Data Preparation

Photon and spacecraft data are all that a user needs for the analysis. For the definition of **LAT data products see the information in the Cicerone**.

The LAT data can be extracted from the **Fermi Science Support Center** web site, as described in section Extract LAT data. Preparing these data for analysis depends on the type of analysis you wish to perform (e.g. point source, extended source, GRB spectral analysis, timing analysis, etc). The different cuts to the data are described in detail in the Cicerone.

Data preparation consists of two steps.

- (gtselect): Used to make cuts based on columns in the event data file such as time, energy, position, zenith angle, instrument coordinates and event class.
- (gtmktime): In addition to cutting the selected events, gtmktime makes cuts based on the spacecraft file and updates the GTI extension.

Here we give an example of how to prepare the data for the analysis of a point source. For your particular source analysis you have to prepare your data performing similar steps, but with the cuts suggested in Cicerone for your case.

1. Event selection with gtselect

In this section, we look at making basic data cuts using gtselect. By default, gtselect prompts for cuts on:

- **Time**
- **Energy**
- **Position** (RA,Dec,radius)
- **Maximum Zenith Angle**

However, by using hidden parameters defined on the command line (or using the 'Show Advanced Parameters' check box in GUI mode), you can also make cuts on:

- **Event class ID**
- **Minimum pulse phase**
- **Maximum pulse phase**

- **Conversion type (-1=both, 0=Front, 1=Back)**

For this example we use data that was extracted using the procedure described in the Extract LAT Data tutorial. The original selection used the following information:

- Search Center (RA, DEC) =(193.98, -5.82)
- Radius = 20 degrees
- Start Time (MET) = 239557417 seconds (2008-08-04 T15:43:37)
- Stop Time (MET) = 255398400 seconds (2009-02-04 T00:00:00)
- Minimum Energy = 100 MeV
- Maximum Energy = 100000 MeV

The LAT operated in survey mode for that period of time. We provide the user with the original photon and spacecraft data files extracted in the Extract LAT data tutorial:

1. L1404301505139948 9F7F38_PH00.fits
2. L140430150513994 89F7F38_PH01.fits
3. L140430150513994 89F7F38_SC00.fits

If more than one file was generated by the FSSC's web site data server, we will need to provide an input file list, in order to use all the event data files in the same analysis. This text file can be generated by typing:

```
ls *_PH* > events.txt
```

This input file can be used in place of a single input events (or FT1) file by placing an @ symbol before the text filename. The output from gtselect will be a single file containing all events from the combined file list that satisfy the other specified cuts.

For point source analysis, it is recommended that you include events with a high probability of being photons. This cut is performed by selecting source class events with the gtselect tool by including the hidden parameter **evclass** on the command line. For LAT Pass 7 data, source events are specified as event class 2 (the default value). This value may change as LAT data processing develops.

Now run gtselect to select the data you wish to analyze. For this example, we consider the source class photons within a 20 degree acceptance cone of the blazar 3C 279. We apply the gtselect tool to the data file as follows:

```
prompt> gtselect evclass=2
Input FT1 file[] @events.txt
Output FT1 file[] 3C279_region_filtered.fits
RA for new search center (degrees) (0:360) [0] 193.98
Dec for new search center (degrees) (-90:90) [0] -5.82
radius of new search region (degrees) (0:180) [180] 20
start time (MET in s) (0:) [0] 239557417
end time (MET in s) (0:) [0] 255398400
lower energy limit (MeV) (0:) [30] 100
upper energy limit (MeV) (0:) [300000] 100000
maximum zenith angle value (degrees) (0:180) [180] 100
Done.
prompt>
```

The filtered data is provided here.

If you don't want to make a selection on a given parameter, just enter a zero (0) as the value.

In this step we also selected the maximum zenith angle value (100 degrees) as suggested in the Cicerone. Photons coming from the Earth limb are a strong source of background. You can minimize this effect with a zenith-angle cut. The value of 100 degrees is the one most commonly used, and provides a 13 degree buffer between your region of interest (ROI) and the Earth's limb. In the next step, gtmktime will remove any time period that our ROI overlaps this buffer region. While increasing the buffer (reducing zmax) may decrease the background rate from albedo gammas, it will also reduce the amount of time your ROI is completely free of the buffer zone and thus reduce the livetime on the source of interest.

Notes:

- The RA and Dec of the search center should match that used in the dataserver selection.
- The radius of the search region selected here must lie entirely within the region defined in the dataserver selection. They can be the same values, with no negative effects.
- The time span selected here must lie within the time span defined in the dataserver selection. They can be the same values with no negative effects.
- The energy range selected here must lie within the time span defined in the dataserver selection. They can be the same values with no negative effects.

gtselect writes descriptions of the data selections to a series of 'Data Sub-Space' (DSS) keywords in the EVENTS extension header. These keywords are used by the exposure-related tools and by gtlike for calculating various quantities, such as the predicted number of detected events given by the source model. These keywords *must* be same for all of the filtered event files considered in a given analysis. gtlike will check to ensure that all of the DSS keywords are the same in all of the event data files. For a discussion of the DSS keywords see the Data Sub-Space Keywords page.

There are multiple ways to view information about your data file. For example:

- You may obtain the value of start and end time of your file by using the **fkeypar** tool. This tool is part of the FTOOLS software package, and is used to read the value of a FITS header keyword and write it to an output parameter file. For more information on the **fkeypar** tool, type: `fhelp fkeypar`
- The gtvcut tool can be used to view the DSS keywords in a given extension, where the EVENTS extension is assumed by default. This is an excellent way to to find out what selections have been made already on your data file (by either the dataserver, or previous runs of gtselect). NOTE: If you wish to viewl the (very long) list of good time intervals (GTIs), you can use the hidden parameter 'suppress_gtis=yes' on the command line. The full list of GTIs is suppressed by default.

```
prompt> gtvcut
Input FITS file[] 3C279_region_filtered.fits
Extension name[EVENTS]
DSTYP1: BIT_MASK(EVENT_CLASS,2,P7REP)
DSUNI1: DIMENSIONLESS
DSVAL1: 1:1

DSTYP2: POS(RA,DEC)
DSUNI2: deg
DSVAL2: CIRCLE(193.98,-5.82,20)

DSTYP3: TIME
DSUNI3: s
DSVAL3: TABLE
DSREF3: :GTI

GTIs:
239557417.494 239558069.093
239559567.98 239563954.084
239565647.986 239569844.085
...
255384024.931 255388191.087
255390039.929 255394247.086
```

```
255396028.929 255398400

DSTYP4: ENERGY
DSUNI4: MeV
DSVAL4: 100:100000

DSTYP5: ZENITH_ANGLE
DSUNI5: deg
DSVAL5: 0:100

prompt>
```

2. Time Selection with gtmktime

You may have noticed that all of these files have a GTI extension in them. Before we look at making selections with the gtmktime tool, we should probably clarify what at **Good Time Interval (GTI)** is:

- Simply stated, a **GTI** is a time range when the data can be considered valid. The GTI extension contains a list of these GTI's for the file. Thus the sum of the entries in the GTI extension of a file corresponds to the time when the data in the file is "good".

How are these interpreted for Fermi?

- The initial list of GTI's are the times that the LAT was collecting data over the time range you selected. The LAT does **not** collect data while the observtory is transiting the Southern Atlantic Anomoly (SAA), or during rare events such as software updates or spacecraft maneuvers.

  Notes:

  - Your object will most likely not be in the field of view during the entire time that the LAT was taking data.
  - Additional data cuts made with gtmktime will update the GTI's based on the cuts specified in both *gtmktime* and *gtselect*.
  - The Science Tools use the GTI's when calculating exposure. If these have not been properly updated, the exposure correction made during science analysis may be incorrect.

gtmktime is used to update the GTI extension and make cuts based on spacecraft parameters contained in the spacecraft (pointing and livetime history) file. It reads the spacecraft file and, based on the filter expression and specified cuts, creates a set of GTIs. These are then combined (logical and) with the existing

GTIs in the Event data file, and all events outside this new set of GTIs are removed from the file. New GTIs are then written to the GTI extension of the new file.

Cuts can be made on any field in the spacecraft file by adding terms to the filter expression using C-style relational syntax:

> !->not, &&-> and, -> or, ==, !=, >, <, >=, <=
>
> ABS(), COS(), SIN(), etc., also work

NOTE: Every time you specify an additional cut on time, ROI, zenith angle, or event class using gtselect, you must run gtmktime to reevaluate the GTI selection.

Several of the cuts made above with *gtselect* will directly affect the exposure. gtmktime will select the correct GTIs to handle these cuts, including the complex cut on zenith angle. The easy way to implement this cut is to exclude time intervals where the buffer zone defined by the zenith cut intersects the ROI from the list of GTIs. In order to do that, run gtmktime and answer "yes" at the prompt:

```
Apply ROI-based zenith angle cut [ ] yes
```

Applying this cut is especially important if your ROI is small (< 20 degrees) bringing your source of interest close to the Earth's limb.

- **Note**: If you are studing a very broad region (or the whole sky) you would lose most (all) of your data when you implement the zenith angle cut. In this case you can allow all time intervals where the cut intersects the ROI, but the intersection lies outside the FOV. To do this, run gtmktime specifying a filter expression defining your analysis region, and answer "no" to the question regarding the ROI-based zenith angle cut:

  ```
  Filter expression [ ] ( ANGSEP(RA_ZENITH,DEC_ZENITH,RA_SCZ,DEC_SCZ) < limb_angle_minus_
  Apply ROI-based zenith angle cut [ ] no
  ```

  Here, RA_of_center_ROI, DEC_of_center_ROI and radius_ROI correspond to the ROI selection made with gtselect, zenith_cut is defined as 100 degrees (as above), and limb_angle_minus_FOV is (zenith angle of horizon - FOV radius) where the zenith angle of the horizon is 113 degrees.

gtmktime also provides the ability to exclude periods when some event has negatively affected the quality of the LAT data. To do this, we select good time intervals (GTIs) by using a logical filter for any of the quantities in the spacecraft file. Some possible quantities for filtering data are:

- DATA_QUAL - quality flag set by the LAT instrument team (1 = ok, 2 = waiting review, 3 = good with bad parts, 0 = bad)
- LAT_CONFIG - instrument configuration (0 = not recommended for analysis, 1 = science configuration)
- ROCK_ANGLE - can be used to eliminate pointed observations from the dataset.
  NOTE: A history of the rocking profiles that have been used by the LAT can be found in the observations section.

The current gtmktime filter expression recommended by the LAT team is:

(DATA_QUAL>0)&&(LAT_CONFIG==1).

NOTE: The "DATA_QUAL" parameter can be set to different values, based on the type of object and analysis the user is interested into (see this page of the Cicerone for the most updated detailed description of the parameter's values). Typically, setting the parameter to 1 is the best option. For GRB analysis, on the contrary, the parameter should be set to ">0".

Here it is an example of running gtmktime on the 3C 279 filtered events file. It is useful to rename the spacecraft file to something easier. Here, we have renamed L14043015051399489F7F38_SC00.fits to spacecraft.fits.

```
prompt> gtmktime
Spacecraft data file[] spacecraft.fits
Filter expression[] (DATA_QUAL>0)&&(LAT_CONFIG==1)
Apply ROI-based zenith angle cut[yes]
Event data file[] 3C279_region_filtered.fits
Output event file name[] 3C279_region_filtered_gti.fits
```

The data with all the cuts described above is provided in this link.

After the data preparation it is advisable to take a look at your data before beginning the detailed analysis. The Explore LATdata tutorial has suggestions on method of getting a quick preview of your data.

Last updated by: D. Davis 05/01/2014