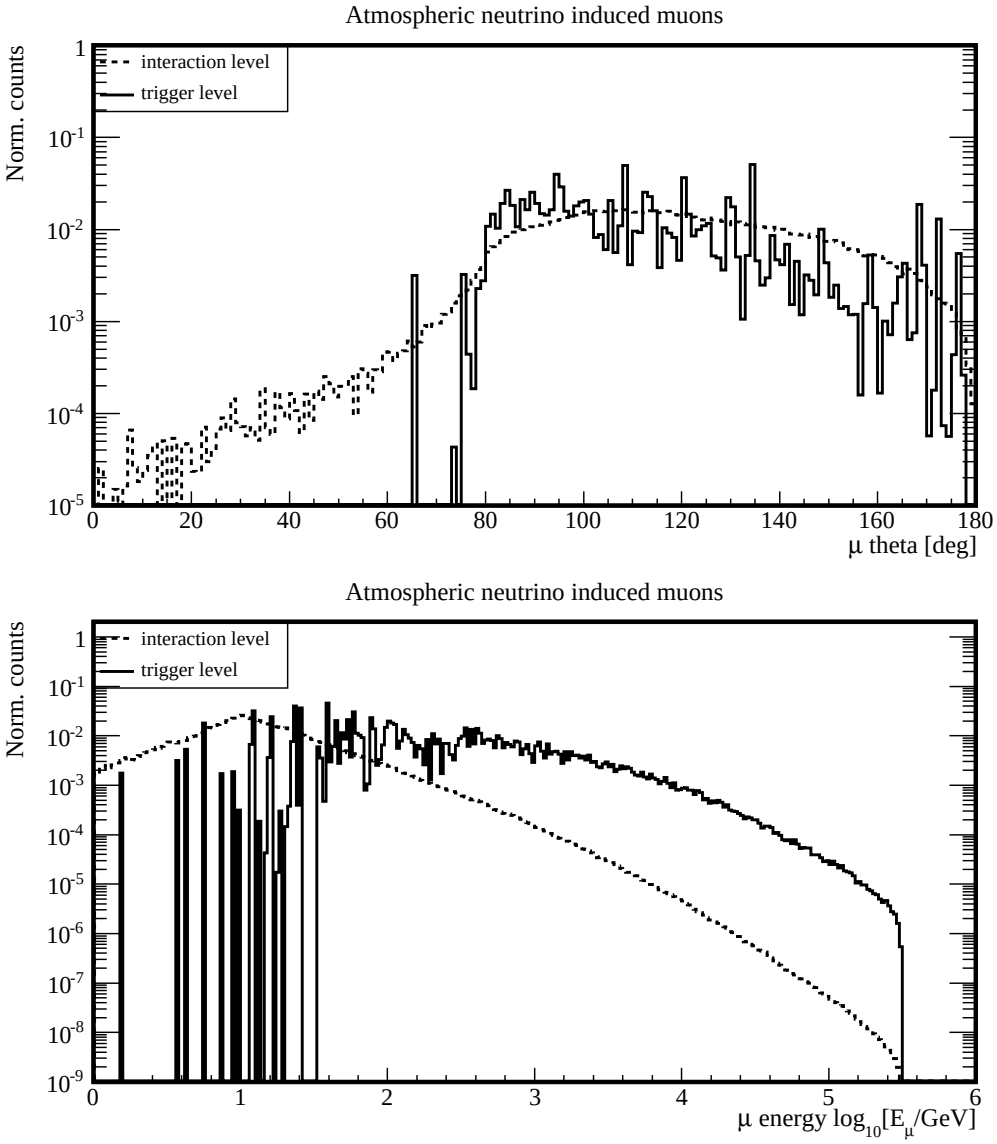


Figure 4.14: *The normalised angular (left) and energy (right) spectra of atmospheric muons produced by atmospheric neutrino interactions, at interaction level (dotted line) and at trigger level (solid line).*

- **Sarastro**:
Führt diese beyden Fremdlinge,
In unsern Prüfungstempel ein:
Bedecket ihre Häupter dann -
Sie müssen erst gereinigt seyn.

Die Zauberflöte - Atto I, Scena XIX
(libretto Emanuel Schikaneder)
Joannes Chrysostomus
Wolfgangus Theophilus Mozart
(1756 - 1791)

5 Data processing and analysis

This chapter consists of several sections related to the event processing and analysis. After giving some basic elements of event reconstruction, we pass to describe event filtering with the aim of removing badly reconstructed tracks, mostly due to the dominant atmospheric muon background. We have divided this event filtering in two main steps. The first, called *low level* filtering, required a condition on only one observed variable, the reconstructed theta angle, to select the event. The second step, the *high level* filtering, consisted in a more refined method which combined several variables to distinguish signal from background. This multivariate approach has been pursued through the classifier called Boosted Decision Trees (BDTs). At the end of the chapter we will delineate our final sample to be used in the hypothesis testing.

5.1 Event processing software

Low level processing of experimental and simulated data was performed by means of the code called `Sieglinde` [158], the AMANDA data processing software. The code dealt with calibration (see sec. 3.2.4 about calibration constants) to low level event selection (first and second level, sec. 5.5.3), including reconstruction and observable calculation (see secs 5.3, 5.4). Two versions of the code were used, the first so-called *classic* was used to process the data of $2001-2004$ up to the first filter level. Whereas, the one called *SLART*, which was available later with a slightly improved reconstruction algorithm, was used to process the second filter level of $2001-2004$ data and the complete low level processing of $2005-2006$ data. The file format of the files handled by *classic* was *f2000* plain text, whereas *SLART* handled either *f2000* or `ROOT` [159] format.

High level processing of experimental and simulated data was performed by means of the `TMVA 3.9.4` code [160], which is a `ROOT`-based software package to perform multivariate analysis (MVA). The training and application of MVA techniques were interfaced with data files, yielded in `ROOT` format by the second filtering step, through some codes developed by the author of this thesis. In this

work we used the Boosted Decision Trees (BDTs) as multivariate technique to perform high level signal-background separation (see sec. 5.5.5).

5.2 Event precleaning

We have pointed out in sec. 3.2.1, that some *background* pulses, induced for instance by the electronics itself, were recorded by the DAQ. The removal of those spurious hits is essential for a good reconstruction of the event, whose algorithm should combine only the relevant Cherenkov light-induced information. Hence, the hit cleaning procedure tried to identify non-particle information by features which discriminated them from detected Cherenkov light photons. This cleaning procedure was performed in several steps, where the hits were tagged as "bad" and thus disregarded in further filtering steps.

Furthermore, a certain kind of unidentified source of noise was outside the detector, which irregularly produced a large amount of non-photon pulses, that ended up to trigger the detector. These so-called non-particle or flare events[1], were identified by several indicators based on their electronic features [161], and further removed from the data sample (see sec. 5.5.4 for more explanations).

5.2.1 Hit cleaning

In this section we describe the different hit cleaning steps.

Bad OMs

The information collected during AMANDA detector monitoring enabled to classify each OM as "good" or "bad". This task was performed by different people, who developed for each year slightly different criteria [162].

In principle an OM was considered bad if it showed too low darknoise rate (hence considered as dead), or on the contrary too high darknoise rate (hence considered as noisy), if it showed instability due to too variable darknoise rate, or for calibration issues. Consequently, all the hits in these bad OMs were discarded. Besides, OMs outside the bulk of the detector (see sec. 3.1) were also removed. These bad OMs were the ones at the bottom of string 4, the top and bottom modules[2] on strings $11-13$, and all the modules on string 17, which got stuck during deployment (see fig. 3.2).

During the austral summers some campaigns were operated to try to identify and eliminate detector problems, like bad-connected modules. Thanks to this,

[1] These events, which created several hits in the detector, were potentially dangerous for ultra high energy neutrinos searches.

[2] Actually, the last 4 modules on string 11 were never deployed.

Year	N. of OMs
2001	513
2002	534
2003	540
2004	535
2005	523
2006	511

Table 5.1: *The total number of Optical Modules, for each year, used for reconstruction.*

some bad OMs were recovered; however it happened also that some previous good OMs were afterwards declared as bad. The total number of good OMs which were used for event reconstruction, for each year, is shown in tab. 5.1.

Time Over Threshold

The *Time Over Threshold* (TOT) is defined as the time between recorded leading (LE) and trailing edges (TE). Then a pulse shape with a short TOT could be an indication that the hits were not due to light in the PMT, but to some electronic artifacts. The TOT distribution for a particular optical module is shown in fig. 5.1. In the picture the main peak for photo-electron induced signals is clearly visible, while a second component with a decreasing exponential behaviour arises at short TOTs.

These hits with a short TOT were then removed applying a selection cut; for OMs with electrical read-out the minimum TOT value required spanned over $75 - 200$ ns (it could change with the year), with a maximum limit of 2000 ns, in case of (rare) hits with large TOT values due to missing TDC trailing edges. Whereas, the minimum TOT value required for OMs with optical readout was set to 5 ns. The TOT distribution shown in fig. 5.1 refers to an OM with electrical read-out, and the vertical dashed line marks its TOT cut value.

Time window

Radioactive decays of K, Th and U isotopes, which could happen in the OM glass sphere and in the PMT material, produced noise pulses; the latter could be also due to thermal noise, but with a lower rate, reduced by South Pole ice temperature. These hits had a random distribution in time, outlined as a noise plateau in fig. 5.2. This picture shows indeed the LE time distribution of uncalibrated (dashed-dotted line) and calibrated (solid line) hits, and we clearly see how the calibrated distribution is shifted towards shorter times, due to the removal of the time delay (see sec. 3.2.4).

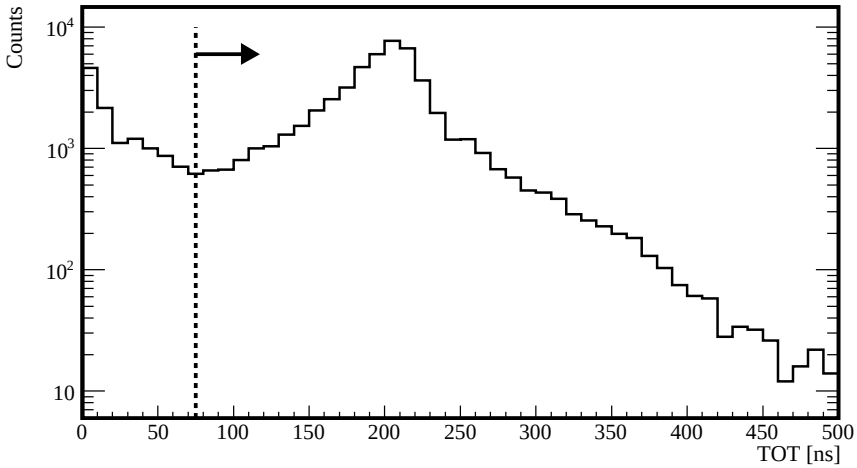


Figure 5.1: *TOT distribution of a chosen OM which refers to experimental data collected during 2004. The noise peak at small TOT values and the peak from real PMT pulses at 200 ns are easy to distinguish. The vertical dashed line and arrow mark the region of selected hits.*

The LE time distribution of uncalibrated hits shows a pronounced peak around the trigger time[1]. It mainly consisted of Cherenkov photons from relativistic particles which crossed the detector in less than 2 μs, triggering it. The AMANDA DAQ recorded all pulses in a time window of 32 μs around the trigger time, so most of the light from single muons should arrive in a short time within the recording interval. Most of the dark noise was then removed by selecting calibrated hits in a time interval [-2.5 μs, 4.5 μs], i.e. a 7 μs time window, around the trigger time. This time window is marked in the fig. 5.2 by the vertical dashed lines and arrows.

Another class of noise pulse due to ionisation of residual gas in the PMT tubes, so-called afterpulse, usually occurred a couple of microseconds after the primary photo-electron pulse; this noise pulse is visible in the picture as a smaller bump after the light peak. The time window cleaning removed also this kind of noise.

Amplitude and isolated hits

Hit amplitudes were settled by the pADC, which remained open for 9.8 μs; the peak amplitude during that time-window was assigned to all hits in that particular OM (see sec. 3.2.4). Hits occurring outside the pADC window, but within the 2.5

[1] The trigger (uncalibrated) time in 2001 – 2004 was set to around 22.5 – 22.8 μs after the opening of the TDC buffer, while in 2005 – 06 it was set to around 10.8 μs.

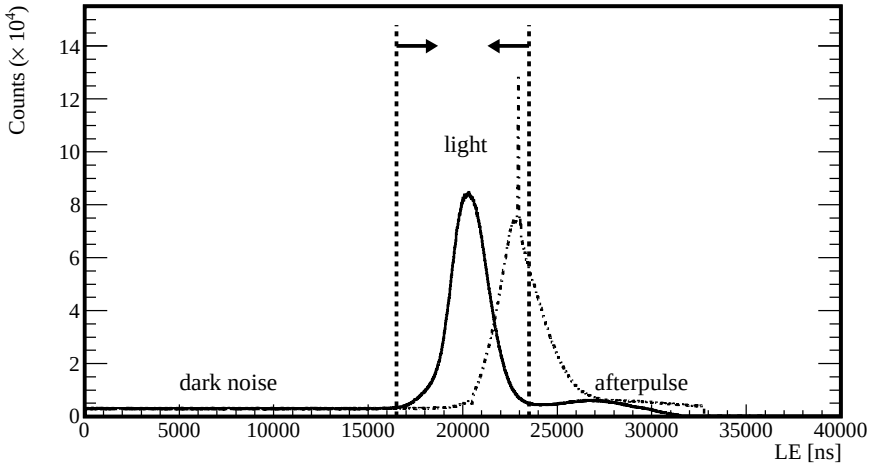


Figure 5.2: *LE time distribution, for all OMs, of uncalibrated (dashed-dotted line) and calibrated (solid line) hits from experimental data collected during 2004. The vertical dashed lines and arrows mark the region of selected hits.*

μs trigger window were assigned to a single photo-electron amplitude. All other hits were assigned to a null amplitude, thus no measured amplitude was associated to them, and therefore they were removed from the event. In conclusion, only hits with a calibrated amplitude in the range $0.1 - 1000$ photo-electrons, were further accepted.

Dark noise could be produced also by isolated hits in space and time. Hence, hits without an associated hit in time i.e. within 500 ns in any of the channels, or without an associated hit in space in any of the channels, i.e traced more than 100 m away from the fired module, were further removed.

Crosstalk

The TOT cleaning removed most of the crosstalk (see sec. 3.2.1) between neighbour-pairs of signal cables. The induced signal was indeed proportional to the derivative of the first signal, thus resulting in a bipolar pulse with a small amplitude and a shorter TOT. However, if the crosstalk was induced by a high amplitude pulse, the simple TOT cut could remove also a large fraction of hits from Cherenkov photons. Hence, to avoid any bias in the track reconstruction, a further improved cleaning step was necessary [163]. For instance, hits from photo-electrons exhibited a non-linear correlation between TOT and amplitude (pADC). This is illustrated in fig. 5.3 for a particular optical module, where some separate populations of hits from crosstalk are visible. Then, a map of OM-pairs that could held crosstalk was constructed, and a fit to the pADC-TOT correlation

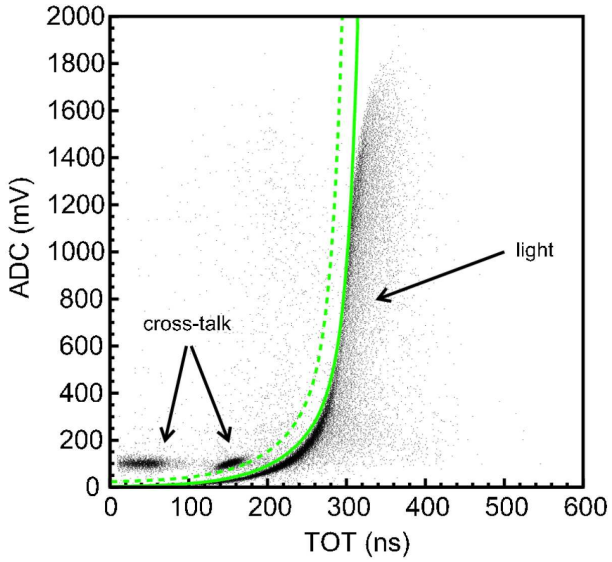


Figure 5.3: *Crosstalk pulses and photo-electron induced pulses form distinct populations in the pADC vs. TDC plot. The correlation between pADC and TOT of light induced photons was fitted (solid green line). Then the shape of this fit was shifted by -20 ns in TOT (dashed green line) and used as selection to remove the crosstalk hits on the left side of this line. Picture taken from [163].*

of real photo-electrons was performed (solid green line in the picture). The fit was then used in order to apply a two-dimensional cut selection; however, the selection cut was shifted by -20 ns in TOT (dashed green line), to avoid any loss of large fraction of good hits, due to fluctuations in the pADC-TOT distribution. Hence, only hits on the right side of this line were further used for high level reconstruction.

5.2.2 Retriggering

Once the hit cleaning was performed, we had to check if those events stripped from spurious hits continued to satisfy either multiplicity or string trigger condition. Hence, the remaining (uncalibrated) hit set was passed through a software retrigger logic. This step removed about 25−30% of the experimental data events, whereas less than 10% of the simulated atmospheric muon events was removed.

5.3 Event reconstruction

After the event precleaning procedure was applied, the basic photon information, like their amplitude and arrival time, were handed to the track recon-

struction procedure. We have explained in secs. 2.2, 2.5 that relativistic charged leptons produce a Cherenkov hallmark in ice. High energy muons produce long tracks, yielding sufficient directional information, which allowed a good reconstruction and angular resolution[1]. On the contrary, electron and tau events produce short tracks, which lack sufficient directional information for an accurate track reconstruction. However, if the generated electromagnetic and hadronic cascades were contained into the active detector, then these events could give a reasonable energy reconstruction; conversely this performance was unfeasible for muons, which left only part of their energy in the detector.

Hence, we stress here that, since our main goal is to search for a neutrino signal induced by neutralino annihilation in the Sun, a good angular resolution is a prerequisite condition, that's why our analysis is focused only on muon events.

Residual time

According to the AMANDA coordinate system (see sec. 4.2), a muon track was defined by a vertex $\mathbf{r}_0 = (x_0, y_0, z_0)$ at the time t_0, with the direction traced by the angles (θ_μ, ϕ_μ) (see fig. 5.4 on the left). The Cherenkov light, emitted by a muon with $\beta \simeq 1$ and travelling along a direction traced by a unit vector $\hat{\mathbf{p}}$, forms a well-defined angle θ_c with respect to $\hat{\mathbf{p}}$; then, a photon emitted at a time t and a point $\mathbf{r}_t$ on the track, could be detected at a time $t_{\rm obs}$ in an OM located at a point $\mathbf{r}_{\rm OM}$ (see fig. 5.4 on the right). Hence, according to the geometry of the picture, this photon was expected to arrive at the OM, in a scattering-free medium, at time:

$$t_{\rm geo} = t + \frac{\hat{\mathbf{p}} \cdot (\mathbf{r}_{\rm OM} - \mathbf{r}_t) + \rho \tan\theta_c}{c_{\rm vac}} \tag{5.1}$$

with ρ the minimum distance track-OM, and $c_{\rm vac}$ the vacuum speed of light[2].

A fundamental variable, used in a reconstruction algorithm based on arrival time, was the *residual time*, $t_{\rm res}$, which is the difference between the observed time of a Cherenkov photon and its expected arrival time, hence:

$$t_{\rm res} \equiv t_{\rm obs} - t_{\rm geo} \tag{5.2}$$

In the ideal case, the $t_{\rm res}$ distribution would be a delta function; however, in the realistic case, i.e. in the experiment, this distribution was broadened and distorted by several effects, which are illustrated in fig. 5.5. The PMT time jitter, which limited the timing resolution σ_j, and the dark noise could generate negative $t_{\rm res}$ values, which would mimic non-physical causality violations. Further, secondary

[1] The pointing accuracy of the reconstruction is limited due to the neutrino-muon scattering angle (see sec. 2.3).

[2] The eq. 5.1 neglects the effect that Cherenkov light propagates with group velocity, however this approximation was valid for AMANDA (see [164] and references therein).

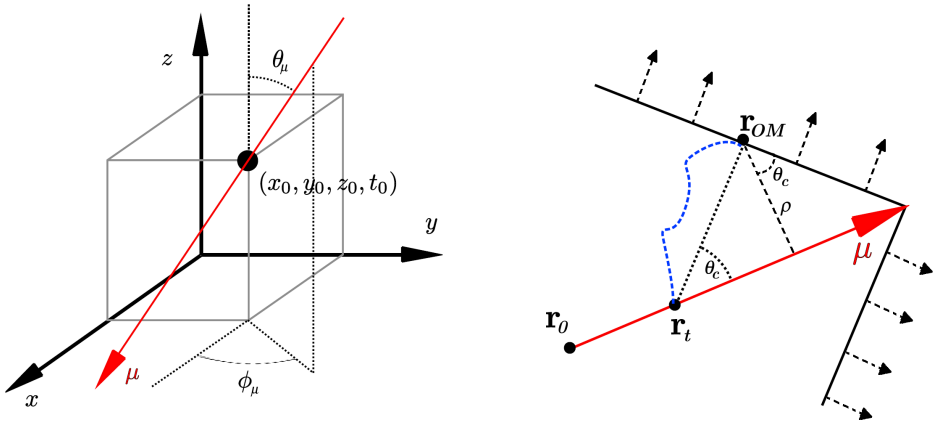


Figure 5.4: *On the left, muon track as defined in the AMANDA coordinate system. It is traced by the particle position at a certain time* $(\mathbf{r}_0, t_o)$ *and by the spherical coordinates* (θ_μ, ϕ_μ)*. On the right, a sketch of track-OM geometry. The muon travels from* $\mathbf{r}_0$ *producing a Cherenkov cone with angle an* θ_c*. The photon emitted at a point* $\mathbf{r}_t$ *could be later detected in an OM located at* $\mathbf{r}_{\mathrm{OM}}$*.*

radiative energy losses along the muon track produced late photons, which arrived after the ideal Cherenkov cone.

Hits with a short residual time, typically in the time window [-25 ns,+75 ns], were called *direct hits*.

The dominant effect on photon arrival times was the scattering in ice, whose effect depended strongly on ρ, the minimum distance OM-track. The distribution of t_{res} depended also on the orientation of the OM with respect to the track, since the encapsulated PMT had a non-uniform angular response. Indeed, OMs facing away from the track could only see light that scattered back towards the PMT face, and on average this effect shifted t_{res} to later times.

5.3.1 First guess method

At the onset of event processing and analysis, the experimental data were dominated by the down-going atmospheric muon flux. Since it was unreasonable to fully reconstruct several billion of events induced by cosmic rays, because of very high CPU-time requirement, a large part of the muon background was then identified by a fast and coarse event reconstruction. This reconstruction method is generally called "first-guess" (FG); events which were not clearly identified as background with this method, were further processed through CPU-time intensive likelihood and topological parameter reconstruction (see sec. 5.3.2).

In this work the fast first-guess reconstruction methods called *Direct Walk* [165], *Direct Wimp* [166] and the more elaborate *JAMS* [167], were used for

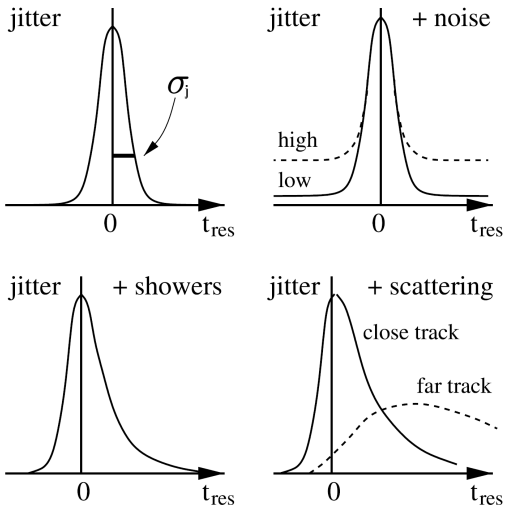


Figure 5.5: *Schematic overview of the t_{res} distribution broadened and distorted by different effects. On top left: by the time jitter. On top right: by the effect of jitter and random noise. On bottom left: by the effect of jitter and secondary cascades along the muon track. On bottom right: by the effect of jitter and scattering. Picture adapted from [164].*

background suppression, whose brief descriptions are given below.

Direct Walk, Direct Wimp

Direct Walk (DW), was a first guess track search algorithm which consisted of four steps. The first step was the search for *track elements* (TRELs): i.e. the search for two coincident direct hits, with their time difference nearly equal to the flight time of the muon. Hence, the algorithm selected pairs of hits which fulfilled the following requirements:

$$\mathcal{D}_{\mathrm{OM}}/c - 30\,\mathrm{ns} < \Delta t < \mathcal{D}_{\mathrm{OM}}/c + 30\,\mathrm{ns} \qquad \&\& \qquad \mathcal{D}_{\mathrm{OM}} > 50\,\mathrm{m} \tag{5.3}$$

where Δt was the time difference of the 2 hits, $\mathcal{D}_{\mathrm{OM}}$ the distance between the OMs, and c the speed of light. The track parameters $(x, y, z, t, \theta, \phi)$ were calculated from the OM positions of the two hits of the TREL. Anyway, if the number of found TRELs exceeded 200, the procedure started again, but with a time limit reduced by 5ns: i.e. 25ns instead of 30ns, and so on. This sharpened the selection in case of too many TRELs, and also reduced the CPU-time requirement for events with a higher hit multiplicity.

The next step was the selection of *track candidates* (CANDs) from the found TRELs. The parameters of these latter had to describe the typical pattern of

a muon track: sufficient number of hits along the track with time residuals according to expectation, and a minimum track length. These quality criteria were calculated from hits associated to the track, in a restricted region of the $\{t_{\rm res}, \rho\}$ plane:

$$-30\,{\rm ns} < t_{\rm res} < 300\,{\rm ns} \qquad \&\& \qquad \rho < 25 \cdot \sqrt[4]{t_{\rm res} + 30\,{\rm ns}} \tag{5.4}$$

where $t_{\rm res}$ is the time residual and ρ the distance OM-TREL. Two cuts on some quality parameters, i.e. number of hits $N_{\rm Hit} > 10$ and spread of the hits along the track $\sigma_L > 20$, were applied to guarantee a minimum of track quality. Then they were combined to a single quality parameter $\mathcal{Q}_{\rm CAND}$, defined as follows:

$$\mathcal{Q}_{\rm CAND} = \min(N_{\rm Hit}, 0.3 \cdot \sigma_L + 7) \tag{5.5}$$

The third step was the *track selection*, performed by picking only those CANDs which fulfilled the condition:

$$\mathcal{Q}_{\rm CAND} \geq 0.7\mathcal{Q}_{\rm max} \tag{5.6}$$

where $\mathcal{Q}_{\rm max}$ was the maximum of all $\mathcal{Q}_{\rm CAND}$ in the event. In the possible case that there were more than one high quality candidates, then a search for a cluster in space for such candidates was performed. For each CAND, the number of CANDs within a cone of 15° was determined, and the cone showing the highest multiplicity was selected.

Finally, the last step was the *track direction selection*: the average of the parameters of the CANDs in the selected cone, was chosen as the first guess hypothesis.

A slightly modified version of the DW algorithm, called Direct Wimp (DWimp), was developed for vertical and low energy tracks, since DW could fail to properly reconstruct those tracks. Hence, in the DWimp algorithm, the requirement for coincident hits was lowered to 35 m for hits on different strings, and to 10 m for hits belonging to the same string. Furthermore, the $N_{\rm Hit}$ requirement was lowered to 8 hits. Hence, DWimp reconstructed less energetic tracks at the cost of more limited angular resolution.

JAMS

The more elaborate JAMS (Just Another Muon Search) was a pattern recognition based first guess method. The basic idea of the JAMS algorithm was that: hits space coordinates (x, y, z) were rotated in a frame $X'Y'Z'$ aligned with the starting point track direction, $\hat{\mathbf{r}}_{\rm ansatz}$, so that (x', y') were in a plane perpendicular to $\hat{\mathbf{r}}_{\rm ansatz}$ and z' along $\hat{\mathbf{r}}_{\rm ansatz}$, then these hits produced a Gaussian cluster in (x', y') plane. Moreover, hits will cluster in time along the muon direction (along z' axis).

Clusters were defined by counting, for each hit, the number of its contiguous hits, where a contiguous hit was defined as the one fulfilling the requirement

$$r = \sqrt{(\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2} < r_{\max} \tag{5.7}$$

where $(\Delta x', \Delta y', \Delta z')$ were the coordinate difference between two hits, and $r_{\max}$ was a user defined value.

The minimum number of hits required by JAMS was 7, in order to keep a cluster for a track hypothesis. Then, a first guess track hypothesis (x, y, z, θ, ϕ) was found, by exploiting the angles from the hypothetical track direction, and the average of the hits on the cluster.

The next step in the JAMS algorithm was to refine the first guess track through a simple likelihood reconstruction. Hence, a quality parameter for each cluster was settled by training a neural network, fed with event topological observables, in order to distinguish high and low quality reconstructions. Then the clusters were sorted by their quality parameter, and the best three track candidates were stored for further analysis steps.

Compared to DW, JAMS performance was slower, but as a way of compensation it was more accurate. Conversely, compared to the full likelihood reconstruction, it was faster but less accurate; anyway since JAMS inspected several directions, it was less subject to wrongly reconstruct coincident muon events than the likelihood reconstruction.

5.3.2 Maximum likelihood method

The likelihood $\mathcal{L}$ for a certain track hypothesis $\mathbf{a} = (\mathbf{r}_0, t_0, \theta, \phi)$, is defined as the product of the probability density functions (p.d.f.) to find an experimental ensemble $\{\mathbf{x}\}$, given the track hypothesis $\mathbf{a}$ as true; for independent components x_i of $\{\mathbf{x}\}$ $\mathcal{L}$ reduces to:

$$\mathcal{L} = \prod_i p(x_i|\mathbf{a}) \tag{5.8}$$

The track hypothesis $\mathbf{a}$ can be varied tuning its parameter space, until the track with the maximum likelihood is found. Hence, this track is selected as the "best-guess" for the true particle direction.

In practice, it is more convenient to minimise the negative logarithm of the likelihood (log-likelihood, or LLH) instead of maximising the likelihood:

$$-\log \mathcal{L} = -\sum_i \log p(x_i|\mathbf{a}) \tag{5.9}$$

The simplest time likelihood function was based on a likelihood constructed from the p.d.f. for arrival times of single photons i at the locations of the hit

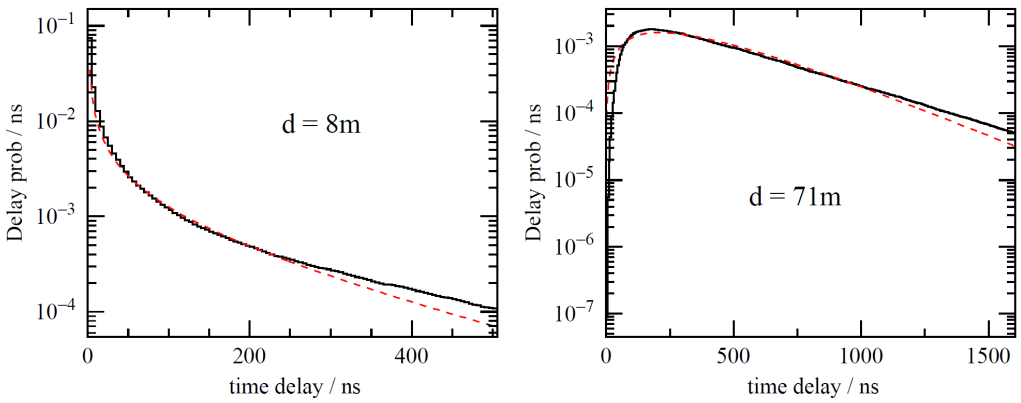


Figure 5.6: *Comparison of the parametrised Pandel function (dashed curves) with the detailed simulation (black histograms) at two distances d from the muon track. On the right, photons which hit distant modules suffered more scattering, with much larger t_{res}. Conversely, on the left, most photons had a small t_{res}, due to the nearness to the OMs. Picture taken from [164].*

OMs[1]

$$\mathcal{L}_{\mathrm{time}} = \prod_{i=1}^{N_{\mathrm{Hit}}} p(t_{\mathrm{res},i}|\mathbf{a}) \tag{5.10}$$

A parametrisation of the arrival time distributions as a function of the closest distance OM-track, ρ, was achieved through the so-called *Pandel function*, which includes scattering and absorption effects [168]

$$p(t_{\mathrm{res}}, \rho) \equiv \frac{1}{N(\rho)} \cdot \frac{\tau^{-(\rho/\lambda)} \cdot t_{\mathrm{res}}^{(\rho/\lambda\,-\,1)}}{\Gamma(\rho/\lambda)} \cdot e^{-\left[t_{\mathrm{res}}\left(\frac{1}{\tau} + \frac{c_{\mathrm{ice}}}{\lambda_a}\right) + \frac{\rho}{\lambda_a}\right]} \tag{5.11}$$

with

$$N(\rho) = e^{-\rho/\lambda_a} \cdot \left(1 + \frac{\tau \cdot c_{\mathrm{ice}}}{\lambda_a}\right)^{-\rho/\lambda}$$

where $c_{\mathrm{ice}} = c_{\mathrm{vac}}/n$ is the speed of light in ice, $\Gamma(\rho/\lambda)$ the Gamma function and $N(\rho)$ a normalisation factor. The parameters used in the above formula, i.e. $\lambda = 33.3$ m, $\lambda_a = 96$ m and $\tau = 557$ ns, were fixed from a fit (dashed curves in fig. 5.6) to detailed Monte Carlo simulations of photon propagation (black histograms in fig. 5.6), using an averaged ice model.

The Pandel parametrisation did not include the electronic jitter, and could diverge, when $\rho < \lambda$, at small t_{res}. The extended *Patched Pandel* distribution

[1] Note that one OM may contribute to the product with several hits.

was then implemented to try to solve these issues, by convolving the Pandel distribution with a Gaussian one, $\{\mathcal{G}\}$, centred at $t_{\text{res}} = 0$ and with the width σ_j from the time jitter. A smooth function was favoured rather then the analytical convolution, because of its slowness, as a transition between the Gaussian and the Pandel distribution, i.e. a third order polynomial $\mathcal{P}(t_{\text{res}}, \rho)$ joining the two distributions [164]. Hence, this patched Pandel distribution was written as follows

$$\hat{p}(t_{\text{res}}, \rho) = \begin{cases} \mathcal{G}(t_{\text{res}}, \rho) & t_{\text{res}} < 0 \\ \mathcal{P}(t_{\text{res}}, \rho) & 0 \leq t_{\text{res}} \leq \sqrt{2\pi}\sigma_j \\ p(t_{\text{res}}, \rho) & t_{\text{res}} > \sqrt{2\pi}\sigma_j \end{cases} \tag{5.12}$$

Noise was added through the introduction of a constant probability in the patched Pandel distribution. The best track hypothesis $\mathbf{a}$, given a set of hits with time residuals $\mathbf{t}_{\text{res}}$ and closest distances $\boldsymbol{\rho}$ with respect to $\mathbf{a}$, was then found by minimising the following expression

$$-\mathcal{L}(\mathbf{t}_{\text{res}}, \boldsymbol{\rho}|\mathbf{a}) = -\sum_{i=1}^{N_{\text{Hit}}} \log \hat{p}(t_{\text{res},i}, \rho_i|\mathbf{a}) \tag{5.13}$$

using the *Simplex* minimisation algorithm [169].

Iterative reconstruction

The likelihood reconstruction method could, however, suffer from tracing local minima instead of the global minimum. An example of how the likelihood function changes with one track parameter, while the other parameters remain fixed, is shown in fig. 5.7. In the same picture a local minimum found by the likelihood minimisation is shown, indicated by a fitted parabola. Symmetries in the detector, especially in theta angle, multiple scattered photons arriving at unforeseen times and uncorrelated random noise hits, could induce local minima; in some other cases the minimiser algorithm could stop at extreme theta angles. A good first-guess track, used as seed for the likelihood reconstruction, could reduce these problems.

The iterative reconstruction was a technique to find the global minimum; it accomplished several consecutive reconstructions with (θ, ϕ) randomised in a cone around the track with the lowest $\{-\log \mathcal{L}\}$ function. This procedure, usually, allows to find the global minimum boosting the iterations, but it had the drawback to increase CPU-time requirement.

We define as the angular resolution the median space angle between true and reconstructed track, which depends on the event ensemble. At the final stage of the event selection, well reconstructed events are kept, therefore their resolution is expected to be better than at the trigger level. So, the angular resolution of

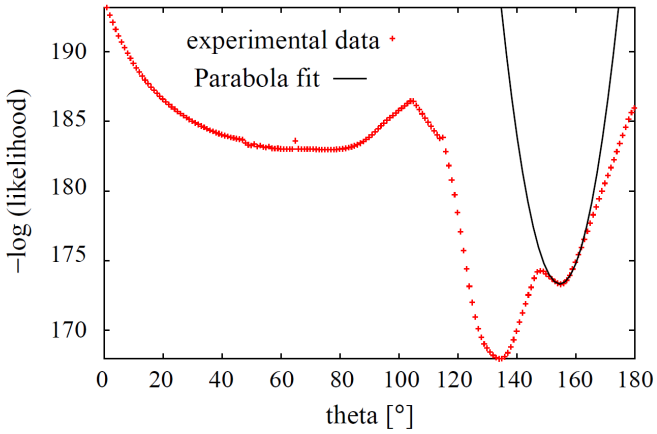


Figure 5.7: *The likelihood space for one experimental event (one-dimensional projection). Each point represents a fit, for which the θ angle was fixed and the other track parameters free to vary, in order to find the best minimum. A local minimum which was found by a gradient likelihood minimisation is indicated by a fitted parabola. Iterative reconstruction methods try to avoid this. Picture taken from [164].*

the likelihood reconstruction at the final cut level was of the order of 3° for high energy neutralino models, degrading with decreasing energy (see tab. 6.1).

Bayesian weighted reconstruction

The maximum likelihood method evaluated all track hypotheses as equally probable, although most of the tracks recorded in data were down-going atmospheric muons. A subjective hypothesis, considering most events as down-going, can be included in the reconstruction method, exploiting in that case a Bayesian approach. We assume that the a priori probability density function of observing a track $\{\mathbf{a}\}$ is given by $h(\mathbf{a})$; Bayes theorem can be used then to calculate the conditional probability density function, $\mathcal{H}(\mathbf{a}|\mathbf{x})$, of observing $\{\mathbf{a}\}$, given an experimental ensemble $\{\mathbf{x}\}$

$$\mathcal{H}(\mathbf{a}|\mathbf{x}) = \frac{\hat{p}(\mathbf{x}|\mathbf{a})h(\mathbf{a})}{\int \hat{p}(\mathbf{x}|\mathbf{a})h(\mathbf{a})d\mathbf{x}} \tag{5.14}$$

The probability density function, $h(\mathbf{a})$, can be evaluated from the theta angle distribution of simulated atmospheric muons, theta being one of the most relevant observables to reject atmospheric muon background. The denominator in equ. 5.14 is a constant, hence it can be neglected in the $\{-\log \mathcal{L}_B\}$ minimisation,

which can be written as follows

$$-\log \mathcal{L}_B(\mathbf{a}|\mathbf{t}_{\rm res}, \boldsymbol{\rho}) = -\left(\sum_{i=1}^{N_{\rm Hit}} \log \hat{p}(t_{{\rm res},i}, \rho_i|\mathbf{a})\right) - \log h(\mathbf{a}) \tag{5.15}$$

The probability density function $h(\mathbf{a})$ can be viewed then as a weight to the standard likelihood, with the result that tracks with poor likelihood will be considered as down-going. A comparison of the logarithm of the likelihood ratios, or $\log \mathcal{L} - \log \mathcal{L}_B$, between the standard fit and the Bayesian fit, can be then used to separate wrong reconstructed muons induced by cosmic rays from true up-going neutrino induced tracks.

5.4 Event observables

In this section we just serve an hors d'oeuvre of the different event observables used in the selection processes, which we will discuss in sec. 5.5, in an attempt to remove the atmospheric muon background.

These observables were divided in three classes: *reconstruction*, *topology* and *hit-reconstruction* (see below); the complete list and their distributions are shown in appendix A.

Reconstruction

The reconstruction algorithms provided some observables, like the muon theta angle (the most natural), or the z coordinate of the reconstructed vertex.

In the previous section we proposed to compare the standard and Bayesian log-likelihood ratio; in the same way we can take now the difference of their reduced log-likelihood[1], $\Delta rLLH$, whose value indicates the difference between the downgoing and all-sky hypotheses; a large value suggests an improbable downgoing hypothesis.

The JAMS first-guess method settled a quality parameter, $\mathcal{Q}_{JAMS}$ (see sec. 5.3.1), on the reconstructed tracks (higher for better reconstructed tracks), which can be used in the event selection.

DWimp, the other first-guess algorithm, provided an internal parameter, $\sigma_\psi^{\rm DWimp}$, which is a measure of the angular resolution of the solution; small values indicate accurate solutions.

[1] The reduced log-likelihood is the ratio of the log-likelihood over the number of degrees of freedom of the fit (five parameters in our case).

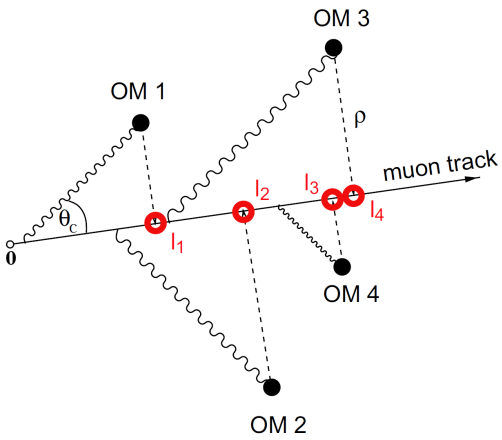


Figure 5.8: *Sketch of the projection of the hits, l_i, along the reconstructed track. Picture adapted from [100].*

Topology

From the hit information some topological observables could be extracted. A simple example is the total number of hits in the event (or the number of hit OMs, or the number of hit strings, not used in this work), which can give an indication of the amount of light deposited in the detector.

The distance to vertical axis, or to the axes origin, of the centre of gravity (COG) of the hits is suitable as well; they can suggest whether the event occurs well inside the detector, or rather outside; this latter case could lead to a low quality reconstruction. Other information arise from the spread of the COG of the hits along the vertical.

Hit-reconstruction

The third class is composed of observables which reflect relations between hits and corresponding reconstructed track. These hits were classified, according to their time residual, as:

- *early*, if $t_{\text{res}} \in$ [-550 ns, -25 ns]
- *direct*, if $t_{\text{res}} \in$ [-25 ns, +75 ns]
- *late*, if $t_{\text{res}} \in$ [+75 ns, +750 ns]

Further, these observables were calculated considering a cylinder around the reconstructed track with radius $\mathsf{r} = 50$ m. OMs further away than $\mathsf{r}_{\text{max}} = 50$ m were disregarded in all the observable calculations.

An indication of the goodness of the reconstruction was given by the projected length of direct hits along the track (see fig. 5.8). Indeed, a large value could suggest that the reconstruction described correctly the hit pattern over large distances.

Cumulative hit distribution along a track was described by a parameter called *Smoothness* [170]. This parameter can be calculated either for the projected length of direct hits, or for their LE distribution. High absolute value of the Smoothness (i.e. close to one) indicated that either hits were missing at OMs, where they would be expected if the track hypothesis was correct, or that hits were found in OMs where none would be expected, thus suggesting a possible wrong reconstruction. The sign of the Smoothness parameter distinguished between the two cases.

The spread of the radial distance of hits around the reconstructed track, can be useful as well. Low energy, fairly reconstructed events, had small average distances; conversely, in wrong reconstructions the hits are not centred around the track, with a large average radial distance.

Another observable that can also be considered is the number of strings with OMs hit (active strings); further, we can consider also active strings with only direct hits, or the ones with only late hits.

OMs without hits near a hypothetical track, or OMs with hits far from the track were unlikely. Upon this, the expected number of hits in the detector can be calculated, and a comparison between the expected and the observed hit distribution can be performed. Further, the ratio of the expected and observed average radial distances, or the separation of the expected and observed hit clouds, could distinguish atmospheric background from up-going neutrino events.

5.5 Event selection

We started this analysis with the aim to search for neutralino induced neutrinos from the Sun, which represent our hypothetical signal. And we have already pointed out over this thesis, that our experiment was background dominated; the atmospheric neutrino background, in particular, could strongly resemble signal events (see e.g. some observable distributions in appx. A.2). Hence, the background rejection should be performed in such a way to minimise as much as possible the loss of signal.

In this section we discuss about the techniques to reduce the atmospheric background, which were implemented into different steps, or filter levels. The first two levels, which can be considered as a part of a "low-level" filter, since they dismiss downgoing events simply putting a requirement on the reconstructed theta angle, are described in secs. 5.5.3.1 and 5.5.3.2.

The final filter level was more sophisticated; it entailed indeed the combination of several event quality observables in a multi-dimensional space. Hence, the application of this method was regarded as a "high-level" filter, and we have used the Boosted Decision Trees (BDTs) as event classifier to perform our multivariate analysis (see sec. 5.5.5).

We remark here, as it will be explained in sec. 5.5.5.2, that all the calculations relative to the effective volumes and the efficiencies for the various neutralino models, refer to 50% of the whole signal Monte Carlo; this particular sub-sample, called *analysis MC sample*, will be used in the last stage of our analysis.

We also remark here, and it will be stressed later, that our analysis is "blind" with respect to the position of the Sun, by not explicitly using the event time or correlation with the Sun position in any of the steps of the event selection. At the end, after optimising the analysis steps, the Sun position is revealed and the final result can be obtained.

5.5.1 Effective volume and efficiency

The effective volume is a suitable tool to inspect the performance of a detector. It can be interpreted as the proper volume of a detector, which has an ideal efficiency ($\varepsilon \equiv 1$) to select each muon produced in neutrino-nucleon interactions, from trigger to final filter level. Naturally, the neutrino and thus muon energy (connected then to the muon range) plays a role in the effective volume calculation. We can write the effective volume as follows

$$V_{\text{eff}} = \frac{N_{\text{obs}}}{N_{\text{gen}}} V_{\text{gen}} \tag{5.16}$$

where N_{obs} is the number of observed events after a selection, out of N_{gen} generated, and V_{gen} the volume containing these generated events. Hence, collecting eqs. 4.5 and 4.6, the above expression can be rewritten as follows

$$V_{\text{eff}} = \frac{\sum_i^N \delta_i V_i(E_\nu, \theta_\nu) w_i(E_\nu)}{\sum_i^N w_i(E_\nu)} \tag{5.17}$$

where $N \equiv N_{\text{gen}}$ and the Kronecker delta, δ_i, could refer to the different event selection (trigger, first level,...).

In the same way we can define an effective area of the detector, which could be compared with the one of a "flat" neutrino detector

$$A_{\text{eff}} = \frac{\sum_i^N \delta_i A_i(\theta_\nu) w_i(E_\nu)}{\sum_i^N w_i(E_\nu)} \tag{5.18}$$

In the progress of this chapter we will calculate the effective volume at different filter levels, for each neutralino model.

$V_{eff(L0)}$	[m^3]	50	100	250	500	1000	3000	5000
ALL	hard	$3.99{\cdot}10^5$	$1.79{\cdot}10^6$	$8.26{\cdot}10^6$	$1.46{\cdot}10^7$	$1.94{\cdot}10^7$	$1.97{\cdot}10^7$	$1.92{\cdot}10^7$
	soft	$1.30{\cdot}10^5$	$3.58{\cdot}10^5$	$1.42{\cdot}10^6$	$2.96{\cdot}10^6$	$4.81{\cdot}10^6$	$7.12{\cdot}10^6$	$7.71{\cdot}10^6$
STD	hard	$1.20{\cdot}10^5$	$1.16{\cdot}10^6$	$6.88{\cdot}10^6$	$1.26{\cdot}10^7$	$1.70{\cdot}10^7$	$1.73{\cdot}10^7$	$1.68{\cdot}10^7$
	soft	$0.06{\cdot}10^5$	$1.14{\cdot}10^5$	$0.93{\cdot}10^6$	$2.23{\cdot}10^6$	$3.86{\cdot}10^6$	$5.92{\cdot}10^6$	$6.46{\cdot}10^6$
STR	hard	$2.79{\cdot}10^5$	$0.63{\cdot}10^6$	$1.38{\cdot}10^6$	$0.20{\cdot}10^7$	$0.24{\cdot}10^7$	$0.24{\cdot}10^7$	$0.24{\cdot}10^7$
	soft	$1.24{\cdot}10^5$	$2.44{\cdot}10^5$	$0.49{\cdot}10^6$	$0.73{\cdot}10^6$	$0.95{\cdot}10^6$	$1.20{\cdot}10^6$	$1.25{\cdot}10^6$

Table 5.2: *Effective volumes at L0 for the different neutralino models (mass and channel), split up for the different trigger selections: STR (exclusive string trigger), STD (inclusive standard multiplicity trigger), ALL (logical sum of the two previous selection).*

The exclusive selection efficiency of a generic filter level with respect to the previous one, concerning the experimental data and the atmospheric background[1] Monte Carlo, is defined as the ratio between the number of events after and before the selection

$$\varepsilon_{L_i} \equiv \frac{n_{L_i}}{n_{L_{i-1}}} \tag{5.19}$$

whereas for the signal Monte Carlo we consider more relevant the ratio of the effective volumes:

$$\varepsilon_{L_i} \equiv \frac{V_{\text{eff}(L_i)}}{V_{\text{eff}(L_{i-1})}} \tag{5.20}$$

5.5.2 Trigger level

We remind here what we explained in sec. 3.2.3, i.e. the condition to record an event revolved around the fulfilment of (at least) one of the triggers, the standard (or multiplicity) and the string correlation trigger[2]. Whilst most of the events satisfied the multiplicity trigger, low energy ones (below 250 GeV) were mostly and exclusively triggered by the string trigger. Hence, we have to point out that the string trigger did good to low mass neutralino models, although what we would expect was a benefit for vertical tracks rather than horizontal tracks, like muons induced by neutrinos from the Sun.

Table 5.2 summarises the effective volumes at trigger level (namely L0), of the simulated signal; for each neutralino model three different classes are outlined: the exclusive string trigger (STR), the standard trigger (STD), and the logical sum of the these two selections (ALL). Looking into this table, we can truly

[1] For atmospheric neutrinos, the proper weight should be taken into account (see sec. 4.2.4).

[2] AMANDA analyses searching for high energy neutrinos ($>$ 1 TeV), discarded string-exclusive triggered events, since they did not give any particular contribution to the final sensitivity.

N_{trig}		2001	2002	2003	2004	2005	2006	$2001-2006$
Exp. data	$(\times 10^9)$	1.456	0.921	1.146	1.219	1.291	1.213	$7.246 \cdot 10^9$
Atm. μ	$(\times 10^9)$	1.372	0.949	1.174	1.248	1.312	1.223	$7.278 \cdot 10^9$
Atm. ν	$(\times 10^3)$	5.815	3.877	4.247	5.065	5.330	4.942	$29.28 \cdot 10^3$

Table 5.3: *In the first part of the table, the annual contribution from all triggers: on the first row, the number of events for the analysis sample experimental data is shown, while in the remaining two, the number of events from the atmospheric muons and neutrino background is shown, rescaled to the live-time. The second part of the table shows the total contribution of the* $2001-2006$ *data-set.*

confirm what is said above, i.e. low energy models like 50 soft, 50 hard, and 100 soft, owe their effective volume to the string trigger, in a fraction of $\sim 95-70$ %; for instance, without the string trigger the 50 GeV soft model would experience a loss of a factor of 20 in efficiency.

In fig. 5.9 the relevant numbers of table 5.2 are plotted (hard channels on the top, while soft channels on the bottom); the different trigger classes are respectively drawn with a solid line (ALL), dashed line (STD) and dotted line (STR, otherwise noted as *string trigger only*).

In fig. 5.10 the effective volumes for each neutralino model (black colour for hard channels, and grey for soft channels) are shown again, but this time different lines represent the different six years of simulated data. We clearly notice, observing the *string trigger only* stream plot (on the bottom of the picture), the effect of string trigger downscaling from 2002 onwards, which reduces by a factor of two the effective volume. For the standard stream (on the top of the picture), we expect no particular variation of the effective volume among the different years.

The number of events at trigger level of the experimental data (the ones belonging to the *analysis sample*, see sec. 4.1.4), are shown in tab. 5.3, along with the number of the simulated atmospheric muons and neutrinos, rescaled to the detector live-time. In the first part of the table, the number of triggered events per year is shown, while in the second part, the total contribution of $2001-2006$ data-set is shown. The amount of events selected for each year, reflects two known cases: the different live-time and the string trigger downscaling from 2002 onwards. Going again into the table, we observe that the experimental data contain about 7 billion of events, which were practically all down-going atmospheric muons. This explanation is what we inferred from the background Monte Carlo simulation, which exhibited a good agreement in rate with respect to experimental data, taking into account the theoretical uncertainties on the primary flux and the experimental uncertainties on the absolute sensitivity of the OMs (see sec. 6.2.1).

To design an accurate Monte Carlo is a challenging task, since some funda-

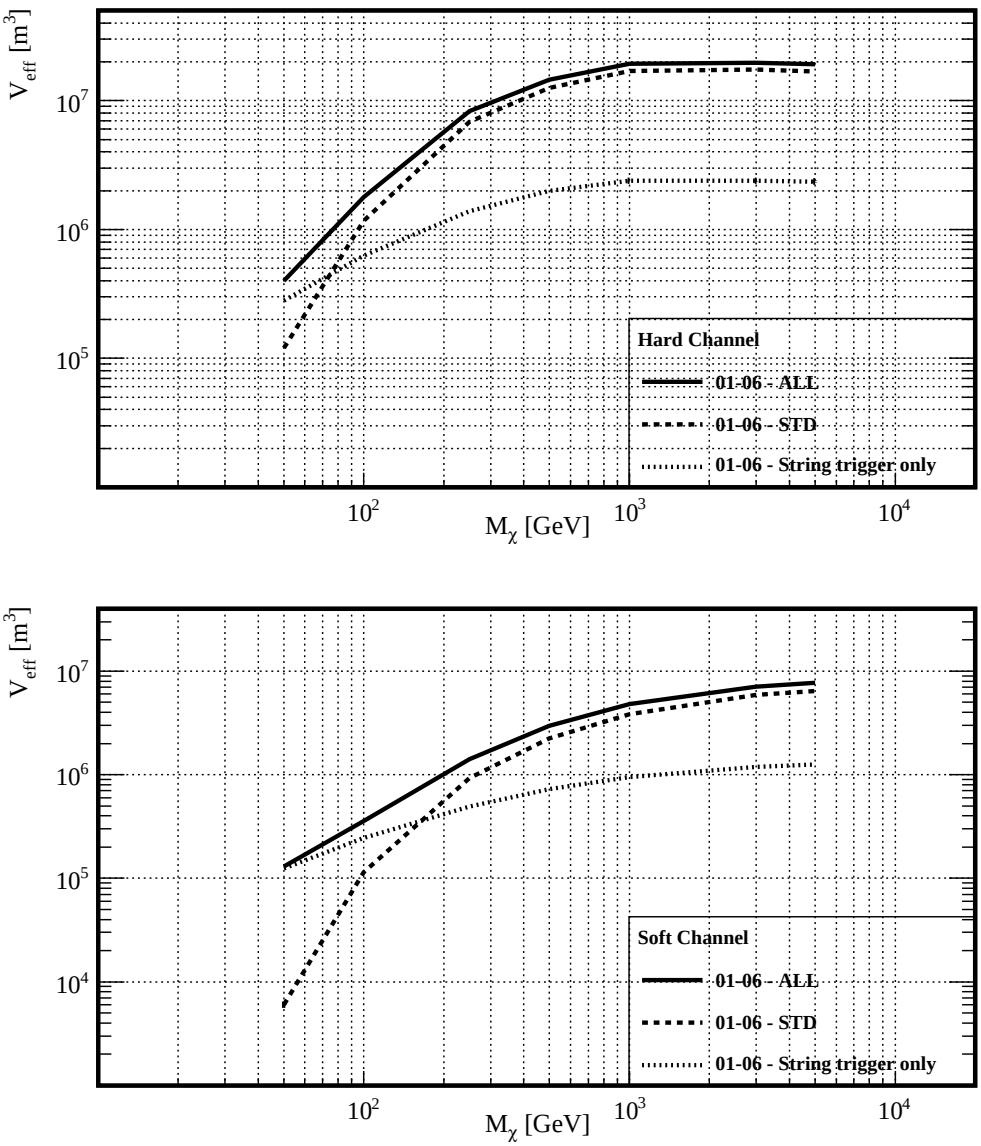


Figure 5.9: *Effective volume at trigger level (or L0) as a function of the neutralino mass for hard (top) and soft (bottom) annihilation channel. The dotted line refers to events triggered only by the string trigger, while the dashed one to events triggered by the standard trigger. The solid line is the logical sum of the two trigger selections. Statistical errors are not visible, covered by the size of the lines*

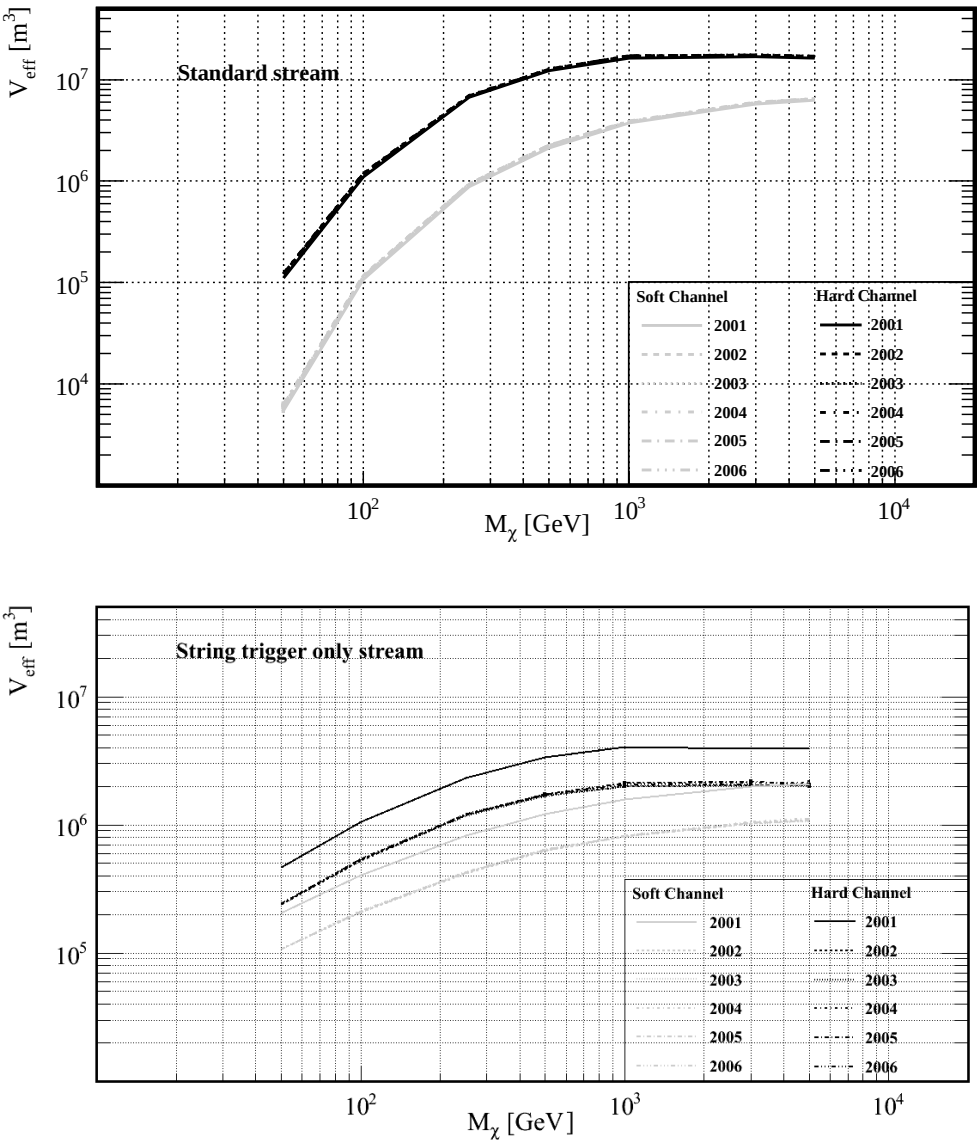


Figure 5.10: *As in fig. 5.9, the effective volumes for each neutralino model at L0 are shown, but split up per data-taking year; top (bottom) plot is the standard (string only) trigger stream, black (grey) lines stand for the hard (soft) channel. Also here, Statistical errors are not visible, since covered by the size of the lines*

mental ingredients can be not really well-known. Anyhow, when we have to check agreement experiment/simulation at later stages (see sec. 6.2.3.1), we should primarily look at the absolute shape of the observable under investigation, to gain a minimum of belief in the simulation. Further, when we will yield the final outcome of this analysis, our estimation of the atmospheric background rate will be based on off-source data and not on Monte Carlo simulation.

5.5.3 Low level filters

The aim of the event selection discussed in this section, is to reject down-going atmospheric muon background, which, as we have seen in the previous section, represented the dominant background. A first prompt selection criterion could be then to demand up-going events; hence, this event selection, which places only a requirement on the theta angle, could be just dubbed as a "low level" filter processing of experimental data and Monte Carlo simulations.

We started this chapter discussing about the precleaning procedure (sec. 5.2), i.e. hit cleaning and calibration of raw data, as a prelude to the event selection debated here. So, actually this is included in this low level filtering, which we want to divide in two main parts: one has been named as *First Level* (L1, sec. 5.5.3.1) and the other one as *Second Level* (L2, sec. 5.5.3.2). As it will be clear over the next subsections, the final requirement to pass L1 was on theta angle reconstructed by the first-guess method (see sec. 5.3.1), while for the still surviving events, the requirement to pass L2 was on theta angle reconstructed by the maximum likelihood method (see sec. 5.3.2).

Both low level filtering parts present two different approaches to the event selection just described; one was developed for the experimental data collected from 2001 to 2004, while the other for the experimental data collected during 2005 up to 2006. Naturally, we followed the same approaches to the Monte Carlo data-set, which we remind was entirely produced by the author of this work (see sec. 4.2)

The low level processing of the experimental data-set, which required intensive CPU-time, was accomplished, in different portion, by several groups of the AMANDA/IceCube Collaboration: DESY-Zeuthen [171], Uppsala [172], Brussels [173] and Madison [174].

5.5.3.1 First level

We have pointed out before, that two different approaches to low level filter characterized the 2001-2004 from the 2005-2006 data-set. In the next paragraphs we precisely discuss about this, focusing on the relevant parts for this work.

Processing of 2001-2004 data-set The complete L1 filter process proceeded in two parts.

The first was performed by the DESY-Zeuthen group (for more details see [100] and references therein), by means of the *classic* version of the `Sieglinde` code. The raw data information were then read from a *f2000* file to be calibrated, with the inclusion of a preliminary selection of good OMs and three hit selections. Next, three first-guess reconstructions were exploited. The "Zeuthen-processing" provided a generic data stream, with the purpose to be used by high energy neutrino analyses performed within the Collaboration. Hence, various event selections of interest to this high energy of searches were defined, which if applied in a later step, would remove all string-exclusive events.

At the end of the process, the data stream were then stored in *f2000* file format.

The second part of the L1 filtering was accomplished by the Uppsala group (more details in [175]). This "Uppsala-processing" added then the necessary information to perform low energy analyses to the data stream, released by the "Zeuthen-processing". The code used was again a classic[1] version of `Sieglinde`. So, once the Zeuthen data stream was read, an improved hit recalibration was done, the flare indicators (see sec. 5.5.4) were calculated, and the retriggering flags attached. Further, a new hit selection, based on a revised list of good OMs and on the introduction of crosstalk cleaning, was added. These final selected hits (actually the first hit of every hit OM) were then handed to a 32-fold iterative patched Pandel log-likelihood reconstruction, using as a seed track the JAMS outcome.

Finally, only events that fulfilled the Zeuthen up-going muon requirement, i.e.

$$\theta_{\mathrm{cDW}} > 70^{\circ} \quad \&\& \quad \mathrm{M24}$$

or events that fulfilled Uppsala up-going muon requirement, which recovered the string-exclusive ones, i.e.

$$\theta_{\mathrm{cDWimp}} > 70^{\circ}$$

were entitled to pass the L1 filtering, and therefore their information were written out in a final *f2000* file. θ_{cDW} was the theta angle as reconstructed by the classic version of DWalk, M24 the 24-fold multiplicity, or namely standard, trigger selection, and θ_{cDWimp} was the theta angle as reconstructed by the classic version of DWimp.

A schematic overview of the relevant parts of the complete L1 processing is shown in tab. 5.4.

[1] Two different versions of the *classic* code were used by the two filtering processes; furthermore another *classic* version was used for simulation. We checked and concluded that the mismatch of the versions did not introduce any bias between experiment and simulation, just as a confirmation from a previous solar Wimp analysis on 2001-2003 data [111].

	L1 2001-2004 Zeuthen-Uppsala (classic Sieglinde)
Hit Selections	OM, TOT cleaning: [HitSel0] OM, TOT, LE cleaning: [HitSel1] OM, TOT, Amp., Isol. cleaning: [HitSel2]
First Guess	classic Direct Walk (using HitSel1): [cDWalk] classic Direct Wimp (using HitSel2): [cDWimp] classic JAMS (using HitSel0): [cJAMS]
Flags	Flare indicators Retriggering (using HitSel0): M24 || String trigger
Hit Selection	OM, TOT, LE, Amp., Isol., XTalk cleaning: [Final HitSel]
Event Selections	$\theta_{\mathrm{cDWalk}} > 70°$: [DWalk 70] $\theta_{\mathrm{cDWimp}} > 70°$: [DWimp 70] DWalk70 && M24 trigger: [UpMu] DWimp70 || UpMu: [Uppsala]
LLH Reconstruction	Selected Final HitSel && Uppsala Patched 1pe Pandel 32 Iterations (seed from cJAMS): [c32JAMS]

Table 5.4: *Schematic overview of the complete L1 filter processing for 2001-2004 data.*

Processing of 2005-2006 data-set This data-set was processed by the Madison group using the *SLART* version of the `Sieglinde` code (more details in [176]). This filtering presented some differences with respect to the 2001-2004 processing in the hit cleaning, in the event selection and in the reconstruction.

The raw data information, in *f2000* format, were read and calibrated, and a good OM selection and four different hit selections were applied.
Unlike $2001-2004$ processing, a different strategy for the crosstalk cleaning was performed and applied early in the processing. Flare indicators and two kinds of retriggering selections were also evaluated, i.e. one considering only M24 (ReTrigM24), and the other considering M24 or string trigger (ReTrigWimp). Thus the events were flagged. Next, after requiring a minimum of 6 hits for each selection, three first-guess reconstructions were exploited. At the end a single (i.e. one search per seed track) Patched Pandel log-likelihood reconstruction was performed using, as a seed, tracks from JAMS and DWalk, although this reconstruction was not used later.

The "Madison-processing" of $2005-2006$ data provided a generic data stream suitable for high and low energy analyses; hence, different event selections were performed, but the event selection which suited our concerns was then

$$\theta_{\mathrm{JAMS}} > 70° \quad || \quad \theta_{\mathrm{DWimp}} > 70°$$

	L1 2005-2006 Madison (SLART Sieglinde)
Hit Selections	OM, XTalk cleaning: [HitSel0] OM, XTalk, TOT, LE, Amp. cleaning: [HitSel1] OM, XTalk, TOT, LE, Amp., Isol. cleaning: [HitSel2] OM, XTalk, TOT, LE, Amp., Isol, first LE hit: [HitSel3]
Flags	Flare indicators Retriggering (using HitSel1): M24 || String trigger: [ReTrigWimp] M24: [ReTrigM24]
First Guess	JAMS (using HiteSel1): [JAMS] Direct Walk (using HiteSel1): [DWalk] Direct Wimp (using HiteSel1): [DWimp]
Event Selections	$\theta_{\mathrm{JAMS}} > 70°$ [JAMS70] $\theta_{\mathrm{DWimp}} > 70°$ [DWimp70] JAMS70 || DWimp70: [L1 Stream]
LLH Reconstruction	Selected HitSel3 && L1 Stream Patched 1pe Pandel Single LLH (seed from JAMS and Direct Walk)

Table 5.5: *Schematic overview of the L1 filter processing for 2005-2006 data.*

Then, summarising, the events entitled to pass this L1 selection were the ones which satisfied the above requirement, and eventually, their information were stored in `ROOT` format files.

A schematic overview of the relevant parts of the complete L1 processing is shown in tab. 5.5.

Discussion The relative efficiency of the L1 event selection, applied to the entire experimental and Monte Carlo data-set, with respect to trigger level (L0), is shown in the second column of tab. 5.8. We notice that about 95% of experimental data were rejected, similarly to the simulated atmospheric muon background. The other background, i.e. the atmospheric neutrinos, and the neutralino-induced neutrinos were instead well kept; except for the lowest energy neutralino model, which lost out more than 30% of its efficiency, other models kept around 90% of their efficiency.

Looking at the simulated atmospheric background, we expected then that about 90% of the upgoing events should be reconstructed just as up-going[1], and indeed they are; whilst about 94% of down-going events were reconstructed as

[1] The Atmospheric neutrino background was simulated between [80°,180°] (see sec. 4.2.1), hence 10% of events should be down-going ones.

downgoing ones. The bulk of the remaining 6% could be misreconstructed horizontal events, mostly triggered by the string trigger, more challenging to reconstruct by the first-guess method.

So, a good reconstruction of string-exclusive events was a challenging task, not only to first-guess method, but to the log-likelihood reconstruction algorithm, from which we should demand for a good point resolution.

5.5.3.2 Second level

The events surviving the L1 filter, were subjected to a maximum log-likelihood reconstruction, with again the purpose to remove downgoing events. As we have already pointed out, the iterative log-likelihood reconstruction was high CPU-time consuming, but it had the advantage to perform a more accurate work with respect to the first-guess method. This was one of the tasks of the second low level filtering, whilst the other was the calculation of event observables (sec. 5.4), which were used later in other event selections, like the precut (sec. 5.5.4) and the high level filter (sec. 5.5.5).

Like for the L1 filtering, two different approaches to the event selection were developed for the 2001 – 2004 data and the 2005 – 2006 data.

The developing of the L2 filtering of the 2001 – 2004 data was accomplished by the Brussels group, and it was successfully applied in a previous AMANDA solar WIMPs analysis with 2001-2003 data [111]. Hence, with the 2001 – 2003 data already at L2, we proceeded on the same line to process 2004 data. The *SLART* `Sieglinde` version was used to process the 2001 – 2004 data-set, both for the improved reconstruction algorithm, and for the output analysis-ready `ROOT` file format.

The processing of 2005 – 2006 data was done by the Madison group, which coherently continued to make use of the *SLART* `Sieglinde` version. We have to mention here, for clarity's sake, that the L2 processing of 2005 – 2006 data was actually called "Level 3" by the Madison group; however, since it accomplished to remove downgoing events by means of a requirement on theta of the log-likelihood reconstruction, we can safely arrange and describe it in the next second paragraph.

Processing of 2001-2004 dataset The L1 calibrated events were read from a *f2000* file, and then the retrigger selection was applied. This selection removed more experimental events than simulated ones, since some electronic artifacts (transient OM, crosstalk) were not simulated.

We rerun in this processing the three first-guess reconstructions but in (improved) SLART style (sJAMS for instance); and before the iterative log-likelihood reconstruction, we applied another selection to speed up the processing, thus de-

	L2 2001-2004 Brussels (SLART Sieglinde)
First Guess	SLART Direct Walk: [sDWalk] SLART Direct Wimp: [sDWimp] SLART JAMS: [sJAMS]
Event Selections	$\theta_{\mathrm{sJAMS}} > 70°$: [sJAMS70] sJAMS70 && Retrigger: [L2b]
LLH Reconstruction	Patched 1pe Pandel 32 Iterations of L2b events (seed from classic and SLART JAMS): [32JAMS]
Event Selections	$\theta_{\mathrm{32JAMS}} > 80°$: [L2c]
LLH Reconstruction	Selection of L2c events: Parabola fit around best 32JAMS track: [32PARA] Bayesian Pandel downgoing track 32 iterations: [32BAYES]
Calculate Observables	Best 32JAMS track used as reference: [Topf32JAMS]

Table 5.6: *Schematic overview of the L2 filtering for 2001-2004 data.*

manding

$$\theta_{\mathrm{sJAMS}} > 70°$$

This was a quite relaxed cut which kept between 80% - 90% of the signal; the experimental and simulated atmospheric muon background were removed by a fraction around 67% - 68%, while the atmospheric neutrino background by a fraction around 10%.

Then the events which passed the above selection were handed to three likelihood reconstructions (see sec. 5.3.2 for details). The first was 32x Iterative Patched Pandel log-likelihood reconstruction (32JAMS); then the selection applied to the reconstructed events was

$$\theta_{\mathrm{32JAMS}} > 80°$$

Thus, the events satisfying the above requirement were passed to the second reconstruction, i.e. a gradient parabolic minimisation fit around the best 32JAMS solution (see e.g. fig. 5.7), whose 1σ area was a measure of the event resolution. The third reconstruction was a Bayesian one, which maximised the same Iterative 32x Patched Pandel function, but including the a priori downgoing track hypothesis; the comparison of its best solution with the one from 32JAMS could help to distinguish upgoing from downgoing events.

So, summarising, the events which finally survived the 32JAMS selection were entitled to pass the L2 filtering.

A schematic overview of the L2 filter for $2001-2004$ data is shown in tab. 5.6.

	L2 2005-2006 Madison (SLART Sieglinde)
Event Selections	$\theta_{\text{JAMS}} > 85°$ [JAMS85] ReTrigWimp && JAMS85 : [L2Wimp] $\theta_{\text{DWalk}} > 80°$ [DWalk80] ReTrigM24 && DWalk80: [L2UpMu]
LLH Reconstruction	Selected (L2UpMu || L2Wimp) events: [L2stream] Patched 1pe Pandel 32 Iterations: (seed from JAMS and Direct Walk): [Muon32] Parabola fit around best Muon32 tracks: [ParaMuon32]
Event Selections	Selected L2Stream events: $\theta_{\text{Muon32}} > 80°$: [L3Muon32]
LLH Reconstruction	Selected L3Muon32 events: Bayesian Pandel downgoing track 64 iterations: [64Bayes]
Calculate Observables	Used best Muon32 track as reference: [TopfMuon32]

Table 5.7: *Schematic view of the L2 filtering for 2005 and 2006 data.*

Processing of 2005-2006 dataset The L1 calibrated events were read from `ROOT` files, then two cuts on theta from first-guess reconstructions were applied, one from JAMS (JAMS85)

$$\theta_{\text{JAMS}} > 85°$$

and the other from Direct Walk (DWalk80)

$$\theta_{\text{DWalk}} > 80°$$

just to speed up the processing, like the one for $2001 - 2004$ data. Next, two event selections were performed, in conjunction with the two different retriggering selections (see the 2nd par. of sec. 5.5.3.1)

ReTrigWimp && JAMS85

and

ReTrigM24 && DWalk80

Then the events which passed at least one of the above selections, were handed to two likelihood reconstructions; the first was an 32x Iterative Patched Pandel log-likelihood reconstruction (Muon32), and the second a gradient parabolic minimisation fit around the best Muon32 solution, whose 1σ area was again a measure of the event resolution.

Next, another event selection was defined, requiring from theta reconstructed by Muon32 the following

$$\theta_{\text{Muon32}} > 80°$$

Passing rate	L1 w.r.t. L0 [ε_{L_1}]	L2 w.r.t. L1 [ε_{L_2}]
Exp. data	0.051	0.055
Atm. μ	0.054	0.052
Atm. ν	0.869	0.792
50 hard	0.739	0.689
100 hard	0.874	0.755
250 hard	0.927	0.814
500 hard	0.934	0.826
1000 hard	0.938	0.832
3000 hard	0.938	0.834
5000 hard	0.938	0.834
50 soft	0.564	0.645
100 soft	0.738	0.694
250 soft	0.871	0.762
500 soft	0.903	0.792
1000 soft	0.918	0.807
3000 soft	0.925	0.816
5000 soft	0.927	0.818

Table 5.8: *Efficiencies for experimental 2001-2006 data and Monte Carlo simulation at L1 (w.r.t. trigger level, or L0) and L2 (w.r.t. L1)*

Further, the events which passed the above selection, were given to a third reconstruction, a 64x Iterative Bayesian one; again, its best solution compared with the one from Muon32 could help to distinguish upgoing from downgoing events.

So, summarising, the events which finally survived the Muon32 selection were entitled to pass the L2 filtering.

A schematic overview of the L2 filter for $2005-2006$ data is shown in tab. 5.7.

Discussion The third column of tab. 5.8 shows the selection efficiency of the L2 with respect to L1 filtering, concerning experimental data and Monte Carlo simulations. We infer then, that about 6% of $2001-2006$ experimental data survived the L2 filtering step, and that about the same fraction of atmospheric muon background passed the filtering. The atmospheric neutrino background survived the L2 step in a fraction of 80%. For high energy neutralino models around 80% to 84% of the signal was kept, while for low energy ones it was around 65% to 76%. The lower efficiency of L2 w.r.t L1 is due, as expected, to the more sophisticated log-likelihood reconstruction compared to the first-guess method.

$V_{eff(L2)}$	[m^3]	50	100	250	500	1000	3000	5000
ALL	hard	$2.03{\cdot}10^5$	$1.18{\cdot}10^6$	$6.23{\cdot}10^6$	$1.12{\cdot}10^7$	$1.51{\cdot}10^7$	$1.54{\cdot}10^7$	$1.50{\cdot}10^7$
	soft	$4.72{\cdot}10^4$	$1.83{\cdot}10^5$	$9.41{\cdot}10^5$	$2.12{\cdot}10^6$	$3.56{\cdot}10^6$	$5.37{\cdot}10^6$	$5.85{\cdot}10^6$
STD	hard	$0.80{\cdot}10^5$	$0.88{\cdot}10^6$	$5.58{\cdot}10^6$	$1.03{\cdot}10^7$	$1.40{\cdot}10^7$	$1.43{\cdot}10^7$	$1.39{\cdot}10^7$
	soft	$0.33{\cdot}10^4$	$0.78{\cdot}10^5$	$7.09{\cdot}10^5$	$1.77{\cdot}10^6$	$3.11{\cdot}10^6$	$4.81{\cdot}10^6$	$5.25{\cdot}10^6$
STR	hard	$1.23{\cdot}10^5$	$0.30{\cdot}10^6$	$0.65{\cdot}10^6$	$0.09{\cdot}10^7$	$0.11{\cdot}10^7$	$1.11{\cdot}10^7$	$0.11{\cdot}10^7$
	soft	$4.39{\cdot}10^4$	$1.05{\cdot}10^5$	$2.32{\cdot}10^5$	$0.35{\cdot}10^6$	$0.45{\cdot}10^6$	$0.56{\cdot}10^6$	$0.60{\cdot}10^6$

Table 5.9: *Effective volumes at L2 for the different neutralino models (mass and channel), split up for the different trigger selections: STR (exclusive string trigger), STD (inclusive standard multiplicity trigger), ALL (logical sum of the two previous selections).*

The total reduction of L2 experimental data and simulated atmospheric muon background, with respect to L0, corresponds to a factor $\sim 3 \times 10^{-3}$; while about 31% of the simulated atmospheric neutrino background was rejected.

High energy neutralino models kept around 76% - 78% of their efficiency w.r.t. L0, while the lowest energy ones kept around 36% - 50%, being a real challenge for the continuation of the analysis.

The effective volume at L2 for the simulated neutralino signal is shown in tab. 5.9; likewise for trigger level, the three different classes (STR, STD, ALL) are outlined for each neutralino model. The numbers in the table, related to low energy models, show that the bulk of events which passed the L2 selection still resulted from exclusive string trigger selection.

In fig. 5.11 the relevant numbers of table 5.9 are plotted (hard channels on the top, while soft channels on the bottom); the different trigger classes are respectively drawn with a solid line (ALL), dashed line (STD) and dotted line (STR).

In fig. 5.12 the effective volumes for each neutralino model (black colour for hard channels, and grey for soft channels) are shown again, but this time different lines represent the different years of simulated data.

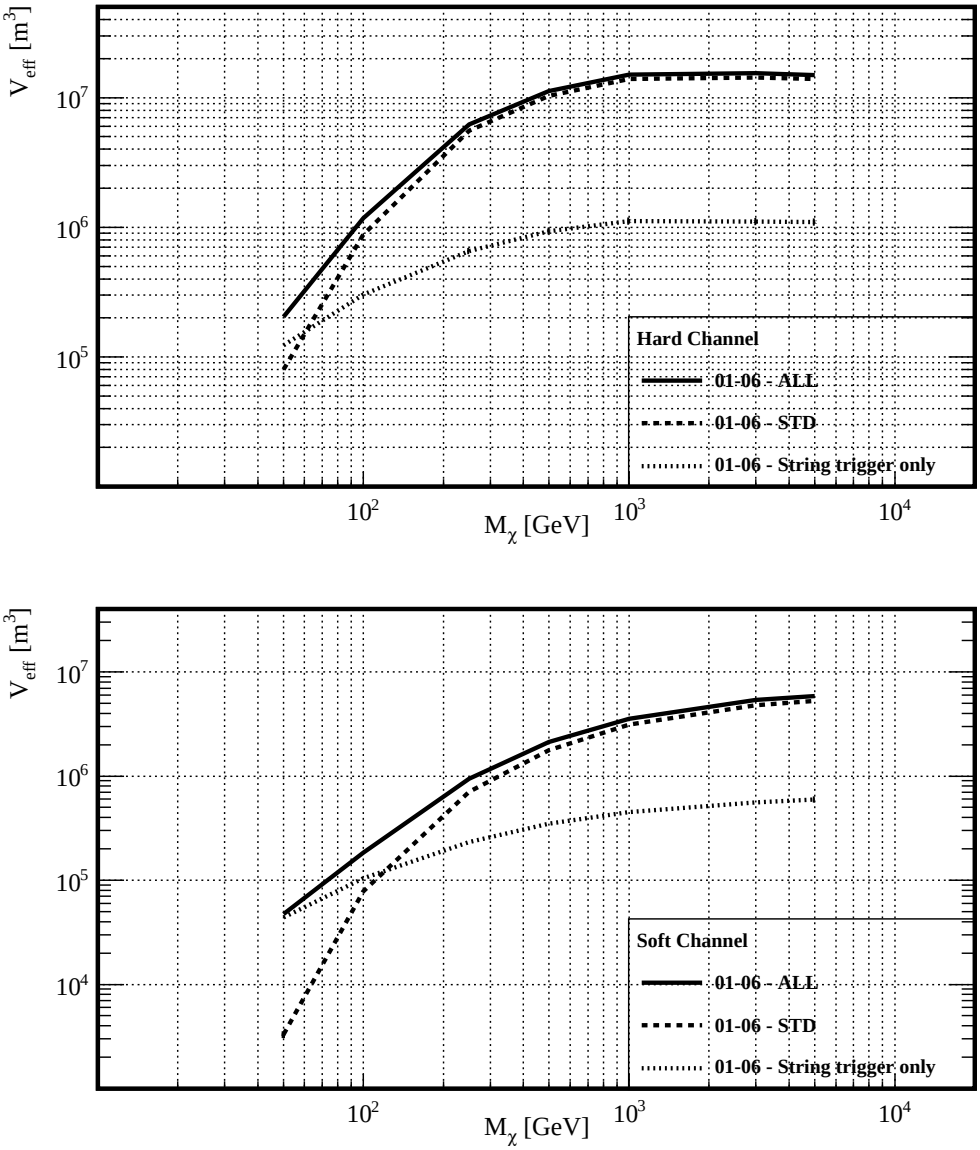


Figure 5.11: *Effective volume at L2 as a function of the neutralino mass for hard (top) and soft (bottom) annihilation channel. The dotted line refers to events triggered only by the string trigger, while the dashed one to events triggered by the standard trigger. The solid line is the logical sum of the two trigger selections. Statistical errors are not visible in some lines, as they are covered by the size of the lines*

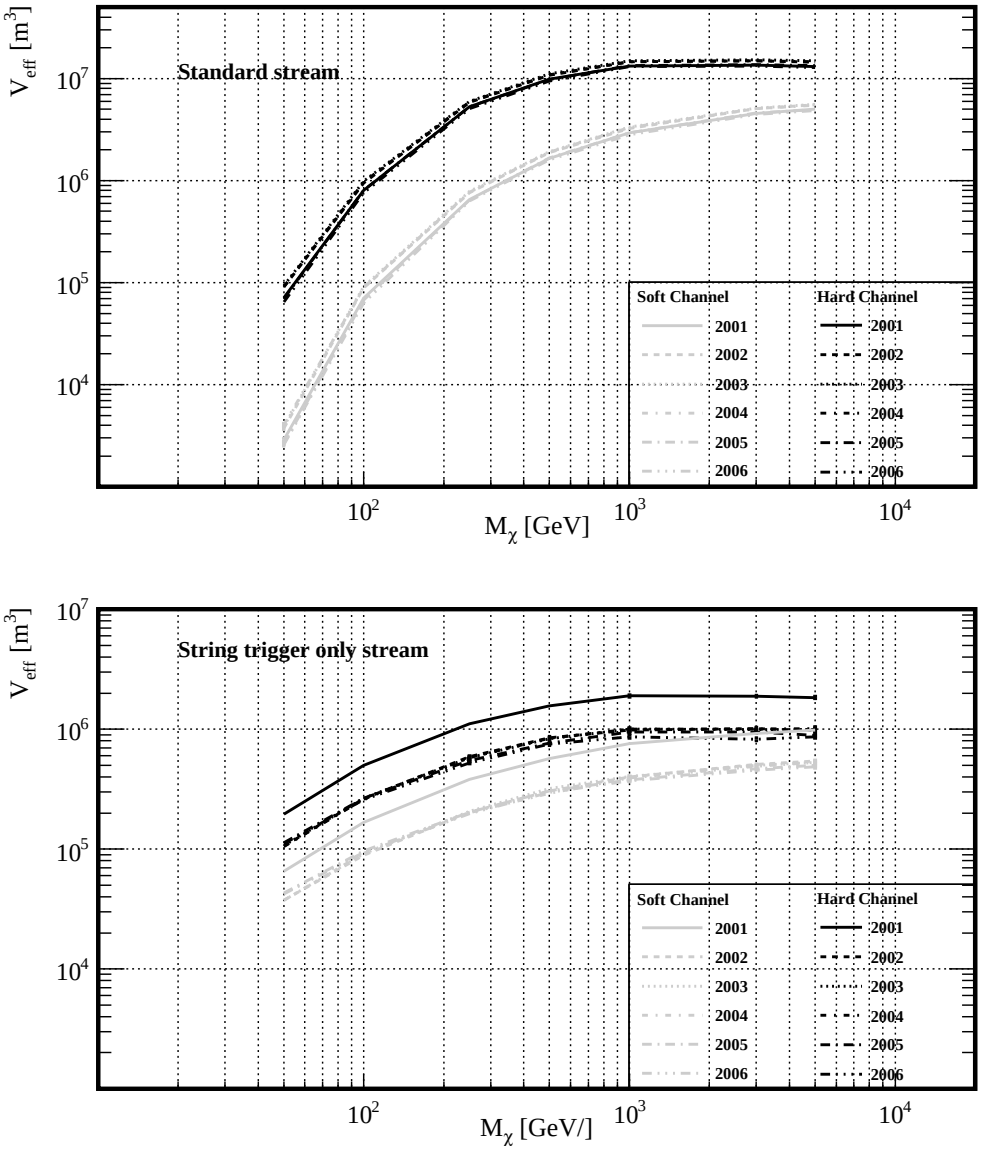


Figure 5.12: *As in fig. 5.11, the effective volumes of each neutralino models at L2 are shown, but split up per data-taking year; top (bottom) plot is the standard (string only) trigger stream, black (grey) lines stand for the hard (soft) channel. Also here, Statistical errors are not visible, but covered by the size of the lines*

Flare indicators	2001	2002	2003	2004	2005	2006
long_noise_1	3	5	5	5	5	5
long_missing_2	3.5	3	7	7	7	7
only_adc_3	18	20	3	10	10	10
nch_dead_4	4	9	4	10	10	10
short_H_5	4	4	6.5	7	7	7
short_M_6	3	5	3.5	10	6	6
missing_ch_7	6	3	3	20	8	8
induc_B10_8	5	4	5	10	6	6
induc_1119_9	3	3	9	10	6	6

Table 5.10: *The flare indicators and their relative cut value, for the different years of the experimental data-set.*

5.5.4 Precuts

Prior to continue the analysis applying the high level filter selection, we had to perform some precuts; some of them were applied only to the experimental data, while others both to experimental data and Monte Carlo simulations.

Precuts on experimental data

The two precuts described here were applied only on experimental data since there were no relative counterparts in the Monte Carlo simulation.

The first precut removed the events (although a negligible amount) triggered only by the SPASE trigger selection (see sec. 3.2.3); the second one removed some non-particle events of unknown origin, the so-called flare events (see sec. 3.2.1). We have looked at the nine flare indicators, as suggested by [161], and applied some cuts to remove non-exponential tails in their distributions. The year-dependent cut values are summarised in table 5.10.

At the end, these two precuts removed less than about 0.5% of the experimental data.

Precuts on experimental data and Monte Carlo simulations

These sequential precuts, applied on both 2001-2006 experimental data and Monte Carlo simulation, were performed in order to remove low quality events, which were mostly triggered by the string trigger. We chose 4 different indicative observables, which are in order: number of hits (N_{Hit}), number of direct hits (N_{dir}), distance from the vertical axis of the centre of gravity of the hits (ρ_{COG}), and reconstructed theta from the log-likelihood reconstruction, (θ_{LLH}). This last powerful cut removed tracks from above the horizon; this means that the exper-

Variables	Cut
Number of hits	$N_{\mathrm{Hit}} > 10$
Number of direct hits	$N_{\mathrm{dir}} > 5$
Distance of the COG of the hits from the Z axis	$\rho_{\mathrm{COG}} < 80$ [m]
Log-likelihood reconstruction of theta angle	$90^0 < \theta_{\mathrm{LLH}} < 130^0$

Table 5.11: *Observables used for the precuts and their cut values (see text for observable explanations).*

imental *BDT optimisation sample* (see sec. 4.1.4) was extremely signal depleted after this cut (since the Sun is in the other hemisphere), and hence suitable for background studies (see sec. 5.5.5.2).

The first cut on N_{Hit} kept events which had more than ten hits, the second cut on N_{dir}, events with more than five direct hits, and the third cut on ρ_{COG} kept events which had hits whose centre of gravity projection in the horizontal plane was less than 80 meters away from vertical (Z) axis. The last cut was applied on theta angle reconstructed by the log-likelihood method, θ_{LLH}: events that were between 90^0 and 130^0 degrees were kept (see for instance the top part of fig. 4.11, even if it is at the trigger level). The four observables and their cut value are summarised in table 5.11, while in fig. 5.13 they are visualised in the same order as the table; the filled grey histogram represents the experimental data, the solid black line the simulated atmospheric muon background, the dotted black line the simulated atmospheric neutrino background, and the dashed line the 100 GeV soft neutralino model. The vertical solid lines and the arrows mark the regions of selected events.

This step removed $\sim 90\%$ of the data, $\sim 92\%$ of the simulated atmospheric muon background, $\sim 55\%$ of the simulated atmospheric neutrino background and around $20\%-40\%$ of the simulated neutralino signals. The precut efficiencies with respect to L2, for experimental data and Monte Carlo simulation, are summarised in table 5.12.

Hence, summing up, after these precuts, the total reduction of experimental data, with respect to L0, corresponds to a factor $\sim 2.7 \times 10^{-4}$, and to a factor $\sim 2.4 \times 10^{-4}$ for the atmospheric muon background; whilst about 69% of the simulated atmospheric neutrino background was rejected.

High energy neutralino models kept around 58% - 60% of their efficiency w.r.t. L0, while the lowest energy ones around 19% - 35%.

The effective volume for the simulated neutralino signal after the precuts is shown in tab. 5.13, again split up in the three different classes (STR, STD, ALL) for each neutralino model. The numbers in the table, related to low energy models, show that the bulk of events which passed the precut selection still resulted

Passing rate	Precuts w.r.t L2 [ε_{precut}]
Exp. data	0.096
Atm. μ	0.084
Atm. ν	0.451
50 hard	0.665
100 hard	0.768
250 hard	0.783
500 hard	0.777
1000 hard	0.775
3000 hard	0.775
5000 hard	0.774
50 soft	0.517
100 soft	0.673
250 soft	0.766
500 soft	0.780
1000 soft	0.782
3000 soft	0.779
5000 soft	0.780

Table 5.12: *Precut efficiencies, w.r.t. L2, for experimental data and Monte Carlo simulations.*

$V_{eff(precut)}$	[m^3]	50	100	250	500	1000	3000	5000
ALL	hard	$1.36{\cdot}10^5$	$9.06{\cdot}10^5$	$4.88{\cdot}10^6$	$8.74{\cdot}10^6$	$1.17{\cdot}10^7$	$1.19{\cdot}10^7$	$1.16{\cdot}10^7$
	soft	$2.45{\cdot}10^4$	$1.24{\cdot}10^5$	$7.20{\cdot}10^5$	$1.65{\cdot}10^6$	$2.78{\cdot}10^6$	$4.18{\cdot}10^6$	$4.56{\cdot}10^6$
STD	hard	$0.64{\cdot}10^5$	$7.26{\cdot}10^5$	$4.55{\cdot}10^6$	$8.31{\cdot}10^6$	$1.12{\cdot}10^7$	$1.15{\cdot}10^7$	$1.11{\cdot}10^7$
	soft	$0.24{\cdot}10^4$	$0.63{\cdot}10^5$	$5.84{\cdot}10^5$	$1.46{\cdot}10^6$	$2.54{\cdot}10^6$	$3.87{\cdot}10^6$	$4.26{\cdot}10^6$
STR	hard	$0.72{\cdot}10^5$	$1.80{\cdot}10^5$	$0.33{\cdot}10^6$	$0.43{\cdot}10^6$	$0.05{\cdot}10^7$	$0.04{\cdot}10^7$	$0.05{\cdot}10^7$
	soft	$2.21{\cdot}10^4$	$0.61{\cdot}10^5$	$1.36{\cdot}10^5$	$0.19{\cdot}10^6$	$0.24{\cdot}10^6$	$0.31{\cdot}10^6$	$0.30{\cdot}10^6$

Table 5.13: *Effective volumes after the precuts for all the neutralino models (mass and channel), split up for the different trigger selections: STR (exclusive string trigger), STD (inclusive standard multiplicity trigger), ALL (logical sum of the two previous selection).*

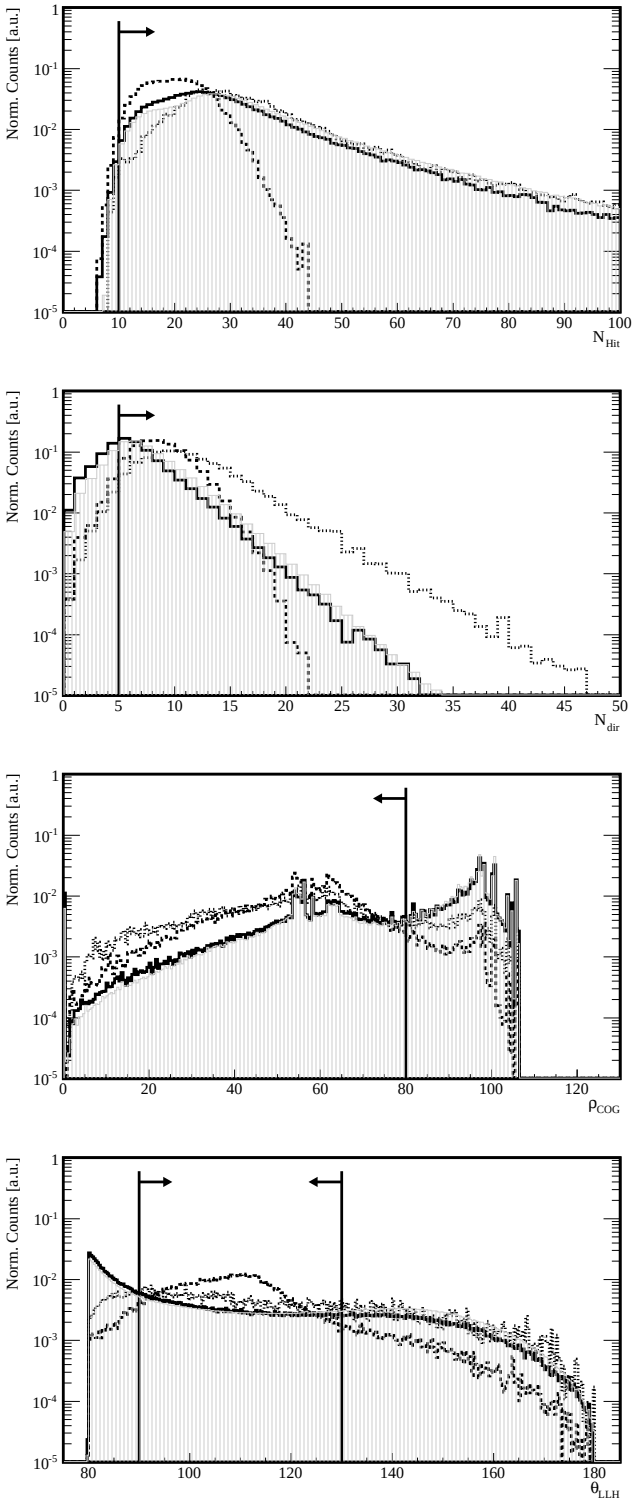


Figure 5.13: *Observables used for the precut step; filled grey histogram are experimental data, solid line are atmospheric muons, dotted line are atmospheric neutrinos, and dashed line represents 100 GeV soft neutralino. Solid vertical line and arrows mark the region of selected events.*

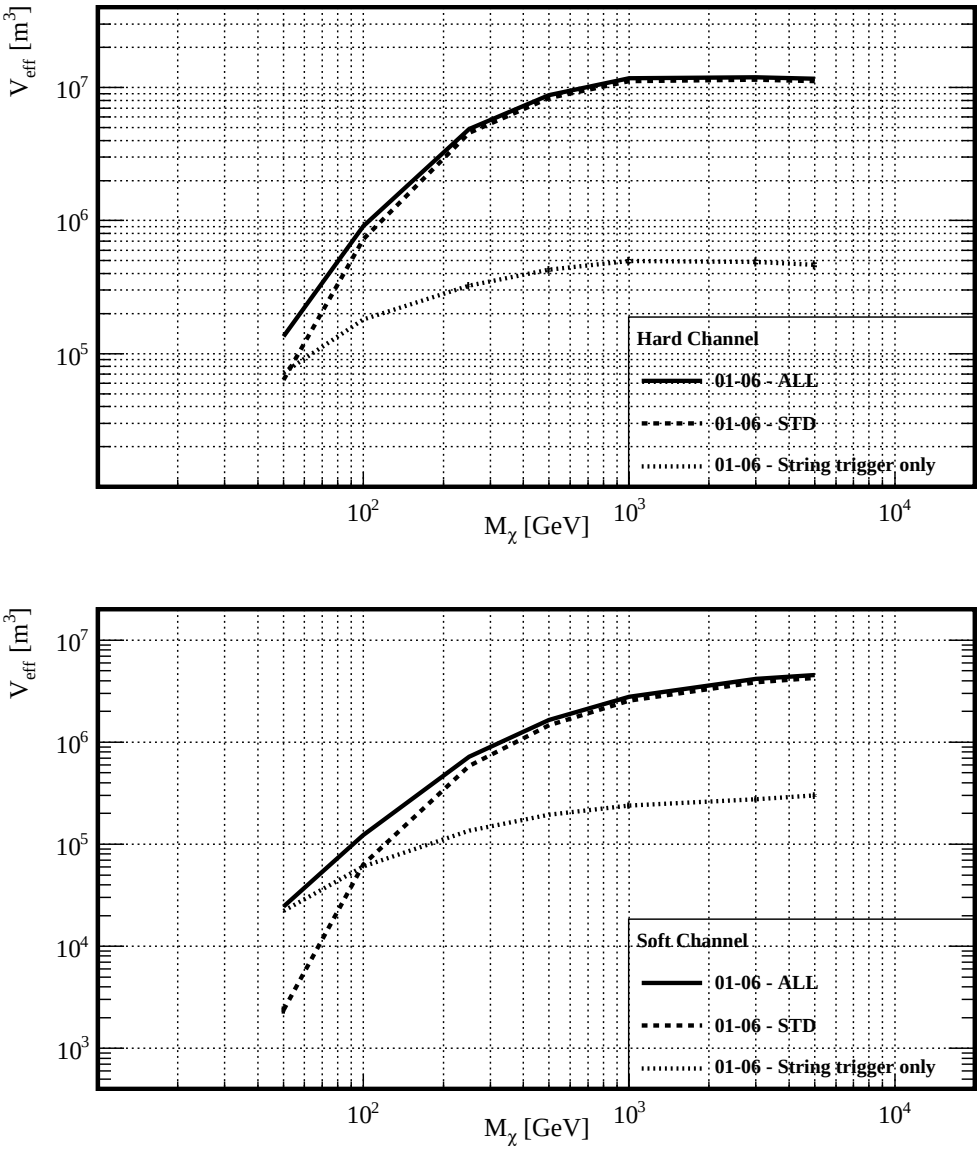


Figure 5.14: *Effective volume after precuts as a function of the neutralino mass for hard (top) and soft (bottom) annihilation channel. The dotted line refers to events triggered only by the string trigger, while the dashed one to events triggered by the standard trigger. The solid line is the logical sum of the two trigger selections. In some lines statistical errors are not visible, covered by the size of the lines.*

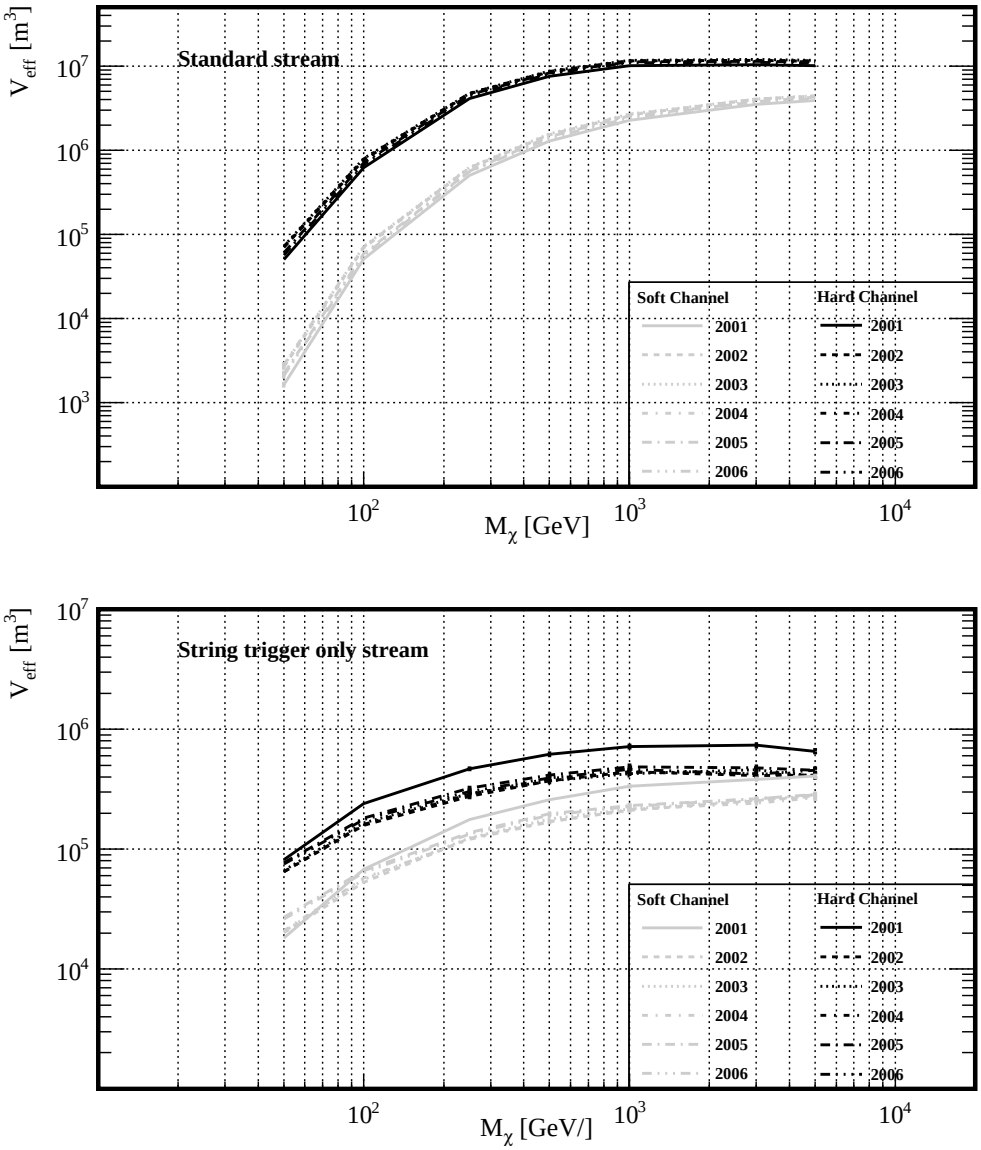


Figure 5.15: *As in fig. 5.9, the effective volumes for each neutralino model after precuts are shown, but split up per data-taking year; top (bottom) plot is the standard (string only) trigger stream, black (grey) lines stand for the hard (soft) channel. Also here, in some lines the statistical errors are not visible, since covered by the size of the lines.*

from exclusive string trigger selection.

In fig. 5.14 the relevant numbers of table 5.13 are plotted (hard channels on the top, while soft channels on the bottom); the different trigger classes are respectively drawn with a solid line (ALL), dashed line (STD) and dotted line (STR).

In fig. 5.15 the effective volumes for each neutralino models (black colour for hard channels, and grey for soft channels) are shown again, but where different lines represent the different six years of simulated data.

5.5.5 High level filter

The Low level filters and precuts removed events which were clearly reconstructed as downgoing ones. However, the experimental data after the precuts[1] were still contaminated by misreconstructed atmospheric muons, making it absolutely hard extracting any neutralino signal from the Sun.

Hence, a further filtering step was needed in order to remove this background, based on some "high quality" discriminating observables, which could be combined in a multidimensional space to perform a "high level" filtering.

We have then used the *Boosted Decision Trees* (BDTs) [177], which is a multivariate method of data classification, to reach the background rejection level (roughly 10^{-7}) demanded for a search for neutralino induced neutrinos from the Sun.

In sec. 5.5.5.1 we will give an overview of the classifier, while in sec. 5.5.5.2 we will describe the optimisation procedure, performed in order to select the final experimental data sample. This filtering step constitutes indeed the last cut before the "hypothesis testing", where the remaining data will be used to test the hypothesis that they contain background and a neutralino signal from the Sun versus a background-only scenario (see next chapter).

5.5.5.1 BDT description

The BDTs represent an extension to a single decision tree (DT) [178], which classifies events grounded on an ensemble of cumulative selection criteria (cuts). These selection criteria define several separate subsets of events, each with a different signal purity.

The training process of a DT, given some signal and background samples, can be outlined as follows. Let's assume there is a certain number of observables, for each event, suitable to discriminate between signal and background. Then, for each observable, the events are ordered by its calculated value; hence, fixing one observable, for each event value the sample is split in two parts (left and

[1] The analysis data sample, after the precuts, included about 2×10^6 events.

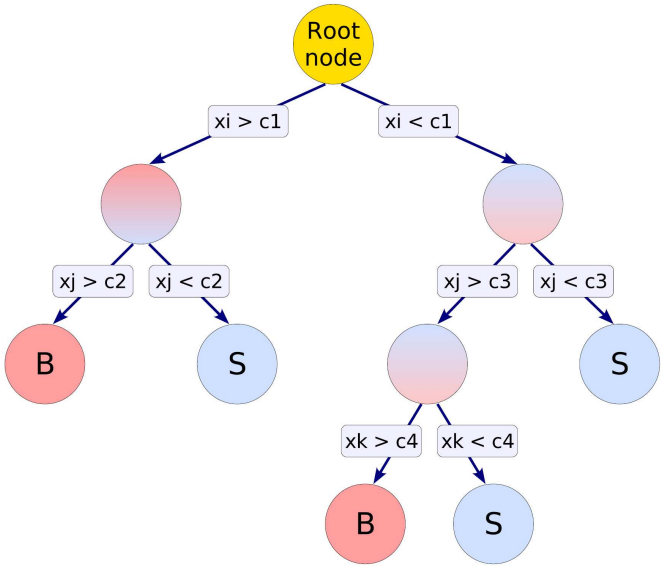


Figure 5.16: *Sketch of a decision tree. Starting from the root node, a sequence of binary splits using the discriminating observables {x} is applied to the sample. Each split uses the observable that at this node gives the best separation between signal and background when being cut on. The same observable may thus be used at several nodes, while others might not be used at all. The leaf nodes at the bottom end of the tree are labelled "S" for signal and "B" for background depending on the majority of events that end up in the respective nodes. Picture taken from [160]*

right), depending on the value of that observable. So, the algorithm picks the splitting value which gives the best separation into one side having mostly signal, and the other mostly background. This is then repeated for each observable in turn, i.e. the selection of the variable and the splitting value which gives the best separation.

At the training start, the event sample was at a "node", and now after splitting there are two samples called "branches". Hence, for each new branch, the process described above is repeated, i.e. finding for each observable the best cut value, and thus the best observable for that branch. The splitting process should continue until a given number of final branches, called "leaves", are obtained, or until each leaf is pure signal or pure background, or has too few events to continue. This description is a little oversimplified; indeed at each stage one picks as the next branch to split, the branch which will give the best increase in the quality of the separation. A simple sketch of a decision tree is drawn in fig. 5.16.

A criterion to define the quality separation between signal and background

can be introduced defining a purity P of the sample in branch, as follows

$$P = \frac{\sum_s w_s}{\sum_s w_s + \sum_b w_b} \tag{5.21}$$

where $w_{s(b)}$ is the signal (background) event weight, and $\sum_{s(b)}$ is the sum over signal (background) events. Hence, $P(1-P)$ is equal to 0 if the sample is pure signal or pure background. For a given branch we can then define the impurity measure, or Gini index

$$\text{Gini} = \left(\sum_{i=1}^{n} w_i\right) P(1-P) \tag{5.22}$$

where n is the number of events on that branch.

A separation quality criterion can be the minimisation of the following expression

$$\text{Gini}_{\text{left}_{\text{son}}} + \text{Gini}_{\text{right}_{\text{son}}} \tag{5.23}$$

and the following expression can be maximised, to settle the increase in quality when a node is split into two branches

$$\text{Gini}_{\text{parent}} - (\text{Gini}_{\text{left}_{\text{son}}} + \text{Gini}_{\text{right}_{\text{son}}}) \tag{5.24}$$

If a leaf will end up with a purity greater than 1/2 (or whatever is set), then it will be called as "signal-leaf", while if the purity is less than 1/2 then it will be a "background-leaf". Hence, events will be classified as "signal-like" if they will land on a signal leaf, conversely as "background-like" on a background leaf, and the resulting tree will be a *decision tree* (DT). These latter are known to be powerful but unstable, i.e. a small change in the training sample can bring a large change in the outcome.

BDTs can overcome this problem, and in the same time be very strong classifier. BDTs are a forest of many DTs, where the classification is performed by averaging the outcome of all the trees in the forest. The growth of the forest occurs iteratively, with the new tree trained with the same sample of events as the previous tree, but with an increased, boosted, weight for the events that were previously misclassified. The boosting algorithm used here is the so-called `AdaBoost` (adaptive boost) [179]. Starting with the original event weights when training the first decision tree, the subsequent tree is trained using a modified event sample where the weights of previously misclassified events are multiplied by a common boost weight α, which is derived from the misclassification rate of the previous tree.

In principle, the splitting of the tree could continue until each leaf node contains only signal or only background events, which could suggest that perfect discrimination is achievable. However, such a decision tree would be strongly overtrained. To avoid overtraining a decision tree must be pruned. This pruning is done using the *cost complexity method* [178], $\mathcal{R}$, where the missclassification rate in each node, r, is compared with that of the subtree below it, $\mathsf{r_d}$. Then the subtree is removed if

$$\mathcal{R} = (\mathsf{r} - \mathsf{r_d})/(\mathsf{n} - 1) < \mathsf{p}$$

where n is the number of son nodes, and p is a user defined pruning strength. The pruning is done by removing the nodes with the lowest $\mathcal{R}$ recursively, until no node with $\mathcal{R} < \mathsf{p}$ is left.

In the literature decision trees are sometimes referred to as the best "out of the box" classifiers. This is because little tuning is required in order to obtain reasonably good results. This is due to the simplicity of the method where each training step (node splitting) involves only a one-dimensional cut optimisation. Decision trees are also insensitive to the inclusion of poorly discriminating input observables.

A ranking of the BDT input observables can be derived by counting how often the observables are 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. This measure of the variable importance can be used for a single decision tree as well as for a forest.

5.5.5.2 BDT optimisation

In the *training* phase of the BDT classifier, we used the second half of our signal Monte Carlo sample, which we called *MC BDT optimisation sample*; we prevented in this way any possible statistical bias on the final outcome of the analysis, which we remind, will be obtained with the first half sample, named *MC analysis sample*.

A subsample of the experimental data, the BDT optimisation sample (periods with the Sun above the horizon, see sec. 4.1.4), was used to describe the background distribution. We did not use the simulated atmospheric muon sample in the training, first of all because we could not produce large samples to avoid statistical fluctuations, due to some time constraints. Further, the simulated background could introduce some bias during the training phase, due to some uncertainties in the model of cosmic ray interactions with the atmosphere. The simulated atmospheric muon and neutrino samples were only used to check the agreement experiment/simulation, and possibly gain an additional confidence in the signal simulation.

Option	Value	Description
`nTrees`	800	Number of trees in the forest
`BoostType`	AdaBoost	Boosting type for tree building
`SeparationType`	Gini Index	Separation criterion applied for the node splitting
`nEventsMin`	10	Minimum number of events in a node where further splitting is stopped
`nCuts`	40	Number of steps in the scan to optimise the cut at a node
`PruneMethod`	CostComplexity	Pruning method
`PruneStrength`	2.5	Amount of pruning: it should be large enough such that overtraining is avoided and needs to be tuned for each analysis

Table 5.14: *Configuration options for the BDT classifier in TMVA package.*

As we said at the beginning of this chapter, we used the `TMVA` package as ground software for BDT training and application. Further, we implemented a code that interfaced this software package, based on `ROOT`, with the experimental data and Monte Carlo simulation files, to extract the needed information. The implemented code was designed flexible enough to support different user defined options, like the handling of input observables and neutralino models to optimise. Several configuration options are available in the `TMVA` package to customise the BDT classifier, which are summarised in tab. 5.14; the more relevant ones were already described in more details in the previous section.

From fig. 5.14, which shows the effective volumes after the precuts, we infer that the detector efficiency was quite similar for neutralino models with mass of 1 TeV and above. Below this value, for the other neutralino models (especially in the low mass region), the differences in efficiency are quite visible. However, in principle this should not prevent to group the fourteen neutralino models into some low and high energy ensembles, choosing then only a couple of templates to perform BDT optimisation. Instead, for this work we chose to optimise all the neutralino models separately, especially because we wanted to verify the impact of this BDT optimisation, almost newly used in AMANDA analyses, to try to improve the efficiency and thus the sensitivity to solar neutralino dark matter. Further, the separate optimisation of all the models did not introduce any extra effort, or time delay, since we had our GRID facilities available.

The BDT algorithm presents the advantage of coping with a large number of

input variables, where more input observables will cover more information which may be useful to improve the separation between signal and background. Some of the observables have more discriminant power with respect to the others, and some observables may have correlations with others; hence it is necessary to select the most sensitive observables to maximise the BDT performance. Hence, for this investigation we have chosen a low and a high energy neutralino template, i.e. 100 GeV hard and 5000 GeV hard, with the purpose to obtain a unique final list of observables suitable for all the neutralino models. We started then to feed the BDT with several observables, whose descriptions are found in sec. 5.4. As we have mentioned at the end of the previous section, a rank can be assigned to the different observables; indeed at the end the optimisation tests, the training algorithm provided a list of ranked observables. Thus a preliminary sorted list of observables was written out after removing the ones with lower rank.

The next step in the investigation was a correlation study[1] between the remaining observables; we demanded for two observables not to be correlated by more than 65%. Between two correlated observables, the one with lower rank was then removed from the preliminary list. We performed this study, looking both at the signal and at the background observable distributions since their correlation patterns are usually different. Hence, we ended up with a final list of 21 observables. The linear correlation coefficients between all input observables, for the 100 GeV hard signal (background), are displayed on the top (bottom) of fig. 5.17, as a different colour gradation. The numbers in the picture correspond to the observables listed in appx. A.1. The plots of the final 21 observables for experimental data, atmospheric background and 500 GeV hard neutralino are shown in appx. A.2. The 21 selected observables show a good agreement experimental data/Monte Carlo background simulation.

The study that produced the final list of observables, was performed considering only one particular year of the complete 2001 – 2006 dataset. Hence, the next step in the observable investigation was to verify their distributions with respect to the other years.

During the period 2001 – 2006, the detector underwent several maintenances or upgrades, like the downscaling of the string trigger from 2002 onwards (see chapter 3). Further, as we explained in sec. 5.5.3, a different approach to low level filters was developed for the 2005 – 2006 experimental data. At the end of this investigation we inferred that some observables showed remarkable differences among three distinct subsamples of data: 2001, 2002 – 2004 and 2005 – 2006, due to the motivations mentioned before. On the top of fig. 5.18 we show a particular observable, the difference in the Bayesian and standard reduced log-

[1] Linear correlations between observables were outlined by `TMVA` package through computing the square-root of the covariance matrix.

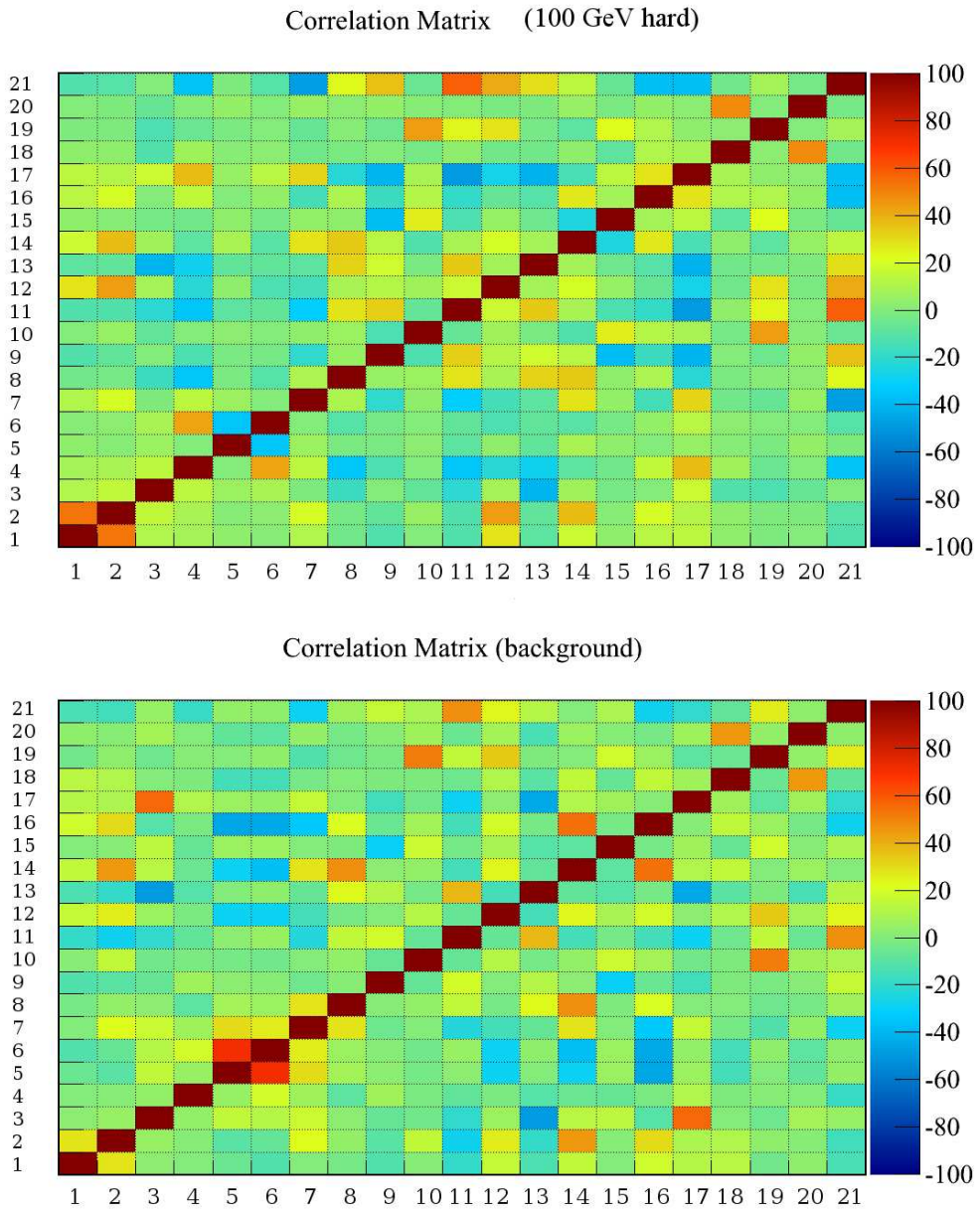


Figure 5.17: *Linear correlation coefficients between the 21 selected observables as a different colour gradation, for the signal (top) and background (bottom) training samples. For the correspondence number - observable, see appx. A.1.*

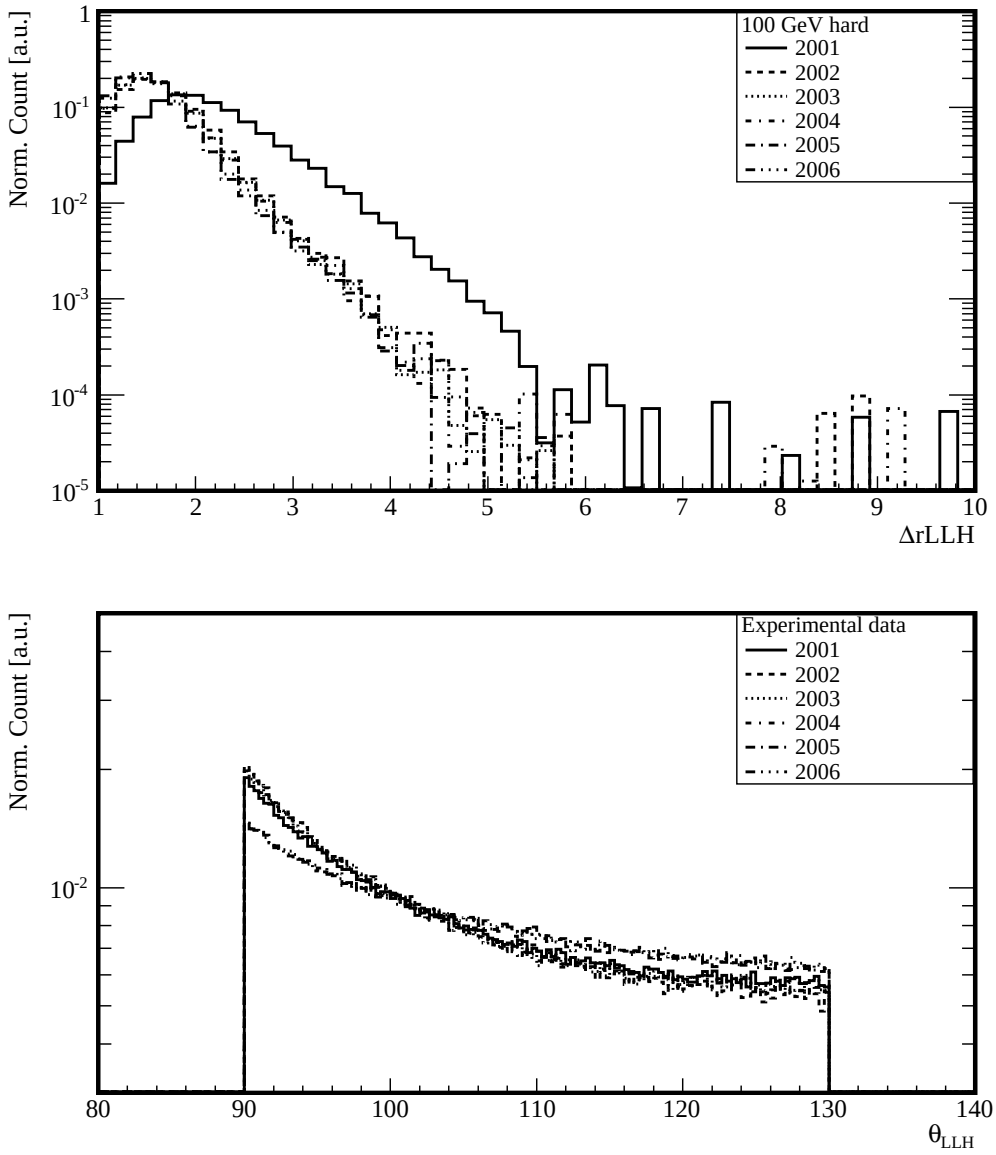


Figure 5.18: *On top, the distribution of the difference in the Bayesian and standard reduced log-likelihood (ΔrLLH) for the 100 GeV hard neutralino model. On the bottom, the theta angle reconstructed by the standard log-likelihood (θ_{LLH}) for the experimental data. Different line styles correspond to different years (see the legend in both pictures).*

likelihood (ΔrLLH) for the 100 GeV hard neutralino model, while on the bottom of the same figure the theta angle reconstructed by the standard log-likelihood (θ_{LLH}) is shown for the experimental data. The ΔrLLH distribution shows the difference between the 2001 and the rest of the data, due to the string trigger downscaling issue, while the θ_{LLH} plot, shows the difference between $2005-2006$ experimental data and the rest of the data sample, due to the different low level filters.

We decided then to optimise three different BDTs, one for each above mentioned subsample of data; i.e. the first BDT optimisation contained only BDT optimisation data and MC from 2001, the second only from $2002-2004$, and the third only from $2005-2006$.

Once these three distinct training phases were finished, the next step was then the *BDT weight application* phase, using the training outcome[1]. The analysis data and signal MC analysis samples, along with the simulated atmospheric background, naturally grouped in the three different subsamples, were then processed to evaluate the BDT performance. As an ending result, this step provided to each event a weight, between [-1,+1], being +1 (-1) for pure signal (background). A plot of the BDT outputs, for experimental data (filled grey area), signal MC (dashed line), atmospheric muons (solid line) and neutrinos (dotted line), is shown as example in fig. 5.19. The atmospheric background (atmospheric muons and neutrinos) BDT output distribution is normalised to the live-time; the signal MC (3000 GeV hard) distribution is also scaled to the live-time.

Figure-of-merit

The events of the BDT optimisation sample, used to train the BDTs, were picked up randomly by the algorithm; actually we did not use the complete experimental data sample, but just part of it, since this sample contained a lot of statistics. On the signal side, we used the 25% of the MC BDT optimisation signal sample to train the BDTs.

The final experimental and MC samples, from which we calculated the sensitivity to signal-induced muons, were then selected by applying a cut on their BDT output value. The best sensitivity can be obtained by choosing a good cut, which will reduce misreconstructed background muons, thus increasing the signal resolution. We have used as figure-of-merit the *Model Discovery Potential* (MDP) as defined in [180]

$$\mathrm{MDP} = \frac{\epsilon_S}{\frac{a^2}{8} + \frac{9b^2}{13} + a\sqrt{n_B} + \frac{b}{2}\sqrt{b^2 + 4a\sqrt{n_B} + 4n_B}} \tag{5.25}$$

[1] We remind here that no particular differences were noticed between the analysis and BDT optimisation experimental data sample (see sec. 4.1.4)

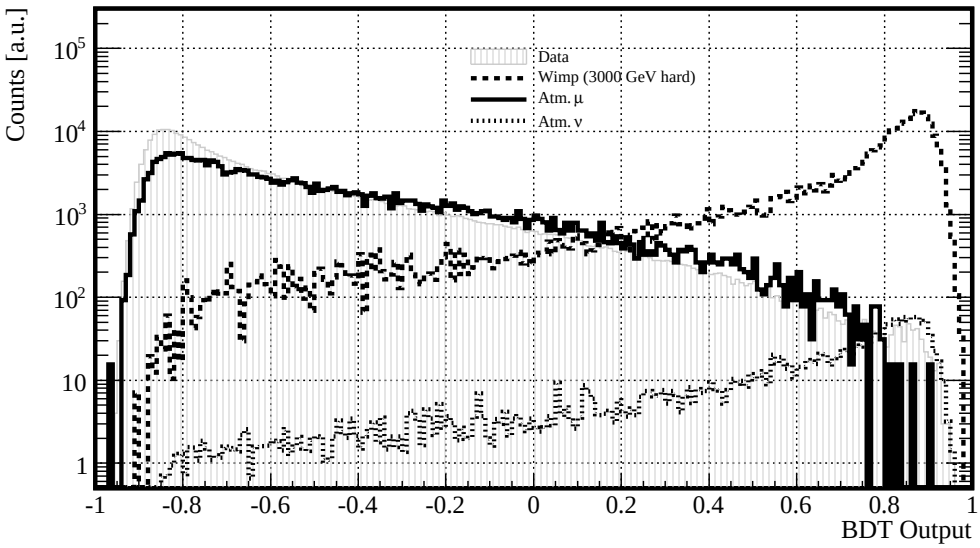


Figure 5.19: *BDT output distributions for experimental data (filled grey area), atmospheric muons (solid line), atmospheric neutrinos (dotted line), and 3000 GeV neutralino hard signal (dashed line).*

which can be maximised to find the best cut, and for this calculation we used the whole BDT optimisation sample. In the above equation ϵ_S is the signal efficiency and n_B the number of background events. This figure-of-merit was meant to minimise the strength of the signal flux needed for a discovery with 5σ significance ($a = 5$) at 90% of C.L. ($b = 1.28$). The above figure-of-merit contains also a suitable feature which does not require the absolute level of the signal flux to optimise the analysis.

In the Gaussian regime this MDP is equivalent to the "standard" significance, i.e. $S/\sqrt{B}$; but for small backgrounds, the MDP does not show asymptotic behaviour, and hence is performing better.

We have to point out that after the BDT cut we have to check also the signal resolution, because a hard cut could degrade it, and have as side-effect the worsening of the sensitivity.

Figure 5.20 is an example of what we have discussed above, where the signal (500 GeV hard) and the background (in this case the 2002−2004 data subsample), are drawn, as a function of the BDT cut value, in dashed and dashed-dotted lines respectively. The figure-of-merit, or sensitivity factor, is also shown as function of the BDT cut value (as a solid line), and the value corresponding to the peak, is then selected as the best cut value.

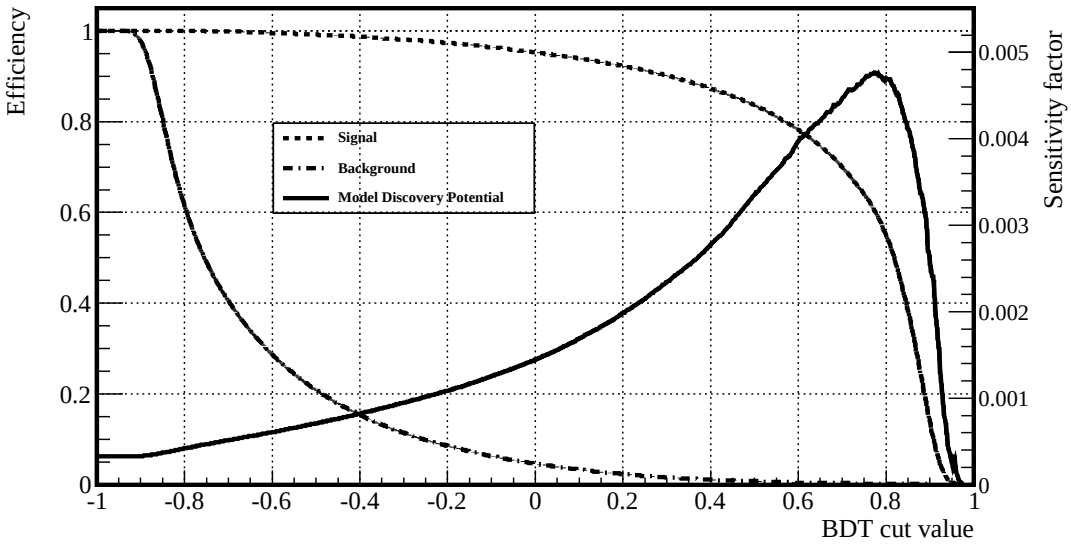


Figure 5.20: *Signal (dashed line) and background (dashed-dotted line) efficiencies as a function of the BDT cut value. The best cut is found in correspondence of the MDP (solid line) peak.*

5.6 Final sample

The final experimental and Monte Carlo analysis data samples were then selected after applying the last event selection, i.e. the BDT cuts. We remind that in total we optimised 14 × 3 cuts, i.e. a cut for each of the fourteen signal models distinct from the three different subsamples.

Hence, considering the subsamples: 2001, 2002 − 2004, and 2005 − 2006, we ended up with 14 different final samples, for the complete 2001 − 2006 dataset, considering each neutralino model.

The numbers concerning the experimental data and atmospheric background events, in the different final samples, are shown in tab. 5.15. We infer then, from the numbers in this table, that the simulated background expectation, even considering the statistical errors, are not consistent with the observed number of events. We will discuss about this issue in the next chapter, in section 6.2, devoted to the treatment of statistical and systematics uncertainties in our analysis. This impinged also on the distributions of the observables after the BDT cuts, which are found in sec. A.3.

The amount of experimental data events surviving the last selection, varied between one to two thousand, depending on the selected neutralino model; this corresponded to a data reduction of $\sim 1 - 3 \cdot 10^{-7}$. We can also conclude that the

Neutralino model selection	exp. data	atm. $\mu + \nu$	atm. μ	atm. ν
50 soft	1521	4064.6 ± 574.7	3880.4 ± 573.0	184.2 ± 44.3
100 soft	1662	3918.7 ± 520.9	3023.2 ± 517.3	895.5 ± 61.6
250 soft	1572	3717.3 ± 371.7	1754.7 ± 363.8	1962.6 ± 76.2
500 soft	2031	3969.8 ± 340.8	1288.7 ± 330.6	2681.2 ± 82.8
1000 soft	2256	4615.0 ± 364.3	1460.2 ± 354.0	3154.8 ± 86.2
3000 soft	2254	4280.7 ± 315.5	1127.1 ± 303.9	3153.6 ±85.0
5000 soft	2024	3973.4 ± 298.5	978.2 ± 287.2	2995.2 ±81.4
50 hard	828	2206.9 ± 356.3	1640.4 ± 352.5	566.5 ± 52.1
100 hard	1253	2791.6 ± 308.7	1039.3 ± 300.6	1752.3 ± 70.5
250 hard	2136	4247.6 ± 300.6	1026.8 ± 288.5	3220.7 ± 84.5
500 hard	2468	4674.3 ± 301.8	1075.5 ± 288.9	3598.8 ± 87.0
1000 hard	2146	4155.7 ± 272.1	802.3 ± 259.2	3353.4 ± 83.1
3000 hard	2501	4865.9 ± 332.6	1217.5 ± 320.7	3648.3 ± 86.9
5000 hard	2369	4292.0 ± 272.3	785.9 ± 258.8	3506.1 ±84.8

Table 5.15: *Number of events remaining after the BDT cut in the experimental data, the total simulated background (atm. μ + atm. ν), the simulated atm. μ and the simulated atm. ν.*

final samples, related to the lower energy neutralino models (50 GeV soft, 100 GeV soft and 50 GeV hard), are mostly dominated by atmospheric muon background. Apart from an intermediate neutralino model, 250 GeV soft, whose atmospheric background contribution was equally shared between muons and neutrinos, the rest of the samples are mostly dominated by atmospheric neutrino background.

The effective volume for each neutralino model after the BDT selection is drawn in fig. 5.21, and in fig. 5.22 the effective volume is shown again (black colour for hard channels, and grey for soft channels), where different line styles represent the different six years of simulated data. The statistical errors on the effective volume are not clearly visible in some lines of the pictures, covered by the size of the same lines; however they are between 1 - 2 %. The numbers related to the effective volume are then arranged in tab. 5.16. From these numbers we infer that, after the last event selection the string trigger still played a key role for the lowest neutralino mass models, while the standard was confirmed as the most efficient trigger for higher neutralino masses.

In tab. 5.17 are summarised the selection efficiencies of the BDT with respect to the precuts, for the experimental data, the simulated atmospheric background and simulated signal, broken up for the different neutralino model analyses.
Hence, we infer that the good performance of the BDT in background rejection

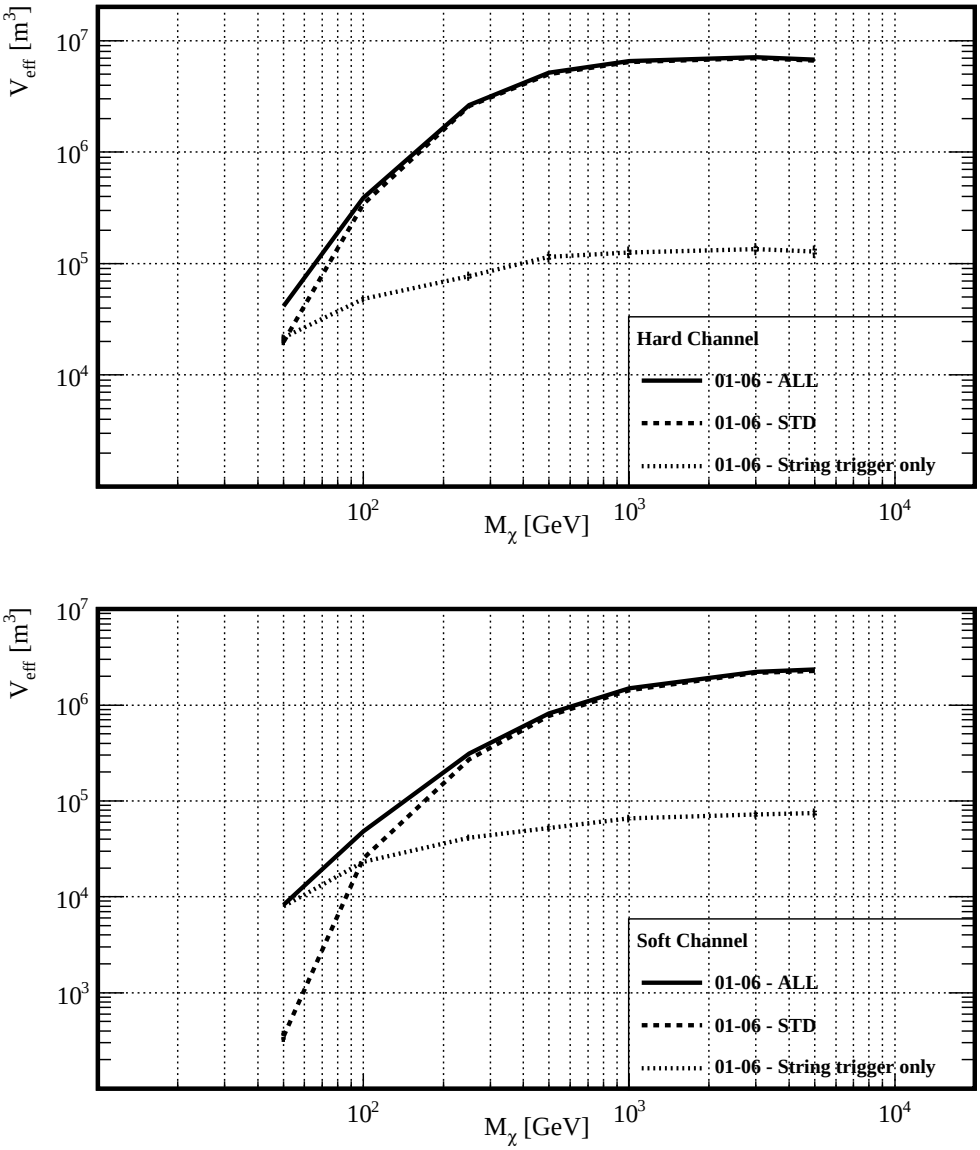


Figure 5.21: *Effective volume after BDT selection as a function of the neutralino mass for hard (top) and soft (bottom) annihilation channel. The dotted line refers to events triggered only by the string trigger, while the dashed one to events triggered by the standard trigger. The solid line is the logical sum of the two trigger selections. In some lines the statistical errors are not visible, but covered by the size of the lines.*

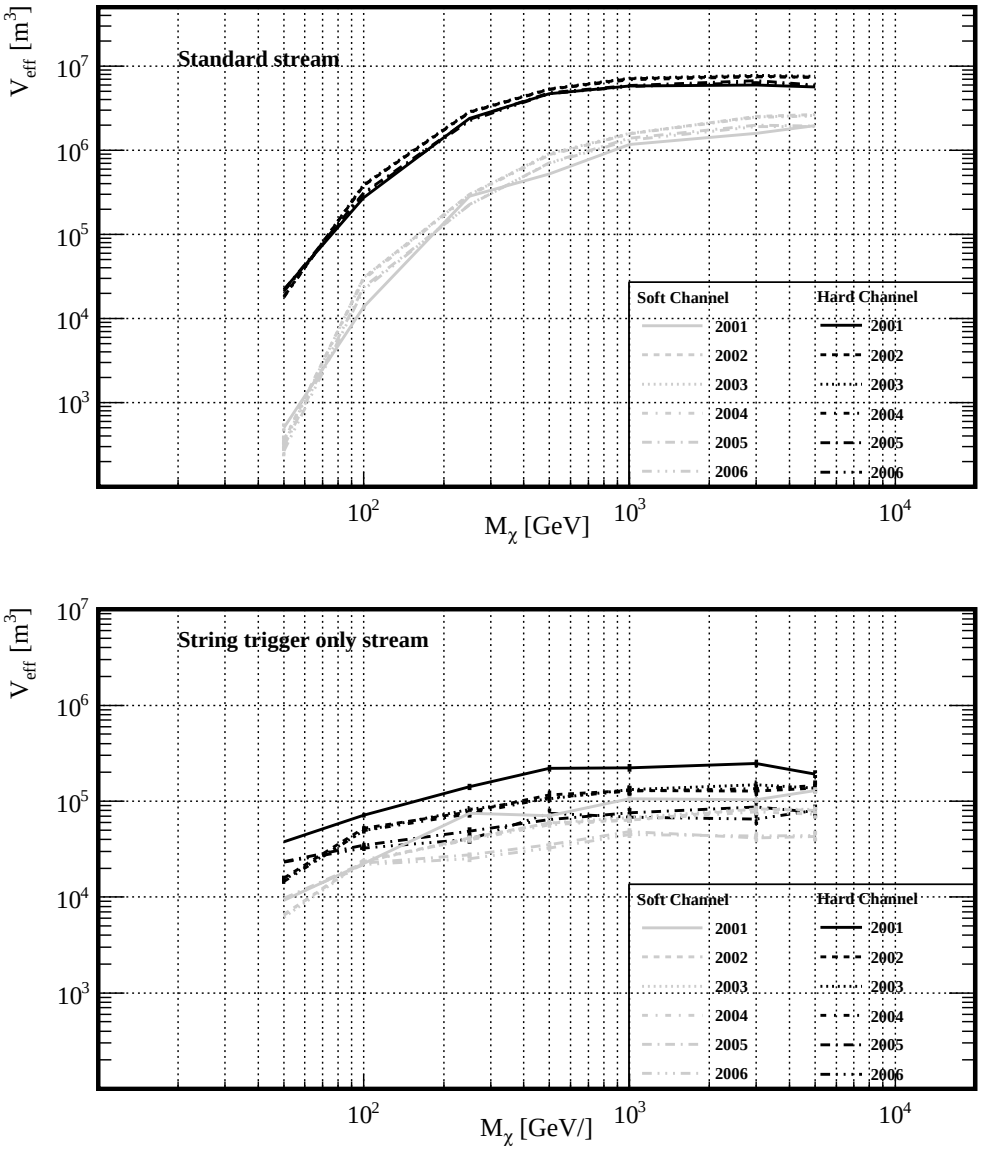


Figure 5.22: *As in fig. 5.21, the effective volumes for each neutralino model after BDT selection are shown, but split up per data-taking year; top (bottom) plot is the standard (string only) trigger stream, black (grey) lines stand for the hard (soft) channel. Also here, in some lines the statistical errors are not visible, but covered by the size of the lines*

$V_{eff(BDT)}$	[m^3]	50	100	250	500	1000	3000	5000
ALL	hard	$4.15 \cdot 10^4$	$3.88 \cdot 10^5$	$2.66 \cdot 10^6$	$5.14 \cdot 10^6$	$6.58 \cdot 10^6$	$7.12 \cdot 10^6$	$6.77 \cdot 10^6$
	soft	$8.29 \cdot 10^3$	$4.81 \cdot 10^4$	$3.12 \cdot 10^5$	$8.21 \cdot 10^5$	$1.50 \cdot 10^6$	$2.23 \cdot 10^6$	$2.35 \cdot 10^6$
STD	hard	$1.99 \cdot 10^4$	$3.40 \cdot 10^5$	$2.58 \cdot 10^6$	$5.03 \cdot 10^6$	$6.45 \cdot 10^6$	$6.99 \cdot 10^6$	$6.64 \cdot 10^6$
	soft	$0.35 \cdot 10^3$	$2.50 \cdot 10^4$	$2.71 \cdot 10^5$	$7.69 \cdot 10^5$	$1.43 \cdot 10^6$	$2.16 \cdot 10^6$	$2.28 \cdot 10^6$
STR	hard	$2.16 \cdot 10^4$	$0.48 \cdot 10^5$	$0.08 \cdot 10^6$	$0.11 \cdot 10^6$	$0.13 \cdot 10^6$	$0.13 \cdot 10^6$	$0.13 \cdot 10^6$
	soft	$7.94 \cdot 10^3$	$2.31 \cdot 10^4$	$0.41 \cdot 10^5$	$0.52 \cdot 10^5$	$0.07 \cdot 10^6$	$0.07 \cdot 10^6$	$0.07 \cdot 10^6$

Table 5.16: *Effective volumes at final level (after BDTs) for all the neutralino models (mass and channel), divided up for different trigger selections: STR, STD, ALL.*

had as drawback a deficit in signal efficiency, with respect to trigger level, numbering to a factor 3 to 16.

We will describe in the next chapter a statistical method to evaluate the signal events content, μ_s, from the experimental dataset. Then if we want to convert this evaluation to a muon flux, we need the effective volume V_{eff} after the last event selection, which multiplied by the live-time (t_{live}) is nothing other than the detector exposure. Indeed, the main physical quantity which we indirectly measured, is the number of signal induced muons per unit volume and time

$$\Gamma_{\nu \to \mu} = \frac{\mu_s}{V_{eff} \cdot t_{live}} \tag{5.26}$$

Passing rate BDT w.r.t. precuts [$\varepsilon_{\mathrm{BDT}}$]

Neutralino model selection	exp. data ($\times 10^{-4}$)	atm. μ ($\times 10^{-4}$)	atm. ν	χ
50 soft	7.89	22.30	0.02	0.34
100 soft	8.63	17.37	0.09	0.39
250 soft	8.16	10.08	0.21	0.43
500 soft	10.54	7.41	0.28	0.49
1000 soft	11.71	8.39	0.33	0.54
3000 soft	11.70	6.48	0.33	0.53
5000 soft	10.50	5.62	0.32	0.51
50 hard	4.30	9.43	0.06	0.31
100 hard	6.50	5.97	0.18	0.42
250 hard	11.09	5.90	0.34	0.54
500 hard	12.81	6.18	0.38	0.59
1000 hard	11.14	4.61	0.35	0.56
3000 hard	12.98	7.00	0.39	0.59
5000 hard	12.29	4.52	0.37	0.58

Table 5.17: *The BDT selection efficiencies w.r.t. the precuts, for the experimental data, the atm. μ, the atm. ν and the neutralino signals (χ).*

- Σωκράτης·
... [152a] ... "πάντων χρημάτων μέτρον" ἄνχρωπον εἶναι, "τῶν μὲν ὄντων ὡς ἔστι, τῶν δὲ μὴ ὄντων ὡς οὐκ ἔστιν."...
- Socrates:
... [152a] ... man is "the measure of all things, of the existence of the things that are and the non-existence of the things that are not." ...

Theætetus
Plato (428/427 bc - 348/347 bc)

6 Signal strength evaluation and inference of some relevant physical quantities

In this chapter we describe a consolidated method, to extract the signal strength from the combined 2001-2006 final sample, which passed all the filtering steps. Next, we investigate all the possible sources of systematic uncertainties which could affect our final results. Then, through our signal estimation, we can infer some relevant physical quantities like the neutrino-to-muon conversion rate, the neutralino annihilation rate in the Sun, the muon flux at the detector, and the elastic neutralino-proton cross section.

6.1 Hypothesis testing

6.1.1 The space angle to the Sun

Well, the source of our hypothetical signal is our Star, the Sun. Hence, to evaluate the number of signal events μ_s from the dataset[1] we could exploit the *space angle* ψ between the source direction, identified by the spherical coordinates $(\theta_\odot, \phi_\odot)$, and that of the reconstructed track, in its turn identified by (θ_μ, ϕ_μ), which is constructed as

$$\psi \equiv \arccos\left(\cos\theta_\mu \cos\theta_\odot + \sin\theta_\mu \sin\theta_\odot \cos(\phi_\mu - \phi_\odot)\right), \tag{6.1}$$

A simple sketch showing the construction of the space angle is outlined in fig. 6.1.

The size of the space angle depends upon two main factors; one of this is the reconstruction method, which has difficulty to convey the true track direction even for high quality events. The reconstructed track is indeed smeared around its true direction, especially in the case of low energy events. The second physical unavoidable factor is the CC scattering angle between the neutrino, which points

[1] We remind here that the analysis MC and experimental data sample were used to derive the final results discussed in this chapter.

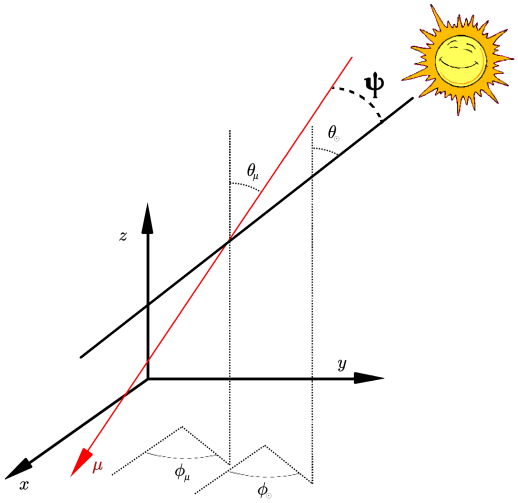


Figure 6.1: *Sketch of the construction of the space angle ψ, between the Sun direction and the reconstructed track.*

directly to the Sun direction, and the ongoing muon; the span of this angle increases with decreasing neutrino energy (see sec. 2.3).

The space angle distribution for two neutralino models, 100 GeV hard and 5000 GeV hard, is shown respectively on the top and on the bottom of fig. 6.2 in dashed line. These distributions have a peak at very low angles, i.e. close to the Sun, and fall rapidly with increasing angle. The 50% quantile, or otherwise noted as the median, of such distributions, can be calculated to estimate the angular resolution. The value of the median for 5000 GeV hard is around 3°, while for 100 GeV it is around 4.5°; for the lowest energy model a cone with an opening angle of about 12° is needed to contain half of the signal events.

In fig. 6.2 the space angle distribution of the background (solid line) is also shown, which was extracted from off-source 2004 experimental data; in sec. 6.1.2 we will discuss about this procedure. Observing the picture, we see that the background distributions for both neutralino model selections are quite similar, and it is more or less the same for the rest of the models. Further, since there is no correlation with the Sun position, these distributions are rather smooth between $20° - 140°$, falling outside these regions because of lack of phase space.

In the so-called *cut-and-count* approach, a model-dependent cone of opening angle ψ_{cone}, where to count the number of hypothetical signal and expected background events, is needed to get the best sensitivity. Thus, some selection methods, like the MDP (see sec. 5.5.5.2) or the *Model Rejection Potential* (MRP) [181], could be exploited to find the optimal opening angle.

However, following this approach, the selected events in the cone are indistinct,

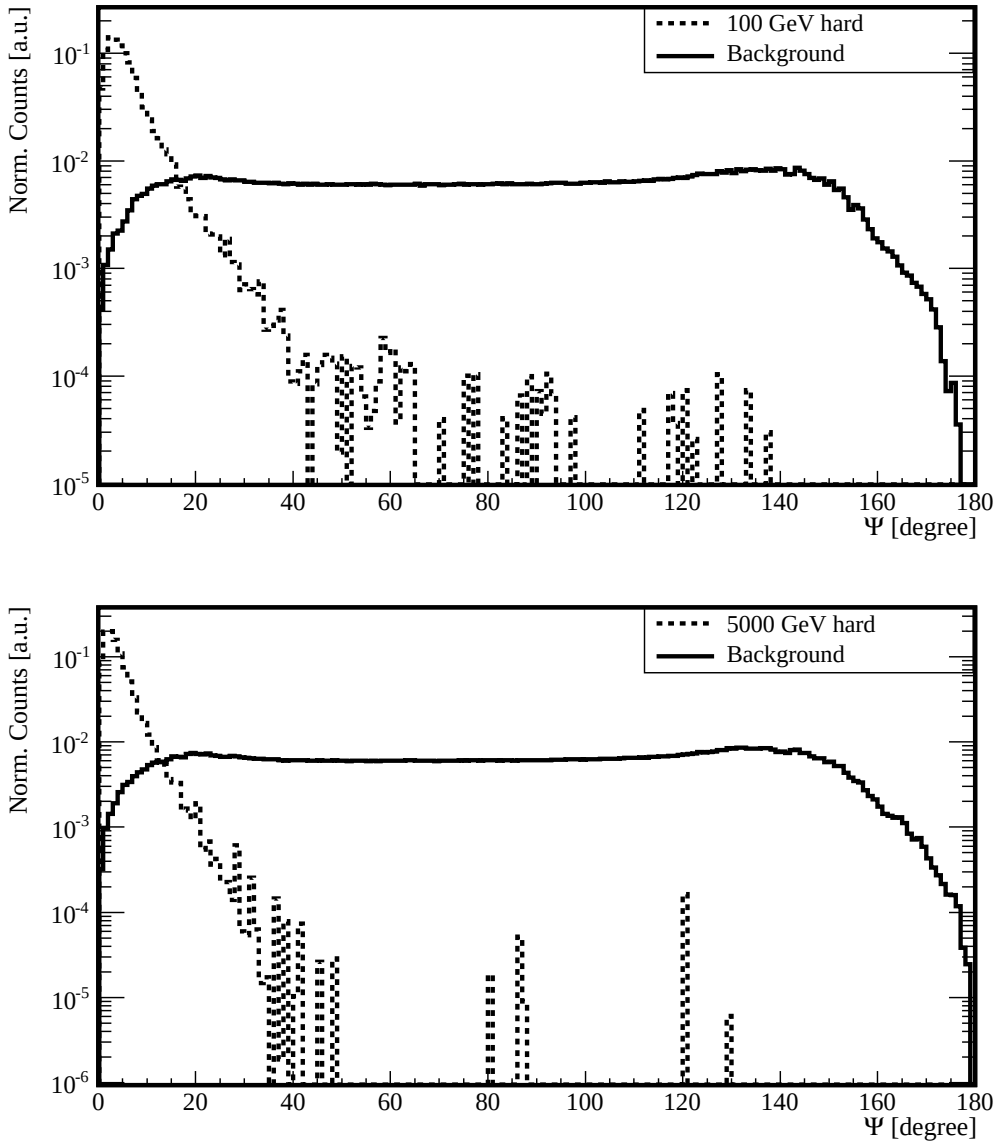


Figure 6.2: *Normalised space angle distribution for the 100 GeV hard (top) and 5000 GeV hard (bottom) models. In both pictures the background refers to experimental data of 2004 (see text).*

while the ones outside it are totally neglected. Hence, if repeated experiments are performed (assuming identical background), and e.g. 10 events within a certain search cone are observed, which are compatible with the number of background expectation, the same confidence interval will be quoted as in another experiment which measures again 10 events, but with 8 of them really close to the Sun direction. The latter case should give a strong hint for the existence of a signal.

Hence, since the above approach could neglect some useful information, we decided to follow another more sophisticated approach. In this method we exploited the shape of the complete space angle distribution, fitting the sum of the known background and signal distributions, with the signal strength as free parameter. The confidence interval was then calculated with a likelihood-ratio test statistic, whose procedure is explained in sec. 6.1.3.

We should mention that the method does not necessarily require the space angle as discriminating distribution. Other distributions can be used as well; however we should expect that the space angle provides the best separation between signal and background, thus improving our sensitivity.

As a conclusive remark, we have to bear in mind the blindness policy during the setting of any of the discussed approaches.

6.1.2 Combining the 2001-2006 experimental data: space angle p.d.f.

The signal space angle distribution, $f_S(\psi)$, has been extracted directly from Monte Carlo signal simulation, for each neutralino model. The discussion about uncertainty sources in the simulation which could affect this distribution and the confidence interval in μ_s, is outlined in sec. 6.2.

The background space angle distribution, $f_B(\psi)$, has been extracted instead by scrambling off-source experimental data. This approach presented some advantages with respect to exploiting simulated atmospheric background; first, because $f_B(\psi)$, being independent of the simulation, was almost systematic uncertainties free. Another advantage was the larger statistics available for the calculations.

During this process the blind policy was carefully fulfilled; indeed working with off-source experimental data meant to take into account only events with reconstructed phi outside the range $[-30°, +30°]$, with respect to the Sun azimuth[1]. This avoided that possible information from signal could invalidate the background pattern.

Further, the space angle of each off-source event was calculated with respect to a fake position of the Sun, identified by $(\theta_{\odot}^{\text{fake}}, \phi_{\odot}^{\text{fake}})$; this is the scrambling procedure. The detector has a theta-dependent acceptance, hence we should keep

[1] About 15% of the events were excluded by this selection.

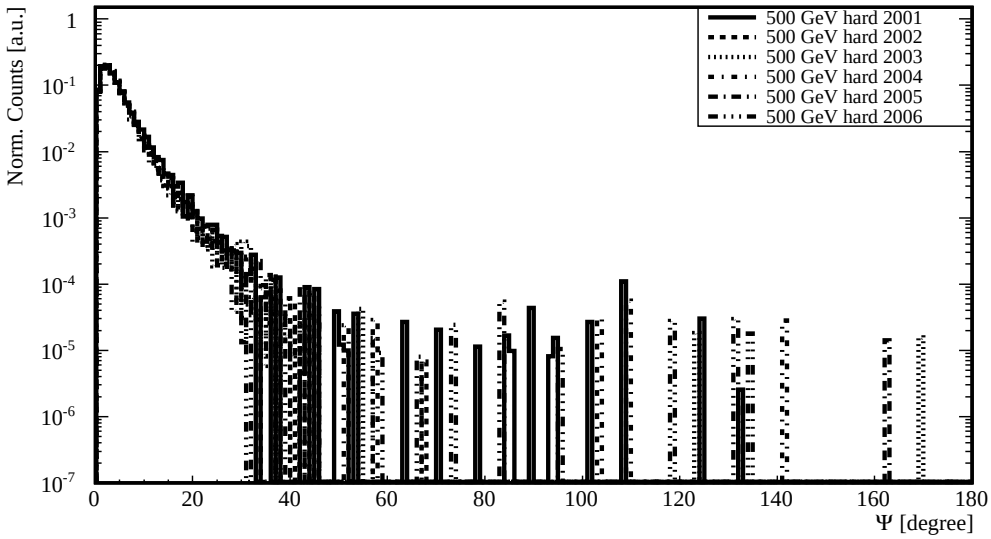


Figure 6.3: *Space angle distribution for neutrinos from the 500 GeV hard neutralino model (different line styles represent different years).*

$\theta_{\odot}^{\text{fake}} = \theta_{\odot}^{\text{true}}$, and scramble only $\phi_{\odot}^{\text{true}}$ to get $\phi_{\odot}^{\text{fake}}$. The space angle background distribution was then constructed by collecting random $\phi_{\odot}^{\text{fake}}$, uniformly extracted between $[0°, 360°]$, and calculating the corresponding space angle between the reconstructed track and the fake Sun direction. To obtain a continuous and smooth space angle distribution without any intermediate empty bins, the procedure was repeated 10000 times for each off-source data event.

The resulting binned[1] $f_S(\psi)$ and $f_B(\psi)$ distributions were normalised to one; in this way they just represented probability density functions (p.d.f.) for signal and background (see e.g. fig. 6.2).

Now, if we want to perform a combined hypothesis test for each neutralino model, which would encompass the complete dataset from 2001 to 2006, we should look at the distributions of the space angle of the signal and the corresponding background of all the years. Figure 6.3 shows the space angle distributions for the 500 GeV hard signal of each year (different line styles cover different years). Figure 6.4 shows the space angle distribution of the background corresponding to the selected neutralino model; also there different line styles cover different years. We have performed such investigation for all the different models, and since the various year dependent signal distributions were in good agreement each other (see indeed fig. 6.3), we summed all the distributions year after year, after rescaling them to the corresponding detector live-time. The same conclusion

[1] We chose 1° as bin width.

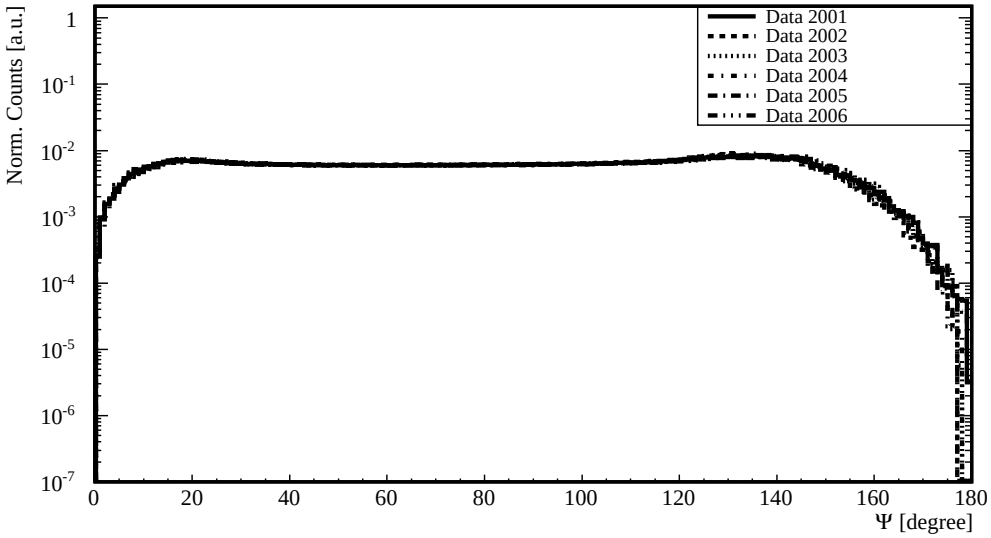


Figure 6.4: *Space angle distribution for the background obtained for the 500 GeV hard neutralino model selection, (different line styles represent different years).*

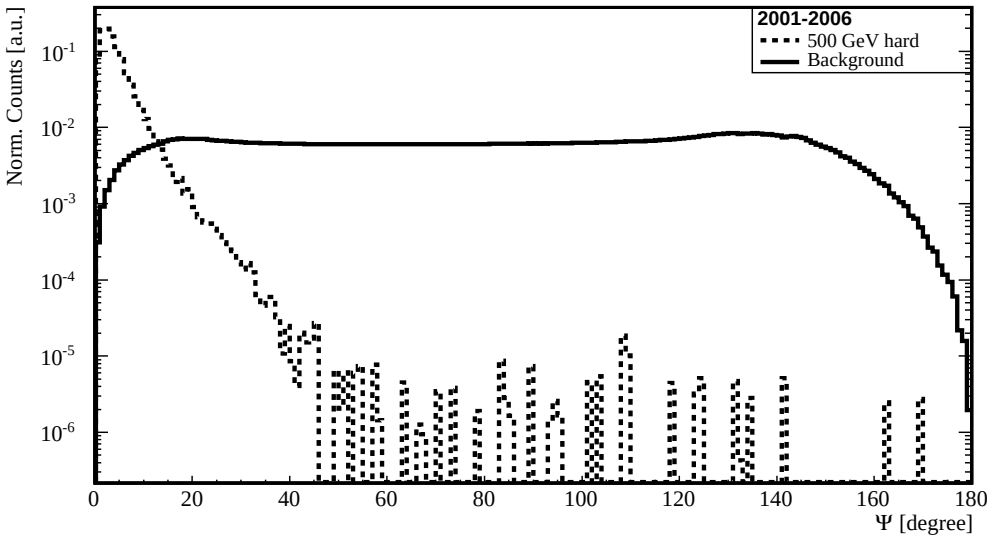


Figure 6.5: *Sum of the signal space angle distributions of fig. 6.3 (dashed line), and the background space angle distributions of fig. 6.4 (solid line).*

	$\widetilde{\psi}$ [degree]						
Neutralino model [GeV]	50	100	250	500	1000	3000	5000
Hard channel	6.8	4.6	3.5	3.2	3.0	3.0	3.0
Soft channel	13.1	6.8	4.4	3.9	3.6	3.4	3.4

Table 6.1: *Median of the combined space angle distributions of all neutralino models.*

we outlined before was valid also for the background (see fig. 6.4); hence we could perform only one combined 2001 – 2006 statistical analysis for each neutralino model. Figure 6.5 shows the summed distributions of the space angle for the 500 GeV hard neutralino model (dashed line) and the corresponding background (solid line).

The calculated median of the different combined signal space angle distributions is shown in tab. 6.1.

6.1.3 Confidence interval for μ_s

Well, recapitulating, in the hypothesis test procedure, the space angle (ψ) distribution is exploited to derive the amount of signal events μ_s compatible with the observed data sample at some confidence level (CL). In this work the confidence level is set to $\alpha = 90\%$. The procedure followed for the hypothesis testing is outlined in this section, together with a series of consistency checks.

Signal content likelihood

For a given signal and background space angle p.d.f., the likelihood of the presence of μ_s signal events in an experiment that observed exactly $n_{\rm obs}$ events, with an ensemble of space angles $\{\psi_i\}$, can be expressed as follows

$$\mathcal{L}(\mu_s) = \prod_{i=1}^{n_{\rm obs}} f(\psi_i|\mu_s), \tag{6.2}$$

where

$$f(\psi_i|\mu_s) = \frac{\mu_s}{n_{obs}} f_S(\psi_i) + \left(1 - \frac{\mu_s}{n_{obs}}\right) f_B(\psi_i) \tag{6.3}$$

is the probability to observe an event with space angle ψ_i when μ_s signal events are present among the total number of observed events $n_{\rm obs}$.

Frequentist construction of confidence intervals

To draw confidence intervals (CI) we will follow the Feldman and Cousins (FC) frequentist approach [182], which suggested this likelihood-ratio test statistic

$$\mathcal{R}(\mu_s) = \frac{\mathcal{L}(\mu_s)}{\mathcal{L}(\widehat{\mu}_s)} \tag{6.4}$$

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

The ordering principle suggested by FC in a frequentist scenario, is based on their test statistic $\mathcal{R}$; this latter is designed as *rank*, since, starting at high values, it ranks the experimental results in order of inclusion in the acceptance interval. Further, it is also a unifying ordering principle, in the sense that the reported CI are determined as one- or two-sided by the experimental data themselves and not by the experimenter. The coverage, essential feature in this framework, is preserved at the same time. A basic ingredient of the FC recipe is the binding of the parameter μ_s to meaningful physical values; hence μ_s, $\widehat{\mu}_s \in [0, n_{\rm obs}]$. Non-physical best fits ($\widehat{\mu}_s < 0$) could indeed arise, for the absence, or for very little amount of signal and for downward fluctuation of the background near the Sun.

Construction of the critical region by means of pseudo-experiments

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

$$[\mathcal{R}^{\alpha}_{\rm crit}(\mu), 1]$$

where $\mathcal{R}^{\alpha}_{\rm crit}(\mu_s)$ is a function of μ_s and it is referred to as the *critical region*. According to Wilks' Theorem [183], in a Gaussian scenario ($n_{\rm obs} \to \infty$), $-2\ln\mathcal{R}$ becomes χ^2-distributed, with in our case one degree of freedom. Hence, we can write

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

However, the actual distribution of $-2\ln\mathcal{R}$ may notably deviate from a χ^2 distribution, for instance near physical boundary regions. Hence, with μ_s too close to the physical boundary, many experiments will quote higher rank than would be attained without the restriction to the physical region. So, the critical value is higher than that of a χ^2 distribution, leading then to over-coverage with respect to the fixed CL.

To guarantee the correct coverage for all the considered μ_s, reiterated pseudo-experiments were performed in order to get, for each μ_s, the $\ln\mathcal{R}$ distribution and its critical value. This procedure (very CPU-time consuming) will be sketched on the next page, and is exactly the same as the one described in [111].

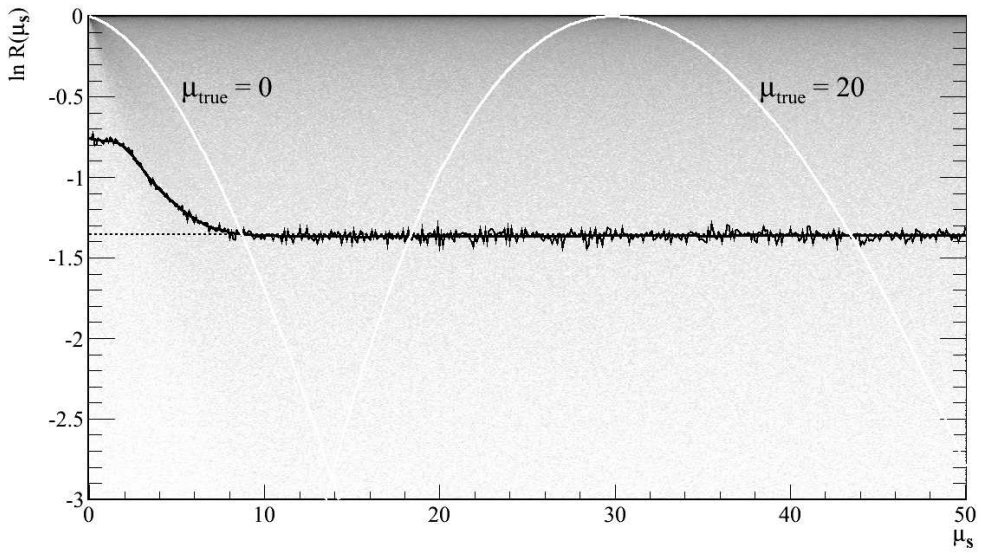


Figure 6.6: *The distribution of* $\ln \mathcal{R}(\mu_s)$ *of 10000 pseudo-experiments is displayed in grey scale. The critical region* $\ln \mathcal{R}_{\text{crit}}(\mu_)$ *at 90% CL is marked by the thin black curve, with its smoothed fit marked by the black thick line. The horizontal black dashed line shows the critical value from the* χ^2 *approximation. The white parabolas represent* $\ln \mathcal{R}(\mu_s)$ *for two pseudo-experiments with* $\mu_{s(\text{true})} = 0$ *(left) and* $\mu_{s(\text{true})} = 20$ *(right).*

1. for each $\mu_s \in [0, 50]$, with step-size $\Delta\mu_s = 0.1$
   (a) for each pseudo-experiment $k = 1, ..., 10000$
      i. given μ_s, sample a set $\{\psi_i\}_k$ with n_{obs} space angles from eq. 6.3
      ii. calculate $\mathcal{L}_k(\mu_s)$ with eq. 6.2
      iii. find $\widehat{\mu}_{s_k}$ with maximum likelihood $\mathcal{L}_k(\widehat{\mu}_{s_k})$
      iv. calculate $\ln \mathcal{R}_k(\mu_s)$
   (b) find the critical value $\ln \mathcal{R}^{\alpha}_{\text{crit}}(\mu_s)$ such that $\ln \mathcal{R}_k(\mu_s) \geq \ln \mathcal{R}^{\alpha}_{\text{crit}}(\mu_s)$ for a fraction α of the 10000 experiments
2. smooth $\ln \mathcal{R}^{\alpha}_{\text{crit}}(\mu_s)$ and fit it with a spline

An example of the outcome of the above mentioned procedure is presented in fig. 6.6, where a two-dimensional distribution of $\ln \mathcal{R}(\mu_s)$ versus μ_s is shown. The black thick line marks the superimposed fit of the critical region (the thin black curve) at 90% CL; the χ^2 approximation (horizontal dashed black line) is valid for $\mu_s > 8$ where $\ln \mathcal{R}(\mu_s) = -0.5 \cdot \chi^2(0.9, 1) = -1.35$.

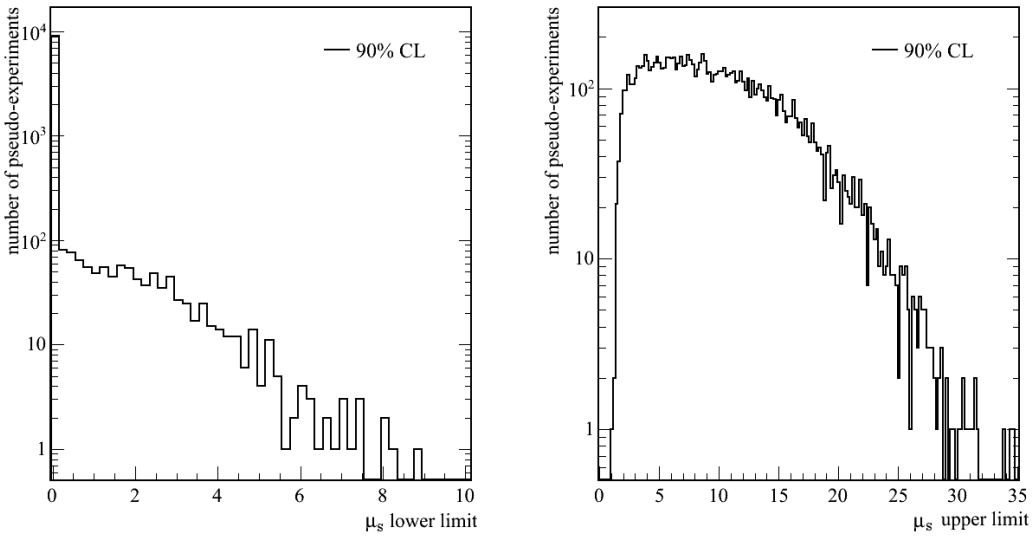


Figure 6.7: *Distribution of lower (left) and upper (right) limits of the confidence intervals at 90% CL, for 10000 pseudo-experiments in the background-only scenario.*

Properties of confidence intervals

Once the critical region is defined, the confidence interval $[\mu^{\alpha}_{\mathrm{s(low)}}, \mu^{\alpha}_{\mathrm{s(up)}}]$ at α CL, for the number of signal events in an observed ensemble of space angles $\{\psi_i\}$, is given by

$$[\mu^{\alpha}_{\mathrm{s(low)}}, \mu^{\alpha}_{\mathrm{s(up)}}] = \{\mu_s | \ln \mathcal{R}(\mu_s) \geq \ln \mathcal{R}^{\alpha}_{\mathrm{crit}}(\mu_s)\} \tag{6.5}$$

Hence, if the rank for the observed μ_s is above the critical value, that μ_s is accepted in the confidence interval.

In fig. 6.6 the $\ln \mathcal{R}(\mu_s)$ for two pseudo-experiments with $\mu_{\mathrm{s(true)}} = 0$ and $\mu_{\mathrm{s(true)}} = 20$ (the two white curves) is also shown. The best physical fit to the generated sets of space angles is found respectively at $\widehat{\mu}_s = 0.0$ and $\widehat{\mu}_s = 29.8$; the first pseudo-experiment produces a one-sided interval with upper limit 8.69 at 90% CL, the second pseudo-experiment produces a two side interval: $[18.3, 43.5]$ at the same CL.

We have already pointed out that in the frequentist approach the coverage of the confidence interval is an essential property. For each pseudo-experiment, with e.g. $\mu_{\mathrm{s(true)}} = 0$, the confidence interval was calculated; then we checked whether $\mu_{\mathrm{s(true)}} = 0$ was inside the interval before storing its lower and upper boundaries. The fraction of covering intervals was calculated and compared with the quoted 90% CL. We calculated then the 50%, 16% and 84% quantiles of

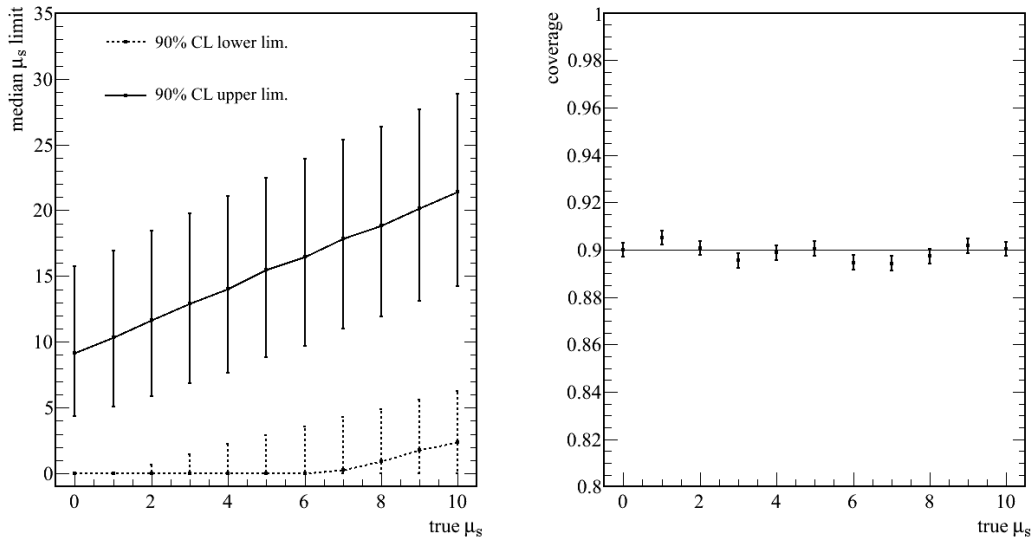


Figure 6.8: *The median interval (left) and coverage (right) of the intervals as a function of* $\mu_{s(true)}$. *Vertical bars represent the* 1σ *spread.*

the lower (upper) limits distribution which represented the median lower (upper) limit $\widetilde{\mu}^{90}_{s(low)}$ ($\widetilde{\mu}^{90}_{s(up)}$) and its 1σ statistical spread.

Figure 6.7 shows the distribution of the lower (left) and upper (right) limits for the background-only scenario ($\mu_{s(true)} = 0$), where a small amount of the pseudo-experiments report lower limits greater than zero. If the coverage is correct, the true value should not be contained in a fraction $1 - 0.9 = 0.1$ of the CI of the pseudo-experiments.

Figure 6.8 shows (on the right) the resultant coverage of the CI for $\mu_{s(true)} \in [0, 10]$, and since no statistical significant deviation from 0.9 was found, we deduce that the intervals were indeed confidence intervals.

Figure 6.8 shows as well (on the left) the median CI and the 1σ spread on its boundaries. The picture is a nice illustration of the FC unified ordering principle which, for strong signal fluxes ($\mu_s > 6$), yields automatically double-sided confidence intervals.

Median upper limits on μ_s and $\Gamma_{\nu\to\mu}$

The median upper limits at 90% CL on μ_s with their 1σ spread, for all the neutralino models, are outlined in tab. 6.2. In the last column of the same table, the median upper limits on the conversion rate (see eq. 5.26) are also shown; we will discuss later in more detail about this quantity.

Once the data are unblinded, in the case of (sad) no neutralino signal scenario,

Neutralino model	$\widetilde{\mu}_s^{90} - 1\sigma$	$\widetilde{\mu}_s^{90}$	$\widetilde{\mu}_s^{90} + 1\sigma$	$\widetilde{\Gamma}_{\nu\to\mu}^{E_\mu>10\,\mathrm{GeV}}$ [Km^{-3} y^{-1}]
50 hard	5.1	10.6	18.2	$1.14 \cdot 10^5$
100 hard	4.3	9.2	15.8	$1.07 \cdot 10^4$
250 hard	4.2	9.1	15.5	$1.55 \cdot 10^3$
500 hard	4.2	9.2	15.3	$8.04 \cdot 10^2$
1000 hard	3.8	8.3	14.1	$5.69 \cdot 10^2$
3000 hard	4.1	8.6	14.5	$5.46 \cdot 10^2$
5000 hard	4.0	8.5	14.4	$5.70 \cdot 10^2$
50 soft	12.4	25.0	41.8	$1.34 \cdot 10^6$
100 soft	6.8	14.6	24.5	$1.37 \cdot 10^5$
250 soft	4.5	9.8	16.5	$1.42 \cdot 10^4$
500 soft	4.6	9.9	16.8	$5.46 \cdot 10^3$
1000 soft	4.5	9.7	16.4	$2.93 \cdot 10^3$
3000 soft	4.3	9.2	15.7	$1.87 \cdot 10^3$
5000 soft	4.1	8.7	14.8	$1.69 \cdot 10^3$

Table 6.2: *For each neutralino model, the median upper limit on the number of signal events with the 1σ spread, and on the conversion rate at 90% CL in the background-only scenario.*

the final upper limits on μ_s, at 90% CL, are expected to be inside the 1σ band with 68% of probability.

6.2 Statistical and systematic uncertainties

As we have mentioned at the end of previous chapter (and we will return to this subject in sec. 6.3.1), the primary physical quantity we are interested to evaluate at the final stage of our analysis, is the volumetric flux or neutrino-to-muon conversion rate $\Gamma_{\nu\to\mu}$ (see eq. 5.26).

The components of this quantity $(\mu_s; V_{\mathrm{eff}}; t_{\mathrm{live}})$, are subject both to statistical and systematic uncertainties, which we will discuss in secs. 6.2.1 and 6.2.2. Actually, the third component of $\Gamma_{\nu\to\mu}$, i.e. t_{live}, is only subject to statistical uncertainty, and in the following section we will show how it is almost negligible if compared with the ones quoted for the remaining quantities. Hence, in sec. 6.2.3 we will consider only these latter and explain how to yield a total uncertainty for each of them.

Finally, in sec. 6.2.4, we will report how the total uncertainty on μ_s and V_{eff} propagates to the physical quantities described in sec. 6.3, like the neutrino-to-muon conversion rate $\Gamma_{\nu\to\mu}$.

6.2.1 Statistical uncertainties

Number of signal event μ_s

The signal and background p.d.f. were extracted from a finite set of events, therefore subject to statistical fluctuations. However, the fluctuations in the signal and background distributions are expected to be small.

The statistical error on μ_s was evaluated by constructing ten times the confidence interval, changing each time the seed of the random number generator. The relative error is then calculated as follows:

$$\frac{\Delta\mu_s}{\mu_s} = \frac{\sigma_{\mu_s}}{\overline{\mu}_s}$$

where σ_{μ_s} is the sample standard deviation and $\overline{\mu}_s$ is the average of the outcomes. This study yielded a relative error of less than $\pm 1\%$.

Effective Volume

The statistical error on V_{eff} depended on the size of the neutralino Monte Carlo sample. The event weight was properly considered in the calculations. The resulting error for higher energy neutralino models was less than $\pm 1\%$, while for the lowest energy models it was less than $\pm 2\%$, since their relative sample was smaller at the final level. Tables 6.3 and 6.4, report the statistical errors on V_{eff} for all the hard and soft neutralino models respectively.

Live-time

In sec. 4.1.2 we have pointed out that the mean statistical error relative to the live-time extracted from each experimental data file was less than 0.1%. Then, considering in total more than $N \simeq 1.5 \cdot 10^5$ experimental data files, the error on the total live-time of the data sample was: $\Delta T_{\text{live}}/T_{\text{live}} < 3 \cdot 10^{-3}$. Hence, the statistical error on the detector live-time was negligible if compared with the one on μ_s and V_{eff}.

6.2.2 Systematic uncertainties

Prominent sources carrying systematic errors are linked, for instance, to the cross section measurement and parametrisation, or to the lack of robust simulation that could efficiently take into account detector calibrations (or other hardware issues), or particle propagation through the ice.

To try to disentangle the effects of systematics due to the detector itself from those related to the surrounding medium, we should make a challenging effort

looking at each systematic source separately, assuming a null correlation between the different sources.

The results we are going to show in this section are related to the systematic errors on the effective volume after the BDT selection, i.e. after the final selection. We assume also that the systematic effects on the CI in μ_s should be quite small; we have to mention that the correct way to work is to include the systematics during the CI calculation, described in sec. 6.1.3. However this is something that should be tackled to improve the technique in the future.

The main sources of systematics errors on the effective volume, i.e. photon propagation and global OMs sensitivity, and in less measure the neutrino oscillation, were evaluated through some dedicated signal Monte Carlo investigations for the various neutralino models. The remaining sources of systematics were taken from the literature.

In these Monte Carlo studies, our baseline set-up to calculate the effective volume V_b is the one that lead to the values in tab. 5.16; the additional samples, processed through the same simulation and analysis chain, as described in sec. 4.2, have only one different component at a time with respect to the baseline set-up; the effective volume is then noted as V_o. The relative error on the effective volume V_b is then:

$$\frac{\Delta V}{V} = \frac{V_o - V_b}{V_b}$$

The main sources of systematics considered in this work are enumerated and discussed below.

1. **Neutrino oscillation**.
   The simulation of neutralinos takes into account neutrino oscillations with parameters illustrated in sec. 4.2.1 (standard set). The measured values of the neutrino oscillation parameters deviate from zero by at least 10σ on average [184]. We evaluated the 1σ error arising from the uncertainty on the oscillation parameters by taking the $1/10^{th}$ of the difference in effective volume for standard and no oscillation samples:

   $$\frac{\Delta V}{V} = \frac{1}{10}\left(\frac{V_{\text{noosc}} - V_b}{V_b}\right)$$

   The results of this study are shown in fig. 6.9. The difference in effective volume between the two samples is shown on the top part of the picture for hard (left) and soft (right) annihilation channels. At the bottom on the same picture this difference is converted to relative errors for hard (left) and soft (left) channels. The 1σ error in the oscillation parameters plays actually a secondary role; at the end we quote a conservative error of $\pm 3\%$ for all the models.

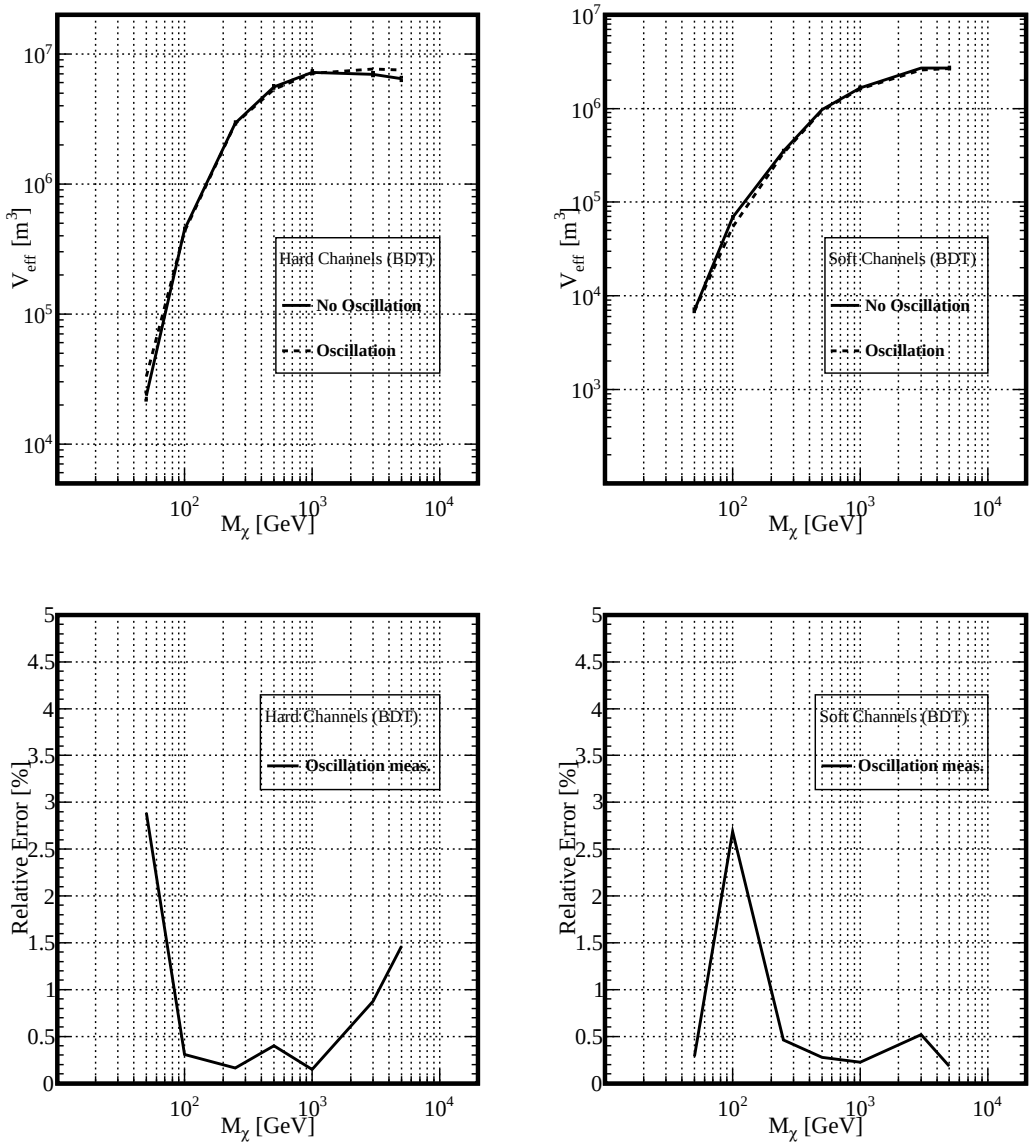


Figure 6.9: *Top: effective volume after BDT selection for samples with (dashed line) and without (solid line) oscillations, for hard (left) and soft (right) neutralino models. Bottom: relative error due to the change in the generator settings, for hard (left) and soft (right) neutralino models.*

2. **Neutrino-nucleon cross section**
   The relative error on the neutrino-nucleon cross section, extracted from [140], yields a systematic uncertainty around ±3% for all the neutralino models.

3. **Tau neutrinos**
   The neutralino Monte Carlo samples, produced for this analysis, did not take into account ν_τ events, since they typically produce showers with too little angular information to survive the event selection. However, outgoing τ leptons can decay to muons in 17% of the cases. They could then survive the event selection, thereby contributing to the total V_{eff}. This effect was estimated to be around $\pm 2-5\%$, being more important for higher energies [185].

4. **Muon propagation in ice**
   The `MMC` code (see sec. 4.2.2), which calculated the energy losses by the propagating muon through the ice, estimated an error on the muon track length of around 1% [105]. This was translated to an error of around 1% on the observed event rate by an analysis searching for astrophysical neutrinos [186]. In this analysis the considered energy range of neutrino induced muons is below the TeV scale, where the ionisation is the prominent process of energy losses, and is much better known than the stochastic energy losses occurring at higher energies. Hence, the error on V_{eff} was expected to be smaller than ±1%.

5. **Ice model and OM sensitivity**

Ice model Another source of systematic error is the uncertainty related to the optical ice properties. As we have seen in sec. 3.2.4, these properties were inferred from calibration of experimental data, which were subject to errors, though.

Two ice models were considered in this systematics study, with which the `Photonics` package (see sec. 4.2.3) was fed: `MILLENNIUM wl06v200` and the more recent `AHAv2` model, which is our baseline ice model. The comparisons of the effective volumes are shown on the top part of fig. 6.10; it is evident from the plots that the relevant differences arise at low energies; from the plots outlined at the bottom of the same picture, we infer that the relative error ranges between ± 3 and 30%.

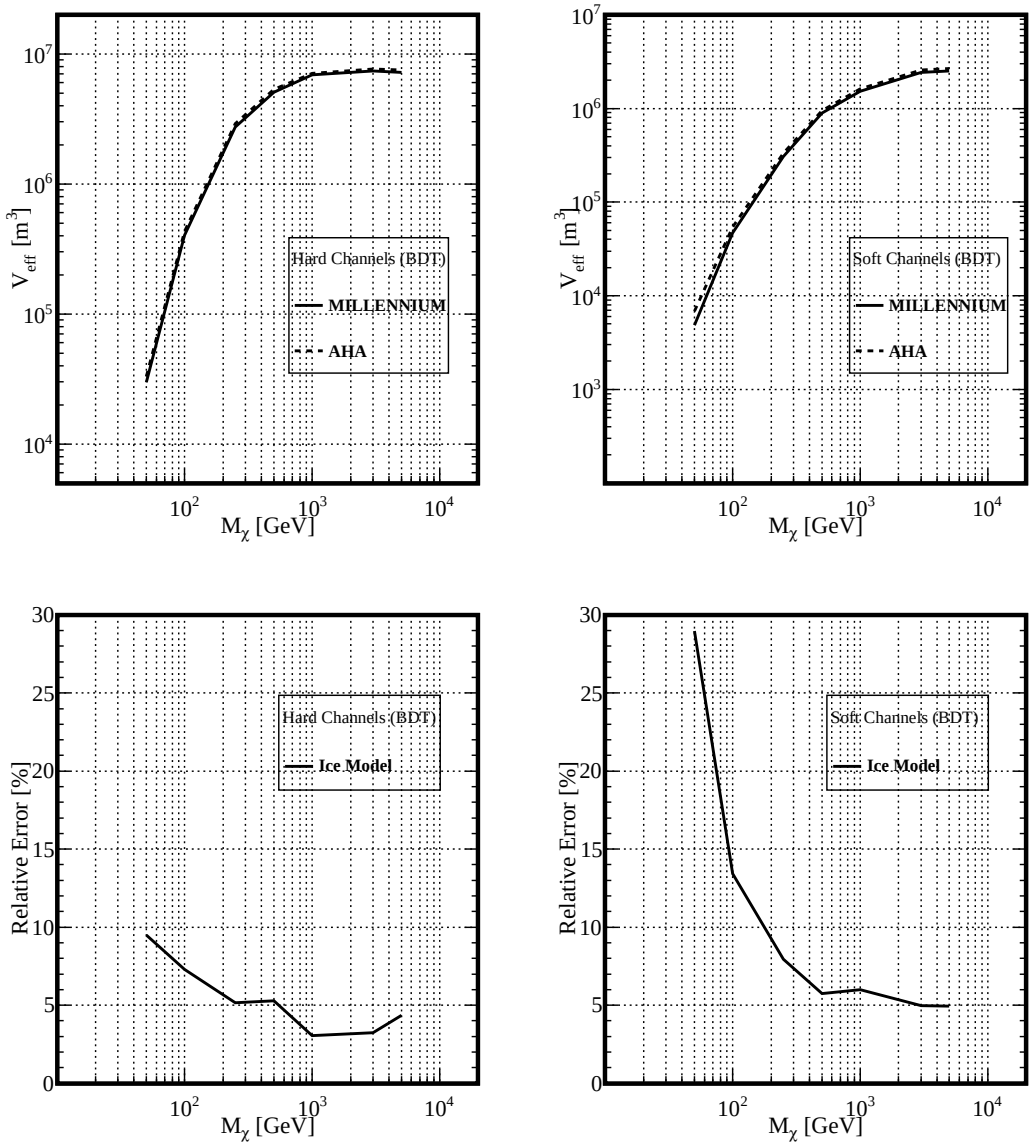


Figure 6.10: *Top: effective volume after BDT selection for samples processed with AHA (dashed line) and MILLENNIUM (solid line) ice models, for hard (left) and soft (right) channels. Bottom: relative error due to the change in the ice model settings, for hard (left) and soft (right) neutralino models.*

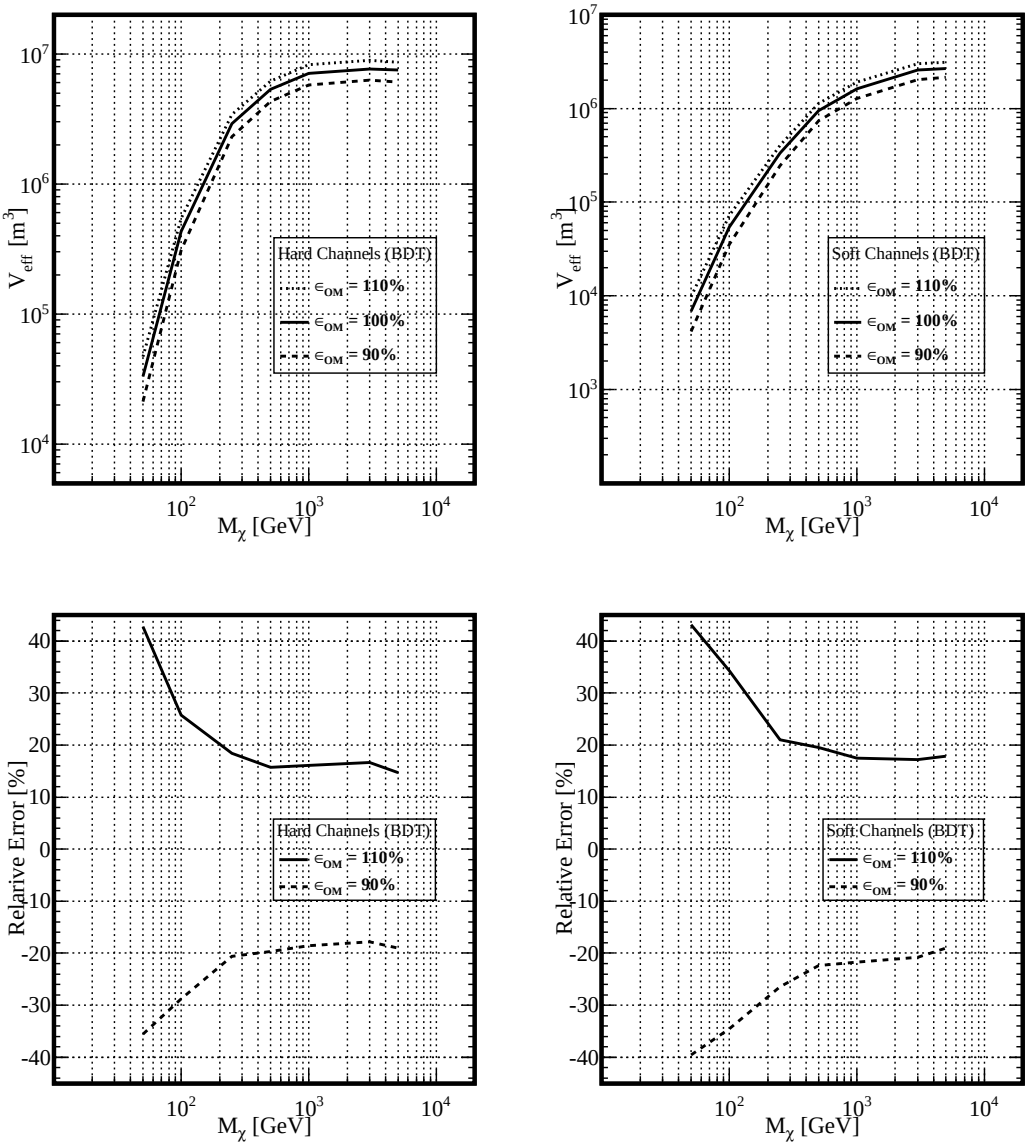


Figure 6.11: *Top: effective volume after BDT selection for three samples with different settings of the global scale of the OM efficiency (±10%), for hard (left) and soft (right) channels. Bottom: relative error due to the change in the global OM efficiency, for hard (left) and soft (right) neutralino models.*

OM sensitivity In the simulation step concerning the detector response (see sec. 4.2.3), the light collection efficiency of each OM was tuned in order to reach a better agreement with experimental data. Actually each OM had its own individual sensitivity due to the electronics and to other agents. The experimental trigger rate varied over the season with a spread of about 10 Hz (see e.g. top of fig. 4.3). In simulations, the change of trigger rate should be larger when the sensitivity of all OMs are scaled up or down at once, i.e. a global shift.

Hence, we simulated two additional Monte Carlo samples with the OM sensitivity globally shifted to $+10\%$ and -10% with respect to the baseline simulation. The comparisons of the calculated effective volumes, shown on the top part of fig. 6.11, indicated that this global shift effect was more evident for low energy neutralino models and for horizontal events.

The relative errors on V_{eff} (see bottom of fig. 6.11), do not show a large asymmetrical behaviour, ranging then symmetrical around $\pm 20 - 40\%$.

The uncertainties from the ice model and from the OM sensitivity were quadratically summed and summarised in tab. 6.3 (hard channel) and tab. 6.4 (soft channel).

6. **Time and geometry calibration**
A dedicated Monte Carlo study [185] evaluated the uncertainties of timing and geometry calibration on V_{eff}; they turned out to decrease the effective volume by less than 5%. Thus we quoted a conservative error of $\pm 5\%$.

6.2.3 Total uncertainties

Under the assumption that all the uncertainties (whether statistical or systematic) on V_{eff} were uncorrelated, its total relative uncertainty is then expressed by the root of the sum of the squared uncertainty of each individual source

$$\left(\frac{\Delta V}{V}\right)_{\text{tot}} = \sqrt{\sum_i \left(\frac{\Delta V}{V}\right)_i^2} \tag{6.6}$$

Tables 6.3 and 6.4 summarise the results for the hard and soft channel respectively. The errors are asymmetric due to the OM sensitivity. For all the other sources we have assumed conservative symmetric errors. From the numbers in the tables we infer also that the dominant contributions come from the ice model and OM sensitivity.

Since we considered for μ_s only the statistical uncertainty be the relevant one, no other sources need to be quadratically summed alongside with it to yield at

Neutralino mass - hard channel	50	100	250	500	1000	3000	5000
Statistical uncertainty [%]	±2	±1	±1	±1	±1	±1	±1
Systematic uncertainty [%]							
1. Neutrino oscillation	±3	±3	±3	±3	±3	±3	±3
2. Neutrino-nucleon cross-section	±3	±3	±3	±3	±3	±3	±3
3. Tau neutrinos	±2	±2	±2	±2	±3	±4	±5
4. Muon propagation in ice	±1	±1	±1	±1	±1	±1	±1
5. Ice + OM sensitivity	$^{+44}_{-37}$	$^{+27}_{-30}$	$^{+19}_{-21}$	$^{+16}_{-20}$	$^{+16}_{-19}$	$^{+17}_{-18}$	$^{+15}_{-19}$
6. Time and geometry calibration	±5	±5	±5	±5	±5	±5	±5
Total uncertainty on V_{eff} [%]	$^{+45}_{-38}$	$^{+28}_{-31}$	$^{+20}_{-22}$	$^{+17}_{-21}$	$^{+18}_{-20}$	$^{+19}_{-20}$	$^{+17}_{-20}$

Table 6.3: *Summary of the relative uncertainties on V_{eff} at the final level for the hard channels.*

the end the total uncertainty on μ_s. Hence (see sec. 6.2.2)

$$\left(\frac{\Delta\mu_s}{\mu_s}\right)_{\text{tot}} = \frac{\sigma_{\mu_s}}{\overline{\mu}_s} \tag{6.7}$$

Neutralino mass - soft channel	50	100	250	500	1000	3000	5000
Statistical uncertainty [%]	±2	±1	±1	±1	±1	±1	±1
Systematic uncertainty [%]							
1. Neutrino oscillation	±3	±3	±3	±3	±3	±3	±3
2. Neutrino-nucleon cross-section	±3	±3	±3	±3	±3	±3	±3
3. Tau neutrinos	±2	±2	±2	±2	±2	±3	±4
4. Muon propagation in ice	±1	±1	±1	±1	±1	±1	±1
5. Ice + OM sensitivity	$^{+52}_{-49}$	±37	$^{+22}_{-28}$	$^{+20}_{-23}$	$^{+18}_{-23}$	$^{+18}_{-21}$	$^{+18}_{-19}$
6. Time and geometry calibration	±5	±5	±5	±5	±5	±5	±5
Total uncertainty on V_{eff} [%]	$^{+52}_{-49}$	±38	$^{+23}_{-29}$	$^{+21}_{-24}$	$^{+19}_{-24}$	$^{+19}_{-22}$	$^{+20}_{-21}$

Table 6.4: *Summary of the relative uncertainties on V_{eff} at the final level for the soft channels.*

6.2.3.1 Further checks on systematic uncertainties

As argued before, the main source of systematic uncertainties are the ice model and OM sensitivity. However are there further possible sources of systematic uncertainties? Did the BDT method introduce some unknown sources of systematic uncertainties? To answer to these questions we should compare data and atmospheric background simulation after the BDT selection.

If we compare only the shapes of the observables, the experimental data and the Monte Carlo background are in agreement; but if we look at the absolute rates, possible sources of systematic uncertainty may show up.

From tab. 5.15 we can infer that actually the BDT method selected more atmospheric Monte Carlo events than experimental data. We see that for the high energy models (from 250 GeV hard on, and 500 GeV soft on), the data were contaminated almost exclusively by atmospheric neutrino induced muons, while for intermediate models (100 GeV hard and 250 GeV soft) the contamination was roughly equally shared between the two sources of atmospheric background. The contamination in the experimental data after the lowest energy BDT selection (50 GeV soft, 100 GeV Soft and 50 GeV hard), was instead almost due to the atmospheric muons.

Following this argumentation, we have evaluated the total uncertainty (due both to statistic and systematic uncertainties), by comparing the number of experimental data events and expected atmospheric neutrino events, which were selected after the BDT optimisation for the high energy neutralino models, as follows

$$\frac{\Delta\mathcal{E}_{tot}}{\mathcal{E}_{tot}} = \frac{N_{data} - N_{atm.\nu}}{N_{atm.\nu}} \tag{6.8}$$

thus, we can write

$$\frac{\Delta\mathcal{E}_{tot}}{\mathcal{E}_{tot}} = \sqrt{\left(\frac{\Delta\epsilon_{stat}}{\epsilon_{stat}}\right)^2 + \left(\frac{\Delta\epsilon_{15pc}}{\epsilon_{15pc}}\right)^2 + \left(\frac{\Delta x_{sys}}{x_{sys}}\right)^2} \tag{6.9}$$

where $\frac{\Delta\epsilon_{stat}}{\epsilon_{stat}}$ is the statistical uncertainty on $N_{atm.\nu}$, $\frac{\Delta\epsilon_{15pc}}{\epsilon_{15pc}}$ the systematic 15% uncertainty on the absolute atmospheric neutrino flux [187], and $\frac{\Delta x_{sys}}{x_{sys}}$ the systematic uncertainty to be evaluated.

We see from eq. 6.8 that the total uncertainty should be negative; indeed, as we said before, we have more simulated neutrino events than experimental data after the BDT selection. From calculation for the high energy models, we find that $\frac{\Delta x_{sys}}{x_{sys}}$ is of the order of the systematic uncertainties quoted for the simulated signal, considering the ice model, the OM sensitivity and the time and geometry calibration (see tabs. 6.3 and 6.4). This means that the uncertainties in the atmospheric neutrino background flux are mainly due to the medium and detector simulation issues. If BDT itself introduced some unknown systematics, then they are almost negligible.

If we apply the above explained method to the intermediate and low energy selections, we should make some change in the eqs. 6.8 and 6.9, to take into

account also the atmospheric muon contamination. The systematic uncertainties found for these models are consistent with the ones quoted for the effective volumes (see tabs. 6.3 and 6.4).

For the models 50 GeV soft, 50 GeV hard and 250 GeV soft, we find values of $\frac{\Delta x_{sys}}{x_{sys}}$ which are of the order -64%, -46% and -44%. We assume that, like for the atmospheric neutrinos, there is an uncertainty on the muon flux of about 15%; the remaining part is due to some unknown effect in the simulation, not visible in the atmospheric neutrino sample. However, we assume that the uncertainty on the effective volume for these models should be closer to the uncertainty quoted for atmospheric neutrinos than for atmospheric muons. Hence, we replace the uncertainties in tabs. 6.3 and 6.4, by taking the average between the above numbers and the ones quoted in the tables. The final uncertainties for these models are: -57%, -42% and -37%. These are indeed the actual values for $\frac{\Delta V}{V}$ which we will put in eq. 6.10 to evaluate the total relative uncertainty on $\Gamma_{\nu\to\mu}$ for these neutralino models (see next section).

6.2.4 Propagation of the total uncertainties

We have already pointed out that the inclusion of systematic uncertainties should be exerted during the construction of the μ_s confidence interval, to guarantee proper coverage of the interval in $\Gamma_{\nu\to\mu}$.

Since our test statistic described in sec. 6.1.3 was not Poissonian, well-established and definitive methods to include systematic uncertainties were not developed yet; thus we have adopted a simple method which was developed and used by other analyses [111, 185]. Through a dedicated toy Monte Carlo study, it was found that

$$\pm\frac{\Delta\Gamma}{\Gamma}=\sqrt{\left(\frac{\Delta\mu}{\mu}\right)^2_{\text{tot}}+\left(\frac{\Delta V}{V}\right)^2_{\text{tot}}\left(\frac{1}{1\mp\left(\frac{\Delta V}{V}\right)_{\text{tot}}}\right)^2} \tag{6.10}$$

was a good approximation of the total relative uncertainty on the conversion rate. This error was asymmetric because the conversion rate did not depend linearly on the effective volume. A small downward (large upward) fluctuation of V_{eff} could lead to a large upward (small downward) fluctuation of $\Gamma_{\nu\to\mu}$.

Since the uncertainty on μ_s is smaller than 1%, the uncertainty on the effective volume completely dominates the total uncertainty on the conversion rate. If we assume that the uncertainties on the factors in eqs. 6.15 and 6.17 are negligible, compared to that on $\Gamma_{\nu\to\mu}$, the uncertainties on Γ_A, Φ_μ and on the SI and SD neutralino-proton cross sections can be derived by multiplying the results with

the scale factor

$$1 \pm \frac{\Delta\Gamma}{\Gamma} \tag{6.11}$$

These scale factors, for each neutralino model, are visualised in tabs. 6.7 and 6.9, respectively for hard and soft channels.

6.3 Physical quantities

In these sections we will illustrate the relevant physical quantities which can be evaluated through our analysis.

The connection between the volumetric flux $\Gamma_{\nu\to\mu}$ and the signal event rate, μ_s/t, is realised once the effective volume of the detector V_{eff} is known. Hence

$$\frac{\mu_s}{t} = \Gamma_{\nu\to\mu} V_{\text{eff}} \tag{6.12}$$

We therefore expect that a number of

$$\mu_s = \Gamma_{\nu\to\mu} V_{\text{eff}} t \tag{6.13}$$

neutrino induced muons are detected during a time t.

We have explained in the previous chapter how to reduce the atmospheric background, and, at the beginning of this chapter, how to evaluate the signal strength, μ_s, from the remaining filtered data. Hence, after estimating the effective volume from Monte Carlo signal simulations, we can infer the volumetric flux (or neutrino-to-muon conversion rate) by means of eq. 6.13.

To derive all these calculations we don't need to known a priori the scale of the incoming flux, but its energy spectrum to optimise the selection criteria and estimate the effective volume. The absolute signal flux is a quantity which we will measure a posteriori.

On the same line we can infer other relevant physical quantities, which are described in the following sections.

6.3.1 Conversion rate

In a background-only scenario, an upper limit can be set to the neutrino-to-muon conversion rate, as follows

$$\Gamma_{\nu\to\mu} \leq \Gamma^{90}_{\nu\to\mu} = \frac{\mu_s^{90}}{V_{\text{eff}} \cdot t_{\text{live}}} \tag{6.14}$$

where t_{live} is the live-time of detector with effective volume V_{eff}, and μ_s^{90} is the upper limit on the signal content at 90% CL (see sec. 6.1.3).

A possible threshold (either from experiment or from simulation) on the muon energy could affect both μ_s^{90} and $V_{\rm eff}$; actually we set a threshold on the energy of simulated neutralino induced muon corresponding to $E_\mu^{\rm thr} = 10$ GeV (see sec. 4.2.1).

6.3.2 Annihilation rate

The volumetric flux is directly proportional to the neutralino annihilation rate in the Sun, Γ_A, by means of

$$\Gamma_{\nu\to\mu} = \frac{\Gamma_A}{4\pi D_\odot^2} \int_0^\infty dE_\nu \sigma_{\nu N} \left(E_\mu \geq E_{thr} \mid E_\nu\right) \rho_N \sum_X B_X \left(\frac{dN}{dE_\nu}\right)_X \tag{6.15}$$

where $D_\odot$ is the distance to the Sun, $\sigma_{\nu N}$ the neutrino-nucleon cross section (above a muon energy threshold, E_{thr}, of 10 GeV), ρ_N the nucleon density at the detector, and B_X the branching ratio for the annihilation channel X with associated neutrino energy spectrum $\left(\frac{dN}{dE_\nu}\right)_X$.

The total neutrino energy spectrum, i.e. the sum of all the contributions from each channel X, weighted by its branching ratio B_X, is the only unknown quantity in the above equation, which is then

$$\left(\frac{dN_\nu}{dE_\nu}\right) = \sum_X B_X \left(\frac{dN_\nu}{dE_\nu}\right)_X \tag{6.16}$$

The above quantity depends on some SUSY parameters, which unfortunately are unknown, i.e. the composition and the mass of the neutralino, which determine branching ratios and energy spectra. Hence, we had to choose a particular mass and branching ratio to translate the measured conversion rate to annihilation rate. In this work we perform an energy dependent optimisation, dependent on the neutralino model, considering neutralino masses between 50 and 5000 GeV. The annihilation in the hard channel (W^+W^-, or $\tau^+\tau^-$ if $m_\chi < m_W$) and in the soft channel ($b\bar{b}$) produce the hardest and the softest neutrino energy spectrum respectively. Any other choice of branching ratios leads to an intermediate energy spectrum; hence the experimental outcome should stay within those two extreme cases. In this way the annihilation rate can be easily compared to theoretical predictions (SUSY, dark matter abundance,...) or to other experiments.

6.3.3 Muon flux

Keeping on this line, a more useful comparison with other experiments can be performed through the calculation of the neutrino induced muon flux above

an energy threshold, E_{thr}, of 1 GeV, which can be expressed as follows

$$\Phi_\mu \left(E_\mu \geq E_{thr} \right) = \frac{\Gamma_A}{4\pi D_\odot^2} \int_{E_{thr}}^{\infty} dE_\mu \frac{dN_\mu}{dE_\mu} \tag{6.17}$$

where $\frac{dN_\mu}{dE_\mu}$ is the neutralino induced muon energy spectrum at the detector, which includes the propagation effects and the energy losses in the medium. More details about equ. 6.17 can be found in reference [188].

6.3.4 Neutralino-proton cross section

If we compare eqs. 2.1 and 2.3 at "full strength", i.e. at the equilibrium[1], we will derive that the annihilation rate is proportional to the total neutralino-nucleon scattering cross section, σ_{tot}, where

$$\sigma_{tot} = \sigma_{\chi H}^{SD} + \sigma_{\chi H}^{SI} + 0.07\sigma_{\chi He}^{SI}$$

Under the hypothesis that the capture rate is fully dominated either by spin-dependent ($\sigma_{tot} \simeq \sigma_{\chi H}^{SD}$) or spin-independent ($\sigma_{\chi H}^{SI} + 0.07\sigma_{\chi He}^{SI}$) scattering, we can extract from the annihilation rate, either the SD or the SI neutralino-proton cross section. More details about the calculations of the proportionality factors between the annihilation rate and the SI or SD cross section can be also found in [188]. We should mention that in all these calculations, some planetary gravitational effects on the capture rate were neglected.

6.4 Final results

In sec. 6.1.3 we develop a hypothesis test and apply it on blinded data to extract the sensitivity (or median upper limit) on μ_s and $\Gamma_{\nu\to\mu}$, in a background-only scenario (see tab. 6.2).

The analysis presented in this work, from the event selection to the handling of systematics uncertainties, was subject to a review within the AMANDA collaboration, where it was carefully checked. Once the analysis procedure had the green light by the collaboration, we were allowed to "unblind" the experimental data, namely to make all the calculations looking at the real position of the Sun, to derive the confidence interval on μ_s. The calculated lower and upper limits, which correspond to different signal model optimisations, along with the best $\hat{\mu}_s$ values, are shown in tab. 6.5 and visualised in fig. 6.12.

[1] Following the argument in [188], we will assume that equilibrium between capture and annihilation rate could happen in the Sun for each different neutralino model considered in this work.

Neutralino model optimisation	μ_s^{low}	μ_s^{up}	$\widehat{\mu}_s$	$\mathcal{P}$-value
50 soft	0.00	31.72	5.94	0.50
100 soft	0.00	21.92	5.93	0.09
250 soft	2.34	22.54	10.21	0.18
500 soft	0.00	11.37	0.51	0.22
1000 soft	0.00	14.19	3.40	0.51
3000 soft	0.00	10.66	0.49	0.31
5000 soft	0.00	10.26	0.49	0.48
50 hard	0.00	16.39	4.29	0.44
100 hard	0.00	16.51	5.14	0.07
250 hard	0.00	15.28	4.23	0.13
500 hard	0.00	13.54	2.41	0.07
1000 hard	0.00	7.43	0.00	0.21
3000 hard	0.00	11.20	0.91	0.14
5000 hard	0.00	9.50	0.00	0.17

Table 6.5: *For each neutralino selection, the lower and upper limit at 90% CL, and the best physical fit $\widehat{\mu}_s$ to unblinded experimental data. The last column shows the $\mathcal{P}$-value (see text).*

The numbers in the table indicate that, apart from the 1000 GeV hard and 3000 GeV hard model selections, a non-zero signal contribution ($\widehat{\mu}_s > 0$), fits the data selected by each of the remaining model optimisations.

The lower limits, apart from the case of 250 GeV soft, meet the physical boundary, i.e. zero. As we can see from fig. 6.12, for most of the neutralino selections, the upper limits show upward fluctuations, if compared with the median upper limits in a background-only scenario (which are projected on tab. 6.2, along with their 1σ spreads). However, this is not the case for the 100 GeV hard and 250 GeV soft selections, where, as we can see in fig. 6.12, the upper limits are above the 1σ median limit, with a factor of respectively 1.8 and 2.3 of the background-only scenario. On the right side of figs. 6.13 and 6.14, the log-likelihood ratio (the FC rank $\mathcal{R}$) distributions are plotted (in solid lines), for these two models, as well as their interception with the 90% critical region (dashed lines).

Well, then how significant are these two upward fluctuations? Are they consistent with the expected background, within some confidence? To try to answer these questions, we investigated the space angle distributions to see how they behave in the domain of the hypothetical signal. These regions, having as indication the mass-dependent angular resolutions (see tab. 6.1), could span between a few to more than ten degrees.

The space angle distributions in the vicinity of the signal, for the 100 GeV

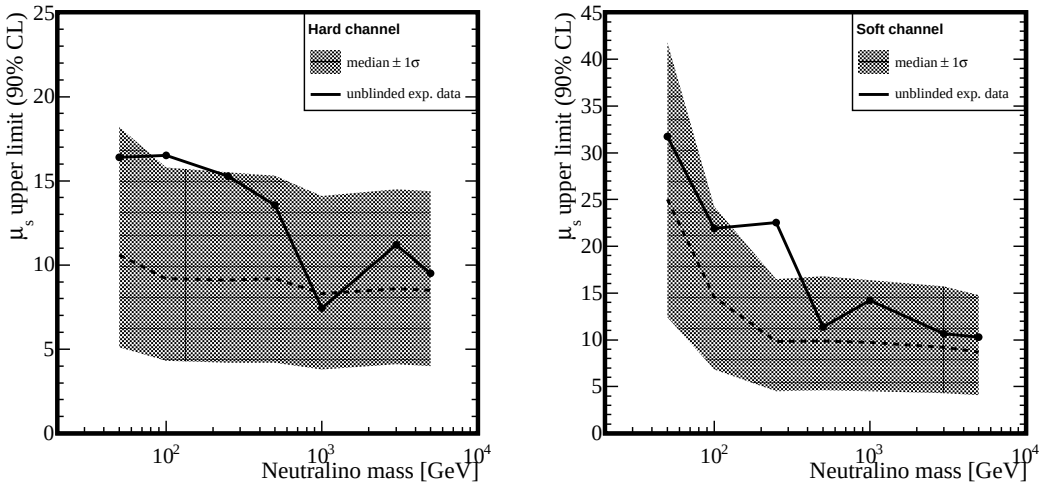


Figure 6.12: *Upper limits on μ_s at 90% CL extracted from unblinded experimental data (markers + solid lines), and the median upper limits with the 1σ spread in the background-only scenario (dashed lines + shaded grey areas). They are reported for all neutralino masses and channels (left hard, right soft).*

	ψ^{90} [degree]						
Neutralino model [GeV]	50	100	250	500	1000	3000	5000
Hard channel	20.6	11.7	8.6	7.9	7.6	7.6	7.5
Soft channel	43.2	21.4	11.6	9.8	9.0	8.6	8.3

Table 6.6: *90% quantile of the space angle distributions of all neutralino models.*

hard and 250 GeV soft model selection, are shown on the left side of fig. 6.13 and fig. 6.14. The complete series of space angle distributions, related to the 14 neutralino models, is reported in appx. B.

To evaluate the probability that the upward fluctuations may originate from statistical fluctuations of the background, we have performed a statistical hypothesis test for the presence of a signal, considering as signal region the 90% quantile of the signal space angle distribution. The calculated quantiles, for each neutralino signal model, are shown in tab. 6.6.

We have followed the procedure reported in [189], where the test statistic is defined as

$$\mathcal{T} = x - y/\tau$$

where x is the number of observations in the signal region, y is the number of observations in the background region and τ is the probability that a background event falls into the background region divided by the probability that it

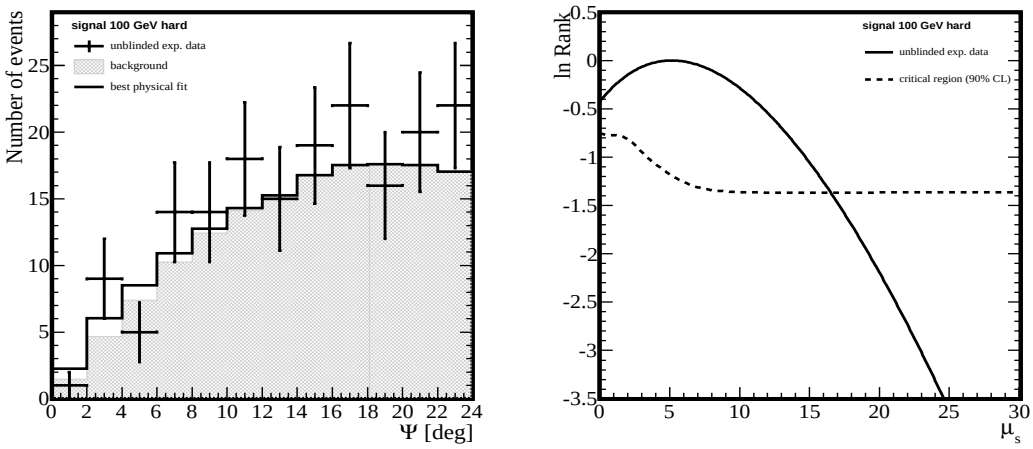


Figure 6.13: *On the left: space angle distribution in the vicinity of the Sun for the 100 GeV hard model. In the picture the unblinded experimental data, the expected background and the best physical signal+background fit to the experiment are shown; the latter two distributions are normalised to the total data set. On the right: the* $\ln \mathcal{R}$ *distribution for the experiment (solid line), and the interceptions with the 90% CL critical region.*

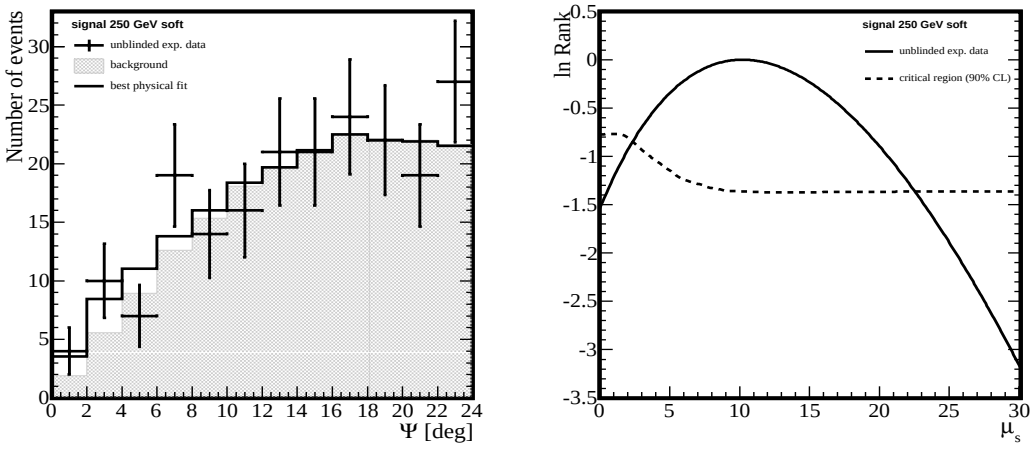


Figure 6.14: *On the left: space angle distribution in the vicinity of the Sun for 250 GeV soft model. In the picture the unblinded experimental data, the expected background and the best physical signal+background fit to the experiment are shown; the latter two distributions are normalised to the total data set. On the right: the* $\ln \mathcal{R}$ *distribution for the experiment (solid line), and the interceptions with the 90% CL critical region.*

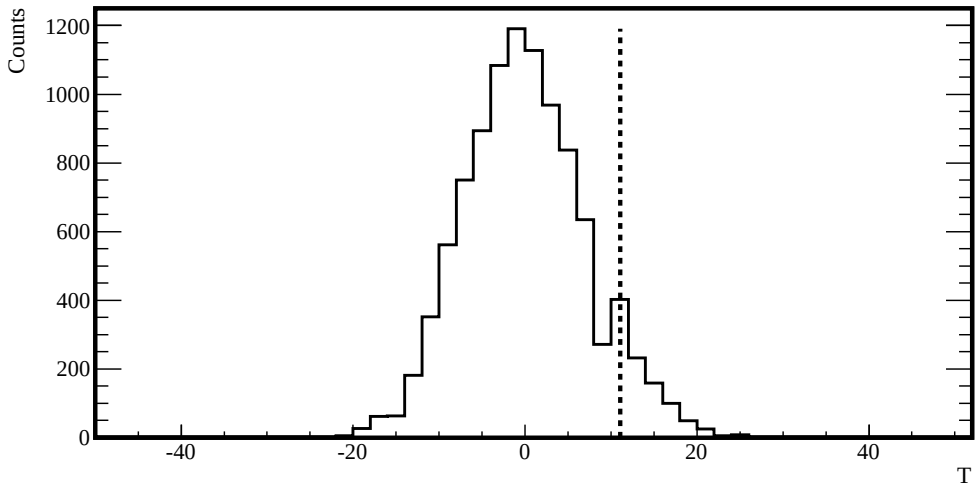


Figure 6.15: *Distribution of the test statistic $\mathcal{T}$ for 10000 pseudo-experiments, for the 100 GeV hard model. The vertical dashed line marks the right part of the histogram to integrate (see text).*

falls into the signal region. We can exploit the background p.d.f., as defined in sec. 6.1.2, to calculate the τ ratio, thus we have all the ingredients to calculate the $\mathcal{T}$ value. Under the null hypothesis all the events observed are background, and we can perform 10000 pseudo-experiments by randomly drawing the $x+y$ number of observations from the background p.d.f., and computing $\mathcal{T}_i$ for each pseudo-experiment. As an illustration of the method, the histogram of fig. 6.15 shows the calculated $\mathcal{T}_i$ for 10000 pseudo-experiments for the 100 GeV hard model. Hence, we can count how many pseudo-experiments have a $\mathcal{T}_i$ greater than $\mathcal{T}$, just integrating the right part of the histogram marked by the dashed vertical line, and then divide this number by 10000; this ratio is the $\mathcal{P}$-value and it can be calculate for each neutralino model selection. The calculated $\mathcal{P}$-values are reported in the last column of tab. 6.5.

Given a certain threshold (usually one quotes 3 or 5 σ), we did not find any significant excess of events in the direction of the Sun, for any of the neutralino models. The upward fluctuations noticed for 100 GeV hard and 250 GeV soft selections, should be then interpreted as upward background fluctuations in the region between 2-4 degrees for the first model, and between 2-4 and 6-8 degrees for the second.

Since the experimental data sample was entirely compatible with background, and no significant excess was observed in the direction of the Sun, we translate the above information to upper limits on the relevant physical quantities introduced in sec. 6.3.

The upper limit on the neutrino-to-muon conversion rate at 90% CL, $\Gamma^{90}_{\nu\to\mu}$, can be derived through eq. 6.14. The translation of eqs. 6.15 and 6.17 to upper limits on the annihilation rate Γ^{90}_{A} and the muon flux Φ^{90}_{μ} were obtained through the procedure explained in [188].

We remind here that the neutrino-to-muon conversion rate is valid for muon energies above 10 GeV, due to our setting in the Monte Carlo neutralino generator used to calculate the effective volumes (see sec. 4.2.1). Hence, to compare our results to the ones of other lower threshold experiments, the muon flux calculation has been rescaled to a 1 GeV muon energy threshold. Our final results are projected on tabs. 6.7 and 6.9 (hard and soft channels), where we also provide the 90% median upper limit (sensitivity) on the muon flux, $\widetilde{\Phi}^{90}_{\mu}$.

The obtained upper limits at 90% CL on the annihilation rate and muon flux are visualised on the top and bottom part of fig. 6.16, in solid (dashed) black lines for hard (soft) channels. These limits are compared to what is expected in the background-only scenario, represented by the grey area in the pictures. We notice that above 1000 GeV, both the hard and soft muon flux limits approach a plateau, where they are roughly within a factor of two from each other. At low masses the difference between the soft and hard channels is more evident, where it reaches a factor of ten for 50 GeV. In these upper limit plots, systematic uncertainties are not included.

As we explained in sec. 6.2.4, scale factors are provided to incorporate the total uncertainties on the measurements. The same scale factors should be used for the conversion rate, the annihilation rate, the muon flux and the neutralino-proton cross section; they can be found in the last column of tabs. 6.7, 6.8 and 6.9, 6.10.

As a final remark, we can conclude that the string trigger played a key role in the efficiencies for the lowest energy neutralino models; without this trigger there would be a tremendous reduction in sensitivity. The string trigger is therefore essential in this low energy region, even for horizontal fluxes.

Wimp Mass (hard channel) [GeV]	$\widetilde{\Phi}_{\mu}^{E_\mu>1\,\mathrm{GeV}}$ [$\mathrm{km}^{-2}\ \mathrm{y}^{-1}$]	$\Gamma_{\nu\to\mu}^{E_\mu>10\,\mathrm{GeV}}$ [$\mathrm{Km}^{-3}\ \mathrm{y}^{-1}$]	$\Gamma_{A}^{E_\mu>1\,\mathrm{GeV}}$ [s^{-1}]	$\Phi_{\mu}^{E_\mu>1\,\mathrm{GeV}}$ [$\mathrm{km}^{-2}\ \mathrm{y}^{-1}$]	err. scale $1\ -/+\ \frac{\Delta\Gamma}{\Gamma}$
50	$1.05\cdot10^{4}$	$1.76\cdot10^{5}$	$6.76\cdot10^{23}$	$1.62\cdot10^{4}$	0.69 / 1.72
100	$1.48\cdot10^{3}$	$1.93\cdot10^{4}$	$7.41\cdot10^{22}$	$2.66\cdot10^{3}$	0.78 / 1.45
250	$4.69\cdot10^{2}$	$2.60\cdot10^{3}$	$5.38\cdot10^{21}$	$7.87\cdot10^{2}$	0.83 / 1.28
500	$3.66\cdot10^{2}$	$1.19\cdot10^{3}$	$2.08\cdot10^{21}$	$5.42\cdot10^{2}$	0.85 / 1.27
1000	$3.25\cdot10^{2}$	$5.11\cdot10^{2}$	$9.70\cdot10^{20}$	$2.91\cdot10^{2}$	0.85 / 1.25
3000	$3.11\cdot10^{2}$	$7.11\cdot10^{2}$	$1.88\cdot10^{21}$	$4.05\cdot10^{2}$	0.84 / 1.25
5000	$3.11\cdot10^{2}$	$6.36\cdot10^{2}$	$1.94\cdot10^{21}$	$3.48\cdot10^{2}$	0.85 / 1.25

Table 6.7: *For the hard channels: median upper limit on the muon flux, and final limits on the conversion rate, annihilation rate and muon flux at 90% CL. The last column shows the systematics scale factors; we will adopt the more conservative one.*

Wimp Mass (hard channel) [GeV]	$\sigma_{\chi p}^{SD}$ [pb]	$\sigma_{\chi p}^{SI}$ [pb]	err. scale $1\ -/+\ \frac{\Delta\Gamma}{\Gamma}$
50	$1.55\cdot10^{-3}$	$5.92\cdot10^{-6}$	0.69 / 1.72
100	$5.98\cdot10^{-4}$	$1.34\cdot10^{-6}$	0.78 / 1.4
250	$2.53\cdot10^{-4}$	$3.16\cdot10^{-7}$	0.83 / 1.28
500	$3.82\cdot10^{-4}$	$3.52\cdot10^{-7}$	0.85 / 1.27
1000	$7.06\cdot10^{-4}$	$5.45\cdot10^{-7}$	0.85 / 1.25
3000	$1.22\cdot10^{-2}$	$8.38\cdot10^{-6}$	0.84 / 1.25
5000	$3.50\cdot10^{-2}$	$2.34\cdot10^{-5}$	0.85 / 1.2

Table 6.8: *For the hard channels: upper limit at 90% CL on the spin-dependent and spin-independent neutralino-proton cross sections. The last column shows the systematics scale factors; we will adopt the more conservative one.*

Wimp Mass (soft channel) [GeV]	$\widetilde{\Phi}_{\mu}^{E_\mu>1\,\mathrm{GeV}}$ [km^{-2} y^{-1}]	$\Gamma_{\nu\to\mu}^{E_\mu>10\,\mathrm{GeV}}$ [Km^{-3} y^{-1}]	$\Gamma_{A}^{E_\mu>1\,\mathrm{GeV}}$ [s^{-1}]	$\Phi_{\mu}^{E_\mu>1\,\mathrm{GeV}}$ [km^{-2} y^{-1}]	err. scale $1 -/+ \frac{\Delta\Gamma}{\Gamma}$
50	$1.60 \cdot 10^5$	$1.70 \cdot 10^6$	$1.95 \cdot 10^{26}$	$2.03 \cdot 10^5$	0.66/ 2.33
100	$1.34 \cdot 10^4$	$2.06 \cdot 10^5$	$6.22 \cdot 10^{24}$	$2.01 \cdot 10^4$	0.72 / 1.61
250	$1.77 \cdot 10^3$	$3.27 \cdot 10^4$	$3.53 \cdot 10^{23}$	$4.07 \cdot 10^3$	0.81 / 1.59
500	$9.09 \cdot 10^2$	$6.29 \cdot 10^3$	$4.20 \cdot 10^{22}$	$1.04 \cdot 10^3$	0.83 / 1.32
1000	$6.09 \cdot 10^2$	$4.28 \cdot 10^3$	$2.04 \cdot 10^{22}$	$8.88 \cdot 10^2$	0.84 / 1.32
3000	$4.67 \cdot 10^2$	$2.17 \cdot 10^3$	$7.57 \cdot 10^{21}$	$5.39 \cdot 10^2$	0.84 / 1.28
5000	$4.33 \cdot 10^2$	$1.98 \cdot 10^3$	$6.32 \cdot 10^{21}$	$5.09 \cdot 10^2$	0.83 / 1.27

Table 6.9: *For the soft channels: median upper limit on the muon flux, and final limits on the conversion rate, annihilation rate and muon flux at 90% CL. The last column shows the systematics scale factors; we will adopt the more conservative one.*

Wimp Mass (soft channel) [GeV]	$\sigma_{\chi p}^{SD}$ [pb]	$\sigma_{\chi p}^{SI}$ [pb]	err. scale $1 -/+ \frac{\Delta\Gamma}{\Gamma}$
50	$4.47 \cdot 10^{-1}$	$1.71 \cdot 10^{-3}$	0.66/ 2.33
100	$5.02 \cdot 10^{-2}$	$1.12 \cdot 10^{-4}$	0.72 / 1.61
250	$1.66 \cdot 10^{-2}$	$2.07 \cdot 10^{-5}$	0.81 / 1.59
500	$7.72 \cdot 10^{-3}$	$7.13 \cdot 10^{-6}$	0.83 / 1.32
1000	$1.49 \cdot 10^{-2}$	$1.15 \cdot 10^{-5}$	0.84 / 1.32
3000	$4.93 \cdot 10^{-2}$	$3.37 \cdot 10^{-5}$	0.84 / 1.28
5000	$1.14 \cdot 10^{-1}$	$7.64 \cdot 10^{-5}$	0.83 / 1.27

Table 6.10: *For the soft channels: upper limit at 90% CL on the spin-dependent and spin-independent neutralino-proton cross section. The last column shows the systematics scale factors; we will adopt the more conservative one.*

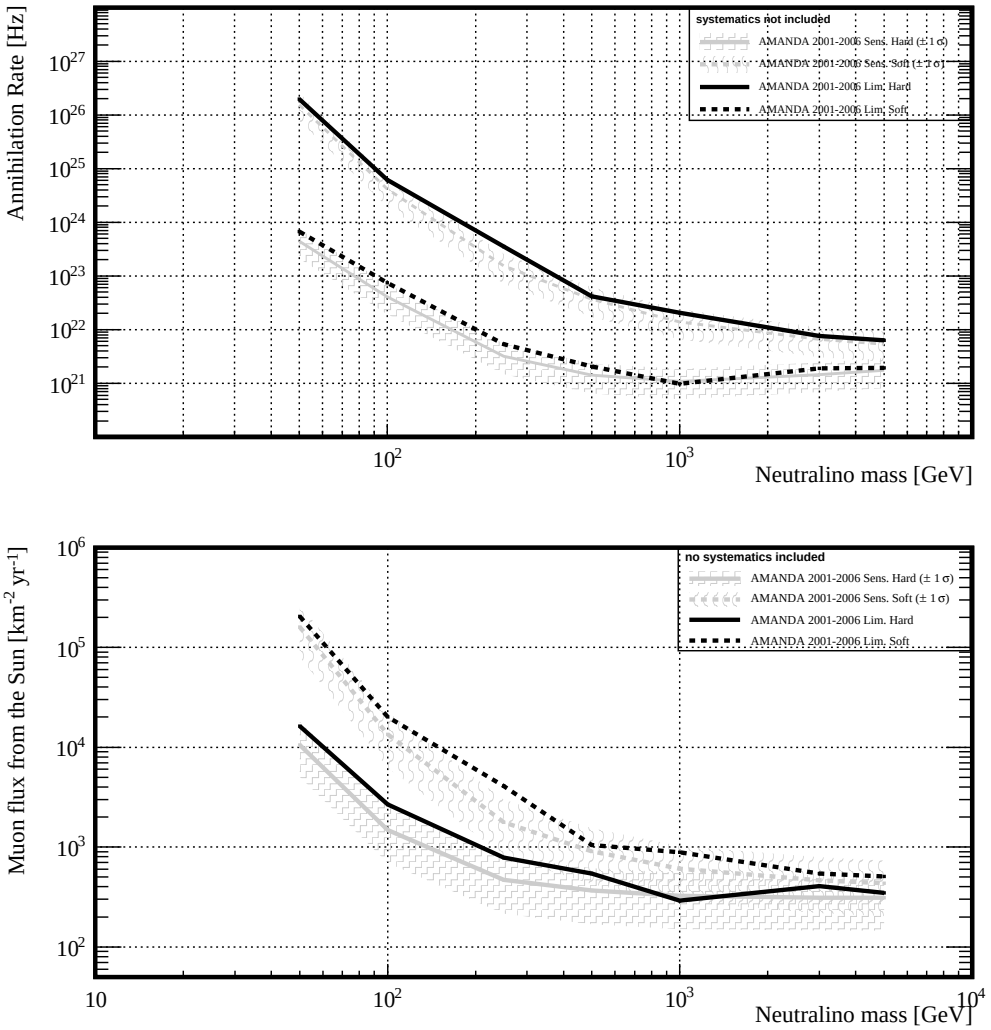


Figure 6.16: *Upper limits, with 90% CL, on the muon flux above 1 GeV from the Sun (bottom) and on the annihilation rate (top), for the background-only scenario (grey area) and for the unblinded experimental data (black lines). The channels are differentiated by the line styles, being solid (dashed) for hard (soft) annihilations. Systematic errors are not included*

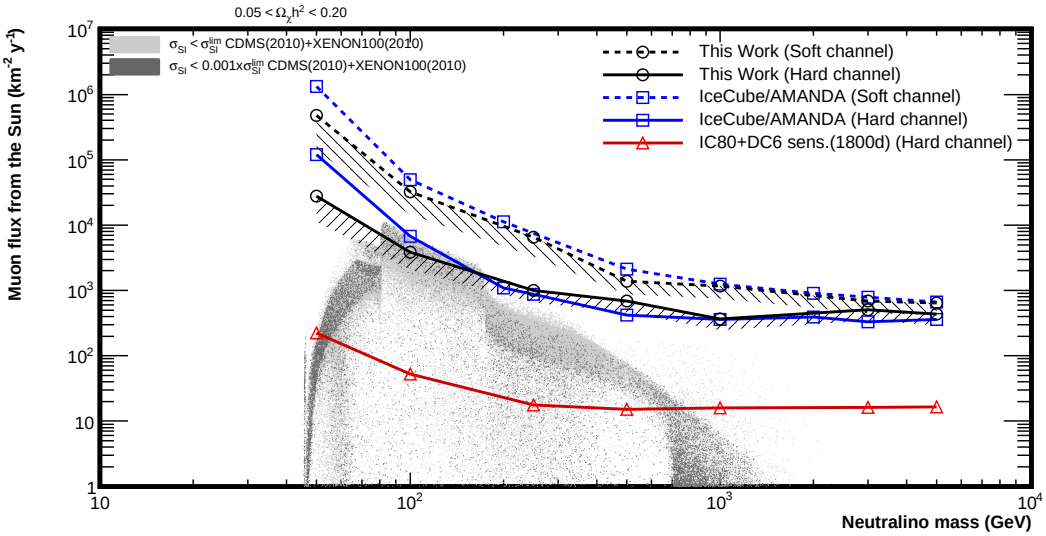


Figure 6.17: *The upper limits for each neutralino model (including systematics) on the muon flux from the Sun at 90% CL are shown as open black circles, and compared to other AMANDA and IceCube searches. The previous AMANDA and IceCube limits are shown as open blue squares. The open red triangles indicate the sensitivity which should be reached by DeepCore after 10 years of data-taking. The lines (solid for hard channel and dashed for soft channel) are there to guide the eyes. The MSSM-7 phase space is indicated by grey markers, which are still allowed by the spin-independent cross section limits set by direct searches.*

6.4.1 Muon flux

As described in sec. 6.3, the 90% CL upper limits on the neutrino-to-muon conversion rate were used to derive upper limits on the neutrino annihilation rates and the resulting muon flux in AMANDA.

In this and the next section we will compare our results with other experimental outcome and theoretical predictions. So far, no significant evidence of a neutralino signal from the Sun was reported by the other indirect search experiments .

The comparison of the upper limits of each neutralino model, at 90% CL and including total uncertainties (statistical + systematics), on the muon flux set by different AMANDA and IceCube searches with our results, are shown in fig. 6.17. The open blue squares represent the best limits extracted from the previous AMANDA and IceCube searches. Our limits in the plot are drawn as open black circles, with a band which shows the size of the total uncertainty. The upper edge of this band, marked by the black line, is our final result with the most conservative uncertainty. The lines (solid for hard channel and dashed for soft channel) are there to guide the eyes.

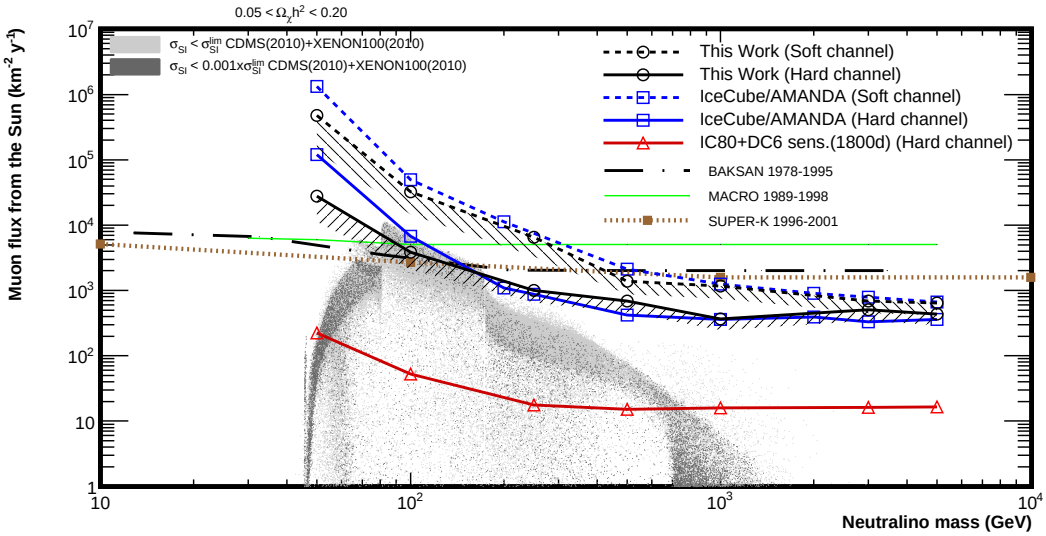


Figure 6.18: *The upper limits for each neutralino model (including systematics) on the muon flux from the Sun at 90% CL are shown as open black circles, and compared to other AMANDA and IceCube searches, along with other indirect searches. The previous AMANDA and IceCube limits are shown as open blue squares. The open red triangles indicate the sensitivity which should be reached by DeepCore after 10 years of data-taking. The lines (solid for hard channel and dashed for soft channel) are there to guide the eyes. The MSSM-7 phase space is indicated by grey markers, which are still allowed by the spin-independent cross section limits set by direct searches.*

Our results are so far the most sensitive AMANDA/IceCube results in the low energy region, i.e. in the whole soft channel domain, and for low energy hard channel models up to 200 GeV; while they are comparable beyond this mass. The previous AMANDA/IceCube analyses adopted different approaches: the one that set, before this work, the best limit in the low energy regions had as filtering simple one-dimensional sequential cuts [111]. It performed better with respect to the other analyses which were optimised for a high energy neutrino selection; one used IceCube data [190] and the other used a dataset selected in a search for point sources with the complete AMANDA data from 2000 to 2006 [132]. With respect to this latter, our method performs even better, if we consider that we had one year less statistics.

In the same picture the marked shaded grey area represents theoretical model predictions which were obtained by scanning over a reduced MSSM phase space with seven free parameters, referred to as MSSM-7 [31, 191], using the `DarkSusy` package. At each point in this multi-dimensional space several quantities are calculated: the neutralino mass, its admixture, the relic density, the annihilation

rate in the Sun, the muon flux from the Sun, the spin-independent and spin-dependent neutralino-nucleon cross section, etc... Then, the information of the spin-independent cross section is compared to the upper limits set by the direct search experiments CDMS [61] and XENON100 [58] (see sec. 1.4.2). The darker grey markers, set instead the expected limits when those experiment will increase their own sensitivity by a factor of 1000.

In fig. 6.18 we added the muon flux upper limits of different indirect search experiments: Super-K [192], BAKSAN [193] and MACRO [194]. Below 100 GeV Super-K, with its low threshold, sets the strongest limits, while above 100 GeV the AMANDA and IceCube neutrino telescopes benefit from their much larger volume.

In both pictures open red triangles are drawn, which represent the foreseen sensitivity, for each neutralino mass annihilating in hard channel, reached by the low energy extension of IceCube called DeepCore [195], considering 10 years of data-taking.

6.4.2 Cross sections

The comparison of the 90% CL upper limits on the SD neutralino-proton elastic cross section with the results from previous AMANDA and IceCube searches is shown on fig. 6.19. It is clear that our results perform better in the low energy region, which is the more interesting to probe several theoretical predictions. The comparison with other experiments is shown in fig. 6.20.

AMANDA and IceCube are most sensitive to the SD cross section since the huge abundance of hydrogen in the Sun favours axial-vector interactions, and thus put some stringent constraint. This is not feasible at the moment for direct detection experiments like CDMS [61], COUPP [68] and KIMS [69]. The grey shaded area shows again the MSSM-7 phase space, that is compatible with the direct search results on the spin-independent scattering, while the darker grey colour marks their future sensitivities.

The open red triangles show the foreseen sensitivity for each neutralino model on hard channels for DeepCore, the low energy extension of IceCube, considering 10 years of data-taking.

The SD and SI cross sections were also obtained using the formalism from [188]. The calculated upper limits at 90% CL on cross sections (both SI and SD) are also shown in tabs. 6.8 and 6.10 for the hard and soft channels.

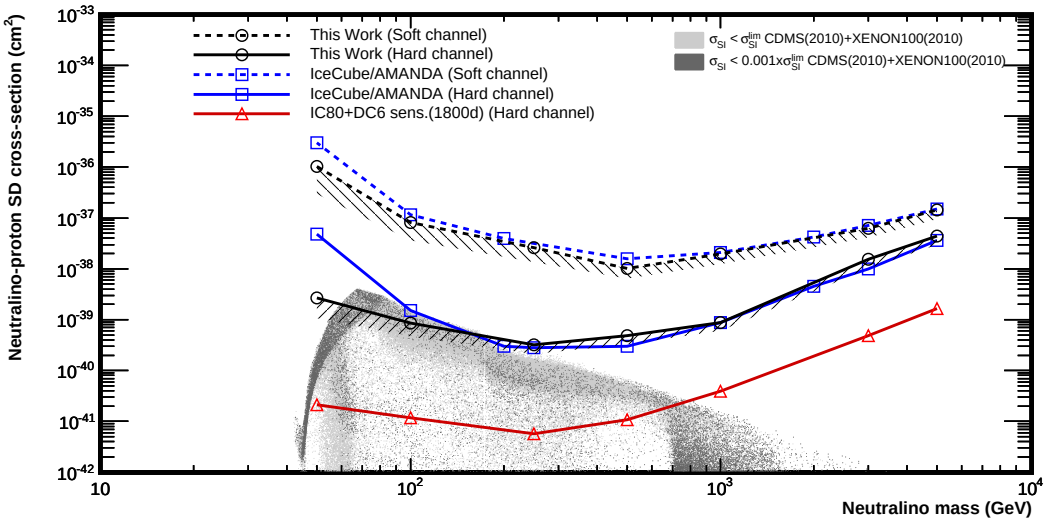


Figure 6.19: *Upper limit at 90% CL on the spin-dependent neutralino scattering cross section on protons, including systematics. The open red triangles indicate the foreseen sensitivity, for each neutralino model, which should be reached by DeepCore after 10 years of data-taking. The grey markers represent the MSSM-7 parameter space not excluded by the spin-independent cross section results from CDMS and XENON100.*

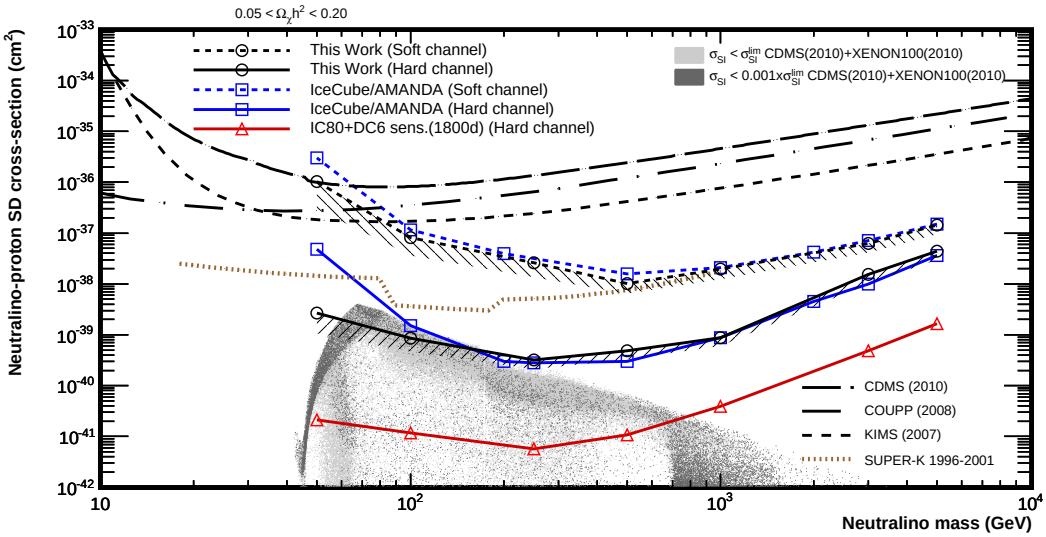


Figure 6.20: *In this picture we added the other bounds which were obtained by direct searches (CDMS, COUPP, KIMS) and by an indirect search for neutrinos in the Sun (Super-Kamiokande). The open red triangles indicate the foreseen sensitivity, for each neutralino model, which should be reached by DeepCore after 10 years of data-taking.*

I'm astounded by people who want to
"know" the Universe when it's hard enough
to find your way around Chinatown.

Woody Allen

(born Allan Stewart Königsberg)

7

Conclusion and outlook

According to several experimental observations, an unknown kind of matter should pervade our Universe, the so-called dark matter. According to modern Cosmology these particles should be stable, heavy and interact only gravitationally or weakly with matter. That is why particles like this are called WIMPs (Weakly Interactive Massive Particles).

One of the best WIMP candidate is the neutralino postulated in the Supersymmetric extension of the Standard Model of particle physics; it could constitute by itself the whole dark matter, or perhaps be one of its ingredients. If neutralinos were gravitationally trapped by the Sun, they could accumulate in its core and start self-annihilating. Hence, a possible way to reveal the neutralino dark matter is by detecting their Standard Model decay products, such as the neutrino. Then a possible excess of neutrinos is expected from the Sun direction over the atmospheric background.

We started this work with the aim to search for neutralino induced neutrinos from the center of the Sun. The data collected by the AMANDA neutrino telescope, which is located at the South Pole, from 2001 to 2006 were exploited for this work, from which we extracted around 812 days of detector live-time which suited the analysis peculiarity.

The main goal of our analysis was the indirect detection of the neutralino, but to reach this point a preparatory work was needed in order to clean the experimental data from the atmospheric background contamination. Due to the Sun position at the South Pole, we expected nearly horizontal tracks of low energy events, which constituted a real challenge for the reconstruction algorithm. Another important aspect was the presence of the string trigger which lowered the threshold to accept these events.

We optimised our event selection in a model depending way considering 14 different neutralino models, i.e. 7 masses, which according to collider and astrophysical constraints are in the range between 50 GeV to 5 TeV, and two annihilation channels which produce two extreme neutrino energy spectra, a soft and a hard one. In order to improve the event selection technique, to disentangle

the signal from the atmospheric background, thus ameliorating our sensitivity, we introduced a multivariate method which combines the information of several variables in a multidimensional phase space. This method improved a lot previous method based on simple one-dimension sequential cuts. The multivariate classifier used in this work was the Boosted Decision Trees (BDTs). We achieved a background reduction of about 10^{-7} and a signal efficiency of about 6% - 35% (depending on the neutralino mass and annihilation channel) with respect to the trigger level. The final experimental data sample contained from around 1000 to 2500 events, depending on the neutralino model optimisation.

The main objective of our analysis, as said before, was the search for an excess of neutrino events from the Sun direction. To evaluate the signal strength from the last remaining data, we adopted a method which exploits the shape information of a discriminating variable, the space angle between the Sun direction and the reconstructed track. This likelihood-ratio test provided confidence intervals in the number of signal events which were physically bounded.

The final outcome of our analysis was that in the final sample no statistically significant excess of events from the direction of the Sun was found. An upper limit with 90% of Confidence Level on the number of signal events was provided for the different neutralino model selections. These upper limits were exploited to infer an upper limit on the neutrino-to-muon conversion rate at the detector, the neutralino annihilation rate in the Sun, the neutrino induced muon flux at the detector and the spin-dependent and spin-independent neutralino-proton cross sections. The 90% upper limits on the muon flux span between $4.35 \cdot 10^2$ and $2.78 \cdot 10^4$ muons km^{-2} yr^{-1} for the hard annihilation channel and $6.46 \cdot 10^2$ to $4.72 \cdot 10^5$ muons km^{-2} yr^{-1} for the soft annihilation channel. The upper limits on the spin-dependent cross section obtained with this work span between 10^{-4} and 1 pb, and are better than those from the direct search experiments.

The above quoted numbers include all the systematic uncertainties. The main sources of uncertainty were the limited knowledge of the Optical Module sensitivity and the unknown effect in the description of the ice properties.

It should be noted that the inclusion of the string trigger was essential for improving the efficiency at the lowest neutralino masses, even if a good reconstruction of such events was pretty difficult.

Overall, our analysis performs very well; the results quoted for the low neutralino mass hard channel models and for all the soft channel models, are so far the most sensitive AMANDA/IceCube results on dark matter. The main reason is the use of the multidimensional method which allows a better separation between signal and background. A search for dark matter performed with IceCube (with only 22 strings) data could not reach the low energy region, since it had a higher energy threshold than AMANDA. Our results, compared with another dark matter analysis performed on AMANDA data from 2000 to 2006 (optimised

for high energy neutrinos), performs even better if we think that we had one year less statistics.

Future improvements on dark matter searches can be definitely achieved. One is the reduction of the systematic uncertainties related to the Optical Module sensitivity and to the ice properties. Especially for low neutralino mass models it is strongly demanded. There is a lot of ongoing work inside the AMANDA/IceCube collaboration, with the development and improvement of new software which goes in the good direction.

Our study of systematics was performed outside the likelihood-ratio frame, however the more appropriate way to settle systematics would be to include them during the confidence interval calculation to guarantee the frequentist coverage of the scaled limits.

As we have mentioned in the Introduction of this work, since 2009 AMANDA has finished its operations. IceCube is the successor in larger scale of AMANDA; its construction started in 2005 and its completion is planned for early 2011. It will consist of 4800 Digital Optical Modules (DOMs) installed on 80 strings between 1450 m and 2450 m below the ice surface, with an instrumented volume of about one-kilometre cube. IceCube was meant to detect high energy neutrino events, and the loss of sensitivity at low energies is compensated by an additional array of six densely instrumented strings, which is called DeepCore. The design of the six dense strings centred around one of the central IceCube strings, provides several advantages compared with AMANDA: 50 out of the 60 DOMs on a DeepCore string are installed in the deep clear ice between $2107 - 2450$ m, below the existing dust-layer, thus improving the reconstruction efficiency and angular resolution due to the longer scattering length for Cherenkov light. The top 6 DOMs in each of the six strings and three concentric rings of the nearest IceCube strings will form the so called "Veto Volume", required to reject the bulk of the downward-going muon background. DeepCore will give the possibility to observe neutrinos from above the horizon thus permitting the increase of the exposure time for neutrinos from dark matter annihilations up to the entire year. The expected sensitivities of DeepCore, after collecting 10 years of data, on the muon flux and on the spin-dependent cross-section, are shown in fig. 6.17 and fig. 6.19 as open red triangles.

As a final remark, we can mention that the filtered dataset used in the last stage of this analysis could still be exploited to pursue other WIMP candidates like the Lightest Kaluza-Klein Particle in the Universal Extra Dimension frame.

A

List and distribution of the observables

In appx. A.1 we will show the list of the 21 observables used to optimise the BDT, broken up in three categories (see sec. 5.4). The distributions of these observables before the BDT cut selection, for the 500 GeV hard neutralino model selection, are shown in appx. A.2. In appx. A.3 the same observables are shown after applying the BDT cut selection.

In the plots the shaded grey area represents the experimental data, the dashed line the neutralino signal, the solid line the atmospheric muon background and the dotted line the atmospheric neutrino background. The Monte Carlo signal and background simulations are normalised to the live-time of the detector.

The references to the observables, before and after the BDT cut, are between parenthesis in each item of appx. A.1. The distributions of the atmospheric background did not include the systematic uncertainties.

The Hit-reconstruction observables were calculated considering a cylinder around the reconstructed track with radius $\rho = 50$ m, and the hits were classified, according to their time residual, as:
early ($t_{res} \in [-550, -25]ns$), *direct* ($t_{res} \in [-25, 75]ns$), *late* ($t_{res} \in [75, 750]ns$)

A.1 Observable list for BDT optimisation

Reconstruction

1. The reconstructed log-likelihood (Pandel) theta angle θ_{LLH} (A.1, A.22)
2. The reconstructed first-guess JAMS theta angle θ_{JAMS} (A.2, A.23)
3. The internal parameter of the Direct Wimp first-guess σ_{ψ}^{DWimp}, which is a measure of the angular resolution of the solution (A.3, A.24)
4. The quality parameter of the JAMS first-guess Q_{JAMS}, which comes from Neural Network trained with some topological variables to separate high and low quality reconstruction (A.4, A.25)
5. The difference between the reduced log-likelihood (Pandel) and the reduced log-likelihood Bayesian reconstruction (A.5, A.26)

$$\Delta rLLH = rLLH_{Bayes} - rLLH_{LLH}$$

6. The Z_{JAMS} coordinate from the first-guess JAMS reconstructed vertex (A.6, A.27)
7. The variable of JAMS first guess σ_{ρ}^{JAMS}, which is the spread of the distance of the hits to the track in a selected cylinder (A.7, A.28)

Topology

8. The distance of the centre of gravity, of the hits selected, from the Z axis ρ_{COG} (A.8, A.29)
9. The distance of the centre of gravity, of the hits selected, from the centre of the axis r_{COG} (A.9, A.30)
10. The spread of the depth of the centre of gravity of the selected hits $\sigma_{Z_{COG}}$ (A.10, A.31)

Hit-reconstruction

11. The number of active strings N_{str}^{50} selected in a cylinder around the track (A.11, A.32)
12. The number of strings $N_{str_{dir}}^{50}$ with direct hits selected in a cylinder around the track (A.12, A.33)

13. The number of strings $N^{50}_{str_{late}}$ with late hits selected in a cylinder around the track (A.13, A.34)

14. The smoothness of the Leading Edge time LE_{smooth} (A.14, A.35)

15. The length of direct hits L_{dir}, i.e the projection of the hit OM (with direct hits) along the reconstructed track; the length is calculated taking the distance of the two outermost of the projected points (A.15, A.36)

16. The smoothness of the length of direct hits L_{smooth} (A.16, A.37)

17. The expected number of hits N^{phit}_{exp} based on the probability to detect a photon given the distance from the track (A.17, A.38)

18. The expected number of hits $N^{phit}_{postexp}$ based on the probability to detect a photon given the distance from the track after the last real hit (A.18, A.39)

19. The expected number of hits N^{phit}_{preexp} based on the probability to detect a photon given the distance from the track before one real hit (A.19, A.40)

20. The distance of the centre of gravity $\rho^{20}_{COG_{(meas-exp)}}$ to the track within a cylinder of radius 20 m. The centre of gravity was defined as the difference of the centre of the gravity of the measured hits minus the centre of gravity of the expected hits (A.20, A.41)

21. The square root of the ratio of the mean (weighted with the number of measured photoelectrons n^{meas}_{pe}) of the squared distance of the OM to the track over the mean (weighted with the number of expected photoelectrons n^{exp}_{pe}) of the squared distance of the OMs to the track $\sqrt{\frac{<\rho^2_{n^{meas}_{pe}}>}{<\rho^2_{n^{exp}_{pe}}>}}$ (A.21, A.42)

A.2 Observable distributions for BDT optimisation

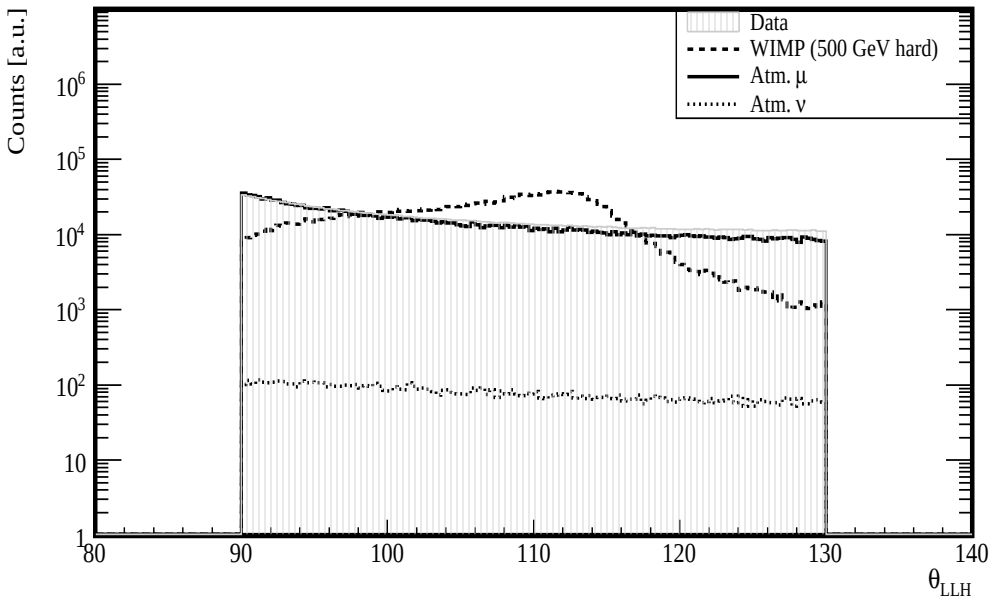


Figure A.1

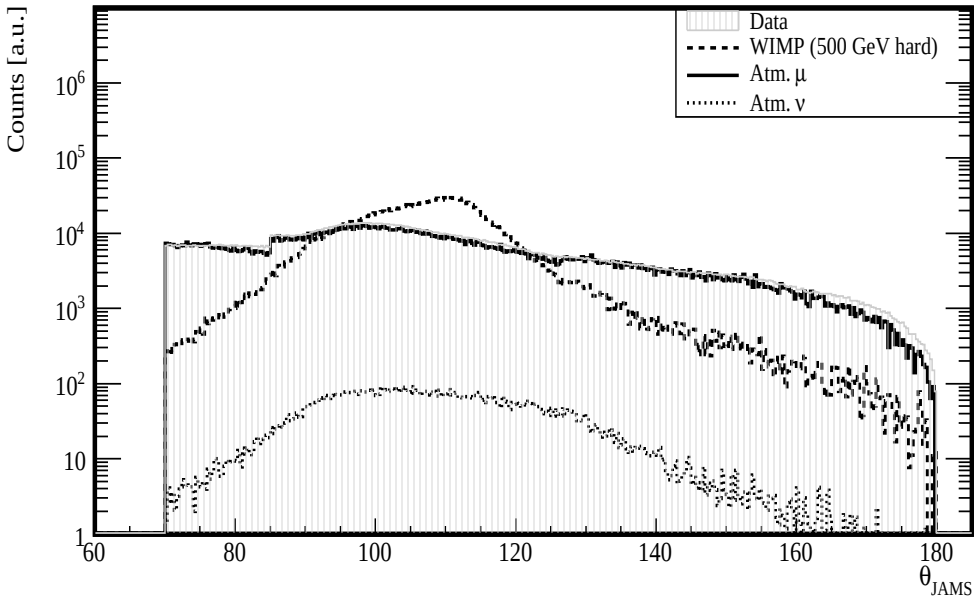


Figure A.2

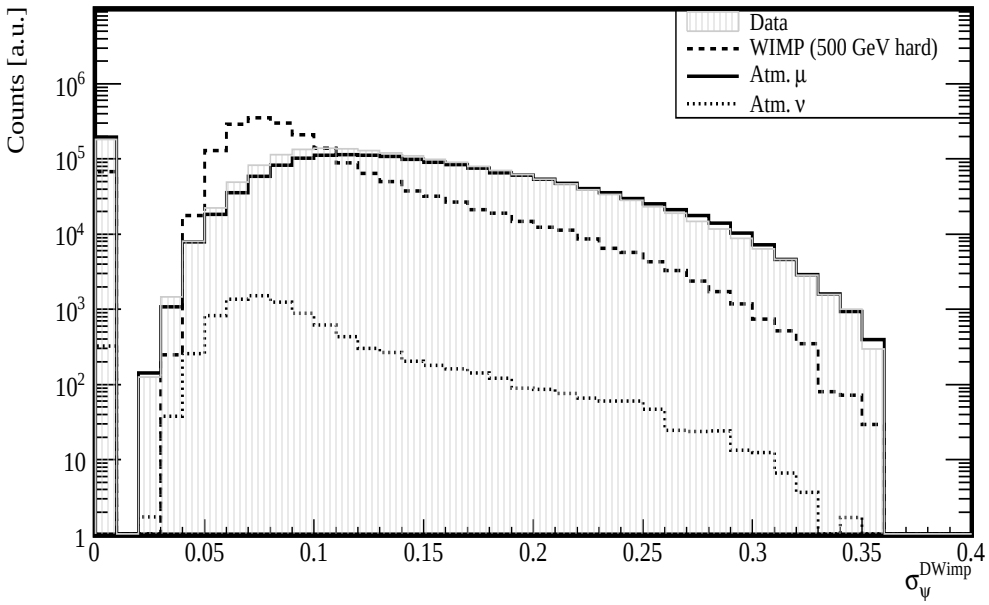


Figure A.3

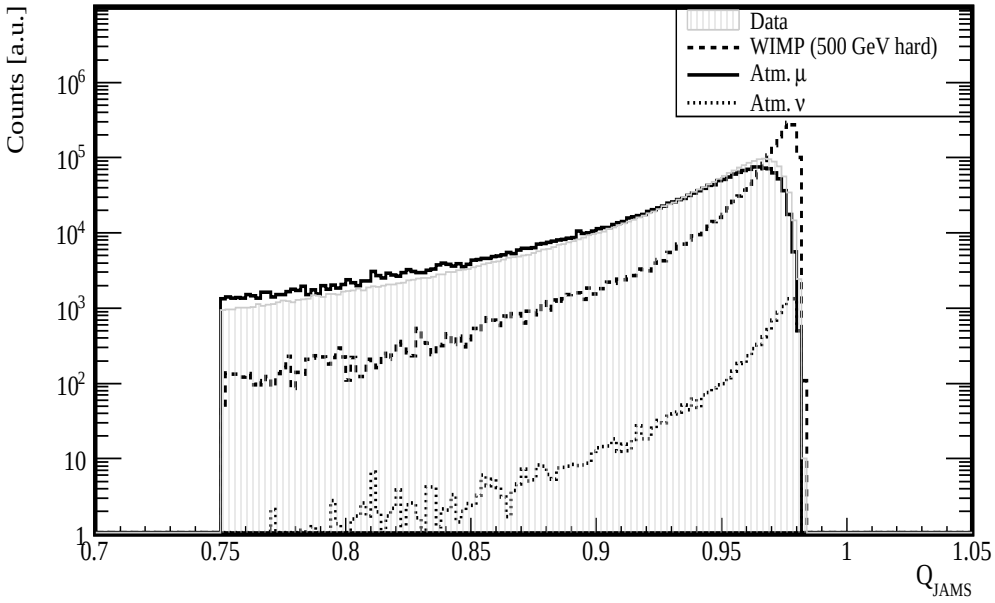


Figure A.4

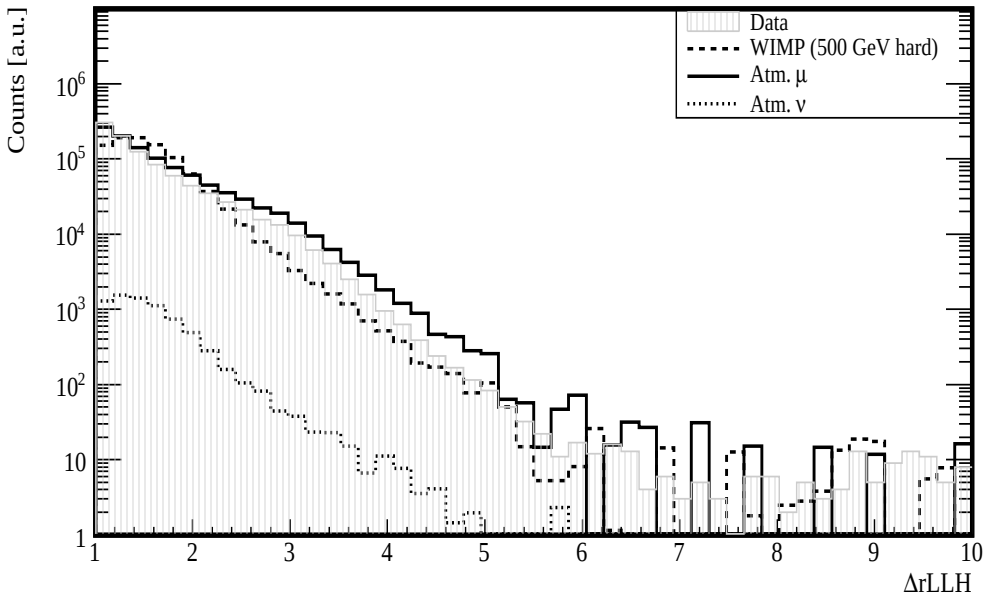


Figure A.5

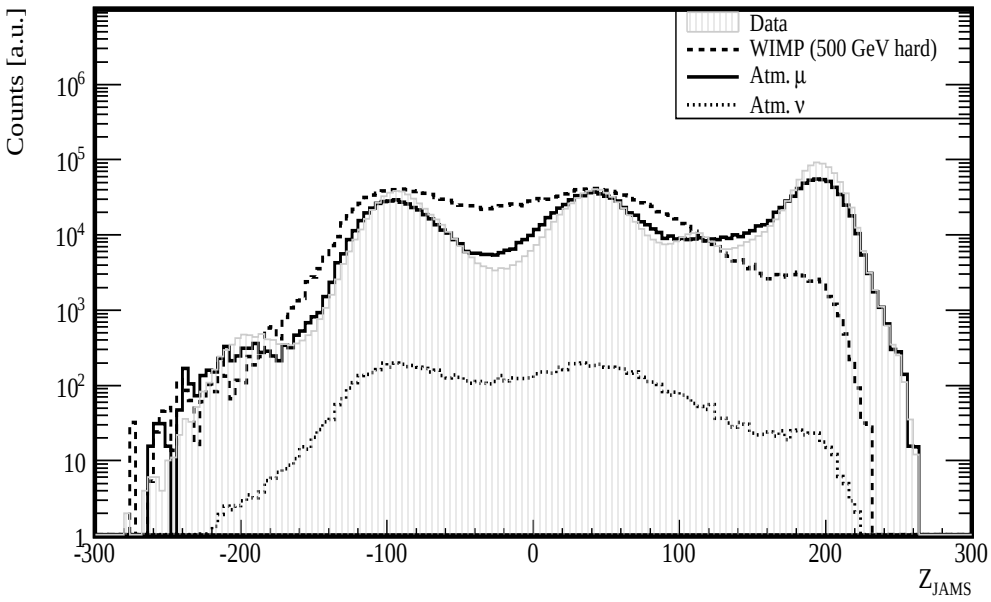


Figure A.6

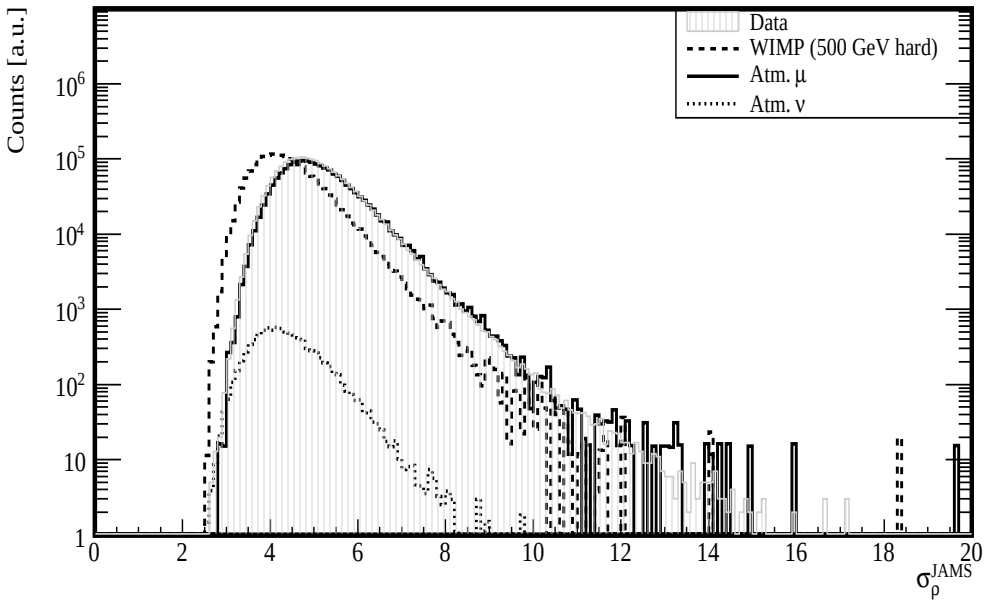


Figure A.7

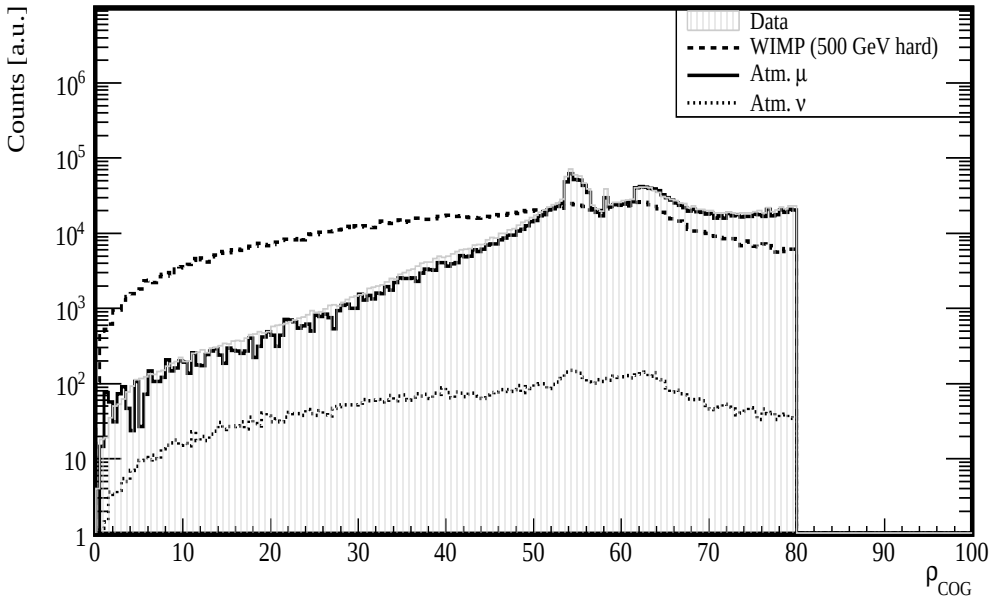


Figure A.8

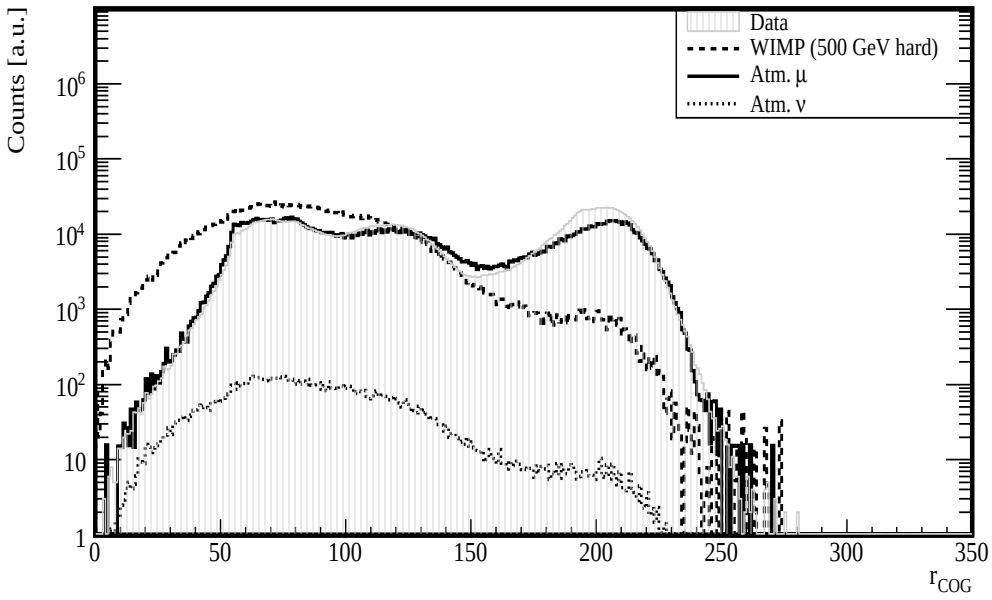


Figure A.9

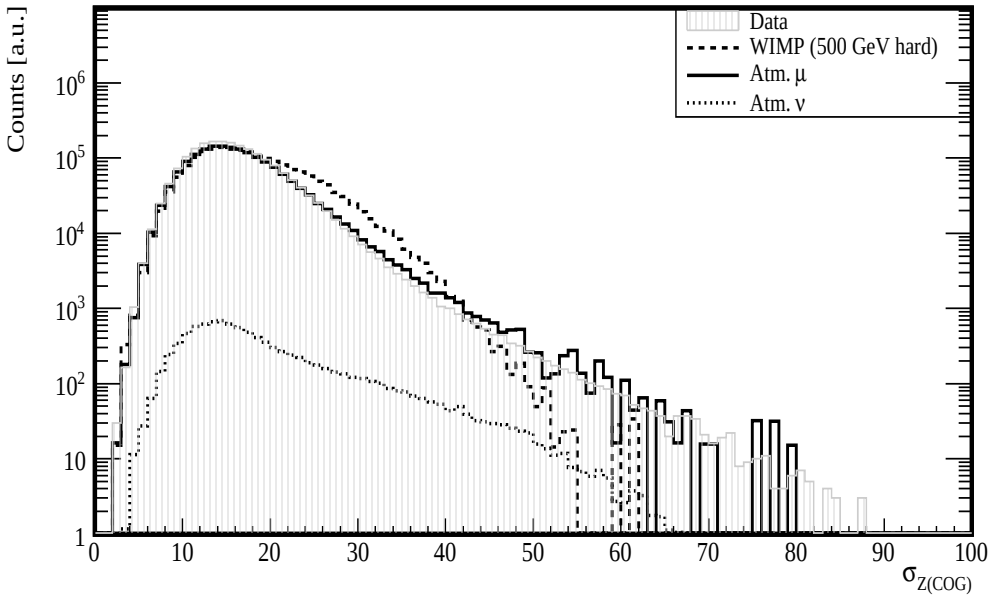


Figure A.10

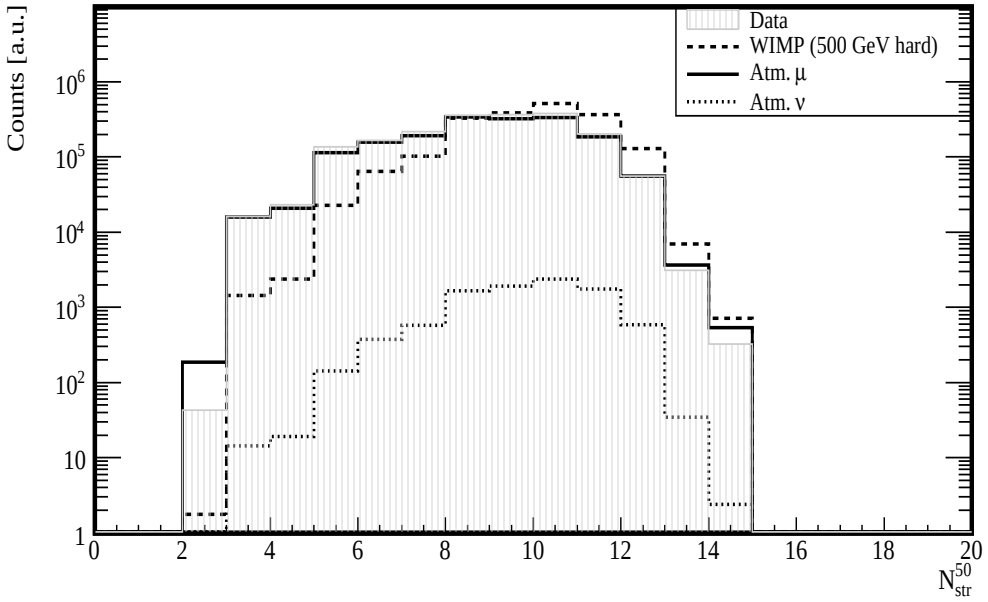


Figure A.11

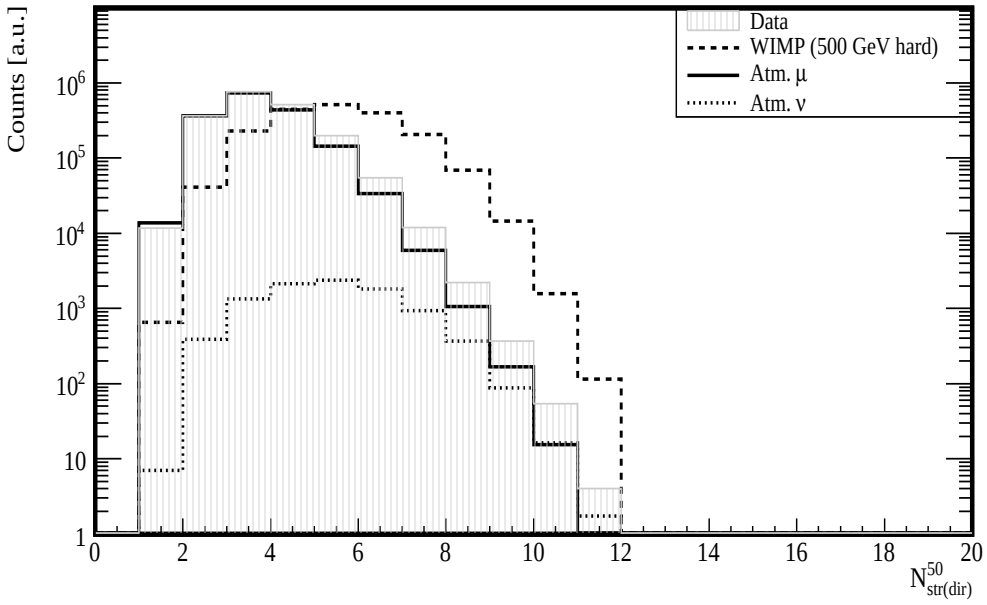


Figure A.12

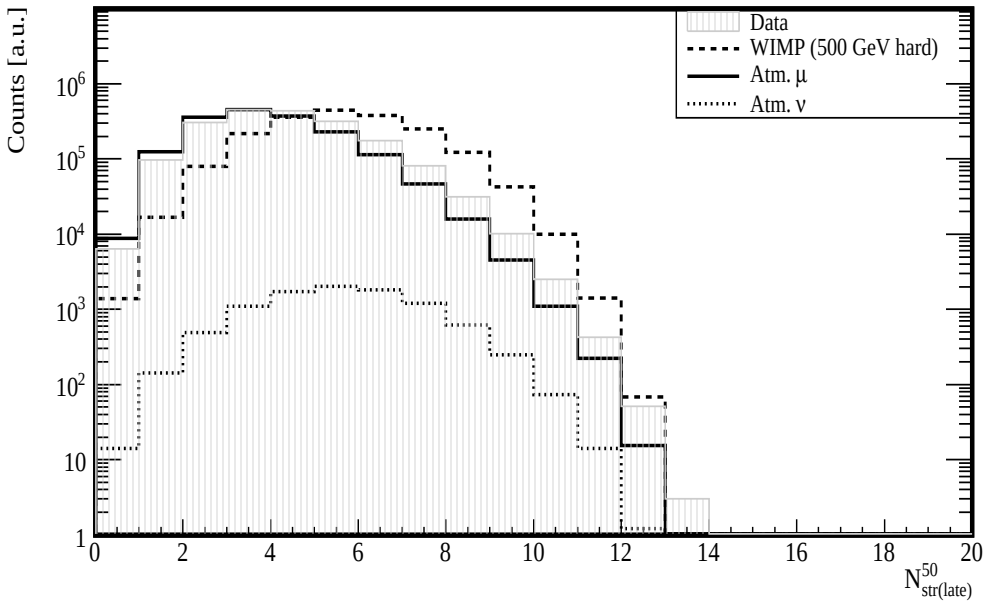


Figure A.13

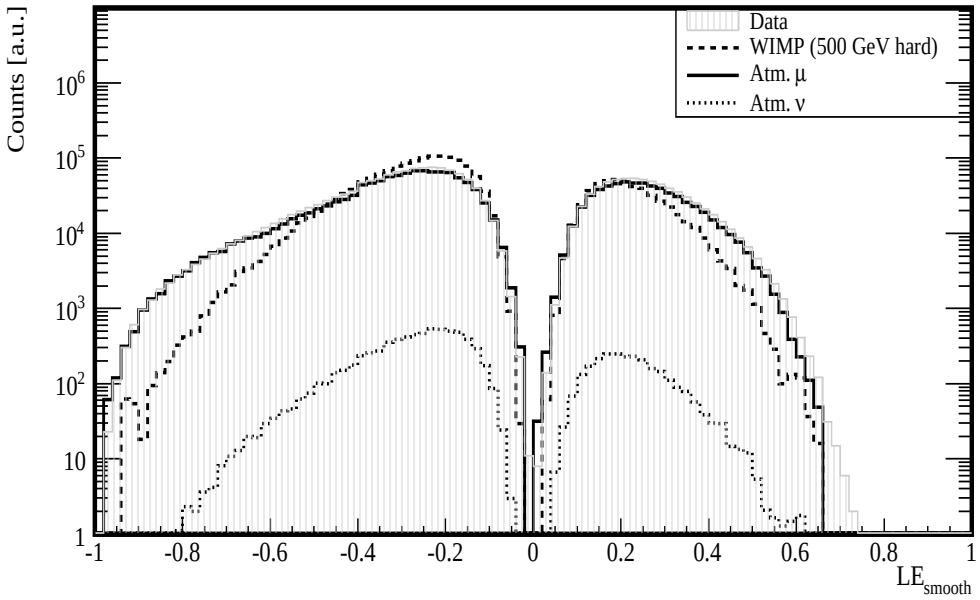


Figure A.14

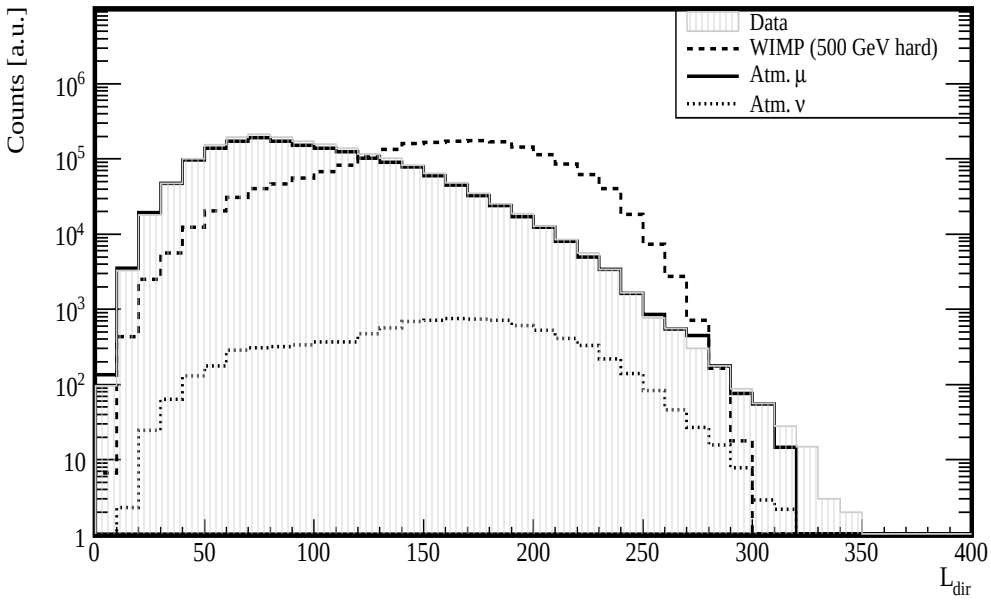


Figure A.15

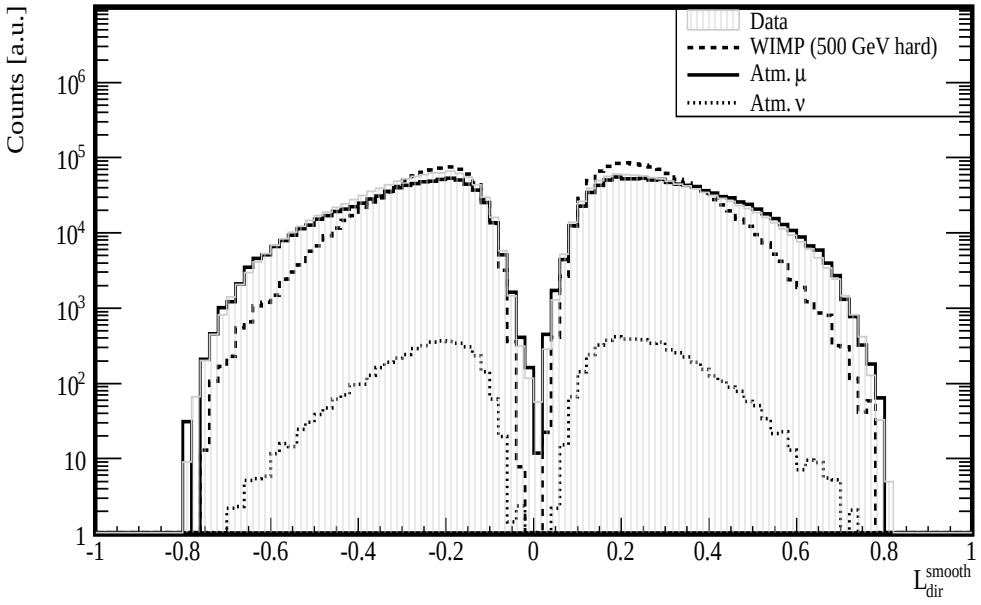


Figure A.16

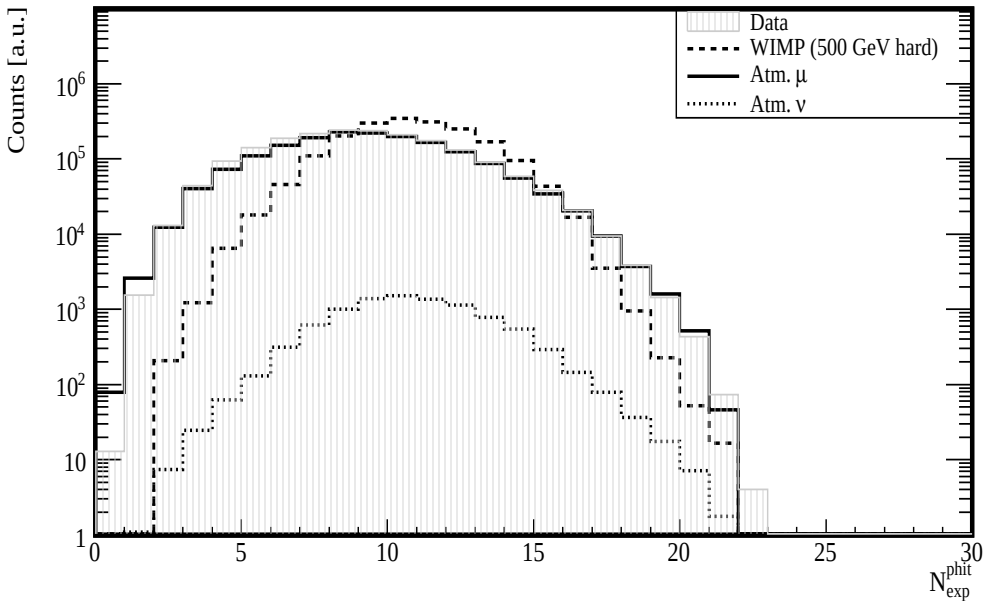


Figure A.17

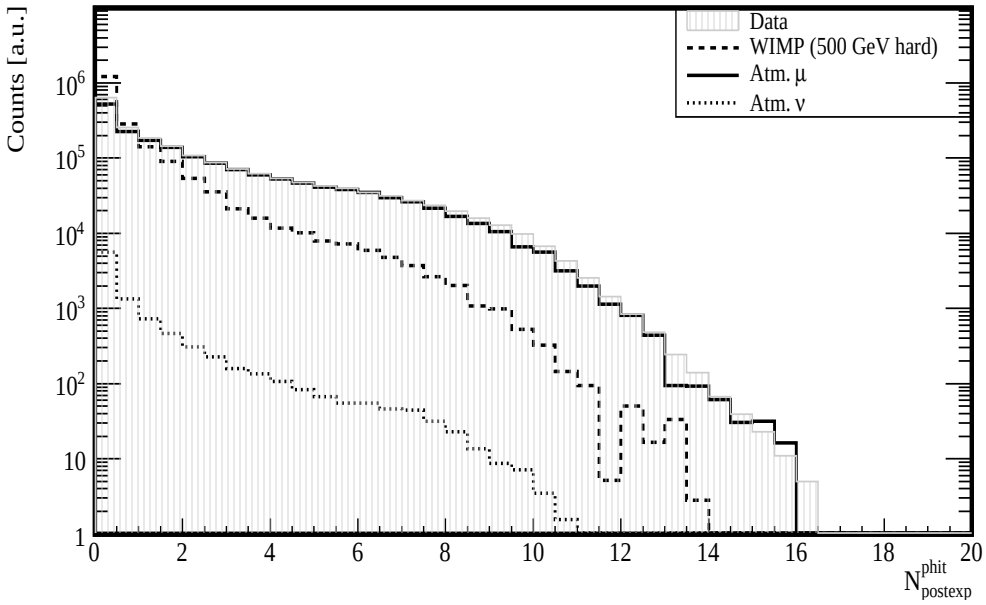


Figure A.18

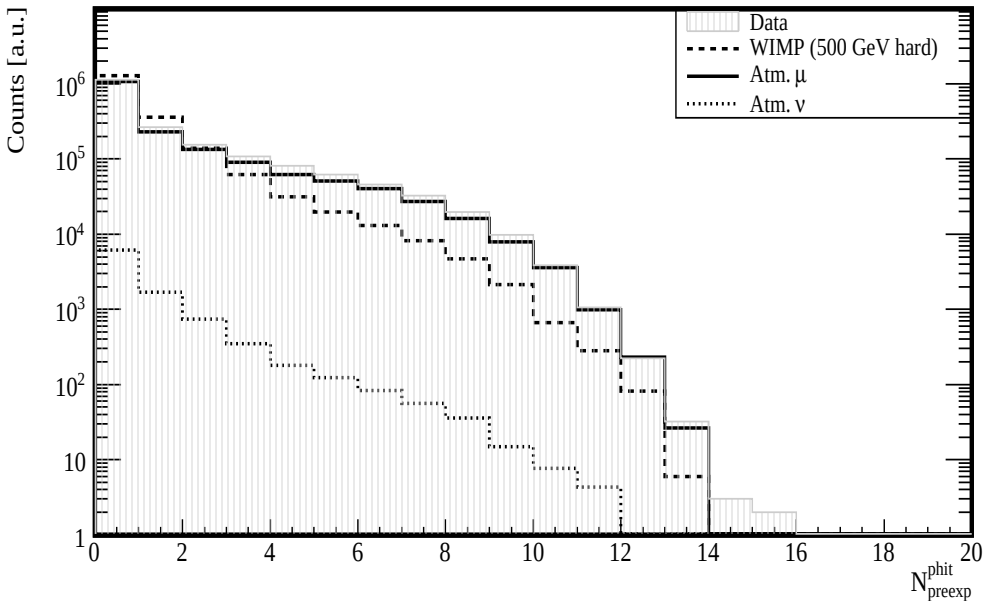


Figure A.19

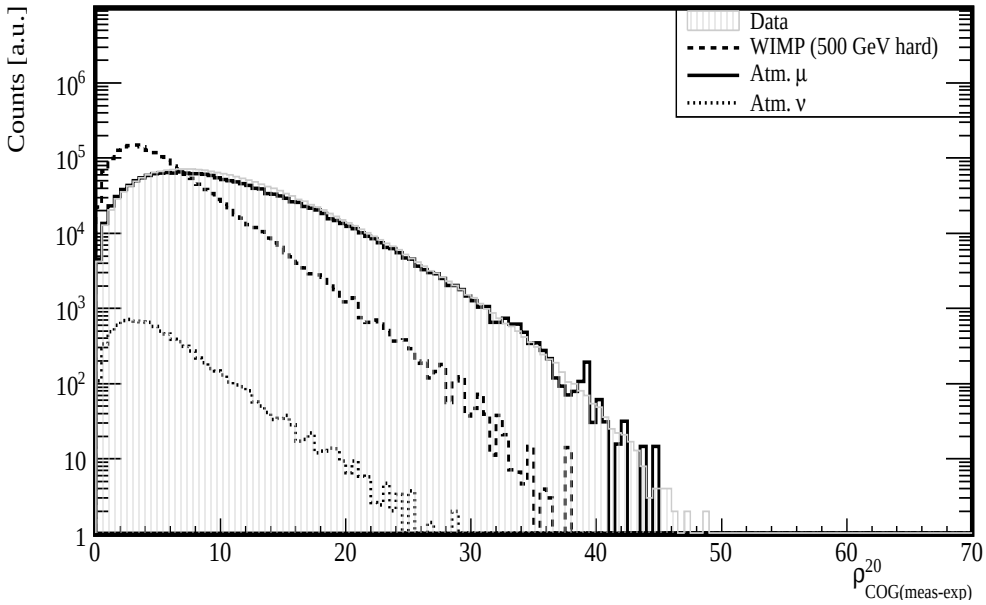


Figure A.20

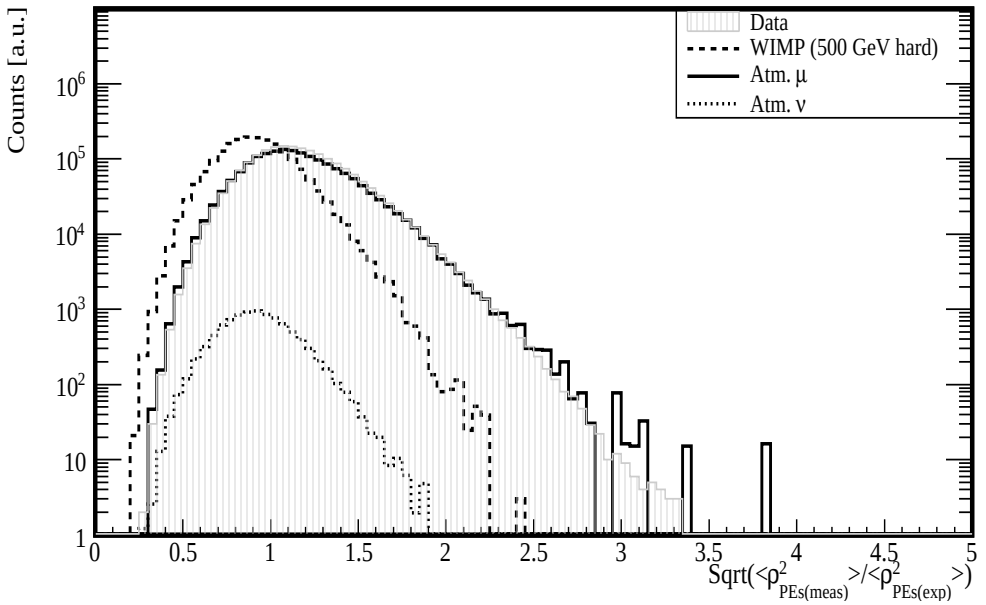


Figure A.21

A.3 Observable distributions after BDT cut

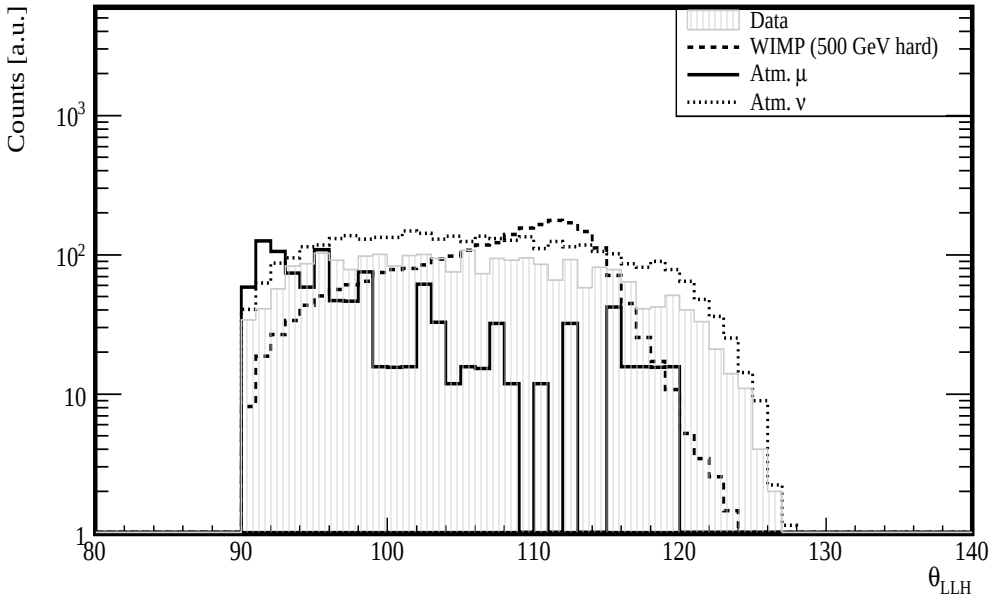


Figure A.22

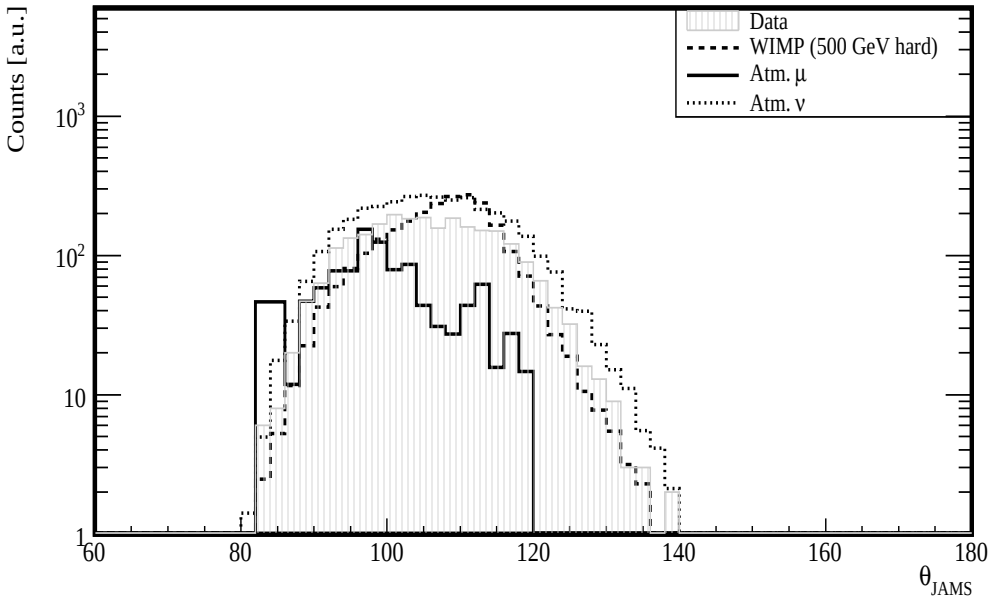


Figure A.23

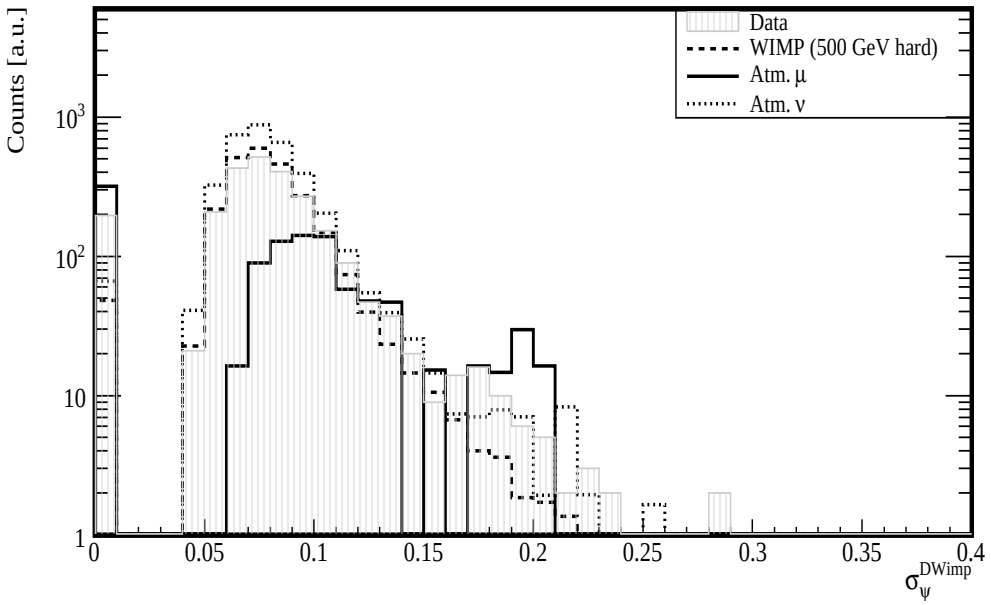


Figure A.24

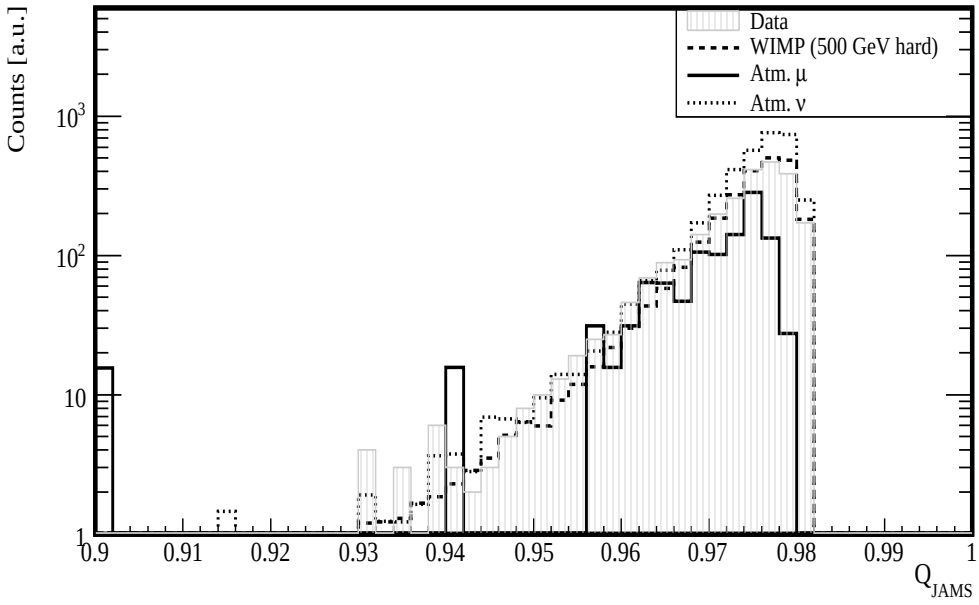


Figure A.25

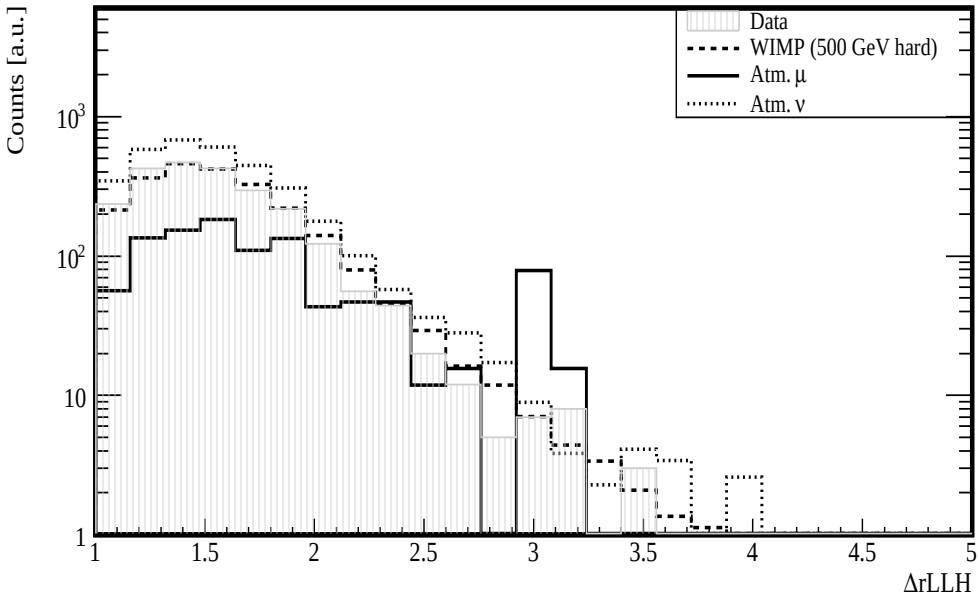


Figure A.26

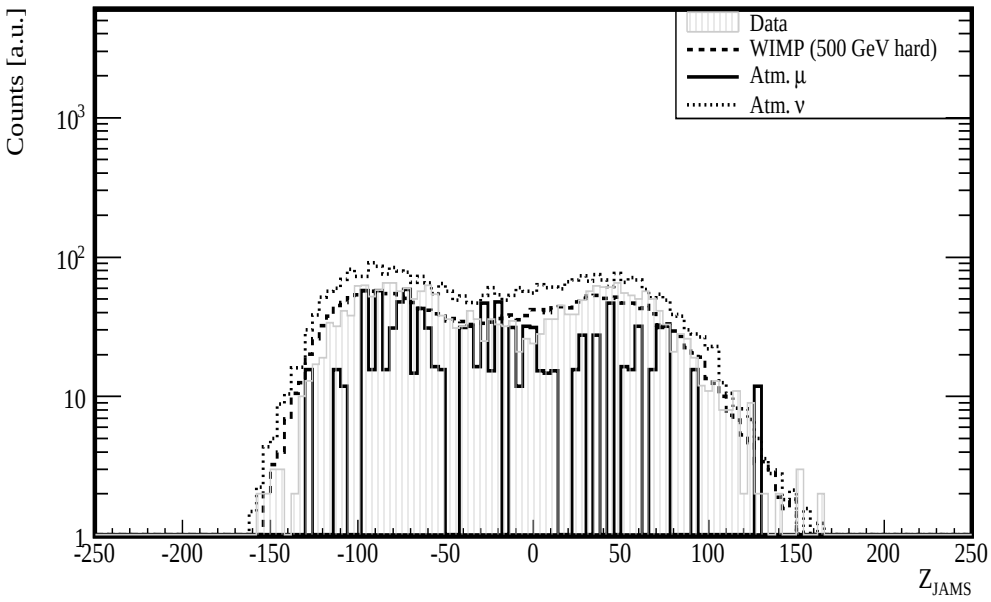


Figure A.27

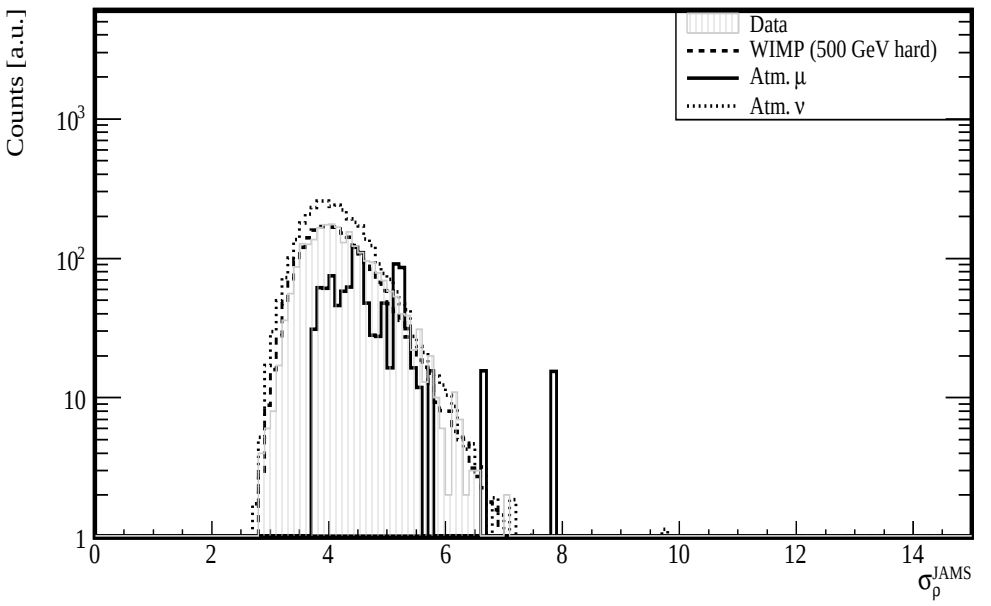


Figure A.28

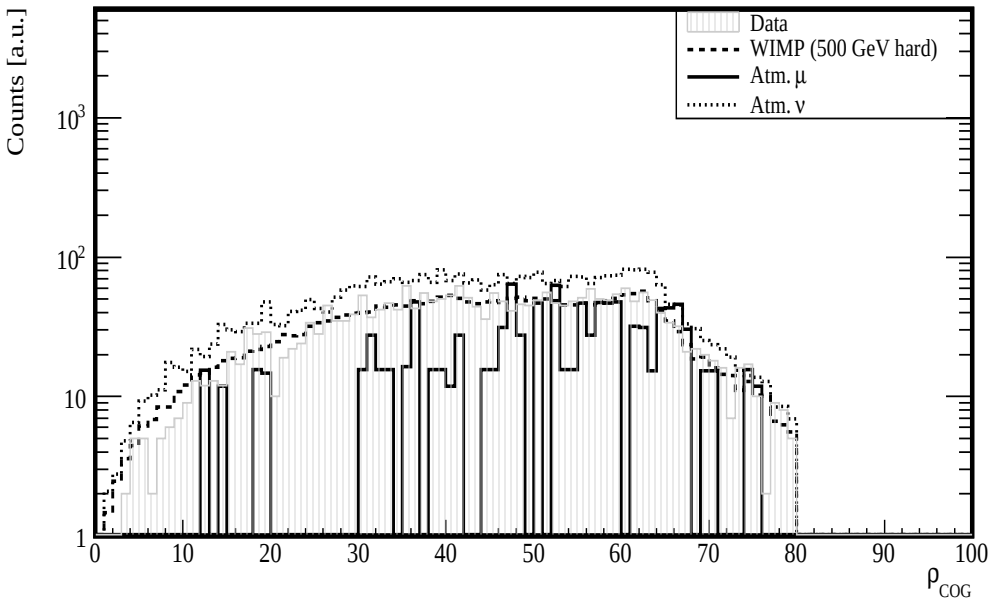


Figure A.29

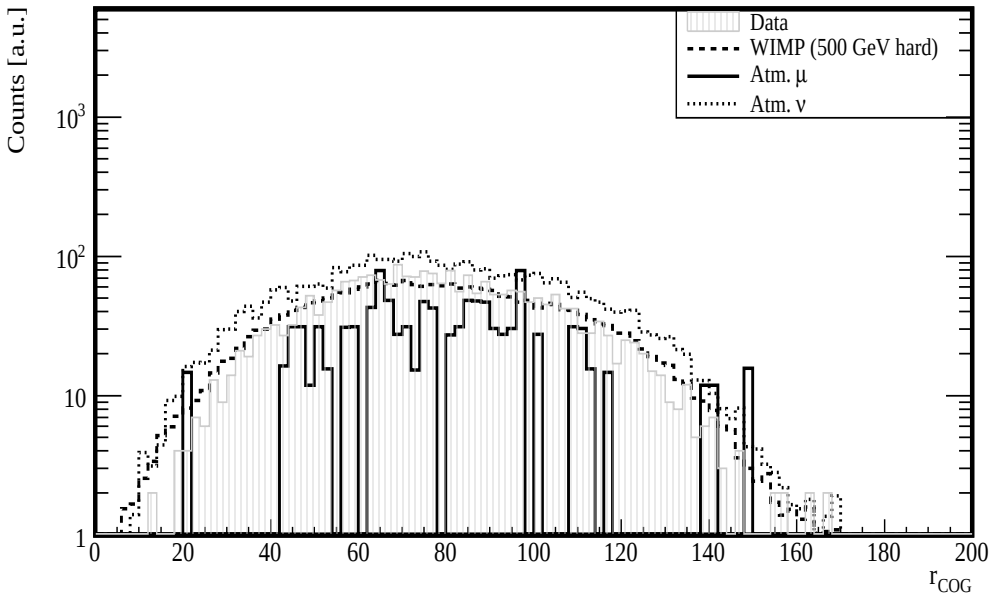


Figure A.30

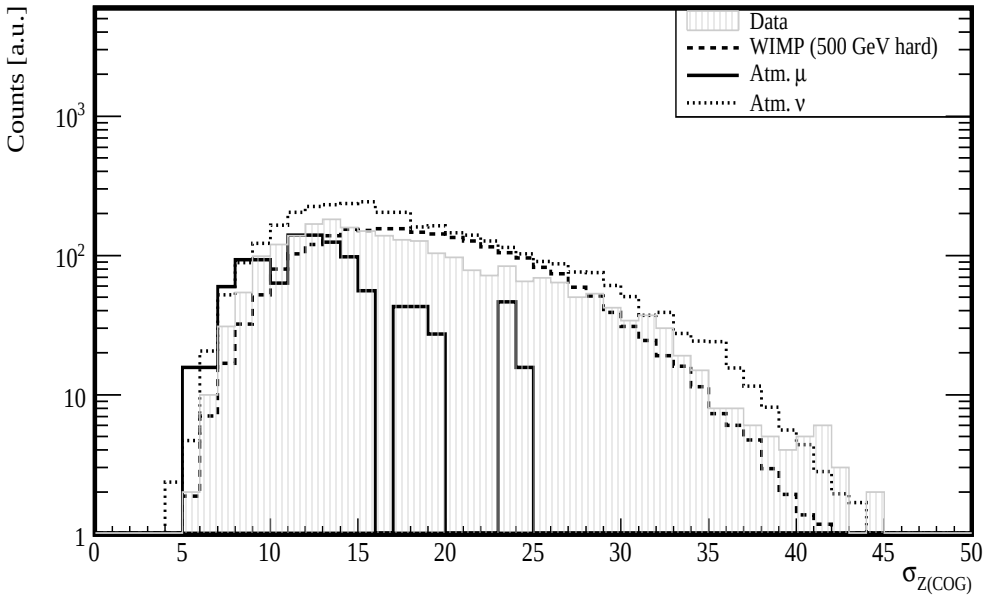


Figure A.31

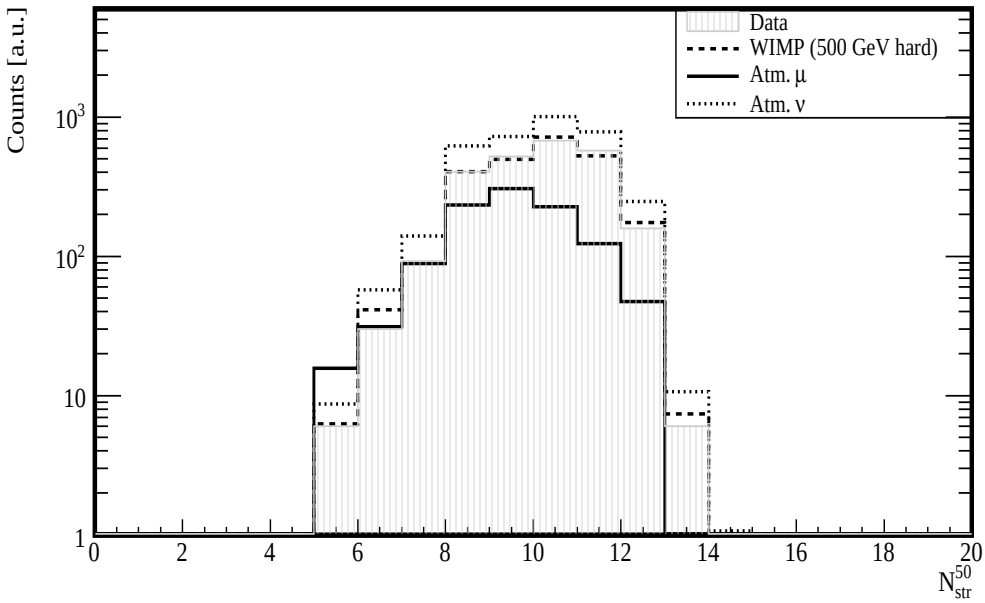


Figure A.32

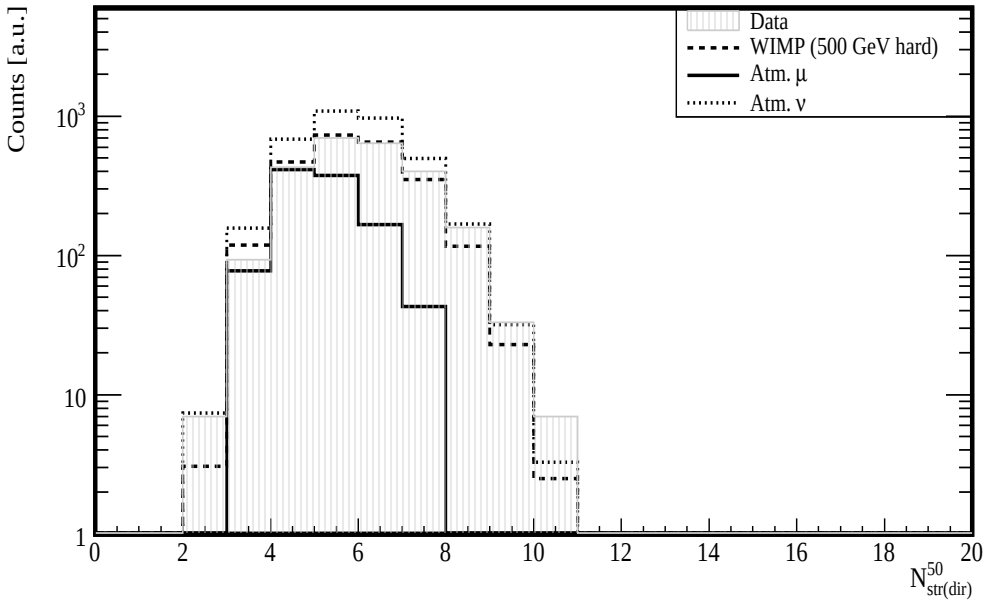


Figure A.33

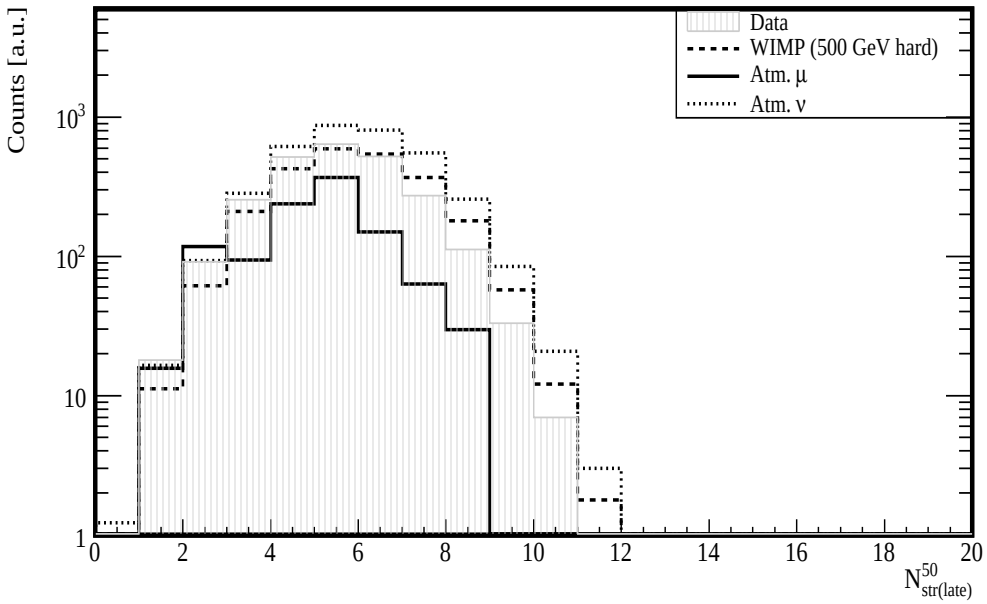


Figure A.34

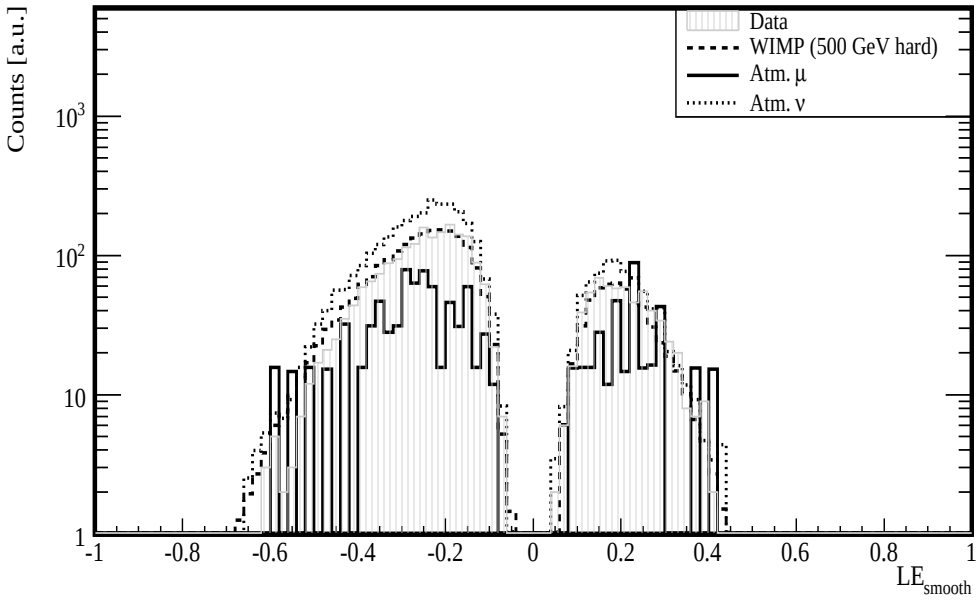


Figure A.35

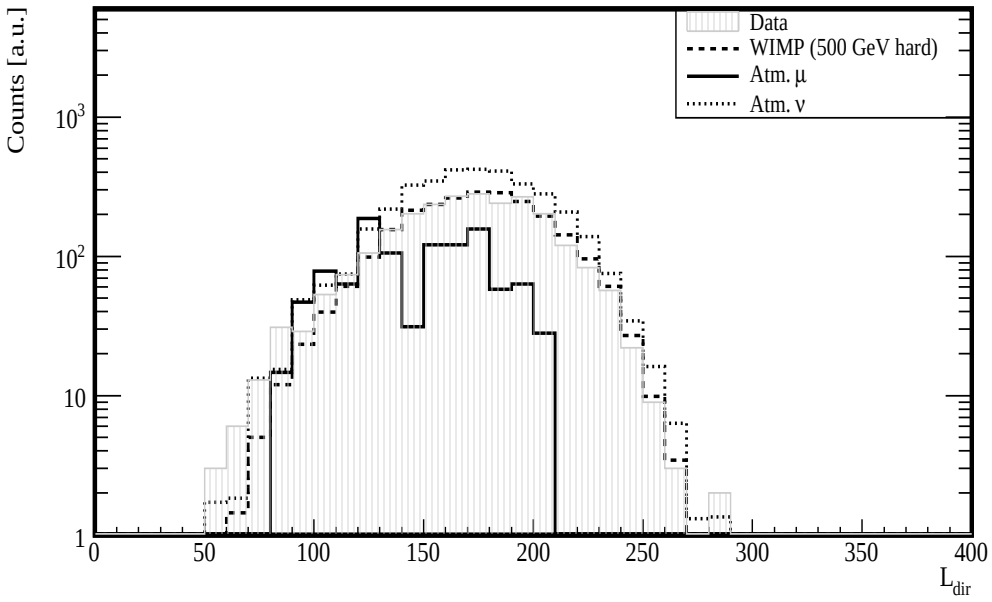


Figure A.36

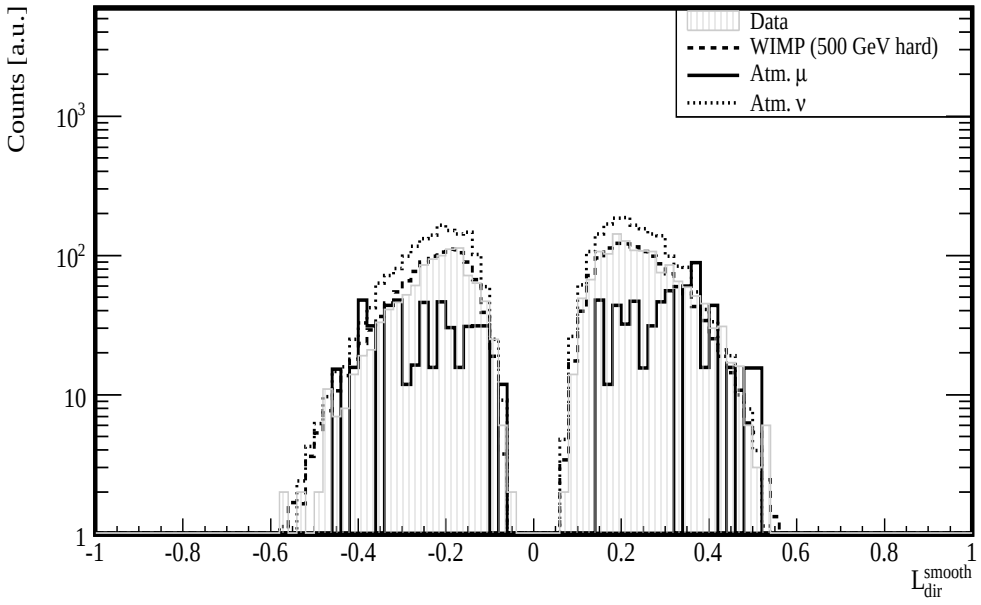


Figure A.37

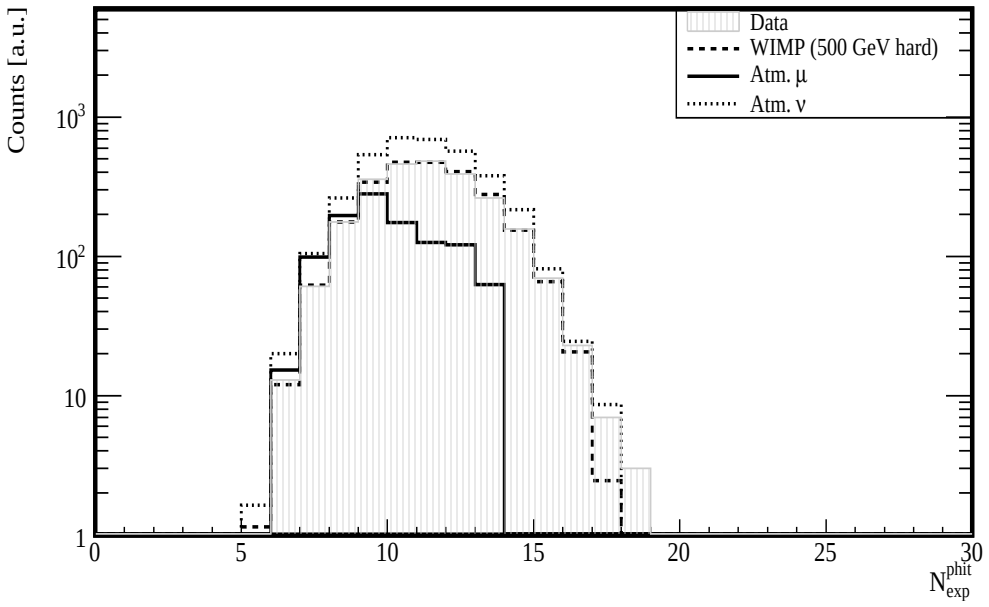


Figure A.38

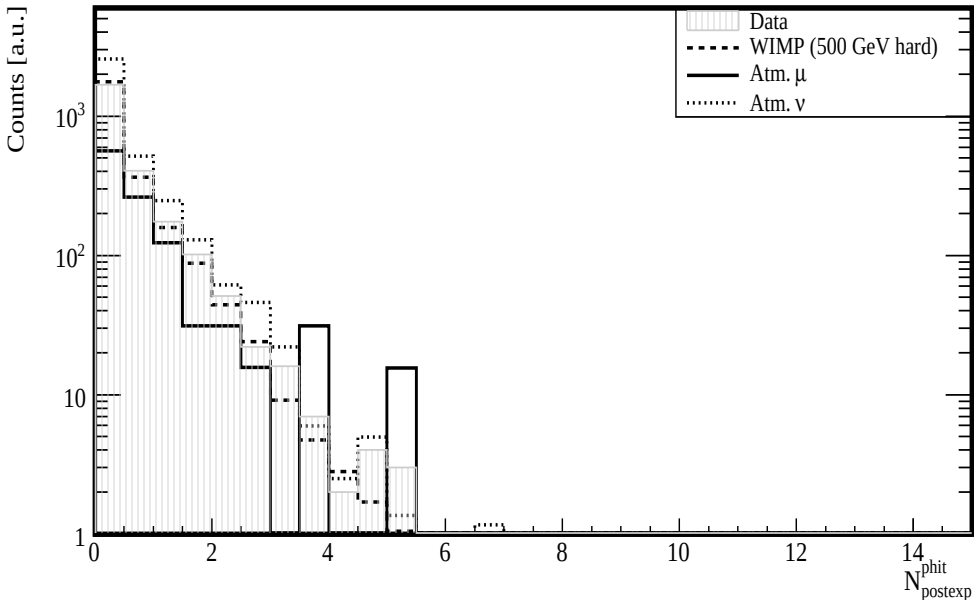


Figure A.39

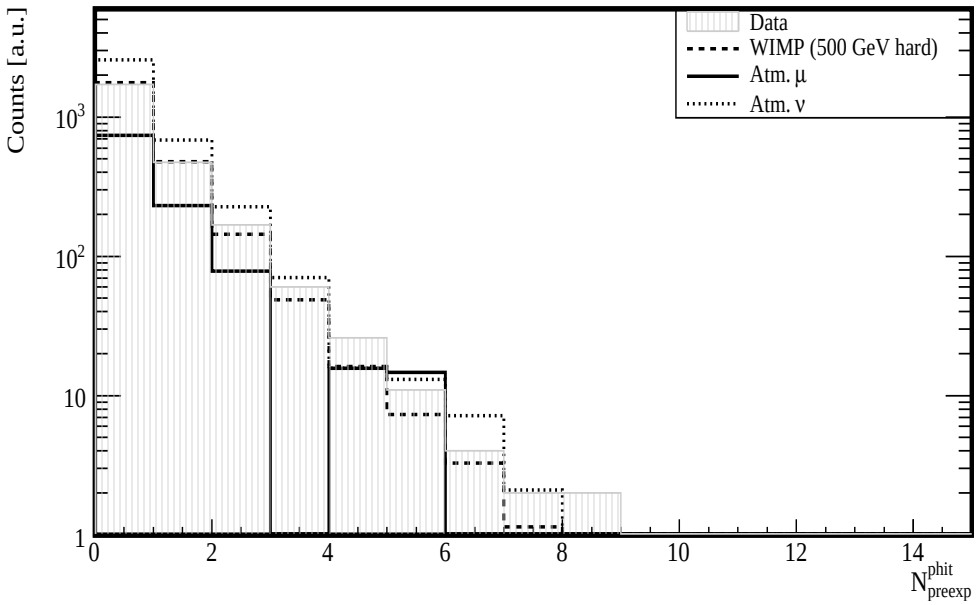


Figure A.40

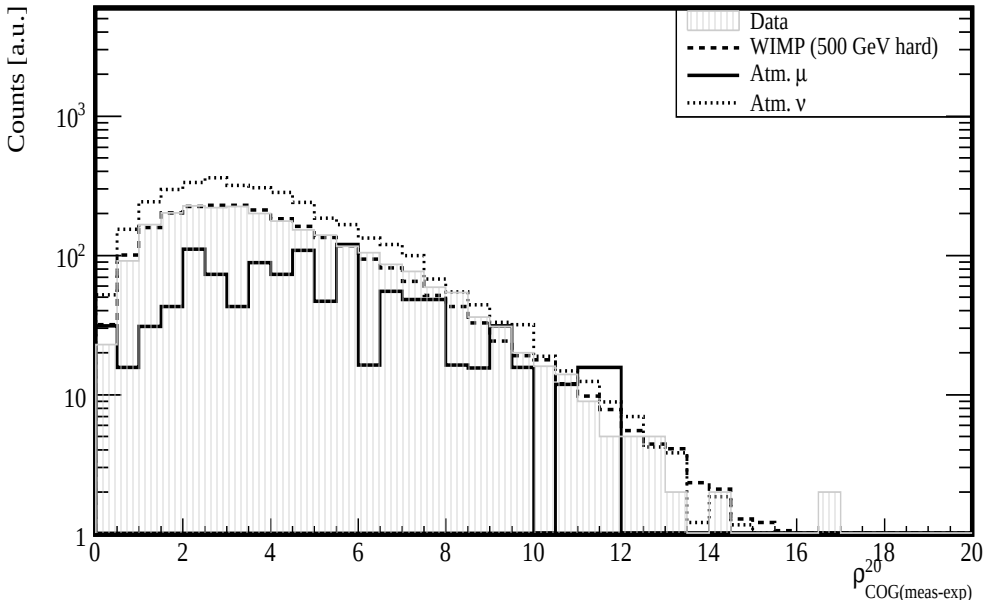


Figure A.41

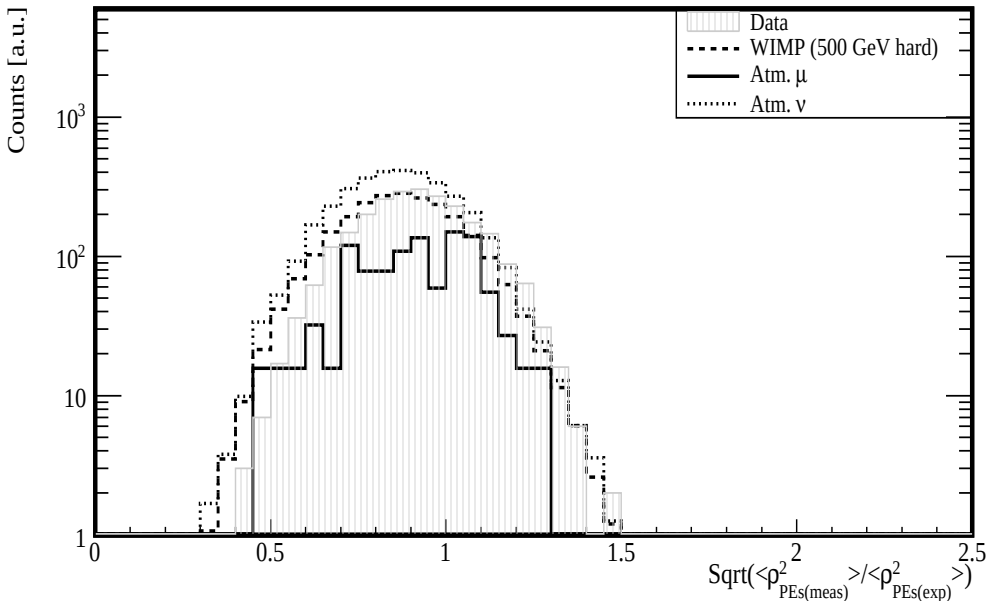


Figure A.42

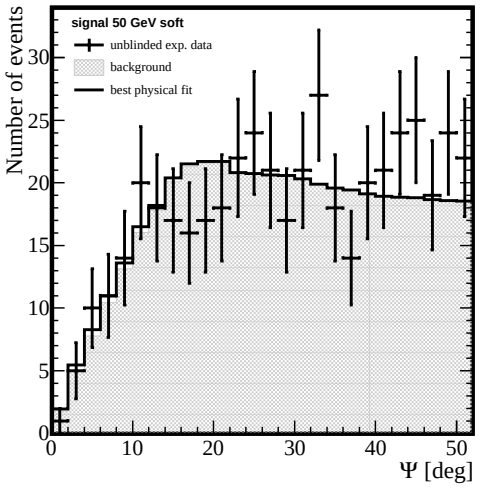
B

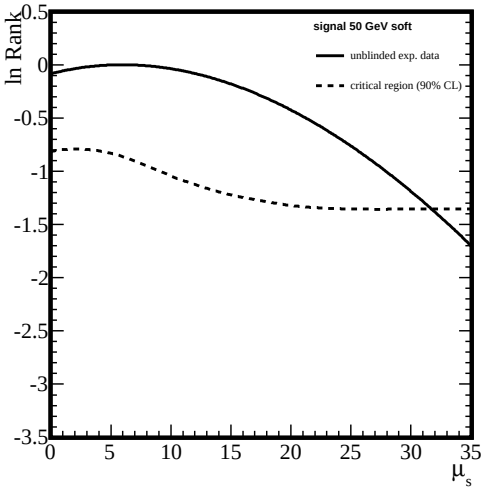
Space angle and log-likelihood ratio distributions

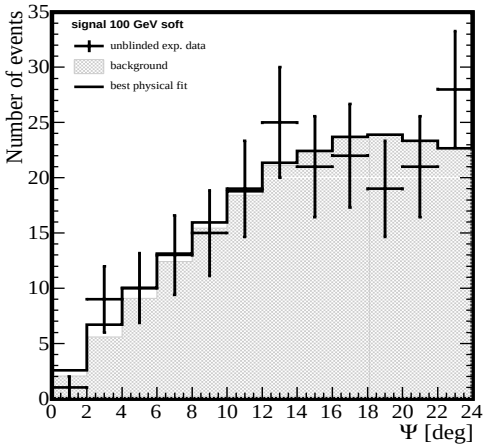
In this appendix, on the left side of next pages, we will show some space angle distributions in the vicinity of the Sun for all the neutralino model selections, in order from soft to hard channels. The space angle distributions of unblinded experimental data and expected background, and the best physical signal+background fit to the experiment are shown in the pictures; the latter two distributions are normalised to the total dataset.

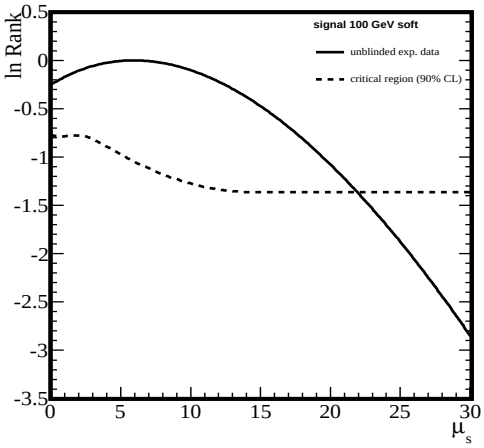
On the right side of the next pages, we will show the log-likelihood ratio distribution $\ln \mathcal{R}$, for the experiment (solid line), and the interceptions with the 90% CL critical region (dashed line), for each neutralino model selection.

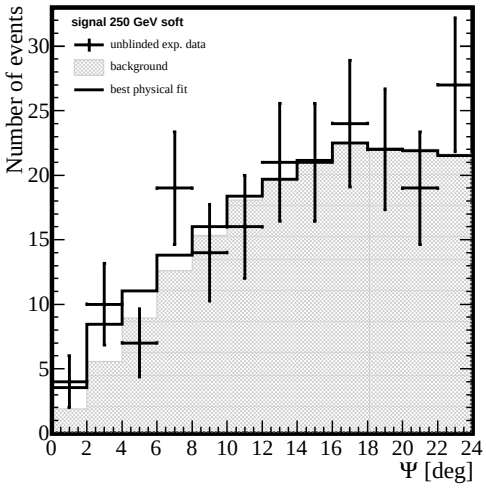
B. SPACE ANGLE AND LOG-LIKELIHOOD RATIO DISTRIBUTIONS

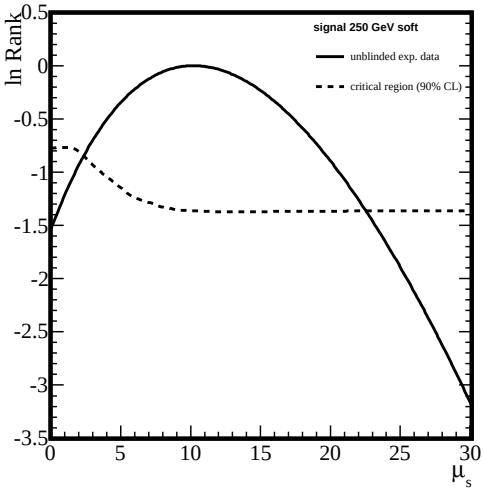


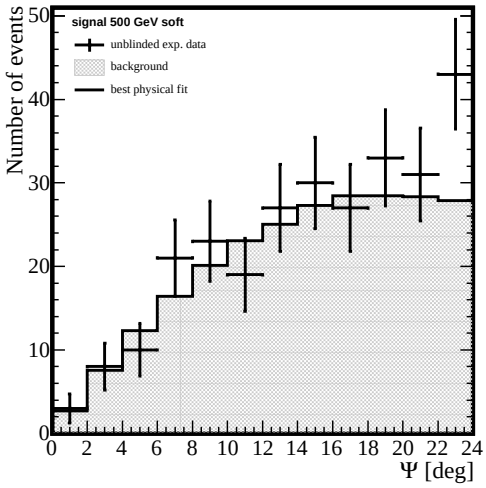


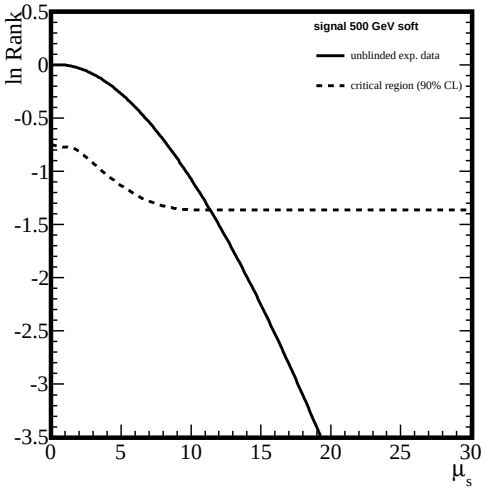


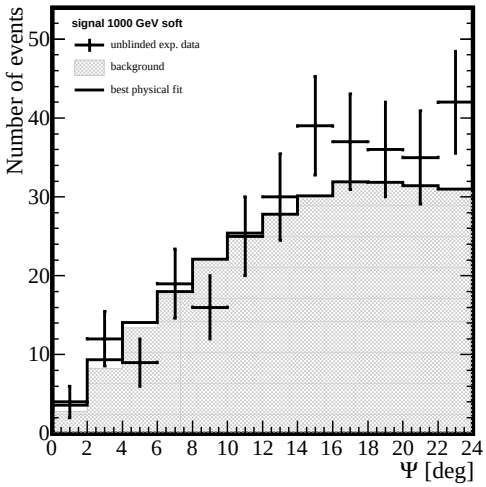


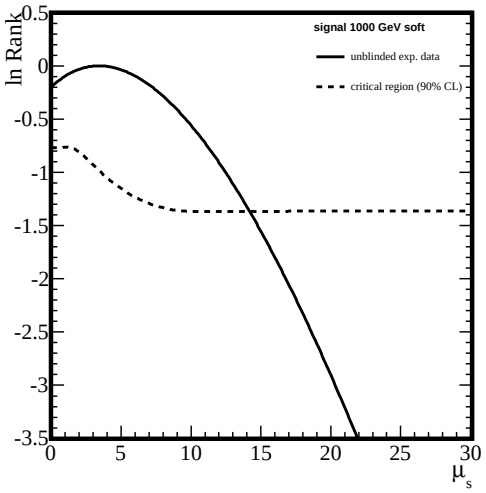


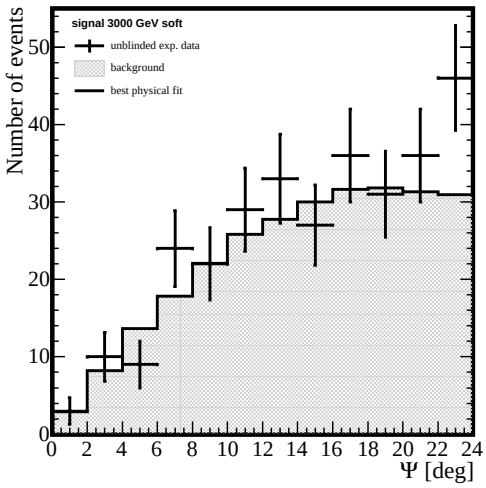


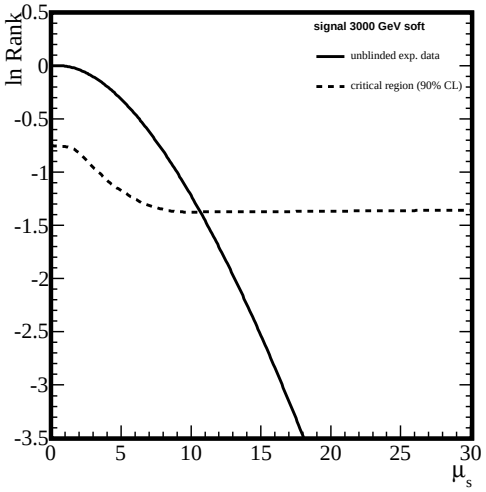

signal 500 GeV soft
unblinded exp. data
background
best physical fit
Number of events
Ψ [deg]


signal 500 GeV soft
unblinded exp. data
critical region (90% CL)
ln Rank
μs

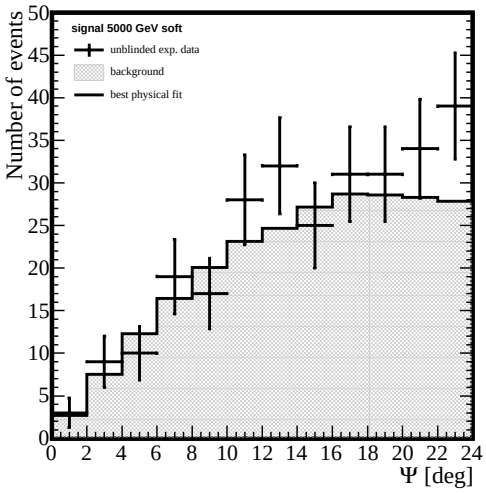

signal 1000 GeV soft
unblinded exp. data
background
best physical fit
Number of events
Ψ [deg]

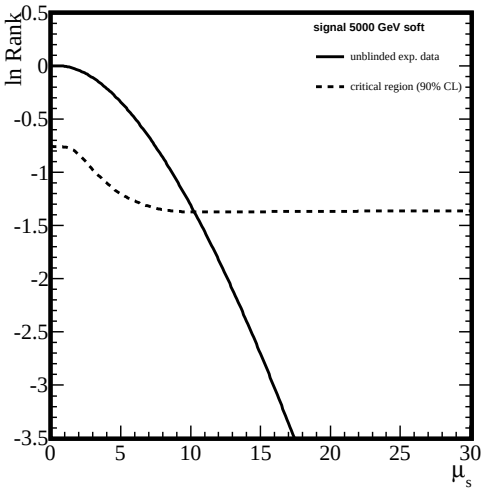

signal 1000 GeV soft
unblinded exp. data
critical region (90% CL)
ln Rank
μs

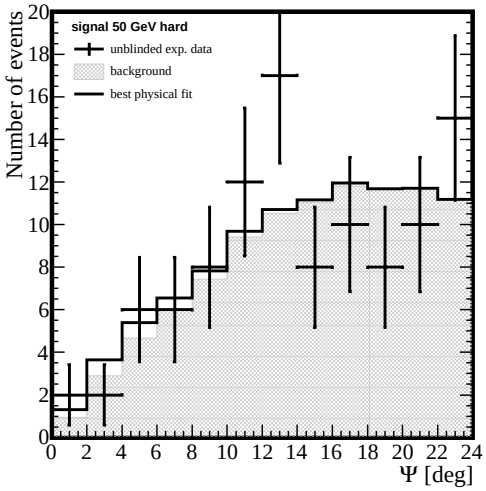

signal 3000 GeV soft
unblinded exp. data
background
best physical fit
Number of events
Ψ [deg]

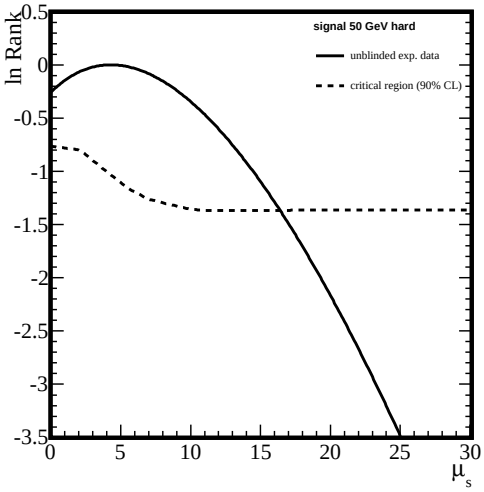

signal 3000 GeV soft
unblinded exp. data
critical region (90% CL)
ln Rank
μs

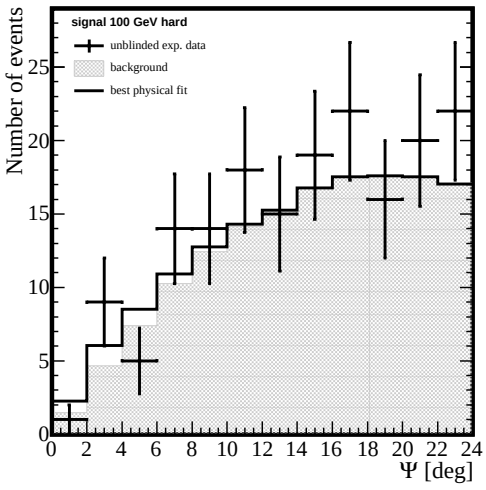
B. SPACE ANGLE AND LOG-LIKELIHOOD RATIO DISTRIBUTIONS

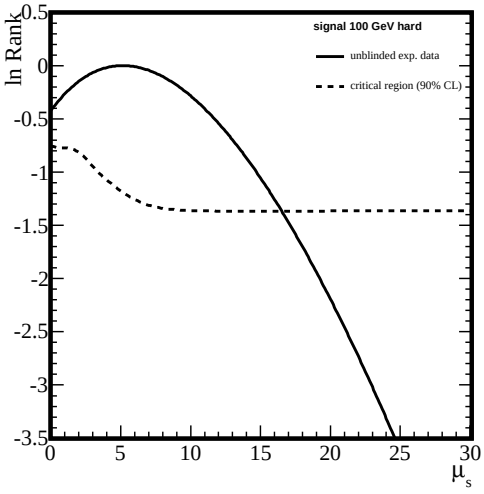


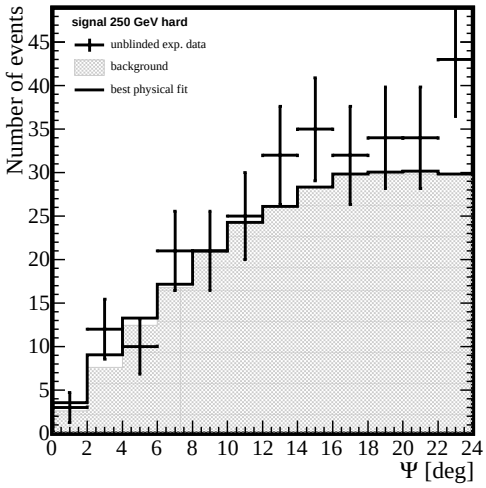


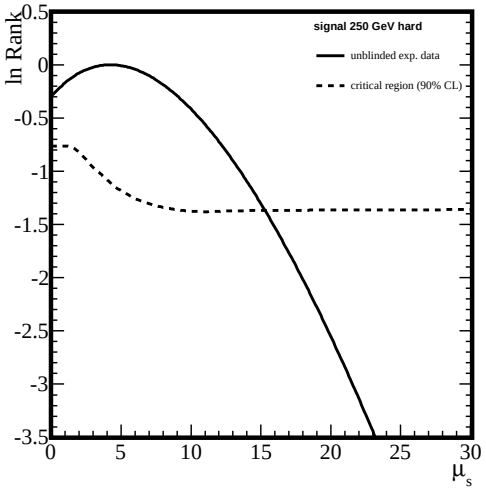


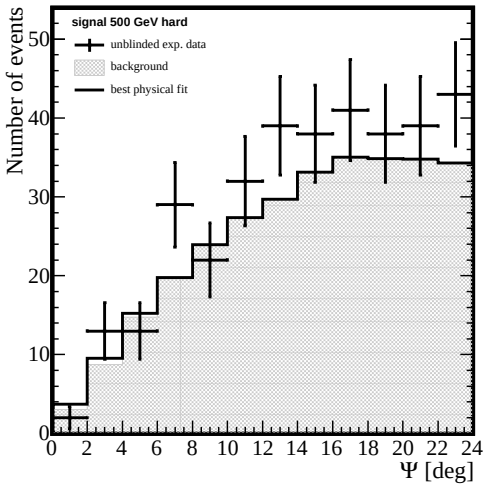


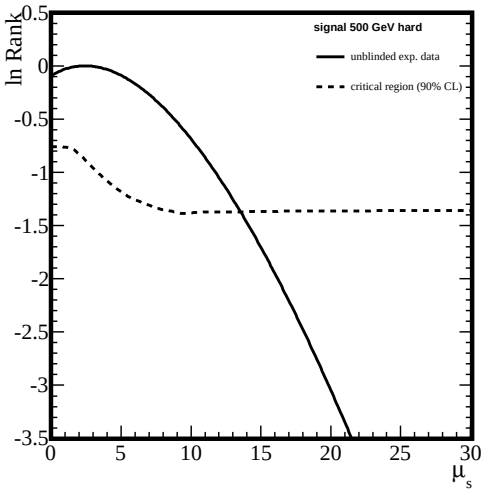


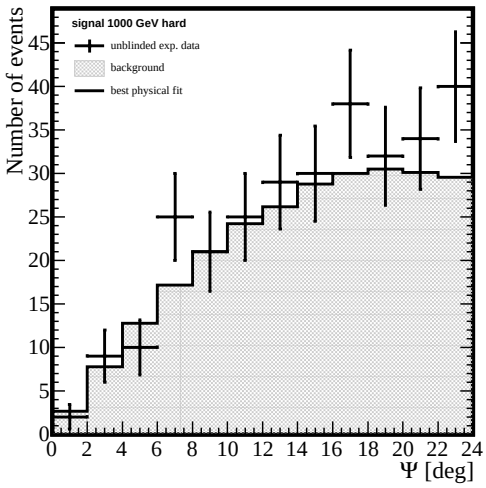


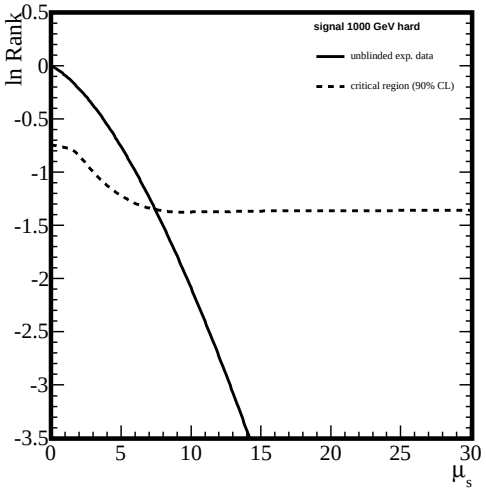

signal 250 GeV hard
unblinded exp. data
background
best physical fit
Number of events
Ψ [deg]

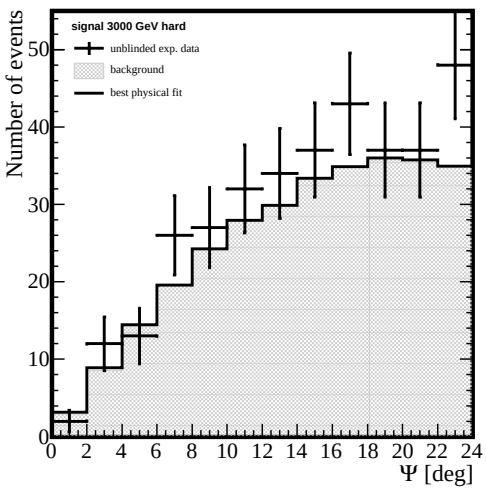

signal 250 GeV hard
unblinded exp. data
critical region (90% CL)
ln Rank
μs

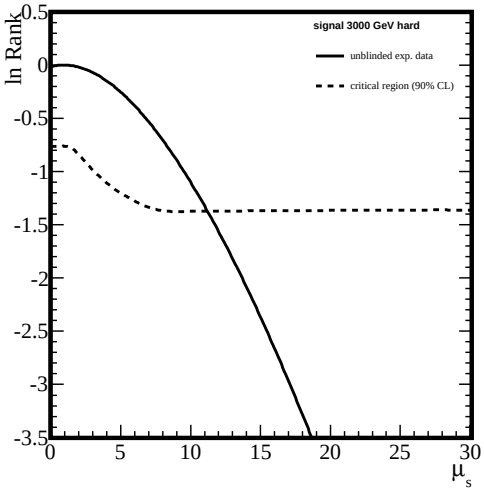

signal 500 GeV hard
unblinded exp. data
background
best physical fit
Number of events
Ψ [deg]

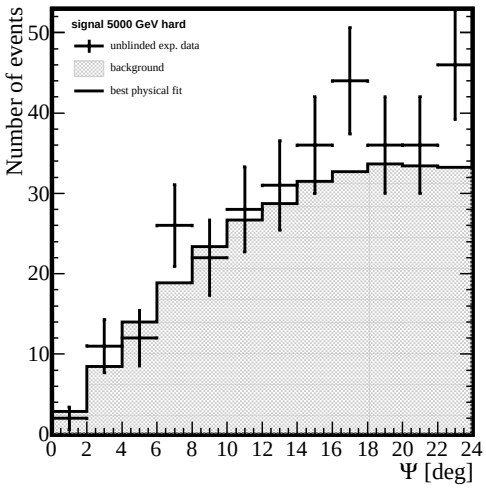

signal 500 GeV hard
unblinded exp. data
critical region (90% CL)
ln Rank
μs

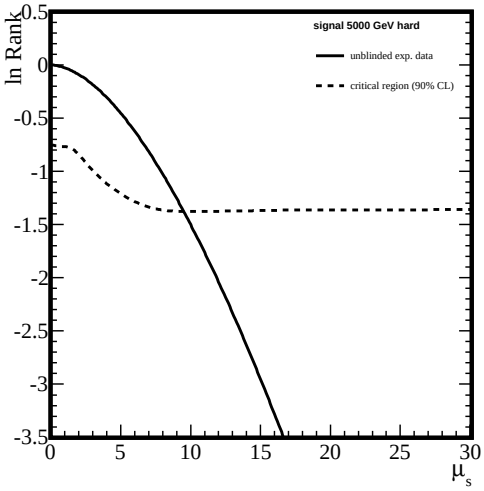

signal 1000 GeV hard
unblinded exp. data
background
best physical fit
Number of events
Ψ [deg]


signal 1000 GeV hard
unblinded exp. data
critical region (90% CL)
ln Rank
μs


signal 3000 GeV hard
unblinded exp. data
background
best physical fit
Number of events
Ψ [deg]


signal 3000 GeV hard
unblinded exp. data
critical region (90% CL)
ln Rank
μs


signal 5000 GeV hard
unblinded exp. data
background
best physical fit
Number of events
Ψ [deg]


signal 5000 GeV hard
unblinded exp. data
critical region (90% CL)
ln Rank
μs

References

[1] Hubble, E. A relation between distance and radial velocity among extra-galactic nebulae. In *Proceedings of the National Academy of Science*, 15, 168–173 (1929). 5

[2] Komatsu, E. *et al.* Five-Year Wilkinson Microwave Anisotropy Probe Oberservations: Cosmological Interpretation. *Astrophys. J. Suppl. Ser.* **180**, 330–376 (2009). 5, 7, 11

[3] Bergström, L. & Goobar, A. *Cosmology and Particle Astrophysics* (Praxing Publishing Ltd, 2004). 5

[4] Jungman, G., Kamionkowski, M. & Griest, K. Supersymmetric dark matter. *Phys. Rept.* **267**, 195 (1996). 6, 9

[5] Bertone, G., Hooper, D. & Silk, J. Particle dark matter: Evidence, candidates and constraints. *Phys. Rept.* **405**, 279 (2005). 6, 12

[6] Ostriker, J. P. & Steinhardt, P. J. Cosmic Concordance. *arXiv:astro-ph/9505066v1* (1995). 7

[7] Komatsu, E. *et al.* Seven-year wilkinson microwave anisotropy probe (WMAP) observations: Cosmological interpretation. *arXiv:1001.4538v2 [astro-ph.CO]* (2010). 7, 12

[8] Begeman, K. G., Broeils, A. H. & Sander, R. H. Extended rotation curves of spiral galaxies: Dark haloes and modified dynamics. *Monthly Notices of the Royal Astronomical Society* **249**, 523 (1991). 8

[9] Navarro, J. F., Frenk, C. S. & White, S. D. *Astrophys. J.* **462**, 563 (1996). 8

[10] Moore, B., Quinn, T., Governato, F., Stadel, J. & Lake, G. *Mon. Not. Roy. Astron. Soc.* **310**, 1147 (1999). 8

[11] Kravtsov, A. V., Klypin, A. A., Bullock, J. S. & Primack, J. R. *Astrophys. J.* **502**, 48 (1998). 8

[12] Ghez, A. M., Klein, B. L., Morris, M. & Becklin, E. E. *Astrophys. J.* **509**, 678 (1998). 8

[13] Ghez, A. M. *et al.* Stellar orbits around the galactic center black hole. *Astrophys. J.* **620**, 744 (2005). 8

[14] Peebles, P. J. E. *Gen. Rel. Grav.* **3**, 63 (1972). 8

[15] Bergström, L., Ullio, P. & Buckley, J. H. *Astropart. Phys.* **9**, 137 (1998). 9

[16] Zwicky, F. Die rotvershiebung von extragalaktischen nebeln. *Helv. Phys. Acta.* **6**, 110 (1933). 9

[17] Markevitch, M. *et al. Astrophys. J.* **567**, L27 (2002). 9

[18] Clowe, D. *et al. Astrophys. J.* **648**, L109 (2006). 9

[19] Klapdor-Kleingrothaus, H. V. & Lewis, G. F. (eds.). *Dark Matter in Astroparticle and Particle Physics* (World Scientific, 2007). 10

[20] Hu, W. & Dodelson, S. *Ann. Rev. Astron. Astrophys.* **40**, 171 (2002). 10

[21] Hu, W., Sugiyama, N. & Silk, J. *Nature* **386**, 37 (1997). 10

[22] Bennett, C. L. *et al. Astrophys. J. Suppl. Ser.* **148**, 1 (2003). 11

[23] Spergel, D. N. *et al. Astrophys. J. Suppl. Ser.* **148**, 175 (2003). 11

[24] Nolta, M. R. *et al.* Five-Year Wilkinson Microwave Anisotropy Probe Observations: Angular Power Spectra. *Astrophys. J. Suppl. Ser.* **180**, 296 (2009). 11

[25] Olive, K. A. TASI lectures on dark matter. *astro-ph/0301505* (2003). 11

[26] Bergström, L. Non-baryonic dark matter: Observational evidence and detection methods. *Rep. Prog. Phys.* **63**, 793 (2000). 12

[27] Primack, J. R. Dark matter and structure formation in the universe. *arXiv:astro-ph/9707285v2* (1997). 12

[28] Amaldi, U., de Boer, W. & Fursteanu, H. *Phys. Lett. B* **260**, 447 (1991). 13

[29] Dimopoulos, S., Raby, S. & Wilczek, F. *Phys. Lett.* **B 112**, 133 (1982). 13

[30] Ellis J. *et al. Nucl. Phys. B* **238**, 453 (1984). 13

[31] Edsjö, J. *Aspects of neutrino detection of neutralino dark matter.* Ph.D. thesis, Uppsala University (1997). 14, 169

[32] Weinberg, S. *Phys. Rev. Lett.* **40**, 223 (1978). 14

[33] Wilczek, F. *Phys. Rev. Lett.* **40**, 279 (1978). 14

[34] Murayama, H., Raffelt, G., Hagmann, C., van Bibber, K. & Rosenberg, L. J. *Eur. Phys. J.* **C 3**, 264 (1998). 14

[35] Dine, M., Fischler, W. & Srednicki. *Phys. Lett.* **B 104**, 199 (1981). 14

[36] Mirizzi, A., Raffelt, G. & Serpico, P. Photon-axion conversion in intergalactic magnetic fields and cosmological consequences. *Lect. Notes Phys.* **741**, 115–134 (2008). 15

[37] Kaluza, T. Zum Unitätsproblem in der Physik. *Sitzungsber. Preuss. Akad. Wiss.* 966–972 (1921). 15

[38] Appelquist, T., Cheng, H. C. & Dobrescu, B. A. *Phys. Rev.* **D 64**, 035002 (2001). 15

[39] Servant, G. & Tait, T. M. P. *Nucl. Phys.* **B650**, 391 (2003). 15

[40] Burnell, F. & Kribs, G. D. *Phys. Rev.* **D73**, 015001 (2006). 15

[41] Kong, K. & Matchev, K. T. *JHEP* **01**, 038 (2006). 15

[42] Kakizaki, M., Matsumoto, S. & Senami, M. *Phys. Rev.* **D74**, 023504 (2006). 15

[43] Gentile, G., Burkert, A., Salucci, P., Klein, U. & Walter, F. *Astrophys. J.* **634**, L145 (2005). 15

[44] Milgrom, M. *Astrophys. J.* **270**, 365 (1983). 15

[45] Bruneton, J. P., Liberati, S., Sindoni, L. & Famaey, B. *JCAP* **03**, 021 (2009). 16

[46] CMS Collaboration. Transverse-momentum and pseudorapidity distributions of charged hadrons in pp collisions $\sqrt{s} = 0.9$ and 2.36 TeV. *JHEP* **2010**, 1–35 (2010). 10.1007/JHEP02(2010)041. 16

[47] URL `http://www.linearcollider.org/`. 17

[48] Feng, J. L. & Peskin, M. E. *Phys. Rev.* **D64**, 115002 (2001). 17

[49] Feng, J. L. *arXiv:1003.0904v1 [astro-ph.CO]* (2010). 17

[50] LEP2 SUSY Working Group. Combined LEP selectron/smuon/stau results, 183–208 GeV. Tech. Rep., LEPSUSYWG/04-01.1 (2004). 17

[51] The Particle Data Group Collaboration, C. Amsler *et al.* Review of Particle Physics. *Phys. Lett.* **B667**, 1 (2008). 18, 29, 30, 33

[52] Allanach, B. C., Belanger, G., Boudjema, F. & Puchov, A. *JHEP* **12**, 020 (2004). 18

[53] Moroi, T., Shimizu, Y. & Yotsuyanagi, A. *Phys. Lett.* **B265**, 79 (2005). 18

[54] Baltz, E. A., Battaglia, M., Peskin, M. E. & Wizansky, T. *Phys. Rev.* **D74**, 103521 (2006). 18

[55] Feng, J. L. *arXiv:hep-ph/0509309v1* (2005). 18

[56] Wasserman, I. *Phys. Rev.* **D 33**, 2071 (1986). 18

[57] Trotta, R., Feroz, F., Hobson, M., Roszkowski, L. & de Austri, R. R. The impact of priors and observables on parameter inferences in the constrained MSSM. *JHEP* **12**, 024 (2008). 19

[58] Aprile, E. *et al.* First Dark Matter Results from the XENON100 Experiment. *Phys. Rev. Lett.* **105**, 131302 (2010). 18, 19, 20, 170

[59] Lebedenko, V. N. *et al. Phys. Rev.* **C79**, 045807 (2009). 18

[60] EDELWEISS Collaboration, E. Armengaud *et al. arXiv:0912.0805v1 [astro-ph.CO]* (2009). 18

[61] CDMS Collaboration, Z. Ahmed *et al. arXiv:0912.3592v1 [astro-ph.CO]* (2009). 18, 170

[62] DAMA Collaboration, R. Bernabei *et al. Eur. Phys. J.* **C56**, 333 (2008). 18

[63] Tucker-Smith, D. & Weiner, N. *Phys. Rev.* **D64**, 043502 (2001). 19

[64] Drobyshevski, E. M. *Mod. Phys. Lett.* **A23**, 3077 (2008). 19

[65] Bernabei, R. *et al. Eur. Phys. J.* **C53**, 205 (2008). 19

[66] Gondolo, P. & Gelmini, G. *Phys. Rev.* **D71**, 123520 (2005). 19

[67] Aalseth, C. *et al.* Results from a Search for Light-Mass Dark Matter with a P-type Point Contact Germanium Detector. *arXiv:1002.4703v2 [astro-ph.CO]* (2010). 19

[68] URL http://www-coupp.fnal.gov/. 20, 170

[69] URL http://q2c.snu.ac.kr/KIMS/KIMS_index.htm. 20, 170

[70] Martin, S. P. *arXiv:hep-ph/9709356v5* (2008). 20

[71] PAMELA Collaboration, O. Adriani *et al. Nature* **458**, 607 (2009). 20

[72] Chang, J. *et al. Nature* **456**, 362 (2008). 20

[73] Fermi LAT Collaboration, A. A. Abdo *et al. Phys. Rev. Lett.* **102**, 181101 (2009). 20

[74] Strong, A. W. *et al. arXiv:0907.0559v1* (2009). 21

[75] Fermi LAT Collaboration, D. Grasso *et al. Astropart. Phys.* **32**, 140 (2009). 21

[76] Zhang, L. & Cheng, K. S. *Astron. Astrophys.* **368**, 1063 (2001). 21

[77] Yuksel, H., Kistler, M. D. & Stanev, T. *Phys. Rev. Lett.* **103**, 051101 (2009). 21

[78] Biermann, P. L. *et al. Phys. Rev. Lett.* **103**, 061101 (2009). 21

[79] Feng, J. L., Kaplinghat, M. & Yu, H.-B. *arXiv:0911.0422v3* (2010). 21

[80] Ibe, M. & bo Yu, H. *arXiv:0912.5425v1* (2009). 21

[81] Zavala, J., Vogelsberger, M. & White, S. D. M. *arXiv:0910.5221v2* (2010). 21

[82] Arvanitaki, A. *et al. Phys.Rev.* **D79**, 105022 (2009). 21

[83] URL http://ams.cern.ch/. 21

[84] URL http://wwwmagic.mppmu.mpg.de/. 21

[85] URL http://veritas.sao.arizona.edu/. 21

[86] URL http://www.mpi-hd.mpg.de/hfm/HESS/. 21

[87] URL http://heasarc.gsfc.nasa.gov/docs/cgro/cossc/egret/. 21

[88] de Boer, W., Sander, C., Zhukov, V., Gladyshev, A. V. & Kazakov, D. I. *Astron. Astrophys.* **444**, 51 (2005). 21

[89] Johannesson, G. A first look at the GeV excess with Fermi LAT. In *Proc. XLIV Rencontre De Moriond* (2009). 21

[90] Halzen, F. & Hooper, D. Prospects for detecting dark matter with neutrino telescopes in light of recent results from direct detection experiments. *Phys. Rev. D* **73**, 123507 (2006). 23

[91] Gould, A. *Astrophys. J.* **321**, 560 – 570 (1987). 24

[92] Barger, V., Keung, W.-Y., Shaughnessy, G. & Tregre, A. High energy neutrinos from neutralino annihilations in the Sun. *Phys. Rev.* **D 76**, 095008 (2007). 24, 25

[93] Markov, M. A. & Zheleznykh, I. M. On high energy neutrino physics in cosmic rays. *Nuclear Physics* **27**, 385 – 394 (1961). 25

[94] Burgess, T. *A search for Solar Neutralino Dark Matter with the AMANDA-II Neutrino Telescope*. Ph.D. thesis, Stockholm University (2008). 27

[95] Crotty, P. High-energy neutrino fluxes from supermassive dark matter. *Phys. Rev. D* **66**, 063504 (2002). 27, 28

[96] Pumplin, J. *et al.* New generation of parton distributions with uncertainties from global QCD analysis. *JHEP* **07**, 012 (2002). 27, 63

[97] Gandhi, R., Quigg, C., Reno, M. H. & Sarcevic, I. Ultrahigh-energy neutrino interactions. *Astroparticle Physics* **5**, 81 – 110 (1996). 28

[98] Perkins, D. H. *Introduction to High Energy Physics* (Addison-Wesley Publishing Compnay, Inc., 1987). 29, 32

[99] Barret, P. H. *et al. Rev. Mod. Phys.* **24**, 133 (1952). 30

[100] Ackermann, M. *Searches for signals from cosmic point-like sources of high energy neutrinos in 5 years of AMANDA-II data*. Ph.D. thesis, Humboldt-Universität zu Berlin (2006). 31, 57, 94, 102

[101] Bethe, H. A. & W, H. *Proc. R. Soc. London Ser.* **A 146**, 83 (1934). 30

[102] Petrukhin, A. A. & Shestakov, V. V. The influence of the nuclear and atomic form factors on the muon bremsstrahlung cross section. *Can. J. Phys.* **46**, S377 (1968). 30

[103] Kokoulin, R. P. & Petrukhin, A. A. In *12th International Cosmic Ray Conference*, vol. 6, A 2436 (1971). 31

[104] Bezrukov, L. B. & Bugaev, E. V. *Sov. J. Nucl. Phys.* **33**, 635 (1981). 31

[105] Chirkin, D. & Rhode, W. Muon Monte Carlo: a high precision tool for muon propagation through matter. *hep-ph/0407075* (2004). 32, 65, 70, 150

[106] Fogli, G. L., Lisi, E., Mirizzi, A., Montanino, D. & Serpico, P. D. Oscillations of solar atmosphere neutrinos. *Phys. Rev.* **D74**, 093004 (2006). 35

[107] Askebjer, P. *et al.* Optical Properties of the South Pole Ice at Depths Between 0.8 and 1 Kilometer. *Science* **267**, 1147–1150 (1995). 38, 50

[108] Wagner, W. *Design and Realisation of a new AMANDA Data Acquisition System with Transient Waveform Recorders.* Ph.D. thesis, Universität Dortmund (2004). 42, 43

[109] Messarius, T. *Entwurf und Realisierung des AMANDA-Softwaretriggers für das TWR-Datenauslese-System.* Ph.D. thesis, Universität Dortmund (2006). 42

[110] Lundberg, J. *On the Search for High-Energy Neutrinos: Analysis of data from AMANDA-II.* Ph.D. thesis, Uppsala University (2008). 44

[111] Hubert, D. *Search with the AMANDA detector for neutralino dark matter in the Sun.* Ph.D. thesis, Vrije Universiteit Brussel (2009). 45, 54, 55, 57, 102, 105, 142, 156, 169

[112] Dickinson, J. E. *et al.* The new south pole air shower experiment: Spase-2. *Nucl. Instrum. Meth.* **A440**, 95–113 (2000). 45

[113] Andres, E. *et al.* The AMANDA neutrino telescope: Principle of operation and first results. *Astropart. Phys.* **13**, 1–20 (2000). 46

[114] Cowen, D. & Hanson, K. Time calibration of the AMANDA neutrino telescope with cosmic ray muons. In *Proceedings of the 27th ICRC* (Hamburg, Germany, 2001). 46

[115] Biron, A. Reconstruction uncertainties due to time calibration errors. AMANDA Internal Report 20001101, DESY-Zeuthen (2000). 46

[116] Woschnagg, K. Preliminary (non-optical) position calibration of strings 14 to 19. AMANDA internal report 20001002, University of California - Berkeley (2000). 47

[117] Woschnagg, K. Matching/combining laser/drill geometries. AMANDA internal note, University of California - Berkeley (1999). 47

[118] Ehrmann, W. U. & Mackensen, A. Sedimentological evidence for the formation of an East Antarctic ice sheet in Eocene/Oligocene time. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **93**, 85–112 (1992). 48

[119] Ackermann, M. *et al.* Optical properties of deep glacial ice at the South Pole. *J. Geophys. Res.* **111**, D13203 (2006). 48, 50, 51

[120] Clark, P. U. *et al.* The Last Glacial Maximum. *Science* **325**, 710–714 (2009). 48

[121] Petit, J. R. *et al.* Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **399**, 429–436 (1999). 48

[122] Price, P. B., Woschnagg, K. & Chirkin, D. Age vs depth of glacial ice at South Pole. *Geophys. Res. Lett.* **27(14)**, 2129–2132 (2000). 48, 49

[123] Hogan, A. W. & Gow, A. J. Occurrence frequency of thickness of annual snow accumulation layers at South Pole. *J. Geophys. Res.* **102(D12)**, 14,021–14,027 (1997). 48

[124] Bramall, N. E. *et al.* A deep high-resolution optical log of dust, ash, and stratigraphy in South Pole glacial ice. *Geophys. Res. Lett.* **32**, L21815 (2005). 48, 49

[125] Mayewski, P. A. *et al.* Major features and forcing of highlatitude northern hemisphere atmospheric circulation using a 110,000- year-long glaciochemical series. *J. Geophys. Res.* **102**, 26,345–26,366 (1997). 49

[126] Miller, S. L. Clathrate hydrates of air in Antarctic ice. *Science* **165**, 489–490 (1969). 49

[127] Price, P. B. Kinetics of conversion of air bubbles to air hydrate crystals in Antarctic ice. *Science* **267**, 1802–1804 (1995). 49

[128] Price, P. B. & Bergström, L. Optical properties of deep ice at the South Pole: Scattering. *Appl. Opt.* **36**, 4181–4194 (1997). 49

[129] Askebjer, P. *et al.* Optical properties of deep ice at the South Pole: absorption. *Appl. Opt.* **36**, 4168–4180 (1997). 50

[130] Olbrechts, P. & Wiebusch, C. On the angular sensitivity of optical modules in the ice. AMANDA internal report 20010102, DESY, Zeuthen (2001). 52

[131] Walck, C. A Study of Life & Death of the AMANDA Detector. Tech. Rep., University of Stockholm (2003). 54

[132] Braun, J. R. *A Maximum-Likelihood Search for Neutrino Point Sources with the AMANDA-II Detector*. Ph.D. thesis, University of Wisconsin - Madison (2009). 57, 169

[133] Kelley, J. L. *Searching for Quantum Gravity with High-energy Atmospheric Neutrinos and AMANDA-II*. Ph.D. thesis, University of Wisconsin - Madison (2008). 57

[134] Ackermann M. & Bernardini B. (for the IceCube Collaboration). An Investigation of Seasonal Variations in the Atmospheric Neutrino Rate with the AMANDA-II Neutrino Telescope. In *29th International Cosmic Ray Conference, Pune, India*, vol. 9 of *HE2.2*, 107–110 (2005). 59

[135] Wissing, H. Climatology from muon rates. In *IceCube Collaboration Meeting, DESY-Zeuthen* (2006). 59

[136] Metropolis, N. The beginning of the Monte Carlo method. *Los Alamos Science* **15**, 125 (1987). 62

[137] The AMANDA Collaboration. *F2000 (version 1.5) AMANDA offline format. An ASCII data format for handling data and MC events* (2001). 62

[138] URL http://www.begrid.be/. 63

[139] URL http://www.sara.nl/index_eng.html. 63

[140] Blennow, M., Edsjö, J. & Ohlsson, T. Neutrinos from wimp annihilations using a full three-flavor monte carlo. *hep-ph/0709.3898* (2007). 63, 150

[141] *WimpEventF2k*. URL http://www.physto.se/~burgess/work/WimpEventF2k. 63

[142] Sjöstrand, T., Mrenna, S. & Skands, P. *PYTHIA 6.4*. JHEP **0605** (2006). 63

[143] Gondolo, P. *et al.* DarkSUSY: Computing supersymmetric dark matter properties numerically. *Journ. of Cosm. and Astropart. Phys.* **0407**, 008 (2004). 63

[144] Bahcall, J., Serenelli, A. & Basu, S. New solar opacities, abundances, helioseismology, and neutrino fluxes. *Astroph. J.* **621**, L85 (2005). 63

[145] Edsjö, J. Calculation of neutrino cross sections and the `nuNevent` neutrino-nucleon scattering Monte Carlo. Tech. Rep., Stockholm University (2005). 63

[146] Wilquet, G. *Masses, mélange et oscillations de neutrinos*. Thèse présentée à l'agrégation à l'enseignement supérieur, Université Libre de Bruxelles (2004). 63

[147] Rizzo, A. `GenN++`: Generator of neutrino-nucleon interactions. Tech. Rep., Vrije Universiteit Brussel (2006). 64

[148] Heck, D., Knapp, J., Capdevielle, J. N., Schatz, G. & Thouw, T. `CORSIKA`: A Monte Carlo code to simulate extensive air showers. Tech. Rep. FZKA-6019, Forschungszentrum Karlsruhe (1998). 68

[149] Chirkin, D. *Cosmic ray energy spectrum measurement with the Antarctic Muon and Neutrino Detector Array (AMANDA)*. Ph.D. thesis, UC Berkeley (2003). 68

[150] T. Antoni *et al.* (KASCADE Collaboration). The cosmic ray experiment KASCADE. *Nucl. Instr. Meth.* **A513**, 490 (2003). 68

[151] Fletcher, R. S., Gaisser, T. K., Lipari, P. & Stanev, T. `SIBYLL`: An event generator for simulation of high energy cosmic ray cascades. *Phys. Rev. D* **50**, 5710–5731 (1994). 68

[152] Hörandel, J. R. On the knee in the energy spectrum of cosmic rays. *Astropart. Phys.* **19**, 193 – 220 (2003). 68

[153] Gazizov, A. & Kowalski, M. Anis: high energy neutrino generator for neutrino telescopes. *astro-ph/0406439* (2004). 69

[154] Lipari, P. Lepton spectra in the earth's atmosphere. *Astropart. Phys.* **1**, 195 (1993). 69

[155] Lundberg, J. *et al.* Light tracking through ice and water–scattering and absorption in heterogeneous media with photonics. *Nucl. Instrum. and Meth. A* **581**, 619 – 631 (2007). 71

[156] *AHA Ice Model*. URL `http://wiki.icecube.wisc.edu/index.php/Aha`. 71

[157] Hundertmark, S. AMASIM neutrino detector simulation program. In *International Workshop on simulation and analysis methods for large neutrino telescope*, DESY-PROC-1999-01 (Zeuthen, Germany, 1999). 71

[158] URL http://internal.icecube.wisc.edu/amanda/software/sieglinde/. 79

[159] *ROOT: An Object-Oriented Data Analysis Framework*. URL http://root.cern.ch. 79

[160] *TMVA: Toolkit for Multivariate Data Analysis with ROOT*. URL http://tmva.sourceforge.net. 79, 119

[161] Pohl, A. C. *A statistical tool for finding non-particle events from the AMANDA neutrino telescope*. Master's thesis, Uppsala University (2004). 80, 112

[162] Ribordy, M. AMANDA-II/2000 data statistics, OM selection and retriggering procedure. AMANDA internal report 20020601, DESY Zeuthen (2002). 80

[163] Taboada, I. *Search for High Energy Neutrino Induced Cascades*. Ph.D. thesis, University of Pennsylvania, State College (2002). 83, 84

[164] Ahrens J. *et al.* Muon track reconstruction and data selection techniques in AMANDA. *Nucl. Instrum. Meth. A* **524**, 169 (2004). 85, 87, 90, 91, 92

[165] Steffen, P. Direct Walk II (Improved version of Direct Walk). AMANDA internal report 20020201, DESY, Zeuthen (2002). 86

[166] Davour, A. *Search for low mass WIMPs with the AMANDA neutrino telescope*. Ph.D. thesis, Uppsala University (2007). 86

[167] Steffen, P. AMANDA pattern recognition (2002). Talk given at AMANDA Collaboration Meeting in Stockholm. 86

[168] Pandel, D. *Bestimmung von Wasser- und Detektorparametern und Rekonstruktion von Myonen bis 100 TeV mit dem Baikal-Neutrinoteleskop NT-72*. Master's thesis, Humboldt-Universität Berlin (1996). 90

[169] Press, W. H., Teukolsky, S. A., Vetterling, W. V. & Flannery, B. P. *Numerical Recipes in C - The art of scientific computing* (Cambridge University Press, 1997), 2nd edn. URL http://www.nr.com. 91

[170] Gaug, M., Niessen, P. & Wiebusch, C. Investigations on Smoothness observables. AMANDA Internal Report 20000201, DESY Zeuthen (2000). 95

[171] URL http://nuastro-zeuthen.desy.de/neutrino_experiments/icecube. 101

[172] URL http://www.isv.uu.se/IceCube. 101

[173] URL http://w3.iihe.ac.be/icecube. 101

[174] URL http://icecube.physics.wisc.edu. 101

[175] URL http://www3.tsl.uu.se/~akesson/wimpfilter.html. 102

[176] URL http://wiki.icecube.wisc.edu/index.php/Offline_filtering_of_the_2006_muondaq_data. 103

[177] Roe, B. P. *et al.* Boosted decision trees as an alternative to artificial neural networks for particle identification. *Nucl. Instrum. and Meth. A* **543**, 577–584 (2005). 118

[178] Breiman, L., Friedman, J., Olshen, R. & Stone, C. *Classification and Regression Trees* (Wadsworth, 1984). 118, 121

[179] Freund, Y. & Schapire, R. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *J. of Computer and System Science* **55**, 119 (1997). 120

[180] Punzi, G. Sensitivity of searches for new signals and its optimization (2003). ArXiv.org:physics/0308063. 126

[181] Hill, G. C. & Rawlins, K. Unbiased cut selection for optimal upper limits in neutrino detectors: the model rejection potential technique. *Astroparticle Physics* **19**, 393 – 402 (2003). 136

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

[183] Wilks, S. S. *Mathematical Statistics* (Wiley, New York, 1962). 2d printing, corrected, 1963. 142

[184] Maltoni, M., Schwetz, T., Tortola, M. A. & Valle, J. W. F. *New J. Phys.* **6** (2004). 148

[185] Abbasi, R. *et al.* Limits to the muon flux from neutralino annihilations in the sun with the icecube 22-string detector. *Phys. Rev. Lett.* **102**, 201302 (2009). 150, 153, 156

[186] Achterberg A. *et al.* Five years of searches for point sources of astrophysical neutrinos with the AMANDA-II neutrino telescope. *Phys. Rev.* **D75**, 102001 (2007). 150

[187] Barr, G. D., Robbins, S., Gaisser, T. K. & Stanev, T. Uncertainties in atmospheric neutrino fluxes. *Phys. Rev.* **D74**, 094009 (2006). 155

[188] Wikström, G. & Edsjö, J. *astro-ph/0903.2986v1* (2009). 159, 164, 170

[189] Rolke, W. & López, A. How to claim a discovery. *PHYSTAT2003* (2003). 161

[190] Wikström, G. *A search for solar dark matter with the IceCube neutrino telescope.* Ph.D. thesis, Stockholm University (2009). 169

[191] URL `http://wiki.icecube.wisc.edu/index.php/DarkSusy_scans`. 169

[192] Desai S. *et al.* Search for dark matter wimps using upward through-going muons in super-kamiokande. *Phys. Rev. D* **70**, 083523 (2004). 170

[193] Boliev M. M. *et al.* Baksan neutralino search. In Klapdor-Kleingrothaus, H. V. & Ramachers, Y. (eds.) *Dark Matter in Astro- and Particle Physics, Dark '96*, 711 (World Scientific, 1997). 170

[194] Ambrosio M. *et al.* Limits on dark matter wimps using upward-going muons in the MACRO detector. *Phys. Rev. D* **60**, 082002 (1999). 170

[195] Resconi, E. (for the IceCube Collaboration). Status and prospects of the IceCube neutrino telescope. *astro-ph/0807.3891v1* (2008). 170

List of Abbreviations

ΛCDM	Lambda-Cold Dark Matter, current concordance model
AHA	Ice optical properties modeller
ALL	Logical sum of STR and STD trigger selection
AMANDA	Antarctic Muon And Neutrino Detector Array
AMASIM	Monte Carlo code simulating AMANDA detector response
AMS	An anti-matter detector to be set on the International Space Station
ANIS	Monte Carlo tool generating all-flavour neutrinos
ATIC	Balloon-borne instrument to measure the energy and composition of cosmic rays
AdaBoost	Adaptive boost algorithm
BAKSAN	Baksan neutrino observatory
BDT	Boosted Decision Tree
BEGrid	The Belgian GRID for research
CC	Charge-current interaction
CDMS II	Direct dark matter detection experiment
CDM	Cold Dark Matter
CI	Confidence Interval
CL	Confidence Level
CMB	Cosmic Microwave Background
COG	Centre of Gravity
CORSIKA	Software package to simulate atmospheric muon background interactions
COUOPP	Direct dark matter detection experiment
CP	Charge conjugation and parity symmetry
CTEQ6	Parton distribution function
CoGeNT	Direct dark matter detection experiment
DAMA/LIBRA	Direct dark matter detection experiment
DAQ	Data acquisition system
DMAD	Discriminator and Multiplicity ADder
DM	Dark Matter
DOM	Digital Optical Module
DT	Decision Tree

DWimp	Direct Wimp first-guess reconstruction
DW	Direct Walck first-guess reconstruction
DarkSUSY	Fortran package for supersymmetric dark matter calculations
DeepCore	IceCube low energy extension
EDELWEISS II	Direct dark matter detection experiment
EGRET	Space-based gamma ray telescope
F2000	AMANDA plain text file format
FC	Feldman and Cousins frequentist approach
FG	First-guess reconstruction method
Fermi LAT	Space-based gamma ray telescope
GALPROP	Software package modelling the cosmic-ray diffusion in the Galaxy
GINI	Generalised inequality index
GISP	Greenland Ice Sheet Project
GPS	Global Position System
GRID	Globalisation and virtualisation of computer infrastructures
GUT	Grand Unified Theory
H.E.S.S.	Ground-based atmospheric Cherenkov telescope
HDM	Hot Dark Matter
HV	High Voltage
ILC	International Linear Collider
IceCube	One-kilometer cube neutrino detector
JAMS	Just Another Muon Search, first-guess reconstruction
KASKADE	Shower core and array detector
KIMS	Direct dark matter detection experiment
KK	Kaluza-Klein
L0	Level 0 filtering, or trigger level
L1	Level 1 filtering
L2	Level 2 filtering
LE	Leading edge
LED	Light Emitting Diode
LGM	Last Glacial Maximum
LHC	Large Hadron Collider
LKP	Lightest Kaluza-Klein Particle
LLH	Log-likelihood reconstruction method
LSP	Lightest Supersymmetric Particle
MACHOs	Massive Compact Halo Objects
MACRO	Deep underground detector at LNGS

MAGIC	Ground-based atmospheric Cherenkov telescope
MAPO	Martin A Pomerantz Observatory
MDP	Model Discovery Potential
MILLENNIUM	Ice optical properties modeller
MMC	Three-dimensional Monte Carlo muon propagator
MOND	MOdified Newtonian Dynamics
MRP	Model Rejection Potential
MSSM	Minimal Supersymmetric extension of the Standard Model
MVA	Multivariate Analysis
NC	Neutral-current interaction
Nd:YAG	Neodymium-doped yttrium aluminium garnet; Nd:$Y_3Al_5O_{12}$
OM	Optical Module
ORB	Optical Receiver Boards
OTDM	Optical Time-Domain Reflectometer
PAMELA	Satellite borne experiment to study of the antimatter
PMT	Photo-multiplier tube
PYTHIA	Montecarlo code to simulate particle interactions
Photonics	Photon tracking tool
Planck	Space observatory experiment
ROOT	An object-oriented data-analysis framework
SARA Matrix	Dutch National High Performance Computing and e-Science Support Center
SD	Spin-dependent
SI	Spin-independent
SLART	ROOT based Sieglinde code
SPASE	South Pole Air Shower Experiment
STD	Standard (or Multiplicity 24, M24) trigger selection
STR	String trigger selection
SUSY	Super Symmetry
SWAMP	SWedish AMPlifier
SYBILL	Monte Carlo modelling the high energy hadronic interactions
Sieglinde	AMANDA Software analysis programme
Super-K	Super-Kamiokande underground experiment
TDC	Time-to-digital converter
TE	Trailing edge
TMVA	Toolkit for Multivariate Data Analysis
TOT	Time Over Threshold

TWR	Transient Waveform Record
UED	Universal Extra Dimensions
UV	Ultraviolet radiation
VERITAS	Ground-based atmospheric Cherenkov telescope
WIMP	Weakly Interacting Massive Particle
WMAP	Wilkinson Microwave Anisotropy Probe
WimpSim	Montecarlo code to simulate neutralino annihilation and interaction
XENON100	Direct dark matter detection experiment
ZEPLIN III	Direct dark matter detection experiment
p.d.f.	Probability Density Function
pADC	Peak analog-to-digital converter
pe	Photo-electron
spe	Single photo-electron

Acknowledgements

Well, this is the last page of my dissertation, however is the most important since I want to say here that I am particular grateful to a lot of people, which help me during the period of my PhD studies in Belgium.

I want to express my deep gratitude to the Interuniversity Institute for High Energies (IIHE) of the Vrije Universiteit Brussel - Université Libre de Bruxelles, without which this work could not start. I want also to acknowledge the FWO-Vlaanderen, which made it possible my trip to the South Pole.

The person towards whom I have a particular debt of gratitude is Catherine De Clercq, my supervisor and co-director of the IIHE, who admit me to the IIHE and gave me the great opportunity to work with her on the dark matter subject, an intriguing puzzle of modern Cosmology. Thanks to Catherine I joined the AMANDA/IceCube Collaboration, and apart going all around the world to participate to conferences and meetings, thanks to her I went to one of the most exotic place of Earth, the South Pole.

I want to thank all the AMANDA/IceCube collaborators at the IIHE: my office mates Sabrina and Mathieu for serious and amusing discussions, Daniel, Nick and Kael for their help. I want to thank also the newcomers of the group, and in particular Mark who translated the abstract of my thesis in Dutch. I will of course not forget to thank my previous office mate Bruny, as well as Philip and Patrick, and in particular Daan with whom I shared the passion for dark matter and from whom I received a lot of help. I wish also to thank the AMANDA/IceCube groups in Mons and in Ghent for nice discussions during our internal meetings. I want to thank all the people in the international AMANDA/IceCube Collaboration and in particular all the WIMP working groups in the Collaboration. I want to send a particular thanks to the people in the Uppsala group (Carlos, Allan, Olga,..) and the Stockholm group (Peo, Klas, Christian, Matthias,..) as well as to Thomas Burgess.

I want to tank all the members of the IIHE, with whom I had good feeling and stimulating interactions, especially at lunch time, where with Stefaan, Olivier and Mateusz we spoke about everything (history, politics, Latin...) but physics. I wish also to remember dear memory of professor Denis Johnson.
The IIHE can be absolutely envied because has the most nice, kind and efficient secretaries of Belgium (or maybe of the Universe...): Marleen and Danielle, as well as it was Rosine. I owe you a big thanks. Without the people of the computing support, my simulations and analysis would not have had success, I wish to send a big thanks to them for this, and in particular to Shkelzen and to Abdel.

Bene, adesso carissima mamma Francesca e carissima sorellona Elena, vorrei con queste poche parole esprimere tutto il mio affetto per voi, che mi avete dato sempre il vostro amore e appoggio specialmente da quando sono andato via dall'Italia. So che per voi non é stato semplice accettare la mia partenza, come lo é stato per me separarmi da voi. Ma noi saremo uniti sempre dovunque io mi trovi.
Che dirvi, vi voglio bene!

Edgar Allan Poe - Eureka

...I propose to take such a survey of the Universe
that the mind may be able really to receive
and to perceive an individual impression.
He who from the top of Ætna
casts his eyes leisurely around,
is affected chiefly by
the extent and diversity of the scene.
Only by a rapid whirling on his heel
could he hope to comprehend the panorama
in the sublimity of its oneness.
But as, on the summit of Ætna,
no man has thought of whirling on his heel,
so no man has ever taken into his brain
the full uniqueness of the prospect;
and so, again, whatever considerations
lie involved in this uniqueness
have as yet no practical existence for mankind.