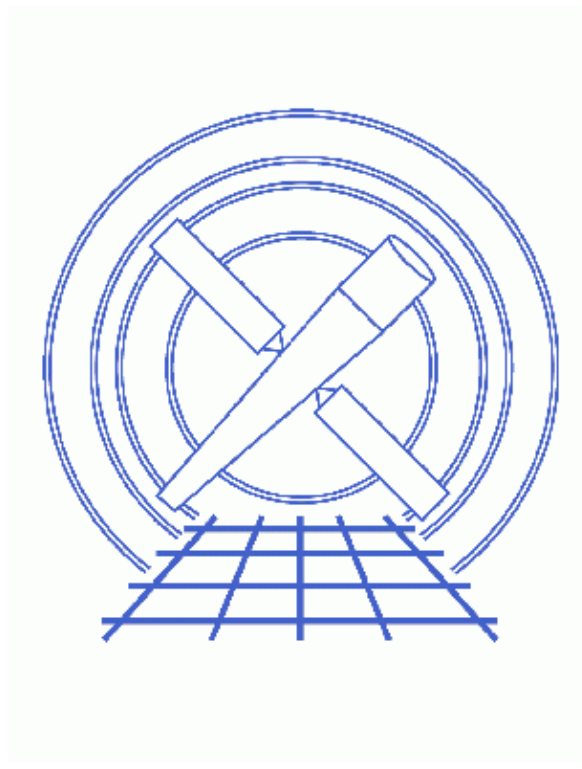


Simulating 1–D Data: the S–lang Script simspec



Sherpa Threads (CIAO 3.4)

Table of Contents

- Getting started
- Obtaining Response Files
- Introduction to simspec
- Download the script
- Running simspec interactively
- Running simspec non-*interactively*
- The simspec usefile parameter
- Addition of a Background file
 - ◆ Choosing a background scale factor
- Parameter files:
 - ◆ simspec
- History
- Images
 - ◆ Plot of the simulated data and fit

Simulating 1–D Data: the S–lang Script simspec

Sherpa Threads

Overview

Last Update: 4 Jan 2007 – updated to use Cycle 9 proposal planning response files

Synopsis:

`simspec` is a command–line script that uses the *Sherpa S–Lang module* to generate and optionally fit a simulated PHA spectrum. The script employs *Sherpa's FAKEIT command*, and automates the procedure described in the [FAKEIT thread](#). It is intended to simplify the process of running simulations (particularly for Chandra observation proposal planning), and allows for script–driven generation of numerous spectra.

Proceed to the [HTML](#) or [hardcopy \(PDF: A4 / letter\)](#) version of the thread.

Getting started

Please follow the "[Sherpa Threads: Getting Started](#)" thread.

Obtaining Response Files

The ready–made response files for proposal planning are available from [the Chandra Proposal Planning CALDB page](#). These files can be used in `simspec` to estimate the spectra of the proposed observations.

The examples in this thread use the ACIS–S default pointing position files:

```
aciss_aimpt_cy09.arf
aciss_aimpt_cy09.rmf
```

The background file used in the examples – `back.pi` – is included in the [sherpa.tar.gz](#) data package.

Introduction to simspec

Input to `simspec` falls into three categories:

- simulation parameters (source model, RMF, ARF, background dataset, exposure time, background scale factor),
- normalization parameters (method, value, source model parameter, lower/upper bound for calculation), and
- fit parameters (source model, lower/upper bound of data filter, confidence–interval estimation)

method)

Note that the syntax for model expression in both simulation and fitting parameters is exactly the same as used in *Sherpa*.

When run, the script prints fit results (best-fit parameters, goodness-of-fit statistics, confidence intervals) to the screen and produces three output files:

- a PHA file containing the simulated spectrum;
- an MDL file ('[ahelp_md1](#)') containing information on source/background datasets, source and instrument models, and source model parameters; and
- a postscript file containing a plot of the simulated data and/or fit (if requested)

The script will optionally group the data before fitting.

Download the script

This thread uses the `simspec` script. The most recent version of `simspec` is v1.1 (09 February 2006):

```
unix% grep Version `which simspec`  
% Version 1.1 (2006/02/09)
```

Please check that you are using the most recent version before continuing. If you do not have the script installed or need to update to a newer version, please refer to the [Scripts page](#).

Running simspec interactively

If `simspec` is run with no command-line arguments, the script will prompt the user for all necessary input parameters (including source-model parameter values). This should make running the script fairly straightforward for first-time users.

The following example demonstrates how to respond to the input prompts in order to create and fit a basic spectrum. (For details on each of the input parameters, see the [parameter listing](#).) **Note that the lines in italic are comments on the parameter settings, not output from the script itself.**

```
unix% simspec  
  Abundances set to Anders & Grevesse  
Model for simulating data (:): xsphabs[abs]*powlaw1d[p1]  
  
  The source model is an absorbed power law ("xsphabs[abs]*powlaw1d[p1]").  
  
Input RMF file (:): aciss_aimpt_cy09.rmf  
Input ARF file (:): aciss_aimpt_cy09.arf  
Input background data file (NONE): back.pi  
  
  The simulation uses an RMF, an ARF, and a background dataset.  
  
Exposure time [s] (0:) (1): 10000  
  
  The simulated dataset has an exposure time of 10000 s.  
  
Background scale factor (0:) (1): 1e-05
```

Fakeit Simulations with simspec – Sherpa

The background scale (1e-05) will be recorded in the BACKSCAL header keyword of the output PHA file; see the [discussion on backgrounds](#) for more information.

Grouping type (NONE|NUM_CTS|ADAPTIVE|SNR|ADAPTIVE_SNR) (NONE): NUM_CTS
Grouping type value (0.0): 20

The simulated data is grouped to 20 counts per bin.

Normalization method (NONE|EFLUX|PFLUX|COUNTS|COUNT_RATE) (NONE): EFLUX
Normalization value (0:) (:): 1e-12
Model parameter to use for normalization (:): p1.ampl
Lower bound for normalization [keV] (0:) (:): 2
Upper bound for normalization [keV] (0:) (:): 10

The data are normalized so that the energy flux, integrated over the range 2 to 10 keV, is 1.0e-12.

Model for fitting data (NONE|SIM|[model_expr]) (:): SIM
Lower bound for fit [keV] (0:))normmin -> 2): 0.5
Upper bound for fit [keV] (0:))normmax -> 10): 8.0
Confidence-interval estimation method (NONE|UNC|COV|PROJ) (NONE):
Output PHA file (:): sim_1.pha
Output MDL file (:): sim_1.mdl.fits
Output postscript plot (NONE): sim_1.ps
Output plot type (DATA|FIT|FIT_DELCHI|FIT_RATIO) (FIT_DELCHI):

Enter simulation source model parameters:
abs.nH parameter value [0.1] 0.0568
p1.gamma parameter value [1] 2.0
p1.ref parameter value [1]
p1.ampl parameter value [1]
WARNING: any applied filters are being deleted!
Write X-Axis: Bin Y-Axis: Counts

Enter fit source model parameters:
abs.nH parameter value [0.0568]
p1.gamma parameter value [2]
p1.ref parameter value [1]
p1.ampl parameter value [0.000389837]

The fit is performed using the same source model and parameter settings employed in the simulation.


FIT RESULTS

LVMQT: V2.0
LVMQT: initial statistic value = 69.4889
LVMQT: final statistic value = 65.1056 at iteration 3
abs.nH 0.0514404 10**22 atoms/cm**2
p1.gamma 1.99571
p1.ampl 0.000366255

Goodness: computed with Chi-Squared Gehrels

DataSet 1: 95 data points -- 92 degrees of freedom.
Statistic value = 65.1056
Probability [Q-value] = 0.984882
Reduced statistic = 0.707669

Done!

Plots of the simulated data and fit are stored in the postscript file, as [shown here](#) . The output MDL file can be read into *Sherpa* as follows:

```

unix% sherpa
...
sherpa> read mdl "sim_1_mdl.fits"

```

This will load the simulated spectrum and background dataset, create the source and instrument models used in the fit, and set the source model parameters to their best-fit values.

Running simspec non-interactively

The previous example can also be run non-interactively by setting all input parameters on the command line (or, equivalently, setting them with the `psset` command). Using the `simspec` in this manner allows the user to script and run multiple simulations efficiently.

In order to prevent `simspec` from prompting for model parameters, the parameter values for the simulation and fit must be specified using the `simplparam` and `fitparams` parameters, respectively, as shown here:

```

unix% simspec simmodel="xsphabs[abs]*powlawld[p1]" rmffile=aciss_aimpt_cy09.rmf arffile=aciss_aimpt_cy0
  bkgfile=back.pi exposure=10000 backscale=1e-05 grouptype=NUM_CTS grouptypeval=20 normtype=EFLUX \
  normval=1e-12 normparam=p1.ampl normmin=2 normmax=10 fitmodel=SIM \
  fitmin=0.5 fitmax=8 paramest=NONE outfile=sim_2.pha mdlfile=sim_2_mdl.fits \
  psplotfile=sim_2.ps plottype=FIT_DELCHI \
  simplparams="abs.nh=0.0568,p1.gamma=2.0" fitparams=SIM

  Abundances set to Anders & Grevesse
WARNING: any applied filters are being deleted!
Write X-Axis: Bin Y-Axis: Counts

-----
FIT RESULTS
-----

LVMQT: V2.0
LVMQT: initial statistic value = 63.9829
LVMQT: final statistic value = 62.076 at iteration 4
      abs.nh  0.081979  10**22 atoms/cm**2
      p1.gamma  2.10285
      p1.ampl  0.00041951

Goodness: computed with Chi-Squared Gehrels

DataSet 1: 98 data points -- 95 degrees of freedom.
Statistic value      = 62.076
Probability [Q-value] = 0.996413
Reduced statistic    = 0.653432

Done!

```

The simspec usefile parameter

The `simspec` script uses the *Sherpa* default values for the optimization method, statistics, and background subtraction, among other things. The `usefile` parameter gives users a means to changing any of these values without altering the `simspec` script itself.

This parameter takes a file containing a list of *Sherpa* commands. The script evaluates each line of this file (first as a *Sherpa* statement and then, if not recognized, as a S-lang statement); the evaluation happens before the models are set up.

Fakeit Simulations with simspec – Sherpa

For example, if one wanted `simspec` to use the [Powell optimization method](#) instead of [Levenberg–Marquardt](#) (the default):

```
unix% cat alter_simspec.txt
method powell
statistic chi dvar
```

Then set the `usefile` parameter to point to this text file:

```
unix% pset simspec usefile=alter_simspec.txt
```

The method used is recorded in the output MDL file:

```
unix% dmlist mdl.fits"[MDL_Models][cols method]" data rows="1:1"
-----
Data for Table Block MDL_Models
-----
ROW      method
-----
1      Powell
```

Addition of a Background file

If a background file was entered to `simspec` then the simulated PHA file contains both the source and background counts. Addition of background counts to the simulations should be properly scaled using the background scale factor; see the [Choosing a background scale factor subsection](#). The default value of `backscale` parameter is set to 1. The total counts are calculated using the expression:

```
Total(i) = S(i) + ((beta_S * t_S) / (beta_B * t_B)) * B(i)
```

where $S(i)$ is the source counts in channel i , $B(i)$ is the background counts in channel i , t_S and t_B are the source and background exposure times respectively, and β_S and β_B are the source and background area scales defined as the ratio of data extraction area to total detector area. For the purpose of the simulations one needs to adjust both the exposure time and the area scales using the `backscale` `simspec` parameter to a desired values given the input background file.

In order to properly fit the simulated PHA data the background model component needs to be either included in the model expression `model_expr` or subtracted from the simulated data. Currently `simspec` subtracts the background by default while fitting the simulated data. The subtraction limits the allowed choice of statistics to [chi_ghrels](#) and [chi_dvar](#).

Choosing a background scale factor

The `backscale` parameter scales the background relative to the simulated spectrum. The value you choose for `backscale` defines the size of the region to which the simulated PHA file corresponds. Suppose you create a background PHA file by extracting it from a real dataset using, for example, [dmextract](#), in a circular region of area $AB = 5000$ pixels. To find the region area in your background file in pixels:

```
unix% dmlist back.pi subspace | grep area
16 sky                      Real4          Field area = 4.63168e+77 Region area = 1256.64
```

The region area is 1256 pixels in this case.

If you want to create a simulated PHA file that corresponds to counts from a source in an area $AS = 10$ pixels, the `backscale` is calculated as `source_area/background_area` or, equivalently, AS/AB :

Fakeit Simulations with simspec – Sherpa

```
backscale = 10 / 1256 = 0.00796178
```

Alternatively, you may be trying to match an existing source PHA file and using the same background file created for it, in which case

```
unix% dmlist sourcepha.pi keys | grep BACKSCAL
```

will show the value of backscale for that file. It should be equal to the ratio of region areas for the source and background files.

If the simulation exposure time and the background exposure times are different, the backscale should compensate for that as well. This is done by using the equation for $Total(i)$ given at the beginning of this section.

Parameters for /home/username/cxcds_param/simspec.par

```
unix% plist simspec
```

Parameters for /home/user/cxcds_param/simspec.par

```
#
# Simulation parameters
#
  simmodel = xsphabs[abs]*powlaw1d[pl] Model for simulating data
  rmffile = aciss_aimpt_cy08.rmf Input RMF file
  arffile = aciss_aimpt_cy08.arf Input ARF file
  bkgfile = back.pi Input background data file
  exposure = 10000 Exposure time [s]
  backscale = 1e-05 Background scale factor
  grouptype = NUM_CTS Grouping type
  grouptypeval = 20 Grouping type value
#
# Normalization parameters
#
  normtype = EFLUX Normalization method
  normval = 1e-12 Normalization value
  normparam = pl.ampl Model parameter to use for normalization
  normmin = 2 Lower bound for normalization [keV]
  normmax = 10 Upper bound for normalization [keV]
#
# Fit parameters
#
  fitmodel = SIM Model for fitting data (NONE|SIM|<model_expr>)
  fitmin = 0.5 Lower bound for fit [keV]
  fitmax = 8 Upper bound for fit [keV]
  paramest = NONE Confidence-interval estimation method
#
# Simulation output
#
  outfile = sim_1.pha Output PHA file
  mdlfile = sim_1.mdl.fits Output MDL file
  psplotfile = sim_1.ps Output postscript plot
  plottype = FIT_DELCHI Output plot type
#
# Model parameters (leave empty to be prompted)
#
  (simparams = ) Parameters for simulation model (DEF|<param_list>)
  (fitparams = ) Parameters for fit model (DEF|SIM|<param_list>)
#
# Sherpa parameters
#
```


Fakeit Simulations with simspec – Sherpa

```
(usefile = )           File of Sherpa commands to run
#
# Other parameters
#
(clobber = no)         Clobber existing output files?
(verbose = 0)          Verbosity level
(mode = ql)
```

History

- 22 Feb 2005 original version, new for CIAO 3.2
- 03 Mar 2005 added Obtaining Response Files section
- 21 Dec 2005 reviewed for CIAO 3.3: no changes
- 08 Feb 2006 `simspec` v1.1 (change to `cxchistory.sl`, one of the scripts called by `simspec`); updated to use Cycle 8 proposal planning response files
- 01 Dec 2006 reviewed for CIAO 3.4: no changes (yet: will be updated to use Cycle 9 proposal planning response files when they are available)
- 04 Jan 2007 updated to use Cycle 9 proposal planning response files

URL: <http://cxc.harvard.edu/sherpa/threads/simspec/>

Last modified: 4 Jan 2007

Image 1: Plot of the simulated data and fit

