spidiffit Users Manual and Explanatory Supplement

A. Strong,
MPE Garching

Issue 2
Version April 17, 2002

Contents

Chapter 1

Introduction

1.1 Scope

The coded-mask imaging γ-ray spectrometer SPI on the INTEGRAL space observatory will detect point sources and diffuse extended emission with an angular accuracy of about 1° over its energy range of 40–8000 keV. The purpose of the software package described here is to represent the measurement in terms of spatial models on the sky, fitting parameters of these models and estimating their uncertainty. This tool is oriented towards large-scale surveys (eg. GCDE), which combine a large set of individual pointings of the spacecraft. It concentrates on fitting spatial as opposed to spectral models, although (through using the different energy channels of the measurement) it will be an important method to generate spectra of diffuse emission.

The principal use will be fitting to maps of physically-based tracers of γ-ray emission. Examples are fitting the 1809 keV ^{26}Al line to free-free or IR emission maps, fitting diffuse continuum γ-ray s to gas (HI, CO) maps. It will also allow fitting spatial functions (e.g. Gaussian, exponential) where there is no 'physical' model available. Fitting components is the best method to get the spectra of diffuse emission when a good spatial (but a poor spectral) model is available. The analysis of measurements with this tool is complementary to deriving results from deconvolved images (e.g. from *spiskymax*) and methods specifically designed for point sources (e.g. *spiros*).

The algorithm performs fitting of raw data (binned counts for many observations) to multi-component models using the full instrument response information. In addition to parameter estimation, a Bayesian statistical analysis is used to include information on model characteristics and other prior information, so that the uncertainties of the fitted parameters are prop-

erly assessed. The prime results of the package are fitted model parameters with their uncertainty; complementing these, the fitted models are given also in the forms of skymaps and spectra.

The package will be referred to as *spidiffit* . The package is closely related to imaging packages: response, convolution, background treatments are common or similar.

A presentation of *spidiffit* with example simulations including the Galactic Deep Exposure (GCDE) can be found at

http://www.gamma.mpe-garching.mpg.de/∼aws/integral.html

1.2 Document Structure

This document has three main parts:

- this introduction
- the software specification for development and testing
- the method/algorithm specification for the scientific part

1.3 Related Documents

1.3.1 Applicable Documents

(These documents drive *spidiffit* and their latest versions must be adhered to in the *spidiffit* package)

1. ISDC Data Specification

1.3.2 Reference Documents

(These documents contain additional and complementary information on the subject and aspects of *spidiffit*)

1. Cambridge MCMC Preprint service : http://www.statslab.cam.ac.uk/ mcmc/
2. Chen M-H, Shao Q-M, Ibrahim, J,G, Monte Carlo Methods in Bayesian Computation, Springer Series in Statistics, 2000, ISBN 0-387-98935-8
3. Gilks W.R., Richardson S., Spiegelhalter D.J., Markov Chain Monte Carlo in Practice, Chapman & Hall 1996, ISBN 0-412-05551-1

4. Gregory P, Loredo T, ApJ, 398, 146-168 (1992). A New Method for the Detection of a Periodic Signal of Unknown Shape and Period

5. Loredo, T: http://www.astro.cornell.edu/staff/loredo/bayes/

6. Loredo T, in Statistical Challenges in Modern Astronomy, ed. E.D. Feigelson and G.J. Babu (New York: Springer-Verlag) pp. 275–297 (1992).

7. Loredo T. Computational Technology for Bayesian Inference (ADASS 1999), see Loredo website

8. Neal, R. (September, 1993) Probablistic Inference using Markov Chain Monte Carlo Methods. Dept. of Computer Science, University of Toronto, available from Cambridge MCMC Preprint service (see above)

9. van Dyk, D A, Connors, A, Kashyap V L, Siemiginowska A, ApJ, 548:224-243, 2001 Analysis of Energy Spectra with Low Photon Counts via Bayesian Posterior Simulation

10. Christensen, N et al., 2001, Class. Quantum Grav. 8, 2677, Bayesian methods for cosmological parameter estimation from cosmic microwave background measurements

11. Knox, L et al., 2001, ApJL 563, L95, The Age of the Universe and the Cosmological Constant Determined from Cosmic Microwave Background Anisotropy Measurements

Chapter 2

Software Specification

2.1 Software Requirements

2.1.1 Functions and Controls

spidiffit uses SPI measured data to estimate parameters of gamma-ray emission models and assess their statistical significances. In detail, this encompasses the functions of:

1. Obtain data
2. Obtain gamma-ray emission models
3. Obtain instrument response
4. Obtain istrumental background model
5. Estimate model parameters
6. Analyze statistical significance
7. Prepare numerical results
8. Prepare map & spectra result files

These functions are controlled in detail through user settings of the parameters:

1. Data range selections
2. Model parameter bounds

3. Type of statistical analysis
4. Details of statistical analysis
5. Flag to generate map & spectral outputs

2.1.2 Interfaces

Input

1. SPI binned event data (as produced by spihist)
2. SPI pointing data
3. SPI energy bounds data
4. SPI deadtimes data
5. SPI background model
6. SPI response (IRF)
7. components skymaps SPI-SKY-IMA-IDX
8. fitting parameters specification

Output

1. Fitting parameters with error estimates
2. Posterior probability distributions of each parameter marginalized over all other parameters
3. Fit quality indicators
4. Resulting model skymap for each component and energy, with error estimates
5. Resulting model skymap summed over components, with error estimates
6. The output skymaps can be used to construct spectra for any sky region covered by the maps.

2.2 Software Architecture and Design

spidiffit is implemented in rather independent sub-modules, which can be tested separately:

- Prepare data, models, and response
- Perform statistical analysis of model parameters
- Prepare outputs and results

2.3 Development and Testing

2.3.1 Relation to other SPI software

spidiffit has much in common with other SPI programs and these can be taken over. It uses the same data and response function R_{jk} as *spiskymax* and *spiskycnv*. Much of the existing structure can therefore be used. The first version of *spidiffit* has been started using *spiskymax* as a basis.

2.3.2 Development Plan

The first implementation may be considered a prototype, for demonstration of the method and its simplified implementation. The second implementation features all above functions for linear models. The third implementation supports also non-linear models. The fourth implementation supports also a set of different priors. The fifth implementation is tbd.

2.3.3 Testing

Standard test cases are defined to measure the science validity and performance of the package. These are based on simulated data for the cases of:

- GCDE exposure pattern fit in 5 spectral bands
- Galactic-plane exposure pattern fit of CO, IR, and ff models in 1809 keV line band

2.3.4 Performance

spidiffit must run on a standard workstation (SUN), however minimum memory requirements are 1GByte (?tbd) and xxx MHz (?tbd). A fit to two 4-parameter spatial model components for all 19 detectors with one background model in 3 spectral bands should take less than 5 minutes.

[From prototyping, a test case with 231 pointings, 19 detectors and 4 image components was solved correctly for the 23 parameters in a few cpu minutes.]

Chapter 3

Method and Algorithms

3.1 Imaging Data and Models

We distinguish image space and data space in the usual way, and define the instrument response as the relation between them. The image is I_j and the expected data is d_k. The expected background is b_k Let R_{jk} be the response of data element k to image element j. Then

$$d_k = \sum_j R_{jk} I_j + b_k$$

An *image model* is a parameterized algorithm for composing an image from components. For a linear model,

$$I_j = \sum_i \theta_i M_{ij}$$

where θ_i are the model parameters.

More generally, the image will still be described by a sum of components, but the image components will be non-linear functions of the parameters (e.g. Gaussian, exponential) and each component is described by several parameters $\overline{\theta}_i$:

$$I_j = \sum_i M_{ij}(\overline{\theta}_i)$$

Similarly the background can be constructed from components of a background model B_{ik}

$$b_k = \sum_i \theta_i B_{ik}$$

where θ_i now introduces background parameters. The sums in the above expressions are over the appropriate subsets of parameters for image and

background model respectively. In this way we can treat image and background model in the same way in the subsequent analysis, and θ_i includes both. The only formal difference between image and background model is that the image is convolved with R_{jk} and the background is not:

$$d_k = \sum_j R_{jk} \sum_{i=1}^{i=N_I} \theta_i M_{ij} + \sum_{i=N_I+1}^{i=N_I+N_B} \theta_i B_{ik}$$

where there are N_I image components and N_B background components, and $N = N_I + N_B$.

In our modelling approach, the measured signal is represented through model components:

$$d_k = \sum_i \theta_i S_{ik}$$

where the sky part is:

$$S_{ik} = \sum_j R_{jk} M_{ij}, \qquad i = 1, N_I$$

and the instrumental-background part is:

$$S_{ik} = B_{ik}, \qquad i = N_I + 1, N_I + N_B$$

In the non-linear case we have to convolve explicitly for each parameter set (i.e. we cannot pre-convolve to get S_{ik}):

$$d_k = \sum_j R_{jk} \sum_{i=1}^{i=N_I} M_{ij}(\overline{\theta}_i) + \sum_{i=N_I+1}^{i=N_I+N_B} \theta_i B_{ik}$$

The procedure above is for a single energy and hence a implies a diagonal response in energy space. However it is simple to generalize to the case of a dataset including many energy channels and parameters for each energy, so that the solution constitutes an *energy deconvolution*. In this case S_{ik} includes the off-diagonal response terms. From the point of view of the analysis technique there is no difference, just S_{ik} is larger (by a factor equal to the square of the number of energy channels), and d_k and θ_i are correspondingly expanded.

3.2 Model Fitting

3.2.1 Approaches

Various approaches are possible based on well-established principles. An important point is that the data are Poisson so that any method must be able to handle such data correctly.

For COMPTEL, EGRET and other γ-ray missions the maximum-likelihood-ratio method has been widely used. The properties of this method and its used for the generation of error-estimates are well understood. It is a classical statistical method.

In recent years much attention has been paid to Bayesian methods, and their advantages are well documented (e.g. Loredo, vanDyk). Bayesian approaches have been incorporated in e.g. Chandra standard packages. Especially in GRB and cosmological statistical analyses it has become the 'method of choice' in many publications (see reference list). The main advantage comes in handling multi-parameter problems where the classical methods have fundamental limitations. In particular the handling of so-called 'nuisance parameters' is in general not possible in classical methods but straightforward in the Bayesian approach. In view of the large amount of literature on this topic explicitly for astronomical applications, and the availability of algorithms (and also software), it is proposed to take advantage of this in providing such an approach for *spidiffit* . Up to a few years ago the method was regarded as impractical for many-parameter problems because of the difficulty of evaluating multi-dimensional integrals, but now methods are readily available (Markov Chain Monte Carlo) which make this both tractable and relatively straightforward to implement.

Last but not least one of the leading Bayesian groups is at the Centre for Interdisciplinary Studies in Garching (neighbouring institute to MPE)

Possibily both approaches should be developed.

3.2.2 Model Fitting: Details

The objective of the model fitting analysis is now to extract information about θ_i in the form of posterior probability distributions, and their moments (mean, standard deviation etc) and any other functions of interest (e.g. the total image).

Fitting Constraint

The likelihood function is:

$$L(D|\bar{\theta}) = \prod_k e^{-d_k} d_k^{n_k}/n_k!$$

where n_k are the measured data (denoted collectively by D).

and the posterior probability $P(\bar{\theta}|D)$ is expressed in terms of the likelihood and the prior probability $Pr(\bar{\theta})$ using Bayes theorem:

$$P(\bar{\theta}|D) = L(D|\bar{\theta})Pr(\bar{\theta})/P(D)$$

where P(D) is known as the *evidence*.

Suitable analytical forms for the prior $Pr(\bar{\theta})$ are given by e.g. van Dyk et al. (2001); these can be used to incorporate the user's knowledge into the problem.

Results

The posterior for one parameter is obtained by marginalizing over the other parameters:

$$P(\theta_i|D) = \int_{i' \neq i} P(\bar{\theta}'_i|D) d^N\theta'$$

and its mean value is

$$< \theta_i|D >= \int \theta_i P(\theta_i|D) d\theta_i = \int \theta_i P(\bar{\theta}|D) d^N\theta$$

with standard deviation

$$\Delta\theta_i|D = sqrt \int (\theta_i - < \theta_i|D >)^2 P(\theta_i|D) d\theta_i = sqrt \int (\theta_i - < \theta_i|D >)^2 P(\bar{\theta}|D) d^N\theta$$

The results are expressed in terms of the posteriors and mean and standard deviations of the parameters, but also more conveniently in terms of the fitted skymaps, i.e. the input maps multiplied by the fitted parameters together with the error estimates.

$$< I_{ij} >=< \theta_i|D > M_{ij}$$

$$\Delta I_{ij} = (\Delta\theta_i|D) M_{ij}$$

The average total map is the sum of the average maps:

$$I_j = \sum_i \theta_i M_{ij}$$

$$< I_j >= \int \sum_i \theta_i M_{ij} P(\bar{\theta}|D) d^N\theta = \sum_i < \theta_i|D > M_{ij}$$

but the error is more complicated to compute since it involves the correlations between the θ_i:

$$\Delta I_j = sqrt \int (I_j - < I_j >)^2 P(\bar{\theta}|D) d^N\theta$$

$$\Delta I_j = sqrt \int (\sum_i \theta_i M_{ij} - \sum_i < \theta_i|D > M_{ij})^2 P(\bar{\theta}|D) d^N\theta$$

$$\Delta I_j = sqrt \int (\sum_i [\theta_i M_{ij} - < \theta_i|D > M_{ij}])^2 P(\bar{\theta}|D) d^N\theta$$

In general we want statistics for a linear function of the parameters:

$$f = \sum w_i \theta_i$$

The full posterior distribution for a given set of w_i must come from the MCMC method (see below) but the standard deviation can be more generally derived from the covariances as follows.

$$\Delta f = sqrt \int (\sum_i [\theta_i w_i - < \theta_i > w_i])^2 P(\bar{\theta}|D) d^N\theta$$

$$(\sum_i [\theta_i w_i - < \theta_i > w_i])^2$$

$$= (\sum_i w_i [\theta_i - < \theta_i >])^2$$

$$= \sum_p \sum_q w_p w_q (\theta_p - < \theta_p >)(\theta_q - < \theta_q >))$$

$$= \sum_p \sum_q w_p w_q (\theta_p - < \theta_p >)(\theta_q - < \theta_q >))$$

$$\Delta f^2 = \int \sum_p \sum_q w_p w_q (\theta_p - < \theta_p >)(\theta_q - < \theta_q >) P(\bar{\theta}|D) d^N\theta$$

$$= \sum_p \sum_q w_p w_q (< \theta_p \theta_q > -2 < \theta_p >< \theta_q > + < \theta_p >< \theta_q >)$$

$$= \sum_p \sum_q w_p w_q (< \theta_p \theta_q > - < \theta_p >< \theta_q >)$$

which shows that to get the standard deviations on any linear function of the parameters it is enough to compute the parameter means $< \theta_i >$ and covariances $< \theta_p \theta_q >$.

In the case of a single parameter $w_p = \delta_{ip}$

$$\Delta f^2 = \Delta \theta_i^2 =< \theta_i^2 > - < \theta_i >^2$$

In the case of uncorrelated parameters $< \theta_p \theta_q > - < \theta_p >< \theta_q >= \delta_{pq}$

$$\Delta f^2 = \sum_i w_i^2 (< \theta_i^2 > - < \theta_i >^2)$$

In the general case, the off-diagonal elements give the correlation between different parameters, and can be either positive or negative, i.e. can either increase or decrease the error in the linear function.

Spectral Aspects

To get a total spectrum we average over regions of the total map for each energy range. The average total map is the sum of the average maps, but the errors on the total map must be computed by an appropiate sum over the total map corresponding to each sample. This can be done by defining the w_i based on the maps M_{ij}, and using the covariance matrix. The weights are defined by

$$w_i = \sum_{j \subset A} M_{ij} / \sum_{j \subset A} 1$$

where A is a mask defining the sky region for which the spectrum is required, and the denominator just normalizes to an average intensity over the region, as required for spectra in photons cm^{-2} sr^{-1} s^{-1}.

Such spectra can be produced for any map regions, after the MCMC run (see below) is complete, so that the MCMC code does not need to use the maps explicitly, and only needs S_{ik}.

MCMC methods

Sampling from $P(\bar{\theta}|D)$ is the main computational problem, and it is proposed to use Markov Chain Monte Carlo (MCMC) methods. A good reference is Gilks et al. (1996), while Chen et al. (2000) is more advanced

and technical but has useful sections. A clear and compact presentation of the basics is in Christensen et al. (2001). Usually we are interested in the posterior for particular parameters with the rest integrated out ('marginalization'). The Metropolis-Hastings MCMC algorithm is simple to implement and does not require any special sampling functions. Data-augmentation algorithms (van Dyk et al. 2001) should also be investigated, but these are considerably more sophisticated.

The *burn-in* phase is important in getting to the right region of the parameter space before starting to compute statistics. Lack of incomplete burn-in is manifested in a slow trend in the statistics with increasing samples, instead of random fluctuations about the converged values.

The posterior, average, standard deviation and covariance matrix are easy to obtain from the sampled set of $\bar{\theta}$.

Both uniform and Cauchy (= tan(π*uniform variate)) proposal distributions are implemented, for both 'local' (changing one parameter at a time) or 'global' (changing all parameters at once) schemes. In general it has been found that the local proposal distribution is more robust and less sensitive to the stepsize than the global scheme. This is presumably because it uses information per parameter so can more easily find an accepted point. For highly correlated basis functions (extreme case: two identical maps) however the local scheme needs a very small step since it can only move at an angle to the major axis and even then does not cover the full region; here the global scheme is much better, since it has a reasonable probability to sample along the major axis. The stepsize must then be correspondingly chosen to reflect the uncertainty in a single parameter.

In all cases the stepsizes for all parameters should be chosen by trial to give an acceptance ratio around 0.23 (NB find references on this topic) and the range 0.15-0.5 is recommended (Gilks et al. 1996, page 55) although the algorithm theoretically converges to the correct result for any acceptance ratio; the value quoted should optimize the rate of convergence. In has been found that a stepsize equal to the expected uncertainty on the parameter is a good trial choice. A possible way to automate this is to use the current estimate of the standard deviation to determine the step (Gilks et al. 1996), but still a starting step is required. The results are evaluated at intervals in the sampling (interval is user-defined) and it is important to check that there is no drift in the average values since this indicates lack of convergence and the result will then be influenced by the starting parameter values.

Chapter 4

Examples

The examples illustrate typical cases, explain the choice of parameters and discuss the results. The complete data for these runs will be made available at TBD.

4.1 ^{26}Al 1.8 MeV line observed in the GCDE

This example shows how to fit the data of the Galactic Centre Deep Exposure (GCDE) to determine the profile of the ^{26}Al 1.8 MeV line, by fitting to a map of 90 GHz free-free emission, which is believed to be a good tracer of this emission. The input map was generated with *gensky* in order to obtain a broad line with FWHM = 5 keV, i.e. significantly larger than the instrumental energy resolution. The goal is to reproduce this broad line in the *spidiffit* spectral analysis.

4.1.1 Parameter file

The parameter file is:

```
debug     ,i,h, 0 ,0,2,"0=silent,1,2 gives more verbose output"

title     ,s,h,"spidiffit v0 992 free-free obs 5 energy ranges 2 equal  gaussians glo
```

debug controls the level of output, can set to 0 except for testing.

title is used to document the run for the user, and is output to the skymap FITS header. display

```
rogroup,s,h,"",,,"R/O Group"
rwgroup,s,h,"",,,"R/W Group"

counts_input_file,  s,h,  "SPI-OBS-DSP.spiskycnv_14.fits[1]",,," input count  file"
pointing_input_file,s,h,  "SPI-OBS-PNT.spisimprep_2.fits[1]",,," input pointing  file
ebounds_input_file, s,h,  "SPI-EBDS-SET.spiskycnv_14.fits[1]",,," input energy bounds
deadtime-dol,s,h,         "SPI-OBS-DTI.spisimprep_2.fits[1]",,,"DTI deadtime/livetime
background_input_file,s,h,"SPI-BMOD-DSP-IDX.spiback.fits[1]"  ,,,"input background fi
source-cat-dol,s,h,"GNRL-REFR-CAT.test_catalogue.fits[1]",,,"input catalogue of sourc
```

The input SPI data is standard, as for *spiskymax* . See the *spiskymax* Manual for details.

```
n_image_parameters,i,h, 1 ,0,100,"number of image parameters to be fit"
image_parameters,s,h,     " 1.0               ",,,  "image reference  parameters"
image_parameters_min, s,h," 0.0               ",,,  "image minimum    parameters"
image_parameters_max, s,h,"5.              ",,,  "image maximum    parameters"
image_parameters_bin, s,h," 0.1             ",,,  "image binning of parameters"
image_parameters_step,s,h," 0.1               ",,,  "image step for   parameters"
```

Defines how many images are to be fitted, and the range (min,max) and binning of the posterior probability distributions. Also the step size to be used for each parameter in the MCMC sampling run. Here there is one image to be fitted, with parameter ranging from 0.0-5.0 and the posterior to be binned in intervals of 0.1. The MCMC sampling step is 0.1. The reference parameters *image_parameters* are used for the initial simulation if *simulate=1*, which is the case here. The input image in each energy range is taken as the basis for the simulation, and since this is assumed to include the required energy dependence, the reference value of 1.0 will cause energy-dependent data to be generated.

```
image-idx_01,s,h,"SPI-SKY-IMA-IDX.gensky_10.fits[1]",,,"input skymap image index file
```

Defines the images to be used for fitting. Here there is only one image, the free-free skymap.

```
n_background_parameters,i,h, 19,0,100,"number of background  parameters to be fit"
 background_parameters,     r,h,     1.0     ,,,  "background reference  parameters"
 background_parameters_min, r,h,0.0            ,,,  "background minimum    parameters
 background_parameters_max, r,h,3.0            ,,,  "background maximum    parameters"
 background_parameters_bin, r,h,0.05           ,,, "background binning of parameters"
 background_parameters_step,r,h,0.01           ,,,"background step for   parameters"
```

Defines the background parameters: there are 19 detectors, the range for the MCMC sampling is 0.0-3.0 (all detectors treated the same), the binning for the posterior is 0.05 and the sampling step for MCMC is 0.01. In addition the simulation will be done with a scaling of the template background (BMOD input file) of 1.0.

```
 irf_input_file,              s,h,"/afs/ipp-garching.mpg.de/mpe/gamma/integral/isdc/d
```

SPI Instrument Response Function file for 19 detectors (available at ISDC under TBD).

```
n_sample, i,h,200000 , 1,1e6,"total number of MCMC samples"
m_sample, i,h,10000    , 1,1e6,"      number of MCMC samples for intermediate evaluat
n_burn_in,i,h,50000   , 1,1e6,"      number of MCMC samples for burn-in"

proposal_distribution,i,h, 11   , 11,22," MCMC proposal distribution, 11,21=uniform 1
```

The total number of MCMC samples is 200000, the results will be evaluated for intermediate printout every 10000 samples, and the burn-in phase (not used for final result) will be 50000 samples. See the description of the MCMC method for explanation of these concepts.

```
image-idx,s,h,"SPI-SKY-IMA-IDX.spidiffit_0.fits(SPI.-SKY.-IMA-IDX.tpl)",,,"output sky
image-int,s,h,     SPI-SKY-IMA.spidiffit_0.fits,,,"output skymap images file"
```

At the end of the run the best-fit skymap will be output for each energy range into this FITS file, in the standard ISDC image format. This is generated by multiplying the input map by the fitted scaling parameters for each energy range. The error estimate is output as a skymap to the same file, obtained by multiplying the input map by the estimated error on the scaling parameters. The output skymaps can be used to generate spectra by integrating over any required region of the sky.

```
source-res-dol,s,h,"SPI-SRCL-RES.spidiffit_00.awsisdc.1.fits(SPI.-SRCL-RES.tpl)",,,"o

skymap_system,              s,h,C,,,"input skymap coordinate system C=celestial, G=G

chi_0, r,  h, 180.   ,,, "Longitude of first pixel (degrees)[image-fov= SURVEY]"
chi_1, r,  h, 200.   ,,, "Longitude of last  pixel (degrees)[image-fov= SURVEY]"
d_chi, r,  h, 1.00 ,,,   "Longitude binsize        (degrees)[image-fov= SURVEY]"
psi_0, r,  h, -10.0,,,   "Latitude  of first pixel (degrees)[image-fov= SURVEY]"
psi_1, r,  h, +10.0,,,   "Latitude  of last  pixel (degrees)[image-fov= SURVEY]"
d_psi, r,  h, 1.00,,,    "Latitude  binsize        (degrees)[image-fov= SURVEY]"
```

The above 8 parameters are not used at present, but will be in future versions.

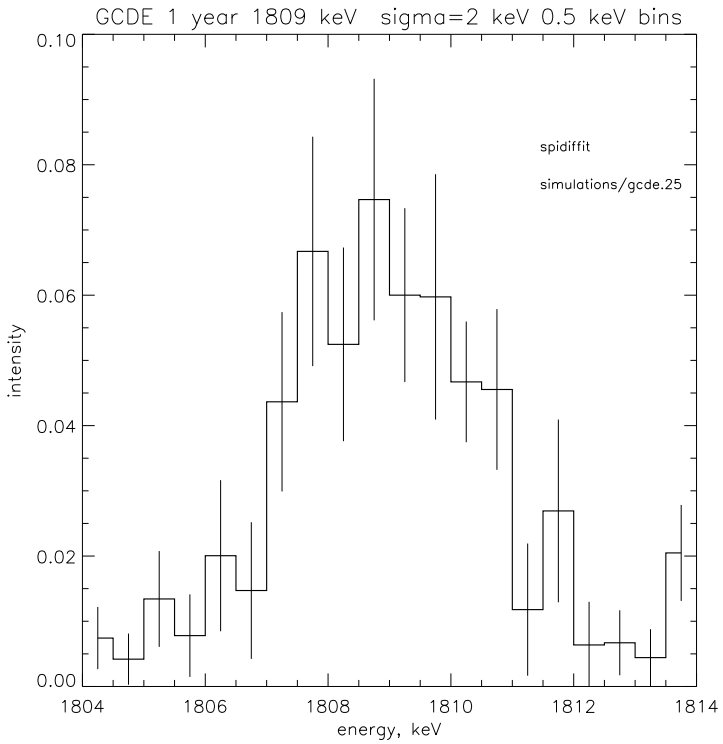
```
energy_range_min,i,h, 1 ,,,"minimum energy range sequence number as in ebounds file:
energy_range_max,i,h,20 ,,,"maximum energy range sequence number as in ebounds file:
```

Allows the user to select a sub-range of the energy ranges present in the input count spectra.

```
simulate ,i,h, 1      , 0,1  ,"0= use SPI.-OBS.-PNT data, 1=simulate"
```

simulate=0 uses the SPI counts data. *simulate*=1 performs an internal simulation based on the parameter default values and the input skymap.

4.1.2 Results: spectrum



This is the resulting spectrum of the 1.8 MeV emission: it exhibits the required broad-line shape as described above.

The output skymaps are obtained in *spidiffit* by multiplying the input free-free map by the fitted scaling factor in each energy. The error bars are obtained from the standard deviations in the same way. The spectral plot is obtained by integration over a selected part of the map (in this case the full map). NB. The creation of the spectral plot is not done in *spidiffit* itself since the choice of integration region needs to be flexible. Here is was done with an idl routine.

4.1.3 Results: parameters, error estimates, covariances

The printed output from *spidiffit* is large and here we extract the final result for one energy range.

```
Log_1  :         +++++++++        analysis for energy range #18:  1812.5-1813.0 keV +
>> mcmc

..................... many intermediate results to trace progress  ........

<< mcmc final sample:
<< final number of samples =200000
<< final number    accepted=194864
<< final fraction  accepted=0.97432

final theta_bar: 3.35898 0.989927 1.07908 1.00029 0.958477 0.991169 0.946898 1.01383
final theta_std: 1.20733 0.0357317 0.0337671 0.0395261 0.0505276 0.0392709 0.0403244
final theta_cov:

12.7404
3.31184 0.981232
3.60093 1.06844 1.16556
3.35211 0.989775 1.07961 1.00213
3.18941 0.949519 1.03485 0.958919 0.921231
3.31199 0.981839 1.06983 0.991369 0.950927 0.983958
3.17793 0.93767 1.02177 0.946937 0.907089 0.93862 0.898242
3.41271 1.00337 1.09379 1.0144 0.971104 1.00468 0.959761 1.02959
3.17176 0.936956 1.02164 0.946906 0.907735 0.938571 0.895959 0.959867 0.897299
3.25875 0.964793 1.05141 0.974752 0.934623 0.966073 0.922699 0.987637 0.922168 0.9503
3.06533 0.900969 0.981897 0.910307 0.872184 0.901954 0.8616 0.922844 0.861491 0.88659
3.31242 0.974767 1.06271 0.985184 0.944148 0.976102 0.932474 0.998692 0.93274 0.95937
3.37283 0.992852 1.08211 1.00306 0.960874 0.994185 0.949973 1.01668 0.949326 0.977237
3.23535 0.953715 1.03955 0.963557 0.923624 0.954494 0.911846 0.977019 0.911901 0.9385
3.17468 0.9345 1.01888 0.944745 0.904647 0.936082 0.894335 0.957713 0.894276 0.919791
3.4022 1.00526 1.09572 1.01628 0.973166 1.00664 0.961825 1.03021 0.961566 0.989526 0.
3.34462 0.988146 1.07765 0.998886 0.956998 0.989249 0.945198 1.01229 0.945122 0.97254
3.36394 0.995252 1.08463 1.00509 0.963456 0.996481 0.951722 1.01833 0.950994 0.979192
3.60941 1.05558 1.15084 1.06715 1.02172 1.05679 1.00998 1.08108 1.00933 1.03855 0.970
3.0841 0.913699 0.995449 0.922693 0.88436 0.915184 0.873608 0.935192 0.873429 0.89883
theta_cov[i,m]-<theta[i]><theta[m]>:

1.45764
-0.0133133 0.00127676
-0.0236928 0.000228816 0.00114022
-0.00783268 -0.00043465 0.000220807 0.00156231
```

```
-0.0300944 0.000696612 0.000578001 0.000168958 0.00255303
-0.0173308 0.000653788 0.000273325 -8.20705e-05 0.000914253 0.00154221
-0.00268236 0.000309598 -1.09766e-05 -0.000231393 -0.000491175 8.40657e-05 0.00162606
0.00726555 -0.000249641 -0.000218608 0.000276675 -0.000630302 -0.000199432 -0.0002347
-0.00806244 -0.000170695 0.000118138 -2.55406e-05 0.000381307 0.000268591 -0.00043361
-0.0135363 0.000415277 0.000174654 0.00028379 0.000884096 0.000485409 0.000239508 -2.
0.0082431 1.34023e-05 -0.000200879 -7.56768e-05 -0.000147593 -0.000131418 -0.00019433
0.00400916 -0.000254562 -0.000127322 -4.05247e-05 0.000102154 -0.000143433 -0.0001664
0.0039232 1.51533e-06 -0.000160704 -0.000182017 -0.000434379 8.80319e-05 0.000278463
-0.000670572 2.55734e-05 -3.20138e-05 -0.000111361 0.000233364 -0.000392158 -0.000389
0.00228927 -0.00043604 -0.000259565 2.65494e-05 -0.000585382 -2.7119e-05 3.73819e-05
-0.00919563 -0.000110701 -0.000195952 0.000390179 -0.000266139 9.98408e-06 0.00015278
-0.00907663 -0.000222384 0.000269783 0.000175134 2.94351e-05 -0.000360166 -0.00020931
-0.0111903 0.000567218 0.000363088 -2.69602e-06 0.000371939 0.00054799 0.000272423 -0
0.0265283 -0.000330423 -0.000170055 0.000188492 -0.000643889 -0.000450356 -3.52116e-0
-0.0145309 0.000499537 3.79791e-06 -6.14954e-05 0.000172732 0.000838839 0.000102653 -
final posterior for parameter #0:  0 0 0 0 37 783 3963 2691 2384 1534 1857 1144 1839
final posterior for parameter #1:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2152 27099 79185
final posterior for parameter #2:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 4139 29634 1
final posterior for parameter #3:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 217 15139 85984
final posterior for parameter #4:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 26793 62530 6280
final posterior for parameter #5:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 642 31922 80109
final posterior for parameter #6:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 29138 75169 7568
final posterior for parameter #7:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1319 12390 54637
final posterior for parameter #8:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1082 13708 92396 7
final posterior for parameter #9:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1741 50503 10089
final posterior for parameter #10:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2590 68356 111237
final posterior for parameter #11:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 351 26587 10763
final posterior for parameter #12:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1078 18680 6495
final posterior for parameter #13:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 9586 58761 9702
final posterior for parameter #14:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 19973 87781 791
final posterior for parameter #15:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 10001 63804 7
final posterior for parameter #16:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 13789 86203 8
final posterior for parameter #17:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 19 7183 88562 7
final posterior for parameter #18:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 5232 7355 518
final posterior for parameter #19:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3199 52853 96791
<< mcmc
```

There are 20 parameters: the scaling factor of the free-free map (parameter 0) plus the 19 background scaling values (parameters 1-19).

The main result is given by the average and standard deviations:

```
final theta_bar: 3.35898 0.989927 1.07908 1.00029 0.958477 0.991169 0.946898 1.01383
final theta_std: 1.20733 0.0357317 0.0337671 0.0395261 0.0505276 0.0392709 0.0403244
```

This shows the value for the scaling-factor is 3.36 ± 1.2, and the background scaling values are all near 1.0 (as expected) with a small error of order 3%. The posterior distributions reflect the fairly large error on the scaling factor and the small error on the background values (since they are determined well from the large amount of data).

As explained in the mathematical section, the covariance gives the degree of correlation between the parameters. Thus the following extract shows that the correlation between the free-free scaling factor and the background is small, as is the correlation between the background values; this is reflected in the corresponding off-diagonal terms.

```
theta_cov[i,m]-<theta[i]><theta[m]>:

1.45764
-0.0133133 0.00127676
-0.0236928 0.000228816 0.00114022
-0.00783268 -0.00043465 0.000220807 0.00156231
```

Chapter 5

Help File

This is a copy of the help file spidiffit.txt which is included with the software delivery and should be consulted for details of the current version.

```
NAME
spidiffit 2.0

USAGE
spidiffit



NB For simulations, this executable depends on previous runs
of  gensky, os_pdefgen, spipoint, spidead, spiback and
spiskycnv.  (More conveniently  simulations can be performed
with spisimprep instead of spipoint and spidead especially for
large survey applications.)

For real data, depends on previous runs of spipoint, spidead,
spiback and spihist.



DESCRIPTION

        Performs fitting of diffuse models to SPI data, using Bayesian
        methods with Monte Carlo Markov Chain evaluation.
```

```
        Full mathematical and algorithmic details are given in the
        ADD diffuse_model_fitting_ADD.ps included with this delivery.

        The models are specified as (multiple-energy) skymaps in the form of
         SPI.-SKY.-IMA-IDX (index) and SPI.-SKY.-IMA (images) datasets.

The instrument response is in the form of IRF (Instrument
Response Function) data as provided by GSFC using GEANT
simulations.

Creates fitting output skymaps in form of  SPI.-SKY.-IMA-IDX (index)
 and SPI.-SKY.-IMA (images) dataset.

       The pointings are defined by a SPI.-OBS.-PNT dataset.

       Deadtime is defined by a SPI.-OBS.-DTI dataset.

       Background is defined by the output of spiback, in the form of
SPI.-BMOD-DSP-IDX (index) and SPI.-BMOD-DSP (data). The
background values for each detector are fitted.

The input count data  may contain multiple energies, in which
case the analysis is performed for a range of energies
which can be specified.


       The statistical output is at present to the "screen" which
       is best sent to a file: spidiffit>spidiffit_results.
       The parameters' mean and standard deviation are output as theta_bar,
       theta_std. The posterior for each parameter is output as
       "posterior for parameter #". Intermediate results are output as
       the MCMC sampling proceeds, as specified by parameter m_sample.
       The final results after n_sample samples for each energy
       can be seen by "grep final spidiffit_results".

CHANGES
Version 1  first version 10 Dec 2001
Version 2  11 Dec 2001 corrected for linux (RAND_MAX); improved output formatting
```

```
PARAMETERS


        debug
         enables debugging. Normally 0.
         0=silent. 1,2 gives more verbose output


        title
         user's description of the run: serves 2 purposes
         1. identifies this run (via the par file)
         2. used to write as comment to FITS header of the output skymap datasets.

rogroup
    read only group for input. If empty, uses the filenames directly.

rwgroup
write group for output.  If empty, no group is created.


        irf_input_file
          input IRF index file which contains the SPI response information
          as generated by GEANT simulations at GSFC.
          Detectors must correpond to the counts and deadtime datasets.

          Singles and singles+doubles are available at present:
          singles:                /isdc/integration/ic/spi/rsp/spi_irf_grp_0005.fits
          doubles+triples:        /isdc/integration/ic/spi/rsp/spi_irf_grp_0006.fits
          singles+doubles+triples:/isdc/integration/ic/spi/rsp/spi_irf_grp_0007.fits


         These IRFs belong  in the IC (Instrument Characteristics) area at ISDC.


        counts_input_file
         input detector count specta file (template SPI.-OBS.-DSP.tpl)

        pointing_input_file
         input pointing file (template SPI.-OBS.-PNT.tpl)
```

```
deadtime-dol
  input deadtime file of type SPI.-OBS.-DTI created by spidead.

ebounds_input_file
 input energy bounds file (template SPI.-EBDS-SET.tpl)

background_input_file
 input background index file (SPI.-BMOD-IDX)

image-idx_01
 input  index for image component #1  (template SPI.-SKY.-IMA-IDX.tpl)
 Contains multiple energies.

 etc...:
image-idx_nn
 input  index for image component #nn  (template SPI.-SKY.-IMA-IDX.tpl)



source-cat-dol
 input catalogue dataset (template GNRL-REFR-CAT)
 This specifies the source positions to be analysed. Select sources using SEL
 [not used in this version]

n_image_parameters
 number of parameters multiplying input images, equal to number of image comp

image_parameters
 values of parameters to be used for data simulation, as vector.

image_parameters_min
 minimum of range of each parameter for posterior distribution, as vector.


image_parameters_max
 maximum of range of each parameter for posterior distribution, as vector.

image_parameters_bin
 binsize for each parameter for posterior distribution, as vector.
```

```
image_parameters_step
 stepsize for each parameter for posterior distribution, as vector.
 The steps used are either uniformly distributed:  theta=image_parameters_ste
                   or given a Cauchy distribution: theta=image_parameters_ste
 where u is uniform in [0,1].
 The type of distribution is controlled by the parameters proposal_distributi


n_background_parameters
 number of parameters multiplying input backgrounds, equal to number of backg

background_parameters
 values of parameters to be used for data simulation, as vector.

background_parameters_min
 minimum of range of each parameter for posterior distribution, as vector.


background_parameters_max
 maximum of range of each parameter for posterior distribution, as vector.

background_parameters_bin
 binsize for each parameter for posterior distribution, as vector.

background_parameters_step
 stepsize for each parameter for posterior distribution, as vector.
 The steps used are either uniformly distributed : theta=background_parameter
                   or given a Cauchy distribution: theta=background_parameter
 where u is uniform in [0,1].
 The type of distribution is controlled by the parameters proposal_distributi


proposal_distribution,i,h,
 MCMC proposal distribution (see ADD for detailed explanation)
  11 = uniform, single parameter ("local")
  21 = uniform, multiparameter   ("general")
  12 = Cauchy,  single parameter ("local")
  22 = Cauchy,  multiparameter   ("general")
 "local" is more effective/robust in most cases (see ADD).
```

```
 n_sample
  total number of MCMC samples after burn-in, per energy range.

 m_sample
   number of MCMC samples for intermediate evaluations and print-out


 n_burn_in
   number of MCMC samples for initial burn-in phase.

 energy_range_min
  minimum energy range sequence number as in ebounds file: 1,2,3...
  energy ranges energy_range_min to energy_range_max are processed.

 energy_range_max
  maximum energy range sequence number as in ebounds file: 1,2,3...


image-idx
 output skymap index file (template SPI.-SKY.-IMA-IDX.tpl)

image-int
 output skymap images file  (template SPI.-SKY.-IMA.tpl)

source-res-dol
 output source catalogue containing flux spectra with error bars (template SPI
 [not used in this version]

skymap_system
  input skymap coordinate system C=celestial, G=Galactic"
  This tells spiskycnv how to treat the input skymap, which does not
  itself specify the basis system. Usually Galactic is used.



 [ following 6 parameters not used in this version]
 chi_0
  Longitude of first pixel (degrees). Centre of pixel (as FITS convention)
```

```
        chi_1
         Longitude of last pixel (degrees).  Centre of pixel (as FITS convention)
        d_chi
         Longitude binsize (degrees).

        psi_0
         Latitude of first pixel (degrees). Centre of pixel (as FITS convention)

        psi_1
         Latitude of last pixel (degrees). Centre of pixel (as FITS convention)

        d_psi
         Latitude binsize (degrees).


        simulate
         0=simulate the data on the basis of the input maps, the response and backgro
         1= use data SPI.-OBS.-DSP detector counts spectra (from spihist)



EXAMPLES


debug     ,i,h, 0 ,0,2,"0=silent,1,2 gives more verbose output"
display   ,i,h, 0 ,0,2,"1,2 displays  skymap using root"
title     ,s,h,"spidiffit v1 231 obs  Gaussian 9 energy ranges 2  gaussians global MC

rogroup,s,h,"",,,"R/O Group"
rwgroup,s,h,"",,,"R/W Group"

counts_input_file,    s,h,  "SPI-OBS-DSP.spiskycnv_14.gcde10.fits[1]",,," input count
pointing_input_file,  s,h,  "SPI-OBS-PNT.spisimprep_2.gcde10.fits[1]",,," input point
ebounds_input_file,   s,h,  "SPI-EBDS-SET.spiskycnv_14.gcde10.fits[1]",,," input ener
deadtime-dol,         s,h,  "SPI-OBS-DTI.spisimprep_2.gcde10.fits[1]",,,"DTI deadtime
background_input_file,s,h,  "SPI-BMOD-DSP-IDX.spiback.gcde10.fits[1]"  ,,,"input back
source-cat-dol,       s,h,  "GNRL-REFR-CAT.test_catalogue.fits[1]",,,"input catalogue
```

```
xn_image_parameters,i,h, 1 ,0,100,"number of image parameters to be fit"
ximage_parameters,s,h,     " 4.0               ",,,  "image reference  parameters"
ximage_parameters_min, s,h," 0.0               ",,,  "image minimum    parameters"
ximage_parameters_max, s,h,"10.0            ",,,  "image maximum    parameters"
ximage_parameters_bin, s,h," 0.1             ",,,  "image binning of parameters"
ximage_parameters_step,s,h," 0.050            ",,,  "image step for   parameters"


 n_image_parameters,i,h, 2 ,0,100,"number of image parameters to be fit"
 image_parameters,s,h,     " 1.5   2.5           ",,,  "image reference  parameters"
 image_parameters_min, s,h," 0.0   0.0           ",,,  "image minimum    parameters"
 image_parameters_max, s,h,"5.   5.          ",,,  "image maximum    parameters"
 image_parameters_bin, s,h," 0.1   0.1          ",,,  "image binning of parameters"
 image_parameters_step,s,h," 0.1   0.1           ",,,  "image step for   parameters


image-idx_01,s,h,"SPI-SKY-IMA-IDX.gensky_9.awsmpe.1.gauss.fits[1]",,,"input skymap im
image-idx_02,s,h,"SPI-SKY-IMA-IDX.gensky_9.awsmpe.3.gauss.fits[1]",,,"input skymap im
image-idx_03,s,h,"SPI-SKY-IMA-IDX.gensky_9.awsmpe.3.gauss.fits[1]",,,"input skymap im
image-idx_04,s,h,"SPI-SKY-IMA-IDX.gensky_9.awsmpe.4.gauss.fits[1]",,,"input skymap im

n_background_parameters,i,h, 19,0,100,"number of background  parameters to be fit"
 background_parameters,     r,h,     1.0     ,,,  "background reference  parameters"
 background_parameters_min, r,h,0.0           ,,,  "background minimum    parameters
 background_parameters_max, r,h,3.0          ,,,  "background maximum    parameters"
 background_parameters_bin, r,h,0.05          ,,, "background binning of parameters"
 background_parameters_step,r,h,0.01          ,,,"background step for   parameters"



irf_input_file,             s,h,"/afs/ipp-garching.mpg.de/mpe/gamma/integral/isdc/da

n_sample, i,h,20000 , 1,1e6,"total number of MCMC samples"
m_sample, i,h,1000    , 1,1e6,"      number of MCMC samples for intermediate evaluati
n_burn_in,i,h,10000   , 1,1e6,"      number of MCMC samples for burn-in"

proposal_distribution,i,h, 11   , 11,22," MCMC proposal distribution, 11,21=uniform 1
```

```
image-idx,s,h,"SPI-SKY-IMA-IDX.spidiffit_1.fits(SPI.-SKY.-IMA-IDX.tpl)",,,"output sky
image-int,s,h,     SPI-SKY-IMA.spidiffit_1.fits,,,"output skymap images file"



source-res-dol,s,h,"SPI-SRCL-RES.spidiffit_1.fits(SPI.-SRCL-RES.tpl)",,,"output catal

skymap_system,              s,h,C,,,"input skymap coordinate system C=celestial, G=G

chi_0, r,  h, 180.   ,,, "Longitude of first pixel (degrees)[image-fov= SURVEY]"
chi_1, r,  h, 200.   ,,, "Longitude of last  pixel (degrees)[image-fov= SURVEY]"
d_chi, r,  h, 1.00 ,,,   "Longitude binsize        (degrees)[image-fov= SURVEY]"
psi_0, r,  h, -10.0,,,   "Latitude  of first pixel (degrees)[image-fov= SURVEY]"
psi_1, r,  h, +10.0,,,   "Latitude  of last  pixel (degrees)[image-fov= SURVEY]"
d_psi, r,  h, 1.00,,,    "Latitude  binsize        (degrees)[image-fov= SURVEY]"

energy_range_min,i,h, 1 ,,,"minimum energy range sequence number as in ebounds file:
energy_range_max,i,h, 1 ,,,"maximum energy range sequence number as in ebounds file:


simulate ,i,h, 1      , 0,1  ,"0= use SPI.-OBS.-DSP data, 1=simulate"
```

```
BUGS
None known yet.

Send all bugs to aws@mpe.mpg.de

AUTHOR

A.W.Strong, MPE
```

```
aws@mpe.mpg.de

SEE ALSO
        diffuse_model_fitting_ADD.ps (included in this delivery)
        SPI ICD
gensky.txt
        spipoint.txt
        spihist.txt
        spiback.txt
        spiskycnv.txt
        spiskymax.txt
        MEMSYS5 User's Manual (to be found at http://www.maxent.co.uk)
```