Improved method of searching for flares of neutral particles from point sources

Jarosław Stasielak[1], Nataliia Borodai[1], Dariusz Góra[1], Marcus Niechciol[2] and Jan Pękala[1]

[1] Institute of Nuclear Physics PAN, Kraków, Poland
[2] Department of Physics, University of Siegen, Siegen, Germany

Motivation

Astrophysical flares may be the sources of some cosmic rays, which, if they are neutral particles, should group into clusters of events correlated in space and time – pointing to their sources.

Search for such clustering in data would provide important evidence for the existence (or set upper limits on flux) of UHE ($E > 10^{17}$ eV) neutral particles.

Standard space-time clustering analysis

To identify flare/flares from a point source, we have to find an excess of events from a particular direction over the background. (J. Braun et al., Astropart. Phys. 29 (2008) 299 + time search)

Maximize the likelihood of possible multiplets in a data sample (doublets, triplets, quadruplets, etc.) and calculate test statistic:

- likelihood that flare consists of **n** signal events within a given multiplet time window ΔT_j :

$$\mathcal{L}(n) = \prod_{i=1}^{N}\left(\frac{n}{N}s_i + (1-\frac{n}{N})b_i\right)$$

- the test statistic: $\mathrm{TS_j(n)} = -2\log\left(\mathcal{L}(0)/\mathcal{L}(n)\right)$

Combined signal PDF $\quad s_i = s_i^{\mathrm{space}} s_i^{\mathrm{time}}$

Gaussian spatial PDF $\quad s_i^{\mathrm{space}} = \frac{1}{2\pi\sigma_i^2}\exp\left(-\frac{|\vec{r}_i - \vec{r}_s|^2}{2\sigma_i^2}\right)$

Heaviside temporal PDF $\quad s_i^{\mathrm{time}} = \frac{H(t_j^{\max} - t_i)H(t_i - t_j^{\min})}{\Delta t_j}$

time span of a given multiplet

Background PDF $\quad b_i = b_i^{\mathrm{space}} b_i^{\mathrm{time}}$

$b_i^{\mathrm{space}} = 1/\Delta\Omega$

$b_i^{\mathrm{time}} = 1/\Delta t_{\mathrm{data}}$

N - number of all events
σ_i - angular reconstruction uncertainty of event i
$\vec{r}_i$, $\vec{r}_s$ - direction to event i and source
$\Delta\Omega$ - solid angle

We obtain estimates of:

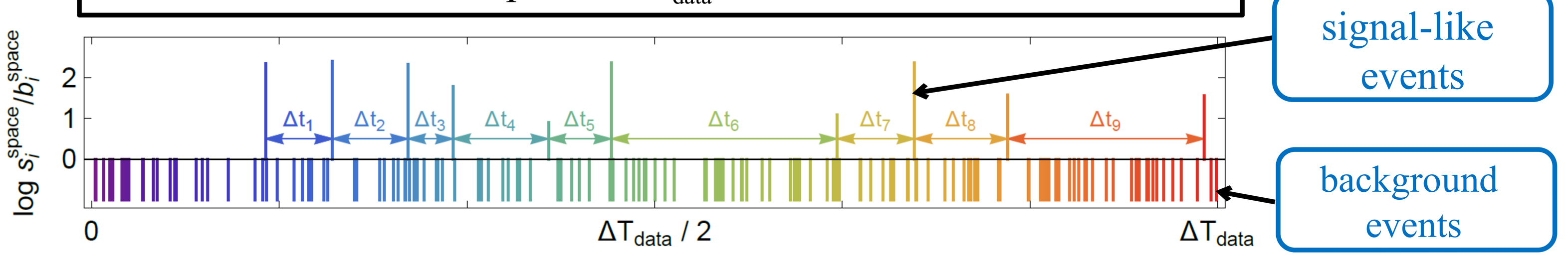
- number of signal events ($\mathbf{n_s}$)
- the flare duration (**ΔT**), i.e. time span $\mathbf{\Delta T_j}$ of the most significant multiplet (multiplet with the highest $\mathrm{TS_j}$)

Improved method of space-time clustering analysis

Stacking method (D. Góra et al. Astropart. Phys. 35 (2011) 201) uses only doublets, consists of 3 steps.

1: Select flare candidates from the data using solely space information:
identify signal-like events based on the ratio of the signal PDF to background PDF > threshold S/B, i.e. $\mathbf{s_i^{space}/b_i^{space} > S/B}$

Data collected over time period ΔT_{data}. Time of events is color coded.



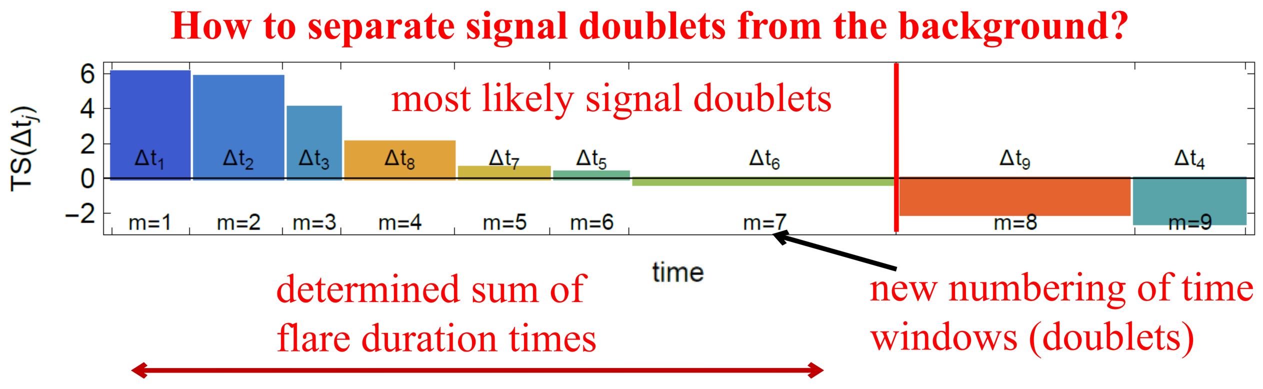
Extract all consecutive doublets to isolate all possible time windows $\mathbf{\Delta t_j}$ that compose the flares contribution.

2: For each doublet j maximize test statistic $\mathbf{TS_{\Delta tj}(n)}$ **(calculate doublet significance)** using standard method with marginalization term to provide more uniform exposure for finding doublets of different widths:

$$\mathrm{TS}_{\Delta t_j}(\mathrm{n}) = -2\log\left[\frac{\Delta T_{\mathrm{data}}}{\Delta t_j}\mathcal{L}(0)/\mathcal{L}(\mathrm{n})\right]$$

Only events within $\mathbf{\Delta t_j}$ interval are taken into account, thus in this step both space and time information is used.
Sort doublets according to the value of $\mathbf{TS_{\Delta tj}}$, i.e. to their significance, and change numbering of doublets introducing multiplicity index **m**.

How to separate signal doublets from the background?



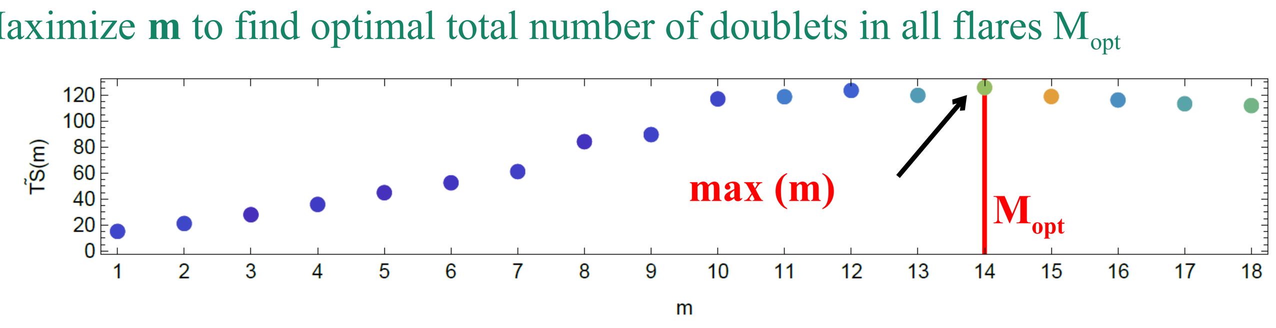
3: Stacking analysis:
one-event signal PDF $\mathbf{s_i}$ is replaced by the weighted sum of signal sub-terms over m doublets:

$$s_i \rightarrow s_i^{\mathrm{tot}}(m) = \Sigma_{j=1}^{m} w_j s_i^j / \Sigma_{j=1}^{m} w_j \qquad \text{with weights } \mathbf{w_j = TS(\Delta t_j)}$$

use standard likelihood and test statistic with stacking term $\mathbf{s_i^{tot}(m)}$:

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

Maximize **m** to find optimal total number of doublets in all flares M_{opt}



M_{opt} determines **total flares duration.** It is **estimated as a sum of most significant (not necessarily consecutive)** doublets: $\Delta T = \Sigma_{m=1}^{M_{opt}} \Delta t_m$

Application of the S_b photon tag to enhance sensitivity for photons search

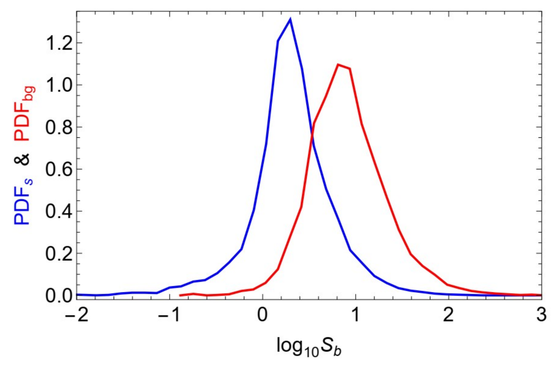
S_b variable is used to discriminate between photons and background showers
(G. Ros et al.. Astropart. Phys. 35 (2011) 140)

$$S_b = \Sigma_{i=1}^{n} S_i\left(\frac{R_i}{1000\ \mathrm{m}}\right)^4$$

detector signal

distance from the shower axis

PDF of $\log_{10} S_b$ variable

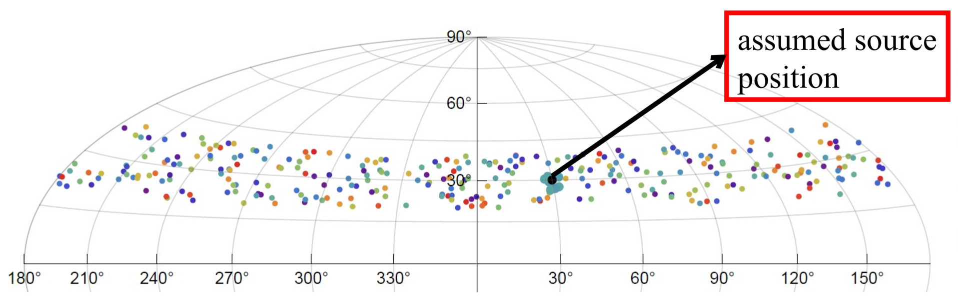


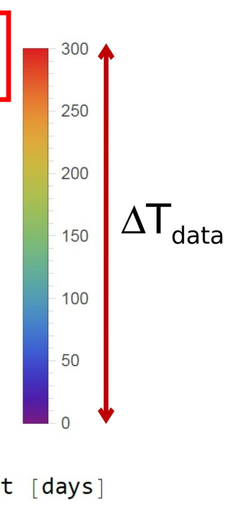
Photon tag: Probability Distribution Functions of $\log_{10} S_b$ for **photons** ($\mathrm{PDF_s}$) and **protons** (background) ($\mathrm{PDF_{bg}}$), make the replacement:

$s_i^{\mathrm{space}} \Rightarrow s_i^{\mathrm{space}} * \mathrm{PDF_s}(S_b)$
$b_i^{\mathrm{space}} \Rightarrow b_i^{\mathrm{space}} * \mathrm{PDF_{bg}}(S_b)$

Monte Carlo test

Randomly generate many sample maps with background and signal events (distributed around an assumed source position and within assumed flares duration $\Delta T_{\mathrm{flare}}$)

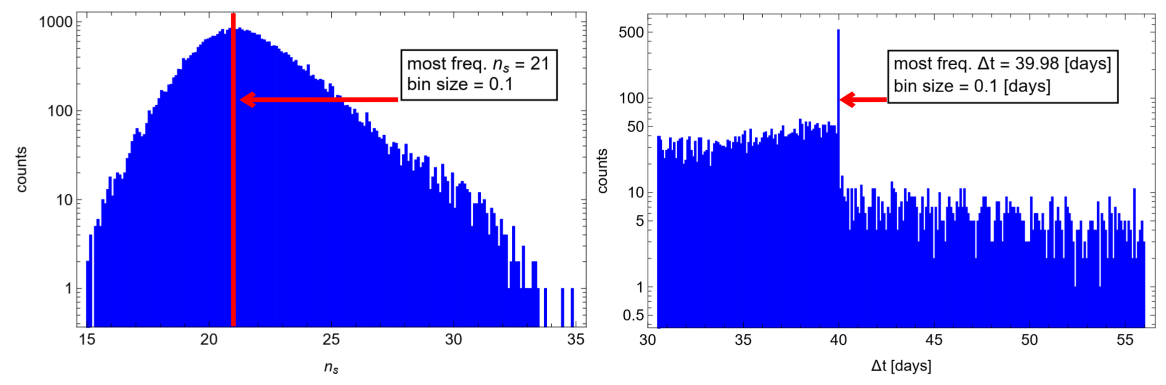




For each map we obtain:

- estimator $\mathbf{n_s}$ of number of signal events $\mathbf{N_s}$
- estimator **ΔT** of the flares duration $\mathbf{\Delta T_{flare}}$
- value of the test statistic for $\mathbf{M_{opt}}$

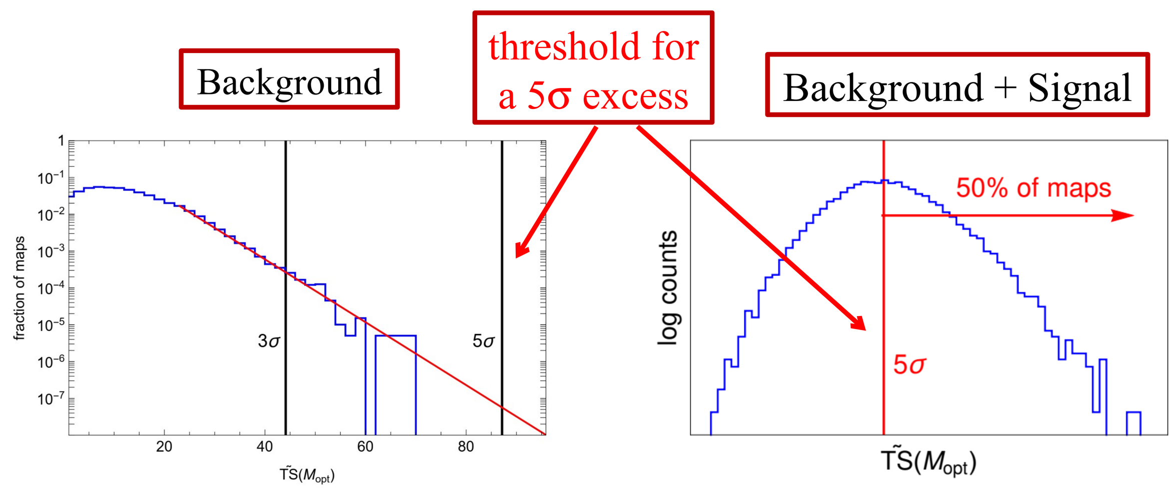
3 flares (duration 20, 10 and 10 days), number of signal events $N_s = 20$:



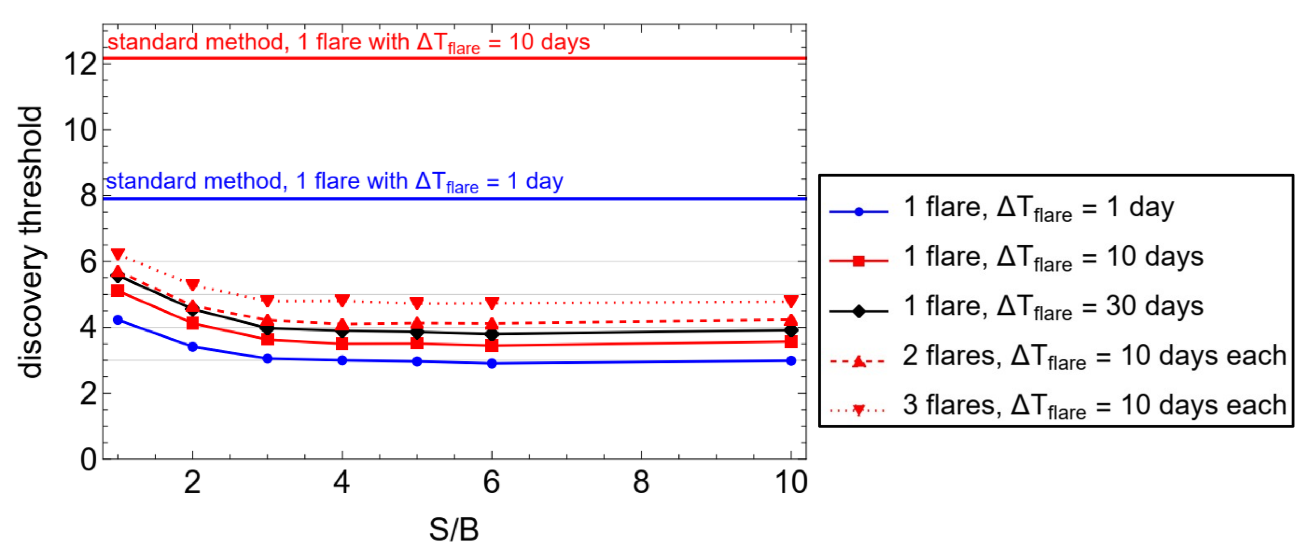
Number of signal events $\mathbf{N_s = 20}$ and total flares duration $\mathbf{\Delta T_{flare} = 40}$ **days** are recovered.

Discovery potential of the stacking clustering method with S_b photon tag

Discovery potential tells us how many signal events are needed to claim discovery of a cluster of events in data. It can be used to compare different methods.
Definition: The discovery threshold is the number of signal events required to achieve a p-value less than 2.87×10^{-7} (one-sided 5σ) in 50% of the maps (i.e. the number of signal events for which median of M_{opt} test statistic distribution is at 5σ threshold).



Discovery thresholds vs threshold for signal-like events S/B



Only a few events are needed to detect flares at a higher S/B threshold for signal-like events.

Summary

Identification of clustering in cosmic-ray data would provide evidence for possible existence of UHE neutral particles and could potentially help identify their sources.

Advantages of the stacking method:
- it is faster than the standard method
- more sensitive to weak flares of any shapes
- able to detect multiple flares.

The stacking method is able to recover the number of signal events and flares duration with small uncertainty.

The stacking method with S_b photon tag requires only 3 to 5 events to discover photon flares!

This work was partly funded by National Science Centre, Grants No. 2020/39/B/ST9/01398 and 2022/45/B/ST9/02163, and by Ministry of Science and Higher Education, Grant No. 2022/WK/12