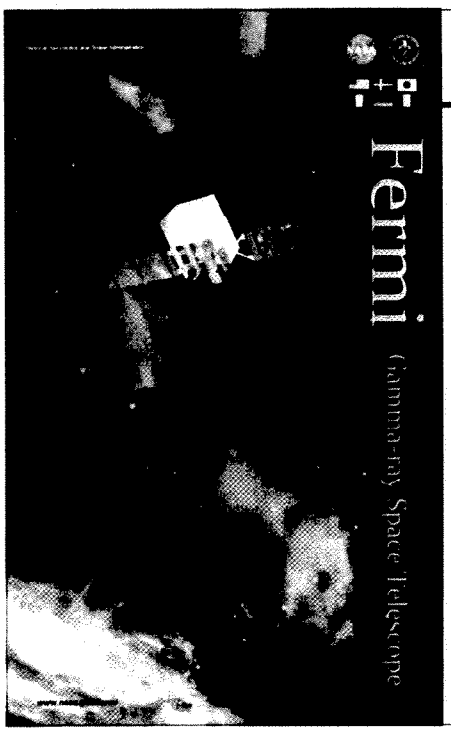


Fermi

Gamma-ray Space Telescope

Analysis Workshop
17 December 2009

Advanced Likelihood

E. Hays

1

Goals

- Quality checks on spectral fitting of point sources
  - Major gotchas
  - Simple checks
  - Models revisited
  - Spectral residuals
  - Spatial residuals
- Useful considerations
  - Impact of region selection
  - Impact of zenith angle selection (relates to above...)
  - Impact of energy selection
  - Impact of spectral model
- Binned vs. Unbinned likelihood

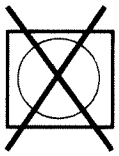
2

Major gotchas

- Parameter estimate depends critically on calculating the proper exposure

selection		livetime	response	minimization
gtselect	**gtmktime**	**gtlcube**	**gtexpmap** **gtsrcmaps**	**gtlike**

- Examples of things that can screw this up
  - fselect, fcopy selections do not update the header keywords used in the exposure calculation
  - Mismatch of data and IRF set
  - Mismatch of initial ROI selection and data cube (binned)
  - Mismatch of calculated diffuse response and model diffuse components - Use different names for different models



3

Likelihood output - simple checks

Did the fit work and does it make sense?

- Did the minimization converge?
- Are the number of predicted photons reasonable?
- Do the parameter values make sense?
  - values hitting limits?
  - source with extremely soft spectrum or hard spectrum?
- Do the parameter errors make sense?
  - Too small? Were enough parameters left free?
  - Larger than the parameter values - with low TS...better luck next time
- Consider the above for target source and field sources
- All of the above become more critical for faint sources, complex regions, time-binned light curves...

4

Likelihood - ROI selection

How big?

- Big enough to constrain model components - source of interest, diffuse emission, nearby sources
- Small enough to avoid significant zenith cut losses to exposure
  - Practical advantage! less photons and less sources => less calculations for unbinned analysis
  - Analysis disadvantage! likelihood is an inclusive modeling strategy
- Recommendations
  - 10 deg for isolated point source (E>100 MeV)
  - Larger regions (15-20 deg) benefit confused sources, aid in separating diffuse at low energy, improve error estimates
- Test it
  - Are fit results reliable for different ROI radii?
  - What is the impact on GTIs?

5

Likelihood model - sources

What should be included?

- All sources that contribute photons to the selected region
  - Bright source list sources within ~10 deg of the ROI boundary - accommodates tail of low energy PSF
  - Same goes for catalog sources once available
- Galactic diffuse model
- Isotropic diffuse model
  - Important for all parts of the sky...provides a home for residual instrumental effects

This is a starting point. Adapt to find what works best for your region and source.

6

Likelihood Model - spectra

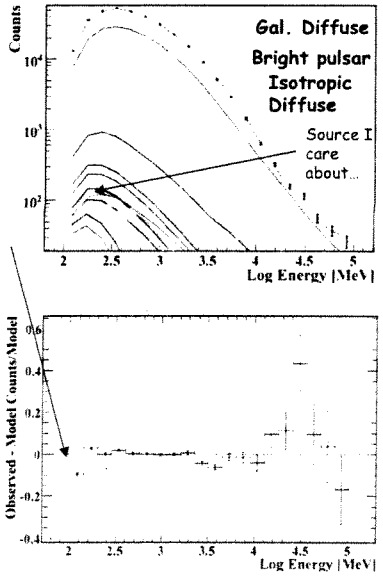
What spectral shape?

- **Power laws are simple and well defined**
  - **For faint sources, difficult to justify more parameters**
- **BUT lots of LAT sources are not simple power laws... some tips to help motivate other spectral forms**
  - **Bright pulsars?**
    - **Try simple exponentially cutoff power laws to improve fits for the pulsar itself *and for nearby sources***
  - **Check the energy distribution for an energy-dependent ROI selection**
  - **Do the power-law fit parameters vary significantly for different minimum energy selections or fits in separate energy bins?**
- **Most accurate and unbiased way to determine spectral parameters and errors is by testing that hypothesis using the likelihood fit**

7

Spectral Residuals

- **Unbinned analysis produces predicted counts and residuals.** Example is a long integration near the Galactic plane and a bright pulsar
- **Discrepancy at low energy is typical**
  - ***Likelihood uses true energy***
- **Discrepancies strongly tied to diffuse model for most analysis**
  - **Diffuse mediates cross talk between your source and neighbors**
  - **Consider relative strength and test impact of model choices and selections on source of interest**



8

Likelihood - reality checks

Is anything missing in the model?

- Visual inspection of count maps and residuals
- Test Statistic maps (unbinned analysis)
  - gttsmap - Tests hypothesis of additional point source over a grid
  - Very Calculation Intensive
    - try small regions (5 deg) and large grid spacing (0.5 deg)
  - Note this can expose deficiencies in the diffuse model in addition to evidence for an additional source
  - Warning: gttsmap is not a tool for localization, gtfindsrc does that
- Predicted and residual count maps (binned analysis)
  - Profiles, radial density, energy dependence

9

Likelihood - useful tests

- Overall consistency - lots of good ways to get at this
- Iteration
  - Consistent results if using output model is fit model?
- Data selection consistency
  - Effects of energy selection?
  - Changes with ROI selection? (Keep in mind this also effects good time selection in combination with zenith cut)
  - Consistency with results in distinct energy bins (ala catalog)
  - Separate analysis of front and back events (using appropriate IRFs, diffuse response, and isotropic model)
  - Effects of time selection
- Fit and Minimization choices
  - Impact of starting parameter values in the model?
  - Fit tolerence? (converging to true minimum?)
  - Effects of optimizer?

10

Binned vs. Unbinned Likelihood

- Unbinned: Treats each photon independently (position, energy)
  - Best theoretical performance
  - More sensitive - important for faint sources
  - Best option for low statistics scenarios - light curves
  - Not for use with spatially extended sources
  - More difficult to diagnose problems in individual source fit
- Binned: Treats the data in bins of position and energy. Minimal criteria - more photons than bins
  - Less computationally intensive than unbinned
  - Handles templates for extended sources
  - Allows more straightforward diagnostics of fit (source maps, spatial profiles, energy dependent comparisons of prediction and model)
  - At highest energies, can run into low statistics even for long integrations

Use of both allows consistency check
(for data sets where both can be reasonably used)

11

gtobssim

- The ultimate test...
  - Can you simulate what you found?

12