The upcoming 4th Fermi-LAT source catalog

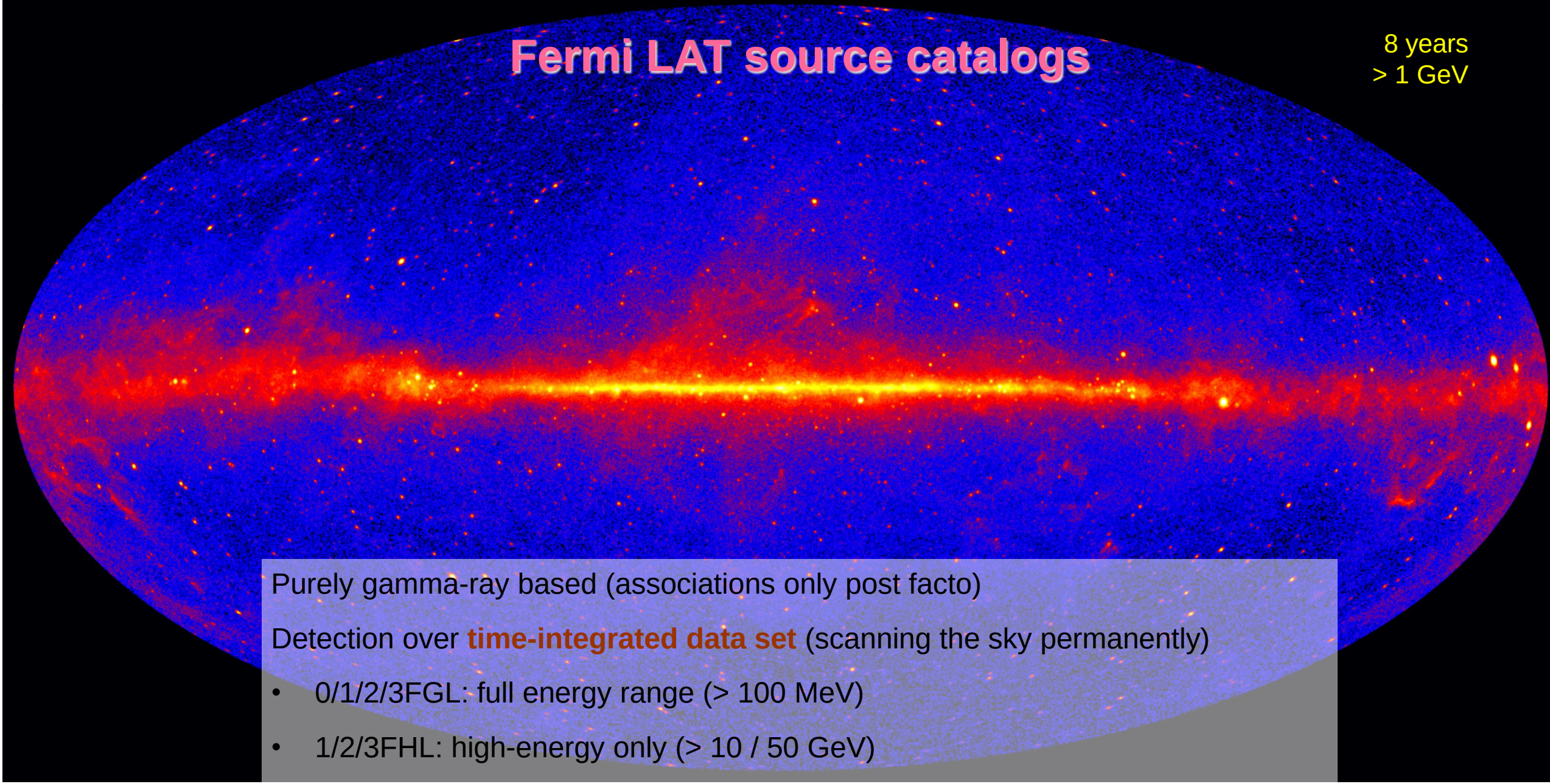
Jean Ballet
AIM, CEA Saclay, France

Toby Burnett, Seth Digel, Benoit Lott
for the LAT collaboration

Very High Energy Phenomena in the Universe
August 17, 2018

Fermi LAT source catalogs

8 years
> 1 GeV



Purely gamma-ray based (associations only post facto)

Detection over **time-integrated data set** (scanning the sky permanently)

- 0/1/2/3FGL: full energy range (> 100 MeV)
- 1/2/3FHL: high-energy only (> 10 / 50 GeV)

Each generation has used **improved data/calibration**: P6 → P7 → P7Rep → P8

Effects of diffuse emission

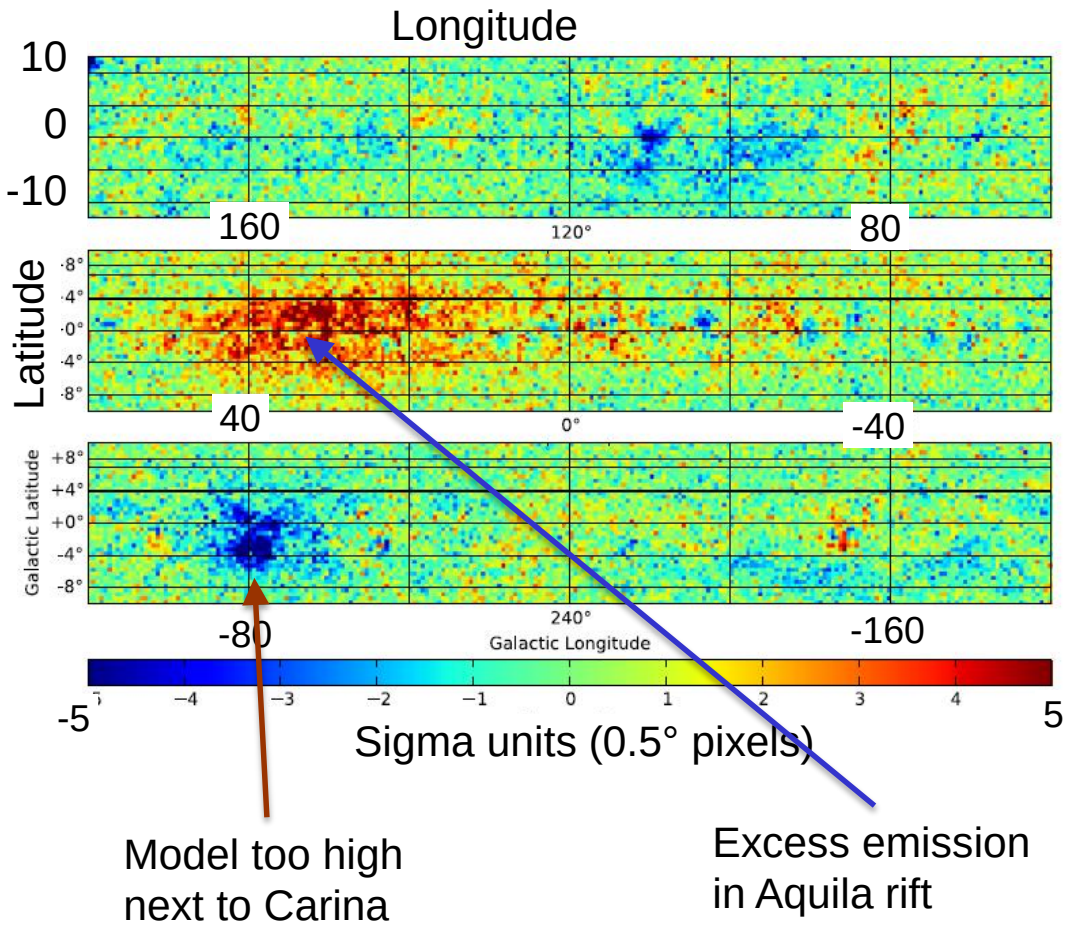
Interstellar emission model (Acero et al 2016, ApJS **223**, 26) **cannot be perfect** (we don't know perfectly the distribution of gas, photons and cosmic rays)

Residuals at level of 3%, both spatially (at all scales) and spectrally

In 3FGL, this impacts sources at same level as statistical errors over the whole Galactic plane

Dominate in Galactic ridge

3FGL (Acero et al 2015, ApJS **218**, 23)



Diffuse emission modeling improvements

follows Acero et al 2016, ApJS **223**, 26

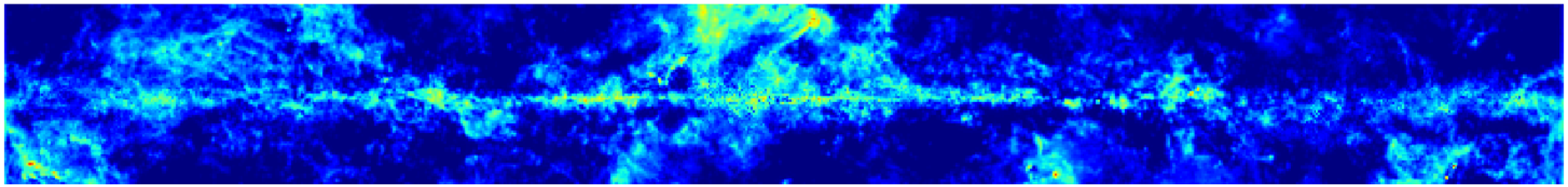
- Refined decomposition of CO (H_2 tracer) and H I into ‘rings’ of Galactocentric distance, including factoring the CMZ from the innermost ring
- Better angular resolution for H I with the new **HI4PI** survey
- Incorporation of **Planck microwave data**, to derive the dark gas component not traced by H I or CO
- Increased freedom for tuning IC model via decomposition into ‘rings’
- Evaluated three models for Loop I
- Re-extracted the Fermi Bubbles
- Tested for a Galactic disk population of unresolved sources

Local gas templates

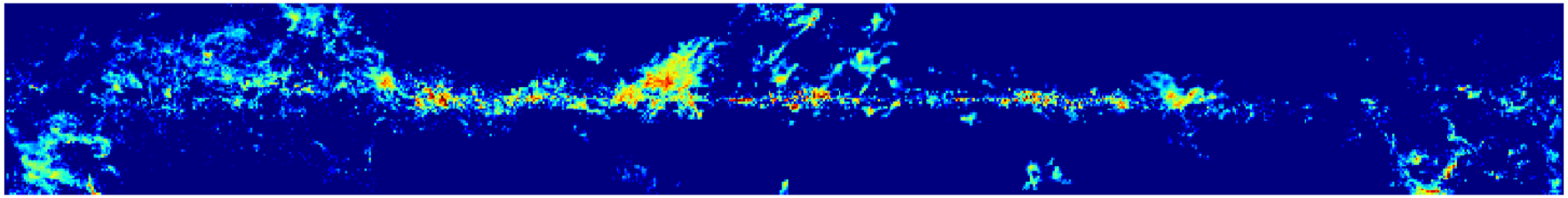
The dark gas is responsible for a lot of the small-scale structure in the ISM (and γ-ray sky)

Artificial structure around massive star-forming regions is greatly improved

Local DNM

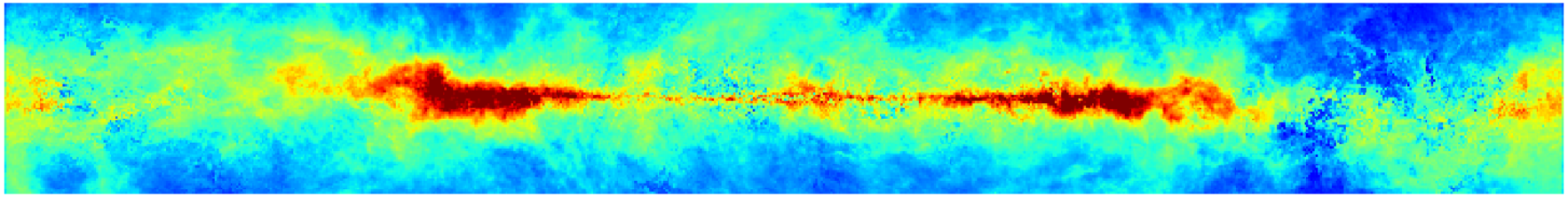


Local H2 (XCO = 1)



Maps for |b| < 20 deg, same (sqrt) scaling

Local H I



Source confusion

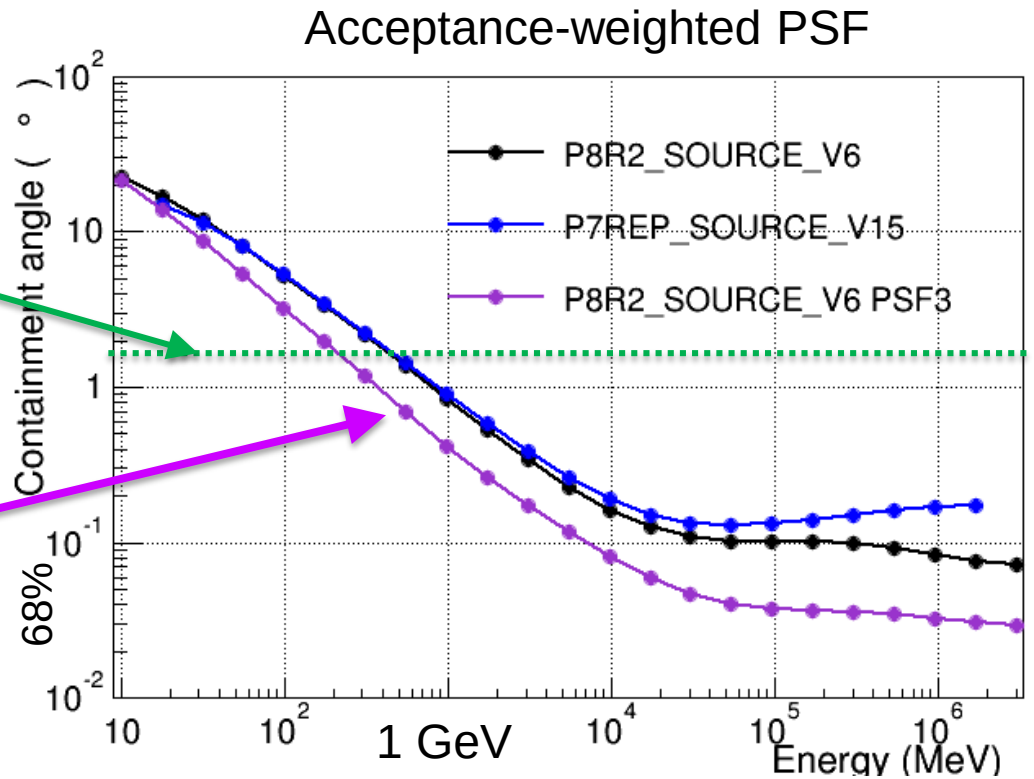
Strong **confusion** at low energy

Average source separation 1.66˚ outside Gal. plane (in recent FL8Y)

Equals R_{68} at 400 MeV

Equals R_{95} at 2 GeV

Somewhat better when selecting 25% best (PSF3) events (200 and 550 MeV respectively)



Each source is correlated with entire sky at some point

Requires **iteration** over Regions of Interest paving the sky

Source confusion

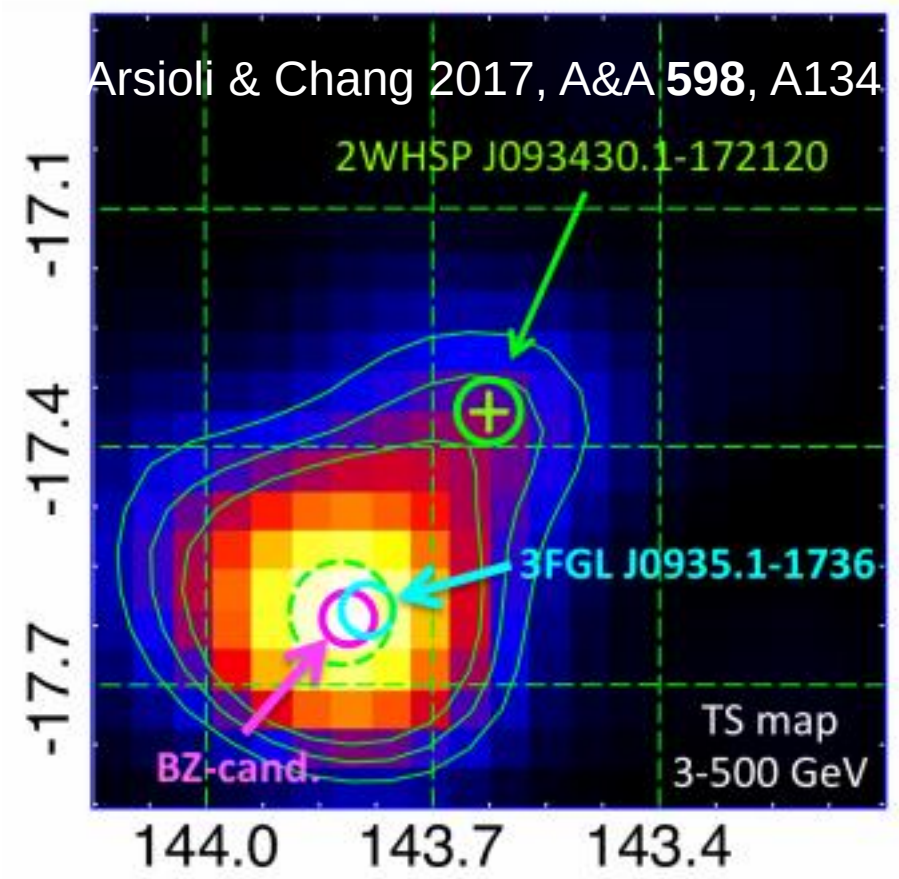
Blazars:

Test Statistic map at right elongated

Probable confusion between two hard blazars ($\Gamma \sim 1.9$) 0.3° apart

The fainter one was subthreshold in 3FGL, now **separate source in FL8Y**

Also frequent confusion between soft bright FSRQ and hard fainter BL Lac whose high-energy photons (best PSF) drive the localization

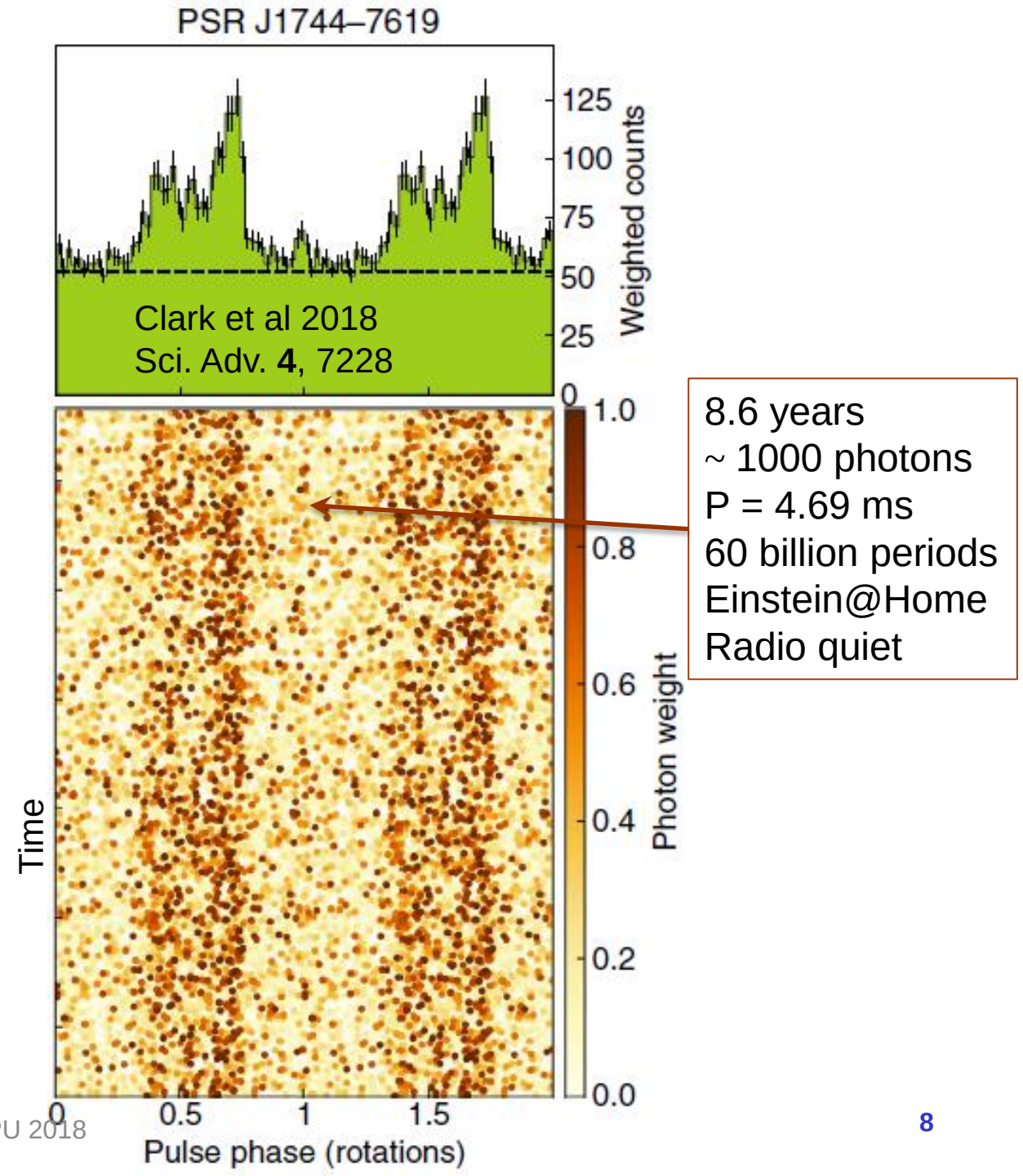


Follow-up of Galactic UNIDs

Pulsars (> 216 in all):
Search for ms/s rotation period
Radio searches are easiest when the pulsar is radio-loud
Otherwise pure γ timing. MSPs are most difficult
> 20 PSR found on top of 3FGL UNIDs

Binaries:
Search for hour/day orbital period
1 more since 3FGL (in LMC)

Clusters of UNIDs in Galactic plane → **extended** sources; 25 in 3FGL, 58 now



Follow-up of extragalactic UNIDs

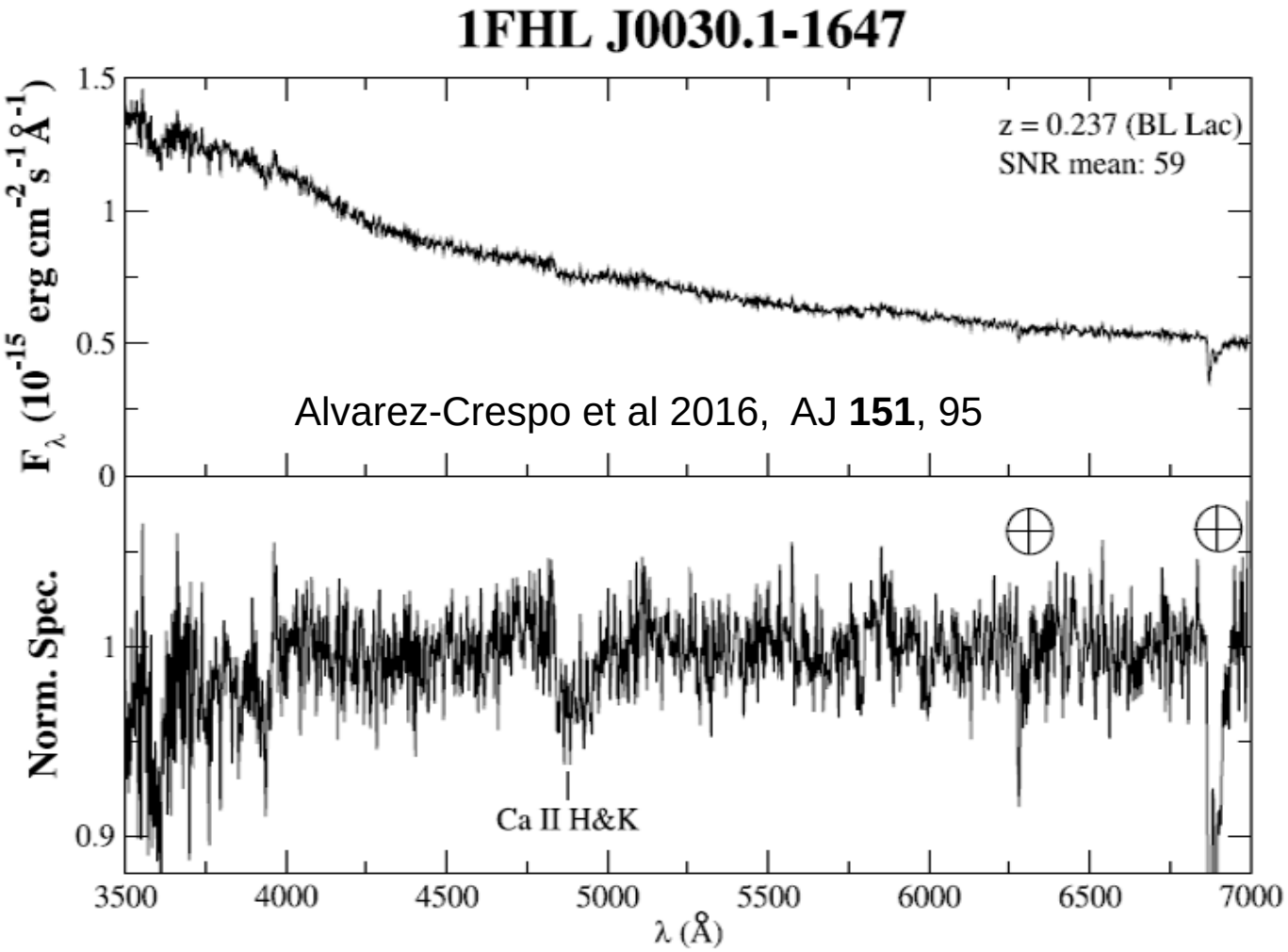
AGN: > 50 BL Lacs & FSRQs found on 3FGL UNIDs

X-rays (Swift) can help localize the source and facilitate optical follow-up

NIR (WISE) colors can also be used to single out blazars

573 **unknown blazars** (BCU) in 3FGL (1/3 of all blazars)

Spectroscopy: getting redshifts of BL Lacs is most difficult



Developments for 4FGL

1. 8 years of Pass 8 data
2. Improved interstellar emission model (in test)
3. Weighted logLikelihood
4. Earth limb suppression
5. Energy dispersion

FL8Y list (~ 5500 sources) made public January 3

- **Will be entirely superseded** by 4FGL with new interstellar emission model (source positions will be recalculated, hence names will all change)
- OK as starting point or background model, but **should not be used directly** (for example for population studies). Cannot be cited.

3FGL vs FL8Y

3FGL		FL8Y
4 years P7Rep	**Data**	8 years **P8**, TS x 2.3 (acceptance)
Front/Back, z < 100°	**Selection**	PSF types, zmax depend on energy
100 MeV – 300 GeV	**Main fit**	100 MeV – 1 TeV
No weights or energy dispersion	**Method**	**Weights**, energy dispersion
gll_iem_v06	**Interstellar**	Idem (**will be updated in 4FGL**)
25	**Extended sources**	58
Cutoff as $\exp\left(-E/E_{\rm cut}\right)$	**Pulsars**	Cutoff as $\exp\left(-a\,E^{2/3}\right)$
Used for PL, PLEC, LP	**Spectral_Index**	PL_Index, LP_Index, PLEC_Index
beta, Exp_Index	**Spectral params**	LP_beta, PLEC_Exp_Index
Test either LP or PLEC	**Spectral shapes**	Always report PL, LP, PLEC params

FL8Y detection and localization

Detection uses TS maps assuming **several spectral shapes**: three power-law (Γ = 1.7, 2.1, 2.4) and one pulsar-like (PLEC Γ = 1.7, E_{cut} = 3 GeV)

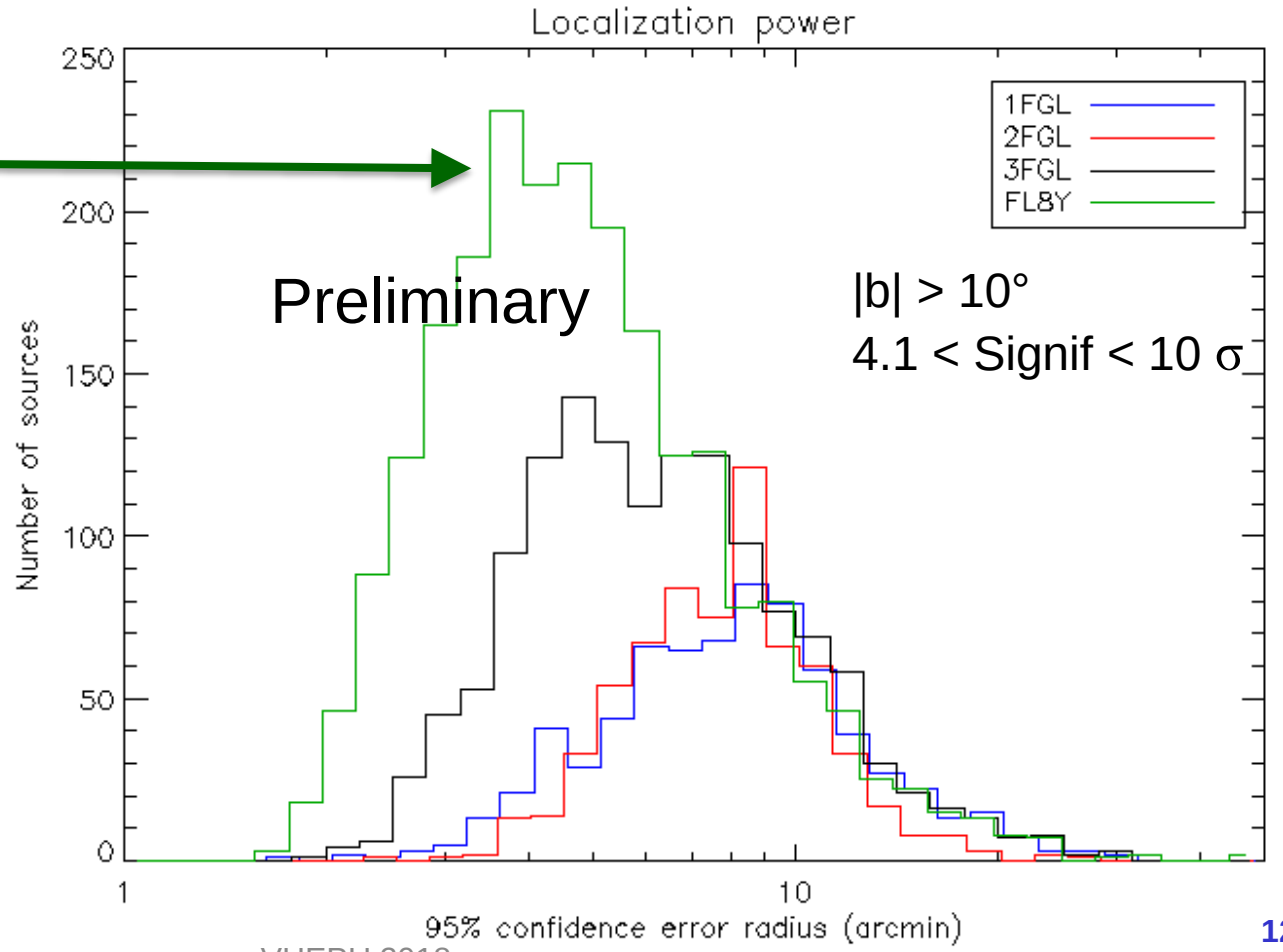
> 13,000 seeds at TS > 10

Localization of faint sources (critical for associations) continues to improve

Median error radius **4.5 arcmin** at 25 < TS < 100

Systematic factor **1.05** on error radius (as in 3FGL)

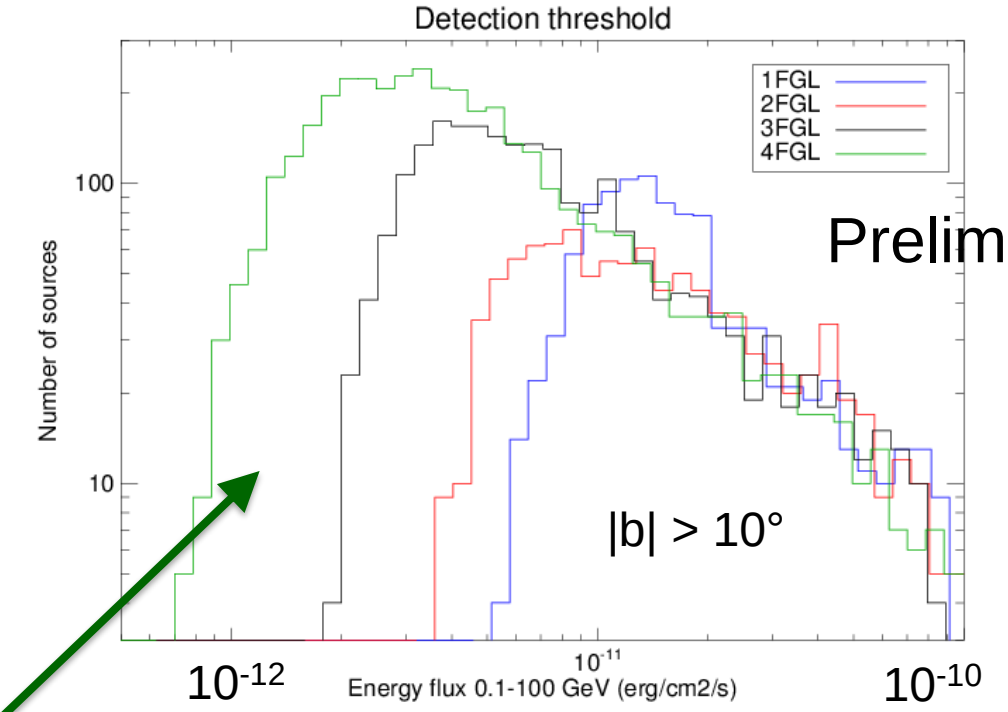
Absolute 95% systematic error: **27 arcsec** (as in 3FHL)



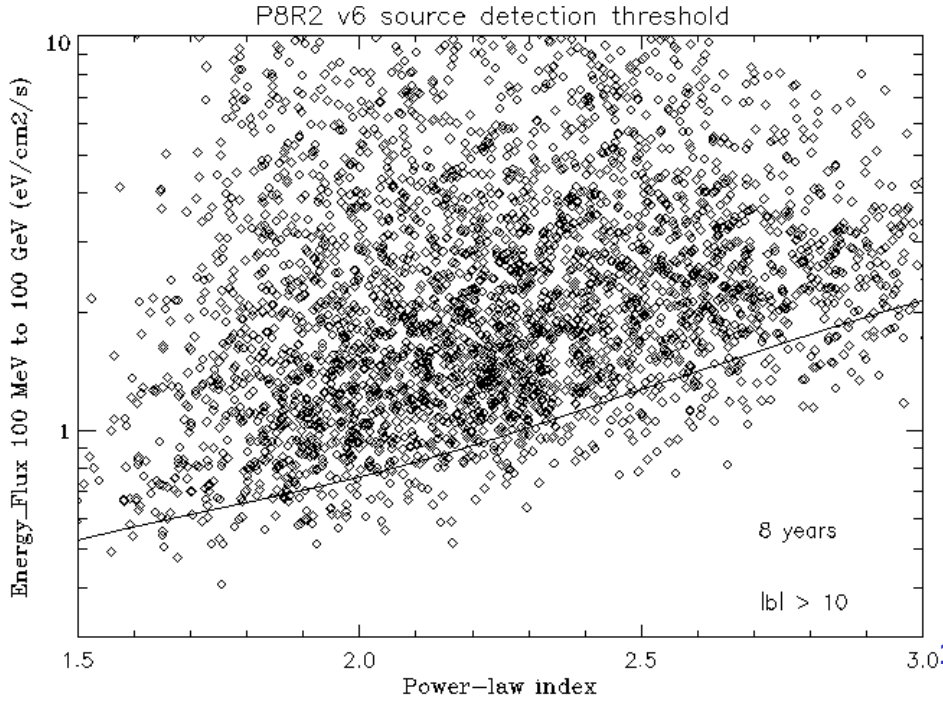
FL8Y characterization

- ✓ Current (**gll_iem_v06**) diffuse model
- ✓ Data-based weights
- ✓ Energy dispersion
- ✓ Fully binned analysis (faster)

- 5,524 sources at TS > 25
- Extragalactic detection threshold around 2 10^{-12} erg/cm^2/s (~ **1 eV/cm^2/s**) in 100 MeV to 100 GeV band
- Depends slightly on spectral shape



Preliminary



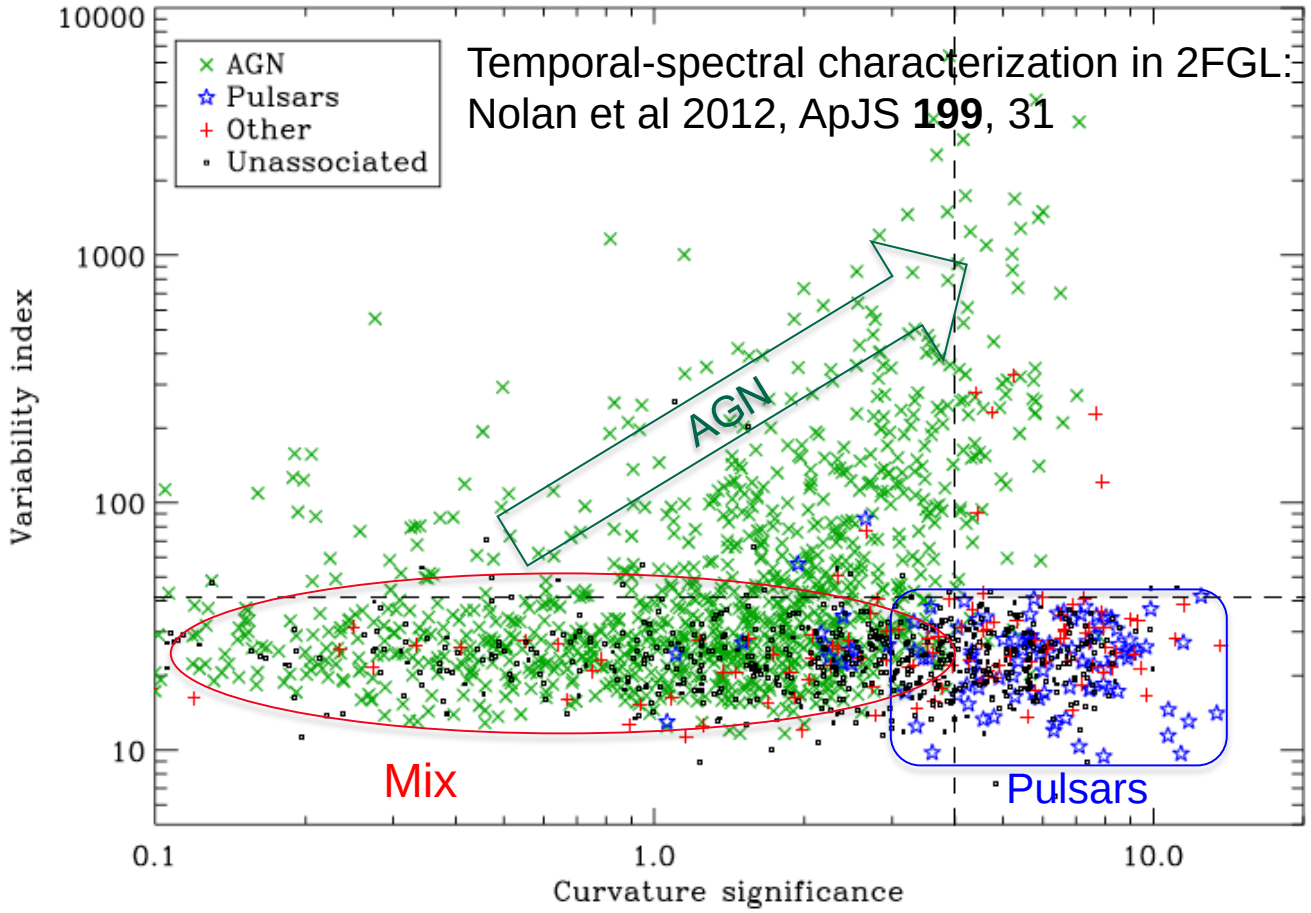
Statistical assessment of UNIDs

Concentrate MW efforts on what is most promising

Use all γ–ray information (particularly variability, spectral curvature) to feed **classification algorithms**, learning from known associations

Logistic regression, classification trees, random forest

Difficulty: Training sample (brighter) has smaller error bars than most UNIDs



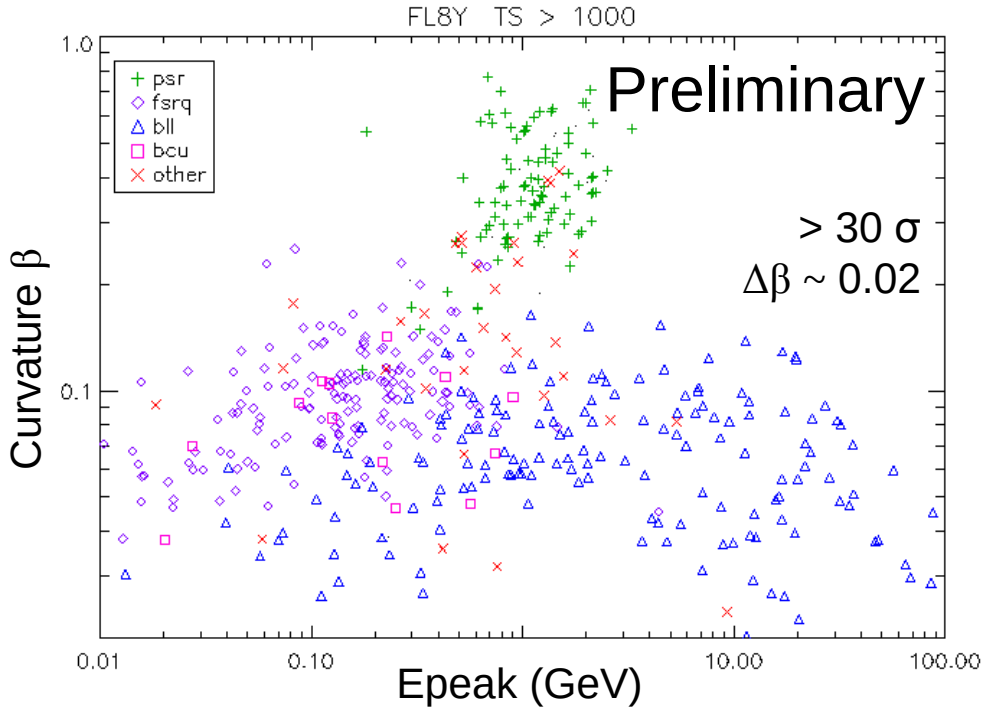
Particularly used to single out **PSR candidates** (minority) against **AGN**

eg Saz-Parkinson et al 2016 (ApJ **820**, 8): rather surprising conclusion that hundreds of unassociated 3FGL sources might be pulsars

Classification from LogParabola parameters

Bright sources: clear separation

- 88 pulsars **(psr)**
- 161 flat-spectrum radio quasars **(fsrq)**
- 172 BL Lacs **(bll)**
- 14 blazars of unknown type **(bcu)**
- 30 **other,** 17 unassoc (dots)



LogParabola: $\ln(\nu F\nu) = K - \beta \ln^2\left(\frac{\nu}{\nu peak}\right)$

Classification from LogParabola parameters

Bright sources: clear separation

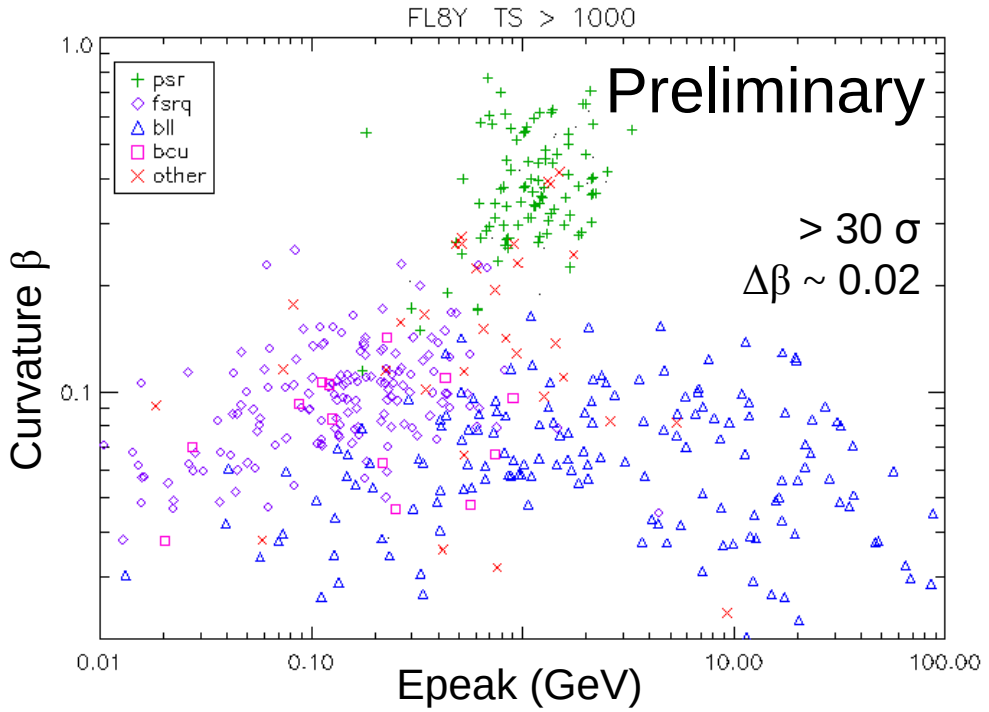
- 88 pulsars **(psr)**
- 161 flat-spectrum radio quasars **(fsrq)**
- 172 BL Lacs **(bll)**
- 14 blazars of unknown type **(bcu)**
- 30 **other,** 17 unassoc (dots)

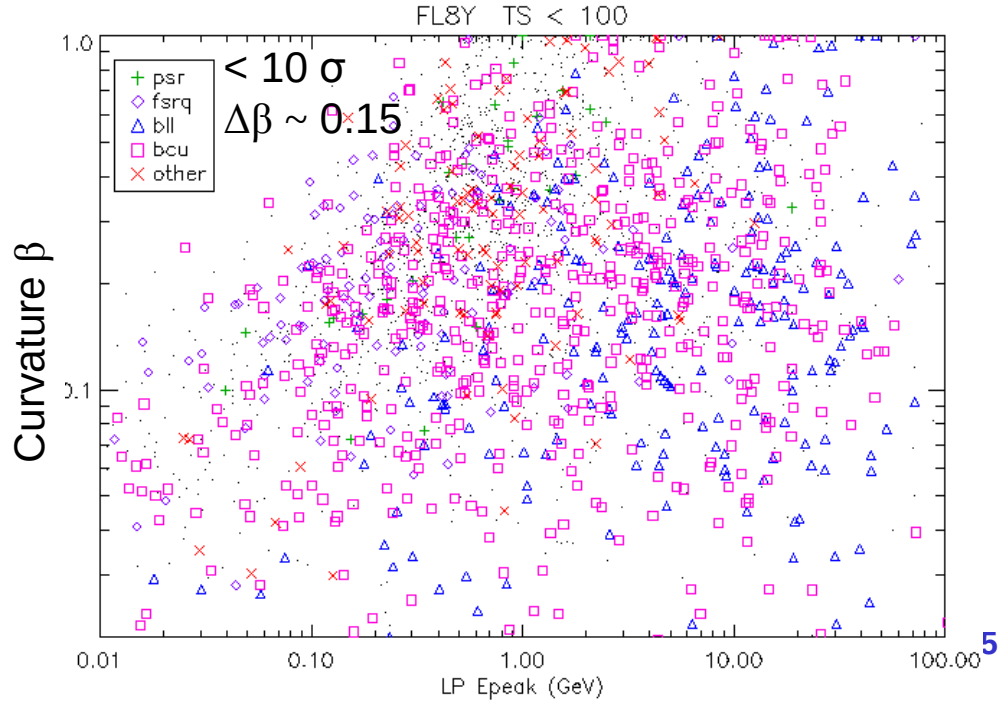
LogParabola: $\ln(\nu F\nu) = K - \beta \ln^2\left(\frac{\nu}{\nu peak}\right)$

But large error bars spoil the picture

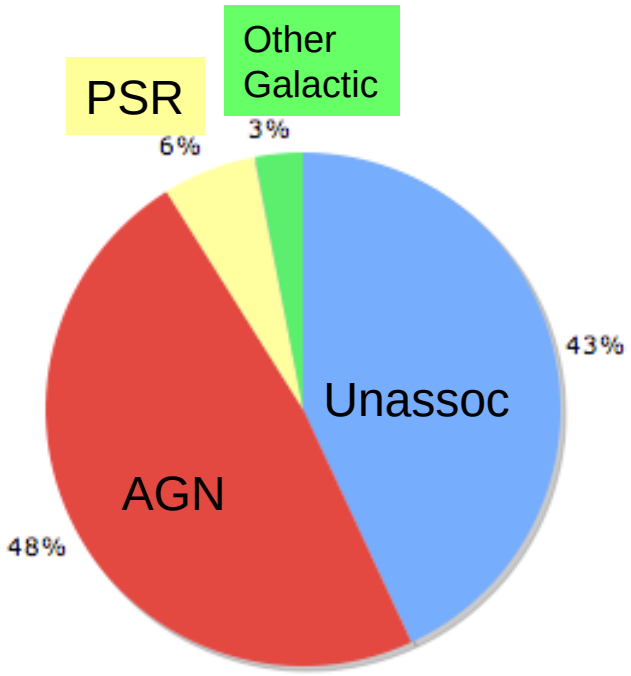
Faint sources: no clear separation

- 31 **psr,** 168 **fsrq,** 245 **BLL/bll**
- 641 **bcu,** 108 **other,** 1525 unassoc

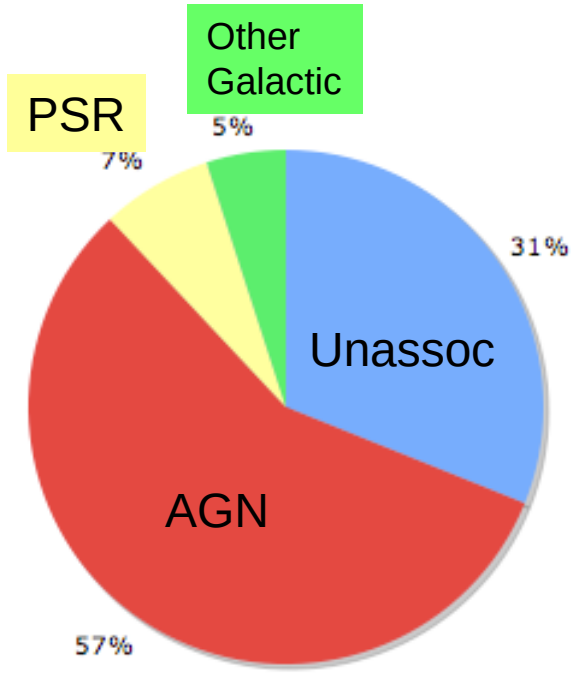




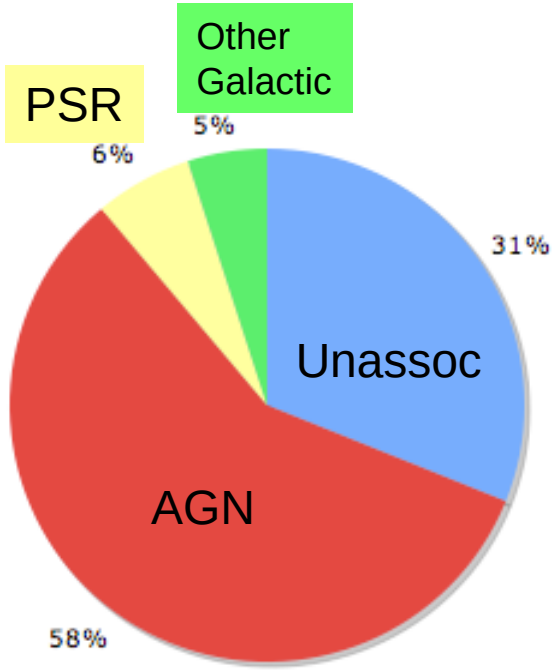
Source association



1FGL
(11 months, 1451 sources)



2FGL
(2 years, 1873 sources)

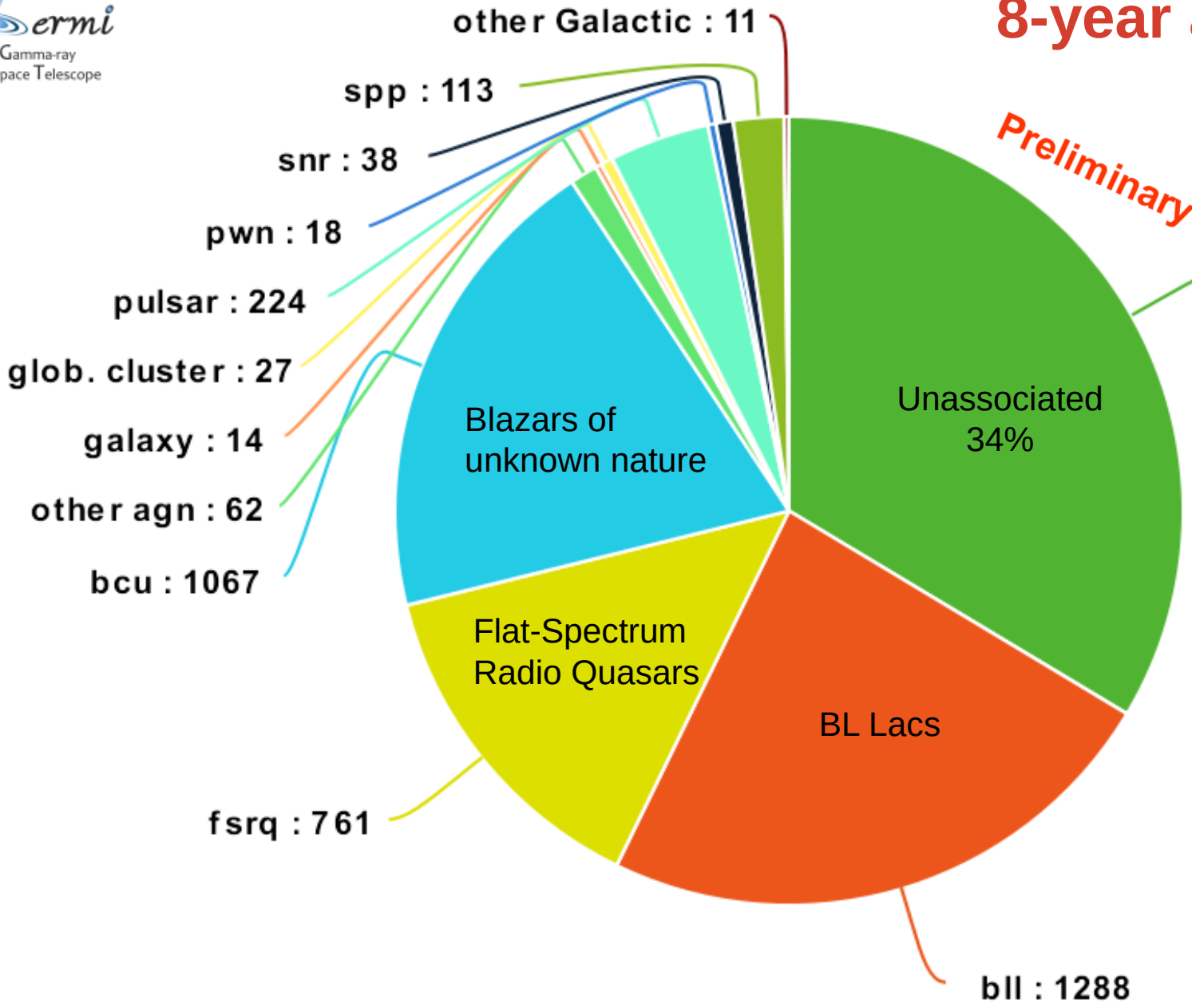


3FGL
(4 years, 3033 sources)

Numbers as in original papers (fewer unassociated now)

Similar fraction of associated sources in 3FGL as in 2FGL, thanks to ongoing effort on deepening counterpart catalogs

8-year association summary

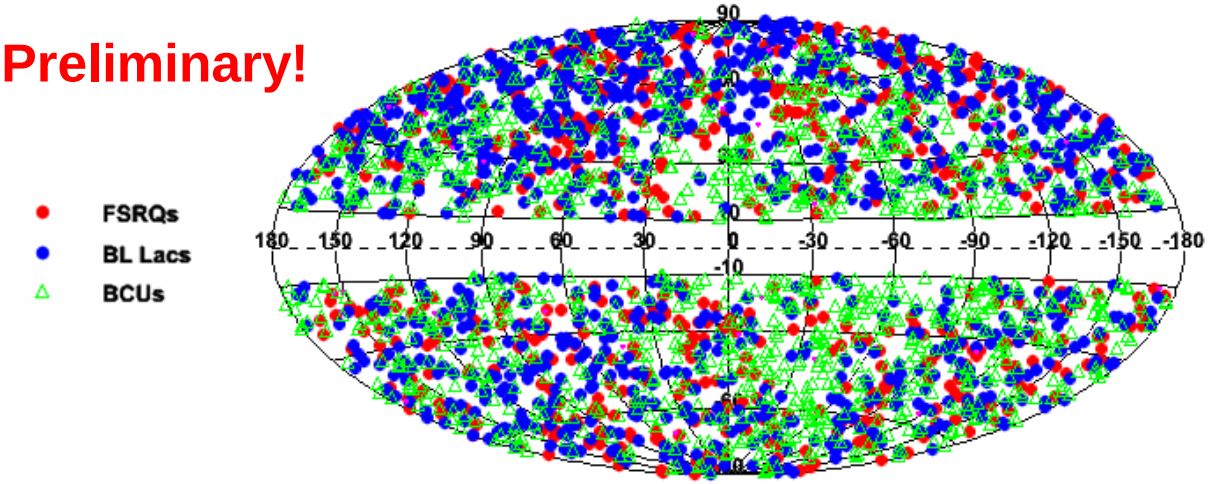


Cat8

Bayesian+Likelihood Ratio associations

4LAC AGN catalog

Preliminary!



Bayesian- +LR- method associations

2745 sources |b|>10° *(3LAC: 1591 sources)*

- **number of FSRQs: 713** ***(+53%)***
- **number of BLLs: 1199** ***(+90%)***
- **number of BCUs: 780** ***(+70%)***

3173 sources (all sky)

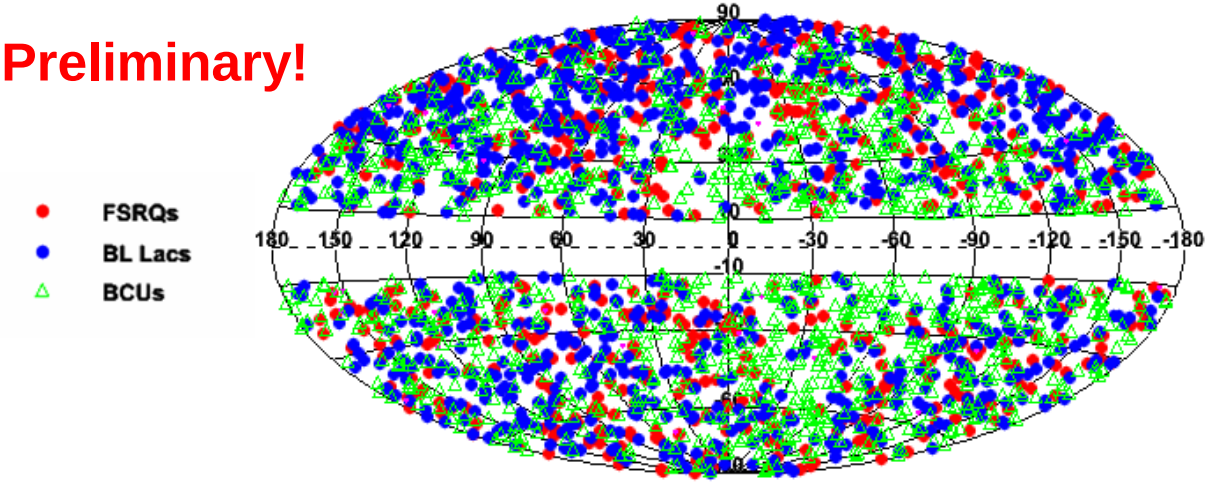
9 NLSy1 (4 new) , 36 RDG, 14 other AGNs

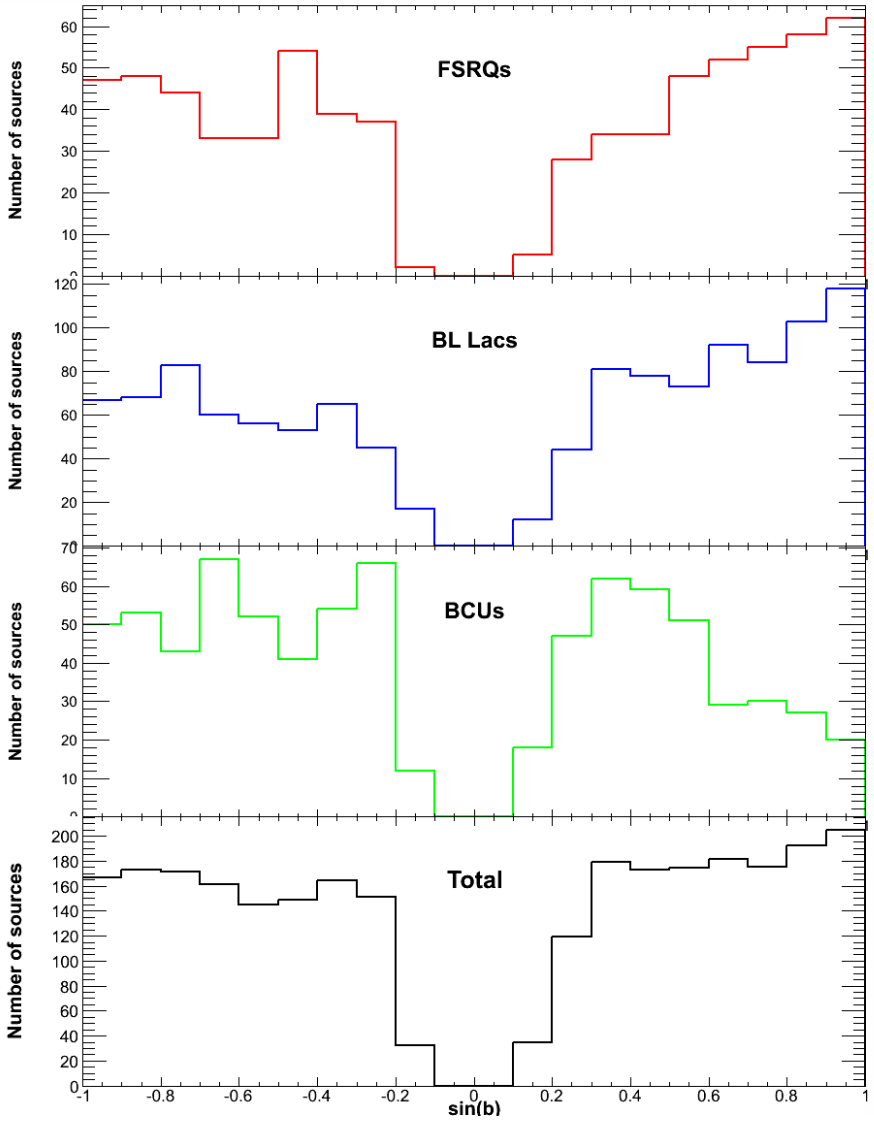
1422 sources in BZCAT

69 TeV sources

4LAC AGN catalog

Preliminary!





Bayesian- +LR- method associations
2745 sources |b|>10° *(3LAC: 1591 sources)*
- **number of FSRQs:** **713** ***(+53%)***
- **number of BLLs:** **1199** ***(+90%)***
- **number of BCUs:** **780** ***(+70%)***

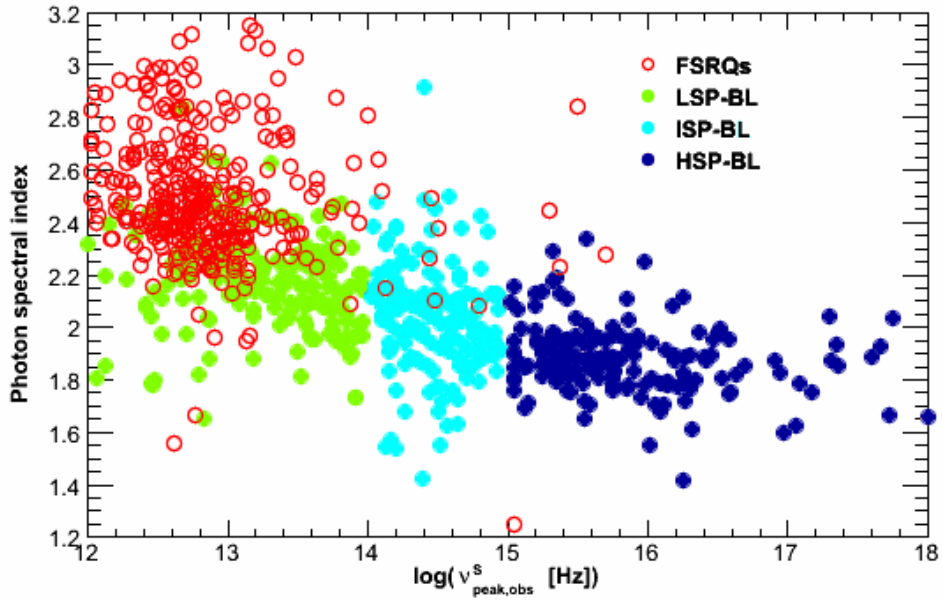
3173 sources (all sky)

9 NLSy1 (4 new) , 36 RDG, 14 other AGNs
1422 sources in BZCAT
69 TeV sources

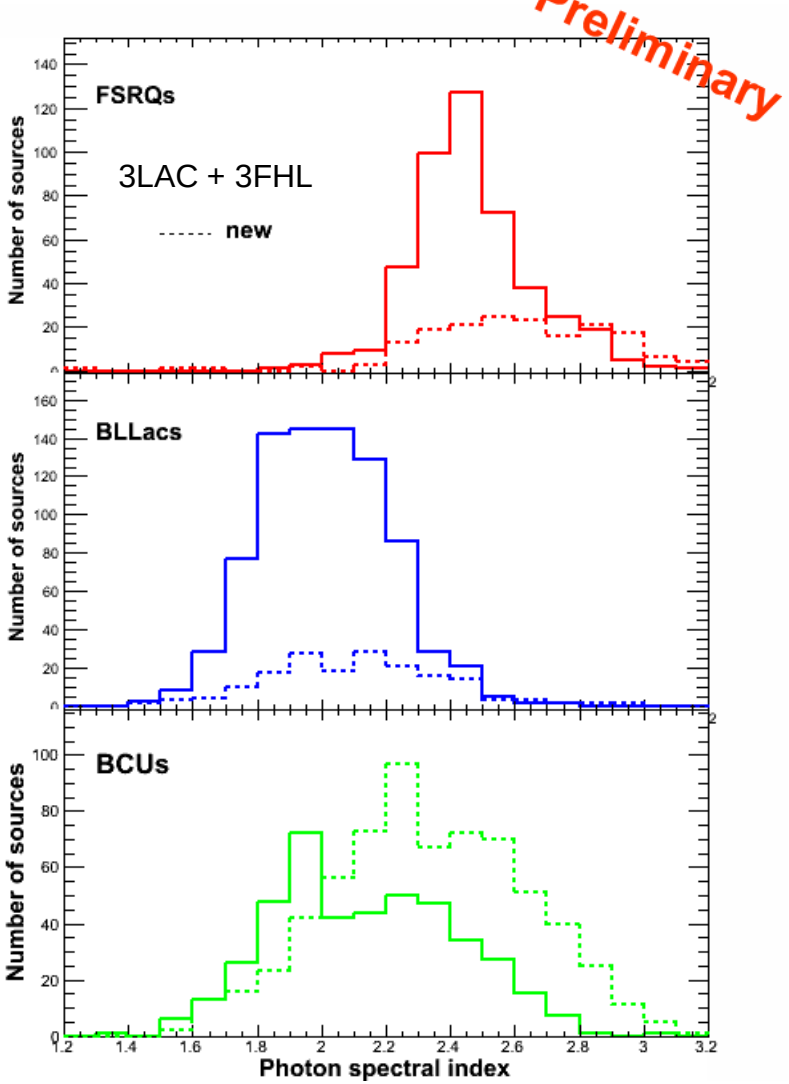
8% deficit of associations in the Southern hemisphere

Photon-index distributions

- Little overlap between FSRQs and BL Lacs, limit at Γ=2.2
- New FSRQs notably softer than 3LAC ones ($\langle\Gamma\rangle$=2.60 vs. 2.46)
- Slightly smaller effect for BL Lacs ($\langle\Gamma\rangle$=2.11 vs. 2.01)
- BCUs index distribution straddling the two classes and extending beyond 2.6. New BCUs softer than 3LAC ones ($\langle\Gamma\rangle$=2.33 vs. 2.14)



Preliminary

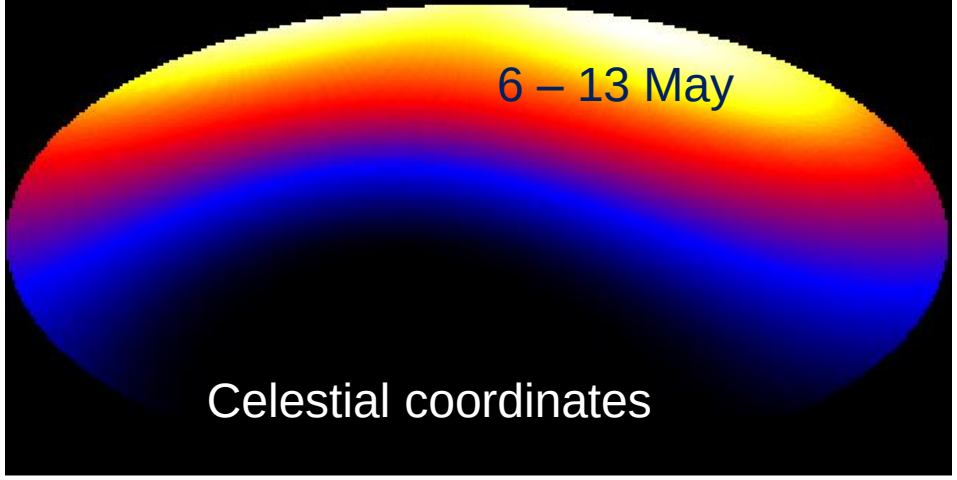


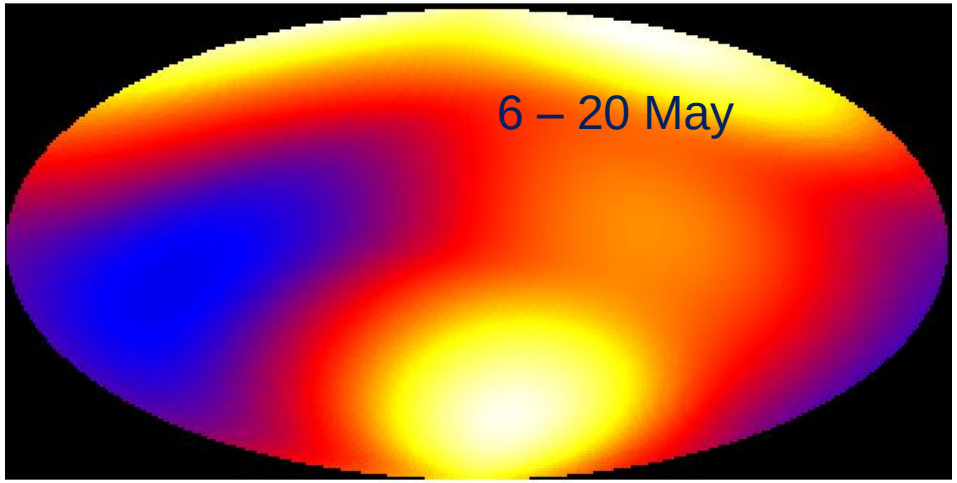
Current *Fermi* operations

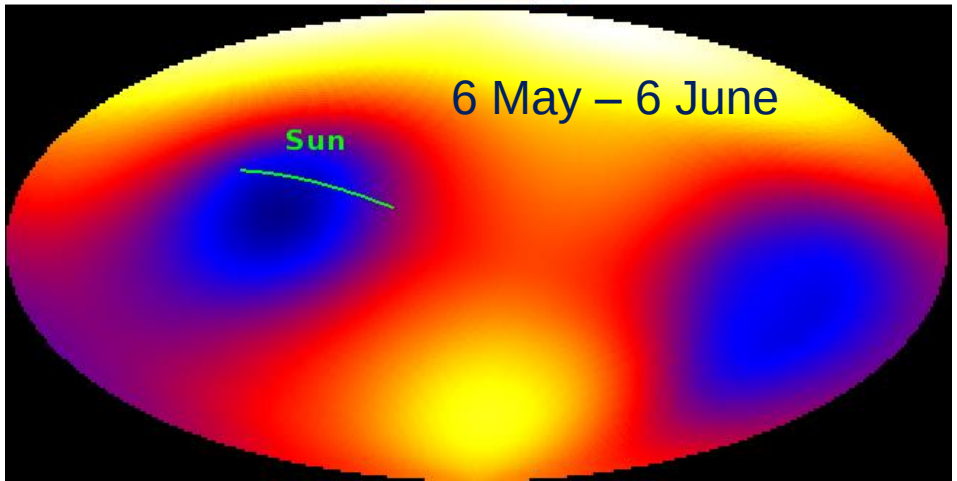
- On March 16, the instruments were automatically powered off after **one of the two solar panels got stuck** (solar panels normally orient toward the Sun as the spacecraft moves)
- The instruments were switched back on in early April, one week more to get back to normal operation temperature
- **Since April 8 the LAT operates normally again**, but remains most of the time at constant rocking angle with respect to zenith (whichever is best for solar panels), switching between +50 and -50° every week or so instead of every orbit
- The main consequence is that **the LAT does not survey the entire sky every 3 hours**, but only a little more than half every 1h30

Current *Fermi* operations

- The orbit's inclination is 25.6°
- The LAT can stay at the same rocking angle for one week (top)
- … but not for two weeks (center)
- Over one month (bottom), exposure minima (20 – 25% of maximum) correspond to the Sun and antiSun.
- Average over one year broadly similar to previous survey mode
- Working to improve sky coverage over daily time scale







Conclusions

- About **5,500 sources** in FL8Y and upcoming 4FGL
- 4FGL will make use of **new diffuse emission model**
- **Weighted logLikelihood** to account for systematics
- Localization better than **5 arcmin** for most sources
- Extragalatic detection threshold of **1 eV/cm^2/s**
- More spectral information, but use with care for faint sources
- Mostly blazars and pulsars, **34% unassociated**
- Need constantly improving all-sky counterpart catalogs to keep up with improving depth

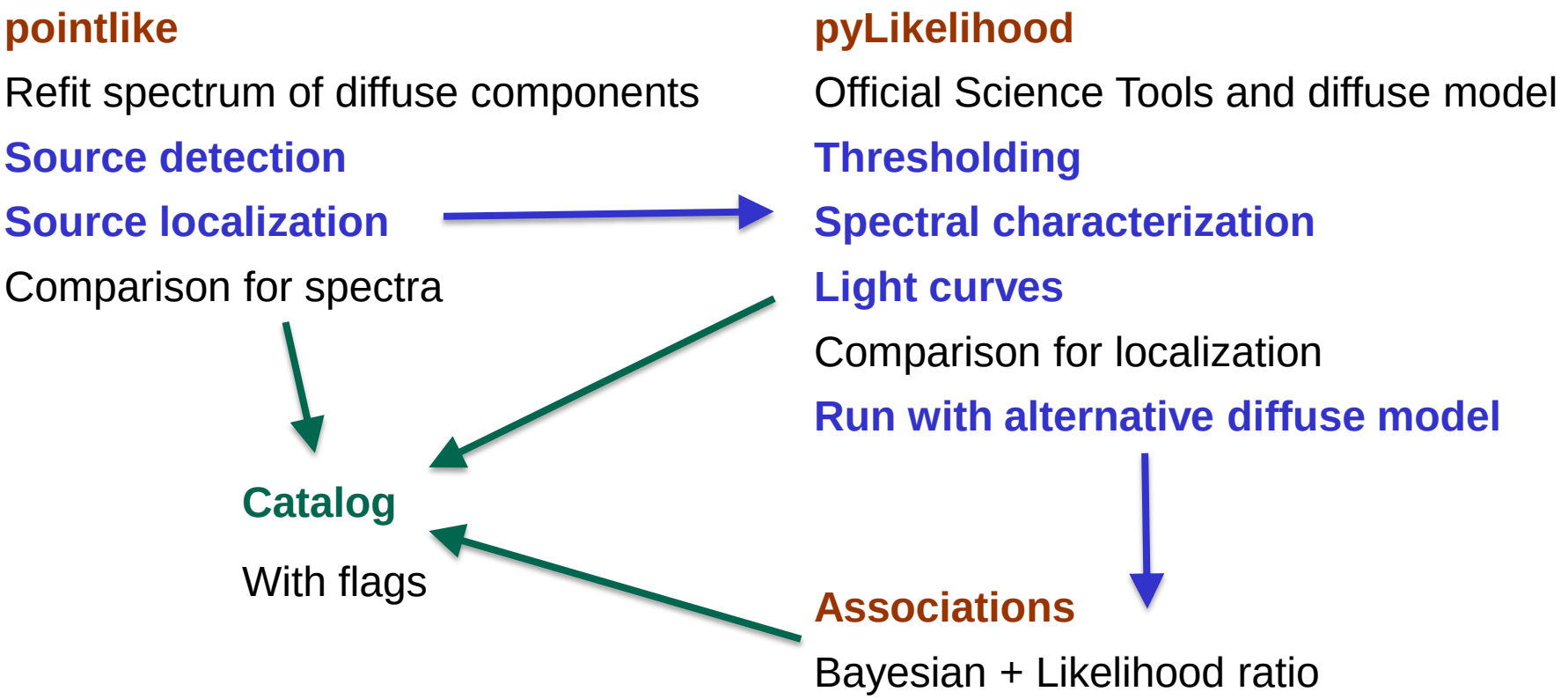
Backup slides

Methodology of the LAT source catalogs

3D maximum likelihood (x,y,E)

Point sources on top of isotropic, interstellar model and extended sources

Report position, significance, association, basic SED and light curve, flags



Weighted logLikelihood

The problem:

- Fermi-LAT data is dominated by imperfectly known diffuse emission
- Point-spread function 1° or worse below 1 GeV
- Large counts under the PSF ➔ systematics dominated at low energy

The proposed solution (J. Ballet at ICRC 2015, J. Ballet & T. Burnett at SCMA 2016)

Weighted logLikelihood: $\mathbf{wlog}L = \Sigma_i w_i (n_i \log M_i - M_i)$

w_i reduces the importance of systematics-dominated areas/energies

Reduce source significance and increase parameter uncertainties

The difficulty: How to define the weights in a proper way

Inspired from χ^2 approach

$w_i = \sigma_i^2 / (\sigma_i^2 + \varepsilon^2 B_i^2) = 1 / (1 + \varepsilon^2 B_i)$ where ε = 3 %

B_i is the relevant (source+background) Poisson counts ($\sigma_i^2 = B_i$)

Weighted logLikelihood 2

$w_i = 1 / (1 + \varepsilon^2 B_i)$ where ε = 3 %

Now how to define B_i ?

$$S(\mathbf{r}, E) = \frac{dB}{dE}(\mathbf{r}, E) \otimes \frac{P(\mathbf{r}, E)}{P(0, E)} \approx \frac{dB}{dE} \pi R_{68}^2(E)$$

Background intensity integrated under PSF (seen by a point source)

$$B_i = N(\boldsymbol{r}_i, E_i) = \int_{E_i}^{2E_i} S(\boldsymbol{r}_i, E) dE$$

Integral above current energy

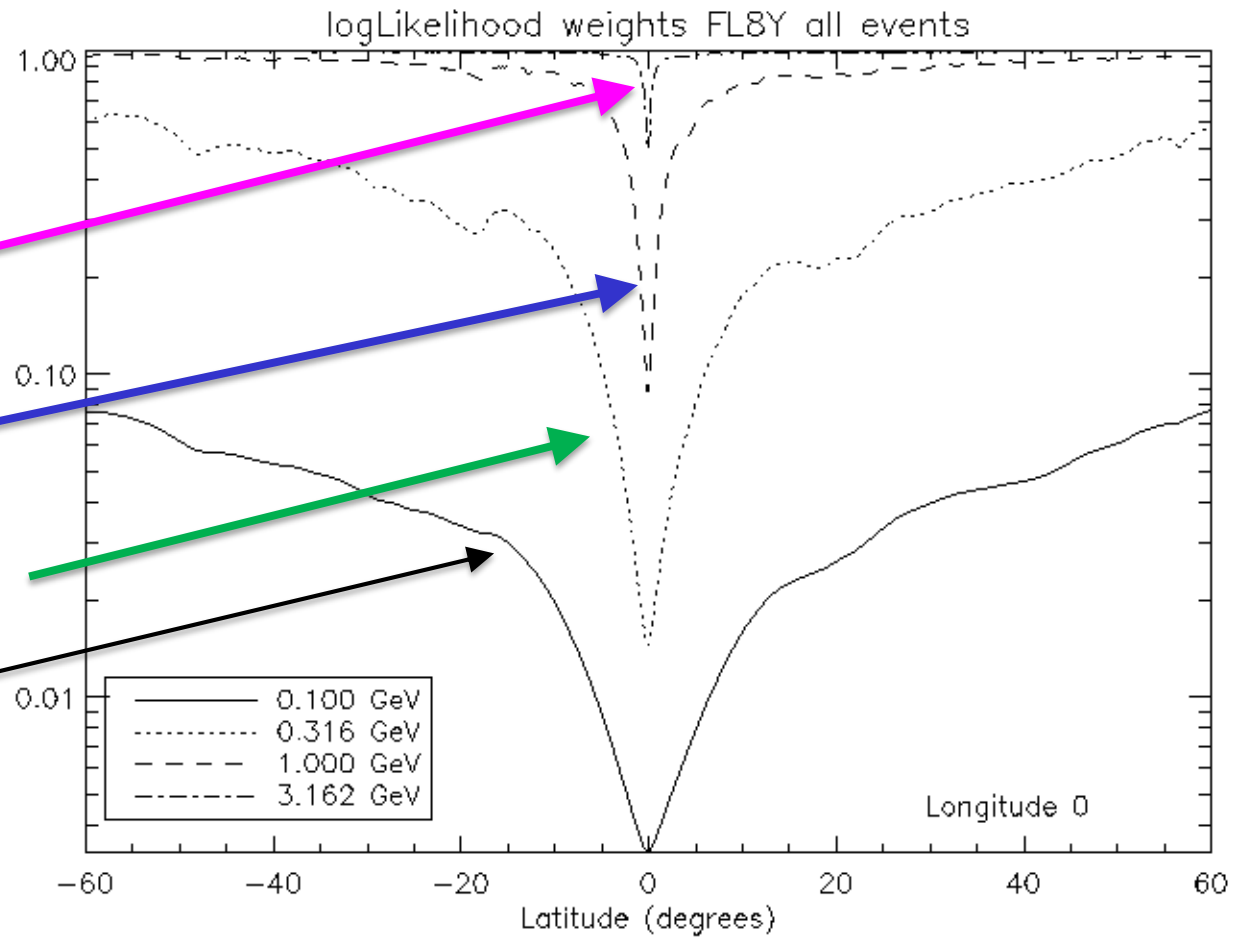
Ad-hoc but **desirable asymptotic limits**, stable against rebinning

$R_{68}(E)^2$ decreases as $E^{-1.6}$ up to 3 GeV so the B_i term decreases very fast

The **weights increase fast with energy**

Calculating the weights

- w_i = 1 everywhere above 10 GeV
- Small effect in the Galactic Ridge at **3 GeV**
- At **1 GeV**, small effect except in the Galactic Ridge
- Strong effect at **300 MeV**
- At 100 MeV, small weights over full sky ➔ useless to keep all data (can keep only best localized PSF3 events)



P8 all events, 8 years

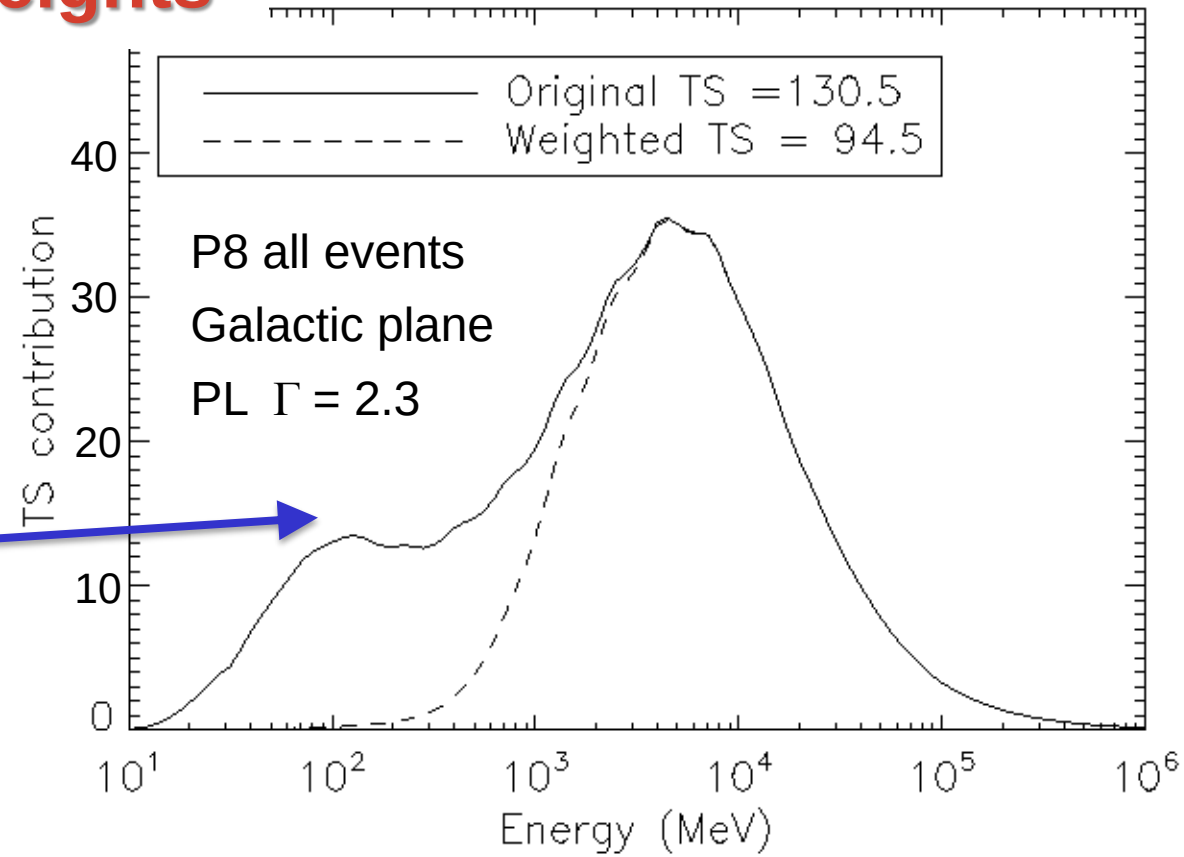
Data-based weights, ε = 3 %

Longitude 0 (through Gal center)

Effect of weights

Standard TS (= 2 ΔlnLike, ~ significance squared) estimate from integral over all energies

The effect of the weights is to **concentrate significance at high energy** where data is more reliable



Model-based weights or

- Background is **interstellar emission model** only
- Original motivation

Data-based weights

- Background is **all data**, common ε
- Fights imperfect modeling of bright sources

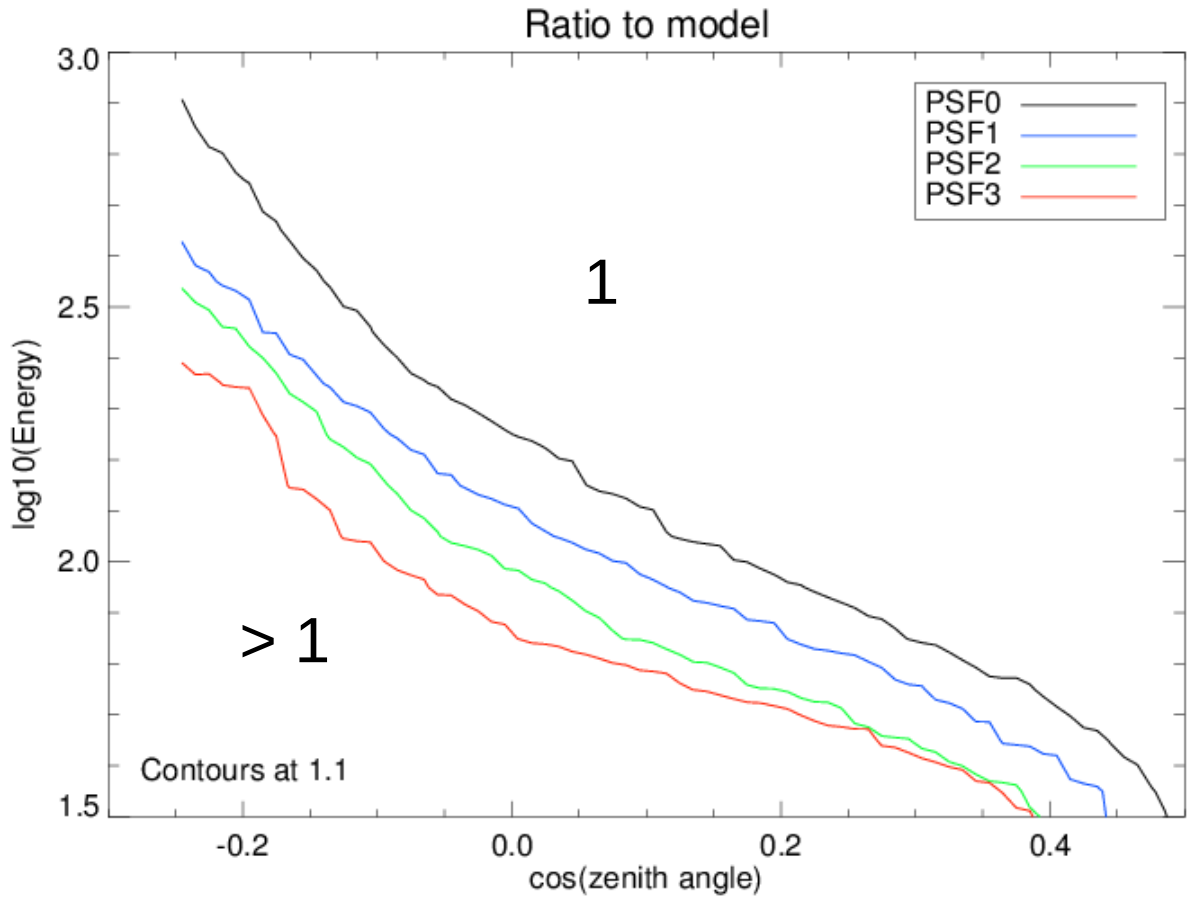
Earth limb suppression

3FGL: Earth limb templates centered on celestial poles, very steep spectra

Earth limb contamination largest for event types with broad PSF (Back, PSF0)

Build exposure map as a function of zenith angle

Compare data per cosz element with expectation from small z



Earth limb suppression

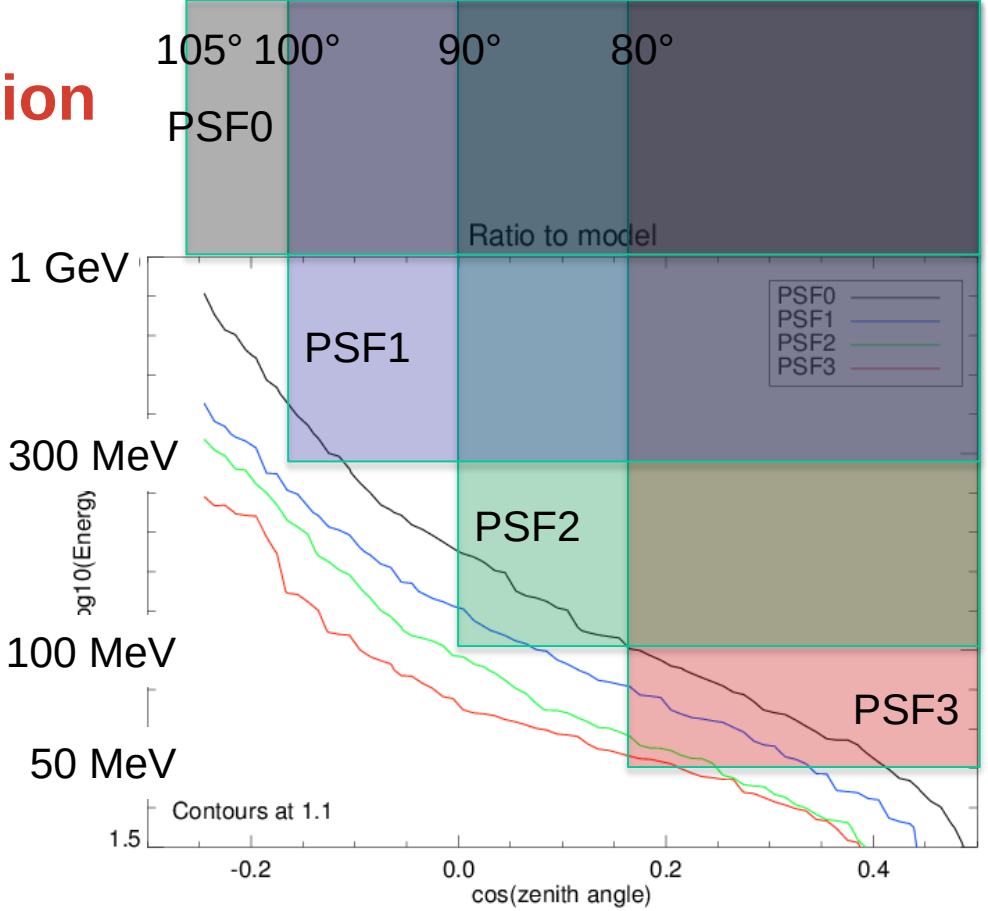
The low-energy sky is limited by systematics anyway

Better solution: **select events with best angular resolution** at low energy

Fights confusion while allowing less stringent cut on zenith angle

Cut on zenith angle when Earth limb contribution becomes larger than 10% of regular photons at that zenith angle

Side-effect: changing cut with energy results in slightly different time intervals



50 – 100 MeV: PSF3 only, z < 80° (not used in FL8Y)

100 – 300 MeV: PSF2+3, z < 90°

0.3 – 1 GeV: PSF1+2+3, z < 100°

> 1 GeV: all events, z < 105°

Energy dispersion

The LAT has $\Delta E/E$ **around 10%** over most of the energy range

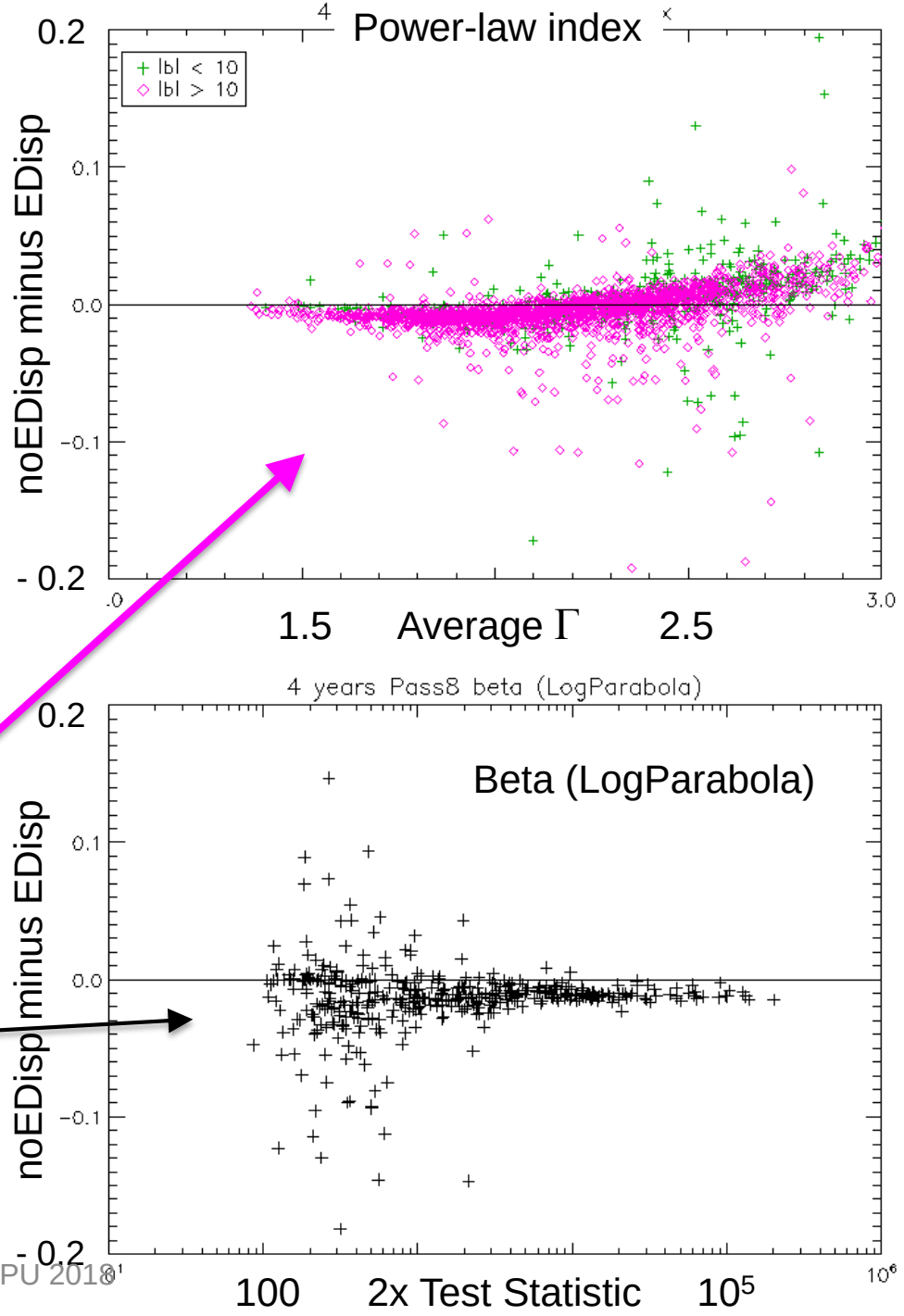
Small effect neglected in 3FGL

Worse below 100 MeV (combined with sharply increasing effective area) and above 500 GeV

Implemented in Science Tools in a simplified way (independent of PSF)

Power-law index distribution slightly narrower (hard sources softer, soft sources harder but only by 0.01-0.02)

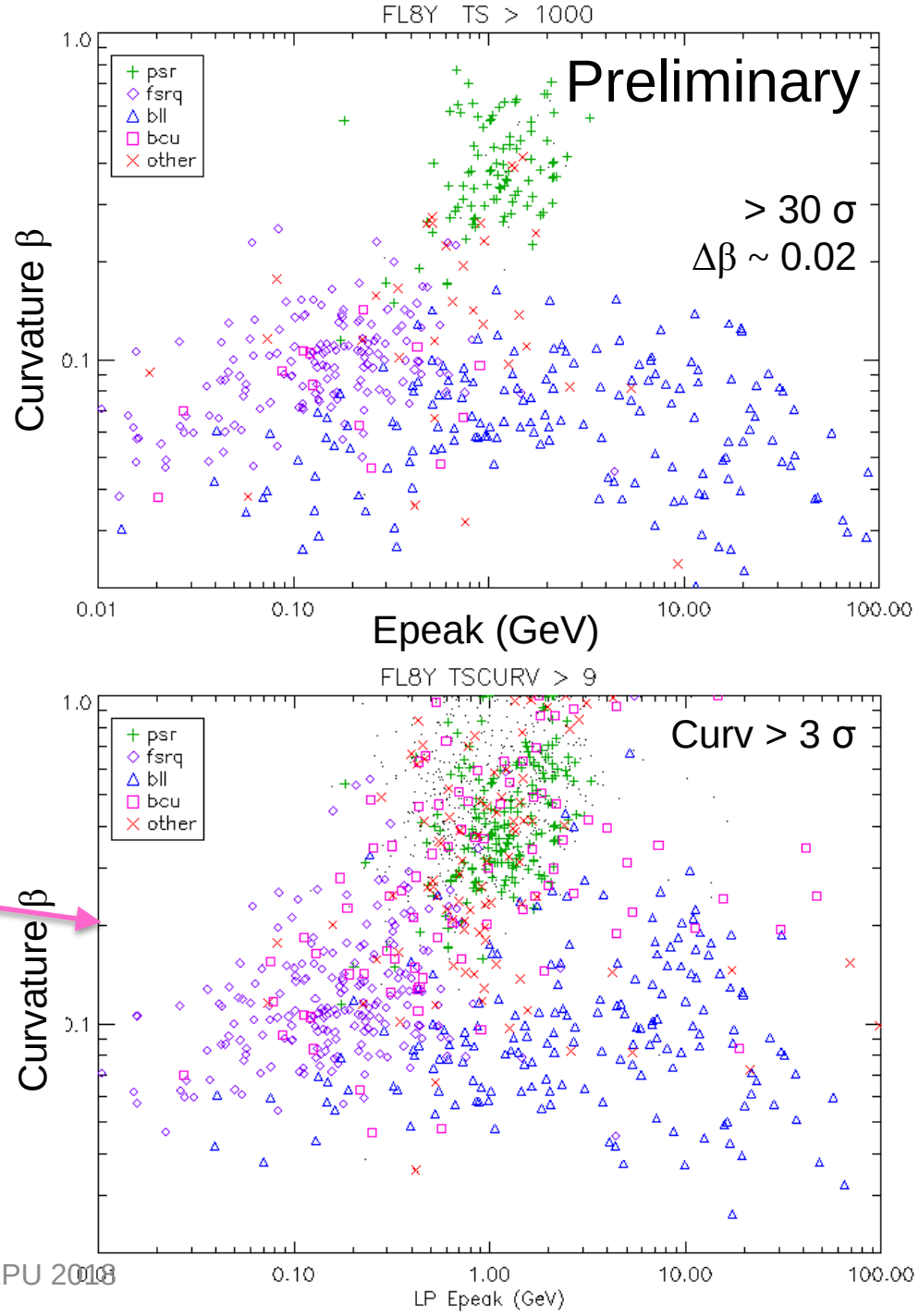
Curved sources **more curved** (energy dispersion broadens spectrum) but β larger by only 0.01 on average



Classification from LogParabola parameters

Why aren't significantly curved bcu in line with fsrq/bll?

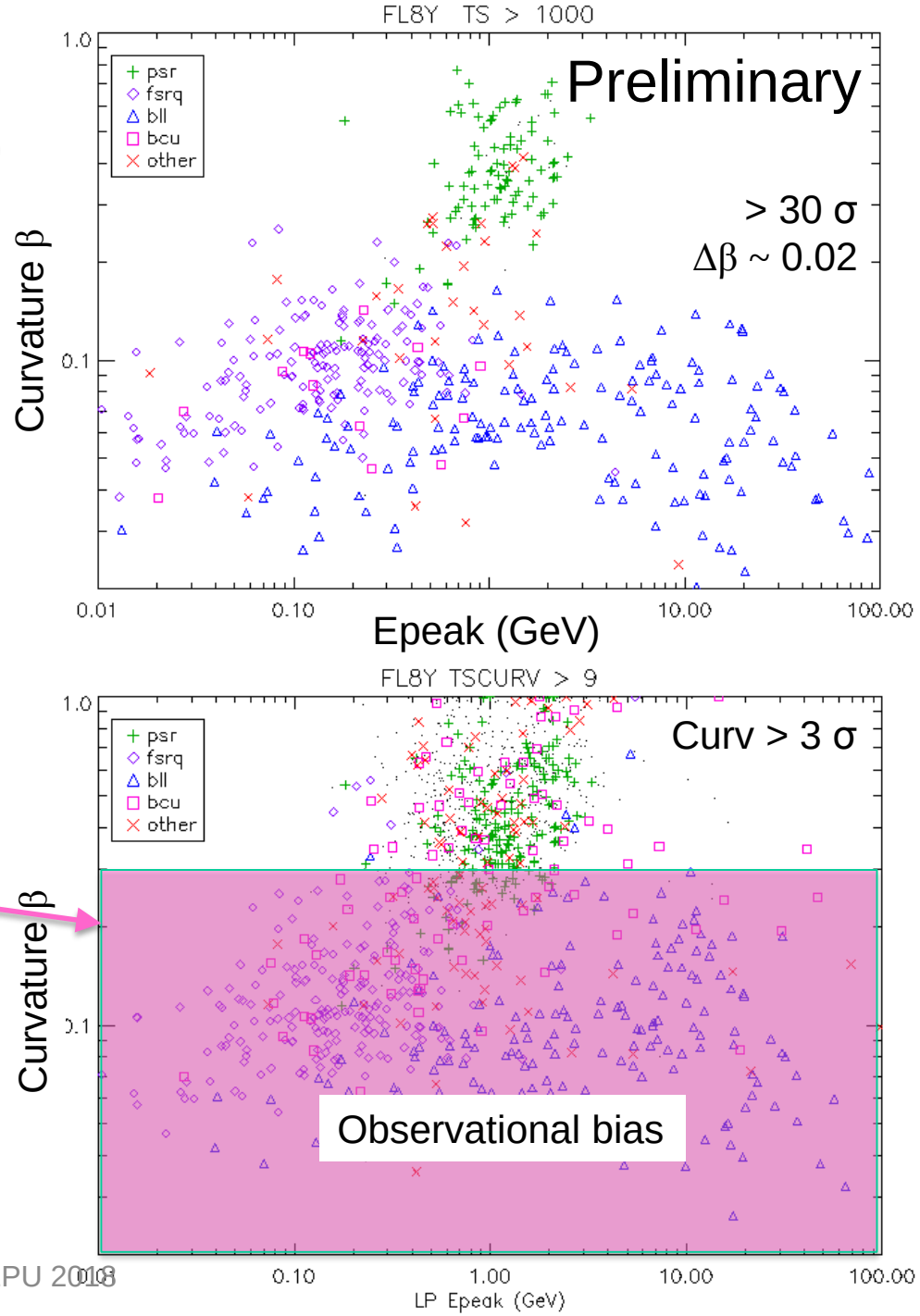
- **PSR/psr**: median TS = 800, Unc_LP_beta ~ 0.04
- **FSRQ/fsrq**: median TS = 2000, Unc_LP_beta ~ 0.02
- **BLL/bll**: median TS = 2000, Unc_LP_beta ~ 0.02
- **bcu**: median TS = 160, Unc_LP_beta ~ 0.10
- **other**: median TS = 340, Unc_LP_beta ~ 0.06
- unassoc: median TS = 88, Unc_LP_beta ~ 0.15



Classification from LogParabola parameters

Why aren't significantly curved bcu in line with fsrq/bll?

- **PSR/psr**: median TS = 800, Unc_LP_beta ~ 0.04
- **FSRQ/fsrq**: median TS = 2000, Unc_LP_beta ~ 0.02
- **BLL/bll**: median TS = 2000, Unc_LP_beta ~ 0.02
- **bcu**: median TS = 160, Unc_LP_beta ~ 0.10
- **other**: median TS = 340, Unc_LP_beta ~ 0.06
- unassoc: median TS = 88, Unc_LP_beta ~ 0.15



Flagged sources

Many flags are not filled in FL8Y.

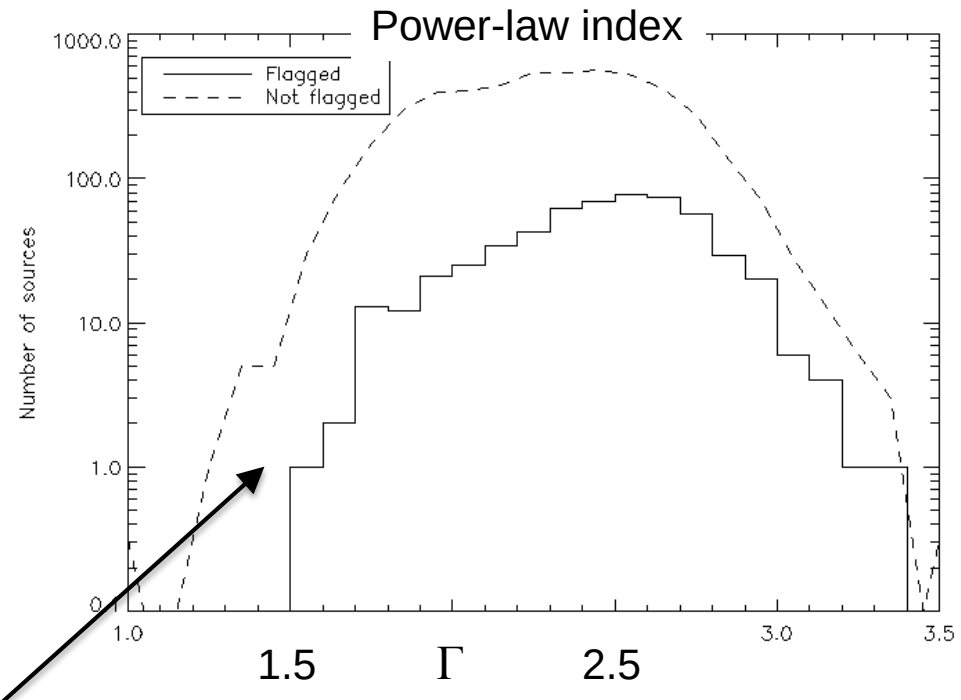
Only 4 active flags corresponding to:

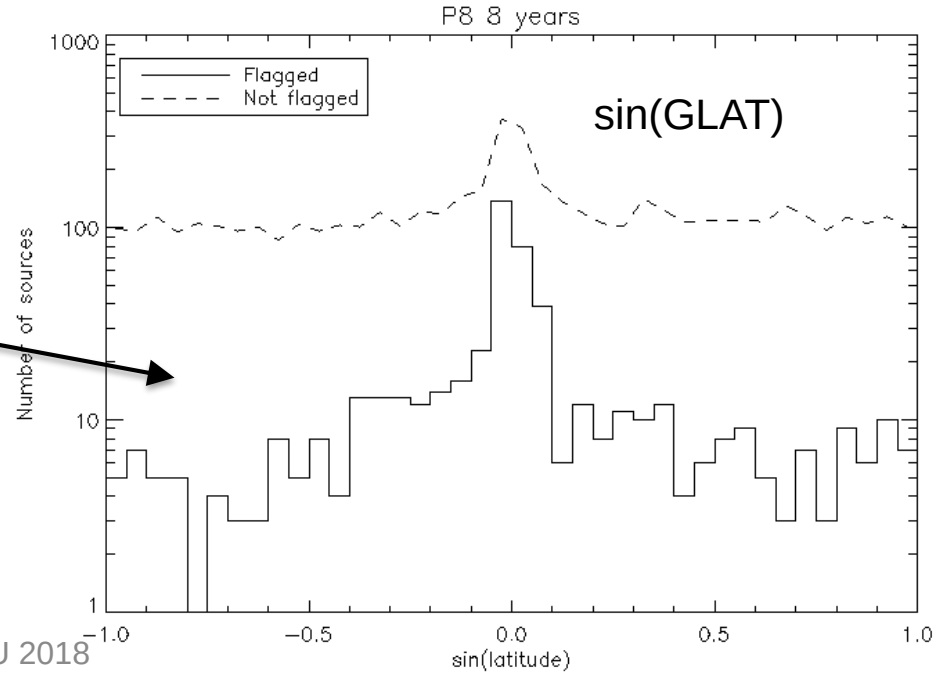
- Comparison pointlike vs gtlike (361)
- Bad localization (119, from pointlike)
- Bad spectral fit (50, from pointlike)
- Highly curved (52, from gtlike)

553 flagged sources in all (10%).

Tend to be softer than average and

In Galactic plane

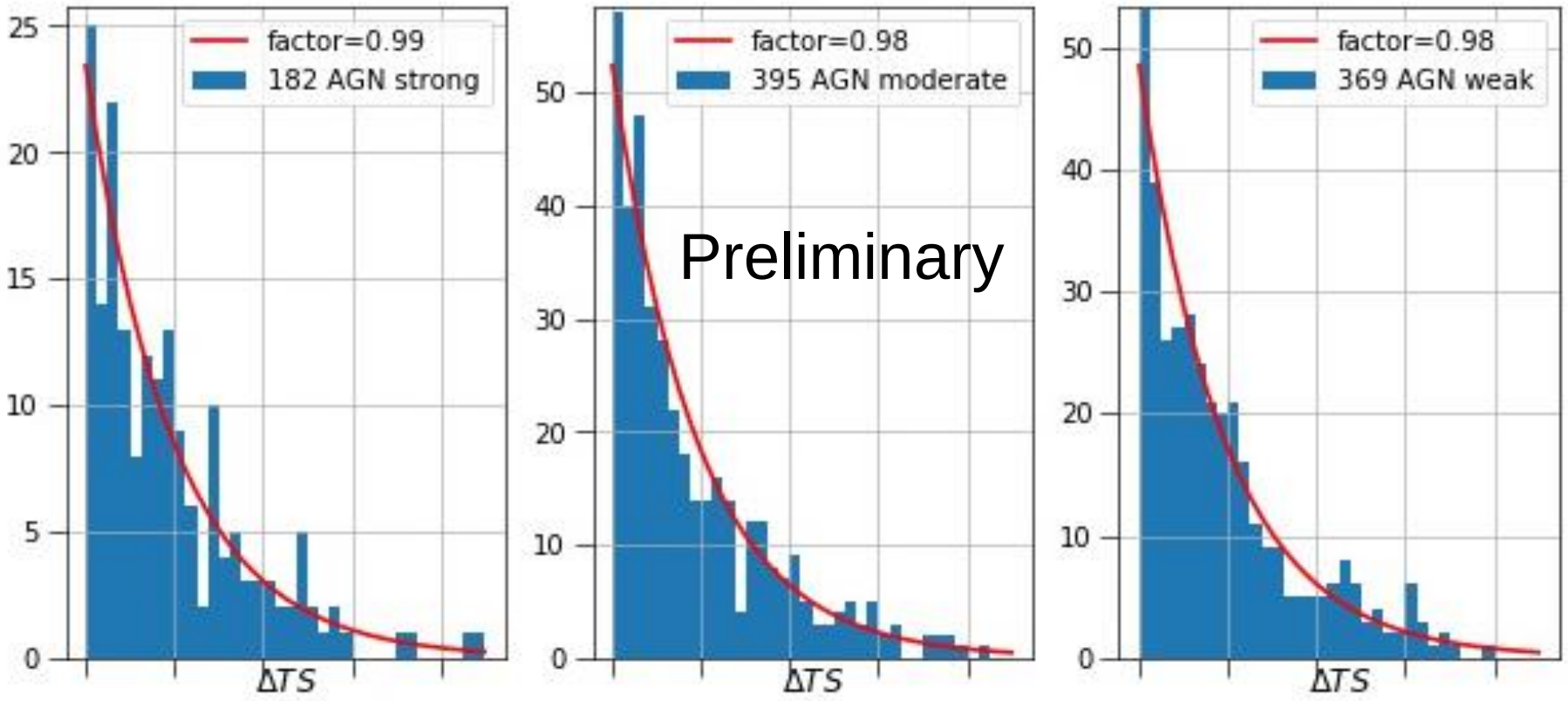




Localization systematics

Systematic factor 1.05 on error radius (as in 3FGL)

Absolute 95% systematic error as reported in 3FHL: 27 arcsec



2x ΔlnLike between best fit and counterpart positions accounting for systematic factors

Fit exponential distribution (χ^2 with 2 dof)

Associations

Based on spatial coincidence only (Bayesian method)

Many catalogs of putative counterparts (per source type: BL Lac, FSRQs, pulsars, …)

Did not include Likelihood ratio results (applied to large MWL surveys: radio, ROSAT)

Identification based on angular extent or correlated variability

65% associated in 3FGL, down to 60% in FL8Y; probably due to limited depth of counterpart catalogs

Association fraction improves with sources significance

Estimated number of false associations: 41 / 3392

22 sources have changed associations from 3FGL to FL8Y

292 3FGL sources missing, close to threshold or split

21 3FHL sources missing, close to threshold + VER J2016+372