PROCEEDINGS OF SCIENCE

Model-independent search for neutrino anisotropies with the ANTARES neutrino telescope

Stefan Geißelsöder for the ANTARES collaboration[*†]
Friedrich-Alexander University of Erlangen-Nürnberg,
Erlangen Centre for Astroparticle Physics
E-mail: stefan.geisselsoeder@fau.de

ANTARES is the largest operational neutrino telescope in the Northern Hemisphere, located in the Mediterranean Sea at a depth of 2475 meter. The direction and energy of the observed particles are reconstructed from the time and amplitude information recorded by the photomultiplier tubes. The collected set of reconstructed events can be analyzed with respect to the spatial, temporal and energy distribution.

The approach shown in this contribution focuses on the spatial distribution, searching unbiasedly for a significant excess of neutrinos with an arbitrary size and shape from any direction in the sky. Techniques originating from the domain of pattern recognition and image processing are used. In contrast to a dedicated search for a specific neutrino emission model this approach is sensitive to a wide range of possible source structures. The result of this method applied to the ANTARES data are presented.

The 34th International Cosmic Ray Conference,
30 July - 6 August, 2015
The Hague, The Netherlands

[*]Speaker.
[†]http://antares.in2p3.fr/

PoS(ICRC2015)1054

1. Neutrino astronomy with ANTARES

Neutrinos are able to traverse through dense matter and are not deflected by galactic or extra-galactic magnetic fields. While these properties make them favorable for astronomy, their detection becomes more complex and requires large volumes. ANTARES [1] is the largest operational neutrino telescope in the Northern Hemisphere , providing a good view on the Galactic Center through the Earth. Despite many dedicated searches [2] [3] [4] [5] focusing on promising candidates for cosmic neutrino sources, no source has been identified statistically significant yet.

2. Model-independent multiscale source search

The model-independent multiscale source search presented here tries to identify regions of arbitrary position, size, shape and internal neutrino distribution in which an excess of neutrino events with respect to the background expectation has been observed. In contrast to the testing of preselected hypotheses, an unbiased approach can also detect unexpected structures. The main drawbacks are higher trial factors than in a dedicated search and possibly a less straight forward interpretation of the result. Since there is no physical model involved, any kind of deviation, for instance uncompensated systematic effects, could be detected, but nevertheless this would be a valuable result.

A discrete spherical grid with 165016 gridpoints, corresponding to a spacing of $\approx 0.5^\circ$, is used to evaluate the directions of the measured neutrinos. Figure 1 shows such a spherical grid with gridpoints in blue and random neutrino events in white. In the example shown in Figure 1, random events with two artificial point sources with 12 and 18 events at a declination of -70° have been added to demonstrate the analysis method. The search evaluates scales from 0.0° up to 90.0°

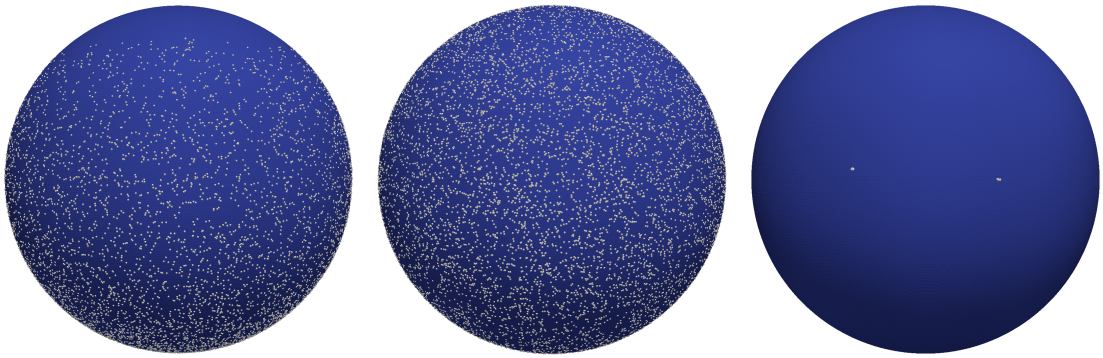

Figure 1: Left: A spherical grid with 12000 random events and two point-like sources. The gridpoints are rendered with a radius of about 0.5°, hence the overlap and form a closed sphere. View on the equator (declination of 0°). Middle: View from below to the south pole (declination of -90°). Right: The same setup displayed without the random events. Since the sphere is a three dimensional model, only the part facing the observer is visible.

in steps of 0.5°. It starts by counting the number of neutrinos located in a ring around each gridpoint with a radius corresponding to the current search scale. The result of this evaluation is one number

for each gridpoint in each scale. The counting is visualized in Figure 2. The results for three scales can be seen in Figure 3.

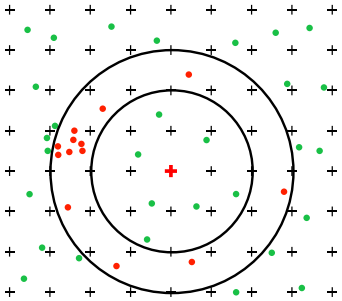

Figure 2: Scheme of the neutrino counting. Crosses mark the gridpoints with a distance of 0.5° between them. Green and red dots are neutrinos. The red cross is the gridpoint that is being evaluated. The current search scale is between the black circles. It is 1.0° (inner circle) to 1.5° (outer circle) in this example. Neutrinos which are found for the current search scale at the current searchpoint are shown in red. The result of the evaluation of this scale at the red gridpoint is 13.

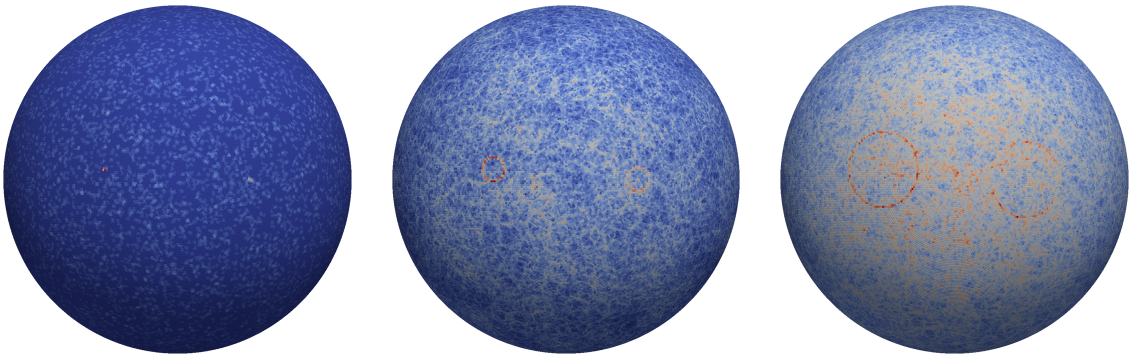

Figure 3: a) The spherical search grid with the number of counted events in a circle between 0.0° and 0.5° around each gridpoint. b) Number of events between 3.0° and 3.5°. c) Number of events between 10.0° and 10.5°. The color scale is readjusted between the different scales.

The next step is to calculate the Poisson probability for each observed value. The number of neutrinos n around each gridpoint has been counted and the expected mean number λ is estimated from scrambled data. With this information the Poisson probability $P(n)$ can be computed. For technical reasons the Poisson probabilities of each gridpoint are then processed as described in formula 2.1.

$$R = log_{10}(\frac{1}{P(x \geq n)}) \tag{2.1}$$

The effect of the computations of this step on the search spheres is shown in Figure 4. Potential source regions containing more neutrinos should be linked to higher values on these spheres, low values can be assumed not to be linked to detectable sources. Separating background from potentially relevant information is called segmentation. In this search this is done by the application

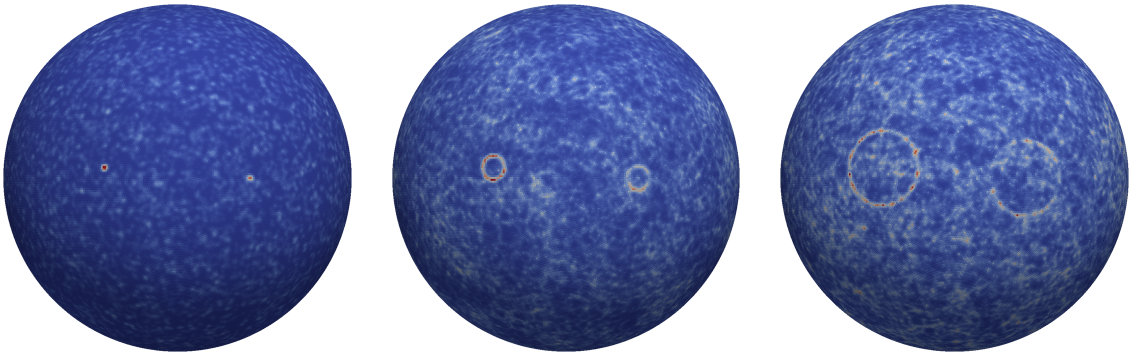

Figure 4: Left: The spherical grid for the scale 0.0° to 0.5° after computation of the Poisson probabilities. Middle: 3.0° to 3.5°. Right: 10.0° to 10.5°.

of a threshold to all R values. The threshold is derived from the histogram of all observed R values. It it set to the beginning of the tail of the distribution. The procedure is visualized in Figure 5. The result after the segmentation is shown in Figure 6. The next step is to reconstruct

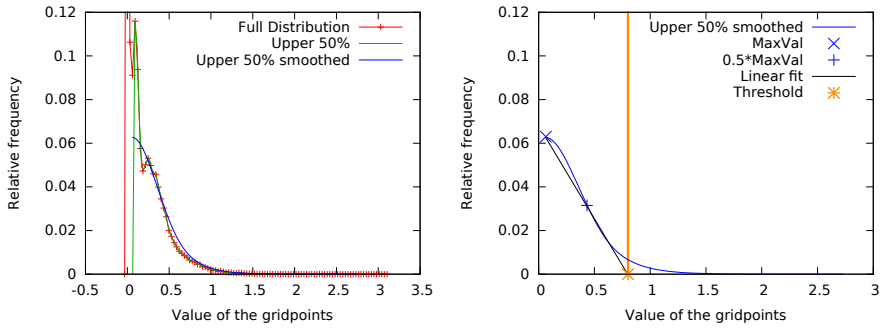

Figure 5: Left: The histogram of R values in red and various intermediate steps of the threshold computation. Right: Various intermediate steps of the threshold computation and the final threshold in orange.

the original location of the neutrinos that caused the detected overfluctuations. For the search sphere with a search distance between 0.0° and 0.5° nothing changes, since the neutrinos have been counted at the location where the information is stored. For all other scales $d > 0.5°$, the information stored at a gridpoint originated from counting neutrinos that are d degrees away. To achieve the remapping of the information to the original location a second grid is initialized with 0.0 values. For each gridpoint p in the original grid, the set p_d of all gridpoints at a distance d around it is computed. For each of the gridpoints within this set, the mean contribution of a gridpoint at this distance to the observed value $\frac{R_p}{size(p_d)}$ is added (in the new grid). Afterwards the new grid contains the corresponding fractions of the overfluctuations mapped back to their origin and this grid is used from there on. The result of these computations is shown in Figure 7. The information where the neutrino distribution had a higher density is automatically encoded in the pattern how the remapped circles around the old gridpoints overlap in the new grid, see middle and right of Figure 7. In order to evaluate the 180 different search scales they have to be combined in some way. The best successfully developed robust solution turned out to be taking the

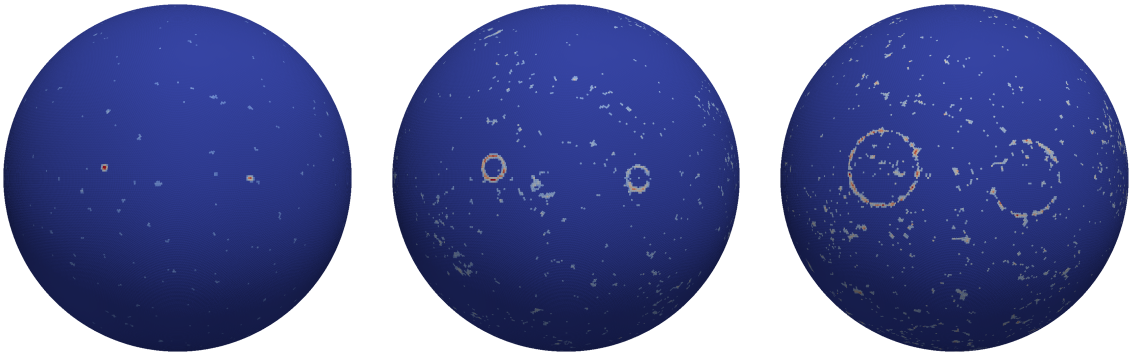

Figure 6: Left: The spherical grid for the scale 0.0° to 0.5° after segmentation. Middle: 3.0° to 3.5°. Right: 10.0° to 10.5°.

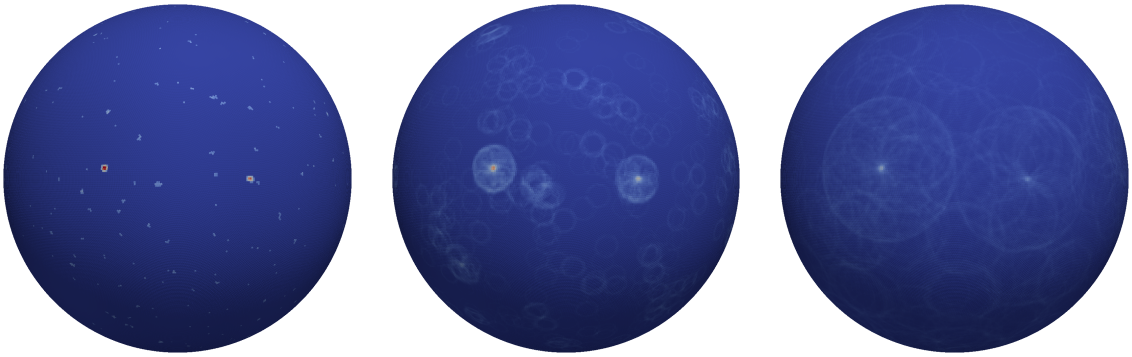

Figure 7: Left: The spherical grid for the scale 0.0° to 0.5° after the remapping. Middle: 3.0° to 3.5°. Right: 10.0° to 10.5°.

sum of all scales. It should be noted that there is the potential to exploit the available information better, for instance by an individual evaluation of the 180 spheres and a more sophisticated combination of the information derived from each. The result of the summation can be seen in Figure 8.

Identifying connected regions with high values on the final sphere, which could be linked to possible neutrino source morphologies and strengths, is again achieved by a segmentation as already described for the individual scales. The same procedure is used, but this time with the additional option to obtain different thresholds by scaling the distance between the previous minimum value x_{min_old} and θ, using a factor α. The new threshold is then given by equation 2.2.

$$\theta_{final} = (\alpha \cdot x_{min\_old} + (1-\alpha) \cdot \theta) \tag{2.2}$$

The effect of different thresholds for segmentation is shown in Figure 9. High positive values for α allow larger extended source structures to be found, hard negative cuts only preserve the high peaks. Multiple values pronounce different aspects

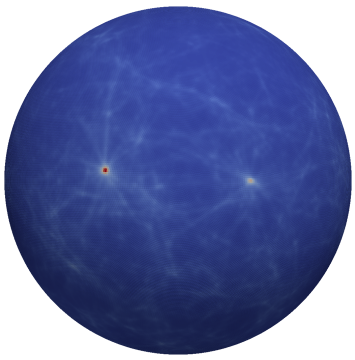

Figure 8: The sum of all 180 search scales.

of the obtained result, but on the other hand they also increase the trial factor for the final result. By heuristic optimization based on a variety of simulated sources, the values for the segmentation have been fixed to α = 0.25 and α = -0.11, since a single value cannot cover the targeted range of sources.

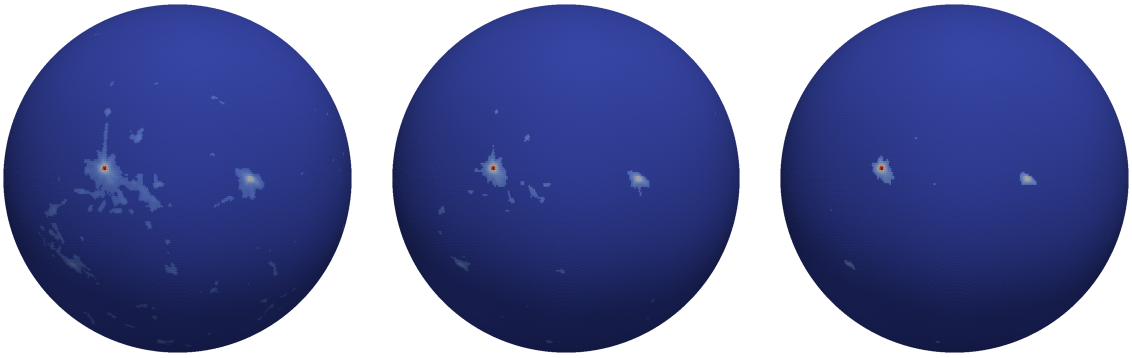

Figure 9: The effect of different α values on the segmentation. Left: +0.25, Middle: +0.05, Right: -0.1.

The gridpoints of the resulting segmented grid are checked for connectedness. A connected group of gridpoints is called a cluster. The next step is to distinguish potentially significant clusters from random accumulations. To achieve this one needs to know the probability how likely a cluster could have been generated by random events. This probability could in theory be determined by pseudo-experiments using scrambled data. A specific cluster shape, position and composition is unlikely to be reproduced, therefore the analysis must rely on more generic attributes to evaluate the significance of a cluster. For instance one can compute the probability for a cluster of the same size or larger, with size measured by the number of gridpoints. Only considering size for a relevance measurement is not sensitive to smaller or even point-like sources. A better metric to find these is for example the maximal value of any gridpoint in the cluster. Many others metrics have been tested, each with different sensitivities to different source characteristics. But if many relevance metrics are used, also a large trial factor has to be considered. Since this search is not intended for a specific source model, the optimized selection had to be done heuristically with a multitude of simulated sources. To specifically detect point sources one would use e.g. the maximal value within a cluster. Justified by the fact that ANTARES has already conducted specialized searches for promising small and point-like sources, the metric size in gridpoints N has been chosen, performing best for large, extended source morphologies. Due to the increased trial factor that comes with more metrics, the second best, the mean value of the $\sqrt{N}$ highest pixels within a cluster, is not included. The significance for each cluster is then derived from pseudo-experiments with scrambled data.

3. Selectfit

A new method for the direction reconstruction of events detected by the ANTARES neutrino telescope called "SelectFit" is introduced here. Instead of using one reconstruction algorithm for the whole sample, Selectfit combines the results of multiple available reconstruction algorithms, trying to select the reconstruction algorithm for each event, which gives the most precise result. This selection is done by a machine learning technique called “Random decision forest” (RDF)

[6]. Selectfit decreases the angular reconstruction uncertainty of a sample of neutrinos or allows to increase the sample size for a fixed angular uncertainty, as illustrated in Figure 10. Since a search for extended objects does not need the same angular precision, which is required for a point source search, less strict quality cuts for the reconstruction can be applied. Together these two aspects increase the number of neutrino events in this analysis compared to the standard ANTARES point-source search.

4. Results

The unblinding of ANTARES data from 2007 to 2012 resulted in 13283 neutrino event candidates. The analysis of these events with the method explained in chapter 2 yielded the preliminary results shown in Figure 11. Using the harder segmentation threshold α = -0.11 no cluster with a significance above 0.8σ has been found. With α = 0.25 a very large structure is found. A wide range of checks for systematic effects that could possibly influence the result has been performed, including for instance the small effect of time variations in the data taking efficiency on the event distributions. After accounting for all known systematic effects, the large structure has a post-trial significance of 2.5σ. It contains the galactic center, which is located in the center of skymaps in galactic coordinates. More details how these structures have formed can be seen in Figure 12, which shows the result of the summation of all scales before the segmentations. One has to keep in mind that the exact borders of these structures are certainly influenced by random fluctuation.

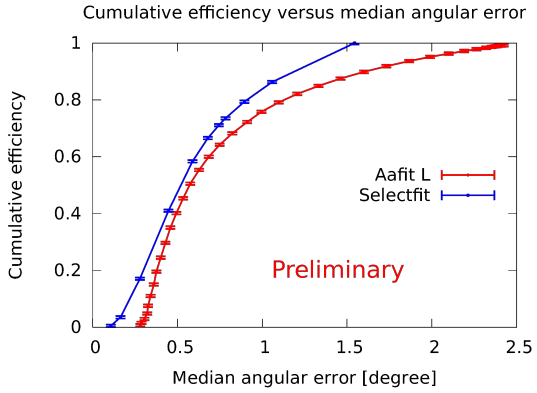

Figure 10: Comparison of the performance of the best individual direction reconstruction algorithm "Aafit" and "Selectfit", the introduced method to efficiently combine multiple reconstruction algorithms.

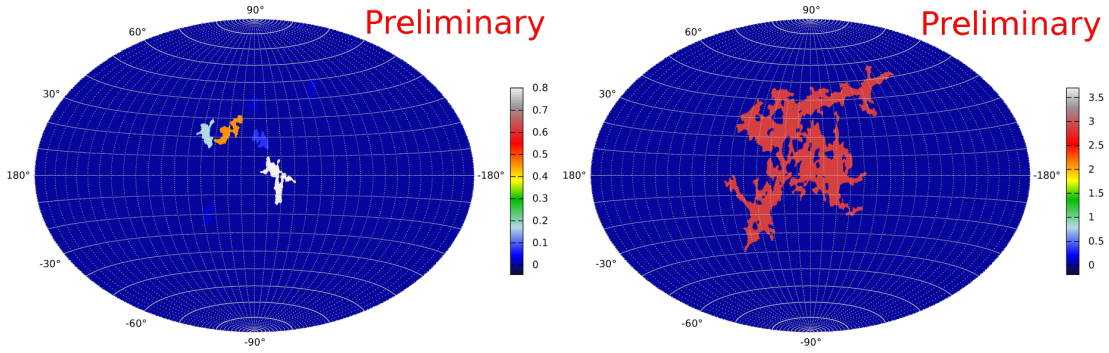

Figure 11: Left: The result on ANTARES data with segmentation using α = -0.11 in galactic coordinates. Right: The result with α = +0.25. The color code of the clusters shows the significance in σ.

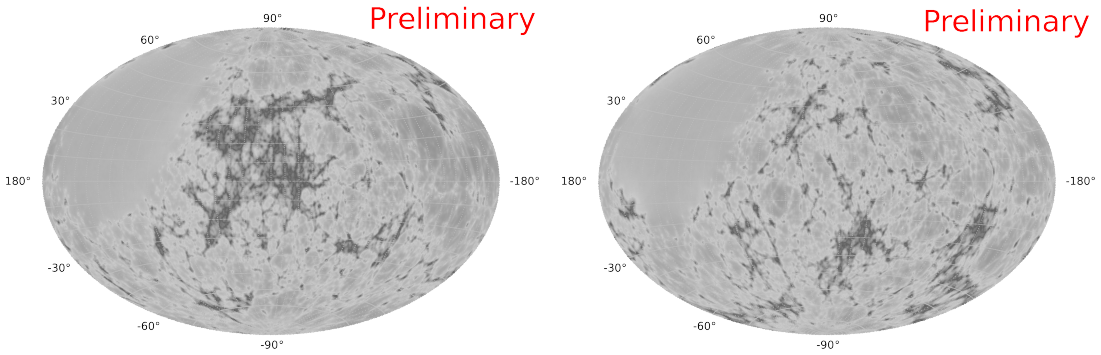

Figure 12: Left: The detailed structure behind the result on ANTARES data in galactic coordinates before the segmentations in arbitrary units. Right: An example for the detailed structure for a random dataset with the same color code for comparison. The observed intensity of the overfluctuation is common, the clustering of their locations is rare.

5. Conclusions

We have devised a new analysis method that is able to detect sources of arbitrary location and morphology without relying on assumptions on the source properties. Applied to ANTARES data this method found a large structure with a post-trial significance of 2.5σ. This preliminary result is consistent with a random background fluctuation.

Even though an unspecific, model-independent analysis is unlikely to obtain a result with high statistical significance due to the high trial factors, these searches provide a good way to become aware of the most interesting structures in data, which then may be worth further, more specific investigations.

References

[1] Clancy James et al, *Highlights from ANTARES, and prospects for KM3NeT.* 34^{th} International Cosmic Ray Conference, ICRC2015-I/476, 2015.

[2] S. Adrian-Martinez et al, *Search for cosmic neutrino point sources with four year data of the ANTARES telescope.* Astrophysics J., 760:53 (2012)

[3] S. Adrian-Martinez et al, *Search for muon neutrinos from gamma-ray bursts with the ANTARES neutrino telescope using 2008 to 2011 data.* A&A 559, A9 (2013)

[4] S. Adrian-Martinez et al, *Searches for Point-like and Extended Neutrino Sources Close to the Galactic Center Using the ANTARES Neutrino Telescope.* The Astrophysical Journal Letters, 786:L5, 2014

[5] S. Adrian-Martinez et al, *A search for neutrino emission from the Fermi bubbles with the ANTARES telescope.* Eur. Phys. J. C (2014) 74:2701

[6] Tin Kam Ho, *The Random Subspace Method for Constructing Decision Forests.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 20:832-844, 1998