Chandra Source Catalog
Requirements

Version 1.1
May 10, 2010

This document is maintained by Ian Evans (ievans@cfa.harvard.edu) on behalf of the *Chandra Source Catalog* project team.

NOTICE

The contents of this document are subject to change. Although this document accurately reflects the CXC's intentions for the first public release of the *Chandra Source Catalog*, that release may not comply with all of the requirements included herein, and compliance with some of the requirements listed herein may be deferred to subsequent catalog releases. The state of the operational implementation or the statistical scientific characterization of these requirements cannot be inferred from this document.

Table of Contents

Part I. Catalog Requirements

Part II. Database Requirements

Part III. Archive Requirements

Part IV. Non-Catalog Deliverables Requirements

Part V. Data Processing Requirements

I. Catalog Requirements

1. Catalog Contents

Each *Chandra Source Catalog* release consist of a predefined "catalog view" of the *Chandra Source Catalog* "database", together with associated archival (file-based) data. Since the principal requirements of interest to the user are those related to a *Chandra Source Catalog* release, the detailed requirements for database objects that are visible in a catalog view are presented in this part of this document.

1.1. Master Source Object

The Master Source Object records the "best-estimate" merged source properties for a source. There is one Master Source Object for each distinct astronomical source in the catalog.

1.1.1. Source Name

The source name shall consist of a character string of the form "**`CXO Jhhmmss.s+ddmmss`**", where **`hhmmss.s`** is the ICRS right ascension in hours, minutes, seconds, and tenths of a second, and **`+ddmmss`** is the signed ICRS declination in degrees, minutes, and seconds. The source name shall be updated in accordance with the recommendations of the *IAU Nomenclature Clearing-House*.

The source name field in the Master Source Object shall be identified by the field name "**`name`**".

1.1.2. Position and Position Errors

Performance Requirement:

The equatorial coordinates of each source shall be determined in the ICRS reference frame with an absolute uncertainty that is not to exceed 1.0 arcseconds (1 σ) for an isolated point source with at least 10 counts located within 3.0 arcminutes of the optical axis, 2.0 arcseconds (1 σ) for an isolated point source with at least 30 counts located within 10.0 arcminutes of the optical axis, and 5.0 arcseconds (1 σ) for an isolated point source with at least 50

counts located within 15.0 arcminutes of the optical axis.

1.1.2.1. Equatorial Coordinates

The best estimate of the ICRS equatorial coordinates for a source shall consist of 5 double precision values that record the ICRS right ascension in decimal degrees, the signed ICRS declination in decimal degrees, the major and minor radii of the 95% confidence level error ellipse of the source position in arcseconds, and the position angle of the major axis of the 95% confidence level error ellipse in degrees.

The right ascension and declination fields in the Master Source Object shall be identified by the field names "**`ra`**" and "**`dec`**", respectively.

The major and minor radii and the position angle of the major axis of the 95% confidence level error ellipse shall be identified by the field names "**`err_ellipse_r0`**", "**`err_ellipse_r1`**", and "**`err_ellipse_ang`**", respectively.

1.1.2.2. Galactic Coordinates

The Galactic coordinates corresponding to the source ICRS equatorial coordinates in section 1.1.2.1 (Equatorial Coordinates) shall consist of 2 double precision values that record the Galactic longitude, *l*, and latitude, *b*, for equinox J2000.0 and epoch J2000.0.

The Galactic latitude and Galactic longitude fields in the Master Source Object shall be identified by the field names "**`gal_b`**" and "**`gal_l`**", respectively.

1.1.3. Source Flags

The source flags highlight significant issues identified during processing that could affect cataloged source properties, or that may be useful search criteria.

1.1.3.1. Extent Flag

The extent flag consists of a Boolean whose value shall be **`TRUE`** if the deconvolved source extent [see section 1.1.4 (Source Extent and Errors)]

is not consistent with a point source at the 90% confidence level in any science energy band with a flux significance [see section 1.2.2.3.2 (Flux Significance)] ≥ 3.0 in any contributing observation. Otherwise, the extent flag shall be **FALSE**.

The extent flag field in the Master Source Object shall be identified by the field name "**extent_flag**".

1.1.3.2. Confusion Flag

The confusion flag consists of a Boolean whose value shall be **FALSE** if the confusion codes [see section 1.2.2.4.2 (Confusion Code)] for all contributing observations do not indicate that multiple sources are identified within the respective source detection regions and do not indicate that the source detection regions overlap other source detection regions, *and* no observations of the source have been rejected from contributing during the merging process because their source ellipses overlap the source ellipses of another source in any observation. Otherwise, the confusion flag shall be **TRUE**.

The confusion flag field in the Master Source Object shall be identified by the field name "**conf_flag**".

1.1.3.3. Pile-up Flag

The pile-up flag consists of a Boolean whose value shall be **FALSE** if none of the values of the pile-up warning field [see section 1.2.2.4.3.1 (Pile-up Warning)] for all contributing ACIS observations and source detection energy bands exceed 0.2 counts per ACIS frame per 3×3 pixel (corresponding to a pile-up fraction ~10%). Otherwise, the pile-up flag shall be **TRUE**.

The pile-up flag field in the Master Source Object shall be identified by the field name "**pileup_flag**".

1.1.3.4. Variability Flag

The variability flag consists of a Boolean whose value shall be **FALSE** if variability is not detected either within any single observation, or between any pair of observations, in any science energy band. Otherwise, the

variability flag shall be **TRUE**.

The criteria for flagging variability within a single observation are identified in section 1.2.2.4.4 (Variability Code). The criteria for flagging variability between two observations are identified in section 1.1.7.1 (Intra-observation Source Variability).

The variability flag field in the Master Source Object shall be identified by the field name "**var_flag**".

1.1.3.5. Streak Source Flag

The streak source flag consists of a Boolean whose value shall be **TRUE** if the streak source flags [see section 1.2.2.4.7 (Streak Source Flag)] for all of the contributing observations are **TRUE** (indicating that all of the contributing observations are ACIS observations *and* that the mean chip coordinates of the source region for each contributing observation fall within a defined region enclosing an identified readout streak). Otherwise, the streak source flag shall be **FALSE**.

The streak source flag field in the Master Source Object shall be identified by the field name "**streak_src_flag**".

1.1.3.6. Saturated Source Flag

The saturated source flag consists of a Boolean whose value shall be **TRUE** if the saturated source flags [see section 1.2.2.4.8 (Saturated Source Flag)] for all of the contributing observations are **TRUE** (indicating that all of the contributing observations are ACIS observations that are likely severely piled-up). Otherwise, the saturated source flag shall be **FALSE**.

The saturated source flag field in the Master Source Object shall be identified by the field name "**sat_src_flag**".

1.1.3.7. Manual Source Region Inclusion Flag

The manual source region inclusion flag consists of a Boolean whose value shall be **TRUE** if the manual source region inclusion flags [see section 1.2.2.4.9 (Manual Source Region Inclusion Flag)] for all of the contributing observations are **TRUE**, indicating that the source region was

manually included in all of the contributing observations. Otherwise, the manual source region inclusion flag shall be **FALSE**.

The manual source region inclusion flag field in the Master Source Object shall be identified by the field name "**man_inc_flag**".

1.1.3.8. Manual Source Region Parameters Flag

The manual source region parameters flag consists of a Boolean whose value shall be **TRUE** if the manual source region parameters flags [see section 1.2.2.4.10 (Manual Source Region Parameters Flag)] for all of the contributing observations are **TRUE**, indicating that the source region parameters were manually modified in all of the contributing observations. Otherwise, the manual source region parameters flag shall be **FALSE**.

The manual source region parameters flag field in the Master Source Object shall be identified by the field name "**man_reg_flag**".

1.1.3.9. Manual Match Flag

The manual match flag consists of a Boolean whose value shall be **TRUE** if the set of source-by-observation objects that contribute to the current master source object region was manually selected or manually modified by human review. (This will typically occur if the source regions for the set of contributing source-by-observation objects cannot be automatically matched.) Otherwise, the manual match flag shall be **FALSE**.

The manual match flag field in the Master Source Object shall be identified by the field name "**man_match_flag**".

1.1.3.10. Inter-observation Variable Hardness Ratio Flag

The inter-observation variable hardness ratio flag consists of a Boolean whose value shall be **TRUE** if one or more of the hardness ratios computed for any source-by-observation object that contributes to the current master source object is inconsistent statistically with the corresponding hardness ratios [see section 1.2.2.7.2 (Hardness Ratios)] computed for all of the other source-by-observation objects that contribute to the current master source object. Otherwise, the inter-observation variable hardness ratio flag shall be **FALSE**.

The inter-observation variable hardness ratio flag field in the Master Source Object shall be identified by the field name "**`var_inter_hard_flag`**".

1.1.4. Source Extent and Errors

1.1.4.1. Deconvolved Source Extent

A parameterization of the best estimate of the deconvolved extent of each source shall be determined in each science energy band from a variance-weighted mean of the deconvolved extent of each source [see section 1.2.2.5.4 (Deconvolved Source Extent)] measured in all contributing observations. The parameterization shall consist of the best estimate values for the 1 σ radius along the major axis, the 1 σ radius along the minor axis, and position angle of the major axis (and associated errors) of a rotated elliptical Gaussian that is convolved with the ray-trace local PSF and derived from the source spatial event distribution.

Performance Requirement:

The parameterization of a source is required exclude a point source model at the 90% confidence level if the source has at least 30 counts, is located within 3.0 arcminutes of the optical axis, and has an apparent 1 σ radius along the major axis > 2.5 arcseconds, or if the source has at least 50 counts, is located within 10.0 arcminutes of the optical axis, and has an apparent 1 σ radius along the major axis > 15 arcseconds, or if the source has at least 100 counts, is located within 15.0 arcminutes of the optical axis, and has an apparent 1 σ radius along the major axis > 30 arcseconds.

If the source is detected in multiple contributing observations at different off-axis angles, then the off-axis angle of the source in the observation for that has the highest flux significance [see section 1.2.2.3.2 (Flux Significance)] shall be the defining off-axis angle for these performance requirements.

The best estimates of the parameter values for a source shall consist of 6 double precision values for each science energy band that record the 1 σ radii along the major and minor axes of the rotated elliptical Gaussian parameterization of the deconvolved source extent in arcseconds, the

position angle of the major axis of the rotated elliptical Gaussian in degrees, and the associated errors.

Note

The rotated elliptical Gaussian parameterization of the source extent will be approximated by a circularly symmetric Gaussian parameterization in the first catalog release. This limitation will be lifted in a future release.

The 1 σ radii along the major and minor axes, position angle of the major axis, and associated error fields in the Master Source Object shall be identified by the field names "**`major_axis`**", "**`minor_axis`**", "**`pos_angle`**", "**`major_axis_err`**", "**`minor_axis_err`**", and "**`pos_angle_err`**", respectively.

1.1.5. Source Fluxes

Source fluxes shall be determined by aperture photometry in each science energy band. Aperture photometry measurements shall be obtained in the source region, and in an elliptical aperture that includes the 90% ECF of the PSF at the source location, with corrections applied for the PSF aperture fraction. The correction factor shall be computed from the ray-trace local PSF under the assumption that the source spatial distribution matches the local PSF.

1.1.5.1. Energy Bands

1.1.5.1.1. ACIS Observations

The energy bands used for ACIS observations are split into two categories. Source detection energy bands are used only in the source detection processing to detect sources; they are chosen to maximize the source detection efficiency for different kinds of sources observed using the back- and front-side illuminated CCDs. Science energy bands are used for all science analyses and are chosen to maximize the scientific utility of the catalog. The process by which the energy bands were chosen is documented in Appendix A.

1.1.5.1.1.1. Source Detection Energy Bands

The source detection energy bands for ACIS are a single *broad* band, covering 0.5–7.0 keV, and *soft*, *medium*, and *hard* bands covering 0.5–1.2 keV, 1.2–2.0 keV, and 2.0–7.0 keV, respectively.

The broad, soft, medium, and hard source detection energy bands are designated "**`B`**", "**`S`**", "**`M`**", and "**`H`**", respectively.

1.1.5.1.1.2. Science Energy Bands

The science energy bands for ACIS are a single *broad* band, covering 0.5–7.0 keV, and *ultra-soft*, *soft*, *medium*, and *hard* bands covering 0.2–0.5 keV, 0.5–1.2 keV, 1.2–2.0 keV, and 2.0–7.0 keV, respectively.

The broad, ultra-soft, soft, medium, and hard science energy bands are designated "**`b`**", "**`u`**", "**`s`**", "**`m`**", and "**`h`**", respectively.

1.1.5.1.2. HRC Observations

The HRC does not have significant energy resolution. There is a single wide energy band for both source detection and science. The wide energy band for HRC includes pulse heights 0 : 254, which correspond approximately to photon energies 0.1–10.0 keV.

The wide energy band is designated "**`w`**" to distinguish it from the broad energy band for ACIS, which has different energy limits.

1.1.5.2. [Deleted]

1.1.5.3. Aperture Photometry Fluxes

1.1.5.3.1. Aperture Source Energy Fluxes

The aperture source energy fluxes and associated confidence limits shall consist of 12 double precision values for each science energy band that record the best estimates of the (background-subtracted) fluxes and two-sided confidence limits in the source region and in an elliptical aperture that includes the 90% ECF of the PSF at the source location, corrected by the PSF aperture fraction, livetime, and exposure, in units of photons s^{-1} cm^{-2} and ergs s^{-1} cm^{-2}. The conversion from photons s^{-1} cm^{-2} to ergs s^{-1}

[1] cm^{-2} is performed by summing the photon energies for each incident source photon and scaling by the local value of the ARF at the location of the incident photon.

Performance Requirement:

> The aperture source energy fluxes in each science energy band determined from the aperture photometry of an isolated point source shall differ from the actual source flux of a point source measured through the same aperture by < 35% at a 1 σ confidence level if the source has at least 30 counts in the energy band and is located within 3.0 arcminutes of the optical axis, or < 30% if the source has at least 50 counts in the energy band and is located within 10.0 arcminutes of the optical axis, or < 25% if the source has at least 100 counts in the energy band and is located within 15.0 arcminutes of the optical axis, under the assumption that the response matrix is diagonal or nearly so.
>
> If the source is detected in multiple contributing observations at different off-axis angles, then the off-axis angle of the source in the observation for that has the highest flux significance [see section 1.2.2.3.2 (Flux Significance)] shall be the defining off-axis angle for these performance requirements.

The aperture source energy fluxes in the source region, and in an elliptical aperture that includes the 90% ECF of the PSF at the source location, and their associated two-sided confidence limits, in the Master Source Object shall be identified by the field names "**photflux_aper**", "**photflux_aper90**", "**photflux_aper_hilim**", "**photflux_aper_lolim**", "**photflux_aper90_hilim**", and "**photflux_aper90_lolim**", respectively, for the photon fluxes, and "**flux_aper**", "**flux_aper90**", "**flux_aper_hilim**", "**flux_aper_lolim**", "**flux_aper90_hilim**", and "**flux_aper90_lolim**", respectively, for the energy fluxes.

1.1.5.3.2. Aperture Model Energy Fluxes

The aperture model energy fluxes and associated confidence limits shall consist of 12 double precision values for each science energy band that record the best estimates of the fluxes and two-sided confidence limits in

the source region, and in an elliptical aperture that includes the 90% ECF of the PSF at the source location, corrected by the PSF aperture fraction, livetime, and exposure, in units of ergs s^{-1} cm^{-2}. The conversion from aperture source count rates [see section 1.2.2.6.3.3 (Aperture Source Count Rates)] in each science energy band to aperture model energy fluxes is performed by scaling from a model spectrum folded through the calibrated response as follows:

> For a source model $\mathcal{F}(E)$ whose integral over the science band is $\mathcal{F}'(band)$, calculate the corresponding band count rate $\mathcal{C}'(band)$ in counts s^{-1} given then effective area calibration $\mathcal{A}(E)$ [and, if available the RMF, $\mathcal{R}(E, band)$] appropriate to the observation; this is $\int \mathcal{F}(E)\mathcal{A}(E)\mathcal{R}(E, band)$ over all energies, or if a diagonal RMF is assumed, $\int \mathcal{F}(E)\mathcal{A}(E)$ over the band. Infer the aperture model energy flux from the measured aperture source count rate $\mathcal{C}(band)$ as $\mathcal{F}(band) = \mathcal{F}'(band)\mathcal{C}(band)/\mathcal{C}'(band)$. The power-law spectral model shall have a fixed spectral index (γ, defined as $F_E \sim E^{-\gamma}$) = 1.7, and a fixed total neutral Hydrogen absorbing column $N_H = N_H(Gal)$ cm^{-2}. The black body spectral model shall have a fixed temperature kT = 1.0 keV, and a fixed total neutral Hydrogen absorbing column $N_H = 3 \times 10^{20}$ cm^{-2}.

Performance Requirement:

> The aperture model energy fluxes in each science energy band determined from the aperture photometry of an isolated point source shall agree with the actual source flux of a point source measured through the same aperture by < 35% at a 1 σ confidence level if the source has at least 30 counts in the energy band and is located within 3.0 arcminutes of the optical axis, or <30% if the source has at least 50 counts in the energy band and is located within 10.0 arcminutes of the optical axis, or <25% if the source has at least 100 counts in the energy band and is located within 15.0 arcminutes of the optical axis, under the assumption that the model spectrum is an accurate representation of the source spectrum.
>
> If the source is detected in multiple contributing observations at

different off-axis angles, then the off-axis angle of the source in the observation for that has the highest flux significance [see section 1.2.2.3.2 (Flux Significance)] shall be the defining off-axis angle for these performance requirements.

The power-law and black body aperture model energy fluxes in the source region, and in an elliptical aperture that includes the 90% ECF of the PSF at the source location, and their associated two-sided confidence limits, in the Master Source Object shall be identified by the field names "**flux_powlaw_aper**", "**flux_powlaw_aper90**", "**flux_bb_aper**", "**flux_bb_aper90**", "**flux_powlaw_aper_hilim**", "**flux_powlaw_aper_lolim**", "**flux_powlaw_aper90_hilim**", "**flux_powlaw_aper90_lolim**", "**flux_bb_aper_hilim**", "**flux_bb_aper_lolim**", "**flux_bb_aper90_hilim**", and "**flux_bb_aper90_lolim**", respectively.

1.1.5.4. Source Significance

The source significance shall consist of a double precision value that records the highest flux significance [see section 1.2.2.3.2 (Flux Significance)] in all contributing observations and science energy bands.

The source significance in the Master Source Object shall be identified by the field name "**significance**".

1.1.6. Spectral Properties

The spectral properties of a source observed by ACIS shall include a set of hardness ratios determined from best estimates of the aperture source energy fluxes [see section 1.1.5.3.1 (Aperture Source Energy Fluxes)] in the source region.

1.1.6.1. Hardness Ratios

Hardness ratios shall consist of 3 double precision values for each pair of science energy bands, excluding the broad and ultra-soft bands, that record the hardness ratios and associated two-sided confidence limits computed from the aperture total counts [see section 1.2.2.6.3.1.1 (Aperture Total Counts)] of the set of source-by-observation objects that contribute to the

current master source object. The hardness ratio for two bands with designations **<x>** and **<y>**, respectively, is defined equivalently as

hard_<x><y> = (*F*(<x>)-*F*(<y>)) / *F*(<broad>),

where ***F*(<x>)** is the aperture source photon flux **(photflux_aper)** in band **<x>**, and ***F*(<broad>)** is the aperture source photon flux in the broad energy band.

The hardness ratios and associated upper and lower confidence limits in the Master Source Object shall be identified by the field names "**hard_<x><y>**", "**hard_<x><y>_hilim**", and "**hard_<x><y>_lolim**", respectively, where **<x>** and **<y>** are the higher and lower energy band designations. For example, the hardness ratio and associated two-sided confidence limits for the hard-to-soft ACIS science energy bands shall be identified by the field names "**hard_hs**", "**hard_hs_hilim**", and "**hard_hs_lolim**", respectively.

1.1.6.2. Spectral Model Fits

If there are at least 150 net (background subtracted) counts in the energy range 0.5–7.0 keV present in the source region [see section 1.2.2.5.1 (Source Region)] in at least one contributing ACIS observation, then power law and thermal (black body) model spectra are fitted to PI event data extracted from the source region, with corrections for the PSF aperture fraction, livetime, and ARF applied when fitting the models.

1.1.6.2.1. Power Law Model Spectral Fit

The power law model spectral fit is performed over the energy range 0.5–7.0 keV as described in section 1.2.4.4 of part V (Spectral Fit). The free parameters to be fitted are the total integrated flux, total neutral Hydrogen absorbing column, and power law photon index.

1.1.6.2.1.1. Power Law Model Photon Index

The power law model photon index and associated confidence limits shall consist of 3 double precision values that record the best estimates of the best-fit power law photon index (α, defined as $F_E \sim E^{-\gamma\alpha}$) and associated two-sided confidence limits from the power law model spectral fits from

all contributing observations.

The power law model photon index and associated two-sided confidence limits in the Master Source Object shall be identified by the field names "**alpha**", "**alpha_hilim**", and "**alpha_lolim**", respectively.

1.1.6.2.1.2. Power Law Model Spectral Fit Flux

The power law model spectral fit flux and associated confidence limits shall consist of 3 double precision values that record the best estimates of the integrated 0.5–7 keV flux and associated two-sided confidence limits derived from the best fitting power law model spectral fits from all contributing observations in units of ergs s^{-1} cm^{-2}.

The power law model spectral fit flux and associated two-sided confidence limits in the Master Source Object shall be identified by the field names "**flux_powlaw**", "**flux_powlaw_hilim**", and "**flux_powlaw_lolim**", respectively.

1.1.6.2.1.3. Power Law Model Spectral Fit N_H

The power law model spectral fit N_H and associated confidence limits shall consist of 3 double precision values that record the best estimates of the best-fit total neutral Hydrogen absorbing column and associated two-sided confidence limits from the power law model spectral fits from all contributing observations in units of 10^{20} cm^{-2}.

The power law model spectral fit N_H and associated two-sided confidence limits in the Master Source Object shall be identified by the field names "**nh_powlaw**", "**nh_powlaw_hilim**", and "**nh_powlaw_lolim**", respectively.

1.1.6.2.1.4. [Deleted]

1.1.6.2.2. Thermal (Black Body) Model Spectral Fit

The thermal (black body) model spectral fit is performed over the energy range 0.5–7.0 keV as described in section 1.2.4.4 of part V (Spectral Fit). The free parameters to be fitted are the total integrated flux, total neutral Hydrogen absorbing column, and black body temperature.

1.1.6.2.2.1. Thermal (Black Body) Spectral Fit Temperature

The thermal (black body) spectral fit temperature and associated confidence limits shall consist of 3 double precision values that record the best estimates of the best-fit temperature (*kT*) in units of keV and associated two-sided confidence limits from the thermal (black body) model spectral fits from all contributing observations.

The thermal (black body) spectral fit temperature and associated two-sided confidence limits in the Master Source Object shall be identified by the field names "**kt**", "**kt_hilim**", and "**kt_lolim**", respectively.

1.1.6.2.2.2. Thermal (Black Body) Spectral Fit Flux

The thermal (black body) spectral fit flux and associated confidence limits shall consist of 3 double precision values that record the best estimates of the integrated 0.5–7 keV flux and associated two-sided confidence limits derived from the best fitting thermal (black body) model spectral fits from all contributing observations in units of ergs s^{-1} cm^{-2}.

The thermal (black body) spectral fit flux and associated two-sided confidence limits in the Master Source Object shall be identified by the field names "**flux_bb**", "**flux_bb_hilim**", and "**flux_bb_lolim**", respectively.

1.1.6.2.2.3. Thermal (Black Body) Spectral Fit N_H

The thermal (black body) spectral fit N_H and associated confidence limits shall consist of 3 double precision values that record the best estimates of the best-fit total neutral Hydrogen absorbing column and associated two-sided confidence limits from the thermal (black body) model spectral fits from all contributing observations in units of 10^{20} cm^{-2}.

The thermal (black body) spectral fit N_H and associated two-sided confidence limits in the Master Source Object shall be identified by the field names "**nh_bb**", "**nh_bb_hilim**", and "**nh_bb_lolim**", respectively.

1.1.6.2.2.4. [Deleted]

1.1.6.3. Galactic N_H

The Galactic N_H shall consist of a double precision value that records the Galactic neutral Hydrogen column density (N_H) in the direction of the source, determined from the interpolated data of Dickey & Lockman (1990; Ann. Rev. Astron. Astrophys., 28, 215) in units of 10^{20} cm^{-2}.

The Galactic N_H in the Master Source Object shall be identified by the field name "**nh_gal**".

1.1.7. Source Variability

1.1.7.1. Intra-observation Source Variability

Intra-observation source variability within any contributing observations shall be assessed according to the highest level of variability seen within any single contributing observation [see section 1.2.2.8 (Source Variability)].

1.1.7.1.1. Intra-observation Kolmogorov-Smirnov (K-S) Test

The intra-observation Kolmogorov-Smirnov (K-S) test probability shall consist of a double precision value for each science energy band that records the maximum value of the Kolmogorov-Smirnov (K-S) test probability, **ks_prob**, that is calculated for each of the contributing observations [see section 1.2.2.8.1 (Kolmogorov-Smirnov (K-S) Test)].

The intra-observation Kolmogorov-Smirnov (K-S) test probability in the Master Source Object shall be identified by the field name "**ks_intra_prob**".

1.1.7.1.2. Intra-observation Kuiper's Test

The intra-observation Kuiper's test probability shall consist of a double precision value for each science energy band that records the maximum value of the Kuiper's test probability, **kp_prob**, that is calculated for each of the contributing observations [see section 1.2.2.8.2 (Kuiper's Test)].

The intra-observation Kuiper's test probability in the Master Source Object shall be identified by the field name "**kp_intra_prob**".

1.1.7.1.3. Intra-observation Variability Probability

The intra-observation variability probability shall consist of a double precision value for each science energy band that records the maximum value of the variability probability, **var_prob**, that is calculated for each of the contributing observations [see section 1.2.2.8.3 (Gregory-Loredo Variability Probability)].

The intra-observation variability probability in the Master Source Object shall be identified by the field name "**var_intra_prob**".

1.1.7.1.4. Intra-observation Variability Index

The intra-observation variability index shall consist of an integer value that records the maximum value of the variability index, **var_index**, that is calculated for each of the contributing observations [see section 1.2.2.8.4 (Variability Index)]. A detailed description of the computation of the variability index, and its interpretation, is documented in Appendix F.

The intra-observation variability index in the Master Source Object shall be identified by the field name "**var_intra_index**".

1.1.7.1.5. Intra-observation Flux Variability

The intra-observation flux variability shall consist of a double precision value for each science energy band that records the maximum value of the flux variability that is calculated from each of the contributing observations by appropriately scaling the count rate variability, "**var_sigma**" [see section 1.2.2.8.5 (Count Rate Variability)], in units of photons s^{-1} cm^{-2}.

The intra-observation flux variability in the Master Source Object shall be identified by the field name "**var_intra_sigma**".

1.1.7.2. Inter-observation Source Variability

Inter-observation source variability between any contributing observations shall be assessed by application of a χ^2 hypothesis test applied to the source region photon fluxes observed in the contributing observations.

1.1.7.2.1. Inter-observation Variability Probability

The inter-observation variability probability shall consist of a double precision value for each science energy band that records the probability that the source region photon flux varied between the contributing observations based on the χ^2 distribution of the photon fluxes and the standard deviations of the individual observations.

The inter-observation variability probability in the Master Source Object shall be identified by the field name "**`var_inter_prob`**".

1.1.7.2.2. Inter-observation Variability Index

The inter-observation variability index shall consist of an integer value that records an index in the range [0, 10] that is derived from the inter-observation variability probability [see section 1.1.7.2.1 (Inter-observation Variability Probability)] to evaluate whether the source region photon flux is constant between the observations. The variability index recorded in the catalog is the maximum value of the variability indices computed for each science energy band.

The inter-observation variability index in the Master Source Object shall be identified by the field name "**`var_inter_index`**".

1.1.7.2.3. Inter-observation Flux Variability

The inter-observation flux variability shall consist of a double precision value for each science energy band that records the absolute value of the difference between the error weighted mean source region photon flux density of all of the contributing observations and the mean source region photon flux density of the single observation for which the absolute value of the difference, divided by the σ for that observation, is maximized, in units of photons s^{-1} cm^{-2}.

The inter-observation flux variability in the Master Source Object shall be identified by the field name "**var_inter_sigma**".

1.1.8. Observation Summary

Summary information about the number of observations and total exposure times for various instrument configurations are recorded.

1.1.8.1. ACIS Observations

The ACIS observation summary information shall consist of 3 integer values that record the total numbers of ACIS imaging, ACIS/HETG, and ACIS/LETG observations, and 3 double precision values that record the total ACIS imaging, ACIS/HETG, and ACIS/LETG exposure times (seconds good time), for all ACIS observations merged to the Master Source Object record for a source.

The total numbers of ACIS imaging, ACIS/HETG, and ACIS/LETG observations, and the total ACIS imaging, ACIS/HETG, and ACIS/LETG exposure times, in the Master Source Object shall be identified by the field names "**acis_num**", "**acis_hetg_num**", "**acis_letg_num**", "**acis_time**", "**acis_hetg_time**", and "**acis_letg_time**", respectively.

1.1.8.2. HRC Observations

The HRC observation summary information shall consist of 3 integer values that record the total numbers of HRC imaging, HRC/HETG, and HRC/LETG observations, and 3 double precision values that record the total HRC imaging, HRC/HETG, and HRC/LETG exposure times (seconds good time), for all HRC observations merged to the Master Source Object record for a source.

The total numbers of HRC imaging, HRC/HETG, and HRC/LETG observations, and the total HRC imaging, HRC/HETG, and HRC/LETG exposure times, in the Master Source Object shall be identified by the field names "**hrc_num**", "**hrc_hetg_num**", "**hrc_letg_num**", "**hrc_time**", "**hrc_hetg_time**", and "**hrc_letg_time**", respectively.

1.2. Source-by-Observation Objects

1.2.1. Observation Object

1.2.1.1. Observation Identification

The observation identification fields shall consist of two integers that record the observation identification (*ObsId*) and observation interval identification (*ObI*) of the observation in which the source was detected.

The *ObsId* and *ObI* in the Observation Object shall be identified by the field names "**`obsid`**" and "**`obi`**", respectively.

1.2.1.2. Pointing Information

1.2.1.2.1. Target Designation

The target designation field shall consist of a character string value that records the target name *specified by the observer* in the original proposal under which the observation was obtained.

Note that the target name specified by the observer will match the target name in the *Chandra* Observation Catalog, but may be unrelated to the current source (or indeed any source in the field of view), and is not required to comply with any naming conventions.

The target name in the Observation Object shall be identified by the field name "**`targname`**".

1.2.1.2.2. Target Coordinates

The target coordinates field shall consist of 2 double precision values that record the target right ascension in decimal degrees and the signed target declination in decimal degrees, *as specified by the observer* in the original proposal under which the observation was obtained.

Note that the target coordinates specified by the observer will match the target coordinates in the *Chandra* Observation Catalog, but may be unrelated to the current source (or indeed any source in the field of view). The accuracy of the target coordinates specified by the observer is

indeterminate.

The target right ascension and declination in the Observation Object shall be identified by the field names "**ra_targ**" and "**dec_targ**", respectively.

1.2.1.2.3. Mean Pointing

The mean pointing of telescope during the observation shall consist of 3 double precision values that record the aspect-averaged ICRS right ascension in decimal degrees, the signed ICRS declination in decimal degrees, and the roll angle of the optical axis of the telescope.

The mean pointing right ascension, declination, and roll angle fields in the Observation Object shall be identified by the field names "**ra_pnt**", "**dec_pnt**", and "**roll_pnt**", respectively.

1.2.1.2.4. Tangent Plane Reference

The tangent plane reference shall consist of 3 double precision values that record the ICRS right ascension in decimal degrees, the signed ICRS declination in decimal degrees, and the roll angle that are the reference for the tangent plane projection of the world coordinates onto the SKY coordinate plane for the observation.

The tangent plane reference right ascension, declination, and roll angle fields in the Observation Object shall be identified by the field names "**ra_nom**", "**dec_nom**", and "**roll_nom**", respectively.

1.2.1.3. Timing Information

1.2.1.3.1. Mission Elapsed Time

The mission elapsed time fields shall consist of 3 double precision values that record the start and stop times of valid data (Good Time Intervals) for the observation, and the total elapsed time of the observation, in seconds Mission Elapsed Time (MET). MET is the standard time system for *Chandra* data, and counts the number of seconds TT since 1998-01-01T00:00:00 TT. Note that the total elapsed time of the observation is *not* equal to the total exposure time in the general case

where there bad time may be present within the observation.

The mission elapsed time start and stop times of valid data, and total elapsed time fields in the Observation Object shall be identified by the field names "**gti_start**", "**gti_stop**", and "**gti_elapse**", respectively.

1.2.1.3.2. FITS Standard (ISO 8601) Format

The FITS standard (ISO 8601) format fields consist of 2 character string values that record the FITS standard (ISO 8601) format equivalents to the start and stop times of valid data, in the format "**<YYYY>-<MM>-<DD>T<HH>:<MM>:<SS>**", where **<YYYY>** is the 4 digit year, and **<MM>**, **<DD>**, **<HH>**, **<MM>**, and **<SS>** are the 2 digit month, day of month, hours, minutes, and seconds TT corresponding to the equivalent MET, *truncated* to integer seconds.

The FITS standard (ISO 8601) format start and stop times of valid data in the Observation Object shall be identified by the field names "**gti_obs**" and "**gti_end**", respectively.

1.2.1.3.3. Modified Julian Date

The modified Julian date fields consist of 2 double precision values that record the MJD (TT) corresponding to the start of valid data, and the reference MJD (TT) corresponding to zero seconds MET (*i.e.*, MJD 50814.0 TT).

The modified Julian date of the start of valid data and the reference MJD in the Observation Object shall be identified by the field names "**gti_mjd_obs**" and "**mjd_ref**", respectively.

1.2.1.4. Instrument Information

1.2.1.4.1. Instrument and Grating Selection

The instrument and grating selection fields consist of 2 character string values that record the instrument selection (**ACIS** or **HRC**), grating selection (**NONE**, **HETG**, or **LETG**), and datamode used to obtain the observation.

The instrument, grating, and datamode in the Observation Object shall be identified by the field names "**instrument**", "**grating**", and "**datamode**" respectively.

1.2.1.4.2. ACIS-specific Configuration Parameters

The ACIS-specific configuration parameters consist of 2 character string values that record the readout mode (**TIMED** for TIMED readout mode, **CONTINUOUS** for CONTINUOUS-CLOCKING readout mode), cycle for alternating exposure (interleaved) mode observations (**P** for the primary cycle, **S** for the secondary cycle), and 1 double precision value that records the CCD frame time (**exptime**) in seconds.

The readout mode, cycle, and frame time in the Observation Object shall be identified by the field names "**readmode**", "**cycle**", and "**exptime**" respectively.

1.2.1.4.3. HRC-specific Configuration Parameters

The HRC-specific configuration parameters consists of a Boolean whose value shall be **TRUE** if the HRC was in a high precision timing mode (all vetoes disabled) for the observation and **FALSE** otherwise.

High precision timing mode in the Observation Object shall be identified by the field name "**timing_mode**".

1.2.1.5. Processing Information

The processing information consists of 3 character string values that record the processing system software and calibration database versions used to create the Level 3 event (evt3) file for the observation, and the creation date/time (UTC) of the Level 3 event (evt3) file for the observation. The latter is recorded in FITS standard (ISO 8601) format [see section 1.2.1.3.2 (FITS Standard (ISO 8601) Format)].

The processing system software version, calibration database version, and Level 3 event (evt3) file creation date/time in the Observation Object shall be identified by the field names "**ascdsver**", "**caldbver**", and "**crdate**" respectively.

1.2.1.6. Observing Cycle

The observing cycle consists of an integer value that records the number of the observing cycle corresponding to the *Chandra* Announcement of Opportunity (AO) in which the observation was scheduled.

The observing cycle in the Observation Object shall be identified by the field name "**`ao`**".

1.2.2. Source Object

1.2.2.1. Source Identification

Each source detected within an observation shall be identified by means of an ordinal source region identifier and an ordinal source identifier within the source region. When combined with 1.2.1.1 (Observation Identification), the region identifier and source identifier (**`1`**, **`2`**, ...) uniquely identify a source/observation pair. A source name is assigned when the individual observation source detection is merged to construct the Master Source Object.

Note

The value of the source identifier shall be restricted to **`1`** in the first catalog release. This restriction will be lifted in a future release.

The region identifier and source identifier in the Source Object shall be identified by the field names "**`region_id`**" and "**`source_id`**" respectively.

1.2.2.2. Position and Position Errors

The equatorial coordinates of a source in each science energy band shall be determined in the ICRS reference frame by analysis of the counts in the source region in that energy band [see section 1.2.2.5.1 (Source Region)].

1.2.2.2.1. Equatorial Coordinates

The ICRS equatorial coordinates of each source in each science energy band shall consist of 5 double precision values that record the ICRS right

ascension in decimal degrees, the signed ICRS declination in decimal degrees, the major and minor radii of the 95% confidence level error ellipse of the source position in arcseconds, and the position angle of the major axis of the 95% confidence level error ellipse in degrees.

Notes

The 95% confidence level error ellipse of the source position will be approximated by a 95% confidence level error circle in the first catalog release. This limitation will be lifted in a future release.

In the first catalog release, the 95% confidence level error ellipse is defined on a tangent plane projection. The 0° position angle reference is defined on that tangent plane to be parallel to the true North direction at the location of the tangent plane reference [see section 1.2.1.2.4 (Tangent Plane Reference)].

The right ascension and declination fields in the Source Object shall be identified by the field names "**`ra`**" and "**`dec`**", respectively.

The major and minor radii and the position angle measured from North through East of the major axis of the 95% confidence level error ellipse shall be identified by the field names "**`err_ellipse_r0`**", "**`err_ellipse_r1`**", and "**`err_ellipse_ang`**", respectively.

1.2.2.2.2. Galactic Coordinates

The Galactic coordinates corresponding to the source ICRS equatorial coordinates in each science energy band in section 1.2.2.2.1 (Equatorial Coordinates) shall consist of 2 double precision values that record the Galactic longitude, *l*, and latitude, *b*, for equinox J2000.0 and epoch J2000.0.

The Galactic latitude and Galactic longitude fields in the Source Object shall be identified by the field names "**`gal_b`**" and "**`gal_l`**", respectively.

1.2.2.2.3. Off-Axis Angles

The location of the source region [see section 1.2.2.5.1 (Source Region)] that includes a source, relative to the optical axis, shall consist of 2 double precision values that record the aspect-averaged off-axis angle, θ, in units of arcminutes, and the azimuthal angle, ϕ, in units of decimal degrees.

The off-axis angle and azimuthal angle in the Source Object shall be identified by the field names "**`theta`**" and "**`phi`**", respectively.

1.2.2.2.4. Mean Chip Coordinates

The location of the source region [see section 1.2.2.5.1 (Source Region)] that includes a source, in chip coordinates, shall consist of 2 double precision values that record the effective **`CHIPX`** and **`CHIPY`** pixel positions corresponding to off-axis angles [see section 1.2.2.2.3 (Off-Axis Angles)] (θ, ϕ).

The averaged **`CHIPX`** and **`CHIPY`** pixel positions in the Source Object shall be identified by the field names "**`chipx`**" and "**`chipy`**", respectively.

1.2.2.3. Source Significance

1.2.2.3.1. Detection Significance

The detection significance shall consist of a double precision value for each source detection energy band that records the value of the source significance of the detection computed by **`wavdetect`** (by dividing the net source counts in the *source detection region* [the source region output by **`wavdetect`**] by the Gehrels error of the sum of the estimated background counts in each pixel of the source region).

The detection significance in the Source Object shall be identified by the field name "**`detect_significance`**".

1.2.2.3.2. Flux Significance

The flux significance shall consist of a double precision value for each science energy band that records the flux significance, defined as

photflux_aper divided by the standard deviation of a 1-dimensional Gaussian distribution that has the same FWHM as the probability density function from which the upper and lower confidence limits (**photflux_aper_hilim** and **photflux_aper_lolim**) are determined, derived from the counts in the source region [see section 1.2.2.5.1 (Source Region)] in that energy band.

The flux significance in the Source Object shall be identified by the field name "**flux_significance**".

1.2.2.4. Source Flags

1.2.2.4.1. Extent Code

The extent code consists of a coded byte whose value shall be 0 if a wavelet transform analysis of counts in the source region [see section 1.2.2.5.1 (Source Region)] is consistent with a point source at the 90% confidence level in all science energy bands. Otherwise, the extent code coding shall be as follows:

1 = Deconvolved source extent [see section 1.2.2.5.4 (Deconvolved Source Extent)] is not consistent with point source at the 90% confidence level in the ultra-soft band,
2 = Deconvolved source extent [see section 1.2.2.5.4 (Deconvolved Source Extent)] is not consistent with point source at the 90% confidence level in the soft band,
4 = Deconvolved source extent [see section 1.2.2.5.4 (Deconvolved Source Extent)] is not consistent with point source at the 90% confidence level in the medium band,
8 = Deconvolved source extent [see section 1.2.2.5.4 (Deconvolved Source Extent)] is not consistent with point source at the 90% confidence level in the hard band,
16 = Deconvolved source extent [see section 1.2.2.5.4 (Deconvolved Source Extent)] is not consistent with point source at the 90% confidence level in the broad band,
32 = Deconvolved source extent [see section 1.2.2.5.4 (Deconvolved Source Extent)] is not consistent with point source at the 90% confidence level in the wide band.
The extent code field in the Source Object shall be identified by the field name "**extent_code**".

1.2.2.4.1.1. [Deleted]

1.2.2.4.2. Confusion Code

The confusion code consists of a coded byte whose value shall be 0 if multiple sources are not associated with a single source detection region in any source detection energy bands, *and* the source detection regions in all source detection energy bands for a source do not overlap the source detection regions for any other source. Otherwise, the confusion code coding shall be as follows:

1 = Multiple sources identified within the source detection region,
2 = Source detection region overlaps another source detection region,
4 = Source detection region overlaps another background region,
8 = Background region overlaps another source detection region,
16 = Background region overlaps another background region.

The confusion code field in the Source Object shall be identified by the field name "**`conf_code`**".

1.2.2.4.3. Pile-up Information (ACIS)

1.2.2.4.3.1. Pile-up Warning

The pile-up warning field consists of a double precision value that records the total number of counts in the broad science energy band per ACIS frame in the source region, normalized by the number of pixels in the source region [see section 1.2.2.5.1 (Source Region)], averaged over the observation. Users can correlate this value with pile-up models to crudely estimate the pile-up fraction sources identified within the source region.

The pile-up warning field in the Source Object shall be identified by the field name "**`pileup_warning`**".

1.2.2.4.4. Variability Code

The variability code consists of a coded byte whose value shall be 0 if variability is not detected either within the observation in any science energy band. Otherwise, the variability code coding shall be as follows:

1 = Intra-observation variability was detected in the ultra-soft (**u**) band,
2 = Intra-observation variability was detected in the soft (**s**) band,
4 = Intra-observation variability was detected in the medium (**m**) band,
8 = Intra-observation variability was detected in the hard (**h**) band,
16 = Intra-observation variability was detected in the broad (**b**) band,
32 = Intra-observation variability was detected in the wide (**w**) band.

The appropriate bit in the variability code corresponding to a specific science energy band shall be set equal to 0 if the variability index [see section 1.2.2.8.4 (Variability Index)] for that science energy band is ≤ 2, otherwise the bit shall be set equal to 1.

The variability code field in the Source Object shall be identified by the field name "**var_code**".

1.2.2.4.5. Chip Edge Code

The chip edge code consists of a coded byte whose value shall be 0 if the source position, source region [see section 1.2.2.5.1 (Source Region)], and background region do not dither off a chip boundary (the edge of the unmasked area of the active region of the ACIS CCD or HRC micro-channel plate segment, as appropriate) during the observation. Otherwise, the chip edge code coding shall be as follows:

1 = Source position dithers off any chip edge during the observation,
2 = Source region dithers off any chip edge during the observation,
4 = Background region dithers off any chip edge during the observation.

The chip edge code field in the Source Object shall be identified by the field name "**edge_code**".

1.2.2.4.6. Multi-chip Code

The multi-chip code consists of a coded byte whose value shall be 0 if the source position, source region [see section 1.2.2.5.1 (Source Region)], and background region do not dither between two chips (ACIS CCDs or HRC micro-channel plate segments) during the observation. Otherwise, the multi-chip code coding shall be as follows:

1 = Source position dithers across two chips during the observation,

2 = Source region dithers across two chips during the observation,
4 = Background region dithers across two chips during the observation,
8 = Source position dithers across more than two chips during the observation,
16 = Source region dithers across more than two chips during the observation,
32 = Background region dithers across more than two chips during the observation.

The multi-chip code field in the Source Object shall be identified by the field name "**multi_chip_code**".

1.2.2.4.7. Streak Source Flag

The streak source flag consists of a Boolean whose value shall be **TRUE** if the observation was obtained using ACIS *and* if the mean chip coordinates [see section 1.2.2.2.4 (Mean Chip Coordinates)] of the source region fall within a defined region enclosing an identified readout streak. Otherwise, the streak source flag shall be **FALSE**.

The streak source flag field in the Source Object shall be identified by the field name "**streak_src_flag**".

1.2.2.4.8. Saturated Source Flag

The saturated source flag consists of a Boolean whose value shall be **TRUE** if the observation was obtained using ACIS *and* if the source is likely severely piled-up to the extent that the source image may be flat-topped or have a central hole. In such cases, most source properties are very likely unreliable (the value of the pile-up warning field [see section 1.2.2.4.3.1 (Pile-up Warning)] is likely unreliable in these cases). Otherwise, the saturated source flag shall be **FALSE**.

The saturated source flag field in the Source Object shall be identified by the field name "**sat_src_flag**".

1.2.2.4.9. Manual Source Region Inclusion Flag

The manual source region inclusion flag consists of a Boolean whose value shall be **TRUE** if the source region was manually included in the

database of source-by-observation objects by human review. (These source regions may typically be located in locally-crowded regions of the sky that violate automated quality assurance local source region number density thresholds.) Otherwise, the manual source region inclusion flag shall be **FALSE**.

The manual source region inclusion flag field in the Source Object shall be identified by the field name "**man_inc_flag**".

1.2.2.4.10. Manual Source Region Parameters Flag

The manual source region parameters flag consists of a Boolean whose value shall be **TRUE** if the source region parameters [see section 1.2.2.5.1 (Source Region)] were manually modified in the database of source-by-observation objects by human review. (These source regions may typically be extremely piled-up so that automated source detection failed to generate reasonable source region parameters.) Otherwise, the manual source region parameters flag shall be **FALSE**.

The manual source region parameters flag field in the Source Object shall be identified by the field name "**man_reg_flag**".

1.2.2.5. Source Extent and Errors

1.2.2.5.1. Source Region

The source region and background region for each source are defined as the single elliptical source region, and single co-located scaled elliptical annular background region that are determined by scaling and merging the individual source detection regions resulting from all of the source detection process (**wavdetect**) spatial scales and source detection energy bands in which the source is detected. The merging and scaling process is described in section 1.1.4.4 of part V (ROI Generation).

The parameter values that define the source region and background region for each source shall consist of 10 double precision values that record the ICRS right ascension of the center of the source and background regions in decimal degrees, the signed ICRS declination of the center of the source and background regions in decimal degrees, the semi-major and semi-minor axes of the source region ellipse and the inner and outer background

region ellipses in arcseconds, and the position angle of the major axes of the region ellipses in degrees.

Notes

In the first catalog release, the source region is defined on a tangent plane projection. The 0° position angle reference is defined on that tangent plane to be parallel to the true North direction at the location of the tangent plane reference [see section 1.2.1.2.4 (Tangent Plane Reference)].

The ICRS right ascension and signed ICRS declination of the center of the source and background regions, the semi-major axis, semi-minor axis, and position angle of the major axis fields that define the source region, the semi-major and semi-minor axis fields that define the inner and outer annuli of the background region, and the position angle of the major axis field that defines the background region in the Source Object shall be identified by the field names "**ra_aper**", "**dec_aper**", "**mjr_axis_aper**", "**mnr_axis_aper**", "**pos_angle_aper**", "**mjr_axis1_aperbkg**", "**mnr_axis1_aperbkg**", "**mjr_axis2_aperbkg**", "**mnr_axis2_aperbkg**", and "**pos_angle_aperbkg**", respectively.

1.2.2.5.1.1. Modified Source Region

The modified source region and modified background region for each source are defined as the intersection of the source region and background region [see section 1.2.2.5.1 (Source Region)] for that source with the field of view [see section 2.1.8 (Field Of View Object) of part III (Archive Requirements)] and excluding any overlapping source regions. The explicit definitions of the modified source region and modified background region for a source are recorded in the region object [see section 2.2.7 (Region Object) of part III (Archive Requirements)] for that source.

The parameter values that define the modified source region and modified background region for each source shall consist of 2 double precision values that record the area of the modified source region and modified background region in square arcseconds.

The area of the modified source region and modified background region in the Source Object shall be identified by the field names "**area_aper**" and "**area_aperbkg**", respectively.

1.2.2.5.2. Source Extent

The source extent is a rotated elliptical Gaussian parameterization of the raw extent of a source, and shall consist of 6 double precision values for each science energy band that record the 1 σ radii along the major and minor axes of the rotated elliptical Gaussian in arcseconds on the detector, the position angle of the rotated elliptical Gaussian in degrees, and the associated errors. The parameterization is computed from a wavelet transform analysis of the counts in the source region [see section 1.2.2.5.1 (Source Region)] in that energy band.

Notes

In the first catalog release, the source extent is defined on a tangent plane projection. The 0° position angle reference is defined on that tangent plane to be parallel to the true North direction at the location of the tangent plane reference [see section 1.2.1.2.4 (Tangent Plane Reference)].

The 1 σ radii along the major and minor axes, position angle of the major axis of the source ellipse, and associated error fields in the Source Object shall be identified by the field names "**mjr_axis_raw**", "**mnr_axis_raw**", "**pos_angle_raw**","**mjr_axis_raw_err**", "**mnr_axis_raw_err**", and "**pos_angle_raw_err**", respectively.

1.2.2.5.3. Point Spread Function Extent

The point spread function extent is a rotated elliptical Gaussian parameterization of the raw extent of the point spread function (PSF) at the location of a source, and shall consist of 6 double precision values for each science energy band that record the 1 σ radii along the major and minor axes of a rotated elliptical Gaussian in arcseconds on the detector, the position angle of the rotated elliptical Gaussian in degrees, and the associated errors, that the detection process would assign to a monochromatic PSF at the location of the source, and whose energy is the effective energy of that energy band. The parameterization of the PSF is

computed from a wavelet transform analysis of the PSF counts in the source region [see section 1.2.2.5.1 (Source Region)] in that energy band and can be compared with the parameterization of the detected source to determine whether the latter is consistent with a point source.

Notes

In the first catalog release, the source region is defined on a tangent plane projection. The 0° position angle reference is defined on that tangent plane to be parallel to the true North direction at the location of the tangent plane reference [see section 1.2.1.2.4 (Tangent Plane Reference)].

The 1 σ radii along the major and minor axes, position angle of the major axis of the point spread function ellipse, and associated error fields in the Source Object shall be identified by the field names "**psf_mjr_axis_raw**", "**psf_mnr_axis_raw**", "**psf_pos_angle_raw**", "**psf_mjr_axis_raw_err**", "**psf_mnr_axis_raw_err**", and "**psf_pos_angle_raw_err**", respectively.

1.2.2.5.4. Deconvolved Source Extent

A parameterization of the deconvolved extent of each source shall be determined in each science energy band. The parameterization shall consist of the best estimate values for the 1 σ radius along the major axis, 1 σ radius along the minor axis, and position angle of the major axis (and associated errors) of a rotated elliptical Gaussian *that is convolved with the ray-trace local PSF* at the location of the source spatial event distribution.

The best estimates of the parameter values for a source shall consist of 6 double precision values for each science energy band that record the 1 σ radii along the major and minor axes of the rotated elliptical Gaussian parameterization of the source extent in arcseconds, the position angle of the major axis of the rotated elliptical Gaussian in degrees, and the associated errors.

Note

The rotated elliptical Gaussian parameterization of the source extent will be approximated by a circularly symmetric Gaussian parameterization in the first catalog release. This limitation will be lifted in a future release.

In the first catalog release, the deconvolved source extent is defined on a tangent plane projection. The 0° position angle reference is defined on that tangent plane to be parallel to the true North direction at the location of the tangent plane reference [see section 1.2.1.2.4 (Tangent Plane Reference)].

The 1 σ radii along the major and minor axes, position angle of the major axis, and associated error fields in the Source Object shall be identified by the field names "**major_axis**", "**minor_axis**", "**pos_angle**", "**major_axis_err**", "**minor_axis_err**", and "**pos_angle_err**", respectively.

1.2.2.6. Source Fluxes

Source fluxes shall be determined by aperture photometry in each science energy band. Aperture photometry measurements shall be obtained in the source region, and in an elliptical aperture that includes the 90% ECF of the PSF at the source location, with corrections applied for the PSF aperture fraction. The correction factor shall be computed from the ray-trace local PSF under the assumption that the source spatial distribution matches the local PSF.

1.2.2.6.1. Energy Bands

1.2.2.6.1.1. ACIS Observations

The energy bands used for ACIS observations are split into two categories. Source detection energy bands are used only in the source detection processing to detect sources; they are chosen to maximize the source detection efficiency for different kinds of sources observed using the back- and front-side illuminated CCDs. Science energy bands are used for all science analyses and are chosen to maximize the scientific utility of the catalog. The process by which the energy bands were chosen is documented in Appendix A.

1.2.2.6.1.1.1. Source Detection Energy Bands

The source detection energy bands for ACIS are a single *broad* band, covering 0.5–7.0 keV, and *soft*, *medium*, and *hard* bands covering 0.5–1.2 keV, 1.2–2.0 keV, and 2.0–7.0 keV, respectively.

The broad, ultra-soft, soft, medium, and hard source detection energy bands are designated "**`B`**", "**`S`**", "**`M`**", and "**`H`**", respectively.

1.2.2.6.1.1.2. Science Energy Bands

The science energy bands for ACIS are a single *broad* band, covering 0.5–7.0 keV, and *ultra-soft*, *soft*, *medium*, and *hard* bands covering 0.2–0.5 keV, 0.5–1.2 keV, 1.2–2.0 keV, and 2.0–7.0 keV, respectively.

The broad, ultra-soft, soft, medium, and hard science energy bands are designated "**`b`**", "**`u`**", "**`s`**", "**`m`**", and "**`h`**", respectively.

1.2.2.6.1.2. HRC Observations

The HRC does not have significant energy resolution. There is a single wide energy band for both source detection and science. The wide energy band for HRC includes pulse heights 0 : 254, which correspond approximately to photon energies 0.1–10.0 keV.

The wide energy band for HRC is designated "**`w`**" to distinguish it from the broad energy band for ACIS, which has different energy limits.

1.2.2.6.2. [Deleted]

1.2.2.6.3. Aperture Photometry Fluxes

1.2.2.6.3.1. Aperture Counts

1.2.2.6.3.1.1. Aperture Total Counts

The aperture total counts shall consist of 4 long integer values for each science energy band that record the total number of source plus background counts measured in the modified source region and the modified background region [see section 1.2.2.5.1.1 (Modified Source Region)], and in the modified elliptical aperture and the modified elliptical

background aperture [see section 1.2.2.6.3.2.1 (Modified Elliptical Aperture)], uncorrected by the PSF aperture fraction.

The aperture total counts in the source region and the background region, and in the modified elliptical aperture and the modified elliptical background aperture, in the Source Object shall be identified by the field names "**cnts_aper**", "**cnts_aperbkg**", "**cnts_aper90**", and "**cnts_aper90bkg**", respectively.

1.2.2.6.3.1.2. Aperture Source Counts

The aperture source counts shall consist of 2 double precision values for each science energy band that record the net number of background-subtracted source counts in the modified source region [see section 1.2.2.5.1.1 (Modified Source Region)], and in the modified elliptical aperture [see section 1.2.2.6.3.2.1 (Modified Elliptical Aperture)], corrected by the appropriate PSF aperture fractions [see 1.2.2.6.3.6 (PSF Aperture Fractions)].

The aperture source counts inferred from the modified source region, and inferred from the modified elliptical aperture, in the Source Object shall be identified by the field names "**src_cnts_aper**" and "**src_cnts_aper90**", respectively.

1.2.2.6.3.2. Elliptical Aperture

The elliptical apertures and elliptical background apertures for each source are defined as the elliptical apertures that includes the 90% ECF of the PSF in each science energy band at the source location used to extract the aperture counts, count rates, and photon and energy fluxes, and the co-located scaled elliptical annular background apertures. The elliptical aperture and elliptical background aperture for each science energy band of each source are co-located with the source region and background region [see section 1.2.2.5.1 (Source Region)] for that source.

The parameter values that define the elliptical aperture and elliptical background aperture for each source shall consist of 8 double precision values for each science energy band that record the semi-major and semi-minor axes of the elliptical aperture and the inner and outer elliptical background apertures in arcseconds, and the position angle of the major

axes of the elliptical apertures in degrees.

Notes

In the first catalog release, the elliptical aperture is defined on a tangent plane projection. The 0° position angle reference is defined on that tangent plane to be parallel to the true North direction at the location of the tangent plane reference [see section 1.2.1.2.4 (Tangent Plane Reference)].

The semi-major axis, semi-minor axis, and position angle of the major axis fields that define the elliptical aperture that includes the 90% ECF of the PSF at the source location, the semi-major and semi-minor axis fields that define the inner and outer annuli of the elliptical background aperture, and the position angle of the major axis field that defines the elliptical background aperture in the Source Object shall be identified by the field names "**mjr_axis_aper90**", "**mnr_axis_aper90**", "**pos_angle_aper90**", "**mjr_axis1_aper90bkg**", "**mnr_axis1_aper90bkg**", "**mjr_axis2_aper90bkg**", "**mnr_axis2_aper90bkg**", and "**pos_angle_aper90bkg**", respectively.

1.2.2.6.3.2.1. Modified Elliptical Aperture

The modified elliptical aperture and modified elliptical background aperture for each source are defined as the intersection of the elliptical aperture and elliptical background aperture [see section 1.2.2.6.3.2 (Elliptical Aperture)] for that source with the field of view [see section 2.1.8 (Field Of View Object) of part III (Archive Requirements)] and excluding any overlapping source regions [see section 1.2.2.5.1 (Source Region)].

The parameter values that define the modified elliptical aperture and modified elliptical background aperture for each source shall consist of 2 double precision values for each science energy band that record the area of the modified elliptical aperture and modified elliptical background aperture in square arcseconds.

The area of the modified elliptical aperture and modified elliptical background aperture in the Source Object shall be identified by the field

names “**area_aper90**” and “**area_aper90bkg**”, respectively.

1.2.2.6.3.3. Aperture Source Count Rates

The aperture source count rates and associated confidence limits shall consist of 6 double precision values for each science energy band that record the background-subtracted source count rates and two-sided count rate confidence limits in the modified source region [see section 1.2.2.5.1.1 (Modified Source Region)], and in the modified elliptical aperture [see section 1.2.2.6.3.2.1 (Modified Elliptical Aperture)], corrected by the appropriate PSF aperture fractions [see 1.2.2.6.3.6 (PSF Aperture Fractions)] and livetime [see 1.2.2.9.1 (Livetime)], in units of Hz.

The aperture source count rates inferred from the modified source region, and inferred from the modified elliptical aperture, and their associated two-sided confidence limits, in the Source Object shall be identified by the field names “**src_rate_aper**”, “**src_rate_aper90**”, “**src_rate_aper_hilim**”, “**src_rate_aper_lolim**”, “**src_rate_aper90_hilim**”, and “**src_rate_aper90_lolim**”, respectively.

1.2.2.6.3.4. Aperture Source Energy Fluxes

The aperture source energy fluxes and associated confidence limits shall consist of 12 double precision values for each science energy band that record the (background-subtracted) fluxes and two-sided confidence limits in the modified source region [see section 1.2.2.5.1.1 (Modified Source Region)], and in the modified elliptical aperture [see section 1.2.2.6.3.2.1 (Modified Elliptical Aperture)], corrected by the appropriate PSF aperture fractions [see 1.2.2.6.3.6 (PSF Aperture Fractions)], livetime [see 1.2.2.9.1 (Livetime)], and exposure, in units of photons s^{-1} cm^{-2} and ergs s^{-1} cm^{-2}. The conversion from photons s^{-1} cm^{-2} to ergs s^{-1} cm^{-2} is performed by summing the photon energies for each incident source photon and scaling by the local value of the ARF at the location of the incident photon.

The aperture source energy fluxes inferred from the modified source region, and inferred from the modified elliptical aperture, and their associated two-sided confidence limits, in the Source Object shall be identified by the field names “**photflux_aper**”,

"**photflux_aper90**", "**photflux_aper_hilim**", "**photflux_aper_lolim**", "**photflux_aper90_hilim**", and "**photflux_aper90_lolim**", respectively, for the photon fluxes, and "**flux_aper**", "**flux_aper90**", "**flux_aper_hilim**", "**flux_aper_lolim**", "**flux_aper90_hilim**", and "**flux_aper90_lolim**", respectively, for the energy fluxes.

1.2.2.6.3.5. Aperture Model Energy Fluxes

The aperture model energy fluxes and associated confidence limits shall consist of 12 double precision values for each science energy band that record the fluxes and two-sided confidence limits in the modified source region [see section 1.2.2.5.1.1 (Modified Source Region)], and in the modified elliptical aperture [see section 1.2.2.6.3.2.1 (Modified Elliptical Aperture)], corrected by the appropriate PSF aperture fractions [see 1.2.2.6.3.6 (PSF Aperture Fractions)], livetime [see 1.2.2.9.1 (Livetime)], and exposure, in units of ergs s^{-1} cm^{-2}. The conversion from aperture source count rates [see section 1.2.2.6.3.3 (Aperture Source Count Rates)] in each science energy band to aperture model energy fluxes is performed by scaling from a model spectrum folded through the calibrated response as follows:

> For a source model $\mathcal{F}(E)$ whose integral over the science band is $\mathcal{F}'(\text{band})$, calculate the corresponding band count rate $\mathcal{C}'(\text{band})$ in counts s^{-1} given then effective area calibration $\mathcal{A}(E)$ [and, if available the RMF, $\mathcal{R}(E, \text{band})$] appropriate to the observation; this is $\int \mathcal{F}(E)\mathcal{A}(E)\mathcal{R}(E, \text{band})$ over all energies, or if a diagonal RMF is assumed, $\int \mathcal{F}(E)\mathcal{A}(E)$ over the band. Infer the aperture model energy flux from the measured aperture source count rate $\mathcal{C}(\text{band})$ as $\mathcal{F}(\text{band}) = \mathcal{F}'(\text{band})\mathcal{C}(\text{band})/\mathcal{C}'(\text{band})$. The power-law spectral model shall have a fixed spectral index (γ, defined as $F_E \sim E^{-\gamma}$) = 1.7, and a fixed total neutral Hydrogen absorbing column N_H = $N_H(\text{Gal})$ cm^{-2}. The black body spectral model shall have a fixed temperature kT = 1.0 keV, and a fixed total neutral Hydrogen absorbing column $N_H = 3 \times 10^{20}$ cm^{-2}.

The power-law and black body aperture model energy fluxes inferred from the modified source region, and inferred from the modified elliptical

aperture, and their associated two-sided confidence limits, in the Source Object shall be identified by the field names "**flux_powlaw_aper**", "**flux_powlaw_aper90**", "**flux_bb_aper**", "**flux_bb_aper90**", "**flux_powlaw_aper_hilim**", "**flux_powlaw_aper_lolim**", "**flux_powlaw_aper90_hilim**", "**flux_powlaw_aper90_lolim**", "**flux_bb_aper_hilim**", "**flux_bb_aper_lolim**", "**flux_bb_aper90_hilim**", and "**flux_bb_aper90_lolim**", respectively.

1.2.2.6.3.6. PSF Aperture Fractions

The PSF aperture fractions shall consist of 4 double precision values for each science energy band that record the fraction of the PSF that is included in the modified source region and the modified background region [see section 1.2.2.5.1.1 (Modified Source Region)], and in the modified elliptical aperture and the modified elliptical background aperture [see section 1.2.2.6.3.2.1 (Modified Elliptical Aperture)].

The PSF aperture fractions in the modified source region and the modified background region, and in the modified elliptical aperture that and the modified elliptical background aperture in the Source Object shall be identified by the field names "**psf_frac_aper**", "**psf_frac_aperbkg**", "**psf_frac_aper90**", and "**psf_frac_aper90bkg**", respectively.

1.2.2.7. Spectral Properties

1.2.2.7.1. Spectral Model Fits

If there are at least 150 net (background subtracted) counts in the energy range 0.5–7.0 keV present in the source region [see section 1.2.2.5.1 (Source Region)] of an ACIS observation, then power law and thermal (black body) model spectra are fitted to the PI event data extracted from the source region, with corrections for the PSF aperture fraction, livetime, and ARF applied when fitting the models.

1.2.2.7.1.1. Power Law Model Spectral Fit

The power law model spectral fit is performed over the energy range 0.5–

7.0 keV as described in section 1.2.4.4 of part V (Spectral Fit). The free parameters to be fitted are the total integrated flux, total neutral Hydrogen absorbing column, and power law spectral index.

1.2.2.7.1.1.1. Power Law Model Photon Index

The power law model photon index and associated confidence limits shall consist of 3 double precision values that record the best-fit power law photon index (α, defined as $F_E \sim E^{-\alpha}$) and associated two-sided confidence limits from the power law model spectral fit.

The power law model photon index and associated two-sided confidence limits in the Source Object shall be identified by the field names "**alpha**", "**alpha_hilim**", and "**alpha_lolim**", respectively.

1.2.2.7.1.1.2. Power Law Model Spectral Fit Flux

The power law model spectral fit flux and associated confidence limits shall consist of 3 double precision values that record the integrated 0.5–7 keV flux and associated two-sided confidence limits derived from the best fitting power law model spectral fit in units of ergs s^{-1} cm^{-2}.

The power law model spectral fit flux and associated two-sided confidence limits in the Source Object shall be identified by the field names "**flux_powlaw**", "**flux_powlaw_hilim**", and "**flux_powlaw_lolim**", respectively.

1.2.2.7.1.1.3. Power Law Model Spectral Fit N_H

The power law model spectral fit N_H and associated confidence limits shall consist of 3 double precision values that record the best-fit total neutral Hydrogen absorbing column and associated two-sided confidence limits from the power law model spectral fit in units of 10^{20} cm^{-2}.

The power law model spectral fit N_H and associated two-sided confidence limits in the Source Object shall be identified by the field names "**nh_powlaw**", "**nh_powlaw_hilim**", and "**nh_powlaw_lolim**", respectively.

1.2.2.7.1.1.4. Power Law Model Spectral Fit Statistic

The power law model spectral fit statistic shall consist of a double precision value that records the value of the χ^2 (data variance) statistic per degree of freedom for the best fitting power law model spectral fit.

The power law model spectral fit statistic field in the Source Object shall be identified by the field name "**`powlaw_stat`**".

1.2.2.7.1.2. Thermal (Black Body) Model Spectral Fit

The thermal (black body) model spectral fit is performed over the energy range 0.5–7.0 keV as described in section 1.2.4.4 of part V (Spectral Fit). The free parameters to be fitted are the total integrated flux, total neutral Hydrogen absorbing column, and black body temperature.

1.2.2.7.1.2.1. Thermal (Black Body) Spectral Fit Temperature

The thermal (black body) spectral fit temperature and associated confidence limits shall consist of 3 double precision values that record the best-fit temperature (*kT*) in units of keV and associated two-sided confidence limits from the thermal (black body) model spectral fit.

The thermal (black body) spectral fit temperature and associated two-sided confidence limits in the Source Object shall be identified by the field names "**`kt`**", "**`kt_hilim`**", and "**`kt_lolim`**", respectively.

1.2.2.7.1.2.2. Thermal (Black Body) Spectral Fit Flux

The thermal (black body) spectral fit flux and associated confidence limits shall consist of 3 double precision values that record the integrated 0.5–7 keV flux and associated two-sided confidence limits derived from the best fitting thermal (black body) model spectral fit in units of ergs s^{-1} cm^{-2}.

The thermal (black body) spectral fit flux and associated two-sided confidence limits in the Source Object shall be identified by the field names "**`flux_bb`**", "**`flux_bb_hilim`**", and "**`flux_bb_lolim`**", respectively.

1.2.2.7.1.2.3. Thermal (Black Body) Spectral Fit N_H

The thermal (black body) spectral fit N_H and associated confidence limits shall consist of 3 double precision values that record the best-fit total neutral Hydrogen absorbing column and associated two-sided confidence limits from the thermal (black body) model spectral fit in units of 10^{20} cm^{-2}.

The thermal (black body) spectral fit N_H and associated two-sided confidence limits in the Source Object shall be identified by the field names "**nh_bb**", "**nh_bb_hilim**", and "**nh_bb_lolim**", respectively.

1.2.2.7.1.2.4. Thermal (Black Body) Model Spectral Fit Statistic

The thermal (black body) model spectral fit statistic shall consist of a double precision value that records the value of the χ^2 (data variance) statistic per degree of freedom for the best fitting thermal (black body) model spectral fit.

The thermal (black body) model spectral fit statistic field in the Source Object shall be identified by the field name "**bb_stat**".

1.2.2.7.2. Hardness Ratios

Hardness ratios shall consist of 3 double precision values for each pair of science energy bands, excluding the broad and ultra-soft bands, that record the hardness ratios and associated two-sided confidence limits computed from the aperture total counts [see section 1.2.2.6.3.1.1 (Aperture Total Counts)]. The hardness ratio for two bands with designations **<x>** and **<y>**, respectively, is defined equivalently as

hard_<x><y> = (*F*(<x>)-*F*(<y>)) / *F*(<broad>),

where ***F*(<x>)** is the aperture source photon flux **(photflux_aper)** in band **<x>**, and ***F*(<broad>)** is the aperture source photon flux in the broad energy band.

The hardness ratios and associated upper and lower confidence limits in the Source Object shall be identified by the field names

"**hard_<x><y>**", "**hard_<x><y>_hilim**", and "**hard_<x><y>_lolim**", respectively, where **<x>** and **<y>** are the higher and lower energy band designations. For example, the hardness ratio and associated two-sided confidence limits for the hard-to-soft ACIS science energy bands shall be identified by the field names "**hard_hs**", "**hard_hs_hilim**", and "**hard_hs_lolim**", respectively.

1.2.2.8. Source Variability

Source variability within an observation shall be assessed by three methods: (1) application of the Kolmogorov-Smirnov (K-S) test to the (unbinned) source region events, (2) application of the Kuiper's test to the source region counts, and (3) computation of the Gregory-Loredo variability probability from the source region counts.

1.2.2.8.1. Kolmogorov-Smirnov (K-S) Test

The Kolmogorov-Smirnov (K-S) test probability shall consist of a double precision value for each science energy band that records the probability that the receipt times of the events within the source region are not consistent with a constant source region flux throughout the observation. The probability shall be computed from a one-sample, two-sided K-S test applied to the unbinned event data, with corrections applied for good time intervals and for the source region dithering across regions of variable exposure during the observation.

The Kolmogorov-Smirnov (K-S) test probability in the Source Object shall be identified by the field name "**ks_prob**".

1.2.2.8.2. Kuiper's Test

The Kuiper's test probability shall consist of a double precision value for each science energy band that records the probability that the receipt times of the events within the source region are not consistent with a constant source region flux throughout the observation. The probability shall be computed from a one-sample Kuiper's test applied to the unbinned event data, with corrections applied for good time intervals and for the source region dithering across regions of variable exposure during the observation.

The Kuiper's test probability in the Source Object shall be identified by the field name "**kp_prob**".

1.2.2.8.3. Gregory-Loredo Variability Probability

The Gregory-Loredo variability probability shall consist of a double precision value for each science energy band that records the probability that the source region flux is not uniform throughout the observation based on the odds ratio calculated from a Gregory-Loredo analysis of the receipt times of the events within the source region.

The Gregory-Loredo variability probability in the Source Object shall be identified by the field name "**var_prob**".

1.2.2.8.4. Variability Index

The variability index shall consist of an integer value for each science energy band that records an index in the range [0, 10] that combines (a) the Gregory-Loredo variability probability [see section 1.2.2.8.3 (Gregory-Loredo Variability Probability)] with (b) the fractions of the multi-resolution light curve output by the Gregory-Loredo analysis [section 2.2.6 of part III (Light Curve Object)] that are within 3 σ and 5 σ of the average count rate, to evaluate whether the source region flux is uniform throughout the observation. A detailed description of the computation of the variability index, and its interpretation, is documented in Appendix F.

The variability index in the Source Object shall be identified by the field name "**var_index**".

1.2.2.8.5. Count Rate Variability

1.2.2.8.5.1. Mean Count Rate

The mean count rate shall consist of a double precision value for each science energy band that records the time-averaged value of the source region count rate derived from the multi-resolution light curve output by the Gregory-Loredo analysis [section 2.2.6 of part III (Light Curve Object)] in units of counts s^{-1}.

The mean count rate in the Source Object shall be identified by the field name "**var_mean**".

1.2.2.8.5.2. Count Rate Standard Deviation

The count rate standard deviation shall consist of a double precision value for each science energy band that records the time-averaged 1 σ statistical variability of the source region count rate derived from the multi-resolution light curve output by the Gregory-Loredo analysis [section 2.2.6 of part III (Light Curve Object)] in units of counts s^{-1}.

The count rate standard deviation in the Source Object shall be identified by the field name "**var_sigma**".

1.2.2.8.5.3. Minimum Count Rate

The minimum count rate shall consist of a double precision value for each science energy band that records the minimum value of the source region count rate derived from the multi-resolution light curve output by the Gregory-Loredo analysis [section 2.2.6 of part III (Light Curve Object)] in units of counts s^{-1}.

The minimum count rate in the Source Object shall be identified by the field name "**var_min**".

1.2.2.8.5.4. Maximum Count Rate

The maximum count rate shall consist of a double precision value for each science energy band that records the maximum value of the source region count rate derived from the multi-resolution light curve output by the Gregory-Loredo analysis [section 2.2.6 of part III (Light Curve Object)] in units of counts s^{-1}.

The maximum count rate in the Source Object shall be identified by the field name "**var_max**".

1.2.2.8.6. Dither Warning Flag

The dither warning flag consists of a Boolean whose value shall be **TRUE** if the highest statistically significant peak in the power spectrum of the

source region count rate derived from the multi-resolution light curve output by the Gregory-Loredo analysis [section 2.2.6 of part III (Light Curve Object)] for the science energy band with the highest variability index [see section 1.2.2.8.4 (Variability Index)] occurs either at the dither frequency of the observation or at a beat frequency of the dither frequency. Otherwise, the dither warning flag shall be **FALSE**.

The dither warning flag in the Source Object shall be identified by the field name "**dither_warning_flag**".

1.2.2.9. Timing Information

1.2.2.9.1. Livetime

The livetime field shall consist of a double precision value that records the livetime for the observation, in seconds. The livetime is the effective exposure time after applying the good time intervals (GTIs) and the deadtime correction factor (**dtcor**).

The livetime in the Source Object shall be identified by the field name "**livetime**".

1.2.2.10. Instrument Information

1.2.2.10.1. Detector Name

The detector name field consist of a character string value that records the name of the detector elements that the bounding box of the background region of interest (ROI) dithers over during the observation. For HRC-I, the recorded values shall be "**HRC-I**", and for HRC-S, the recorded values shall be "**HRC-S**". For ACIS, the recorded value shall be "**ACIS-<n>**", where **<n>** is the string representation of the set of ACIS CCDs over which the bounding box of the background region of interest (ROI) dithers, identified by CCD number (0–9) in numeric order. For example, if the bounding box of the background region of interest (ROI) dithered across the boundary of all 4 ACIS-I chips (I0–I3), then the recorded value would be "**ACIS-0123**", and if the bounding box of the background region of interest (ROI) dithered across the boundary of ACIS-S chips S3 and S4, then the recorded value would be "**ACIS-78**".

The detector name in the Source Object shall be identified by the field name "**`detector`**".

2. Criteria for Source Inclusion

Master Source Objects and Source-by-Observation Objects that are included in the database shall only be candidates for inclusion in the catalog if they meet certain requirements, listed in this section. Candidate sources that meet these requirements are included in the catalog by following the procedures documented in section 5 (Release Process).

2.1. Master Source Objects

Each Master Source Object that is included in the catalog shall *only* include contributions from Source-by-Observation Objects that meet the requirements for inclusion in the catalog.

2.1.1. Source Significance

The Master Source Object for a source must have a minimum flux significance [**`significance`** in section 1.1.5.4 (Source Significance)] of 3.0 in at least one science energy band to be a candidate for inclusion in the catalog.

2.2. Source-by-Observation Objects

All Source-by-Observation Objects that contribute to the merge processing used to generate a Master Source Object *that is included in the catalog* shall be included in the catalog. Source-by-Observation Objects that *do not* contribute to the merge processing used to generate a Master Source Object that is included in the catalog may be present in the database, but shall not be included in the catalog.

All sub-objects (*e.g.*, Observation Object, Source Object) of the Source-by-Observation Object for a given source and observation shall be treated as a single unit when applying this requirement.

2.2.1. Flux Significance

The Source-by-Observation Object for a source and observation must

have a minimum flux significance [**flux_significance** in section 1.2.2.3.2 (Flux Significance)] in the Source Object of 3.0 in at least one science energy band to be a candidate for inclusion in the catalog.

2.3. Observation Selection

2.3.1. [Deleted]

2.3.2. Extended Source Fields

Observations that include sources that are extended on 5 arcminute or larger size scales shall not be included in the catalog.

2.4. Quality Assurance Process

Only Master Source Objects and Source-by-Observation Objects that meet the requirements of part VI (Quality Assurance Requirements) shall be candidates for inclusion in the catalog.

3. Catalog Characterization (Statistical)

3.1. Source Detection Efficiency

Characterization of source detection efficiency shall be performed for each instrument (ACIS-I, ACIS-S, HRC-I, and HRC-S) assuming that the nominal instrument aim-point and default SIM offset positions are used for the observation. The characterization shall apply to the complete source detection process, including any post-detection steps that affect the resulting source list.

Where the characterization depends on the source spectral energy distribution, the characterization shall be performed assuming an unabsorbed power-law spectral energy distribution with power law indices $\gamma = 1.7$, where $F_E = E^{-\gamma}$, and a thermal spectral energy distribution with a black-body temperature $kT = 1.0$ keV.

Where the characterization is performed per energy band, then the source shall be assumed to have a monochromatic spectrum with an energy equal to the equivalent monochromatic energy of the energy band.

3.1.1. Limiting Sensitivity

Characterization shall record the limiting flux of a point source that can be detected at the catalog flux significance limit [see section 2.1.1 (Source Significance)] as a function of off-axis angle, total background integrated over the duration of the observation, and instrument configuration.

Off-axis angles from 0 to 15 arcminutes shall be characterized. Total (X-ray plus instrumental) backgrounds, which include contributions of 1 times the average instrumental background rate derived from stowed observations for ACIS, and 1 times the nominal instrumental background rate for HRC, shall be characterized, with ACIS back-illuminated and front-illuminated CCDs considered separately.

3.1.1.1. Sky Coverage

Characterization shall record the total sky coverage in square degrees sensitive to detection of a point source at the catalog flux significance limit [see section 2.1.1 (Source Significance)] as a function of source flux and instrument configuration. ACIS back-illuminated and front-illuminated CCDs considered separately.

3.1.2. Completeness

Characterization shall record the fraction of point sources that are detected at the catalog flux significance limit [see section 2.1.1 (Source Significance)] as a function of source flux, off-axis angle, total background integrated over the duration of the observation, and instrument configuration.

Off-axis angles from 0 to 15 arcminutes shall be characterized. Total (X-ray plus instrumental) backgrounds, which include contributions of 1 times the average instrumental background rate derived from stowed observations for ACIS, and 1 times the nominal instrumental background rate for HRC, shall be characterized, with ACIS back-illuminated and front-illuminated CCDs considered separately.

3.1.3. False Source Rate

Characterization shall record the number of false sources per square

degree that are erroneously detected at the catalog flux significance limit [see section 2.1.1 (Source Significance)] as a function of source flux, off-axis angle, total background integrated over the duration of the observation, and instrument configuration. Image artifacts (e.g., readout streaks, unscreened bad pixels, ...) that are not rejected by the detection process shall be included in the determination of the false source rate.

Off-axis angles from 0 to 15 arcminutes shall be characterized. Total (X-ray plus instrumental) backgrounds, which include contributions of 1 times the average instrumental background rate derived from stowed observations for ACIS, and 1 times the nominal instrumental background rate for HRC, shall be characterized, with ACIS back-illuminated and front-illuminated CCDs considered separately.

3.2. Photometric Uncertainty

3.2.1. Aperture Photon Flux Uncertainty

Characterization shall record the uncertainty in measured aperture photon fluxes of un-piled-up point sources as a function of source flux, off-axis angle, aperture, total background integrated over the duration of the observation, and instrument configuration.

Off-axis angles from 0 to 15 arcminutes shall be characterized. Source fluxes from 10 to 100,000 counts, and total (X-ray plus instrumental) backgrounds, which include contributions of 1 times the average instrumental background rate derived from stowed observations for ACIS, and 1 times the nominal instrumental background rate for HRC, shall be characterized, with ACIS back-illuminated and front-illuminated CCDs considered separately.

3.3. Astrometric Uncertainty

3.3.1. Absolute Source Position Uncertainty

Characterization shall record the uncertainty in measured absolute source positions of un-piled-up point sources in the ICRS reference frame as a function of source flux, off-axis angle, total background integrated over the duration of the observation, and instrument configuration.

Off-axis angles from 0 to 15 arcminutes shall be characterized. Source fluxes from 10 to 100,000 counts, and total (X-ray plus instrumental) backgrounds, which include contributions of 1 times the average instrumental background rate derived from stowed observations for ACIS, and 1 times the nominal instrumental background rate for HRC, shall be characterized.

3.3.2. Relative Source Position Uncertainty

Characterization shall record the uncertainty in measured relative source positions of pairs of un-piled-up point sources in the ICRS reference frame as a function of source fluxes, off-axis angles, total background integrated over the duration of the observation, and instrument configuration.

Source pairs with separations of 0 to 30 arcminutes, and source off-axis angles from 0 to 15 arcminutes shall be characterized. Source fluxes from 10 to 100,000 counts, and total (X-ray plus instrumental) backgrounds, which include contributions of 1 times the average instrumental background rate derived from stowed observations for ACIS, and 1 times the nominal instrumental background rate for HRC, shall be characterized.

3.4. Source Extent

Characterization shall record the uncertainty in the measured source extent parameterization for point sources and extended sources with an elliptical flux distribution as a function of source flux, off-axis angle, total background integrated over the duration of the observation, and instrument configuration.

Point sources and extended sources with size scales up to 30 arcseconds, located at off-axis angles from 0 to 15 arcminutes, shall be characterized. Source fluxes from 10 to 100,000 counts, and total (X-ray plus instrumental) backgrounds, which include contributions of 1 times the average instrumental background rate derived from stowed observations for ACIS, and 1 times the nominal instrumental background rate for HRC, shall be characterized.

3.4.1. Close Source Pairs

Characterization shall record the uncertainty in measured source extent parameterizations for close pairs of un-piled-up point sources as a function of source separations, fluxes, off-axis angles, total background integrated over the duration of the observation, and instrument configuration.

Source pairs with separations from 20% to 150% of the 90% circular Encircled Counts Fraction (ECF) radius, and source off-axis angles from 0 to 15 arcminutes shall be characterized. Source fluxes from 10 to 100,000 counts, and Total (X-ray plus instrumental) backgrounds, which include contributions of 1 times the average instrumental background rate derived from stowed observations for ACIS, and 1 times the nominal instrumental background rate for HRC, shall be characterized.

3.5. Source Variability

Characterization shall record the uncertainty in true variability as a function of the variability index (for values > 0). Characterization shall record the uncertainty in the lack of variability for the variability index value = 0.

4. Traceability

4.1. Revision History Tracking

4.1.1. Data History

The history of every data element within the catalog shall be maintained, including a record of date and time of the insertion, revision, or deletion of any data element. Changes to the catalog shall only be performed according to the requirements of section 5 (Release Process).

4.1.2. Catalog Releases

Each catalog release shall correspond to a view of the state of the catalog corresponding to a specified (predefined) date and time. New releases of the catalog shall take place according to the requirements of section 5 (Release Process).

5. Release Process

5.1. Timing and Notification

Notification of the planned timing of catalog releases shall be provided to the user community at least 1 month prior to the scheduled date of the release.

Other than in exceptional circumstances, catalog releases shall not be scheduled to occur between the release of a *Chandra* Announcement of Opportunity (AO) and the AO deadline.

5.2. Frequency

Catalog releases shall occur no more frequently than every 6 months, with no more than every year being the goal after 2 years following the first release of the catalog.

5.3. Atomicity of Releases

From the point of view of the user, each catalog release shall occur in an atomic manner. User views of a partially updated catalog shall not be permitted.

5.4. Characterization

The statistical characterization of the properties of the contents of a catalog release [see section 3 (Catalog Characterization (Statistical))] shall be provided to the user community at the same time as the catalog release, with the exception that the catalog release shall not be held up if a preliminary statistical characterization is available for all of the statistical properties, and the characterization of each property is largely complete, well-understood, and stable.

II. Database Requirements

1. Database Contents

Each *Chandra Source Catalog* release consist of a predefined "catalog view" of the *Chandra Source Catalog* "database", together with associated archival (file-based) data. Since the principal requirements of interest to the user are those related to a *Chandra Source Catalog* release, the detailed requirements for database objects that are visible in a catalog view are presented in part I (Catalog Requirements) of this document, which should be referenced for further information. This part of this document identifies any differences between database and catalog requirements, and specifies any requirements that apply only to the database.

1.1. Non-applicable Catalog Requirements

The following catalog requirements do not apply to the database.

1.1.1. Performance Requirements

Paragraphs identified as "*Performance Requirement*" in section 1 (Catalog Contents) of part I (Catalog Requirements) do not apply to the database.

1.1.2. Criteria for Source Inclusion

Section 2 (Criteria for Source Inclusion) of part I (Catalog Requirements) does not apply to the database.

1.1.3. Catalog Characterization (Statistical)

Section 3 (Catalog Characterization (Statistical)) of part I (Catalog Requirements) does not apply to the database.

1.1.4. Traceability

Section 4 (Traceability) of part I (Catalog Requirements) does not apply to the database.

1.1.5. Release Process

Section 5 (Release Process) of part I (Catalog Requirements) does not apply to the database.

1.2. Database-specific Requirements

1.2.1. Criteria for Source Inclusion

Master Source Objects and Source-by-Observation Objects shall only be included in the database if they meet certain requirements, listed in this section.

1.2.1.1. Quality Assurance Process

Only Master Source Objects and Source-by-Observation Objects that meet the requirements of part VI (Quality Assurance Requirements) shall be included in the database.

1.2.2. Traceability

1.2.2.1. Revision History Tracking

1.2.2.1.1. Data History

The history of every data element within the database shall be maintained, including a record of date and time of the insertion, revision, or deletion of any data element.

III. Archive Requirements

1. Archive Contents

1.1. Data Format

1.1.1. FITS Files

The data objects defined in section 2 (Data Object Definitions) shall be recorded in data files that are formatted in accordance with the requirements of the NOST FITS standard, version 2.0 or later. The headers of these files shall comply with the requirements of the ASC FITS File Designers' Guide, version ASC-FITS-2.1.0, as amended.

FITS file contents and header metadata contents shall include all information needed to enable the files to be manipulated by data analysis tools appropriate to the type of data product that are part of the CXC *CIAO* data analysis package. Data products shall support compatible use with other astronomical data analysis software to the extent possible given implementation constraints, provided that such compatibility is not inconsistent with the requirements of the tools that comprise the *CIAO* data analysis package.

2. Data Object Definitions

2.1. Full Field Objects

A set of full field objects is created for each observation that is processed through the Level 3 pipeline that is described in part V. The full field objects are identified by the observation identification (*ObsId*) and observation interval identification (*ObI*) of the observation [see section 1.2.1.1 of part I (Observation Identification)], and by cycle [see section 1.2.1.4.2 of part I (ACIS-specific Configuration Parameters)] for ACIS alternating exposure (interleaved) mode observations.

2.1.1. Full Field Event Object

The full field event object consists of single FITS format event file for each observation that results from Level 3 processing described in part V. The observation event file will have been reprocessed through

acis_process_events or **hrc_process_events** (as appropriate) to apply the latest instrument calibrations, and will have has the standard event status and event grade (ACIS only) filters applied. In addition, the ACIS-S4 CCD will have the **destreak** algorithm applied. Good time intervals will be revised to eliminate periods of background flares.

The full field event object FITS file shall be named

<i><s><obs>_<obi>N<v>[_<c>]_evt3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.1.1.1. Good Time Intervals

The good time intervals (GTIs) record the start and stop times (in MET seconds) of intervals during which time the quality of the event data is usable. The GTIs associated with the full field event object [see section 2.1.1 (Full Field Event Object)] are included as additional HDUs in the full field event object FITS format event file. There is a single GTI for an event file containing HRC data. For an event file containing ACIS data, there is a GTI for each ACIS CCD included in the full field observation.

2.1.2. Full Field Image Object

The full field image object consists of a single FITS format image file for each observation and science energy band. The full field image object consists of a set of image equivalents of the full field event object [section 2.1.1 (Full Field Event Object)] filtered by the appropriate science energy band, blocked at multiple resolutions in SKY coordinates, and background-subtracted. The images are recorded in flux units of photons s^{-1} cm^{-2} by subtracting the full field background image [see section 2.1.3 (Full Field Background Image Object)] in the appropriate science energy band from the filtered and blocked full field event object event data, and dividing the result by the full field exposure map [section 2.1.4 (Full Field Exposure Map Object)]. Each image shall be 2048 $\times$ 2048 blocked SKY pixels in size. The blocking factors shall match those

used in the source detection process (**wavdetect**), and are instrument-specific. They shall be 1, 2, and 4 for ACIS (-I or -S), 2, 5, and 12 for HRC-I, and 2, 5, 12, and 25 for HRC-S. The images shall be recorded in successive FITS HDUs, in order of decreasing field coverage.

The full field image object FITS file shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<b>_img3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<b>** is the energy band designation, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.1.2.1. Full Field Image

The full field image consists of a set of JPEG format image files for each observation. For ACIS observations, a multicolor images are generated by combining the image equivalents of the full field event object [section 2.1.1 (Full Field Event Object)] filtered by the soft, medium, and hard science energy bands, blocked at multiple resolutions in SKY coordinates, according to the algorithm of Lupton et al. (2004, PASP, 116, 133). For HRC observations, the images are generated from the image equivalent of the full field event object [section 2.1.1 (Full Field Event Object)] filtered by the wide science energy band, blocked at multiple resolutions in SKY coordinates. In both cases, the images are exposure corrected by dividing the filtered and blocked full field event object event data by the full field exposure map [section 2.1.4 (Full Field Exposure Map Object)].

The full field image JPEG files shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<f>_img3.jpg

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<c>** is the cycle, and **<f>** is the blocking factor. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations. Each image shall be 1024 × 1024 pixels in size.

2.1.3. Full Field Background Image Object

The full field background image object consists of a single FITS format image file for each observation and science energy band. The full field background image object consists of a set of background counts images in the appropriate science energy band, blocked at multiple resolutions in SKY coordinates. The image is recorded in units of counts. Each background counts image shall be 2048 × 2048 blocked SKY pixels in size. The blocking factors shall match those used for the corresponding images that comprise the full field image object [see section 2.1.2 (Full Field Image Object)]. The background counts images shall be recorded in successive FITS HDUs, in order of decreasing field coverage.

The full field background image object FITS file shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<b>_bkgimg3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<b>** is the energy band designation, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.1.4. Full Field Exposure Map Object

The full field exposure map object consists of single FITS format file for each observation and science energy band that results from Level 3 processing described in part V. A full field instrument map is generated at the equivalent monochromatic energy for each science energy band, and a set of exposure maps, blocked at multiple resolutions in SKY coordinates, are computed by applying the aspect histogram sampled at 0.5 arcsec resolution as described in part V. Each exposure map shall be 2048 × 2048 blocked SKY pixels in size. The blocking factors shall match those used for the corresponding images that comprise the full field image object [see section 2.1.2 (Full Field Image Object)]. The exposure maps shall be recorded in successive FITS HDUs, in order of decreasing field coverage.

The full field exposure map object FITS file shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<b>_exp3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<b>** is the energy band designation, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.1.5. Full Field Sensitivity Map Object

The full field sensitivity map object consists of single FITS format file for each observation and science energy band that records the limiting sensitivity for a point source to satisfy the minimum flux significance criterion specified in section 2.2.1 (Flux Significance) of part I. The limiting sensitivity shall be determined in units of photons s^{-1} cm^{-2}. The sensitivity map shall be 2048 × 2048 blocked SKY pixels in size, with blocking factor of 4 for ACIS (-I or -S), 12 for HRC-I, and 25 for HRC-S.

The full field sensitivity map object FITS file shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<b>_sens3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<b>** is the energy band designation, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.1.6. Aspect Histogram Object

The aspect histogram object consists of a single FITS format file for each observation that records the histogram of aspect pointing and roll offsets during the observation. The aspect histogram is used in the Level 3 pipeline that is described in part V to compute the response matrix averaged over the observation.

The aspect histogram object FITS files shall be named

<i><s><obs>_<obi>N<v>[_<c>]_ahst3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval

identification, **<v>** is the data product version number, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.1.7. Bad Pixels Object

The bad pixels object consists of a single FITS format file for each observation that records the bad pixels information for the observation. The bad pixels information is used in the Level 3 pipeline that is described in part V to flag for removal the effects of detector bad pixels during the observation.

The bad pixels object FITS files shall be named

<i><s><obs>_<obi>N<v>[_<c>]_bpix3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.1.8. Field Of View Object

The field of view object consists of a single FITS format file for each observation that records the region of the sky that was imaged by the detector during the observation, including the effects of spacecraft dither.

The field of view object FITS files shall be named

<i><s><obs>_<obi>N<v>[_<c>]_fov3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.2. Source Region Objects

A set of source region objects is created for each detected source region in each observation that is processed through the Level 3 pipeline that is described in part V . The source region objects are identified by the

region identifier [see section 1.2.2.1 of part I (Source Identification)], observation identification (*ObsId*) and observation interval identification (*ObI*) of the observation [see section 1.2.1.1 of part I (Observation Identification)], and by cycle [see section 1.2.1.4.2 of part I (ACIS-specific Configuration Parameters)] for ACIS alternating exposure (interleaved) mode observations.

2.2.1. Source Region Event Object

The source region event object consists of a single FITS format event file for each observation and detection source region. The source region event file is the result of applying a rectangular SKY region filter, that has a size equal to the bounding box of the background region of interest (ROI) corresponding to the detected source region, to the corresponding full field event object [see section 2.1.1 (Full Field Event Object)].

The source region event object FITS files shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<r>_regevt3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<r>** is the region identifier, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.2.1.1. Good Time Intervals

The good time intervals (GTIs) record the start and stop times (in MET seconds) of intervals during which time the quality of the event data is usable. The GTIs associated with the source region event object [see section 2.2.1 (Source Region Event Object)] are included as additional HDUs in the source region event object FITS format event file. There is a single GTI for an event file containing HRC data. For an event file containing ACIS data, there is a GTI for each ACIS CCD included in the source region.

2.2.2. Spectrum Object

2.2.2.1. PHA Spectrum Object

The PHA spectrum object consists of single FITS format file for each observation and detection source region that records the low resolution PI spectrum of the events extracted from the source region for an ACIS observation, resulting from Level 3 processing described in part V. The PI spectrum is constructed using weighted responses for the areas of the detector over which the source region dithers. A low resolution PI spectrum of the events extracted from the background region is recorded in the FITS HDU immediately following the source PHA spectrum HDUs.

The PHA spectrum object FITS file shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<r>_pha3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<r>** is the region identifier, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.2.2.2. Auxiliary Response File Object

The auxiliary response file object consists of single FITS format file for each observation and detection source region that records the weighted ARF for the areas of the detector over which the source region for an observation dithers.

The auxiliary response file object FITS file shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<r>_arf3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<r>** is the region identifier, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.2.2.3. Redistribution Matrix File Object

The redistribution matrix file object consists of single FITS format file for each observation and detection source region that records the weighted RMF for the areas of the detector over which the source region for an ACIS observation dithers.

The redistribution matrix file object FITS file shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<r>_rmf3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<r>** is the region identifier, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.2.3. Source Region Image Object

The source region image objects consist of a set of FITS format image files, one for each science energy band, for each observation and detection source region. Each source region image object is the image equivalent of the source region event object [section 2.2.1 (Source Region Event Object)] filtered by the appropriate science energy band, and blocked in SKY coordinates at the lowest blocking factor (i.e., highest spatial resolution) used in the source detection process (**wavdetect**) that includes the location of the source region. The images are recorded in flux units of photons s^{-1} cm^{-2} by dividing the filtered and blocked source region event object event data by the source region exposure map [section 2.2.4 (Source Region Exposure Map Object)].

The source region image object FITS files shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<r><b>_regimg3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<r>** is the region identifier, **<b>** is the energy band designation, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.2.3.1. Source Region Image

The source region images consist of a set of one or more JPEG format image files for each observation and detection source region. For both ACIS and HRC observations, single band source region images are generated from the image equivalent of the source region event object [section 2.2.1 (Source Region Event Object)] filtered by each of the science energy bands defined for the instrument, and blocked in SKY coordinates at the lowest blocking factor (i.e., highest spatial resolution) used in the source detection process (**wavdetect**) that includes the location of the source region. In addition, for ACIS, a multicolor source region image is generated by combining the image equivalents of the source region event object [section 2.2.1 (Source Region Event Object)] filtered by the soft, medium, and hard science energy bands, blocked in SKY coordinates at the same resolution as the corresponding single band source region images, according to the algorithm of Lupton et al. (2004, PASP, 116, 133). In all cases, the image is exposure corrected by dividing the filtered and blocked source region event object event data by the source region exposure map [section 2.2.4 (Source Region Exposure Map Object)].

The single band source region image JPEG files shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<r><b>_regimg3.jpg

and the multicolor source region image JPEG file shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<r>_reg3img3.jpg

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<r>** is the region identifier, **<b>** is the energy band designation, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.2.4. Source Region Exposure Map Object

The source region exposure map object consists of single FITS format event file for each observation, detection source region, and science energy band that results from Level 3 processing described in part V. A full resolution instrument map that covers the detected source region

bounding box is generated at the equivalent monochromatic energy for each energy band, and the exposure map is computed by applying the aspect histogram sampled at 0.5 arcsec resolution and blocking in SKY coordinates, as described in part V. The blocking factor shall match that used for the corresponding full field image object [see section 2.2.3 (Source Region Image Object)].

The source region exposure map object FITS files shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<r><b>_regexp3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<r>** is the region identifier, **<b>** is the energy band designation, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.2.5. Point Spread Function Object

The point spread function objects consist of a set of FITS format image files, one for each science energy band, for each observation and detection source region. Each point spread function object records the normalized PSF computed at the equivalent monochromatic energy of the appropriate science energy band, sampled on the same pixel grid as the source region image object [see section 2.2.3 (Source Region Image Object)].

The point spread function object FITS files shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<r><b>_psf3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<r>** is the region identifier, **<b>** is the energy band designation, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.2.5.1. Point Spread Function Image

The point spread function images consist of a set of one or more JPEG format image files for each observation and detection source region. Point

spread function images are generated from the point spread function objects [section 2.2.5 (Point Spread Function Object)] in each of the science energy bands defined for the instrument, sampled on the same pixel grid as the source region image [see section 2.2.3.1 (Source Region Image)].

The point spread function image JPEG files shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<r><b>_psf3.jpg

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<r>** is the region identifier, **<b>** is the energy band designation, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.2.6. Light Curve Object

The light curve objects consist of a set of FITS format files, one for each science energy band, for each observation and detection source region. Each light curve object records the the multi-resolution light curve output by the Gregory-Loredo analysis of the receipt times of the source events within the source region. A background light curve with identical time-binning to the source light curve is derived from an analysis of the events within the background region. The background light curve is recorded in the FITS HDU immediately following the source light curve HDUs. A detailed description of the computation of the multi-resolution light curve is documented in Appendix F.

The light curve object FITS files shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<r><b>_lc3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<r>** is the region identifier, **<b>** is the energy band designation, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

2.2.7. Region Object

The region object consists of a single FITS format region file for each observation and detection source region that records the region description for the modified source region and the corresponding modified background region [see section 1.2.2.5.1.1 (Modified Source Region) of part I (Catalog Requirements)], in consecutive HDUs.

The region object FITS file shall be named

<i><s><obs>_<obi>N<v>[_<c>]_<r>_reg3.fits

where **<i>** is the instrument designation, **<s>** is the data source, **<obs>** is the observation identification, **<obi>** is the observation interval identification, **<v>** is the data product version number, **<r>** is the region identifier, and **<c>** is the cycle. The optional discriminator identified in square brackets is included only for ACIS alternating exposure (interleaved) mode observations.

IV. Non-Catalog Deliverables Requirements

1. Sensitivity Analysis Tool

The sensitivity analysis tool described in these requirements shall be included in the *CIAO* data analysis system.

1.1. Limiting Sensitivity

The sensitivity analysis tool shall provide the means to allow the user to determine the limiting sensitivity for including a point source in the catalog (part I) or, optionally, in the database (part II), in any source detection energy band or science energy band at a user-specified location that is included in the field of view of one or more observations that is included in the catalog.

The limiting sensitivity shall be determined in units of photons s^{-1} cm^{-2}, unless the user specifies a source spectrum. In the latter case, the limiting sensitivity shall be determined in units of ergs s^{-1} cm^{-2}.

1.2. Source Flux Upper Limit

The sensitivity analysis tool shall provide the means to allow the user to determine the flux upper limit in any source detection energy band or science energy band above a local user-specified background region for a user-specified source region centered at a user-specified location that is included in the field of view of one or more observations that is included either in the catalog (part I) or, optionally, in the database (part II). The user-specified source and background regions shall be restricted to be no larger than TBD.

The source flux upper limit shall be determined in units of photons s^{-1} cm^{-2}, unless the user specifies a source spectrum. In the latter case, the source flux upper limit shall be determined in units of ergs s^{-1} cm^{-2}.

2. Mosaics

2.1. Sensitivity Mosaic

2.2. Sky Mosaics

V. Data Processing Requirements

1. Per-Observation Pipelines

    1.1. Source Detection Pipeline

        1.1.1. Observation Calibration

        1.1.2. Data Cleaning

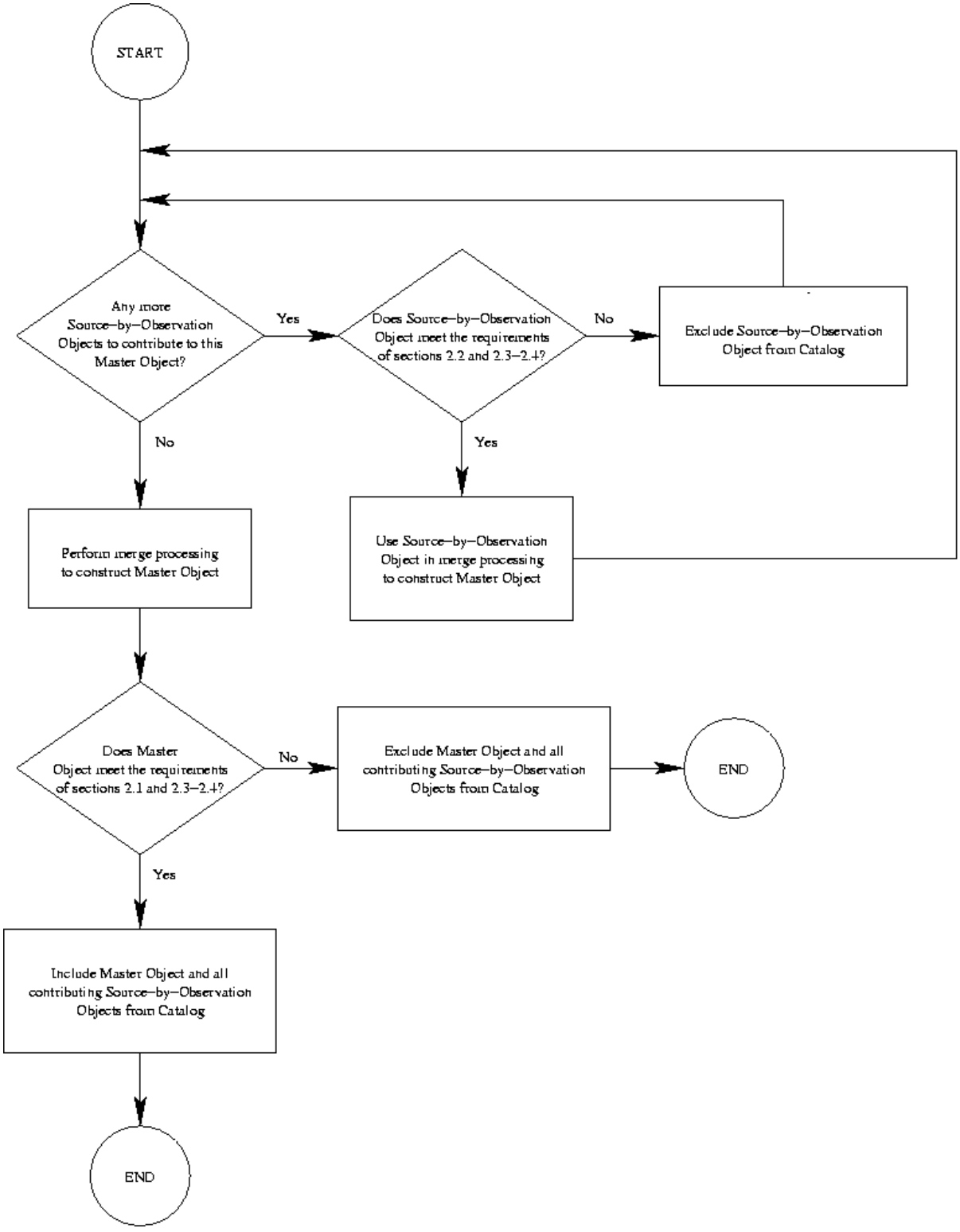


Figure V.2.1: High level flow used for determining which Master objects and Source-by-Observation Objects are included in the catalog, and which Source-by-Observation objects contribute to the merge processing that generates the catalog version of a Master Object.

VI. Quality Assurance Requirements

1. Automated Quality Assurance

    1.1. Validation Steps

2. Manual Quality Assurance

    2.1. Trigger Criteria

    2.2. Manual Actions

VII. User Interface Requirements

1. Basic Architecture

1.1. User Catalog Model

The user's view of the catalog shall be modeled as two tables (views), one for the Master Source Catalog (MSC) and one for the Per-Observation Source Catalog (OSC), where each row represents a source and each column a quantity or parameter that is officially part of the catalog; in addition, there will be columns containing (links to): included data objects; references to the sister catalog; TBD objects and services. Choice of celestial coordinate systems shall be restricted ICRS and Galactic.

1.2. Interface Components

There are four main functional components to the interface. The first three MUST have a web-based version, although it is conceivable that for full functionality of the third one a down-loadable version MAY be preferable for professional users.

1.2.1. Query Interface

The query interface allows users to make database queries to the catalog tables. It SHALL work by transmitting an SQL or SQL-like query through a GET and/or POST operation. The objective is that this interface:

- allow queries to be initiated from a web browser, as well as from the command line
- allow saving, restoring, and editing of queries
- become compatible with the IVOA Table Access Protocol (TAP) that is based on the Astronomical Data Query Language (ADQL)

It is not to say that users need to know and use SQL/ADQL; see below [section 2.1.1 (Query Interface)]. All queries MUST be logged.

1.2.2. Query Results Interface

Results SHALL be returned, at the user's discretion, in HTML or RDB. RDB is sufficiently general that it satisfies the needs of users who want ASCII or tab-delimited tables, and has the advantage that it is easily read

into databases and spreadsheets; in particular, this means that such tables can also be used as uploaded user files. The entries SHALL contain links that allow inspection of individual sources [section 1.2.3 (Individual Source Interface)]. In addition, the user SHALL be able to request selected data objects to be returned collectively in a tar file or individually through links in results tables.

1.2.3. Individual Source Interface

For inspecting individual source properties more closely, there SHALL be an interface, activated from the Query Results Interface as a separate window or as a stand-alone down-loadable tool, that presents all known properties of a single source, including links to related sources (between MSC and OSC), images and plots of associated data objects, full-field images, other services, etc. In this interface the user SHALL have access to bi-directional links between source records and full-field, exposure-corrected, background-subtracted images.

1.2.4. API

The API SHALL allow client-based tools to access the catalogs and include the functionality of the first three interfaces.

1.3. Cross-correlation

Cross-correlation of catalogs (or joins between tables), whether they be residing elsewhere on the Internet or provided by the client, is an essential part of catalog-based research. There are three levels involving increasing complexity: between MSC and OSC; between CSC and a client-provided table; and between CSC and one or more catalogs on the internet (*e.g.*, in Vizier). We hope to be able to rely on VO services for this last type of cross-correlations.

2. First Release Minimum Requirements

The UI properties listed in this section MUST be included in the first release of the User Interface.

2.1. Interfaces

2.1.1. Query Interface

The CSC MUST support a CGI-like interface that allows clients to provide an SQL query through GET and/or POST. Ideally, the SQL SHOULD be ADQL compliant, but for the first release it SHALL be sufficient to implement the subset contained in Sybase's SQL, allowing SELECT, FROM, WHERE, ORDER. The interface MUST also allow specification of output format. It MUST check the query and return a friendly error message if the query is not valid or not legal. This will allow command-line queries. There MUST be a web interface that allows clients to:

- construct SQL queries in a user-friendly manner (*i.e.*, it should not look like SQL)
- save queries locally in SQL form and ASCII format
- restore saved queries
- submit a query

This interface MUST relay error messages, if any, and MAY provide feedback on estimated response time or other information of interest to the client. All queries MUST be logged, retaining relevant information such as the query itself, date and time, IP address, success or failure, size of response, response time.

2.1.2. Query Results Interface

As described in section 1.2.2 (Query Results Interface), results SHALL be returned, at the user's discretion, in HTML or RDB. In addition, the user SHALL be able to request selected data objects to be returned collectively in a tar file or individually through links in results tables. Data object columns that have jpeg versions MAY display those postage stamp images to represent the links. The entries SHALL contain links that allow inspection of individual sources [section 2.1.3 (Individual Source Interface)].

2.1.3. Individual Source Interface

TBD.

2.1.4. API

Software tools will be able to interact with the catalog (query and retrieve data objects) through a combination of the query interface [section 2.1.1 (Query Interface)] and the results interface [section 2.1.2 (Query Results Interface)], using RDB format. This is somewhat crude, but one could well imagine a general-purpose wrapper to be written that will make this task less odious at the application level.

3. Highly Desirable Properties

The features described in this section are highly desirable. They should be included in the first release if they are relatively simple to implement with little effort.

3.1. Interface Components

3.1.1. Query Interface

3.1.1.1. Upper Limits

The user SHOULD have access to sensitivity information and be able to derive upper limits under specific (TBD) constraints.

3.1.1.2. Name Resolvers

The query interface SHOULD provide access to name resolvers for an auxiliary query. The supported name resolvers SHALL be SIMBAD and NED. The user SHOULD be able to select either or both, and, in the latter case, the order in which they are consulted.

3.1.2. Query Results Interface

3.1.2.1. Output Formats

VOTable SHOULD be added to the output options.

3.1.2.2. Linking to External Services

The UI SHOULD allow the option of linking results to DataScope or a similar external service. DataScope allows querying astronomical

archives (images, catalogs) for information pertinent to a particular position; what is involved is inserting the source's position into the DataScope URL.

3.2. Cross-correlation

3.2.1. User-Provided Cross-match Data

The user SHOULD be able to upload a table of positions and, optionally, errors for a cross-match query. RDB SHALL be an allowed format. The interface MAY allow cross-matching on more parameters.

4. Longer Term Requirements

The features in this section have a lower priority at this time. Presumably, the order of implementation will largely be determined by the amount of effort required and the emergence of new capabilities. It is also expected that this list will be subject to revision due to experience gained and external developments.

4.1. Interface Standards

4.1.1. ADQL Compatibility

Expand the interface to the database to full ADQL compliance. This entails in particular the implementation of ADQL-specific functions such as XMATCH and REGION, and ADQL equation scripting.

4.1.2. API

The users shall be provided an API access capability allowing standard web service access to the catalog and data objects.

4.2. Interface Formats

4.2.1. Server Access

Ongoing development, TBD.

4.2.2. Output Formats

Enhancements TBD.

4.3. Functionality

4.3.1. Integration of Internal Services

Provide links to Chandra proposal abstracts and bibliography database.

4.3.2. Integration of External Services

Integrate linkage to NED, SIMBAD, Vizier, USNO-B.

4.3.3. Cross-Matching

Provide cross-matching functionality against external catalogs. If the interface is ADQL-compliant, it should be possible to achieve this functionality through VO services like Open SkyNode with little effort.

4.3.4. Links to VO Applications

The user shall be able to load data from selected results columns into VO tools like VOPlot and TOPCAT.

4.3.5. Integration of User-Provided Data

The user shall be able to cross-match queries with user-provided (uploaded) catalogs and to augment returned results with such catalog data.

4.3.6. Custom Processing

The user shall be able to request flux densities to be returned for a user-defined band, using event or pha data.

4.3.7. Versioning

The user shall be able to query previous versions of the catalog. This capability shall include full traceability of source identity through catalog versions.

Appendix A

Chandra Source Catalog Energy Bands

Michael L. McCollough, CXC/SAO

January 26, 2007

Abstract: Given below is the rational for deciding the energy bands and their associated effective energy (mono-energy used for exposure maps, etc.) to be used in Chandra Source Catalog. A brief discussion of the problems and issues that needed to be addressed are given.

I. Energy Bands

For the Chandra Source Catalog (CSC) we desire to find the optimal energy bands to use for source detection and flux measurements. For ACIS and HRC we need a broad bands which will cover the entire range in energy to which the telescope and detectors are sensitive. Additional for ACIS observations, for which there are spectral resolution capabilities, it is also desirable to have multiple narrow bands. Based on prior experience with the data it was decided that there was adequate spectral resolution and sensitivity for three additional bands (soft, medium, and hard) which would also be used for images, source detection, flux measurements and hardness ratios. An additional fourth band for super soft sources was also found to be desirable. Below is a review of how the energy boundaries of the various bands were determined.

A. Energy Bands for ACIS

For ACIS several issues need to considered in determining the bands to be used. Given below is a summary of the initial inputs, previous work, telescope and detector response information, and rational used in deciding the energy boundaries.

1. Initial Issues and suggestions:

Below are a list of initial issues and suggestions which were used to start the evaluation process for the energy boundaries to be used for ACIS CSC catalog entries.

a. Effective Area: There is a need to avoid large changes in the effective area (in the HRMA and detectors) in the middle of the bandpass whenever possible. Such changes can possibly have a negative impact on calculating such things as exposure maps. In some cases (the broad band and soft band) this may not be possible.

b. Signal vs. Noise: There is a need to avoid extending the bandpasses too low (soft) or too high (hard). Integrating the bandpass where there is no signal has the risk of adding noise to the measurement while not adding any signal.

c. Iridium Edge: There is an iridium M-edge in the 2–2.5 keV area. This would make a natural break

between the medium and hard bands.

d. Chip Differences: The front illuminated ACIS chips have very little sensitivity below 0.3 keV and the back illuminated ACIS chips go down to about 0.1 keV. Possibly a compromise between the two is desirable.

e. High End: It is likely that the hard (and broad) band should go out to at least 7 keV (to include the Fe K lines). But by the time you get to 10 keV you have very little signal and are likely just adding noise. Some compromise between these energies is desirable.

2. Previously used bands:

In the table below is a summary of the bands that have been used at the CXC and were found in a brief survey (10 different articles in ApJ) of the literature.

Survey	***ChaMP***	***Antennae XRBs***	***Antennae Soft Diffuse***	***Most Common (Various ApJ)***
Soft	0.3–0.9 keV	0.3–1.0 keV	0.3–0.65 keV	0.3–1.0 keV
Medium	0.9–2.5 keV	1.0–2.5 keV	0.65–1.5 keV	1.0–2.0 keV
Hard	> 2.5 keV	2.5–7.0 keV	1.5–6.0 keV	2.0–7.0 keV
Broad		0.3–7.0 keV		

In the table below is the range of energy boundaries found in the various ApJ articles.

Soft (lower)	***Soft/Medium***	***Medium/Hard***	***Hard (upper)***
0.1–0.3 keV	0.5–1.1 keV	2.0 keV	6.0–8.0 keV

3. XMM bands:

For CSC it should be noted that the shape of the XMM effective area curve is very similar to Chandra's up to ~6 keV. The XMM mirrors has gold edges while Chandra has Iridium, but they are quite close in energy (e.g. 2.0 keV vs. 2.2 keV; the edges are quite complex and spread out). Thus similar energy bands to those used by XMM may expect and desirable for the CSC.

a. XMM current source catalog (Serendipitous Source Catalogue: 1XMM):

This catalog uses the following energy bands bands:

i. Basic energy bands:

0.2–0.5 keV

0.5–2.0 keV

2.0–4.5 keV

4.5–7.5 keV

7.5–12.0 keV

ii. Broad energy bands:

0.2–2.0 keV

2.0–12.0 keV

0.2–12.0 keV

0.5–4.5 keV

(see http://xmmssc-www.star.le.ac.uk/newpages/UserGuide_1xmm.html#TabBands)

b. XMM's next version of a source catalog (2XMM):

Will uses the following energy bands bands:

i. Basic energy bands:

0.2–0.5 keV

0.5–1.0 keV

1.0–2.0 keV

2.0–4.5 keV

4.5–12.0 keV

b. Broad energy bands:

0.2–12.0 keV

(private communication: Clive Page)

4. Effective area, quantum efficiency, and contamination:

In determining the energy bands it is important to examine how the effective area and quantum efficiency of the detector vary as a function of energy. One also has to address the issue of contamination of ACIS and how it impacts the sensitivity of the the detector as a function of energy and time. A secondary issue is the role the spectrum of the observed sources may play in determining the bands to be used.

a. Effective area and quantum efficiency: In Fig. 1 is a plot of the product of the effective area with the quantum efficiency of a back illuminated chip (S3: solid line) and a front illuminated chip (I3: dashed line). One can note the following:

i. One can see the edges due to Iridium with the largest drop occurring around 2 keV.

ii. There is a strong C-K edge at around 0.3 keV.

iii. One can see the marked difference to response between the back illuminated and front illuminated chips below 1 keV.

b. Contamination: In Fig.2 are the same curves as Fig.1 except they have been multiplied by a transmission factor which is determined by the amount of contamination on a given chip. The important things to note are the very deep C edge and the large overall lost in response below 1 keV (especially in the back illuminated chip). One should keep in mind that the L3 catalog will span the entire range of contamination from none to the most recently measured value.

c. Spectral weighting: In Figs.3 and 4 are the same curves as Fig.1 and 2 respectively, except they have been multiplied by a spectral weighting function of the form:

$$\left(\frac{E}{E_0} \right)^{-\alpha}$$

where E is the energy, E_0 is the normalization energy (taken to be 1 keV here), and α is a power law index which is taken to be 1.0 for typical objects observed with Chandra. This weighting will give a better sense of the contribution of Chandra sources various regions of the spectrum. The most notable effect is increase the effective areas at low energy and decrease the high energy cutoff to around 7.5 keV.

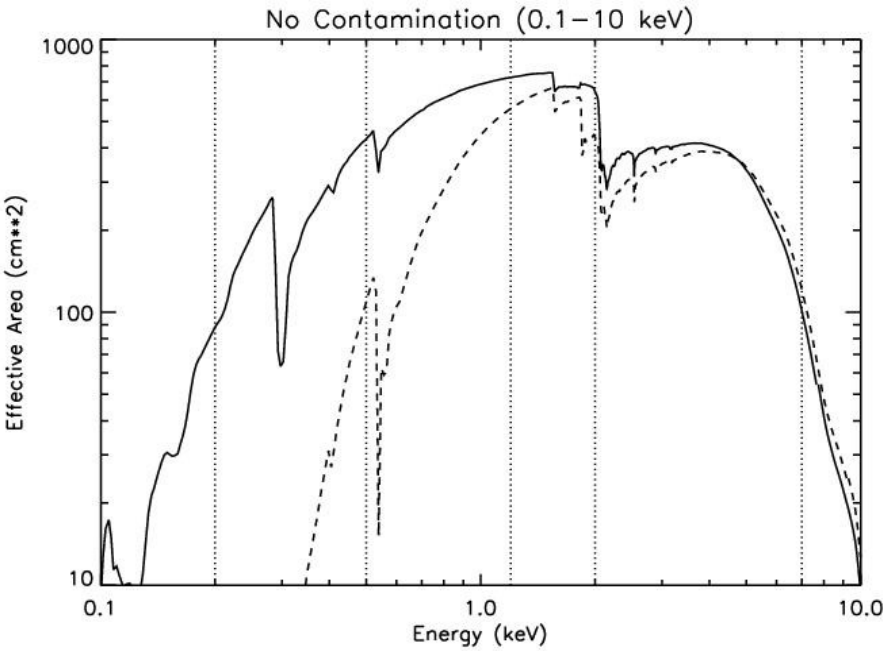


Fig. 1: This is a plot of the product of the effective area of the telescope with quantum efficiency of the detector. The solid line is the back illuminated chip (S3) and the dashed line is the front illuminated chip (I3). The dotted lines are the energy boundaries of the CSC bands.

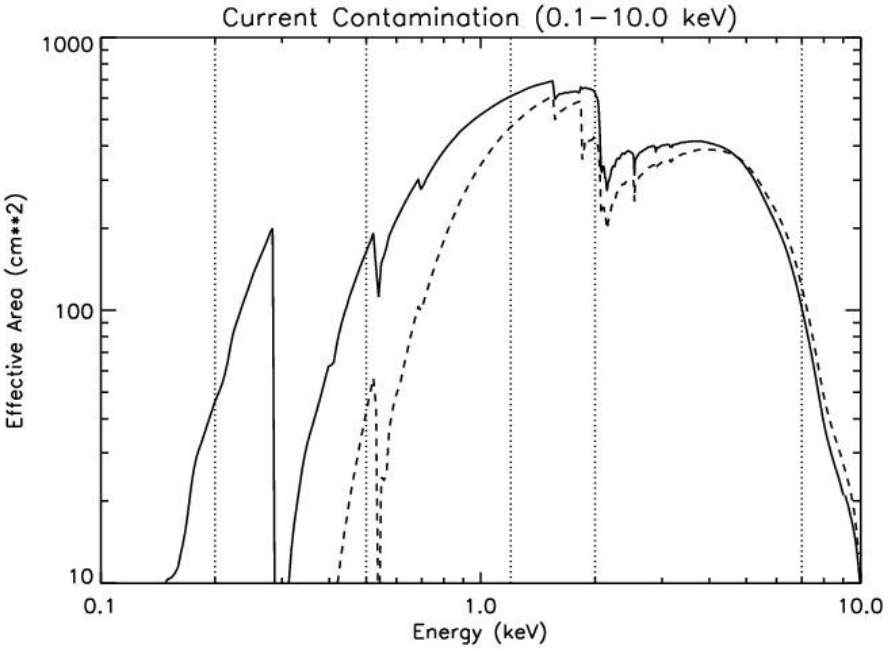


Fig. 2: This is a plot of the product of the effective area of the telescope with quantum efficiency of the detector with the current reduction due contamination included. The solid line is the back illuminated chip (S3) and the dashed line is the front illuminated chip (I3). The dotted lines are the energy boundaries of the CSC bands.

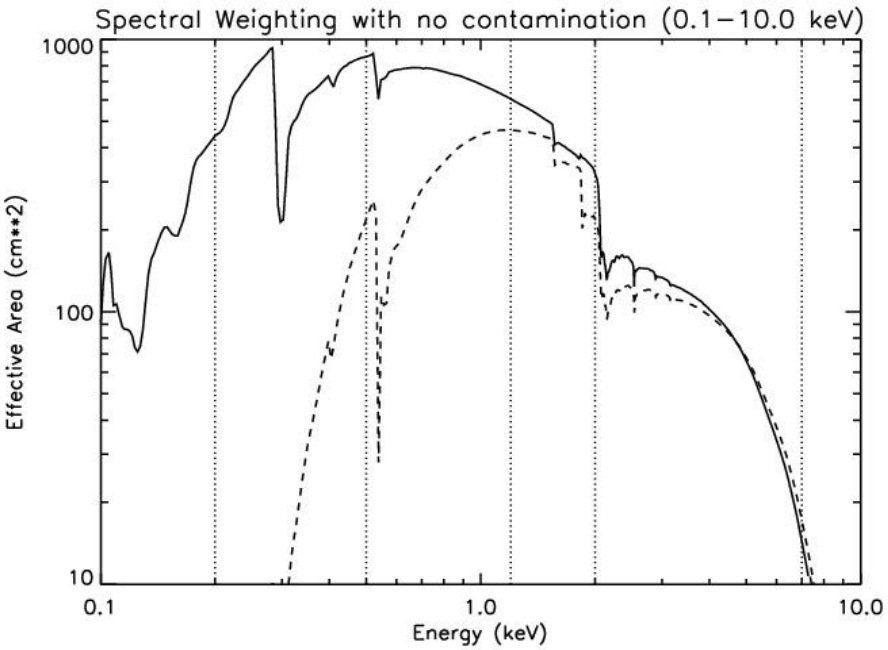


Fig. 3: This is a plot of the product of the effective area of the telescope with quantum efficiency of the detector times a spectral weighting function (see text). The solid line is the back illuminated chip (S3) and the dashed line is the front illuminated chip (I3). The dotted lines are the energy boundaries of the CSC bands.

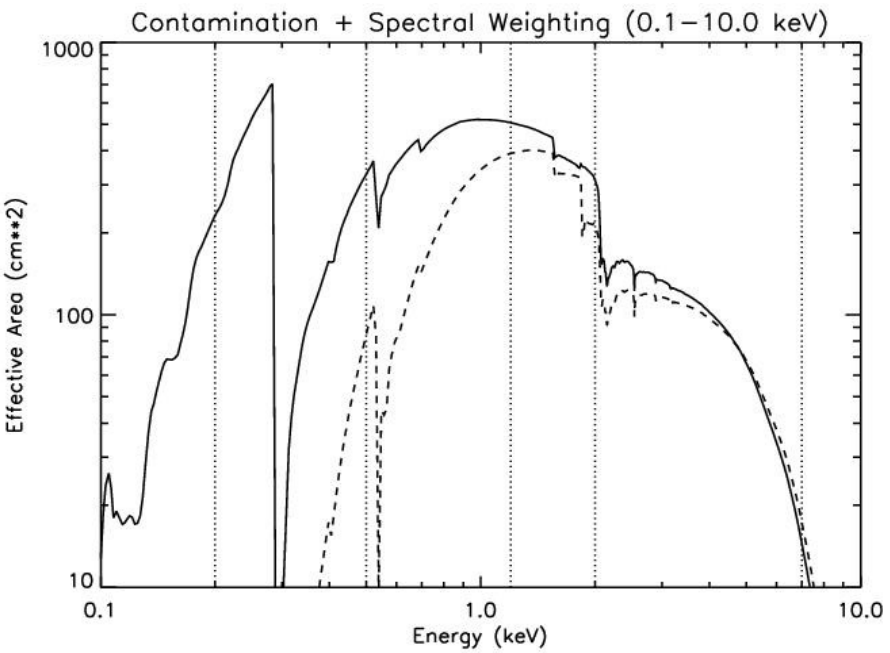


Fig. 4: This is a plot of the product of the effective area of the telescope with quantum efficiency of the detector times a spectral weighting function (see text) with the current reduction due contamination included. The solid line is the back illuminated chip (S3) and the dashed line is the front illuminated chip (I3). The dotted lines are the energy boundaries of the CSC bands.

5. Synthetic Color-Color Plots:

To better access which energy bands should be used, for the CSC, synthetic color-color plots were created using PIMMS. (The initial scripts for these simulations were provided by A. Tennant.) Three models were used:

(1) absorbed Power Law;

(2) absorbed Raymond-Smith;

(3) absorbed Blackbody.

What was sought was a set of bands which, for a reasonable range of parameters, would fill the color-color plot and hence serve as a useful diagnostic in understanding the nature of the sources being detected. This also would allow one to differentiate between various spectral models. For the color-color plots three bands were used: hard (H); medium (M); and soft (S). There was also a total (T) band which corresponds to the broad band to be used. The colors were defined as:

$$\frac{(M-S)}{T} \quad \text{(soft color)}$$

$$\frac{(H-M)}{T} \quad \text{(hard color)}$$

For the simulations runs both the I3 and S3 chip were run so that differences of the front and back illuminated chips could be evaluated.

a. Simulation Results: Several different sets of bands were run and the results examined. The following conclusions were reached:

i. An initial set of bands which include a very soft band (0.2–0.5 keV) were a poor diagnostic for most of the likely sources to found in the CSC. There were questions about the calibration in this soft band as well. For this reason the lower bound was taken to be 0.5 keV. But see section b. below for a further discussion of this issue.

ii. Because of the strong change in effective area at 2.0 keV this was a natural break for the M to H band.

iii. To determine break between the S to M we did several different simulations. The results were a break at 1.2 keV which tends to optimization the desire to fill the color-color plot (as a function of parameter space) while still having close to an equal amount of effective area in each band. See section 6 for more of a discussion of this point.

iv. We also did several simulations to evaluate the high energy cutoff. It was desirable to go up to at least 7.0 keV to include the Fe complex found in many X-ray sources in 6.4–6.9 keV region. There were some concerns that other lines above 7.0 keV might contribute to the hard band. But various simulations showed any such contribution was negligible. Thus an upper bound of 7.0 keV was chosen.

The simulated color-color plots of the chosen bands are shown in Figs. 5–7 for the ACIS-S detector.

b. Super Soft Sources: There was still a concern about the detection and analysis of Super Soft Sources. For this reason simulations were run for a low temperature Blackbody using the 0.2–0.5 keV band along with the new S and M bands. In this case the total (T) band is from 0.2–2.0 keV. The results for ACIS-S are shown in Fig. 8. It appears that this band could be used in the search and analysis of Super Soft Sources. It was decided that this Ultra-Soft (U) band would be useful and should be kept.

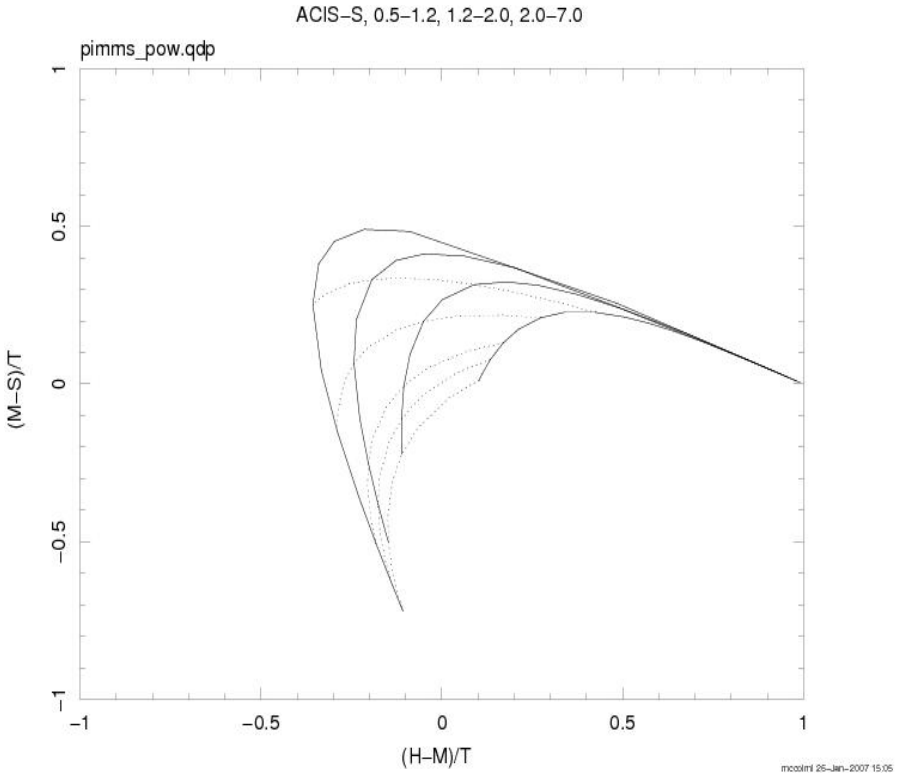


Fig. 5: This a synthetic color-color plot made for an absorbed power law model using PIMMS for the chosen CSC energy bands. The fixed spectral Index values used were 1, 2, 3, and 4. The fixed values of N_h *used were* $1.0\times10^{20}\,cm^{-2}$, $1.0\times10^{21}\,cm^{-2}$, $2.0\times10^{21}\,cm^{-2}$, $5.0\times10^{21}\,cm^{-2}$ *, and* $1.0\times10^{22}\,cm^{-2}$. *The solid lines represent lines constant spectral index and the dotted lines are lines of constant* N_h .

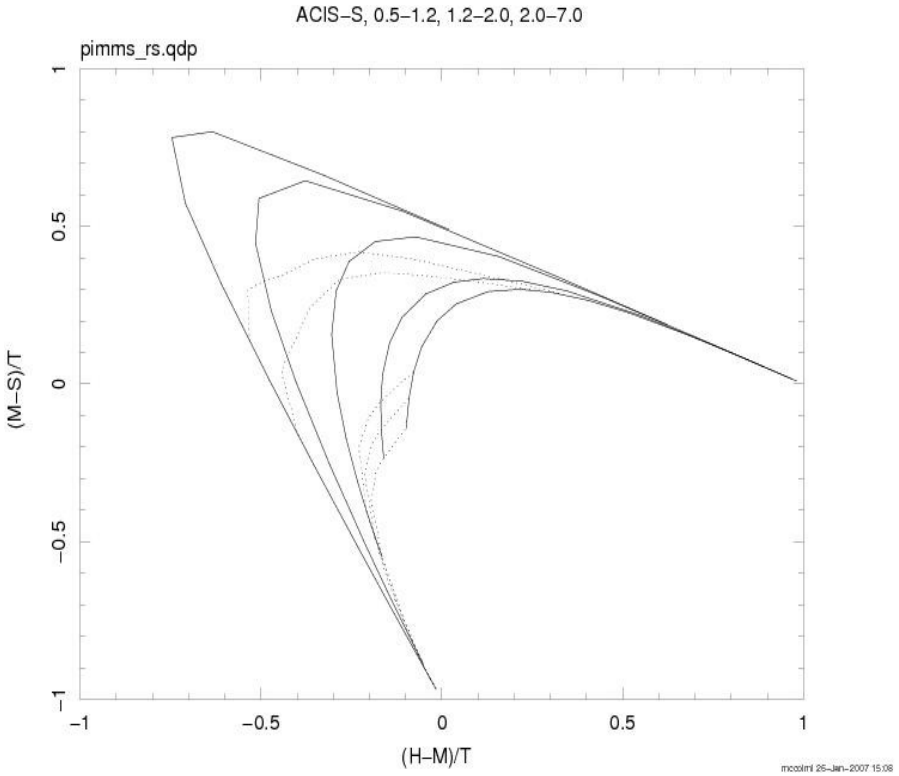


Fig. 6: This a synthetic color-color plot made for an absorbed Raymond-Smith model using PIMMS for the chosen CSC energy bands. The fixed temperature values used were 0.25, 0.5, 1, 2, and 4 (in keV) . The fixed values of N_h used were $1.0\times10^{20}\,cm^{-2}$, $1.0\times10^{21}\,cm^{-2}$, $2.0\times10^{21}\,cm^{-2}$, $1.4\times10^{22}\,cm^{-2}$, and $1.75\times10^{22}\,cm^{-2}$. The solid lines represent lines constant temperature and the dotted lines are lines of constant N_h .

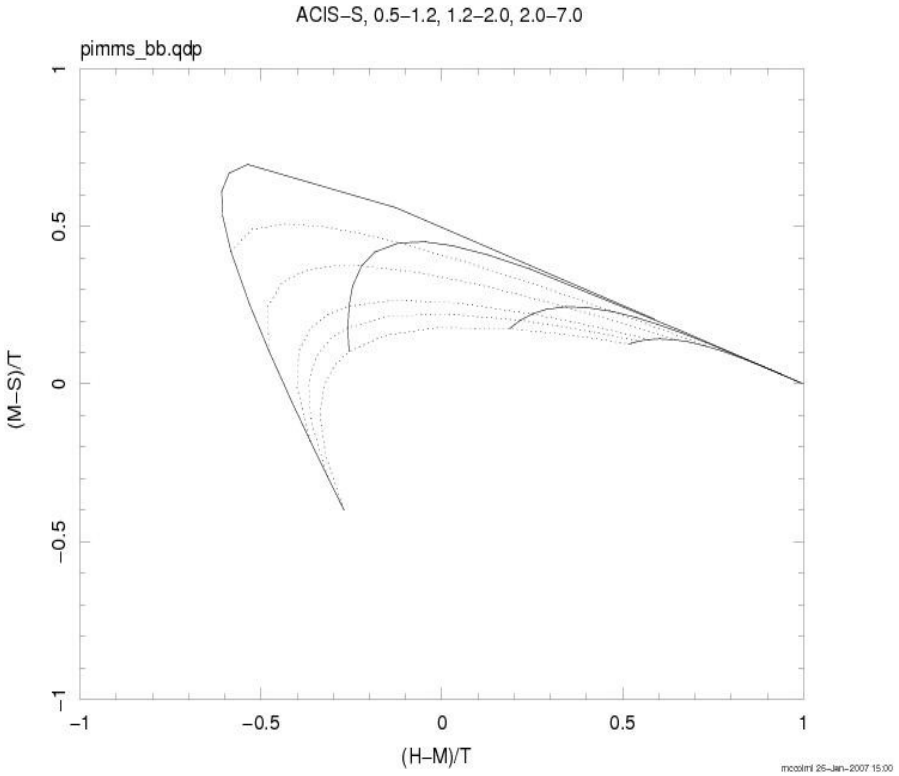


Fig. 7: This a synthetic color-color plot made for an absorbed Blackbody model using PIMMS for the chosen CSC energy bands. The fixed temperature values used were 0.25, 0.5, 1, and 2 (in keV) . The fixed values of N_h *used were* $1.0\times10^{20}\,cm^{-2}$, $1.0\times10^{21}\,cm^{-2}$, $2.0\times10^{21}\,cm^{-2}$, $5.0\times10^{21}\,cm^{-2}$, and $1.0\times10^{22}\,cm^{-2}$. The solid lines represent lines constant temperature and the dotted lines are lines of constant N_h .

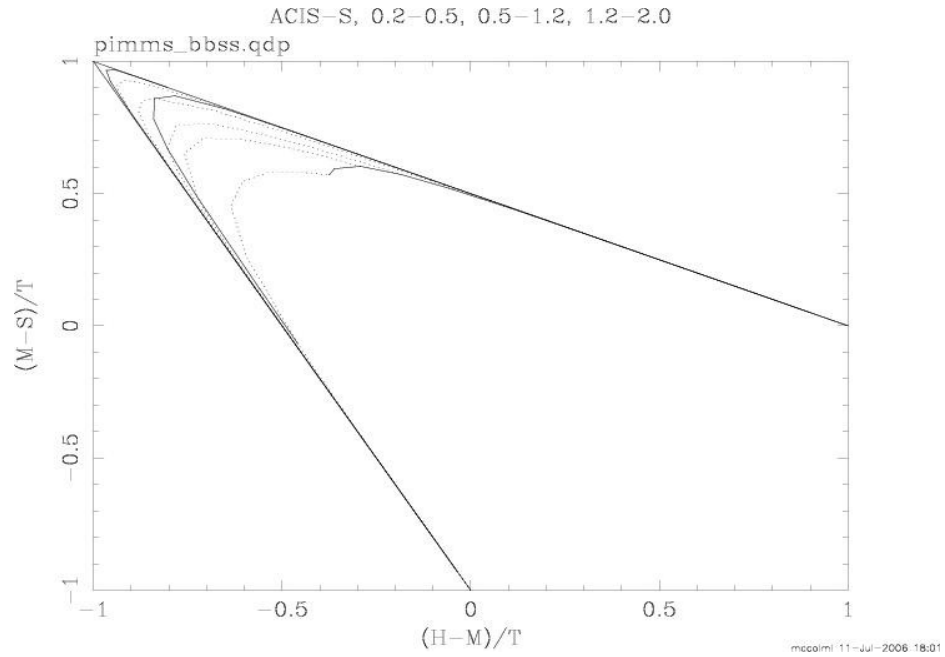


Fig. 8: This a synthetic color-color plot made for an absorbed Blackbody model using PIMMS using the Ultra-Soft, Soft, and Medium CSC bands. The fixed temperature values used were 20, 50, 100, and 250 (in eV) . The fixed values of N_h used were $1.0\times10^{20}\,cm^{-2}$, $1.0\times10^{21}\,cm^{-2}$, $2.0\times10^{21}\,cm^{-2}$, $5.0\times10^{21}\,cm^{-2}$, and $1.0\times10^{22}\,cm^{-2}$. The solid lines represent lines constant temperature and the dotted lines are lines of constant N_h .

6. Equal Effective Area:

To access where to put the break energy for the S to M bands it was found desirable that each band have roughly the same effective area. To estimate the energy for this we summed QE weighted effective area for the 0.5–2.0 keV band. We then created two bands (low and high) for which summed QE weighted effective area was calculated and divided by the total for the 0.5–2.0 keV band. This is plotted in Fig. 9 as a function of break energy between the two bands for both ACIS-I and ACIS-S. The break energy of 1.2 keV fits the criteria of have relatively equal effective areas between the two bands.

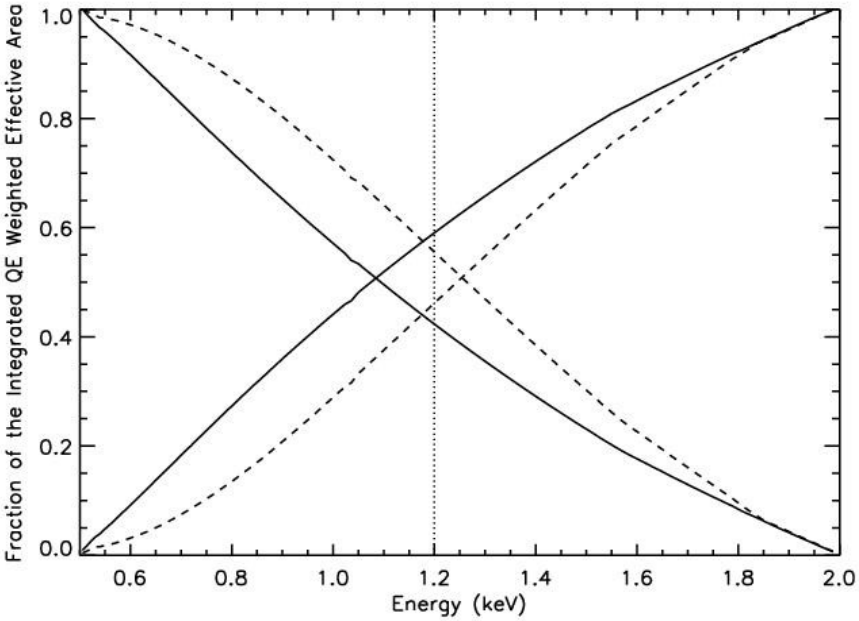


Fig. 9: To estimate the Soft to Medium band break energy for the 0.5–2.0 keV band the QE weighted effective area was summed for the band. Then two bands (low and high) were created for which summed QE weighted effective area was calculated and divided the total for the entire band. This represent the fraction of the total effective area for the band contained in each of these sub bands. These bands are plotted as function of break energy. The solid lines are for an ACIS-S chip and the dashed lines are for an ACIS-I chip. The dotted line represents the Soft to Medium break energy chosen for the CSC.

7. CSC Energy Bands:

Based on the above sections we have chosen the bands given in the table below to use for CSC. The rational follows:

Bands	***Broad***	***Ultra-Soft***	***Soft***	***Medium***	***Hard***
Energies	0.5–7.0 keV	0.2–0.5 keV	0.5–1.2 keV	1.2–2.0 keV	2.0–7.0 keV

a. Broad (0.5–7.0 keV): The simulations indicate that bands going below 0.5 keV does not serve as a useful diagnostic for most sources likely to be found in the CSC. As a result a value of 0.5 keV is recommend for the low end of the broad band. At the high end there is a rapid fall off. One needs to go to at least 7.0 keV to get the Fe K lines. Simulations indicate that very little if any useful information is gained by going about 7.0 keV. In fact increased background included in the band by going to higher energies may prove detrimental. As a result the upper cutoff was chosen to be 7.0 keV.

b. Ultra-Soft (0.2–0.5 keV): It was felt that for very soft sources that an additional low energy band was needed. For the front illuminated chips there is virtually no response below 0.3 (contamination only makes this worse). For the back illuminated chips there is response out to 0.1 keV but this is also

greatly curtailed by contamination. Thus a compromise value of 0.2 keV was chosen. Since this band will be used to search for soft sources we do not want to go too high in energy and a value 0.5 keV was chosen to match the soft band boundary. This value will be just below the O-K edge. The only real issue with this band is that the C-K edge will be in the middle of this band and will be deep in contaminated observations.

c. Soft (0.5–1.2 keV): The lower bound is set as noted above. The break between was chosen to be an optimization between equal effective area and fill the color-color plot as a function of parameter space. As a result a value of 1.2 keV was chosen.

d. Medium (1.2–2.0 keV): The low bound is set as noted above. The natural upper boundary is a 2.0 keV. One could make arguments for going 0.5 keV either side of this value. But making it 2.0 keV appears a reasonable natural boundary.

e. Hard (2.0-7.0 keV): The limits discussed in *a*. and *d*. set the limits of this band.

B. Energy Band for HRC

Since the HRC has limited spectral resolution a single broad band from 0.1–10 keV was chosen. All images, source detections, and fluxes are based on this band pass. The effective area ($HRMA \times quantum\ efficiency$) for HRC-I and HRC-S is shown in Fig. 10. In Figs.11 is the same plot multiplied by the spectral weighting function given in section **A.**

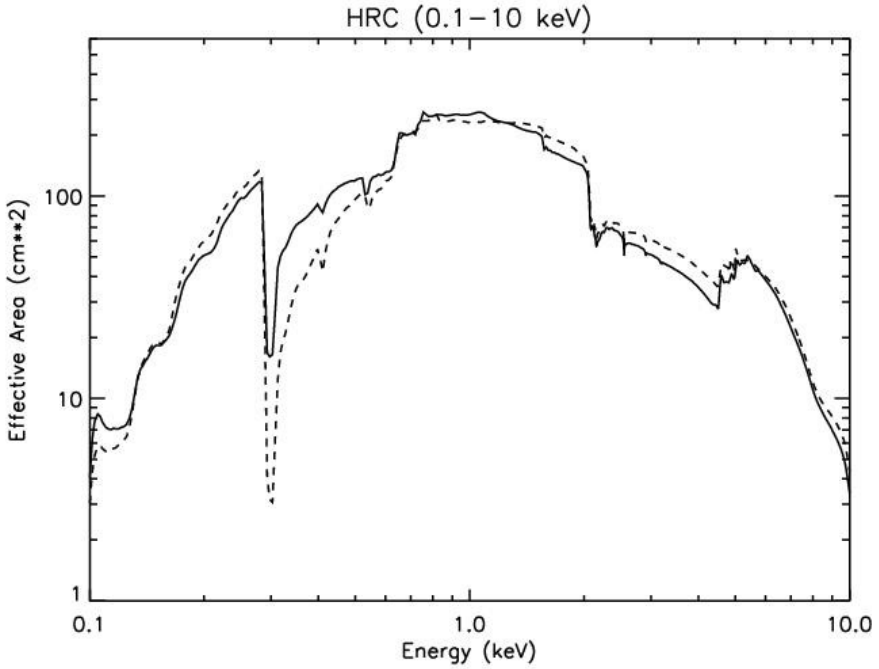


Fig. 10: This is a plot of the product of the effective area of the telescope with quantum efficiency of the detector. The solid line is the HRC-S detector and the dashed line is the HRC-I detector.

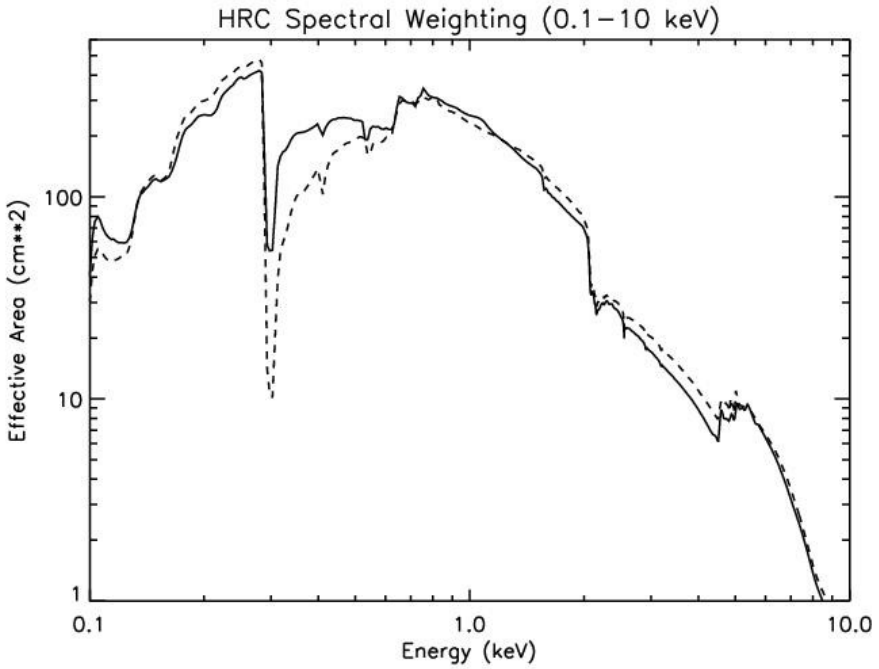


Fig. 11: This is a plot of the product of the effective area of the telescope with quantum efficiency of the detector times a spectral weighting function (see text). The solid line is the HRC-S detector and the dashed line is the HRC-I detector.

II. Effective Band Energy

For calculating exposure/instrument maps, psf, etc. for the CSC bands one needs to have an effective single energy to represent the band. Below is given the method of calculating the effective energy of each band. A number of cases are considered and the final values are given.

A. Method of Calculation

We calculate that effective energy for each band using the following relation:

$$E_{eff} = \frac{E_b}{E_n}$$

where the weighted energy E_b and the normalization E_n are given by:

$$E_b = \sum E_i A_i Q_i C_i S_i \Delta E_i$$

and

$$E_n = \sum A_i Q_i C_i S_i \Delta E_i$$

Where:

E_i : Energy of interval *i* of the band being considered.

A_i : Effective Area of the telescope (HRMA) of interval *i* of the band being considered.

Q_i : Detector quantum efficiency for interval *i* of the band being considered.

C_i : Reduction in transmission due to the build up of contamination on ACIS for interval *i* of the band being considered.

S_i : Spectral weighting function of the form $(E_i/E_0)^{-\alpha}$ where E_0 is the normalization energy (taken to be 1 keV) and α is the spectral index.

ΔE_i : The width of energy interval *i* of the band being considered.

In both E_b and E_n the sum is performed over the sampling of the energy band in question.

B. Results for ACIS and HRC Bands

The effective energies for each ACIS band were calculated for the following cases:

1. E_a : The energy determined by only weighting by effective area of the telescope. Done for both ACIS and HRC.
2. E_I : The energy determined by weighting by the effective area of the telescope and quantum efficiency of the I3 chip (ACIS) or HRC-I.
3. E_{Ic} : The energy determined by weighting by the effective area of the telescope, the quantum

efficiency of the I3 chip, and the maximum value of the transmission reduction due to contamination. (ACIS only)

4. E_{Is} : The energy determined by weighting by the effective area of the telescope, the quantum efficiency of the I3 chip or HRC-I, and a spectral weighting (α = 1).
5. E_{Ics} : The energy determined by weighting by the effective area of the telescope, the quantum efficiency of the I3 chip, the maximum value of the transmission reduction due to contamination, and a spectral weighting (α = 1). (ACIS only)
6. E_{S} : The energy determined by weighting by the effective area of the telescope and quantum efficiency of the S3 chip or HRC-S.
7. E_{Sc} : The energy determined by weighting by the effective area of the telescope, the quantum efficiency of the S3 chip, and the maximum value of the transmission reduction due to contamination. (ACIS only)
8. E_{Ss} : The energy determined by weighting by the effective area of the telescope, the quantum efficiency of the S3 chip or HRC-S, and a spectral weighting (α = 1).
9. E_{Scs} : The energy determined by weighting by the effective area of the telescope, the quantum efficiency of the S3 chip, the maximum value of the transmission reduction due to contamination, and a spectral weighting (α = 1). (ACIS only)

Cases	***0.5–7.0 keV***	***0.2–0.5 keV***	***0.5–1.2 keV***	***1.2–2.0 keV***	***2.0–7.0 keV***	***HRC***
E_{a}	3.04 keV	0.35 keV	0.85 keV	1.59 keV	4.22 keV	3.14 keV
E_{I}	3.38 keV	0.44 keV	0.95 keV	1.58 keV	4.27 keV	2.77 keV
E_{Ic}	3.49 keV	0.45 keV	0.98 keV	1.59 keV	4.28 keV	
E_{Is}	2.36 keV	0.44 keV	0.91 keV	1.55 keV	3.86 keV	1.44 keV
E_{Ics}	2.54 keV	0.45 keV	0.95 keV	1.56 keV	3.86 keV	
E_{S}	3.04 keV	0.38 keV	0.88 keV	1.59 keV	4.11 keV	2.77 keV
E_{Sc}	3.21 keV	0.36 keV	0.92 keV	1.60 keV	4.12 keV	
E_{Ss}	1.99 keV	0.36 keV	0.84 keV	1.55 keV	3.69 keV	1.44 keV
E_{Scs}	2.22 keV	0.33 keV	0.88 keV	1.56 keV	3.70 keV	

The above cases which most likely represent what will be found in the CSC catalog are those found with spectral weighting (with and without contamination). One also has to weight somewhat between how many sources will be found in an front illuminated vs. a back illuminated chip. With that in mind we chose the following as effective energies for the chosen bandpasses.

Bands	***0.2–7.0 keV***	***0.2–0.5 keV***	***0.5–1.2 keV***	***1.2–2.0 keV***	***2.0–7.0 keV***	***HRC***
Energy	2.3 keV	0.4 keV	0.92 keV	1.56 keV	3.8 keV	1.5 keV

These will be used in creating quantities that need a single energy in order to calculate (instrument/exposure maps, psfs, etc.).

Appendix B

To: L3 Distribution
From: F. Primini
Subject: Optimum Soft Energy Band Boundaries for Super-Soft Sources
Date: 3/2/2006

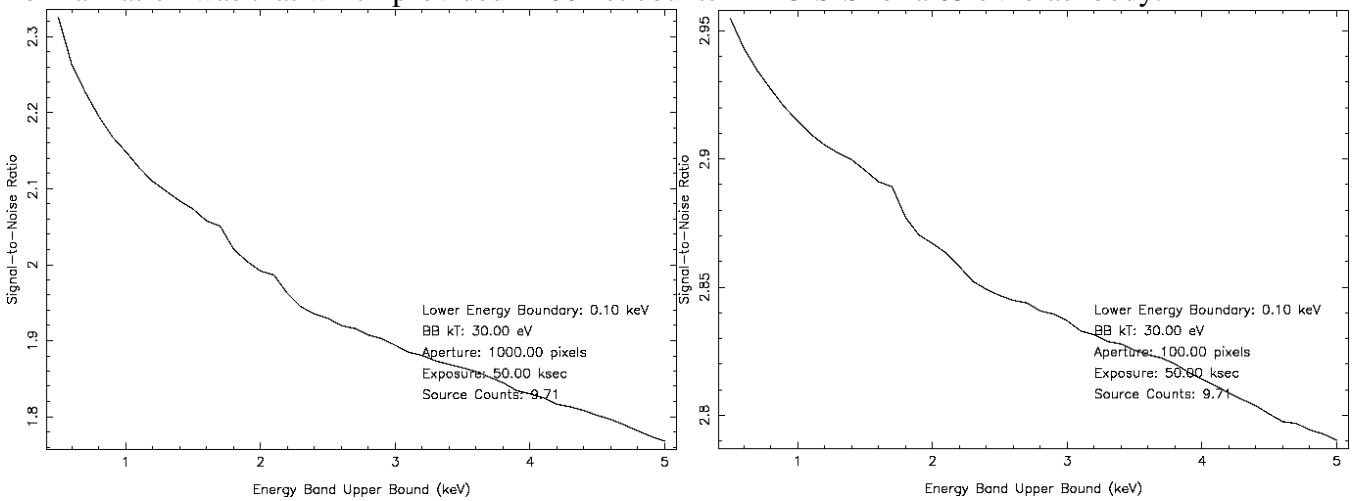
It has been suggested that the current L3 soft energy band boundaries (0.2–0.5 keV) are not optimum *source detection* bands for super-soft sources. To investigate this question, I have simulated super-soft spectra in ACIS-S and ACIS-I, using *fakeit* in *ISIS*, and computed simple source count signal-to-noise ratios for a range of soft band boundaries. To model super-soft spectra, I used blackbody spectra with kT=30, 65, and 100 eV , and added background obtained from appropriate blank-sky datasets scaled to the source exposure time and aperture size. I used a source exposure time of 50 ksec. Typical model spectra are shown in Figure 1.

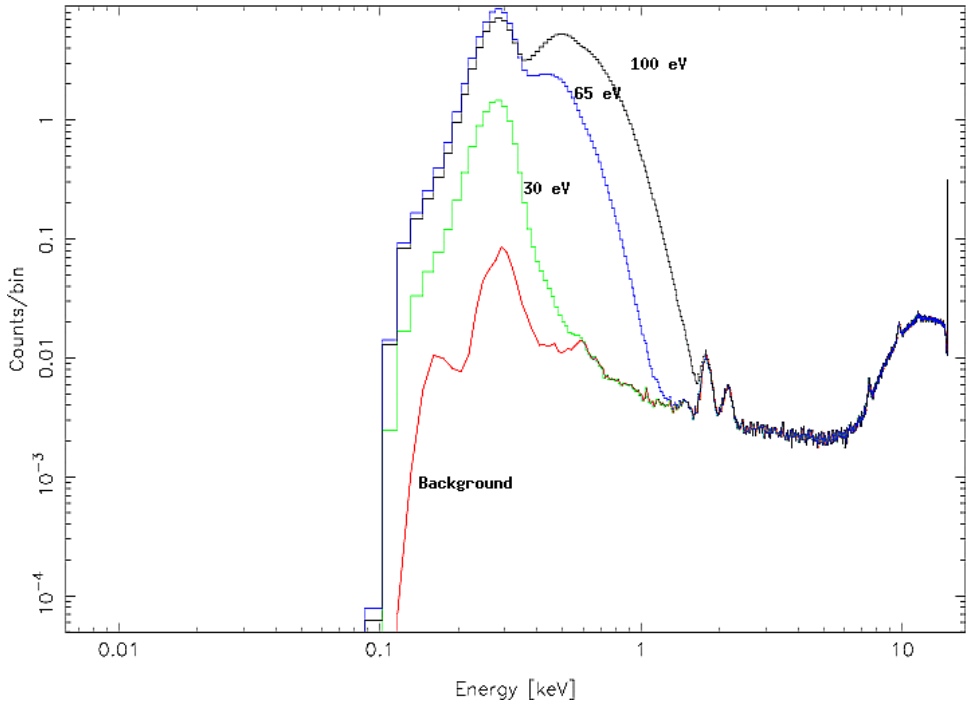
I considered two cases: a small 100 pixel source aperture, appropriate for an on-axis source, and a larger 1000 pixel aperture, appropriate for a source ~8' off-axis. For each aperture size and blackbody temperature, I simulated 1000 spectra and computed SNR's in bands with lower energy boundaries of 0.1, 0.2, and 0.3 keV, and upper energy boundaries from 0.5 to 5 keV. Final SNR's were simple means of those computed in each of the 1000 simulations. The results are shown in Figures 2–4 (ACIS-S) and 5 (ACIS-I). I did not compute a full set of ACIS-I results since they appeared similar to the ACIS-S results.

Only for the very softest spectra (kT~30 eV) does the current soft band yield the highest SNR, although it's not clear that it's the optimum. For the higher temperature blackbodies, the optimum energy band has an upper bound of ~1 keV. There is little difference between the lower bounds of 0.1 and 0.2 keV, although a lower bound of 0.3 keV definitely reduces the SNR values. However, for small aperture sizes (i.e., on-axis), the SNR decays very slowly as the upper energy boundary increases beyond 1 keV. This simply reflects the fact that for these apertures, the spectrum is source dominated, even at 50 ksec. For the large aperture size, the SNR decays more rapidly as the upper boundary increases, indicating that these spectra are beginning to be background-dominated. However, even for these spectra, the decrease in SNR from the optimum upper bound of 1 keV to higher energies is not large.

On the basis of these tests, I conclude that except for sources with very soft spectra or very large off-axis angles, there is no strong reason to optimize the soft band to detect super-soft sources, since they will most likely be detected in the broad band anyway.

Fig. 1: Model super-soft source spectra in ACIS-I (top) and ACIS-S (bottom) for a range of blackbody temperatures. Background spectra were obtained from appropriate blank-sky datasets, scaled to the exposure time (50 ksec.) and aperture size (100 pixels) of the source. In all cases, the blackbody normalization was that which provided ~100 net counts in ACIS-S for a 65 eV blackbody.





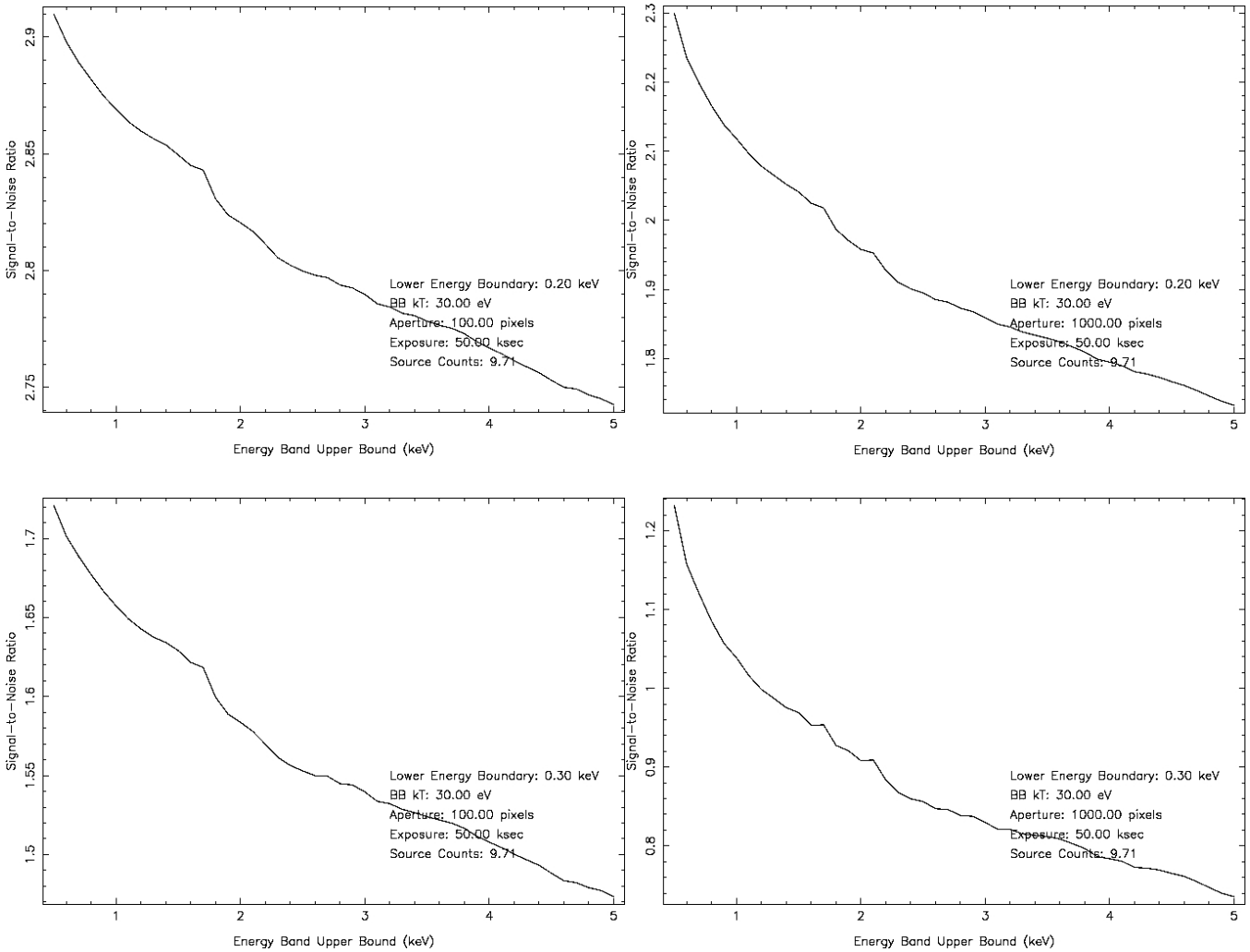


Fig. 2: SNR vs. soft band upper boundary for a 30 eV blackbody on-axis and ~8' off-axis in ACIS-S.

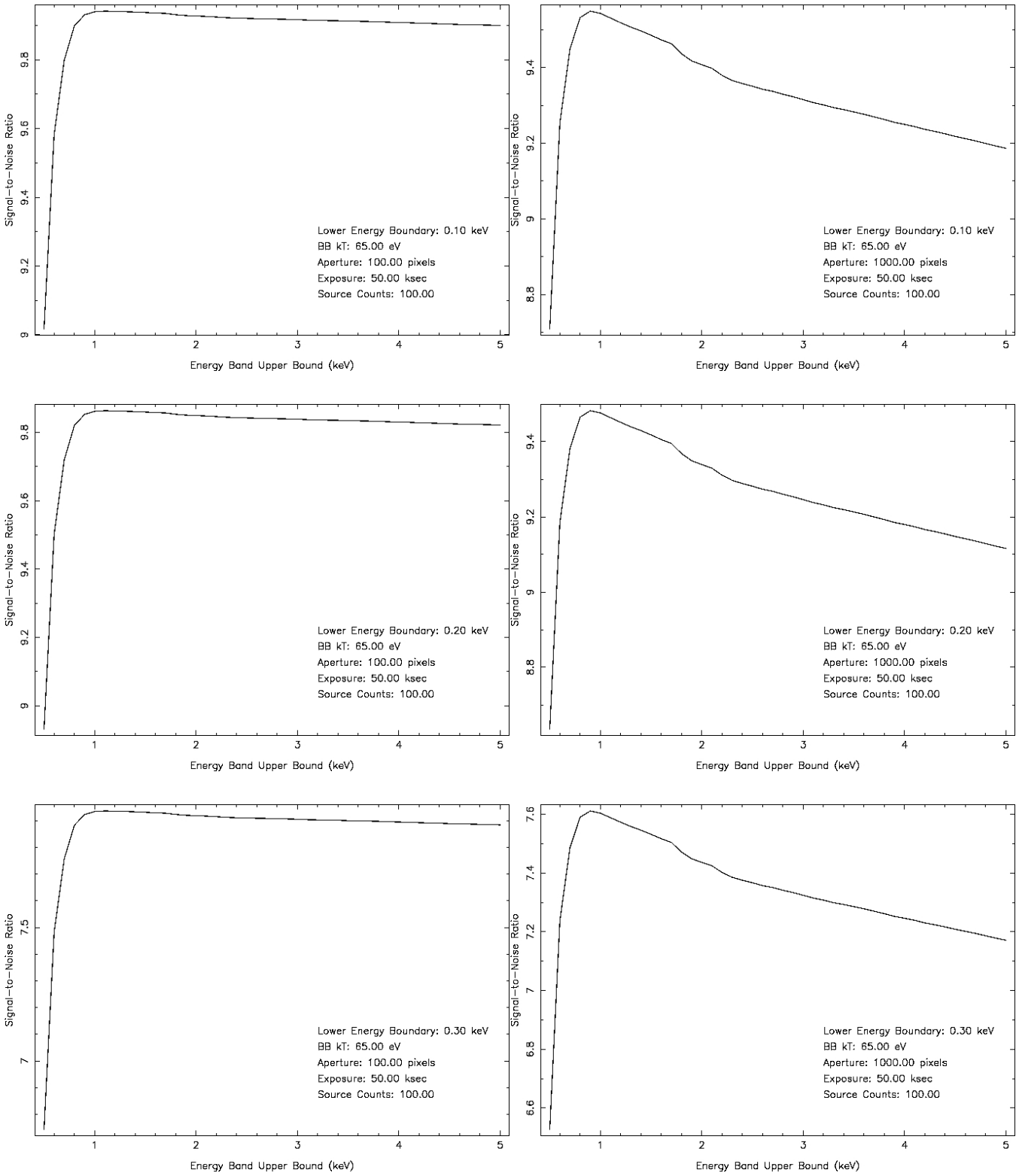


Fig. 3: SNR vs. soft band upper boundary for a 65 eV blackbody on-axis and ~8' off-axis in ACIS-S.

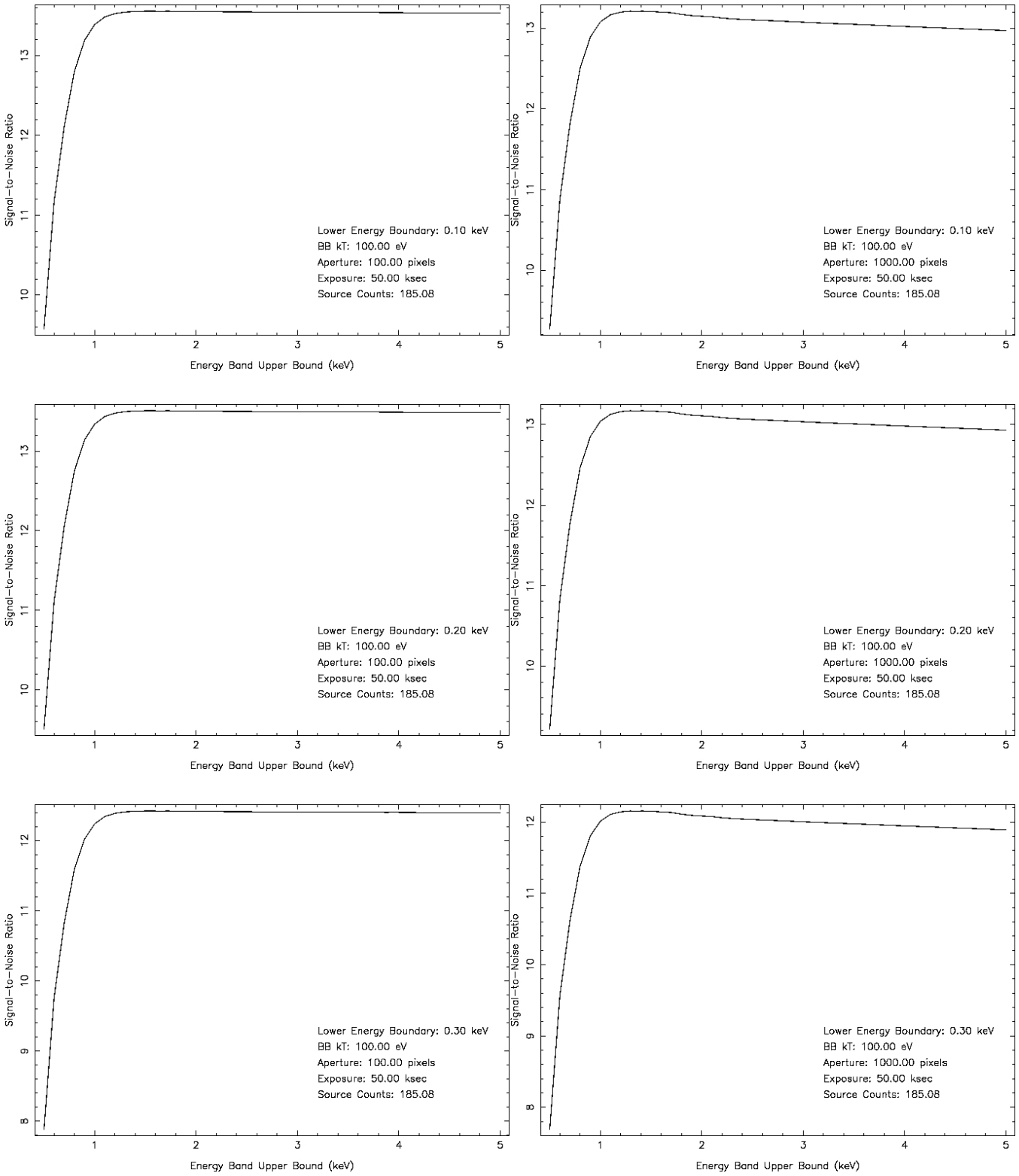


Fig. 4: SNR vs. soft band upper boundary for a 100 eV blackbody on-axis and 8' off-axis in ACIS-S.

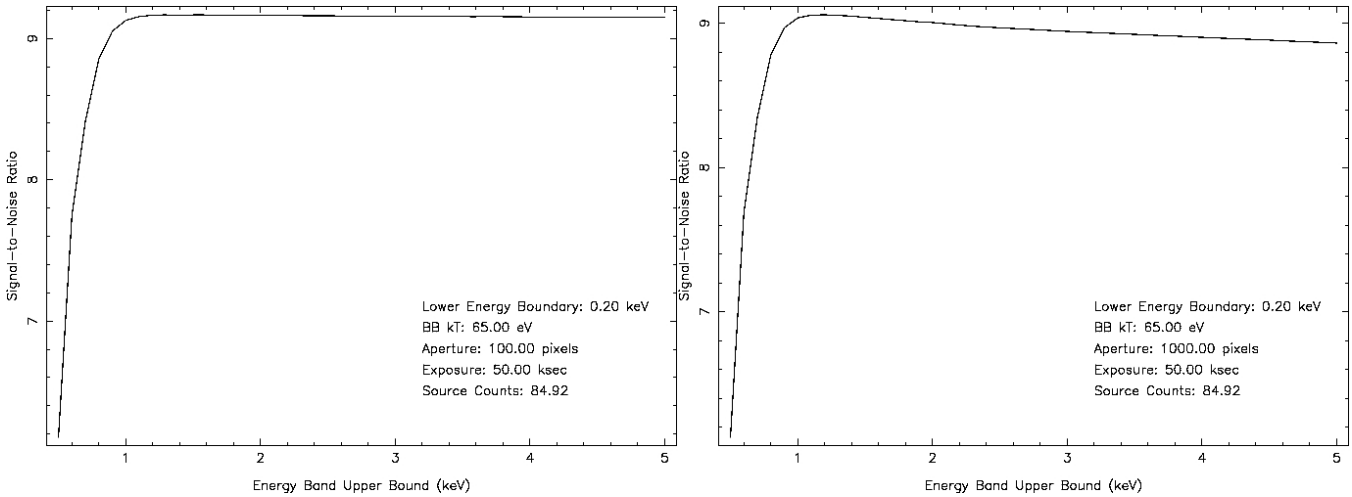


Fig. 5: SNR vs. soft band upper boundary for a 65 eV blackbody on-axis and ~8' off-axis in ACIS-I.

Appendix C
Default Spectral Fit Parameters for the Chandra Source Catalog

Michael L. McCollough CXC/SAO
January 24, 2007

For the Chandra Source Catalog we need a default set of spectral fit parameters to be used for flux determinations in case where there are too few counts available to make a reliable spectral fit to the data. For the Chandra Source Catalog we need the following: (a) count threshold for spectral fitting; (b) the parameters for an absorbed power law fit; (c) the parameters for an absorbed Blackbody fit.

Using an XMM study which used the XMM1 catalog the following conclusions were arrived at:

Spectral Fitting Threshold: An XMM comparison of the detect counts found in EPIC to the input counts used in simulations for these observations found a systematic bias in the number of counts observed below 250 counts (a larger number then expected). This is case of an **Eddington bias** (also referred to as **Malmquist bias**). Thus to avoid problems with this systematic effect it is recommended that a limiting value of at least 250 counts be used for the threshold for spectral fits.

Power Law Fit: This same XMM study performed fits to a large number of sources (all over 250 counts) with an absorbed power law. The resultant fit parameters were a spectral index of 1.80 ± 065 and a 1 sigma range of the column density of $N_h = 00 - 9.8 \times 10^{20} cm^{-2}$. These values are in keeping with the default values used for the 1XMM catalog of an index of 1.7 and a column density of $N_h = 3.0 \times 10^{20} cm^{-}$. It is recommend that these 1XMM values be used for as the default fit parameter (for sources below 250 counts) for the Chandra Source Catalog.

Blackbody Fit: The XMM studies did not consider a thermal fit to the sources. The source for which this fit will have the most relevance are galactic. Looking at galactic sources we find the following ranges:

- *Normal Stars:* Average of around 1 keV with a range 1–10 keV.
- *RS CVn Stars:* Average X-ray spectrum of 1.58 keV with a range of 0.83–4.6 keV.
- *Cataclysmic Variables:* Have a soft component of 30–100 eV and a hard component of 1–50 keV.

- *X-Ray Binaries (Neutron Stars):* Typical 1–4 keV bremsstrahlung spectrum.
- *X-Ray Binaries (Black Holes):* Typical 1–2 keV spectrum during high state.
- *Super Soft Sources:* A range of 10–200 eV.

Based on these wide range of values a kT of 1.0 keV is recommended with a column density of $N_h = 3.0 \times 10^{20} cm^{-2}$ to be used as the default values for the Chandra Source Catalog.

To determine some of these parameters we relied on an XMM study (using the 1XMM catalog): XMM-SOC-CAL-TN-0023 "A statistical evaluation of the EPIC flux calibration" (April 2003).

Appendix D

Date: Feb. 5, 2004
To: L3 Distribution
From: F. Primini
Subject: Computing Flux and Flux Significance

1 Introduction

Here are simple aperture photometry formulae for computing photon flux and flux significance for a point source. For "flux significance" I will adopt the traditional "signal-to-noise ratio" definition, i.e., $Flux/\sigma_{Flux}$, where σ_{Flux} is the 1-sigma error in flux. Strictly speaking, it is only in the limit of Gaussian statistics that a single "1-sigma error" has meaning, but I'll defer discussion of this point to the end of the memo, and in what follows leave formulae for σ_{Flux} in general terms.

2 Definitions

$\mathbf{x_0,y_0}$	Position of Source (centroid)
S	Source Flux (ph cm^{-2} s^{-1})
n	Counts in Source Aperture
$\mathbf{A_S}$	Area in pixels of Source Aperture
$\mathbf{E_S}$	Average Exposure ($cm^2 s$) in Source Aperture
α	Fraction of PSF enclosed in Source Aperture, i.e., $\alpha = \int_{Source\ Aperture} psf(x_0, y_0, x, y)\, dxdy$
B	Background Density (counts $pixel^{-1}$)[1]
m	Counts in Background Aperture
$\mathbf{A_B}$	Area of Background Aperture
$\mathbf{E_B}$	Average Exposure in Background Aperture
β	Fraction of PSF enclosed in Background Aperture, i.e., $\beta = \int_{Background\ Aperture} psf(x_0, y_0, x, y)\, dxdy$

3 Simple Case: Isolated Source, $\beta = 0$

The situation is shown in Figure 1. I have drawn a detached background aperture for clarity; in reality, it may be attached to the source aperture. The assumption $\beta = 0$ means that no source counts are

scattered into the background aperture, and could be realized by setting the background aperture radius to a large value (e.g., > r_{99}). Whether this would be practical or desirable is open to debate.

By inspection, the flux for the source is then given by

$$S=\left(\frac{1}{\alpha E_S}\right)\left(n-\frac{A_S}{A_B}m\right)$$

and the error on the flux is given by

$$\sigma_S=\left(\frac{1}{\alpha E_S}\right)^{\square}\left(\sigma_n^2+\left(\frac{A_S}{A_B}\right)\sigma_m^2\right)$$

4 A More Complicated Case: Isolated Source, $\beta>0$

In this case, the background aperture is sufficiently close to the source that some source counts are scattered into it. Now, S must be determined by solving two simultaneous linear equations:

$$n=\alpha E_S S+A_S B$$
$$m=\beta E_B S+A_B B$$

or

$$S=\frac{n-\frac{A_S}{A_B}m}{\alpha E_S-\beta E_B\frac{A_S}{A_B}}$$

Note this reduces to the equation for S in the previous section on setting $\beta=0$. For simplicity, let

$$a=\frac{1}{\alpha E_S-\beta E_B\frac{A_S}{A_B}}$$

and

$$b=\frac{\frac{A_S}{A_B}}{\alpha E_S-\beta E_B\frac{A_S}{A_B}}$$

we can then write

$$S=an-bm$$

and

$$\sigma_S^2=a^2\sigma_n^2+b^2\sigma_{\square m}^{\square}$$

5 General Case: Multiple, Overlapping Sources

Consider the case indicated in Figure 2. Here, each source aperture includes contributions not only from the source it encloses and background, but also possibly from nearby sources. For laziness sake, I've drawn all regions as simple, independent circles and annuli; in real life they may overlap, and will have to be adjusted by excluding parts to ensure that the counts n_1, n_2, m_1, and m_2 are statistically independent.

Generalizing the definitions for α and β in Section 2 to the case of multiple sources:

α_{ij} Fraction of PSF for source i enclosed in source aperture for source j;

β_{ij} Fraction of PSF for source i enclosed in background aperture for source j ;

E_{S_i} Average Exposure ($cm^2 s$) in Source Aperture i;

A_{S_i} Area in pixels of Source Aperture i ;

E_{B_j} Average Exposure ($cm^2 s$) in Background Aperture j;

A_{B_j} Area in pixels of Background Aperture j ;

We can then write, for the case of two sources as in figure 2:

$$n_1 = \alpha_{11} E_S \; S_1 + \alpha_{21} E_{S_1} S_2 + A_{S_1} B$$
$$n_2 = \alpha_{12} E_{S_2} S_1 + \alpha_{22} E_{S_2} S_2 + A_{S_2} B$$
$$m_1 = \beta \quad E_{B_1} S_1 + \beta_{21} E_{B_1} S_2 + A_{B_1} B$$
$$m_2 = \beta_{12} E_{B_2} S_1 + \beta_{22} E_{B_2} S_2 + A_{B_2} B$$

But the last two equations can be summed into one, namely,

$$\sum_j m_j = \sum_j \beta_{1j} E_{B_j} S_1 + \beta_{2j} E_{B_j} S_2 + \sum_j A_{B_j} B$$

where the sums are over the number of background regions. We now have 3 simultaneous linear equations, which we can solve for S_1, S_2, and B. The solutions will again be of the form

$$S_i = a_i n_1 + b_i n_2 + c_i m$$
$$\sigma_{S_i}^2 = a_i^2 \sigma_{n_1}^2 + b_i^2 \sigma_{n_2}^2 + c_i^2 \sigma_m^2$$

The expression of a, b, and c in terms of the various α , β , etc. is left as an exercise for the reader. In general, for N sources, there will be N+1 simultaneous linear equations to solve, and S_i and $\sigma_{S_i}^2$ will be linear combinations of the counts n_i and $\sigma_{n_i}^2$, respectively. However, it probably doesn't pay to extend the analysis to N > ~ 2–3 since the assumption that B is constant will probably be violated.

6 Determining σ_n^2

All that remains is to specify how to calculate quantities of the form σ_n^2, where n is small, so that Poisson statistics apply. As mentioned before, using a single number for "1 σ error" is only valid in the Gaussian limit, where $\sigma^2 \sim$ n, but this is a poor approximation for small n. However, if we interpret σ as the half-size of the two-sided 68.27% confidence limit about the true value, we can extend the concept more accurately for small n cases. The popular Gehrels approximation (Gehrels 1986, ApJ, 303, 336)

$$\sigma_n \approx 1+\sqrt{n+3/4}$$

is a better approximation than $\sqrt{n}$ (this is the approximation used in the Penn State Photometry code), but it overestimates the 68% confidence region. This is because the region is not symmetric about the true value. In fact, Gehrels provides two formulae, one for the lower bound, λ_l, and one for the upper bound, λ_u (see equations 7 and 11 in his paper; S=1 for the 68% confidence region):

$$\lambda_u = n+\sqrt{n+3/4}+1$$
$$\lambda_l = n-\sqrt{n- \ /4}$$

The following table illustrates the differences in the approximations for various values of n:

n	$\sigma_n=\sqrt{n}$	$\sigma_n=1+\sqrt{n+3/4}$	$\sigma_n=\frac{\lambda_u-\lambda_l}{2}=\frac{\sqrt{n+3/4}+\sqrt{n-1/\ }+1}{2}$
10	3.16	4.28	3.70
30	5.48	6.55	6.00
50	7.07	8.12	7.59
100	10.0	11.0	10.5

I suggest we use this last approximation for σ_n.

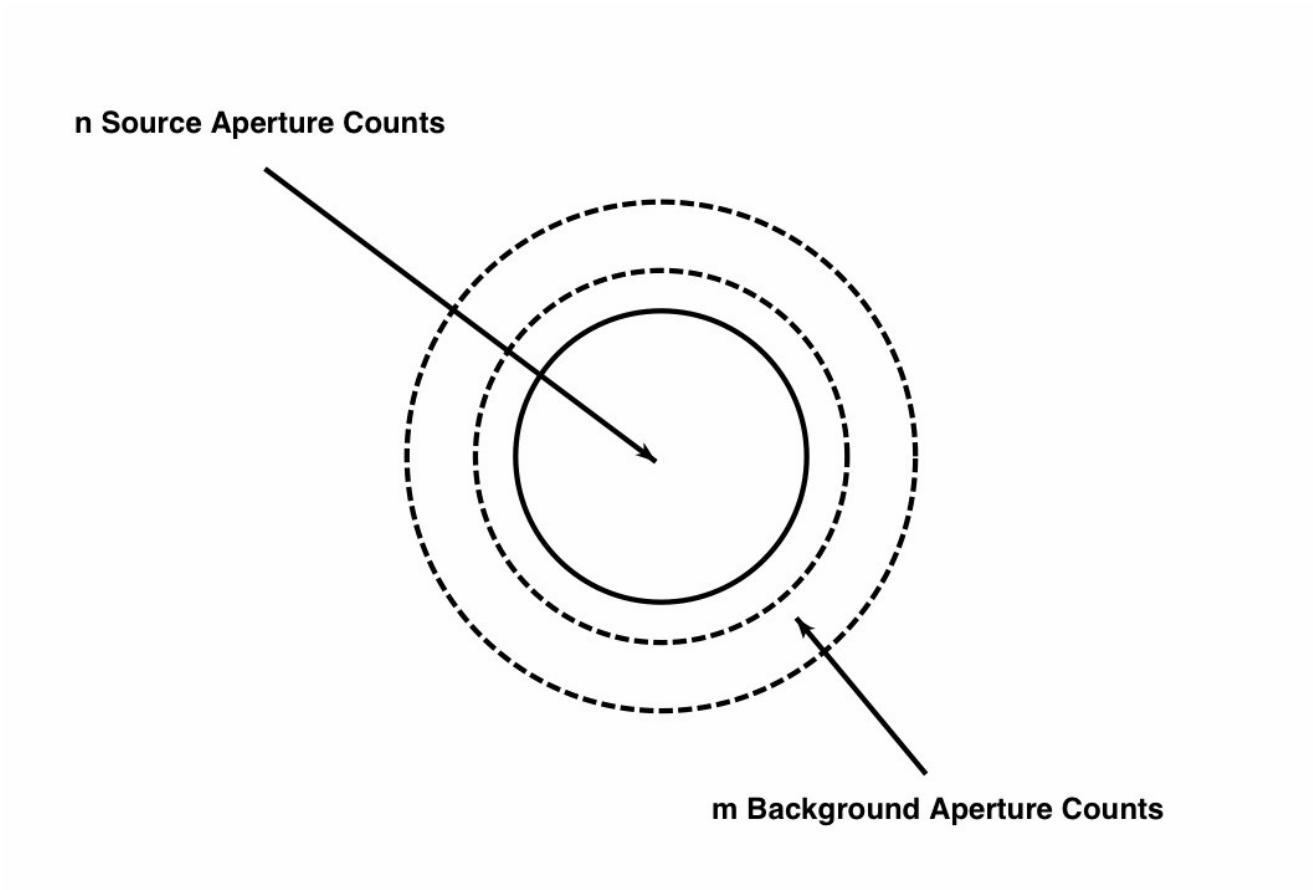


Figure 1.: A Single Isolated Point Source

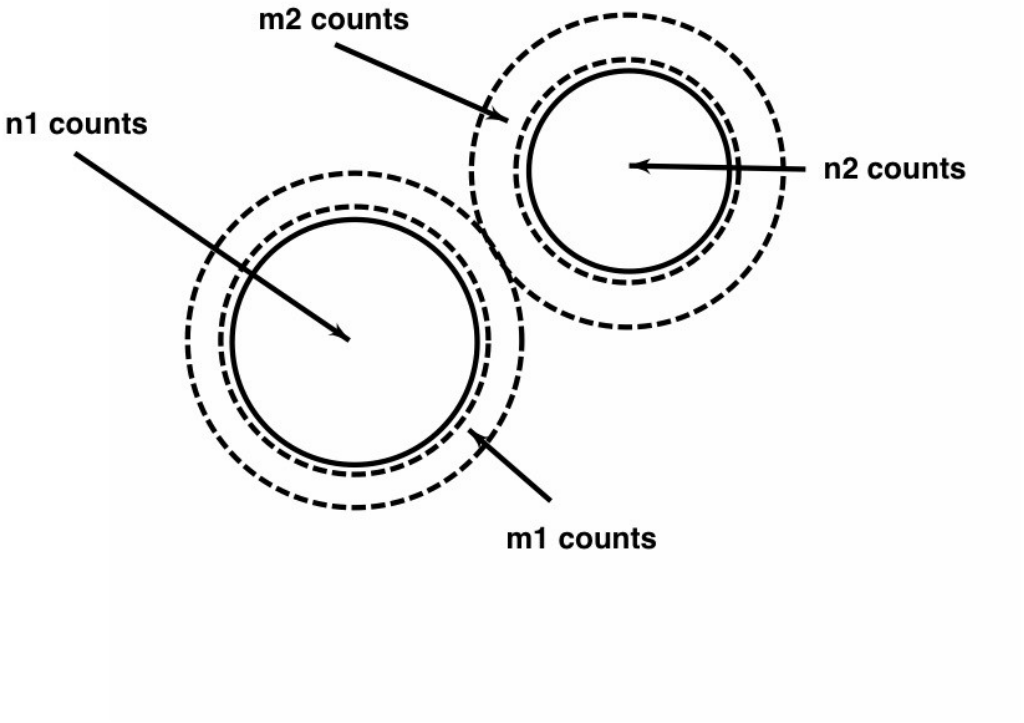


Figure 2.: Multiple Overlapping Sources

Footnotes

[1]Since the background contains both cosmic and instrumental components, it shouldn't be expressed in the same units as source flux. I'm assuming here that the background is essentially flat (at least over the scale of the region of interest) and dominated by the non-cosmic component.

Appendix E

Memorandum

To: L3 Distribution
From: F. Primini
Subject: Revised Specifications for Computing Aperture Photometry Quantities
Date: June 6, 2007

Introduction

This is a revised specification for computing the aperture photometry quantities, including errors, described in section 1.2.2.6.3 of the Chandra Source Catalog Requirements, v. 0.5. The basic paradigm remains the same as that described in the original aperture flux specification (https://icxc.harvard.edu/soft/schedule/L3/Sci_Docs/significance.ps), but I have generalized the formalism to deal explicitly with rates, photon fluxes, and energy fluxes, and I have recast the mathematics slightly to conform with the approach used by Vinay Kashyap for computing the Bayesian Posterior Probability Distribution for the source quantities. A memo describing the mathematical details of Vinay's approach accompanies this memo, and sample code for calculating the probability distribution is included as an appendix here. This probability distribution is used in the error calculations.

Deleted Quantities

In the context of the current algorithm for computing fluxes, it does not make sense to compute net source counts, rate, or flux, corrected for psf fraction, separately in source and background apertures. Such quantities are identified in section 1.2.2.6.3 by extension **`aperbkg`** (and **`aperbkg_err`** for errors). All such quantities should be deleted from the requirements.

Definitions

n Counts in source aperture

m Counts in background aperture

A_s Number of pixels in source aperture

A_B Number of pixels in background aperture

α Fraction of PSF in source aperture

β Fraction of PSF in background aperture

T_s Exposure time in source aperture (sec)

T_B Exposure time in background aperture (sec). For generality, this is defined independently from T_s , although it usually has the same value.

E_s Average exposure map value in source aperture ($cm^2 - sec$)

E_B Average exposure map value in background aperture ($cm^2 - sec$)

$\bar{\epsilon}_s$ Average photon energy in source aperture ($\bar{\epsilon}_s = \sum_{i=1}^{n} \epsilon_i / n$)

$\bar{\epsilon}_B$ Average photon energy in background aperture ($\bar{\epsilon}_B = \sum_{i=1}^{m} \epsilon_i / m$)

Basic Formalism

It is assumed that the source and background apertures are distinct (although they are not necessarily simple shapes) so that n and m are statistically independent. The general approach is to express n and m as sums of contributions from source and background quantities, yielding two simultaneous linear equations which can be solved for the source quantity. For reasons which I hope will become clear shortly, I chose to write these as follows:

$$n = fs + cb = fs + b'$$
$$m = gs + rcb = gs + rb'$$

Here, the intermediate quantities c and b are introduced for consistency in describing source and background units, but do not enter into the final calculations.

Again, the equations for n and m are cast in this form so that we can easily take advantage of Vinay's formalism for computing probability distributions. The actual meanings of the parameters f, g, c, and r will depend on the particular source quantity we're interested in, e.g., net counts, rate, etc.

The quantity of interest is s, and its Gaussian error, σ_s are given, in general, by

$$s = \frac{rn - m}{rf - g} \quad , \qquad \sigma_s^2 = \frac{r^2 n + m}{(rf - g)^2} \quad .$$

The determination of s and its error for the different source quantities net counts, net rate, net photon flux, and net energy flux, then reduces to the proper determination of the parameters f, g and r. These are described below.

Net Counts (Requirements Section 1.2.2.6.3.1.2)

Here, the units of s and b are simply *counts* and *counts*$-pix^{-2}$, and f, g, c, and r can be written in terms of our earlier definitions as

$$f = \alpha$$
$$g = \beta$$
$$c = A_S$$
$$r = \frac{A_B}{A_s} \quad .$$

Net Rate (Requirements Section 1.2.2.6.3.3)

The units of s and b are now $counts-sec^{-1}$ and $counts-sec^{-1}-pix^{-2}$, and f, g, c, and r can be written in terms of our earlier definitions as

$$f=\alpha T_s$$
$$g=\beta T_B$$
$$c=A_S T_S$$
$$r=\frac{A_B T_B}{A_s T_s} \quad .$$

Photon Flux (Requirements Section 1.2.2.6.3.4)

The units of s and b are now $photons-cm^{-2}-sec^{-1}$ and $photons-cm^{-2}-sec^{-1}-pix^{-2}$, and f, g, c, and r can be written in terms of our earlier definitions as

$$f=\alpha E_s$$
$$g=\beta E_B$$
$$c=A_S E_S$$
$$r=\frac{A_B E_B}{A_s E_s} \quad .$$

Net Energy Flux (Requirements Section 1.2.2.6.3.4)

The units of s and b are now $ergs-cm^{-2}-sec^{-1}$ and $ergs-cm^{-2}-sec^{-1}-pix^{-2}$, and f, g, c, and r can be written in terms of our earlier definitions as

$$f=\alpha E_s/\bar{\epsilon}_s$$
$$g=\beta E_B/\bar{\epsilon}_B$$
$$c=A_S E_S/\bar{\epsilon}_s$$
$$r=\frac{A_B E_B \bar{\epsilon}_s}{A_s E_s \bar{\epsilon}_B} \quad .$$

In this case, since $\bar{\epsilon}_s$ and $\bar{\epsilon}_B$ appear in the denominators of some quantities, we need to be able to handle the exception when they are 0. In those cases, we should set them equal to the midpoint energy of the energy band in use.

Computation of Errors

In all cases, once f, g, and r are known, the maximum likelihood estimator for s can be computed from

$$s_{MLE} = \frac{rn - m}{rf - g} \quad .$$

The confidence region on s is computed from the posterior probability distribution for s. The computation of this distribution is described in the accompanying memo, "Background Marginalized X-Ray Source Intensity", by V. Kashyap. The distribution we're interested in is described in eq. 26. Sample **s-lang** code for computing this distribution given n, m, f, g, and r is provided in the appendix. The basic routine is **s_pdf**, whose arguments are n, m, r, f, and g, (note n = **C** and m = **B** in the documentation). The distribution is provided as an array with a binsize of 0.1 in s in the code.

To determine the energy bounds, we should find the mode of the distribution (which should be close to s_{MLE}) and sum probabilities in alternating bins on either side of the mode until we achieve the desired confidence level. The values of s in the final bins determine the error bounds. If $s = 0$ is reached before the confidence level is achieved, the integration should continue only on the positive side of the mode, and the flux should be flagged as an upper limit.

Appendix — S-Lang Code for Computing Posterior Probability Distribution

```
#! /usr/bin/env slsh
_auto_declare =1;
_traceback=1;

()=evalfile("/Users/fap/bin/slang.sl");
require("gsl");

% For now, get variables from command line;

C = integer(__argv[1]);
B = integer(__argv[2]);
r = atof(__argv[3]);
f = atof(__argv[4]);
g = atof(__argv[5]);

define s_pdf(C,B,r,f,g){

 % Compute the posterior probability distribution for source counts or
intensity ;
 % marginalized over background, assuming non-informative gamma priors
(alpha=1, beta-0) ;

 variable i,j,k;
 variable C1 = C+1;
 variable B1 = B+1;
 variable CB1= C+B+1;

 variable rf_minus_g = r*f-g;
 variable one_plus_r = 1.0+r;
```

```
 variable K = Double_Type[C1,B1];
 variable S = Double_Type[C1,B1];

 % Compute MLE value for s, to estimate reasonable range for pdf ;

 variable S_mle = (r*C-B)/rf_minus_g;
 if(S_mle<=0.0) S_mle=30;

 variable s = [0.0:3*S_mle:0.1];
 variable ps= Double_Type[length(s)];

 for(k=0;k<=C;,k++){
  for(j=0;j<=B;j++){

   K[k,j]=rf_minus_g*f^k*g^j*r^(B-j)*gamma(CB1-k-j);
   K[k,j]/=(gamma(k+1)*gamma(C1-k)*gamma(j+1)*gamma(B1-j)*one_plus_r^(CB1-k-j));

  }
 }

 for(i=0;i<length(s);i++){

  for(k=0;k<=C;,k++){
   for(j=0;j<=B;j++){

    S[k,j]=s[i]^(k+j)*exp(-1.0*s[i]*(f+g));

   }
  }

  ps[i] = sum(_reshape(S*K,[(C+1)*(B+1)]));

 }

 % Renormalize psf just in case;

 return s,ps/(0.1*sum(ps));
% return s,ps;

} % end of s_pdf;

define s_gaussian(C,B,r,f,g){

 % Compute the classical Gaussian probability distribution for source counts or
intensity ;

 variable i,j,k;
 variable C1 = C+1;
 variable B1 = B+1;
 variable CB1= C+B+1;

 variable rf_minus_g = r*f-g;
 variable one_plus_r = 1.0+r;
```

```
 variable K = Double_Type[C1,B1];
 variable S = Double_Type[C1,B1];

 % Compute MLE value for s, to estimate reasonable range for pdf ;

 variable S_mle = (r*C-B)/rf_minus_g;
 variable sigma_s = sqrt((r*r*C+B)/(rf_minus_g*rf_minus_g));

 variable S_tmp = S_mle;
 if(S_tmp<=0.0) S_tmp=30;

 variable s = [0.0:3*S_tmp:0.1];
 variable ps= Double_Type[length(s)];

 ps = (1.0/(sqrt(2.0*PI)*sigma_s))*exp(-1.0*(s-S_mle)*(s S_mle)/
(2.0*sigma_s*sigma_s));

 return s,ps;
} % end of s_gaussian;

(s,ps) = s_pdf(C,B,r,f,g);

mode_s_pdf = s[where(ps==max(ps))];
()=printf("Posterior Probability Distribution Mode:\t%f\n",mode_s_pdf[0]);
norm_s_pdf = sum(ps)*0.1;
()=printf("Posterior Probability Distribution Normalization:\t%f\n",norm_s_pdf);

fp=fopen("S_Posterior_PDF","w+");

for(i=0;i<length(s);i++) ()=fprintf(fp,"%f\t%f\n",s[i],ps[i]);

()=fclose(fp);

(s,ps) = s_gaussian(C,B,r,f,g);

mode_s_gaussian = s[where(ps==max(ps))];
()=printf("Gaussian Probability Distribution Mode:\t%f\n",mode_s_gaussian[0]);
norm_s_gaussian = sum(ps)*0.1;
()=printf("Gaussian Probability Distribution Normalization:\t
%f\n",norm_s_gaussian);
S_mle = (r*C-B)/(r*f-g);
()=printf("Max. Likelihood Estimate:\t%f\n",S_mle);

fp=fopen("S_Gaussian_PDF","w+");

for(i=0;i<length(s);i++) ()=fprintf(fp,"%f\t%f\n",s[i],ps[i]);

()=fclose(fp);
```

Memorandum

To: L3 Distribution
From: F. Primini
Subject: Avoiding Numerical Instabilities in Computing Aperture Fluxes
Date: June 22, 2007

Introduction

In my June 6, 2007 memo, I described the calculation of background-marginalized posterior probability distributions for source intensity quantities (i.e., net counts, rates, photon and energy fluxes). The implementation has indicated that the calculation of the probabilities for the photon and energy fluxes are numerically instable. I describe here an alternate mathematical formulation which avoids these instabilities.

Difficulties in the Original Formulation

Recall the original formula for the probability (eq. 26 in Vinay's memo "Background marginalized X-ray source intensity"):

$$p(s|CB)\,ds = ds\,\frac{(rf-g)}{\Gamma(C+1)\Gamma(B+1)}\sum_{k=0}^{C}\sum_{j=0}^{B}\frac{f^k g^j r^{B-j} s^{k+j} e^{-s(f+g)}\Gamma(C+1)\Gamma(B+1)\Gamma(C+B-k-j+1)}{\Gamma(k+1)\Gamma(C-k+1)\Gamma(j+1)\Gamma(B-j+1)(1+r)^{C+B-k-j+1}}$$

If C or B, the counts in source and background apertures, are large, the above expression encounters difficulties in evaluating the Γ function since $\Gamma(n+1)=n!$. Additional difficulties occur if the aperture correction factors f and g are also large. This latter condition is almost always true in computing the probabilities for photon or energy fluxes, since f and g scale with exposure map values. In the test cases run so far, the problems encountered have indeed been in the calculation of the fluxes.

Alternative Formulation

The above difficulties can be addressed by re-writing and simplifying the expression for $p(s|CB)$:

$$p(s|CB)\,ds = ds\,(rf-g)\sum_{k=0}^{C}\sum_{j=0}^{B}\frac{(fs)^k e^{-fs}}{k!}\frac{(gs)^j e^{-gs}}{j!}\frac{r^{B-j}\Gamma(C+B-k-j+1)}{\Gamma(C-k+1)\Gamma(B-j+1)(1+r)^{C+B-k-j+1}}$$

Here, I have taken advantage of the above definition of $\Gamma(n+1)=n!$. The first two terms in the summation can now be recognized as Poisson probability functions, which I'll write as

$$P_{pois}(k|fs)=\frac{(fs)^k e^{-fs}}{k!}\quad,$$

and similarly for $P_{pois}(j|gs)$. The probability distribution for s can now be written as

$$p(s|CB)\,ds = ds\,(rf-g)\sum_{k=0}^{C}\sum_{j=0}^{B}\frac{P_{pois}(k|fs)\,P_{pois}(j|gs)\,r^{B-j}\,\Gamma(C+B-k-j+1)}{\Gamma(C-k+1)\,\Gamma(B-j+1)\,(1+r)^{C+B-k-j+1}} \quad .$$

Of course, the Γ and Poisson functions can still encounter difficulties for large arguments, but there are well-known techniques for mitigating these. For example, many special functions libraries (including the GNU Scientific Library) contain routines for evaluating both the Γ and $\log\Gamma$ functions, the latter being used for large arguments. The above expression can then be written

$$p(s|CB)\,ds = ds\,(rf-g)\sum_{k=0}^{C}\sum_{j=0}^{B}P_{pois}(k|fs)\,P_{pois}(j|gs)\,e^{((B-j)\ln(r)+\ln\Gamma(C+B-k-j+1)-\ln\Gamma(C-k+1)-\ln\Gamma(B-j+1)-(C+B-k-j+1)\ln(1+r))}$$

Finally, algorithms exist for evaluating P_{pois} without evaluating $n!$ for large n . I include example S-Lang code for this, as well as example code for computing $P(s|CB)$ in the appendix. I've tested this code on cases which failed earlier and obtained good results.

Appendix — S-Lang Code that implements the revised computation of $P(s|CB)$

```
% pois(n,xmu) computes the Poisson probability of obtaining n for a mean xmu,
% without evaluating n! Explicitly.

define pois(n,xmu){

 variable i;
 variable sum=1.0;
 variable prb=exp(-1.0*xmu);
 for(i=1;i<=n;i++) sum*=xmu/double(i);
 prb*=sum;
 return prb;

}

% s_pdf_alt1 computes P(s|CB) without evaluating powers of very large numbers

define s_pdf_alt1(C,B,r,f,g,dp){

 % Compute the posterior probability distribution for source counts or intensity ;
 % marginalized over background, assuming non-informative gamma priors (alpha=1,
beta-0) ;

 variable i,j,k;
 variable C1 = C+1;
 variable B1 = B+1;
 variable CB1= C+B+1;

 variable rf_minus_g = r*f-g;
 variable one_plus_r = 1.0+r;

 variable K = Double_Type[C1,B1];
 variable S = Double_Type[C1,B1];

 % Compute MLE value for s, to estimate reasonable range for pdf ;
```

```
 variable S_mle = (r*C-B)/rf_minus_g;
 if(S_mle<=0.0) S_mle=30;
 variable ds = S_mle*dp;

 variable s = [0.0:3*S_mle:ds];
 variable ps= Double_Type[length(s)];
 variable fs, gs, pois_fs, pois_gs, tmps;


 for(i=0;i<length(s);i++){
 ps[i]=0.0;
 fs = f*s[i];
 gs = g*s[i];

  for(k=0;k<=C;,k++){
   pois_fs=pois(k,fs);
   for(j=0;j<=B;j++){
       pois_gs=pois(j,gs);
       tmps =(B-j)*log(r)+lngamma(CB1-k-j)-lngamma(C1-k)-lngamma(B1-j)-\
             (CB1-k-j)*log(one_plus_r);
       ps[i]+= exp(tmps)*pois_fs*pois_gs;

   }
  }
  ps[i]*=(rf-g);
 }

 % Renormalize psf just in case;

 return s,ps/(ds*sum(ps)),ds;

} % end of s_pdf_alt1;
```

Appendix F

Effectiveness of the Gregory-Loredo Algorithm for Detecting Temporal Variability in Chandra Data

Arnold Rots, CXC/SAO

2005-09-29

Abstract

We describe application of the Gregory-Loredo algorithm for detecting temporal variability in Chandra data. We have performed a test on 118 sources spanning the intensity range of 5 to 24000 photons over 102000 s.

We conclude that the G-L algorithm, when combined with a secondary criterion, is extremely robust, yielding a reliable variability indicator as well as a light curve with optimal resolution.

Introduction

This note describes using the Gregory-Loredo algorithm (1992, ApJ **398**, 146) to detect temporal variability in sources identified in the L3 pipeline (intra-ObI only), based on the event files.

Briefly, N events are binned in histograms of m bins, where m runs from 2 to m_{max}. The algorithm is based on the likelihood of the observed distribution n_1, n_2, …, n_m occurring. Out of a total number of m^N possible distributions the multiplicity of this particular one is $N!/(n_1!.n_2!\ ...\ n_m!)$. The ratio of the latter to the former provides the probability that this distribution came about by chance. Hence the inverse is a measure of the significance of the distribution. In this way we calculate an odds ratio for m bins versus a flat light curve. The odds are summed over all values of m to determine the odds that the source is time-variable. For more details, see the paper.

The method works very well on event data and is capable to deal with data gaps. We have added the capability to take into account temporal variations in effective area. As a byproduct, it delivers a light curve with optimal resolution.

Although the algorithm was developed for detecting periodic signals, it is a perfectly suitable method for detecting plain variability by forcing the period to the length of the observation.

Implementation

We have implemented the G-L algorithm as a standard C program, operating on simple ASCII files for ease of experimentation.

Input data consist of a list of event times and, optionally, good time intervals with, optionally,

normalized effective area (i.e., 1.0 for full exposure).

Two output files are created: odds ratios as a function of m and a light curve file which includes $\pm 3\sigma$ curves.

Usage:

[-i]	input file with event times (infile) [stdin]
[-om]	results file: probabilities as a function of m (outfile) [stdout]
[-ntrng]	maximum number of (good) time intervals allowed (integer, >= 0) [200]
[-n]	maximum number of events to be accepted (integer, >= 0) [70000]
[-tb]	Start of time range (-1 if not used) (double) [-1]
[-te]	End of time range (-1 if not used) (double) [-1]
[-rfrac]	fraction of events to be included in subsample (double, 0.000 to 1.000) [1]
[-rseed]	seed for random subsample selection (integer) [1]
[-olc]	resulting output file with light curve (outfile) [stdout]
[-log]	If yes, prints running log on standard out (boolean) [false]
[-mmin]	the minimum number of model bins to use (integer, 2 to 3000) [2]
[-mmax]	the maximum number of model bins (integer, 2 to 3000) [see below]
[-U]	display this message

If m_{max} is not explicitly specified, the algorithm is run twice. The first time all values of m are used, up to the minimum of 3000 and $(t_e - t_b) / 50$; i.e., variability is considered for all time scales down to 50 s which is about 15 times the most common ACIS frame time. The sum of odds $S(m) = \text{sum}\ (O(i),\ i = m_{min} .. m) / (m - m_{min} + 1)$ is calculated as a function of m and its maximum is determined. Then the algorithm is run again with m_{max} set to the highest value of m for which $S(m) > \max(S) / \sqrt{e}$.

The light curve that is generated by the program essentially consists of the binnings weighed by their odds ratios and represents the most optimal binning for the curve. The standard deviation σ is provided for each point of the light curve.

The program provides information on the total odds ratio O (or, rather, its $^{10}\log$), the corresponding probability of a variable signal, the m value with the maximum odds ratio and the odds-weighted first moment of m, as well as the characteristic time scales represented by these two values.

There is an ambiguous range of probabilities: $0.5 < P < 0.9$, and in particular the range between 1/2 and 2/3 (above 0.9 all is variable, below 0.5 all is non-variable). For this range we have developed a secondary criterion, based on the light curve, its average σ, and the average count rate. We calculate the fractions f_3 and f_5 of the light curve that are within 3σ and 5σ, respectively, of the average count rate. If $f_3 > 0.997$ AND $f_5 = 1.0$ for cases in the ambiguous range, the source is deemed to be non-variable.

Finally, the program assigns a *variability index*:

Variability Index	Condition	Comment
0	$P <= 1/2$	Definitely not variable
1	$1/2 < P < 2/3$ AND $f_3 > 0.997$ AND $f_5 = 1.0$	Not considered variable
2	$2/3 <= P < 0.9$ AND $f_3 > 0.997$ AND $f_5 = 1.0$	Probably not variable
3	$0.5 <= P < 0.6$	May be variable
4	$0.6 <= P < 2/3$	Likely to be variable
5	$2/3 <= P < 0.9$	Considered variable
6	$0.9 <= P$ AND $O < 2.0$	Definitely variable
7	$2.0 <= O < 4.0$	Definitely variable
8	$4.0 <= O < 10.0$	Definitely variable
9	$10.0 <= O < 30.0$	Definitely variable
10	$30.0 <= O$	Definitely variable

The code is structured such that all I/O is executed though three function calls. It is a simple matter to replace these three functions, e.g., to change to output in FITS format.

Test Results

The program was run on all 118 sources found by *wavdetect* in ObsId 635. The total time span of the observation was 102 ks and the sources varied between 5 and 24000 counts. The average time to run the program was 1.5 s per source. 71 sources were found to be variable with an odds ratio > 1.0 (probability > 0.5). Visual inspection of the light curves of all 118 sources found 54 that are variable, though there are a few borderline cases on either side of the divide.

Examples of output files and a summary of the test results are presented in the Appendix.

Analysis

The following table summarizes the number of variable sources detected, the false, and the missed detections, as a function of odds ratio and probability range.

Odds ratio range	Probability range	Good detections	False detections	Missed with secondary criterion	False with secondary criterion
1.0 – 2.0	1/2 – 2/3	2	7	1	0
2.0 – 9.0	2/3 – 0.9	5	10	2	0
> 9.0	> 0.9	47	0	0	0

Using G-L just by itself is problematic in the probability range 0.5 – 0.9, considering the required trade-off between missed detections and false detections. We solved this problem by designing the secondary criterion that is based on the fraction of the light curve within 3σ and 5σ. The final result is that three borderline variable sources are missed (with emphasis on *borderline*), but there are no spurious detections. However, any user who is concerned about missing potential candidates should be encouraged to inspect all sources with a variability index greater than 0.

The G-L algorithm, as expected, is pleasantly insensitive to the shape of the light curve, something that is a known problem with the current implementation of the K-S test. It is also not over-interpreting the data in low count rate sources, requiring a statistically significant deviation from a flat distribution before yielding an odds ratio greater than one.

The light curves (and the $\pm 3\sigma$ curves), in providing precisely the desired resolution, are of the kind that we would want to include in the L3 product package.

Attached are ten figures, highlighting the different types of cases. Dashed lines represent the $\pm 3\sigma$ curves. Figs. 1 – 7 provide a typical cross-section of the variable sources. Fig. 8 shows one of the cases that cannot very well be handled by simple statistics: it fails all criteria, but there appears to be a definite trend; obviously, the judgment as to what is "definite" is subjective and we do not consider the source variable. To some extent, the example in Fig. 5 is in that category, too.

Finally, the test has brought to light the issue of time variable exposure (or effective area). It appears in a number of sources with characteristic times that are harmonics of one of the two dither periods (707 and 1000 s). Figs. 9 and 10 show that this can be properly taken care of by providing the program with normalized effective area as a function of time.

Conclusion

We conclude that G-L provides a robust algorithm for detecting temporal variability that is insensitive to the type and shape of variability and that takes properly into account the uncertainties in the count rate, requiring a statistically significant departure from a flat count rate for it to declare variability. The light curves provided by the program appear to be near-optimal for what we intend to present to users.

The addition of the secondary criterion results in a reliable test, though careful users may want to inspect the light curves of all sources with a non-zero variability index.

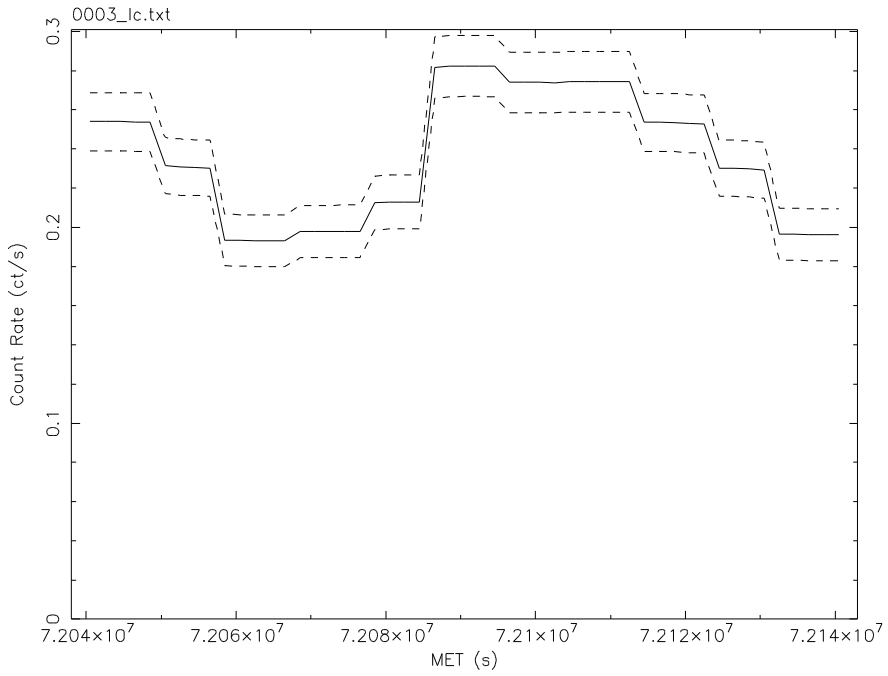


Fig. 1. Source 3: 24093 counts. Even though the SNR is high, the resolution of the light curve is fairly low since higher resolution is not warranted by its shape.

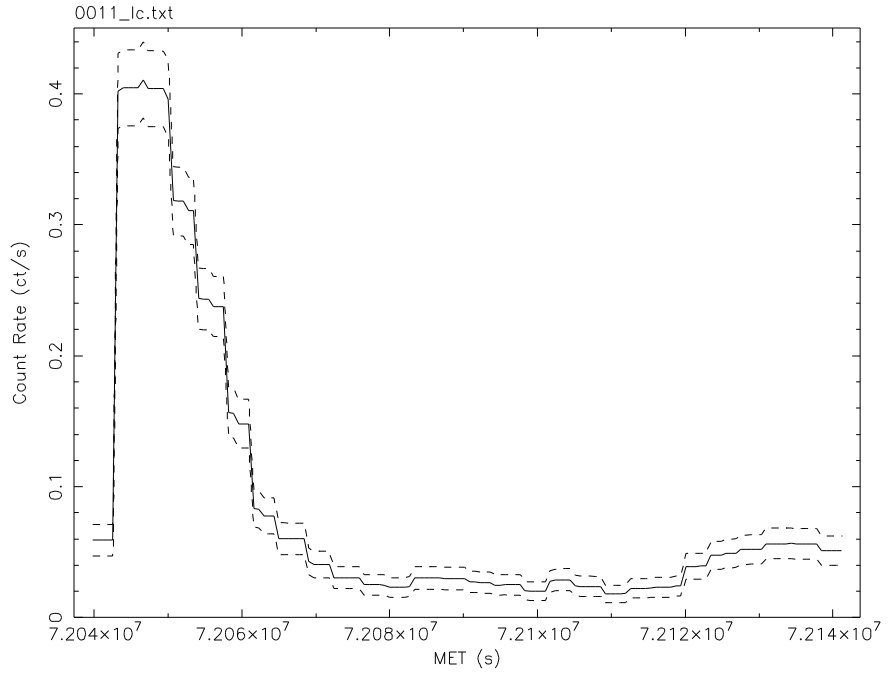


Fig. 2. Source 11: 8697 counts. The timescale of the changes in this source are very much shorter than in source 3; hence the resolution is higher.

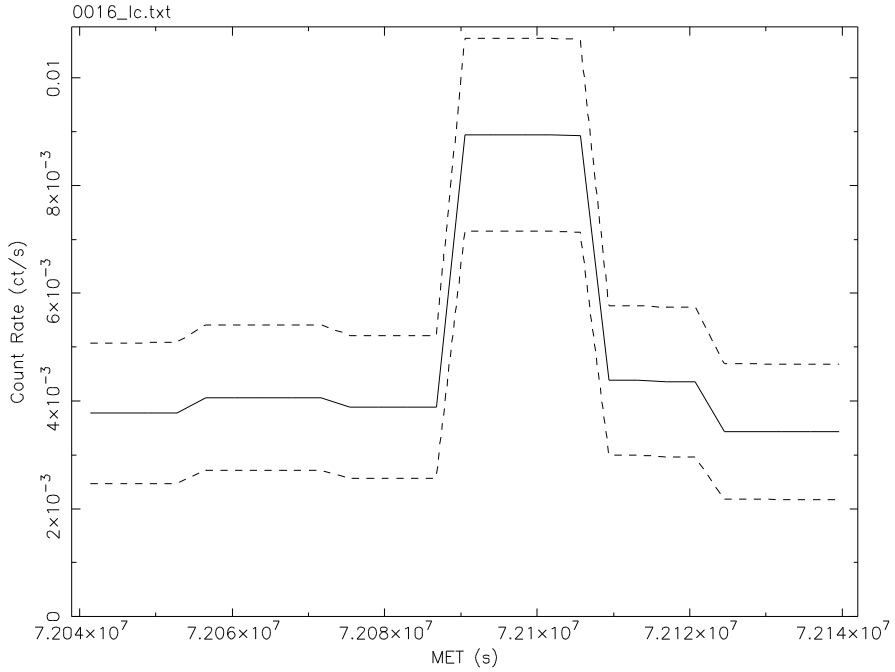


Fig. 3. Source 16: 484 counts. The odds ratio was high (^{10}log(odds)=8.6).

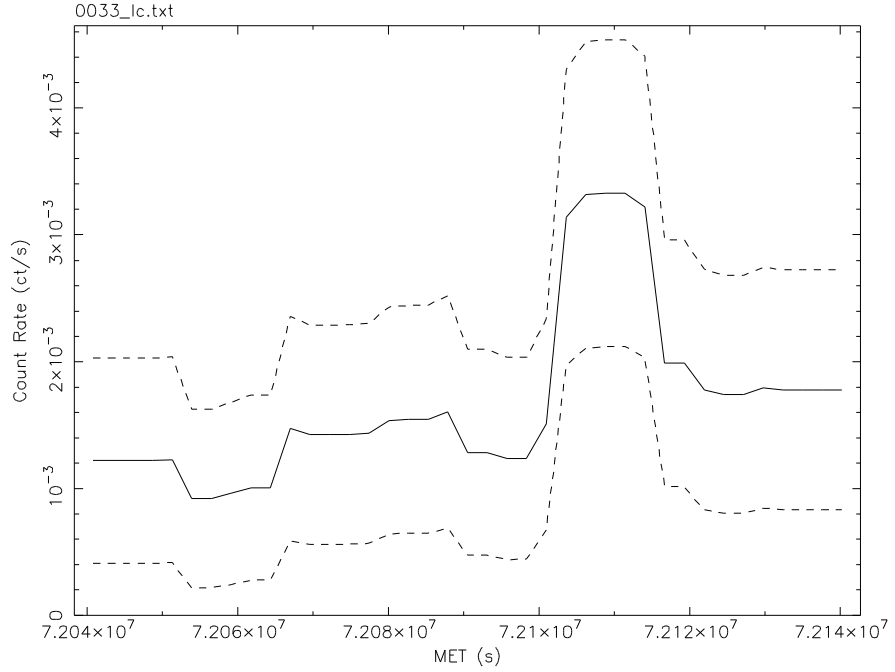


Fig. 4. Source 33: 171 counts.

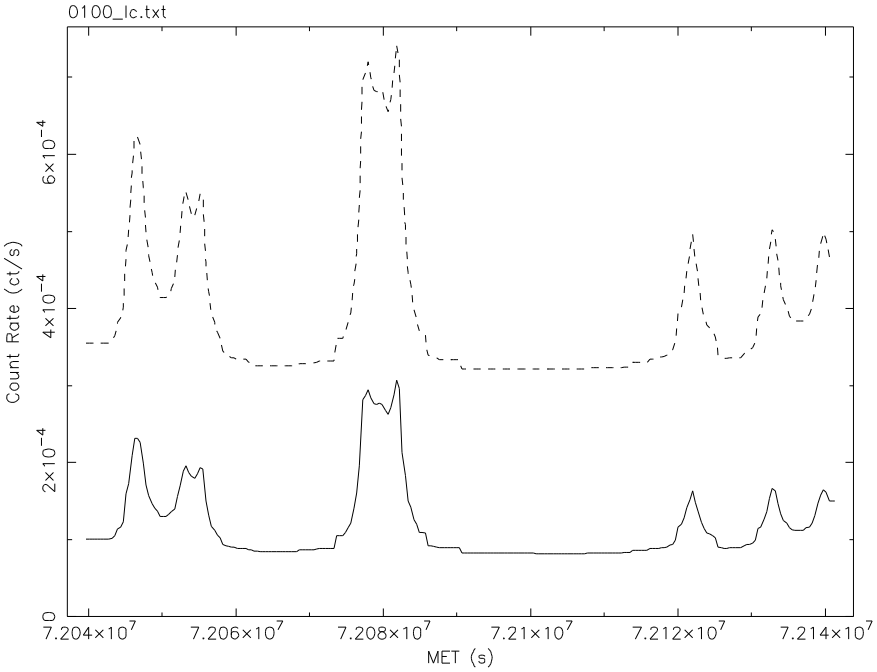


Fig. 5. Source 100: 12 counts. The odds ratio is only 2.2 and on the basis of the 3σ fraction it should be rejected, but it certainly looks variable, although it is a borderline case.

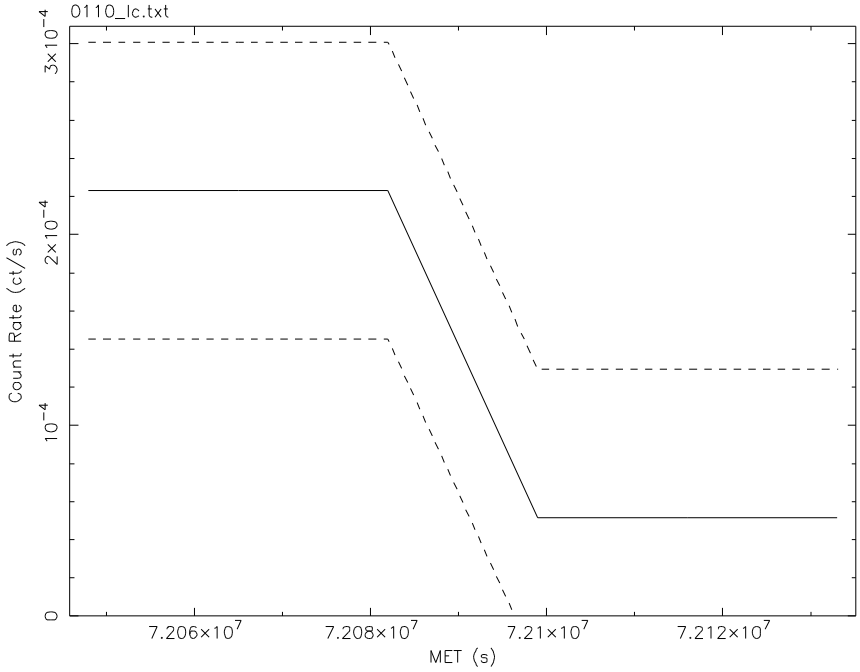


Fig. 6. Source 110: 14 counts. A low count rate, but variable, nevertheless.

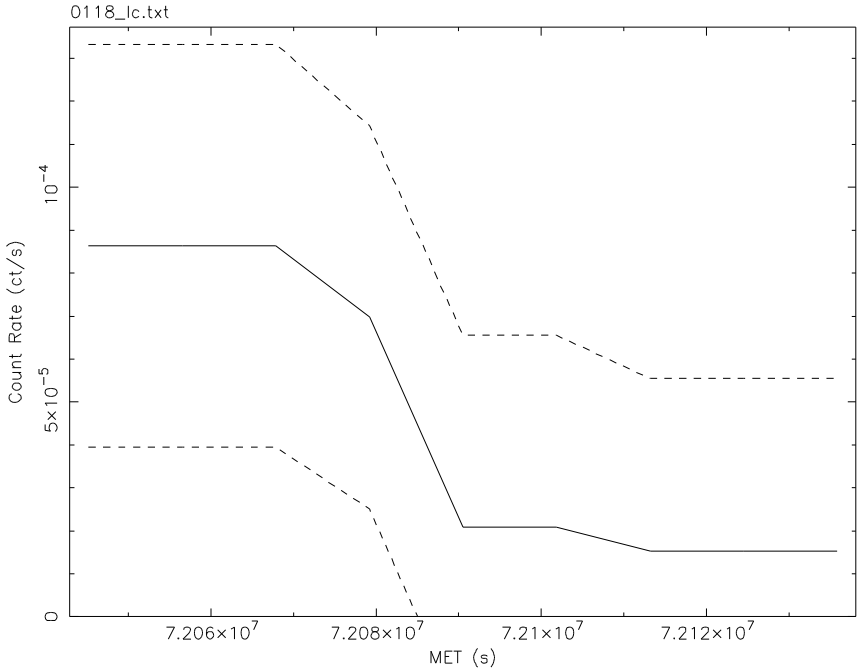


Fig. 7. Source 118: 5 counts. With only 5 counts I would not stake my reputation on this one. f_3 backs that up: probably not variable.

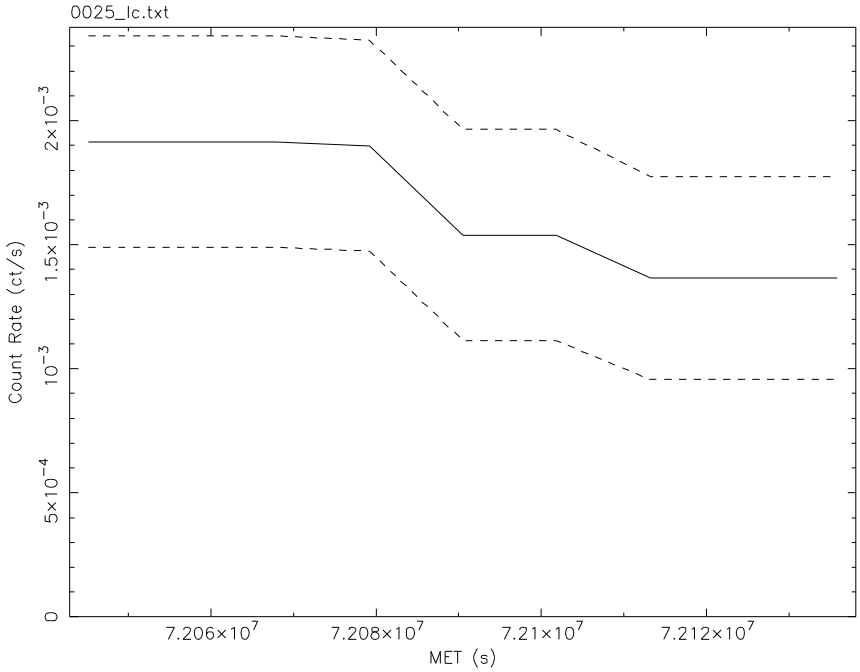


Fig. 8. Source 25: 170 counts. This is an example of a borderline case where there is no statistically significant change while there is yet an unmistakable trend. I still would not consider this source variable.

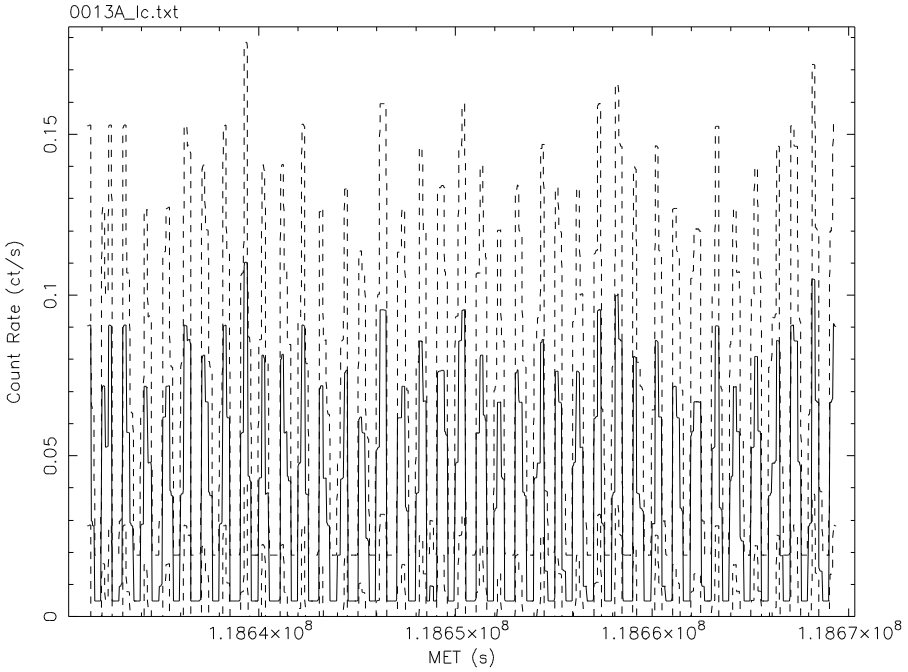


Fig. 9. Source 13 of ObsId 1575 (1376 counts), showing clearly the 707 s dither period.

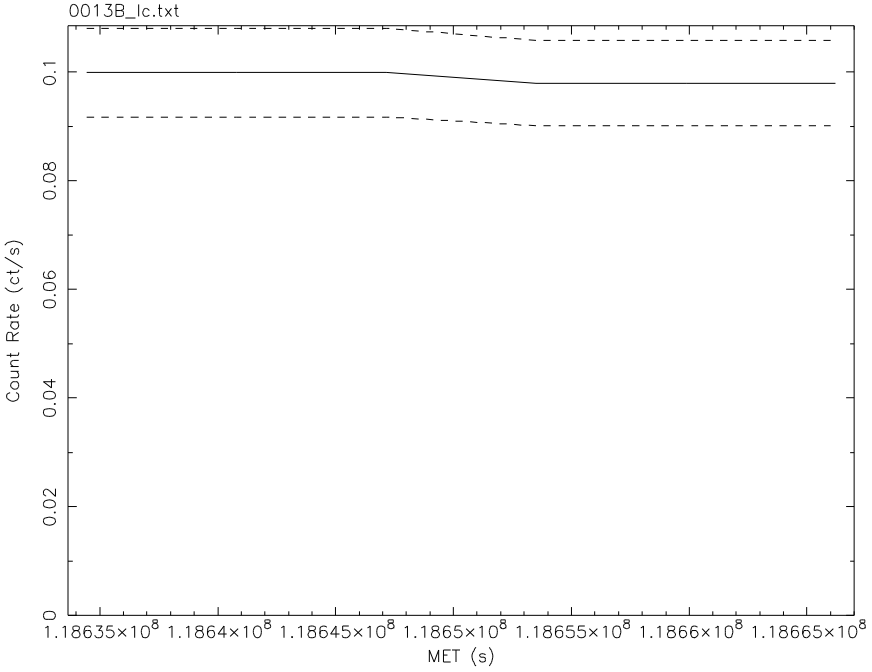


Fig. 10. Source 13 of ObsId 1575 taking into account the normalized effective area as a function of time.

Appendix

Example of output file with odds ratios as a function of *m*.

```
# >>> G L V A R Y <<<
#
# Data file: 0001_
# Time range: 72039530.664659 to 72141499.614917
# Time zero point: 0.000000
#
# Run with 8674 SINGLE events, total integration time 101968.950258 sec
# Total time span covered: 101968.950258 sec
# 10Log ( Odds for variable signal ): 39.753
# Probability of a variable signal : 1.000
# mmin = 2, mmax = 14
# Fraction of light curve within 3 sigma of average rate: 0.333333, 5 sigma:
0.880952
# Variability index: 10
#
# First moment over m, characteristic time: 9.00 11331.7
# m with maximum odds, characteristic time: 9 11329.9
#
# m Probability Sum(Odds)
#
2 0.000000000000 0.000000
3 0.000000000358 0.000000
4 0.000000000000 0.000000
5 0.000000000000 0.000000
6 0.001739450095 0.004523
7 0.000000000000 0.003769
8 0.000000000000 0.003230
9 0.997008295938 1.622965
10 0.000000000000 1.442636
11 0.000000000000 1.298372
12 0.001252253609 1.181818
13 0.000000000000 1.083333
14 0.000000000000 1.000000
```

Example of light curve output file.

```
# >>> G L V A R Y <<<
#
# Data file: 0001_
# Time range: 72039530.664659 to 72141499.614917
# Time zero point: 0.000000
#
# Run with 8674 SINGLE events, total integration time 101968.950258 sec
# Total time span covered: 101968.950258 sec
# 10Log ( Odds for variable signal ): 39.753
# Probability of a variable signal : 1.000
# mmin = 2, mmax = 14
# Fraction of light curve within 3 sigma of average rate: 0.333333, 5 sigma:
0.880952
# Variability index: 10
#
# Time <F> sigma <F>-3*sigma <F>+3*sigma
#
72040744.58 0.074508293 0.002434692 0.067204217 0.081812369
72043172.42 0.074508293 0.002434692 0.067204217 0.081812369
72045600.25 0.074508293 0.002434692 0.067204217 0.081812369
72048028.09 0.074502112 0.002434583 0.067198362 0.081805863
72050455.92 0.074502112 0.002434583 0.067198362 0.081805863
72052883.75 0.073710951 0.002423006 0.066441932 0.080979970
72055311.59 0.073710951 0.002423006 0.066441932 0.080979970
72057739.42 0.073710337 0.002423000 0.066441338 0.080979336
72060167.26 0.073710337 0.002423000 0.066441338 0.080979336
72062595.09 0.075820101 0.002453685 0.068459046 0.083181156
72065022.93 0.075825399 0.002453778 0.068464064 0.083186734
72067450.76 0.075825399 0.002453778 0.068464064 0.083186734
72069878.60 0.075825399 0.002453778 0.068464064 0.083186734
72072306.43 0.075825399 0.002453778 0.068464064 0.083186734
72074734.27 0.082169759 0.002542570 0.074542048 0.089797470
72077162.10 0.082169759 0.002542570 0.074542048 0.089797470
72079589.93 0.082169759 0.002542570 0.074542048 0.089797470
72082017.77 0.082176529 0.002542683 0.074548480 0.089804577
72084445.60 0.082176529 0.002542683 0.074548480 0.089804577
72086873.44 0.081033740 0.002527049 0.073452594 0.088614886
72089301.27 0.081033740 0.002527049 0.073452594 0.088614886
72091729.11 0.081020337 0.002526882 0.073439691 0.088600982
72094156.94 0.081020337 0.002526882 0.073439691 0.088600982
72096584.78 0.073987792 0.002427082 0.066706545 0.081269039
72099012.61 0.073977932 0.002426912 0.066697197 0.081258667
72101440.45 0.073977932 0.002426912 0.066697197 0.081258667
72103868.28 0.073977932 0.002426912 0.066697197 0.081258667
72106296.11 0.073977932 0.002426912 0.066697197 0.081258667
72108723.95 0.123074792 0.003017199 0.114023193 0.132126390
72111151.78 0.123074792 0.003017199 0.114023193 0.132126390
72113579.62 0.123074792 0.003017199 0.114023193 0.132126390
72116007.45 0.123040355 0.003016742 0.113990131 0.132090580
72118435.29 0.123040355 0.003016742 0.113990131 0.132090580
72120863.12 0.093064133 0.002684086 0.085011876 0.101116390
72123290.96 0.093064133 0.002684086 0.085011876 0.101116390
72125718.79 0.093011353 0.002683583 0.084960604 0.101062103
```

```
72128146.63 0.093011353 0.002683583 0.084960604 0.101062103
72130574.46 0.088265162 0.002623430 0.080394871 0.096135453
72133002.29 0.088260601 0.002623359 0.080390524 0.096130679
72135430.13 0.088260601 0.002623359 0.080390524 0.096130679
72137857.96 0.088260601 0.002623359 0.080390524 0.096130679
72140285.80 0.088260601 0.002623359 0.080390524 0.096130679
```

Summary of Test Results

Source	Counts	Var Inx	log(Odds)	Probab-ility	f3	f5
0029	257	0	-1.318	0.046	1.0000	1.0000
0030	700	0	-1.314	0.046	1.0000	1.0000
0088	251	0	-1.165	0.064	1.0000	1.0000
0086	250	0	-1.071	0.078	1.0000	1.0000
0012	1156	0	-1.042	0.083	1.0000	1.0000
0064	374	0	-0.958	0.099	1.0000	1.0000
0040	116	0	-0.937	0.104	1.0000	1.0000
0027	2710	0	-0.769	0.146	0.8000	1.0000
0077	89	0	-0.689	0.170	1.0000	1.0000
0045	72	0	-0.669	0.176	1.0000	1.0000
0068	157	0	-0.653	0.182	1.0000	1.0000
0076	69	0	-0.401	0.284	1.0000	1.0000
0063	55	0	-0.314	0.327	1.0000	1.0000
0025	170	0	-0.313	0.327	1.0000	1.0000
0079	56	0	-0.303	0.333	1.0000	1.0000
0073	33	0	-0.274	0.347	1.0000	1.0000
0039	76	0	-0.265	0.352	1.0000	1.0000
0089	35	0	-0.259	0.355	1.0000	1.0000
0083	106	0	-0.245	0.363	1.0000	1.0000
0099	37	0	-0.233	0.369	1.0000	1.0000
0057	34	0	-0.221	0.376	1.0000	1.0000
0046	94	0	-0.177	0.399	1.0000	1.0000
0069	30	0	-0.175	0.400	1.0000	1.0000
0075	100	0	-0.155	0.412	1.0000	1.0000
0106	58	0	-0.155	0.412	1.0000	1.0000
0061	31	0	-0.150	0.415	1.0000	1.0000
0091	48	0	-0.147	0.416	1.0000	1.0000
0047	160	0	-0.137	0.422	1.0000	1.0000
0087	36	0	-0.136	0.423	1.0000	1.0000
0090	20	0	-0.123	0.430	1.0000	1.0000
0071	28	0	-0.104	0.440	1.0000	1.0000
0067	18	0	-0.102	0.442	1.0000	1.0000
0092	18	0	-0.099	0.443	1.0000	1.0000
0094	38	0	-0.098	0.444	1.0000	1.0000
0111	17	0	-0.091	0.448	1.0000	1.0000
0085	20	0	-0.079	0.455	1.0000	1.0000
0097	14	0	-0.068	0.461	1.0000	1.0000
0113	14	0	-0.061	0.465	1.0000	1.0000
0093	13	0	-0.060	0.466	1.0000	1.0000
0105	12	0	-0.053	0.470	1.0000	1.0000
0070	49	0	-0.037	0.478	1.0000	1.0000
0098	24	0	-0.031	0.482	1.0000	1.0000
0050	55	0	-0.018	0.490	1.0000	1.0000
0065	184	0	-0.018	0.490	0.9997	1.0000
0082	34	0	-0.018	0.489	1.0000	1.0000
0038	124	0	-0.013	0.492	1.0000	1.0000
0114	7	0	-0.009	0.495	1.0000	1.0000
0108	9	1	0.012	0.507	1.0000	1.0000
0112	9	1	0.030	0.518	1.0000	1.0000
0109	12	1	0.033	0.519	1.0000	1.0000

0101	46	1	0.036	0.521	1.0000	1.0000
0117	17	1	0.145	0.583	1.0000	1.0000
0095	25	1	0.153	0.587	1.0000	1.0000
0107	39	1	0.161	0.592	1.0000	1.0000
0034	104	1	0.295	0.664	1.0000	1.0000
0103	28	2	0.334	0.683	1.0000	1.0000
0100	12	2	0.349	0.691	1.0000	1.0000
0115	21	2	0.371	0.701	1.0000	1.0000
0044	162	2	0.385	0.708	1.0000	1.0000
0102	33	2	0.475	0.749	1.0000	1.0000
0036	314	2	0.486	0.754	0.9998	1.0000
0081	222	2	0.504	0.762	0.9995	1.0000
0104	82	2	0.509	0.763	1.0000	1.0000
0118	5	2	0.582	0.793	1.0000	1.0000
0072	131	2	0.624	0.808	0.9993	1.0000
0066	63	2	0.646	0.816	1.0000	1.0000
0048	136	2	0.739	0.846	1.0000	1.0000
0051	167	3	0.071	0.541	0.8095	1.0000
0018	377	5	0.651	0.817	0.6667	1.0000
0084	16	5	0.849	0.876	0.9838	1.0000
0024	542	5	0.855	0.877	0.8000	0.9111
0031	287	6	1.006	0.910	0.3333	1.0000
0110	14	6	1.079	0.923	0.0000	1.0000
0059	64	6	1.082	0.924	0.7778	1.0000
0096	93	6	1.220	0.943	0.9995	1.0000
0033	171	6	1.243	0.946	0.8718	0.8974
0042	222	6	1.265	0.948	0.9997	1.0000
0058	530	6	1.308	0.953	0.7917	1.0000
0060	58	6	1.468	0.967	0.9136	0.9753
0074	47	6	1.787	0.984	0.9663	1.0000
0049	67	7	2.250	0.994	1.0000	1.0000
0037	100	7	2.269	0.995	0.9103	0.9359
0080	135	7	2.919	0.999	0.8333	0.8611
0028	336	7	2.972	0.999	0.8333	0.8333
0043	149	7	3.081	0.999	0.9352	0.9769
0035	129	7	3.111	0.999	0.8730	0.9524
0041	2040	7	3.154	0.999	0.7222	0.7222
0062	38	7	3.164	0.999	0.8718	0.8718
0116	14	7	3.434	1.000	0.9989	0.9993
0052	98	7	3.498	1.000	0.0000	1.0000
0032	291	8	4.030	1.000	0.4815	1.0000
0023	220	8	4.096	1.000	0.9682	1.0000
0056	68	8	4.968	1.000	0.9444	0.9921
0055	69	8	5.490	1.000	0.6190	0.8095
0021	659	8	6.178	1.000	0.3333	0.3333
0053	57	8	6.292	1.000	0.9815	0.9815
0015	4847	8	8.525	1.000	0.0000	0.6667
0016	484	8	8.626	1.000	0.8148	0.8148
0054	138	9	11.108	1.000	0.9799	0.9856
0019	3247	9	21.207	1.000	0.5000	0.6667
0014	1323	9	29.815	1.000	0.7424	0.9091
0026	199	10	30.290	1.000	0.9603	0.9683
0001	8674	10	39.896	1.000	0.3333	0.8810
0078	328	10	50.032	1.000	0.9000	0.9333
0007	1115	10	58.955	1.000	0.5333	0.8000
0020	1025	10	72.159	1.000	0.7667	0.9167
0003	24093	10	77.424	1.000	0.1569	0.4314

0006	2831	10	79.677	1.000	0.3333	0.6889
0004	6775	10	93.660	1.000	0.1282	0.4872
0022	651	10	117.251	1.000	0.1961	0.2941
0008	4445	10	142.491	1.000	0.8756	0.9813
0009	1106	10	181.623	1.000	0.2444	0.4444
0010	1102	10	231.745	1.000	0.8989	0.9579
0017	785	10	367.387	1.000	0.1149	0.1724
0013	1401	10	523.414	1.000	0.1560	0.2766
0005	2992	10	627.994	1.000	0.0575	0.0575
0011	8697	10	2158.204	1.000	0.0333	0.0333
0002	17450	10	5767.780	1.000	0.0000	0.0172

Appendix G

PSF Characterization Summary

Margarita Karovska, CXC/SAO

31 January 2007

I. Statistical Errors in Low-count PSFs

We carried out an extensive study of the statistical errors in the PSF calculated with SAOsac. The simulations were carried out using ChaRT/SAOsac for an axis-source at 5 monochromatic energies. The number of counts in these simulations varied from ~10 to 10^6. Each simulation was run 20 times with a different random number inputs. The ECFs and PSF sizes were calculated for each of these, and the statistical error were estimated.
The results show that the accuracy in the PSF is dominated by the statistical errors (for sources with few hundred counts) rather then by the limited accuracy of the HRMA model or the instrumental effects. The statistical error is over 50% for sources with few dozen counts.

II. Density vs. Output Total Counts

We investigated the relation between the ChaRT/SAOsac input density values, and the totaloutput PSF counts in the simulated PSFs. The aim is to calibrate the density needed as an input for SAOsac simulations of PSFs per source for the L3 pipeline. Simulations were run for several monochromatic energies corresponding the selected L3 energy bands [e.g. monochromatic PSFs centered on the center energy of the soft (0.4 keV), medium (1.25 keV), and the hard (3.8 keV) bands]. These simulations we carried out for counts ranging from few tens to ~10^6 counts. The results were summarized in a "look-up" table that will be used in determining the input density in the SAOsac L3 simulation for each detected source.

III. Characterization of Systematic ECF Error for On- and Off-axis Sources

Simulations were carried out in order to characterize the systematic error in the ECF estimates for PSFs obtained using a monochromatic energy instead using a source spectrum. For example, SAOsac PSF simulations were carried out using a spectrum for an absorbed power law with $N_H = 3\times10^{20}$ and $\gamma = 1.7$ for 2×10^6, 4×10^5, and 2×10^3 counts. We calculated the ECFs and compared the results with the ECFs calculated from simulations of monochromatic PSFs at 2.3 keV (center of L3 broad band). The monochromatics PSF simulations were obtained using SAOsac with 5 input densities. The resulting PSFs contained total counts ranging from 30 to 3×10^5 counts. Comparison of the calculated ECFs for the monochromatic sources with those calculated using a source was carried out at 50%, 90%, and 95% ECFs. The results are very similar at 50% ECF (errors <5%), however, they show significant differences at 90%, and 95% (e.g. ~20%). Although these differences cannot be ignored for very bright sources, for most of the sources containing less then ~1000, the errors are dominated by the differences

in the ECFs caused by low-count statistics.
We conclude that for most of the sources in our catalog using monochromatic PSFs calculated in the band center will be a satisfactory approximation.
We also carried out a study using PSF simulations of off-axis sources to test if we can use monochromatic approximation for estimating the ECFs in the L3 pipeline. For example, we run PSF simulations for a source at 10' off-axis angle, using a PL spectrum (as in the case of on-axis source) and a monochromatic PSFs at 2.3 keV (center of L3 broad band). We used a high counts simulations (10^5 and 10^6 photons) to identify what is the difference in the case where the PSF counts are not dominated by the statistics. We compared the 50%, 90%, and 95% ECFs and find that the difference is ~5%. This supports the conclusion from the on-axis study, that for the purposes of the catalog (release 1) it should be acceptable to use the monochromatic PSF approximation.

IV. Differences in Simulated and Observed Off-axis Sources: ECFs and Centroids

We carried out a comparison between simulated sources at various off-axis angles with a set of off-axis calibration observations of ArLac and other bright unresolved sources. SAOsac PSF simulations were carried out for sources on-axes, and at 14', 6' off-axis angles using the ArLac spectrum and a set of exposure times. The input exposure time was also selected to match the number of counts in the Ar Lac cal data (~4000 counts). We calculated ECFs and centroids using the simulated PSFs and the observed data. The results show that the differences in 90% ECFs between the observations and the simulated PSFs are less then 5% in most of the cases. The difference in centroids is ~3.5" at 14' off-axis, and ~2" at 6' off axis.
This is acceptable for the purpose of the catalog.

V. Convolution Factors for HRC and ACIS

We carried an extensive study of the simulated PSF comparison with on- and off-axis observations of bright sources observed with HRC-I and ACIS-I. The goal was to determine the convolution factor for the HRMA simulated PSFs that we need to use to simulate a realistic PSF, and derive source extent for the sources detected in the L3 pipeline.