Statistical Issues in High-Energy Gamma-Ray Astronomy for GLAST

S. W. Digel
Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA 94305, USA

This paper describes the statistical issues involved in analyzing data from high-energy gamma-ray telescopes, at levels from event reconstruction to correlations of populations of astrophysical sources. Some motivation for attempting to do astronomy with high-energy gamma rays is also given, along with some of the constraints implied by operating the instrument in orbit. Specific attention is given to the Large Area Telescope (LAT) under development for launch in late 2006 on the Gamma-ray Large Area Space Telescope (GLAST) mission.

1. INTRODUCTION

Gamma-ray astronomy has developed only relatively recently owing to many technical challenges in detecting gamma rays. In the energy range below ~50 GeV, gamma-ray detectors must be in space. (At higher energies, air showers from interactions of gamma rays with the upper atmosphere can be detected from the ground.) The missions that have flown to date with sensitivity in the >20 MeV range (Table 1), in particular the Energetic Gamma-Ray Experiment Telescope (EGRET) on the Compton Gamma-Ray Observatory, have revealed a remarkable variety of astrophysical sources of high-energy gamma rays, and plausible prospective source classes remain to be discovered with future missions, in particular the Large Area Telescope (LAT) on the Gamma-ray Large Area Space Telescope (GLAST) [1, 2], which promises a great increase in sensitivity.

1.1. Sources of Celestial Gamma Rays

Astrophysical sources of gamma rays are nonthermal, accelerating particles in shocks, e.g., in jets and supernova remnants, and in the intense fields of pulsars (rotating magnetized neutron stars). Gamma rays are produced in these sources by high-energy electrons via bremsstrahlung scattering on nucleons, or inverse Compton scattering low-energy photons, or in the case of sufficiently strong magnetic fields via synchrotron or curvature radiation. They are also produced in pion decay from high-energy proton-nucleon interactions. The universe is essentially transparent to gamma rays, and their observation provides a unique, direct probe of these processes in nature. (At ~TeV energies and cosmological distances, attenuation does occur by γ-γ interactions on the cosmic microwave background.) The known or prospective classes of celestial gamma-ray sources for the next generation of instruments are described briefly below.

1.1.1. Diffuse gamma-ray emission

Interactions of Galactic cosmic rays with interstellar gas and low-energy photons make the Milky Way itself a diffuse source of high-energy gamma rays. The intensity is greatest at low Galactic latitudes, owing to the concentration of the interstellar gas and sources of cosmic rays in the spiral arms of the Milky Way.

1.1.2. Active galaxies

With EGRET, a class of active galaxies called blazars was discovered to be a powerful source of gamma rays [3, 4]. Approximately 70 were identified in EGRET data. The generally-accepted interpretation is that blazars have jet emission associated with a massive black holes in their nuclei, and that the jets are closely aligned with the line of sight to the earth. Other active galaxies, with less favorable alignment of their jets, are also known gamma-ray sources, although much less intense; the only example from EGRET is Centaurus A [5].

1.1.3. Gamma-Ray Bursts

These are extremely bright, short-lived sources, most of which have been identified with some kind of cataclysmic explosions in star-forming galaxies at cosmological distances.

1.1.4. Pulsars

In the Milky Way, a subset of rotation-powered pulsars comprise a well-established class of gamma-ray sources, with approximately 9 identified in the EGRET data. The usual method of discovery is to phase fold the gamma rays according to timing information derived from radio observations (see Sec. 5.2). At least one gamma-ray pulsar, Geminga, is not a radio pulsar; searching for periodicity without a timing ephemeris is discussed in Section 5.2.

1.1.5. Other classes

In general, Galactic sources are associated with tracers or remnants of massive star formation—pulsars (possibly radio quiet), binary pulsars, millisecond pulsars, supernova remnants, plerions (filled center supernova remnants powered by a pulsar), OB/WR associations, microquasars, microblazars, and isolated black holes have all been proposed as sources of gamma rays in the Galaxy.

Table I High-Energy Gamma-Ray Astronomy Missions

Instr.	**Years**	$\theta_{0.1}$[a]	θ_{10}[b]	**Energies (GeV)**	$A_{eff}\Omega$ **cm^2 sr**	**No. γ-Rays**
OSO-3	1967–68	18°	-	>0.05	1.9	621
SAS-2	1972–73	7	-	0.03–10	40	$\sim 1\times 10^4$
COS-B	1975–82	7	-	0.03–10	40	$\sim 2\times 10^5$
EGRET	1991–00	5.8	0.5°	0.03–10	750	1.4×10^6
AGILE	2005–	4.7	0.2	0.03–50	1500	4×10^6/yr
AMS	2005[c]–	-	0.1	1–300	500	$\sim 2 \times 10^5$/yr
LAT	2007–	3.5	0.1	0.02–300	25,000	1×10^8/yr

[a] Angular resolution at 0.1 GeV
[b] Angular resolution at 10 GeV
[c] Scheduled for the 16th shuttle mission once launches resume.

Outside the Milky Way, Galaxy clusters and starburst galaxies are prospective new classes of gamma-ray sources. EGRET detected the Large Magellanic Cloud in the light of its diffuse emission.

EGRET detected impulsive GeV emission from intense, X-class flares that occurred near solar maximum in 1991 [6].

Figure 1 shows the gamma-ray sky seen by EGRET and Figure 2 shows a simulated all-sky image from the planned one-year sky survey with the LAT. Projections are that the LAT will detect several thousand gamma-ray sources. Owing to the scanning coverage of the sky survey, the LAT will also provide extremely well sampled light curves.

1.2. History of high-energy gamma-ray astronomy in space

Celestial gamma rays were first detected by OSO-3, which saw the diffuse emission of the Milky Way in 1968 [7], and there have been three missions since of increasing size and resolution. Table I summarizes the past and upcoming missions; all have used pair conversion trackers (see Sec. 2). OSO-3 used plastic scintillators as the tracking material. SAS-2, COS-B, and EGRET had wire grid spark chambers for tracking. Upcoming missions will use silicon strip detectors, which have the advantages of finer pitch (for better angular resolution), orders of magnitude faster readout (for limiting the dead time), and no reliance on expendables (like spark chamber gas).

Three missions in Table I are under development. The LAT, under development for launch by NASA in early 2007 on the GLAST mission, will provide a great increase in sensitivity . The development of the LAT is being supported by NASA, DOE; CEA and IN2P3 in France; ASI, CNR, and INFN in Italy; and institutions in Japan and Sweden. Astro-rivelatore Gamma a Immagini LEggero (AGILE), under development by ASI [8], is a smaller instrument of generally similar design that is planned for launch approximately one year before GLAST. The Alpha Magnetic Spectrometer (AMS) is a cosmic-ray experiment to be launched on the space shuttle for installation on the ISS [9]. It will have sensitivity to gamma rays in the range >1 GeV.

AGILE and the LAT have very large fields of view relative to preceding instruments because they will not rely on time-of-flight (TOF) scintillators below the tracking section to discriminate upward-moving, i.e., background, events. The field of view is limited because events must cross both the tracker and the TOF system in order to be accepted. The trade off is a greater event rate, and the need to rely on post processing to reject upward-moving events.

2. DETECTION OF HIGH-ENERGY GAMMA RAYS

At X-ray energies, photons can be focused with grazing incidence mirrors, but gamma rays cannot be focused similarly. The collecting area of a gamma-ray telescope is therefore directly related to the physical size of the detector, which is not the case in X-ray astronomy.

X-ray detectors more or less can count individual X-rays. Gamma-ray detectors convert the gamma rays to positron-electron pairs, then track their trajectories through the instrument (see Fig. 3) and measure their energies to infer the directions and energies of the gamma rays. Most conversions of the gamma rays happen in heavy metal (W in the case of the LAT) foils interleaved with the tracking layers (silicon strip detectors for the LAT). The trade off for including conversion foils to increase the probability of conversion is that the electron and positron tend to scatter on passage through the foils in subsequent layers, decreasing the accuracy with which their directions and energies can be determined. As a result, gamma-ray telescopes have much poorer angular resolution than X-ray instruments, typically measured in degrees rather than arcseconds (Table I).

The LAT has a modular design, arranged as a 4×4 grid of independent towers, each with a tracker (TKR) and calorimeter (CAL) section (Fig. 3). The TKR section of each tower has 18 tracking planes, each with two layers of silicon strip detectors, one for measuring x coordinates and the other for y, with W foils interleaved between the planes. The CAL section of each tower has CsI(Tl) crystals arranged as a hodoscope: 8 layers of 12 crystals each, with the orientations of the layers alternating between the x and y directions. Each end of each log is instrumented with PIN photodiodes to detect the scintillations. The anticoincidence detector (ACD) surrounds the top and sides of

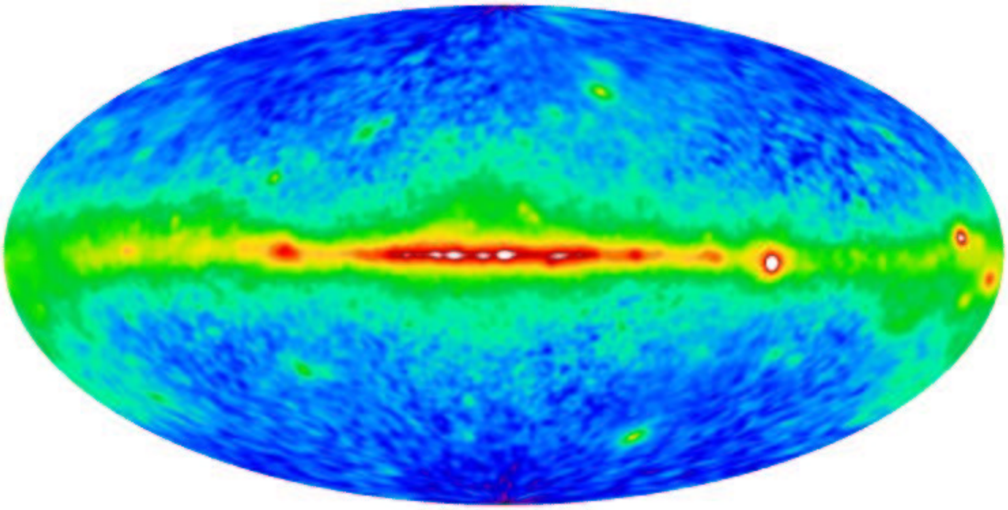

Figure 1: Intensity of gamma-ray emission >100 MeV observed by EGRET, displayed in false color. The Aitoff projection is in Galactic coordinates, and the bright band across the center of the image is the diffuse emission from the Galactic plane. The bright point sources at low latitude are rotation-powered pulsars. Many of the bright sources removed from the plane are blazars.

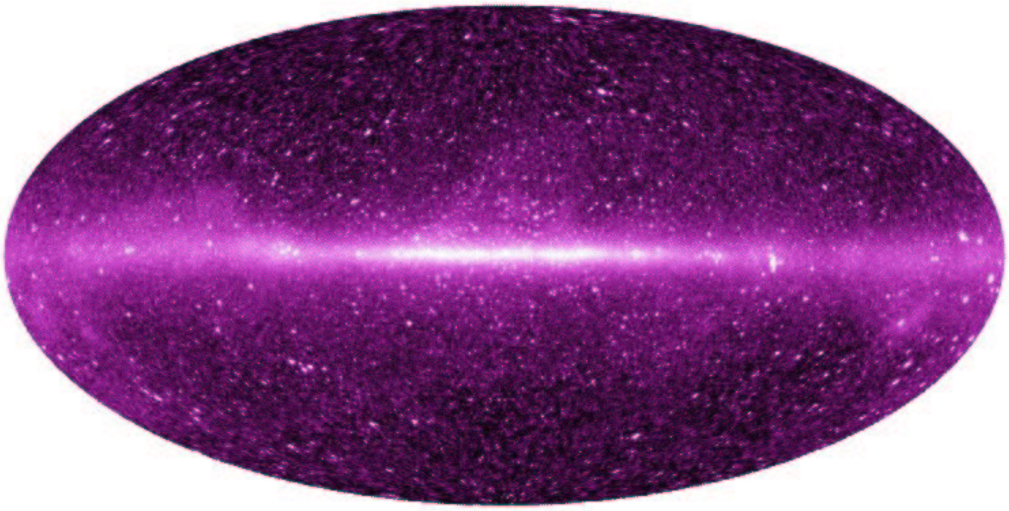

Figure 2: Simulated gamma-ray intensity observed by the LAT during the planned one-year sky survey. The energy range shown here is >1 GeV, where the angular resolution and effective collecting area of the LAT are much greater than for EGRET. The model of the sky includes the cataloged EGRET sources as well as populations of fainter sources and the diffuse emission of the Milky Way [10]

the array; it registers the passage of charged particles and therefore is used in anticoincidence with the TKR and CAL in forming the trigger for gamma rays.

One advantage of a pair production tracking detector is that the field of view is enormous, ∼2.2 sr for the LAT. The LAT can observe many targets simultaneously and does not need to be pointed at a particular target. In fact, to increase the overall observing efficiency, the standard operating mode will have the LAT continuously scan the sky. This avoids the loss of observing time due to earth occultations and also limits the need to detect (and reject on board) the bright background of albedo gamma rays from interactions of cosmic rays in the upper atmosphere.

2.1. Design Issues for LAT Data Handling

A number of design compromises must be made for a gamma-ray detector to be operated in space. Most importantly, the collecting area is limited by the size of the rocket fairing, the mass is limited by the lift capacity, the power by the feasible solar cell and radiator capacities, and the data rate to the ground by allocations of telemetry bandwidth. The charged par-

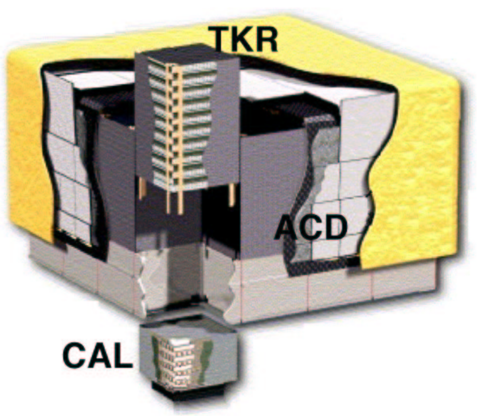

Figure 3: Cutaway view of the LAT. One of the sixteen towers is shown with its TKR module on top of the CAL module. The ACD is an array of plastic scintillator tiles that cover the towers. Surrounding the ACD is a thermal blanket and micrometeoroid shield. The overall dimensions are 1.8 × 1.8 × 0.75 m.

ticle background in orbit is intense; the orbit-averaged trigger rate will be approximately 3 kHz for the LAT. The actual rate of triggers from celestial sources will be $\sim$2 Hz. The telemetry bandwidth is sufficient to send event data ($\sim$10 kbits per event) at an average rate of $\sim$30 Hz, so efficient filtering of the data in flight is essential.

The combination of a signal:background event rate ratio of $< 10^{-3}$, the need to reconstruct gamma-ray information from tracks and energy depositions in the LAT, the resulting limited angular resolution, the bright and structured diffuse gamma-ray emission from the Milky Way, and low fluxes of celestial point sources provide ample motivation for careful treatment of the data at every step of the analysis of LAT data.

3. LOW-LEVEL ANALYSIS

3.1. Nature of the Data

Readout of the LAT is triggered by the occurrence of hits in 3 successive X-Y TKR planes in tower, or a large energy deposition in the CAL. Simple, robust algorithms are used to filter the data. As described above, on board filtering of the data is required by the available average telemetry rate. The data are lists of the s that were hit, measurements of light output from the ends of the CsI(Tl) logs in the CAL, and a list of the tiles of the ACD that were hit.

3.2. Event Reconstruction

In ground processing, reconstruction of events (interactions of a cosmic ray or gamma ray in the LAT) starts with grouping the hits (silicon strips that registered a charged particle that trigger a readout) in the TKR into clusters, because adjacent strips can register the passage of the same particle. A pattern recognition algorithm is applied to associate the clusters into ‘tracks’, with preference for finding the longest, straightest tracks. The current algorithm is combinatorial (i.e., brute force).

The identified track or tracks are then fit via Kalman filtering (e.g., [11]). This defines the best estimate of the initial direction of the charged particle. This process is iterative with analysis of the energy depositions in the CAL, which is used for estimating the overall energy of the event. The energy information is used to evaluate the scattering angles expected in each tracker plane. Multiple scattering in the W conversion foils is quite non-Gaussian. In principle, this is a problem for the method. However, the uncertainty in the energy determination is great enough that it (rather than the non-Gaussian tails of multiple scattering) dominates the uncertainty of estimated scattering angles. The assumption implicit in the current analysis is that in these circumstances, the Kalman filtering method is applicable; more study of the validity probably would be prudent. An approach for track reconstruction that uses concepts from particle filtering is being investigated as a potential alternative [12]. It is more challenging computationally but should be able to explicitly take into account processes like multiple scattering.

Reconstructed tracks are analyzed to define the conversion point of the gamma ray and its initial direction. An example reconstruction is shown in Figure 4. At higher energies, the positron and electron tracks may not separate at the resolution of the tracker, so the vertex and the estimated initial direction come from analysis of a single track.

The energy deposition in the CAL in general must be corrected to account for partial containment of the showers. The CAL is only 8.5 radiation lengths deep (owing to the constraint on the mass that can be placed in orbit), and even at moderately large inclination angles, significant corrections are required. Two approaches are being evaluated, shower profiling and last-layer correlation. The development of the showers in the CAL can be reconstructed with coarse resolution, owing to the hodoscopic design discussed in Sec. 2. Intrinsic fluctuations in energy deposition as showers develop limit the resolution achievable by these techniques. At energies >100 GeV, typical energy containment is less than 40%; showers are still developing at the point that they leave the CAL and any correction scheme necessarily involves a large extrapolation.

If multiple tracks are found, the best (straightest, highest energy) tracks are checked for intersection. The estimated energies and initial directions of the two tracks are used to calculate the energy and direc-

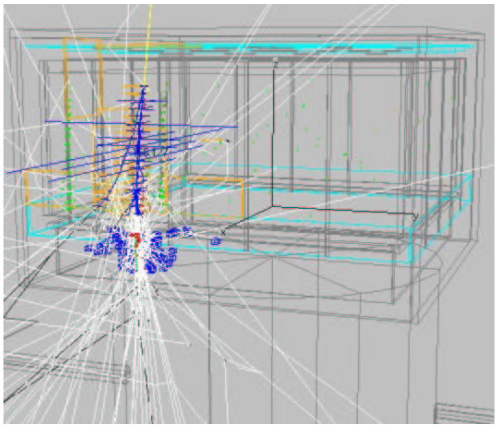

Figure 4: Simulated interaction of a 1 GeV gamma ray in the LAT. The LAT is indicated by a wire frame outline that also includes a schematic, cylindrical spacecraft. The reconstructed tracks are indicated in blue. The white lines are soft (X-ray) photons and the ACD tiles that are hit are outlined in orange. The CsI logs in the CAL with significant energy depositions are also indicated. Courtesy T. Usher.

tion of the incident gamma-ray.

3.3. Event Classification

Once an event is reconstructed, the final classification needs to be made. Fundamentally, the classification discriminates between charged particles and gamma rays, although sub-classifications also will be made. For example, heavy cosmic rays that do not undergo nuclear interactions in the CAL will be flagged for use in calibration of the CAL in flight. Also, gamma rays with especially well-measured energies or directions will be flagged.

Through extensive Monte Carlo simulation of the instrument, informed by beam tests of prototypes, useful diagnostics for discriminating cosmic rays from gamma rays have been identified. By far the most powerful is the intersection of the reconstructed event direction with a tile of the ACD that recorded a hit (passage of a charged particle). Other cuts are not as obvious, and not completely orthogonal. The production of a 'clean' gamma-ray data set is vitally important, owing to the orders-of-magnitude greater intensity of cosmic rays than celestial gamma rays. Currently, the classification of events is implemented using a classification tree trained with simulated data. Each node in the tree applies a single test (e.g., one based on projected distance to the nearest ACD tile that was hit). The result from traversing the tree is the probability that the event is a gamma ray; the probabilities are defined from the results of passing simulated events through the tree.

Decision trees are also used to identify the events that probably have well measured energies and directions. As mentioned above, multiple scattering in the tracker unavoidably causes long 'tails' in the point-spread function (PSF). The tails can confound the analysis of sky regions of high source density or intense diffuse emission, and to the extent that the events in the tails can be identified (and ignored) the cost to the effective collecting area may be worth the trade off for these circumstances.

One issue with this approach is the stability of the classification trees. For a classification tree analysis, a small change in input quantities can dramatically affect the path through the tree, and the resulting classification. More recently developed methods, such as boosted decision trees (e.g., Friedman, this volume), do not suffer this shortcoming and may be adopted for the classification of LAT events.

4. HIGH-LEVEL ANALYSIS

High-level analysis of LAT data is gamma-ray astronomy, the detection and characterization of celestial sources of gamma rays. Generally, for the reasons described above, the characterization will be via model fitting, where the parameters of a model quantify what we are trying to learn from the data. This approach has a long history in gamma-ray astronomy. Pollock et al. [13] introduced the maximum likelihood method for model fitting for analysis of data from COS-B, and the same approach was used for EGRET [14]. Technically, the method is extended maximum likelihood, because the number of gamma rays is itself random variable.

For likelihood analysis, the detector is represented by its response functions, high-level descriptions of how the point-spread function, energy resolution, and effective collecting area depend on energy, direction (relative to the instrument coordinate system), and other measurable quantities, like plane of conversion of the gamma ray, and the results of the event classification trees. This high-level description of the LAT, derived from Monte Carlo studies and accelerator beam tests, abstracts the instrument, the event data, reconstruction, and particle background rejection into what is needed for modelling the sky. The likelihood function is the probability of the data given the model. The response functions relate a model defined on the sky (sources of given spectra, positions, etc.) to the data space of measured energies, directions, etc., taking into account the pointing and live time history of the LAT for the period of interest.

For analysis of LAT data we may encounter practical limitations to the evaluation of the likelihood function. In maximizing the likelihood, changes in $\ln L$ of $\sim$1 are significant. Owing to the breadth of the PSFs at low energies, the region of the sky in a typical analysis will be of $\sim 15^\circ$ across and may contain hundreds of thousands of gamma rays (in an analysis of a one-year time frame). Numerical accuracy will have to be carefully maintained in the evaluation of the likelihood function.

The source models may also contain dozens of parameters (source positions, spectral indicies, scaling factors for diffuse emission). Only a fraction of these may be adjustable in any given analysis—e.g., coordinates of known sources may be held fixed—but even so maximization of the likelihood function will be a multidimensional optimization. In principle, this is manageable, but optimizations will be most reliable with good initial guesses for parameters. An implementation of the Expectation Maximization (EM) algorithm [15] is being explored for possible use in speeding up likelihood optimizations of models for LAT data. In this approach, gamma rays are provisionally assigned to specific sources in the model, based on the current values of the parameters of the model. Next, values of the parameters are optimized source by source, requiring likelihood evaluations for only a fraction of the gamma rays at a time. Source assignments are then updated, and the whole process is iterated. For models with a large number of sources, this approach potentially offers a tremendous advantage in computation time.

4.1. Nonparametric source detection

The fundamental limitation of likelihood analysis is that it does not answer a question that you are not asking. Also, as mentioned above, the method is computationally intensive and subject to limitation of numerical accuracy. A practical (fast and accurate) nonparametric method for detecting sources would have a great deal of appeal, of course. Even if a method provides just a useful starting point for detailed likelihood analysis to derive parameter estimations, it could be very useful. Several methods are under consideration for analysis of LAT data, including wavelet transformations (continuous and discrete), independent component analysis, general multiresolution image deconvolution, and a multidimensional extension of the Bayesian blocks algorithm (see Sec. 5.3). To the extent that they depend on knowledge of the instrument response functions for filtering, these methods may have difficulties due to the scanning observing mode of the LAT, which effectively mixes response functions for each source.

A further complication to the analysis of celestial gamma-ray sources in the LAT data is the brightness of the Earth's limb in gamma rays. These 'albedo' gamma rays produced in cosmic-ray interactions in the upper atmosphere have been characterized with data from SAS-2 [16]. The emission is quite intense relative to the gamma-ray sky, although fairly soft. The intensity exhibits a strong east-west variation, and also depends on solar activity. Even for the routine scanning sky coverage of the LAT, the horizon is never far from the field of view. The current plan is to exclude from the high-level analysis regions of the sky at large zenith angles ($>\sim110^\circ$). The cuts necessarily will be made based on measured zenith angle, and owing to the relatively strong dependence of the PSF on angle, must be more conservative at lower energies. The cuts on zenith angle complicate both the data selection and the calculation of exposure.

4.2. Characterization of sources

By whatever means a gamma-ray source is detected, characterization of the source means determining the confidence region for its location on the sky, and measuring its spectrum and variability. If a searched-for source is not detected, a meaningful upper limit for its flux should also be determined. For EGRET, these were evaluated by applying the likelihood ratio test and appealing to Wilks' theorem for interpretation of the results [14]. The likelihood ratio test was used to compare source models, e.g., one with a given point source with its maximum likelihood position and one with that source shifted somewhat in position, and Wilks' theorem was used to relate the likelihood ratios to significance levels. The interpretations of significances in the EGRET data was backed up by Monte Carlo simulations.

Recently, Protassov et al. [17] have pointed out that often in astronomy the likelihood ratio test is misapplied to circumstances where one of the parameters (e.g., source flux) is on the border of the range on which it is defined (like 0 in the case of source flux). For questions like this, Protassov et al. propose evaluating Bayesian posterior predictive P-values. The procedure to be implemented for determination of confidence ranges and upper limits in routine analysis of LAT data is still being evaluated.

4.3. Identification of gamma-ray sources

By standards of astronomy at other wavelengths, the positions of gamma-ray sources are measured very poorly. The majority of the EGRET sources are unidentified, $\sim$170 out of 271 in the Third EGRET Catalog [4], largely for this reason (see Sec. 2 and Fig. 5). For EGRET, 95% confidence contours for source locations were typically 1-2° across. The number of potential counterparts is so large that no compelling case for any particular counterpart can be made on the basis of positional coincidence.

For the LAT we plan to adopt an objective procedure for identifying potential counterparts, taking advantage of all of the information that we can derive, such as variability (especially correlated variability). For the LAT, source location regions will be much

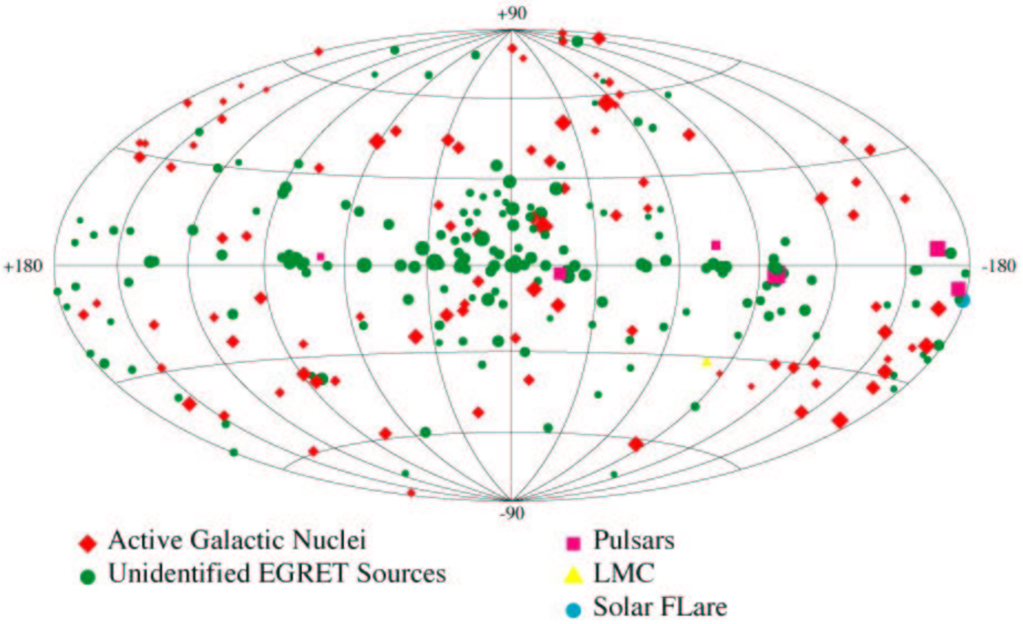

Figure 5: Sky locations of the sources in the Third EGRET Catalog [4]. Larger symbols indicate greater fluxes, scaled logarithmically. A concentration toward the Galactic plane is evident. The majority of the sources are unidentified.

smaller, on the order of several arcminutes for typical sources. This is still relatively large for counterpart searches, but will certainly make the problem easier. Mattox et al. [18, 19] made a Bayesian analysis of potential blazar counterparts to unidentified EGRET sources that used positional correlations of EGRET sources with radio continuum sources, as well as the flux and spectral index distributions of radio sources already known to be blazars. Sowards-Emmerd et al. [20] introduced a 'figure of merit' for assessing source counterparts that also includes X-ray spectral information. The figure of merit essentially includes weighting factors based on the X-ray and radio characteristics of known blazars. Establishing a new class of sources is more difficult, as statistical results for other members of the population are not available.

Variability is a common characteristic of high-energy gamma-ray sources. Blazars undergo episodic flares, during which fluxes can increase by factors of several on time scales of hours or less. Gamma-ray pulsars are periodic sources, typically with periods of hundreds of milliseconds; integrated over many periods their gamma-ray fluxes are quite steady. Indeed the brighter EGRET pulsars were used as calibration sources in flight. So, for unidentified sources, measures of variability can be used to distinguish between blazars and pulsars.

For candidate pulsars for which ephemeris data are available, generally from monitoring observations in the radio, epoch folding the gamma rays is a well-established way to search for gamma-ray pulsations. Well-defined statistical techniques have been applied to such pulsation searches [21]. The sensitivities of the tests are limited by the lack of a 'template' for pulsations. Some tests are most powerful for detecting sinusoidal variations, for example.

For suspected pulsars of unknown timing parameters, period searching is in principle possible, but hampered by many complications; see [22]. As is apparent from Fig. 6, the gamma-ray pulsations have no standard template. Before timing searches, the arrival times of the gamma rays need to be corrected for the arrival time variations due to the changes in the position of the spacecraft. If the direction of the source on the sky is not known very accurately (and it will likely not be), then uncertainties in the arrival time correction accumulate quickly. Phase drifts owing to the unknown spin-down rate (as large as 10^{-13} Hz s^{-1} for a young pulsar) can also become significant over the days or weeks required to accumulate enough gamma rays from a given source. So a period search is effectively multidimensional, including the coordinates of the prospective pulsar and its spin-down rate.

4.4. Source Identification

Positional coincidence alone is in general not adequate to establish the identification of a gamma-ray source with a counterpart detected at other wavelengths. The accuracy of position determinations with the LAT will typically be at the few arcminute level (depending on source spectrum and diffuse intensity). This is inadequate, owing to the high density of potential counterparts (e.g., the NRAO VLA Sky Survey

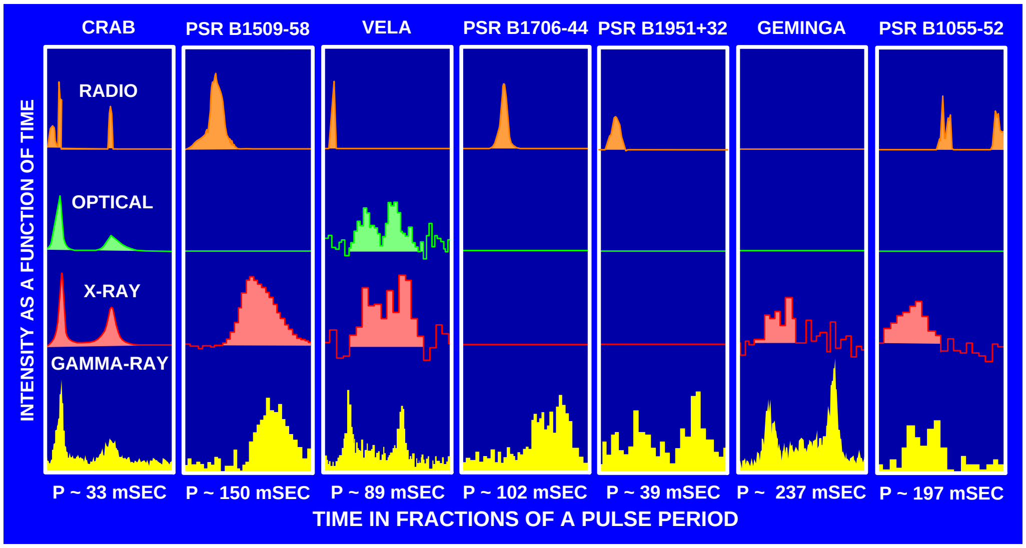

Figure 6: Composite of light curves at different wavelengths for many of the pulsars detected by EGRET. (Source: D. J. Thompson)

has >50 sources deg^{-2} [23]), without additional information that supports the identification.

Useful information can be applied based on the characteristics of either the γ-ray source or of the potential counterparts. For example, blazars have been established to be associated with flat-spectrum radio sources [14, 18].

Correlated variability is a powerful technique for identification, the prototypical example being γ-ray pulsars. For suspected counterparts with known ephemerides, statistical tests have been developed to evaluate whether the γ-ray source is pulsing with the same period. Gregory & Loredo [24] presented a Bayesian approach for periodicity searching that relies only on the assumption that the pulse profile can be assumed to be stepwise continuous; in their analysis, more complicated profiles, i.e., those with greater numbers of steps are naturally discouraged in favor of simpler profiles.

With any method of searching for periodicity, pulsations must be detected against the background of non-pulsed emission, e.g., from an associated nebula and diffuse interstellar emission, and the 'signal-to-noise' may be optimized by making PSF-dependent cuts the γ-rays included in the searches, taking advantage of the narrowing of the PSF at high energies to reduce the fraction of γ-rays of diffuse origin. An alternative approach that has been proposed instead of cutting events is to weight them according to the widths of the corresponding PSFs.

Gamma-ray bursts (GRBs) are well-known as brief, intense, and intensely variable gamma-ray sources, in recent years firmly established to be at cosmological distances and associated with galaxies with extensive massive star formation. Typical GRBs are brightest in the ~few MeV range, and from EGRET relatively little is known about their higher-energy behavior. This is primarily due to the large dead time per trigger (~0.2 s) of the EGRET spark chamber tracker; the dead time per event for the LAT likely will be significantly less than 0.1 ms.

That there are two populations of GRBs, short duration with hard spectra and long duration with relatively soft spectra, has long been established [25]. Typical durations are ~0.3 s for the short bursts vs. ~30 s for the long bursts, and the short bursts have spectral indexes harder by an increment of ~0.5. To date, only the long-soft population have been able to have counterpart identifications via rapid follow up observations, because the localization of the short-hard bursts with (typically) hard X-ray detectors is very poor. One goal for science with the LAT is to obtain excellent positions for the hard-spectrum bursts and GRB 'trigger' algorithms are being explored. By the nature of the LAT (see Sec. 2.1), some on board processing is required to make provisional reconstruction and classification of events on board. The trigger algorithms look for clusters of events in direction and time, using a moving time window. The on board algorithm will be tuned carefully via Monte Carlo simulations, and the LAT is designed to send GRB notifications to the ground using the demand access Tracking and Data Relay Satellite (TDRS) system, which promises latencies measured only in seconds.

The time profiles of GRBs typically consist of a train of pulses of different profiles (the details of which also depend on energy). Objective decomposition of GRB profiles into intervals with constant event rates (i.e., within which the variations of the event rate are not statistically significant), with proper attention to the background noise level, can be achieved using the Bayesian Block algorithm developed by Scargle [26]. The algorithm cannot sort out overlapping pulses, an unsolved problem in the general case, but still can provide useful characterizations of time profiles of GRBs. It does not require the events to be binned in time, so no minimum time scale for detecting variability is imposed by the method.

4.5. Population studies

Several related approaches have been used to demonstrate that the low-latitude EGRET gamma-ray sources are correlated with tracers of massive star formation, without needing to claim identifications for any particular unidentified source. Kaaret & Cottam [27] fitted Gaussian profiles to the latitude and longitude distributions of low-latitude unidentified EGRET sources and generated random sets of point sources consistent with these distributions. For each set, the number of sources lying within 1° of an OB association was counted. The probability of a chance association of as many as 16 sources out of 25, as observed with the actual EGRET sources, was evaluated from the distribution as 6.1 $\times 10^{-5}$.

Romero et al. [28] applied a somewhat different technique to evaluate the significance of positional correlations between the distributions of low-latitude EGRET gamma-ray sources and Wolf-Rayet stars, OB associations, and supernova remnants. The observed distributions of positional offsets was compared with the offsets obtained for the sources scrambled in longitude and latitude in such a way that their latitude distribution was exactly maintained. The conclusion that correlations were statistically significant was especially strong for SNR. Of course, Wolf-Rayet stars, OB associations and supernova remnants necessarily have fairly similar distributions, and may in fact interact to produce gamma-ray sources [29]

Grenier [30] investigated the distribution of unidentified EGRET sources by evaluating the log N-log S (flux distribution) for the sources and comparing the distribution of expected detections on the sky for various assumed intrinsic spatial distributions of the sources. This likelihood approach naturally compensates for sensitivity variations owing to different depths of exposure and intensities of diffuse emission across the sky. The results indicated a significant correlation with tracers of dense interstellar gas and star formation. Correlation with the population of radio pulsars was notably weaker, although most radio pulsars have lifetimes as gamma-ray sources much shorter than as radio sources and pulsars are known to have large proper motions.

A stacked source analysis can be used to study whether a population of putative gamma-ray sources can be detected collectively, even if individual sources are not bright enough for detection. Stacking the data for a population of sources means coadding counts and exposure centered on each source position. (Generally, unrelated nearby point sources must be excised.) This technique was recently applied by Reimer et al. [31] to investigate whether nearby, X-ray bright clusters of galaxies are also gamma-ray sources; see Fig. 7. The coadded exposure for 58 clusters resulted in an upper limit of ~6 $\times 10^{-9}$ cm^{-2} s^{-1} for the average flux above 100 MeV, approximately 8 times more sensitive than the upper limit for any of the galaxy clusters individually.

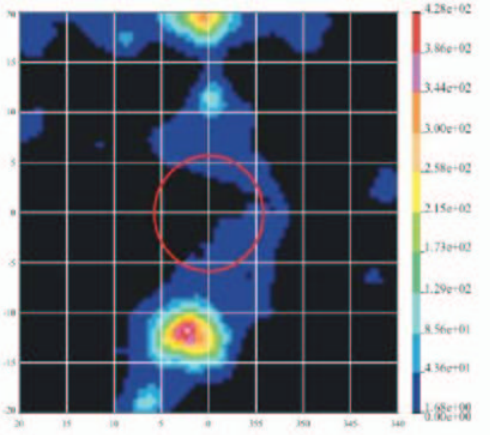

Figure 7: Source location (likelihood test statistic) map for the stacked EGRET counts and exposure for 58 X-ray bright galaxy clusters. No significant emission is seen at the composite source location at the center of the field [31].

5. CONCLUSIONS

The LAT instrument on GLAST will have revolutionary sensitivity and should revolutionize gamma-ray astronomy. Fulfilling the promise of the instrument will require careful statistical treatment of the data at all levels, from event reconstruction and classification, to source detection, source identification, and population studies. This relates to the detection method for gamma rays, their low intrinsic fluxes, and the scanning observing mode that will be routine for the LAT. Some classes of sources, in particular blazars and gamma-ray bursts, will require triggers for near-real time alerts. Searches for periodic emission from pulsars will also be needed.

References

[1] `http://glast.gsfc.nasa.gov`
[2] `http://glast.stanford.edu/`

[3] C. von Montigny, et al., "High-Energy Gamma-Ray Emission from Active Galaxies: EGRET Observations and Their Implications", ApJ, 440, 525, 1995

[4] R. C. Hartman, et al., "The Third EGRET Catalog of High-Energy Gamma-Ray Sources", ApJS, 123, 79, 1999

[5] P. Sreekumar, et al., "GeV emission from the nearby radio galaxy Centaurus A", APh, 11, 221, 1999

[6] G. Kanbach, et al., "Detection of a long-duration solar gamma-ray flare on June 11, 1991 with EGRET on COMPTON-GRO", A&AS, 97, 349, 1993

[7] W. L. Kraushaar, et al., "High-Energy Cosmic Gamma-Ray Observations from the OSO-3 Satellite", ApJ, 177, 341, 1972

[8] http://agile.mi.iasf.cnr.it/

[9] http://ams.pg.infn.it/ams-italy/ams.htm

[10] S. D. Hunter, et al., "EGRET Observations of the Diffuse Gamma-Ray Emission from the Galactic Plane", ApJ, 481, 205, 1997

[11] B. B. Jones, "A Search for Gamma-Ray Bursts and Pulsars, and the Application of Kalman Filters to Gamma-Ray Reconstruction", Ph. D. thesis, Stanford Univ. 1998, http://arxiv.org/abs/astro-ph/0202088

[12] R. D. Morris and J. Cohen-Tanugi, "An Analysis Methodology for the Gamma-ray Large Area Space Telescope", Proc. MaxEnt 23, "Bayesian and Maximum-Entropy Methods", Jackson Hole, WY, in press, 2004

[13] A. M. T. Pollock, et al., "Search for gamma-radiation from extragalactic objects using a likelihood method", A&A, 94, 116, 1981

[14] J. R. Mattox, et al., "The Likelihood Analysis of EGRET Data", ApJ, 461, 396, 1996

[15] A. P. Dempster, et al., "Maximum likleihood from imcomplete data via the EM algorithm", J. Royal Stat. Soc., Ser B., 39(1), 1, 1977

[16] D. J. Thompson, G. A. Simpson, and M. E. Ozel, "SAS 2 observations of the earth albedo gamma radiation above 35 MeV", JGR,, 86, 1265, 1981

[17] R. Protassov, et al., "Statistics, Handle with Care: Detecting Multiple Model Components with the Likelihood Ratio Test", ApJ, 571, 545, 2002

[18] J. R. Mattox, et al., "The Identification of EGRET Sources with Flat-Spectrum Radio Sources", ApJ, 481, 95, 1997

[19] J. R. Mattox, et al., "A Quantitative Evaluation of Potential Radio Identifications for 3EG EGRET Sources", ApJS, 135, 155, 2001

[20] D. Sowards-Emmerd, R. W. Romani, and P. F. Michelson, "The Gamma-Ray Blazar Content of the Northern Sky", ApJ, 590, 109, 2003

[21] O. C. de Jager, "On periodicity tests and flux limit calculations for gamma-ray pulsars", ApJ, 436, 239, 1994

[22] A. M. Chandler, et al., "A Search for Radio-Quiet Gamma-Ray Pulsars", ApJ, 556, 59, 2001

[23] J. J. Condon, et al., "The NRAO VLA Sky Survey", AJ, 115, 1693, 1998

[24] P. C. Gregory and T. J. Loredo, "A new method for the detection of a periodic signal of unknown shape and period", ApJ, 398, 146, 1992

[25] C. Kouveliotou, et al., "Identification of Two Classes of Gamma-Ray Bursts", ApJ, 413, L101, 1993

[26] J. D. Scargle, "Studies in Astronomical Time Series Analysis. V. Bayesian Blocks, a New Method to Analyze Structure in Photon Counting Data", ApJ, 504, 405, 1998

[27] P. Kaaret and J. Cottam, "Do the Unidentified EGRET Sources Lie in Star-forming Regions?", ApJ, 462, L35, 1996

[28] G. E. Romero, P. Benaglia, and D. F. Torres, "Unidentified 3EG gamma-ray sources at low galactic latitudes", A&A, 348, 868, 1999

[29] T. Montmerle, "On gamma-ray sources, supernova remnants, OB associations, and the origin of cosmic rays", ApJ, 231, 95, 1979

[30] I. A. Grenier, "Spatial distribution of the unidentified EGRET sources off the galactic plane", AdSpR, 15, 573, 1995

[31] O. Reimer, "EGRET Upper Limits on the High-Energy Gamma-Ray Emission of Galaxy Clusters", ApJ, 588, 155, 2003