Funded by
Deutsche Forschungsgemeinschaft
German Research Foundation

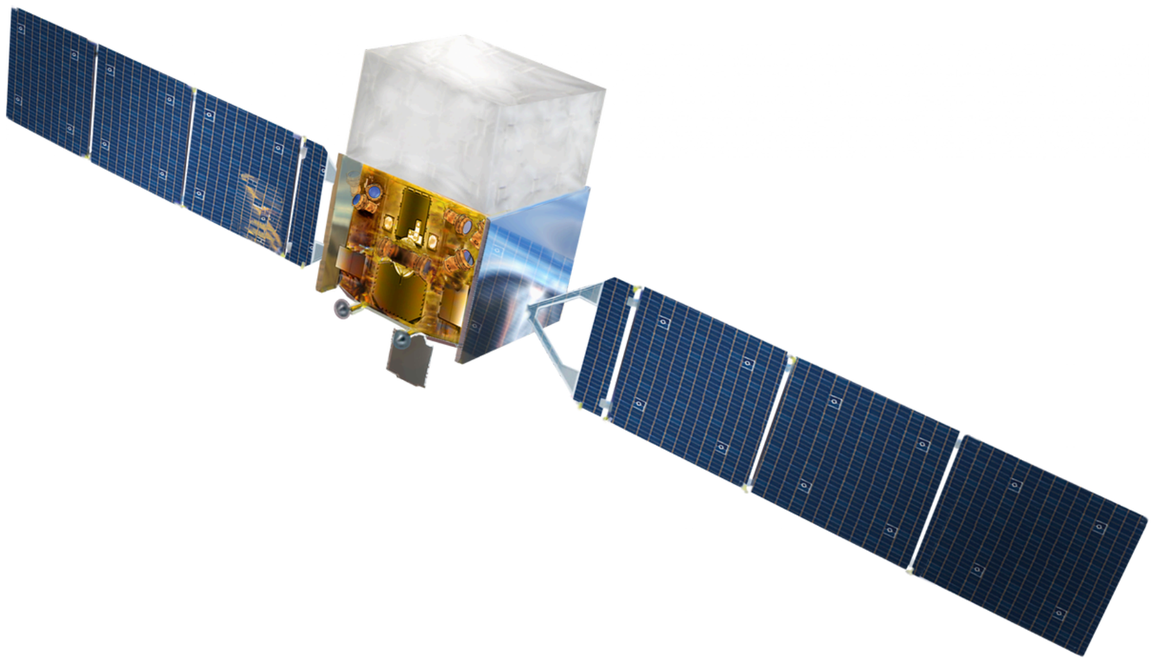
# Clustering Analysis of Fermi-LAT Unidentified Point Sources

*8th Heidelberg International Symposium on High-Energy Gamma-Ray Astronomy*
3rd September 2024

**Giovanni Cozzolongo**, A. Mitchell, S. Spencer, D. Malyshev, T. Wach, T. Unbehaun

ECAP, FAU Erlangen-Nürnberg

# Fermi Gamma-ray Space Telescope

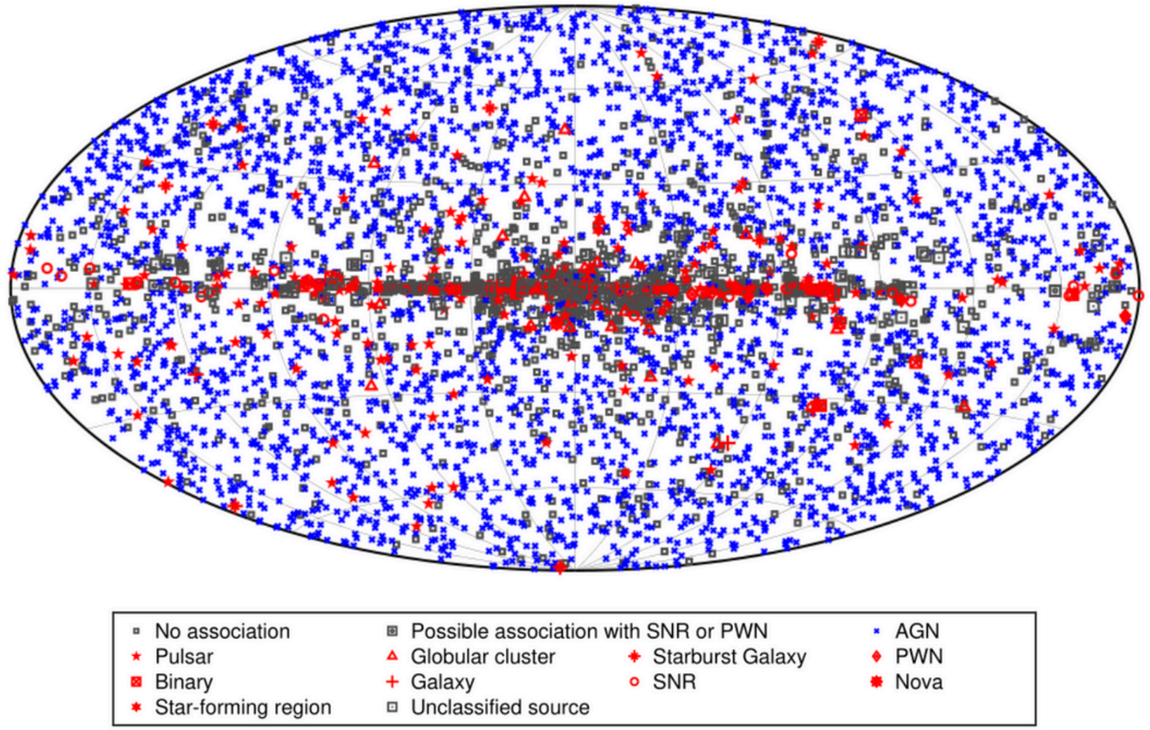


- **Large Area Telescope** (LAT)
- **Pair conversion** telescope.
- Detects photons in an energy range **from 20 MeV to over 300 GeV**.
- Covers the **entire sky in 3 hours**.
- Measures of **time**, **energy** and **direction** of incident photons.

https://arxiv.org/abs/2307.12546

- LAT **14-year** Source Catalog (4FGL-DR4).

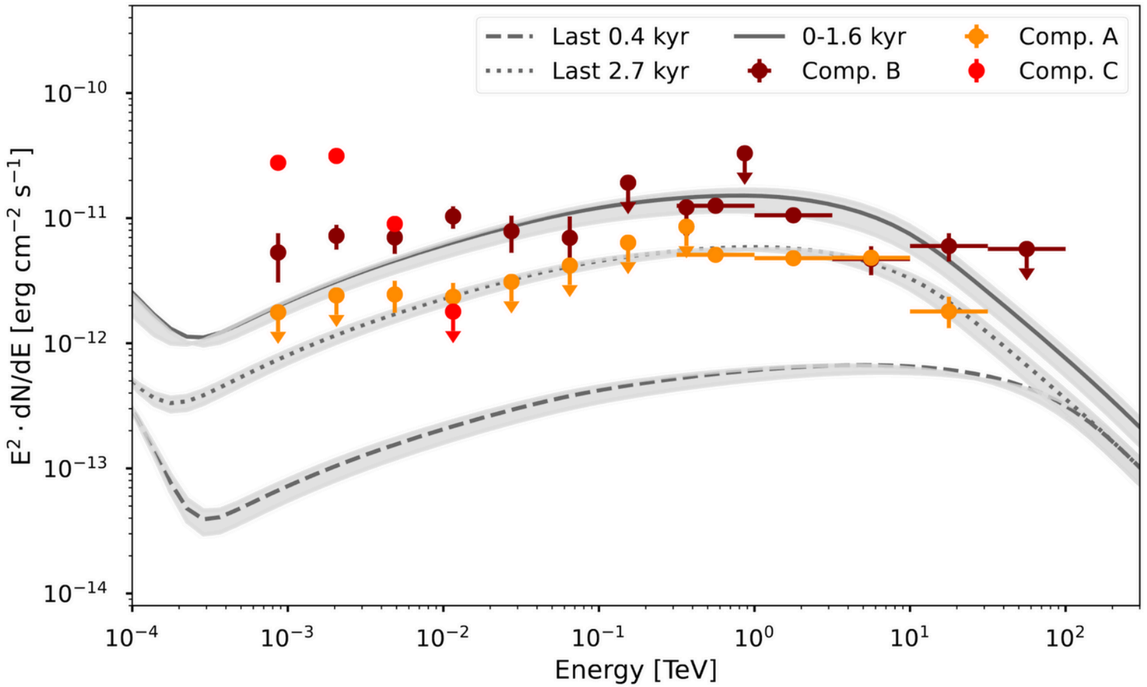
# Fermi Point Source Catalog




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

- **4FGL-DR4** catalog contains 7195 sources.
- There are 2065 **unidentified sources**.
- There are 81 **extended sources**
- Are there **extended sources erroneously described as groups of point sources**?

# The case of 4FGL J1813.1-1737e

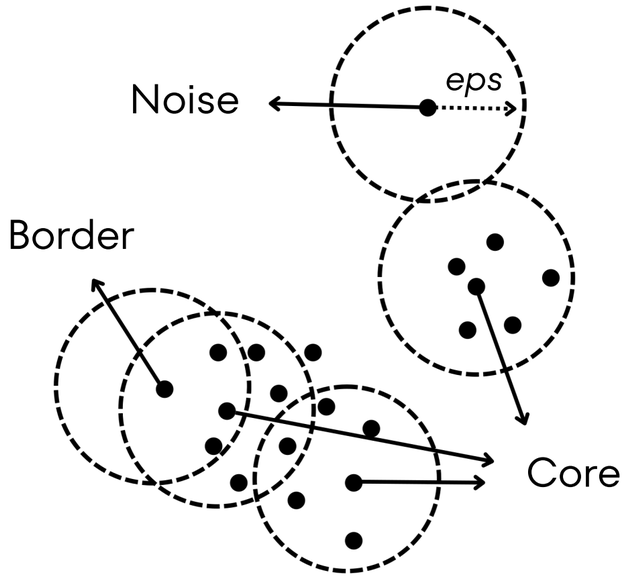


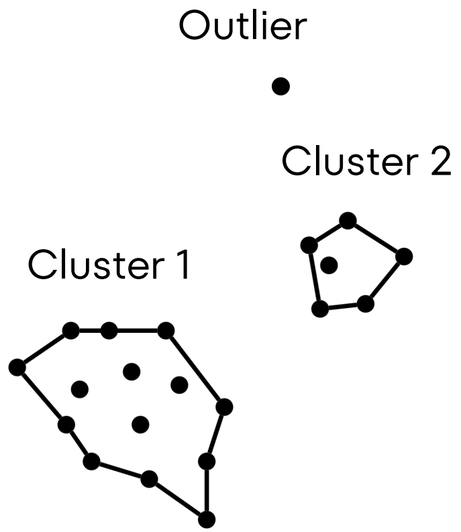

F. Aharonian et al. (2024) A&A, 686, A149

https://arxiv.org/abs/2403.16802

- **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).

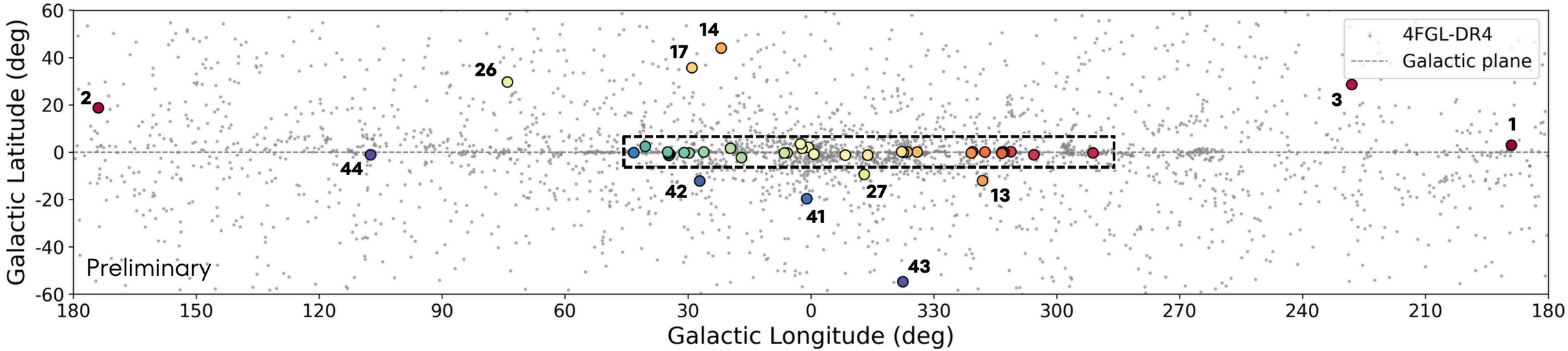
# Spatial clustering



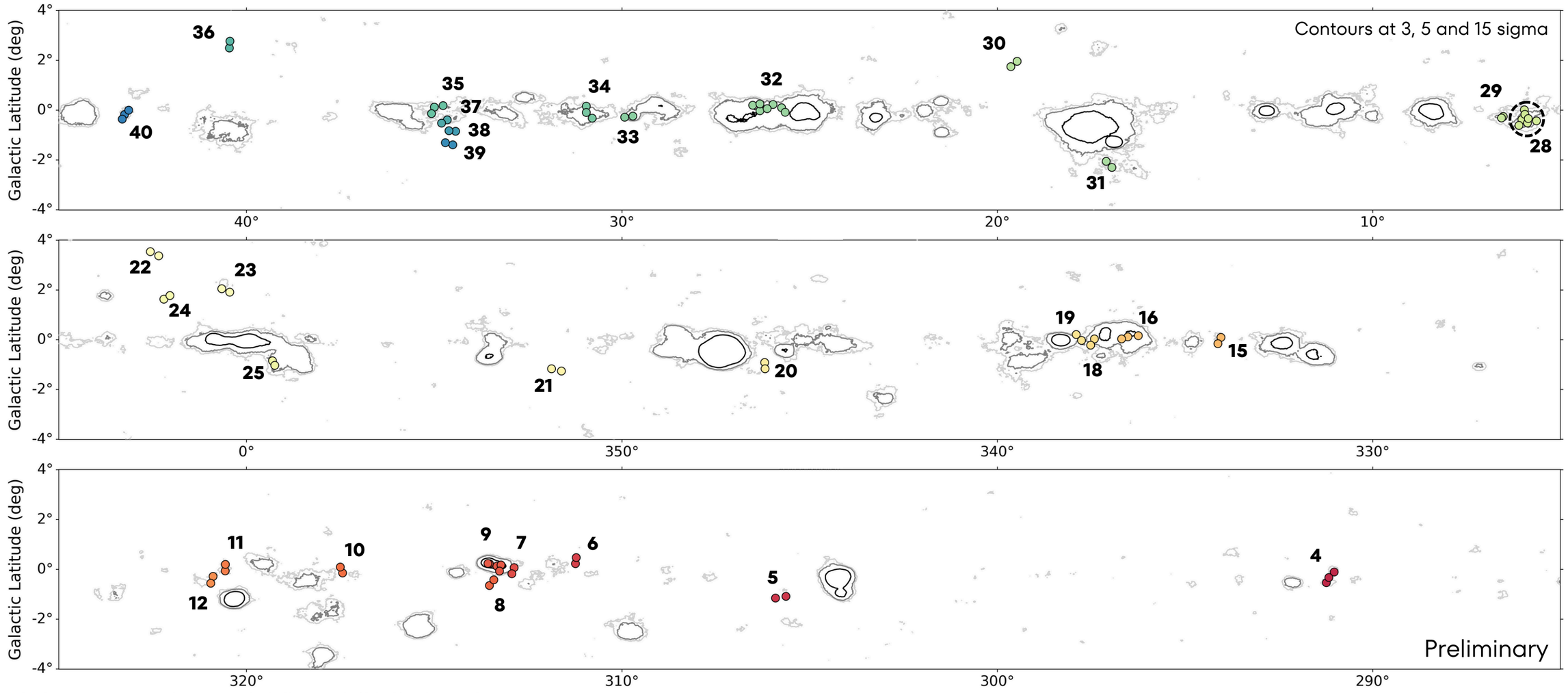

- Apply **unsupervised machine learning** to the 4FGL catalog.
- Used **DBSCAN** (Density-Based Spatial Clustering of Applications with Noise).
- DBSCAN **searches for clusters of points** and classifies the points into core, border or noise points (Ester et al., 1996).

# Fermi-LAT clusters map



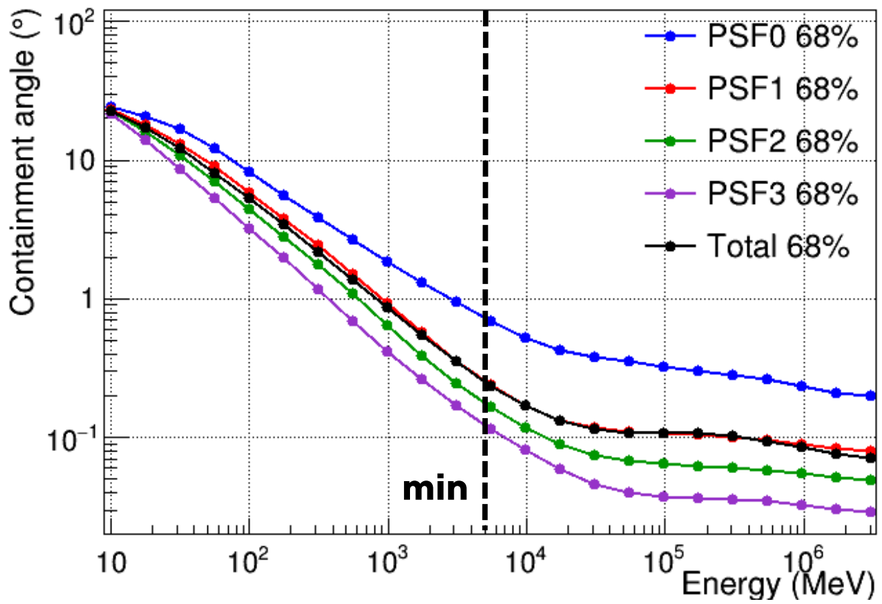

- Set **eps = 0.005 rad** ≈ 0.3 deg and **minPts = 2**.
- Included only **unassociated sources** and sources associated with **pulsars, supernova remnants**, and **active galaxies** (7030 sources in total).
- Found **44 clusters (106 sources)**, each including at least one unidentified source.
- There are **mostly clusters of size 2**, with some up to 7.

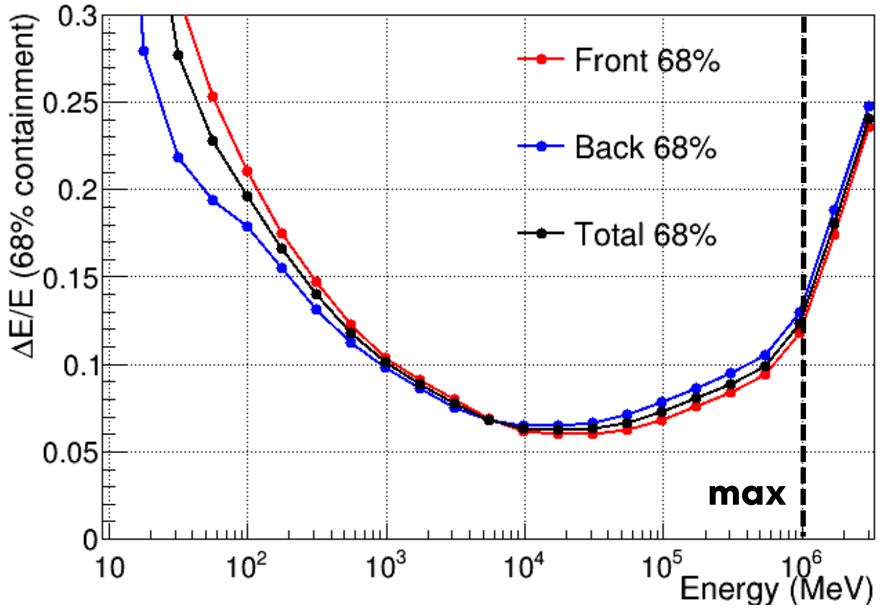
# HGPS (HESS Galactic Plane Survey) contours map




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

# Spatial binning




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).

# Test Statistics

- Likelihood nomenclatures:
  - **No-source** hypothesis: $\mathrm{L}_0$
  - **Single point** hypothesis: $\mathrm{L}_{\mathrm{ps}}$

- Test Statistic definitions:
  - **Extended source TS**: $\mathrm{TS} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_0)$
  - **Source extension TS**: $\mathrm{TS}_{\mathrm{ext}} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_{\mathrm{ps}})$
  - **N point sources TS**: $\mathrm{TS}_{\mathrm{Npts}} = 2\ \log(\mathrm{L}_{\mathrm{Npts}}/\mathrm{L}_{\mathrm{ps}})$

- Followed these criteria (Ackermann et al. 2017, ApJ, 843, 139):
  - **Claim a detection**: $\mathrm{TS} > 25$
  - **Define a source as extended**: $\mathrm{TS}_{\mathrm{ext}} > 16$
  - **Find the preferred model** (AIC test): $\mathrm{TS}_{\mathrm{ext}} > \mathrm{TS}_{\mathrm{Npts}} - 2\Delta k$

# Test Statistics

- Likelihood nomenclatures:
  - **No-source** hypothesis: $\mathrm{L}_0$
  - **Single point** hypothesis: $\mathrm{L}_{\mathrm{ps}}$

- Test Statistic definitions:
  - **Extended source TS**: $\mathrm{TS} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_0)$ ←
  - **Source extension TS**: $\mathrm{TS}_{\mathrm{ext}} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_{\mathrm{ps}})$
  - **N point sources TS**: $\mathrm{TS}_{\mathrm{Npts}} = 2\ \log(\mathrm{L}_{\mathrm{Npts}}/\mathrm{L}_{\mathrm{ps}})$

- Followed these criteria (Ackermann et al. 2017, ApJ, 843, 139):
  - **Claim a detection**: $\mathrm{TS} > 25$
  - **Define a source as extended**: $\mathrm{TS}_{\mathrm{ext}} > 16$
  - **Find the preferred model** (AIC test): $\mathrm{TS}_{\mathrm{ext}} > \mathrm{TS}_{\mathrm{Npts}} - 2\Delta k$

# Test Statistics

- Likelihood nomenclatures:
  - **No-source** hypothesis: $\mathrm{L}_0$
  - **Single point** hypothesis: $\mathrm{L}_{\mathrm{ps}}$

- Test Statistic definitions:
  - **Extended source TS**: $\mathrm{TS} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_0)$
  - **Source extension TS**: $\mathrm{TS}_{\mathrm{ext}} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_{\mathrm{ps}})$ ←
  - **N point sources TS**: $\mathrm{TS}_{\mathrm{Npts}} = 2\ \log(\mathrm{L}_{\mathrm{Npts}}/\mathrm{L}_{\mathrm{ps}})$

- Followed these criteria (Ackermann et al. 2017, ApJ, 843, 139):
  - **Claim a detection**: $\mathrm{TS} > 25$
  - **Define a source as extended**: $\mathrm{TS}_{\mathrm{ext}} > 16$
  - **Find the preferred model** (AIC test): $\mathrm{TS}_{\mathrm{ext}} > \mathrm{TS}_{\mathrm{Npts}} - 2\Delta k$

# Test Statistics

- Likelihood nomenclatures:
  - **No-source** hypothesis: $\mathrm{L}_0$
  - **Single point** hypothesis: $\mathrm{L}_{\mathrm{ps}}$

- Test Statistic definitions:
  - **Extended source TS**: $\mathrm{TS} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_0)$
  - **Source extension TS**: $\mathrm{TS}_{\mathrm{ext}} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_{\mathrm{ps}})$
  - **N point sources TS**: $\mathrm{TS}_{\mathrm{Npts}} = 2\ \log(\mathrm{L}_{\mathrm{Npts}}/\mathrm{L}_{\mathrm{ps}})$ ←

- Followed these criteria (Ackermann et al. 2017, ApJ, 843, 139):
  - **Claim a detection**: $\mathrm{TS} > 25$
  - **Define a source as extended**: $\mathrm{TS}_{\mathrm{ext}} > 16$
  - **Find the preferred model** (AIC test): $\mathrm{TS}_{\mathrm{ext}} > \mathrm{TS}_{\mathrm{Npts}} - 2\Delta k$

# Test Statistics

- Likelihood nomenclatures:
  - **No-source** hypothesis: $\mathrm{L}_0$
  - **Single point** hypothesis: $\mathrm{L}_{\mathrm{ps}}$

- Test Statistic definitions:
  - **Extended source TS**: $\mathrm{TS} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_0)$
  - **Source extension TS**: $\mathrm{TS}_{\mathrm{ext}} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_{\mathrm{ps}})$
  - **N point sources TS**: $\mathrm{TS}_{\mathrm{Npts}} = 2\ \log(\mathrm{L}_{\mathrm{Npts}}/\mathrm{L}_{\mathrm{ps}})$

- Followed these criteria (Ackermann et al. 2017, ApJ, 843, 139):
  - **Claim a detection**: $\mathrm{TS} > 25$ ←
  - **Define a source as extended**: $\mathrm{TS}_{\mathrm{ext}} > 16$
  - **Find the preferred model** (AIC test): $\mathrm{TS}_{\mathrm{ext}} > \mathrm{TS}_{\mathrm{Npts}} - 2\Delta k$

# Test Statistics

- Likelihood nomenclatures:
  - **No-source** hypothesis: $\mathrm{L}_0$
  - **Single point** hypothesis: $\mathrm{L}_{\mathrm{ps}}$

- Test Statistic definitions:
  - **Extended source TS**: $\mathrm{TS} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_0)$
  - **Source extension TS**: $\mathrm{TS}_{\mathrm{ext}} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_{\mathrm{ps}})$
  - **N point sources TS**: $\mathrm{TS}_{\mathrm{Npts}} = 2\ \log(\mathrm{L}_{\mathrm{Npts}}/\mathrm{L}_{\mathrm{ps}})$

- Followed these criteria (Ackermann et al. 2017, ApJ, 843, 139):
  - **Claim a detection**: $\mathrm{TS} > 25$
  - **Define a source as extended**: $\mathrm{TS}_{\mathrm{ext}} > 16$ ←
  - **Find the preferred model** (AIC test): $\mathrm{TS}_{\mathrm{ext}} > \mathrm{TS}_{\mathrm{Npts}} - 2\Delta k$

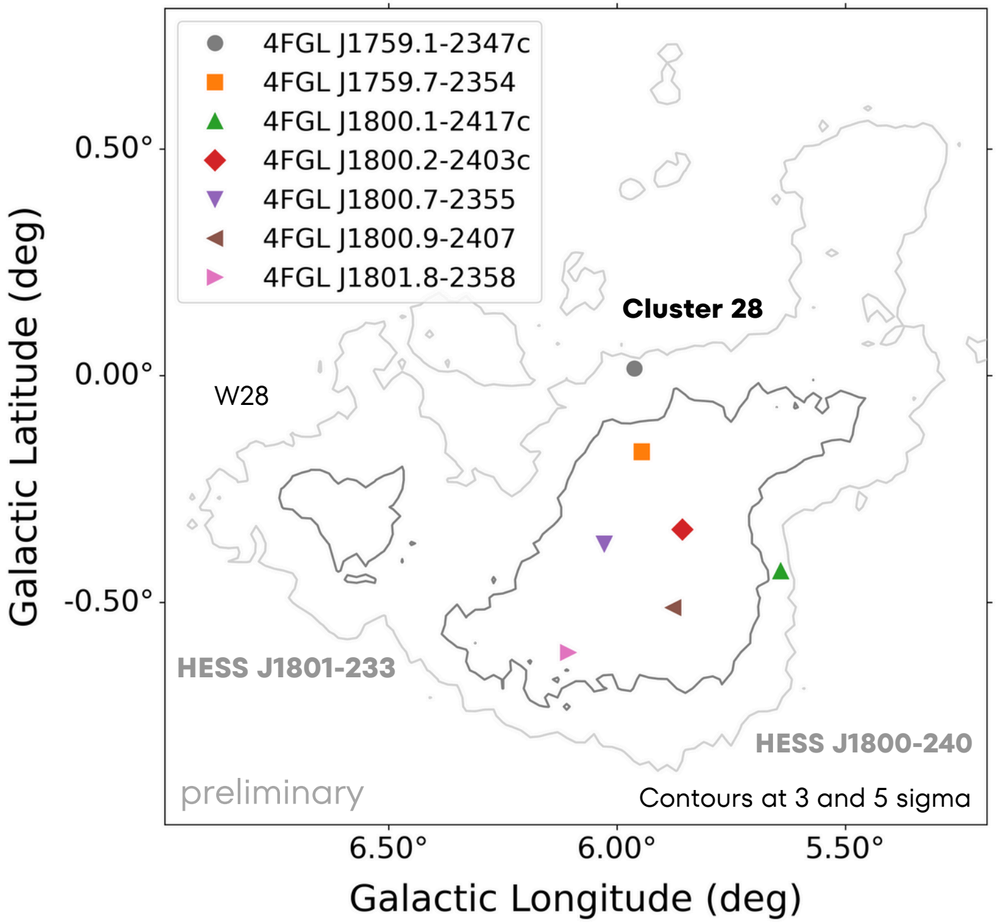
# Test Statistics

- Likelihood nomenclatures:
  - **No-source** hypothesis: $\mathrm{L}_0$
  - **Single point** hypothesis: $\mathrm{L}_{\mathrm{ps}}$

- Test Statistic definitions:
  - **Extended source TS**: $\mathrm{TS} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_0)$
  - **Source extension TS**: $\mathrm{TS}_{\mathrm{ext}} = 2\ \log(\mathrm{L}_{\mathrm{ext}}/\mathrm{L}_{\mathrm{ps}})$
  - **N point sources TS**: $\mathrm{TS}_{\mathrm{Npts}} = 2\ \log(\mathrm{L}_{\mathrm{Npts}}/\mathrm{L}_{\mathrm{ps}})$

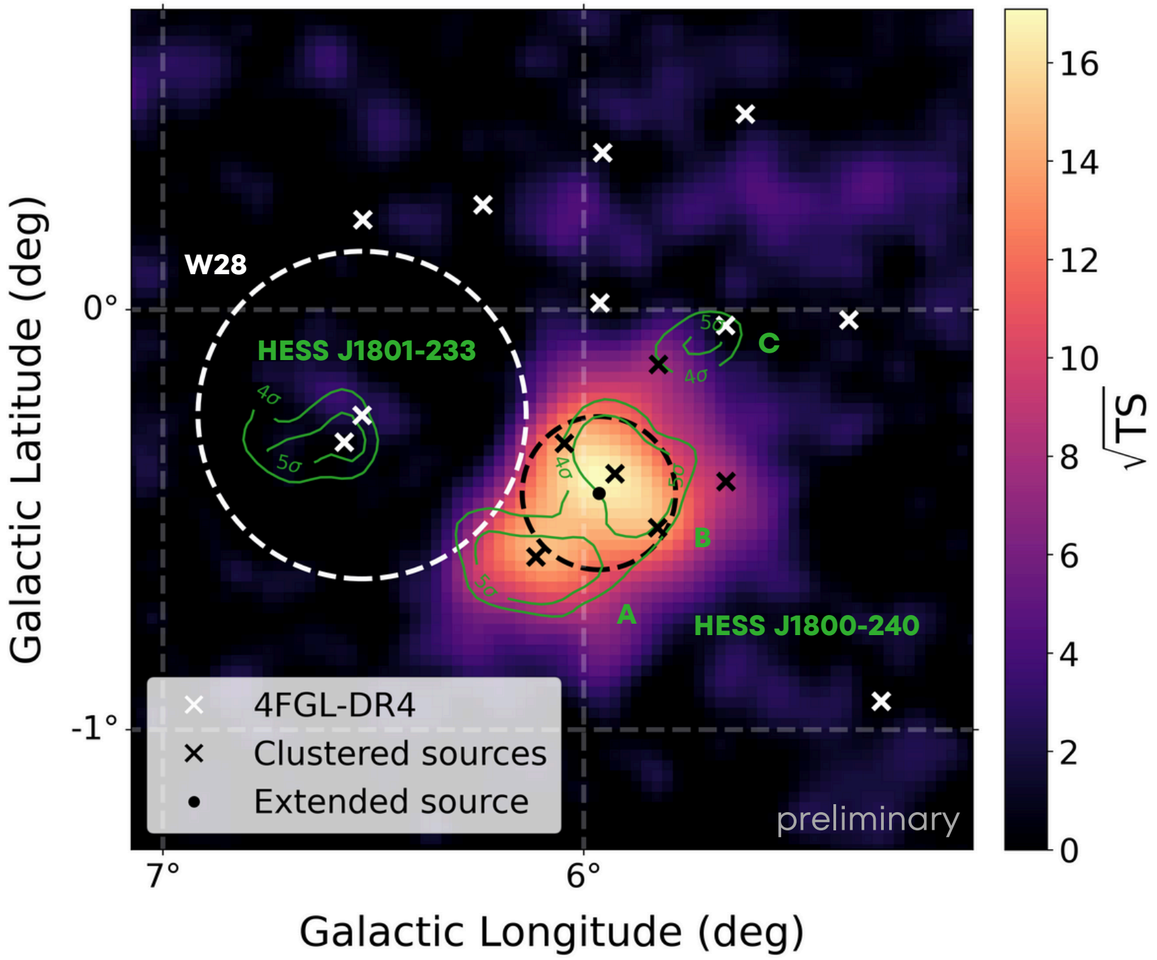
- Followed these criteria (Ackermann et al. 2017, ApJ, 843, 139):
  - **Claim a detection**: $\mathrm{TS} > 25$
  - **Define a source as extended**: $\mathrm{TS}_{\mathrm{ext}} > 16$
  - **Find the preferred model** (AIC test): $\mathrm{TS}_{\mathrm{ext}} > \mathrm{TS}_{\mathrm{Npts}} - 2\Delta k$ ←

# Cluster 28 example



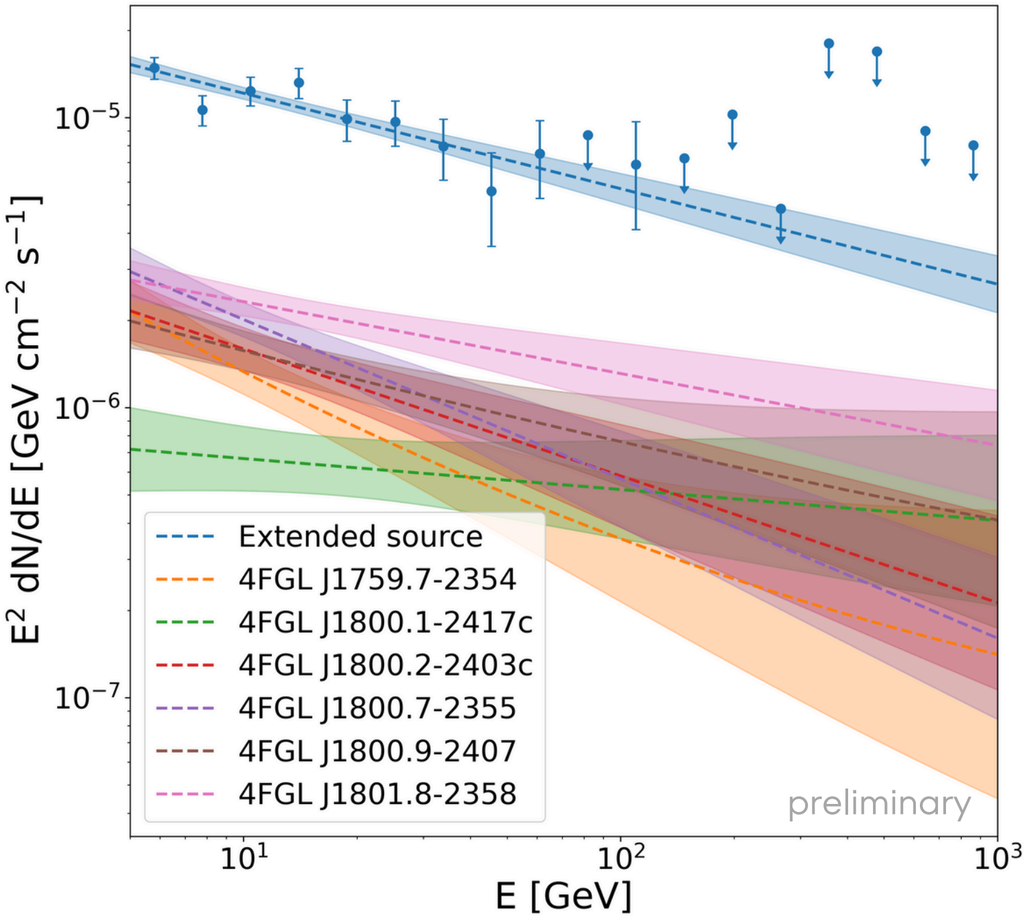

- There are **seven unassociated sources** (grouped together in a cluster by the DBSCAN algorithm) coincident with a significant residual in HGPS.
- For the following analysis, we used **Fermipy** v1.2.3 (Wood et al., 2017).
- We **fit an extended source model to the cluster**, excluding 4FGL J1759.1-2347c for which a point source model is preferred.

# Cluster 28 residuals map




- **TS results**:
  - $\mathrm{TS} = 747$
  - $\mathrm{TS}_{\mathrm{ext}} = 407$
  - $\mathrm{TS}_{\mathrm{6pts}} = 378$
- **TS criteria**:
  - ✓ $\mathrm{TS} > 25$
  - ✓ $\mathrm{TS}_{\mathrm{ext}} > 16$
  - ✓ $\mathrm{TS}_{\mathrm{ext}} > \mathrm{TS}_{\mathrm{6pts}} - 34$
- **Spatial model**:
  - type: *RadialGaussian*
  - $\mathrm{GLON} = (5.96 \pm 0.01)\,\mathrm{deg}$
  - $\mathrm{GLAT} = -(0.44 \pm 0.01)\,\mathrm{deg}$
  - $\mathrm{r}_{68\%} = (0.26 \pm 0.01)\,\mathrm{deg}$

# Cluster 28 energy spectrum




- **Spectrum type**:
  - PowerLaw: $\frac{dN}{dE} = N_0 \left( \frac{E}{E_0} \right)^{\Gamma}$
- **Spectrum parameters**:
  - $\mathrm{N_0} = (2.6 \pm 0.4) \times 10^{-11} \, \frac{1}{\mathrm{MeV^2 \, cm^2 \, s}}$
  - $\Gamma = -2.33 \pm 0.06$
  - $\mathrm{E_0} = 1000 \, \mathrm{MeV}$
- **Curvature test** relative to *PowerLaw*:
  - *LogParabola*: $1.152$
  - *PLSuperExpCutoff4*: $0.909$

# Summary

- **Spatial clustering** with the DBSCAN algorithm for unidentified point sources from the 4FGL-DR4 catalog **to search for potentially mis-identified extended sources**.

- We use **Fermipy to compare a collection of point sources model and a single extended source model** for clusters in the Galactic plane.

- The example of Cluster 28 shows that, **potentially, there are extended sources instead of clusters of unassociated Fermi-LAT sources**

# Outlook

- **Include systematic errors** (e.g., uncertainties in the Galactic diffuse emission, on the shape of the extended source, and in our knowledge of the Fermi-LAT IRFs).
- **Vary the radius** from 0.3 up to e.g. 0.5 degrees.
- Explore on **TeV/MWL context**.
- Look at the clusters in depth (e.g. **joint fits of Fermi-LAT and H.E.S.S. data** in Gammapy (Donath et al., 2023).
- Extend this analysis to **incorporate eROSITA data** (see the poster “Towards a joint X-ray and gamma-ray analysis of Pulsar Wind Nebulae with Gammapy” by K. Egg).

# Thank you for listening!

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

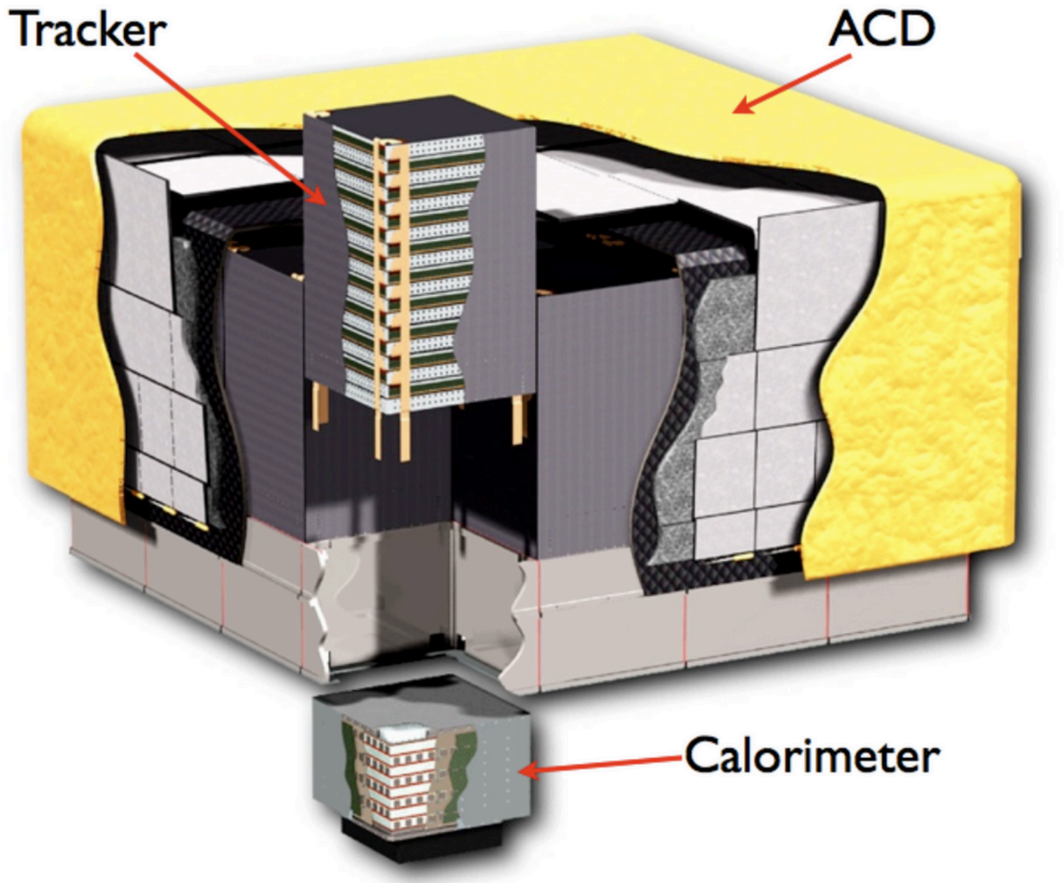
# Back-Up (1)

- **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

# Back-Up (2)

- 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)

# Back-Up (3)




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).