Search for High Energy neutrinos from the Galaxy with the ANTARES neutrino telescope

Timothée Grégoire

January 25, 2017

Laboratoire APC, Paris

1 Event Selection

The event selection is adapted from (Thèse de Tino ou papier de Javier, selon...)

An event is selected as a track-like event if it is reconstructed up-going to reject atmospheric muons. It should also have a good reconstruction quality and an angular error estimate smaller than 1°.

(Je ne parle pas des triggers, devrais-je ?)

To be selected as shower-like, an event should not have been selected as a track-like event to avoid overlapping. Then, the rejection of atmospheric muons is the main goal of the shower event selection. When a muon is reconstructed by the shower algorithm, the reconstructed position often lie far away from the detector, so the reconstructed position is selected to be contained in the detector and the value of the M-estimator used to fit this position is required to be not too high. The rejection of atmospheric muons is also done selecting events reconstructed as up-going, using a dedicated likelihood function and a random decision forest algorithm. Their is also a selection on the angular error estimator.

2 Search Method

When the signal searched for is well defined and that the probability density functions of observables are known, which is the case here, the likelihood ratio test has maximum power. It has already been applied to various searches in neutrino astronomy and in particular on ANTARES data ([1] citations) on reduced portions of the sky for individual point-like or extended sources at the degree scale (à vérifier). It is adapted here to a full sky search where the all sky signal map is computed according to the KRA$_\gamma$ model presented in section ??. The data are considered to be a mixture of signal and background events, so the likelihood is defined as:

$$\mathcal{L}_{sig+bkg} = \prod_{\mathcal{S}\in\{tr,sh\}} \prod_{i\in\mathcal{S}} [\mu^{\mathcal{S}}_{sig} \cdot pdf^{\mathcal{S}}_{sig}(E_i, \vec{x_i}) + \mu^{\mathcal{S}}_{bkg} \cdot pdf^{\mathcal{S}}_{bkg}(E_i, \vec{x_i})] \tag{1}$$

Where E_i is the reconstructed energy of an event i and $\vec{x_i}$ its reconstructed direction. For each type of event $\mathcal{S}$ (track or shower), the number of background events $\mu^{\mathcal{S}}_{bkg}$ corresponds to the total number of events $\mu^{\mathcal{S}}_{tot}$ minus the number of signal events $\mu^{\mathcal{S}}_{sig}$, which is fitted by maximising the likelihood. μ_{sig} must have a positive value. The maximisation of the likelihood is done using the TMinuit algorithm within the ROOT framework. The signal and background probability density functions of an event are defined as:

$$pdf_{sig}(\alpha_i, \delta_i, E_i) = \mathcal{M}^{\mathcal{S}}_{sig}(\alpha_i, \delta_i) \cdot \mathcal{E}^{\mathcal{S}}_{sig}(E_i, \alpha_i, \delta_i) \tag{2}$$

$$pdf_{bkg}(z_i, \delta_i, E_i) = \mathcal{M}^{\mathcal{S}}_{bkg}(\delta_i) \cdot \mathcal{E}^{\mathcal{S}}_{bkg}(E_i, z_i) \tag{3}$$

where $\mathcal{M}^{\mathcal{S}}$ is the probability density function to be at a certain position in the sky. It is in equatorial coordinate (right ascension α_i and declination δ_i) for the signal (fig. 1). Only the declination is used for the background as the right ascension distribution is flat because of the earth rotation. A random combination of two spline parametrizations is used as can be seen on figure 3. $\mathcal{E}^{\mathcal{S}}$ is the probability density function to have a certain energy. In the signal case, it depends on the equatorial coordinates as the energy spectrum of the KRA$_\gamma$ model depends on the position in the sky. To gain CPU time, $\mathcal{E}_{sig}$ is considered to be $pdf(E_i, \delta_i) \cdot pdf(E_i, \alpha_i)$ using the distributions represented on figure 2. For the background, it depends on the corresponding zenith to account for systematics (fig. 4). All these distributions are made from MC simulations except for $\mathcal{M}^{\mathcal{S}}_{bkg}(\delta_i)$.

Figures/timothee_morpho_sh_aitoff.png

Figures/timothee_morpho_tr_aitoff.png

Figure 1: Probability density function of the reconstructed position of a signal event in equatorial coordinates. **Left:** Shower-like events, **Right:** Track-like events.

Figures/timothee_2Dspect_sh_dec.png

Figures/timothee_2Dspect_tr_dec.png

Figures/timothee_2Dspect_sh_ra.png

Figures/timothee_2Dspect_tr_ra.png

Figure 2: Signal distribution of the energy estimator in function of the reconstructed (**Above:**) declination and (**Below:**) right ascension. **Left:** Shower-like events, **Right:** Track-like events.

Figures/timothee_sindec_sh_withoutMC.png

Figures/timothee_sindec_tr_withoutMC.png

Figure 3: Background distribution of the sine of the reconstructed declination. The red and green curves are two different spline parametrisations of the data used to get $\mathcal{M}_{bkg}$. **Left:** Shower-like events, **Right:** Track-like events.

Figures/timothee_bgr_spect2D_sh.png

Figures/timothee_bgr_spect2D_tr.png

Figure 4: Background distribution of the energy estimator versus the reconstructed zenith. **Left:** Shower-like events, **Right:** Track-like events.

Figures/timothee_pdf_TS.png

Figure 5: Probability density function of the test statistic for different fluxes corresponding to a mean of {0, 11, 22, 33, 44} events passing the cuts. Note that the binning is not constant.

Figures/timothee_UL_vs_TS_fitted.png

Figure 6: UL that can be putted in function of the value of the test statistic, fitted by $(ax+b)\cdot(1-c\cdot e^{dx})$.

The logarithm of the likelihood can be rewritten as:

$$log\mathcal{L}_{sig+bkg} = \sum_{\mathcal{S}\in\{tr,sh\}} \sum_{i\in\mathcal{S}} log[\mu_{sig}^{\mathcal{S}} \cdot \mathcal{M}_{sig}^{\mathcal{S}}(\alpha_i,\delta_i) \cdot \mathcal{E}_{sig}^{\mathcal{S}}(E_i,\delta_i,\alpha_i) + (\mu_{tot}^{\mathcal{S}} - \mu_{sig}^{\mathcal{S}}) \cdot \mathcal{M}_{bkg}^{\mathcal{S}}(\delta_i) \cdot \mathcal{E}_{bkg}^{\mathcal{S}}(E_i,z_i)] \quad (4)$$

Then to have a better capacity to choose between background or signal hypotheses, the likelihood to have some signal is weighted against the likelihood to have only background in our data. So a test statistic which is the logarithm of the likelihood ratio is build:

$$TS = log(\mathcal{L}_{sig+bkg}) - log(\mathcal{L}_{bkg}) \quad (5)$$

with $\mathcal{L}_{bkg} = \mathcal{L}_{sig+bkg}(\mu_{sig}^{sh} = \mu_{sig}^{tr} = 0)$.

To choose between different hypotheses, it is necessary to get the probability density function of the test statistic for each hypothesis so for each value of the KRA$_\gamma$ flux (including a null flux, the background-only hypothesis). To do so, pseudo-experiments are produced for each possible value of μ_{sig}^{sh+tr}. A pseudo-experiment is a set of events that is generated using the probability density functions of the energy estimator and angular position (in local coordinates for the background to account for the effects of the detector symmetries). 100,000 pseudo-experiments are produced for a value of $\mu_{sig}^{sh+tr} = 0$ and 10,000 pseudo-experiments for each other value of μ_{sig}^{sh+tr} in [1, 55].

For each pseudo-experiment the test statistic is computed, so for each number of signal events passing the cuts, the distribution of these test statistics gives the probability density function $pdf_{\mu_{sig}^{sh+tr}}(TS)$. The interesting quantity is the flux, so knowing that a flux will produce a mean of n events, $pdf_{\mu_{sig}^{sh+tr}}(TS)$ is converted into the probability density function to have a certain flux Φ, $pdf_\Phi(TS)$:

$$pdf_\Phi(TS) = pdf_{<\mu_{sig}^{sh+tr}>=n}(TS) = \sum_i P(i|n) \cdot pdf_{\mu_{sig}^{sh+tr}=i}(TS) \quad (6)$$

with P a Poissonian.

Some of these probability density functions can be seen on the figure 5.

Then the anti-cumulatives of these probability density functions (fig. 7(a)) are computed. It gives the probability to have a test statistic bigger than a certain value for each hypothesis. Let's call P the p-value, i.e. the probability that the background has a test statistic bigger than TS_P. This is given by the plot represented on figure 7(b) with P in y axis and TS_P the corresponding value in x.

Then, the median upper limit at 90% confidence level, $\mathrm{UL}(\mathrm{TS}_{0.5})$, corresponds to the value of the flux which has more than 90% probability to have a test statistic bigger than $\mathrm{TS}_{0.5}$.

The average upper limit is:

$$< \mathrm{UL} >= \sum_P \mathrm{UL}(\mathrm{TS}_P) \times pdf_{bkg}(\mathrm{TS}_P) \quad (7)$$

with UL(TS) shown on figure 6 and pdf_{bkg}(TS) is the distribution filled in yellow on figure 5.

The discovery probability at 3σ correspond to the probability to have a test statistic bigger than $\mathrm{TS}_{3\sigma}$ for the signal hypothesis.

Figure 7: Anti-cumulative of the probability density function of the test statistic. It gives the probability to have a test statistic bigger than the value in x axis.

Figures/timothee_antiCumTS.png

(a) Different curves corresponding to different fluxes.

Figures/timothee_antiCumTS_bg.pdf

(b) Only the background hypothesis with the y axis in logarithm scale and a zoom in x axis. The vertical lines show the value of the test statistic above which there is a probability smaller than 2σ (left line) or 3σ (right line) to have a bigger test statistic.

3 Sensitivity estimation

There is 208 shower-like events and 7300 track-like events in our data. The mean number of events that is expected from the KRA$_\gamma$ model version with the cut at $5 \cdot 10^7$ GeV is $< \mu_{sig} > = 13.720$ events with 20.3% of showers. The energy range is between 2.2 and 258 TeV for the showers and between 402 GeV and 230 TeV for the tracks.

The sensitivity obtained, defined as the median upper limit at 90% confidence level, is 1.166 times the KRA$_\gamma$ flux (Φ_{KRA_γ}), which corresponds to a mean of 16 events and the average upper limit is $1.026 \cdot \Phi_{KRA_\gamma}$, which correspond to a mean of 14.08 events. If the model is correct, there is 13.4% probability to claim a 3σ discovery. Figure 8 illustrate the evolution of the discovery probability in function of the percentage of the KRA$_\gamma$ flux and figure 7(b) gives the probability to have a test statistic bigger than the one in x axis in the background-only hypothesis.

For the version with the cut at $5 \cdot 10^6$ GeV, a mean of 11.623 events is expected with 19.7% of showers. The median upper limit is $1.55 \cdot \Phi_{KRA_\gamma}$, which corresponds to a mean of 18 events and the average upper limit is 1.354 times KRA$_\gamma$ flux, which correspond to a mean of 15.73 events. The probability to claim a 3σ discovery of this flux is 6.78%.

4 Results

The background parametrisation being a combination of two splines using a random number r_{dec}, the test statistic and the number of fitted events are the average for 5000 values of r_{dec}. For the version of the KRA$_\gamma$ model with the cut at $5 \cdot 10^7$ GeV the data give an average test statistic of 0.650. The number of fitted tracks is $< \mu_{\text{fitted tracks}} >= 0.00$ and showers $< \mu_{\text{fitted showers}} >= 3.18$. The p-value is obtained using the value of the test statistic and the curve represented on figure 7(b), it is 0.456. The upper limit is obtained from the fit of the curve of the upper limit in function of the test statistic (fig. 6), it is $1.268 \cdot \Phi_{KRA_\gamma}$. The corresponding probability density function for the background-only hypothesis is 0.13 and 0.19 for the signal hypothesis.

For the version of the KRA$_\gamma$ model with the cut at $5 \cdot 10^6$ GeV the data give $< TS >= 0.420$, $< \mu_{\text{fitted tracks}} >= 0.00$ and $< \mu_{\text{fitted showers}} >= 2.67$, the p-value is 0.581 and the upper limit is $1.330 \cdot \Phi_{KRA_\gamma}$. The probability density function at this test statistic is 0.24 for the background-only hypothesis and 0.26 for the signal hypothesis.

Figure 9 show that fitting few shower events and no tracks can happen for background or signal hypotheses.

References

[1] Michael T. & KM3NeT, 2015, Neutrino point source search including cascade events with the ANTARES neutrino telescope, in proceedings, 34 th International Cosmic Ray Conference (ICRC 2015)

Figures/timothee_disco.png

Figure 8: Discovery probability at 3σ in function of the signal neutrino flux in percentage of the KRA_γ flux.

Figures/timothee_nsig_ratio.png

Figure 9: Distribution of the ratio of the number of fitted shower events over the total number of fitted events. The peaks correspond to a number of fitted showers (left), tracks (right) or both (middle) equal to zero.