# LAT Automated Science Processing for Gamma-Ray Bursts

James Chiang

GLAST Science Support Center SLAC, 2575 Sand Hill Rd, Menlo Park CA 94025

Abstract. The LAT Instrument Science Operations Center (ISOC) will perform various tasks to support coordination of multiwavelength observations for transient sources. In this paper, we describe the prototype implementation of the Automated Science Processing (ASP) for the detection and analysis of gamma-ray bursts (GRBs) in LAT and GBM data. The GRB-related tasks include: position refinement using LAT data given initial GBM or GCN locations, spectral analysis using LAT data alone, joint spectral fitting with GBM data, gamma-ray afterglow detection and characterization, and blind searches for prompt burst emission in LAT data.

## **IMPLEMENTATION OVERVIEW**

These analyses will be performed on the ground as part of the Level-2 processing in the ISOC data pipeline. Since the event and spacecraft data will have been converted to FITS format in the Level-1 processing, we may use LAT ScienceTools and HEASARC FTOOLs to perform the bulk of the ASP analyses, with Python scripts driving the various ScienceTool and FTOOL applications.

In addition to the standard suite of LAT analysis software, some custom analysis tools are required. In order to determine the start and stop times for the prompt GRB emission in the LAT data, we apply the Bayesian Blocks algorithm [1] to the LAT event arrival times. This algorithm has been implemented in C++ and exposed to Python using SWIG [6]. For the GRB blind search task, we apply a version of the algorithm that has been developed for the on-board GRB detection in the LAT flight software. We will describe the ASP version of this algorithm in the final section of this paper. As with all of the ISOC processing, from Level-0 through Level-2, the ASP analysis task execution will be managed by the ISOC Pipeline-II.

## **POSITION REFINEMENT**

GRB detections by the GBM or other sources such as Swift or the Interplanetary Network will arrive at the LAT ISOC in the form of GCN Notices [3]. The GRB start time from these Notices will be used to determine the time interval over which the LAT data for the burst will be extracted. The LAT extraction time interval is 200 seconds bracketing the GCN burst start time. This will allow the ASP analysis to capture any high energy gamma-ray precursor that may be present on a 100 second time scale. The Bayesian Block algorithm is run on these data, and the first and last change points are

Contributed to Symposium on Gamma-ray Bursts in Connections with the GLAST-LAT Collaboration Meeting, 09/01/2006--9/1/2006, Stockholm, Sweden



**FIGURE 1.** Left: Counts map of a simulated LAT burst from DC2. The black cross and circle are the GBM location and error circle. The red cross indicates the LAT best-fit location. Center: Counts light curve. The yellow region indicates the burst interval as determined by the Bayesian Blocks analysis. Right: Test-statistic (TS) map showing the best-fit position, and the 68, 90, and 99% C.L. contours.

adopted as the burst start and stop times in the LAT data. The data are then re-extracted for this subinterval and fed to a program that performs the position refinement.

In the source position refinement, a point source modeled with a power-law photon spectrum is posited at each candidate source location, and the algorithm minimizes the negative log-likelihood of that point source over the sky coordinates to find the optimal source location. At each candidate position, the photon index and flux are fit using the energy-dependent point spread function (PSF) and instrument effective area in the likelihood analysis. Since the extraction interval will typically be tens of seconds, in this and all subsequent prompt burst analyses, we ignore any diffuse emission from Galactic or extragalactic sources as well as charged particle backgrounds. These latter sources of background will be significantly curtailed by the on-board filter and by the ground-processing cuts that are designed to reduce the residual charged particle background to be an order-of-magnitude smaller than the extragalactic diffuse emission measured by the EGRET instrument [5]. These ground-processing cuts distinguish the ASP analysis from an on-board analysis since the residual backgrounds will be much larger for the on-board data.

Figure 1 shows the results of the GRB refinement task applied to a simulated burst in GLAST DC2 data.<sup>1</sup> The leftmost figure is a counts map of the LAT events that were associated with the simulated burst. The black cross and circle show the GBM location estimate and error circle, and the red cross shows the best fit position from the ASP position refinement analysis. The center plot shows the counts light curve for this burst, and the yellow region indicates the burst duration found from the Bayesian Blocks analysis. The rightmost plot shows the best fit position and error contours calculated from the LAT data.

<sup>&</sup>lt;sup>1</sup> DC2 is the second Data Challenge conducted by the LAT collaboration. In these exercises, realistic models of the gamma-ray sky are created and run through a complete simulation of the instrument in orbit, including charge particle backgrounds, GEANT simulation of particle interactions, reconstruction, etc..



**FIGURE 2.** Left: Counts spectrum and power-law model fit to simulated LAT prompt burst data from DC2. Right: Fractional residuals from the fit.

### SPECTRAL ANALYSIS

Once the prompt burst duration and position have been ascertained via the preceding analysis, the data are again re-extracted using the Bayesian Blocks start and stop times and within a 15 degree acceptance cone centered on the best-fit source position. The size of the acceptance cone is nominal here and may change as the analysis is refined with further testing. A standard unbinned likelihood analysis is then applied to these data to determine the best fit spectral parameters. Figure 2 shows the results of fitting a power-law spectrum to the LAT data for this burst. The best fit flux for this burst is  $6.3(\pm 1.1) \times 10^{-3}$  photons cm<sup>-2</sup>s<sup>-1</sup> over the energy range 20 to  $2 \times 10^{5}$  MeV, and the best-fit photon index is  $\Gamma = -2.0 \pm 0.1$ .

When GBM data are available, joint fits will be performed using Xspec [7]. The broader combined spectral range afforded by the GBM and LAT, from 8 keV to 200 GeV, will allow more complex spectral models, such as the Band function, to be fit.

## GAMMA-RAY AFTERGLOW ANALYSIS

The afterglow analysis proceeds in a manner similar to standard LAT point source analyses. The time scales for the afterglow analysis are sufficiently long that diffuse emission from the Galactic and extragalactic components must be included in the analysis. In addition, it will likely be necessary to account for the residual charged particle background; and if any strong point sources are nearby, e.g., the Vela pulsar, those sources must also be modeled.

A 15 degree acceptance cone centered on the best fit position from prompt burst analysis is used for the spatial event selection, and an interval starting at the end of the prompt burst duration and ending 5 hours later is used for the temporal extraction.



**FIGURE 3.** Left: Counts spectrum and power-law model fit to simulated LAT afterglow burst data from DC2. The uppermost solid curve is the sum from the three spectral components in the model. The next highest curve (at 1 GeV) is the Galactic diffuse emission, followed by the extragalactic diffuse emission and the burst afterglow emission. Right: Fractional residuals from the overall fit.

Note that this time duration and acceptance cone size for the afterglow data are nominal and will be adjusted as our knowledge of any gamma-ray afterglow component grows. It will also be necessary to perform this analysis on a number of different time scales and intervals in order to characterize any spectral variability in the afterglow emission. In figure 3, we show the results of a spectral analysis of the afterglow emission for our simulated burst.

### **BLIND SEARCHES IN LAT DATA**

A blind search algorithm for GRB emission will be applied to all Level-1 data in the ISOC processing, and specifically, it will be applied to data obtained from each satellite downlink, spanning roughly 3 hours. There will be some buffering of the data from previous neighboring downlink to capture bursts that may straddle the downlink boundaries. Furthermore, this analysis will be in addition to GBM/GCN triggered analyses and any on-board detection scheme. As noted, the blind search algorithm that is used is based on the on-board algorithm [2], but the particle background rates for the on-board analysis will be substantially higher.

The ASP algorithm begins by partitioning the data into consecutive, non-overlapping groups of typically 20 to 100 events. The size of the data segments will be adjusted based on our growing knowledge of GRB prompt emission in the LAT band. However, 20 will likely be the minimum to ensure statistical significance for a detection associated within a given data segment. For each data segment, the log-likelihood that a burst is present is evaluated and is based on the spatial and temporal properties of the events.

To assess the likelihood of clustering in time, we compute the log-likelihood under the null hypothesis, i.e., that the event rate is constant in time:

$$\log \mathscr{L}_{t} = \sum_{i} \log(1 - \exp[-r_{bg}(t_{i+1} - t_{i})])$$
(1)

Here the  $t_i$ s are the event arrival times. The quantity  $r_{bg}$  is the mean event rate over an appropriate time interval; for the ASP, we use the event rate averaged over the current downlink.

For the spatial part, we consider two methods for computing the contribution to the log-likelihood. The first follows the implementation used in the on-board version. In this method, a cluster finding algorithm is applied to the events in a given data segment. Each measured photon direction is initially treated as a candidate burst direction, and the number of events lying within an acceptance cone radius of  $17^{\circ}$  about that direction is used as a measure of the clustering. <sup>2</sup> The mean direction for events in the largest such cluster is computed, and the log-likelihood under the null hypothesis (no clustering) is evaluated using

$$\log \mathscr{L}_p = \sum_i \log(1 - \hat{n} \cdot \hat{n}_i) \tag{2}$$

where  $\hat{n}$  is the candidate source position and  $\hat{n}_i$  are the measured directions for events in that cluster.

The second method uses the energy-dependent LAT PSF to evaluate the log-likelihood that there is in fact a point source at a candidate location:

$$\log \mathscr{L}_p = -\max_j \left[ \sum_{i \neq j} \log P(\hat{n}_j \cdot \hat{n}_i; E_i, \theta_j) \right]$$
(3)

where P(...) is the PSF from the LAT instrument response functions,  $E_i$  is the measured energy of the photon, and  $\theta_j$  is the angle between the candidate source direction and the LAT z-axis. For this method, we simply use each photon direction,  $\hat{n}_j$ , as the candidate photon direction and form the sum over all other directions  $\hat{n}_i$  in the data segment. Note that since we are considering the alternative hypothesis that there is a point source, we take the *negative* maximum log-likelihood. For either method, the total log-likelihood is then

$$\log \mathscr{L} = \log \mathscr{L}_t + \log \mathscr{L}_p \tag{4}$$

Table 1 lists the GRB-like transients that were included in the DC2 simulations and that had > 5 events detected in the LAT. All but three of these transients are GRBs that triggered in the GBM. The other three comprise a "primordial black hole" (PBH) that produced photons only in the LAT energy range, an X-class Solar Flare that was comparable to the 1991 June 11 event detected by EGRET [4], and a GRB that did not trigger the GBM. Also listed in the table are the -log-likelihood values

 $<sup>^2</sup>$  This acceptance cone radius has been optimized for the on-board algorithm, and we simply adopt it here. However, since the background rates differ greatly between this and the on-board analysis, further study may reveal a better value for the ground-based analysis.

events.				
Time (MET)	Туре	N <sub>det</sub>	$-\log \mathscr{L}_{cone}$	$-\log \mathscr{L}_{\rm psf}$
221142014	GRB	101	254	185
221260539	GRB	12	149	64
221276802	PBH	15	205	169
222322394	GRB	12	154	48
222367682	GRB	37	282	240
222801583	GRB	13	<112	<19
223103093	Solar Flare	4669	229	150
223132662	GRB	11	131	30
223711708	GRB	6	<112	<19
224077473	GRB	407	320	268
224596229	GRB	36	225	155
224910123	GRB*	38	205	203

**TABLE 1.** Simulated GRB-like transients with > 5 LAT-detected events

\* untriggered in GBM

that were obtained using the two different spatial analysis methods. Figure 4 shows the distributions of log-likelihood values for these two methods. The trigger thresholds (dotted vertical lines) correspond roughly to 5-sigma detections.

Strong Solar flares such as the 1991 June 11 event present a special case as the duration of these flares is of order hours, and so the log-likelihood for such flares would exceed a nominal threshold for GRB detection for many data segments, the size of which are tuned to the time scales of GRBs. In order to handle such bright, longer time scale transients, we impose a trigger condition which is only reset after the log-likelihood goes back below threshold and after an artificially imposed deadtime (10<sup>3</sup>s nominally). The latter condition ensures that a flare which peaks close to threshold does not cause too many triggers. Figure 5 illustrates these conditions.

Note that the plots in Figure 4 include the log-likelihood values for all data segments (except those associated with the Solar flare), and so there will be more events in the tail than there are transients listed in Table 1 because a strong burst (e.g., with > 20 LAT-detected events) will typically span more than one data segment. However, in order not to have a too misleading number of events in the tails, the plots in Figure 4 exclude the log-likelihood values for those associated with the epoch of the Solar flare (223102000 < t < 223106000 MET).

#### REFERENCES

- 1. Jackson, B., et al., IEEE Signal Processing Letters 12, 105-108 (2005).
- Bonnell, J. T., & Norris, J. P., http://glast.gsfc.nasa.gov/science/grbst/pubs/ GRBTrigsLocs\_2.pdf (2004)
- 3. http://gcn.gsfc.nasa.gov/
- 4. Kanbach, G., et al., A&ASS 97, 349-353 (1993).
- 5. Sreekumar, P., et al., ApJ 494, 523-534 (1998).
- 6. http://www.swig.org/
- 7. http://heasarc.gsfc.nasa.gov/docs/xanadu/xspec/



**FIGURE 4.** Plots of the log-likelihood distributions (equation 4) applied to the DC2 data with 20 events per data segment. The dotted vertical lines indicate the thresholds used to identify candidate transient events. The thresholds shown correspond roughly to 5-sigma detections based on Gaussian fits to the main peak of the respective distribution. Left: Acceptance cone-based likelihood, threshold = -112. Right: PSF-based calculation, threshold = 19.



FIGURE 5. The acceptance cone log-likelihood calculation during the epoch of the DC2 Solar flare.