Description of the NIRSpec linearity correction and saturation check reference files

Abstract:

This document describes the format of and the data and algorithms used to create the NIRSpec linearity correction and saturation check reference files that are input for the STScI ramps-to-slopes pipeline.

Change log:

Version	Date	Description of changes
1	Nov 1, 2013	Initial version for build 2 pipeline
2	Nov 18, 2014	Update to reflect changes in reference files for build 3 pipeline

1 INTRODUCTION

The STScI processing pipeline needs reference files in order to compute count rate images from the input up-the-ramp data, a process often referred to as "ramps to slopes". This document describes the algorithms and data used to create the linearity correction and saturation check reference files delivered to STScI.

The linearity correction file provides the polynomial coefficients in order to correct for the integrated counts in NIRSpec exposures for the non-linear response of the detectors. The saturation check files give the number of counts up to which the non-linearity is considered to be correctable. Pixel values above the saturation threshold should be considered saturated and unusable for science.

2 REFERENCE FILE FORMAT

All reference files are stored in fits format as defined by Hanisch et al. (2001). The primary data array is always empty and the actual reference data is stored in image extensions. In the following sections we briefly describe the primary header, the image extensions, and the data quality maps.

2.1   Primary header and extensions

Each reference file has an empty primary data array, with the primary header containing all the necessary keywords according to Giardino (2013) and STScI (2014). The `REFTYPE` keyword has the value 'LINEARITY' for the linearity correction reference files and 'SATURATION' for the saturation check reference files. Furthermore, the primary headers contains more (non-mandatory) keywords related to the creation of the reference files.

The actual reference data is always stored in image extensions. Tables 1 and 2 below lists the extensions in the linearity and saturation reference files, respectively.

#	Name	Dimensions	Type	Description
1	COEFFS	NCOL×NROW×NC	-32	Linearity correction coefficients data cube
2	DQ	NCOL×NROW	16	Linearity correction quality bit map
3	DQ_DEF	binary fits table	N/A	Meaning of data quality bits/values

Table 1: Number, name, dimensions, type, and description of the extensions in the linearity correction reference files. NCOL and NROW correspond to the dimension of the detector area read-out (2048×2048 pixels for full frame). NC gives the number of coefficients of the linearity correction polynomial. The given type denotes the BITPIX value of the data, i.e. -32 corresponds to single precision float and 16 to a short integer (unsigned), respectively.

#	Name	Dimensions	Type	Description
1	SCI	NCOL×NROW	16	Pixel saturation threshold map
2	DQ	NCOL×NROW	16	Linearity correction quality bit map
3	DQ_DEF	binary fits table	N/A	Meaning of data quality bits/values

Table 2: Number, name, dimensions, type, and description of the extensions in the saturation check reference files. Description as for Table ?? above.

The `COEFFS` extension in the linearity correction reference file holds the polynomial correction coefficients for each detector pixel. The linearity correction algorithm uses a polynomial of the form

$$F_c = c_0 + c_1 \times F + c_2 \times F^2 + c_3 \times F^3 + \ldots + c_{n-1} \times F^n \tag{1}$$

where F_c is the corrected count number (in DN), c_0, c_1, c_2, $\ldots$, and c_{n-1} are the polynomial correction coefficients (n in total), and F is the uncorrected count number (in DN). Here, uncorrected means before linearity correction, but after (super) bias subtraction and reference pixel correction. In principle, n can be any number, but for the delivered reference files the polynomial degree is always four. Therefore, the dimension of the `COEFFS` image extension in the delivered linearity correction files is 2048×2048×5. In order to linearity correct window mode exposures, the appropriate detector area is cut-out and used.

The `SCI` extension in the saturation check reference files holds the pixel-by-pixel saturation information as a 2-D array of integers. Pixel values that exceed the saturation values are considered saturated and should not be used to determine the count rate later. As this saturation detection is one of the first steps in the ramps-to-slopes pipeline, the saturation values are before (super) bias subtraction or reference pixel correction.

The `DQ` extension holds the data quality bit map as a 2-D array of integers, with the `DQ_DEF` extension describing their meanings. The currently defined values are given in the next section.

2.2   Data quality maps

Data quality maps hold information about bad pixel or unwanted behavior as bit values in a two dimensional map. The currently defined values for the linearity correction and saturation check reference files are listed below in tables 3 and 4, respectively.

Bit	Value	Name	Description
0	1	DO_NOT_USE	Bad pixel, do not use (for whatever reason). Currently never set when creating reference files.
1	2	NONLINEAR	Pixel is (highly) non linear and linearity correction is deemed unreliable. Therefore, for the current reference files, the coefficients will be set so that no linearity correction is performed at all (like for NO_LIN_CORR below).
2	4	NO_LIN_CORR	No linearity correction available, i.e. the coefficients in equation 1 are $c_0 = 0$, $c_1 = 1$, and $c_{2...n-1} = 0$.

Table 3: Bit, value, name, and description of the data quality map bitplanes used in the linearity correction reference files. These are stored in the DQ_DEF extension as a binary fits table.

Bit	Value	Name	Description
0	1	DO_NOT_USE	Bad pixel, do not use (for whatever reason). Currently never set when creating reference files.
1	2	NO_SAT_CHECK	No saturation check data available, e.g. for low dynamic range pixels. The saturation value in the `SCI` extension is set to 65535 (ADC limit) in those cases. Name was previously NO_SAT_CHCK (mind the missing 'E').

Table 4: Bit, value, name, and description of the data quality map bitplanes used in the saturation check reference files. These are stored in the DQ_DEF extension as a binary fits table.

3   REFERENCE FILE CREATION

3.1   Linearity correction and saturation data

In order to determine the polynomial coefficients for linearity correction, one needs illuminated exposures that i) integrate slowly up the ramp, ii) reach full well or ADC saturation for all pixels eventually, and iii) have stable illumination over time. This data cannot be obtained at NIRSpec level, due to light being blocked in the aperture plane by the MSA and slit masks. Therefore, the necessary exposures have been obtained at FPA level testing at the Goddard DCL. There are 25 exposures in total with 88 groups each and a count rate of approximately 100 DN/s.

These exposures with counts going slowly up the ramp are then averaged the following way:

1. The super bias obtained under dark conditions is subtracted from each group.

2. Each group and output of input exposures is reference pixel subtracted. This means, that for each group/frame, each output (four in total), and odd/even columns separately, the average of the four top and bottom row pixels is computed. These averages are then subtracted from all pixels in the appropriate output and odd/even columns.

3. The reference pixel subtracted exposures are averaged group by group, pixel by pixel, using a sigma-clipped mean (3-σ outlier rejection, 3 iterations). The sigma-clipped mean is used in order to remove outliers from the stack that might be, for example, due to cosmic ray hits.

The resulting high signal-to-noise data cube is then used as input for the linearity correction coefficients determination program.

The linearity correction is determined on a pixel-by-pixel basis. For each pixel, the first groups with a total count level of less than 10,000 DN (above the bias) are used to estimate the count rate at zero counts. This is done by fitting an exponential function

$$x_i = a + \alpha(1 - \exp(-i/\beta)) \tag{2}$$

to the data, where i is the group number starting at 1 and x_i is the measured data value (in DN) at the given group, for all $x_i \leq 10000$. The slope b at the zeroth group ($i = 0$) is then

$$b = \frac{\alpha}{\beta}. \tag{3}$$

The predicted counts for a fully linear pixel are then

$$y_i = a + b \times i \tag{4}$$

with a and b from equations 2 and 3, respectively. In order to determine the correction coefficients from equation 1, we fit

$$y_i = c_0 + c_1 \times x_i + c_2 \times x_i^2 + c_3 \times x_i^3 + c_4 \times x_i^4, \tag{5}$$

starting using all groups, i.e. $i = 1 \ldots 88$. As this will likely result in using saturated data, the fit will be poor and the reduced χ^2 will be large. In order to find the saturation threshold and obtain a good fit, the number of used groups will be reduced, removing data from the end of the exposure, and the fit in equation 5 repeated. This is done until χ^2 rises again or is below 1.5 times the minimum χ^2 obtained for any range of i. The count value of the last group for which the goodness of the fit matches the criteria given above is taken as the saturation limit, and its value plus the bias is stored in the SCI extension of the saturation check reference file. As we are only interested in the change of slope due to the linearity correction, the fist coefficient is set to zero $c_0 = 0$ in all cases.

3.2   Data quality flags

The algorithms to create the data quality bits for the linearity correction reference files that are given in table 3 are described below.

The NONLINEAR bit is set, if the χ^2 value for the fit from equation 5 is unusually high or the derived correction at maximum counts (i.e. the saturation limit) is large. A χ^2 is considered unusual, if it is more than five times greater than the standard deviation of all χ^2 for the full detector. Linearity correction is considered large, if $F_c/F - 1 > 0.5$ for any value F below saturation. Currently, the linearity correction coefficients are set to $c_1 = 1$ and $c_2 \ldots_4 = 0$ in the above cases, i.e. the linearity correction is switched off.

The NO_LIN_CORR bit is set, if the coefficients could not be determined, e.g. due to a too low dynamic range of the pixel (last group in illuminated exposure below 10000 DN). This is for example true for all reference pixels. In this case, the linearity correction coefficients are set to $c_1 = 1$ and $c_2 \ldots_4 = 0$, and therefore $F_c = F$ (linearity correction switched off, like above).

The NO_SAT_CHECK bit is set, if no reliable saturation limit could be determined. This is again true for low dynamic range pixels and reference pixels. In these cases, the value in the SCI extension of the saturation check will be equal to 65535.

4   DELIVERED REFERENCE FILES

In this section we list the delivered linearity correction reference files. The reference files are derived from data taken at the Goddard DCL during FPA104 characterization. Because the saturation levels now include the bias, there is no need to derive new saturation check reference files due to the detector degradation and the resulting increase in hot pixels. One must note that for hot pixels that exhibit field-enhanced emission, the linearity correction will be incorrect, due to the intrinsic nonlinearity of this effect.

Table 5 on the next page lists the reference files provided in this delivery.

4.1   Data used

Table 6 gives the NIDs of the exposures that were used to create the reference files. This information, as well as the full filename of all input files, is also available in the primary header of the reference files.

#	Filename	Use after
1	nirspec_linc_nrs1_f_03.00.fits	
2	nirspec_linc_nrs2_f_03.00.fits	2001-01-01T00:00:00
3	nirspec_satc_nrs1_f_01.00.fits	
4	nirspec_satc_nrs2_f_01.00.fits	

Table 5: The number, filename, and use after date for the delivered linearity correction reference files. The value of the "use after" column corresponds to the USEAFTER keyword in the primary fits header.

Number	Data used
1 to 4	FPA104 characterization data obtained at Goddard DCL. NIDs 5390 to 5414 (all 88 group full frame exposures under illumination).

Table 6: The data used to create the reference files listed in table 5.

5 REFERENCES

Giardino, G. 2013, NIRSpec reference data product specification, NIRSpec Technical Note NTN-2013-007, ESA/ESTEC

Hanisch, R. J., Farris, A., Greisen, E. W., et al. 2001, Astronomy & Astrophysics, 376, 359

STScI. 2014, JWST Calibration Reference Files: File Formats for the Build 3 Pipeline, https://confluence.stsci.edu/display/JWST/, STScI