



JWST Noise Floor. II. Systematic Error Sources in JWST NIRCам Time Series

Everett Schlawin¹, Jarron Leisenring¹, Michael W. McElwain², Karl Misselt¹, Kenneth Don¹, Thomas P. Greene³,
Thomas Beatty¹, Nikolay Nikolov⁴, Douglas Kelly¹, and Marcia Rieke¹

¹Steward Observatory 933 North Cherry Avenue, Tucson, AZ 85721, USA

²NASA Goddard Space Flight Center Exoplanets and Stellar Astrophysics Laboratory, Greenbelt, MD 20771, USA

³NASA Ames Research Center Space Science and Astrobiology Division, Moffett Field, CA 94035, USA

⁴Space Telescope Science Institute 3700 San Martin Drive, Baltimore, MD 21218, USA

Received 2020 October 7; revised 2020 December 8; accepted 2021 January 4; published 2021 February 11

Abstract

The James Webb Space Telescope (JWST) holds great promise for characterizing atmospheres of transiting exoplanets, potentially providing insights into Earth-sized planets within the habitable zones of M-dwarf host stars if photon-limited performance can be achieved. Here, we discuss the systematic error sources that are expected to be present in grism time-series observations with the NIRCам instrument. We find that pointing jitter and high-gain antenna moves in addition to the detectors' subpixel crosshatch patterns will produce relatively small variations (less than 6 parts per million, ppm). The time-dependent aperture losses due to thermal instabilities in the optics can also be kept to below 2 ppm. To achieve these low noise values, it is important to employ a sufficiently large (more than $1''.1$) extraction aperture. Persistence due to charge-trapping will have a minor (less than 3 ppm) effect on the time series 20 minutes into an exposure and is expected to play a much smaller role than it does for the Hubble Space Telescope WFC3 detectors. We expect detector temperature fluctuations to contribute less than 3 ppm. In total, our estimated noise floor from known systematic error sources is only 9 ppm per visit. We urge caution, however, because unknown systematic error sources could be present in flight and will only be measurable on astrophysical sources such as quiescent stars. We find that reciprocity failure may introduce a perennial instrument offset at the 40 ppm level, so corrections may be needed when a multi-instrument multi-observatory spectrum is stitched together over wide wavelength ranges.

Unified Astronomy Thesaurus concepts: Exoplanet atmospheres (487); Exoplanet atmospheric composition (2021); Red noise (1956); Astronomical detectors (84); Space telescopes (1547); Pointing error (1272)

Supporting material: data behind figure

1. Introduction

The James Webb Space Telescope (JWST) will be a powerful new tool for characterizing exoplanets, especially in transiting systems. The transmission and emission spectra of planets will permit new avenues for measuring the composition, temperature profiles, and maps in a variety of planets (e.g., Beichman et al. 2014; Barstow et al. 2015; Greene et al. 2016b; Howe et al. 2017; Schlawin et al. 2018).

It is critical, however, that the JWST perform close to its photon noise limit in order to fully characterize exoplanet atmospheres with small scale heights relative to the radii of their host stars. This is especially true if the JWST is to be used to search for biosignatures in habitable-zone terrestrial planets, such as the TRAPPIST-1 system (e.g., Barstow & Irwin 2016; Krissansen-Totton et al. 2018; Pidhorodetska et al. 2020). While there would be enough photons to detect an ozone feature in 30 transits of TRAPPIST-1 d (Barstow & Irwin 2016), there could be systematic errors that complicate the measurement. Systematic errors can come from astrophysical sources such as unocculted starspots (e.g., Rackham et al. 2018). They can also occur due to observatory and instrument artifacts from the JWST itself.

We focus here on time-series applications of the JWST NIRCам, which also have applicable lessons for other near-infrared instruments that use HgCdTe detectors on the JWST or ground-based observatories. NIRCам is the Near-infrared Camera on JWST, capable of imaging two $2' \times 2'$ fields of view simultaneously with a short-wavelength filter (between 0.6 and $2.3 \mu\text{m}$) and a long-wavelength (LW) filter (between 2.4 and

$5.0 \mu\text{m}$; Rieke et al. 2005). NIRCам has a grism pupil element that can be used on the LW channel for spectroscopy (Greene et al. 2016a, 2017). In the grism time-series mode, the short-wavelength channel can be used with weak-lens imaging or with a dispersed Hartmann sensor, which is discussed as a science enhancement (Schlawin et al. 2017).

The time-series observation modes of NIRCам can capture changes in the flux of science target(s) to capture astrophysical variable phenomena, such as planets that eclipse or transit their host stars, rotational modulations, or pulsation modes in stars. The time-series modes point the telescope as stably as possible at the target(s) without dithering, repointing, or moving mechanisms in order to measure the flux as precisely as possible. This is in contrast to other standard JWST modes that dither (move) the telescope to mask bad pixels, reduce pixel-sampling errors, and reduce flat-field errors. With time-series modes, the goal is not absolute flux accuracy or unpixelated images, but rather precision measurements of flux changes. The resulting data from time-series modes consist of a series of integrations of the same field where it is possible to study time variations in the pixels.

NIRCам has four short-wavelength detectors and one LW detector over each of its $2' \times 2'$ fields of view, totaling 10 detectors in all. The fields of view are split into the A and B side and are numbered A1 through A4 and B1 through B4 for the short-wavelength channels and A5 and B5 for the LW channels. Time-series imaging uses 5 of the detectors on the B side for high-precision photometry. The grism time series uses the grism pupil element to disperse the spectrum at a resolution of $\sim 1100\text{--}1700$ (Greene et al. 2017) on the A5 detector while

imaging simultaneously with the A1 through A4 detectors. We expect the grism time series to be the most frequently employed time-series mode on NIRC*am* and the science to come from the spectrum, so we default most calculations in this paper to the A5 detector.

Suissa et al. (2020) find that the JWST noise floor is a critical factor in detecting H₂O features of Earth-sized ocean-bearing worlds around M dwarfs. For the Transiting Exoplanet Survey Satellite (TESS) input catalog, it is not possible to detect water vapor in water-bearing worlds if the absolute noise floor exceeds 5 ppm. However, if the noise floor is 3 ppm, there are about eight exoplanet systems for which the JWST could detect 2.7 μ m water-vapor features in less than 100 hr of accumulated exposure time. Thus, the noise floor is a critical parameter in future searches for habitable conditions and the presence of atmospheres on Earth-sized worlds.

In Schlawin et al. (2020), hereafter Paper I, we discussed the random errors that increase the noise of JWST NIRC*am* time-series observations above the photon noise. Here, we discuss the systematic errors that can increase the noise in JWST time series, again with a focus on the NIRC*am* instrument. We summarize important results from Paper I at the end of Section 2. The NIRS*pec* and NIRS*ss* time-series modes may have similar systematic errors because they have the same types of detectors and similar telescope-pointing jitter. However, there will be some differences because NIRS*pec* bright-object time series uses a slit (Ferruit et al. 2014) and NIRS*ss* single-object slitless spectroscopy has a unique cylindrical lens to spread the light in the cross-dispersion direction (Doyon et al. 2012; Albert et al. 2014).

While photon noise and the random errors discussed in Paper I will fall as $1/\sqrt{N}$ for N integrations, systematic errors may fall more slowly with N . If the systematic errors cause correlations between integrations, they will fall less rapidly than $1/\sqrt{N}$ or may not decrease at all with N if integrations are perfectly correlated. Systematic errors can introduce nonastrophysical time-dependent behavior in JWST light curves or produce a deterministic offset between the true and the measured signal. Often, systematics can be functions of other variables such as the telescope pointing, wave-front errors in the JWST optics, the temperature of a detector, or the charge-trap state in the detector.

Section 2 lists the known systematic effects that can degrade the precision of time-series measurements compared to the photon noise limit. Sections 3 through 7 provide simulations and analysis of these effects to provide some quantitative estimates of how much they might contribute to JWST NIRC*am* time series. We conclude in Section 8 that the known systematic errors we study contribute ~ 9 ppm or smaller effects, but more tests in flight will be necessary to verify performance and develop mitigation strategies.

2. List of Known Systematic Effects

In Paper I we described the random effects in NIRC*am* time series that increase the scatter in light curves. In addition to these random effects, systematic effects can also produce correlated noise between integrations due to other variables such as temperature and telescope pointing. Here, we consider systematic effects that can introduce longer timescale variations and impact the transit or eclipse depths of planets.

1. **Intrapixel sensitivity:** Telescope-pointing drifts and jitter can impact the amount of flux measured if the response within pixels is not perfectly uniform. This has a particularly strong impact on the time series from the Spitzer IRAC instrument (Ingalls et al. 2016). It can also appear on JWST detectors because the subpixel flat field is not uniform. Crosshatching patterns on HgCdTe detectors, which all JWST near-infrared instruments use, exist at the subpixel level (Shapiro et al. 2018; Ninan et al. 2019). The intrapixel sensitivity is more pronounced when the point-spread function (PSF) is undersampled and there are pointing variations, such as with the Spitzer IRAC imager.
2. **Variable aperture losses:** A photometric or spectroscopic light curve will vary in time if the amount of flux from a star within its assigned extraction aperture varies. The JWST is expected to have wave-front error variations due to slight motions of the mirror surfaces with thermal expansion and contraction that result in subtle changes in the JWST PSF. If the amount of flux contained within an extraction aperture (sometimes referred to as encircled energy) varies on hour-long timescales across a transit or eclipse, this can cause spurious variations in planet transit or eclipse-depth measurements.
3. **Detector temperature fluctuations:** The JWST detectors are sensitive to temperature changes on the focal plane arrays. For example, laboratory tests show that 100 mK temperature fluctuations can result in $\sim 80e^-$ changes in the detector bias offsets that are not corrected for by reference pixels (Hall 2006). NIRC*am* detectors are actively thermally controlled to keep temperature fluctuations to $\lesssim 1$ mK. However, during cryovac ground-testing, the LW A detector (A5) exhibited ~ 20 -minute oscillations with an amplitude of ~ 15 – 20 mK when the primary temperature monitor control (TMC) was used. Switching to the redundant TMC, an independent heater and Cernox temperature sensor for A5 showed no observable oscillations; this configuration will be employed in orbit. If necessary, the effect of temperature fluctuations can be calibrated to $\pm 1e^-$ for excursions less than ± 50 mK (Hall 2006).
4. **Charge-trapping:** A major correlated error source on the Hubble Space Telescope's Wide Field Camera 3 (HST WFC3) is the ramp effect (Berta et al. 2012). The ramp effect is due to charge trapped within the detector's depletion region, initially lowering the measured flux, and then released after detector reset, thereby increasing the measured flux most prominently at the beginning of an integration (Zhou et al. 2017). The JWST NIRC*am* instrument has very low rates of persistence ($< 1 \text{ DN s}^{-1}$) at 100 s after saturation on all detectors (Leisenring et al. 2016), and the JWST will observe continuously without the Earth occultations experienced by the HST, so it is expected that charge-trapping will be a minor concern for the JWST HgCdTe detectors. This allows pixels to reach a steady state, assuming that the pointing jitter is within a pixel, as expected for the JWST.
5. **Reciprocity failure:** The response of a HgCdTe detector can change with the brightness of a source, which is called reciprocity failure (Biesiadzinski et al. 2011). The brightness of a star and planet system does not change significantly with time, so it will not significantly affect the precision of a light curve, but it might affect accuracy

because the filter bandpass and stellar spectral energy distribution change the count rates with wavelength. It is possible that reciprocity failure could introduce a systematic offset, slope, or curvature that depends on the count rate of the detector.

6. Ghost images and contaminating light: There are 17 optical surfaces in the NIRCcam LW channel that can create spurious reflections or ghost images. Ground-based testing revealed a ring-like ghost on grism images with up to 1.5% of the value as the average flux in a grism aperture. This is a second-order effect because it contaminates the measured flux by 1.5% of the difference in transit depth. So, for a 100 ppm variation in transit depth with wavelength, we expect about 1.5 ppm differences from the peak part of ghost images. This would be present in all images, so it could produce a systematic offset but not a time-variable signature.

The systematic errors discussed here should be compared with the random error sources discussed in Paper I. We found in Paper I that detector-noise sources can exceed the photon noise for short integrations and thus can affect science on bright targets. The main challenge for reducing the detector noise is that it is highly correlated along pixels in the fast-read direction parallel to the spectrum. When the flux from many pixels is averaged in wavelength, the read noise stays constant, rather than falling as $\sqrt{N_{\text{px}}}$, where N_{px} is the number of pixels, as would occur if each pixel were independent. For the shortest integrations where there are only two read samples, the $1/f$ detector noise can exceed 830 ppm in bins that are $0.17 \mu\text{m}$ wide and 14 spatial-pixels tall, compared to 390 ppm for the photon noise.

Fortunately, the detector read noise can be reduced by adjusting the observing and extraction strategy. Operationally, a GRISM mode is being discussed that disperses the spectrum perpendicular to the $1/f$ noise correlations, permitting them to be dramatically reduced with background subtraction. The read noise can be mitigated by using a very narrow (3 pixel or $0''.19$) aperture or a covariance weighting scheme that takes into account pixel noise correlations. The covariance weighting scheme can reduce the contribution due to read noise to 230 ppm per read pair sample. This very narrow aperture, however, could be more susceptible to systematic effect 1 (intrapixel sensitivity) and systematic effect 2 (variable aperture losses) discussed above. If many groups are averaged together, this will drop the read noise by a factor of $\sqrt{N_{\text{read}}}$, where N_{read} is the number of nondestructive reads. For a typical ~ 1 hr transit duration, a subarray size of 2048×256 , and five groups, there are a total of 2290 read samples per pixel. This will reduce the 230 ppm random noise for two read samples by $\sqrt{2290/2}$ or 7 ppm in the transit depth. We therefore are interested in effects that are of order 7 ppm. For bright targets that use even smaller subarrays such as the 2048×64 subarray, we are interested in systematic errors that are a factor of about $\sqrt{4}$ smaller or 3.5 ppm.

3. Subpixel Crosshatching

The NIRCcam detectors, most notably the A5 detector used in grism time series, have a crosshatching pattern in their flat-field profile. Figure 1 shows a zoom-in of the A5 detector flat field calibration image from Cryogenic Vacuum Testing 3 (CV3) at NASA Goddard. The crosshatch pattern extends to the subpixel

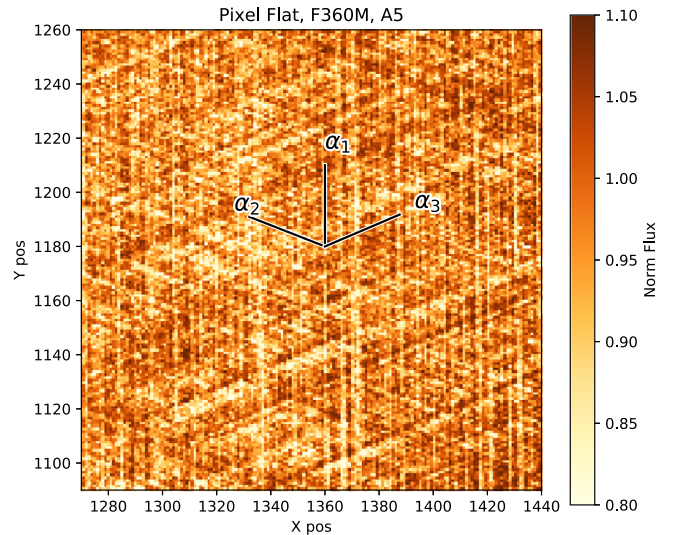


Figure 1. The A5 detector flat field has a pronounced crosshatch pattern at $\alpha_1 = 90^\circ$, $\alpha_2 = 158^\circ$ and $\alpha_3 = 23^\circ$ CCW from the positive X direction. The crosshatch pattern is visible in a zoom-in of the pixel flat field and exists at the subpixel level as well. When the image of a star moves on top of this crosshatch pattern, it will create flux variations.

level and thus can cause flux variations when the centroid of an image or spectrum moves, even on a subpixel level. In Appendix A, we characterize the crosshatch pattern and its connection to the crystallographic structure of the HgCdTe detector material in more detail. Observers interested in the direct consequences of subpixel effects, rather than the details of the calculation, can skip to Section 3.4.

We also briefly compare the amplitude of the crosshatch signal in our A5 detector with published measurements of the crosshatch pattern and intrapixel sensitivity for other H2RG detectors. Shapiro et al. (2018) examine the amplitude of the crosshatch pattern in a candidate detector for the Euclid mission that has particularly strong crosshatch features. The amplitude of variation in the flat-field signal for the candidate Euclid detector is about 1.5–2 times the level of the NIRCcam A5 detector. We also note that Hardy et al. (2008, 2014) measured the intrapixel response of a detector produced for the fine-guidance sensor of the JWST. Hardy et al. (2014) find that on average, there is no sensitivity across a pixel because the flux is distributed to neighboring pixels. So long as one uses a large extraction aperture to include neighboring pixels, there is no average intrapixel response function affecting all pixels. However, there are defects or large drops in the intrapixel map with amplitudes of $\sim 10\%$ that could result in photometric variations. The map in Hardy et al. (2008) of earlier-generation JWST detectors shows long linear features that may be related to crosshatch patterns. The map in Hardy et al. (2014) does not show a strong crosshatch pattern, but the amplitude of the defects is similar to the crosshatch pattern dips we measure from the NIRCcam A5 flat field. We therefore expect that our analysis of the crosshatch pattern will give similar amplitude effects as the Hardy et al. (2014) measurements. On-flight measurements of the subpixel sensitivity of NIRCcam will reveal if there similar defect structures at the location of the grism time-series field point.

3.1. Crosshatch Pattern Modeling

We model the crosshatch pattern in frequency space because it has a smoother pattern as a function of spatial frequency than

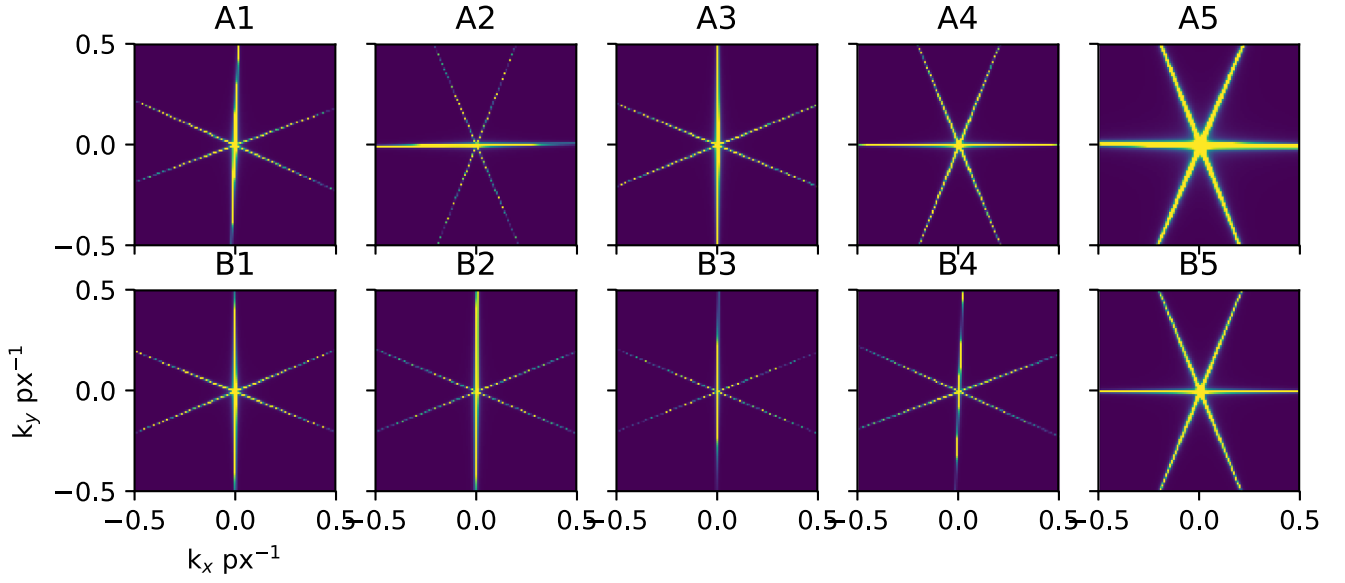


Figure 2. Best-fit power spectra of all the NIRC detectors show that the LW detectors A5 and B5 show the most pronounced crosshatch patterns. The primary crosshatch tends to be aligned closely but not exactly with either the X- or Y-axis. The images have all been normalized to the same scale to show the relative crosshatch amplitudes of different detectors.

in spatial coordinates, as seen in Figure 16. We divide the 2D Fourier power spectral density by the number of pixels in an input image. This was experimentally determined to give a power spectral density that does not change with the dimensions of the input image.

We model the 2D Fourier power spectral density $f(k_{x,i}', k_{y,i}')$ for one angle θ_i as

$$f(k_{x,i}', k_{y,i}', a_i, b, c) = a_i \exp(-|k_{x,i}'|/b) \left(\frac{1}{(0.5c)^2 + k_{y,i}'^2} \right), \quad (1)$$

where k_x' and k_y' are rotated frequency coordinates in the parallel and perpendicular directions, a is the amplitude, b is the parallel exponential constant, and c is the Lorentzian full width at half maximum of the perpendicular dependence. For each of the three angles, there is a set of rotated coordinates centered at $(k_{x,0}, k_{y,0}) = (0,0)$ following

$$k_x'(\theta_i) = k_x \cos \theta_i + k_y \sin \theta_i \quad (2)$$

and

$$k_y'(\theta_i) = -k_x \sin \theta_i + k_y \cos \theta_i. \quad (3)$$

We include three angles so that the total power spectral density is

$$f(k_x, k_y) = \sum_{i=1}^3 f_i(k_{x,i}'(\theta_i), k_{y,i}'(\theta_i), a_i, b, c) + d \exp(-k_r/e_b). \quad (4)$$

This model has eight free parameters: three angles (θ_i), three amplitudes (a_i), a joint parallel exponential constant (b), and a joint Lorentzian width (c). Finally, we include an exponential term that fits the broad azimuthally symmetric background to all the power spectra, where d is the amplitude of the radial “background,” $k_r = \sqrt{k_x^2 + k_y^2}$ is the radial frequency, and e_b is the radial background exponential constant. We only fit the

region of the power spectral density above frequencies of 0.05 px^{-1} to focus on the high-frequency component of the flat field where the crosshatch is most prevalent. The lower spatial frequencies describe broader structures such as epoxy voids or illumination gradients that are less relevant for time-series observations.

For the F300M filter and A5 detector, we find $\theta_1 = 0^\circ 906 \pm 0^\circ 001$, $\theta_2 = 68^\circ 605 \pm 0^\circ 001$, $\theta_3 = 113.075 \pm 0.001$, $a_1 = 2.774 \pm 0.002 \times 10^{-7}$, $a_2 = 1.569 \pm 0.002 \times 10^{-7}$, $a_3 = 1.744 \pm 0.002 \times 10^{-7}$, $b = 0.345 \pm 0.0005 \text{ px}^{-1}$, $c = 2.05 \pm 0.07 \times 10^{-2} \text{ px}^{-1}$, $d = 6.78 \pm 0.0110^{-3}$, and $e_b = 0.26 \pm 0.01 \text{ px}^{-1}$. In other words, the separations between the vectors are $67^\circ 699$, $67^\circ 831$, and $44^\circ 470$. The uncertainties were simply derived from the diagonals of the covariance matrix and may be underestimated.

The NIRC 10 sensor chip assemblies (SCAs) have very similar crosshatch angles with average separations of $44^\circ 5$ and $67^\circ 7$ – $67^\circ 8$, as seen in Figure 2. While the absolute orientations relative to pixel axes change, the relative angles of the three crosshatch axes are within $0^\circ 15$ from the average. They all have one primary crosshatch axis oriented either nearly along the X pixels or the Y pixels, with two flanking axes at $\sim 67^\circ 7$ to either side. The LW arrays have a more pronounced crosshatch than the short-wavelength arrays. The NIRC 10 B5 array primary crosshatch direction is most closely aligned with a pixel axis (the Y direction in the Fourier domain) of any SCA. Figure 2 shows the best-fit crosshatch 2D power spectra.

3.2. Position of the SUBGRISM Array

The crosshatch pattern varies in amplitude and angular dependence from one physical location on the detector to the next. The A5 detector has more pronounced crosshatching toward the middle of the detector that falls off toward the perimeter. On the other hand, the B5 detector has more pronounced crosshatching toward the boundary. We fit the crosshatch pattern to three regions of the A5 detector, corresponding to those used for the NIRC 10 grism time-series mode as well as another considered position:

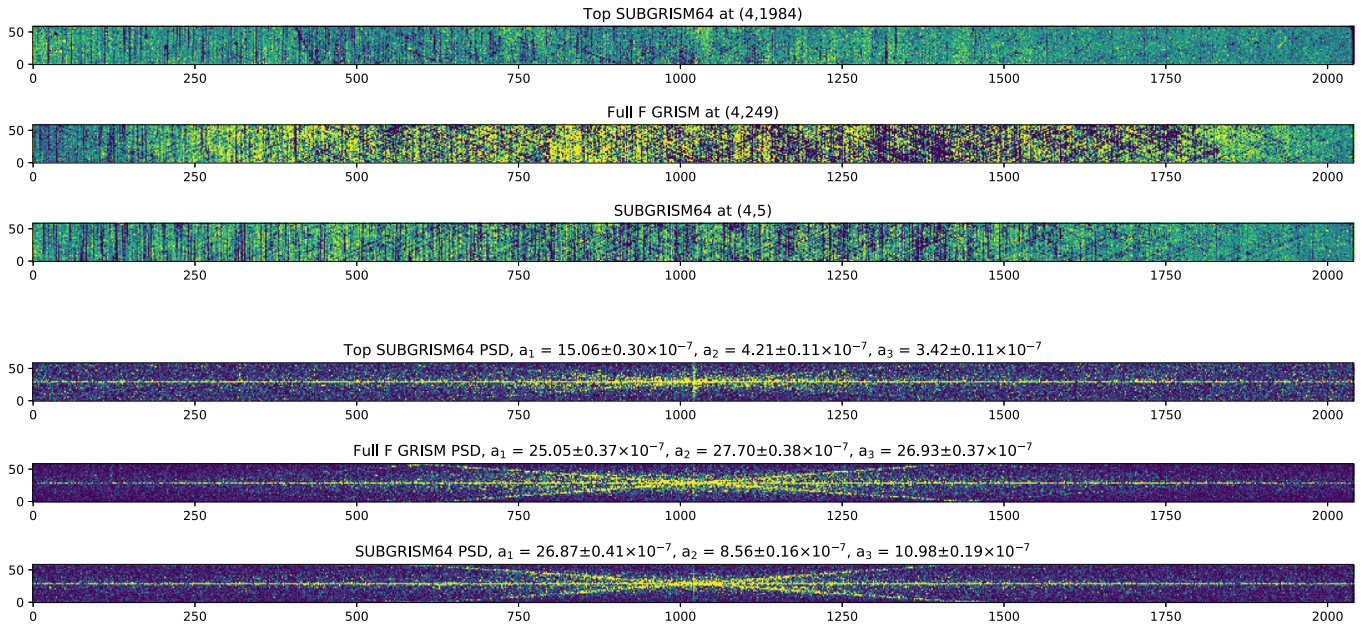


Figure 3. The top of the A5 array has the lowest level of crosshatching and the middle has the highest. This is visible in the flat-field images (top three panels) and their power spectral density functions (bottom three panels) for three different sections of the A5 detector: a theoretical “aperture” position at the top of the array (first plot from the top), the full-frame position near $Y = 512$ (second plot from the top) and the nominal SUBGRISM64 at the bottom of the array (third plot from the top). The bottom left corner for each subarray is listed in parentheses in the title of the flat-field plot, and the best-fit amplitudes are listed in the title of the power spectral density plot.

1. The 2040×64 SUBGRISM position at the bottom of the array where the bottom left corner is (4, 5), which excludes an additional illuminated row (see below).
2. A 2040×64 theoretical subarray at the top of the A5 detector where the bottom left corner is (4, 1984).
3. A 2040×64 cutout of the full-frame detector that is centered on the grism time-series field point where the bottom left corner is (4, 249).

These three regions are depicted in Appendix B. The SUBGRISM256 and SUBGRISM128 modes will place the target at the same location as the 2048×64 SUBGRISM64 position, so results would apply to all three subarrays.

The three regions considered are shown in Figure 3. The corner coordinates of these regions are specified in raw detector pixels. We exclude reference pixels (which form a 4-pixel-wide boundary around the detector) and an additional row at the bottom and top of the array because these illuminated rows are outliers in the flat field. The first position is nearly the same as the SUBGRISM64 subarray position in the grism time-series mode, except that the SUBGRISM64 includes the first five rows, which are excluded here from the crosshatch analysis. The field point used for the SUBGRISM64 mode is expected to be the same as the field point that will be used for the SUBGRISM128 and SUBGRISM256 subarrays. This field point is $X = 468$, $Y = 35$ for the F322W2 filter and $X = 1097$, $Y = 35$ for the F444W filter. The position at the top of the A5 detector has the smallest amplitude of the crosshatch pattern in all three directions. The best-fit amplitudes are used as input to the jitter simulations in Section 3.3. We find that the maximum flux change for a 0.1 pixel shift is 118 ppm for the nominal grism position, 108 ppm for the top of the array, and 153 ppm at the full-frame grism position.

3.3. Subpixel Crosshatching Calculation

The subpixel crosshatch pattern can introduce time-variable noise as the PSF moves with telescope-pointing jitter and long-

timescale drifts. NIRCcam observations are expected to have a root mean square deviation of 6.0 mas in each axis when measured in 15 s intervals over a 10,000 s observation.⁵

We simulate the effect of this jitter on NIRCcam time-series observations with the following steps:

1. Fit the existing flat field to a crosshatch model in the Fourier domain for a given filter.
2. Extrapolate the crosshatch model to higher frequencies (30 px^{-1}) to estimate the subpixel structure.
3. Simulate an oversampled PSF using `webbpsf` (Perrin et al. 2014).
4. Create a simulated image and its photometric flux.
 - (a) Shift the PSF along the X or Y direction.
 - (b) Multiply the PSF by the simulated flat field.
 - (c) Bin the simulated images into the native pixel size.
 - (d) Divide by the pixel-to-pixel flat field (as would be done in a pipeline).
 - (e) Measure the flux within photometric extraction aperture with the same centroid shift applied.
5. Repeat step 4 for a scan along the X direction across ± 6 mas.
6. Repeat step 4 for a scan along the Y direction across ± 6 mas.

As discussed, in Section 3.1 we fit Equation (4) to the measured power spectral density of the F300M filter for the frequencies above 0.05 px^{-1} . We evaluate Equation (1) for frequencies in the oversampled image from 0 to 30 px^{-1} and multiply this by the number of pixels and the square of the oversampling factor (i.e., 60^2) to convert from the scale-invariant power spectral density (PSD) to the simulated PSD. We then assign random phases uniformly from 0 to 2π for the complex Fourier plane because the complex phase distribution

⁵ See https://jwst-docs.stsci.edu/jwst-observatory-hardware/jwst-pointing-performance#JWSTPointingPerformance-Pointing_stabilityPointingstability.

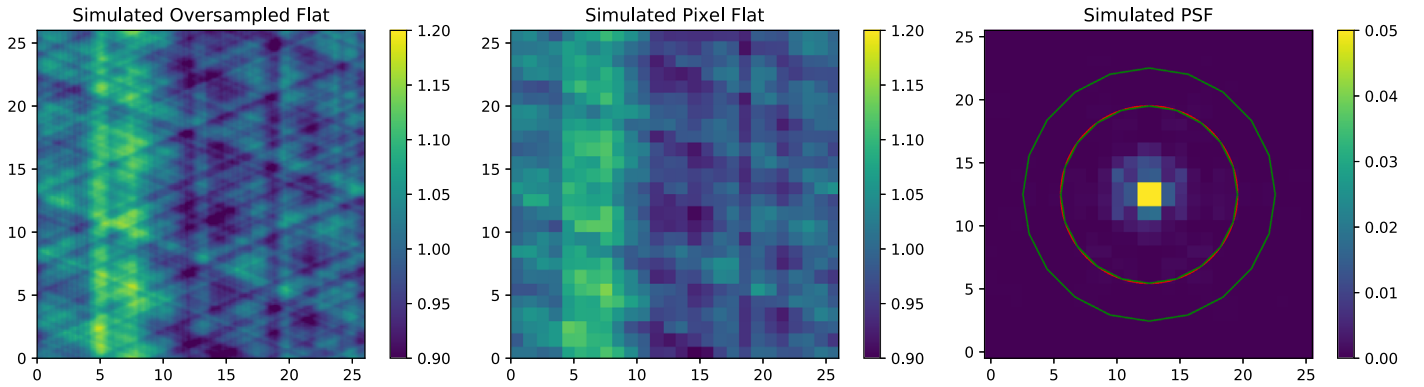


Figure 4. The steps of the subpixel crosshatch photometry simulation are shown here for the A5 detector. We simulate an oversampled flat (left) by extrapolating the power spectrum to frequencies above the pixel-sampling level. This simulated flat is binned to the native pixel resolution to produce a pixel flat field (middle) that would be used by a standard pipeline that has no subpixel corrections. A *WebbPSF* PSF is created at the oversampled resolution, multiplied by the oversampled flat field and then binned to the native resolution (right). Finally, the simulated image is divided by the pixel flat field as would be performed in a standard pipeline reduction. The simulated PSF shows the source and background extraction apertures used in the simulation, which are centered on the PSF.

of the original image is similar to uniform. Finally, we take the real part of the inverse Fourier transform to create an oversampled image, as shown in Figure 4. For comparison to the original flat in Figure 1, we binned the oversampled flat field to the native LW pixel (1.0 native pixels = 60.0 oversampled pixels $\approx 0''.063$). This simulated pixel flat field with dimensions of 26×26 LW pixels has a peak of 1.15 to valley of 0.90 with a robust (outlier-rejected) standard deviation of 0.05. For comparison, the middle of the original F300M flat for the A5 detector has a robust standard deviation 0.06.

We next calculate a PSF using *webbpsf* (Perrin et al. 2014), oversampled by a factor of 60 for the F300M filter as a representative wavelength for the F322W2 filter. *webbpsf* includes an optional blurring due to high-frequency pointing jitter. We set this Gaussian blurring parameter to 1 mas to simulate a worst-case scenario where the pointing drift is dominated by long-timescale behavior with minimal high-frequency jitter. This is a worst-case scenario because the sharpest images will result in the highest subpixel sensitivity. We multiply the oversampled *webbpsf* source by the simulated oversampled flat and bin this to native pixel resolution to create a simulated observation, as shown in Figure 4. This simulated observation is divided by the binned simulated flat, as would be done by a standard pipeline flat-field correction.

We repeat steps 4(a) through 4(e) of the observation simulation by shifting the oversampled PSF by subpixel amounts. The subpixel shifts of the oversampled PSF are performed with `scipy.ndimage.shift`. For each shifted PSF, we multiply this by the subpixel crosshatch pattern and bin the result. These simulated operations represent the pointing drift on the subpixel scale.

We calculate aperture photometry with a circular extraction aperture using a radius of 7 pixels and a background annulus from 7 to 10 pixels on each of the observations. The aperture is recentered by the same position as the shift direction. For each subpixel pointing drift, the aperture sum is subtracted by the background flux average per pixel multiplied by the pixel area of the source-extraction aperture. While the simulations here include no background flux, we use the background subtraction to simulate the operations applied by a photometry pipeline. These are also expected to be similar to a spectroscopic pipeline for a narrow spectral region or spectral line. The

differential flux between the centered and the shifted PSF is shown in Figure 5.

3.4. Subpixel Crosshatching Scan Results

The subpixel crosshatch structure does indeed create flux variations with image motion, as shown in Figure 5. The amplitude of the flux changes is potentially up to 150 ppm. This is significant when compared to atmospheric features of giant planets ($\lesssim 100$ ppm) or the transit depth of an Earth-like planet transiting a Sun-like star (84 ppm).

Fortunately, the subpixel crosshatch systematic is a smooth function of pointing drift, shown in Figure 5. A polynomial function can be fit to this subpixel dependence, and then the centroid of each integration can be inserted as an argument to the function to provide a correction as a function of time. Alternatively, a Gaussian process regression or pixel-level decorrelation (PLD; Deming et al. 2015) could be applied. Centroiding is possible in imaging and spectroscopic modes using PSF fitting or cross-correlation. When the LW grism time-series mode is enabled, short-wavelength (SW) imaging data will automatically be collected simultaneously, enabling the SW centroids to be used to track the motion of the dispersed grism image. The SW centroids should be better than ~ 1 mas using a Gaussian fit to the a central spot in a weak-lens image. Furthermore, the pixels for time-series modes can be characterized in detail because they will be reused for all time-series observations in a given mode. Target acquisition is expected to achieve a centroiding accuracy $\lesssim 10$ mas ($\lesssim 0.15$ LW pixels) for unsaturated target acquisition.⁶ This ensures that time-series observations will reliably return to the same location within the same set of pixels with every visit.

We find that the flux variations due to the crosshatch pattern are very sensitive to the source-extraction aperture radius. As in Figure 5 (right), we calculate a scan of subpixel shifts along the Y direction from -6.3 mas to $+6.3$ mas (for a total length of 0.2 LW pixels at the 63 mas px^{-1} plate scale). Next, we calculate the standard deviation of flux across this scan to characterize its variability. We repeat this calculation for source apertures from 5 to 20 LW px to explore the variability as a function of aperture size. As shown in Figure 6, there is a clear

⁶ See <https://jwst-docs.stsci.edu/near-infrared-camera/nircam-operations/nircam-target-acquisition/nircam-time-series-imaging-target-acquisition>.

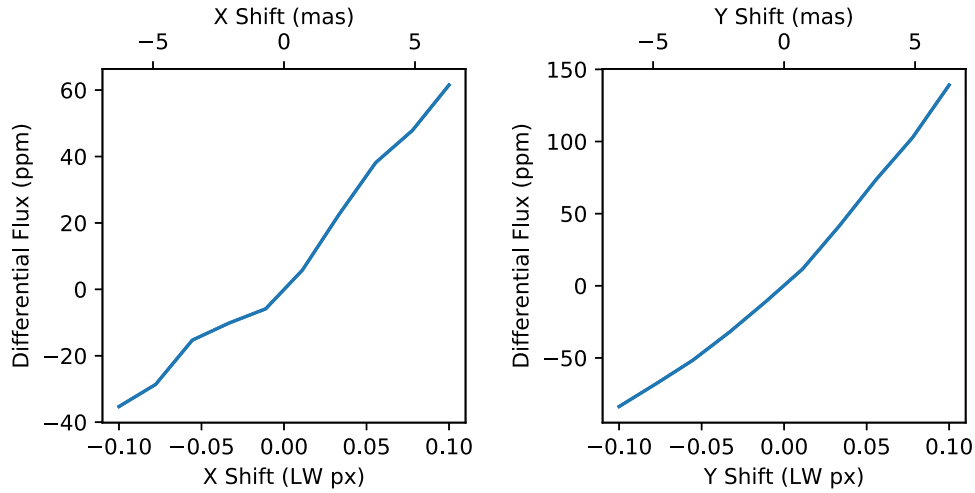


Figure 5. The simulated subpixel crosshatch pattern from Figure 4 for the F300M filter and the A5 detector will cause spurious flux changes with small pointing drifts. Here, we show scans in the X and Y directions of ± 6.3 mas for the ~ 63 mas/pix plate scale of the LW detector. These are correctable with smooth functions such as polynomials.

threshold of about 8 LW px or $0''.5$ below which the flux variability grows strongly. This is likely due to the exponential nature of the crosshatch pattern power spectrum described in Equation (1). Therefore, it is important to extract an extraction aperture greater than 8 LW px to reduce the sensitivity of NIRCam photometry or spectroscopy to the detector crosshatch patterns. This ensures a lower amplitude of variability with pointing jitter as well as a smoother function of centroid position. We note that this is in tension with the issue of $1/f$ noise discussed in Paper I because the read noise favors small apertures or strong weights on the brightest pixels.

3.5. Time-series Simulation

The JWST Attitude Control System team provided a simulation of the expected pointing performance based on ground-test data, which is shown in Figure 7. The pointing is tracked and corrected by the Fine Guidance Sensor (FGS) instrument, but small pointing residuals and drifts on-sky remain for the NIRCam instrument. These pointing residuals include slow ($< 1 \text{ Hz}$) and fast ($> 1 \text{ Hz}$) disturbances that are below and above the fine-guidance control bandwidth, respectively.

The slow disturbances include contributions from thermal distortion from the star trackers and the telescope, as well as fuel slosh that occurs when the telescope is repointed. The thermal-distortion component is affected by repointings of the telescope that change the solar illumination on the star trackers, which is mostly affected by pitch changes but also includes effects from roll changes. The time constant for the thermal-distortion component is expected to be about an hour. The fuel slosh is excited at the beginning of a visit when slewing to the target and stopping at that new attitude. For both of these disturbances, there is an impact on the roll about the guide star that is not corrected for by the low-bandwidth coarse pointing loop.

The fast disturbances come from the reaction wheels, the pulsed MIRI cryocooler, science instrument mechanisms, and high-gain antenna (HGA). These fast disturbances occur at rates $> 1 \text{ Hz}$ that are outside the FGS bandwidth. The reaction wheel jitter is time variable due to changes in wheel speeds as momentum accumulates, whereas the MIRI cryocooler jitter is

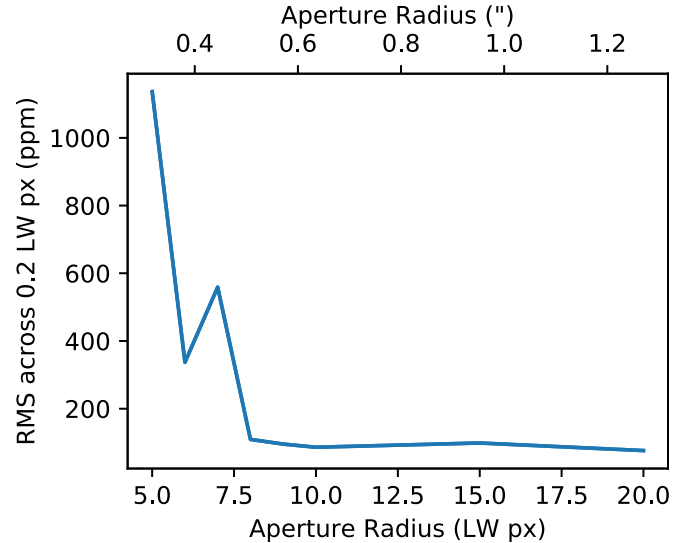
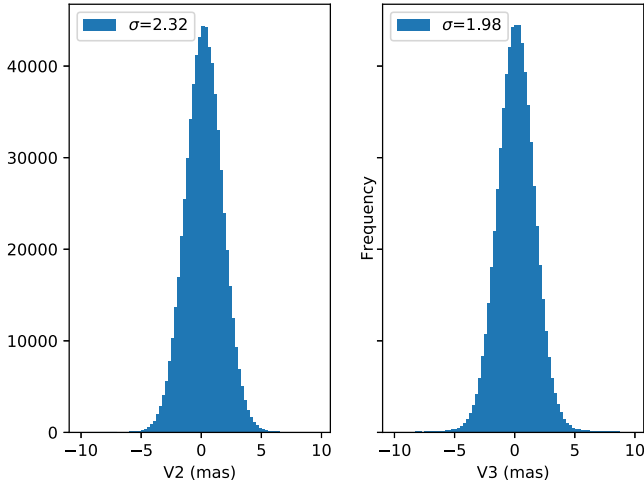
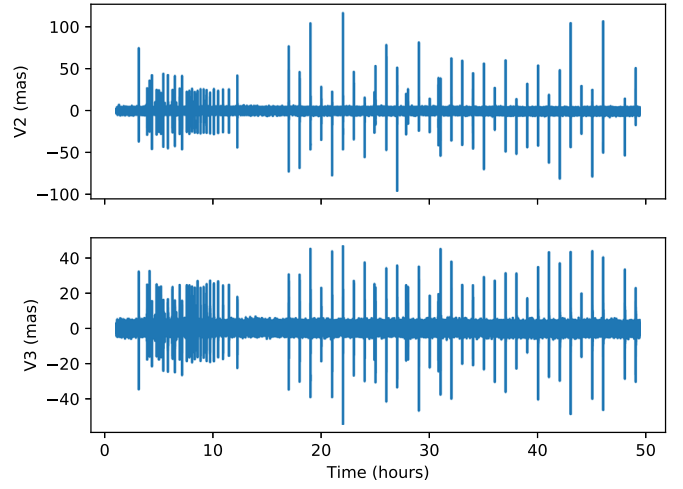


Figure 6. It is important to use sufficiently large ($\gtrsim 8 \text{ LW pixels}$) extraction apertures because the flux variability due to jitter grows significantly for smaller radii. Here, we simulate a -6.3 mas to $+6.3$ mas scan in the Y direction in 10 steps, as shown in Figure 5 (right), and measure the standard deviation of the differential flux. The experiment is repeated for a range of source-extraction aperture radii, where a clear threshold in behavior is visible at about 8 LW pixels .

pulsed at a constant frequency that is tuned to avoid resonances with the deployed dynamics of the observatory. The reaction wheels and MIRI cryocooler are passively attenuated but will still lead to a blurring of the PSF over an integration. The science instrument mechanism and HGA disturbances are intermittent only and managed by the operational procedures. In the case of transiting exoplanet observations, it is preferred to stay under fine-guidance control to keep the star on the same pixels even when mechanisms or the HGA is moved. When the HGA is adjusted, there are larger disturbances, as shown in Figure 7. These HGA moves are designed to keep the antenna pointing at ground-based antennas for telemetry and data download. The errors due to HGA moves are predicted to dampen quickly on $\lesssim$ one-minute e -folding timescales, as shown in Figure 8. This means that for 1-minute integration times,

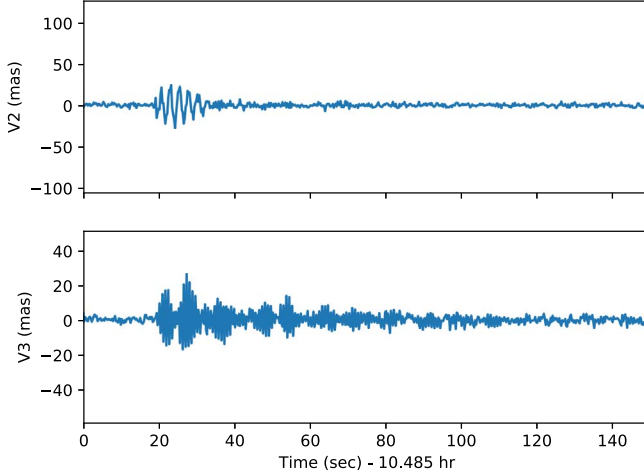


Pointing Histograms

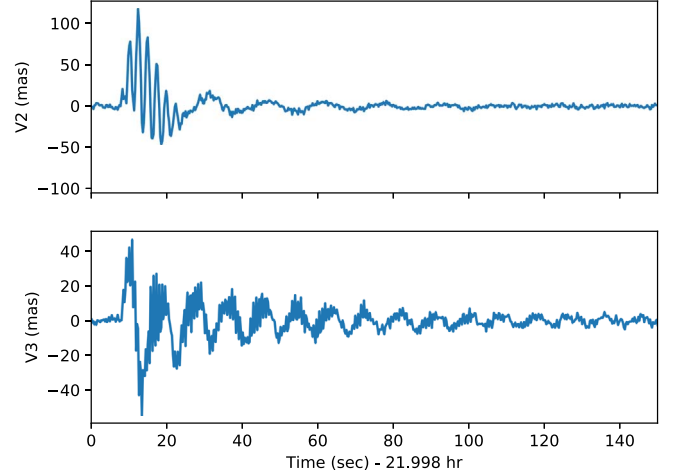


Pointing Time Series

Figure 7. The typical pointing stability is expected to be extremely stable to within a standard deviation of 2.4 mas (0.04 LW pixels), as seen in the histograms (left). There are occasional jumps due to HGA repointings (right). This pointing model only shows the line-of sight variations and does not include thermal distortions or the high-frequency (>1 Hz) jitter. V2 and V3 are the observatory axes perpendicular to the telescope boresight axis.



Pointing Series Zoom 1



Pointing Series Zoom 2

Figure 8. Occasional pointing errors due to HGA repointings are expected to be quickly corrected for by the fine-steering mirror with e -folding timescales shorter than one minute.

only about 1% of integrations will be affected by HGA moves, and the data for these moves could be discarded.

In addition to the line-of-sight disturbances we discussed, there is interboresight motion due to the thermal distortion of the star trackers and the telescope. The interboresight motion results in a centroid shift of sources on the science instruments. While the V2/V3 plane is sensed and corrected for by the fine-guidance control loop, there is a small residual uncompensated roll about the V1 axis due to telescope and star tracker distortions. The prediction from observatory models is that the uncorrected interboresight motion can be approximated as spatially linear in the V2/V3 plane and has a time constant that is 1 hr (P. Maghami 2021, private communication). We therefore model the disturbance with an exponential function,

$$V_{\text{thermal}} = 2.55 \text{ mas}(1 - e^{-t/\tau}), \quad (5)$$

where V_{thermal} is the thermal drift that is in a random linear direction in the V2/V3 plane, the magnitude of the shift is

2.55 mas, and τ is the exponential constant that is set to be 3600 s (P. Maghami 2021, private communication). In reality, the time constant for the star tracker assembly and optical telescope element are different, but this represents a worst case for the slow drift. This thermal term is then added to the above line-of-sight jitter simulation to produce an estimate of the pointing variations for the NIRCcam instrument.

The 2.55 mas thermal drift amplitude in Equation (5) was derived using the worst-case lever arm from a guide star to any point in the field of view of NIRCcam A-side ($\sim 10'06$) and can vary across the focal plane. The boresight term is mainly due to the observatory roll, and is proportional to the magnitude of the vector from guide star and science target star. For example, a variation in the field point of $1'$ along the same line would result in a reduction of image motion of around $2.55 \text{ mas} \times 1'/10'06 = 0.25 \text{ mas}$. The magnitude of pointing errors for MIRI, NIRSpec, and NIRISS are expected to be a factor of 1.6, 1.1 and 1.2 larger, respectively (P. Maghami 2021, private communication).

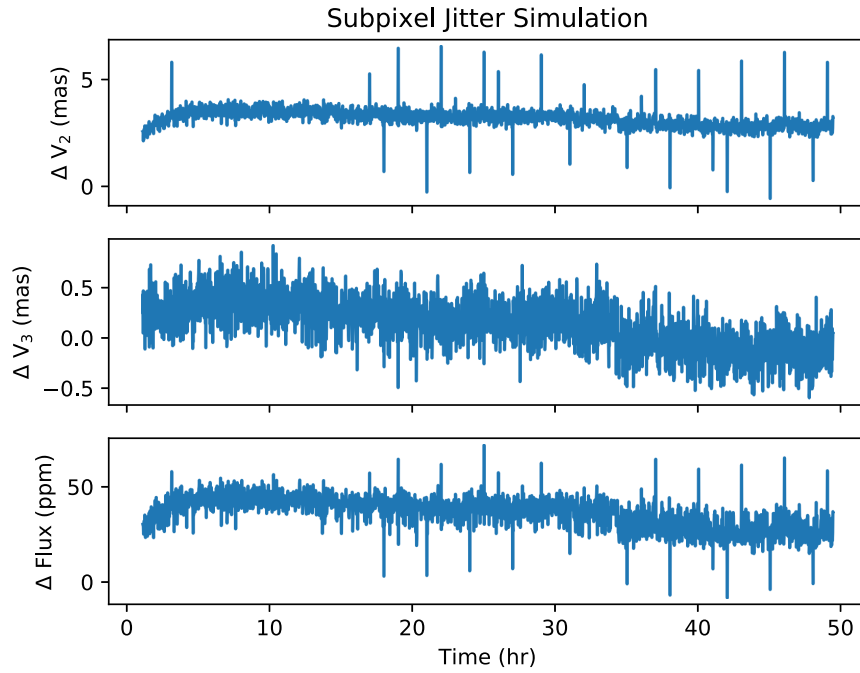


Figure 9. The pointing time series from Figure 7 is binned into a one-minute cadence and a thermal term is added using Equation (5), resulting in a one-minute cadence time series (top two panels) and used as input shifts to the crosshatch model. The very precise pointing expected by the JWST will produce very small flux variations (bottom panel) on the crosshatch pattern, with an expected standard deviation of only 6 ppm.

(The data used to create this figure are available.)

We simulate the change in flux as a function of position using our oversampled flat field, as in Figure 5. First, we bin the subpixel jitter time series into one-minute time bins to represent one-minute integrations, as shown in Figure 9. We then use the positions to shift the PSF on the oversampled flat field as listed at the beginning of this section. Next, we perform photometry on the simulated images with a circular aperture and background annulus.

The resulting flux variations on each integration are very small compared to factors such as the $1/f$ noise. The standard deviation of the time series due to small line-of-sight telescope jitter on top of the subpixel crosshatching pattern is only 6 ppm. Therefore, the pointing jitter is expected to have a small affect on NIRCам time-series stability. Furthermore, centroid motion measurements can be used to correct the flux variations due to subpixel jitter with a polynomial or Gaussian process regression. Small aperture sizes that are most affected by subpixel motions (below $0''.5$ in radius) would benefit the most from a centroid correction function. Another possibility for modeling the crosshatch pattern is a fit to the Fourier amplitude and phase, as done for the flat field, but in the subpixel regime. The spacecraft telemetry can be used to identify HGA moves to assess the short ($\lesssim$ one-minute e -folding) centroid vibrations it causes and remove that data if desired.

We note that the 6 ppm value will be wavelength and position dependent, so it is just a representative value at $3.0\ \mu\text{m}$. This flux variation can vary with both the wavelength dependence of the PSF and the location of a given wavelength on the detector on the spatially dependent crosshatch pattern. For example, we showed in Section 3.2 that at three different field points, the flux variations for a 0.1 pixel shift varied from 108 ppm to 153 ppm. Therefore, the noise due to pointing jitter

of 6 ppm is expected to vary by about 50%, so it can range from 3 ppm to 9 ppm depending on the particular wavelength.

4. Variable Aperture Losses

The JWST was designed to keep the optical performance near the diffraction limit and minimize disturbances to the wave front. However, there are small optical path differences (OPDs) from an ideal wave front that can change with time. The mirrors can undergo slight deviations in position, shape, and angle due to the thermal expansion and contraction of the observatory components. Furthermore, there can be oscillations or vibrations of the observatory components in response to mechanical disturbances such as reaction wheels and the cryocooler. These variations in the OPD will result in different PSF images within the NIRCам instrument. A changing PSF can adversely affect time-series observations by varying the amount of flux contained within a fixed extraction aperture, much like it does for ground-based observing at faster and more dramatic timescales with atmospheric seeing variations.

We assume that the JWST begins by observing an unrelated object for a different science program at a hot attitude or pitch angle of -2deg and is then slewed to an exoplanet system to a cold attitude with a significantly different pitch of 45deg . This large slew would represent the worst case for thermal differences in the time series. Figure 10 shows the time evolution of the OPD wave-front error (WFE) in response to this slew.

We use a parameterized version of Northrop Grumman and NASA Goddard thermal models incorporated into *webbpsf*. These models account for changes in the primary mirror Zernike components from the three main contributions shown in Figure 10: (1) temperature perturbations that occur on the Optical Telescope Element (OTE) backplane as the telescope

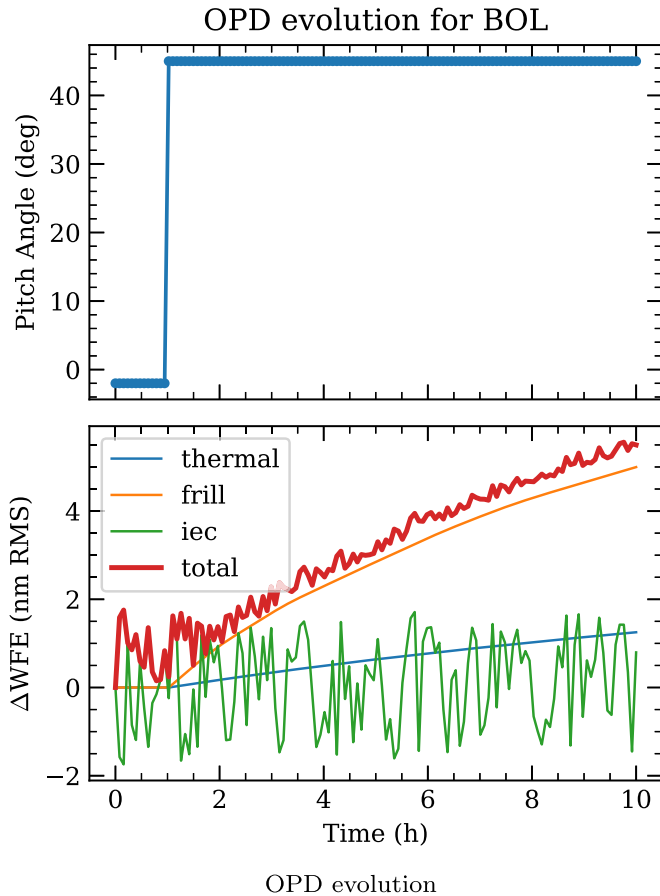


Figure 10. The WFE evolves in time due to thermal changes of the observatory. Here, we assume a single slew to the exoplanet science target (top panel), which causes slow telescope and frill changes (blue and orange lines). In addition to these variations, the IEC heaters (green line) cause short-timescale variations in the wave front. Values assume performance for BOL.

pitch angle changes during a slew (blue), (2) tensioning of the primary mirror frill structure (orange), (3) and effects of the ISIM Electronics Compartment (IEC) heaters (green). The total effect of all these WFE sources is shown in red. This model is based on measurements from cryogenic vacuum testing at NASA Johnson where the telescope was cooled from ~ 300 K to ~ 40 K. The time constant for the thermal changes is about 5–6 days, so a typical exoplanet transit or eclipse observation ($\sim$ hours) will not reach thermal equilibrium. We assume beginning-of-life (BOL) conditions where the sunshield has had no degradation due to micrometeoroid impacts.

We calculate the PSFs for each simulated OPD using `webbpsf` (Perrin et al. 2014) to examine how this can impact the time series. Figure 11 shows the differences in the PSFs at different epochs of the time evolution. In the beginning, we assume that the primary mirror and frill are in thermal equilibrium so the only wave-front disturbances come from the IEC heaters, which change the core of the PSF ($\lesssim 5$ LW px or $\lesssim 0''.32$), as shown in Figure 11, left. After 1 hr in our simulation, the telescope is slewed and the thermal and frill changes become important. Now, the PSF differences are most pronounced at intermediate distances (~ 10 LW px or $0''.63$).

We next examined how the PSF variations manifest as time-variable extraction aperture losses. We calculated aperture photometry using the `photutils` package and experimented with a range in aperture radii from 4 to 23 LW px ($0''.25$ – $1''.5$).

We centered the apertures by fitting a 2D Gaussian. The resulting changes in flux are shown in Figure 12 in parts per million.

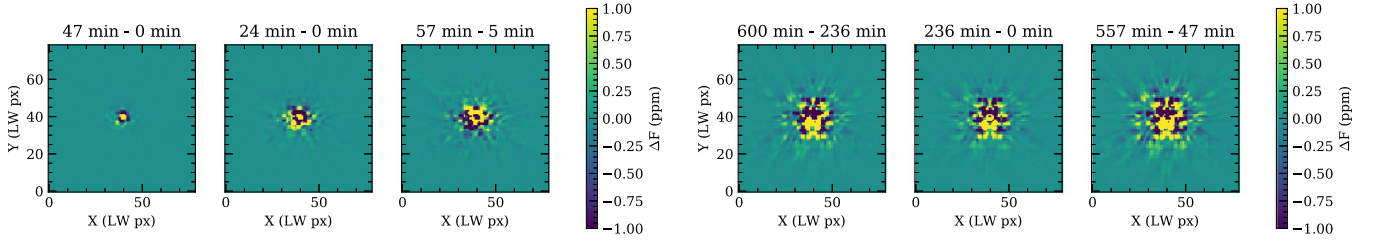
In the variable wave-front time series, the largest variations are from slow thermal variations of the wave front. From the start of the slew at 1.0 hr to the end of the simulation at 10.0 hr, the variations can be as much as 55 ppm. The thermal variations cause smooth changes in the aperture flux, so they can be removed with linear or quadratic detrending. This will likely be smaller than the host star flux variability. For example, the typical solar variability on minute- to hour-long timescales is 50 ppm or larger, as measured by the Total Irradiance Monitor (Kopp et al. 2005). The stellar variability tends to grow larger with later spectral types (Rackham et al. 2019).

We find that the variable aperture losses also are smaller for large photometric apertures. For small extraction apertures ($\lesssim 6$ LW px or $\lesssim 0''.4$), there is also a 5–10 ppm variability due to the IEC heaters, as seen in Figure 12 because the IEC heaters affect the core of the PSF. The thermal variations affect the PSF farther out, so an aperture radius larger than 10 LW px ($0''.63$) is needed to minimize the thermal disturbances. With sufficiently large apertures (> 18 px), the time-variable aperture losses go down to 2 ppm. We separated the IEC component and thermal components in Figure 12 by choosing images dominated by those components. For the IEC-dominated component, we compare the fluxes for images 6 and 1 (before the slew or thermal changes happen) and for the thermal-dominated component, we compare the first and last images. We note that the aperture-loss experiments assume a circular fixed aperture with uniform pixel weights, but another weighting scheme might be applied, such as the covariance weights in Paper I.

The variable aperture losses are modest ($\lesssim 60$ ppm) for all apertures tested here, but could be reduced with more advanced photometry. One method commonly used for ground-based observations is to scale the aperture radius by the full width at half maximum (FWHM). Scaling apertures with the size of the PSF have successfully reduced noise with Spitzer IRAC light curves (Knutson et al. 2012; Ingalls et al. 2016). Here, we estimate the FWHM by fitting a Gaussian to every simulated PSF. This does reduce the thermal PSF disturbances by ~ 5 ppm for an aperture radius that is 7.4 times the FWHM versus an aperture radius that is fixed at 16 px. However, an aperture scaled by the FWHM is less helpful for the fast IEC heater variations with small aperture radii.

5. Charge-trapping

The HST WFC3 shows noticeable ramps with each HST orbit, where the detected flux rises approximately as $1 - Re^{-t/\tau}$ as function of time t , where R is the ramp amplitude and τ is the exponential time constant (e.g., Berta et al. 2012), as seen in Figure 13 for an example light curve. The ramp effect has been explained and modeled as the capture and release of electrons in detector charge traps (Zhou et al. 2017). These charge traps (also known as persistence) are responsible for latent images, where a bright exposure followed by darker illumination shows a residual image whose emission rate decays with time (Smith et al. 2008; Leisenring et al. 2016; Tulloch 2018). We present a schematic for the charge-trapping effect in Appendix C.



IEC heater PSF differences

Thermal PSF differences

Figure 11. Wave front errors will change the PSF with time, shown here for the F444W filter. PSF differences are largest for the core ($\lesssim 5$ LW px) due to the integrated science instrument module electronics compartment heater (left), and they are largest at intermediate distances (~ 10 LW px) due to thermal variations (right). The PSF differences shown here are selected from specific times in the entire simulation to highlight the morphology and variation that are possible when the PSF changes.

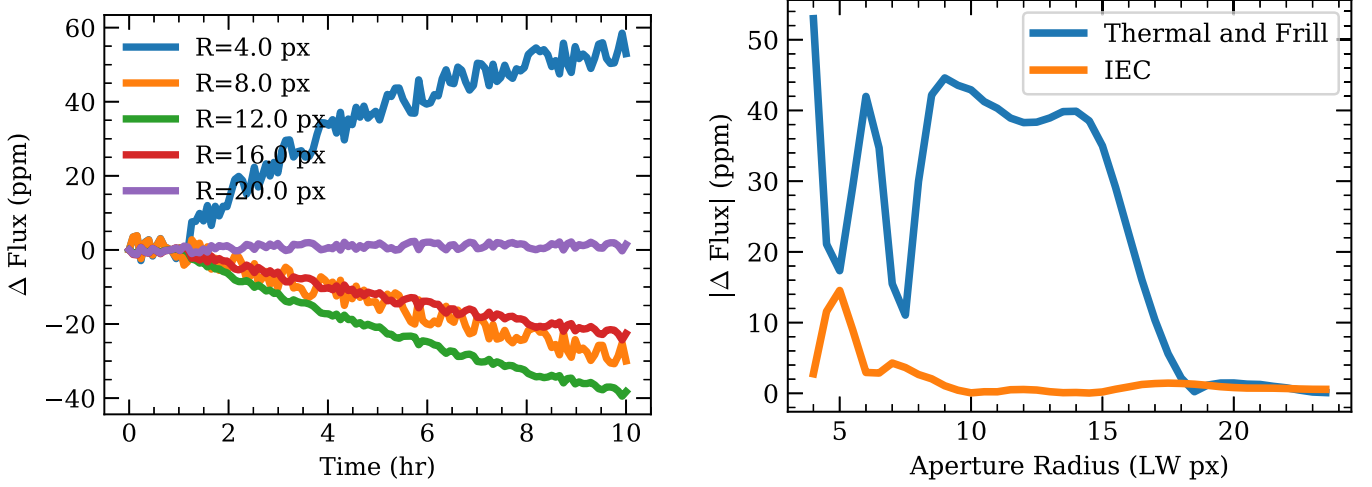


Figure 12. Wave front errors will cause flux losses within an extraction aperture, causing variations in time series. Left: The flux will vary smoothly along the telescope thermal timescale of 5-6 days, but small apertures are sensitive to the IEC heaters. Right: Small apertures are more sensitive to WFE, but above 18 pixels, the time variations are less than 2 ppm.

Fortunately, the JWST NIRCcam charge-trapping effects are much smaller than for the HST WFC3. Leisenring et al. (2016) find that the persistence rate after saturating the detectors and turning off the lamp is $0.5\text{--}5 \text{ DN s}^{-1}$ or $0.9e^- \text{ s}^{-1}$ to $27e^- \text{ s}^{-1}$ at the 30 s mark, depending on the detector. After 1000 s, the persistence level falls to $10^{-2} e^- \text{ s}^{-1}$ for the A5 detector, which is the detector used for grism time series. For a pixel that fills to about 60% well depth in a 20 s integration, the count rate for that pixel will be $3200 e^- \text{ s}^{-1}$. Therefore, the persistence level compared to the signal level at 1000 s into a time series will be about 3 ppm. This gives an order-of-magnitude estimate for the ramp effect, assuming that the charge-capture timescale is not significantly longer than the charge-release timescale.

Figure 13 shows a comparison between a light curve from the HST WFC3 and a laboratory test of a NIRCcam-like detector. The HST WFC3 observation comes from secondary-eclipse observations of the CoRoT-1 system (GO program 12181), taken with the G141 Filter and a GRIMS128 subarray aperture. The exposure was taken with SPARS10 MULTI-ACCUM mode with 16 samples for an exposure time of 100.65 s. As seen in Figure 13, there is a pronounced ramp with each HST orbit. The first orbit has an especially large ramp compared to the others, and thus the first orbit is typically discarded in light-curve analysis (e.g., Kreidberg et al. 2014;

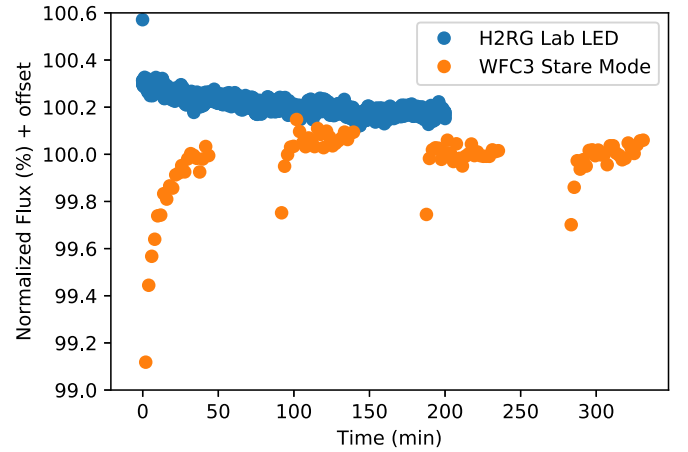


Figure 13. The JWST H2RG detectors will have much smaller ramp amplitudes than the HST WFC3 detector. The WFC3 light curve (orange points) shows a ramp-up behavior with every orbit due to charge-trapping (e.g., Zhou et al. 2017). The Earth occultations with every HST orbit prevent the detector from reaching steady state. The JWST, on the other hand, will observe uninterruptedly to approach steady state and use H2RG detectors that have smaller charge-trapping amplitudes. A Lab-LED illuminated H2RG detector light curve (blue points) shows much less pronounced charge-trapping behavior. The variations within this H2RG light curve are dominated by the lamp output and not by the detector itself.

Evans et al. 2016; Wakeford et al. 2017). The first orbit, however, can be modeled with charge-trapping (e.g., Zhou et al. 2017).

For the laboratory test of a NIRC*am*-like detector, we use a NIRC*am* detector that was not selected for the flight instrument, but has many of the same noise behaviors. The detector is controlled with SIDECAR ASIC electronics, again with an unused NIRC*am* part in a dewar at the University of Arizona. An exposure was commanded in RAPID read mode and a 2048×64 subarray. There were eight groups per integration with four output amplifiers, resulting in an integration time of 2.72 s. For this test, a $5.2 \mu\text{m}$ cutoff detector is illuminated by a $3.4 \mu\text{m}$ LED source with 1.5 mA of current. The current and integration setting of this lamp result in a maximum pixel value of about 60% well capacity and thus are below saturation or strong nonlinearities.

The exposure is started before the lamp is turned on to study the charge-trapping ramp-up immediately following illumination. The laboratory H2RG LED illumination test shows much smaller ramp-up behavior than HST WFC3. In fact, there is no detectable ramp-up behavior at all because the lamp varies with a standard deviation of 370 ppm, so any ramp behavior is below that level. Furthermore, the JWST does not have to contend with Earth occultations, so the detector has the full time of the baseline (typically >1 hr) to reach steady state. As seen in Figure 13, there is a gradual decrease in flux following the lamp illumination. These effects are likely related to the LED lamp output responding to small temperature drifts rather than charge-trapping itself because the measured flux variations can change direction and magnitude with repeated experiments. Based on the 3 ppm persistence level estimate and 370 ppm upper limit from laboratory tests, we conclude that ramp-up behaviors will play a much smaller role for the JWST NIRC*am* as they do for the HST WFC3. A time-series test during NIRC*am* on-orbit commissioning will help evaluate the timescale and amplitude for any settling behaviors such as charge-trapping.

6. Detector Temperature Variations

The NIRC*am* detectors will be actively controlled to an expected precision of $\lesssim 1$ mK as measured in cryogenic vacuum testing of the flight hardware. This is achieved with a focal plane heater and temperature sensor. We estimate here how 1 mK temperature fluctuations could manifest as time-variable signals. We note that measurements from Hall (2006) show that 100 mK variations result in 80 DN variations in the detected absolute counts. Interpolating linearly to smaller temperature fluctuations, there will be 0.8 DN ($0.8 e^-$ for the gain assumed in Hall 2006) per 1 mK temperature variation. If this were to unevenly affect the background and source pixels in a set of time-series extraction apertures, it might adversely affect the time series, but only if it shows up in the slope images.

We perform another test with the same University of Arizona laboratory dewar as discussed in Section 5 to test the effect of detector temperature on exoplanet time series of illuminated slope images with background subtraction applied. We mount two $5.2 \mu\text{m}$ cutoff detectors in a focal plane array (FPA) that can accommodate four detectors at once. They are both illuminated by the same $3.4 \mu\text{m}$ LED to simultaneously monitor the flux as the FPA changes temperature. The exposure

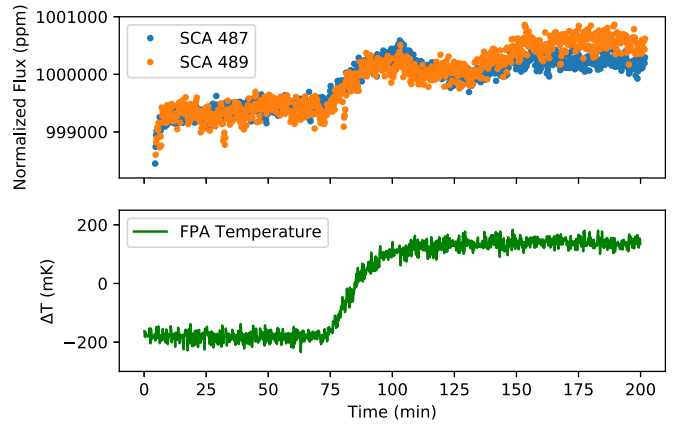


Figure 14. The laboratory time series of two NIRC*am*-like detectors and electronics shows an increase in signal at higher detector temperatures. The temperature was adjusted by 330 mK here by design to see the effect in addition to the variability of the LED lamp. However, in flight, the temperature variations are expected to be 1 mK, resulting in only 3 ppm variations. The dip at ~ 130 minutes is most likely due to small current fluctuations in the LED. The initial five-minute ramp is also likely due to the LED, because it was observed to change directions and magnitude depending on the test and lamp current. FPA telemetry can be used to calibrate temperature variations in flight.

consists of 1000 integrations of a 2048×256 STRIPE subarray with four outputs, eight GROUPS in RAPID readout mode (one frame per group), and an integration time of 10.7 s. This will more directly translate into the time variability experienced in exoplanet light-curve measurements than the absolute counts in Hall (2006). The detectors are illuminated by an LED while the temperature are simultaneously adjusted with an FPA heater. The detector temperature is monitored on the FPA with a Cernox sensor. We adjust the set point so that there is a 330 mK increase in the FPA temperature and measure the time series over 200 minutes to study the variations. We make this large temperature change and interpolate to 1 mK because some LED lamp variations impede our ability of measuring flux changes much below 300 ppm without detrending models and assumptions about the noise.

Figure 14 shows the change in temperature as the heater is powered and also the resulting flux change on two different detectors. We measure the flux with a rectangular extraction aperture of 106×142 and 79×180 pixels on the 487 and 489 detectors, respectively, so the results represent the average pixel behavior. The fluxes of the two apertures are subtracted by horizontally offset background apertures that are masked from direct illumination by the LED. For the 330 mK increase in temperature, both detectors (numbered 487 and 489) show approximately the same 900 ppm increase in normalized flux. We note that the illumination levels of the two detectors are different because the peak counts per integration were 25,000 DN/px and 39,000 DN/px for detectors 487 and 489. This suggests that the temperature variations result in fractional changes to output photometry or spectroscopy instead of an absolute change in DN. There are variations in the lamp output on the two detectors due to current instability or the temperature of the LED lamp itself in addition to the S curve due to the temperature increase.

Next, we linearly interpolated this fractional flux increase from 330 mK variations down to 1 mK, i.e., a rate of 2.7 ppm per mK temperature variation. We therefore expect detector temperature fluctuations to be a relatively small part of the

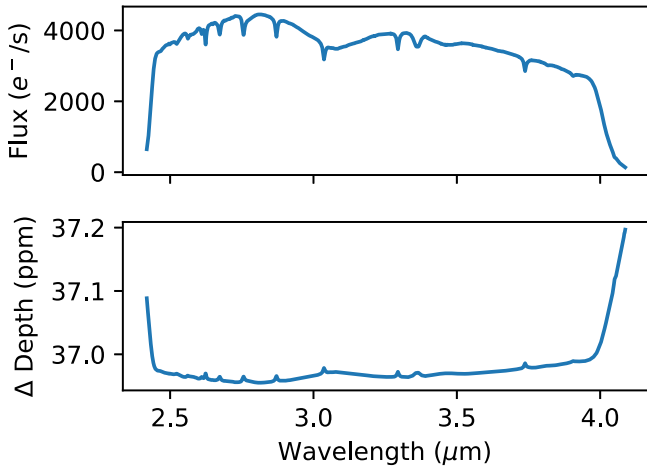


Figure 15. Reciprocity failure will change the transit depth measured by a HgCdTe detector like the one measured in Hill et al. (2010). The top panel shows an input A0 V spectrum signal and the bottom shows the change in transit depth caused by reciprocity failure, which adds a 37 ppm offset to the transit depth, but sub-ppm variations across the spectrum.

time-series error budget. A caveat is that one of the temperature monitor controls (TMC1) for the A5 detector used for grism time series can result in 20 mK oscillations in temperature. Fortunately, there is a redundant controller TMC2 that keeps the variations to 1 mK. In the event that the TMC2 fails, calibrations must be applied to TMC1, which could potentially see 54 ppm variability with temperature. Alternatively, the B5 detector could be used, but with a different lower transmission grism and software adjustments required to operate the subarrays for the B side of NIRCам.

7. Reciprocity Failure and Nonlinearity

The HgCdTe pixels are nonlinear with respect to both the percentage of their well capacity and also with the incident photon rate. They are nonlinear with respect to the percentage of well capacity filled because the photodiode becomes less sensitive to incoming photons as the depletion region shrinks. This nonlinearity with the depletion region size (percentage of well capacity or well fraction) is calibrated and corrected for in the pipeline using a constant-flux source measured with different integration times. It can also be mitigated by keeping the brightest pixels safely below the saturation level of the detector.

The nonlinearity with incident photon rate is called reciprocity failure (Hill et al. 2010; Biesiadzinski et al. 2011) and has not been as well characterized on the JWST detectors. Here, we estimate the difference in flux that could occur if the NIRCам detector behaves like the HgCdTe 1.7 μm cutoff detector measured by Hill et al. (2010), shown in Figure 4 of that paper. For the input signal, we assume an A0V host star, which has strong hydrogen features and a deep transit depth of 2%.

The resulting change in transit depth due to reciprocity failure is shown in Figure 15. The main effect of reciprocity failure is an offset in the transit depth. For the 2% (20,000 ppm) transit depth considered here, there is an offset of 37 ppm due to reciprocity failure. The signal level changes in stellar absorption features, with the Raleigh-Jeans shape of the continuum and at the edges of the filter bandpass, but this changes the transit depth by less than

Table 1
Estimates of Known Systematic Errors in NIRCам Time Series

Noise Source	Time Variability (ppm)	Perennial Offset (ppm)
Pointing Jitter on Crosshatch Pattern	6	...
Time-variable Aperture Losses	2	...
Charge Traps	3	3
Detector Temperature Fluctuations	3	...
Nonlinearity and Reciprocity Failure	...	37
Ghost Images	...	1
Total	9 ^a	37

Note.

^a Our estimates of the known error sources indicate very small systematic levels. We urge caution in using this as a standard because there will likely be new, unknown error sources that contribute to the JWST time-series errors. We assume that the pointing jitter is correlated with aperture loss but that charge traps and detector temperature are independent of all other noise terms.

1 ppm at the wavelengths of the stellar features. Therefore, the main effect of reciprocity failure is an offset in transit depth that could affect the intercomparison between different instruments or facilities. We expect that residual nonlinearities that are not corrected by the pipeline will have a similar effect on the signal.

8. Conclusions and Combined Noise Floor

The JWST has the potential to transform our understanding of transiting planets with extremely high-precision light-curve measurements. The JWST can collect enough photons to observe atmospheric gases and even biosignatures in optimistic scenarios with no astrophysical noise (e.g., Krissansen-Totton et al. 2018). However, there are electronic $1/f$ noise effects (discussed in Paper I) and systematic effects (this paper) that can limit the JWST ability to achieve the photon limit. We have discussed the systematic errors in this work that can appear in time-series observations.

We combine the known sources of systematic error here to give an estimated noise floor for the NIRCам time-series modes. Our systematic noise sources, as discussed in Sections 3 through 7, are summarized in Table 1. We categorize the noise into two flavors that can impact the science. The first is time variability that can occur over an exoplanet transit light curve or similar measurement and will likely be different from one visit to the next. The second is a perennial or permanent offset that will appear in all visits in a similar way and could potentially affect interinstrument comparisons when a spectrum is stitched together over a wide wavelength range from the visible to the near-infrared.

The largest of the known time-variable noise sources is due to pointing jitter on top of the intrapixel sensitivity. Detector crosshatching structures extend down to the subpixel level where they will not be corrected for in standard flat-fielding procedures. Fortunately, the pointing drifts are expected to be very small (2 mas) other than brief and infrequent HGA moves, so the standard deviation of flux due to pointing jitter is expected to be only ~ 6 ppm. The variations in flux at the subpixel level appear smooth enough that they could likely be calibrated as a function of position. Time-variable aperture losses play a smaller role than the subpixel crosshatching. We estimate how heating and cooling of various observatory components can change the PSF and find that they will likely appear as 2 ppm variability. We note that the subpixel crosshatching and time-variable aperture losses grow for

apertures smaller than $1''$, so wider apertures are necessary for these estimates to apply.

Charge traps and detector temperature variations will also play a small role in time-variable noise. We estimate that charge traps will affect the JWST at the 3 ppm level or smaller, which is less than for the HST due to low charge-trap densities and because the JWST will observe continuously without Earth occultations. Detector temperatures will be controlled to 1 mK levels and thus keep the flux stable to $\lesssim 3$ ppm based on interpolation of laboratory measurements.

We combine the errors assuming that the pointing jitter and time-variable aperture losses are highly correlated and that all other errors are statistically independent. The pointing jitter and time-variable aperture losses are both related to thermal variations and displacement of instrument optics that have similar time constants. We therefore assume that these two error sources add linearly. The remaining errors are unrelated phenomena, so we assume they add in quadrature. The final estimate of the time variability is then 9 ppm per visit. This indicates that the main factor impacting NIRCam grism time series will be correlated electronic noise, as discussed in Paper I, and that the systematic noise floor will be less important. The $1/f$ noise discussed in Paper I with the GRISMR disperser can be as high as 1000 ppm per integration, so it remains the largest challenge to high precision. Efforts to mitigate $1/f$ noise are underway, such as the GRISM C disperser, which is discussed as a science enhancement for the observatory.

We urge caution when assuming the low noise floor of 9 ppm, however. Laboratory measurements are plagued by instabilities in light sources, where the current and temperature are difficult to control at the ppm level. Of the laboratory tests measured at NASA centers and the University of Arizona laboratory, the most stable light curve had a standard deviation 320 ppm over 20 minutes after a linear trend was removed. The real noise floor will only be measured in flight on a more stable source such as a quiescent star while using the real flight configuration. There will be a commissioning exercise of an eclipsing binary to provide insight about the noise floor and stability of the NIRCam grism time-series mode.

While we estimate a noise floor of 9 ppm and urge caution about unknown sources of error, corrections can be made to reduce the standard deviation of time series. The error sources we discussed are potentially correctable by monitoring state variables of the instrument and telescope such as antenna moves, the image centroid, the charge-trap state, detector temperature, and the PSF shape. Additionally, some of these noise sources (such as small detector temperature fluctuations or IEC heaters) will be independent from one visit to the next. We urge cautious estimates of the noise for the first JWST cycle, but future cycles may achieve better performance.

We note that two noise sources can potentially produce a perennial offset in the transmission spectrum: the charge-trapping effect, and nonlinearity. The charge-trapping, though small, will behave similarly from one visit to the next and can introduce an exponential-like curvature to the time series. Fortunately, this curvature is small (3 ppm) and will be correctable with algorithms such as RECTE (Zhou et al. 2017). Detector nonlinearities and reciprocity failure will also occur in all visits because the illumination of the detectors will be the same. Reciprocity failure can potentially introduce a ~ 40 ppm spectral offset from the true value. Detector nonlinearities and reciprocity failure can potentially

be calibrated with data from the absolute calibration program. However, care must be taken when performing interinstrument comparisons across JWST instruments or other observatories. We note that the systematic errors discussed in this work are more broadly applicable to instruments and detectors beyond NIRCam and the JWST. The intrapixel sensitivity effects and crosshatching pattern discussed in Section 3 can occur on many different detectors. Generally, the intrapixel sensitivity effects are most pronounced when a PSF is smaller than 2 pixels (e.g., short-wavelength Spitzer channels and JWST NIRSpec at short wavelengths; Ferruit et al. 2014; Ingalls et al. 2016), but less important when the PSF spans over many pixels (e.g., JWST NIRISS Doyon et al. 2012). We therefore expect that NIRSpec prism/grating observations will have a higher subpixel sensitivity than the estimated 6 ppm value estimated for NIRCam, but a lower sensitivity than for JWST NIRISS. They will also be modified by the detector-dependent and spatially dependent crosshatch pattern and any other subpixel detector structure. Thus, the impact of crosshatching on observations that use HgCdTe detectors will depend on the optical design as well as the amplitude of the crosshatch pattern.

The variable aperture losses due to the thermal behavior of the telescope discussed in Section 4 are also relevant for other space-based observatories. For example, high-precision time-series observations are proposed for LUVOIR (The LUVOIR Team 2019), the Origins Space Telescope (Meixner et al. 2019), and the Nancy Grace Roman Telescope (Gaudi et al. 2019), which all could have long thermal settling timescales. We expect the numbers in this paper to serve as useful starting points in estimating the noise in these observatories. Finally, the crosshatching pattern, charge-trapping effects, detector temperature fluctuations, and nonlinearity effects occur on all HgCdTe detectors, which are used widely in astronomy (e.g., WIRCam, HPF, ARCoIRIS, and MOSFIRE; Baril et al. 2006; McLean et al. 2012; Mahadevan et al. 2014; Schlawin et al. 2014), so high-precision measurements will be aided by image stability, temperature control, and nonlinearity corrections.

Thanks to Peiman Maghami for providing the line-of-sight pointing-jitter simulation and information on interboresight motion for use in estimating the subpixel sensitivity as well as explaining sources of pointing error. Thanks to Knicole Colon for useful discussions on the effect of pointing jitter on time-series stability. Thanks to Yifan Zhou for providing some charge-trapping timescales for the HST. Rafia Bushra did extensive analysis of cryogenic vacuum testing that provided valuable insights about the various systematic errors that JWST can encounter. Thanks to Sam Myers for bringing the Suissa et al. (2020) paper to our attention. Thanks to John Stansberry for providing information on the subarray field points and for tireless work to enable and improve the grism time-series mode. Thanks to Chaz Shapiro and Eric Huff for providing information on the Euclid detector crosshatch pattern. Funding for E. Schlawin is provided by NASA Goddard Space Flight Center. T. Greene acknowledges funding from the NASA JWST project in WBS 411672.05.05.02.02. We thank the anonymous referee for their helpful suggestions to improve the paper and its readability.

Facilities: JWST(NIRCam), NASA Goddard Space Environment Simulator, NASA Johnson Chamber A.

Software: astropy package (Astropy Collaboration et al. 2013), pynrc (Leisenring 2019), webbpsf (Perrin et al. 2014),

numpy (van der Walt et al. 2011), scipy (Virtanen et al. 2020), pysynphot (Lim et al. 2015), ncdhas, photutils (Bradley et al. 2016), matplotlib (Hunter 2007).

Appendix A

Detailed Characterization of Subpixel Crosshatching

Here, we primarily study the subpixel behavior of the A5 detector because it is used for the NIRCam grism time-series spectroscopy mode, covering $2.4\text{--}5.0\ \mu\text{m}$ with a plate scale of $63\ \text{mas}\ \text{px}^{-1}$ (Greene et al. 2017). The flat field of the A5 detector has a pronounced crosshatch pattern, as shown in Figures 1 and 16. The patterns are located at $23^\circ.1$, $90^\circ.9$, and $158^\circ.6$ counterclockwise (CCW) from the $+X$ direction of the detector. The angles between these lines are $67^\circ.8$, $67^\circ.7$ and $44^\circ.5$. These angles and patterns can be analyzed with a 2D power spectrum, as shown in Figure 16.

The angles of the crosshatch patterns are determined by the crystal pattern of HgCdTe. HgCdTe has a zincblende structure with tetrahedral bond angles where each Hg or Cd atom is surrounded by four Te atoms (Gemain et al. 2012). HgCdTe detectors are manufactured using molecular beam epitaxy upon a substrate, a process that can result in topological defects with

peak-to-valley amplitudes of $5\text{--}20\ \text{nm}$ in height variations (Chang et al. 2008).

The surface morphology of the HgCdTe crystal shows that the crosshatch patterns are oriented along the intersection of the (211) growth plane of the crystal and the eight HgCdTe slip planes. The relative angles of the HgCdTe slip planes and (211) growth plane are $44^\circ.42$, $67^\circ.79$, and $67^\circ.79$ (Chang et al. 2008), very close to the observed crosshatch angles. A projection of zincblende structure is shown in Figure 16 at the same orientation of the power spectrum structure. The similarity between the crosshatch patterns in the flat field and the topological variations observed in Chang et al. (2008) leads to the likely conclusion that the surface variations lead to quantum efficiency variations. Thus, the crosshatch pattern is most likely related to the crystal lattice structure of the HgCdTe substrate and not to the pixel circuitry or readout electronics.

The crosshatch structure of the detectors extends down to the subpixel level, so it will not be fully corrected with a flat-field division. This subpixel structure has been imaged with microscopy on a candidate Euclid HgCdTe detector (Shapiro et al. 2018). Atomic force microscopy shows topological features that are approximately $1.2\ \mu\text{m}$ in width (Chang et al. 2008) compared to the $18\ \mu\text{m}$ pixel sizes.

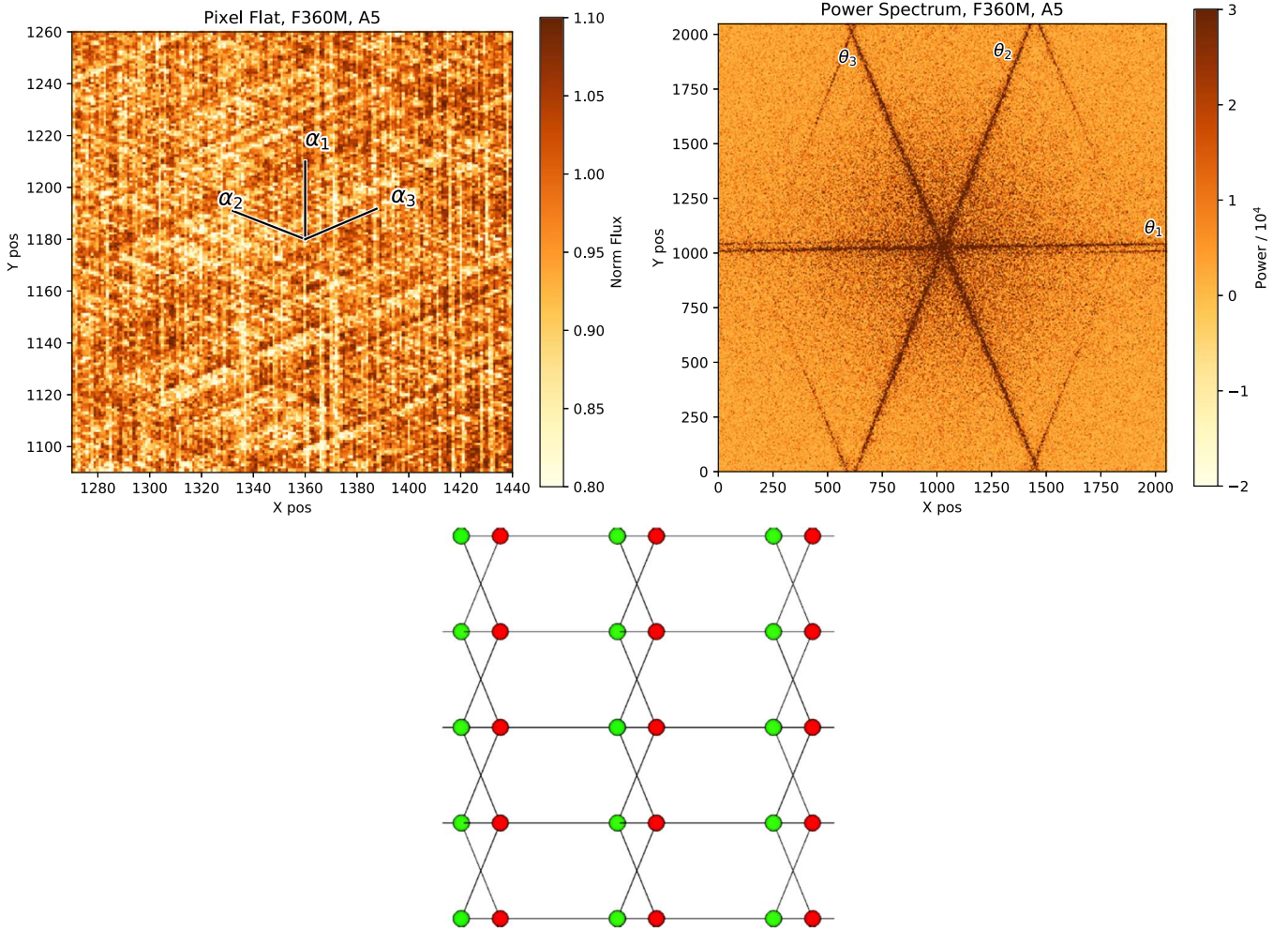


Figure 16. The A5 detector flat field has a pronounced crosshatch pattern at $\alpha_1 = 90^\circ.9$, $\alpha_2 = 158^\circ.6$ and $\alpha_3 = 23^\circ.1$ CCW from the positive X direction (top left plot). The flat field from Figure 1 is reproduced here. The crosshatch pattern is best analyzed with a 2D Fourier power spectrum (top right), which shows a continuum of frequencies aligned with the three crosshatch angles. Note that the lines in the frequency plane (θ_i) are perpendicular to the crosshatch directions in the image plane (α_i). These three relative crosshatch angles ($67^\circ.8$, $67^\circ.7$ and $44^\circ.5$) are similar to a projection of the tetrahedral (zincblende) lattice structure of HgCdTe (bottom), $67^\circ.8$, $67^\circ.8$, $44^\circ.4$). The green circles represent either Cd or Hg, and the red circles represent Te.

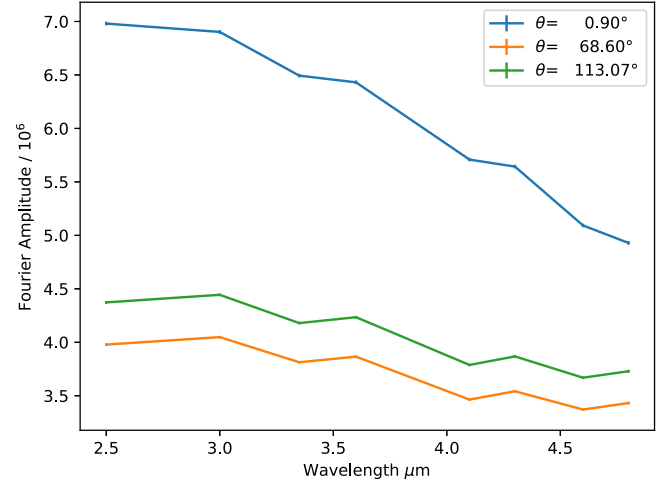
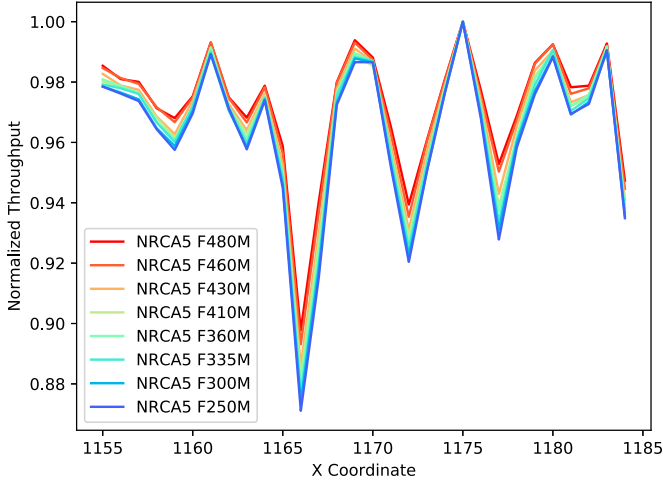


Figure 17. The crosshatch pattern changes as a function of wavelength likely because the shorter wavelengths resolve the structure better. In other words, the crosshatch pattern becomes slightly blurred at longer wavelengths due to diffraction. The shorter wavelengths have deeper crosshatch troughs in a flood-illuminated flat field (left). The flat fields are shown here for the A5 detector, the same as shown in Figure 16. The sharper crosshatch features at shorter wavelengths show up as larger Fourier amplitudes (right).

The width of the structures can also be estimated from the behavior for crosshatch lines as they cross pixel boundaries (Ninan et al. 2019). We estimate that the crosshatch pattern oriented at 90.9° crosses 1 horizontal pixel for every 64 vertical pixels. While crossing the boundary, there are ≈ 38 rows where a sharp feature in the crosshatch pattern spans 2 pixels instead of 1. The sharp feature is traced by the pixels that drop significantly below the mean flat-field value in a local region. If the sharp feature in the crosshatch pattern has a tophat shape, then these 38 rows where the pattern spans two pixels imply a tophat full width of 0.6 pixels or a physical width of $10.8 \mu\text{m}$ for an $18 \mu\text{m}$ pixel pitch. This is more than twice the estimate from Ninan et al. (2019) for an HgCdTe used on the Habitable Zone Planet Finder (HPF) instrument. We expect that the width of the crosshatch pattern varies among detectors or that a tophat function is a poor approximation of the actual subpixel response of this detector. Within NIRCcam detectors, there are large variations in the strength and orientation of crosshatch features.

The crosshatch pattern is wavelength dependent, as seen in Figure 17. Here, the flat field for the A5 detector as well as the 2D Fourier amplitude are shown, which are the same data as used in Figure 16. The throughput variations are largest for short wavelengths (better resolving crystal structures) and smallest for the long wavelengths. This is visible both in the throughput cross section of the flood-illuminated flat field and the Fourier amplitude of the crosshatch as a function of wavelength. We find a steeper wavelength dependence for the crosshatch pattern near 0.9° from horizontal in the frequency domain or 0.9° from vertical in the length domain.

Appendix B Map of Regions

In Section 3.2 we described three regions on the A5 detector in which a dispersed grism image with varying degrees of crosshatch amplitude might be placed. Figure 18 provides a map of these three regions on top of the flat-field image for reference.

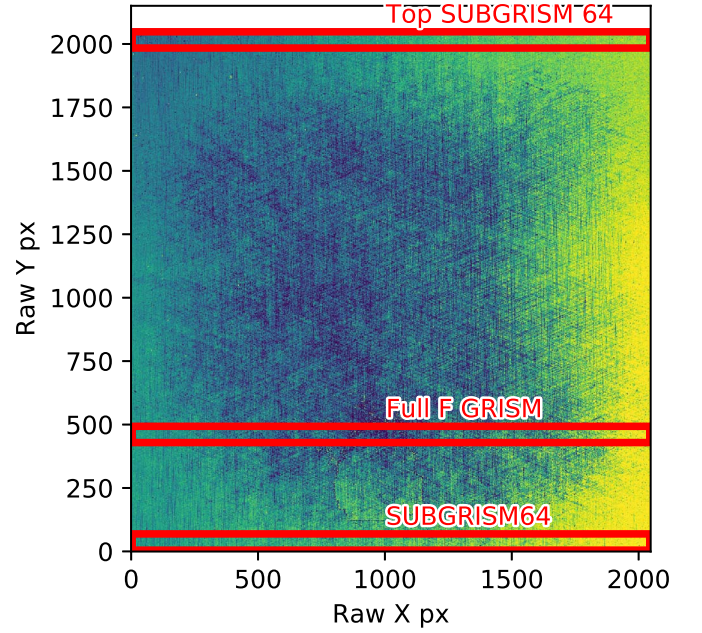


Figure 18. The three regions considered in Section 3.2 are shown in a full-frame flat-field image of the A5 detector.

Appendix C Charge-trapping Schematic

All HgCdTe detectors show a signal after bright illumination even when the illumination is removed; this is called persistence. The physical mechanism that explains persistence is that charges are trapped in the depletion region after illumination by photons as the detector well fills (Smith et al. 2008). This was specifically measured for the NIRCcam detectors in Leisenring et al. (2016). After the reverse-bias voltage is applied (a reset) in an ideal detector, the depletion region will be devoid of mobile charge carriers (electrons and holes). However, in real detectors, some charges are trapped within the depletion region. During a future integration, these charges will be released and shrink the size of the depletion region, recording a spurious signal (data numbers) for that pixel

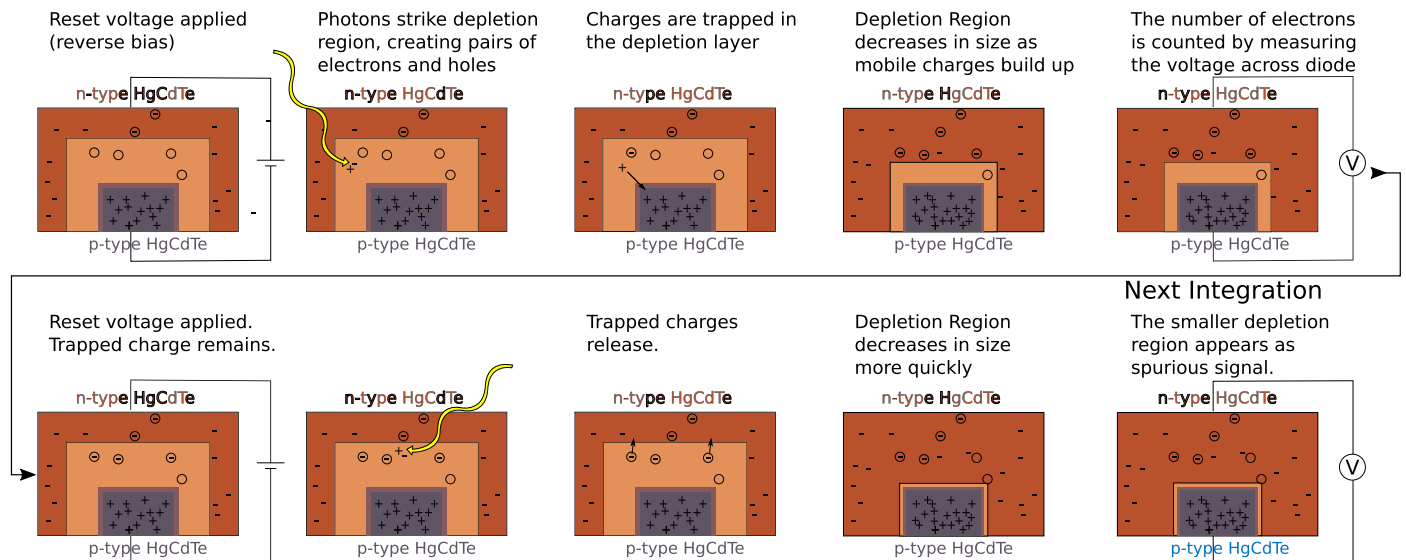


Figure 19. A cartoon schematic of a H2RG detector with charge traps inspired by Smith et al. (2008), Tulloch (2018), and Leisenring et al. (2016). Charged traps for negative and positive carriers (depicted as circles) will capture charge before it can flow to the undepleted regions of the detector. The charge traps do not immediately empty with a detector reset, but instead release at later times; this causes a spurious signal. This spurious signal (the shrinking of the depletion region) causes a voltage change in future integrations that appears as the same as a true signal from incoming photons.

not related to the external illumination of the current integration.

Figure 19 shows a schematic of a detector that has charge traps in the substrate. After illumination by a source, these traps will fill with charge that does not migrate across the depletion layer. When the reset voltage is applied to the detector, the trapped charge remains within the depletion layer. In the second integration in the schematic, charge is released, which reduces the size of the depletion layer faster than the incoming photons would otherwise. The consequence of charge traps is that there is a deficiency in measured charge on the first integration and an excess of measured charge on the second integration. Therefore, a constant astrophysical signal can be measured as a time-varying signal due to the charge-trapping effect. Over timescales longer than the charge-release time, the detector will reach steady state in the functional form of the latent signal, assuming the observed scene stays constant. As discussed in Section 5, the flight NIRCcam detector shows low levels of persistence compared to previous generation detectors used in HST.

ORCID iDs

Everett Schlawin <https://orcid.org/0000-0001-8291-6490>

Jarron Leisenring <https://orcid.org/0000-0002-0834-6140>

Michael W. McElwain <https://orcid.org/0000-0003-0241-8956>

Thomas P. Greene <https://orcid.org/0000-0002-8963-8056>

Marcia Rieke <https://orcid.org/0000-0002-7893-6170>

References

- Albert, L., Doyon, R., Kuzmenko, P. J., et al. 2014, *Proc. SPIE*, 9151, 915158
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
- Baril, M. R., Ward, J., Teeple, D., et al. 2006, *Proc. SPIE*, 6269, 62690Z
- Barstow, J. K., Aigrain, S., Irwin, P. G. J., Kendrew, S., & Fletcher, L. N. 2015, *MNRAS*, 448, 2546
- Barstow, J. K., & Irwin, P. G. J. 2016, *MNRAS*, 461, L92
- Beichman, C., Benneke, B., Knutson, H., et al. 2014, *PASP*, 126, 1134
- Berta, Z. K., Charbonneau, D., Désert, J.-M., et al. 2012, *ApJ*, 747, 35
- Biesiadzinski, T., Lorenzon, W., Newman, R., et al. 2011, *PASP*, 123, 958
- Bradley, L., Sipocz, B., Robitaille, T., et al. 2016, *astropy/photutils*: v0.3, Zenodo, doi:10.5281/zenodo.164986
- Chang, Y., Becker, C. R., Grein, C. H., et al. 2008, *JEMat*, 37, 1171
- Deming, D., Knutson, H., Kammer, J., et al. 2015, *ApJ*, 805, 132
- Doyon, R., Hutchings, J. B., Beaulieu, M., et al. 2012, *Proc. SPIE*, 8442, 84422R
- Evans, T. M., Sing, D. K., Wakeford, H. R., et al. 2016, *ApJL*, 822, L4
- Ferruit, P., Birkmann, S., Böker, T., et al. 2014, *Proc. SPIE*, 9143, 91430A
- Gaudi, B. S., Akeson, R., Anderson, J., et al. 2019, *BAAS*, 51, 211
- Gemain, F., Robin, I. C., Brochen, S., et al. 2012, *JEMat*, 41, 2867
- Greene, T. P., Chu, L., Egami, E., et al. 2016a, *Proc. SPIE*, 9904, 99040E
- Greene, T. P., Kelly, D. M., Stansberry, J., et al. 2017, *JATIS*, 3, 035001
- Greene, T. P., Line, M. R., Montero, C., et al. 2016b, *ApJ*, 817, 17
- Hall, D. N. B. 2006, *ASSL*, 336, 491
- Hardy, T., Baril, M. R., Pazder, J., & Stilburn, J. S. 2008, *Proc. SPIE*, 7021, 70212B
- Hardy, T., Willott, C., & Pazder, J. 2014, *Proc. SPIE*, 9154, 91542D
- Hill, R. J., Malumuth, E., Foltz, R., et al. 2010, *Proc. SPIE*, 7742, 774222
- Howe, A. R., Burrows, A., & Deming, D. 2017, *ApJ*, 835, 96
- Hunter, J. D. 2007, *CSE*, 9, 90
- Ingalls, J. G., Krick, J. E., Carey, S. J., et al. 2016, *AJ*, 152, 44
- Knutson, H. A., Lewis, N., Fortney, J. J., et al. 2012, *ApJ*, 754, 22
- Kopp, G., Lawrence, G., & Rottman, G. 2005, *SoPh*, 230, 129
- Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, *Natur*, 505, 69
- Krissansen-Totton, J., Garland, R., Irwin, P., & Catling, D. C. 2018, *AJ*, 156, 114
- Leisenring, J. 2019, *pynrc*, 0.8.0dev, <https://pynrc.readthedocs.io/en/latest/>
- Leisenring, J. M., Rieke, M., Misselt, K., & Robberto, M. 2016, *Proc. SPIE*, 9915, 99152N
- Lim, P. L., Diaz, R. I., & Laidler, V. 2015, *pysynphot*: Synthetic photometry software package, Astrophysics Source Code Library, ascl:1303.023
- Mahadevan, S., Ramsey, L. W., Terrien, R., et al. 2014, *Proc. SPIE*, 9147, 91471G
- McLean, I. S., Steidel, C. C., Epps, H. W., et al. 2012, *Proc. SPIE*, 8446, 84460J
- Meixner, M., Cooray, A., Leisawitz, D., et al. 2019, arXiv:1912.06213
- Ninan, J. P., Mahadevan, S., Stefansson, G., et al. 2019, arXiv:1903.06614
- Perrin, M. D., Sivaramakrishnan, A., Lajoie, C.-P., et al. 2014, *Proc. SPIE*, 9143, 91433X
- Pidhorodetska, D., Fauchez, T., Villanueva, G., Domagal-Goldman, S., & Kopparapu, R. K. 2020, *ApJL*, 898, L33
- Rackham, B. V., Apai, D., & Giampapa, M. S. 2018, *ApJ*, 853, 122
- Rackham, B. V., Apai, D., & Giampapa, M. S. 2019, *AJ*, 157, 96
- Rieke, M. J., Kelly, D., & Horner, S. 2005, *Proc. SPIE*, 5904, 1
- Schlawn, E., Greene, T. P., Line, M., Fortney, J. J., & Rieke, M. 2018, *AJ*, 156, 40

- Schlawin, E., Herter, T. L., Henderson, C., et al. 2014, *Proc. SPIE*, **9147**, [91472H](#)
- Schlawin, E., Leisenring, J., Misselt, K., et al. 2020, *AJ*, **160**, 231
- Schlawin, E., Rieke, M., Leisenring, J., et al. 2017, *PASP*, **129**, 015001
- Shapiro, C., Huff, E., & Smith, R. 2018, *Proc. SPIE*, **10709**, [1070936](#)
- Smith, R. M., Zavodny, M., Rahmer, G., & Bonati, M. 2008, *Proc. SPIE*, **7021**, [70210J](#)
- Suissa, G., Mandell, A. M., Wolf, E. T., et al. 2020, *ApJ*, **891**, 58
- The LUVOIR Team 2019, arXiv:[1912.06219](#)
- Tulloch, S. 2018, arXiv:[1807.05217](#)
- van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *CSE*, **13**, 22
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, **17**, 261
- Wakeford, H. R., Sing, D. K., Kataria, T., et al. 2017, *Sci*, **356**, 628
- Zhou, Y., Apai, D., Lew, B. W. P., & Schneider, G. 2017, *AJ*, **153**, 243