ISTITUTO NAZIONALE DI FISICA NUCLEARE

Sezione di Genova

INFN-12-11/GE
6th July 2012

A STUDY ON SPARKING EVENTS IN THE DATA ANALYSIS OF THE ANTARES DETECTOR

Matteo Sanguineti[1], Marco Anghinolfi[1], Evgeniya Davydova[2], V. Kulikovskiy [1]

[1)]*Università degli Studi di Genova, Dipartimento di Fisica and INFN-Sezione di Genova, Italy*
[2)] *Lomonosov Moscow State University, Russia*

Abstract

The ANTARES detector, completed in 2008, is the largest neutrino telescope in the Northern hemisphere. It is located at a depth of 2.5 km in the Mediterranean Sea, 40 km off the Toulon shore. The scientific scope of the experiment is very broad, being the search for astrophysical neutrinos as the main goal. In this note we analyse some events reconstructed with an anomalous high number of hits on PMTs. Likely these noise events are produced by light from electric discharge on PMT bases. We propose a method to distinguish these events from high-energy muon events.

PACS.: 07.05.Kf

*Published by **SIDS–Pubblicazioni***
Laboratori Nazionali di Frascati

1 INTRODUCTION

The ANTARES Collaboration has constructed a neutrino telescope in the Mediterranean Sea$^{1)}$. The main aim of the project is the search for high-energy neutrinos of astrophysical origin. This is achieved by the detection of Cherenkov photons induced by the passage of relativistic charged particles resulting from neutrino interactions in the material surrounding the detector. The most important channel is the charged current interactions producing muons. Other signatures, such as the cascades produced both in charged and neutral currents are also detected. These photons are detected in a large three-dimensional array of 885 photomultipliers (PMTs) installed on twelve vertical lines anchored on the sea bed at a depth of 2475 m.

The data acquisition system of the detector is based on the ”all-data-to-shore” concept, in which signals from the photomultipliers above a given threshold are digitized and sent to shore for processing. Here a computer farm filters the data with specific triggers and possible muon tracks are reconstructed from the hits. Data are organized into runs of the duration of few hours. These runs are used for all physics analysis performed in the collaboration (search for point and diffuse source, measure of flux from GRB and supernovae, etc…).

The consistency check of events from different runs before any analysis is fundamental, so the ANTARES collaboration produced a complex process for data quality assessment. Many parameters are used to evaluate the quality of the track reconstruction:

- Λ is the likelihood value of the reconstructed track divided by number of degree of freedom (maximum likelihood is obtained with AAfit algorithm$^{2)}$)
- β is the angular error of the reconstructed track (AAfit)
- χ^2 is the chi square value of the reconstructed track (chi square is obtained with BBfit algorithm$^{3)}$)

The aim of this study is the development of an algorithm that can recognize events reconstructed with an anomalous number of hit on PMTs. Likely this events are not physical but are due to the light emission by electrostatic discharge from the phototube.

This work is really important in ANTARES analysis because also high-energy events have a high number of hits, so the distinction of these two kinds of signal is fundamental.

2 THE ANTARES RECONSTRUCTION ALGORITHM

The reconstruction of muon tracks is based on the arrival time of the Cherenkov photons on the PMTs: two different algorithms are used in the ANTARES detector:

- AAfit algorithm consists of multiple fitting steps. The final step is based on a full likelihood description of the detected hits, which also accounts for background light$^{2)}$.
- BBfit algorithm is based on χ^2 minimization of the reconstructed track$^{3)}$.

Figure 1 shows an event reconstructed as a downgoing muon with BBfit algorithm.

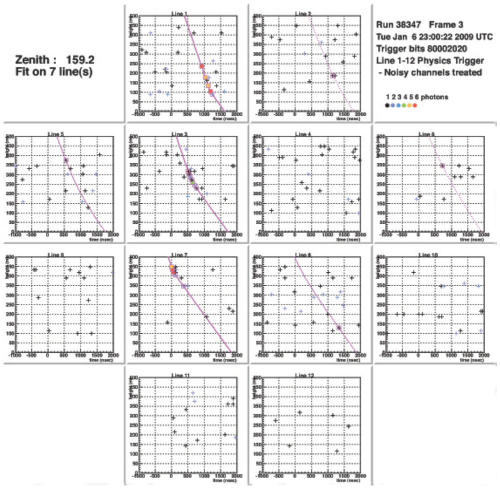

FIG. 1: Typical event reconstruction in the ANTARES detector.

Besides the already mentioned quality parameters, both algorithms give various track information: azimuth and zenith of the track, number of hits on PMTs used reconstructing the event, the arrival time, etc…

The number of hits is really important because it is correlated with the energy of the event so this parameter is used for the energy estimation of the tracks. However the sparking events could look like high-energy events causing errors in the analysis. That is why we require a good algorithm selection of sparking events.

Figure 2 shows a typical distribution of the number of hits per events used in the AAfit reconstruction strategy (three years diffuse flux analysis).

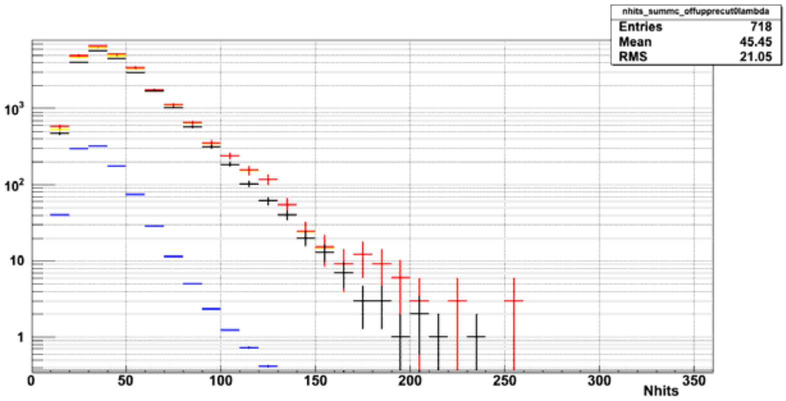

FIG. 2: Number of hits distribution (Λ >-6, β<1°, zenith>90°). Black points are the data, yellow points are atmospheric muons, blue points are atmospheric neutrinos and red points are the sum of atmospheric neutrinos and muons.

There is a good agreement between Monte Carlo simulation and data, except in the region above nhits□100, where Monte Carlo overestimates data. On average events are reconstructed with 45 hits.

In the following section the new sparking runs selection algorithm will be described.

3 SELECTION OF SPARKING RUNS

A small number of runs contain events with exceptionally high hit multiplicity. Nevertheless a rare occurrence in time, a sparking PMT could produce a signal with a similar profile to that of a neutrino-induced shower: it is therefore important be able to distinguish the sparking noise events from the real one.

A study on sparking runs of data production “2011-05” was performed for the period 2008-2010.

First we have created a simple algorithm to exclude from the analysis the runs where noise-sparking events are dominant. In this approach we have separated the events reconstructed with a number of hits fewer than 200 with respect to the events reconstructed with a number of hits greater than 200.

Then we have calculated the following ratio:

$$S=\frac{\sum\limits_{events\ with\ nhits<200} n_{hits}^2}{\sum\limits_{events\ with\ nhits>200} n_{hits}^2}$$

In this formula we calculate the square of the number of hits in order to give more weight to the events with high number of hits. Using this simple algorithm, in fact, the ratio S turns out to be different for sparking runs (typically 0.01-1) and from normal runs (typically greater then 100).

The algorithm was tested using known sparking runs found in previous analysis using a spherical shape events fit$^{4)}$. Our algorithm recognized these runs. In addition to that some unknown sparking runs were found:

- run 42746 is clearly sparking (Figure 3).
- run 34665 is probably sparking (also run 34663 is known as sparking)(Figure 4).
- some low statistics runs, which have some unusual events with high values of number of hits, have been found (42513,42746,42915,42919,46980, 51036, 53508, 53851)(Figure 5).

On average the percentage of normal runs evaluated as sparking by our algorithm is about 1 ‰.

For an immediate comparison the same distribution is shown for a normal run (Figure 6)

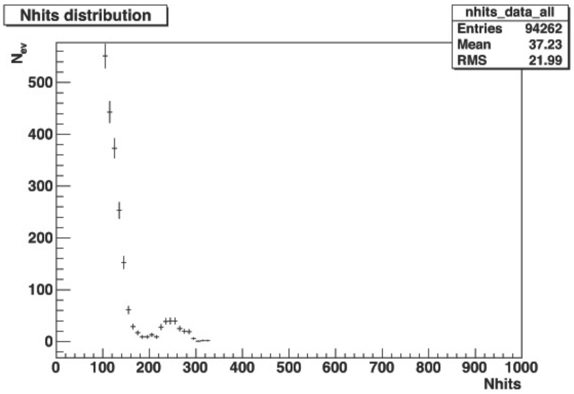

FIG. 3: Number of hits distribution of run 42746.

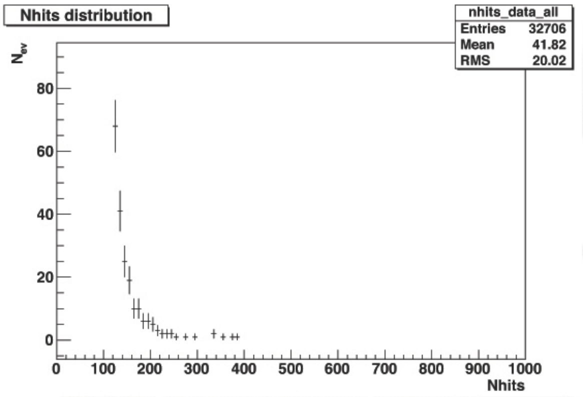

FIG. 4: Number of hits distribution of run 34663.

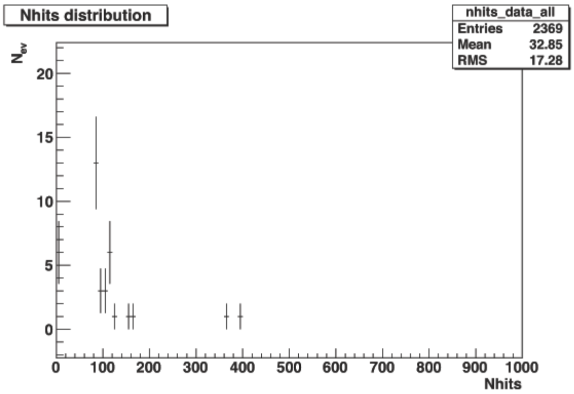

FIG. 5: Number of hits distribution of run 51036.

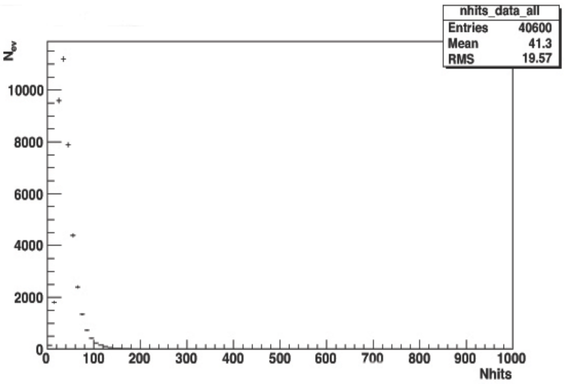

FIG. 6: Number of hits distribution of a normal run.

This algorithm seems to be more effective rather than the spherical shape fit. However events with high number of hits in low statistics runs are really critical; it is not easy to say if they are sparking events or high-energy events. So a new way to identify these kind of events have been sought.

To start with new effective correlations between the number of hits and other track parameters has been studied.

Figure 7 shows the correlation between number of hits and Λ for a clearly sparking run.

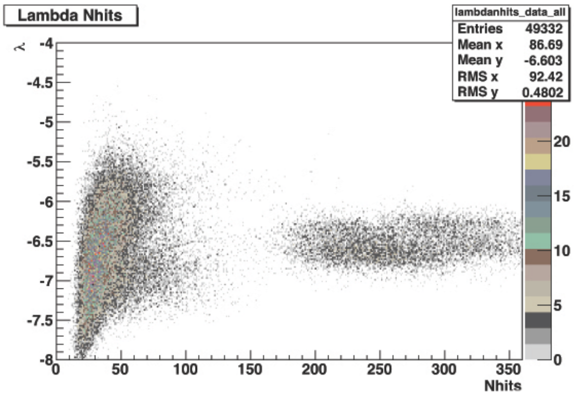

FIG. 7: Correlation between number of hits and Λ for a run clearly sparking.

The value of Λ for sparking events is around -6.5 (right part of figure 7), not so different from typical events.

This behaviour is certainly not physical because such events are not present in the Monte Carlo of the run shown in figure 8, the reduced number of events in the simulation being optimized for a sensible comparison.

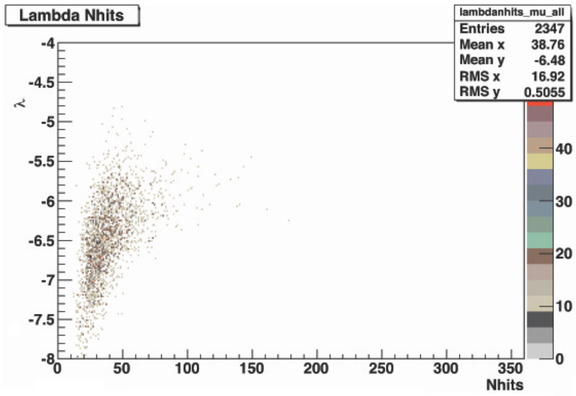

FIG. 8: Correlation between number of hits and Λ for a run clearly sparking (Monte Carlo simulation)

A similar result is obtained in the correlation between the number of hits and the angular reconstruction error β(Figure 9).

The angular error of sparking tracks is comparable to typical tracks. However, using the bbfit χ^2 , a considerable distinction can be seen between the two types of events. (Figure 10).

Selecting tracks with χ^2 less than 10, we eliminate most of sparking events and part of events clearly badly reconstructed.

This selection can be helpful, especially when few sparking events in many different runs can make a cumulative effect over the long term. This is demonstrated in figure 11, where the correlation between Λ and the number of hits is shown in a three years data taking period excluding sparking runs from the list and without cut on χ^2. The events in the red circle are probably sparking, because they have the same Λ and the number of hits of clearly noise-events in sparking runs.

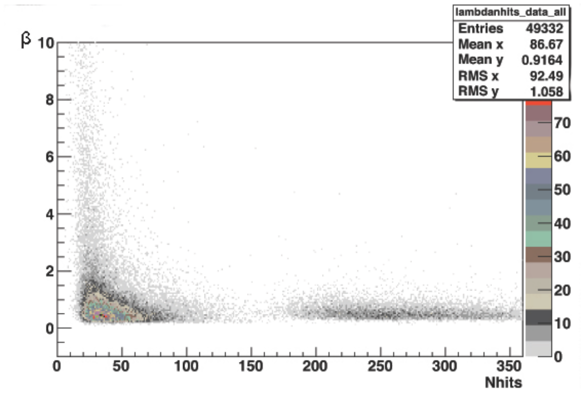

FIG. 9: Correlation between number of hits and β for a run clearly sparking.

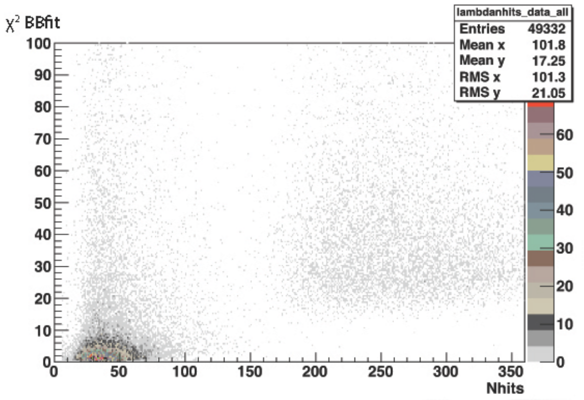

FIG. 10: Correlation between number of hits and β for a run clearly sparking.

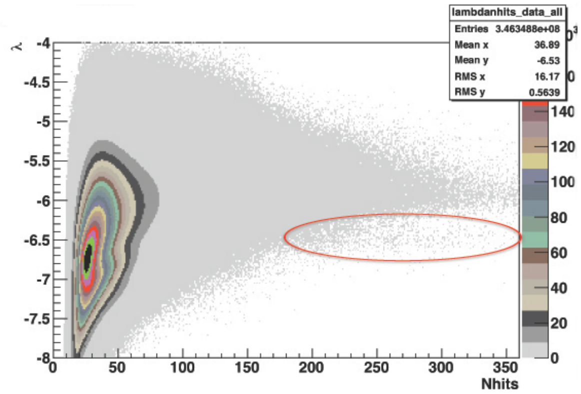

FIG. 11: Correlation between number of hits and Λ for a run clearly sparking (three years excluding sparking runs and without cut on χ^2).

This region around Λ= -6.5 and number of hits from 200 to 350 (red circle), is absent in the MC simulation. (Figure 12)

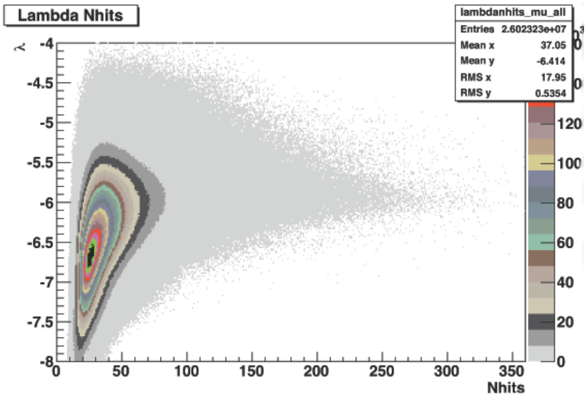

FIG. 12: Monte Carlo simulation of correlation between number of hits and Λ for a run clearly sparking (three years excluding sparking runs and without cut on χ^2).

However, selecting only events with χ^2 less then 10, most of events in the red circle of figure 11 are removed, as shown in figure 13.

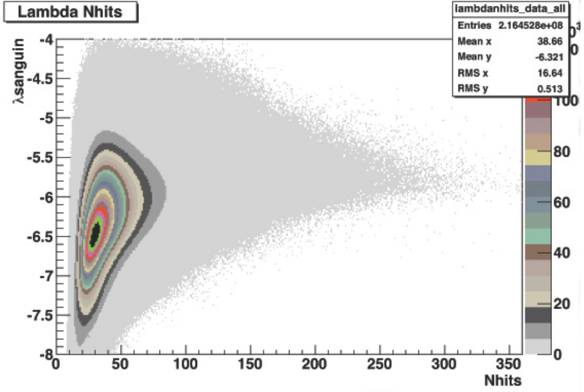

FIG. 13: Correlation between number of hits and Λ for a run clearly sparking (three years excluding sparking runs and with χ^2>10).

We have also studied the distribution of these sparking events as a function of the run number to check if they are grouped in time or have a uniform distribution. The result, shown in Figure 14-15 for the 2008-2010 period, indicates that these sparking events occur almost uniformly in time.

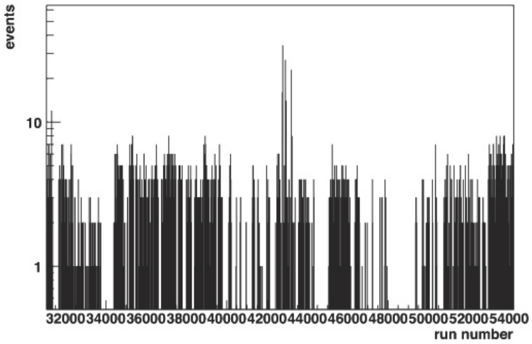

FIG. 14: Number of anomalous events in each run (three years excluding sparking runs and only events with χ^2>10, number of hits >250)

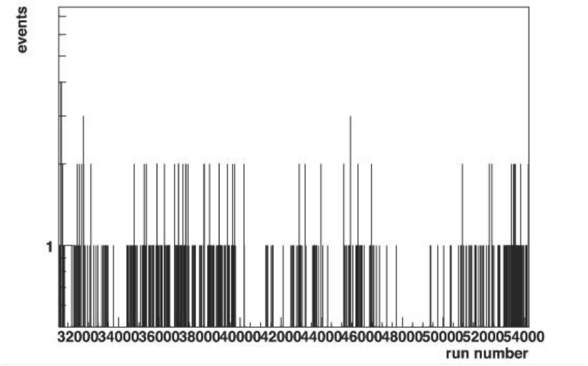

FIG. 15: Number of anomalous hits in each run (three years excluding sparking runs and only events with $\chi^2 > 10$ and number of hits >350)

Then we looked at the topology of this events (Figure 14-15) with high number of hits and $\chi^2 > 10$. Some different kind of events were found and are presented on following figures (16-19)

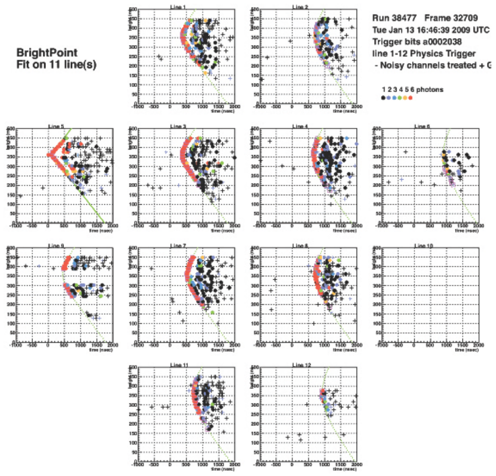

FIG. 16: Topology of an anomalous event found in “no sparking” runs

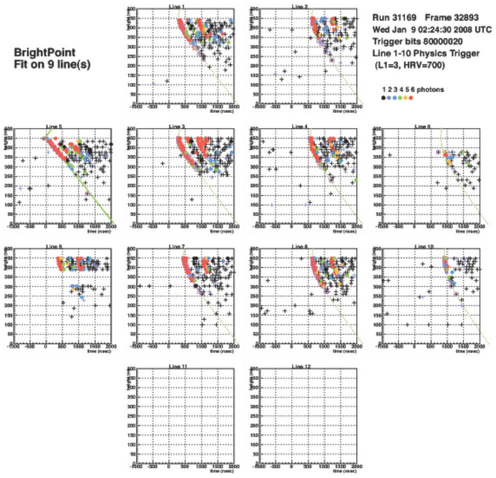

FIG. 17: Topology of an anomalous event found in “no sparking” runs

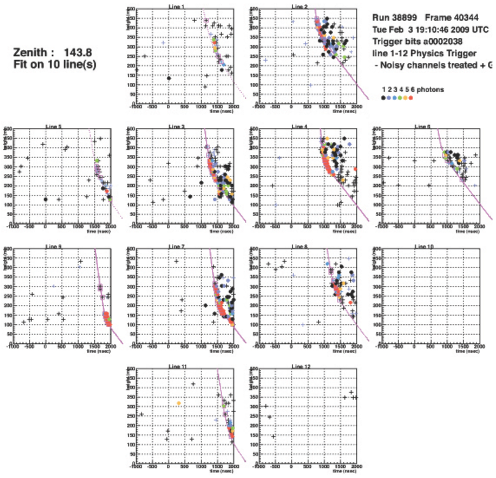

FIG. 18: Topology of an anomalous event found in "no sparking" runs.

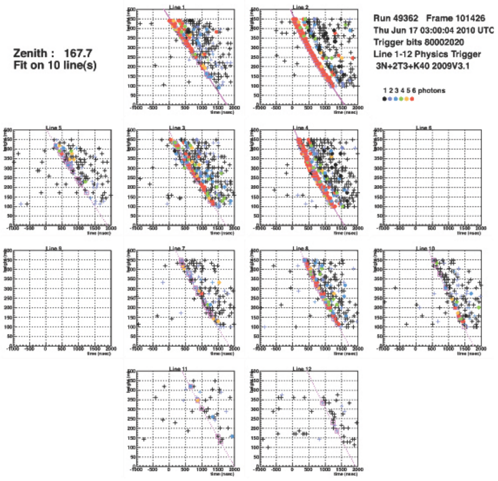

FIG. 19: Topology of an anomalous event found in "no sparking" runs.

The track in figure 14 looks to be a typical sparking event, while the interpretation of the other figures is not clear: some are reconstructed as bright points, but other as tracks.

Another particular aspect of the sparking runs is the direction of events. In several sparking run we have noticed a preferred suspicious direction as regard zenith and azimuth angles: the reason seems to be ascribed to the reconstruction mechanism but we have no clear evidence for that. This effect is well evident in figures 20 and 21.

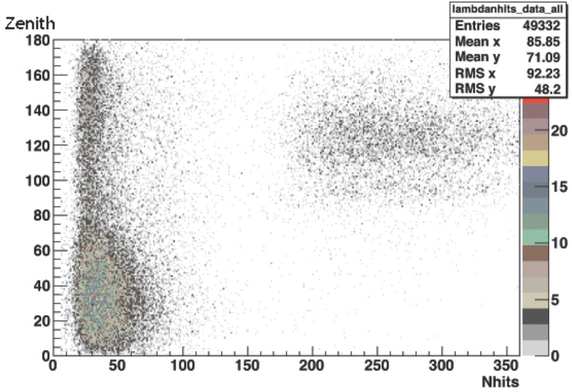

FIG. 20: Correlation between number of hits and zenith angle for a run clearly sparking.

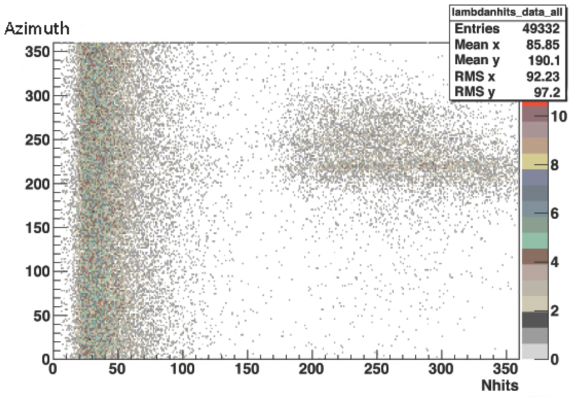

FIG. 21: Correlation between number of hits and azimuth angle for a run clearly sparking.

This feature isn't intuitively explained and requires further investigation.

6 CONCLUSIONS

The aim of this note was to investigate the so called 'sparking' events in the ANTARES data and possibly to find an algorithm to isolate them. Some goal have been achieved:

- A new algorithm for sparking run detection was developed
- Some new sparking runs candidates were found
- Evidence of sparking events and other anomalous events in "no sparking" runs was found.
- A cut on χ^2 (Bbfit strategy) was proposed in order to remove sparking events in "no sparking" runs

But there are other problems which are still open:

- Interpretation of anomalous events topology
- Understandings of sparking events angular correlation, despite those events are far apart in time.

These last considerations indicate that deeper analysis possibly based on the topology of the single sparking event has to be investigated.

For the moment we suggest to:

- exclude the new sparking runs candidates from all physics analysis
- include the χ^2 cut into the "standard" events selection for all physics analysis

7 REFERENCES

(1) ANTARES collaboration, ANTARES: the first undersea neutrino telescope, arXiv:1104.1607.

(2) A. Heijboer, Track reconstruction and point source searches with ANTARES, ANTARES-06-08-2004.

(3) Jürgen Brunner, The BBfit Reconstruction algorithm, ANTARES-SOFT-2009-003.

(4) Laura Core, private comm.