Model independent search for transient multimessenger events with AMON using outlier detection methods

Timothée Grégoire on behalf of the AMON group

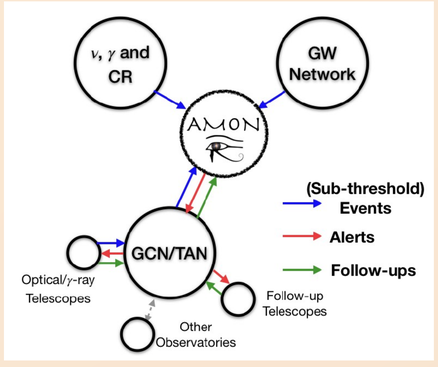
Abstract

The Astrophysical Multimessenger Observatory Network (AMON) receives subthreshold data from multiple observatories in order to look for coincidences. Combining more than two datasets at the same time is challenging because of the range of possible signals (time windows, energies, number of events...). However, outlier detection methods can circumvent this issue by identifying any signal divergent from the background (scrambled data).

We propose to use these methods to make a model independent combination of the subthreshold data of neutrino and gamma ray experiments. Using the python outlier detection (PyOD) package, it allows us to test several methods from a simple ``k-nearest neighbours'' algorithm to a more sophisticated Generative Adversarial Active Learning neural networks which generates data points to better discriminate inliers from outliers.

AMON

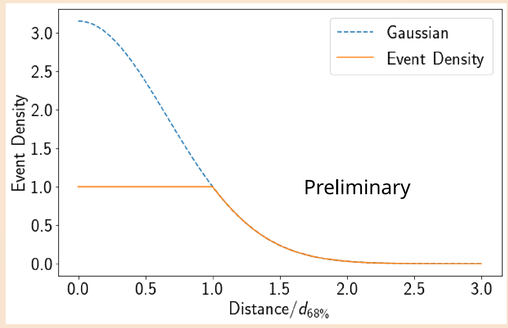
- The **Astrophysical Multimessenger Observatory Network** (AMON) receives in **realtime**:
  - **IceCube** subthreshold singlets and high energy “Gold and Bronze” tracks as well as cascades, **ANTARES** tracks, **HAWC** hotspots and HAWC bursts, **Fermi-LAT** data
- Subthreshold datasets are combined in realtime
  - Statistically significant signals are send publicly to the **Gamma-ray Coordinate Network (GCN)**
  - Small field of view instruments can **follow up** the alerts



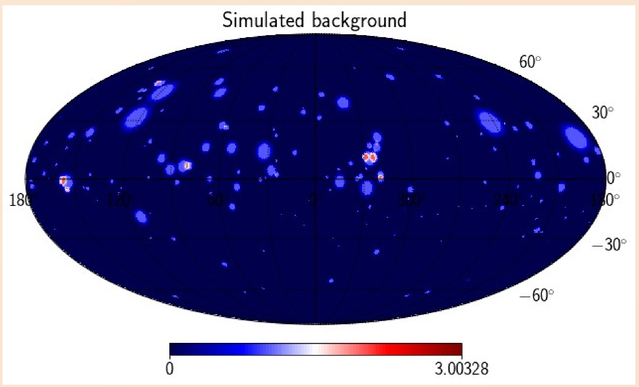
- Data are stored on the AMON servers for **archival analyses**
- Study **the most energetic phenomena** in the universe and answer **fundamental questions of astrophysics and cosmology**

The Search Method

- Search for **coincident signal in multiple datasets**
  - AMON is designed to **combine more than two datasets**
  - Method mostly **independent from the datasets** used as inputs
- It is not feasible to simulate realistically all possible signal combinations
  - Multiple datasets, many models, many unknowns...
- **Outlier detection methods**
  - Learn the background and classify divergent data points as signal
  - Model independent search
- Background is easy to simulate by scrambling the data
- Produce **skymaps of event density** per time step of 6h from events list
- Event density
  - 1 inside the 68% error contour
  - Gaussian otherwise



Skymap of the event density of a simulated background for illustrative purpose



- Input data points of the algorithm correspond to each pixels of each time steps
- Each data point contains
  - **N event densities** for the N datasets to combine
  - **Altitude, azimuth** of the pixel seen from 0°N, 0°E
- **Test of several** outlier detection **algorithms** from PyOD library
  - K-nearest neighbours, Principal Component Analysis, AutoEncoder, Multiple Objective Generative Adversarial Active Learning...
- **Signal simulation** to choose best algorithm
  - Inject coincident events in multiple datasets
  - Not representative of all possible signal
    - Cannot use it to quantify the sensitivity
    - Gives a proof of concept

Status and Perspectives

- We plan to use this method for the combination of five datasets:
  - the **ANTARES tracks**, **IceCube singlets**, **HAWC hotspots**, **HAWC bursts** and **Fermi-LAT** data
- Test other time steps and event density definitions
- On the longer term
  - Run this analysis in **real time**
  - Add **more datasets**