The road to 4FGL

Jean Ballet

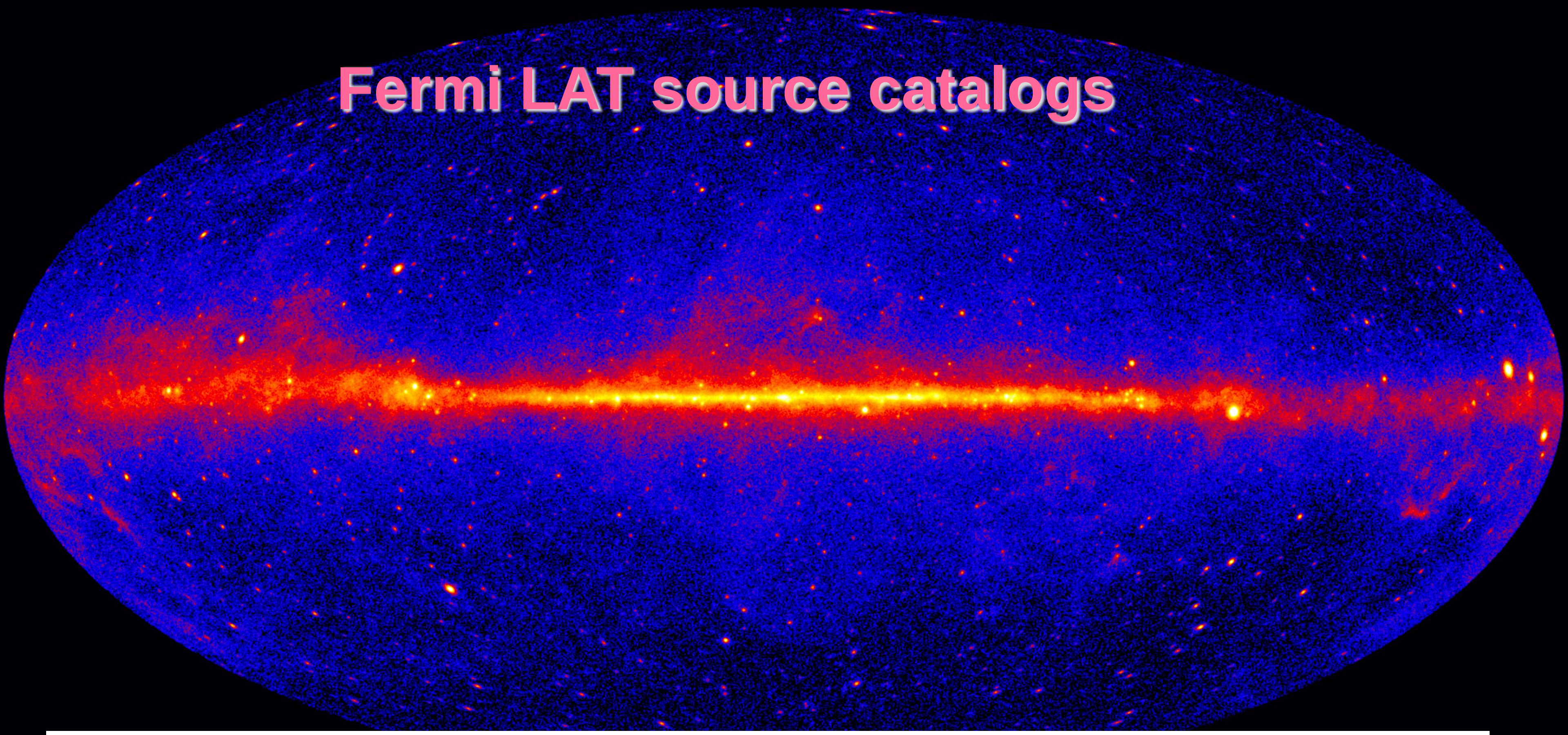
AIM, CEA Saclay, France

Toby Burnett, Seth Digel

and the LAT collaboration

Fermi symposium, Garmisch-Partenkirchen
October 19, 2017

Fermi LAT source catalogs



Purely gamma-ray based (associations only post facto)

Concentrate on **persistent** sources, detection over time-integrated data set

- 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

Handling source confusion

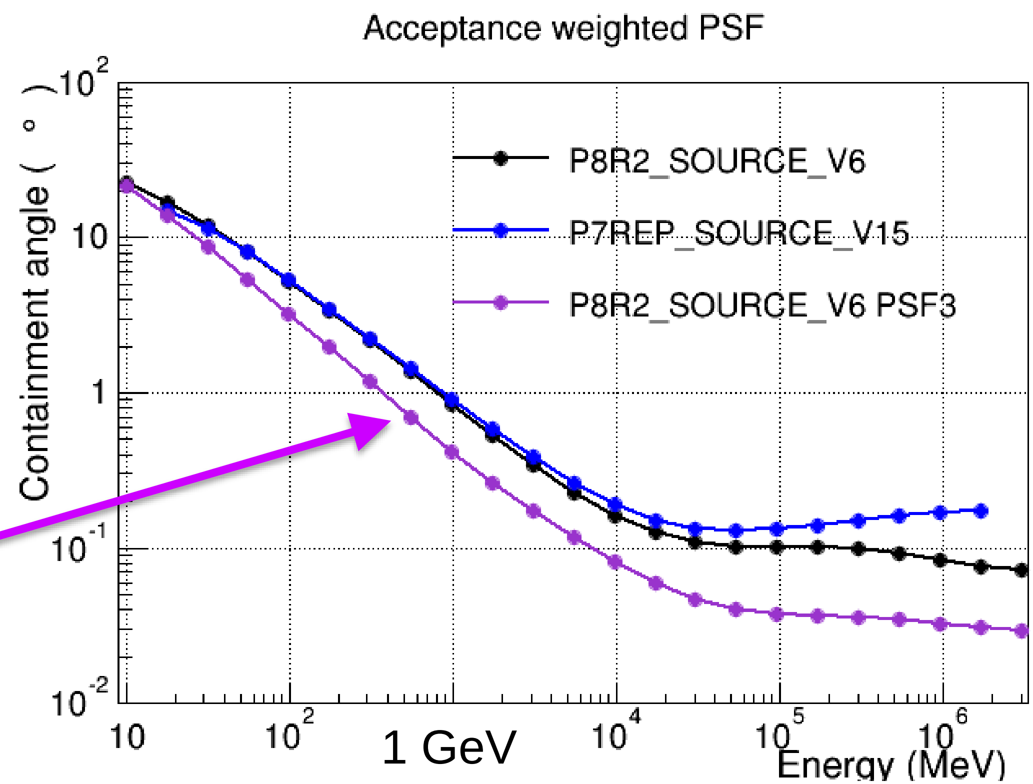
Strong **confusion** at low energy

Average source separation 2.2° (3FGL) outside Gal. plane

Equals R_{68} at 300 MeV

Equals R_{95} at 1.2 GeV

Somewhat better with Pass 8 PSF3 event type



Each source is correlated with entire sky at some point

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

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

pointlike

Refit spectrum of diffuse components

Source detection

Source localization →

Comparison for spectra

pyLikelihood

Official Science Tools and diffuse model

Thresholding

Spectral characterization

Light curves

Comparison for localization

Run with alternative diffuse model

Associations

Bayesian + Likelihood ratio

Catalog

With flags

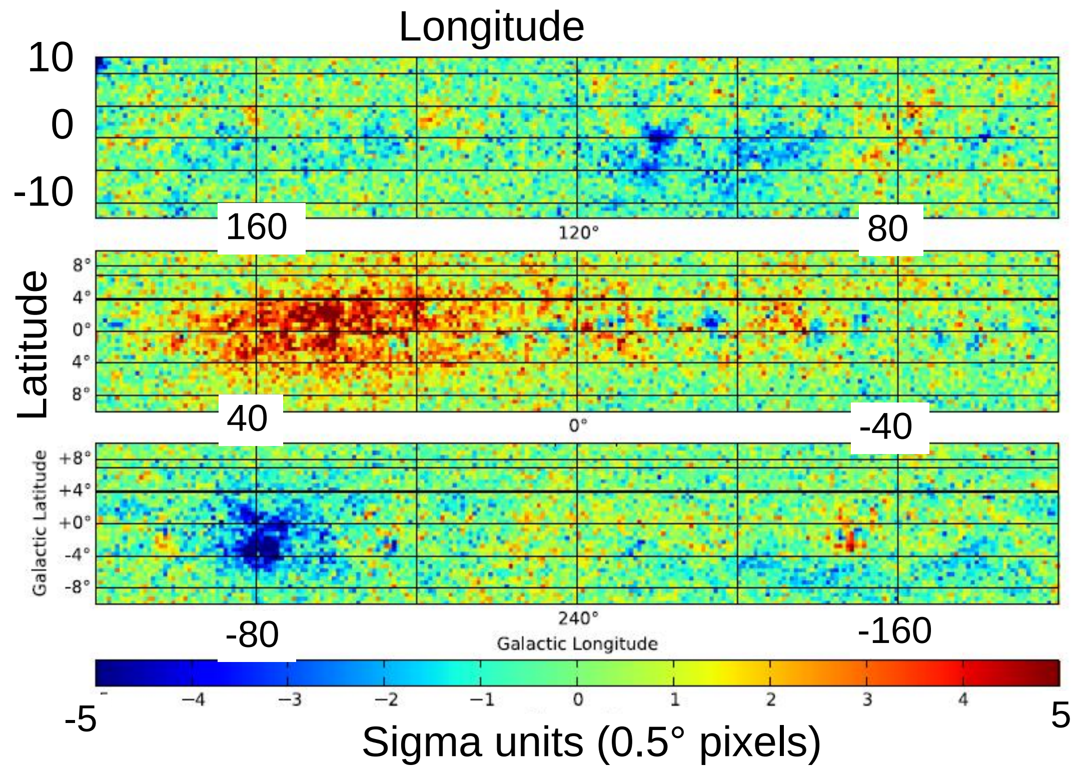
Limits of the 3FGL source catalog

Interstellar emission model (Acero et al 2016, ApJS **223**, 26) is not perfect.

Residuals at level of 2 – 3%, both in space and energy

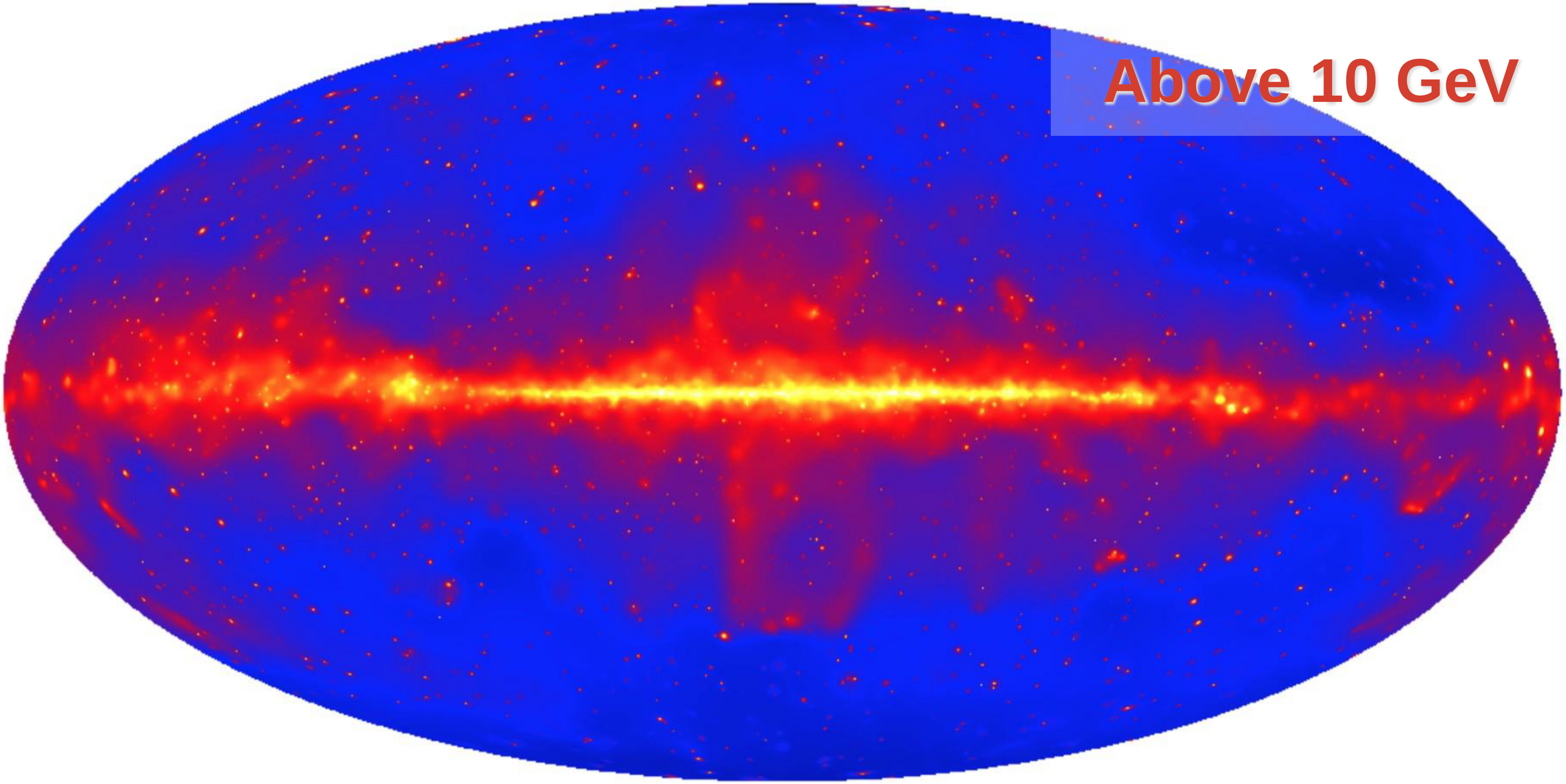
Impact sources at same level as statistical errors over the whole Galactic plane

Dominate in Galactic ridge



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

4 years of **P7Rep** data > 100 MeV, 3033 sources



Narrow PSF, **much simpler** analysis

Many fewer events (700,000)

Galactic diffuse emission not as dominant, except in Ridge

D. Thompson on Monday

3FHL (7 years, P8)

> 1500 sources

Ajello et al 2017, ApJS **232**, 18

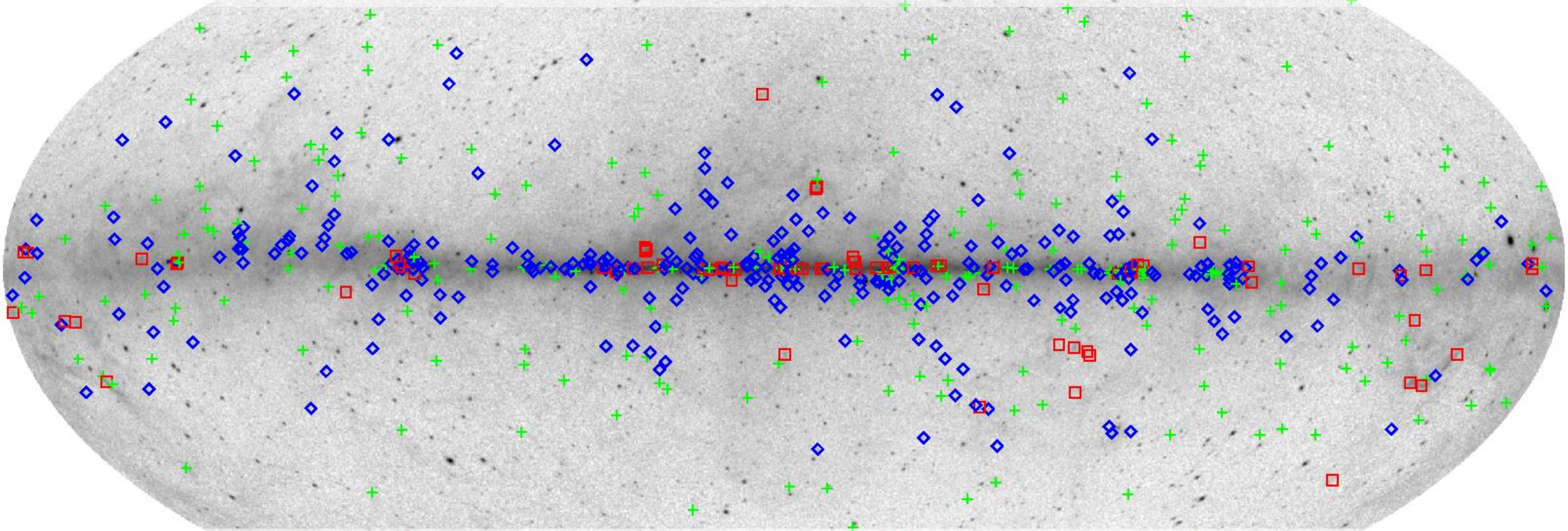
Developments for 4FGL

Why did it take so long?

1. Improved interstellar emission model
2. Weighted logLikelihood
3. Earth limb suppression
4. Energy dispersion

Interstellar Emission Model for 4FGL: Motivations

- We need a model of diffuse emission to increase the sensitivity to point and small extended sources, and to more accurately characterize them
- Systematic uncertainties in the Galactic diffuse emission model limit the precision with which sources can be characterized



Green: All flagged

Blue: Unassoc, strongly dependent on diffuse model

Red: Interstellar clump (c)

– **The potential for spurious sources related to Galactic diffuse emission has been recognized**

– **Multiwavelength data to improve the situation are now available**

Approach to Modeling Diffuse Emission

follows Acero et al 2016, ApJS 224, 8

- **Diffuse emission is produced by cosmic-ray interactions with interstellar gas (pi0 decay and Bremsstrahlung) and photons (inverse Compton scattering)**
  - **Optical depths are typically quite small**
  - **Cosmic rays are relatively smoothly distributed**
- **These imply that a model could be considered a linear combination of templates corresponding to different regions of the Galaxy or different phases of the ISM**
  - **With gamma-ray spectral information related to the proton/electron spectra in the vicinity**
- **Developing a model becomes identifying these templates and fitting them to the data**
  - **Maximum likelihood analysis (using Gardian*)**
  - **Need to iterate with source detection and characterization**
    - **The pointlike step of the Catalog pipeline**

* See Ackermann et al. (2012; arXiv:1202.4039)

Modeling Diffuse Emission (2)

- A fundamental challenge is that none of the templates tracing ISM or photon components is known even approximately perfectly
  - Atomic gas, molecular gas, dark gas, photon radiation field
- In addition, not every needed component has a multiwavelength counterpart
  - Fermi bubbles, other potential extended regions
- So part of identifying suitable templates is testing potential alternatives

Diffuse Emission Modeling Improvements for 4FGL

presentation by S. Digel this morning

- **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, new nonlinear relations between dust optical depth and column density*, and adaptations for metallicity gradients, 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**

* Presentation by Q. Remy this morning

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

w_i reduces the importance of systematics-dominated areas/energies

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

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

Weighted logLikelihood

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

Now how to define $\boldsymbol{B_i}$?

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

Integral under PSF

$$B_i = \mathrm{N}(\mathbf{r}_\mathrm{i}, E_i) = \int_{E_i}^{E_{\max}} S(\mathbf{r}_\mathrm{i}, E)\, dE$$

Integral above current energy

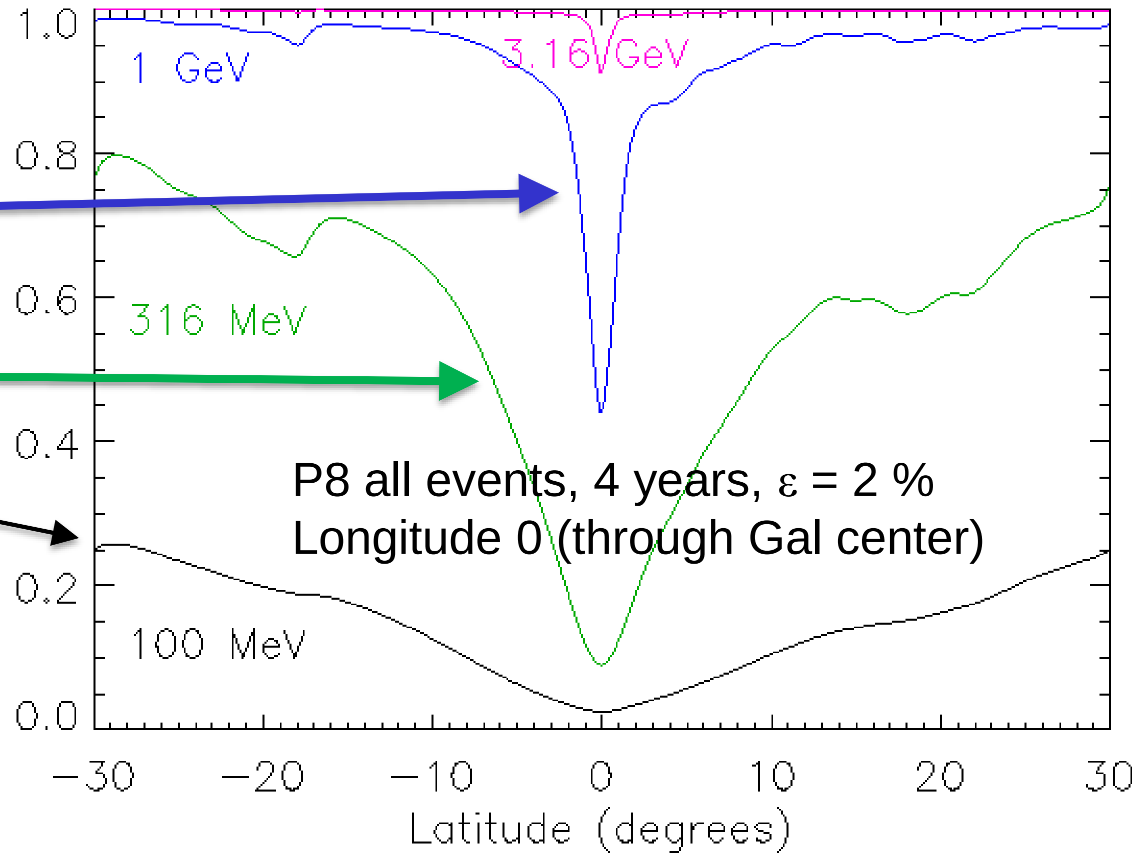
Ad-hoc but **desirable asymptotic limits**, stable against rebinning

$R_{68}(E)^2$ 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 **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

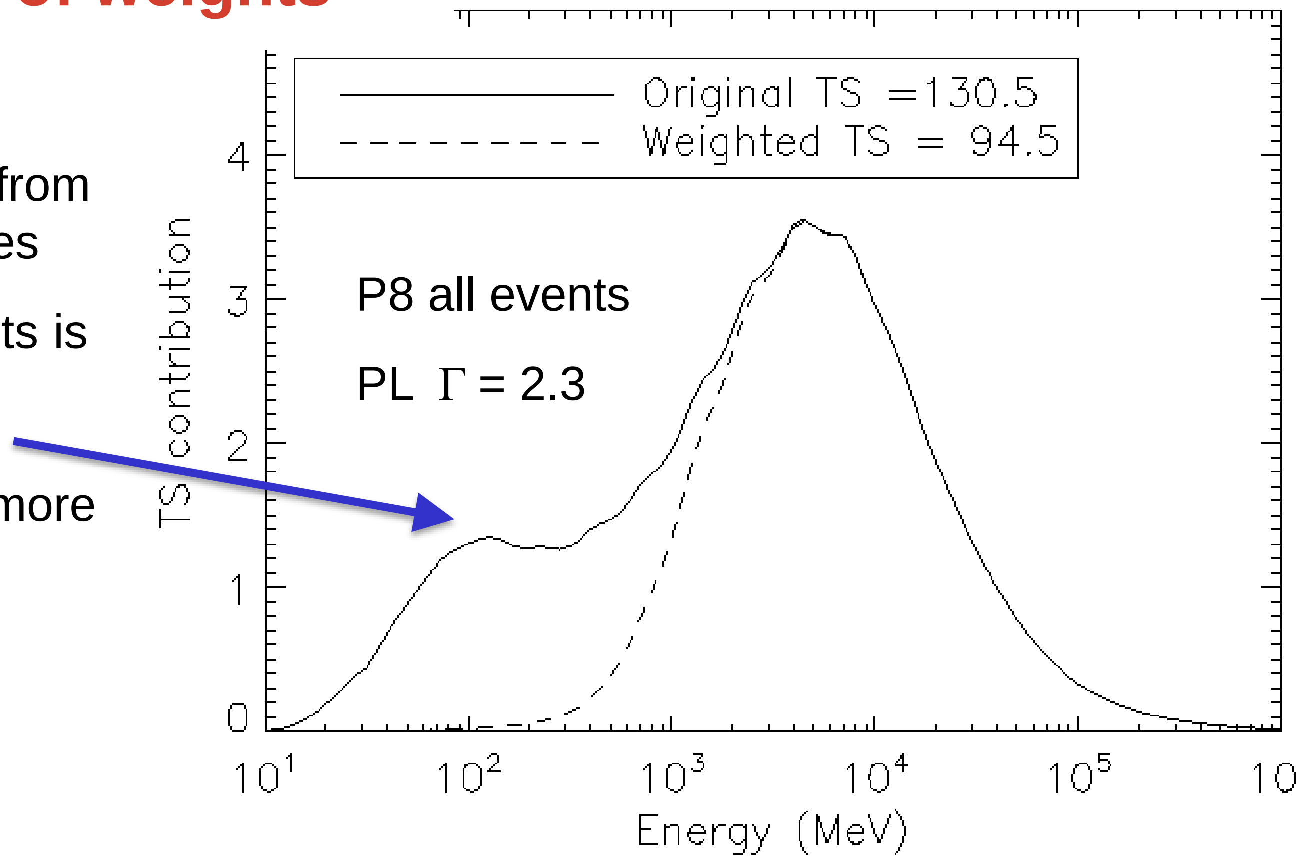


Applications in E. Charles' presentation

Effect of weights

Standard TS 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** **Data-based weights**

Model-based weights
- Background is **interstellar emission model** only
- Original motivation

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

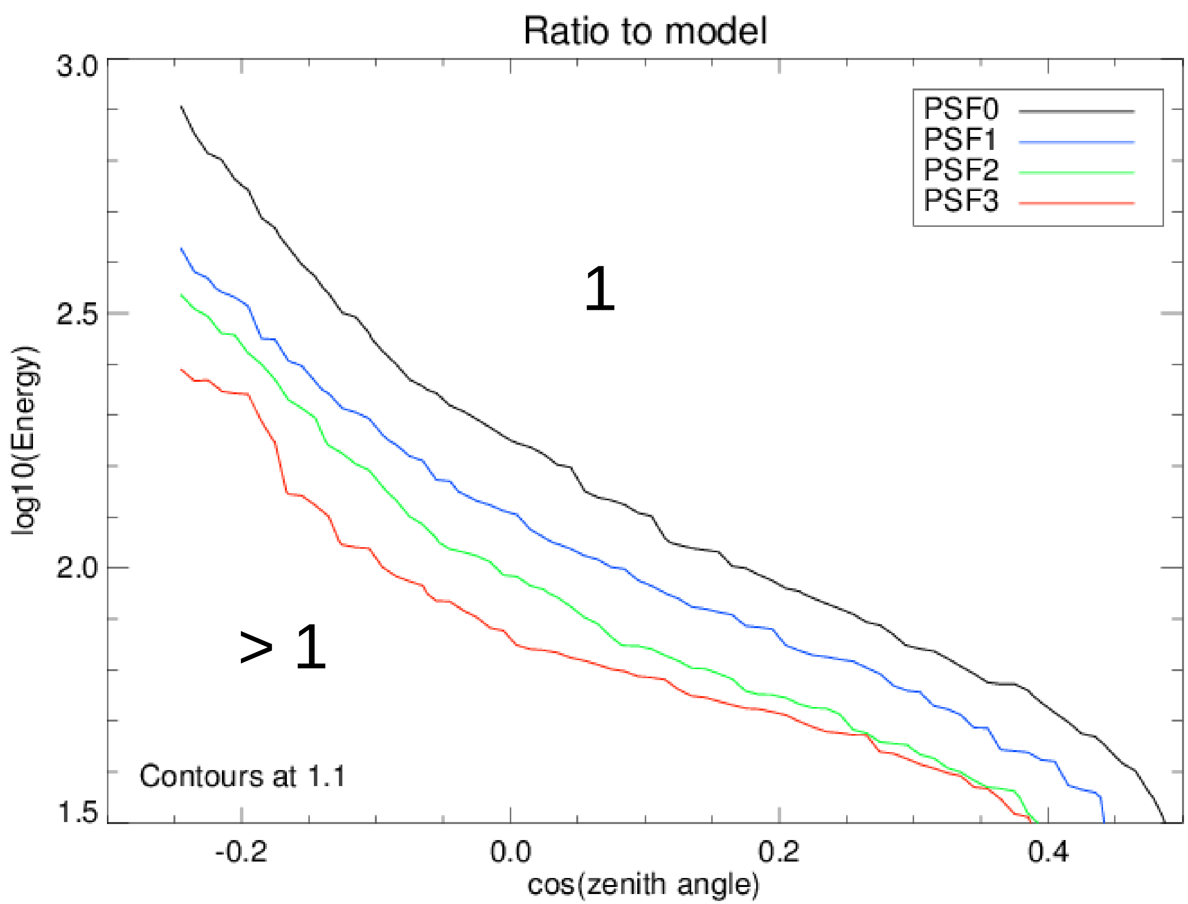
Earth limb suppression

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

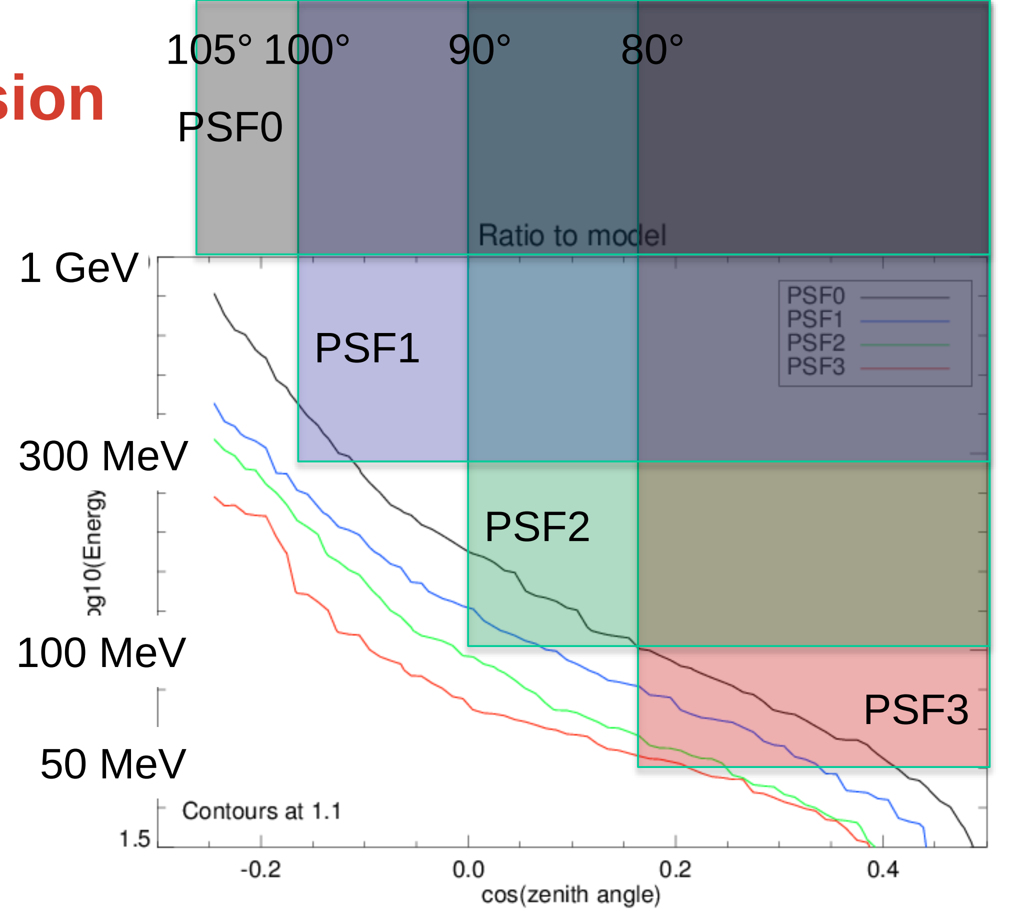
The low-energy sky is limited by systematics anyway

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°

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 ΔE/E **around 10%** over most of the energy range

Small effect neglected in 3FGL

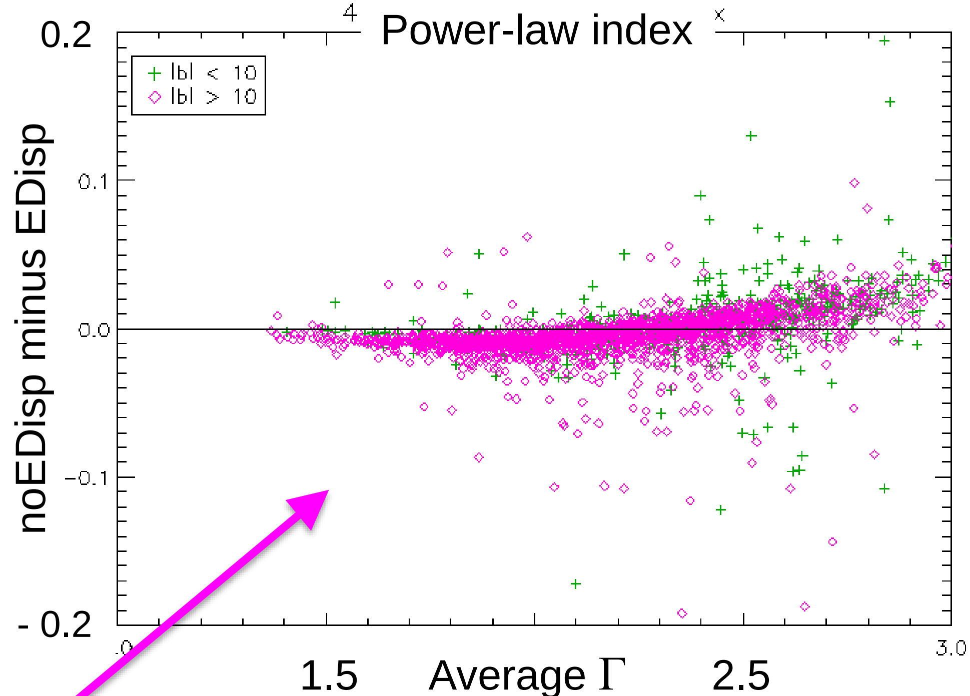
Worse below 100 MeV (combined with sharply increasing effective area) and above 500 GeV

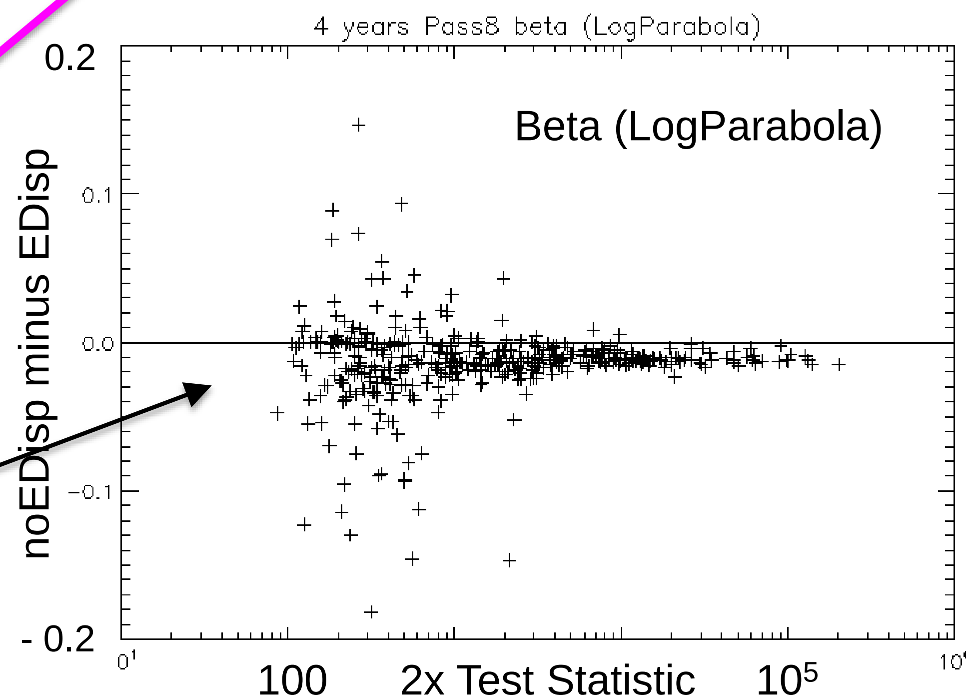
Implemented in Science Tools in a simplified way (independent of PSF)

Tested on 3FGL sources

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





3FGL vs 4FGL

3FGL		4FGL
4 years P7Rep	**Data**	8 years **P8**, TS x 2.5 (time x eff. area)
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**	**New** version
25	**Extended sources**	58 (will add FHES - R. Caputo's talk)
0.1 - 0.3 - 1 - 3 - 10 - 100 GeV	**SED bands**	0.05 - 0.1 - 0.3 - 1 - 3 - 10 – 30 - 300
48 x 1 month	**Light curves**	48 x 2 months
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

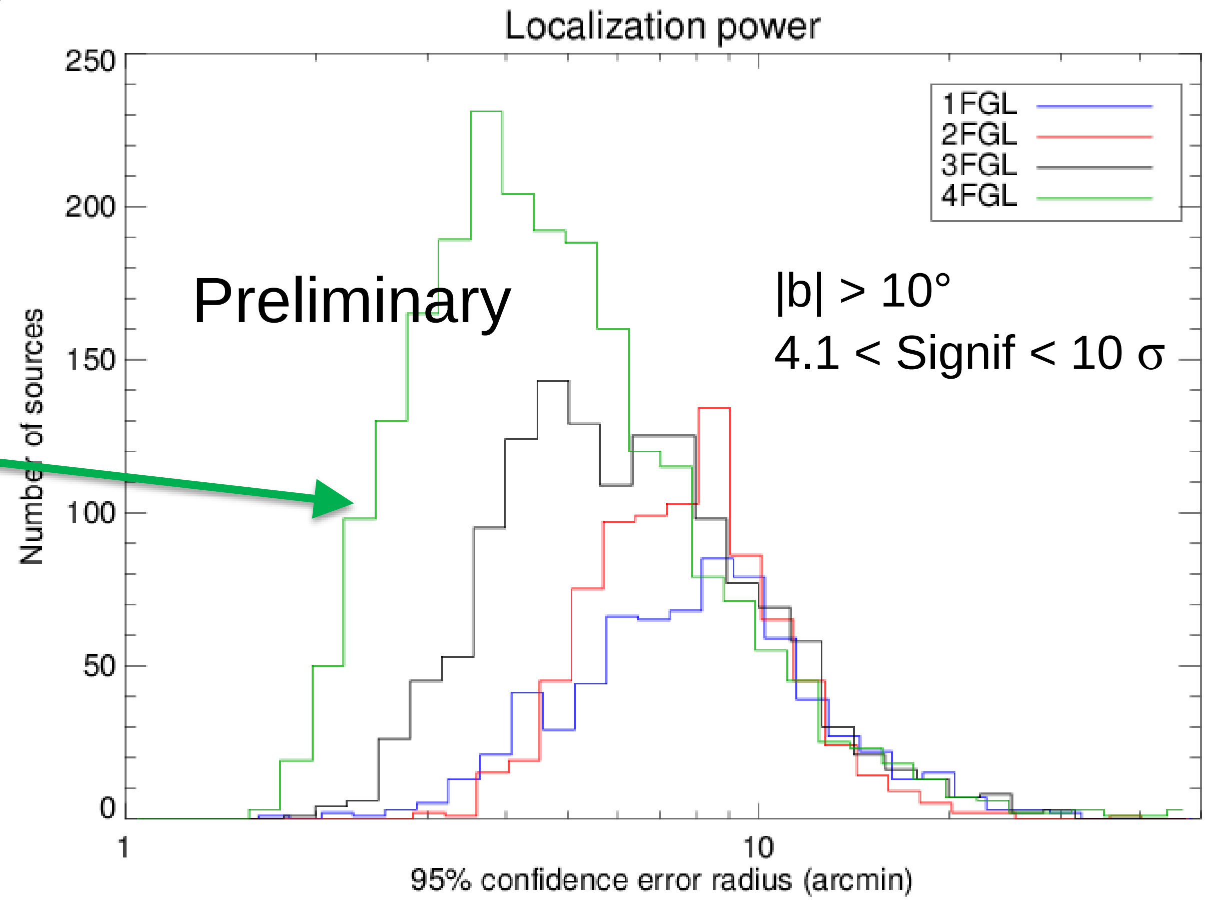
pointlike early results

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)

> 10,000 seeds at TS > 10

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

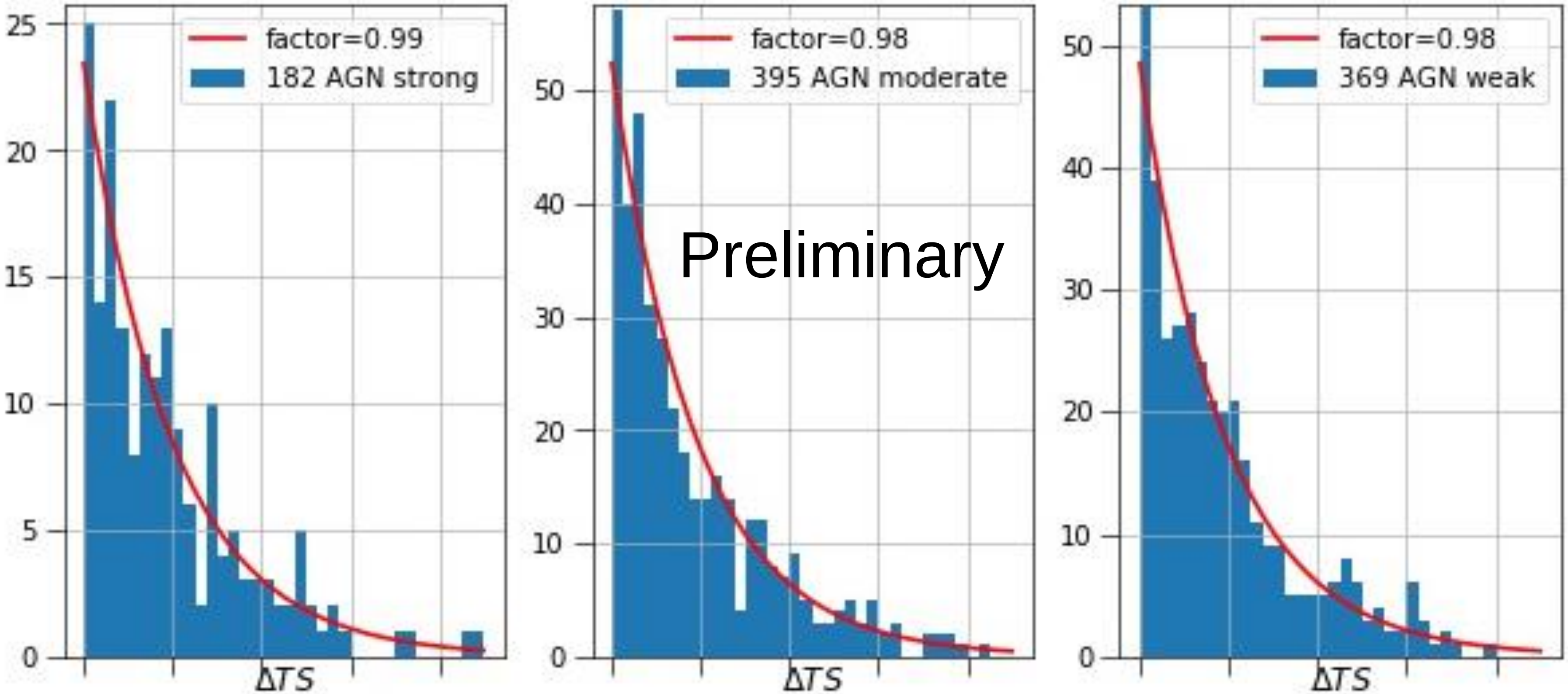
Median error radius at 25 < TS < 100 is **4.5 arcmin**



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)

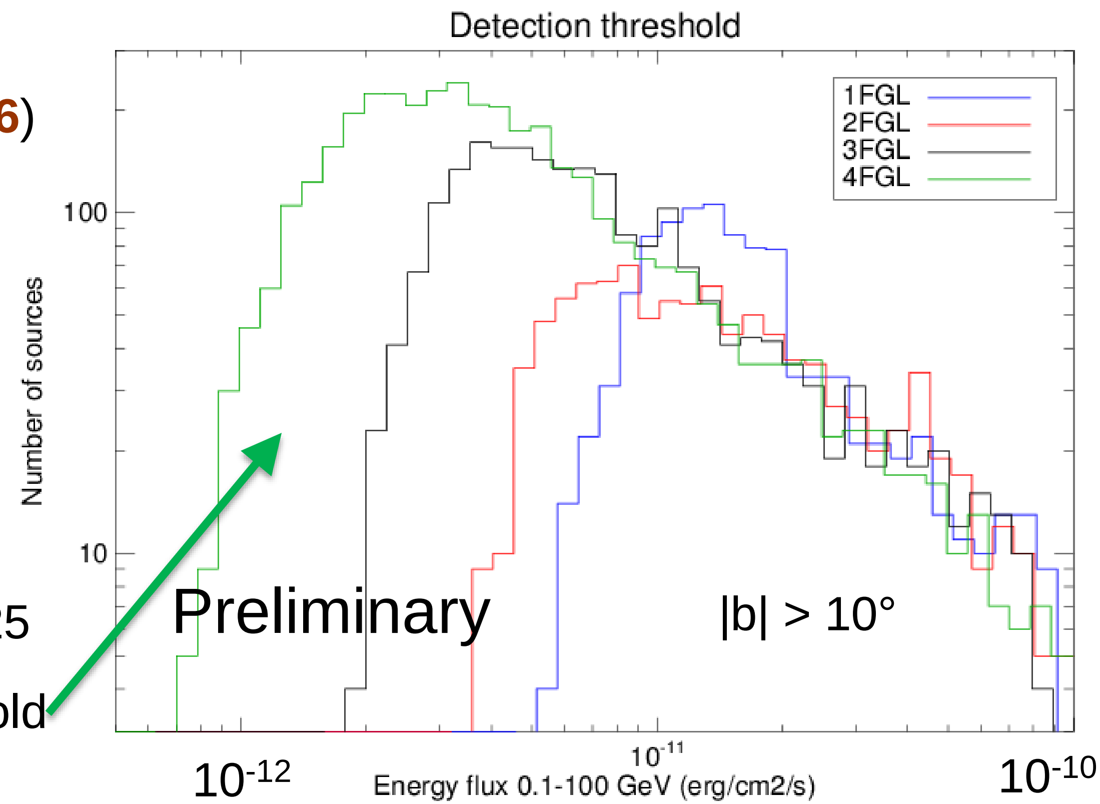
pyLikelihood and associations

First run with current (**gll_iem_v06**) diffuse model

- ✓ Data-based weights
- ✓ No energy dispersion
- ✓ Fully binned analysis (faster)

Very preliminary numbers

- About 5,500 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



About 40% **unassociated** sources from Bayesian method only

Fraction was 36% in 3FGL

Summary

- Main Fermi-LAT catalog (3FGL) is starting to get out of date
- Source reliability limited by **systematics** at low energy
- Improving the **interstellar emission model**
- Adding explicit formalism (**weights**) to account for systematics in the significance and errors
- 4FGL: 8 years of Pass 8 data
- More than 5,000 sources
- Early availability (not full contents) at **end of 2017**