Analyzing low-energy *Fermi*-LAT data

The source catalog approach

Jean Ballet

DAp/AIM, CEA Saclay, France

for the Fermi summer school 2022

Easy analysis above 1 GeV

- PSF $R_{68} < 1°$, $R_{95} < 2.5°$
- Uncertainties dominated by statistical fluctuations
- Isotropic + Galactic background

logLikelihood approach: $$\log L = \Sigma_i (n_i \log M_i - M_i)$$

10x10° square ROI enough (pixels < 0.1° to optimize high-energy events)

Include background sources 2° beyond (3° for the brightest ones)

Binned exposure map over the entire sky, in celestial coordinates (simpler interpolation). The pixels need not be as small as in the images (0.5° or even 1° is fine)

Difficulties below 1 GeV

- PSF is broader
  - Correlations between sources
  - Faint sources are background-dominated

- Many events
  - Binned likelihood
  - Uncertainties can be dominated by systematics

- Earth limb contamination

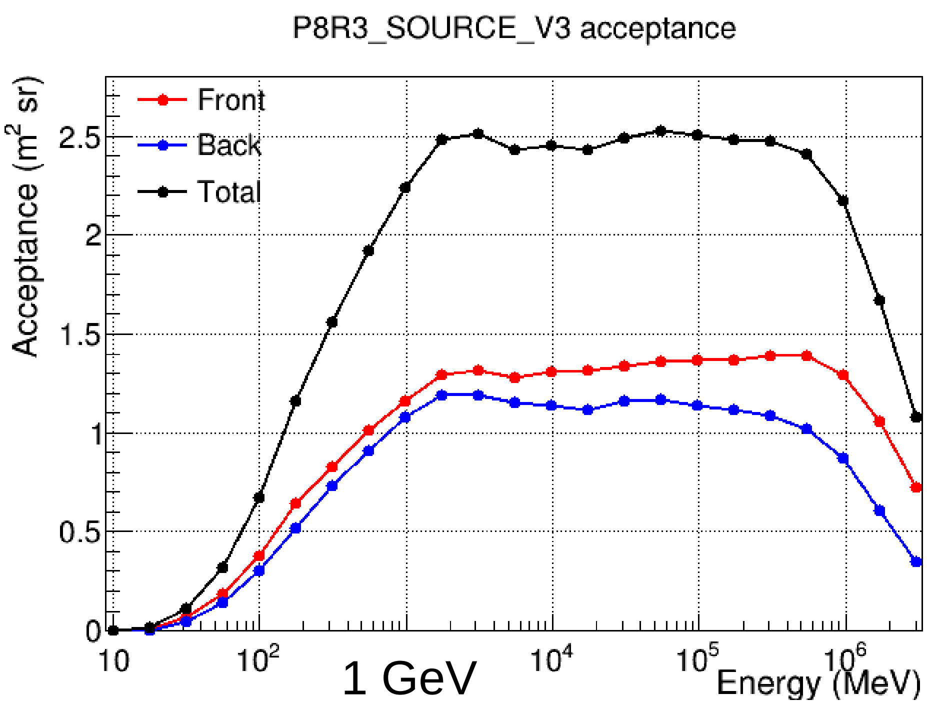
- Effective area increases fast with energy

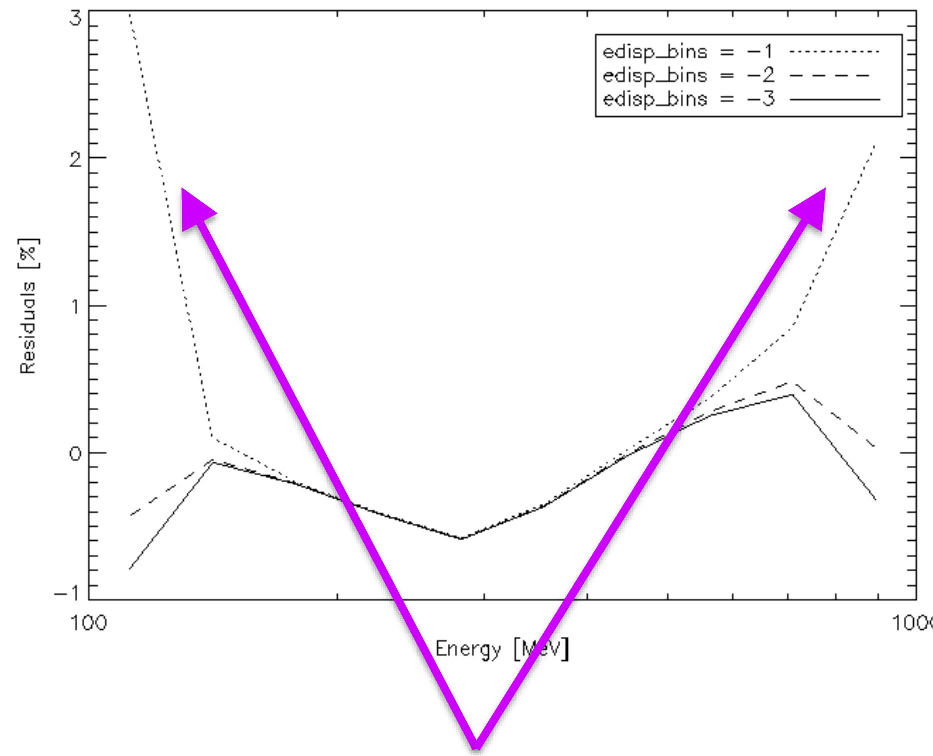
All those effects get **worse toward low energies** and are particularly acute below 100 MeV

Addressed in the 4FGL and 4FGL-DR3 papers

The effective area and acceptance

Specific actions to maintain precision





Fast variations below 1 GeV:
Use at least 10 bins per decade
in counts and exposure maps

Avoid spectral residuals at both ends due to energy dispersion

edisp_bins = –2 (at least 0.2 decade)

The point-spread function

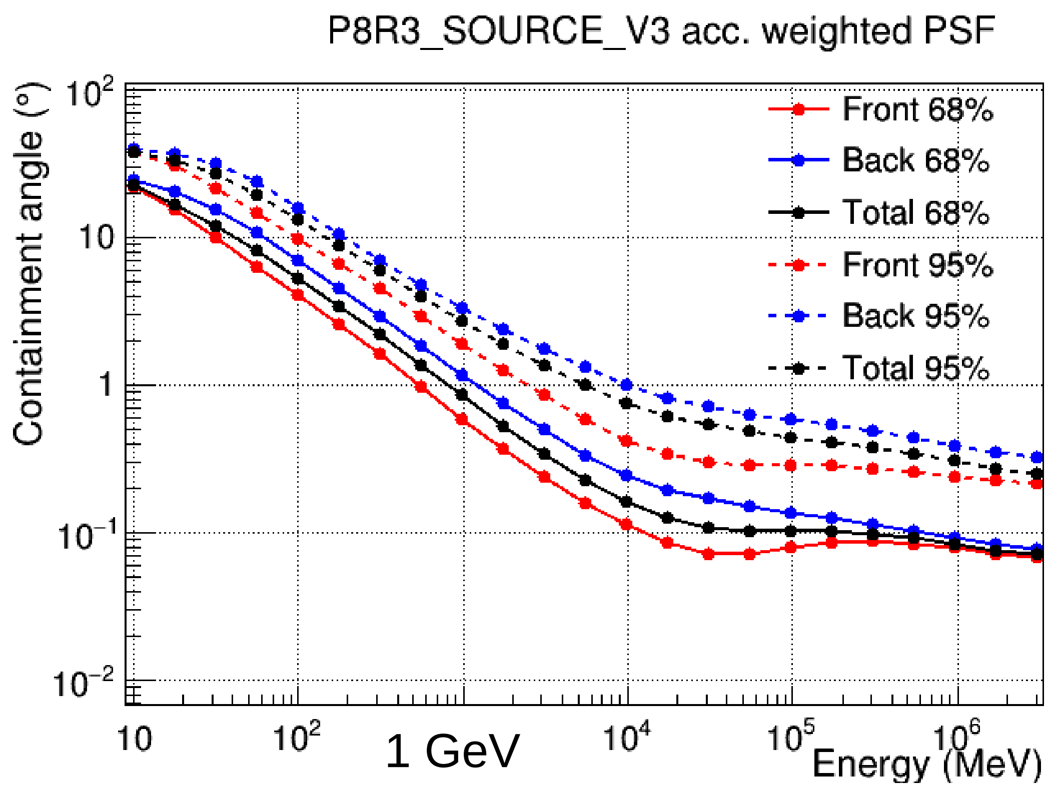
At 1 GeV, R_{68}=0.8° R_{95}=2.6°

At 316 MeV, R_{68}=2.1° R_{95}=5.8°

At 100 MeV, R_{68}=5.2° R_{95}=12.8°

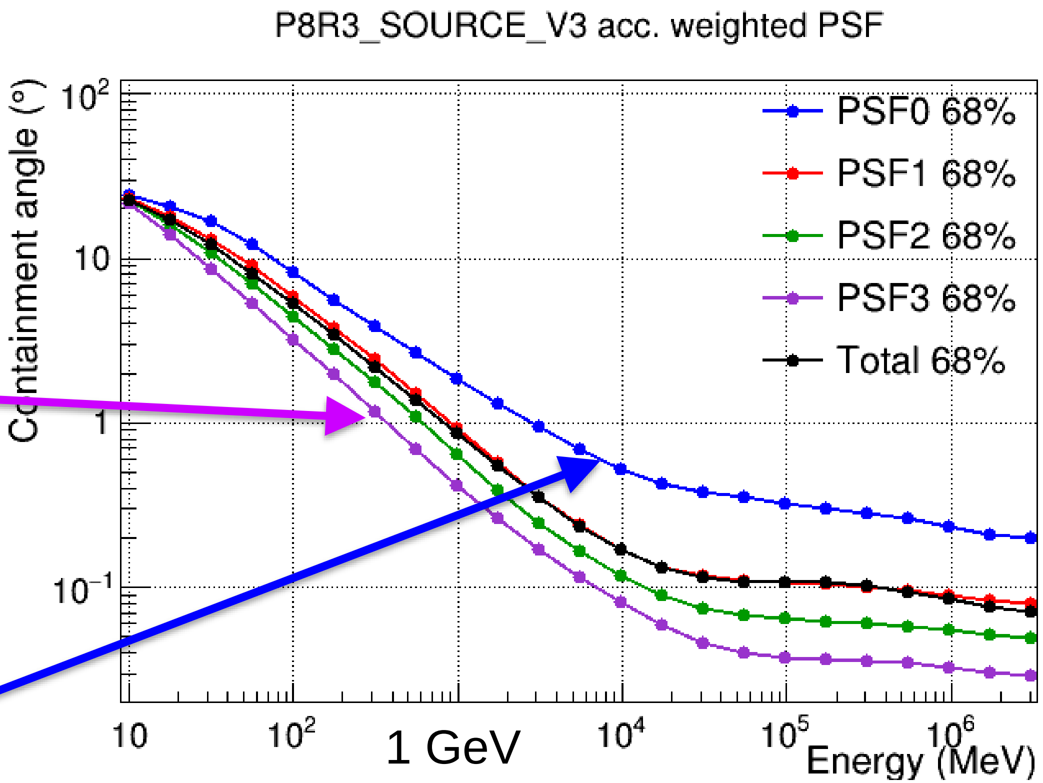
Everything stems from there

- Earth limb contamination
- Source confusion
- Background dominance



The point-spread function 2

PSF event types were invented at Pass 8 to optimize angular resolution



Better with PSF3 event type:

At 316 MeV, R_{68}=1.1° R_{95}=2.7°

At 100 MeV, R_{68}=3.2° R_{95}=7.1°

Worse with PSF0 event type

At 316 MeV, R_{68}=3.8° R_{95}=8.5°

At 100 MeV, R_{68}=8.2° R_{95}=17.5°

When working on long exposures, it is more important to **optimize the PSF** than to include as many events as possible

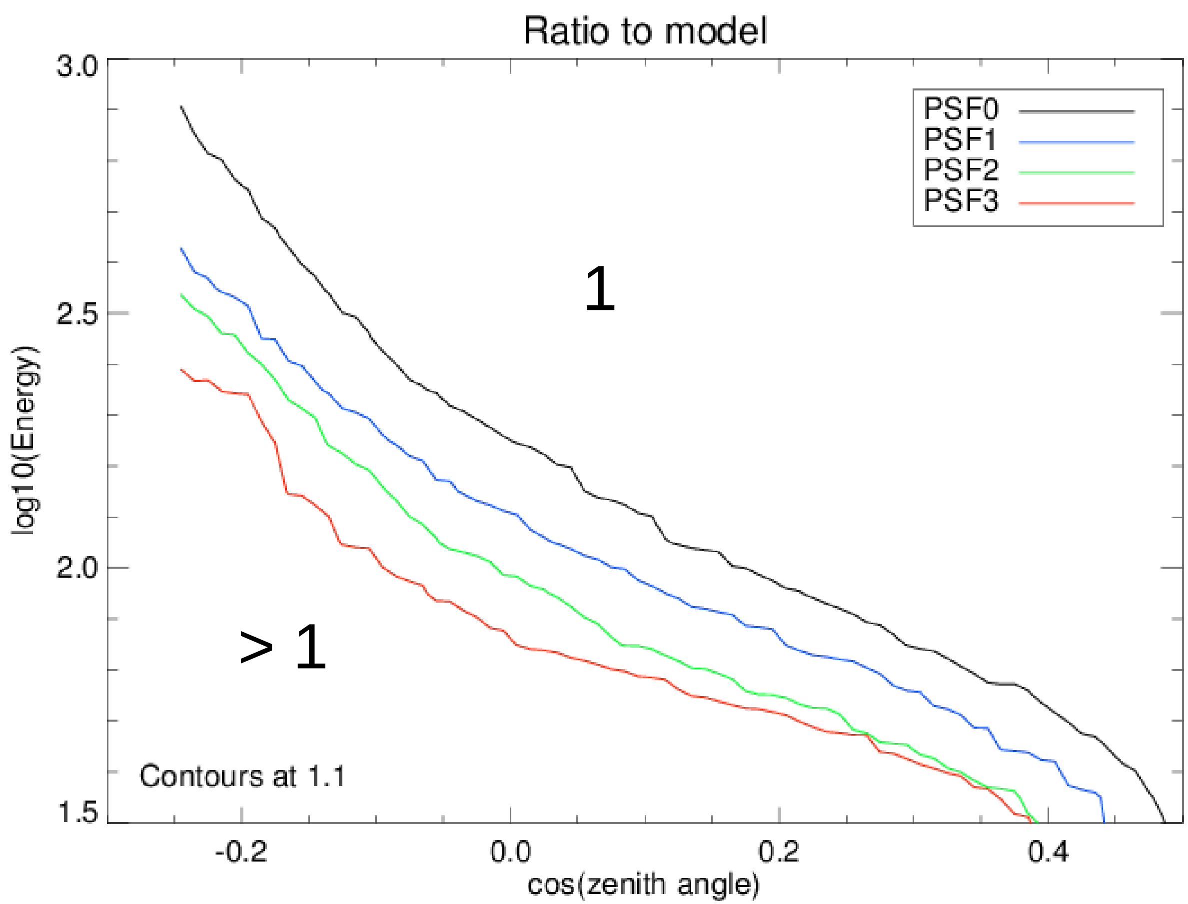
Earth limb suppression

Earth limb contamination depends mostly on the PSF

Build exposure map as a function of zenith angle

Compare data per cosz element with expectation from small z

Contours at 10% contamination



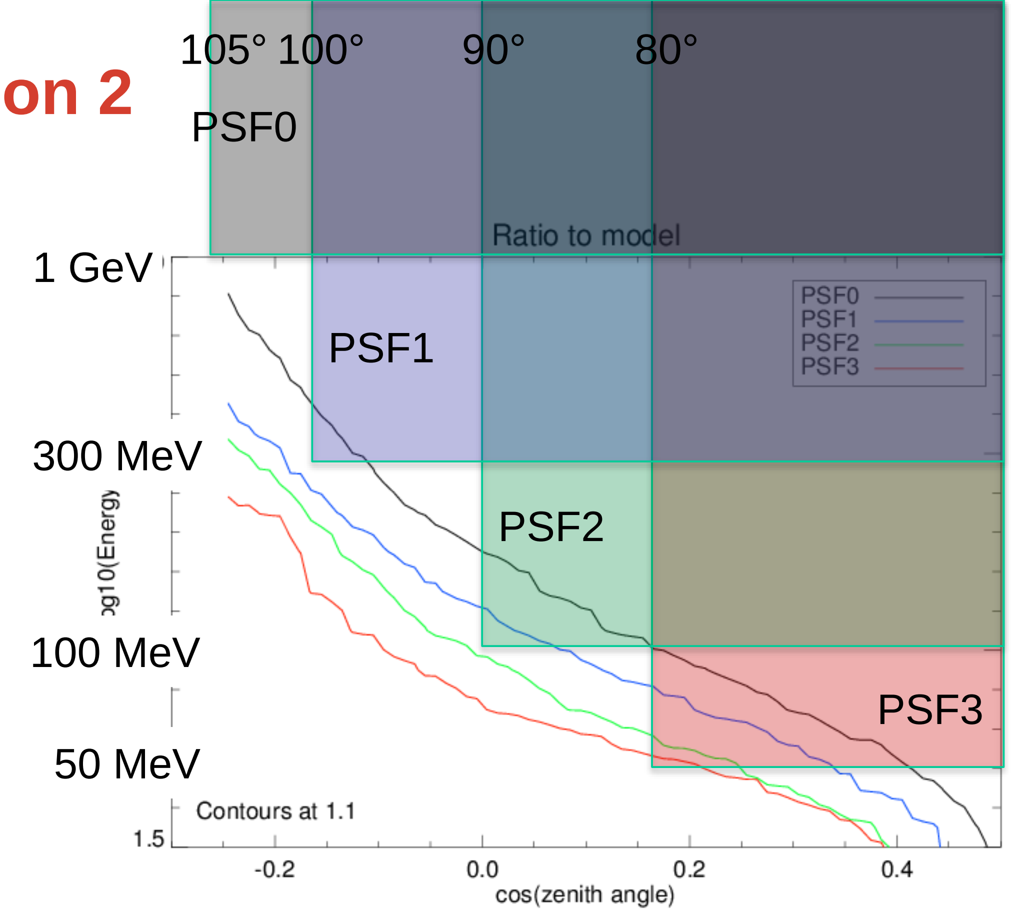
Earth limb suppression 2

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 when Earth limb contribution gets larger than 10% of regular photons at that zenith angle (much less after integrating on zenith angle)



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°

Earth limb suppression 3

The selection described in the previous slide, used since 4FGL, is **more restrictive than the FSSC recommendation** (Zmax = 90° for all events)

https://fermi.gsfc.nasa.gov/ssc/data/analysis/documentation/Cicerone/Cicerone_Data_Exploration/Data_preparation.html

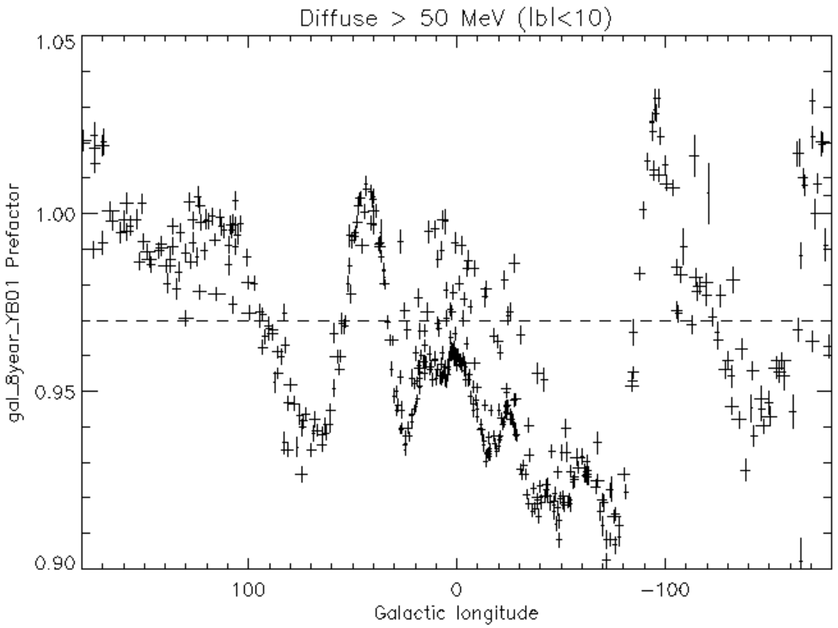
The gradual event type / zenith angle selection is immune to the Earth limb contamination, but entails a number of consequences:

- Changing Zmax with energy results in **slightly different time intervals**
- Using files extracted over narrow energy bands is more sensitive to the edisp_bins setting for energy dispersion (set to –2 for 10 bins per decade)
- It increases the number of files (but not the overall volume, since the same distinction can be used to optimize ROI and pixel size)

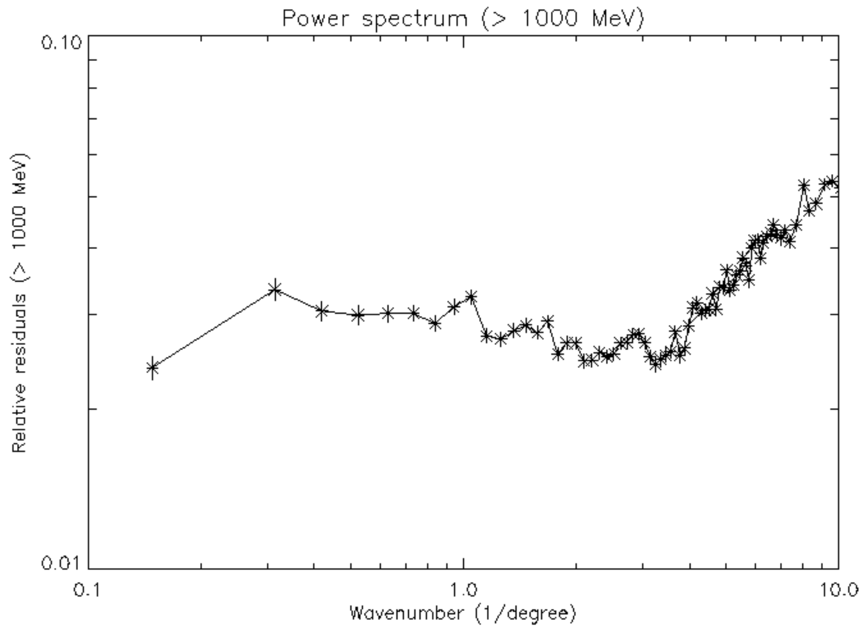
Diffuse systematics

Characterized in the 3FGL paper to be at the level of **3% both spectrally and spatially** (see also the weights page)

The diffuse model changed at 4FGL but the **residuals are at the same level**



10 years, diffuse normalization in each ROI along the Galactic plane



4 years, ratio of residuals to total in Fourier space in large sky region

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

Equivalent to reducing n_i and M_i by the same w_i factor

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 ε = 3 %

Appendix B of the 4FGL paper

Weighted logLikelihood 2

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

Now how to define B_i ?

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

Integral under PSF (for point sources)

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

Integral above current energy
E_{max} set to 2 E_i

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

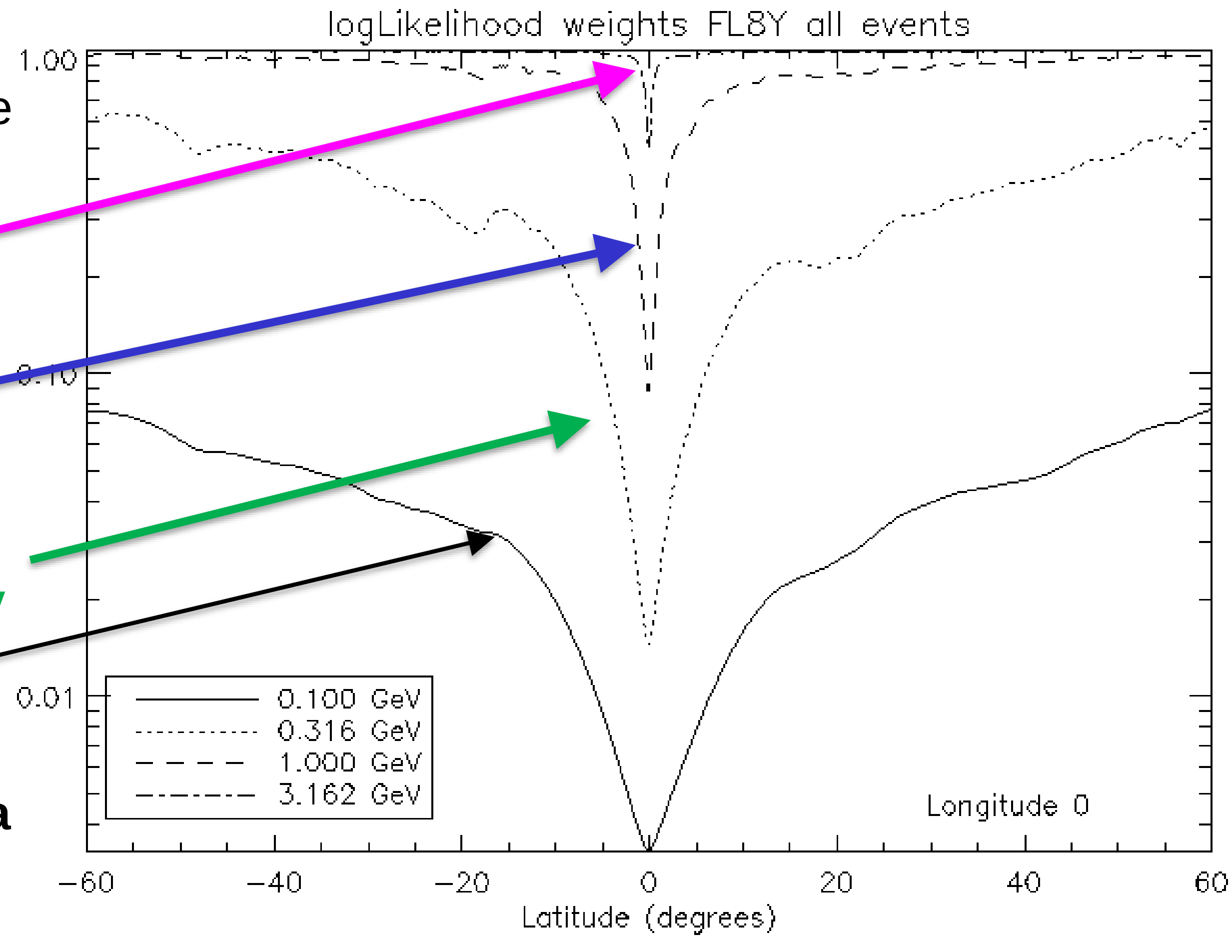
$R_{68}(E)^2$ decreases as $E^{-1.6}$ up to 3 GeV so the B_i term decreases very fast

The **weights increase fast with energy**

Chose to define the **background from the total observed counts**, not only the diffuse emission. Allows accounting for uncertainties on the bright sources

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**



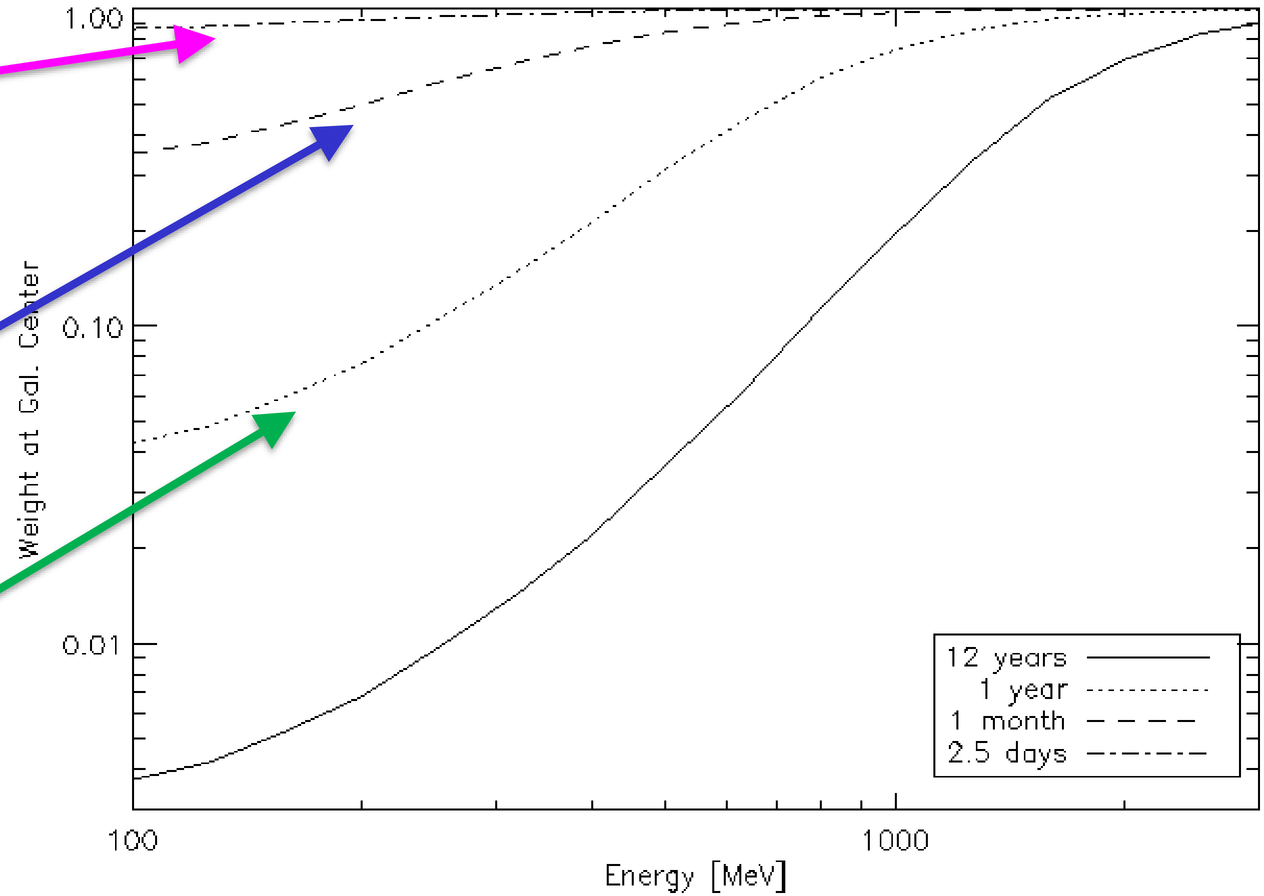
P8 all events, 8 years

Data-based weights, ε = 3 %

Longitude 0 (through Gal center)

Weights with time

- At a few days, systematics are negligible even at GC
- At 1 month, important up to 300 MeV in Gal Ridge but negligible elsewhere
- At 1 year, important over most of the Gal plane, similar at |b| = 20° to 1 month at GC



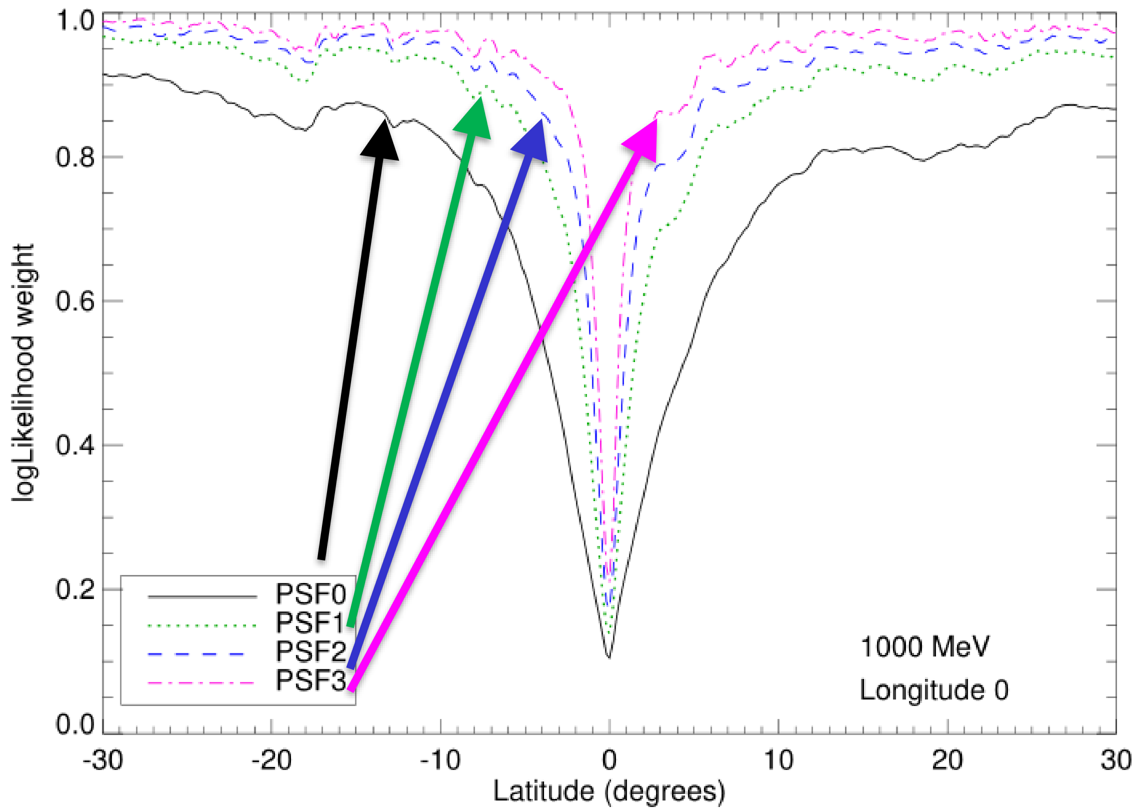
All events, 12 years, at Gal center

$\boldsymbol{B_i}$ is simply proportional to exposure

Data-based weights, ε = 3 %

Weighted log-likelihood

- PSF0 (worst PSF) more attenuated, and over a broader latitude range, than the others
- How to weight event types when they are **used together**, (summed likelihood) starting from the weights applicable to each type **used separately**
- Add factor α such that the correct weights are recovered if all event types are the same



$$w_k = \frac{\alpha}{1 + \varepsilon^2 N_k} \quad N_{\min} = \min_k N_k$$

$$K_{\mathrm{tot}} = \sum_k \left(\frac{N_{\min}}{N_k}\right)^2 \quad \alpha = \frac{1 + \varepsilon^2 N_{\min}}{1 + \varepsilon^2 N_{\min} K_{\mathrm{tot}}}$$

Pass 8, 8 years

Data-based weights, ε = 3 %

Computing the weights

Implemented by Eric Charles in the Fermi Tools

It combines three executables:

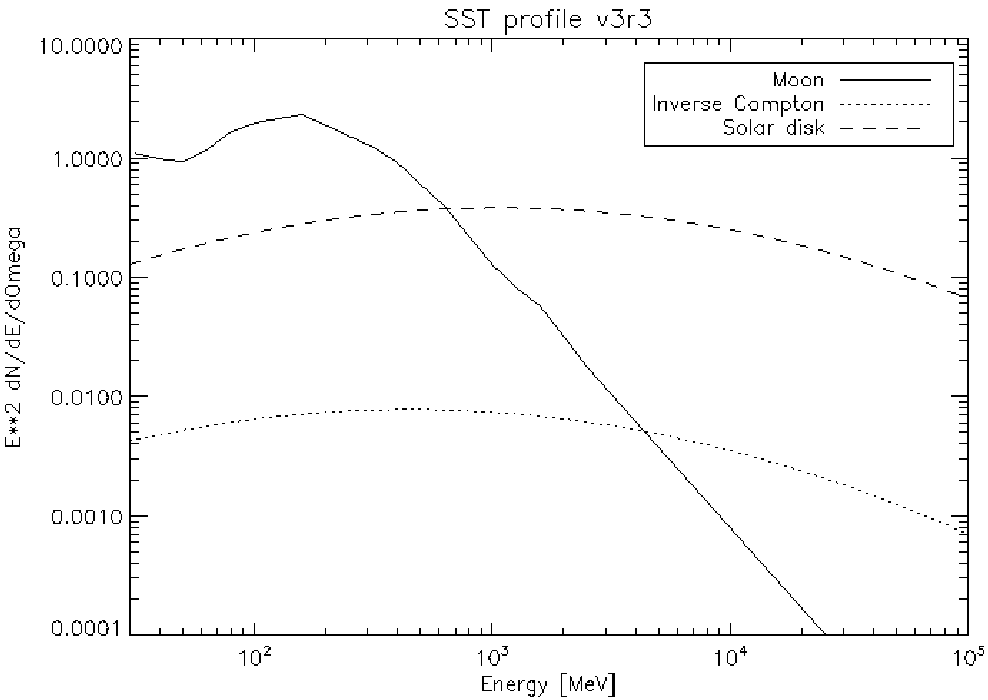
- *gteffbkg* (get the suitably integrated $\boldsymbol{B_i}$)
- *getalphabkg* (get the α when several components)
- *gtwtsmap* (gets the weights themselves)

I recommend computing the **weights over the entire sky** once and for all (in HealPix). In any case, like the exposure, they must be computed beyond the data map (the sky does not stop at the data map)

When using several components, requires a weights file for each

Sun and Moon model

- ✓ The **Moon** is quite bright at low energy
- ✓ The Sun and Moon model must be computed over the **same period as the data** (the Sun moves over the year and its coverage varies a lot since the Solar Array Drive Anomaly).
- ✓ Use G. Johannesson's Solar System Tools for this
- ✓ This works only for the quiet Sun, must eliminate solar flares as well (via GTI)



SunMoon models for 8, 10 and 12 years are available at the FSSC in the Catalog pages

Diffuse parameters

Nothing special at low energy

Standard approach involves

- Free normalization of the isotropic
- Power-law modulation (norm and index) of the Galactic
- Possibly priors (see Sect 3.3 and Table 3 of the DR3 paper) to prevent the minority contributor to reach unphysical values

Correlations with nearby sources

Fitting a source properly requires accounting for **correlations with neighbors**

Strong up to R_{68}, still present at R_{95} (strong sources)

At 100 MeV, R_{68}=3.2° R_{95}=7.1° even with PSF3 event type selection

There are always **many sources within that distance**

Bright sources even further away can impact the fit of the diffuse parameters

This requires knowing the nearby sources in the first place

It is **very difficult to detect new sources in LAT data below 1 GeV**.
Better rely on higher energies, variability, MWL info or existing catalogs

Cannot assume that the sources are as in the catalog (**most are variable**)

Leave most source normalizations free, spectral index for the bright ones

Faint hard sources can be ignored at low energy

Sources outside the ROI must be fixed (ideally optimized in specific ROI)

RoI extraction

Extract data at least 7° (PSF3 R_{95} at 100 MeV) beyond region of interest (to cover the PSF well)

Do not go beyond 20x20° map (just more complications)

The catalog process uses a **complex map definition** (item 5 of Sect 3.2 and Table 2 of the 4FGL paper), in which both the map size and the pixel size change for each component

This is **unnecessary** when working only below 1 GeV (a pixel size of 0.2° is small enough, so the files will never get very big)

It is **better to use the same data map at all energies**, also allowing to fit reliably the spectral parameters of sources over the entire map if need be

Consider background sources at least 7° (same as above) beyond the data map

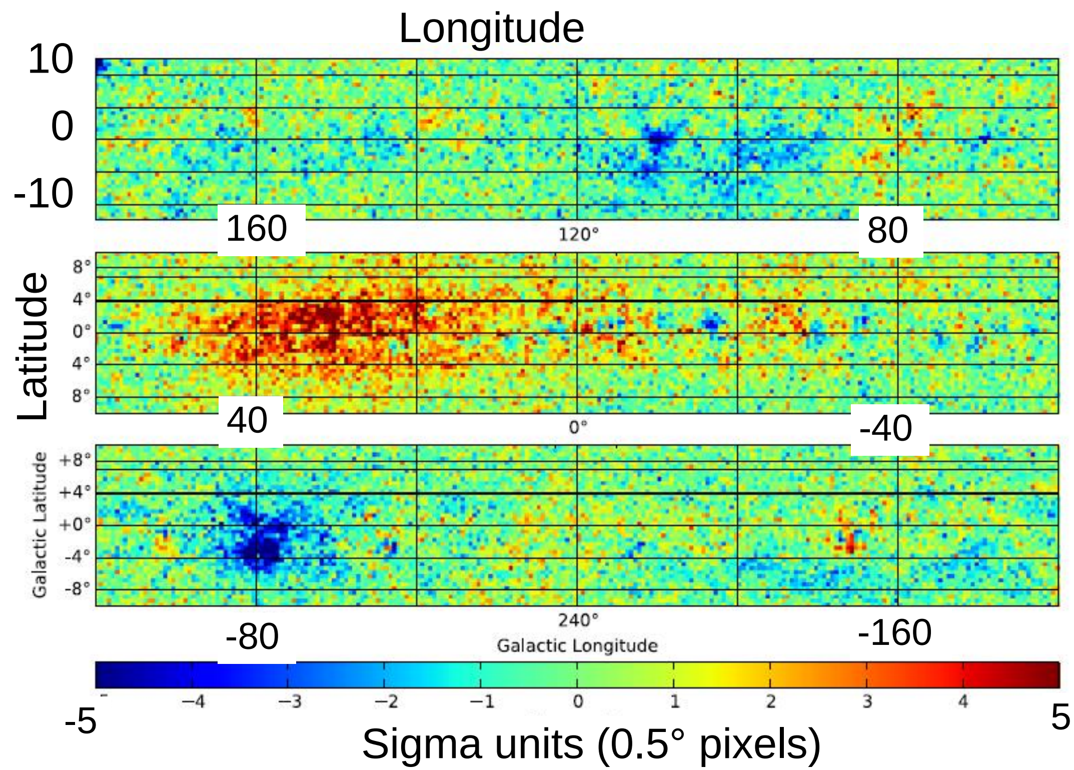
Beware of faint soft power-law sources in the catalog, which can be overestimated at 100 MeV (if their real spectrum is curved)

Check residuals

For large areas of the sky (such as the Galactic plane), can directly build residuals by comparing model (including all sources) and data

Can multiply by weights directly (at cube level)

A more powerful tool is P. Bruel's **PS maps**. See his paper and get the gtpsmap.py tool at the FSSC



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

4 years of **P7Rep** data, 3033 sources

0.5° pixels

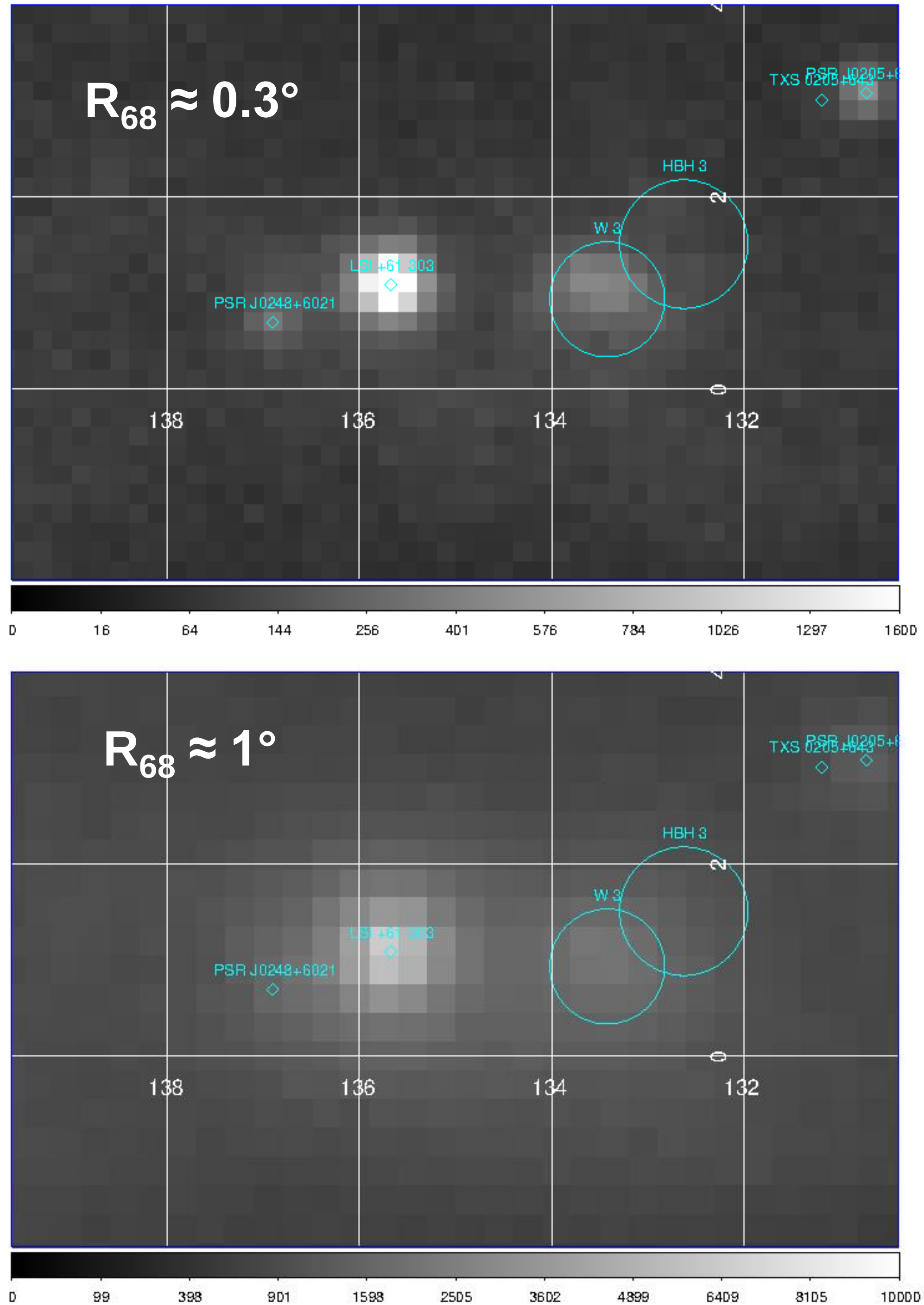
Effect of weights

LS I +61 303

Bright γ-ray binary in outer Galactic plane

glon = 135.7°

close to the **W3** nebula touching the **HB3** SNR



Effect of weights 2

Without weights:

- LS I: TS = 104,450; β = 0.090+/-0.007
- W3: TS = 3,267; β = 0.52+/-0.05

With weights:

- LS I: TS = 16,580; β = 0.09+/-0.02
- W3: TS = 1,259; β = 0.58+/-0.19