Spectral analysis of point-like sources with EPIC

Matteo Guainazzi (XMM-Newton SOC, mguainaz@xmm.vilspa.esa.es)

November 27 2001

Abstract

In this paper, some basic principles of point-like sources spectral analysis with the EPIC cameras are summarized. A brief outline of the most commonly used functionalities of *evselect* and *xmmselect*, is presented. Items such as the background subtraction, pile-up effect mitigation, and response matrices are also discussed.

1 The XMM-Newton extractor

One defines *extractor* a software machine, able to appropriately bin columns of an event list to generate scientific products. In the context of this paper, we will focus our attention to the generation of spectra. The SAS spectral extractor is run through the task *evselect*, and its GUI interface *xmmselect*. There is a one-to-one correspondence between *evselect* line commands and *xmmselect* interactive procedures. Whenever quoting a specific extraction command is necessary, we will refer to the former.

2 The main *xmmselect* window

Running `xmmselect` from the command-line or the SAS GUI-interface yields the activation of the main *xmmselect* window. This window is actually constituted by 4 separate widgets/panels. ¿From top to bottom:

1. a widget, where data screening selection expressions can be specified/edited
2. a list of the columns in the event list file `EVENTS` extension. Each record of this panel corresponds to one column, and contains 4 active elements:
   - square checkboxes, which may be used to select/deselect the pair of `EVENTS` columns, on which an image shall be accumulated
   - round checkboxes, which may be used to select/deselect the `EVENTS` column, on which an image, histogram, OGIP spectrum or light curve[1] shall be accumulated
   - one rectangular box named after each of the `EVENTS` column
   - a pair of editable widgets, reporting the minimum (`min`) and maximum (`max`) value of the corresponding column

[1] An "OGIP spectrum" ("light curve") is an histogram on the energy (time), whose format is compliant with the specification of the NASA "Office of Guest Investigator Program" (`http://heasarc.gsfc.nasa.gov/docs/lhea/668.html`). OGIP-compliant files can be analyzed with LHEASOFT packages, such as, *e.g.*, XSPEC (XANADU)

Clicking on a rectangular box causes a selection expression of the form: `(COLUMN in [min:max])` to be logically "ANDed" to the selection expression

3. the `1D region` and `2D region` boxes. Clicking one of them causes the regions defined in the current `grace` or `ds9` session (cf. Sect. 3.1 and 3.2 below) to be interpreted and logically "ANDed" to the current selection expression, respectively.

4. boxes to choose the scientific product to be accumulated, and to launch the accumulation. Together with the scientific products already listed above, one may also choose to extract *filtered event lists*.

3 Interactive definition of 1- and 2-D selection regions

One of the most common tasks in the generation of customized spectra is the definition of regions in a 1-D (*e.g.*: time) or 2-D (*e.g.*: sky plane) parameter space, from which good events are accumulated. In this Section we will briefly outline how this can be interactively done, exploiting the interaction between *xmmselect* and the software packages `grace` (1-D data plotting) and `ds9` (image viewer/analyzer).

3.1 Definition of a time selection expression with *xmmselect*

If the accumulation of a light curve is started, *xmmselect* launches the main *evselect* window. The user can define in the 5 different panels of this window all the available *evselect* accumulation parameters. Clicking "Run", *evselect* starts accumulating the light curve. Once the accumulation is completed, a `grace` window is launched, and the accumulated light curve displayed. `grace` is a flexible, multi-purpose plotting package, which includes several graphic and data analysis functionalities. A detailed description of these functionalities is beyond the scope of this paper. Its "Define region" function allow to select/exclude interactively regions of the plotted light curve, according to a wide range of different options (within/outside a vertical/horizontal range, within a polygon etc.). Once a user has defined the time intervals, on which s/he wants to accumulate her/his spectra, a click on the `1-D region` button of the main *xmmselect* window will force *xmmselect* to interpret the selected intervals as a *selectlib* expression and parse them - as logical AND - to the selection expression *xmmselect* widget. For example: if a user has selected on the `grace` displays light curve an horizontal range comprised between $T = 9.5 \times 10^7$ s and $T = 9.6 \times 10^7$ s, the corresponding *selectlib* selection expression will appear as follows:

```
(TIME in [9.5E+07:9.6E+07])
```

Any spectra later accumulated with the modified selection expression will contain only events included in the selected time intervals

3.2 Definition of an extraction region with *xmmselect*

The launch of an image accumulation from the main *evselect* window starts a `ds9` session. `ds9` is an advanced software package to display and analyze X-ray images. A detailed description of `ds9` and its functionalities is beyond the scope of this paper. One of the `ds9` menu allows the user to select/exclude regions on the displayed image. Several different shapes (circle, box, ellipses, polygons etc.) and effects (rotation, digital or cursor resize) are available. Multiple or complex regions can be defined as well. Once the user is happy, with a given selection region choice, clicking on the `2-D region` button in the main *xmmselect* window forces *xmmselect* to interpret the selected region(s) as a *selectlib* expression and parse them - as logical AND - to the selection expression *xmmselect* widget. For instance: if a circular region around the position of null sky pixels and radius equal to 100 pixels has been defined in `ds9`, the corresponding *selectlib* expression will appear as follows:

```
((X,Y) in circle(0,0,100))
```

Any spectra later accumulated with the modified selection expression will contain only events belonging to the selected sky region.

4 Camera-dependent screening criteria

On-flight calibration experience has allowed to identify two families of camera-dependent screening criteria, which shall be normally applied to the accumulation of a spectrum in addition to any time, spatial or other scientifically-driven screening criteria:

1. `#XMMEA_EM` (MOS) and `#XMMEA_EP` (pn). This is a bitwise *selectlib* selection function defined in the header of the EPIC event lists. It allows a user to automatically discard "bad events" in terms of bad rows, edge effects, spoiled frames, GATTI rejection, cosmic ray or particle events, diagonal events, events beyond thresholds etc. The users must be aware that such expressions are not applied either by the SAS or by the PPS reduction pipeline. They need therefore to be applied on *all* EPIC event lists to extract clean scientific products;
2. pattern selection. In the MOS cameras, a pattern selection in the range 0–12 ensures the maximization of the signal-to-noise ratio against not X-ray events. Although this is the recommended choice, analysis of spectra extracted from single events only (*i.e.*: `PATTERN==0`) is supported and calibrated as well. Single event spectra may be used to mitigate the pile-up effects in bright sources. For the pn cameras, analysis is supported for spectra extracted from single (*i.e.*: `PATTERN==0`), double (*i.e.*: `PATTERN` in the range 1 to 4) and "single+double" (`PATTERN<=4`) events.

5 Background subtraction issues

There is no unique recipe to extract the "optimal" background spectrum. The users are invited to experience different background accumulation techniques and compare the relevant results to gain confidence on the robustness of their scientific results. For most practical cases, however, background spectra can be conveniently chosen from source-free regions *of the same chip* where the source of interest lays. Good geometry choices are: a circular region approximately at the same off-axis angle (in the MOS) or on the same chip raw coordinates row (in the pn); an annulus around the source.

Alternatively, blank sky field files are available for retrieval in the XMM-Newton calibration pages (`http://xmm.vilspa.esa.es/ccf`). Source (+ background) and background spectra need to be extracted from the same regions in detector coordinates, and using the same selection expressions. Blank sky field background spectra do not have the `EXPOSURE` keyword properly set by *evselect*/*xmmselect*. The value of this keyword needs to be therefore updated on the accumulated spectrum, according to the specifications in Lumb (2001).

XMM-Newton experiences sometimes high flaring background periods, most likely due to the focusing onto the detector focal plane of soft protons accelerated by magnetic reconnection phenomena in the upper layers of the Earth's magnetosphere. These intervals need to be excluded as well from the spectrum accumulation, following this recipe:

- extract a high-energy ($E > 10$ keV), single-event light curve, using the following selection expression:

  ```
  "(PI>10000)&&(PATTERN==0)&&#XMMEA_E[MP]"
  ```

- evaluate an appropriate threshold (`${threshold}`) to define the "flaring" intervals
- create a Good Time Interval file, including only the light curve bins whose counts fall below the threshold:

  ```
  tabgtigen table="high_energy_curve.fits" gtiset=gti.fits
            expression="COUNTS<${threshold}"
  ```

- apply the file `gti.fits` to any subsequent spectrum accumulations, adding to the selection expression the string: "... `gti(gti.fits,TIME)` ..."

6 How to deal with pile-up?

Pile-up is the accumulation of n events onto the same pixel during the time a CCD frame is read out. The accumulated events are interpreted as one single event, whose energy is equal to the sum of the energies of the accumulated events. The effect of pile-up is therefore two-fold: it produces both flux loss (one event registered instead of many) and spectral distorsion (energies are shifted bluewards).

Pile-up starts becoming a problem above a count-rate threshold, which depends on the camera and the instrumental mode (via the frametime). If the EPIC cameras are operated in Full Frame Mode, pile-up is significant ($> 5\%$) for broadband count rate > 0.7 and 8 s^{-1} for the MOS and the pn cameras, respectively. Analogous thresholds for other EPIC modes can be found in Ehle et al. (2001). Please note that these numbers need to be regarded as indicative only. One needs to assess the effect of pile-up on her/his own data to achieve confidence on the corresponding scientific results.

In the most extreme cases, pile-up is easy to identify, because it produces a "hole" at the core of the PSF. In less extreme cases, still a comparison of the source radial profile with the theoretical PSF (as after the CCF and viewable through the SAS task *calview*) may indicate loss of core photons, which is indicative of pile-up. However, a much safer indicator is the pattern distribution function. For pn data only, the SAS task *eppatplot* extracts such a distribution from an event list, and compares it with the theoretical model for Full Frame Mode. This allows the user to evaluate (still in a semi-quantitative way) the energy range, where pile-up does not constitute a problem.

To cure data affected by pile-up, two main guidelines can be recommended:

- extract spectra from single events only
- excise the pile-up affected PSF core. Experiments demonstrate that such an operation does not introduce significant ($\Delta\Gamma < 0.02$) spectral distorsion in sources up to a (not-excised) pn count rate of about 20 s^{-1}.

7 EPIC response matrices

As customary in X-ray astronomy, spectra are generated in counts. In order to convert counts into physical units, the detector transfer function needs to be provided to the spectral analysis tools (*e.g.*: XSPEC). Such an information is included in the *response matrix*.

Response matrices for X-ray astronomy are yielded by the matrix product of two factors:

- a matrix, describing for each energy E_0 the probability distribution function that an incoming photon with energy E_0 is detected at a channel corresponding to energy E. This *redistribution matrix* is normally characterized by the extension `*.rmf` (although some of the XMM-Newton response files do not unfortunately follow this naming convention)
- a vector, describing the instrument (optics+detector) *effective area* as a function of energy. Standard units are cm^{-2}. The file containing this vector is called also "ancillary file", and normally characterized by the extension `.arf`

The matrix product of an effective area vector with a redistribution matrix yields the *response matrix*. The corresponding files have normally the extension `*.rsp`.

EPIC response matrices can be accessed/generated following three different methods:

- using a standard ("pre-canned") response matrix files, available through the web page `http://xmm.vilspa.esa.es/ccf`. These matrices are provided and regularly updated by the EPIC Hardware Team, and are well suited for the analysis of on-axis, point-like sources. Different matrices are available according to different cameras (MOS1, MOS2, pn), filters (thin, medium, thick), pattern selection ("0" and "0-12" for the MOS cameras; "0" or "singles", "1-4" or "doubles" and "0-4" or "singles+doubles"), and source positions in raw chip coordinates (pn only; the gain correction is dependent on the distance from the readout node). Standard matrices are produced under the assumption of an infinite radius extraction region. The flux determined with them needs therefore to be corrected for finite-region flux loss effects.
- generating a redistribution matrix via the SAS task `rmfgen` and an effective area vector via the SAS task `arfgen` (in this order!). Such tasks need a spectral file as input. They ensure in principle full control over the effects determining the response matrix, allow the usage of source specific detector maps, and calculate the source-dependent Point Spread Function energy correction. A description of the SAS-task generated response matrices and their performances can be found in Saxton & Siddiqui (2001)
- generating an effective area vector via the SAS task `arfgen`, using a standard redistribution matrix as input to `arfgen` to consistently calculate the energy grids

Standard matrices normally follow more closely the evolution of the on-flight calibrations. However, the difference between them and the SAS-generated response matrices is now virtually negligible for most practical cases.

8 Is the quality of my spectral fitting good enough?

It is of course very difficult to provide the users with a general answer to this question. The answer depends critically on the specific scientific goals of each observer. Nonetheless, the users may find some reference information about the expected accuracy of the XMM-Newton calibrations from two main sources:

- the CCF *calibration notes*, which report the expected accuracies associated with each individual CCF constituent
- the SAS *validation reports*. A set of mostly Performance Verification data has been defined at an early stage of the mission. Whenever a new SAS release is delivered, SAS is run on these test datasets to verify:
  - whether the analysis of all the instrument modes is supported
  - the accuracy of the XMM-Newton calibrations

SAS validation reports are regularly (and hopefully frequently) updated, to reflect any improvements in the on-flight calibrations. They can be found at the following URL:

`http://xmm.vilspa.esa.es/public/xmm_sas_sv_top.html`

Acknowledgments

Comments from M.Audard, C.Erd and M.Santos-LLeo are gratefully acknowledged.

References

Ehle M., et al., 2001, XMM-Newton Users' Handbook, available at:
`http://xmm.vilspa.esa.es/user/A02/uhb/xmm_uhb.html`

Lumb D., 2001, XMM-SOC-CAL-TN-0016, available at:
`http://xmm.vilspa.esa.es/docs/documents/CAL-TN-0016-1-0.ps.gz`

Saxton R.D., Siddiqui H., 2001, XMM-SOC-PS-TN-43, available at:
`http://xmm.vilspa.esa.es/docs/documents/PS-TN-0043-2-0.ps.gz`