Methods for detecting UHE photons through direction-time clustering algorithms

C. Priyadarshi[1], J. Stasielak[1], N. Borodai[1], J. Pękala[1], D. Góra[1], M. Niechciol[2]
[1]Institute of Nuclear Physics Polish Academy of Sciences, PL-31342 Krakow, Poland
[2]Center for Particle Physics Siegen, University of Siegen, Siegen, Germany

References

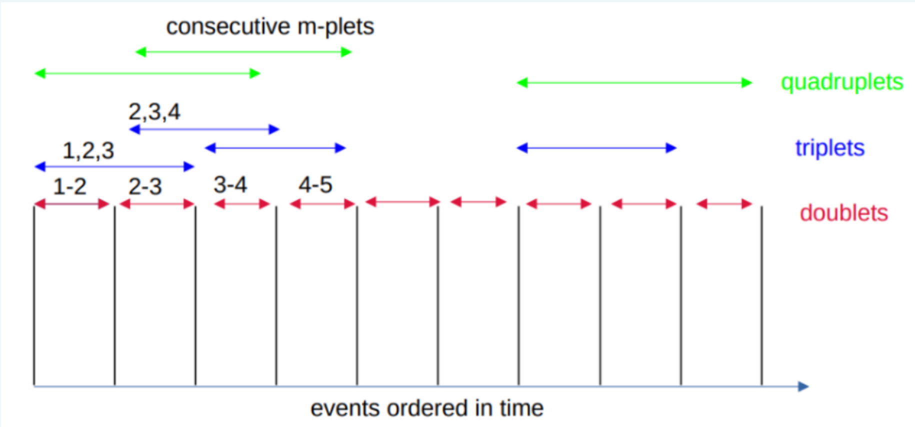
Motivation

Astrophysical flares are among the prime candidates for the production of ultra-high-energy (UHE, $E > 10^{17} eV$) cosmic rays. The search for the sources of UHE neutral particles offers distinct advantages over those of UHE charged particles, as the former traverse cosmic distances undeflected by magnetic fields, and thus may be detected as groups of clustered events, correlated temporally and directionally.

Standard clustering method

This algorithm uses the unbinned likelihood method [1] + a search for time correlation. This is carried out in the following steps,

- For a dataset with N events and spread across ΔT period, if there are n flare signal events, evaluate the space-time PDF of both signal, s_i, and background, b_i events,



- Evaluate the likelihood for each multiplet (doublets, triplets, quadruplets, etc) combination in the dataset,

$$\mathcal{L}(\mathbf{n}, \mathbf{\Delta t_j}, \vec{r_s}) = \prod_{i=1}^{N} \left(\frac{n}{N} \mathbf{s_i} + \left(1 - \frac{n}{N}\right) \mathbf{b_i} \right)$$

- and maximize it to get the test statistic,

$$\mathbf{TS(n)} = -2 \log \left[\frac{\mathcal{L}(0, \Delta t_j, \vec{r_s})}{\mathcal{L}(n, \Delta t_j, \vec{r_s})} \right]$$

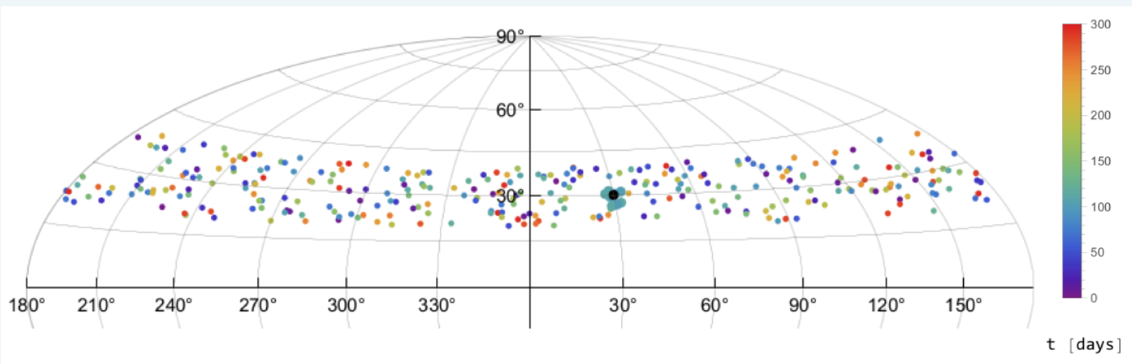
Combined signal PDF: $\mathbf{s_i} = \mathbf{s_i^{space}} \cdot \mathbf{s_i^{time}}$

Combined background PDF: $\mathbf{b_i} = \mathbf{b_i^{space}} \cdot \mathbf{b_i^{time}}$

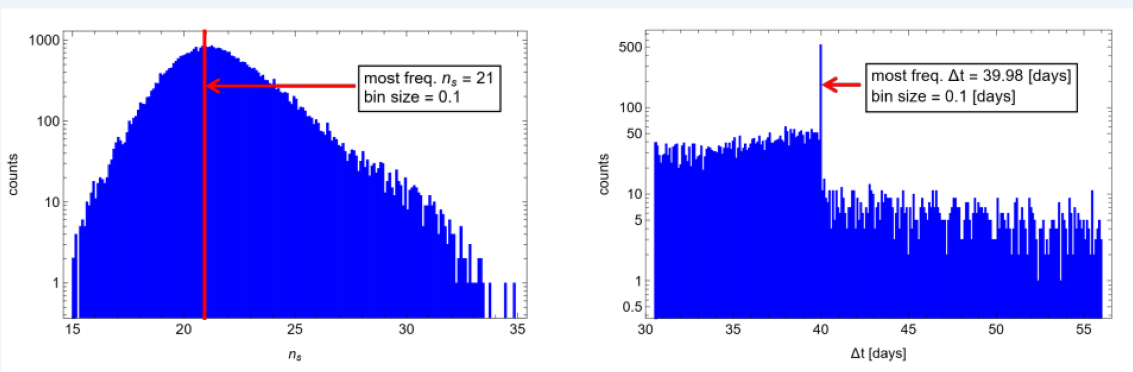
where, the angular resolution and the direction of each event are σ_i and r_i, within a time window of $\Delta t_j = t_j^{max} - t_j^{min}$.

Monte Carlo test

We randomly generated many scrambled sky maps with background and signal events, uniformly distributed over a sky region around an assumed flare location and duration.



For each map, we obtained the distributions of the estimators, $\mathbf{n_s}$ for the number of signal events, $\mathbf{\Delta T_{flare}}$ for the flare duration, and the value of the test statistic for $\mathbf{M_{opt}}$. The tests for all the algorithms can recover these estimators with good accuracy.

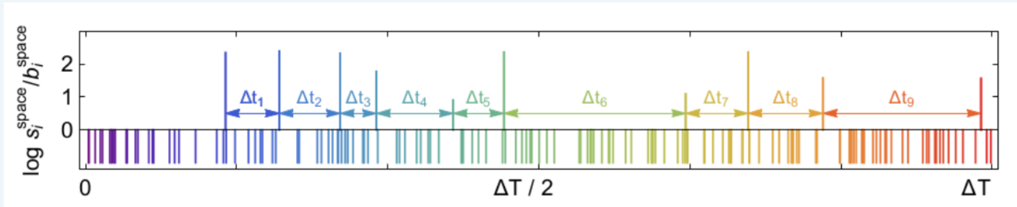


An example of 3 flares with a total of 20 signal events and 40 (20+10+10) days flare period.

Doublet stacking method

This method [2] adds the following steps to the standard method,

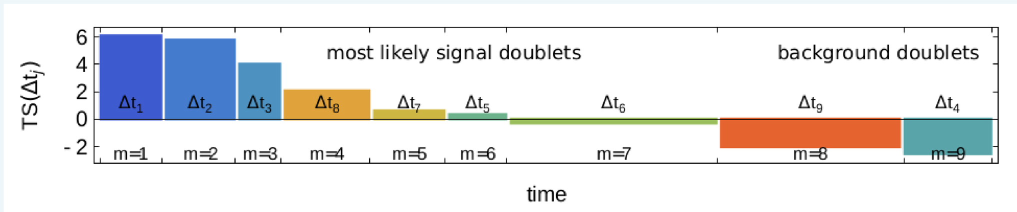
1: Selecting flare event candidates, only based on spatial information - $\mathbf{s_i^{space}/b_i^{space} > S/B}$ threshold



2: For each doublet j, maximize the $\mathbf{TS_{\Delta t_j}(n)}$ using the standard method, with an additional *marginalization* term to provide a more uniform exposure for finding doublets of different widths,

$$TS_{\Delta t_j}(n) = -2 \log \left[\frac{\Delta T}{\Delta t_j} \frac{\mathcal{L}(0)}{\mathcal{L}(n)} \right]$$

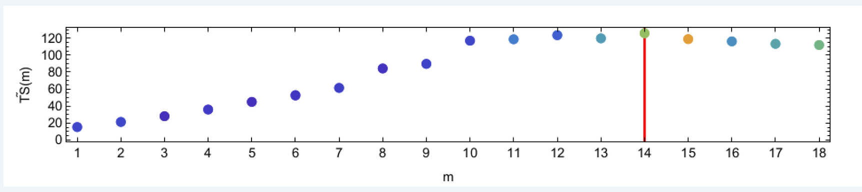
Next, sort the doublets according to the $TS_{\Delta t_j}$ and change the numbering of the doublets, introducing a multiplicity index, **m**.



3: Replace each event's signal PDF with a *weighted* sum of signal sub-terms over m doublets,

$$\mathbf{s_i} \rightarrow \mathbf{s_i^{tot}(m)} = \sum_{j=1}^{m} w_j s_i^j / \sum_{j=1}^{m} w_j \quad \text{where } w_j = TS_{\Delta t_j}$$

$$\Rightarrow TS \rightarrow \widetilde{TS}(m) = -2 \log[\mathcal{L}(0)/\mathcal{L}(n, m)]$$



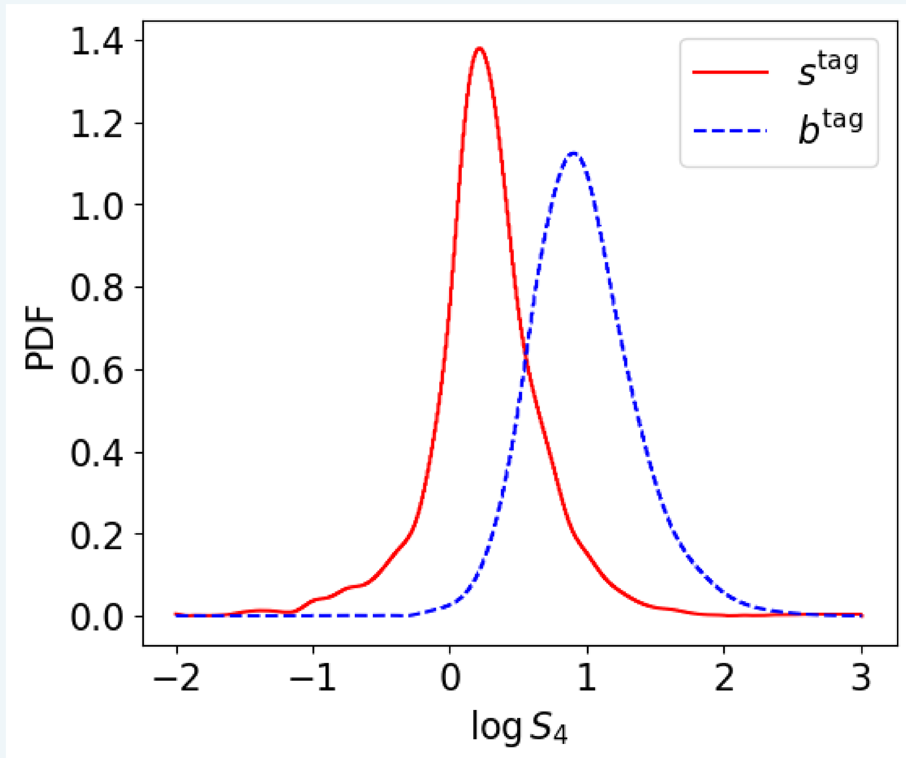
Maximize **m** to find the optimal total number of doublets in all flare/s, $\mathbf{M_{opt}}$, which determines the total flare duration. It is the number of most significant (**not necessarily consecutive**) doublets: $\Delta T = \sum_{m=1}^{M_{opt}} \Delta t_m$

Application of photon tag, $\mathbf{S_b}$

A common variable used to discriminate between photon- and hadron-initiated showers is the $\mathbf{S_b} = \sum_k S_k (\frac{R_k}{1000m})^b$ [3, 4], where S_k is the signal in the k-th detector, and R_k is the distance of the k-th detector from the shower axis.

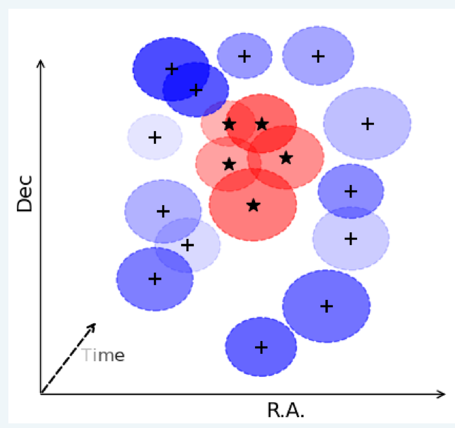
$$\mathbf{s_i^{space}} \Rightarrow \mathbf{s_i^{space}} \cdot \mathbf{s_i^{tag}} = \mathbf{s_i^{space}} \cdot \mathbf{s^{tag}}(\log \mathbf{S_4^i})$$

$$\mathbf{b_i^{space}} \Rightarrow \mathbf{b_i^{space}} \cdot \mathbf{b_i^{tag}} = \mathbf{b_i^{space}} \cdot \mathbf{b^{tag}}(\log \mathbf{S_4^i})$$



Python package, UHECluster

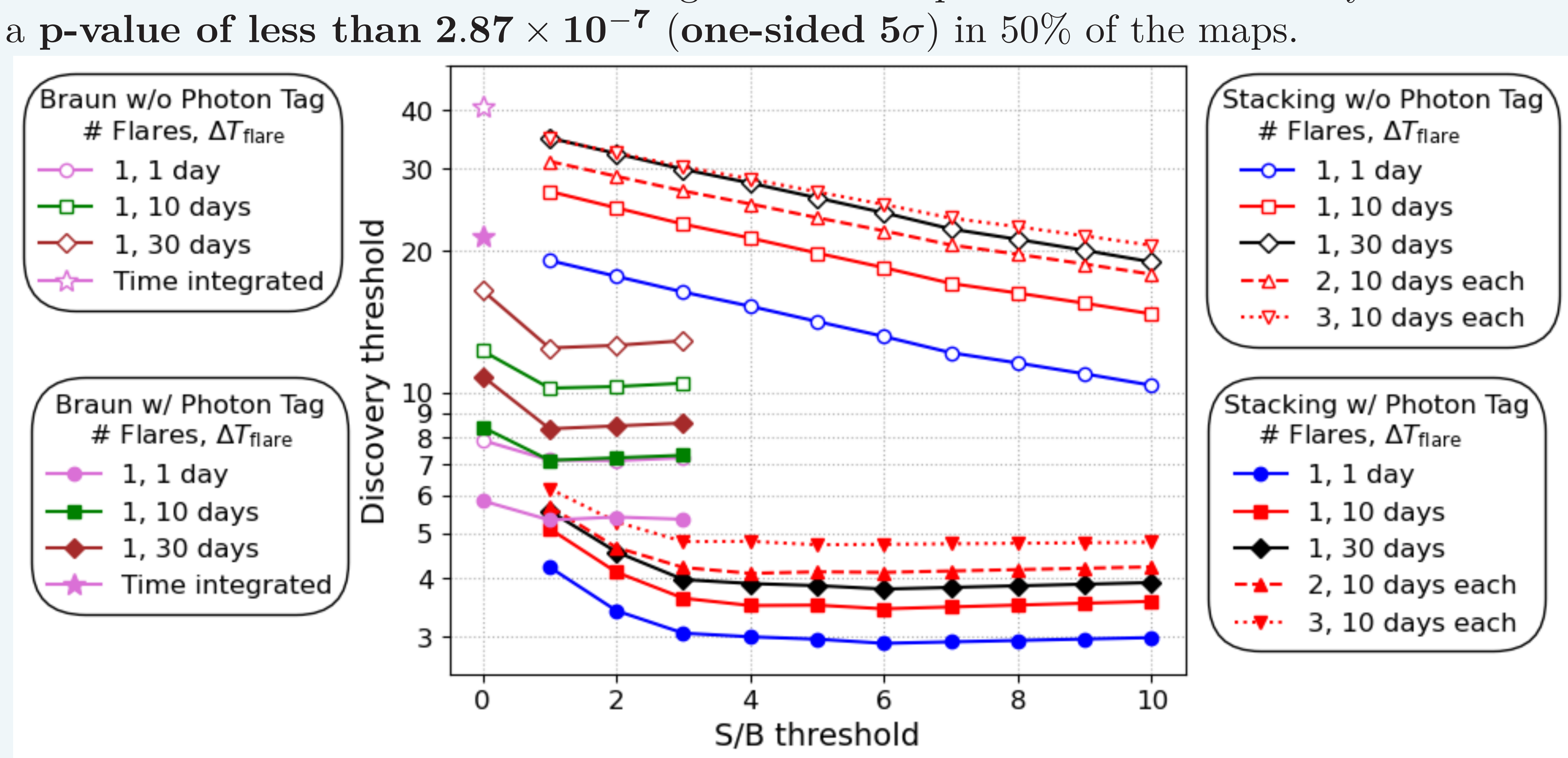
We created a Python package *UHECluster*, hosted at *gitlab*, to develop a uniform modular platform for running all the algorithms as simple Python scripts[5]. The results obtained from using the standard "Braun" method, with the application of the S_4 photon tag and a time-integrated (only directional correlation), are presented below, alongside earlier results obtained from the stacking method.



The different methods included in the package are categorized based on 1) core algorithm, 2) selection of background data, 3) modifications to the likelihood, and 4) modifications to the TS estimation.

Discovery potential of the clustering methods

Discovery potential is the **minimum** number of signal events required to claim discovery. This claim is generally assumed for a **p-value of less than 2.87** $\times$ $\mathbf{10^{-7}}$ (**one-sided 5**σ) in 50% of the maps.



Fewer events are required to claim discovery, even at a higher S/B threshold!

References and Acknowledgements

[1] Jim Braun et al. 2008. DOI: 10.1016/j.astropartphys.2008.02.007.

[2] D. Góra et al. 2011. DOI: 10.1016/j.astropartphys.2011.07.007.

[3] G. Ros et al. 2011. DOI: 10.1016/j.astropartphys.2011.06.011.

[4] The Pierre Auger Collaboration. 2022. DOI: 10.3390/universe8110579.

[5] Jaroslaw Stasielak et al. 2025. DOI: 10.22323/1.501.0400.

[6] Nataliia Borodai. 2025. DOI: 10.22323/1.501.0199.

The authors would like to thank their colleagues from the Pierre Auger Collaboration for all the fruitful discussions. We want to acknowledge the support in Poland from the National Science Centre, grants No. 2020/39/B/ST9/01398 and 2022/45/B/ST9/02163 as well as from the Ministry of Science and Higher Education, grant No. 2022/WK/12

Summary

We compare the results obtained from different combinations of the main clustering algorithms and demonstrate the advantages of using the one with smarter modifications. The stacking method is,

- ***faster*** than the standard method
- more ***sensitive*** to weak or **multiple** flares of any shape
- further enhanced in its performance with the addition of a ***better*** signal discriminator[6]

A discovery of UHE-neutral particles would set another milestone in the ever-advancing multi-messenger astronomy and broaden our understanding of astrophysics.