LAT Data Simulation and Likelihood Analysis

Jim Chiang
NASA/GSFC, Code 661, Greenbelt MD 20771
JCA/Physics Dept., University of Maryland, Baltimore County, Baltimore MD 21250
SLAC, 2575 Sand Hill Rd, Menlo Park CA 94025

Observation Simulation

- Time scales:
  - orbital/rocking — 95 min ($\sim$ 30–40 min chunks of continuous exposure to a single sky position)
  - SAA passages — 24 hrs
  - orbit precession — $\sim$ 55 days
- IRFs (GLAST25)
  - PSF: $\sigma \sim 10^{\circ}(E/30\text{ MeV})^{-3/4}$ (with long tails)
  - eff. area (for $i = 0$): $\sim 10^3$ cm^2 at 30 MeV, $\sim 10^4$ cm^2 at 30 GeV
- Other effects included: inclination and zenith angle cuts
- Not simulated: changing detector modes, live-time effects, energy dispersion

Observation Simulation (cont.)

- 3EG sources — 3C 279, 3C 273, Crab, Vela
- Galactic and extragalactic diffuse ($\Gamma = 2.1$)
- periodic variability — $P = 10^4$ sec (non-pulsar)
- "flares" — triangle & step functions
- spectral models: power-laws only (so far)
- Method:
  - Compute orbit solution with 35° step rocking, $28\overset{\circ}{.}5$ orbital inclination
  - Loop over (point) sources and for each time step compute i, $A_{\rm eff}$, psf, #photon events
  - Draw apparent event energies and positions

Likelihood Calculation

- "This is not X-ray astronomy" (PLN)
  `http://giants.stanford.edu/~pln/glastlike-paper.pdf`:
  paucity of photons, scanning, ~ 2 sr FOV, $\sigma_{\rm psf} \sim 0^\circ\!.1$–$10^\circ$
  $\Rightarrow$ each photon has its own response
- Likelihood model and total response R

$$M(E', \hat{p}', t) = \int dE\, d\hat{p}\, D(E'; E, \hat{p}, \vec{L}(t)) P(\hat{p}'; E, \hat{p}, \vec{L}(t)) A(E, \hat{p}, \vec{L}(t)) S(E, \hat{p}) \quad (1)$$

$$\equiv \int dE\, d\hat{p}\, R(E', \hat{p}', t; E, \hat{p}) S(E, \hat{p}). \quad (2)$$

- Source model

$$S(E, \hat{p}) = \sum_i s_i(E) \delta(\hat{p} - \hat{p}_i) + S_G(E, \hat{p}) + S_{\rm eg}(E, \hat{p}). \quad (3)$$

Likelihood Calculation (cont.)

- Need to maximize

$$\log \mathcal{L} = \sum_j \log M(E'_j, \hat{p}'_j, t_j) - N_{\mathrm{pred}}, \tag{4}$$

where the predicted number of photons is

$$N_{\mathrm{pred}} = \int dE'\, d\hat{p}'\, dt\, M(E', \hat{p}', t). \tag{5}$$

- Storage requirements for the total response may require ~ 40 GB storage (PLN)
$\Rightarrow$ compute $R_k(E', \hat{p}', t; E, \hat{p})$ on the fly

Decomposing $log\mathcal{L}$

- The abc's (and d's) of computing $\log \mathcal{L}$:

$$a_{ij} \equiv \int dE\, s_i(E)\, R(E'_j, \hat{p}'_j, t_j; E, \hat{p}_i) \tag{6}$$

$$b_j \equiv \int dE\, d\hat{p}\, [S_G(E, \hat{p}) + S_{\rm eg}(E, \hat{p})]\, R(E'_j, \hat{p}'_j, t_j; E, \hat{p}) \tag{7}$$

$$c_i \equiv \int dE\, s_i(E) \int dE'\, d\hat{p}'\, dt\, R(E', \hat{p}', t; E, \hat{p}_i) \tag{8}$$

$$d \equiv \int dE\, d\hat{p}\, [S_G(E, \hat{p}) + S_{\rm eg}(E, \hat{p})] \int dE'\, d\hat{p}'\, dt\, R(E', \hat{p}', t; E, \hat{p}). \tag{9}$$

The log-likelihood is

$$\log \mathcal{L} = \sum_j \log \left(\sum_i a_{ij} + b_j \right) - \sum_i c_i - d. \tag{10}$$

  - 10^6 photons$\times 10^2$ point sources $\sim 10^8$ terms
  - If only parameters of one or two sources are varied for each trial value of $\log \mathcal{L}$, then only $\lesssim 2\%$ of terms need to be recomputed for each function evaluation

- Angular integrals — ROI ($\hat{p}'$) $\sim 15°$ radius; Source Region ($\hat{p}$) $\gtrsim 30°$ radius

- Time integrals $\Rightarrow$ decomposing c's and d's further

Preliminary Likelihood Results: 3C 279 and 3C 273

- Two steady point sources, no background, 55 days of data, 7500 events
- No diffuse emission (only a's and c's computed)
- $\sigma_{\rm psf} \propto E^{-3/4}$
  $\Rightarrow$ fitted spectral index depends on psf accuracy
  - $\sigma_{\rm psf}(10{\rm GeV}) \lesssim 0\overset{\circ}{.}1 \Rightarrow 10^3$ abscissa points per 30°
  - Quantify sensitivity ($\Gamma_1 = 1.96$; $\Gamma_2 = 2.58$):

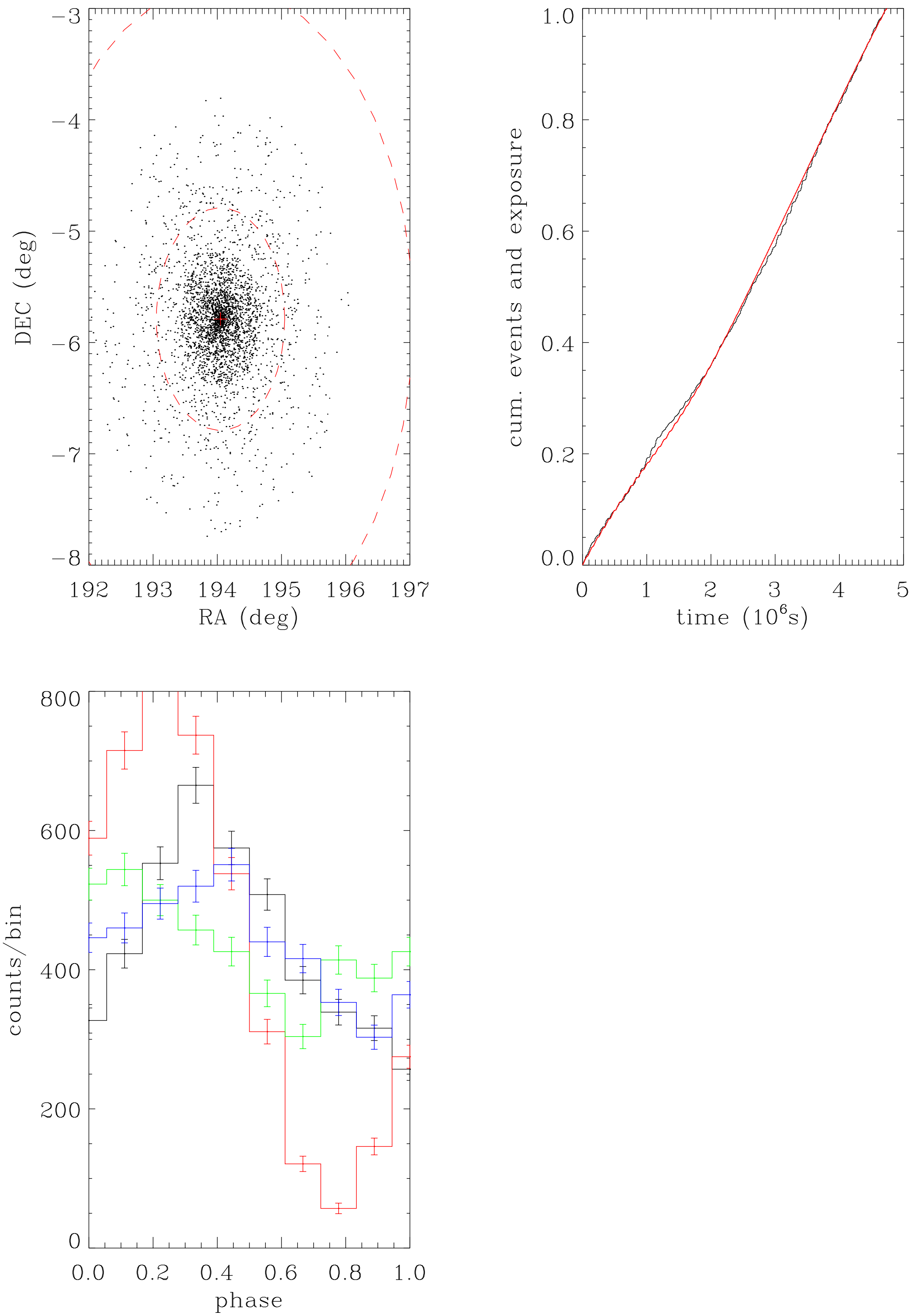
	Γ_1	Γ_2	$\langle N_{\rm eval}({\rm tol} = 10^{-2})\rangle$
$0.95\,\sigma_{\rm psf}$	1.91 ± 0.01	2.33 ± 0.04	74
$\sigma_{\rm psf}$	1.93 ± 0.01	2.37 ± 0.05	72
$1.05\,\sigma_{\rm psf}$	1.95 ± 0.01	2.41 ± 0.05	71

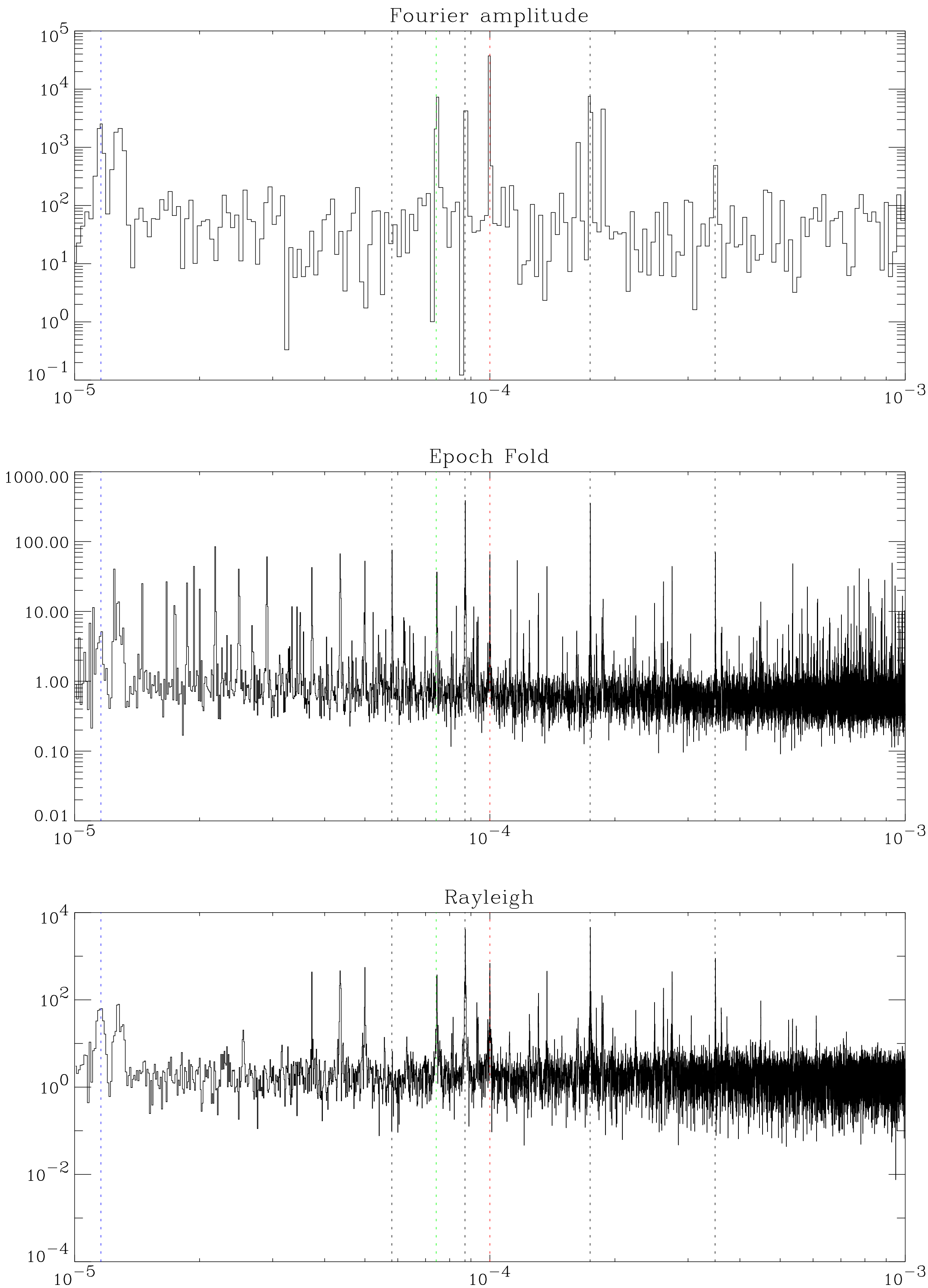
Optimization Issues

- Poisson statistics $\Rightarrow$ Levenberg-Marquardt is out
- Non-derivative methods
  - Nelder-Mead, Powell (both inefficient)
  - Require careful parameter scaling
- Derivative methods
  - Newton, Quasi-Newton, etc.
  - Analytic derivatives available for fluxes and spectral indices, but not for positions $\Rightarrow$ LIKE's "fine" TS maps
  - How can we fit for positions and spectral parameters simultaneously? (evaluate derivatives numerically?)
- Need methods for dealing with *a priori* constraints, e.g., non-negative fluxes
  - Enforce maximum step size
  - Penalty functions ($+\log\mathcal{L}$) outside boundaries

To Do

- Restrict to ROI and Source Region (i.e., do angular integrals)
- Diffuse emission
- More point sources (i.e., allow for freezing and thawing of parameters, etc.)
- Find a good optimization package
- User interface (Pat's `glike.py` GUI)


DEC (deg)
RA (deg)
cum. events and exposure
time (10^6 s)
counts/bin
phase



Fourier amplitude
10^5
10^4
10^3
10^2
10^1
10^0
10^{-1}
10^{-5}
10^{-4}
10^{-3}
Epoch Fold
1000.00
100.00
10.00
1.00
0.10
0.01
10^{-5}
10^{-4}
10^{-3}
Rayleigh
10^4
10^2
10^0
10^{-2}
10^{-4}
10^{-5}
10^{-4}
10^{-3}