MIT Kavli Institute   Chandra X-Ray Center

MEMORANDUM
April 11, 2012

To: Jonathan McDowell, SDS Group Leader
From: Glenn E. Allen, SDS

Subject: Afterglow and hot-pixel spec
Revision: 2.10
URL: http://space.mit.edu/CXC/docs/docs.html#aft
File: /nfs/cxc/h2/gea/sds/docs/memos/afterglow_spec_2.10.tex

1 Afterglows and Hot pixels

1.1 Description

A cosmic-ray "afterglow" is produced when a large amount of charge is deposited on a CCD by a cosmic ray. Most of the charge is clocked off of the CCD in a single frame. However, a small amount can be captured in charge traps that release the charge relatively slowly. As a result, a sequence of events can appear in a single detector pixel over a few frames as the trapped charge is released.

To date, two algorithms have been used by the CXC to identify cosmic-ray afterglows. The first algorithm was implemented in the CIAO tool `acis_detect_afterglow` and used for pipeline processing from the summer of 2000 to the fall of 2004. This algorithm searches for occasions when events are detected in two or more consecutive frames on the same CCD pixel. While the events are flagged as potential cosmic-ray afterglows and excluded from Level 2 event-data files, the corresponding pixels are not included in the observation-specific bad-pixel file. This algorithm finds many afterglow events, but at the expense of discarding X-ray events associated with real astrophysical sources. The fraction of the source events that are discarded depends on the brightness and variability of a source.

In an attempt to minimize the loss of source events, a second algorithm was developed and implemented in the CIAO tool `acis_run_hotpix`, which is a script that executes the tools `acis_find_hotpix`, `acis_classify_hotpix` and `acis_build_badpix`. The second algorithm searches for detector pixels that have an unusually large number of events that occur over a short period of time. Suspicious pixels are added to the observation-specific bad-pixel file only if the neighboring pixels do not have a significant excess of events. This condition helps insure that events associated with dithered sources are not discarded. Events associated with afterglows are flagged and excluded from Level 2 event-data files. The newer algorithm has been used for pipeline processing (and reprocessing) since the fall of 2004. While it is relatively gentle on astrophysical sources, it does let some afterglows "slip through the cracks." The afterglow detection efficiency depends on the number of events in the afterglow. The efficiency declines quickly as the number of events in an afterglow drops below about eight.

This spec describes, in part, a third afterglow-detection algorithm, which is implemented in the tool `acis_find_afterglow`. Like the second algorithm, it is designed to avoid discarding events associated with real astrophysical sources. It is also designed to enhance the detection efficiency for afterglows that have as few as four events. The principal change between the second and third afterglow-detection algorithms is that the third algorithm searches for afterglows using the events in a short, sliding time window instead of using the events from the entire duration of an observation. The algorithm searches in three dimensions instead of two. Another change is that the algorithm is designed to try to prevent bright sources in the field of view from reducing the afterglow detection efficiency.

The latter change has also been incorporated into a new hot-pixel detection algorithm, which is included here. Since the tool `acis_find_afterglow` includes both the afterglow and hot-pixel detection algorithms, it supersedes the tools `acis_find_hotpix` and `acis_classify_hotpix`. It also supersedes `acis_detect_afterglow` because it is sensitive to afterglows with four or more events.

1.2 Input

1. A Level 1 event-data file (acis*evt1.fits)
2. A Level 1 observation-specific bad-pixel file (acis*bpix1.fits)
3. A Level 1 mask file (acis*msk1.fits)
4. A Level 1 exposure statistics file (acis*stat1.fits)

1.3 Output

1. An updated observation-specific bad-pixel file

1.4 Parameters

1. `infile,s,a,"",,,"Name of input event-data file"`
2. `outfile,s,a"",,,"Name of output bad-pixel file"`
3. `badpixfile,s,a,"",,,"Name of input bad-pixel file"`
4. `maskfile,s,a,"",,,"Name of input mask file"`
5. `statfile,s,a,"",,,"Name of input exposure-statistics file"`
6. `expnowindow,i,h,10,1,100,"Number of frames in the sliding time window"`
7. `probthresh,r,h,0.001,1.0e-10,0.1,"Minimum post-trials significance of potential afterglows (e.g., 1 sigma = 0.159, 90% = 0.1, 2 sigma = 0.0228, 99% = 0.01 and 3 sigma = 0.00135)"`
8. `cntthresh,i,h,4,2,10,"Minimum number of events in an afterglow"`
9. `regwidth,i,h,7,3,255,"Size of reference region (e.g., 7 pixels × 7 pixels)"`
10. `nfpixreg,i,h,32,16,256,"Size of region used to calculate the fluence"`
11. `nfrepeat,i,h,10,1,30,"Number of iterations during the calculation of the fluence"`
12. `tolerance,r,h,1.0e-15,1.0e-16,1.0e-6,"Tolerance"`
13. `runhotpix,b,h,yes,,,"Search for hot pixels?"`
14. `clobber,b,h,no,,,"Overwrite output file if it exists?"`
15. `verbose,i,h,0,0,5,"Amount of messages produced (0=none, 5=a lot)"`
16. `mode,s,h,"ql",,,`

1.5 Processing

In the standard ACIS pipeline, the afterglow-detection algorithm is used after the bias(es) has been searched for bad bias values, after the bias-parity error file(s) has been searched for bad pixels and the "FEP0" problem, and before the event data is searched for hot pixels. The afterglow and hot-pixel detection algorithms are summarized below.

Verify that the `infile`, `badpixfile`, `maskfile`, and `statfile` exist. If `clobber=no`, then verify that the `outfile` does not exist. Verify that the `infile` has `READMODE=TIMED`. The afterglow and hot-pixel detection algorithms are not appropriate for `READMODE=CONTINUOUS`. Verify that the values of the parameters `expnowindow`, `probthresh`, `cntthresh`, `regwidth`, `nfpixreg`, `nfrepeat`, and `tolerance` are in the valid ranges for these parameters. Note that `regwidth` must be an odd number. The only valid values for the parameter `nfpixreg` are 16, 32, 64, 128, and 256, values for which a node can be sub-divided into an integer number of equal-sized regions.

1.5.1 Afterglows

1. Exclude "invalid" pixels[1] from the search.

2. To improve the performance of the algorithm, perform more than one pass through the data. In the first pass, potential afterglow events are identified as suspicious using a minimum set of criteria. Events i and j may be part of an afterglow if the following four conditions are satisfied.

$$\begin{aligned} \texttt{CCD\_ID}_i &= \texttt{CCD\_ID}_j, & (1)\\ \texttt{CHIPX}_i &= \texttt{CHIPX}_j, & (2)\\ \texttt{CHIPY}_i &= \texttt{CHIPY}_j, \text{ and} & (3)\\ |\texttt{EXPNO}_i - \texttt{EXPNO}_j| &\leq \texttt{expnowindow}. & (4)\end{aligned}$$

3. During a subsequent pass through the data, the ith, $(i+1)$th, $\ldots$, $(i+n)$th events on a pixel are identified as an afterglow if each consecutive pair of events in the set satisfies equations 1–4 and if

$$\begin{aligned} N_{\rm evt}^{\rm aft} &\geq \texttt{cntthresh}, & (5)\\ P_{\rm post}^{\rm aft} &< \texttt{probthresh}, \text{ and} & (6)\\ P_{\rm post}^{\rm ref} &\geq \texttt{probthresh}, & (7)\end{aligned}$$

where the post-trials probabilities[2]

$$P_{\rm post}^{\rm aft} = 1 - \left(1 - P_{\rm pre}^{\rm aft}\right)^{N_{\rm trial}^{\rm aft}}, \qquad (9)$$

[1]Here an invalid pixel is one that has `SAMP_CYC` = 0 in the `maskfile`, that has `STATUS` bit 15 (streak) set in the `infile`, but only for the duration of the streak, or that has one or more of the following `STATUS` bits set in the `badpixfile`.

Bit	Description	Notes
0	bad pixel	
1	bad columns	
2	bias-parity error	only for the duration of the error
3	bias = 4095	
4	bias = 4094	
11	mid-chip node boundary	
13	FEP0 problem	only for the duration of the error
15	afterglow	only for the hot-pixel algorithm and only for the duration of the afterglow
16	bad bias value	

Note that the `STATUS` bits are numbered from 0 to 31. It is not necessary to ignore pixels that have bias values of 4096 (i.e., are missing data) because biases with such problems are adjusted on the ground. If they are not adjusted, then all events on pixels with a bias value of 4096 are discarded.

[2]Equations 9 and 10 should be computed as shown only if $N_{\rm trial} P_{\rm pre} \geq 0.693148$. Otherwise, if $N_{\rm trial} P_{\rm pre} < 0.693148$, then use

$$P_{\rm post} = \frac{N_{\rm trial} P_{\rm pre}}{1!} - \frac{N_{\rm trial}\left(N_{\rm trial}-1\right) P_{\rm pre}^2}{2!} + \frac{N_{\rm trial}\left(N_{\rm trial}-1\right)\left(N_{\rm trial}-2\right) P_{\rm pre}^3}{3!} - \ldots - \frac{N_{\rm trial}\left(N_{\rm trial}-1\right)\ldots\left(N_{\rm trial}-15\right) P_{\rm pre}^{16}}{16!} \qquad (8)$$

to avoid some concerns about numerical precision. Equation 8 has a relative precision of about 2×10^{-16} or better.

$$P_{\rm post}^{\rm ref} = 1 - \left(1 - P_{\rm pre}^{\rm ref}\right)^{N_{\rm trial}^{\rm ref}}, \tag{10}$$

the pre-trials probability $P_{\rm pre}^{\rm aft}$ is given by the series[3]

$$P_{\rm pre}^{\rm aft} = \left[\frac{1}{2}\frac{\left(N_{\rm bgd}^{\rm aft}\right)^{N_{\rm evt}^{\rm aft}}}{N_{\rm evt}^{\rm aft}!} + \left(\sum_{n=N_{\rm evt}^{\rm aft}+1}^{\infty} \frac{\left(N_{\rm bgd}^{\rm aft}\right)^{n}}{n!}\right)\right] e^{-N_{\rm bgd}^{\rm aft}}, \tag{12}$$

$N_{\rm evt}^{\rm aft}$ is the number of events in the potential afterglow, the number of background events for the potential afterglow

$$N_{\rm bgd}^{\rm aft} = \frac{F}{\texttt{SAMP\_CYC}_{\rm aft}}\left(\frac{N_{\rm frame}^{\rm aft}}{N_{\rm frame}^{\rm tot}}\right), \tag{13}$$

$\texttt{SAMP\_CYC}_{\rm aft}$ is the sample cycle for the pixel on which the potential afterglow occurred, $N_{\rm frame}^{\rm aft}$ is the number of valid frames[4] in the afterglow, $N_{\rm frame}^{\rm tot}$ is the total number of valid frames for the CCD, the number of trials[5] $N_{\rm trial}^{\rm aft}$ is estimated to be

$$N_{\rm trial}^{\rm aft} = \sum_k N_{{\rm pix},k}^{\rm ccd}\left(N_{{\rm frame},k}^{\rm tot} - \texttt{expnowindow} - 1\right), \tag{14}$$

$N_{{\rm pix},k}^{\rm ccd}$ is the number of valid pixels[1] for the kth CCD (i.e. $N_{{\rm pix},k}^{\rm ccd} = 1024 \times 1024$ less the number of invalid pixels), $N_{{\rm frame},k}^{\rm tot}$ is the total number of valid frames for the kth CCD, the number of trials $N_{\rm trial}^{\rm ref}$ is the number of candidate afterglows that satisfy equations 1–6, the pre-trials probability $P_{\rm pre}^{\rm ref}$ that the event fluence in the reference region is consistent with the event fluence on the entire node (i.e., that the potential afterglow or hot pixel is not part of a dithered source) is given by the series[3]

$$P_{\rm pre}^{\rm ref} = \begin{cases} \left[\frac{1}{2}\frac{\left(N_{\rm bgd}^{\rm ref}\right)^{N_{\rm evt}^{\rm ref}}}{N_{\rm evt}^{\rm ref}!} + \left(\sum_{n=N_{\rm evt}^{\rm ref}+1}^{\infty} \frac{\left(N_{\rm bgd}^{\rm ref}\right)^{n}}{n!}\right)\right] e^{-N_{\rm bgd}^{\rm ref}}, & \text{if } N_{\rm evt}^{\rm ref} > 0 \\ 1 & \text{if } N_{\rm evt}^{\rm ref} = 0 \end{cases} \tag{15}$$

$N_{\rm evt}^{\rm ref}$ is the total the number of events in the reference region,[6] $N_{\rm bgd}^{\rm ref}$ is given by

$$N_{\rm bgd}^{\rm ref} = \frac{F}{\texttt{SAMP\_CYC}_{\rm ref}} N_{\rm pix}^{\rm ref}, \tag{16}$$

$\texttt{SAMP\_CYC}_{\rm ref}$ is the sample cycle for the pixels in the reference region,[7] $N_{\rm pix}^{\rm ref}$ is the number of valid pixels[8]

[3] Each term in the series is of the form $\mu^n/(n!\exp\mu)$. In practice, the series does not extend to infinity. The last term is the one for which $n = N$, where N is the smallest integer that satisfies the relation

$$\frac{\mu^N}{N!} < \texttt{tolerance}\left[\frac{1}{2}\frac{\mu^{n_0}}{n_0!} + \frac{\mu^{n_1}}{n_1!} + \ldots + \frac{\mu^{N-1}}{(N-1)!}\right]. \tag{11}$$

[4] Here an invalid frame for a CCD is one that is not listed in the `statfile`. For `TIMED` mode observations, frames with `EXPNO` < 3 are invalid. Note that a frame does not have to include an afterglow event to be included in $N_{\rm frame}^{\rm aft}$. For example, if a pixel has afterglow events in frames 100, 101, 104, 107, 109, 113, and 119, and if all of the frames from 100 to 119 are valid, then $N_{\rm frame}^{\rm aft} = 20$.

[5] This estimate is an upper limit on the number of trials. The actual number of trials includes only the number of "independent" searches. Since adjacent windows in the sliding `EXPNO` window overlap, they are not independent. A lower limit on $N_{\rm trial}^{\rm aft}$ can be obtained by calculating the number of nonoverlapping windows. This value is smaller than equation 14 by a factor of about ($\texttt{expnowindow} + 1$). Since a precise value for $N_{\rm trial}^{\rm aft}$ can be difficult to determine, equation 14 is used because it yields the most conservative (i.e., the largest) number of trials.

[6] $N_{\rm evt}^{\rm ref}$ does not include the events on the central pixel of the region and the events that lie on a different node from the central pixel (if the region overlaps more than one node).

[7] Equation 16 is valid only if all of the valid pixels in the reference region have the same sample cycle. If, for example, the reference region contains subsets A and B with $N_{\rm pix,A}^{\rm ref}$ and $N_{\rm pix,B}^{\rm ref}$ valid pixels and sample cycles $\texttt{SAMP\_CYC}_{\rm ref,A}$ and $\texttt{SAMP\_CYC}_{\rm ref,B}$, respectively, then equation 16 becomes

$$N_{\rm bgd}^{\rm ref} = F\left(\frac{N_{\rm pix,A}^{\rm ref}}{\texttt{SAMP\_CYC}_{\rm ref,A}} + \frac{N_{\rm pix,B}^{\rm ref}}{\texttt{SAMP\_CYC}_{\rm ref,B}}\right). \tag{17}$$

[8] $N_{\rm pix}^{\rm ref}$ does not include the central pixel of the region (i.e. the pixel on which the potential afterglow occurred), pixels that

in the `regwidth` pixel $\times$ `regwidth` pixel reference region surrounding the pixel with the potential afterglow, and the nominal background fluence F is computed as follows.

(a) For each `nfpixreg` pixel $\times$ `nfpixreg` pixel region l of the node,

$$F_l = \texttt{SAMP\_CYC}_l \frac{N^l_{\rm evt}}{N^l_{\rm pix}}, \tag{18}$$

where $\texttt{SAMP\_CYC}_l$ is the sample cycle for region l,[9] $N^l_{\rm pix}$ is the total number of valid pixels[1] in the region, and $N^l_{\rm evt}$ is the total number of events on these pixels during the entire observation.

(b) Select the regions where $N^l_{\rm pix}$ is greater than zero, F_l is greater than zero, and F_l is less than two times the mean value of the F_ls.[10]

(c) Set $F_{\rm med}$ equal to the median of the values of F_l selected in step (b).

(d) Set F_σ equal to the standard deviation[11] of the values of F_l selected in step (b).

(e) From the full set of regions for the node, select those where $N^l_{\rm pix}$ is greater than zero, F_l is greater than zero, F_l is greater than or equal to $F_{\rm med} - 2F_\sigma$, and F_l is less than $F_{\rm med} + 2F_\sigma$.[10]

(f) Set $F_{\rm med}$ equal to the median of the values of F_l selected in step (e).

(g) Set F_σ equal to the standard deviation[11] of the values of F_l selected in step (e).

(h) Repeat steps 3e–3g an additional `nfrepeat` $- 1$ times (i.e. these steps are performed a total of `nfrepeat` times).

(i) Set F equal to the value of $F_{\rm med}$ from the last iteration of step (f).

4. Each potential afterglow that satisfies the criteria in equations 1–7 is written to the `outfile` with

$$\texttt{TIME} = \texttt{TIME}_{\rm start} - \texttt{TIMEPIXR} \times \texttt{TIMEDEL} - \texttt{FLSHTIME} \tag{21}$$

and

$$\texttt{TIME\_STOP} = \texttt{TIME}_{\rm stop} + (1 - \texttt{TIMEPIXR}) \times \texttt{TIMEDEL}, \tag{22}$$

where $\texttt{TIME}_{\rm start}$ and $\texttt{TIME}_{\rm stop}$ are the `TIME`s in the `statfile` that are associated with the start and stop `EXPNO`s of the afterglow and `TIMEDEL`, `TIMEPIXR`, and `FLSHTIME` are keywords in the `statfile`.

5. The contents of the `badpixfile` are also copied to the `outfile`.

1.5.2 Hot pixels

1. Exclude "invalid" pixels[1]

lie on a different node from the central pixel, and any other invalid pixels.[1]

[9]Equation 18 is valid only if all of the valid pixels in the region have the same sample cycle. If, for example, the region contains subsets A and B with $N^l_{\rm evt,A}$ and $N^l_{\rm evt,B}$ events on $N^l_{\rm pix,A}$ and $N^l_{\rm pix,B}$ valid pixels and sample cycles $\texttt{SAMP\_CYC}_{l,{\rm A}}$ and $\texttt{SAMP\_CYC}_{l,{\rm B}}$, respectively, then equation 18 becomes

$$F_l = \frac{\texttt{SAMP\_CYC}_{l,{\rm A}} N^l_{\rm evt,A} + \texttt{SAMP\_CYC}_{l,{\rm B}} N^l_{\rm evt,B}}{N^l_{\rm pix,A} + N^l_{\rm pix,B}}. \tag{19}$$

[10]If no regions satisfy these conditions for one or more nodes of a CCD, then a warning message is produced and no afterglows or hot pixels are written to the `outfile` for the CCD. However, processing of the other CCDs, if any, proceeds in the usual manner and the `outfile` may contain afterglows and hot pixels for them.

[11]Here,

$$F_\sigma = \left[\frac{1}{N_l} \sum_l (F_l - F_{\rm med})^2 \right]^{1/2}, \tag{20}$$

where N_l is the number of regions in the sum.

2. If there are multiple events in an observation that have the same values of the coordinates CCD_ID, CHIPX, and CHIPY, then the pixel is identified as hot for the entire duration of an observation if

$$P^{\rm hot}_{\rm post} < \texttt{probthresh} \text{ and} \tag{23}$$
$$P^{\rm ref}_{\rm post} \geq \texttt{probthresh}, \tag{24}$$

where the post-trials probability $P^{\rm hot}_{\rm post}$ is given by

$$P^{\rm hot}_{\rm post} = 1 - \left(1 - P^{\rm hot}_{\rm pre}\right)^{N^{\rm hot}_{\rm trial}}, \tag{25}$$

the pre-trials probability $P^{\rm hot}_{\rm pre}$ is given by the series[3]

$$P^{\rm hot}_{\rm pre} = \left[\frac{1}{2}\frac{\left(N^{\rm hot}_{\rm bgd}\right)^{N^{\rm hot}_{\rm evt}}}{N^{\rm hot}_{\rm evt}!} + \left(\sum_{n=N^{\rm hot}_{\rm evt}+1}^{\infty} \frac{\left(N^{\rm hot}_{\rm bgd}\right)^{n}}{n!}\right)\right] e^{-N^{\rm hot}_{\rm bgd}}, \tag{26}$$

$N^{\rm hot}_{\rm evt}$ is the total number of events on the potential hot pixel for the observation,

$$N^{\rm hot}_{\rm bgd} = \frac{F}{\texttt{SAMP\_CYC}_{\rm hot}}, \tag{27}$$

$\texttt{SAMP\_CYC}_{\rm hot}$ is the sample cycle for the potential hot pixel, the number of trials

$$N^{\rm hot}_{\rm trial} = \sum_k N^{\rm ccd}_{{\rm pix},k}, \tag{28}$$

$N^{\rm ccd}_{{\rm pix},k}$ is the number of valid pixels[1] for the kth CCD (i.e. $N^{\rm ccd}_{{\rm pix},k} = 1024 \times 1024$ less the number of invalid pixels), $P^{\rm ref}_{\rm post}$ is given by equation 10, where $N^{\rm ref}_{\rm trial}$ is the number of candidate hot pixels that satisfy equation 23 and the probability $P^{\rm ref}_{\rm pre}$ that the event fluence in the reference region is consistent with the event fluence on the entire node is given by equation 15.

3. Each potential hot pixel that satisfies the criteria in equations 23 and 24 is written to the outfile with

$$\texttt{TIME} = \texttt{TIME}_{\rm start} - \texttt{TIMEPIXR} \times \texttt{TIMEDEL} - \texttt{FLSHTIME} \tag{29}$$

and

$$\texttt{TIME\_STOP} = \texttt{TIME}_{\rm stop} + (1 - \texttt{TIMEPIXR}) \times \texttt{TIMEDEL}, \tag{30}$$

where $\texttt{TIME}_{\rm start}$ and $\texttt{TIME}_{\rm stop}$ are the TIMEs in the statfile that are associated with the first and the last valid EXPNOs, respectively, for the CCD that contains the hot pixel, and TIMEDEL, TIMEPIXR, and FLSHTIME are keywords in the statfile.

4. The contents of the badpixfile are also copied to the outfile.

Once the afterglow and hot-pixel detection algorithms have been used, the tool acis_build_badpix is used to mark the pixels adjacent to such pixels as bad and the tool acis_process_events is used to set the appropriate STATUS bit for events associated with afterglows (bit 16 of 0–31) and hot pixels (bit 4 of 0–31).

1.6 Caveats

1. Since the algorithms in this spec are designed to prevent the events associated with bright sources from being discarded, it is not possible to find afterglows or hot pixels associated with such sources.

2. The choices of default values for the parameters expnowindow, probthresh, cntthresh, regwidth, nfpixreg, nfrepeat, and tolerance may not be optimum.

3. Although it may not be optimum to do so, the afterglow and hot-pixel detection algorithms are applied separately to the primary and secondary data for interleaved mode observations.

4. The algorithms are not applied to the data for continuous-clocking mode observations.

5. If two hot pixels are adjacent to one another, the tool acis_find_afterglow may miss one of them.

6. Should pixels that have multiple afterglows be labeled as hot?

7. Are the occurrences of strings of events that have the same values for EXPNO, CCD_ID, and CHIPX, but different values of CHIPY, associated with cosmic rays? Likewise, are the occurrences of strings of events that have the same values for EXPNO, CCD_ID, and CHIPY, but different values of CHIPX, associated with cosmic rays?[12]

8. The tool acis_build_badpix can combine adjacent afterglows into a single bad region, whether the afterglows overlap in time or not. Separate afterglows should not be combined with one another (and have the adjacent bad-pixel STATUS bit set) unless they overlap in time.

[12]In an email from Catherine Grant dated April 11, 2012, she notes that "The vertical strings of events are generally due to cosmic ray afterglows. The horizontal strings are from cosmic ray events that happened in the serial registers. In both cases, since you're only seeing things that made it through the event finding algorithm, you won't always see the whole feature, just whatever looked "event-like". You'll generally only see these on the FI CCDs, not the BIs. Watch Bev's raw frame movies and you'll see both kinds (http://space.mit.edu/home/fergason/acis/raw.html). The vertical afterglows happen frequently and are easier to see, but the horizontal ones are there as well."