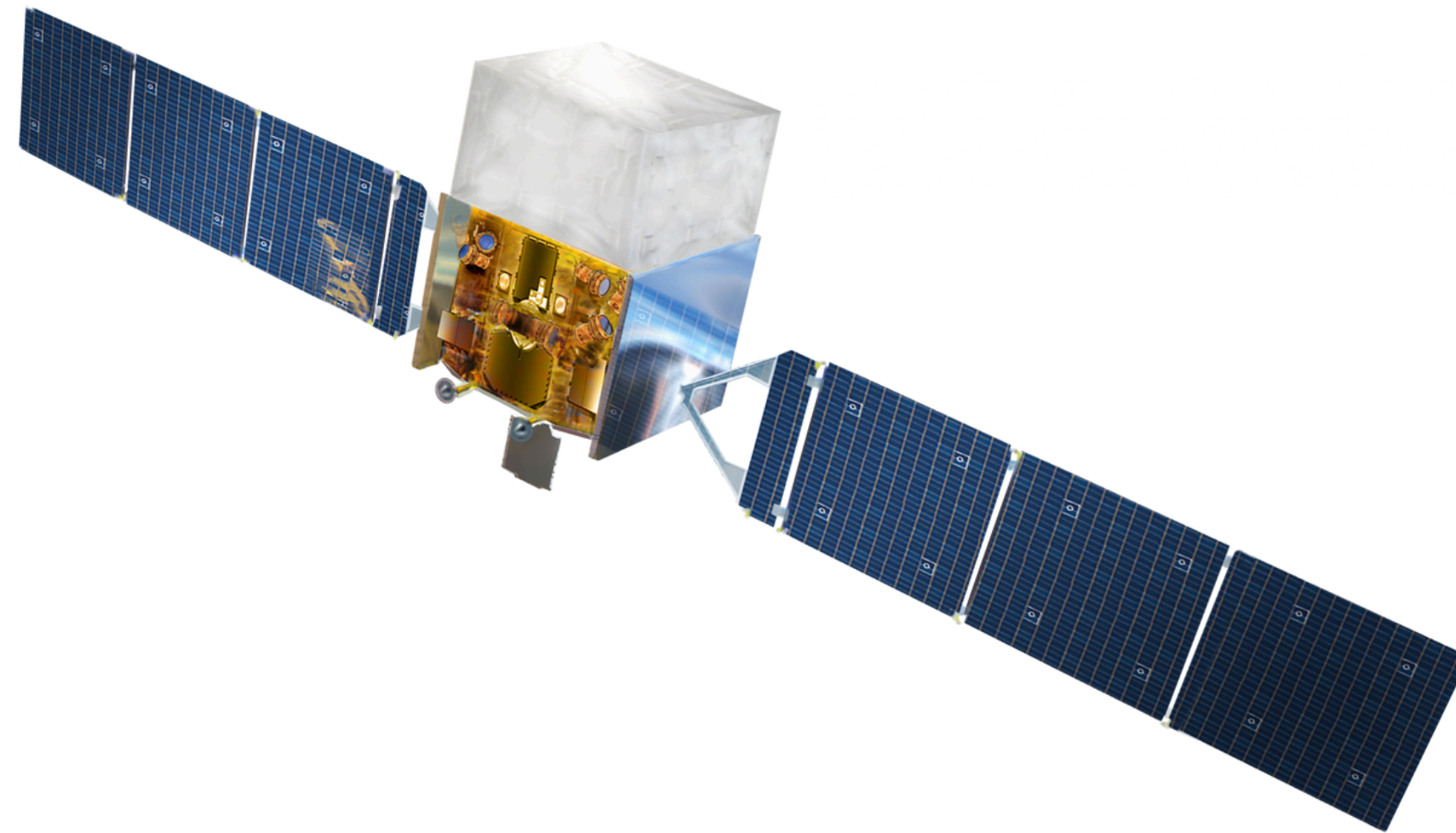


Identifying potentially mis-modelled extended sources in 4FGL using clustering analysis

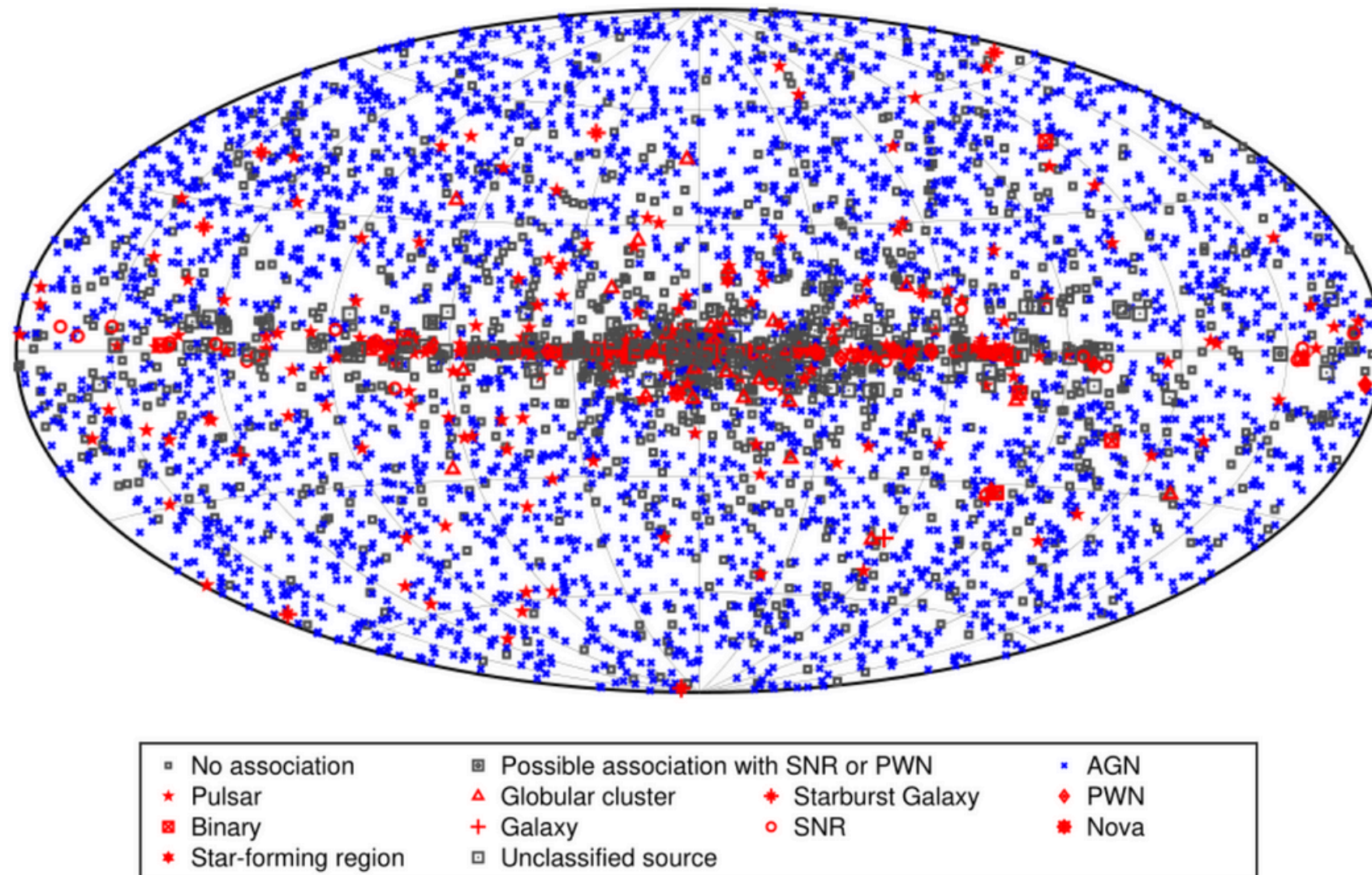
Giovanni Cozzolongo, A. Mitchell, S. Spencer, D. Malyshev, T. Unbehauen

Astroparticle School 2025



- **Large Area Telescope (LAT).**
- **Pair conversion** telescope.
- Detects photons in an energy range **from 20 MeV to over 300 GeV.**
- Covers the **entire sky in 3 hours.**
- Measures of **time, energy** and **direction** of incident photons.
- LAT **14-year** Source Catalog (4FGL-DR4).

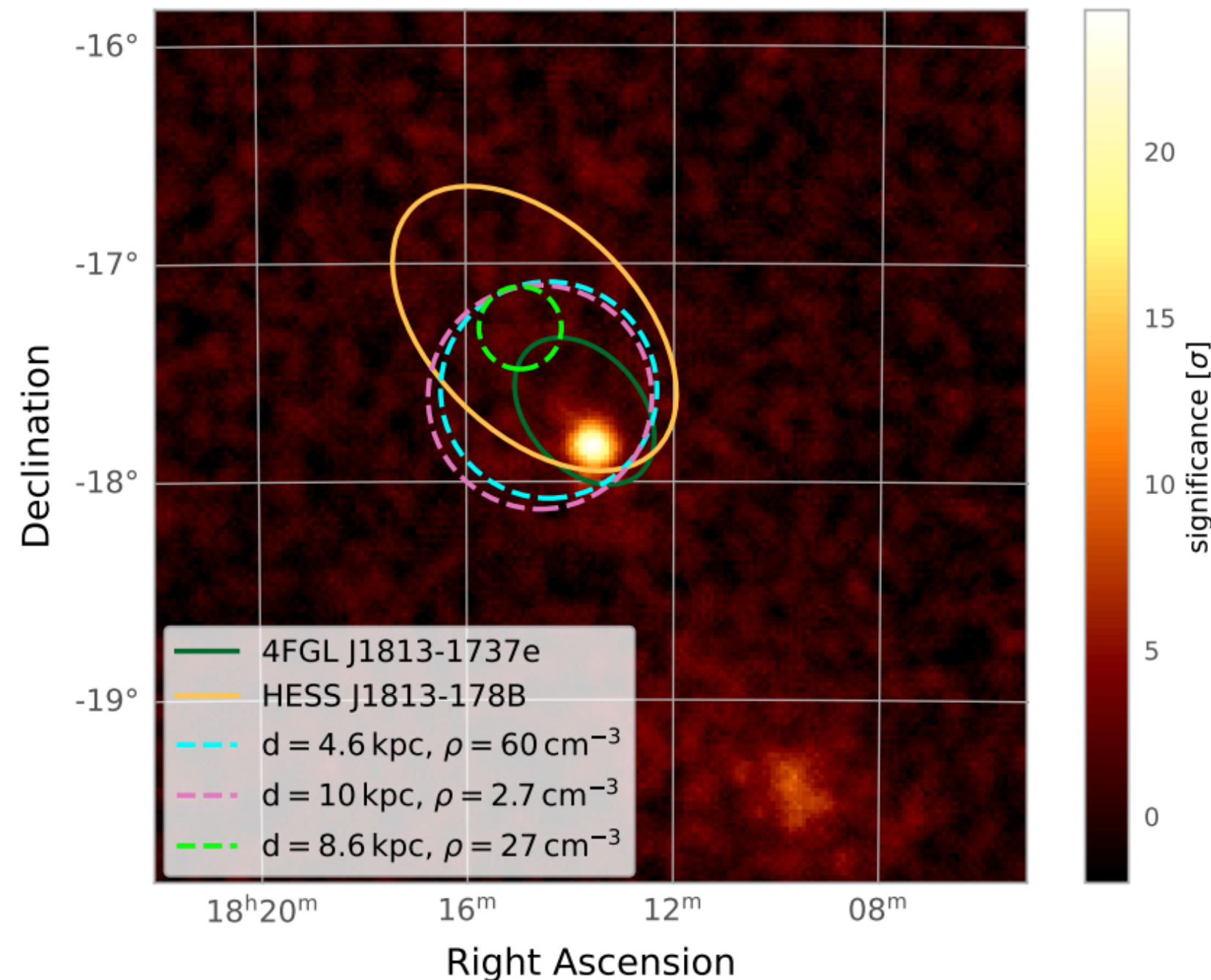
<https://arxiv.org/abs/2307.12546>



Credits: S. Abdollahi et al. (2020) ApJS, 247, 33.

- 7194 **4FGL-DR4 sources**.
- 2065 **unidentified sources**.
- 81 **extended sources**.
- Extended sources may be **misidentified as point source clusters**.
- May **match HESS extended sources**.

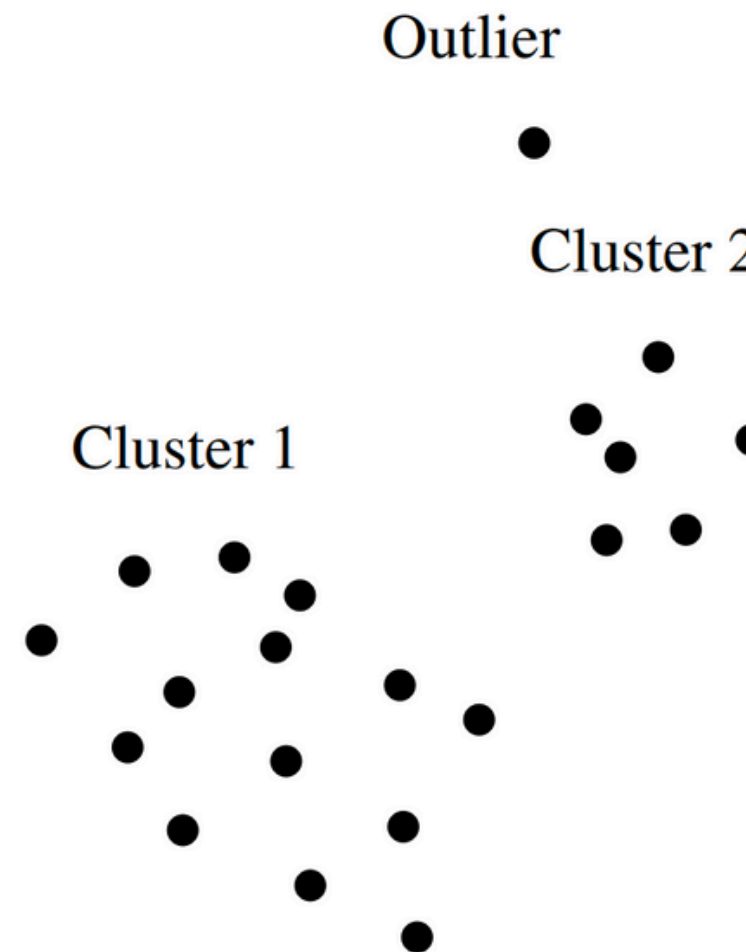
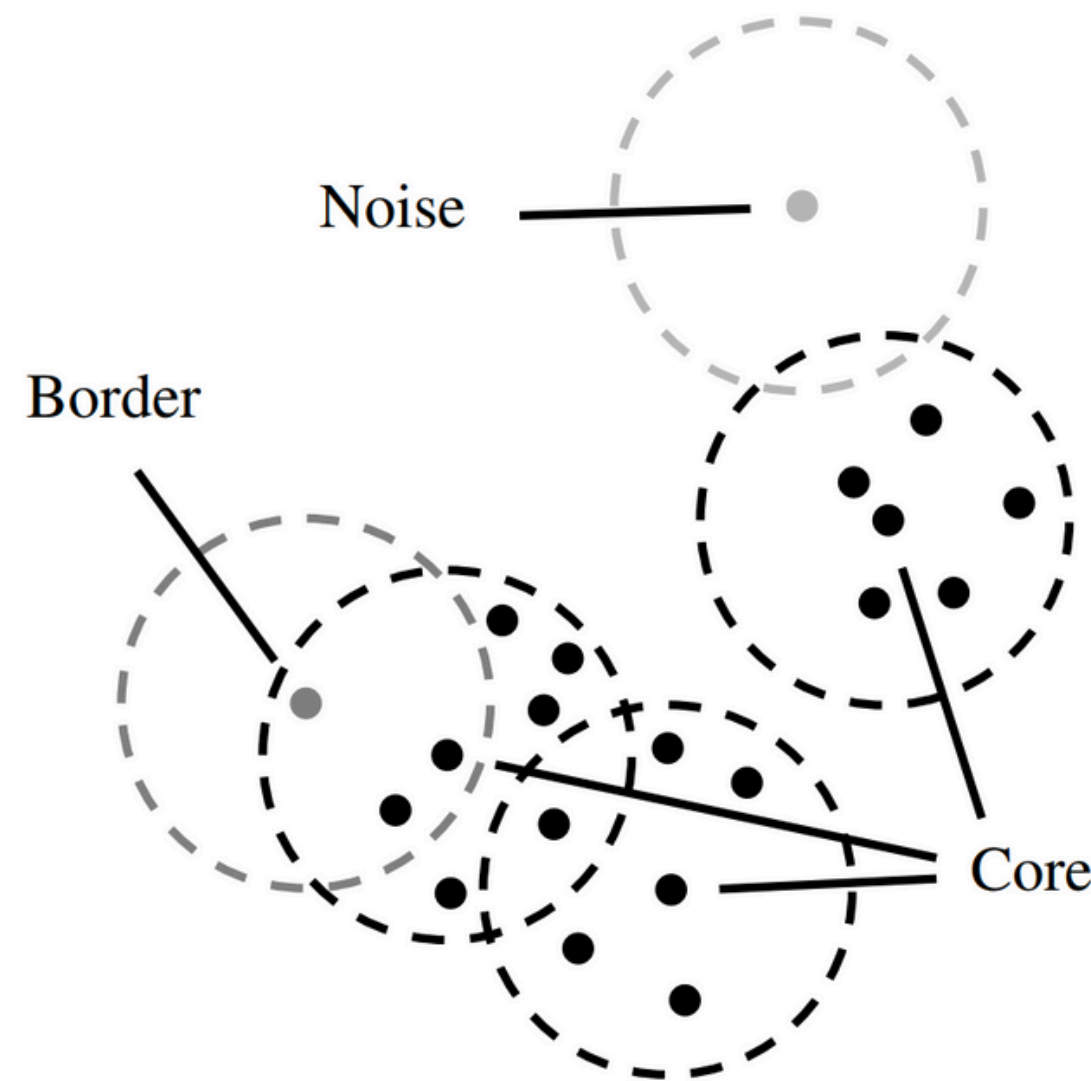
The case of 4FGL J1813.1-1737e



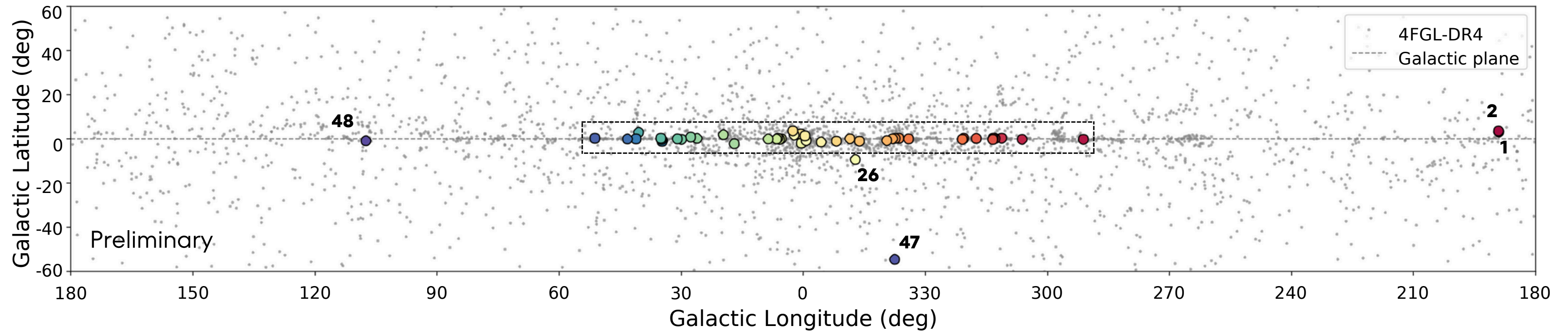
Credits: F. Aharonian et al. (2024) A&A, 686, A149

- **Two unidentified 3FGL point sources** in the region of HESS J1813-178, associated with a γ -ray PWN (Acero et al., 2015).
- **Extended morphology fits data better** than two point sources (Araya, 2018).
- **Comparable spectral indices** measured at GeV and TeV (Araya, 2018).
- **Fermi-LAT and H.E.S.S.** data can be described by a single source model (F. Aharonian et al. 2024, A&A, 686, A149).

<https://arxiv.org/abs/2403.16802>



- Apply **unsupervised machine learning** to the 4FGL catalog.
- Used **DBSCAN** (Density-Based Spatial Clustering of Applications with Noise).
- DBSCAN **searches for clusters of points** and classifies the points into core, border or noise points (Ester et al., 1996).

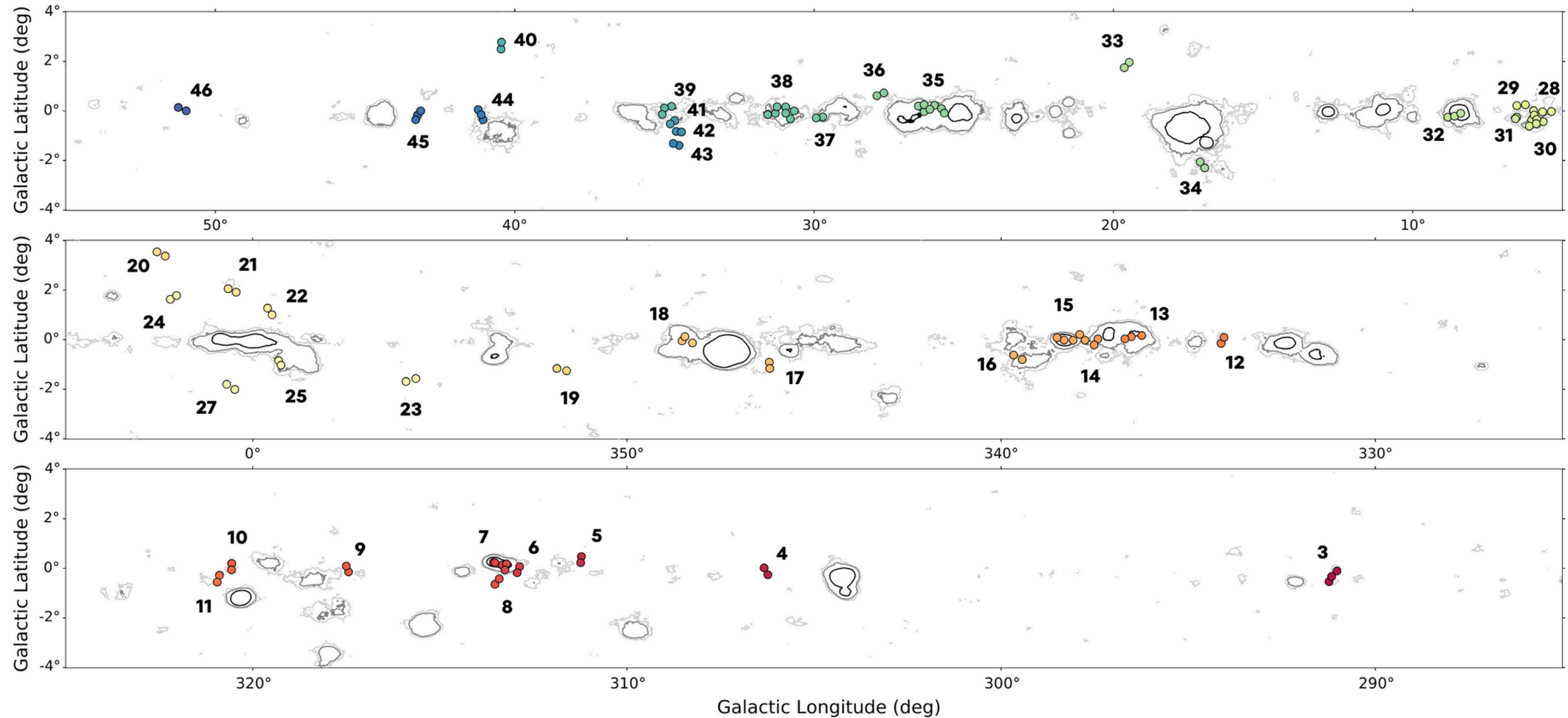


- Set **eps = 0.005 rad** ≈ 0.3 deg and **minPts = 2**.
- Included only **unassociated sources** and sources associated with **pulsars**, **supernova remnants**, and **active galaxies** (7030 sources in total).
- Found **48 clusters (124 sources)**, each including at least one unidentified source.
- There are **mostly clusters of size 2**, with some up to 7.

HGPS (HESS Galactic Plane Survey)



ERLANGEN CENTRE
FOR ASTROPARTICLE
PHYSICS



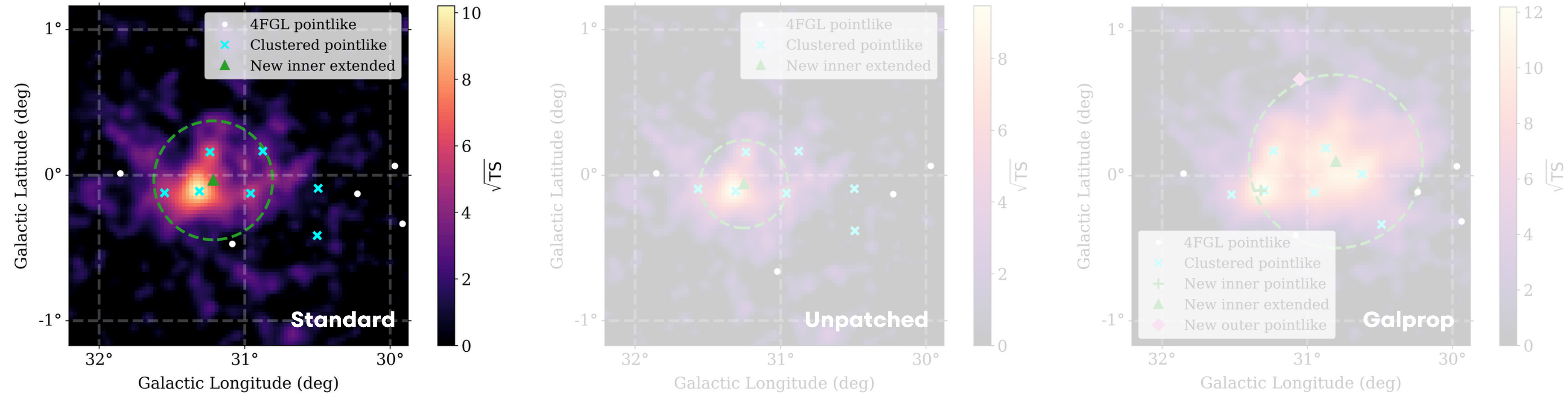
HGPS contours are from H. Abdalla et al. (2018) A&A, 612, A1

- Likelihood nomenclatures:
 - **No-source** hypothesis: L_0
 - **Single point** hypothesis: L_{ps}
 - **Many points** hypothesis: L_{Npts}
 - Extended hypothesis: L_{ext}
- Test Statistic definitions:
 - **Extended source TS:** $TS = 2 \log(L_{ext}/L_0)$
 - **Source extension TS:** $TS_{ext} = 2 \log(L_{ext}/L_{ps})$
 - **N point sources TS:** $TS_{Npts} = 2 \log(L_{Npts}/L_{ps})$
- Followed these criteria (Ackermann et al. 2017, ApJ, 843, 139):
 - **Claim a detection:** $TS > 25$
 - **Define a source as extended:** $TS_{ext} > 16$
 - **Find the preferred model** (AIC test): $TS_{ext} > TS_{Npts} - 2\Delta k$

- Likelihood nomenclatures:
 - **No-source** hypothesis: L_0
 - **Single point** hypothesis: L_{ps}
 - **Many points** hypothesis: L_{Npts}
 - Extended hypothesis: L_{ext}
- Test Statistic definitions:
 - **Extended source TS:** $TS = 2 \log(L_{ext}/L_0)$
 - **Source extension TS:** $TS_{ext} = 2 \log(L_{ext}/L_{ps})$
 - **N point sources TS:** $TS_{Npts} = 2 \log(L_{Npts}/L_{ps})$
- Followed these criteria (Ackermann et al. 2017, ApJ, 843, 139):
 - **Claim a detection:** $TS > 25$
 - **Define a source as extended:** $TS_{ext} > 16$
 - **Find the preferred model (AIC test):** $TS_{ext} > TS_{Npts} - 2\Delta k$

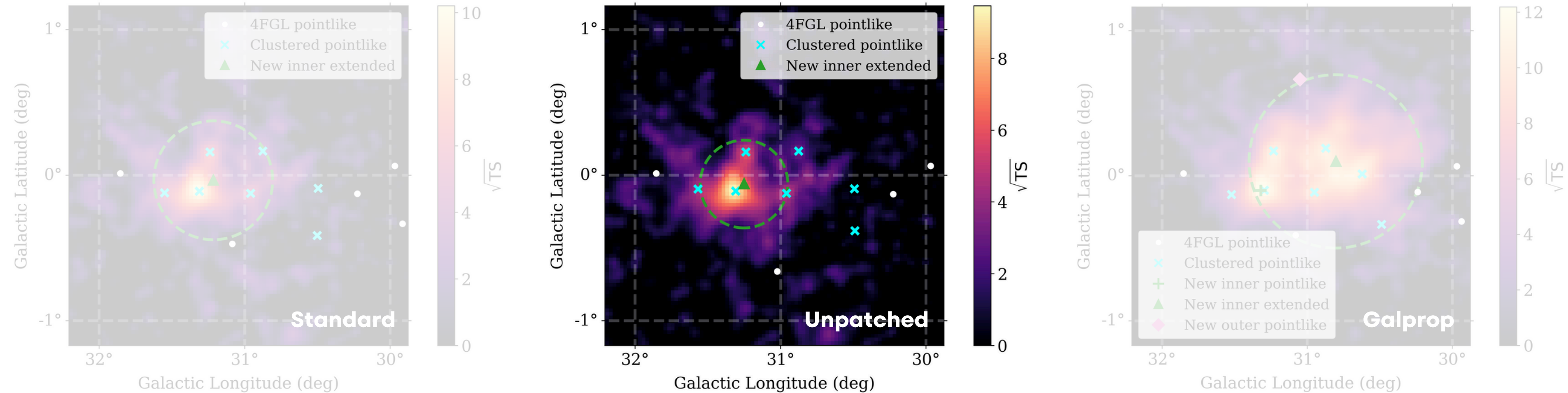
- Likelihood nomenclatures:
 - **No-source** hypothesis: L_0
 - **Single point** hypothesis: L_{ps}
 - **Many points** hypothesis: L_{Npts}
 - Extended hypothesis: L_{ext}
- Test Statistic definitions:
 - **Extended source TS:** $TS = 2 \log(L_{ext}/L_0)$
 - **Source extension TS:** $TS_{ext} = 2 \log(L_{ext}/L_{ps})$
 - **N point sources TS:** $TS_{Npts} = 2 \log(L_{Npts}/L_{ps})$
- Followed these criteria (Ackermann et al. 2017, ApJ, 843, 139):
 - **Claim a detection:** $TS > 25$
 - **Define a source as extended:** $TS_{ext} > 16$
 - **Find the preferred model** (AIC test): $TS_{ext} > TS_{Npts} - 2\Delta k$

Cluster 38 residuals maps



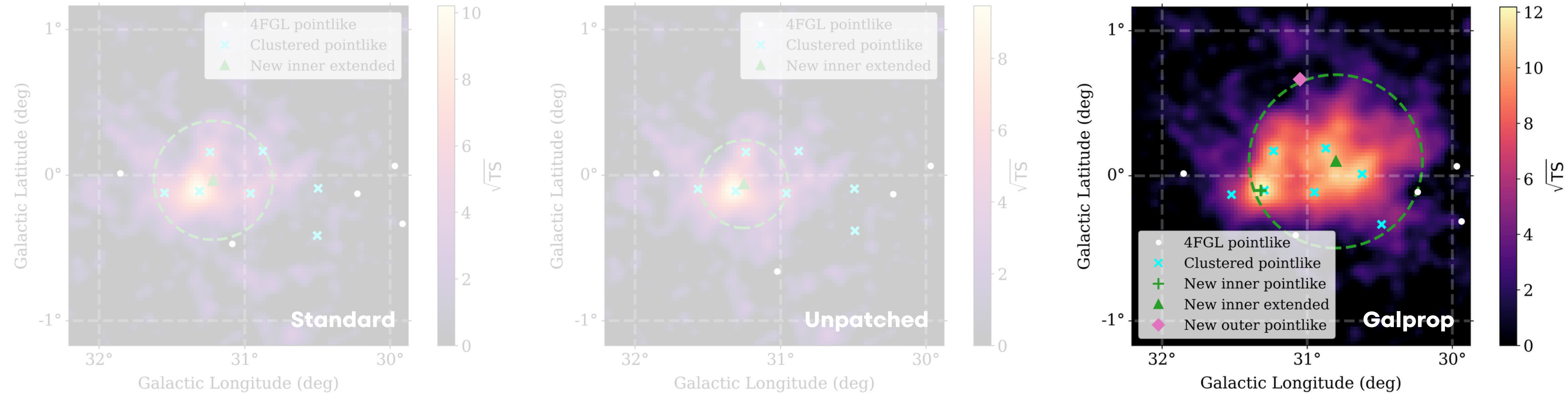
- The **patch** accounts for residual emission that is not well represented by the standard interstellar emission templates (e.g. Fermi Bubbles).
- **GALPROP** is a numerical code for calculating the propagation of relativistic charged particles and the diffuse emissions produced during their propagation.

Cluster 38 residuals maps



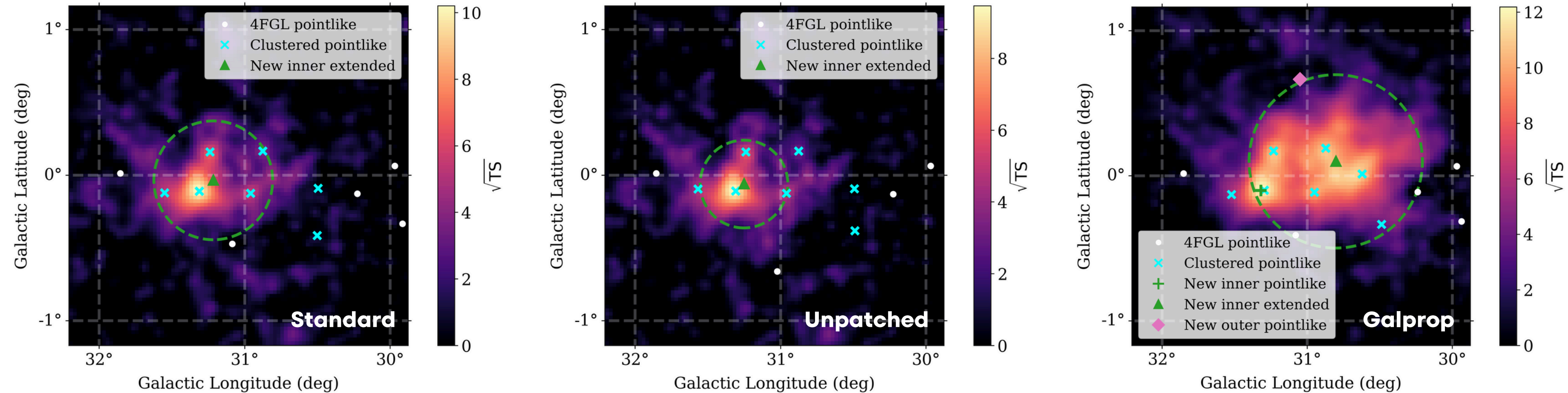
- The **patch** accounts for residual emission that is not well represented by the standard interstellar emission templates (e.g. Fermi Bubbles).
- **GALPROP** is a numerical code for calculating the propagation of relativistic charged particles and the diffuse emissions produced during their propagation.

Cluster 38 residuals maps



- The **patch** accounts for residual emission that is not well represented by the standard interstellar emission templates (e.g. Fermi Bubbles).
- **GALPROP** is a numerical code for calculating the propagation of relativistic charged particles and the diffuse emissions produced during their propagation.

Cluster 38 residuals maps



- The **patch** accounts for residual emission that is not well represented by the standard interstellar emission templates (e.g. Fermi Bubbles).
- **GALPROP** is a numerical code for calculating the propagation of relativistic charged particles and the diffuse emissions produced during their propagation.

Results of the promising clusters

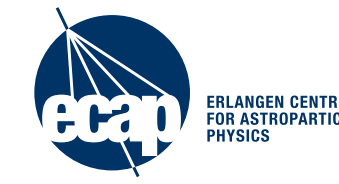
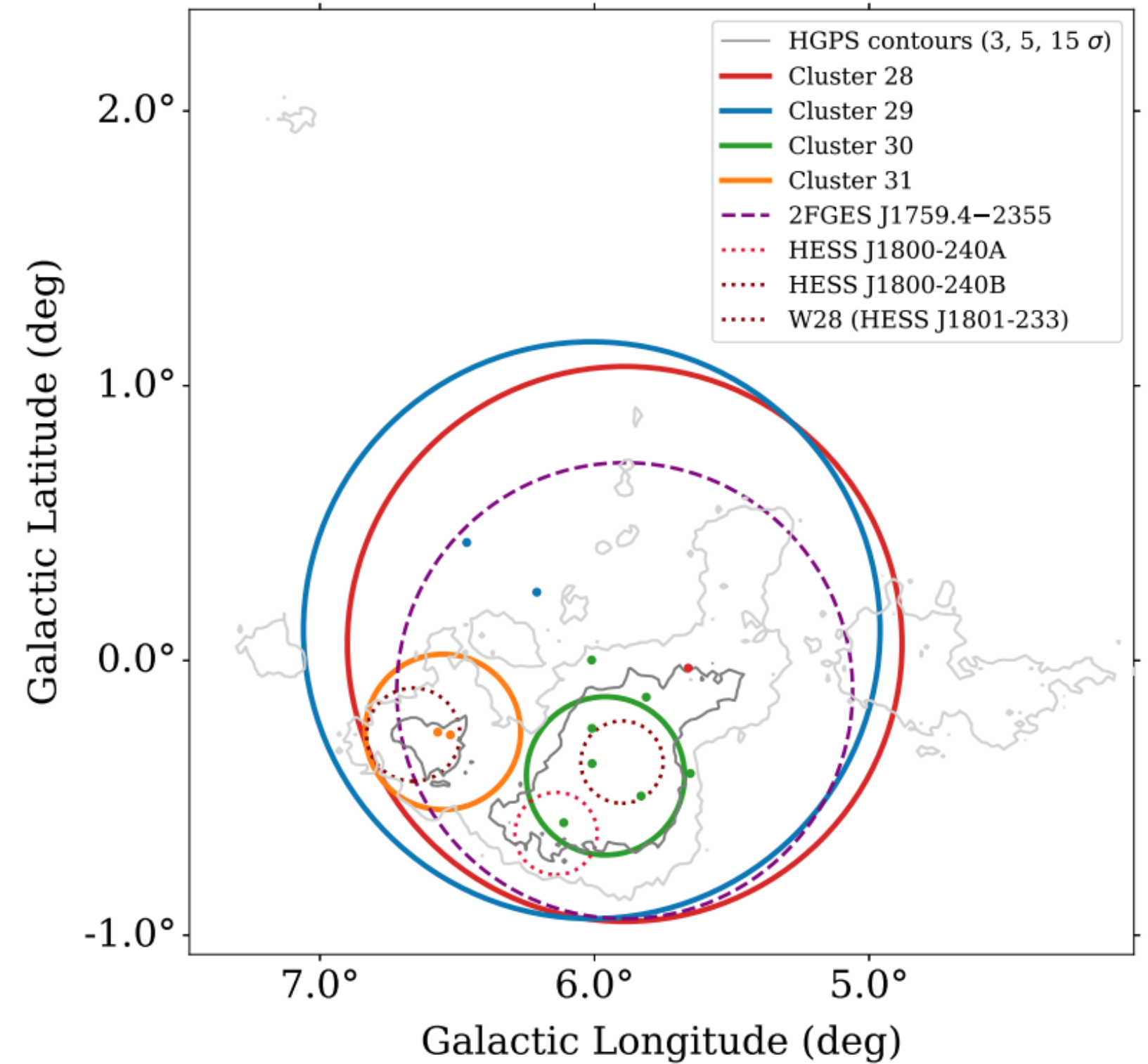
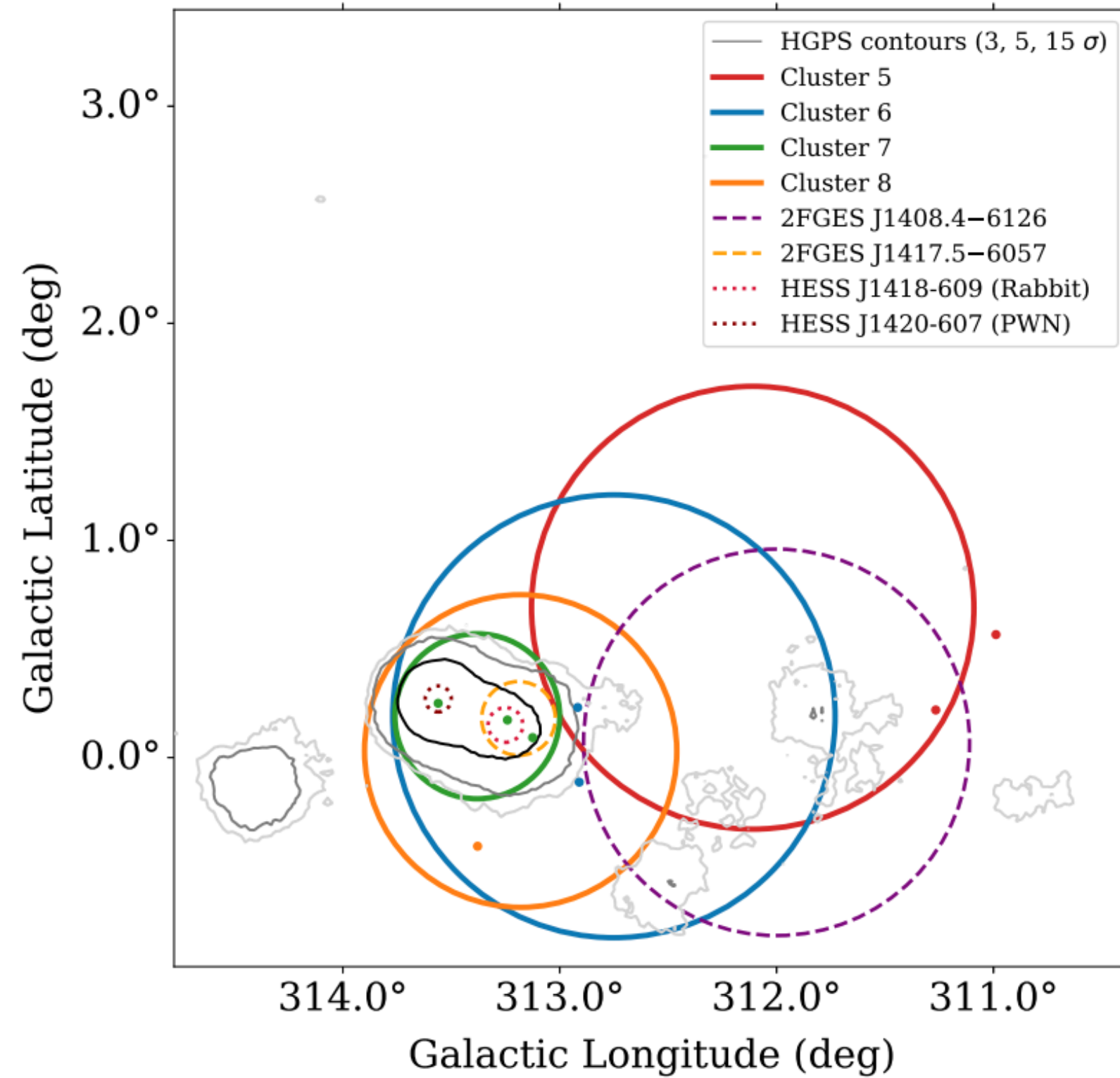


Table 3: Properties of significant clusters with test statistic values. In particular, ΔAIC is equal to the value of AIC for each model minus the AIC value for the best-fit mode. Flag definitions: (1) mass center of original clustered source positions is outside the 68% containment radius of extended source found using standard diffuse model; (2) mass center of original clustered source positions is outside the 68% containment radius of extended source found using unpatched model; (3) mass center of original clustered source positions is outside the 68% containment radius of extended source found using GALPROP model; (4) additional point sources identified beyond the extended source.

Cluster ID	TS _{original}	TS _{extended}	TS _{extension}	Params Orig	Params Cluster	ΔAIC	Flags
2	78	128	49	16	12	58	
4	8	61	65	17	13	61	1, 3
5	20	102	115	21	17	90	
6	22	181	173	25	19	171	2
7	355	420	150	28	21	78	4
8	18	113	102	18	17	96	2, 3
11	33	58	48	15	11	32	2, 3
12	43	70	27	19	14	37	3
14	65	176	68	20	18	115	4
16	20	81	63	13	13	61	
17	33	143	140	15	11	118	
23	17	98	92	13	12	82	2, 3
28	92	242	154	20	24	141	4
29	56	184	137	25	20	138	
30	797	802	452	43	17	56	1, 2, 3
31	0	1511	546	25	23	1515	4
34	52	181	167	15	10	139	
38	69	266	62	40	10	256	2
39	58	191	56	27	17	153	1, 2
41	0	2271	760	25	23	2275	4
42	0	280	81	20	15	290	
43	99	112	45	16	12	20	
45	2823	2816	19	20	16	0	1, 2, 3, 4
46	193	197	79	14	10	11	1, 2
48	58	89	51	14	10	38	3

2FGES and HGPS matches



- **Spatial clustering** with DBSCAN **to find misidentified extended sources.**
 - Comparing **point sources vs single extended source models** with Fermipy.
 - **Vary the radius** from 0.3 to 0.5 degrees.
 - **Matching** the clusters with sources from **2FGES** and **HGPS.**
 - Explore on **TeV/MWL context.**
 - Presented at **γ-2024.**
 - Presented at **HESS Collaboration Meeting.**
 - **Paper draft** almost ready.
-
- **Joint Fermi-LAT and H.E.S.S. fits** in Gammapy.
 - **Incorporate eROSITA** data.

*Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen Centre for Astroparticle Physics
Nikolaus-Fiebiger-Str. 2, Lehrstuhl für Physik 91058 Erlangen, Germany
giovanni.cozzolongo@fau.de*

- **Binning**

- Spatial bin size: 0.025 deg
- ROI width: 6 deg
- Source ROI width: 10
- Energy bins per Decade: 8

- **Event class**

- P8R3_SOURCE

- **Event types**

- FRONT + BACK
- PSF0, PSF1, PSF2, PSF3

- **Models**

- Catalog: gll_psc_v34.fit
- Galactic Diffuse: gll_iem_v07.fits

- **IRFs**

- P8R3_SOURCE_V3_PSF*_v1

- **Energy**

- min: 5 GeV
- max : 1 TeV

- **Time**

- min: 246823875
- max: 681004805

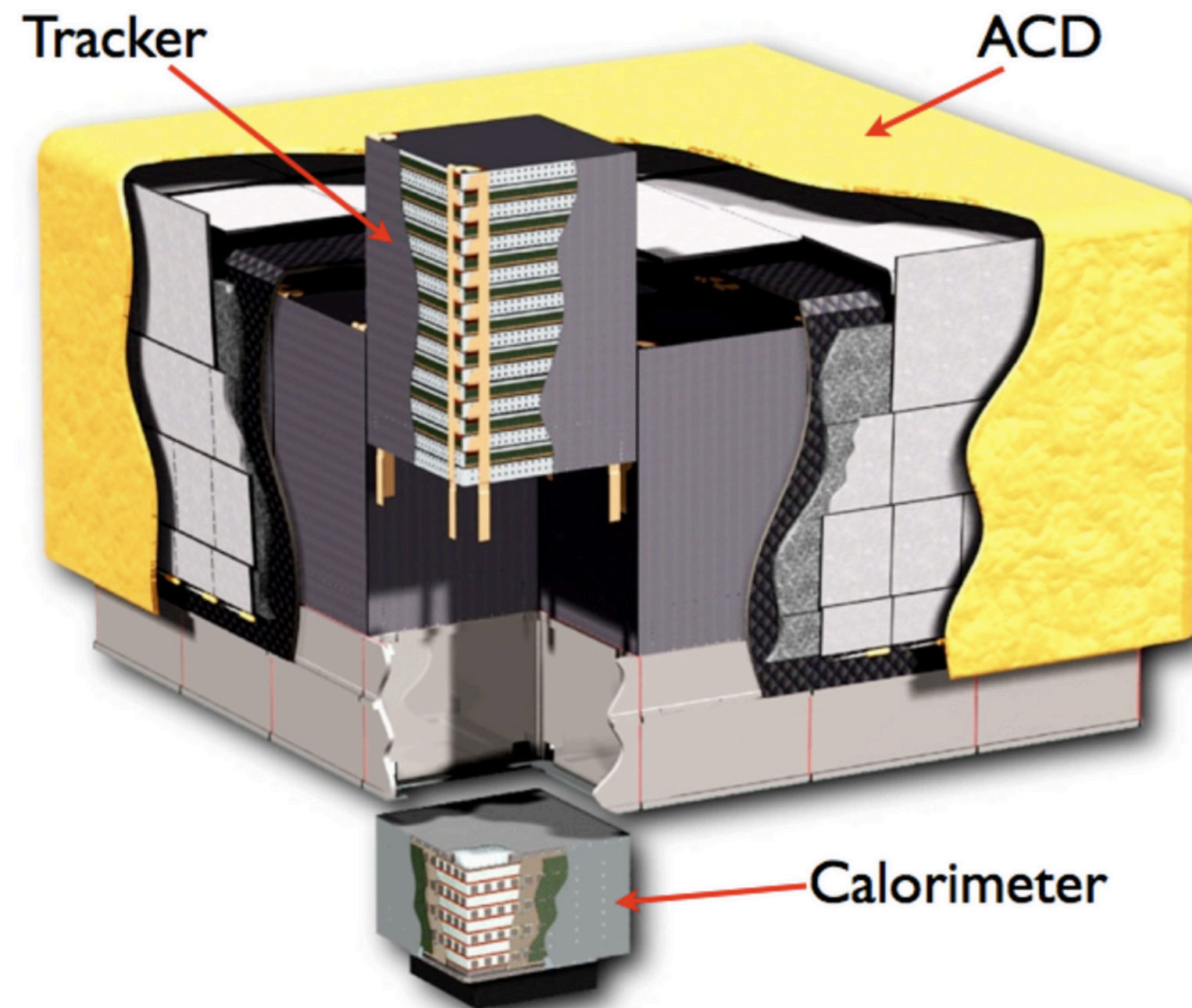
- **Filters**

- DATA_QUAL>0
- LAT_CONFIG==1

- **Zenith angle cut**

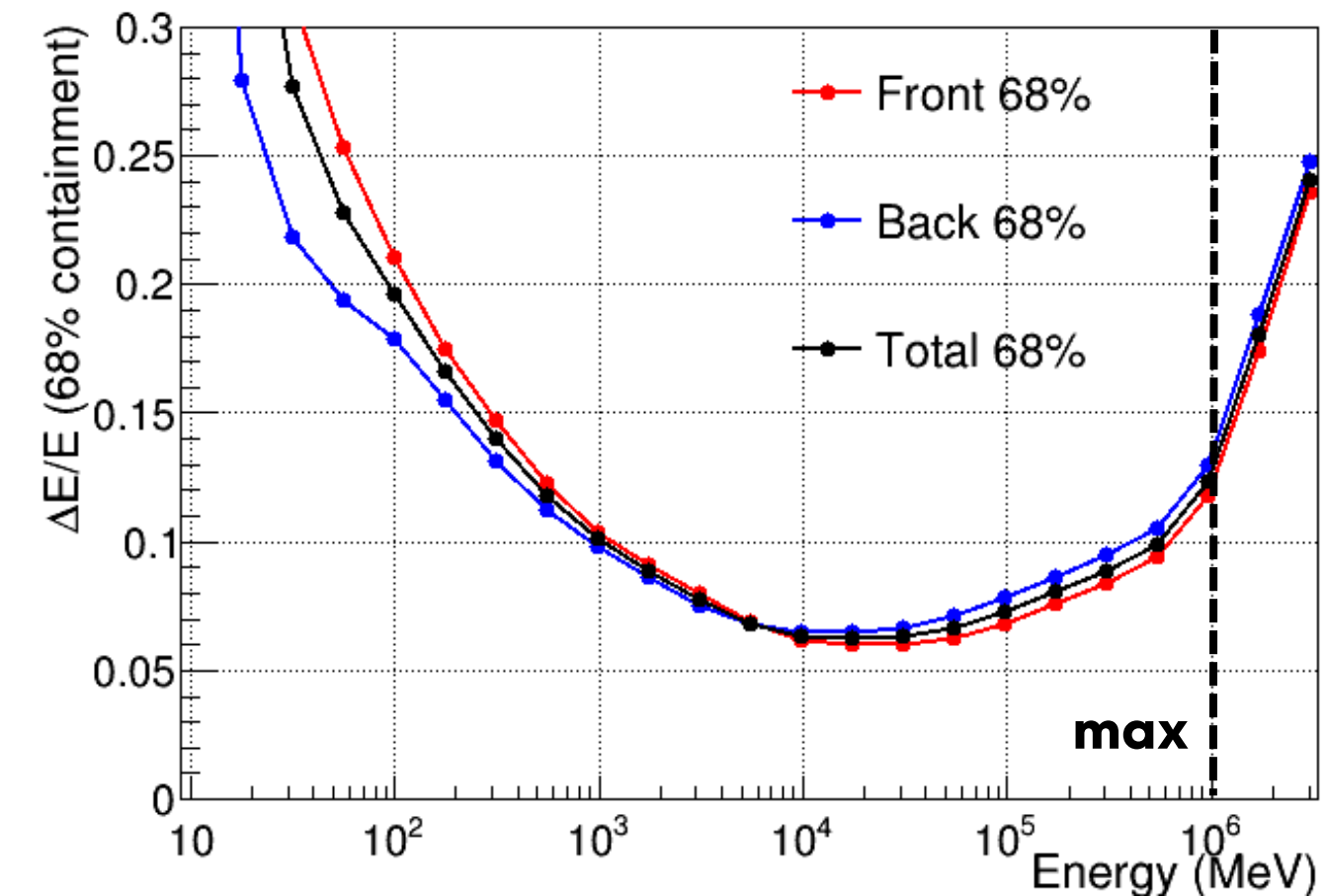
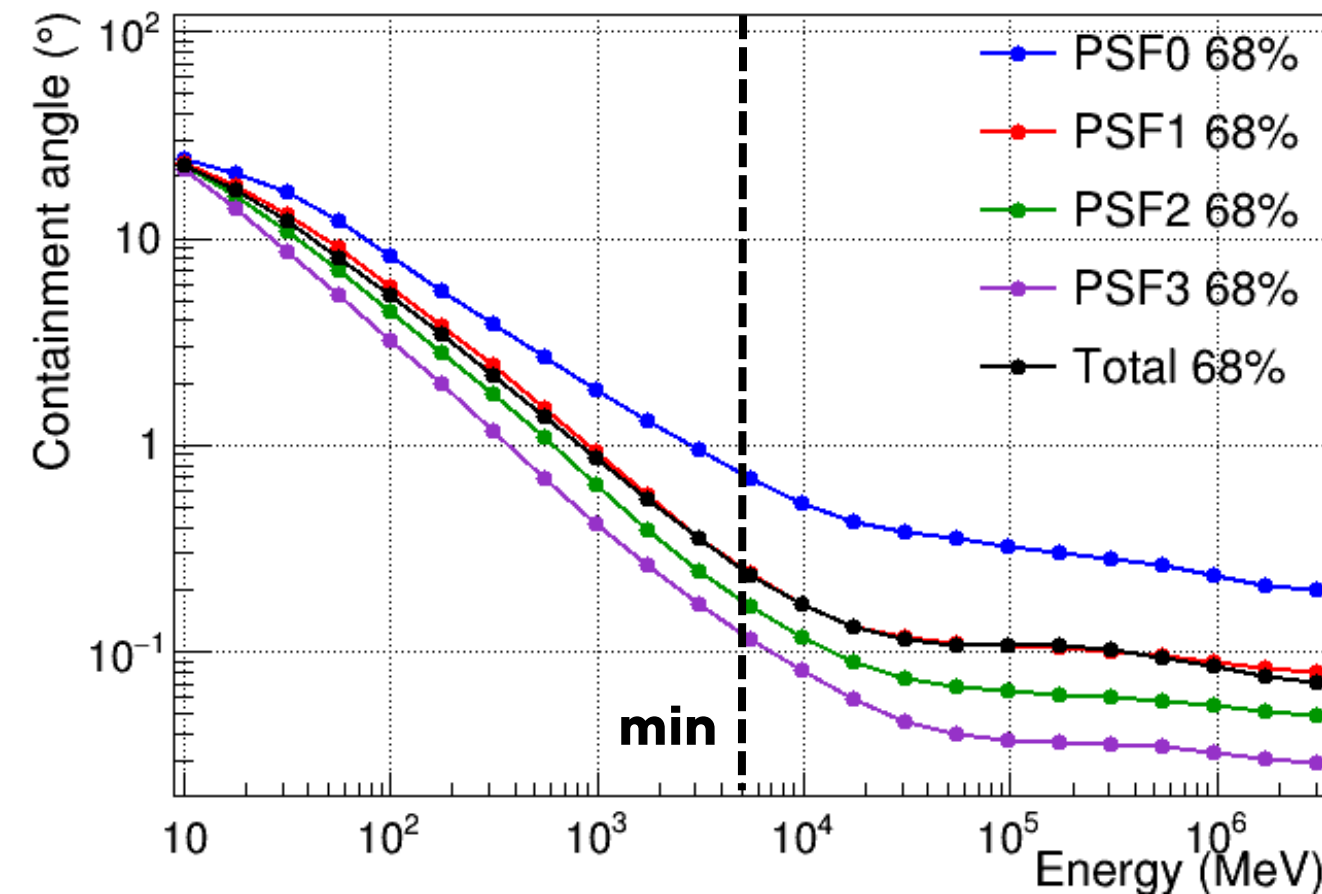
- 90 deg

- GC: galactic center
- **PSR**: young pulsars
- **MSP**: millisecond pulsars
- **PWN**: pulsar wind nebula
- **SNR**: supernova remnant
- **SPP**: supernova remnant / pulsar wind nebula
- GLC: globular cluster
- SFR: star-forming region
- HMB: high-mass binary
- LMB: low-mass binary
- BIN: binary
- NOV: nova
- **BLL**: Lac type of blazar
- **FSRQ**: FSRQ type of blazar
- RDG: radio galaxy
- **AGN**: nonblazar active galaxy
- SSRQ: steep spectrum radio quasar
- CSS: compact steep spectrum radio source
- **BCU**: blazar candidate of uncertain type
- NLSY1: narrow-line Seyfert 1
- SEY: seyfert galaxy
- SBG: starburst galaxy
- GAL: normal galaxy (or part)



W. B. Atwood et al. (2009) ApJ, 697, 1071

- Fermi Gamma-ray Space Telescope.
- Launched by NASA on June 11, 2008.
- **Pair conversion** telescope.
- Detect photons in an energy range **from 20 MeV to over 300 GeV**.
- Field of view 2.4 steradian (20% sky).
- Covers the **entire sky** in 3 hours.
- Measures of **time, energy** and **direction** of incident photons.
- 4 × 4 array of identical towers each one including a **tracker/converter** and a **calorimeter** module;
- **Anticoincidence detector (ACD)**.



Credits: Fermi-LAT Collaboration, 2021

https://www.slac.stanford.edu/exp/glast/groups/canda/lat_Performance.htm

- **PSF event types** are based on the quality of the reconstructed direction.
- Set **minimum energy** to 5 GeV to achieve PSF of less than 0.1 degrees.
- Set **maximum energy** to 1 TeV, achieving Energy Dispersion below 15%.
- Can use **spatial bin size** of 0.025 deg (Ackermann et al. 2018, ApJS, 237, 32).

