Unbinned Maximum Likelihood for LAT Data: Results I

J. Chiang (GSFC-UMBC)

Last Revised November 1, 2002

1 Preliminaries: A Basic Test

One of the simplest non-trivial tests of the likelihood method is fitting data for two (somewhat) nearby point sources that have different fluxes and spectral indices. I simulated point sources at the positions of 3C 279 and 3C 273, using the P1234 3EG fluxes and photon spectral indices. The relevant data from the 3EG catalog are

Object	RA	Dec	Flux	Index
3C 279	193.98	-5.82	74.2	1.96
3C 273	187.25	2.17	15.4	2.58

Flux is in units of $10^{-8}\,\mathrm{ph\,cm^{-2}\,s^{-1}}$ and corresponds to energies ≥ 100 MeV. For these source parameters, ~ 7800 photons are detected between 30 MeV and 300 GeV. In Figure 1, some attributes of the photons from one of these simulations are plotted. The simulation time is 826 orbits corresponding to one precession period.

In order to make this test very basic, both the Galactic and extragalactic diffuse background were set to zero in these simulations. Furthermore, no region-of-interest (ROI) cut was imposed initially, i.e., all photon events from the sources were included in the likelihood calculation (cf. §4); and for the first set of tests, the energy resolution was taken to be infinite. These latter two simplifications make the evaluation of the predicted number of photons extremely easy to compute. For each source i, the term giving the predicted number of observed photons in the $\log\mathcal{L}$ decomposition (see eqs. 4, 5, & 11, LikeMemo 0) becomes

$$c_i \equiv \int dE\, s_i(E) \int dE'\, d\hat{p}'\, dt\, R(E', \hat{p}', t; E, \hat{p}_i) \tag{1}$$

$$= \int dE\, s_i(E) \int dt\, A(E, \hat{p}_i, \vec{L}(t)), \tag{2}$$

where E and $\hat{p}$ are true photon energy and direction, E' and $\hat{p}'$ are apparent energy and direction, A is the effective area, s_i is the source spectrum, and $\vec{L}(t)$ describes the telescope orbit and attitude as a function of time. Note that the integral over t in the second relation is just the exposure to the sky position $\hat{p}_i$ as a function of energy. The other relevant terms in the $\log\mathcal{L}$ decomposition are even easier to evaluate:

$$a_{ij} \equiv \int dE\, s_i(E)\, R(E'_j, \hat{p}'_j, t_j; E, \hat{p}_i) \tag{3}$$

$$= s_i(E_j)\, P(\hat{p}'_j; E'_j, \hat{p}_i, \vec{L}(t_j)) A(E'_j, \hat{p}_i, \vec{L}(t_j)), \tag{4}$$

where P is the point spread function, and the subscript j indexes the photon events.

2 Sensitivity to the Accuracy of the PSF

Contrary to my previous findings, the aforementioned likelihood calculation appears to be entirely *insensitive* to deficiencies in modeling the point spread function accurately, at least under the assumption of infinite energy resolution. My earlier likelihood calculations had an error that roughly amounted to assuming zero energy resolution, i.e., $E/\Delta E = 0$ instead of $E/\Delta E = \infty$, as assumed in the simulations. In reality, because of the expected finite energy resolution of the LAT, the capability to measure spectra should in fact depend on how accurately the PSF is known, albeit probably only very weakly (see §3).

In order to quantify the importance of the accuracy of the PSF, a series of simulations were performed in which the Gaussian widths used to characterize the PSF were scaled by varying degrees: $\sigma_{\rm psf} \rightarrow$ Scale Factor $\times\, \sigma_{\rm psf}$. A maximum likelihood fit was performed on each dataset. In Table 1, we show results from both the zero and infinite energy resolution cases.

Table 1: PSF Accuracy Tests

Scale Factor	$E/\Delta E = \infty$		$E/\Delta E = 0$	
	$\Gamma_{\rm 3C\ 279}$	$\Gamma_{\rm 3C\ 273}$	$\Gamma_{\rm 3C\ 279}$	$\Gamma_{\rm 3C\ 273}$
0.90	1.96 ± 0.01	2.60 ± 0.04	1.91 ± 0.01	2.48 ± 0.05
0.95	1.96 ± 0.01	2.59 ± 0.02	1.94 ± 0.01	2.54 ± 0.06
1.00	1.96 ± 0.01	2.58 ± 0.03	1.97 ± 0.02	2.64 ± 0.05
1.05	1.96 ± 0.01	2.58 ± 0.02	1.99 ± 0.02	2.72 ± 0.05
1.10	1.96 ± 0.01	2.58 ± 0.04	2.02 ± 0.02	2.82 ± 0.07

It is worth noting that surprisingly accurate spectral information can be ascertained from fitting the PSF alone, even in the absence of any energy information, provided we know the PSF to better than $\sim 5\%$ (and there is no diffuse emission).

3 Energy Dispersion

The energy resolution of the LAT is specified to be nominally $\sim 10\%$ on-axis over the entire energy range of the instrument. Lacking any information on inclination dependence, I have assumed, for this set of tests, a Gaussian function for the energy redistribution with width $\sigma_e = E/10$, independent of inclination:

$$D(E'; E, \hat{p}, \vec{L}(t)) \equiv \frac{1}{\sqrt{2\pi}\sigma_e} \exp\left(-\frac{(E'-E)^2}{2\sigma_e^2}\right). \tag{5}$$

The $\log \mathcal{L}$ components become

$$a_{ij} = \int_{E_{\rm min}}^{E_{\rm max}} dE\, s_i(E)\, P(\hat{p}'_j; E, \hat{p}_i, \vec{L}(t_j)) A(E, \hat{p}_i, \vec{L}(t_j)) D(E'_j; E, \hat{p}, \vec{L}(t_j)) \tag{6}$$

$$c_i = \int_{E_{\rm min}}^{E_{\rm max}} dE\, s_i(E) \int dt\, A(E, \hat{p}_i, \vec{L}(t)) \int_{E'_{\rm min}}^{E'_{\rm max}} dE' D(E'; E, \hat{p}, \vec{L}(t)) \tag{7}$$

$$= \int_{E_{\rm min}}^{E_{\rm max}} dE\, s_i(E) \int dt\, A(E, \hat{p}_i, \vec{L}(t)) \times \frac{1}{2}\left[{\rm erfc}\left(\frac{E'_{\rm min} - E}{\sqrt{2}\sigma_e}\right) - {\rm erfc}\left(\frac{E'_{\rm max} - E}{\sqrt{2}\sigma_e}\right)\right] \tag{8}$$

Here $E'_{\rm min}$ and $E'_{\rm max}$ are the upper and lower energy cuts on the photon selection. The complementary error function is defined in the usual way,

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-t^2)\, dt. \tag{9}$$

Note that the integrals over the true energy E in equations 6–8 should formally have limits $(E_{\rm min}, E_{\rm max}) = (0, \infty)$. In practice, $E_{\rm min}$ and $E_{\rm max}$ are the minimum and maximum true photon energies for which the response functions are reliably known. In order to avoid biases at the upper and lower energy cuts, it should be sufficient to choose $E'_{\rm min} = E_{\rm min} + n_{\rm sig}\sigma_e$ and $E'_{\rm max} = E_{\rm max} - n_{\rm sig}\sigma_e$, where $n_{\rm sig}$ is some integer $\gtrsim 3$ corresponding to an appropriate confidence limit. However, for the present calculations, we have used $n_{\rm sig} = 0$.

Table 2: Energy Dispersion Tests

Scale Factor	$E/\sigma_e = 10$		$E/\sigma_e = \infty$	
	$\Gamma_{\rm 3C\ 279}$	$\Gamma_{\rm 3C\ 273}$	$\Gamma_{\rm 3C\ 279}$	$\Gamma_{\rm 3C\ 273}$
0.90	1.97 ± 0.01	2.62 ± 0.03	1.95 ± 0.01	2.54 ± 0.03
1.00	1.97 ± 0.01	2.62 ± 0.03	1.95 ± 0.01	2.54 ± 0.03
1.10	1.97 ± 0.01	2.62 ± 0.04	1.95 ± 0.01	2.54 ± 0.03

Table 2 shows the results for a set of maximum likelihood calculations with $E/\sigma_e = 10$, as used in the simulations, and for one with $E/\sigma_e = \infty$, i.e., identical to the likelihood calculations in §2. Even without taking into account the energy dispersion, the spectral results are fairly robust. This suggests that we may not need to know the energy dispersion very accurately and may be able to characterize it using a simple analytic form such as a Gaussian function. Using a Gaussian representation makes evaluation of the integral over apparent energies E' in equation 7 extremely efficient since an analytic approximation can be used to compute $\mathrm{erfc}(x)$ (NR §6.2). In neither scheme do the results demonstrate any sensitivity to the accuracy of the PSF. Figure 2 shows a sample fit to one realization of the Monte Carlo simulation using $E/\sigma_e = 10$.

4 ROI Cuts

Using linear combinations of Gaussian functions to represent the PSF allows for a relatively straight-forward accommodation of an acceptance cone ROI cut. The properly normalized two-dimensional Gaussian function in spherical coordinates is

$$p(\mu, \phi; \sigma) = \frac{1}{2\pi\sigma^2} \frac{\exp\left[(\mu - 1)/\sigma^2\right]}{1 - \exp(-2/\sigma^2)}, \tag{10}$$

where $\mu \equiv \cos\theta$.[1] Let $\theta_{\rm roi}$ be the radius of the region-of-interest, and let ψ_i be the angle between the center of the ROI and the position of source i. The fraction of the PSF, P_i, for source i that is contained within the ROI is

$$\int_{\rm ROI} P_i(\mu, \phi) d\Omega = 2\pi \int_{\mu_+}^1 P_i(\mu, \phi) d\mu - 2 \int_{\mu_+}^{\mu_-} \phi_{\rm min} P_i(\mu, \phi) d\mu \tag{11}$$

[1]Note that in the limit of small θ, equation 10 is proportional to the 2D Gaussian function in the Euclidean plane, $p(\theta) \propto \exp(-\theta^2/2\sigma^2)$, with θ as the polar radius. In all of our PSF calculations, both for the observation simulation and the likelihood analysis, we will use equation 10 to represent the 2D Gaussian functions, adopting the GLAST25 PSF parameters without modification.

where $\mu_+ = \cos(\theta_{\rm roi} + \psi_i)$, $\mu_- = \cos(\theta_{\rm roi} - \psi_i)$, and

$$\phi_{\rm min} = \cos^{-1}\left(\frac{\mu\cos\psi_i - \cos\theta_{\rm roi}}{(1-\mu^2)^{1/2}\sin\psi_i}\right). \tag{12}$$

Imposing an ROI cut, the c_i terms in $\log\mathcal{L}$ become

$$\begin{aligned} c_i &= \int_{E_{\rm min}}^{E_{\rm max}} dE\, s_i(E) \int dt\, A(E,\hat{p}_i,\vec{L}(t)) \int_{E'_{\rm min}}^{E'_{\rm max}} dE'\, D(E';E,\hat{p},\vec{L}(t)) \\ &\quad \times \int_{\rm ROI} d\hat{p}'\, P(\hat{p}';E,\hat{p}_i,\vec{L}(t)) \qquad (13) \\ &= \int_{E_{\rm min}}^{E_{\rm max}} dE\, s_i(E) \left[\int_{E'_{\rm min}}^{E'_{\rm max}} dE'\, D(E';E)\right] \\ &\quad \times \int dt \left[A(E,\hat{p}_i,\vec{L}(t)) \int_{\rm ROI} d\Omega P_i(\mu,\phi;\sigma_1(E,\hat{p}_i,\vec{L}(t)),\sigma_2(E,\hat{p}_i,\vec{L}(t))\right] \qquad (14) \end{aligned}$$

Equation 13 is the general expression for a non-varying source, whereas equation 14 is specialized to the present case (i.e., $\sigma_e = E/10$, independent of source inclination, etc.). Using the two-Gaussian PSF representation

$$P_i(\mu,\phi;\sigma_1,\sigma_2) = w\, p(\mu,\phi;\sigma_1) + (1-w)\, p(\mu,\phi;\sigma_2), \tag{15}$$

we have

$$\int_{\rm ROI} P_i(\mu,\phi) = w \int_{\rm ROI} p(\mu,\phi;\sigma_1) + (1-w) \int_{\rm ROI} p(\mu,\phi;\sigma_2), \tag{16}$$

where

$$\begin{aligned} \int_{\rm ROI} p(\mu,\phi;\sigma) &= \frac{1}{1-\exp(-2/\sigma^2)} \left\{1 - \exp\left(\frac{\mu_+ - 1}{\sigma^2}\right)\right. \\ &\quad \left. -\frac{1}{\pi\sigma^2}\int_{\mu_+}^{\mu_-} d\mu \exp\left(\frac{\mu-1}{\sigma^2}\right) \cos^{-1}\left(\frac{\mu\cos\psi_i - \cos\theta_{\rm roi}}{(1-\mu^2)^{1/2}\sin\psi_i}\right)\right\}. \qquad (17) \end{aligned}$$

The last expression is rather unwieldy and requires numerical integration. Since it must be evaluated twice for each source at each time step, it is useful to compute the μ-integral for each source i on a grid of σ values prior to integrating over time. The values of the μ-integrals can then be interpolated on this grid when performing the time integral. If the source positions are to be varied, then the grid must cover the σ-ψ plane.

Table 3: ROI Cut Tests

$\theta_{\rm roi}$	$\Gamma_{\rm 3C\ 279}$	$\Gamma_{\rm 3C\ 273}$	Notes
15°	1.96 ± 0.01	2.66 ± 0.04	See Figure 1
20°	1.96 ± 0.01	2.63 ± 0.04	”
30°	1.97 ± 0.01	2.63 ± 0.03	”
∞	1.97 ± 0.01	2.62 ± 0.03	Same as §3 results
20°	1.94 ± 0.01	2.52 ± 0.03	PSF integrals set to unity

In Table 3, results for the ROI calculations are presented. The first three entries have ROI cuts applied to both the data and the likelihood calculations. The fourth entry has no ROI cut in either case and is the same as the results reported in §3 (with a PSF scale

factor of unity). The final entry has a $\theta_{\rm roi} = 20^\circ$ cut applied to the data, but the calculation of the c_i's from §3 is used, i.e., $\theta_{\rm roi}$ is set to infinity in the likelihood evaluations. Because the Gaussian widths of the PSFs go roughly as $\sigma \propto E^{-3/4}$, ROI cuts preferentially remove lower energy photons for sources within the ROI. Accordingly, the fitted spectra for these sources are biased to be harder for that final entry. For these tests, the ROI was centered on 3C 279. Since 3C 273 is closer to the edge of the ROI, a larger fraction of its soft photons are removed by the ROI cut, and its fitted spectrum is hardened to a greater degree than that of 3C 279.

5 To Do

The tests discussed here are extremely simple. Nonetheless, they proved to be useful in shaking out some bugs in the implementation. Some next steps, roughly in order:

- Test other optimization methods. Brent's implementation, PRAXIS.F, of Powell's method is currently used. Methods that use derivatives of the objective function as well as Peter Freeman's version of Powell that accounts for upper and lower limits on parameter values will be investigated.
- Diffuse emission.
- More complex source spectra.
- Other extended sources.
- Fit for source position.

References

[1] Chiang, J., 2002, `http://lheawww.gsfc.nasa.gov/~jchiang/SSC/like_0.ps` (LikeMemo 0)

[2] Numerical Recipes 1992 (NR)

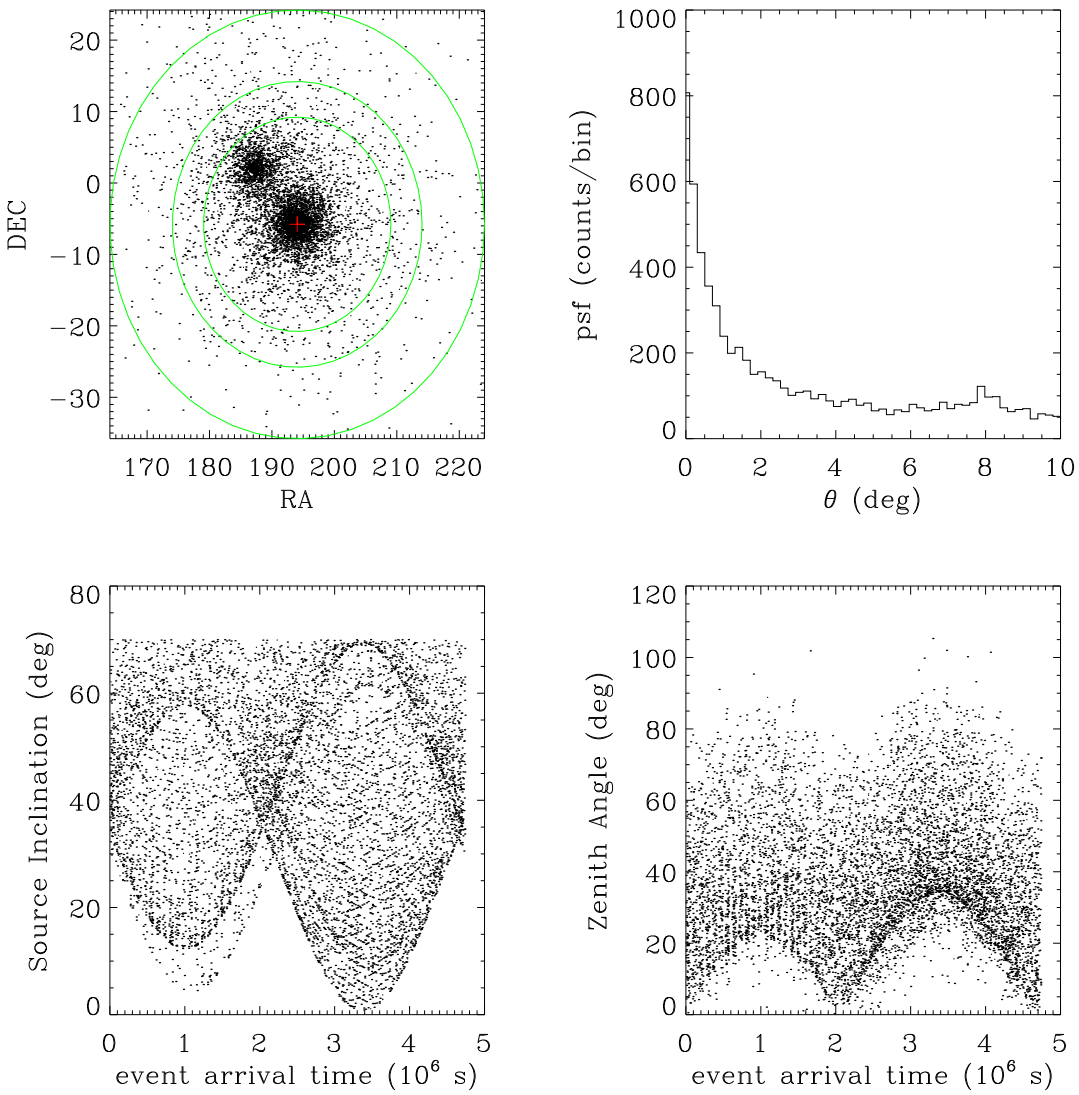

Figure 1: Photon data from an 826 orbit simulation of 3C 279 (red cross in upper left plot) and 3C 273. The green circles are (roughly) the 15°, 20°, and 30° ROI acceptance cones used in §4. The PSF profile centered on 3C 279 is shown in the upper right plot. Photons from 3C 273 compose the bump at $\theta = 8°$.

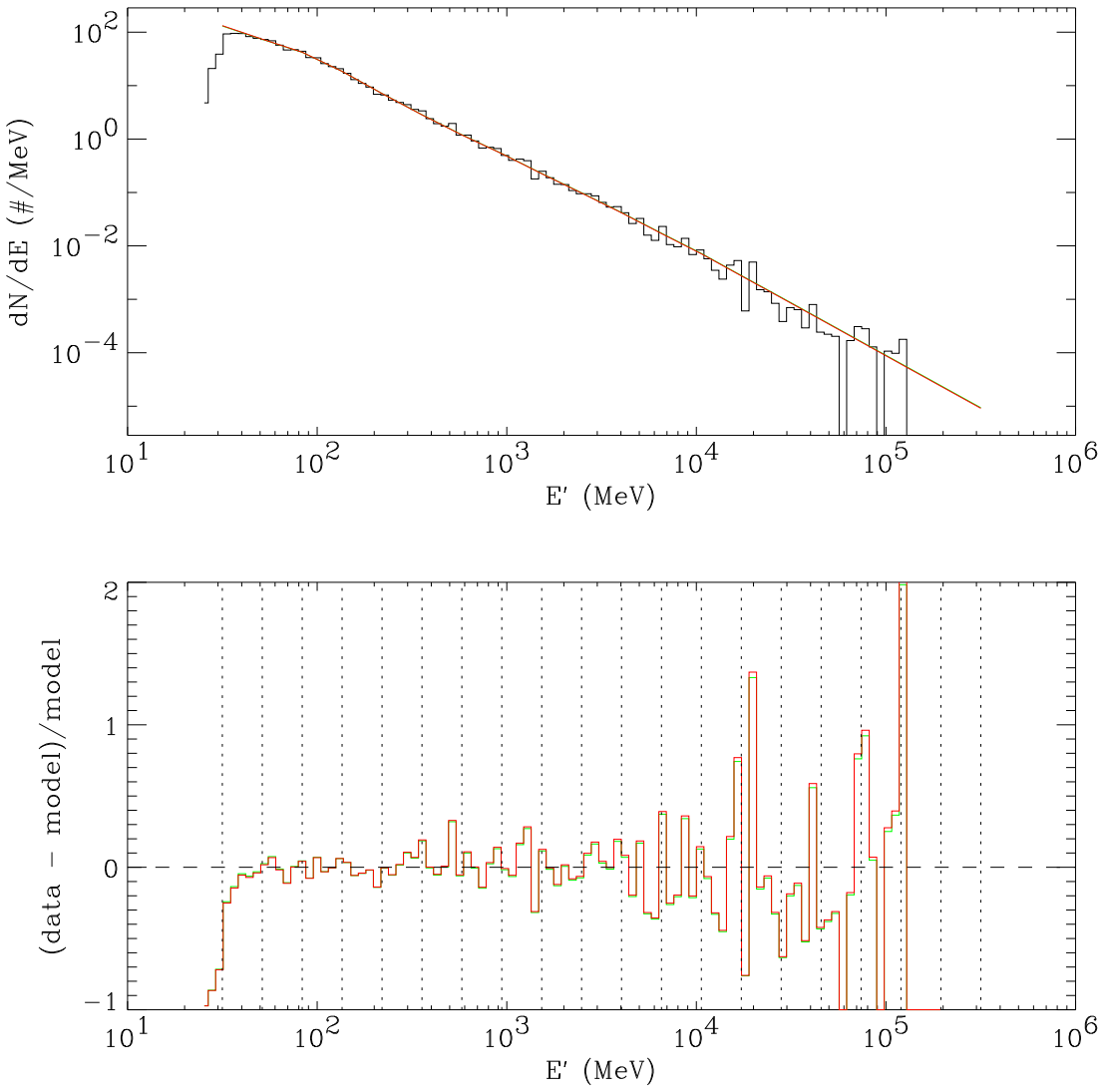

Figure 2: The histogram in the upper plot is the counts spectrum for the entire data set. The red curve is the spectrum obtained by convolving the model through the instrument response using the parameters from the maximum likelihood calculation and the green curve (almost completely obscured by the red curve) is the model obtained using the input parameters for the simulation. The lower plot shows the relative magnitude of the residuals. The dotted vertical lines show the abscissa locations of the energy integral quadratures.