STScI | SPACE TELESCOPE SCIENCE INSTITUTE

JWST TECHNICAL REPORT

Title: Frame 0 Analysis for NIRCam Integrations	Doc #: JWST-STScI-006203, SM-12 Date: 12 March 2018 Rev: -
Authors: A. Rest, A. Canipe, B. Hilbert Phone: (410) 338-4358	Release Date: 2 May 2018

1 Abstract

We describe how frame 0 can be used to increase the dynamic range of data read out with grouped readout modes. We discuss the different noise sources, and the best ways to minimize these noise sources, in particular any systematic bias offset. We show that attention needs to be paid not only to signals that saturate in the first group, but also to the ones that saturate in the second group. Finally, we evaluate whether or not it is advantageous to add frame 0 to the ramp-fitting and provide our recommendation for how to implement these changes in the pipeline.

2 Introduction

In order to decrease data volume and thus the requirements on the download bandwidth, NIRCam offers observers the option to average multiple non-destructive readouts (frames) into each group of an exposure. See Table 2-1 for a list of averaging scenarios. This averaging is done on the spacecraft and cannot be undone; observers will not have access to the original unaveraged frames. Figure 2-1 shows a ramp using the SHALLOW4 readout pattern in which 4 individual frames are averaged into one group, and one frame is skipped.

One complication of this grouping strategy is how to deal with saturated pixels. For instance, using the example above where four frames are averaged into the first group, a pixel that saturates in the 2nd, 3rd, or 4th frame renders the first group's flux measurement useless, despite the fact that the signal in the original first frame is valid (see the red line in Figure 2-1 for an illustration). In this case then, the pixel would have no good flux measurements.

To remedy this, the first frame (frame 0) is saved separately, with the goal to use frame 0 to calculate the flux in pixels where the grouped data is saturated in the initial group. This can significantly increase the dynamic range, e.g. by $\Delta m = 2.5 log_{10}(8) = 2.26$ magnitudes for readout modes like DEEP8 in which 8 frames are averaged, or $\Delta m = 1.5$ for the readout pattern shown in Figure 2-1.

Operated by the Association of Universities for Research in Astronomy, Inc., for the National Aeronautics and Space Administration under Contract NAS5-03127

Table 2-1 A list of available NIRCam MULTIACCUM readout patterns. The number of samples is the number of frames plus the number of frames that are skipped.

Readout Patterns	Samples (per group)	Frames Averaged (in each group)
RAPID	1	1
BRIGHT1	2	1
BRIGHT2	2	2
SHALLOW2	5	2
SHALLOW4	5	4
MEDIUM2	10	2
MEDIUM8	10	8
DEEP2	20	2
DEEP8	20	8

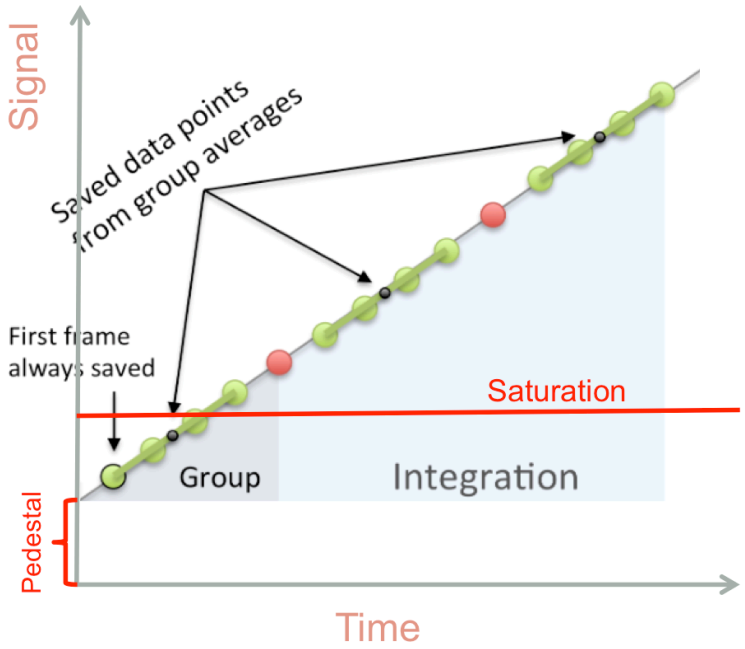


Figure 2-1 Illustration of a SHALLOW4 readout pattern that averages 4 frames and skips one frame to create each group. This is one of the standard NIRCam readout modes listed in Table 2-1.

3 Method

Currently, for a pixel which has only one valid group, the slope s is calculated by dividing the flux f_0 in the first group by its time Δt instead of calculating a slope through a line fit as shown in Equations 1 and 2.

$$s = f_0 \quad (1)$$

$$f_0 = f + f_{ped} \quad (2)$$

Here, f is the true accumulated flux, and f_{ped} is a pedestal, e.g. due to an imperfect reduction of a bias offset. Therefore, the frame 0 image can be run through the same JWST Science

Calibration Pipeline (hereafter referred to as the pipeline) steps as a normal ramp, with the exception of the jump detection step (alternatively, the jump detection step could internally not do anything to a frame 0 image, in which case the exact same recipe can be used for both the normal ramp and the frame 0 image).

3.1 Noise Sources

When the slope s is determined by fitting a line, any pedestal f_{ped}, e.g. due to the imperfect removal of the bias offset, will only cause a non-zero fitted y-intercept, and it therefore does not introduce any noise into the slope measurement itself. However, when the slope s is calculated with the frame 0 image using Equation 1, then any pedestal f_{ped} will bias the slope measurement. Here, we quantify how these noise sources impact the slope measurements of a given pixel and of an aperture of N pixels.

$$\sigma_{f0,Pix}^2 = \sigma_P^2 + \sigma_{RD}^2 + \sigma_{SB,stat}^2 + \sigma_{SB,sys}^2 \quad (3)$$

$$\sigma_P^2 = f/g \quad (4)$$

Equations 3 and 4 show the various noise contributors to a pixel in frame 0, where g is the gain and the different noise sources are explained in Table 3-1.

We examine the noise contribution from each source in a group where the pixel reaches saturation at the end of the group. The first three terms are random in nature. For readout patterns like DEEP8, which average 8 frames, the flux values f range from 50000/8 = 6250 to 50000 ADU, assuming a saturation value of 50000 ADU. Thus, the Poisson σ_P of the flux f ranges from 50 to 140 ADU, assuming a gain g = 2.5. Even the lowest value of σ_P = 50 is high compared the typical values of $\sigma_{SB,stat} \approx 20$ ADU, and therefore the random noise in the frame 0 calculation are dominated by σ_P.

The systematic noise introduced by the superbias subtraction can be divided into subcomponents

$$\sigma_{SB,sys}^2 = \sigma_{SBO}^2 + \sigma_{SB,1/f}^2 + \sigma_{SB,odd/even}^2 \quad (5)$$

where σ_{SBO}^2 is a constant bias offset for the full image, $\sigma_{SB,1/f}^2$ is any residual left from imperfect $1/f$ removal using the reference pixels, and $\sigma_{SB,odd/even}^2$ is any residual left from imperfect removal of the bias structure in the odd and even columns.

We separate these noise sources because their different types of spatial correlation propagate differently when several neighboring pixels are averaged, e.g. when they fall into the same aperture used for photometry. This is a realistic scenario; for example, it is likely for a saturated star that several neighboring pixels need to have their slope calculated with the frame 0 method. We calculate the noise if N pixels reduced by the frame 0 method are averaged using Equation 6.

$$\sigma_{f0,N\,Pix}^2 = \frac{1}{N}\left(\sigma_P^2 + \sigma_{RD}^2 + \sigma_{SB,stat}^2\right) + \sigma_{SBO}^2 + \frac{1}{\sqrt{N}}\sigma_{SB,1/f}^2 + \frac{|N_{odd} - N_{even}|}{N}\sigma_{SB,odd/even}^2 \quad (6)$$

The first three noise sources are Poissonian in nature, and therefore their variances decrease by a factor of N. However, the last three potential noise sources decrease by smaller factors. Most notably, if there is a small systematic offset σ_{SBO} left after superbias and reference pixel correction, it can become the dominant noise source since it does not decrease with N. In the case where the $1/f$ noise is incompletely removed, we can assume that a given row has the same systematic offset $\sigma_{SB,1/f}$ and therefore the remaining variance decreases only by $\sqrt{N}$ and not by

N. Similarly, if there is a systematic difference $\sigma_{SB,odd/even}$ between odd and even columns, it only cancels out if there is an equal number of odd and even pixels. Let's assume there are M more odd than even pixels. Then a total offset of $M \times (1/N)\ \sigma_{SB,odd/even}$ is added to the flux value. In the extreme case of all saturated pixels either even or odd, it is the same as σ_{SBO}.

Table 3-1 List of the different kinds of noise sources when calculating the slope s with Equation 1. Typical noise source values are listed in the 2nd column when available.

Noise Source	Values (ADU)	Description
σ_P^2		Poisson noise in f
σ_{RD}^2	7	Readnoise
$\sigma_{SB,stat}^2$	15	Statistical kTC noise in superbias
$\sigma_{SB,sys}^2$		Systematic noise sources in superbias
σ_{SBO}^2		Constant bias offset for full image
$\sigma_{SB,1/f}^2$		Row by row offsets due to incomplete removal of $1/f$ noise
$\sigma_{SB,odd/even}^2$		Odd/even column offset due to incomplete removal

3.2 Data

For the analysis in the following sections, we use simulated images generated using the NIRCam imaging simulator tool on GitHub (Hilbert 2018, in prep), listed in Table 3-2. One of the intermediate products of the image simulator is a count rate image showing simulated sources added, including the overall background. This product allows us to directly compare the pipeline output slope image with the input slopes from the simulator. Additionally, the simulator creates a ZEROFRAME FITS file extension in the output uncalibrated image that can be extracted and separately run through the pipeline for frame 0 analysis (as the pipeline does not yet recognize the ZEROFRAME extension). The first exposure (`nrca1_47Tuc`) is used in the following section to compare the output slopes for the full ramp and frame 0. These are simulations of the 47 Tuc (NGC 104) globular cluster using detector NRCA1 and the RAPID readout pattern. The rest of the exposures (`nrca1_inrate`) in the table are simulated "flat field" exposures with different input count rates generated to increase the number of pixels available for statistical calculations on the slope. These are used in Section 4.

Table 3-2 Exposures described in Section 3.2 that are used in Section 4. They are all generated using the NIRCam image simulator.

File	Description
`nrca1_47Tuc_RAPID_rate.fits`	Simulated RAPID ramp (4 groups) run through the pipeline

File	Description
nrca1_bkg10e-_RAPID_rate.fits nrca1_bkg100e-_RAPID_rate.fits nrca1_bkg190e-_RAPID_rate.fits nrca1_bkg561e-_RAPID_rate.fits nrca1_bkg940e-_RAPID_rate.fits nrca1_bkg1500e-_RAPID_rate.fits nrca1_bkg2246e-_RAPID_rate.fits nrca1_bkg3500e-_RAPID_rate.fits nrca1_bkg4020e-_RAPID_rate.fits nrca1_bkg4501e-_RAPID_rate.fits nrca1_bkg4750e-_RAPID_rate.fits nrca1_bkg5003e-_RAPID_rate.fits nrca1_bkg5200e-_RAPID_rate.fits nrca1_bkg5400e-_RAPID_rate.fits nrca1_bkg5800e-_RAPID_rate.fits nrca1_bkg6000e-_RAPID_rate.fits nrca1_bkg6200e-_RAPID_rate.fits nrca1_bkg6500e-_RAPID_rate.fits nrca1_bkg7000e-_RAPID_rate.fits nrca1_bkg7200e-_RAPID_rate.fits nrca1_bkg7400e-_RAPID_rate.fits nrca1_bkg7800e-_RAPID_rate.fits nrca1_bkg8000e-_RAPID_rate.fits	Simulated flat field RAPID ramps with different input count rates run through the pipeline

4 kTC Noise

The noise due to the thermal generation of electrons within the detector material is called kTC noise. The spatially uncorrelated pattern of kTC noise is different from integration to integration, but the same from group to group within a single integration. Therefore, kTC noise on the order of 15 ADUs (~ 37 electrons) is introduced when a superbias is subtracted. In general, this noise source is not relevant since it is the same group to group; it only affects the y-intercept and not the slope when the ramp is fitted. However, in the case of frame 0 with only one data point, this noise goes directly into the slope measurement itself. As mentioned in Section 3.1, the signal in frame 0 can range from 6,000-50,000 ADUs in a saturated pixel, read out with a DEEP8 readout pattern. For kTC noise in the range of 15 ADUs, the error introduced is small, only up to 0.25%. However, since these are pixels with a lot of signal and therefore very small uncertainties, this would increase the uncertainty by up to a non-negligible 7%, assuming a gain of 2.5 e-/pixel.

5 Bias Offset

In theory, there should be no systematic bias offset $\sigma_{SB,sys}$ left after superbias and reference pixel correction. Here we check if it is true, and what effect it has on the calculated slopes. The notation used in the following analysis is described in Table 5-1. We extracted the ZEROFRAME extension from the simulated 47 Tuc exposure and ran both the full ramp and frame 0 through the pipeline to generate slope images. We then compared the input and output slopes ($\Delta Rate$) for frame 0 and the full ramp, using one of the intermediate products of the simulator. The median of the difference between the output and input slopes for $F0_{orig}$ and I_{orig} is -0.055 and 0.051, respectively, and is consistent with zero. When we artificially introduce an offset σ_{SBO} of 15 ADU (roughly the kTC noise) to the simulated exposures (I_{offset} and $F0_{offset}$), we find that the slopes in $F0_{offset}$ significantly change, while the slopes in I_{offset} do not change (Figure 5-1 and Table 5-2). This shows the importance of having no systematic bias offset in frame 0, something that is of limited importance to the normal ramp fitting.

Table 5-5-1 This table lists the notation used in Section 5.

Symbol	Description
I_{orig}	original full ramp slope image run through pipeline
I_{offset}	slope image from image with artificial 15 ADU offset
I_{corr}	slope image from I_{offset} corrected by the median of the PEDESTAL
$F0_{orig}$	original frame 0 slope image run through pipeline
$F0_{offset}$	frame 0 slope image from image with artificial 15 ADU offset
$F0_{corr}$	frame 0 slope image from $F0_{offset}$ corrected by the median of the PEDESTAL
$\Delta Rate$	pipeline output slope - simulator input slope

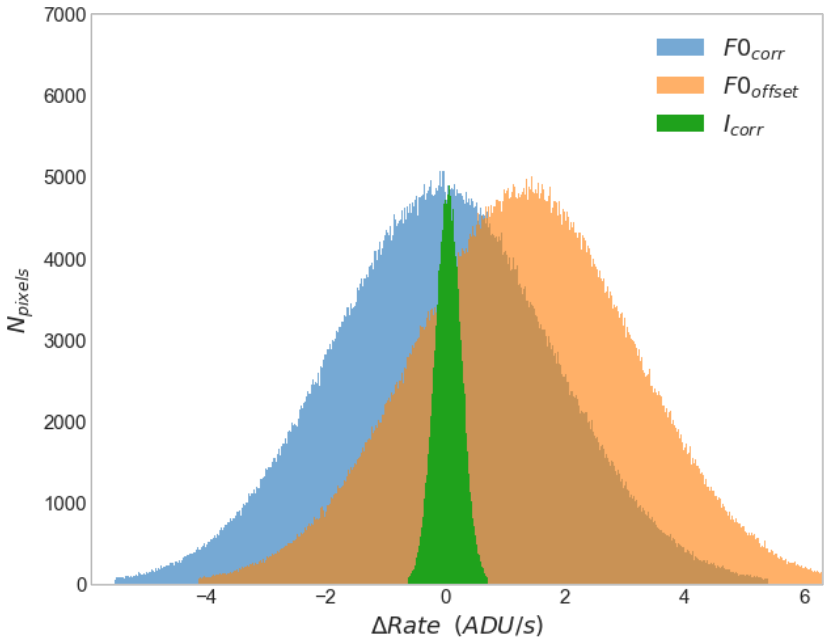


Figure 5-5-1 Image showing the difference between the output slope and the input slope for the $F0_{offset}$ (orange), $F0_{corr}$ (blue), and I_{corr} (green) images. Because it is virtually the same as $F0_{corr}$, $F0_{orig}$ is not shown in this plot.

Table 5-5-2 Differences between output and input slopes for original slope images and images with an artificial bias of 15 ADU added.

Symbol	$\Delta Rate$
I_{orig}	0.05141
I_{offset}	0.05142
$F0_{orig}$	-0.05490
$F0_{offset}$	1.43920

5.1 Pedestal Correction Step

One way to guard against any systematic noise $\sigma_{SB,sys}$ is by analyzing the PEDESTAL extension of the image after ramp-fitting. For our test, we use the image from the previous section into

which we have artificially introduced a bias offset σ_{SBO} (orange histogram in Figure 5-1). We calculate the median PEDESTAL, and find a value of 14 ADU (consistent with the added value of 15 ADU). We subtract this value from the both the regular and frame 0 ramps, and rerun the ramp-fitting pipeline step (I_{corr} and $F0_{corr}$). Figure 5-1 shows the difference between the input and output slopes ($\Delta Rate$) for both the regular and frame 0 image before and after the correction. Not surprisingly, the median of the differences in slopes for the regular image I_{offset} and I_{corr} is unchanged and is consistent with zero, while the median of the differences in slope for frame 0 $F0_{offset}$ goes from an offset of 1.4 ADU/s from zero (different from $F0_{orig}$ by a factor of ~30) back to a value consistent with zero (-0.05 ADU) for $F0_{corr}$.

Most pixels in the 47 Tuc image used for this analysis contain only background signal and are not saturated. Therefore, we look now at the simulated flat field images where we used a range of slope values as inputs to the simulator to create 2048 x 2048 pixels with roughly the same count rates. As we did before, we artificially introduced an offset σ_{SBO} of 15 ADU to these exposures. Using this method, we were able to compare the simulated input rates with the pipeline output rates for most of the pixels on the detector to get a robust measurement of the median $\Delta Rate$.

Figure 5-2 shows the median of the differences in input and output slopes ($\Delta Rate$) of $F0_{corr}$, $F0_{offset}$, I_{corr}, and I_{offset} for bins determined by the flux. To magnify trends in the data, we scaled $\Delta Rate$ by the input rate in Figure 5-3. As expected, for $F0_{corr}$ and I_{corr} $\Delta Rate$ is close to zero. However, $\Delta Rate$ for $F0_{offset}$ is consistently off by ~1.4 ADU/sec due to the introduced bias σ_{SBO}, confirming that for any input rate it will be important to correct for systematic biases that are present in an exposure after the reference pixel correction. For I_{offset}, σ_{SBO} does not have a significant effect on $\Delta Rate$ for input rates $\lesssim$ 2000 ADU/sec; however, as the rate increases more pixels become saturated in the second group and only the first group is used to calculate the slope. This is discussed further in Section 5.2.

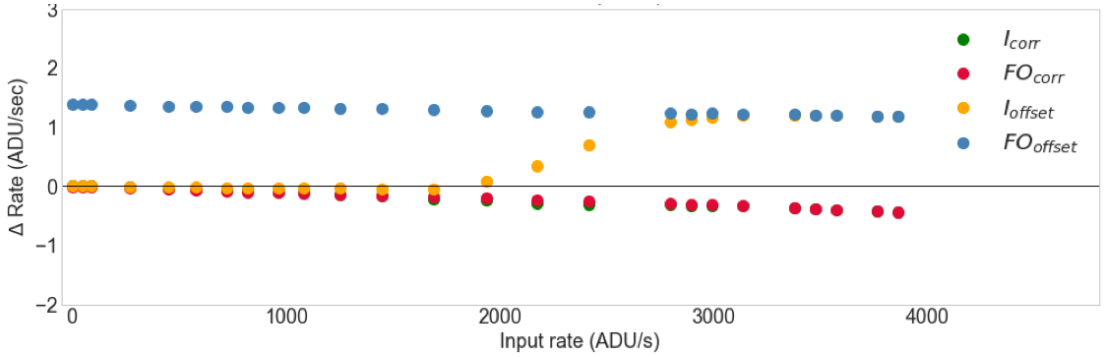


Figure 5-5-2 Plot showing the flux dependence of $\Delta Rate$ for I_{offset}r, $F0_{offset}$, I_{corr}, and $F0_{corr}$. The slight negative trend of the values is due to a small 0.01% offset in the $\Delta Rate$ values shown in Figure 5-3.

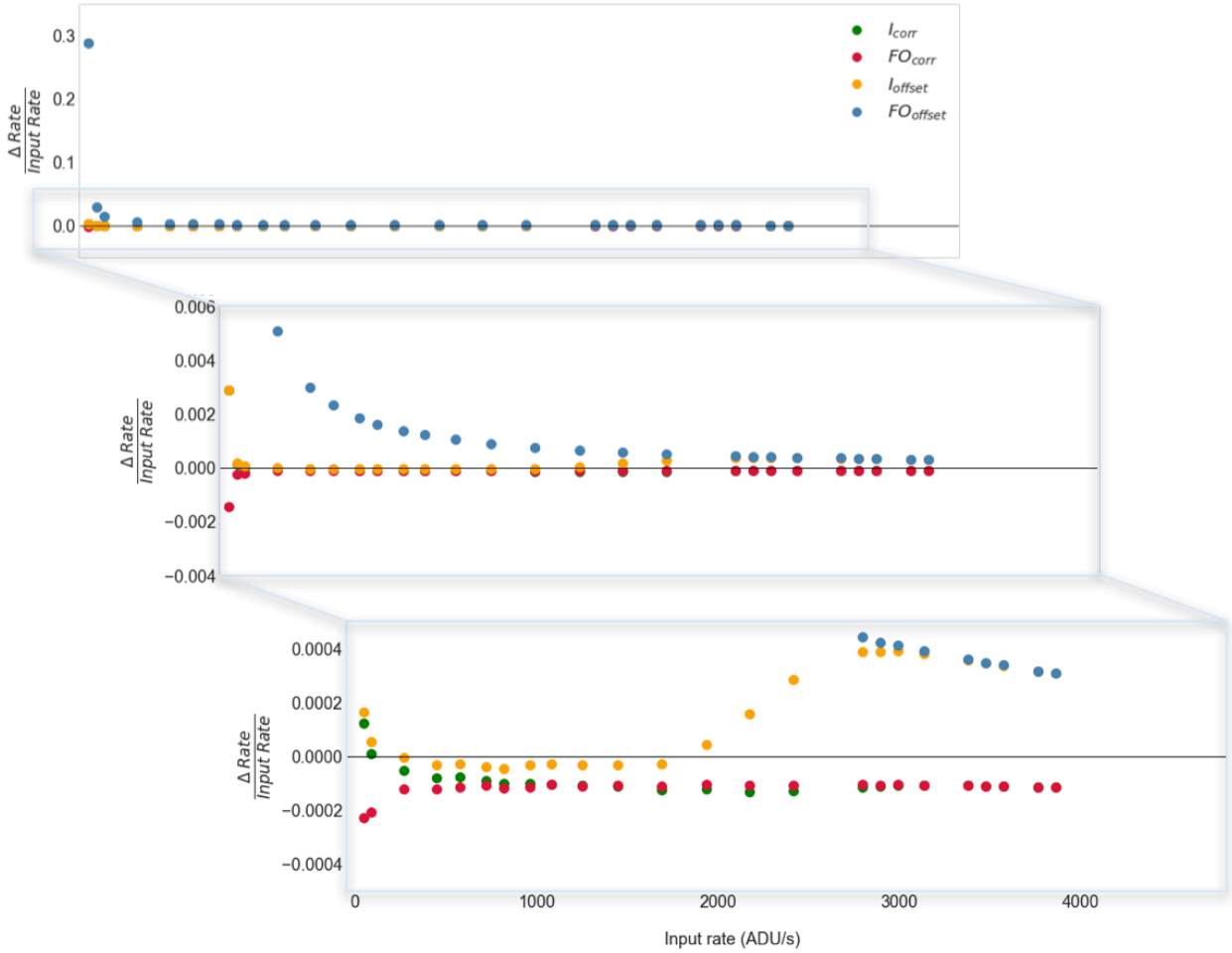


Figure 5-5-3 Plot of $\Delta Rate$ scaled by the input rate for I_{offset}, $F0_{offset}$, I_{corr}, and $F0_{corr}$. The insets are zoomed-in slices of the top panel. The small 0.01% offset in the bottom panel is discussed briefly in Section 5.1. The increase in $\Delta Rate$ for I_{offset} is discussed in Section 5.2.

The scattering in $\Delta Rate$ that is evident for input rates $\lesssim$ 200 ADU/sec in Figure 5.3 is due to noise sources in the input dark exposures used by the simulator to construct the images. We verified that using different dark exposures from the same test campaign show similar positive or negative scattering behavior for low input rates. There is also a roughly consistent 0.01% offset in $\Delta Rate$ for all input rates, which is still under investigation (though it is likely related to the simulator or the reference files used by the simulator to construct the exposures).

5.2 Saturation in 2nd group

For pixels that saturate in the second group, then in the current pipeline only the first group is used to calculate the slope. It therefore is the same case as the slope calculation for frame 0, i.e. it is not a linear fit, but just a simple rate calculation. Thus, it has the same additional noise sources $\sigma_{SB,stat}$ and $\sigma_{SB,sys}$ as frame 0. This is apparent for I_{offset} in Figures 5-2 and 5-3, as $\Delta Rate$ for input rates $\gtrsim$ 2000 ADU/sec becomes significantly larger than 0 due to an increasing number of saturated pixels. Figure 5-4 shows the number of saturated pixels in each group for different input rates on the left axis and $\Delta Rate$ for I_{offset} on the right axis. As more pixels become saturated in the second group, the output slope for I_{offset} is determined instead by the first

group, which makes the introduced bias offset σ_{SBO} increasingly more significant. For our RAPID ramps, the first group of I_{offset} is equivalent to $F0_{offset}$, so $\Delta Rate$ is the same for I_{offset} and $F0_{offset}$ when the rate is high enough that all pixels saturate in the second group.

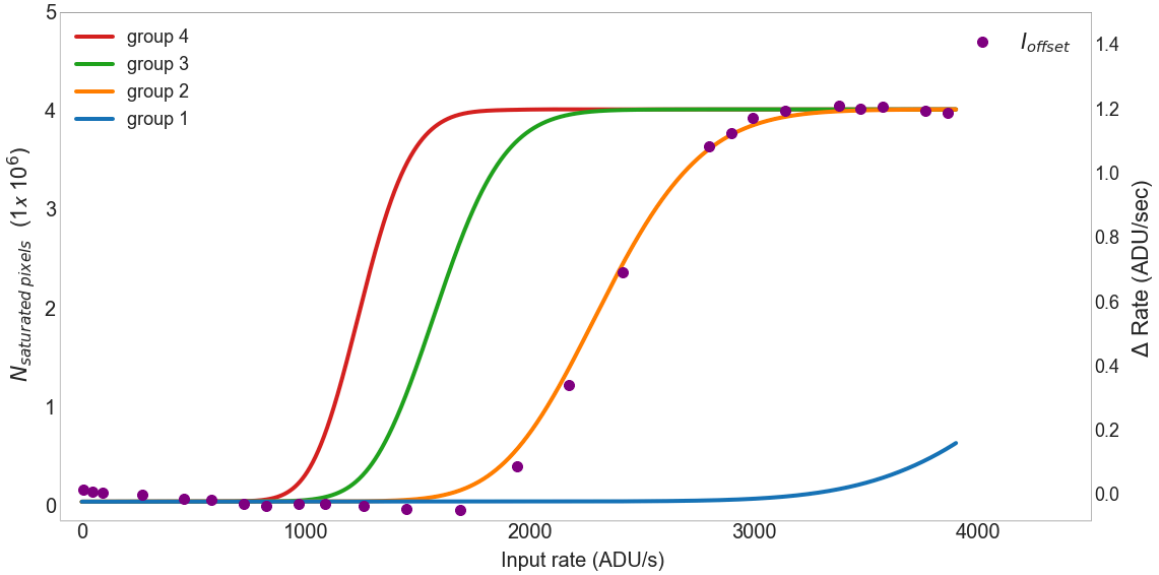


Figure 5-5-4 The number of saturated pixels in each group for different input rates (left axis), plotted simultaneously with $\Delta Rate$ for I_{offset} (right axis). As the input rate increases, more pixels become saturated and the output slope for I_{offset} is determined using the first group.

Additionally, there is no way to know if a cosmic ray (CR) hit within the first group (and after the first frame), but in such cases frame 0 may still be trusted. We therefore suggest using frame 0 in combination with the first group to calculate the slope. This not only gets rid of the σ_{SB} noise sources, but it also allows a method to identify pixels that seem to be hit by a cosmic ray: the pixels should be flagged if the fitted pedestal is different from zero by more than 3 times the kTC noise $\sigma_{SB,stat}$.

6 Ramp Fitting with Frame 0

We want to evaluate whether or not it is advantageous to add frame 0 to the ramp fitting in general, rather than only using it in the special cases mentioned in the previous sections.

6.1 Noise

The effect of including frame 0 in the ramp fitting is most likely largest in readnoise-limited or background-limited cases (e.g., exposures with narrow filter bands, spectroscopy, or deep exposures in sparse fields). For example, some GTO/ERS proposals use the DEEP8 readout pattern with only 3 or 4 groups.

Estimating the noise in NIRCam ramps is complicated, since the fluxes in the groups are correlated. Therefore, for simplicity, we consider how the noise would change when calculating the mean count in a dark frame (which should of course be zero). We can calculate the mean in two ways: a straight mean $\overline{M}$, or a weighted mean $\overline{WM}$, where the weight is the inverse variance. In addition, we can include frame 0 in the calculation. We estimate the uncertainties for these calculations using the usual error propagation as

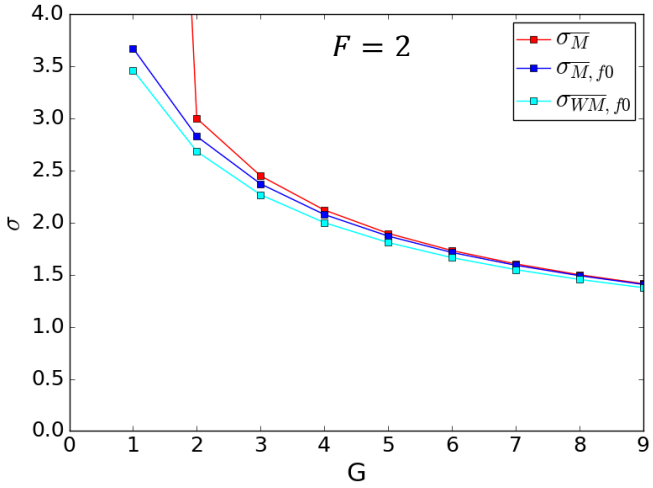
$$\sigma_{\bar{M}}^2 = \begin{cases} \sigma_{kTC}^2 + \frac{1}{GF}\sigma_{RD}^2 & for\ G = 1 \\ \frac{1}{GF}\sigma_{RD}^2 & for\ G > 1 \end{cases} \quad (7)$$

$$\sigma_{\bar{M},f0}^2 = \frac{1}{(G+1)^2}\left(\sigma_{RD}^2 + \frac{G}{F}\sigma_{RD}^2\right) \quad (8)$$

$$\sigma_{\overline{WM}}^2 = \frac{1}{GF\frac{1}{\sigma_{RD}^2}} = \frac{1}{GF}\sigma_{RD}^2 \quad (9)$$

$$\sigma_{\overline{WM},f0}^2 = \frac{1}{\frac{1}{\sigma_{RD}^2} + GF\frac{1}{\sigma_{RD}^2}} = \frac{1}{GF+1}\sigma_{RD}^2 \quad (10)$$

where G is the number of groups, and F is the number of frames per group. Ramp fitting in the readnoise-dominated regime is essentially the same as taking the average, but with one additional free parameter. We therefore can assume that the noise properties and propagation are very similar for both the mean and the linear fit in these circumstances, and we can use this to determine how important it is to include frame 0 in the ramp fitting. Figure 6-1 shows how the noise depends on G for different frames per group F.



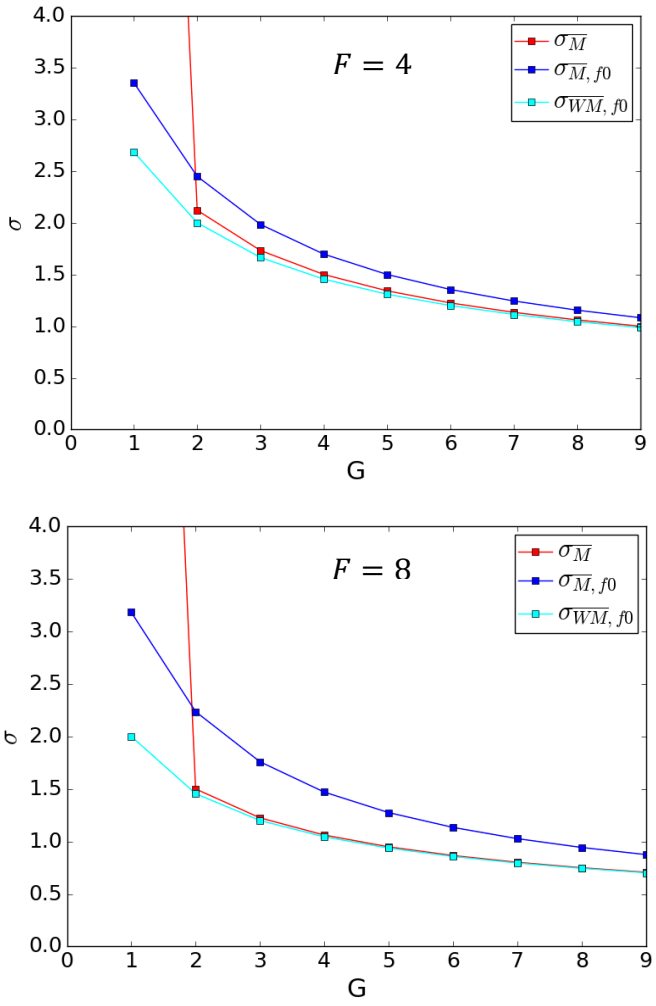


Figure 6-1 Plots showing noise σ versus number of groups G for 2 frames per group F (top), 4 frames per group F (middle), and 8 frames per group F (bottom).

We note the following:

- Due to the kTC noise, $\sigma_{\overline{M}}^2$ is very large for $G = 1$. This can happen in two cases: (1) A pixel saturates in the second group. This, however, is not a readnoise-dominated case, but rather the kTC noise is relatively small compared to the Poisson noise of the flux (see previous sections). (2) A cosmic ray hits the second group. In this case the true uncertainty of the 'slope' calculated with only the first group is dominated by the kTC noise, which is several factors higher than the readnoise!
- Including frame 0 in the average $\overline{M}$ that is not weighted by the variance increases the noise.
- Including frame 0 in the weighted average $\overline{WM}$ decreases the noise, but only by a few percent.

6.2 Cosmic Ray Detection

It is expected that a large fraction of the pixels will have cosmic ray hits in deep exposures. Therefore, cosmic ray rejection and correction is an important part of the reduction. One of the biggest challenges is to determine which groups are affected by a cosmic ray, in particular when

the total number of groups is small. If frame 0 is included in the ramp fitting, it effectively increases the number of groups by one with respect to CR detection. We suspect that including frame 0 will significantly improve cosmic ray detection and correction for integrations with less than 6 groups. This needs to be tested as soon as the optimal cosmic ray rejection threshold in the jump detection step of the pipeline has been determined for NIRCam.

7 Implementation Recommendations

Following the analysis described in the previous sections, we suggest the following implementations into the vanilla version of the JWST calibration pipeline:

- For all pixels that are saturated in the first group but not saturated in frame 0, the slopes derived from frame 0 should be used in order to significantly increase the dynamic range of the image.
- The accuracy of the slopes determined from frame 0 directly depends on the superbias correction: if there is any unanticipated issue with the superbias correction (e.g., a non-subtracted offset of 50 or 100 ADU), it will introduce a large systematic bias into the measurements. Therefore, we recommend keeping track of the median and standard deviation of the PEDESTAL extension. If there are any significant deviations from the usual, expected values, then the frame 0 measurements (or measurements solely based on the first group) cannot be trusted.

We suggest the following implementations into the optimal pipeline:

- One of the main goals for the optimal pipeline is the correct propagation of uncertainties. Therefore, for all pixels for which only 1 single value is used to calculate the 'slope', the kTC noise should be added to the uncertainties. This is particularly important when the noise in the pixel is readnoise-dominated; for example, when the second group is hit by a cosmic ray.
- Frame 0 should be added to the ramp fitting in general, or at least in all cases for which only a small number of groups are available for the linear fit, such as an early cosmic ray hit. This will mitigate the effect of the kTC noise, and help identify cosmic ray hits between frame 0 and the last frame in first group.
- A correction could be applied if the median of the PEDESTAL extension is significantly different to zero, as described in section 5.1

8 References

Hilbert, B., 2018, *The NIRCam Data Simulator Handbook,* STScI-006219