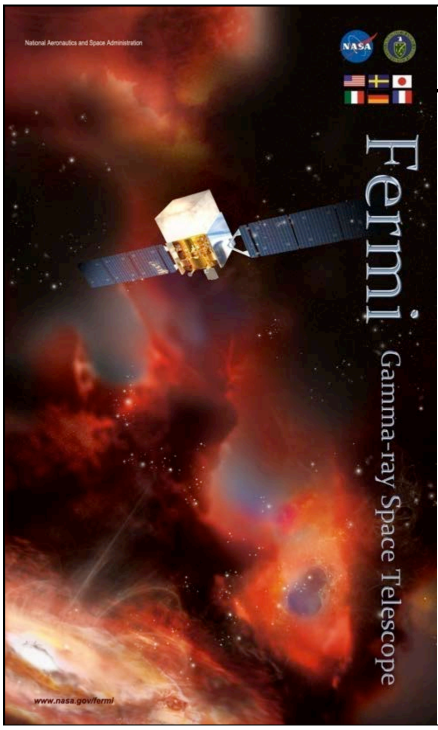


Tutorial on Fermi-LAT data analysis
the case of 3c454.3

Francesco Longo
(University of Trieste and INFN Trieste)

Material from P.Bruel, M.Razzano, S.Buson and R.Desiante

PhD course 2017 – Padova

Outline

- Overview of the Fermi Large AreaTelescope
  - How it works
  - LAT data
  - LAT performance
- Fermi Science Tools
  - General Introduction
- Maximum Likelihood Overview
  - Source modeling
- One study case:
  - 3c454.3: analysis tutorial

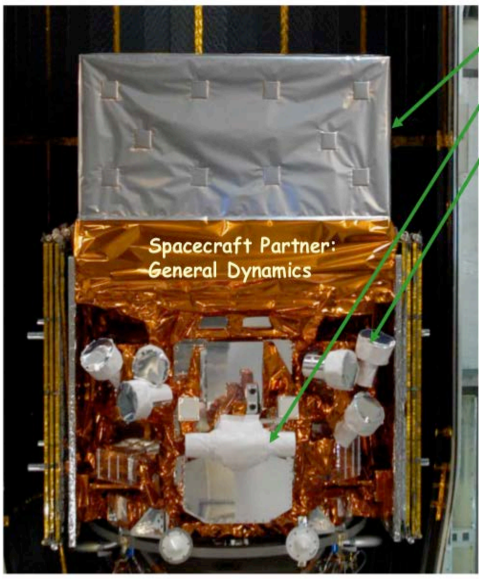
2

Organization

- **Today**
  - Introduction to the LAT data analysis
  - LAT data introduction
  - LAT data exploration
  - LAT data preparation
  - Preparation of Sky models

- **Tomorrow**
  - Likelihood analysis of LAT data
  - Hands-on on an Extragalactic source
    - General LAT analysis with gtlike
    - Light curve with aperture photometry
    - SED (possibly)

The observatory



Large AreaTelescope (LAT)
20 MeV - >300 GeV

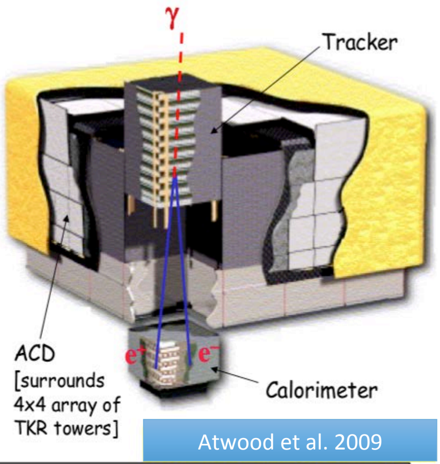
Gamma-ray Burst Monitor (GBM)
NaI and BGO Detectors
8 keV - 30 MeV

KEY FEATURES

- Huge field of view
  - –LAT: 20% of the sky at any instant; in sky survey mode, expose all parts of sky for ~30 minutes every 3 hours. GBM: whole unocculted sky at any time.
- Huge energy range, including largely unexplored band 10 GeV - 100 GeV. Total of >7 energy decades!
- Large leap in all key capabilities. Great discovery potential.

How the LAT works

- <u>Precision Si-strip Tracker (TKR)</u> Measure the photon direction; gamma ID.
- <u>Hodoscopic CsI Calorimeter (CAL)</u> Measure the photon energy; image the shower.
- <u>Segmented Anticoincidence Detector (ACD)</u> Reject background of charged cosmic rays; segmentation removes self-veto effects at high energy.
- <u>Electronics System</u> Includes flexible, robust hardware trigger and software filters.

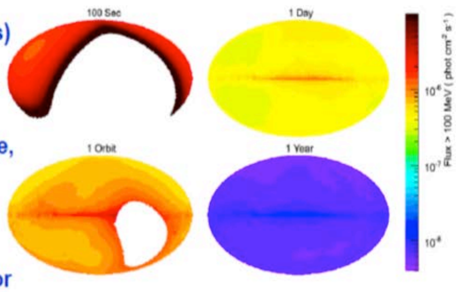


<u>Systems work together to identify and measure the flux of cosmic gamma rays with energy 20 MeV - >300 GeV.</u>

5

Operating Mode

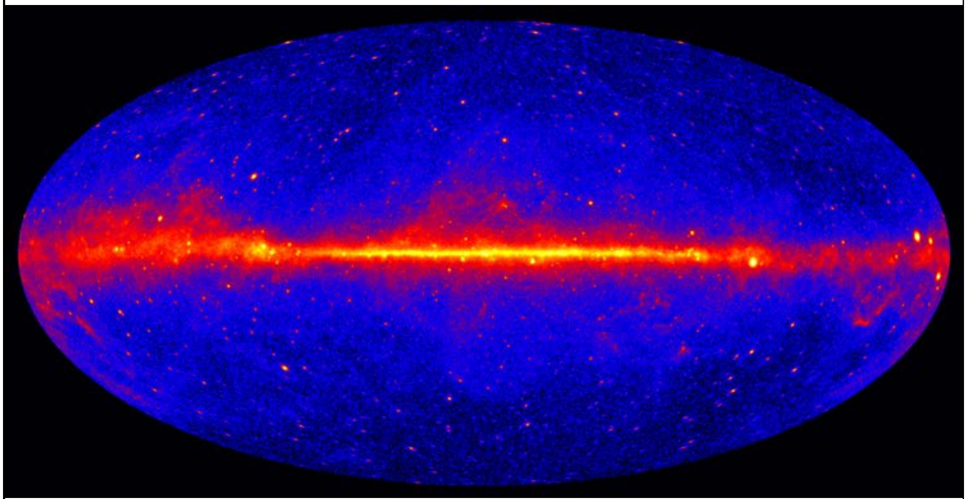
- Primary observing mode is Sky Survey
  - Full sky every 2 orbits (3 hours)
  - Uniform exposure, with each region viewed for ~30 minutes every 2 orbits
  - Best serves majority of science, facilitates multiwavelength observation planning
  - Exposure intervals commensurate with typical instrument integration times for sources
  - EGRET sensitivity reached in days



- Pointed observations when appropriate (selected by peer review in later years) with automatic earth avoidance selectable. Target of Opportunity pointing.
- Autonomous repoints for onboard GRB detections in any mode.

6

The Fermi Sky



7

What do you need for the analysis

LAT DATA ARE PUBLIC!!

- **Data ...of course!**
  - LAT detected events
  - Spacecraft related stuff
  - Extras (Backgrounds , catalog sources, timing..)
- **Fermi Science Tools**
  - http://fermi.gsfc.nasa.gov/ssc/data/analysis/software/
- **Other ancillary tools**
  - ftools, HEASOFT, DS9 etc..
  - http://heasarc.gsfc.nasa.gov/docs/software.html
- **Lots and lots of scripts!**
  - Fermi tools are already scriptable
  - You can also use your favourite scripting language... but ..
  - Science Tools can be imported as Python modules!

8

LAT data analysis references

Fermi Data Analysis: starting points

Fermi Science Support Center: http://fermi.gsfc.nasa.gov/ssc/

Fermi Newsletters: http://fermi.gsfc.nasa.gov/ssc/resources/newsletter/

Fermi Data Access: http://fermi.gsfc.nasa.gov/cgi-bin/ssc/LAT/LATDataQuery.cgi

Fermi Science Tools Reference Manual:
http://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/references.html

Fermi Analysis Threads:
http://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/
http://fermi.gsfc.nasa.gov/ssc/data/analysis/documentation/Cicerone/

Fermi - LAT Likelihood Algorithm description
http://fermi.gsfc.nasa.gov/ssc/data/analysis/documentation/Cicerone/Cicerone_Likelihood/
Cash W. 1979, ApJ 228, 939
Mattox J. R. et al 1996, ApJ 461, 396
Protassov et al. 2002, ApJ 57, 545

LAT Performance Page: http://www-glast.slac.stanford.edu/software/IS/glast_lat_performance.htm
The Large Area Telescope on the Fermi Gamma-Ray Space Telescope Mission, W.B. Atwood, et. al., ApJ, 2009, 695, 1071.
The On-orbit Calibrations for the Fermi Large Area Telescope, A.A. Abdo, et al. arXiv:0904.2226v1

How to access LAT Data

http://fermi.gsfc.nasa.gov/ssc/data/analysis/

LAT Data

•**LAT data products can be downloaded by the FSSC website**

–**LAT Data server** http://fermi.gsfc.nasa.gov/cgi-bin/ssc/LAT/LATDataQuery.cgi

–**Archive of weekly files**

–ftp://legacy.gsfc.nasa.gov/fermi/data/

•**Two main data products (stored in FITS format)**

–**Events file (FT1)**

–**i. e. “what the LAT sees”**

– (photons, their energy, coordinates, time, event classes etc..)

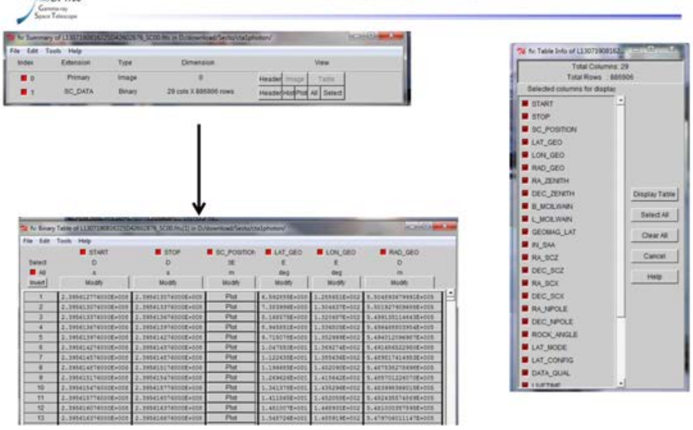
–**Spacecraft files (FT2)**

–**i. e. “where the LAT is”**

– (position, angles..)

FT1: photons, photons...

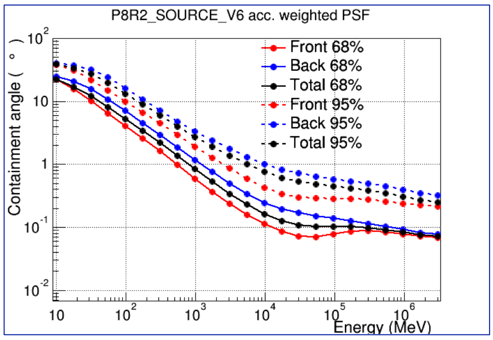
FT2: where is Fermi ?



spacecraft Data

Extras - IRFs

- •LAT Instrument Response Functions (IRFs)
  - –Point Spread Function (PSF)
  - –Effective Area
  - –Energy Resolution



- •Highly dependent on energy
- •and arrival direction of incident
- •photon

- •Fundamental for analysis!!

- •http://www.slac.stanford.edu/exp/glast/groups/canda/lat_Performance.htm

Pass8 (?)...

New Pass8 data

http://fermi.gsfc.nasa.gov/ssc/data/analysis/documentation/Pass8_usage.html

Pass 8 introduction

What does Pass mean?

- Each pass corresponds to a version of the Fermi LAT data
- It implies a whole package:
  - Intrument simulation
  - Reconstruction code
  - Event selection
  - Instrument Response Functions (IRFs)
  - Systematic uncertainties
  - Isotropic template (which includes the cosmic-ray residual background)
  - And sometimes more (Galactic diffuse model, Earth limb template, Sun+Moon template)
- It's only when we have validated the whole package that we can release it to the public.

Pass8 introduction

From Pass 6 to Pass 8

- Pass 6 (launch time)
  - Pass 6 reconstruction
  - Pass 6 selection
  - Based on pre-launch instrument simulation
  - First data revealed the issue of out-of-time pile-up (aka ghosts)
  - New: instrument simulation with ghosts -> correct IRFs
- Pass 7
  - Pass 6 reconstruction
  - New: Pass 7 selection optimized with simulations with ghosts
- Pass 8
  - New: improved instrument simulation
  - New: Pass 8 reconstruction, as ghost-proof as possible
  - New: Pass 8 selection

Pass8 Introduction

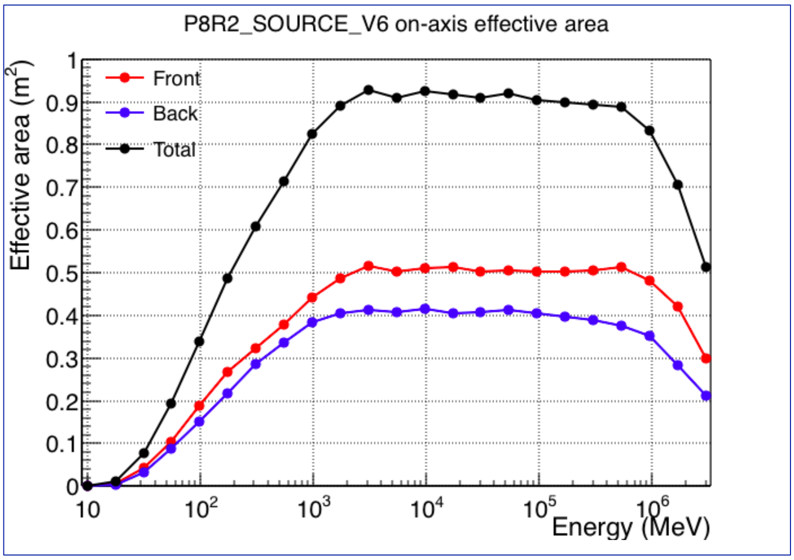
Pass 8 improvements

- Ghost handling
  - Tracker: ignoring ghost hits
  - ACD: partial deghosting
  - Calorimeter: clustering and cluster classification
- Improved direction measurement
  - Tree-based track finder
- Improved energy measurement
  - Extension of the energy range: from ~10 MeV to ~3 TeV
- Improved track/ACD matching information
  - Using the uncertainty of the tracker direction
- Improved event selection
  - Using the ROOT TMVA package (tmva.sourceforge.net)
- Additional sub-classes of events

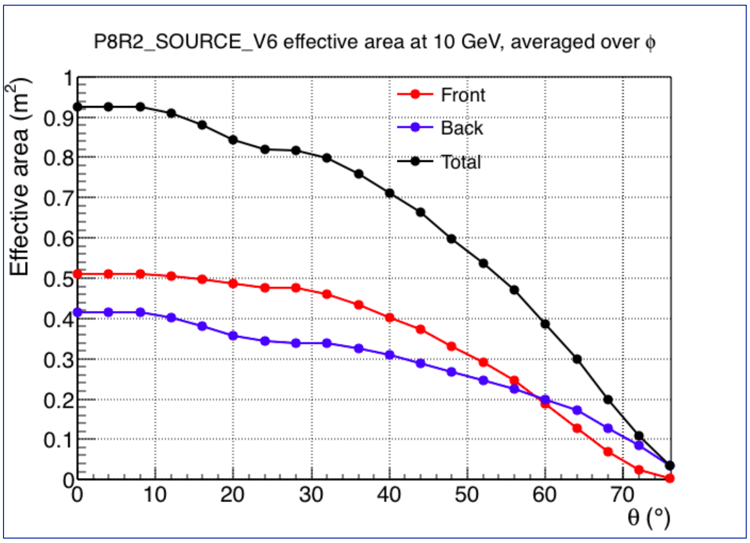
Pass8 performance

http://www.slac.stanford.edu/exp/glast/groups/canda/lat_Performance.htm

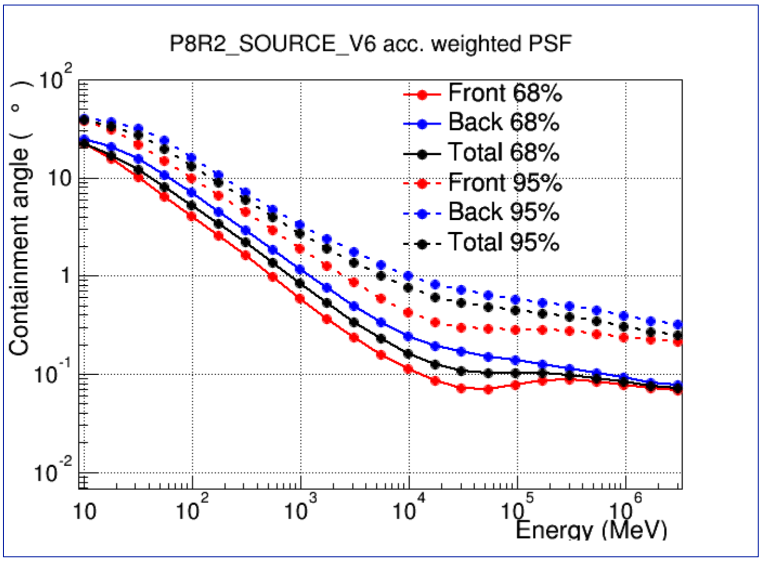
Effective Area



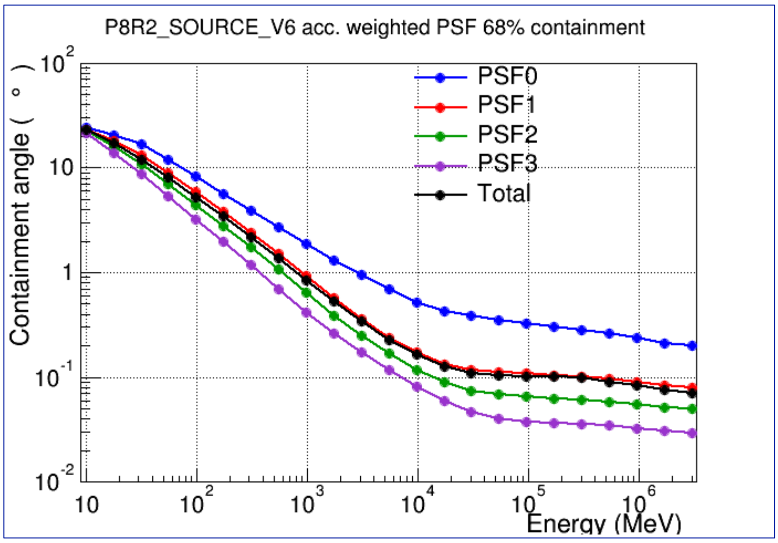
Pass8 Effective Area



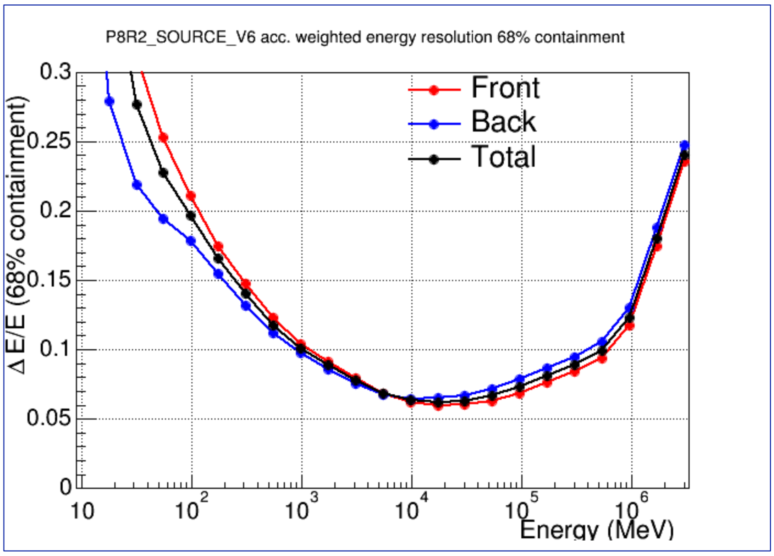
Point Spread Function



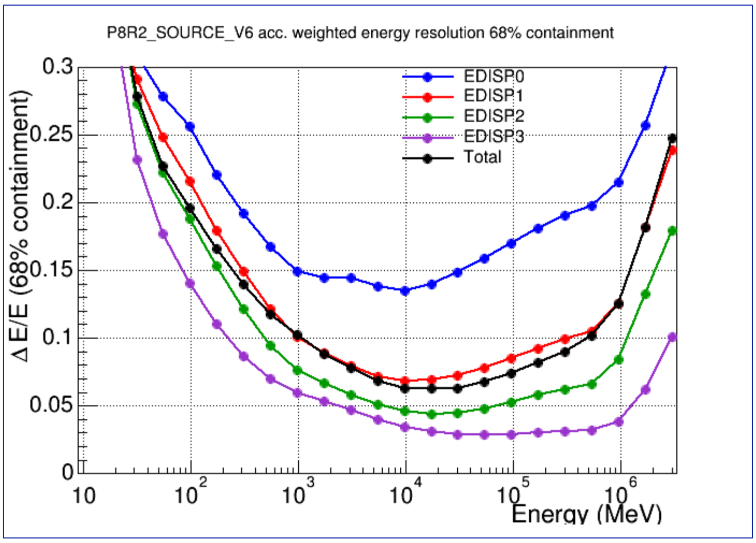
Point Spread Function



Energy resolution



Energy resolution



Extras

- **Diffuse models (.txt & FITS files)**
  - **To correctly take into accounts the galactic and extagalactic backgrounds**
  - http://fermi.gsfc.nasa.gov/ssc/data/access/lat/BackgroundModels.html

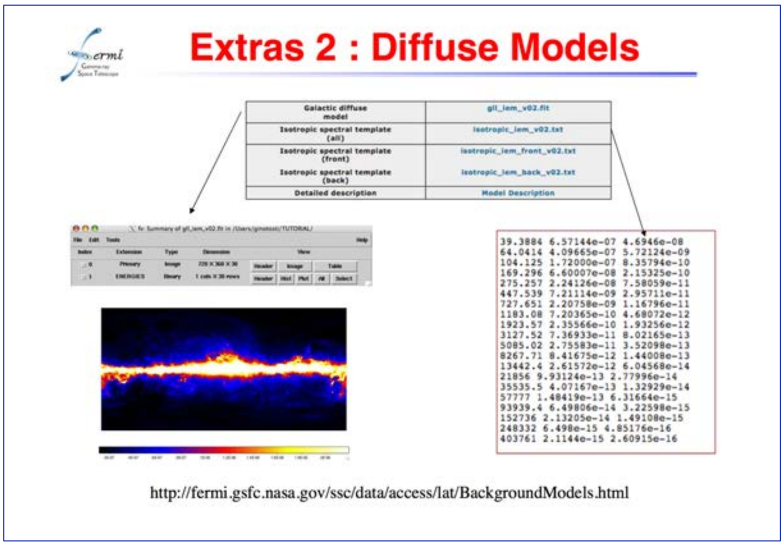
- **Source Catalogs (3FGL, 3FHL, PSRs, GRB, SNR …)**

- **Region of Interest model definition (stored in XML files)**
  - More on this later …

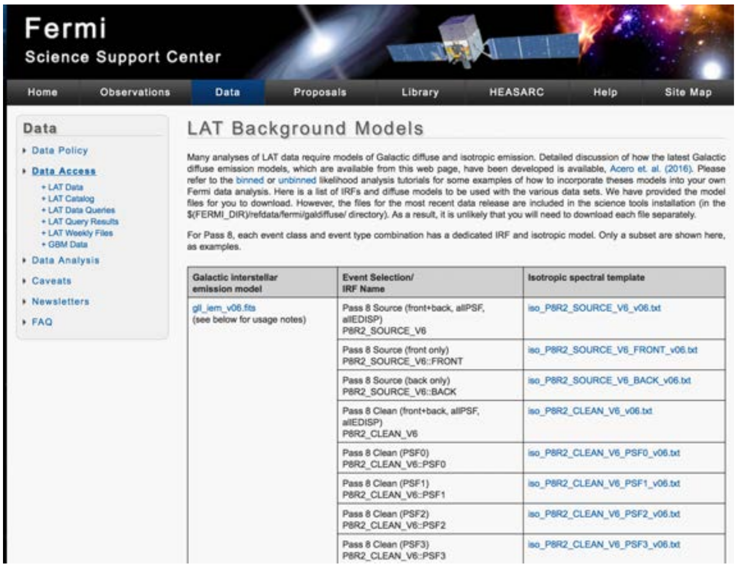
- **Choose the proper combination of**
  - event classes based on the probability of being a photon and background level
    - TRANSIENT (for very short events)
    - SOURCE (suggested for source analysis)
    - CLEAN and ULTRACLEAN (lowest particle contamination – for diffuse sources analysis)
  - Event types based on conversion point or PSF or Energy reconstruction

36

LAT background models



LAT background models



Fermi
Science Support Center

Home | Observations | Data | Proposals | Library | HEASARC | Help | Site Map

Data
- Data Policy
- Data Access
  - LAT Data
  - LAT Catalog
  - LAT Data Queries
  - LAT Query Results
  - LAT Weekly Files
  - GBM Data
- Data Analysis
- Caveats
- Newsletters
- FAQ

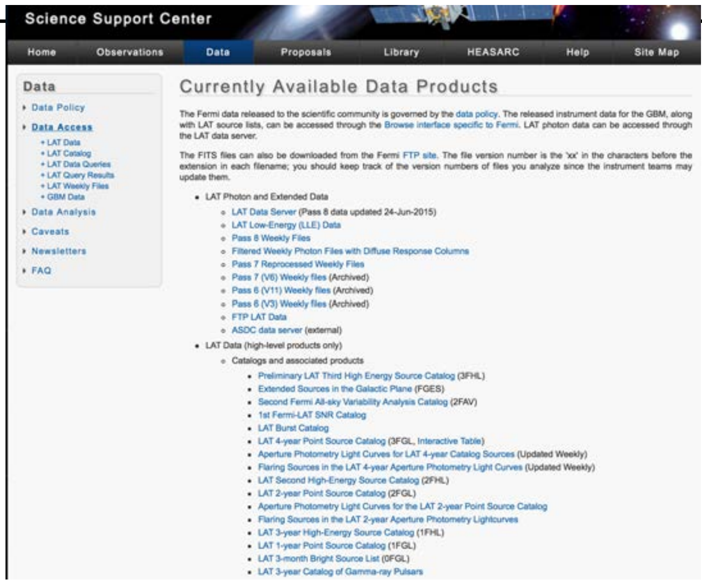
LAT Background Models

Many analyses of LAT data require models of Galactic diffuse and isotropic emission. Detailed discussion of how the latest Galactic diffuse emission models, which are available from this web page, have been developed is available, Acero et. al. (2016). Please refer to the binned or unbinned likelihood analysis tutorials for some examples of how to incorporate theses models into your own Fermi data analysis. Here is a list of IRFs and diffuse models to be used with the various data sets. We have provided the model files for you to download. However, the files for the most recent data release are included in the science tools installation (in the $(FERMI_DIR)/refdata/fermi/galdiffuse/ directory). As a result, it is unlikely that you will need to download each file separately.

For Pass 8, each event class and event type combination has a dedicated IRF and isotropic model. Only a subset are shown here, as examples.

Galactic interstellar emission model	Event Selection/ IRF Name	Isotropic spectral template
gll_iem_v06.fits (see below for usage notes)	Pass 8 Source (front+back, allPSF, allEDISP) P8R2_SOURCE_V6	iso_P8R2_SOURCE_V6_v06.txt
	Pass 8 Source (front only) P8R2_SOURCE_V6::FRONT	iso_P8R2_SOURCE_V6_FRONT_v06.txt
	Pass 8 Source (back only) P8R2_SOURCE_V6::BACK	iso_P8R2_SOURCE_V6_BACK_v06.txt
	Pass 8 Clean (front+back, allPSF, allEDISP) P8R2_CLEAN_V6	iso_P8R2_CLEAN_V6_v06.txt
	Pass 8 Clean (PSF0) P8R2_CLEAN_V6::PSF0	iso_P8R2_CLEAN_V6_PSF0_v06.txt
	Pass 8 Clean (PSF1) P8R2_CLEAN_V6::PSF1	iso_P8R2_CLEAN_V6_PSF1_v06.txt
	Pass 8 Clean (PSF2) P8R2_CLEAN_V6::PSF2	iso_P8R2_CLEAN_V6_PSF2_v06.txt
	Pass 8 Clean (PSF3) P8R2_CLEAN_V6::PSF3	iso_P8R2_CLEAN_V6_PSF3_v06.txt

LAT catalogs



Science Support Center

Home | Observations | Data | Proposals | Library | HEASARC | Help | Site Map

Data
- Data Policy
- Data Access
  - LAT Data
  - LAT Catalog
  - LAT Data Queries
  - LAT Query Results
  - LAT Weekly Files
  - GBM Data
- Data Analysis
- Caveats
- Newsletters
- FAQ

Currently Available Data Products

The Fermi data released to the scientific community is governed by the data policy. The released instrument data for the GBM, along with LAT source lists, can be accessed through the Browse interface specific to Fermi. LAT photon data can be accessed through the LAT data server.

The FITS files can also be downloaded from the Fermi FTP site. The file version number is the 'xx' in the characters before the extension in each filename; you should keep track of the version numbers of files you analyze since the instrument teams may update them.

- LAT Photon and Extended Data
  - LAT Data Server (Pass 8 data updated 24-Jun-2015)
  - LAT Low-Energy (LLE) Data
  - Pass 8 Weekly Files
  - Filtered Weekly Photon Files with Diffuse Response Columns
  - Pass 7 Reprocessed Weekly Files
  - Pass 7 (V6) Weekly files (Archived)
  - Pass 6 (V11) Weekly files (Archived)
  - Pass 6 (V3) Weekly files (Archived)
  - FTP LAT Data
  - ASDC data server (external)
- LAT Data (high-level products only)
  - Catalogs and associated products
    - Preliminary LAT Third High Energy Source Catalog (3FHL)
    - Extended Sources in the Galactic Plane (FGES)
    - Second Fermi All-sky Variability Analysis Catalog (2FAV)
    - 1st Fermi-LAT SNR Catalog
    - LAT Burst Catalog
    - LAT 4-year Point Source Catalog (3FGL, Interactive Table)
    - Aperture Photometry Light Curves for LAT 4-year Catalog Sources (Updated Weekly)
    - Flaring Sources in the LAT 4-year Aperture Photometry Light Curves (Updated Weekly)
    - LAT Second High-Energy Source Catalog (2FHL)
    - LAT 2-year Point Source Catalog (2FGL)
    - Aperture Photometry Light Curves for the LAT 2-year Point Source Catalog
    - Flaring Sources in the LAT 2-year Aperture Photometry Lightcurves
    - LAT 3-year High-Energy Source Catalog (1FHL)
    - LAT 1-year Point Source Catalog (1FGL)
    - LAT 3-month Bright Source List (0FGL)
    - LAT 3-year Catalog of Gamma-ray Pulsars

https://fermi.gsfc.nasa.gov/ssc/data/access/

Event types

P8R2 Event Type Name	Event Type Partition	Event Type Value (evtype)
FRONT	Conversion Type	1
BACK	Conversion Type	2
PSF0	PSF	4
PSF1	PSF	8
PSF2	PSF	16
PSF3	PSF	32
EDISP0	EDISP	64
EDISP1	EDISP	128
EDISP2	EDISP	256
EDISP3	EDISP	512

Event classes

P8R2 IRF name	Event Class (evclass)	Class Hierarchy	Photon File	Extended File
P8R2_ULTRACLEANVETO_V6	1024	Standard	X	X
P8R2_ULTRACLEAN_V6	512	Standard	X	X
P8R2_CLEAN_V6	256	Standard	X	X
P8R2_SOURCE_V6	128	Standard	X	X
P8R2_TRANSIENT010_V6	64	Standard		X
P8R2_TRANSIENT020_V6	16	Standard		X
P8R2_TRANSIENT010E_V6	64	Extended		X
P8R2_TRANSIENT020E_V6	8	Extended		X
P8R2_TRANSIENT015S_V6	65536	No-ACD		X

Recommendations

Event Selection Recommendations (P8R2)

Analysis Type	Minimum Energy (emin)	Maximum Energy (emax)	Max Zenith Angle (zmax)	Event Class (evclass)	IRF Name
Galactic Point Source Analysis	100 (MeV)	500000 (MeV)	90 (degrees)	128	P8R2_SOURCE_V6
Off-plane Point Source Analysis	100 (MeV)	500000 (MeV)	90 (degrees)	128	P8R2_SOURCE_V6
Burst and Transient Analysis (<200s)	100 (MeV)	500000 (MeV)	100 (degrees)	16	P8R2_TRANSIENT020_V6
Galactic Diffuse Analysis	100 (MeV)	500000 (MeV)	90 (degrees)	128	P8R2_SOURCE_V6
Extra-Galactic Diffuse Analysis	100 (MeV)	500000 (MeV)	90 (degrees)	1024	P8R2_ULTRACLEANVETO_V6
Impulsive Solar Flare Analysis	100 (MeV)	500000 (MeV)	100 (degrees)	65536	P8R2_TRANSIENT015S_V6

Caveats

Caveats About Analyzing LAT Pass 8 Data

These caveats are relevant for the P8R2 version of the Pass 8 photon dataset. They are an updated version of previous sets of caveats for analysis of Pass 7 reprocessed (P7REP), Pass7 (P7_V6) and Pass 6 (P6_V3 and P6_V11) event selections and Instrument Response Functions (IRFs).

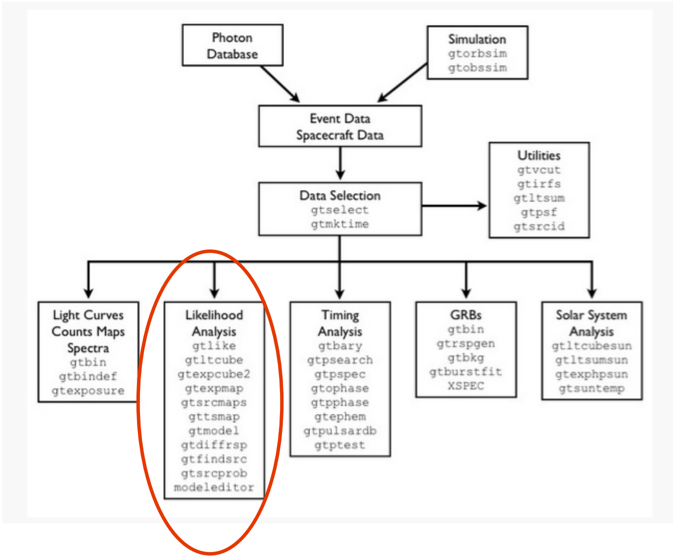
The LAT team is still working to validate all aspects of Pass 8 data and analysis. As a consequence it is expected that, in the coming year, the range of application of Pass 8 data will be increased, the tools and files will be improved and the systematic uncertainties will be decreased. These caveats will be modified accordingly.

The P8R2_V6 IRFs are defined between 5.62 MeV and 3.16 TeV but that does not mean that they have been fully validated over this whole energy range. Following the Pass 7 validation effort, the LAT team has started performing studies in order to check the consistency and precision of the instrument simulation and the instrument response representation provided by the IRFs. These studies are based on the analysis of Vela ($E < 10$ GeV), bright AGN (3 GeV $< E <$ 100 GeV) and the Earth limb ($E > 10$ GeV). They include:

- comparing the distributions of reconstructed quantities between data and the simulation of the instrument for a given selection (e.g. SOURCE class)
- comparing the fraction of events of a loose selection (e.g. TRANSIENT020 class) that are accepted in a more stringent selection (e.g. SOURCE class) in data and as predicted by the instrument simulation or the IRFs
- comparing the fraction of events in each event type for a given selection in data and as predicted by the instrument simulation or the IRFs
- comparing the radial profiles of a point source in data and as predicted by the IRFs

http://fermi.gsfc.nasa.gov/ssc/data/analysis/LAT_caveats.html

Overview of Fermi Science Tools



44

Maximum Likelihood Overview

45

Maximum likelihood technique

Given a set of observed data:

→ Produce a model that accurately describes the data, including parameters that we wish to estimate

→ derive the probability (density) for the data given the model (PDF)

→ treat this as a function of the model parameters (likelihood function)

→ maximize the likelihood with respect to the parameters - ML estimation.

46

Maximum likelihood ingredients

Data: $X = \{x_i\} = \{x_1, x_2, ..., x_N\}$

Model parameters: $\Theta = \{\theta_j\} = \{\theta_1, \theta_2, ..., \theta_M\}$

Likelihood: $\mathcal{L}(\Theta|X) = P(X|\Theta)$

Conditional probability rule for independent events: $P(A, B) \underset{\text{CPR}}{=} P(A)P(B|A) \underset{\text{Independence}}{=} P(A)P(B)$

For independent data:

$$P(X|\Theta) = P(\{x_i\}|\Theta) = P(x_1|\Theta)P(x_2, .., x_N|\Theta) = \cdots$$
$$= P(x_1|\Theta)P(x_2|\Theta) \cdots P(x_N|\Theta) = \prod_i P(x_i|\Theta)$$

$$\mathcal{L}(\Theta|X) = \prod_i P(x_i|\Theta)$$

$\mathcal{L}$ is the product of the probability of observing the detected counts in each bin.

47

Maximum likelihood estimation

Parameters can be estimated by maximizing likelihood.

→ Easier to work with log-likelihood:

$$\ln \mathcal{L}(\Theta) = \ln \mathcal{L}(\Theta|X) = \sum_i \ln P(x_i|\Theta)$$

Estimates of $\{\hat{\theta}_k\}$ from solving simultaneous equations:

$$\left.\frac{\partial \ln \mathcal{L}}{\partial \theta_j}\right|_{\{\hat{\theta}_k\}} = 0$$

For one parameter, if we have: $\mathcal{L}(\theta) \sim e^{-\frac{(\theta-\hat{\theta})^2}{2\sigma_\theta^2}}$ **Gaussian approximation**

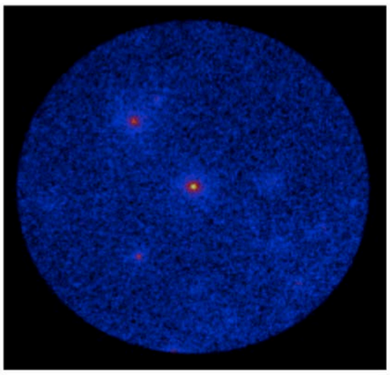
then: $$\left.\frac{\partial^2 \ln \mathcal{L}}{\partial \theta^2}\right|_{\hat{\theta}} = -\frac{1}{\sigma_\theta^2}$$

so 2nd derivative is related to "errors"

48

The Challenge

- Gamma-ray data is a list of counts (photons) reconstructed in the detector.
- Qualitative exploration of the data suggests the presence of sources (spatial clustering).
- Quantitative analysis requires evaluating the significance of a 'model' of our region.



Count Map

49

Fermi Gamma-ray Space Telescope

The Procedure

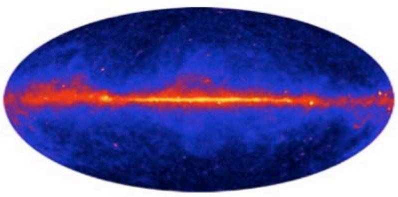
- Basically the initial '**model**' is a guess of the various parameters of the sources in our region:
  - location
  - spectral shape
  - flux
  - etc.
- The guess can be easier if you have a starting point like the 2FGL (3FGL).
- We quantify (using the Likelihood Method) the statistical significance of the model and vary the parameters to determine the most likely parameter values.

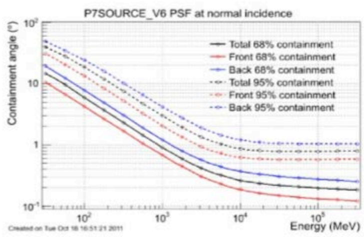
50

Fermi Gamma-ray Space Telescope

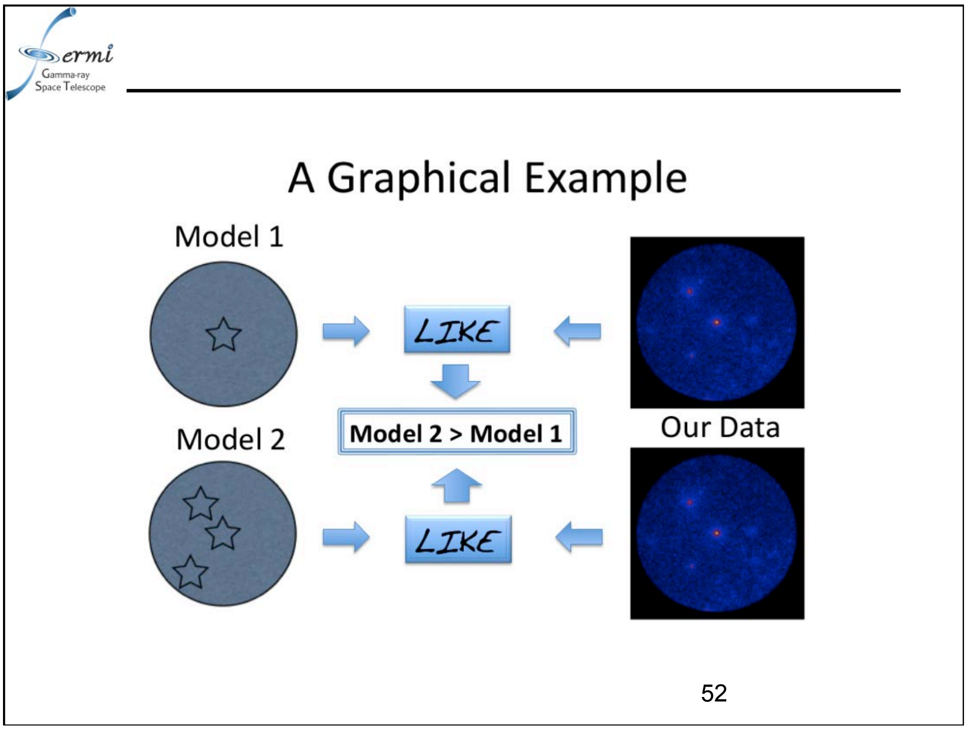
Why Model Fitting?

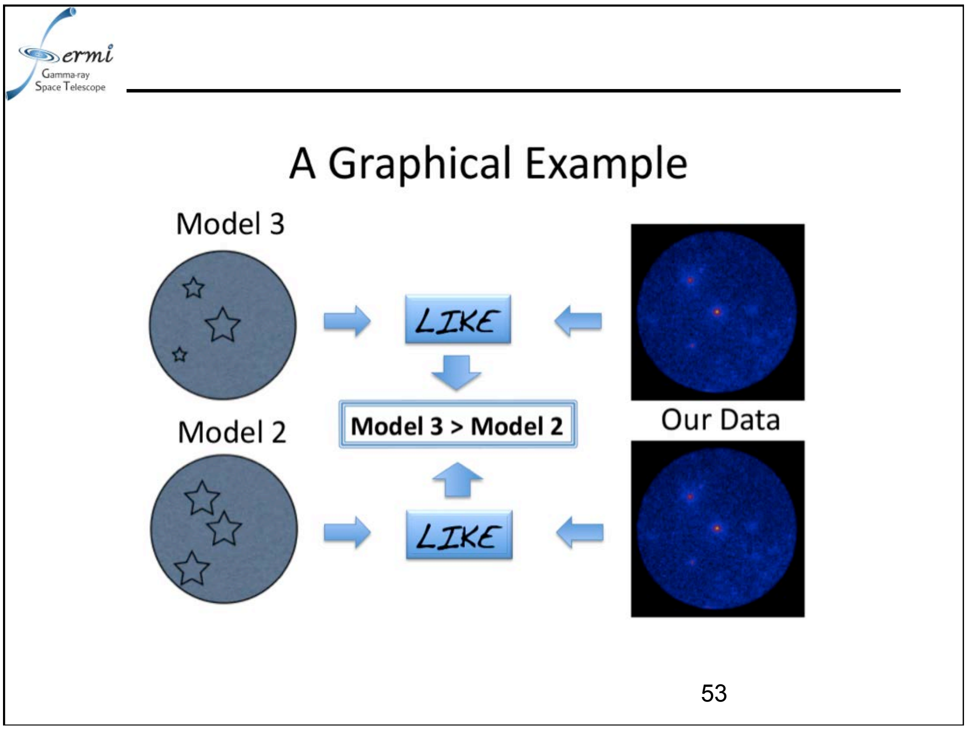
- We use the Likelihood method because the LAT data are limited by statistics, a bright diffuse background and a broad PSF.
- The model construction defines the questions we want to ask which means:
  - It will not answer a question you are not asking (ie. unknown parameters).
- The Likelihood will not tell you if a fit is 'good'. If the model does not represent your data well, the results will also not represent reality well.





51


Fermi
Gamma-ray
Space Telescope
A Graphical Example
Model 1
LIKE
Model 2 > Model 1
Our Data
Model 2
LIKE
52


Fermi
Gamma-ray
Space Telescope
A Graphical Example
Model 3
LIKE
Model 3 > Model 2
Our Data
Model 2
LIKE
53

Likelihood Analysis

- The likelihood $\mathcal{L}$ is the probability of obtaining your data given an input model.
- In our case, the input model is the distribution of gamma-ray sources on the sky and includes their intensity and spectra.
- One will maximize $\mathcal{L}$ to get the best match of the model to the data. Given a set of data, one can bin them in multidimensional (energy, sky pixels, ...) bins.
- $\mathcal{L}$ is the product of the probabilities of observing the detected counts in each bin.

$$\mathcal{L} = \Pi p_k$$

54

Write L as a function of the source model

- The source model is folded with the IRFs in order to obtain the predicted number of counts

- The IRFs can be decomposed into three funtions: Effective Area (proj area of the detector * efficency), Energy Dispesion, Point Spread Function

- Small number of counts in each bin --> Poisson distribuiton

- Bin size infinitesimally small

- Assume only steady source for standard analysis

The function to maximize is:

$$\log \mathcal{L} = \sum_j \log M(E'_j, \hat{p}'_j, t_j) - N_{\rm pred}$$

where the sum is performed over photons in the ROI. The predicted number of counts is $N_{\rm pred} = \int_{\rm ROI} dE' d\hat{p}' \, dt M(E', \hat{p}', t)$

55

Fermi Gamma-ray Space Telescope

Test Statistic

$$TS = -2 \log \frac{\mathcal{L}_0}{\mathcal{L}_1} \underset{N \to \infty}{\to} \chi^2_{m-h}$$

- $\mathcal{L}_0$: Null hypothesis max likelihood, h parameters
- $\mathcal{L}_1$: Alternative hypothesis max likelihood, m parameters
- $m-h$: non fixed parameters

- In the limit of a large number of counts, Wilk's Theorem states that the TS for the null hypothesis is asymptotically distributed as χ_n^2 where n is the number of parameters characterizing the additional source.
- As a basic rule of thumb, the square root of the TS is approximately equal to the detection significance for a given source.

56

Fermi Gamma-ray Space Telescope

Summarizing

- Observed a photon from a location, at a time, with an energy.
- Assume a model: $S(E,\hat{p},t) = \sum_i s_i(E,t)\delta(\hat{p}-\hat{p}_i) + S_G(E,\hat{p}) + S_{eg}(E,\hat{p}) + \sum_l S_l(E,\hat{p},t),$ (Point Sources; Galactic & EG diffuse Sources; Other Sources)
- Calculate the probability of that photon being detected assuming our model: $M(E',\hat{p}',t) = \int_{SR} dE d\hat{p} R(E',\hat{p}',t;E,\hat{p})S(E,\hat{p},t)$
- Calculate the total number of predicted counts assuming our model.

$$N_{pred} = \int_{ROI} dE' d\hat{p}'\, dt M(E',\hat{p}',t)$$

- Adjust the model until this is maximized:

$$\log \mathcal{L} = \sum_j \log M(E'_j, \hat{p}'_j, t_j) - N_{pred}$$

- Calculate the TS:

$$TS = -2 \log \frac{\mathcal{L}_0}{\mathcal{L}_1} \underset{N \to \infty}{\to} \chi^2_{m-h}$$

57

Fermi Gamma-ray Space Telescope

Keep in mind that ..

- Many variables may be calculated BEFORE selecting the models
- Very important to have a reliable model
- Absolute value of likelihood meaningless!
  - Likelihood function has no meaning itself, e.g., it is not a probability. Its usefulness comes from theorems such as the LRT.
- Comparison between model w/ and w/o source to reject null hypothesis = no source (TS large → reject null hypothesis)

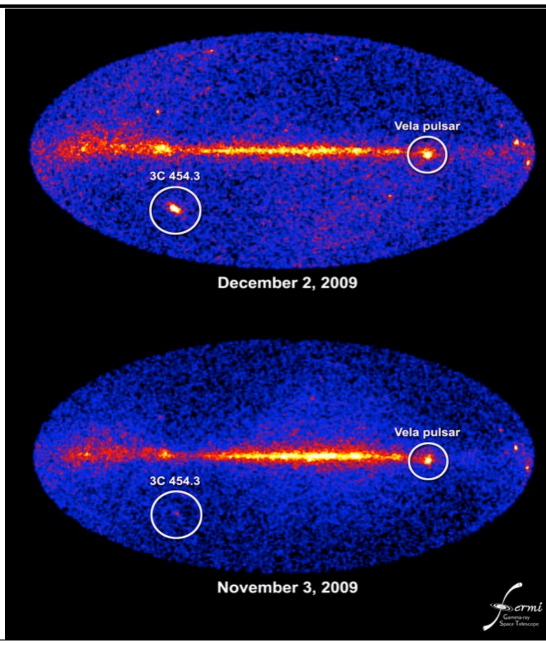
58

Fermi Gamma-ray Space Telescope

Analysis Tutorial

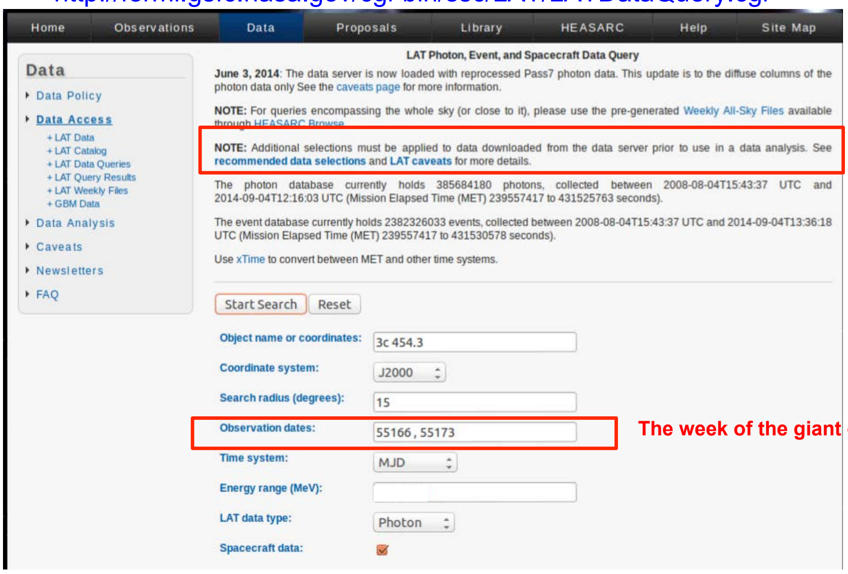
70

Blazar one of … 3c454.3's record flares!



How to download data

http://fermi.gsfc.nasa.gov/cgi-bin/ssc/LAT/LATDataQuery.cgi



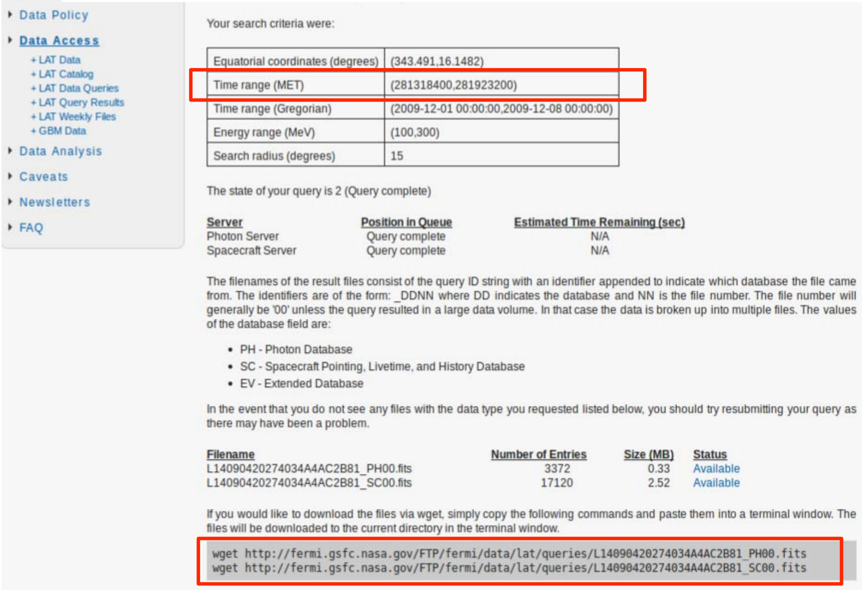
The week of the giant outburst!!

Fermi Gamma-ray Space Telescope

→ Download both spacecraft and photon data ←

Take note of the start and stop MET

follow the link



- Data Policy
- Data Access
  - + LAT Data
  - + LAT Catalog
  - + LAT Data Queries
  - + LAT Query Results
  - + LAT Weekly Files
  - + GBM Data
- Data Analysis
- Caveats
- Newsletters
- FAQ

Your search criteria were:

Equatorial coordinates (degrees)	(343.491,16.1482)
Time range (MET)	(281318400,281923200)
Time range (Gregorian)	(2009-12-01 00:00:00,2009-12-08 00:00:00)
Energy range (MeV)	(100,300)
Search radius (degrees)	15

The state of your query is 2 (Query complete)

Server	Position in Queue	Estimated Time Remaining (sec)
Photon Server	Query complete	N/A
Spacecraft Server	Query complete	N/A

The filenames of the result files consist of the query ID string with an identifier appended to indicate which database the file came from. The identifiers are of the form: _DDNN where DD indicates the database and NN is the file number. The file number will generally be '00' unless the query resulted in a large data volume. In that case the data is broken up into multiple files. The values of the database field are:

- PH - Photon Database
- SC - Spacecraft Pointing, Livetime, and History Database
- EV - Extended Database

In the event that you do not see any files with the data type you requested listed below, you should try resubmitting your query as there may have been a problem.

Filename	Number of Entries	Size (MB)	Status
L14090420274034A4AC2B81_PH00.fits	3372	0.33	Available
L14090420274034A4AC2B81_SC00.fits	17120	2.52	Available

If you would like to download the files via wget, simply copy the following commands and paste them into a terminal window. The files will be downloaded to the current directory in the terminal window.

```
wget http://fermi.gsfc.nasa.gov/FTP/fermi/data/lat/queries/L14090420274034A4AC2B81_PH00.fits
wget http://fermi.gsfc.nasa.gov/FTP/fermi/data/lat/queries/L14090420274034A4AC2B81_SC00.fits
```

Fermi Gamma-ray Space Telescope

gtselect (select data)

[/home/]$ gtselect evclass=128 evtype=3
Input FT1 file[ph.fits]
Output FT1 file[filtered.fits]
RA for new search center (degrees) (0:360) [343.494812]
Dec for new search center (degrees) (-90:90) [16.1495]
radius of new search region (degrees) (0:180) [15]
start time (MET in s) (0:) [281318400]
end time (MET in s) (0:) [281923200]
lower energy limit (MeV) (0:) [100]
upper energy limit (MeV) (0:) [500000]
maximum zenith angle value (degrees) (0:180) [90]

Events with a high prob. to be gammas

Setting the max ZA, filter gammas from albedo events (gamma from the Earth that can be a significant source of background)

```
> gtselect evclass=128 evtype=3  infile=ph.fits outfile=filtered.fits \
ra=343.49  dec=16.15 rad=15 tmin=281318400 tmax=281923200  \
emin=100 emax=500000 zmax=90
```

Note: all analysis steps are scriptable via explicit assign parameters on command-line. Look at the manual for details. 74

gtmktime (cut the bad time intervals)

[/home/]$ gtmktime
Spacecraft data file[sc.fits]
Filter expression[(DATA_QUAL>0)&&(LAT_CONFIG==1)]
Apply ROI-based zenith angle cut[no]
Event data file[filtered.fits]
Output event file name[filtered_gti.fits]

Filter out events collected while passing in SAA and other low-quality events

Use ZA to filter only proper GTIs

75

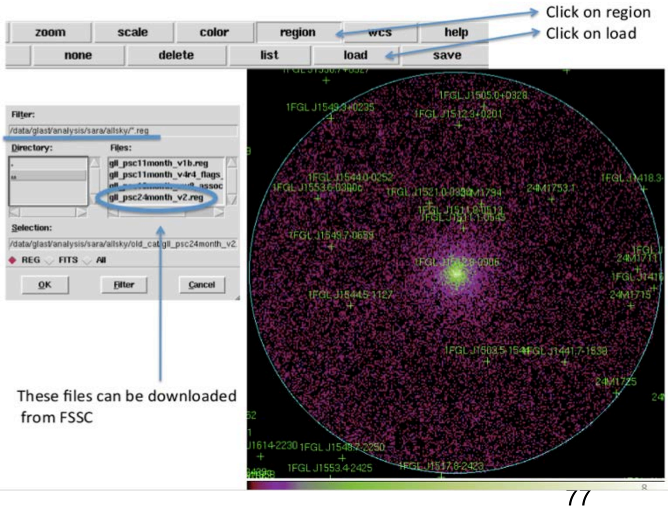
gtbin (Counts Map)

[/home]$ gtbin
Type of output file (CCUBE|CMAP|LC|PHA1|PHA2|HEALPIX) [CMAP]
Event data file name[filtered_gti.fits]
Output file name[CMAP.fits]
Spacecraft data file name[sc.fits]
Size of the X axis in pixels[120]
Size of the Y axis in pixels[120]
Image scale (in degrees/pixel)[0.25]
Coordinate system (CEL - celestial, GAL -galactic) (CEL|GAL) [CEL]
First coordinate of image center in degrees (RA or galactic l) [343.494812]
Second coordinate of image center in degrees (DEC or galactic b) [16.1495]
Rotation angle of image axis, in degrees[0]
Projection method e.g. AIT|ARC|CAR|GLS|MER|NCP|SIN|STG|TAN: [AIT]

76

Look at the counts map

Use DS9 to look at the counts map of your ROI and check for close-by sources
> ds9 CMAP.fits &



77

gtbin (Light Curve)

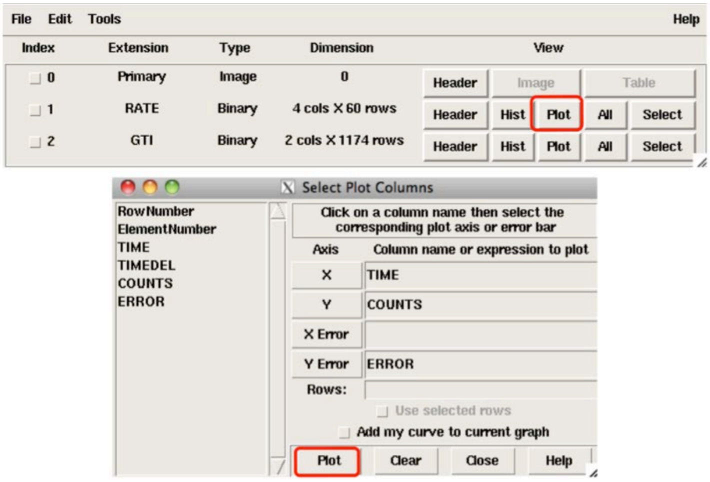
[/home]$ gtbin
Type of output file (CCUBE|CMAP|LC|PHA1|PHA2|HEALPIX) [LC]
Event data file name[filtered_gti_smallROI.fits] → NB selected at 1 deg
Output file name[LC.fits]
Spacecraft data file name[sc.fits]
Algorithm for defining time bins (FILE|LIN|SNR) [LIN]
Start value for first time bin in MET[281318400]
Stop value for last time bin in MET[281923200]
Width of linearly uniform time bins in seconds[86400]

Take care that just one source falls inside the gtselect'ed ROI
(and take out albedo and other low-quality gammas via gtmktime)

78

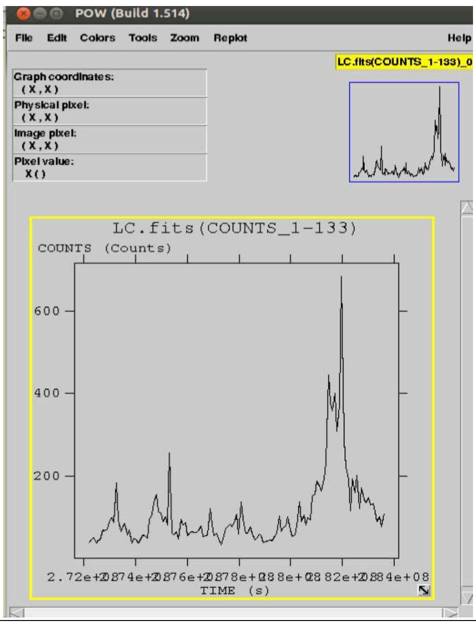
Light-curve: a quick-look

Use FitsView to look at the lightcurve:
> fv LC.fits &



79

Light-curve: a quick-look



Rough estimate of light curve
(background not subtracted!)

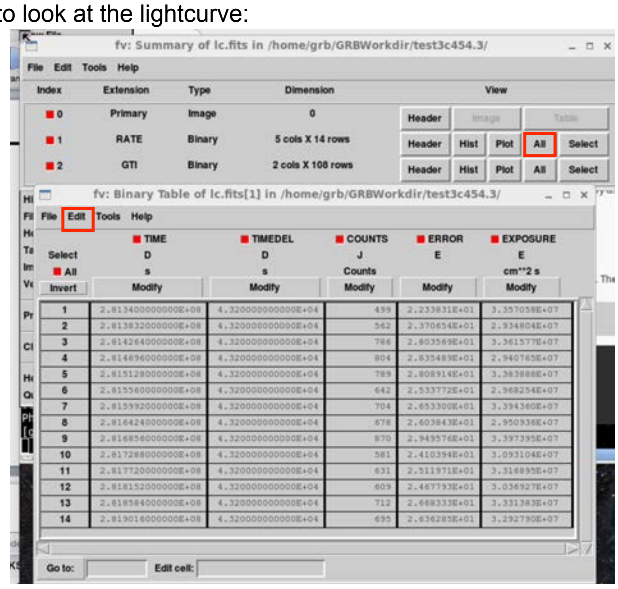
80

gtbin – II (Light Curve)

[/home]$ gtexposure
Light curve file[] lc.fits
Spacecraft file[] sc.fits
Response functions[CALDB]
Source model XML file[none]
Photon index for spectral weighting[-2.1]

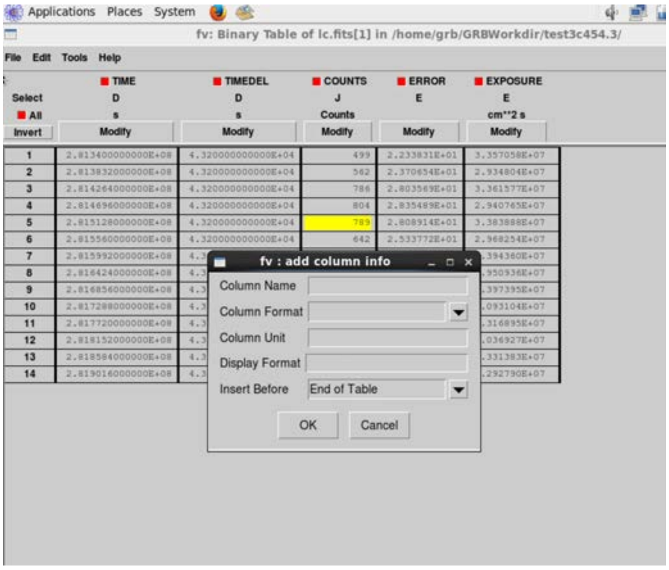
Light-curve: calculate fluxes

Use FitsView to look at the lightcurve:
> fv LC.fits &



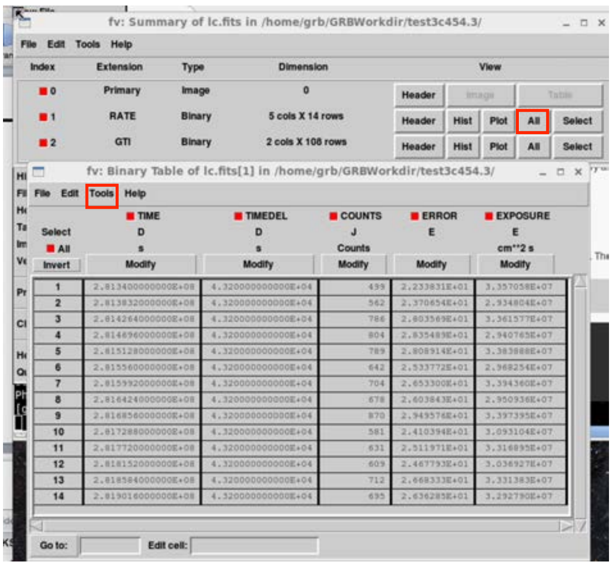
Light-curve: calculate fluxes

Insert new column
Calculate rate
Calculate rate errors



Light-curve: calculate fluxes

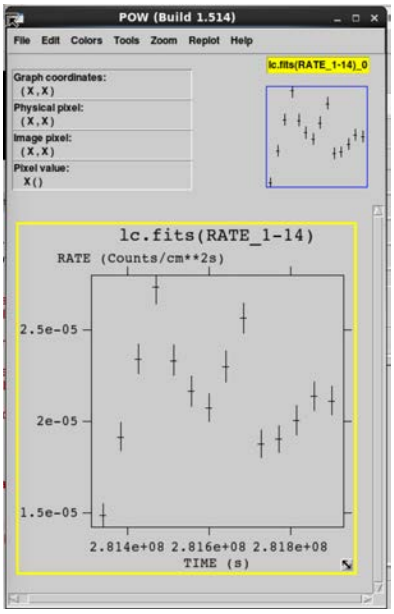
Insert new column
Calculate rate
Calculate rate errors



Light-curve: calculate fluxes

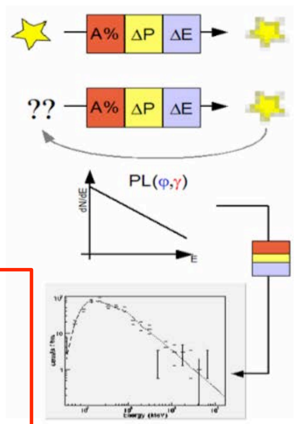
Insert new column
Calculate rate
Calculate rate errors

Light curve



Perform the fit: the likelihood approach

In high energy gamma rays it is never possible to really isolate a source because of limited statistics and strong and structured background.
Therefore statistical techniques have to be applied.
The most used method is the likelihood analysis based on the Poisson statistics.
The method requires to assume a model for the signal detected by the telescope.



- Assume a **model**
- Model **convolved** with Instrument responce Function (IRF)
- Maximizing likelihood find the best set of parameters that reproduce the observed spectrum

87

Perform the fit: the likelihood approach

- Absolute value of likelihood meaningless!
- Comparison between model w/ and w/o source to reject H_0 = no source
- Many variables may be calculated BEFORE selecting the models
- IRFs depend on inclination angle:
  Livetime Cube: seconds in $\Delta\Omega$ with a given z, the time that the LAT observed a given position on the sky at a given inclination angle
  Exposure Map: integration of the effective area over the FoV weighted by the livetime over a position-energy grid, $N_{model} = \int \Phi_{model}(\Omega, E(t)) \times A_{LAT}(\Omega, E)$

88

Likelihood 1st step: gtltcube

[/home]$ gtltcube
Event data file[filtered_gti.fits]
Spacecraft data file[sc.fits]
Output file[ltCube.fits]
Step size in cos(theta) (0.:1.) [0.025]
Pixel size (degrees)[1]
……..

The "livetime cube" must be re-calculated when a new time-interval or a new ZA is selected

89

Likelihood 2nd step: gtexpmap

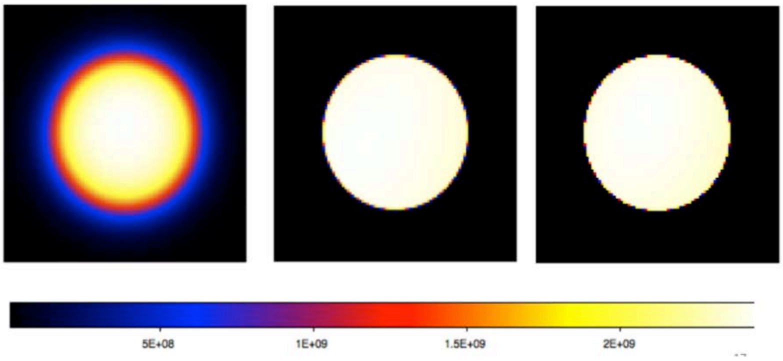
[/home/]$ gtexpmap
Event data file[filtered_gti.fits]
Spacecraft data file[sc.fits]
Exposure hypercube file[ltCube.fits]
output file name[expMap.fits]
Response functions[CALDB]
Radius of the source region (in degrees)[30]
Number of longitude points (2:1000) [120]
Number of latitude points (2:1000) [120]
Number of energies (2:100) [20]
Computing the ExposureMap using ltCube.fits
...

Add 15° to the ROI radius

90

gtexpmap

Quick check with DS9: fields must be homogenous



91

Likelihood 3rd step: the XML model

5. Create a source model XML file

The gtlike tool reads the source model from an XML file. The model file contains your best guess at the locations and spectral forms for the sources in your data. A source model can be created using the model editor tool, by using the user contributed tool **make3FGLxml.py** (available at the user-contributed tools page), or by editing the file directly within a text editor.

Here we cannot use the same source model that was used to analyze six months of data in the Unbinned Likelihood tutorial, as the 2-year data set contains many more significant sources and will not converge. Instead, we will use the 3FGL catalog to define our source model by running make3FGLxml.py. To run the script, you will need to download the current LAT catalog file and place it in your working directory:

```
prompt> make3FGLxml.py gll_psc_v16.fit 3C279_binned_gti.fits -o 3C279_input_model.xml
-G $FERMI_DIR/refdata/fermi/galdiffuse/gll_iem_v06.fits -g gll_iem_v06
-I $FERMI_DIR/refdata/fermi/galdiffuse/iso_P8R2_SOURCE_V6_v06.txt
-i iso_P8R2_SOURCE_V6_v06 -s 120 -p TRUE -v TRUE
This is make3FGLxml version 01r0.
The default diffuse model files and names are for pass 8
and assume you have v10r00p05 of the Fermi Science Tools or higher.
Creating file and adding sources from 3FGL
Added 312 point sources, note that any extended sources in ROI were modeled as point sources
because psForce option was set to True
prompt>
```

Note that we are using a high level of significance so that we only fit the brightest sources and we have forced the extended sources to be modeled as point sources. This only affects the lobes of Centarus A which are just outside the FOV.

It is also necessary to specify the entire path to location of the diffuse model on your system. The resulting XML model contains 312 sources. Clearly, the simple 4-source model we used for the 6-month Unbinned Likelihood analysis would have been too simplistic.

This XML file uses the spectral model from the 3FGL catalog analysis for each source. (The catalog file is available at the LAT 4-yr Catalog page.) However, that analysis used a subset of the available spectral models. A dedicated analysis of the region may indicate a different spectral model is preferred. For more details on the options available for your XML models, see:

- Descriptions of available Spectral and Spatial Models
- Examples of XML Model Definitions for Likelihood

Likelihood 3rd step

./make3FGLxml.py gll_psc_v16.fit filtered_gti.fits -o 3c454.3.xml -G /home/grb/software/GlastExt/diffuseModels/v2r0/gll_iem_v06.fits -g gll_iem_v06 -l /home/grb/software/GlastExt/diffuseModels/v2r0/iso_P8R2_SOURCE_V6_v06.txt -i iso_P8R2_SOURCE_V6_v06 -s 120 -p TRUE -v TRUE

Likelihood 3rd step: the XML model

•Backgrounds

```
<!-- Diffuse Sources -->
<source name="galactic_background"   type="DiffuseSource">
<spectrum type="PowerLaw">
<parameter free="1" max="10" min="0" name="Prefactor" scale="1" value="1"/>
<parameter free="0" max="1" min="-1" name="Index" scale="1.0" value="0"/>
<parameter free="0" max="2e2" min="5e1" name="Scale" scale="1.0" value="1e2"/>
</spectrum>
<spatialModel file="gll_iem_v06.fits" type="MapCubeFunction">
<parameter free="0" max="1e3" min="1e-3" name="Normalization" scale="1.0" value="1.0"/>
</spatialModel>
</source>
<source name="extragalactic_background" type="DiffuseSource">
<spectrum file="iso_P8R2_SOURCE_V6_v06.txt" type="FileFunction">
<parameter free="1" max="10" min="1e-2" name="Normalization" scale="1" value="1"/>
</spectrum>
<spatialModel type="ConstantValue">
<parameter free="0" max="10.0" min="0.0" name="Value" scale="1.0" value="1.0"/>
</spatialModel>
</source>
```

94

XML model

•Typical source entry for an assumed powerlaw

```
{<source name="......" type="PointSource">...
 </source>
                    Your sources here
```

```
<source name="3c454.3" type="PointSource">
<spectrum type="PowerLaw2">
<!-- Source is in ROI center -->
<parameter error="0.00" free="1" max="1000" min="1e-06" name="Integral" scale="1e-04"
value="1.000"/>
 <parameter error="0.00" free="1" max="0" min="-5" name="Index" scale="1" value="-2.000"/>
                <parameter free="0" max="3e6" min="20" name="LowerLimit" scale="1"
value="100."/>
      <parameter free="0" max="3e6" min="20" name="UpperLimit" scale="1" value="300000."/>
      </spectrum>
<spatialModel type="SkyDirFunction">
<parameter free="0" max="360.0" min="-360.0" name="RA" scale="1.0" value="343.494812"/>
<parameter free="0" max="90" min="-90" name="DEC" scale="1.0" value="16.149500"/>
</spatialModel>
</source>
```

95

XML model

Test different models... power law * HE exponential cut-off

```
<source name="3c454.3" type="PointSource">
<spectrum type="PLSuperExpCutoff">
<parameter free="1" max="1000" min="1e-05" name="Prefactor" scale="1e-07"
value="1"/>
<parameter free="1" max="0" min="-5" name="Index1" scale="1" value="-1.7"/>
<parameter free="0" max="1000" min="50" name="Scale" scale="1" value="200"/>
<parameter free="1" max="30000" min="500" name="Cutoff" scale="1" value="3000"/>
<parameter free="0" max="5" min="0" name="Index2" scale="1" value="1"/>
</spectrum>
```

•Look here for source model definition and XML model definitions:
http://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/xml_model_defs.html
http://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/source_models.html

•Useful python script to load 2FGL sources that belongs to your ROI in your XML file model (make3FGLxml.py)
http://fermi.gsfc.nasa.gov/ssc/data/analysis/user/

96

Diffuse response

[/home/]$gtdiffrspEvent data file[]filtered_gti.fits
Spacecraft data file[] sc.fits
Source model file[] 3c454.3.xml
Response functions to use[] CALDB

97

Finally… gtlike performig the actual fit

[/home/]$ gtlike plot=yes
Statistic to use (BINNED|UNBINNED) [UNBINNED]
Spacecraft file[sc.fits]
Event file[filtered_gti.fits]
Unbinned exposure map[expMap.fits]
Exposure hypercube file[ltCube.fits]
Source model file[../xml_models/_3c454.3_model_ROI15.xml]
Response functions to use[P7REP_SOURCE_V15]
Optimizer (DRMNFB|NEWMINUIT|MINUIT|DRMNGB|LBFGS) [NEWMINUIT]

Typically use DRMNGB/DRMNFB to find a rough estimate of the likelihood maxima and refine later on with **MINUIT (or NEWMINUIT)**

98

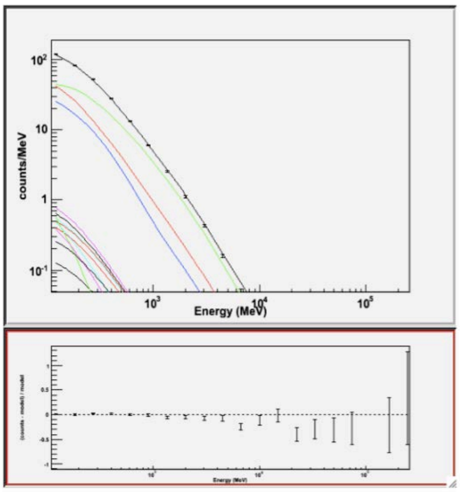
Likelihood output

{'3c454.3': {'Integral': '0.146106 +/- 0.00271733',
'Index': '-2.29973 +/- 0.017189',
'LowerLimit': '100',
'UpperLimit': '300000',
'Npred': '4171.85',
'ROI distance': '0',
'TS value': '17548.4',
'Flux': '1.46192e-05 +/- 2.7178e-07',
…
extragalactic_background': {'Normalization': '1.20197 +/- 0.23541',
'Npred': '643.953',
'Flux': '0.000170707 +/- 3.34331e-05',
},
'galactic_background': {'Prefactor': '0.739969 +/- 0.251827',
'Index': '0',
'Scale': '100',
'Npred': '357.929',
'Flux': '0.000215978 +/- 7.35023e-05',

gtlike creates two output files:
1) results.dat: fit results
2) counts_spectra.fits: the counts in a proper energy binning

99

Plot



Solid lines follows the order as they are listed in the file results.dat:
black) ROI fit
red) 1st source (pks1510)
green) 2nd source (galactic)
blue) 3rd source (istropic)
Magenta) 4th source
Cyan) 5th source
...

100

Comparison of different models

Powerlaw * HE exp cut-off

{'3c454.3': {'Prefactor': '0.39194 +/- 0.00793161',
'Index1': '-2.12802 +/- 0.03056',
'Cutoff': '5495.55 +/- 934.857 (MeV)
'Npred': '4157.04',
'ROI distance': '0',
'TS value': '17604.2',
'Flux': '1.41693e-05 +/- 2.72878e-07'
(ph cm-2 s-1)

----> Comparing TS values for different models!
For this source, in this time interval, the model with the HE exponential cutoff is favoured with respect to the Simple Powerlaw

You can repeat analyses by yourself also following instructive and complete Tutorials on the FSSC web-site:
- Standard Likelihood: http://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/likelihood_tutorial.html
- PyLike: http://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/python_tutorial.html

101

SED modeling

Generate Spectral Points

To generate spectral points to plot on top of the butterfly that we just produced, you need to go back to the data selection part and use gtselect (filter in python) to divide up your data set in energy bins and run the likelihood fit on each of these individual bins. Luckily, there's a script that does this for us that we'll employ to get a final result. This script also generates the butterfly plot we just produced so you won't have to redo that again and it also does a lot of checks on the data to make sure that everything is going ok. If you've made it this far, you might be a little curious as to why we didn't jump right into using this tool but now you're in a position to play with the python tools and make them do what you want them to. The script is also much more careful in handling units and saving everything to files than we have been in this interactive session.

So download the likeSED.py user contributed tool (it's in the SED_scripts package) and load it up into python. You can find information on the usage of this tool on the same page where you downloaded it. It can be used to generate an SED plot for both binned and unbinned analyses but we're only going to work on a binned analysis here.

```
>>> from likeSED import *
This is likeSED version 12.1, modified to handle Pass 7 selections.
```

Now you need to create a likeInput object that takes our unbinnedAnalysis object, the source name and the number of bins we want as arguments. We're going to make 9 bins here like in the paper and we're also going to make custom bins and bin centers. You can have the module chose bins and bin centers for you (via the getECent function) but we're going to do it this way so we can have the double-wide bin at the end. We're also going to use the 'fit2' file (the result of our fit on the full dataset with the NewMinuit optimizer) for the fits of the individual bins but we need to edit it first to make everything fixed except for the integral value of PG 1553+113. Go ahead and do that now. We're basically assuming that the spectral index is the same for all bins and that the contributions of the other sources within the ROI are not changing with energy.

```
>>> inputs = likeInput(like,'3FGL J1555.7+1111',nbins=9)
>>> inputs.plotBins()
```

SED modeling

This last step will take a while (approximately 30 minutes) because it's creating an expmap and event file for each of the bins that we requested (look in your working directory for these files). Once it is done we'll tell it to do the fit of the full energy band and make sure we request that it keep the covariance matrix. Then we'll create a likeSED object from our inputs, tell it where we want our centers to be and then fit each of the bands. After this is done, we can plot it all with the Plot function.

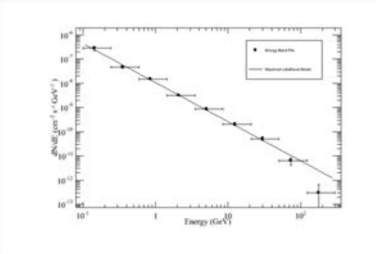
```
>>> inputs.fullFit(CoVar=True)
Full energy range model for 3FGL J1555.7+1111:
3FGL J1555.7+1111
Spectrum: PowerLaw
32 Prefactor: 5.928e-01 2.586e-02 1.000e-04 1.000e+04 ( 1.000e-11)
33 Index: 1.621e+00 2.674e-02 0.000e+00 1.000e+01 (-1.000e+00)
34 Scale: 1.448e+03 0.000e+00 3.000e+01 5.000e+05 ( 1.000e+00) fixed

Flux 0.1-300.0 GeV 7.2e-08 +/- 4.2e-09 cm^-2 s^-1
Test Statistic 3141.46
>>> sed = likeSED(inputs)
>>> sed.getECent()
>>> sed.fitBands()
-Running Likelihood for band0-
-Running Likelihood for band1-
-Running Likelihood for band2-
-Running Likelihood for band3-
-Running Likelihood for band4-
-Running Likelihood for band5-
-Running Likelihood for band6-
-Running Likelihood for band7-
-Running Likelihood for band8-
>>> sed.Plot()
```

This produces the following three plots: a TS plot, a counts spectrum, and a butterfly plot (statistical errors only):

SED modeling

The TS plot shows the TS value for each bin in the spectrum produced by likeSED. Notice that the last bin is double-wide.



Counts spectrum produced by likeSED along with the spectral points.