


# A catalog of gamma-ray source candidates from 16 years of Fermi-LAT data at $|b| > 10^\circ$

F. Akhondi, M. Soor, and H. Hedayati Kh. 

Department of Physics, K. N. Toosi University of Technology, PO Box 15875-4416, Tehran, Iran

Received 19 August 2025 / Accepted 29 December 2025

## ABSTRACT

**Context.** Sixteen years of continuous observations with the Fermi Large Area Telescope (LAT) provide an unprecedented opportunity to search for faint or previously unrecognized  $\gamma$ -ray emitters, particularly at high energies where the angular resolution improves and the diffuse background is reduced.

**Aims.** We present a new catalog of  $\gamma$ -ray source candidates identified using a data-driven, clustering-based analysis of Pass 8 LAT data above 3 GeV at Galactic latitudes of  $|b| > 10^\circ$ .

**Methods.** The pipeline integrates the Slewing MEthod for Finding Core, version 3 (SIMEFIC III), for photon denoising with the density-based spatial clustering of applications with noise (DBSCAN) algorithm.

**Results.** We detect 3112  $\gamma$ -ray cluster candidates, of which 2932 are associated with known  $\gamma$ -ray sources. Crossmatching yields 37 counterparts in 1CGH, 23 in CRATES, 21 in Roma-BZCAT, and 17 in MST-12yr. The remaining 370 clusters have no known  $\gamma$ -ray counterpart and are proposed as new source candidates.

**Key words.** methods: data analysis – methods: numerical – methods: statistical – catalogs

## 1. Introduction

The Fermi Gamma-ray Space Telescope, launched in 2008, has revolutionized gamma-ray astronomy by surveying the entire sky every  $\sim 3$  hours (approximately two orbits) and collecting unprecedented volumes of high-energy astrophysical data (Ackermann et al. 2012). Its primary instrument, the Large Area Telescope (LAT) (Atwood et al. 2009), detects gamma rays from  $\sim 20$  MeV to  $>1$  TeV, recording each photon's arrival time, direction, and energy (Abdollahi et al. 2022).

The LAT is a pair-conversion telescope comprising three subsystems:

- Tracker (converter): incoming gamma rays interact with high- $Z$  material, converting to electron-positron pairs. The trajectories of these charged particles enable reconstruction of the original gamma-ray direction.
- Calorimeter: measures the energy of the pair by absorbing and sampling the electromagnetic shower.
- Anticoincidence detector (ACD): surrounds the tracker and vetoes charged-particle backgrounds via coincident signal rejection.

When analyzing photon directions (sky coordinates), LAT data produce 2D sky maps where point sources manifest as localized excesses of gamma-ray events. Traditional point-source detection in LAT data – similar to its predecessor EGRET (Kanbach et al. 1989) – relies on maximum likelihood estimation. A source is considered significantly detected if the test statistic (TS) exceeds 25, where TS is defined as

$$TS = 2 \ln(\mathcal{L}/\mathcal{L}_0), \quad (1)$$

with  $\mathcal{L}$  denoting the maximum likelihood of the model including the source and  $\mathcal{L}_0$  the likelihood without it (Mattox et al.

1996). Following the Fermi-LAT convention and the  $TS$ - $\sigma$  relation discussed in Arsioli et al. (2025, footnote 3), a detection threshold of  $TS \approx 25$  corresponds to a significance of approximately  $4\sigma$ , while  $TS \approx 12$  corresponds to  $\sim 3\sigma$  when considering two degrees of freedom.

However, the LAT's high sensitivity and the increasing source density within small regions of interest (ROIs) render likelihood-based model fitting computationally prohibitive (Starck et al. 2009). As a result, non-parametric methods have gained traction in astrophysics over the past two decades due to their computational efficiency and minimal model assumptions. Several algorithms adapted for LAT data leverage local photon density, including minimum spanning trees (MST), as described by Campana et al. (2008), and DBSCAN (Tramacere & Vecchio 2013). While faster than likelihood-based methods, these techniques require careful background suppression and hyperparameter tuning.

In this study, we detect gamma-ray point sources by applying the SIMEFIC III denoising algorithm to LAT data prior to clustering (see Hedayati Kh. et al. (2024); Soor et al. (2025) for details). By effectively reducing background contamination, SIMEFIC III facilitates the use of simpler clustering algorithms that are less sensitive to hyperparameter selection.

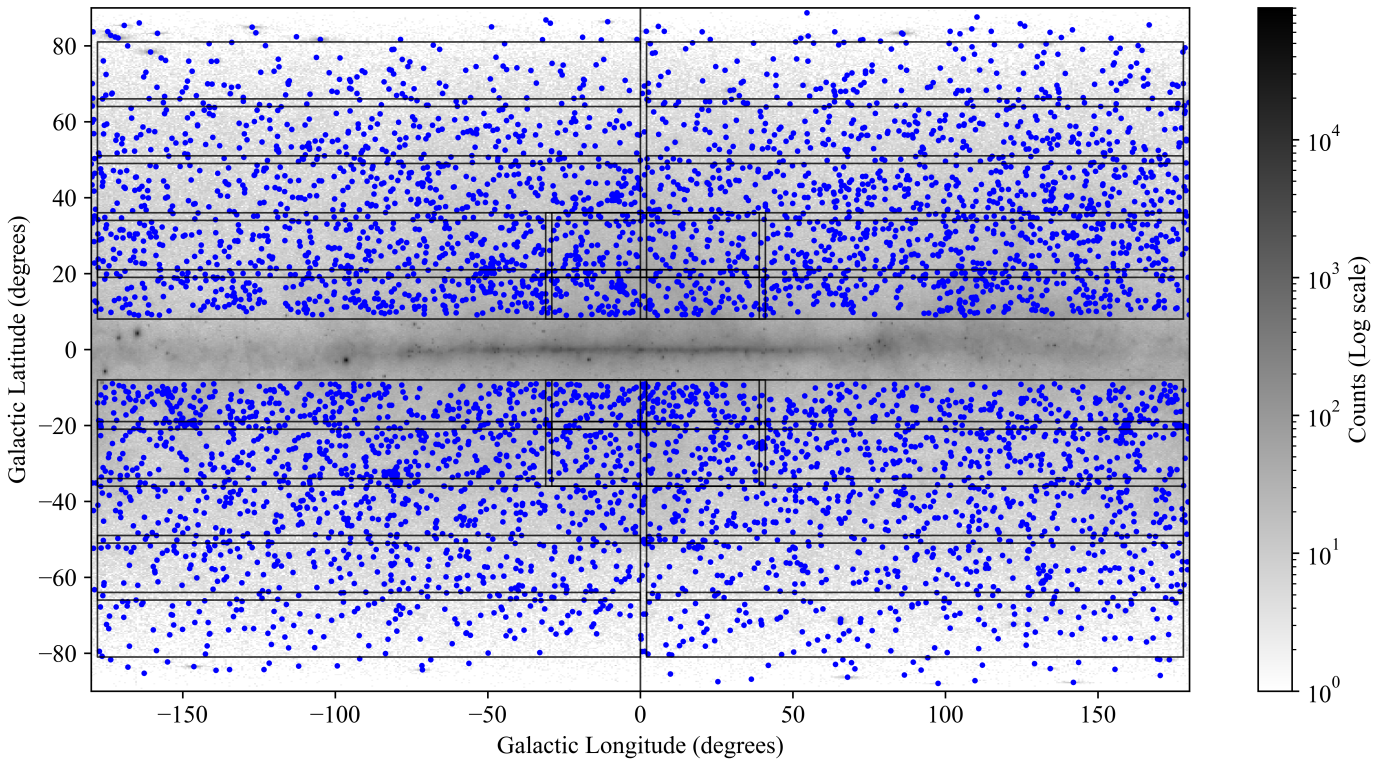
In Section 2, we outline our methodology for analyzing the Fermi-LAT data. Section 3 provides an overview of the DBSCAN algorithm, discussing its implementation in our analysis as well as the parameters that were carefully tuned to optimize source detection. Section 4 provides an overview of the SIMEFIC III. In Section 5, we present the results of our analysis, including a detailed breakdown of identified gamma-ray sources, their classifications, and associations with existing catalogs. Finally, in Section 6, we conclude with a discussion of our findings.

\* Corresponding author: hedayati@kntu.ac.ir

**Table 1.** Comparison of energy thresholds in Fermi-LAT analyses.

Energy threshold	PSF (68% containment)	Relative count rate (normalized to 1 GeV)	Primary use case
0.1 GeV	$\sim 5.0^\circ$	Very high (10 $\times$ )	Diffuse emission studies
1.0 GeV	$\sim 0.8^\circ$	Reference (1 $\times$ )	General source detection
3.0 GeV	$\sim 0.3^\circ$	Moderate (0.4 $\times$ )	Crowded field analysis
10.0 GeV	$< 0.1^\circ$	Low (0.1 $\times$ )	Point source localization

**Notes.** PSF values represent 68% containment radii for combined front+back converting events, averaged over incidence angles (SLAC National Accelerator Laboratory n.d.).



**Fig. 1.** Photon count map showing the 30 overlapping analysis regions, based on the 9Y-MST tiling scheme with additional fields added near the Galactic plane ( $10^\circ < |b| < 20^\circ$ ). Each region overlaps adjacent fields by  $2^\circ$  to reduce edge effects during source detection. The inner Galactic plane ( $|b| \leq 10^\circ$ ) is excluded from the analysis due to intense diffuse gamma-ray emission. DBSCAN clusters are denoted by blue circles.

## 2. Event selection and sky partitioning

We analyzed 16 years of Fermi-LAT Pass 8 data (August 4, 2008, to August 4, 2024) from the Fermi Science Support Center (FSSC) archive. We selected Source-class events (`evclass = 128`) with energies of  $E > 3$  GeV, balancing angular resolution and photon statistics (see Table 1). At 3 GeV, the LAT point spread function (PSF) improves to  $\sim 0.3^\circ$  (68% containment) compared to  $0.8^\circ$  at 1 GeV, while retaining  $\sim 40\%$  of the photon flux, providing an optimal compromise for detecting faint sources in crowded regions. The following event selection criteria were applied:

- Event class: Source (`evclass = 128`).
- Event type: Both front- and back-converting events (`evtype = 3`).
- Zenith angle:  $\leq 105^\circ$  (to minimize Earth limb contamination).

- Data quality: `DATA_QUAL==1 && LAT_CONFIG==1 && |ROCK_ANGLE| > 52^\circ`.

To account for the nonuniform photon distribution across the sky, we divided the gamma-ray sky into 30 ROIs:

- 22 regions follow the 9Y-MST tiling (Campana et al. 2018).
- 8 additional regions along the Galactic plane ( $10^\circ < |b| < 20^\circ$ ) for enhanced coverage.

Each ROI has a  $2^\circ$  overlap with neighboring regions to mitigate edge effects. The Galactic plane ( $|b| < 10^\circ$ ) was excluded due to: (1) dominant diffuse gamma-ray emission, (2) significant background contamination, and (3) challenges in reliable point source identification in complex emission regions.

The final dataset contains 3 198 820 photons with  $E > 3$  GeV, roughly 10 $\times$  more photons than the 324 133 photons above 10 GeV in the 12Y-MST catalog (Campana & Massaro 2021). Figure 1 shows the ROI layout, while Table 2 provides the region parameters in detail.

**Table 2.** 30 sky regions with a 2° overlap.

Num.	Region		Solid angle sr	Photon number		K		Denosing step		Total clusters	
	$l$ (°)	$b$ (°)		N	S	N	S	N	S	N	S
1.	$0^\circ < l < 360^\circ$	$ b  > 79^\circ$	0.1154	28 649	18 091	10	8	3	2	20	6
2.	$0^\circ < l < 182^\circ$	$64^\circ <  b  < 81^\circ$	0.2824	70 544	44 275	7	6	3	4	30	43
3.	$178^\circ < l < 362^\circ$	$64^\circ <  b  < 81^\circ$	0.2955	63 674	43 131	7	6	3	4	20	53
4.	$0^\circ < l < 182^\circ$	$49^\circ <  b  < 66^\circ$	0.5045	126 461	96 246	8	7	4	3	40	162
5.	$178^\circ < l < 362^\circ$	$49^\circ <  b  < 66^\circ$	0.5101	112 153	101 351	7	7	4	2	110	54
6.	$0^\circ < l < 182^\circ$	$34^\circ <  b  < 51^\circ$	0.6923	203 405	183 021	9	9	3	3	113	98
7.	$178^\circ < l < 362^\circ$	$34^\circ <  b  < 51^\circ$	0.6999	174 319	188 468	8	9	3	3	389	145
8.	$0^\circ < l < 41^\circ$	$19^\circ <  b  < 31^\circ$	0.1876	96 238	74 004	14	12	2	3	21	30
9.	$39^\circ < l < 182^\circ$	$19^\circ <  b  < 31^\circ$	0.6544	260 937	204 731	11	10	3	4	208	287
10.	$178^\circ < l < 331^\circ$	$19^\circ <  b  < 31^\circ$	0.7002	176 861	256 541	8	11	3	4	262	181
11.	$329^\circ < l < 362^\circ$	$19^\circ <  b  < 31^\circ$	0.5010	79 189	64 250	14	13	6	5	25	30
12.	$0^\circ < l < 41^\circ$	$9^\circ <  b  < 21^\circ$	0.1445	125 324	89 192	20	16	6	3	6	18
13.	$39^\circ < l < 182^\circ$	$9^\circ <  b  < 21^\circ$	0.504	355 172	268 450	17	14	5	5	185	152
14.	$178^\circ < l < 331^\circ$	$9^\circ <  b  < 21^\circ$	0.5392	216 294	295 615	11	14	2	6	94	63
15.	$329^\circ < l < 362^\circ$	$9^\circ <  b  < 21^\circ$	0.1163	97 985	72 052	19	16	7	7	147	120

**Notes.** The total number of photons, solid angle, and the corresponding K for both the north and south Galactic regions are provided. The best denosing step and the corresponding number of clusters are also provided.

### 3. SIMEFIC III algorithm overview

The SIMEFIC III algorithm is a denosing technique designed for datasets with numerous localized intensity peaks embedded in substantial background noise. Instead of modifying the intrinsic characteristics of these peaks, it suppresses interstitial noise, thereby making their locations more distinct and easier to identify. Although not originally developed specifically for gamma-ray photon maps, its application to such data has yielded promising results.

SIMEFIC III works by quantifying local density variations to identify regions of locally elevated intensity that deviate from the expected background behavior. Through this approach, the algorithm selectively suppresses background fluctuations, while preserving key signal features.

By reducing spurious noise and emphasizing genuine local extrema, SIMEFIC III prepares data for downstream analysis. For example, when paired with clustering methods such as DBSCAN, the denoised dataset enhances source detection reliability by decreasing false positives, especially in regions with complex spatial structures.

For a detailed discussion of the SIMEFIC III denosing method, refer to [Hedayati Kh. et al. \(2024\)](#), which outlines its theoretical foundations, performance in high-noise environments, and use as a preprocessing tool. An enhanced, iterative version of SIMEFIC III that applies multiple denosing passes to further improve signal extraction is presented in our follow-up study ([Soor et al. 2025](#)).

The denosing algorithm was applied iteratively to each of the 30 analysis regions. After each iteration, photons with IDs greater than the average were removed, progressively reducing background noise. This process continued until an optimal denosing step was identified for each region. The selected step numbers – representing the best trade-off between noise suppression and source retention – are listed in [Table 2](#). The optimal step for each region was defined as the one in which the remaining clusters exhibited the highest correspondence with known

Fermi-LAT sources having a TS greater than 25. Specifically, the optimal step maximized the number of clusters lying within a minimal angular distance of such Fermi sources (see [Soor et al. \(2025\)](#) for details).

### 4. DBSCAN clustering

Density-based spatial clustering of applications with noise (DBSCAN; [Ester et al. \(1996\)](#)) is a clustering algorithm that utilizes two key parameters: epsilon ( $\epsilon$ ), which defines the neighborhood radius around each point, and minimum points (MinPts), which specifies the minimum number of neighboring points required for a point to be classified as a core point. The algorithm categorizes points as core, border, or noise and links core points to form clusters of high density.

We applied the DBSCAN algorithm to identify photon clusters in Fermi-LAT data. The performance of DBSCAN is sensitive to the choice of  $\epsilon$  and MinPts, and these parameters must be tuned based on the specific characteristics of the dataset. According to [Tramacere & Vecchio \(2013\)](#), the parameter  $\epsilon$  is related to the PSF of the LAT detector. Although our analysis uses photons with energies above 3 GeV, we adopt  $\epsilon = 0.1^\circ$  following [Campana et al. \(2022\)](#), who associate this scale with the Fermi-LAT PSF at 10 GeV. This conservative choice favors the detection of compact, high-confidence clusters and helps suppress spurious associations in crowded fields. The MinPts parameter, denoted as  $K$ , was estimated for each region using the method described by [Campana et al. \(2022\)](#):

1. Average photon density:  $K$  was initially estimated based on the average photon density,

$$K_{\text{ave}} = \frac{N_{\text{tot}} \Omega_\epsilon}{\Omega_{R_{\text{extr}}}}, \quad (2)$$

where  $\Omega_{R_{\text{extr}}}$  is the solid angle of the extraction region,  $\Omega_\epsilon$  is the solid angle of a circle with radius  $\epsilon$ , and  $N_{\text{tot}}$  is the total number of photons within the extraction region.

**Table 3.** Astronomical catalogs used for crossmatching.

Catalog	Description	Article	Source type	Matches
CRATES	A large sample of bright, compact, flat-spectrum radio sources	Healey et al. (2007)	Radio sources	23
TeVcat	An online database of Very High Energy (TeV) gamma-ray sources	Wakely & Horan (2008)	Gamma-ray sources	–
WIBRaLS	A catalog of radio-loud candidate $\gamma$ -ray emitting blazars assembled from WISE sources	D’Abrusco et al. (2014)	Blazars	8
Roma-BZCAT	Fifth edition of the Roma-BZCAT catalog of blazars based on multifrequency surveys	Massaro et al. (2015)	Blazars	21
2WHSP	A large list of high-synchrotron-peaked blazars (HSPs)	Chang et al. (2017)	Blazars	6
3HSP	A sample of extreme and high-synchrotron-peaked (EHSP/HSP) blazars and candidates	Chang et al. (2019)	Blazars	–
2BIGB	Gamma-ray analysis of EHSP and HSP blazars using 11-year Fermi-LAT data	Arsioli et al. (2020)	Blazars	–
MST-12yr	A catalog of MST sources identified in 12-year Fermi-LAT data	Campana & Massaro (2021)	Gamma-ray sources	17
4FGL-DR4	Data Release 4 of the Fermi-LAT catalog based on 14 years of observations	Ballet et al. (2024)	Gamma-ray sources	–
1CGH	Gamma-ray detections above 10 GeV based on 16 years of Fermi-LAT observations	Arsioli et al. (2025)	Gamma-ray sources	37

**Notes.** Summary of the catalogs used for crossmatching with photon clusters identified using SIMEFIC III followed by clustering with DBSCAN. The last column lists the number of positional matches found in our analysis.

2. Background density estimation: To estimate the background density,  $D_{\text{bkg}}$ , we ran DBSCAN with  $K = K_{\text{ave}} + K_{\text{frac}} \sqrt{K_{\text{ave}}}$ , where  $K_{\text{frac}} = 3.5$ .
3. Background  $K$  Estimation: The background  $K$  was then calculated as

$$K_{\text{bkg}} = D_{\text{bkg}} \Omega_{\epsilon}. \quad (3)$$

4. Final  $K$  value: the final value of  $K$  used in DBSCAN was computed as

$$K = K_{\text{bkg}} + K_{\text{frac}} \sqrt{K_{\text{bkg}}}. \quad (4)$$

For each photon field, the final estimated value of  $K$  is listed in Table 2. This value was held constant throughout the DBSCAN iterations for denoising.

To assess the statistical significance of the clusters identified by DBSCAN, we applied the likelihood ratio method introduced by Li & Ma (1983):

$$S = \sqrt{2 \left( N_{\text{on}} \ln \left( \frac{2N_{\text{on}}}{N_{\text{on}} + N_{\text{off}}} \right) + N_{\text{off}} \ln \left( \frac{2N_{\text{off}}}{N_{\text{on}} + N_{\text{off}}} \right) \right)}, \quad (5)$$

where  $N_{\text{on}}$  represents the number of photons in the cluster region, and  $N_{\text{off}}$  is the number of background photons. This expression provides a quantitative measure of cluster significance, facilitating the discrimination between genuine sources and background fluctuations.

We adopted a threshold of  $S = 2$  for source significance, rejecting clusters with  $S < 2$  as noise. The SIMEFIC algorithm was applied iteratively to the dataset, and DBSCAN was reapplied in each iteration to refine the cluster selection.

## 5. SIMEFIC III catalog

To evaluate the performance of our point-source detection pipeline, we built a preliminary catalog of gamma-ray sources based on DBSCAN clusters extracted from SIMEFIC III-denoised photon data. This catalog includes 3112 clusters that are spatially compact, statistically significant, and consistent with the expected characteristics of astrophysical gamma-ray sources.

For each identified cluster, we determined its centroid coordinates (Galactic longitude and latitude, as well as right ascension and declination). The centroid was computed using a SIMEFIC-weighted average of photon positions, as described in our previous work Soor et al. (2025), ensuring a more accurate representation of the spatial distribution of significant local maxima within the cluster.

To suppress contamination from background noise, we applied a conservative photon multiplicity threshold of  $N_{\gamma} \geq 4$ , based on both simulations of the isotropic background and empirical studies of low-count Poisson fluctuations.

### 5.1. 4FGL-DR4 associated clusters

We crossmatched the resulting source catalog with the Fermi-LAT Fourth Source Catalog Data Release 4 (4FGL-DR4; Ballet et al. 2024), which lists 7194 sources detected between 50 MeV and 1 TeV. For this comparison, we adopted a matching radius of  $0.15^\circ$ , corresponding to the median PSF containment radius of the LAT above  $\sim 3$  GeV:  $\sim 0.1^\circ$  for front-converting and  $\sim 0.2^\circ$  for back-converting events. This choice balances completeness against reliability, preventing the loss of true associations at  $< 0.1^\circ$ , while minimizing false matches at  $> 0.3^\circ$ , in line with 4FGL-DR2/DR4 and MST-12yr methodologies for robust  $> 1$  GeV source associations.

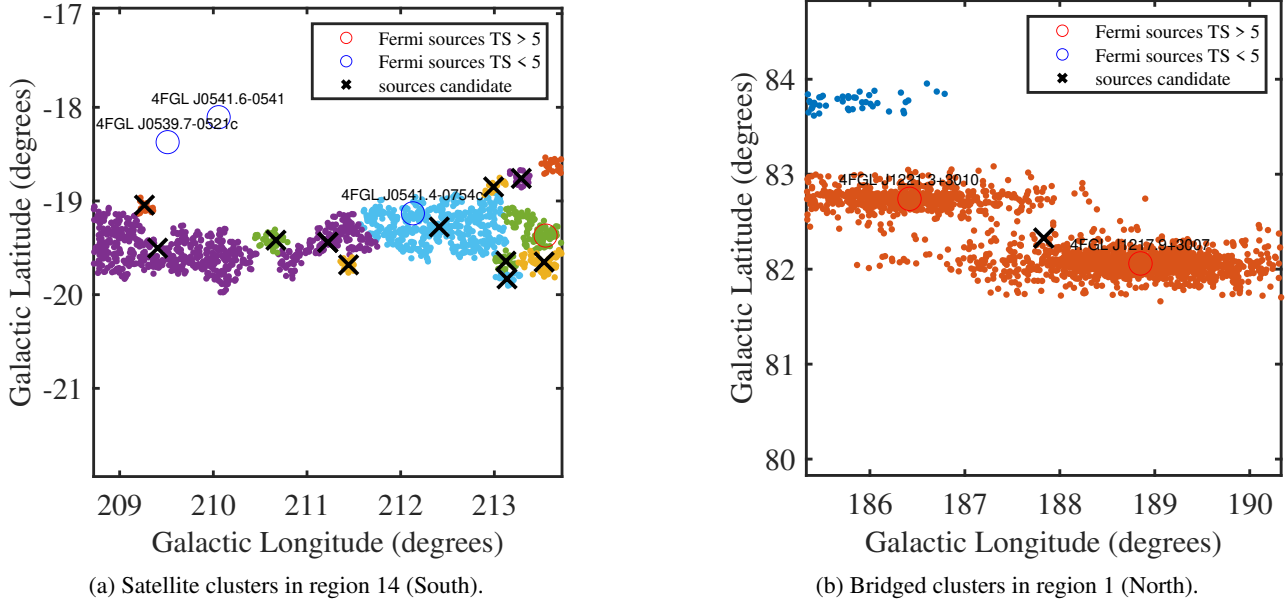
Using this criterion, 2822 clusters were associated with 4FGL-DR4 sources. Unmatched clusters were flagged as candidate new detections and further evaluated to exclude spurious signals arising from instrumental artifacts or diffuse background fluctuations. As an additional validation step, photon maps were visually inspected.

It is important to note that our analysis covers 16 years of Fermi-LAT observations (August 2008–August 2024), whereas 4FGL-DR4 is based on the first 14 years. Consequently, some of our detections may appear with enhanced significance or may be absent from 4FGL-DR4 due to the increased photon statistics in our dataset.

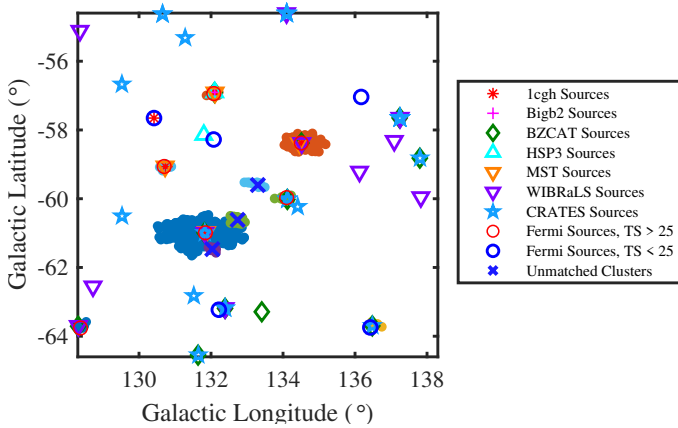
### 5.2. New gamma-ray source candidates

To broaden source identification, we crossmatched our catalog with the additional  $\gamma$ -ray and multiwavelength catalogs listed in Table 3, using the same  $0.15^\circ$  radius adopted for the 4FGL-DR4 associations. The crossmatch yielded 112 counterparts, with the distribution of matches summarized in Table 3.

A subset of 370 DBSCAN-identified clusters showed no positional association with 4FGL-DR4 sources or any other known gamma-ray or multiwavelength catalog and were flagged as new gamma-ray source candidates. These clusters constitute promising targets for follow-up analyses in future Fermi-LAT data releases and multiwavelength studies, particularly in regions with limited prior coverage or high source confusion. As with 4FGL-DR4 associations, candidate clusters were required to satisfy  $N_{\gamma} \geq 4$  and were visually inspected on photon maps to



**Fig. 2.** Examples of special cluster morphologies. DBSCAN clusters are shown in different colors with black crosses marking centroids. 4FGL-DR4 sources with  $TS > 25$  and  $TS < 25$  are shown as red and blue circles, respectively. Panel a: fragmented extended source producing satellite clusters. Panel b: single DBSCAN cluster encompassing two nearby 4FGL-DR4 sources, a typical bridged case.



**Fig. 3.** Example slice from one of the 30 analyzed regions, showing a portion of Region 4 (South) after the third denoising step. The central source candidate, SIM06\_J0112+0254, is clearly visible near the center of the image, while two additional sources associated with satellite clusters – SIM08\_J0110+0149 and SIM08\_J0107+0104 – appear in the lower left area. The central source remains detectable up to the sixth denoising step, whereas the two satellite clusters persist until the eighth step.

exclude artifacts arising from diffuse backgrounds or instrumental effects.

Figure 3 shows a  $10^\circ \times 10^\circ$  region centered at Galactic coordinates ( $l = 133.3^\circ, b = -59.6^\circ$ ), representing a subsection of Region 4 (South). After applying the denoising procedure at the third iteration, the DBSCAN algorithm (minimum cluster size of four photons) was executed. The resulting clusters are displayed in distinct colors. For each cluster, a search within a  $0.15^\circ$  radius of its centroid was performed to identify possible counterparts in various gamma-ray and multiwavelength catalogs, as summarized in the accompanying catalog list.

Each candidate source is assigned a unique identifier of the form SIM\_# JHHMM.m+DDMMaa, where SIM denotes a SIMEFIC-identified source and # refers to the denoising stage at which the source was last detected before potential removal in later iterations. The source name encodes its J2000 coordinates using a standard convention.

The full catalog (Gamma\_SimCAT\_v1.0), providing a comprehensive summary of all remaining clusters (i.e., new source candidates not associated with 4FGL-DR4 or any other known catalog), is available at the CDS (see Data availability section). For each cluster, it lists the centroid coordinates, the number of constituent photons ( $N_\gamma$ ), the cluster significance ( $S$ ), the associated region index, the mean photon energy ( $\langle E \rangle$ ), the cluster classification type (e.g., Near Fermi, Bridged, Satellite, or None), and, when applicable, the angular separation from the nearest Fermi catalog source. Although the parameters  $N_\gamma$ ,  $S$ , and  $\langle E \rangle$  are not analyzed in detail within this work, they are included for completeness and to facilitate future studies. These quantities provide additional context for interpreting cluster properties and may help in subsequent analyses, such as assessing spectral characteristics, variability, or cross-identification with multiwavelength observations.

### 5.3. Special clusters

#### 5.3.1. Bridged clusters

In regions of elevated and spatially variable photon density—particularly near the Galactic plane-adjacent sources—were sometimes not resolved as separate detections due to their proximity. As a result, multiple nearby sources were occasionally merged into a single extended cluster. Such cases, referred to as bridged clusters (Campana & Massaro 2021), occur when overlapping photon distributions prevent clean separation by density-based clustering. An example is shown in Figure 2.

To identify bridged clusters systematically, we examined a diagnostic region with a radius twice the cluster’s characteristic size, centered on its centroid. If two or more 4FGL-DR4 sources

**Table 4.** Mean photon count ( $\langle N_\gamma \rangle$ ) and detection significance ( $\langle S \rangle$ ) for the four cluster categories.

Category	Near-Fermi	Bridged	Satellite	New candidates
$\langle N_\gamma \rangle$	30.1	542.4	30.9	15.9
$\langle S \rangle$	4.0	22.0	3.1	2.5

**Notes.** Bridged clusters exhibit higher values in both metrics, while satellite and near-Fermi sources show overlapping properties. New candidates are the faintest on average, highlighting the need for additional discriminative features.

were located within this region, the cluster was flagged as a bridge (bridge). In some cases, blended structures can be disentangled by reapplying the denoising and clustering algorithms in a smaller ( $10^\circ \times 10^\circ$ ) field with adjusted parameters.

### 5.3.2. Satellite clusters

Satellite clusters, first described by Campana et al. (2018), arise from the fragmentation of extended sources into multiple smaller clusters. These fragments are typically artifacts of the clustering process rather than genuine point sources. An example is shown in Figure 2.

Following the approach of Campana & Massaro (2021), we flagged pairs of clusters as satellites (sat) if their angular separation was less than the sum of their radii, increased by 10%. As with bridged clusters, a refined analysis in a smaller region with adjusted parameters often helps to recover the true underlying structure.

### 5.3.3. Near 4FGL-DR4 clusters

Clusters classified as Near 4FGL-DR4 (designated as NF, short for “Near Fermi” in the CDS catalog) lie just beyond the  $0.15^\circ$  matching radius from a known 4FGL-DR4 source but remain close enough that a possible association cannot be ruled out. Some of these clusters may represent the same underlying astrophysical source; however, since their separation exceeds the adopted threshold, definitive classification requires case-by-case evaluation through detailed spatial and spectral analysis.

### 5.4. Comparative properties of special cluster categories

We compared the average photon count and detection significance across four cluster categories: bridged clusters, satellite clusters, near-Fermi clusters, and new candidates, with results summarized in Table 4. Bridged clusters stand out with markedly higher values in both metrics, consistent with their interpretation as blends of multiple nearby sources that were not cleanly separated during clustering. This distinct profile justifies treating them as a separate class in subsequent source detection and modeling efforts.

By contrast, satellite and near-Fermi clusters show similar averages, suggesting that photon count and significance alone are insufficient to reliably distinguish between them. New candidate sources exhibit the lowest averages, appearing as fainter and less prominent clusters, which may indicate either intrinsically weak sources or limited photon statistics. These trends underscore the need for additional discriminative features in future classification schemes and for dedicated follow-up analyses to confirm the astrophysical nature of the new candidates.

## 6. Conclusion

In this study, we have presented a comprehensive catalog of gamma-ray source candidates derived from 16 years of Fermi-LAT data, focusing on the sky regions with  $|b| > 10^\circ$ . By applying the SIMEFIC III denoising algorithm combined with the DBSCAN clustering method, we identified a total of 3112 candidate sources, significantly advancing our understanding of the high-energy gamma-ray sky. Among these, 370 candidates are previously unreported detections with no counterparts in existing catalogs, representing a rich set of potential astrophysical objects for further investigation.

We successfully associated 2822 of the detected sources with known entries in the 4FGL-DR4 catalog and established additional matches with other multiwavelength catalogs. These associations validate our methodology and demonstrate the effectiveness of our approach in revealing significant gamma-ray sources. Our results highlight the value of adaptive data analysis techniques, particularly in regions where a high source density complicates traditional detection methods.

This work not only contributes to ongoing cataloging efforts in high-energy astrophysics but also emphasizes the importance of follow-up observations of the newly identified candidates. These sources offer exciting opportunities for future research, especially in regions previously considered challenging due to strong gamma-ray emissions.

### Data availability

The catalog table associated with this article is available at the CDS via <https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/707/A61>.

*Acknowledgements.* The authors thank the anonymous referee for constructive comments that helped improve the clarity and quality of this manuscript.

## References

- Abdollahi, S., Acero, F., Baldini, L., et al. 2022, *ApJS*, 260, 53  
 Ackermann, M., Ajello, M., Albert, A., et al. 2012, *ApJS*, 203, 4  
 Arsioli, B., Chang, Y.-L., & Musiimenta, B. 2020, *MNRAS*, 493, 2438  
 Arsioli, B., Chang, Y.-L., & Ighina, L. 2025, *MNRAS*, 539, 1458  
 Atwood, W. B., et al. 2009, *ApJ*, 697, 1071  
 Ballet, J., Bruel, P., Burnett, T. H., Lott, B., & Collaboration, T. F.-L. 2024, *Fermi Large Area Telescope Fourth Source Catalog Data Release 4 (4FGL-DR4)*  
 Campana, R., & Massaro, E. 2021, *A&A*, 652, A6  
 Campana, R., Massaro, E., Gasparrini, D., et al. 2008, *MNRAS*, 383, 1166  
 Campana, R., Massaro, E., & Bernieri, E. 2018, *A&A*, 619, A23  
 Campana, R., Massaro, E., Bocchino, F., et al. 2022, *MNRAS*, 515, 1676  
 Chang, Y.-L., Arsioli, B., Giommi, P., & Padovani, P. 2017, *A&A*, 598, A17  
 Chang, Y.-L., Arsioli, B., Giommi, P., et al. 2019, *A&A*, 632, A77  
 D’Abrusco, R., Massaro, F., Paggi, A., et al. 2014, *ApJS*, 215, 14  
 Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. 1996, *KDD*, 96, 226  
 Healey, S. E., Romani, R. W., Taylor, G. B., et al. 2007, *ApJS*, 171, 61  
 Hedayati Kh., H., Soor, M., & Akhondi, F. 2024, *ApJ*, 968, 37  
 Kanbach, G., et al. 1989, *Space Sci. Rev.*, 49, 69  
 Li, T., & Ma, Y. 1983, *ApJ*, 272, 317  
 Massaro, E., Maselli, A., Leto, C., et al. 2015, *Ap&SS*, 357, 1  
 Mattox, J. R., et al. 1996, *ApJ*, 461, 396  
 SLAC National Accelerator Laboratory. n.d., Fermi-LAT Performance, [https://www.slac.stanford.edu/exp/glast/groups/canda/lat\\_Performance.htm](https://www.slac.stanford.edu/exp/glast/groups/canda/lat_Performance.htm), accessed: 2025-08-16  
 Soor, M., Akhondi, F., & Hedayati Kh., H. 2025, *MNRAS*, 537, 730  
 Starck, J.-L., Fadili, J. M., Digel, S., et al. 2009, *A&A*, 504, 641  
 Tramacere, A., & Vecchio, C. 2013, *A&A*, 549, A138  
 Wakely, S. P., & Horan, D. 2008, *ICRC*, 3, 1341