

Changes to the COS Extraction Algorithm for Lifetime Position 3

Charles R. Proffitt¹, K. Azalee Bostroem^{1,2}, Justin Ely¹, Deatrick Foster¹, Svea Hernandez^{1,3}, Philip Hodge¹, Robert I. Jedrzejewski¹, Sean A. Lockwood¹, Derck Massa⁴, Molly S. Peeples¹, Cristina M. Oliveira¹, Steven V. Penton¹, Rachel Plesha¹, Julia Roman-Duval¹, Hugues Sana¹, David J. Sahnou¹, Paule Sonnentrucker¹, Joanna M. Taylor¹

¹*Space Telescope Science Institute, Baltimore, MD*

²*University of California, Davis, CA*

³*Department of Astrophysics/IMAPP, Raboud University, Nijmegen, NL*

⁴*Sigma Space Corporation, 4600 Forbes Boulevard, Lanham, MD*

25 September 2015

ABSTRACT

The COS FUV Detector Lifetime Position 3 (LP3) has been placed only 2.5" below the original lifetime position (LP1). This is sufficiently close to gain-sagged regions at LP1 that a revised extraction algorithm is needed to ensure good spectral quality. We provide an overview of this new "TWOZONE" extraction algorithm, discuss its strengths and limitations, describe new output columns in the XID files that show the boundaries of the new extraction regions, and provide some advice on how to manually tune the algorithm for specialized applications.

Contents

- 1. Introduction (page 2)
- 2. The New Extraction Algorithm (page 4)
 - 2.1 Spectral Image Straightening (TRCECORR) (page 6)
 - 2.2 Spectral Alignment (ALGNCORR) (page 7)
 - 2.3 Defining the Extraction Zones (page 10)
 - 2.4 Creating the Images Used for Extraction (page 11)
 - 2.5 Extracting the Spectrum (page 13)
 - 2.6 Comparing the TWOZONE and BOXCAR Extractions (page 14)
 - 2.7 New Columns in the X1D Output Files (page 15)
- 4. Controlling the New Algorithm (page 17)
 - 4.1 Header Keyword Switches (page 17)
 - 4.2 The TRACETAB Reference File (page 18)
 - 4.3 The TWOZXTAB Reference File (page 18)
 - 4.4 The PROFTAB Reference File (page 19)
- 5. Customizing the TWOZONE Extraction (page 22)
- References (page 24)
- Appendix: Changes between CalCOS 3.0 and 3.1 (page 24)

1. Introduction

The Cosmic Origins Spectrograph (COS) far-ultraviolet (FUV) channel uses photon counting micro-channel plate (MCP) detectors together with cross-delay line (XDL) anodes for measuring the time and location of photons incident on a CsI photocathode. Each of two independent detector segments, designated as FUVA and FUVB, consists of a stack of three micro-channel plates with a high voltage field across them. When a photon impacts the photocathode on the top plate of a micro-channel stack, the resulting photo-electron generates a cascade of electrons that emerges from the bottom, the position of which is determined by the pulse timing measured in the delay line anodes. The pulse height amplitude (PHA) generated by each event is also recorded. The relation between the recorded PHA value and the number of electrons in the charge cloud is approximately $n_e = 10^{(PHA/20.5+6.22)}$ (McPhate 2010, private communication).

The delay line design of the COS detector requires operating the MCP stacks with relatively high gain, with each initial photo-electron typically resulting in two to ten million electrons at the bottom of the stack. This causes considerable charge to be extracted from the glass of the MCP pores. As a result, as the detector is used and charge is extracted, the number of electrons extracted per incident photon declines, particularly in heavily used detector locations (Sahnou et al. 2011). When the PHA value declines below 3, (about 2 million electrons per event), it becomes more difficult both to separate photon events from background events and to properly measure the location of the original incident photon on the photocathode. This leads to regions of degraded quantum efficiency, i.e., “holes”, in the incident spectrum that cannot be properly calibrated. While increasing the voltage across the plates can temporarily restore the gain, once the maximum allowed voltage has been reached it is necessary to relocate the spectrum to new, pristine detector location to preserve the quality of the observed data.

COS FUV Lifetime Position 1 (LP1), located near the center of the FUV detector, was in use from the beginning of COS operations in 2009 until July 23, 2012. At that time the FUV spectra were relocated to LP2, at a position +3.5” or about +41 detector pixels above the starting location (Osten et al. 2013). On February 9, 2015 the COS FUV spectra for most modes were relocated again to LP3. However, to preserve spectral resolution, which is better closer to the detector center (see Sahnou et al. 2014), and to conserve the available detector area for subsequent lifetime positions, LP3 was located much closer to LP1, at an offset of only –2.5” or –33 detector pixels. Note that COS FUV spectra have cross dispersion widths that vary with setting and which tend to increase towards shorter wavelengths. The G130M 1055 and 1096 CENWAVE settings have widths as large as 65 pixels. For new LP3 position, this is close enough to the gain sagged regions at LP1 that these two modes were not moved, but instead left at LP2.

For many of the other modes that were moved, the separation from LP1 is still small enough that the standard boxcar pipeline extraction would have rejected and discarded large wavelength regions due to overlap with previously sagged LP1 regions, (i.e., regions where 5% or more of the sensitivity is lost due to the low gain). The original "BOXCAR" algorithm rejects the entire wavelength bin if one of these sagged pixels is included in the extraction height, even if only at the edge of the extraction region where it has little impact on the extracted spectrum. However, the location of LP3 was strategically selected so that sagged-pixels located at the edges of the extraction zone do not affect the spectral quality or flux accuracy of LP3 science spectra. To avoid unnecessarily discarding columns affected by such pixels, a new algorithm, referred to as the "TWOZONE" algorithm, has therefore been implemented that avoids discarding wavelength bins when a gain-sagged region is far enough out in the wings of the profile that it does not have a significant effect on the extracted flux.

Below, we provide more detailed information about the new algorithm as implemented in the COS calibration software CalCOS 3.1, which is expected to be released in the latter part of 2015. We include discussions of the new reference file types needed to support this calibration, and of how the extraction can be altered by editing data file keywords or reference files. Differences between this version and the

initial release in early 2015 of the TWOZONE algorithm in CalCOS 3.0 are detailed in an appendix of this document. Additional information about CalCOS and the pipeline processing of COS data can be found in the COS Data Handbook (Massa & York et al. 2013; Fox et al. 2015)

2. The New Extraction Algorithm

The CalCOS pipeline implementation of the BOXCAR algorithm uses a fixed extraction region for each mode that is generously sized to include all of the flux from a point source target after allowing for the expected acquisition centering errors. The new TWOZONE algorithm instead uses a smaller variable height extraction region that is adjusted to conform to the expected cross-dispersion profile at each wavelength. The tighter fit of the extraction region to the profile requires a number of new calibration steps, which can be summarized as follows:

1. The TRCECORR step of CalCOS straightens the 2D spectral image to remove residual localized geometric distortion in the cross-dispersion (Y) direction
2. The ALGNCORR step of CalCOS shifts the observed spectrum in the cross-dispersion direction to align it with a reference point source profile chosen to match the grating and central wavelength of that exposure
3. When the X1DCORR step of CalCOS is run with the new algorithm, by setting the header keyword XTRACALG=TWOZONE, two regions are defined on the detector, the height and location of which vary as a function of wavelength. These zones are defined in terms of the enclosed energy as measured in the reference point-source cross dispersion profile. An example of these zone boundaries is illustrated in Figure 1.
 - One region, which we refer to as the “OUTER ZONE”, defines the region over which events will be summed to calculate the 1D extracted count rates. For all FUV modes at LP3, this is currently set to include the central 99% of the enclosed energy.
 - A second region, the “INNER ZONE”, defines the region that will be checked for bad data quality flags. This is currently set to include the central 80% of the profile’s energy. For most COS FUV profiles, the 80% enclosed energy contour also does a good job of separating the profile’s core, where the intensity of light is largest, from the more diffuse and extended wings.

The resulting extraction gives net count rates, and therefore fluxes, very close to that obtained using the older BOXCAR algorithm. However, with the new algorithm wavelength bins, (i.e. XFULL values), are only rejected if “bad” pixels are in the core of the cross dispersion profile where most of the light falls. This allows the spectrum to be successfully extracted even if gain-sagged regions begin to overlap the far wings of the spectral profile. In addition, since the outer zone’s height is usually smaller than the old BOXCAR extraction height, less detector background counts are included in the extraction region. This will improve the signal-to-noise, especially for faint sources. Below we will explain each of the steps introduced above in greater detail.

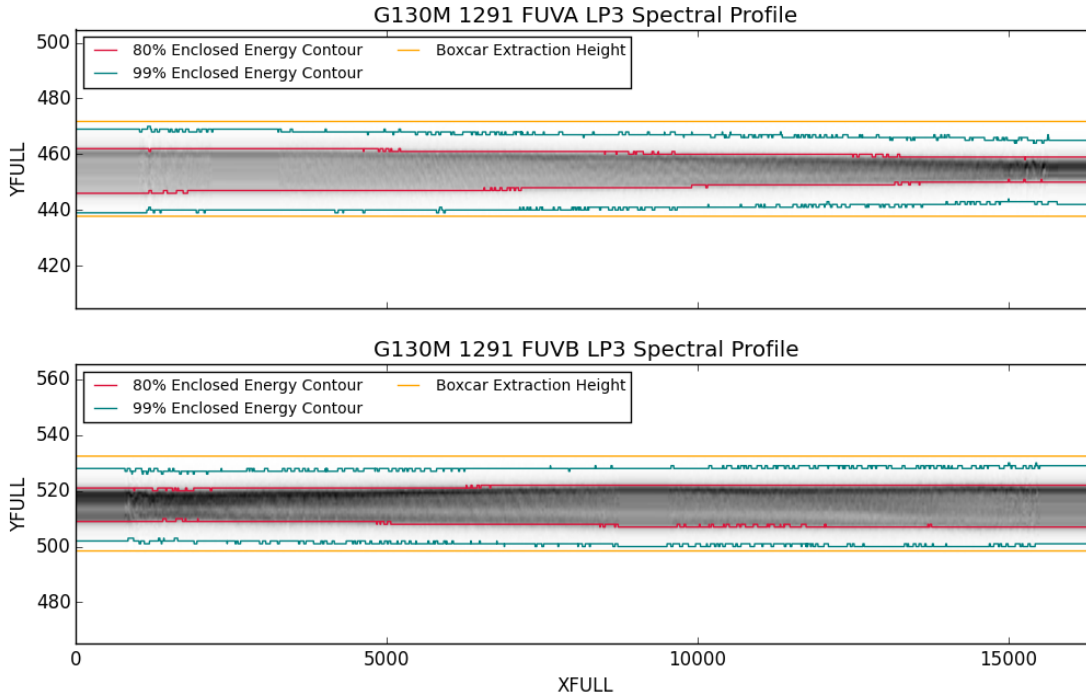


Figure 1: Images of the reference profiles for the G130M 1291 FUVA and FUVB segments are shown, together with the 99% and 80% enclosed energy contours. Also shown for comparison is a box of the same height as used for the BOXCAR extraction (yellow).

A number of new reference file types were introduced to support the new extraction. The parameters controlling the new extraction algorithm are in a new table specified in the "TWOZXTAB" keyword. The reference file storing the trace correction vectors used for straightening the spectral image is given in the TRACETAB keyword, and the file containing the 2D reference profile shapes for each CENWAVE is specified in the PROFTAB keyword. The structure and content of these reference files are described in detail in section 4 of this document.

At the current time, the new TWOZONE algorithm is only being used to calibrate data obtained at LP3. COS FUV data obtained at other lifetime positions, including any new observations at the G130M 1055 and 1096 settings which remained at LP2, are still calibrated by default with the BOXCAR algorithm, as are all COS NUV spectra.

To apply the TWOZONE algorithm to LP1 or LP2 FUV data would require the construction and delivery of new TWOZXTAB, TRCETAB, and PROFTAB reference files appropriate for those locations. Applying the new algorithm to COS NUV spectra would require additional changes to the CalCOS code itself, but the extent of such changes has not been evaluated.

The choice of which algorithm to use is handled automatically in the MAST pipeline by setting appropriate keywords in the file headers based on the detector and lifetime position used. Apart from a few new informational columns in the x1d data files, the formats for all output data products are as before. So for users who simply wish to adopt the pipeline-calculated flux and error as a function of wavelength that is tabulated in the x1d or x1dsum files, the changes should be transparent.

The new algorithm is, however, not optimized to calibrate data of extended targets and the delivered reference files are not optimized for data obtained with the BOA aperture. Users with these kinds of observations may need to perform customized extractions and carefully evaluate their data for possible artifacts.

A note about coordinate systems: In CalCOS, the originally measured RAWX and RAWY coordinates of each photon event are first corrected for thermal and geometric distortions to produce locations in XCORR and YCORR coordinates. These coordinates can be thought of as being fixed with respect to the physical structure of the micro-channel plates and photocathode, and it is in this coordinate system that detector defects in the bad-pixel mask (BPIXTAB), and the low gain regions in the gain sag table (GSAGTAB) are tabulated. Additional corrections to event locations are then also calculated to take into account variations in how the optics projected the spectrum onto the detector for that particular observation, (i.e., corrections from the WAVECORR step), and also for the Doppler shifts induced by the changing velocity of HST relative to an external target as HST orbits the Earth (DOPPCORR). These additional corrections go into the determination of the final XFULL and YFULL coordinates that are assigned to each event. The XFULL value for a given setting should correspond to the geo-centric wavelength of each incident photon while the YFULL location corresponds to spatial location perpendicular to the dispersion. The TRCECORR and ALGNCORR steps of the new algorithm do not change the XCORR, YCORR, or XFULL locations of events, but do add additional corrections to the final YFULL locations.

2.1 Spectral Image Straightening (TRCECORR)

Although CalCOS applies a 2D geometric correction to the raw event locations in the GEOCORR calibration step, noticeable distortion remains in the corrected images. In particular, there remain significant small-scale distortions in the cross-dispersion (Y) direction that appear to be fixed on the physical detector, and which remain approximately consistent over heights of several tens of pixels. Removing this residual Y distortion results in significantly smoother cross-dispersion profiles that are easier to align with template point source profiles. Using high signal-to-noise observations of each mode, we derived the contents of the TRACETAB reference file. In the TRACETAB reference file, we define the "trace" center location in each detector column i to be the flux weighted centroid $Y_i^c = \sum_j (n_{ij} Y_{ij}) / \sum_j n_{ij}$ of the incident counts in column i , where n_{ij} is the number of counts at each location Y_{ij} and j is the index of the row. Trace locations for a number of modes are shown in Figure 2. The differences between these trace locations and the median trace location as a function of XCORR are tabulated in the TRACETAB reference table. This difference vector is subtracted from the YFULL location of all events, except those that fall into the WCA region or outside the detector active area. In Figure 3, we can see the effect of applying the trace correction on the shape of the 2D spectral image. Application of the TRCECORR step also brings the flux-weighted centroid for all columns to the same YFULL location, which will be important for the ALGNCORR step.

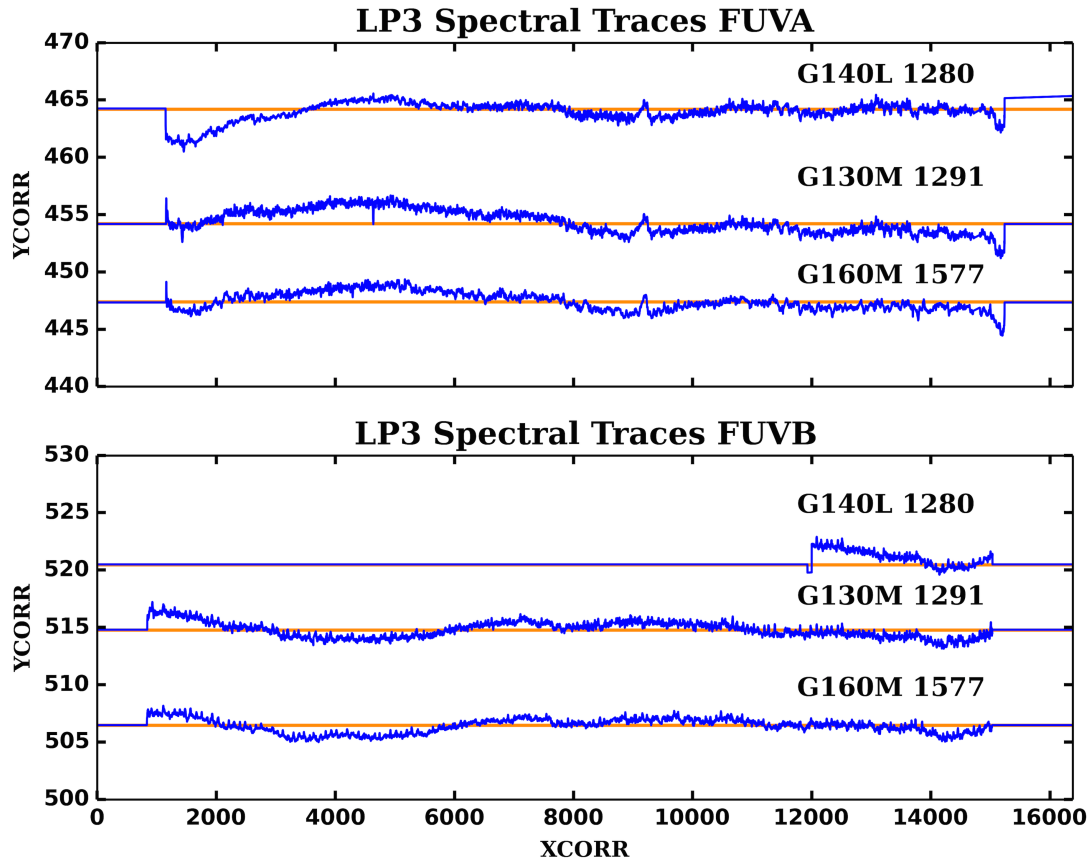


Figure 2: Spectral trace locations for one CENWAVE setting of each of the COS FUV gratings are shown. Trace locations for the other CENWAVE settings are very similar. The difference between the measured YCORR center as a function of XCORR (blue lines) and the median center (orange line) is tabulated for each CENWAVE, and subtracted from the event locations to straighten the science images.

2.2 Spectral Alignment (*ALGNCORR*)

In the ALGNCORR step, the spectrum is aligned in the cross-dispersion direction with a reference profile appropriate for that OPT_ELEM (grating), CENWAVE, APERTURE, and in principle also the Lifetime Position, although at the current time we only supply reference profiles at LP3. The offset between the observed spectrum and the reference profile is first calculated by collapsing both the observed and reference spectral images along the dispersion direction to produce 1D cross-dispersion profiles. The observed and reference 1D profiles are background-subtracted in a consistent manner, and their centroids are calculated. The offset is then the difference between the observed and reference centroid. The alignment is performed by subtracting this offset from the YFULL coordinates of all events outside the WCA aperture. The alignment process is illustrated in Figure 4.

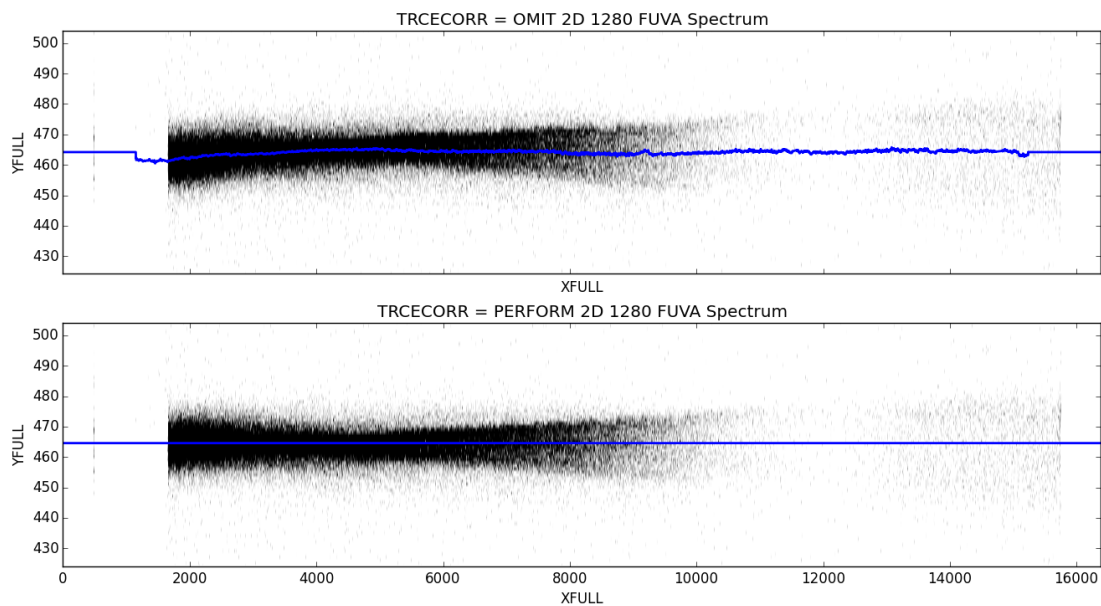


Figure 3: A spectral image for G140L FUV is shown without (above) and with (below) the TRCECORR correction applied. In the upper panel, the blue line shows the trace taken from the reference file and shifted to align with the data.

The ALGNCORR step is intended to be used after the TRCECORR step. Since the TRCECORR step has already removed variation of the YFULL centroid location as a function of wavelength, the calculated centroid location for the 1D cross-dispersion profile should not depend of the source spectral energy distribution; the centroids for red and blue spectra should be consistently defined. It is assumed that TRCECORR has corrected for any spectral structure that is fixed with respect to the physical detector, and that the tabulated reference profiles for a given setting are simply functions of wavelength and spatial offset.

When summing the observed 2D image along the dispersion direction to produce a 1D profile, columns in the observed image that are contaminated by airglow lines or which are flagged with a DQ value that matches one of the bits set in SDQFLAGS are excluded from the collapsed spectrum. An exception is that DQ values flagging gain sagged regions (DQ=8192) are not excluded from the alignment even when this bit is set in SDQFLAGS. This is done to allow the alignment step to work properly even when very wide gain sagged regions exist in the far wing of the profile. The same columns are excluded from the collapse of both the observed and reference profiles.

Geo-coronal emission can result in a separate image of the 2.5" diameter PSA aperture being projected onto the detector for each geo-coronal emission line. The exclusion of the contamination from these airglow regions is necessary to prevent the spectrum of a faint source from being aligned using the sky light in these aperture images rather than with the actual target's spectrum. The dispersion relation is used to calculate the expected locations on the detector of the wavelengths corresponding to the geo-coronal Lyman α (1215.67), O I (1302.168, 1304.858, 1306.029, 1355.598, 1358.512), and N I (1199.55, 1200.233, 1200.710) lines. Any column within 500 X pixels of Lyman α or within 200 X pixels of any other geo-coronal line is excluded from the collapsed cross dispersion profile produced during the alignment step. Currently these wavelengths and sizes of these geo-coronal regions are written into

the CalCOS code and cannot be adjusted by the users. Note that these columns are only excluded from the alignment calculation; during the actual extraction they will be treated in the same way as other detector regions.

Below we describe in detail the steps of the alignment process:

After the observed and reference spectral images are collapsed along the dispersion direction, the centroid of the observed 1D profile is calculated. The parameters HEIGHT, B_SPEC, B_BKG1, B_BKG2, and BHEIGHT are taken from the row of the TWOZXTAB reference file that matches the grating (OPT_ELEM), CENWAVE, and APERTURE of the observation. The flux weighted centroid of the 1D observed profile is calculated as

$$Y_C = \sum_j (n_j - b) Y_j / \sum_j (n_j - b),$$

where b is the average of the counts in two background regions located in the wings of the cross dispersion profile at locations B_BKG1 and B_BKG2 of size BHEIGHT, and n_j is the total number of counts at each location Y_j of the 1D cross-dispersion profile. The vertical region of size HEIGHT used for the calculation is then re-centered at the calculated Y_C value and the background regions are adjusted by the same amount. The calculation is iterated until the change in Y_C between iterations is less than 0.005 pixels or 5 iterations have been completed. An estimate of the Poisson uncertainty in the value of Y_C is also calculated assuming that the error squared is given by

$$\sigma_c^2 = \sum_j (n_j + \sigma_b^2) (Y_j - Y_c)^2 / (\sum_j (n_j - b))^2,$$

although since the background has been already averaged over several Y values, in practice the error in the background, σ_b^2 , can and is neglected in the calculation of σ_c^2 .

If after five iterations, the calculation has not converged or if the calculated Poisson error of the centroid location is greater than the YERRMAX value in the TWOZXTAB, then it is assumed that the ALGNCORR step has failed to find the spectrum and no ALGNCORR correction will be applied. In this case the ALGNCORR keyword in the output files is set to “SKIPPED”, but the extraction will proceed under the assumption that the observed spectrum was already aligned with the reference profile but is too faint for the algorithm to detect. Tests using a wide variety of COS FUV data sets suggest that a YERRMAX value of about 0.8 pixels does a good job of distinguishing detections from blank images in most cases, and this value is currently used for all modes in the TWOZXTAB reference file. If the centroid location of the observed spectrum was successfully found, ALGNCORR is set to “COMPLETE” and a centroid calculation is also done for the reference profile. The spectrum offset is then set to difference between the observed and reference profile centroids:

$$SP_OFF = Y_c (\text{observed data}) - Y_c (\text{reference}),$$

and this value will be *subtracted* from the YFULL positions of all the individual observed photon events to bring the measured centers of the observed and the

reference profiles into alignment in YFULL coordinates. However, the positions of any events that fell into the WCA aperture region on the detector are not adjusted. These calculations are done separately for each detector segment and the resulting offsets are listed in the SP_OFF_A or SP_OFF_B keyword of the fits file extension header.

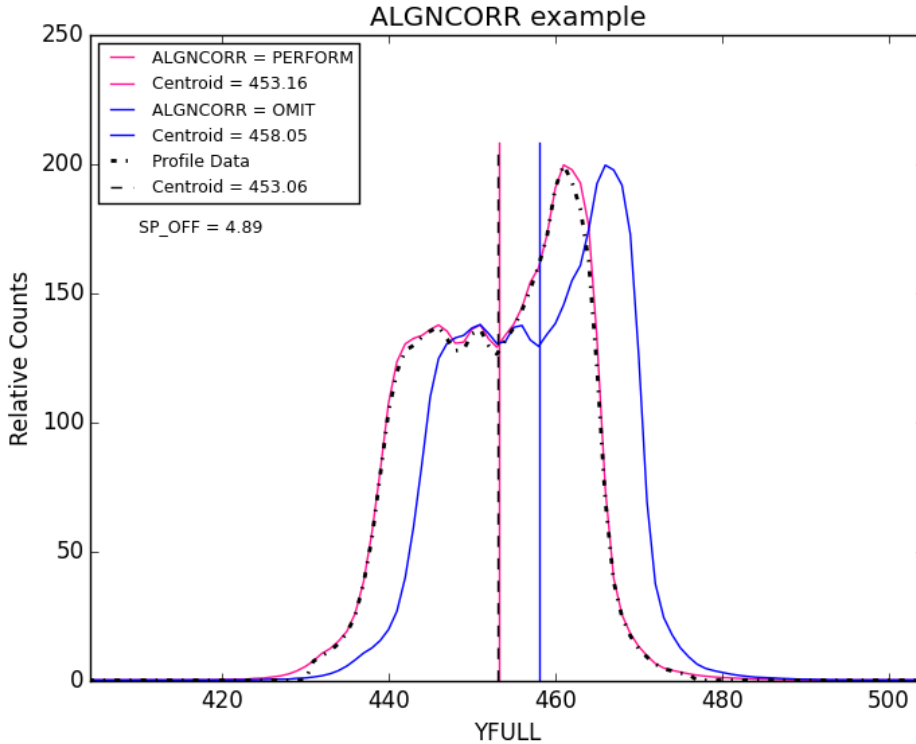


Figure 4: This figure shows the collapsed cross dispersion profile of a G130M 1222 FUVA with (pink line) and without (blue line) the correction from the ALGNCORR step. The broken line shows the collapsed cross dispersion reference profile to which the observed profile was aligned. The flux weighted centroid values for each of the profiles are also marked.

2.3 Defining the Extraction Zones

Once the observed and the reference profile are aligned, the YFULL pixel values for the upper and lower boundaries of the inner and outer zones are calculated. As discussed below, these boundary calculations for the pixel locations depend only on the tabulated 2D reference profile and the specified enclosed energy fractions, and do not directly use the observed spectral image.

For these calculations, the profile is first clipped to a box of size HEIGHT centered on the flux-weighted center of the profile, calculated during the ALIGNCORR phase. It is assumed that 100% of the profile's enclosed energy curve is contained within this box and any wings extending beyond this are ignored. In each column of the image, the YFULL locations are found below which the fraction of the total profile enclosed energy matches the LOWER_OUTER, LOWER_INNER, UPPER_INNER, and UPPER_OUTER values taken from the TWOZXTAB. Currently, for all modes for which TWOZONE extraction is supported, these have values of 0.005, 0.1, 0.9, and 0.995 respectively. Fractional "outer" pixel values are rounded up to the nearest

integer, while the “lower” values are rounded down. This procedure results in the outer zone containing $\sim 99\%$ of the profile’s enclosed energy while the inner zone contains the central $\sim 80\%$ of the enclosed energy. Due to the rounding to whole pixels, the actual enclosed energy values will vary slightly from column-to-column, and for the outer zone these are tabulated in the ACTUAL_EE vector, which is also output in the X1D file. The height of the extraction region at each XFULL location is saved in the NUM_EXTRACT_ROWS vector. The inner and outer boundaries are reported in the Y_LOWER_OUTER, Y_LOWER_INNER, Y_UPPER_INNER, Y_UPPER_OUTER vectors in the x1d files. Examples of how the final zones conform to one of the reference profiles are shown in Figures 1 and 5.

Background regions are defined centered on the B_BKG1 and B_BKG2 locations, each with the size given by the BHEIGHT value taken from the TWOZXTAB.

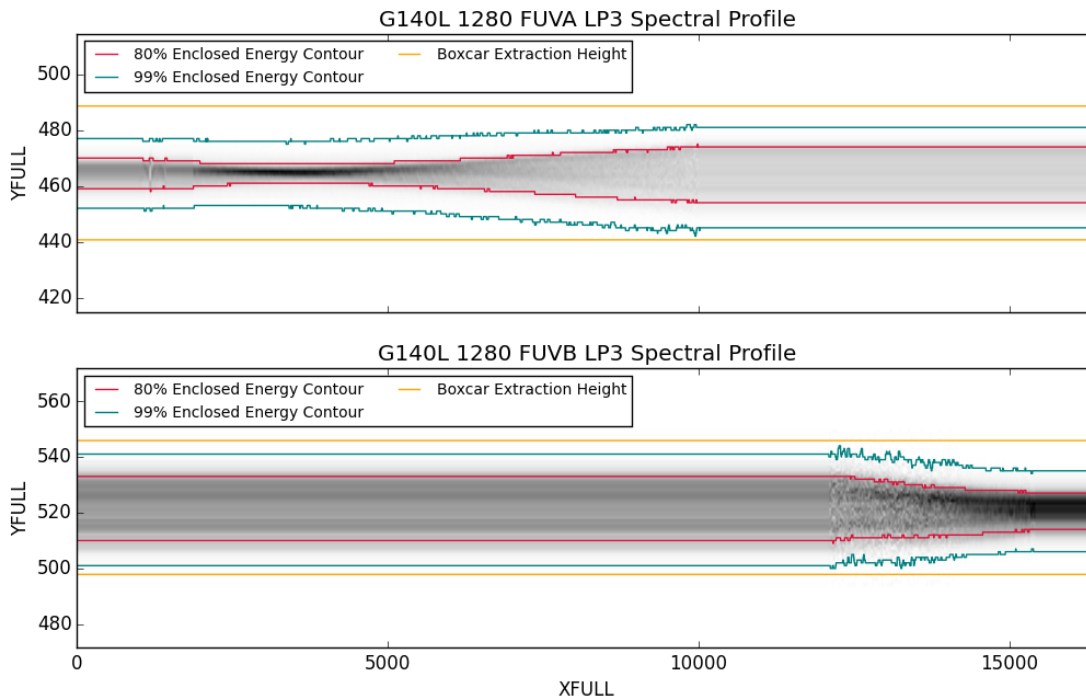


Figure 5: Images of the reference profiles for the G140L 1280 FUVA and FUVB segments are shown, together with the 99% and 80% enclosed energy contours. Also shown for comparison is a box of the same height as used for the BOXCAR extraction (yellow). Note that for this setting there are some regions where there is insufficient signal to clearly define the profile shape, and the last well measured profile is simply extended to cover the remainder of the detector.

2.4 Creating the Images Used for Extraction

For the TWOZONE extraction, CalCOS creates “Counts” and “Flat Fielded“ (FLT) images and output files in the same way as is done for the BOXCAR extraction. Each pixel in the “counts” image is simply the sum of the events that fall into that pixel according to their XFULL and YFULL coordinates. Of course in the case of the TWOZONE extraction the TRCECORR and ALGNCORR steps will have modified the YFULL coordinates of the individual events. The FLT image is similar except that the “EPSILON” values for each event that includes the effects of flat-field and dead-time corrections are applied.

The DQ image that gives the map of data quality values at each pixel location needs to be constructed differently for the two algorithms. The BPIXTAB reference file defines a number of rectangular regions, the locations of which are specified in XCORR and YCORR coordinates, along with the data quality flag to be applied to each of these regions. The GSAGTAB similarly specifies the rectangular regions where gain sag has caused a local sensitivity loss of more than 5% (i.e., where the modal gain is below 3). The regions specified in the GSAGTAB depend on the observation date and the high voltage setting used, but once the flagged regions appropriate for an individual observation are identified, they can be combined with the fixed regions taken from the BPIXTAB.

The DQ image array used in the extraction needs to be in the corrected XFULL and YFULL coordinates. Even for the case of the old BOXCAR algorithm, the correction between the CORR and FULL coordinates is not a simple shift, as both the correction estimated from the WAVECORR step and the Doppler correction for HST orbital motion can vary during the exposure, causing different shifts to be applied to different events. The bad pixel regions need to be smeared out over the range of these possible shifts to properly account for all the different XCORR, YCORR locations and any associated DQ flags that could have potentially contributed to an individual XFULL, YFULL pixel. As a result, one bad pixel in the XCORR-YCORR detector coordinates could map into several pixels in XFULL-YFULL coordinates.

Note that it is not possible to simply accumulate the DQ array from the DQ values attached to the events in the same way that the FLT or counts images are made, as the sparse nature of COS FUV data provides insufficient sampling. As an extreme example, for a region flagged as a dead spot there are likely to be no events sampling that region that could be used to transfer that DQ flag to the output XFULL YFULL coordinates.

For the BOXCAR algorithm, it is sufficient to shift and stretch each of the rectangular DQ regions using the range of SHIFT1 and SHIFT2 values found from the WAVECORR together with the range of X shifts found from the application of the DOPPCORR correction. This new set of shifted and expanded rectangular regions can then be used to construct the DQ image array for the observation.

For the TWOZONE algorithm, the differential column-by-column shifts applied by the TRCECORR step will distort the individual rectangular regions given in the BPIXTAB, GSAGTAB, and SPOTTAB reference files. So it is no longer sufficient to simply shift and expand the individual regions. Instead a full 2D image in XCORR-YCORR coordinates is first made from the BPIXTAB, GSAGTAB, and SPOTTAB and then distorted by applying the TRCECORR correction plus the ALGNCORR shift. This DQ map is then mapped to the final coordinates by dithering the image over the full range of shifts found by the WAVECORR plus DOPCORR steps, and then combining this stack of dithered images with a bitwise OR function to produce the final combined DQ image.

2.5 Extracting the spectrum

The GROSS count rate and the “effective count rate” (E_{CR}) vectors are created by simply summing each column in the counts and FLT images respectively from the lower to the upper boundary of the outer zone. A BACKGROUND_PER_PIXEL vector is also generated by summing the counts measured at each XFULL location in the background regions, dividing by the height of the background regions (BHEIGHT), and then smoothing with a square function of width BWIDTH. Since in the TWOZONE extraction the vertical size of the extraction region varies from column to column, the final BACKGROUND vector is then the product of this BACKGROUND_PER_PIXEL and NUM_EXTRACT_ROWS vectors to account for the column-to-column variation in the size of the extraction region. Both the BACKGROUND and BACKGROUND_PER_PIXEL are reported in the x1d file.

The NET vector is then calculated from these quantities in a similar way as is done for the BOXCAR extraction, except that the final NET vector is scaled to 100% enclosed energy.

$$\text{NET}[i] = (\text{GROSS}[i] - \text{BACKGROUND}[i]) \times (E_{CR}[i]/\text{GROSS}[i]) / (\text{ACTUAL_EE}[i])$$

The error in this net count rate is then calculated, properly taking into account the various scaling factors applied in the formula above, (see the COS Data Handbook for full details). Finally the FLUXCORR step is called to convert the NET and error in the net to the final FLUX and ERROR vectors output to the X1D file.

The 1D data quality or DQ vector is created by combining the DQ flags in the individual pixels of each column with a bitwise OR function, *but only over the range of the inner zone*. DQ flags that only fall in the outer zone are usually ignored. An exception is made if any DQ values in the outer zone have bits set that match bits set in the SDQOUTER header keyword; if any pixels in the outer zone are flagged with one of these values, it will be included in the extracted DQ vector. A DQ_BOTH vector that combines all the DQ flags over both the outer and inner region is also calculated, but subsequent steps of the calculation do not currently use this. Each wavelength bin in the extracted DQ vector is then compared to the SDQFLAGS (serious data quality flags) header keyword. Any wavelengths bits set in the DQ vector that match the values included in SDQFLAGS are considered as bad and discarded from inclusion in the X1DSUM files. These missing wavelengths will often be filled in when data from multiple FPPOS positions are combined to construct the final X1DSUM file.

For COS FUV data the SDQFLAGS keyword will usually be set to 2 (hot spots) + 8 (poorly calibrated region) + 16 (very low response regions with a >80% depression) + 128 (pixel out of bounds) + 8192 (gain sag holes) = 8346, while SDQOUTER will be set to 2 (hot spots). So gain sag holes and very low response regions will only cause a wavelength bin to be rejected if they are in the inner zone, while a hot spot will cause a wavelength to be rejected if the hot spot overlaps either the inner or the outer zone.

2.6 Comparison of the TWOZONE and BOXCAR extractions

The resulting differences between the new TWOZONE algorithm and the original BOXCAR algorithm are illustrated in Figure 6, which compares the wavelength bins that are rejected by each of the algorithms for a G130M 1291 observation taken at LP3.

The large extraction region of the BOXCAR algorithm results in significant overlap with the gain-sagged regions left behind from the detector usage at the original LP1. In each column where the gain sag region overlaps the extraction region, the entire wavelength bin will be marked as bad, resulting in a number of holes in the extracted spectrum.

The normal recommendation for COS FUV observations is to take a set of four such exposures, each at a different FP-POS offset. This shifts the detector XCORR location at which each wavelength falls, allowing many of the holes to be filled in the combined exposure at slightly reduced signal-to-noise. However, when the bad columns are caused by previous gain sag at LP1 the spacing of the holes will often match the FP-POS offset and some gaps in wavelength coverage will remain in the combined product. For this exposure, this would occur when using the BOXCAR extraction, especially in the region centered near XFULL=5000. However, in this case, the bad detector regions are far enough in the wings of the spectrum that including these “bad” bins would have no noticeable effect on the combined spectrum. The TWOZONE algorithm handles this by only flagging a wavelength bin as bad when a pixel in the inner zone is flagged as bad. This results in far fewer gaps in the extracted spectrum.

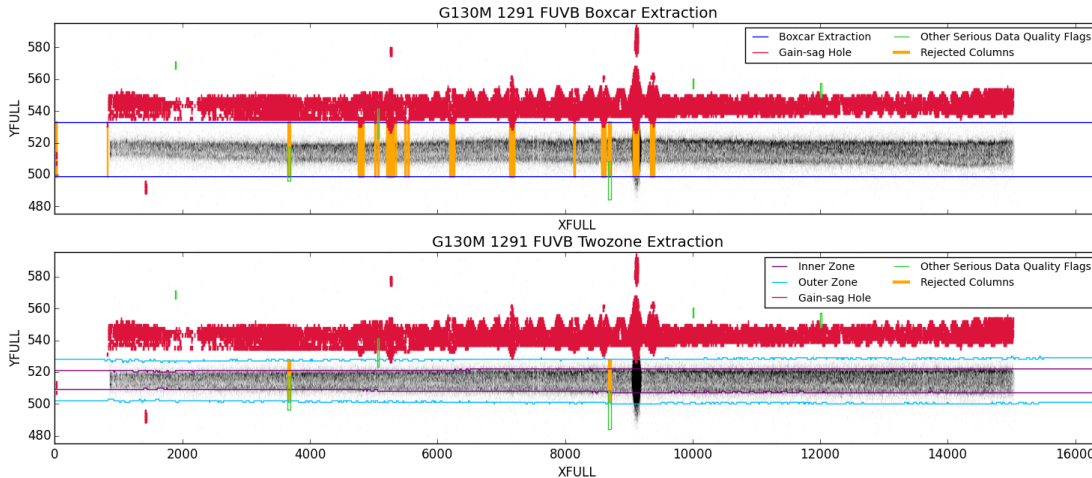


Figure 6: A comparison of the BOXCAR (upper panel) and TWOZONE (lower panel) extraction algorithms is shown for the FUVB segment of a G130M 1291 observation at LP3. The observed events are shown using a grey-scale image. Regions suffering more than 5% local sensitivity loss due to gain sag are marked in red, while other serious data quality flags in the extraction region are outlined with green boxes. The yellow bars show the wavelength bins that are rejected as bad by each of the algorithms. Note that since the TWOZONE algorithm only rejects wavelength bins when a bad pixel region lies within the inner zone, it rejects far less of the spectrum than does the BOXCAR algorithm.

It is important to consider how large the flux errors that result from ignoring bad detector regions in the outer zone will typically be. The worse case would be a completely dead region that completely spanned one side of the outer zone, but just missed penetrating the inner zone leading to an unflagged 10% artificial dip in spectrum. However, COS FUV observations are usually averaged over 4 FP-POS positions which put different wavelengths at different detector locations, so in the final summed products even such an extreme feature would be diluted by a factor of four, leaving only a 2.5% dip in the spectrum. This would only be detectable at 3σ with a S/N of 100. In practice, most cases will be much less severe than this, as the regions most damaged by gain sag at LP1 are well away from even the outer zone boundaries. More typical will be the overlap of the edge of a gain-sagged region with a small portion of the outer zone. If a region with a 20% DQE loss overlaps only 5% of the point-source profile energy, then the expected error in the flux measurement is only 1% at each FP-POS position, and could not be detected with the limited S/N of COS spectra. For all COS modes, the 80% enclosed energy boundary also does a good job of encompassing the steepest regions of the cross-dispersion profile where small misalignments of the observed and reference profile can produce significant changes in the enclosed energy.

2.7 New Columns in the X1D Output Files

Table 1 lists the columns in the x1d output files that are produced for each segment of each individual exposure. The columns listed in blue are the new columns produced by CalCOS versions 3.0 and later that were not included in earlier versions. These new columns are included for all COS data whether reduced with the TWOZONE or BOXCAR algorithm. They are intended to make it easier to understand how the extraction was performed, but subsequent calibration steps do not make use of them. Note that these new columns are included only in the X1D files for individual exposures and are not included in the X1DSUM files that combine the results for multiple exposures.

The Y_LOWER_OUTER and Y_UPPER_OUTER vectors give the YFULL pixel locations of the outer zone boundaries. All pixels from Y_LOWER_OUTER through Y_UPPER_OUTER are included in the extraction sums. The Y_LOWER_INNER and Y_UPPER_INNER boundaries similarly mark the pixel locations of the area used for the data quality (DQ) flag combination. Overplotting these vectors on the flat-fielded spectral images (FLT files), provides a useful way of verifying the success of the alignment algorithm, illustrating which detector pixels were included in the extraction, and which were included as part of the DQ combination (Figure 7). Since the sums are done over integer pixels, the actual fraction of the enclosed energy of the profile that is enclosed in the extraction box in each column will differ slightly from the nominal amount given by difference between the UPPER_OUTER and LOWER_OUTER values that were taken from the TWOZXTAB. The actual amount is tabulated in the ACTUAL_EE vector.

Because the extraction height will vary from column to column with the TWOZONE algorithm, the total BACKGROUND that is subtracted from the GROSS to derive the NET will also fluctuate with the height of the extraction region and so the X1D file

also includes the BACKGROUND_PER_PIXEL vector that gives the mean detector background per pixel (i.e., per 1 pixel high box), which corresponds to BACKGROUND divided by the number of rows in the background region.

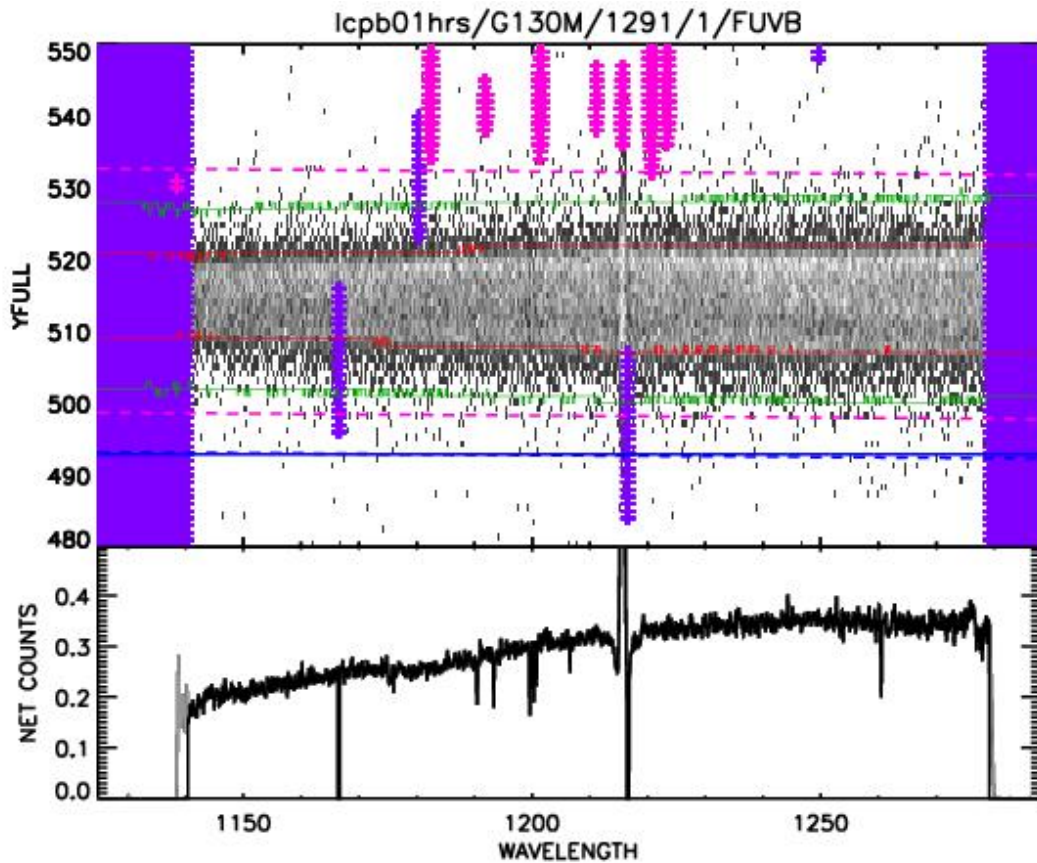


Figure 7: (top) FLT image of G130M/1291/FUVB exposures (grey scale). The inner and outer boundaries are shown as red and green lines (the BOXCAR extraction zone is shown as a pink dashed line for comparison). The purple and pink crosses correspond to bad and gain-sag pixels respectively. The blue solid lines indicate the background regions. (Bottom) Corresponding extracted net counts from the x1d spectrum.

Table 1: Columns in the x1d output file. The items in blue show the new columns that were added to the x1d file in CalCOS version 3.0 and later.

Column Name	Data Type	Units	Description
SEGMENT	CH*4	-	-
EXPTIME	D	s	-
NELEM	I	-	Number of elements in array
WAVELENGTH	D[16384]	Angstroms	-
FLUX	E[16384]	erg/s/cm ² /Å	-
ERROR	E[16384]	erg/s/cm ² /Å	-
GROSS	E[16384]	counts/s	-
GCOUNTS	E[16384]	counts/s	-
NET	E[16384]	counts/s	Normalized sum of counts over extraction region
BACKGROUND	E[16384]	counts/s	Total background subtracted from each column
DQ	I[16384]	-	DQ combined over inner region
DQ_WGT	E[16384]	-	-
DQ_OUTER	I[16384]	-	DQ combined over both regions
BACKGROUND_PER_PIXEL	E[16384]	counts/s	Background per pixel
NUM_EXTRACT_ROWS	I[16384]	pixels	Extracted height in each column
ACTUAL_EE	E[16384]	counts/s	Normalization factor in each column
Y_LOWER_OUTER	E[16384]	pixels	Pixel location of boundary
Y_UPPER_OUTER	E[16384]	pixels	Pixel location of boundary
Y_LOWER_INNER	E[16384]	pixels	Pixel location of boundary
Y_UPPER_INNER	E[16384]	pixels	Pixel location of boundary

4. Controlling the New Algorithm

A number of new keyword switches and reference files were introduced to support the new algorithm. Note that the delivered TWOZXTAB, TRACETAB, and PROFTAB reference files described below are only intended for use with data taken at LP3. The data in each of these reference files would need to be re-derived before application to other lifetime positions.

4.1 Header Keyword Switches

Several new keywords are now used to control the extraction algorithm:

1. **TRCECORR** – PERFORM or OMIT
2. **ALGNCORR** – PERFORM or OMIT
3. **XTRCTALG** – BOXCAR or TWOZONE

Currently the MAST pipeline will only use the new algorithm for FUV spectra taken at LP3 using the PSA and BOA apertures. Data taken at LP1 or LP2 will continue to use the BOXCAR algorithm as they have in the past. When using the old BOXCAR algorithm, the MAST pipeline will set the header keywords **TRCECORR** and **ALGNCORR** to "OMIT", and **XTRCTALG** to "BOXCAR". When using the new "TWOZONE" algorithm, **TRCECORR** and **ALGNCORR** are set to "PERFORM" and **XTRCTALG** to "TWOZONE". While it is possible to manually set other

combinations of these keywords, this is not normally recommended as it may result in an inaccurate calibration with the current reference files.

The offsets calculated by the ALGNCORR step for each detector segment are stored in the SP_OFF_A and SP_OFF_B keywords in the extension header of the x1d file. These keywords give the value subtracted from the YFULL location of each event during the ALGNCORR step. When manually recalibrating data with CalCOS, a user can force the ALGNCORR step to apply a different shift by setting the keywords SP_SET_A and/or SP_SET_B in the extension header of the input rawtag or corrtag file to the desired offset value.

4.2 The TRACETAB Reference File

The TRACETAB reference file (suffix *trace.fits*) gives the variation of the Y centroid of the spectrum (TRACE) from its median (TRACE_YLOC) value as a function of the XCORR (column) location along the dispersion direction in COS FUV data. The table contains arrays selected on the SEGMENT, OPT_ELEM, CENWAVE and APERTURE columns. During the TRCECORR calibration step, the TRACE array is interpolated to the XCORR value of each event to give the value of the shift to be included in the calculation of the final YFULL value for that event. This correction has the effect of removing residual geometric distortions that cause the spectral centroid to oscillate up and down by several pixels over the spectrum.

Table 2: Columns in the TRACETAB Reference File

Column Name	Data Type	Units	Description
SEGMENT	CH*4	-	Segment Name FUVA or FUVB
OPT_ELEM	CH*8	-	Grating name
CENWAVE	I	Angstrom	Central wavelength
APERTURE	CH*4	-	PSA or BOA
DESCRIP	CH*28		Pedigree information
TRACE_YLOC	R	pixel	YCORR location of center of trace (median)
TRACE	D[16384]	pixel	Y position correction array
ERROR	D[16384]	pixel	Trace profile error array

4.3 The TWOZXTAB Reference File

The TWOZXTAB reference file (suffix *_2zx.fits*) specifies a number of the parameters used in the new algorithm, including the area of the detector to search for the source spectrum, and the location and size of the background regions to use. These quantities are similar to values used in the XTRACTAB reference file that controls the BOXCAR extraction. However, in the TWOZXTAB the HEIGHT specifies the size of the initial region to search for centering the spectrum before the zone boundaries described above are specified, while B_SPEC gives the location at which to center this initial search region. Another difference is that the XTRACTAB allows separate heights, BHEIGHT1 and BHEIGHT2, to be specified for the two background regions, while the TWOZXTAB specifies a single BHEIGHT used for both extraction regions

When the ALGNCORR step calculates the flux-weighted center of the observed spectrum, it also calculates an estimate of the Poisson uncertainty in that location. When this error estimate becomes large, the calculated centroid is no longer a good estimate of the actual spectral location, possibly because no spectrum is visible in the image. So, when the Poisson error for the centroid location exceeds the quantity YERRMAX, the measured location is not trusted and no shift is applied by the ALGNCORR step. This causes the extraction to be done at a location centered on the unshifted reference profile. Tests using a wide variety of COS FUV data sets suggest that a YERRMAX value of 0.8 pixels does a good job of distinguishing detections from blank images for most cases, and this value is currently used for all modes in the TWOZXTAB reference file.

The TWOZXTAB also specifies how to calculate the boundaries illustrated in Figures 1 and 5 and the lower panel of Figure 6 that are used for the extraction and data quality flagging by defining the enclosed energy fraction that defines each of these boundaries. Currently for all modes where the new algorithm is applied, LOWER_OUTER = 0.005 and UPPER_OUTER = 0.995, which includes 99% of the enclosed energy as defined by each wavelength column of the reference profile in the extraction region, while LOWER_INNER = 0.1 and UPPER_INNER = 0.9 so that only bad pixel flags that fall within the central 80% of the enclosed energy are included in the DQ vector. However, in principle different values could be chosen for each GRATING/CENWAVE/SEGMENT combination.

Table 3: Columns in the TWOZXTAB reference file

Column Name	Data Type	Units	Description
SEGMENT	CH*4	-	Segment Name FUV A or FUV B
OPT_ELEM	CH*5	-	Grating name
CENWAVE	I	-	Central wavelength
APERTURE	CH*3	-	PSA or BOA
B_SPEC	D	Pixel	YFULL location of the middle of extraction region
HEIGHT	I	Pixel	Full height of extraction region
B_BKG1	D	Pixel	Center in YFULL of 1st background region
B_BKG2	D	Pixel	Center in YFULL of 2nd background region
BHEIGHT	D	Pixel	Full height of the background extraction regions
BWIDTH	I	Pixel	Width of boxcar smoothing for the background along the dispersion
LOWER_OUTER	D	-	Enclosed energy fraction defining lower edge of the total extraction zone
UPPER_OUTER	D	-	Enclosed energy fraction defining upper edge of the DQ combination zone
LOWER_INNTER	D	-	Enclosed energy fraction defining upper edge of the DQ combination zone
UPPER_INNTER	D	-	Enclosed energy fraction defining upper edge of the DQ combination zone
YERRMAX	D	Pixel	Maximum acceptable error for flux weighted centroid position
PEDIGREE	CH*8	-	-

4.4 The PROFTAB Reference File

The profile table gives the 2D profile of a point source as a function of column number (XFULL) in COS FUV data. It is assumed that the small-scale detector dependent structure that is included in the TRACETAB reference file has been removed prior to the generation of these profiles. The flux-weighted centroid of the resulting profiles should therefore be the same in all columns.

The row from the PROFTAB used by CalCOS is selected by matching the SEGMENT, OPT_ELEM, CENWAVE and APERTURE columns to the settings of the particular observation. In practice, profiles have been only tabulated for the PSA aperture. In the delivered reference file, the APERTURE column is set to 'ANY' to allow the PSA profiles to also be applied to observations taken through the BOA aperture. It is unclear how this affects the quality of the resulting calibration; however, COS FUV spectral observations with the BOA aperture are very rarely performed.

To save space, the profiles are tabulated only over a portion of the full 1024 pixel height of the COS FUV detector arrays. The ROW_0 column specifies to which row in the full sized array the first row in the table corresponds (starting from 0). The CENTER column specifies the YFULL coordinate of the centroid of the profile in the full 1024 high array.

An image of the reference profile for G140L 1280 is shown in Figure 5, and sample profiles for G130M and G160M are shown in Figures 8 and 9. A more detailed description of how the trace and profile reference files were created will be presented in a forthcoming ISR by Ely et al. (2015).

Table 4: Columns in the PROFTAB Reference File

Column Name	Data Type	Units	Description
SEGMENT	CH*4	-	Segment Name FUV A or FUV B
OPT_ELEM	CH*8	-	Grating name
CENWAVE	I	Angstrom	Central wavelength
APERTURE	CH*4	-	PSA or BOA
DESCRIP	CH*8	-	Pedigree information
CENTER	R	pixel	Profile centroid
ROW_0	I	pixel	Row offset of profile array
PROFILE	R[16385, 339]	-	Profile in (XFULL, YFULL), offset by ROW_0

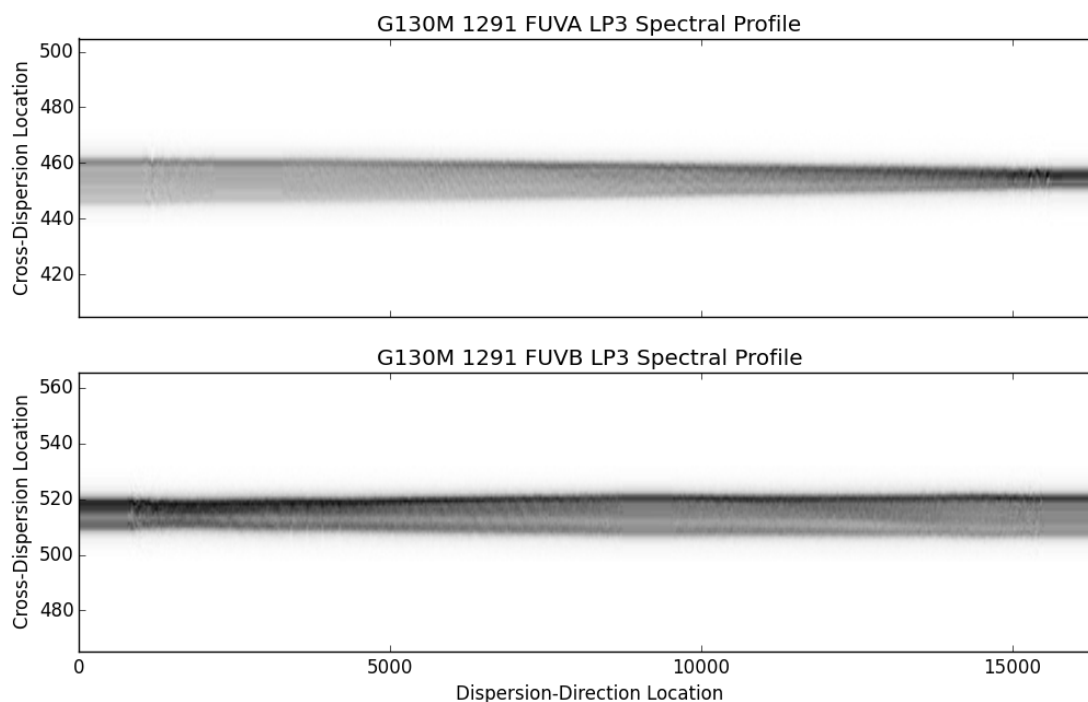


Figure 8: Images of the reference profiles for the G130M 1291 setting at LP3 on the FUVA and FUVB segments are shown – note the very asymmetrical and highly wavelength dependent nature of the cross-dispersion profile. At most wavelengths, the G130M cross dispersion profile is double peaked due to the astigmatism in the cross-dispersion direction.

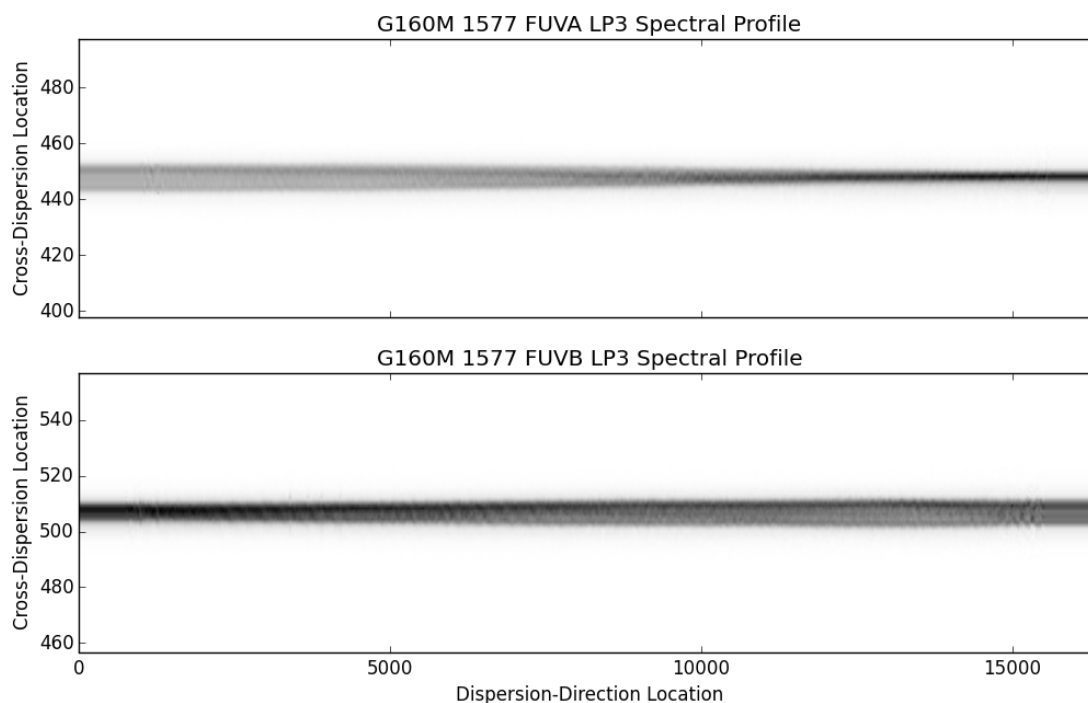


Figure 9: Images of the reference profiles for the G160M 1577 setting at LP3 on the FUVA and FUVB segments are shown – note the highly wavelength dependent nature of the cross-dispersion profile. The COS FUV cross-dispersion profiles vary considerably between wavelength settings, and tend to become narrower at longer wavelengths.

5 Customizing the TWOZONE Extraction

The new TWOZONE extraction algorithm and the associated pipeline reference files are optimized for the spectral extraction of bright point source spectra. There are, however, a number of circumstances under which a customized extraction might yield better results.

- For extended sources the use of the TWOZONE algorithm may lead to an underestimate of the measured flux and/or poor alignment of the extraction region with the source.
- For very faint point or extended sources, the ALGNCORR step may not always be able to reliably measure the position of the source, and may incorrectly default to assuming that the target is already aligned with the reference profile.
- For faint point sources it may be possible to significantly improve the signal-to-noise ratio by reducing the extraction height to minimize the included detector background.
- In some cases it may be useful to check the reliability of the extraction by varying the size of the inner and outer zones to see how changes affect individual spectral features.

If the user simply wishes to use the BOXCAR extraction in place of the TWOZONE algorithm, XTRCTALG should be set to “BOXCAR”, and TRCECORR and ALGNCORR should be set to “OMIT”. This will use the larger extraction regions defined for that algorithm in the XTRACTAB reference file. However, for observations at LP3 there may be significant overlap with the gain-sagged regions near LP1, and this may affect the quality of the spectra. It may be useful to also adjust the B_SPEC and HEIGHT parameters in the XTRACTAB reference file to fine-tune the location and size of the BOXCAR extraction region.

Both the XTRACTAB, which is used with the BOXCAR extraction and during the WAVECORR step for both algorithms, and the TWOZXTAB, which is used with the TWOZONE extraction algorithm, contain columns named HEIGHT and B_SPEC. For the BOXCAR algorithm, these parameters together with the SLOPE column directly control the size and location of the extraction region. For the TWOZONE algorithm, the HEIGHT and B_SPEC numbers instead control the size and initial location of the region used for the ALGNCORR step. The HEIGHT column in the TWOZONE algorithm is also used to define the cross-dispersion width of the reference profile that is assumed to include 100% of the enclosed energy. The actual extraction region at each wavelength is adjusted so that the enclosed energy fraction of the reference profile matches the values given in the LOWER_OUTER and UPPER_OUTER columns of the TWOZXTAB. For example, for the default values of LOWER_OUTER=0.005 and UPPER_OUTER=0.995, at each wavelength the extraction region is adjusted so that the central 99% of the encircled energy as measured from the reference profile is included. Fractional pixel locations are rounded outwards, and the final extracted net count rate and flux values will be scaled for the exact encircled energy fraction in each column by dividing the background subtracted count rate by ACTUAL_EE.

To force a spectral extraction using the TWOZONE algorithm to sum over a region that contains only the central 80% of the reference profile's encircled energy, the user would need to change LOWER_OUTER to 0.1 and UPPER_OUTER to 0.9 in the appropriate row of the TWOZXTAB prior to recalibration of the data. For very faint sources where the detector background is larger than the light in the wings of the source profile, the use of a narrower extraction region may significantly increase the signal-to-noise ratio in the final extracted spectrum.

Values of 0 or 1 for the enclosed energy boundaries have a special meaning. Setting the LOWER_OUTER to a value of 0 forces the extraction to start at the bottom of the region defined by a rectangular box of size HEIGHT, while setting the UPPER_OUTER to 1 forces it to end at the upper boundary. This allows the use of a rectangular extraction box rather than the wavelength dependent extraction region based on the profile shape that is normally used for the TWOZONE algorithm. As is the case with the BOXCAR algorithm, it may be useful to tune the B_SPEC and HEIGHT values when setting the TWOZONE algorithm to use a fixed rectangular region, although in this case the values of B_SPEC and HEIGHT in the TWOZXTAB file will need to be edited rather than the values in the XTRACTAB.

The LOWER_INNER and UPPER_INNER columns in the TWOZXTAB behave very similarly to the "OUTER" boundaries, except that they are used to control the region over which data quality flags are combined rather than the region over which counts are summed. The user can also adjust these values by editing the rows of the TWOZXTAB reference file.

The background regions in the TWOZONE algorithm can also be modified. To change where the background regions are located or the height of the background regions, edit the background centers (B_BKG1 and B_BKG2) and the background height (BHEIGHT) in the TWOZXTAB. The background regions should not be placed directly above the spectrum at LP3, as that is where LP1 was located, and the detector is therefore very gain-sagged in that location. Users should also avoid having the background regions overlap the WCA (Wavelength Calibration Aperture) region where the PtNe line lamp flashes for wavelength calibration fall. The WCA locations and sizes can be found in the XTRACTAB reference file. Currently the ALGNCORR step only supports selection of the same height for both background regions via the BHEIGHT column; different heights for the two background regions are not supported.

The user can also override the shifts calculated by ALGNCORR. This can be useful if the automatic algorithm failed to properly center the target. To do this, the user should change the value of the keyword SP_SET_A, (for detector segment FUVB), or SP_SET_B, (for FUVB), to the desired offset value prior to running CALCOS. This will cause the specified value be used in place of the SP_OFF_A or SP_OFF_B value calculated by the ALGNCORR algorithm. These keywords should be set in the first [1] extension header of the rawtag or corrtag file used as input for CALCOS. *Note that the SP_OFF_* and SP_SET_* keywords give the amount by which the uncorrected spectrum is offset from the reference profile, so a **positive** value of the SP_SET_A or SP_SET_B will move the observed spectral image **down** in YFULL coordinates.*

Change History For COS ISR 2015-03

Version 1: 25 June 2015 – Original Document

References

- Ely, J., et al. 2015, Technical Instrument Report COS 2015-xx (in progress)
Fox, A., et al. 2015, “COS Data Handbook”, Version 3.0, (Baltimore: STScI)
Massa, D. and York, B. et al. 2013, “COS Data Handbook”, Version 2.0, (Baltimore: STScI)
Osten, R. et al. 2013, Instrument Science Report COS 2013-16, (Baltimore: STScI)
Sahnow, D. J. et al. 2011, “Gain Sag in the FUV Detector of the Cosmic Origins Spectrograph”, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2011SPIE.8145E..0QS (also ISR COS 2011-05)
Sahnow, D. J. et al. 2013, Instrument Science Report COS 2013-13, (Baltimore: STScI)

Appendix: Changes between CalCOS 3.0 and CalCOS 3.1

This ISR describes the TWOZONE extraction algorithm as it will be implemented in version 3.1 of CalCOS, which is expected to be released in the late summer or early fall of 2015. An earlier version (CalCOS 3.0) of the TWOZONE algorithm had been released in early 2015 to support the initial move of COS FUV spectra to LP3. The differences between this original version 3.0 and the newer version 3.1 described above are documented below.

In version 3.0, the sums over the extraction region in each column were done from Y_LOWER_OUTER to Y_UPPER_OUTER-1, i.e., the pixel at Y_UPPER_OUTER was not included. In version 3.1 the sums are now done from Y_LOWER_OUTER to Y_UPPER_OUTER including both boundary pixels. This now changes the value NUM_EXTRACT_ROWS column from Y_LOWER_OUTER-Y_UPPER_OUTER to Y_LOWER_OUTER-Y_UPPER_OUTER+1. A similar change was made for the inner zone so that the pixel at Y_UPPER_INNER is now included when combining the DQ flags in the inner zone.

Some of the other new information columns output to the X1D file were incorrectly populated in CalCOS 3.0; these will be corrected in the version 3.1 release. These new columns are also included when the BOXCAR extraction algorithm is performed, but in version 3.0 for data using the BOXCAR algorithm the Y values for the upper and lower boundaries were set to the scalar value at column 0 and did not include the ~2 pixel tilt of the region used for the BOXCAR extraction. The X1D output vector called BACKGROUND_PER_ROW in 3.0 will be renamed BACKGROUND_PER_PIXEL in 3.1; for data reduced with the BOXCAR algorithm, this vector was not correctly populated in version 3.0.

In addition to the incorrect population of some of the new informational columns described above, a few remaining issues with the improved extraction algorithm in CalCOS 3.0 were identified. None of these issues compromised the calibration of data at LP3. Fixes for these issues will be included in version 3.1.

CalCOS version 3.0 did not properly shift the flagging of the bad pixel regions for the shifts calculated by the TRCECORR and ALGNCORR steps (see section 2.4 above). To ensure that this bug did not cause bad wavelength columns in the extracted spectrum to be erroneously flagged as good, the bad pixel map in the LP3 region was adjusted to be more conservative in the flagging of bad regions by expanding the vertical extent of these bad regions by 6 pixels in the region of the detector near LP3. This workaround will be removed from the bad pixel table once version 3.1 has been released to the community and is in use in the MAST pipeline.

In version 3.1 ALGNCORR will be changed so that it no longer rejects columns flagged as containing gain sagged pixels (DQ value 8192). This allows the alignment algorithm to work well even when the edge of the initial alignment region overlaps a severely gain-sagged part of the detector. Columns containing gain sagged regions in the inner zone are still flagged as bad during the extraction step and excluded from inclusion in the X1DSUM files.

Version 3.1 introduces a new SDQOUTER keyword to allow selected DQ flags to force rejection of a wavelength bin even if only detector pixels in the outer zone are affected. A new “_spot” reference file, which allows bad pixels to be specified over limited time intervals, was also implemented. Together these changes will allow exclusion of transient hot spots on the detector from the extracted net and flux vectors.

Other minor issues in version 3.0 included incorrect format specifications for some columns in the output fits files, and a bug that prevented successfully running CalCOS if any of the enclosed energy fractions specified in the TWOZXTAB were set to either 0 or 1.