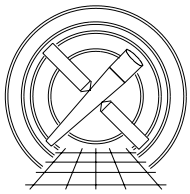

MIT Kavli Institute **Chandra X-Ray Center**

MEMORANDUM
June 15, 2011

To:	Jonathan McDowell, SDS Group Leader
From:	Glenn E. Allen, SDS
Subject:	`DELTOCLK spec`
Revision:	1.2
URL:	http://space.mit.edu/CXC/docs/docs.html#deltoclk
File:	/nfs/cxc/h2/gea/sds/docs/memos/deltoclk_spec_1.2.tex

1 Description

Occasionally the values of `DELTOCLK` in exposure-records files are anomalously high or anomalously low. Inaccurate values of `DELTOCLK` lead to inaccurate pulse heights. Therefore events that are processed using anomalous values have a `STATUS` bit set to one and are discarded from Level 2 event-data files. Note that the input data are the values of `OVRCLOCK` from exposure-statistics files instead of the values of `DELTOCLK` from exposure-records files because the difference between the two sets of values for a node is a constant (the initial overclock) and because the exposure-statistics files are available to users who may want to reprocess data themselves.

2 Input

1. A Level 1 event-data file (acis*evt1.fits)
2. An exposure-statistics file (acis*stat1.fits)

3 Output

1. An updated event-data file
2. An updated exposure-statistics file

4 Parameters

1. `infile,f,a,"",,,"Name of input event-data file"`
2. `outfile,f,a"",,,"Name of output event-data file"`
3. `expstatsfile,f,a,"",,,"Name of exposure-statistics file to be updated"`

4. numframes,i,h,5,3,1001,"Nominal number of frames in the sliding time window (must be odd)"

5. maxframegap,i,h,11*,1,1001,"Maximum difference between consecutive frame numbers in window"

6. minnumframes,i,h,3,3,1001,"Minimum number of frames in window"

7. deltoclkthresh,i,h,3†,1,4096,"Minimum offset in adu that is considered anomalous"

5 Processing

1. Verify that the infile and expstatsfile exist. If clobber=no, then verify that the outfile does not exist. Verify that the values of the parameters numframes, maxframegap, minnumframes and deltoclkthresh are in the valid ranges for these parameters. Note that numframes must be an odd number.

2. If the data are being reprocessed by a user, then perform the following two steps.

   (a) Set STATUS bit 10 (of 0-31) to zero (i.e. unset it) for all of the events in the outfile. Subsequent processing sets this bit to one for the appropriate events.

   (b) Remove the contents of the HDU "BADOCLK" in the expstatsfile. This HDU is updated appropriately in subsequent processing.

3. Perform the following steps separately for each node of each CCD.

   (a) For each valid‡ frame k, compute the expected value M_k of OVRCLOCK. In most cases, M_k is the median§ of the values of OVRCLOCK for the frames from $k-n$ to $k+n$, inclusive, where

$$n = \frac{\texttt{numframes} - 1}{2}. \tag{1}$$

   There are some cases that require special treatment.

   i. If one or more of the frames immediately before and/or immediately after frame k are invalid, then use the nearest valid frames, not the invalid frames to compute the median. For example, if numframes $= 5$, then the OVRCLOCK data for the consecutive valid frames i, j, k, l, and m are used to compute the median, provided that $\texttt{EXPNO}_i < \texttt{EXPNO}_j < \texttt{EXPNO}_k < \texttt{EXPNO}_l < \texttt{EXPNO}_m$ and that

$$\texttt{EXPNO}_j - \texttt{EXPNO}_i \leq \texttt{maxframegap}, \tag{2}$$
$$\texttt{EXPNO}_k - \texttt{EXPNO}_j \leq \texttt{maxframegap}, \tag{3}$$
$$\texttt{EXPNO}_l - \texttt{EXPNO}_k \leq \texttt{maxframegap}, \text{ and} \tag{4}$$
$$\texttt{EXPNO}_m - \texttt{EXPNO}_l \leq \texttt{maxframegap}. \tag{5}$$

*The largest DTYCYCLE used to date is 10. Furthermore, as shown in Figure 1, consecutive frame numbers typically differ by eleven or less.

†As shown in Figure 4, the algorithm is relatively insensitive to the value of deltoclkthresh for values greater than or equal to three.

‡Here, a frame is valid if:

the value of EXPNO for the frame is greater than the value of EXPNO for the preceding frame (if any),

the value of EXPNO for the frame is less than the value of EXPNO for the subsequent frame (if any), and

there is OVRCLOCK data available for the frame.

If one or more of these conditions are not satisfied, then the frame is considered invalid. Invalid frames are not included in the "BADOCLK" HDU of the expstatsfile. Any events that may occur during invalid frames do not have STATUS bit 10 set to one.

§If the number of elements in the set for which the median is being computed is odd, then the median is the middle value of the set after the set has been sorted (e.g. the median of [1,1,1,4,5,6,7] is 4). If the number of elements is even, then the median is the mean of the middle two values of the sorted set. If the mean is not an integer, then the mean is truncated to obtain an integer (e.g. the median of [1,1,1,4,5,6] is 2, not 2.5).

If equation 2 is not satisfied, then exclude the data for frame i from the computation of the median. If equation 3 is not satisfied, then exclude the data for frames i and j. Similarly, if equation 4 is not satisfied, then exclude the data for frames l and m and if equation 5 is not satisfied, then exclude the data for frame m.

ii. At or near the beginning of the data, there may not be n valid frames preceding frame k. For example, for the first valid frame, there are no valid frames that precede it. In this case, the median is computed using the data for the first frame and for the n subsequent, valid frames (provided that the subsequent frames do not violate the `maxframegap` constraint). For the second valid frame, the median is computed using the data for the first and second valid frames and for the n subsequent, valid frames. (Again, the `maxframegap` constraint applies.)

iii. Likewise, at or near the end of the data, there may not be n valid frames following frame k.

iv. There must be at least `minnumframes` valid frames in the set of data used to compute M_k. If there are fewer than `minnumframes` valid frames, due to violations of the `maxframegap` constraint and/or due to the lack of data at the beginning or the end of a data set, then set M_k equal to the value of `OVRCLOCK` for frame k (i.e. there is too little data to determine whether the value of `OVRCLOCK` is anomalous).[¶] A consequence of this condition is that an `infile` must have at least `minnumframes` valid frames of data for a node in order to search for anomalous values of `OVRCLOCK` for the node.

(b) If the value of `OVRCLOCK` for a valid frame is greater than or equal to M_k + `deltoclkthresh` or is less than or equal to M_k − `deltoclkthresh`, then the value of `OVRCLOCK` is anomalous. In this case,

i. modify the `outfile` to set `STATUS` bit 10 (of 0-31) to one for all of the events during frame k for the particular CCD and node in question and

ii. update the "`BADOCLK`" HDU of the `expstatsfile` to include the `EXPNO`, `CCD_ID`, and `NODE_ID` information for the frame, CCD, and node in question.

6 TBD

1. Recreate Figures 2–4 (especially 4) to determine what the default value of `maxframegap` should be.

2. Plot the number of anomalies as a function of `TIME` and `CCD_ID` and `NODE_ID`.

[¶] In this case, frame k is not included in the "`BADOCLK`" HDU of the `expstatsfile` and any events that occur during frame k do not have `STATUS` bit 10 set to one.

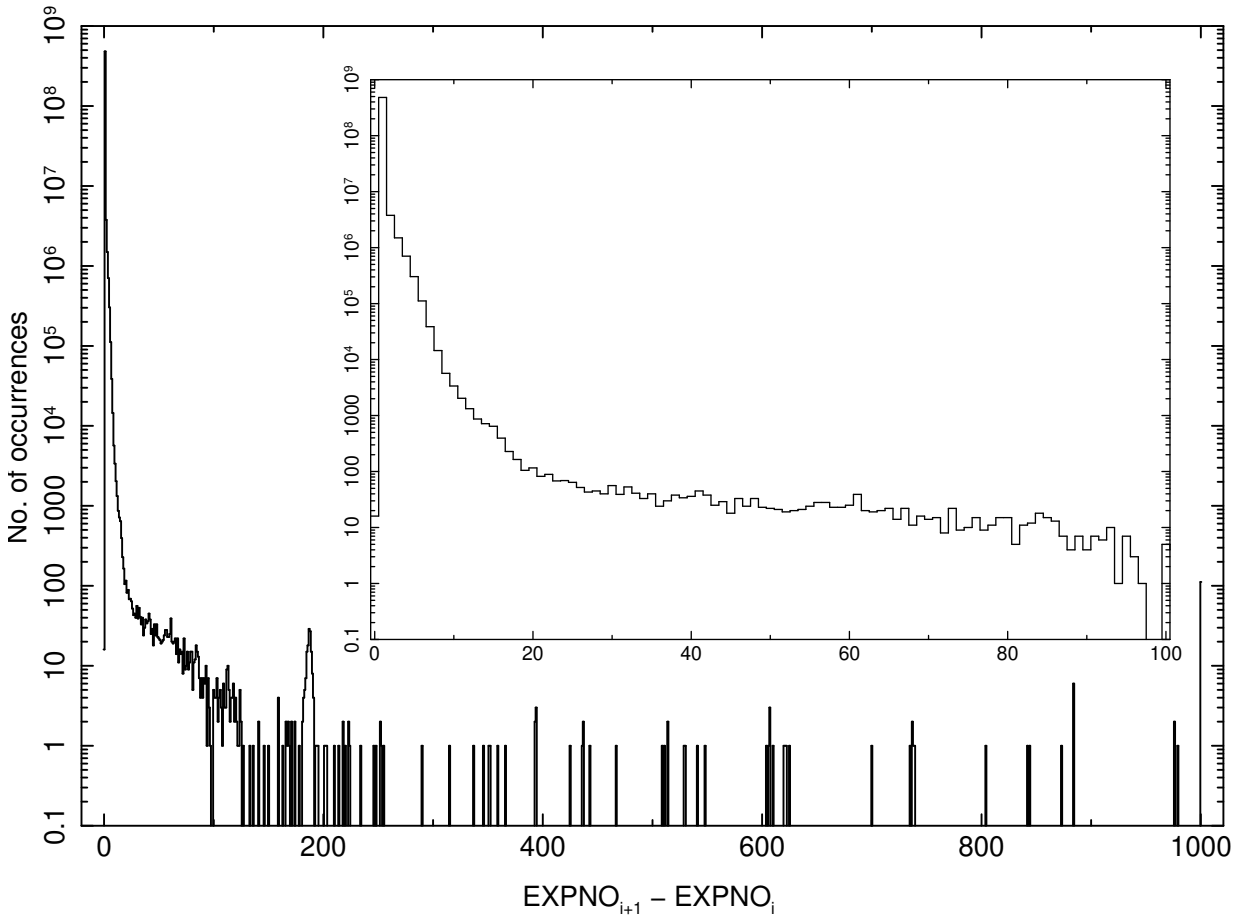

Figure 1: A histogram of the differences between consecutive frame numbers for 63,778 exposure-records files. A separate histogram was computed for each node of each file. The histogram in this figure is a sum of each of these histograms. If a difference is greater than 1000, then the difference is set to 1000. If a difference is less than 0, then the difference is set to 0. The inset histogram is a plot of the same data, except that it includes only the data in the range from 0 to 100. The histogram includes all frames, whether they are valid or not.

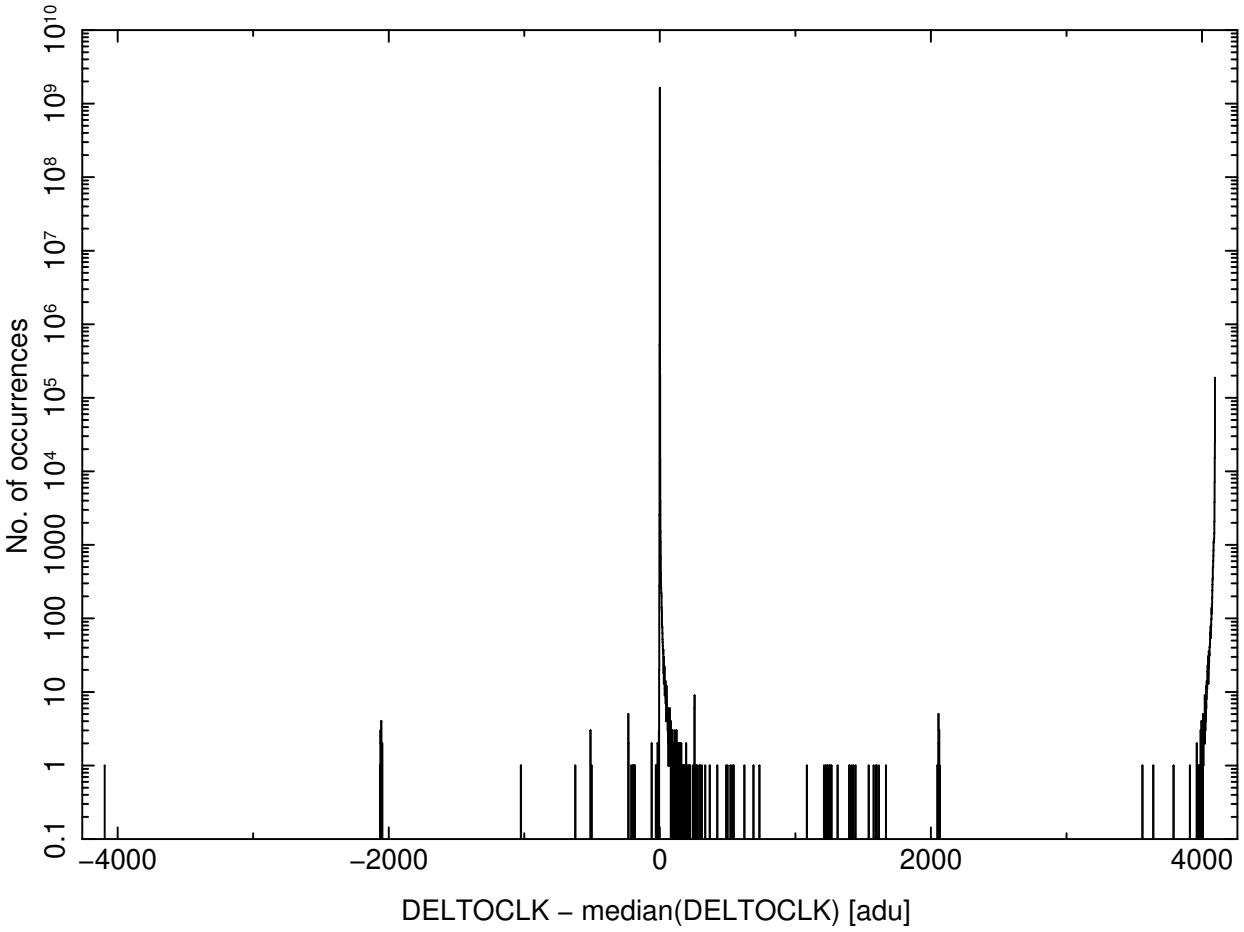

Figure 2: A histogram of the differences between the values of `DELTOCLK` and the values of the median `DELTOCLK` for 63,778 exposure-records files. A separate histogram was computed for each node of each file. The histogram in this figure is a sum of each of these histograms. The values of the median were computed using a `numframes` = 5, `maxframegap` = ∞, and `minnumframes` = 3. If a value of `DELTOCLK` differs by more than 4096 adu from the corresponding median, then the difference is set to 4096 or −4096 as appropriate.

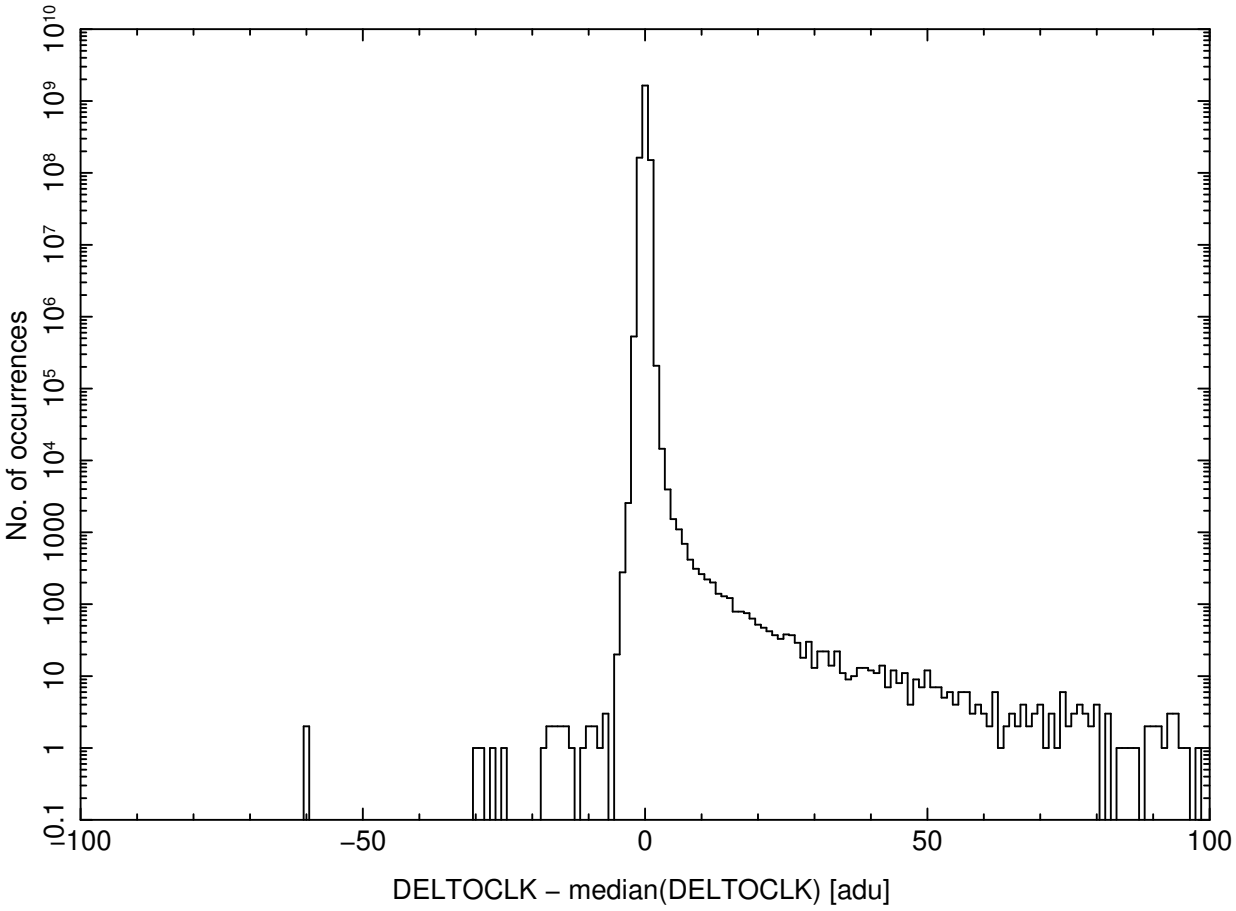

Figure 3: This figure is identical to Figure 2, except that it includes only the data in the range from -100 to 100 adu.

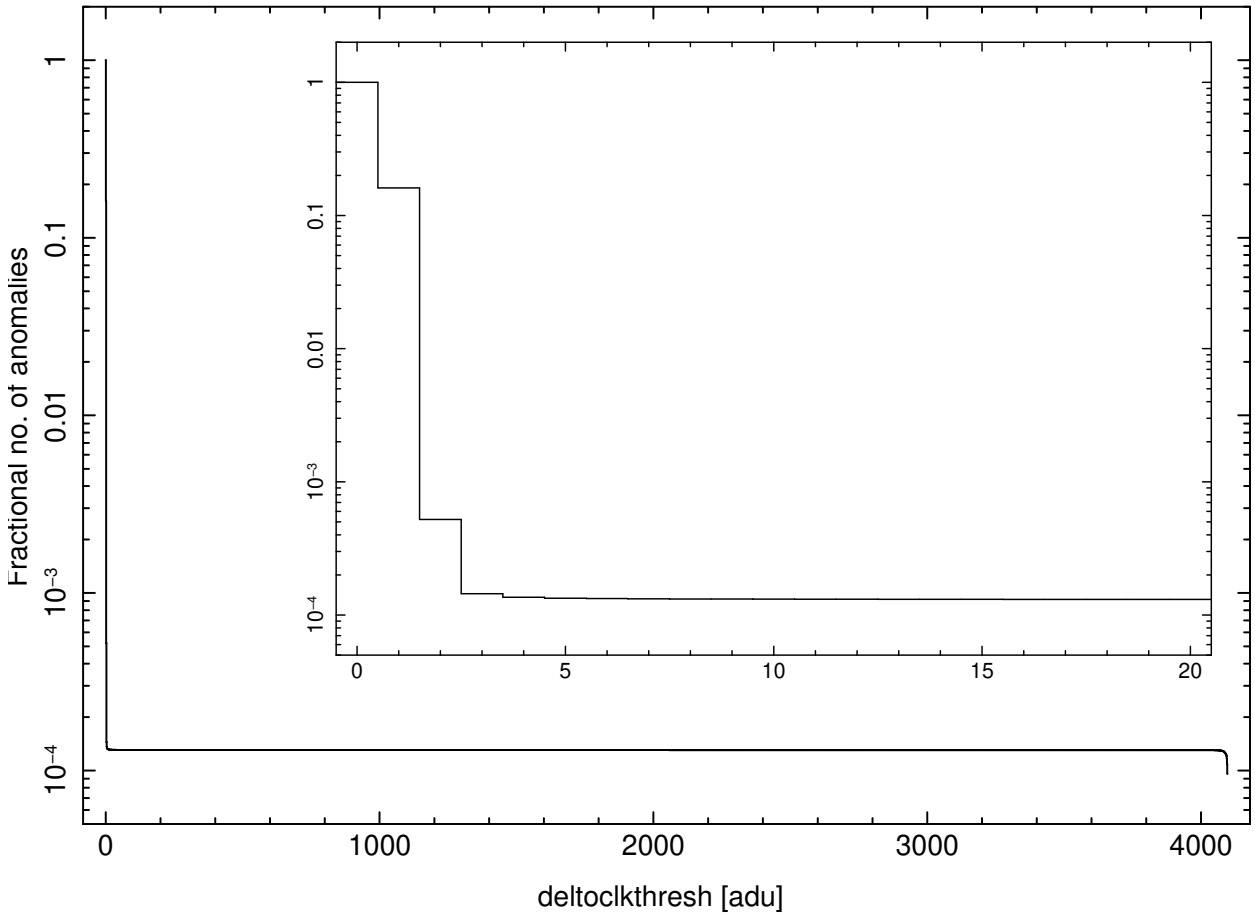

Figure 4: A plot of the fraction of the values of `DELTOCLK` that are identified as anomalous as a function of the parameter `deltoclkthresh`. This plot was produced using the data that is shown in Figure 2. The inset histogram is a plot of the same data, except that it includes only the data in the range from 0 to 20.