Chandra X-Ray Center

Grating Data Products:
Level 2 to CXC Archive Interface Control Document

(`http://hea-www.harvard.edu/asclocal/icd/TBD`)[1]

Revision 1.4—28 May 1999

Submitted: ______________________ ________

David Huenemoerder — Date

Grating Scientists, CXC Science Data Systems

______________________ ________

John Houck — Date

Grating Scientists, CXC Science Data Systems

Concurred: ______________________ ________

Arnold Rots — Date

CXC Data Systems Archive Scientist, Pipelines

Concurred: ______________________ ________

Janet De Ponte — Date

CXC Data Systems Group Leader, Pipelines

Concurred: ______________________ ________

Martin Elvis — Date

Manager, CXC Science Data Systems

Approved: ______________________ ________

Harvey Tananbaum — Date

Director, CXC

[1]Current DRAFT version can be found at: `http://space.mit.edu/ASC/docs/ICD_TGL2.ps.gz`.
The top-level ICD index is at `http://hea-www.harvard.edu/asclocal/sds/development/ICD/index.html`

Document and Change Control Log			
Date	**Version**	**Section**	**Status**
23 Apr 98	0.0	all	Initial Draft, for review by J.Houck
4 May 98	1.0	all	Update for review
8 July 98	1.1	all	Update for release; ARF deleted from Type II PHA file. Type II PHA header refined. Added spectral measurement products.
9 August 98	1.2	4.2, 4.5	Small changes to column specifications; new rows in REGION extension to describe background.
22 October 98	1.2(DRAFT)		Added changebars. Updated unresolved issues (pending changes).
11,26 February 99	1.3	all	– Added language to support Type I PHA. – Added GTI HDU to PHA files. – Removed arf, rmf products: they have separate ICD. – Deleted ARF section; redundant w/ ExpMap ICD. – Removed summary products: TBS by USG in another ICD?
28 May 99	1.3	all	Misc edits for release. CO#'s: 0017–0024,

Unresolved Issues

The following is a list of unresolved, un-reviewed, or un-implemented items:

1. 980709: Summary products deferred to USG (postscript, gif, html, TBD).
2. Need to specify grating redistribution matrix function
3. Need to add details for L2 event files.
4. Need to add zero-order PHA spectrum to L2 grating spectrum file. (for next revision, 1.5, Aug99).

Contents

List of Tables

List of Figures

1 Introduction

This document describes the interface to be employed in transferring the products of grating (HETG, the High Energy Transmission Grating, or LETG, the Low Energy Transmission Grating) obervations Standard Data Processing from the CXC Level 2 processing pipeline to the CXC Data Archive, according to the requirements stipulated in the "ASC Data System Requirements" (Applicable Document 2).

1.1 Purpose

TG (generically referring to HETG and LETG Transmission Grating instruments) Level 2 processing, described in Applicable Document 2, consists of processing TG Level 1.5 event files and ancillary products (which are described in Applicable Document 7 into a binned-spectral FITS file. This document describes the structure and content of the resulting files. In addition, the Level 2 pipeline produces ancillary calibration products (effective area and redistribution files), measurements (feature detection and characterization), and summary information products.

1.2 Scope

This interface shall apply to all TG-specific data products that are generated by CXC Level 2 pipelines and distributed to the CXC Data Archive (see Applicable Document 2 and the "ASC Data System Software Design," Applicable Document 3) during the course of the AXAF mission.

1.3 Applicable Documents

The Applicable Documents required for background and detail on grating Level 2 products are as follows:

1. ASC Coordinates: Rev 5.0
   `http://hea-www.harvard.edu/asclocal/sds/docs/jcmdocs/ps/SDS02.ps`
2. ASC AMO-2400 (SE03)
   ASC Data System Requirements (ASC.302.93.0008)
3. ASC AMO-2401 (DS01)
   ASC Data System Software Design (ASC.500.93.0006)
4. ASC FITS File Designers' Guide (ASC-FITS-2.0.0)
   `http://hea-www.harvard.edu/~arots/asc/fits/ascfits.ps`
5. Level 1 to CXC Archive Interface Control Document (Rev. 1.6)
   `http://hea-www.harvard.edu/asclocal/icd/HRC_Level1/hrcl1icd_16.ps`
6. ACIS Data Products:
   Level 1 to ASC Archive Interface Control Document (Rev. 1.9)
   `http://hea-www.harvard.edu/asclocal/icd/ACIS_Level1/acis_l1_1.9.ps`
7. Grating Data Products:
   Level 1.5 to CXC Archive Interface Control Document
   `http://space.mit.edu/ASC/docs/ICD_L1.5.ps.gz`

8. Guide to AXAF Data Products (content.txt; v 2.1), section "FITS File Naming Convention for the AXAF Archive"
`http://hea-www.harvard.edu/~arots/asc/fits/content.txt`

9. FITS REGION Binary Table Design (ASC-FITS-REGION-1.0)
`http://hea-www.harvard.edu/~arots/asc/fits/region.ps`

10. The OGIP Spectral File Format (OGIP/92-007)
`http://legacy.gsfc.nasa.gov/docs/heasarc/ofwg/docs/spectra/ogip_92_007/`

11. Exposure Map Data Products to Archive ICD
`http://hea-www.harvard.edu/asclocal/icd/EXPMAP_PRODUCTS/Exp_DP_ICD.ps`

12. Analysis Reference Data for Exposure Maps and Grating L1.5-2 Pipelines ICD
`http://http://hea-www.harvard.edu/asclocal/icd/ARD/Rev1.2/ARD_ICD1.2.ps`

1.4 Data Content Summary

TG data sets generated by the Level 2 processing pipeline shall consist of data files conforming to the FITS format coforming to conventions described in the "ASC FITS File Designers' Guide" (Applicable Document 4, and references therein). These files contain header keyword entries and binary table (BINTABLE) extensions. Following rules outlined in Applicable Document 4, all these files will contain a null primary header followed by a main binary table (the "principal HDU") and auxiliary extensions ("auxiliary HDU").

Any other types of files will either be of types in common use (e.g., PostScript), or fully described here.

1.5 Recipients and Utilization

The primary recipients of TG Level 2 data products, via distribution from the archive, are AXAF observers, who will utilize these data products for scientific data analysis. The CXC may also make use of specific Level 2 data products for instrument calibration, instrument and/or spacecraft monitoring and trends analysis, and validation and verification of the Level 0, Level 1, Level 1.5, and Level 2 software and of the data products themselves.

1.6 Pertinent Relationships with Other Interfaces

Changes to the definition of Level 1.5 data products, as described in Applicable Document 7 may affect the Level 2 data products described in the current document.

1.7 Assumptions and Constraints

1.7.1 Missing or Undefined Data

The IEEE value, "NaN" (Not A Number) is used for missing real values. Note that this is not explicitly encoded as a TNULL keyword value in the table headers; it is implicit in the FITS standard.

1.7.2 Time Dependent Data

For each TG science event run reported in the AXAF telemetry stream, Level 2 processing shall generate a set of product files as described in Section 1.9. The products depend upon possibly time-dependent calibration data for coordinate transformations and region definitions.

1.7.3 Observation Intervals

The natural subdivision of TG processing prior to Level 2 is the Observation Interval (ObI). Each ObI may span several TG science runs, which are the atomic unit of the scientific instrument's telemetry. TG data prior to Level 2 will be processed by ObI. The Level 2 pipeline merges ObIs. Each ObI (or other time-interval) can in principal produce its own Level 2 products, but the pipeline will process merged ObIs.

1.7.4 Un-specified FITS Bintable Columns

Any particular file *may* contain arbitrary additional columns, added at other stages of processing, or by a User. Though they are unspecified, they are permitted. Level 2 software (analysis tools) should ignore them or simply pass them through to output, as directed by Level 2 tool specifications. Their presence should, however, *not* invalidate the product as TG Level 2.

1.7.5 HDU Identifier Keywords

Unless explicitly noted, the `HDUNAME` and `EXTNAME` header keywords will be identical.

1.8 Products Not Covered

TG Level 2 products that are used for maintenance and diagnostic purposes (those that are not supplied to the user for scientific data analysis), or which are generic AXAF Level 2 products, are not currently included within the interface defined by this document.

1.9 Labeling and Identification

The data files generated by the Level 2 processing pipeline shall be assigned external names as defined by the convention specified in "FITS File Names for the AXAF Archive", Applicable Document 8). The relevant default values for instrument field (`<i>`) are `aciss` or `hrcs`; for data source (`<s>`) is "f" for flight (or possibly "s" for simulation). The appropriate designation for the processing level and sublevel is "2". Level 2 products are listed in the following table:

`*_evt2.fits`	Merged events	required
`*_pha2.fits`	Binned spectra, Level 2, contents are "Type II 'PHA' or "Type I 'PHA' "	required

`*_peaks2.fits`
`*_fit2.fits`
`*_ident2.fits`
	Spectral measurements.	Required. Strong features in spectrum.

(The "*" designates an appropriate label sub-string as defined by the CXC FITS file-naming conventions.)

2 Overview of Files

The "ASC FITS Designers' Guide" (Applicable Document 4) defines and lists general header components for the primary header and for all binary table extensions. Since grating data are always obtained with either the HRC or ACIS detectors, file content is also described in detail by the HRC and ACIS Level 1 Interface Control Documents (Applicable Documents 5 and 6). Level 1.5-added content is described in the Level 1.5 ICD (Applicable Document 7).

Events: Level 2 event files are filtered concatenations of each Level 1.5 file. GTI extensions are merged, region extensions are concatenated.

To be updated, Rev 1.5, 9/99.

Spectrum: The basic product is a binned spectrum. This product is based on the OGIP Type II PHA Spectrum (see "The OGIP Spectral File Format", Applicable Document 10. New columns have been added to describe the source, spectrum part, and order. Column names have been taken from the Level 1.5 definition. The ACIS intrinsic energy region (PHA, or PI) is specified by calibration tables (named in the header) and by processing history (parameters).

Spectral Response: The spectral response is the effective area averaged over the obsevation. This accounts for variations in efficiency with position on the detector, including source position and position variation with time (dither) and detector gaps, as well as spectral order. It depends on the spatial/spectral extraction region (in particular the cross-dispersion aperture and the PI-order-sorting region, if ACIS), which are accounted for in the Level 2 exposure map calculation. This is traditionally referred to as an Auxiliary Response File (ARF), which is equivalent to the exposure map divided by time ([cm^2] vs energy). It is a source-specific function for CXO grating observations. Since a finer grid than the counts spectrum is generally desired, it will be provided as a separate file, instead of including the information as another column in the spectrum file.

The spectral response, or ARF, is defined in detail in the Exposure Map Data Products ICD (Document 11).

Spectral Redistribution: The redistribution describes the energy-dependent convolution kernel of the line-spread-function. It depends on the source position via the off-axis angle. Hence, it is in general different for every source, but varies slowly near the optical axis. It is provided as traditional OGIP Response Matrix Files (RMF), each of which may includes multiple orders.

Region: The binned spectra are selected from a spatial/spectral and CCD-Energy (scaled PHA) region of the Level 1.5 file. The region is given in grating diffraction coordinates (one coordinate in the dispersion direction, and one in the cross-dispersion direction; see Applicable Documents 1 and 7). The spatial/spectral extraction region mask is by default encoded as a region extension (see "FITS REGION Binary Table Design (ASC-FITS-REGION-1.0)", Applicable Document 9) in the Level 2 FITS file. The energy region is not encoded explicitly, but is implicit in calibration files and processing history.

Measurements: The binned spectra have features such as emission lines, absorption lines, and edges. Standard measurement products derived from these spectra will be collected into FITS bintables. Those spectral measurements which can be done without significant scientific judgement (such as strong emission line detection and characterization) are done automatically by the Level 2 pipeline. The specific quantities will evolve with experience, and are tentatively:

- line wavelength and uncertainty
- line flux and uncertainty
- line width and uncertainty
- line ID
- continuum spectrum and uncertainty

Other measurements can only be done via interactive analysis.

TBD by USG: summary products

Summary: At the end of Level 2 processing, summary products are produced for quick-look inspection and assessment by the end-user, and also by CXC validation and verification processing. Products are tentatively:

- Postscript plots and tables of spectra, images, results (e.g., field image, spectral regions, PHA-position plots, binned spectra, total counts vs source, counts vs order vs source, fluxed spectra)
- HTML interface to summary products
- GIF versions of images and plots

Products may exist both as Postscript files which can be viewed or directly printed, and/or hypertext files for easy interactive viewing.

Summary products are not included further in this document.

3 Event Files (*_evt2.fits)

Data Product Summary

ASC Data Product	Events
Instrument(s)	HETG, LETG
Level	2
Scientist/SDS	D. Huenemoerder
Filetype	Events
Created by tool	TBD
Used by tool(s)	various
Sample file	TBS

To be completed for next revision, 1.5, Aug99.

- HDU table
- header merging rules

4 Spectral Histogram Files (*_pha2.fits)

Data Product Summary

ASC Data Product	Binned Spectrum
Instrument(s)	HETG, LETG
Level	2
Scientist/SDS	D. Huenemoerder
Filetype	PHA Type II, or PHA Type I
Created by tool	tgextract
Used by tool(s)	various
Sample file	TBS

CXO binned grating spectrum files are stored as "PHA" files. Here, "PHA" does not refer to the content, but to the form of the file, as defined by OGIP conventions, originally conceived for spectral files from a Pulse Height Analyzer (which produced signals pulse-height bins with Pulse Height Amplitudes).

During Level 1.5 processing, event coordinates undergo several transformations. Given a zero-order centroid in sky coordinates, a region angle, and region width, events within each region are given several new coordinates. Proper interpretation of the spectrum requires consistent selection on several coordinates. These coordinates are carried over into the Level 2 spectrum file as attributes of the spectrum.

TG_SRCID is the source number, as output from the detection algorithm. A *TG_SRCID* of 0 means background — the photon has not been associated with a detected source. It is permitted to store binned spectra from multiple sources in one Type II file. However, for archival use it is advantageous to split files by source when this can be done uniquely. Hence, pipeline binned spectral products will be per source.[2]

TG_PART identifies the region of the spectrum of which the photon is a part. This can have only specific values as follows:

- **0:** zero order;
- **1:** HEG part of the spectrum;
- **2:** MEG part;
- **3:** LETG photon;
- **4–7:** LETG/HESF parts of the spectrum;
- **99:** background.

TG_M: The diffraction order, if resolved. This is a small integer, not likely to be of absolute value greater than 10 for HETG/ACIS-S. The value, 99, is used as a flag for "unresolved" or background photons.

For LETGS photons (detector is HRC, *TG_PART* is 3-7), *TG_M* will be either $+1$ or -1 (since HRC cannot resolve the orders via a pulse-height value).

[2] It must be kept in mind that if there are multiple sources in HETG observations the orders from different sources cross and photons in those regions may not be unambiguously resolved. If the region is scientifically important, the spectra must be analyzed jointly starting with the Level 1.5 event-list.

The physically meaningful geometric limit set by the ACIS-S array with HETG is about ± 62. This is for a 10 keV photon observed with the aim-point at one end of the ACIS-S array, and the detected photon at the other end. It is not a likely scenario, but is used to define allowed physically conceivable data ranges.

4.1 HDU Components

The following table describes the file structure by Header-Data Unit number, type, extension name, content, and HDU classes. An asterisk (*) denotes the principal HDU.

HDU	HDU Type	EXTNAME	CONTENT	HDUCLASS	Description
0	PRIMARY	N/A	N/A	N/A	NULL
1 (*)	BINTABLE	SPECTRUM	SPECTRUM	OGIP SPECTRUM TOTAL COUNT	Binned spectra, vs order vs source, grid, and ancillary information. A Type II file is the default; Type I is supported.
2	BINTABLE	REGION	TGMASK2	ASC REGION TG	Spectral-spatial region over which the spectrum was binned.
3	BINTABLE	GTI	GTI	OGIP STANDARD	Spectral-spatial region over which the spectrum was binned.
4	BINTABLE	SPECTRUM	SPECTRUM	OGIP SPECTRUM TOTAL COUNT	Zero-order PHA spectrum. (TBD 08/99)
5	BINTABLE	REGION	ZOMASK	ASC REGION	Spatial region over which the zero-order PHA spectrum was binned. (TBD 08/99)

Zero-order details to be supplied and reviewed for next revision, 1.5, Aug99.

The structure of the HDUs, as specified by the ASC FITS Guide (Document 4), is

- Null Primary HDU:

  image mandatory
  null configuration control
  short timing
  short observation

- Principal table HDU:

  bintable mandatory
  table coordinates
  full configuration control
  full observation
  full timing

- Auxiliary table HDU:

bintable mandatory
table coordinates
short configuration control
short timing
short observation

4.1.1 Naming Type I and II PHA Files

Type II PHA files are named according to the CXC FITS guidelines (Document 8). For Type I files, the optional discriminator field ("`<f>_`") is used to indicate the source identifier, spectrum part, and order. For example, for source 2, MEG part, order plus 3, the substring would be "`02Mp03_`". An example filename for ACIS-S, flight, starting time of 123456789, version 1, source 2, HEG minus first order, level 2 Type I PHA file would be: `acissf123456789N001_02Hm01_pha2.fits`.
The correscponding Type II (which holds all sources, orders, and parts) file-name is: `acissf123456789N001_pha2.fits` .
(Note that the "2" in "`pha2`" is for "Level 2", and *not* Type II.)

4.2 Column Specifications: Type II PHA

The contents of the FITS columns of the binned spectrum extension are given in the following tables These generally conform to the "OGIP Type II PHA" definition (See Document 10), with some additions appropriate for AXAF grating spectra. Some columns are optional, but are listed for complete compatibility with the OGIP standard; those that probably won't be used are in a separate table.

Columns which are constant may be promoted to header keywords. Some are mutually exclusive; for example, if a `BACKGROUND` array column is present, then `BACKFILE` does not need to be specified. These various conditions will be explained below. The primary point of following the OGIP Type II specification is to maintain compatibility with XSPEC.

The length of the vector columns (specified by `TFORM`*n*, such as for "COUNT" or "FLUX") is given in the table as 16384. This is a *representative maximum length*! It could in principal be longer or shorter. For LETGS, for example, binning at a single resolution element per bin will give about 6700 bins, and HETGS about 3400. Analysis tools can specify arbitrary binning. Pipelines will default to 0.5-1.0 resolution-element bins. Custom analysis may produce files only containing a few bins around lines of interest, or even arrays discontinuous in wavelength to hold several regions of interest.

The pipelines will bin the events into counts *vs.* wavelength, and fluxes into the units of photons $\mathrm{cm^{-2}\,s^{-1}\,bin^{-1}}$ (or photons $\mathrm{cm^{-2}\,s^{-1}\,Å^{-1}}$) *vs.* Å, and so these are the units and coordinates specified in the column descriptions. However, custom analysis may convert wavelengths to energy units in the Level 1.5 event list, and then bin into counts *vs.* energy. Likewise, analysis could produce flux in units of ergs instead of photons. These are all permitted in a spectrum file, as long as consistency is maintained within the grid columns (BINLO, BINHI) and units column.

TTYPE	TUNIT	TFORM	TLMIN	TLMAX	TDBIN	TNULL	Comment
SPEC_NUM		1I	1	*	N/A	0	Spectrum number (arbitrary running index).
ROWID		64A	N/A	N/A	N/A	NONE	Arbitrary string identifier
TG_M		1I	−62	62	N/A	99	Diffraction order (m)
TG_PART		1I	0	99	N/A	N/A	Spectral component (HEG, MEG, LEG, HESF parts)
TG_SRCID		1I	1	*	N/A	0	Source ID, output by detect.
ANCRFILE		64A	N/A	N/A	N/A	NONE	Ancillary response file name (ARF).
RESPFILE		64A	N/A	N/A	N/A	NONE	Grating Redistribution (LSF) file name (RMF).
BACKFILE		64A	N/A	N/A	N/A	NONE	Background file name.
CHANNEL		nI	1	n	1	0	Vector of spectral bin numbers.
COUNTS	count	nI	0	*	1	-1	Counts array (a spectrum).
STAT_ERR	count	nE	0.0	*	N/A	NaN	Statistical uncertainty ("error") on count column.
QUALITY		nI	0	5	N/A	N/A	Quality flag; 0 =>good, 1 =>bad, 2 =>dubious, 3,4/spare, 5 =>bad/user
GROUPING		nI	-1	1	N/A	255	Grouping vector; +1 => channel is new bin, −1 =>continuing bin, 0 =>undefined.
BACKGROUND	count	nI	0	*	1	-1	Background count vector.
BIN_LO	angstrom	nD	0.0	400.0	1.0E-04	NaN	Bin boundary, "left" edge.
	degrees	nD	-2.0	2.0	8.0E-05	NaN	
	mm	nD	-300	300	0.012	NaN	
	keV	nD	0.0	14.	1.0E-03	NaN	
	eV	nD	0.0	14000.	1.0	NaN	
	pixel	nJ	1	65536	1	0	
BIN_HI	angstrom	nD	0.0	400.0	1.0E-04	NaN	Bin boundary, "right" edge
	degrees	nD	-2.0	2.0	8.0E-05	NaN	
	mm	nD	-300	300	0.012	NaN	
	keV	nD	0.0	14.	1.0E-03	NaN	
	eV	nD	0.0	14000.	1.0	NaN	
	pixel	nJ	1	65536	1	0	
FLUX	photon /s /cm**2 / [bin|Å]	nE	0.0	N/A	N/A	NaN	Flux; ergs or photons per bin
FLUX_ERR	(same as FLUX)	nE	0.0	N/A	N/A	NaN	Uncertainty on flux.

Some columns are maintained for back-compatibility with `XSPEC`, or other software which relies on the standard OGIP spectral file specification. The columns (or keywords) which will probably not be used are given in the following table.

TTYPE	TUNIT	TFORM	TLMIN	TLMAX	TDBIN	TNULL	Comment
BACKSCAL		1E	0	N/A	N/A	*NaN*	Background scale factor.
CORRFILE		64A	N/A	N/A	N/A	none	Correction file name.
CORRSCAL		1E	0	N/A	N/A	*NaN*	Correction scale factor.
SYS_ERR	count	nE	0.0	TBD	N/A	*NaN*	Systematic uncertainty on count vector column.
AREASCL		1E	0.0	N/A	N/A	*NaN*	Area scaling factor, if ANCRFILE needs a correction.

Table notes:

"N/A" means "Not Applicable". The "TNULL" value of "*NaN*" stands for the IEEE "Not A Number" (which is not actually encoded into the header, but is implicit in the FITS convention). These are the columns defined by the OGIP Type II specification, but which are unlikely to be needed for AXAF grating spectra.

* For items with no physical legal maximum, the TLMAX value is limited by the data-type to max(TFORM).

FLUX and FLUX_ERR columns are optional. They are only produced when a spectral response is applied to the counts.

Additional descriptions of table columns:

SPEC_NUM Required. Spectrum number.

ROWID Optional; Arbitrary string identifier

TG_M Required; diffraction order (m)

TG_PART Required; Spectral component (HEG, MEG, LEG, ...)

TG_SRCID Required; source ID. If the Level 1.5 file has been split by source, then this will be a constant for the Level 2 file, and it will become a header keyword (`TG_SRCID` $= n$).

ANCRFILE Required; Ancillary response file name (ARF). This is the output of the exposure map calculation. The file referred to will be an analog of this file in that each row here will have a corresponding row in the spectral response file.

RESPFILE Required. Redistribution file name (RMF). This is a redistribution from wavelength to position, and may contain one or more orders, and be valid for one or more sources.

BACKFILE Optional; background file. Mutually exclusive with the BACKGROUND column.

CHANNEL Required; Vector of spectral "channel" numbers.

COUNT Required; Counts array (a spectrum).

STAT_ERR Optional; statistical uncertainty ("error") on count column. If the keyword POISSERR exists and has the value `T`, then this column need not be present. (Note that this condition applies to all rows.)

QUALITY Optional; Quality flag. The interpretation is as follows:

0: Good;

1: bad – never use;

2: dubious, recommend ignoring;

5: user-specified as bad;

Other values: spare.

If all values are 0, then this should be promoted to a header keyword (`QUALITY=0`) and the column deleted (note: this must be true for *all* rows).

GROUPING Optional. Grouping vector. If the data are un-grouped, then a header keyword is set (`GROUPING = 0`) and the column is not present. Otherwise, the interpretation is:

+1: channel starts a new group;

−1: channel is part of a continuing group;

0: grouping is undefined.

(Note that if *any* row is grouped, then all rows must have the grouping vector present.)

BACKGROUND Optional; background count vector. Mutually exclusive with the BACKFILE column.

BINLO Required; Bin boundary, low coordinate. By default, the bin units are Ångstroms. Custom analysis may produce other units.

BINHI Required; Bin boundary, high coordinate. Default units are Ångstroms. Custom analysis may produce other units, but the units of BINHI and BINLO must be the same.

FLUX Optional; Flux; Default units are photons $\mathrm{cm^{-2}\,s^{-1}\,bin^{-1}}$ (or photons $\mathrm{cm^{-2}\,s^{-1}\,Å^{-1}}$), but custom analysis may specify other units. Flux is determined by dividing the count-rate by a suitable effective area.

FLUX_ERR Optional. Uncertainty on flux. Only present if the FLUX column is defined.

BACKSCAL Optional; background scale factor

CORRFILE Optional; "correction" file

CORRSCAL Optional; "Correction" scale factor

SYS_ERR Optional; Systematic uncertainty on count vector column. If present, these values may get added in quadrature to the statistical errors by processing software (such as XSPEC).

AREASCL This is an optional area-scaling factor. It is most likely a constant and equal to 1.0, and so should be a header keyword (if present at all). It is only for XSPEC compatibility.

4.3 Column Specifications: Type I PHA

The Type I PHA format is supported for storage of single-order spectra and for back-compatibility with existing software. The format follows straightforwardly from that of the Type II file: an array column of `TFORM =` n`D` becomes a column vector having n rows of `TFORM = 1D`. Other array columns in the row become column vectors. Other scalar columns become header keywords.

4.4 Keywords Describing Coordinate Systems

All columns either need no coordinate system, or are already in physical units. There are no `TCRPX`n`,` `TCRVL`n, etc., keywords required.

4.5 Keywords Describing Calibration Data

Resolution of HETGS (HETG + ACIS-S) photons into diffraction orders requires application of the calibration from wavelength to pulse-height (PI, for Pulse-Invariant signal), and information on the distribution of the response. To do this, a special table is read, which was in turn generated from the CCD response matrices (RMFs), for a given fractional enclosed energy. Subsequent processing will require knowledge of this filter, so keywords are stored in the event extension header, as defined by:

```
COMMENT Order-sorting lookup tables:
COMMENT   Tables of Energy_lo and Energy_hi vs E used for order-sorting,
COMMENT   constructed for a specific fractional enclosed probability
COMMENT Reference ICD for file definition is: ARD ICD v1.1.
COMMENT
ACISDE='acis_95_V0.0.fits'          /E_lo, E_hi vs energy table vs CCD
```

The format of the order-sorting tables given in the "Analysis Reference Data for Exposure Maps and Grating L1.5-2 Pipelines ICD" (Document 12).

4.6 Grating Spectral Region Definitions

During Level 1.5 processing, for each ObI, a regional mask is created for each observation, as defined by the zero-order sky coordinates output from source detection and width and angle parameters. The source regions are designed to be of generous dimensions, such that the L1.5 product is a superset of photons likely to be selected for analysis or Level 2 processing, thus avoiding re-running the L1.5 pipeline for alternate selection criteria. It is defined in sky (X, Y) coordinates (typically a right-ascension and declination tangent-plane system) and is used to filter the events for transformation to grating diffraction coordinates.

During Level 2 processing, additional filtering may be applied to sub-select within the Level 1.5 region, but more naturally in grating coordinates, in which the zero-order is at the origin.

The generic definition of an CXC FITS region is given in the "FITS REGION Binary Table Design" (Applicable Document 9). The region HDU will be a single `REGION` auxiliary extension, whose header is comprised of these components:

mandatory `REGION`-content header,
table coordinate system and ranges,
short configuration control,
short timing, and
short observation information.

The components which differ substantially from the event file's other extensions (events, GTI) are the mandatory `REGION` keywords, and the table coordinate system. A region specification is in general specific to the ObI.

The `REGION` definition for Level 1.5 spectral mask are comprised of parts described by a circle for zero-order and rotated boxes for diffracted orders. There are generally multiple components for each spectrum and possibly multiple sources per observation. Level 2 processing may apply additional filtering within the Level 1.5 region. A schematic of the HETG, LETG, and HESF regions in sky coordinates is shown in Figure 4.6.

The interpretation of the `COMPONENT` value is shown in Table 2. The interpretation of the coordinate vectors for grating region shapes, as given in the FITS region specification (Applicable Document 9) is:

`BOX`: First element of each X and Y vectors specify the center of a rectangle. The first element of `ROTANG` specifies counter-clockwise rotation of the rectangle with respect to X and Y axes. The center of the rotation is the center of the rectangle (given by X, Y). ROTANG is always zero for this specification.

Other elements (if present) of X, Y are undefined. The first two elements of R specify the full-widths of X and Y, respectively.

The diffraction coordinates must be tied to the region generic "X" and "Y" tokens so that the `SHAPE` and `R` columns can be properly interpreted. The mechanism for this is via "Data Model" keywords:

```
MFORMn  = 'TG_LAM, TG_D'       / wavelength vs cross-disp angle
MTYPEn  = 'wav-pos '           / Wavelength-vs-position coords
```

where n is the $n^{\rm th}$ coordinate pair (and here would be 1, since there are no others). These keywords specify that `TG_LAM` is to be treated as X and `TG_D` is to be treated as Y, according to the region file specification. There are other coordinates which could be used for the dispersion, such as `TG_R` and `TG_MLAM`. As an example with the former, the header coordinate keywords would become:

`MTYPE` TBD.

```
MFORMn  = 'TG_R, TG_D'       / wavelength vs cross-disp angle
MTYPEn  = 'pos '           / Wavelength-vs-position coords
```

4.7 Volume, Size, and Frequency Estimates

TBS.

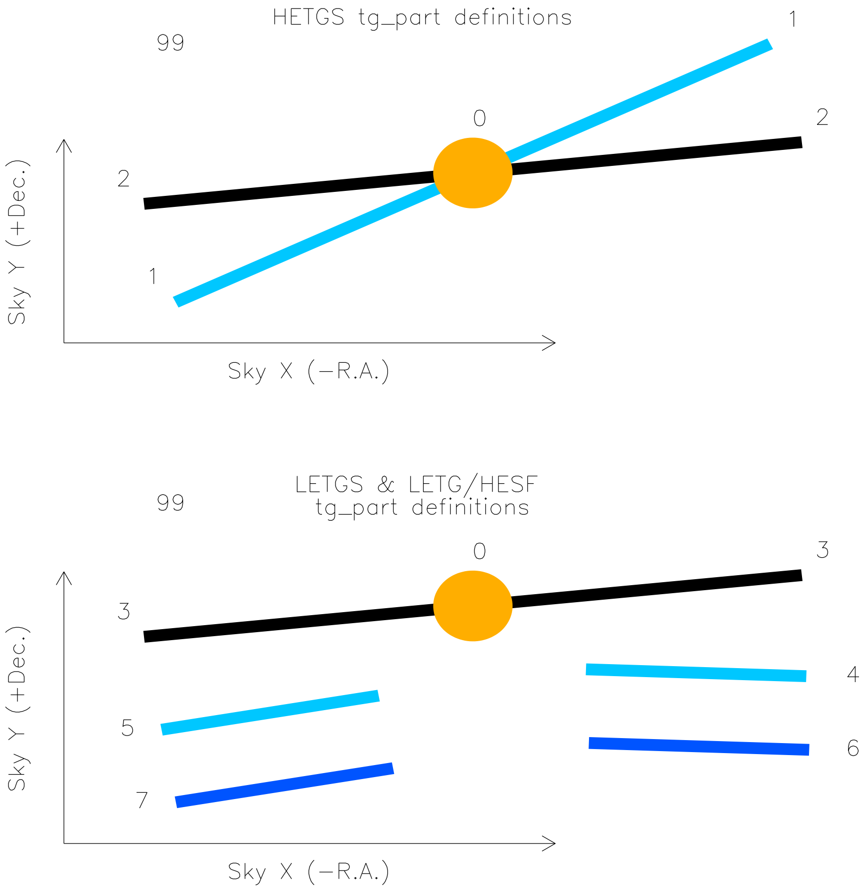

Figure 1: *This is a schematic representation of the grating region showing the different parts of the spectrum. On the top is the HETG specification, and on the bottom, LETG. For the LETG, the parts, 4–7 are optional, only being required if the HESF is in the beam. Scales are arbitrary. There has been a small rotation applied, to emphasize the fact that in sky coordinates, the spectrum can have any orientation. In Level 2 processing, region parts are obtained from a Level 1.5 event file column, and carried into the Level 2 product. Within each part, Level 1.5 processing assigns dispersion and cross-dispersion coordinates.*

TTYPE	TUNIT	TFORM	TLMIN	TLMAX	Comment
SPEC_NUM	-	1I	1	*	Spectrum number, which points the the row in the SPECTRUM extension to which the region on this row applies.
ROWID	-	64A	N/A	N/A	Source or a background region? (SOURCE — BACKGROUND_UP — BACKGROUND_DOWN)
SHAPE	-	16A	-	-	Shape of region: must be `BOX` for grating Level 2.
TG_LAM	angstrom	1E	0	400.	Dispersion coordinate vector for `SHAPE`.
TG_D	deg	1E	-2.	2.	Cross-dispersion coordinate vector for `SHAPE`.
R	N/A	2E	0	400	Radius vector for `SHAPE`. The interpretation of coordinates and units depends upon the `SHAPE`.
ROTANG	deg	1E	-360	360	Rotation angle for SHAPE, in degrees (usually 0.0 for grating Level 2).
COMPONENT	-	I	0	7	Component number that `SHAPE` belongs to (default is 1). Equivalent to `TG_PART`.
INCLUDE	-	I	0	1	Inclusion (1; default) or exclusion (0).
TG_SRCID	-	I	1	*	Source identification number.
TG_M	-	I	-62	62	Diffraction order.

Table 1:
FITS grating region file binary table contents. A TLMAX *of "*" means that the maximum is only limited by the data-type:* TLMAX = max(TFORM).

GRATING	COMPONENT	Meaning
HETG	0	Zero-order (optional)
	1	HEG arm
	2	MEG arm
LETG	0	Zero-order (optional)
	3	LEG arm
	4	HESF (Drake Flat) region 1 (optional)
	5	HESF (Drake Flat) region 2 (optional)
	6	HESF (Drake Flat) region 3 (optional)
	7	HESF (Drake Flat) region 4 (optional)

Table 2: *Region table component interpretation. Grating is either HETG or LETG. By default, all diffracted components present are binned.*

5 Spectral Measurements Files

The quick-look tools are intended to provide the user with a quick introduction to potentially complex, high-resolution X-ray spectra (see section ??). They perform automated detection of strong spectral features, automated fitting of profile functions to detected features, and automated search of a line database to provide candidate identifications for the detected features. Note that the output from these tools is *not suitable for precise quantitative comparison to physical models*, nor was it intended to be. Quantitative analysis of high resolution X-ray spectra requires human guidance (see section ??) and is intended to be done using interactive software such as the CXC Data Visualizer.

5.1 Detailed Interface Specifications

The primary input file is a binned spectrum (histogram) file based on the OGIP Type II PHA spectrum file as described in section ??. The input may also consist of one or more spectra stored in OGIP Type I PHA files. Bin edge coordinates in the input file (Type I/II PHA files) can be given in Angstroms, mm, pix, deg, keV or eV. Bin values may be in counts, photons s^{-1} cm^{-2} angstrom^{-1} or ergs s^{-1} cm^{-2} angstrom^{-1}. The primary reference data is the spectral line database as described in the Spectroscopy Database Interface Control Document. The grating line spread function may also be referenced in a future release; at this writing, it is not used. The output file formats are defined in detail in the following sections.

In the following format definitions, the grating-related keywords TG_SRCID, TG_PART, TG_M, etc. are not listed explicitly, but it is assumed that these essential keywords will be carried forward from the input files to the output files for all three tools.

5.1.1 Labeling and Identification

For custom processing, the output files generated by the TG Level 2 quick-look tools shall be assigned external names at the discretion of the user.

For pipeline application, the output file names are

Tool	File Name
`specdetect`	*_peaks2.fits
`specmeasure`	*_fit2.fits
`specidentify`	*_ident2.fits

Table 3: *File names for pipeline application of spectral scanning tools.*

5.1.2 Substructure Definition and Format

The "CXC FITS Designers' Guide" (Applicable Document 4) defines and lists header components for the primary header and for all binary table extensions. Since grating data are always obtained with either the HRC or ACIS detectors, file content is also described in detail by the HRC and ACIS Level 1 Interface Control Documents (Applicable Document 6).

5.1.3 Output File Types

Because input binned spectrum files may contain multiple spectra (histograms), the output format for these tools provides support for the processing of multiple spectra. The following types of output are produced:

Detect: contains a list of detected emission lines generated by `specdetect` along with an estimate of the global continuum shape in the input spectrum. The line list and estimated continuum histogram are placed in separate FITS extensions. The user may optionally specify the continuum shape by supplying a continuum histogram on input.

Fit: contains a list of best-fit parameters generated by `specmeasure`. These parameters usually represent the line profile and the associated local continuum shape.

Identification: contains line identifications generated by `specidentify`. For each observed line, the tool extracts a list of lines from the spectral line database which 1) lie within some user-defined tolerance of the rest wavelength of the detected peak and which are 2) predicted to be relatively strong at the user-specified collisional ionization equilibrium temperature.

At this writing, the quick-look tools process only emission lines and continuum. Nevertheless, where possible, we often use the term "spectral feature" rather than "emission line" to allow for future upgrades which may include processing of absorption lines, absorption edges, recombination features and blends. In many cases, it is not convenient to retain this level of generality in the discussion of file format details; in those cases we specialize our terminology to that of emission lines. We have tried to make the discussion as clear as possible and to apply appropriate levels of generality. Nevertheless, please be aware of this conflict in terminology and try to interpret this document appropriately.

5.2 Detect:

Data Product Summary

ASC Data Product	Detected Features (*_peaks2.fits)
Instrument(s)	HETG, LETG
Level	2
Scientist/SDS	J. Houck
Filetype	FITS bintable
Created by tool	`specdetect`
Used by tool(s)	various
Sample file	TBS

`specdetect` scans each input spectrum and generates a list of emission lines and an estimate of the global continuum shape (unless the continuum shape is user-supplied). For details on the input parameters and the algorithms used, see the `specdetect` manual page. The detect output for a single spectrum may occupy one or two FITS extensions. One extension contains the list of detected emission lines and the second extension contains the global continuum used by the detect algorithm. A single file may contain output for several input spectra.

5.2.1 Detect HDU Components

The following table (Table 4) describes the file structure by Header-Data Unit number, type, extension name, content, and HDU classes. An asterisk (*) denotes the principal HDU.

HDU	Type	EXTNAME	CONTENT	HDUCLASS	Description
1	NULL	N/A	N/A	N/A	
2 (*)	BINTABLE	SLINELIST	SLINELIST	ASC SLINELIST	List of emission lines detected in the input spectra.
3	BINTABLE	SCONTINUUM	SCONTINUUM	ASC SCONTINUUM	Continuum functions used in line detection.

Table 4: *The Header-Data Unit structure of the* `specdetect` *output FITS file.*

When multiple extensions with the same EXTNAME are present in the same file, the EXTVER keyword will be used to distinguish extensions and the HDUNAME of each extension will have the integer value of EXTVER appended.

The structure of the HDUs, as specified by the CXC FITS Guide (Document 4), is

- Null Primary HDU:
  - image mandatory
  - null configuration control

HDU	Type	EXTNAME	EXTVER	HDUNAME	Description
1	NULL	N/A	N/A	N/A	
2	BINTABLE	SLINELIST	1	SLINELIST1	Line list for spectrum 1
3	BINTABLE	SCONTINUUM	1	SCONTINUUM1	Continuum for spectrum 1
4	BINTABLE	SLINELIST	2	SLINELIST2	Line list for spectrum 2
5	BINTABLE	SCONTINUUM	2	SCONTINUUM2	Continuum for spectrum 2

Table 5: *The FITS extension structure of the* `specdetect` *output FITS file.*

  - short timing
  - short observation

- Principal table HDU:
  - bintable mandatory
  - table coordinates
  - full configuration control
  - full observation
  - full timing

- Auxiliary table HDU:
  - bintable mandatory
  - table coordinates
  - full configuration control
  - full observation
  - full timing

5.2.2 Keywords Describing Processing History

The following keywords are used to record the parameter settings in effect when the `specdetect` output was generated. These settings determine the detection thresholds and continuum estimation algorithm. If the continuum is input by the user, the keyword CONTMETH will have the value 'USER'. In this case, the keywords WINDOWSZ, ENDMETHD, LOCUT and HICUT are not needed and may be omitted.

```
CONTFILE='NONE'                     /Input continuum histogram file
EMTHRESH=                       4.0 /Emission line detect threshold (sigma)
ABTHRESH=                       0.0 /Absorption line detect threshold (sigma)
CONTMETH='MEDIAN'                   /Continuum calculation method
WINDOWSZ=                        50 /Size of running window (# bins)
ENDMETHD='DEFAULT'                  /Continuum end segment treatment
LOCUT=                          0.2 /Low end of median filter
HICUT=                          0.8 /High end of median filter
```

Table 6: Keywords describing `specdetect` line table extension.

5.2.3 Line Table Column Specifications

The line table contains a list of detected emission lines, giving an estimated center position, amplitude and FWHM for each along with an estimate of the continuum level and continuum slope at the position of each line.

Physical units on the various quantities are carried over from the input file. Where possible, the physical units of the bin coordinates are generically referred to by XX and the physical coordinates of the bin values are referred to by YY. The first derivative with respect to the bin coordinate (e.g. the continuum slope) then has units generically referred to by YY/XX. Similarly, the TLMIN and TLMAX values are generically referred to by XXmin, XXmax, YYmin and YYmax. Allowed units and value ranges (TUNIT, TLMIN, TLMAX) for XX and YY are listed in Table 7.

TTYPE	TUNIT	TFORM	TLMIN	TLMAX	TDBIN	TNULL	Comment
INDEX		I	1	16384	1	0	Feature number
POSITION	XX	D	XXmin	XXmax	N/A	NaN	Feature position
FWHM	XX	D	0.0	N/A	N/A	NaN	Line full width at half maximum
AMPLITUDE	YY	E	YYmin	YYmax	N/A	NaN	Height above continuum
CONTINUUM	YY	E	YYmin	YYmax	N/A	NaN	Estimated continuum level
CONTSLOPE	YY/XX	E	N/A	N/A	N/A	NaN	Estimated continuum slope

Table 7: Contents of the `specdetect` line table extension. “N/A” means “Not Applicable”. “*NaN*” is the IEEE standard for “Not A Number”.

Additional descriptions of table columns:

INDEX This index value is arbitrarily assigned to each feature and is used to track the feature through subsequent processing.

POSITION Approximate feature position as determined by the detect algorithm.

FWHM Approximate full-width at half-maximum as determined by the detect algorithm.

AMPLITUDE Height of detected peak above the continuum.

CONTINUUM Approximate continuum level at POSITION as determined by the detect algorithm.

CONTSLOPE Approximate continuum slope at POSITION as determined by the detect algorithm.

5.2.4 Keywords Describing Processing History

Except for the CONTENT keyword, all history keywords used in the line table HDU should be repeated in the continuum table HDU. The following additional keywords may also appear in the continuum table HDU.

5.2.5 Continuum Table Column Specifications

The continuum table contains the continuum spectrum which was used by the detect algorithm along with a copy of the continuum subtracted spectrum. The histogram bins are the same as those in the input spectrum.

TTYPE	TUNIT	TFORM	TLMIN	TLMAX	TDBIN	TNULL	Comment
CHANNEL		I	1.0	16384	1	0	Bin index
BINLO	XX	D	XXmin	XXmax	N/A	NaN	Bin boundary, low coordinate
BINHI	XX	D	XXmin	XXmax	N/A	NaN	Bin boundary, high coordinate
CONTINUUM	YY	E	YYmin	YYmax	N/A	NaN	Continuum level
NET	YY	E	N/A	N/A	N/A	NaN	Continuum subtracted spectrum

Table 8: Contents of `specdetect` continuum table extension. “N/A” means “Not Applicable”. “*NaN*” is the IEEE standard for “Not A Number”.

Additional descriptions of table columns:

CHANNEL Vector of spectral “channel” numbers.

BINLO Bin boundary, low coordinate. By default, the bin units are Ångstroms. Custom analysis may produce other units.

BINHI Bin boundary, high coordinate. Default units are Ångstroms. Custom analysis may produce other units, but the units of BINHI and BINLO must be the same.

CONTINUUM Continuum value used by the detect algorithm, either estimated internally, or user-supplied. The units should be the same as the input spectrum.

NET The result of subtracting CONTINUUM from the input spectrum.

5.3 Fit:

Data Product Summary

ASC Data Product	Fit Parameters (*_fit2.fits)
Instrument(s)	HETG, LETG
Level	2
Scientist/SDS	J. Houck
Filetype	FITS bintable
Created by tool	`specmeasure`
Used by tool(s)	various
Sample file	TBS

`specmeasure` processes an input spectrum using the output from `specdetect` to generate a list of best fit parameters and associated confidence limits. A single set of parameters is listed for each detected feature.

5.3.1 Fit HDU Components

The following table (Table 9) describes the file structure by Header-Data Unit number, type, extension name, content, and HDU classes. An asterisk (*) denotes the principal HDU.

HDU	Type	EXTNAME	CONTENT	HDUCLASS	Description
1	NULL	N/A	N/A	N/A	
2 (*)	BINTABLE	SLINEFIT	SLINEFIT	ASC SLINEFIT	List of fit parameters for features detected in an input spectrum.

Table 9: *The Header-Data Unit structure of the* `specmeasure` *output FITS file.*

When multiple extensions with the same EXTNAME are present in the same file, the EXTVER keyword will be used to distinguish extensions and the HDUNAME of each extension will have the integer value of EXTVER appended.

HDU	Type	EXTNAME	EXTVER	HDUNAME	Description
1	NULL	N/A	N/A	N/A	
2	BINTABLE	SLINEFIT	1	SLINEFIT1	fit parameters spectrum 1
3	BINTABLE	SLINEFIT	2	SLINEFIT2	fit parameters spectrum 2

Table 10: *The FITS extension structure of the* `specmeasure` *output FITS file.*

The structure of the HDUs, as specified by the CXC FITS Guide (Document 4), is

- Null Primary HDU:
  - image mandatory

  - null configuration control
  - short timing
  - short observation
- Principal table HDU:
  - bintable mandatory
  - table coordinates
  - full configuration control
  - full observation
  - full timing

5.3.2 Keywords Describing Processing History

The `specmeasure` parameter file values used to generate the model fits are recorded in the following keywords.

```
NPEAKS  =                          37 /Number of peak fits in this table
FUNCNAME='LorentzInt'                 /Name of fit function
BUNITS  ='counts'                     /Physical units of fit function
FITREG  =                         2.0 /Fit region: line_ctr +/- fit_reg * fwhm_max
FITMETH ='simplex'                    /fit method
FITSTAT ='chi parent'                 /fit statistic
CONFLEV =                        68.3 /Confidence level
```

Table 11: Keywords describing the `specmeasure` parameter table extension.

5.3.3 Measure Table Column Specifications

Physical units on the various quantities are carried over from the input file. Where possible, the physical units of the bin coordinates are generically referred to by XX and the physical coordinates of the bin values are referred to by YY. The first derivative with respect to the bin coordinate (e.g. the continuum slope) then has units generically referred to by YY/XX. Similarly, the TLMIN and TLMAX values are generically referred to by XXmin, XXmax, YYmin and YYmax. Allowed units and value ranges (TUNIT, TLMIN, TLMAX) for XX and YY are listed in Table 7.

TTYPE	TUNIT	TFORM	TLMIN	TLMAX	TDBIN	TNULL	Comment
INDEX		I	1	16384	1	0	Observed peak index
CHISQU		E	0.0	N/A	N/A	NaN	Value of chi-squared fit statistic
NFITBINS		I	0	16384	N/A	0	Number of spectral bins used in profile fit
POS	XX	E	XXmin	XXmax	N/A	NaN	Best fit peak position
FHWM	XX	E	0.0	N/A	N/A	NaN	Best fit Lorentzian FWHM
SIGMA	XX	E	0.0	N/A	N/A	NaN	Best fit Gaussian sigma
NORM	YY	E	N/A	N/A	N/A	NaN	Best fit profile normalization
POLY0	YY	E	N/A	N/A	N/A	NaN	Constant term in polynomial continuum
POLY1	YY/XX	E	N/A	N/A	N/A	NaN	Slope of polynomial continuum
POS_LO	XX	E	XXmin	XXmax	N/A	NaN	Lower confidence limit on fit peak position
POS_HI	XX	E	XXmin	XXmax	N/A	NaN	Upper confidence limit on fit peak position
FHWM_LO	XX	E	0.0	N/A	N/A	NaN	Lower confidence limit on fit Lorentzian FWHM
FHWM_HI	XX	E	0.0	N/A	N/A	NaN	Upper confidence limit on fit Lorentzian FWHM
SIGMA_LO	XX	E	0.0	N/A	N/A	NaN	Lower confidence limit on fit Gaussian sigma
SIGMA_HI	XX	E	0.0	N/A	N/A	NaN	Upper confidence limit on fit Gaussian sigma
NORM_LO	YY	E	N/A	N/A	N/A	NaN	Lower confidence limit on fit normalization
NORM_HI	YY	E	N/A	N/A	N/A	NaN	Upper confidence limit on fit normalization
POLY0_LO	YY	E	N/A	N/A	N/A	NaN	Lower confidence limit on constant term in polynomial continuum

Table 12: Contents of `specmeasure` line table extension. "N/A" means "Not Applicable". "*NaN*" is the IEEE standard for "Not A Number".

TTYPE	TUNIT	TFORM	TLMIN	TLMAX	TDBIN	TNULL	Comment
POLY0_HI	YY	E	N/A	N/A	N/A	NaN	Upper confidence limit on constant term in polynomial continuum
POLY1_LO	YY/XX	E	N/A	N/A	N/A	NaN	Lower confidence limit on slope term in polynomial continuum
POLY1_HI	YY/XX	E	N/A	N/A	N/A	NaN	Upper confidence limit on slope term in polynomial continuum

Table 13: (continued) Contents of `specmeasure` line table extension. "N/A" means "Not Applicable". "NaN" is the IEEE standard for "Not A Number".

Additional descriptions of table columns:

FWHM Optional: Lorentzian full-width at half maximum. Present only if the line profile fit includes a Lorentzian.

FWHM_LO Optional: Lower confidence limit on Lorentzian full-width at half maximum. Present only if the line profile fit includes a Lorentzian.

FWHM_HI Optional: Upper confidence limit on Lorentzian full-width at half maximum. Present only if the line profile fit includes a Lorentzian.

SIGMA Optional: Gaussian width parameter. Present only if the line profile fit includes a Gaussian.

SIGMA_LO Optional: Lower confidence limit on the Gaussian width parameter. Present only if the line profile fit includes a Gaussian.

SIGMA_HI Optional: Upper confidence limit on the Gaussian width parameter. Present only if the line profile fit includes a Gaussian.

POLY0 Optional: Constant term in polynomial continuum. Present only if the line profile fit includes a polynomial continuum.

POLY0_LO Optional: Lower confidence limit on constant term in polynomial continuum. Present only if the line profile fit includes a polynomial continuum.

POLY0_HI Optional: Upper confidence limit on constant term in polynomial continuum. Present only if the line profile fit includes a polynomial continuum.

POLY1 Optional: Slope of polynomial continuum. Present only if the line profile fit includes a polynomial continuum.

POLY1_LO Optional: Lower confidence limit on slope of polynomial continuum. Present only if the line profile fit includes a polynomial continuum.

POLY1_HI Optional: Upper confidence limit on slope of polynomial continuum. Present only if the line profile fit includes a polynomial continuum.

5.4 Identification:

Data Product Summary

ASC Data Product	Feature (*_ident2.fits)	Identifications
Instrument(s)	HETG, LETG	
Level	2	
Scientist/SDS	J. Houck	
Filetype	FITS bintable	
Created by tool	`specidentify`	
Used by tool(s)	various	
Sample file	TBS	

Given a list of detected spectral features from `specdetect` with positions and amplitudes refined by `specmeasure`, `specidentify` matches detected features with features in the spectral database. Matching features may be filtered using a number of criteria which are described in the tool manual page. Several candidate identifications may be associated with each detected feature.

5.4.1 Identify HDU Components

The following table (Table 14) describes the file structure by Header-Data Unit number, type, extension name, content, and HDU classes. An asterisk (*) denotes the principal HDU.

HDU	Type	EXTNAME	CONTENT	HDUCLASS	Description
1	NULL	N/A	N/A	N/A	
2 (*)	BINTABLE	LINEID	LINEID	ASC LINEID SLINELIST	List of line identifications

Table 14: *The Header-Data Unit structure of the* `specidentify` *output FITS file.*

When multiple extensions with the same EXTNAME are present in the same file, the EXTVER keyword will be used to distinguish extensions and the HDUNAME of each extension will have the integer value of EXTVER appended.

HDU	Type	EXTNAME	EXTVER	HDUNAME	Description
1	NULL	N/A	N/A	N/A	
2	BINTABLE	LINEID	1	LINEID1	line identifications for spectrum 1
3	BINTABLE	LINEID	2	LINEID2	line identifications for spectrum 2

Table 15: *The FITS extension structure of the* `specidentify` *output FITS file.*

The structure of the HDUs, as specified by the CXC FITS Guide (Document 4), is

- Null Primary HDU:
  - image mandatory
  - null configuration control
  - short timing
  - short observation
- Principal table HDU:
  - bintable mandatory
  - table coordinates
  - full configuration control
  - full observation
  - full timing

5.4.2 Keywords Describing Processing History

The `specidentify` parameters used to control the line identification algorithm are recorded in the following keywords. These parameters are used mainly as filters to simplify the task of identification and to reduce the volume of data returned from the spectral database.

```
LINELIST='db_line.fits'                /Spectral line database
DOPPUNIT='redshift'                    /Source Doppler shift units
DOPPSHFT=                          0.6 /Source Doppler shift
LINETOL =                          0.0 /Fractional position error tolerance
TEMPUNIT='keV'                         /Source temperature unit
TEMPTURE=                          5.0 /Source temperature
DENSITY =                       1.0e-3 /Source ion number density [cm^-3]
FILTER  ='Mg, Si, S, ~Fe'              /Line list filter string
```

Table 16: Keywords describing `specidentify` line table extension.

5.4.3 Identification Table Column Specifications

For each observed feature, this table contains a list of items in the spectral database which fit the search criteria. Each item is listed by name (element, ion and transition specification) along with a predicted strength (if available).

The format below indicates that 100 lines were identified. This number is intended to be a representative value only; the actual number could be larger or smaller and should be determined by the tool at run-time. For each list of matching lines, a *representative maximum* list length of 100 matches per feature is given in the table. We anticipate that most reasonable searches will return far fewer candidate identifications! However, alternate uses of this format could make use of lists which are longer or shorter.

TTYPE	TUNIT	TFORM	TLMIN	TLMAX	TDBIN	TNULL	Comment
INDEX		I	1	16384	1	0	Observed peak index
POSOBS	angstrom	E	0.0	400.0	N/A	NaN	Observed peak position
POSREST	angstrom	E	0.0	400.0	N/A	NaN	Observed peak rest position
NMATCHES		I	0	16384	N/A	0	Number of matching lines in the database
MINDEX		100I	0	16384	N/A	0	db line index
MPOS	angstrom	100E	0.0	400.0	N/A	NaN	db line position
MPOS_ERR	angstrom	100E	0.0	400.0	N/A	NaN	uncertainty in db line position
MSTR	erg cm^3 sec^{-1}	100E	0.0	N/A	N/A	NaN	db line emissivity
MSTR_ERR	erg cm^3 sec^{-1}	100E	0.0	N/A	N/A	NaN	uncertainty in db line emissivity
MELEM		100A2	N/A	N/A	N/A	‘’	db element label
MIONCHG		100I	1	100	N/A	999	db ion charge
MLABEL		100A40	N/A	N/A	N/A	‘’	db line transition label
MLOWER_LEV		100I	1	N/A	N/A	0	index of lower energy level
MUPPER_LEV		100I	1	N/A	N/A	0	index of upper energy level

Table 17: Contents of `specidentify` line table extension. “N/A” means “Not Applicable”. “*NaN*” is the IEEE standard for “Not A Number”.

Additional descriptions of table columns:

MINDEX Optional: Integer index identifying a particular line in the line list used. Present only if the input line list provides unique integer indices for each line (e.g. the SPEX line list does, but the APED list does not).

MPOS The line wavelength as listed in the reference line list. This value should be compared with the rest wavelength of the observed feature. It is important to note that, because wavelengths in the reference line list may be subject to significant errors, the accuracy or inaccuracy of the matchup between particular observed and reference wavelengths should not be the sole factor in assigning a final line identification.

MPOS_ERR Optional: The estimated uncertainty in the line wavelength listed in the database if known (likely to be poorly understood).

MSTR The predicted intensity of an emission line assuming collisional ionization equilibrium in a solar abundance plasma at a particular temperature and density, using a particular ionization balance calculation. The details of the computed value depend on the reference line list used.

MSTR_ERR Optional: The estimated uncertainty in the predicted emission line intensity (if available). It is very likely that this uncertainty is itself poorly understood and probably does not consistently account for all the uncertainties in all the data that went into the calculation.

MLOWER_LEV Optional: The integer index of the lower energy level involved in producing this emission line (if available). Note that the existence of such an integer index implies the existence of a corresponding energy level table (kept in the spectroscopy database). If the energy level table is changed, all the associated line list data may change as well.

MUPPER_LEV Optional: The integer index of the upper energy level involved in producing this emission line (if available). See also MLOWER_LEV.

5.5 Volume, Size, and Frequency Estimates

TBS

5.6 Example Files

TBS