

# **I/O performance**

## **Input from LST-1 analysis**

**Daniel Morcuende**  
**Gammapy coding sprint | Oct 2023**

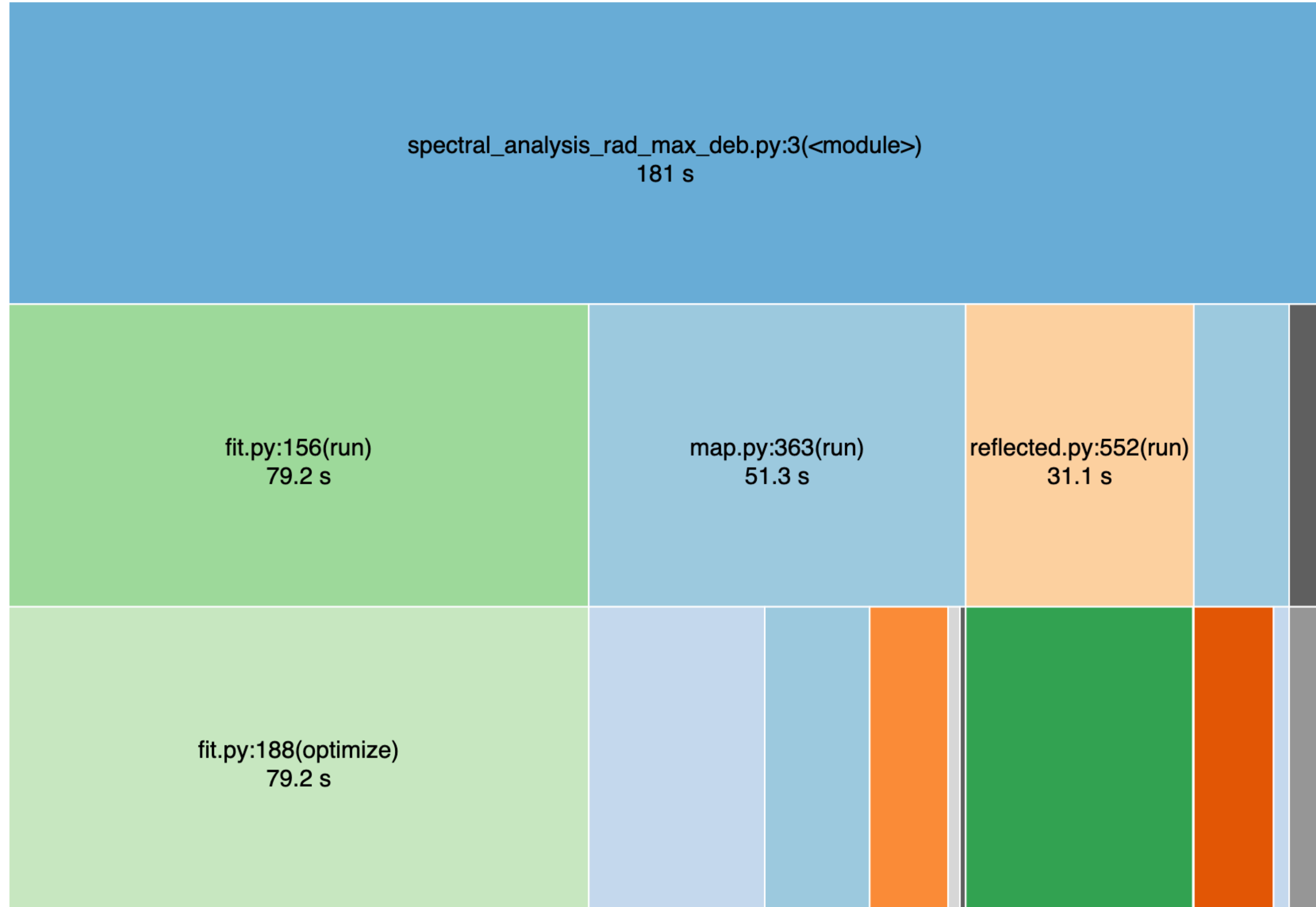
# Reading data in

- Based on the analysis performed for the Crab LST-1 paper with **Gammapy v0.20**
- Started working with zipped files > ~30 min to read in 35 hours of observations
  - A large amount of time spent unzipping DL3 input files
- We instead started to use unzipped DL3 files -> ~3 min to read in the same dataset

# Reading data in (+1d spectral analysis)

Gammapy v1.1

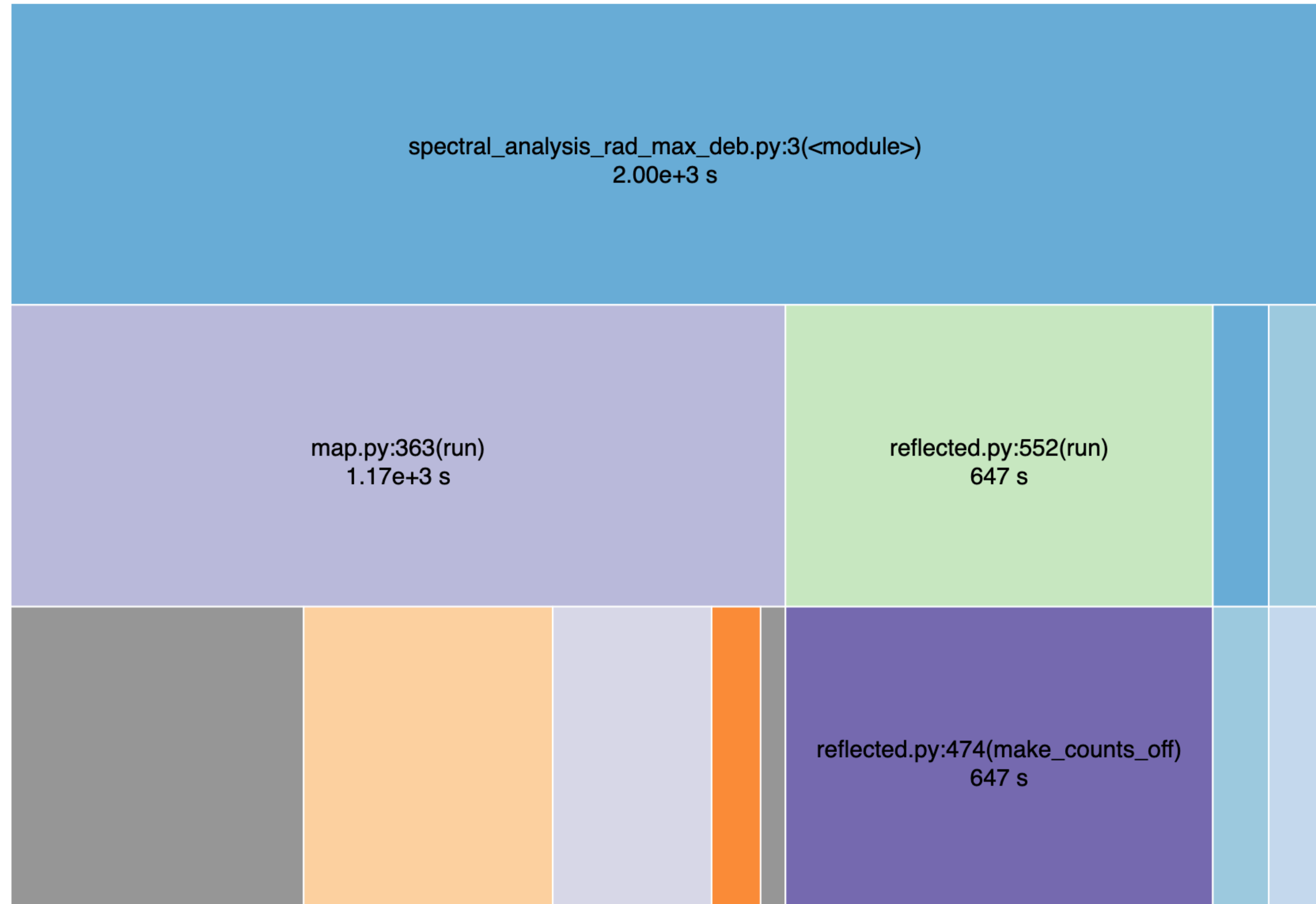
- Unzipped
- 7.9G of DL3 files
- 3 min



# Reading data in (+1d spectral analysis)

Gammapy v1.1

- Gzipped
- 7.0G of DL3 files
- x10 time!



# Reading data in (+1d spectral analysis)

gzip  
read and seek

| ncalls        | tottime  | percall   | cumtime | percall   | filename:lineno(function)                               |
|---------------|----------|-----------|---------|-----------|---|
| 1             | 0.002017 | 0.002017  | 1998    | 1998      | spectral_analysis_rad_max_deb.py:3(<module>)            |
| 2265/1        | 0.01866  | 0.01866   | 1998    | 1998      | ~:0(<built-in method builtins.exec>)                    |
| 344567        | 0.07467  | 2.167e-07 | 1860    | 0.005399  | fits.py:123(__get__)                                    |
| 4097          | 0.1336   | 3.26e-05  | 1860    | 0.454     | fits.py:74(load)  |
| 52248783      | 40.56    | 7.764e-07 | 1766    | 3.38e-05  | gzip.py:462(read)                                       |
| 45215         | 18.3     | 0.0004048 | 1626    | 0.03595   | _compression.py:120(seek)                               |
| 148030        | 0.07665  | 5.178e-07 | 1626    | 0.01098   | gzip.py:376(seek)                                       |
| 149206/148032 | 0.07963  | 5.379e-07 | 1626    | 0.01098   | ~:0(<method 'seek' of '_io.BufferedReader' objects>)    |
| 52247609      | 1523     | 2.915e-05 | 1523    | 2.915e-05 | ~:0(<method 'decompress' of 'zlib.Decompress' objects>) |
| 78347/43681   | 0.05523  | 1.264e-06 | 1274    | 0.02918   | hdulist.py:330(__getitem__)                             |
| 26117         | 0.09825  | 3.762e-06 | 1272    | 0.04869   | hdulist.py:1253(_read_next_hdu)                         |
| 22018         | 0.04796  | 2.178e-06 | 1271    | 0.05773   | base.py:333(readfrom)                                   |
|               |          |           |         |           |   |

# Lazy Fits loading

Changed cache=True in In `Observations` (suggested by Regis)

```
aeff = LazyFitsData(cache=False)
edisp = LazyFitsData(cache=False)
psf = LazyFitsData(cache=False)
bkg = LazyFitsData(cache=False)
_rad_max = LazyFitsData(cache=False)
_events = LazyFitsData(cache=False)
_gti = LazyFitsData(cache=False)
_pointing = LazyFitsData(cache=True)
```

Unzipped: 140 s

Zippped: 480 s

# Profiling

| ncalls      | tottime  | percall   | cumtime | percall   | filename:lineno(function)                               |
|-------------|----------|-----------|---------|-----------|---|
| 2736/1      | 0.02286  | 0.02286   | 479.1   | 479.1     | ~:0(<built-in method builtins.exec>)                    |
| 1           | 0.006654 | 0.006654  | 479.1   | 479.1     | spectral_analysis_rad_max_deb_zip.py:3(<module>)        |
| 117         | 0.003958 | 3.383e-05 | 361.8   | 3.093     | map.py:363(run)   |
| 344567      | 0.07433  | 2.157e-07 | 342.2   | 0.0009932 | fits.py:123(__get__)                                    |
| 702         | 0.02492  | 3.551e-05 | 342.1   | 0.4874    | fits.py:74(load)  |
| 9377811     | 7.432    | 7.925e-07 | 323     | 3.444e-05 | gzip.py:462(read)                                       |
| 7745        | 3.272    | 0.0004224 | 293.8   | 0.03793   | _compression.py:120(seek)                               |
| 25663       | 0.01613  | 6.286e-07 | 293.8   | 0.01145   | gzip.py:376(seek)                                       |
| 25901/25665 | 0.01454  | 5.665e-07 | 293.8   | 0.01145   | ~:0(<method 'seek' of '_io.BufferedReader' objects>)    |
| 9377573     | 277      | 2.954e-05 | 277     | 2.954e-05 | ~:0(<method 'decompress' of 'zlib.Decompress' objects>) |

# Longer time to load data #4680

- Problem with the H.E.S.S. dataset reported, probably also related.
  - Suggested to activate the lazy loading



# Other concerns

- Joint vs stacked analysis flux points calculation

