ESTEC

European Space Research
and Technology Centre
Keplerlaan 1
2201 AZ Noordwijk
The Netherlands
www.esa.int

PERFORMANCE REPORT

Mimicking the observation of an exoplanet transit: analysis of a 3-hour exposure acquired at ISIM-CV3 and assessment of systematic effects

Prepared by G. Giardino, S. Birkmann, P. Ferruit, B. Dorner, B. Rauscher

Reference ESA-JWST-SCI-NRS-RP-2016-012
Issue/Revision 1.0
Date of Issue April 10, 2017
Status Released

APPROVAL

Title **Mimicking the observation of an exoplanet transit**: analysis of a 3-hour exposure acquired at ISIM-CV3 and assessment of systematic effects	
Issue Number 1	**Revision Number** 0
Author G. Giardino, S. Birkmann, P. Ferruit, B. Dorner, B. Rauscher	**Date** April 10, 2017
Approved by	**Date of Approval**
M.Sirianni	April 10, 2017

CHANGE LOG

Reason for change	Issue Nr.	Revision Number	Date

CHANGE RECORD

Issue Number 1	Revision Number 0		
Reason for change	**Date**	**Pages**	**Paragraph(s)**

DISTRIBUTION

Name/Organisational Unit
SCI-ODJ
ESTEC

Contents

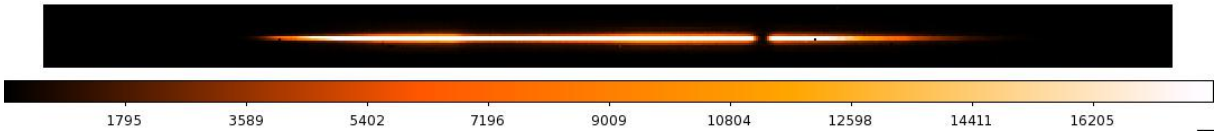

Figure 1: Count-rate image of the test source used during ISIM cryogenic campaign to mimic the observation of a bright source in NIRSpec BOTS mode. The light is dispersed by the PRISM. Units are Counts/s.

1 INTRODUCTION

There are a number of instrumental effects and calibration uncertainties that can increase the actual noise level of a NIRSpec observation of an exoplanet transit above the noise floor of the instrument as determined by the detector read noise and the signal shot noise. To assess the stability of our system and the impact of systematic effects for this instrument mode we performed a dedicated test during ISIM-CV3, i.e. the cryo-vacuum campaign of the JWST Integrated Science Instrument Module (ISIM), that took place at Goddard Space Flight Center during the winter 2015-2016 (Kimble et al. 2016).

2 NIRSPEC OBSERVATIONS OF A BRIGHT SOURCE DURING ISIM-CV3

To evaluate the response of NIRSpec to an exposure of a bright target over a time scale typical of an exoplanet transit, we acquired a $\sim$3 hour-long exposure of a bright source. The exposure was taken with NIRSpec at the operating temperature of $\sim$ 40 K, using the PRISM, and consists of 12,000 integrations of three groups (of one frame) each, in detector window mode, with subarray size 512$\times$ 32, corresponding to a frame read-out time of $t_g = 0.226$ s, i.e. to an effective integration time of $t_{\rm int} = (n_g - 1)t_g = 0.45$ s per integration. The count-rate image of the dispersed source in one integration is shown in Fig. 1.

The light-source consisted of an LED with a PSF size similar to that of JWST, with no requirements in terms of flux stability. The test set-up was such that the source underwent spatial jitter and a drift of more than 30 mas over the 3-hour exposure. We traced the source displacement over all the integrations measuring the baricenter position of the spectral trace, for the source movements in cross-dispersion direction (i.e. y-direction), and the center of a prominent carbon-dioxide absorption feature at 4.2 μm, for displacements in dispersion direction (i.e. x-direction). These measurements are shown in Fig. 2, note that the amplitude of the drift is more than four times the pointing stability requirement for JWST, which has to be better than 7 mas (1σ) over 10,000 s.

The first step of NIRSpec data analysis consists in producing count-rate images from the up-the-ramp individual integrations, so 12,000 count rate images were produced from the raw data by our standard ramp-to-slopes pipeline. We do not expect the source luminosity to be stable over the exposure time-scale and it is clear that the STD over the 12,000 integrations of the illuminated pixels in the count-rare images is dominated by the overall total counts variations as traced by the white light curve, that is the total signal in the detector subarray for each integration, as shown in the top-panel of Fig. 4 (normalized to the signal average over the all exposure).

Therefore, after having produced the count-rate image for each integration, assigned a wavelength to each pixel in the image, and spatially collapsed the 2D-data set, the spectrum so derived for each integration was normalized by the relative white-light curve value (as given in Fig. 4). After this correction, it is apparent that the STD of each pixel over the 12,000 spectra is dominated by the contribution by the correlated noise from the detector read-out electronics. This noise component manifests itself in low level cross-dispersion stripes and bands visible in the count-rate images. The presence of these fluctuations in Teledyne HAWAII detectors is well known and the special read-out mode NRSIRS2, in conjunction with a Wiener filtering algorithm named IRS^2 (to be applied by the ramp-to-slope processing pipeline), has been developed to minimize the impact of this $1/f$-noise source in full-frame mode – see Moseley et al. (2010); Rauscher et al. (2012). IRS^2 exploits the signal recorded by the detector reference pixels and in each pixel reference channel to track the temporal fluctuations

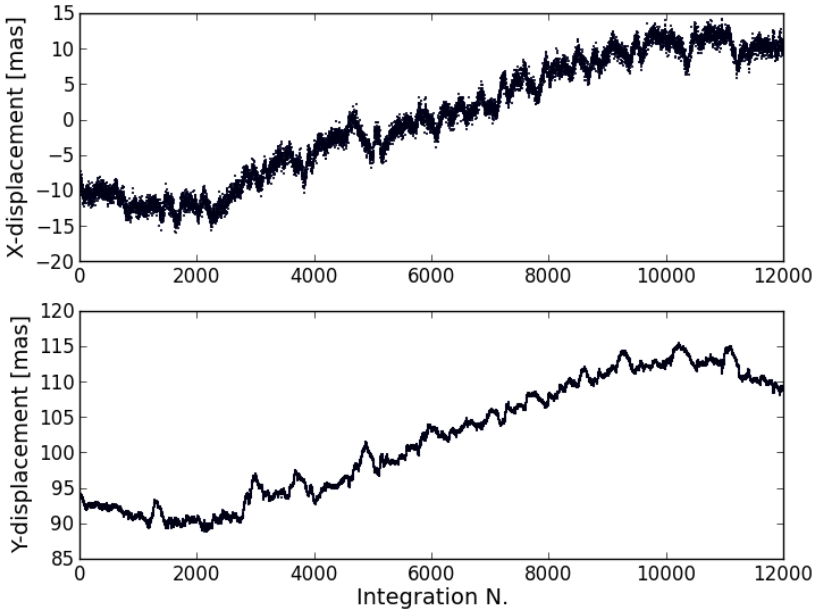

Figure 2: Jitter and drift movement of the LED source that was used to mimic the observation with NIRSpec of a bright star for a 1.5 hour-long exposure during the cryogenic campaign. Note that the pointing stability of JWST is required to be better than 7 mas (1σ) over 10,000 s.

of the read-out electronics and filter it out.

Reference pixels are not available when using detector subarrays, however in the case of the observation of a point source with a subarray of size 32 pixel in spatial direction, there is a significant number (≥ 10) of pixels in each column that are not illuminated and can be used to track the $1/f$-noise components. To filter out this noise components from the count-rate images we used a simple algorithm based on computing the median count-rate of the dark pixels in a column (i.e. all pixels with projected spatial coordinate $y > |0.5|S_y$, where S_y is the slit size in spatial direction), and subtracting this value from the count-rate of all the pixels in that column. The operation is repeated for each column on every count-rate images in the 12,000 sequence.

Despite its simplicity, this "destriping" algorithm is very effective in suppressing the $1/f$-noise. Our ramp-to-slope pipeline provides, together with the count-rate image of each integration, the statistical variance for each pixel, computed from that pixel count-rate and the variance of the read-noise[1] using the expression given Eq. 1 of Rauscher et al. (2007, 2010). Starting from this value we can derive the expected STD of our signal (due to statistical fluctuations only), in each pixel of the collapsed subarray for a given integration time. As illustrated in Fig. 3, before destriping, the measured STD of the total electron count in each integration is 3-4 times the expected value, depending on the incident signal (panel *b*), while after filtering out the $1/f$-noise component, the observed STD closely matches the ideal level in the absence of signal (see panel *c* for $\lambda < \ 1.0\ \mu$m and $\lambda > 4.8\ \mu$m). Nevertheless, a large excess of noise remains in the pixels with signal and in particular around the steepest spectral gradients.

This excess noise is due to the interplay between the under-sampled PSF, whose FWHM is comparable to the pixel size, and the spatial movements of the source. To correct for this effect, after flat-fielding, we performed a bi-linear fit of each pixel signal (operating on the 2d-spectral trace) as a function of the source position in x and y in each integration (as plotted in Fig. 2) and subtracted the fitted value from the pixel value. As shown in Panel *d* of Fig. 3, this correction brings the measured STD over 9,000 integration in line with expected value,

[1] as measured from CDS noise in dark exposures

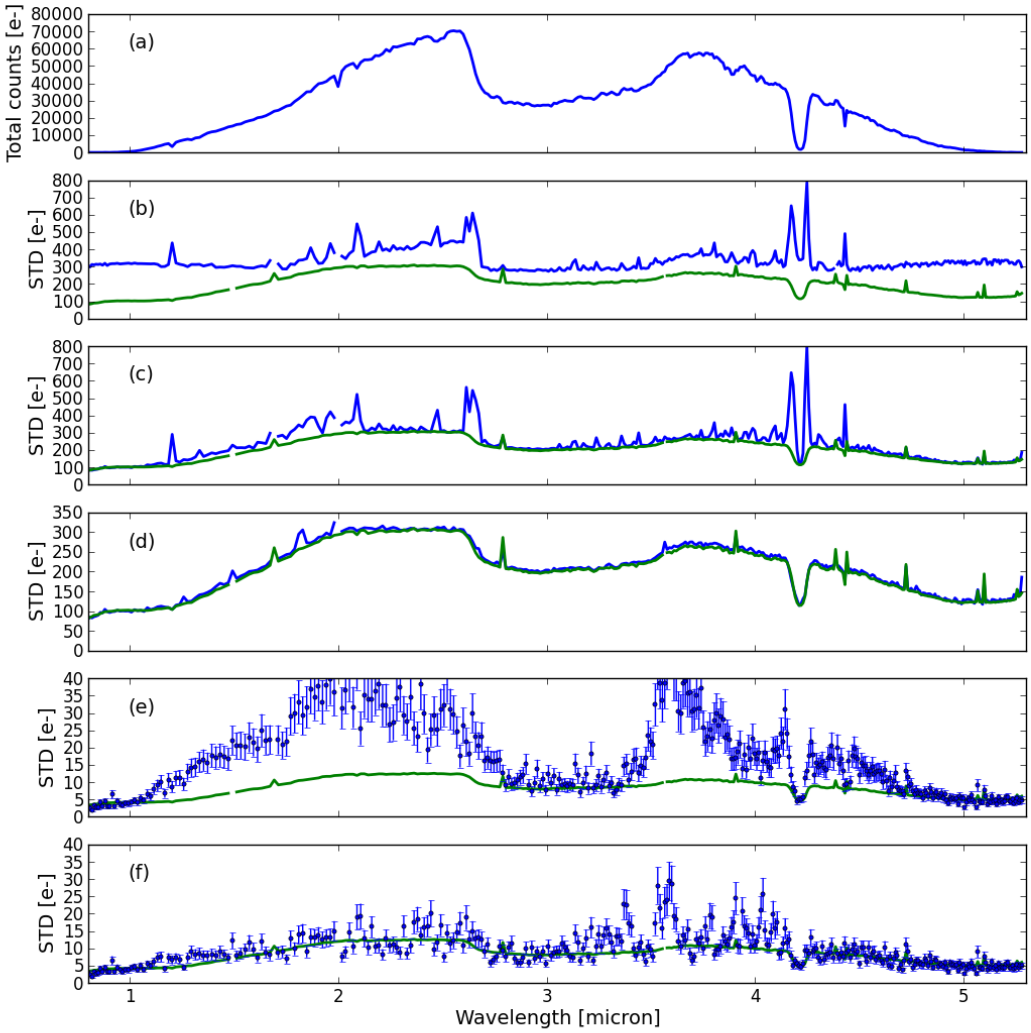

Figure 3: The (spatially) collapsed spectrum of the bright point source observed with NIRSpec during ISIM-CV3 (Panel *a*) and STD deviation (all other panels) of each pixel in the collapsed trace. (Panel *b* - *f*): in blue the measured STD over the last 9,000 integrations of the 3-hour exposure, in green the expected level of STD from the statistical noise sources of an ideal instrument. Panel *b* the STD of the raw data, white-light curve normalized; Panel *c* STD as *b*, but after image destriping ($1/f$-noise removal using un-illuminated pixels) and having flat-fielded; Panel *d* as *c*, but after having de-correlated each pixel signal by the source spatial shifts in x and y direction; Panel *e* as *d*, but here the STD is computed over 15 (temporal bins) of 600 integrations each allowing us to reach higher sensitivity - the error bar on the STD is also given. Panel *f* as *e* but rather than normalizing by the white light-curve here the data were corrected by a 5-color light curve (see text).

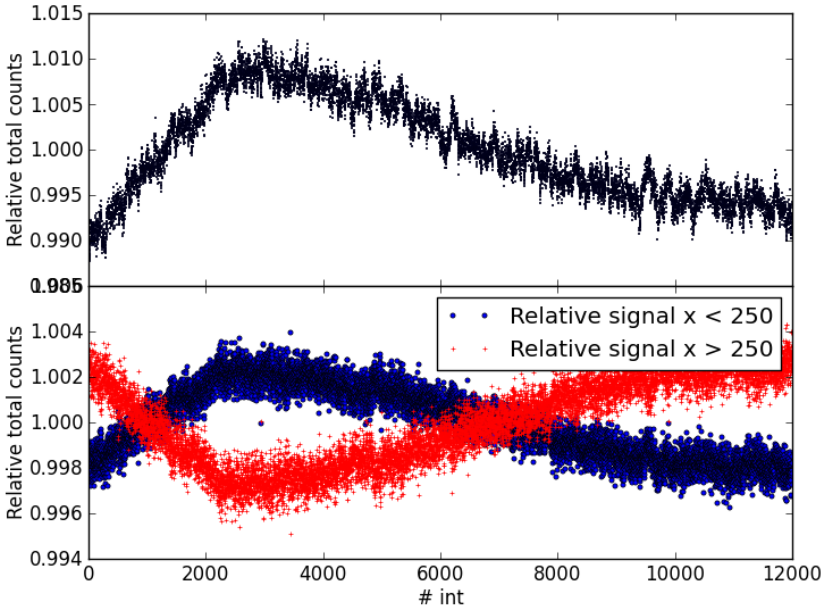

Figure 4: Top Panel: the relative white-light curve obtained by summing up the signal over all pixels in the trace in each integration, and normalizing by the average value over all integration. Bottom Panel: the signal measured in each integration by summing up the signal over all pixels in the trace with coordinate in dispersion direction, x, < 250 (blue dots) and $x > 250$ (red-cross) normalized by the white light curve, so that the values trace the source flux variation in the blue and red part of the spectrum relative to the total flux.

i.e. between 200 and 300 e^- for signal greater than 30,000 e^-. The first 3,000 integration of the exposure were not included in this analysis because the test source had just been switched on and it exhibited a more prominent (wavelength dependent) drift that the white-light-curve normalization is unable to correct for (see Fig. 4 and discussion below).

In order to reach higher sensitivity we binned the 9,000 integrations into 15 bins of 600 each, corresponding to bins of 4.5 minutes. For this integration time the expected STD of a signal between 30,000−70,000 e^- is 8−12 e^-, corresponding to a noise-to-signal ration (N/S) of less than 300 ppm per pixel, however, as shown in Panel *e* of Fig. 3, the measured STD are in some wavelength ranges in excess of 2 times these values. While we cannot exclude that other instrumental effects are causing (part of) the excess noise at this level of sensitivity, the STD measured here is very likely dominated by the flux instability of the light source.

We evaluated the relative variations of the blue and red part of the spectrum by taking the total signal in the blue half of the subarray (i.e. $x < 250$) and in the red-part ($x > 250$); as shown in Fig. 4 (bottom-panel), from integration 3,000 onwards, the relative amplitude of the blue and red light curve are monotonically varying by $\sim$ 0.6 % relative to the total flux drift. There are various reasons to be confident that here we are genuinely looking at the lamp flux variations as opposed to trends in the detector response. After filtering out the $1/f$-component from the detector noise, the STD of the dark pixels is fully consistent with the expected value as based on the read noise and dark current, excluding intrinsic variations, of the additive type, of the detector response. Beside the PRISM data discussed here, we also acquired a 3-hour exposure with grating G235H dispersing the light from the same source onto both NIRSpec detector (operated in window-mode with window of size 2048×32). Cross correlating the white light curves from the data from SCA491 to SCA492, we find highly correlated data with a time lag of $\sim$ 2 sec, corresponding to the time-gap in the start of the exposure in the two detectors. This excludes fluctuation of a multiplicative type (i.e. proportional to the flux) in each

individual detector or response variations synchronous in both detectors such as the type that could be possibly be driven by drifts in operating temperature.

To test the assumption that the excess noise seen here is due to the color changes of the lamp, we subdivided the collapsed spectral data into 5 bins of 40 pixels each and computed the total signal in each bin for each integration (thereby producing 5 light curves of 9,000 points). Then, before computing the (temporally binned) STD for each pixel, the time-series for each pixel was normalized by a light curve as derived by interpolating between the two spectrally adjacent light curves, integration per integration. The values of the STD for each pixel derived in this way are shown in the bottom panel of Fig. 3, where it can be seen that the measured STD is now better in line with the expected value.

The 5-color-light curve normalization allows us to reach the expected sensitivity level of 200−300 ppm, for total counts of 70,000−30,000 e^-, although higher residual noise is left in the wavelength interval 3.4−4 μm, where the 5-color correction does not perform so well. It is clear from this analysis that the poor flux stability of the test source limits significantly our ability to probe noise of systematic origin in NIRSpec data using this time series. Nevertheless, the fact that the residual noise above the floor level can be reduced here by a simple tracking of the source flux in 5 wavelength bins, give us confidence in the stability of NIRSpec over the three-hour time scale typical of a transit observation and the ability of the instrument to reach the noise-floor of 200 ppm in less than 5 minutes of integration. After correction of the source's color drifts, the median of the difference between ideal and measured STD of each pixel over $\sim$ 200 pixels with signal greater than 30,000 e^-, is $\sim$ 2 e^-, indicating that NIRSpec performance in terms of noise level is within 20% of the ideal value, for integration times of 5 minutes.

3 OTHER SOURCES OF SYSTEMATIC NOISE

The analysis above shows that when observing a bright target for a time-scale typical of an exoplanet transit, the dominant source of systematic noise in the NIRSpec data are the $1/f$-noise component from the detector electronics and, in conjunction with significant pointing jitter and drift, the signal variations in individual pixels due to an under-sampled PSF and the source movement. The analysis also identified the processing steps that effectively reduce the impact of these two noise components, resulting in data points with noise level within 20% of the noise-floor of the ideal instrument, for integration times up to $\sim$ 5 minutes.

Among other possible sources of systematic noise that could limit the instrument performance in the observation of exoplanets the combination of pointing jitter and drift with intrapixel sensitivity variations and variable slit- and diffraction-losses had been identified as possibly the most significant. The effect of pointing jitter and drift combined with variable slit- and diffraction-losses was evaluated using NIRSpec Instrument Performance Simulator (IPS) – see Piqueras et al. (2008); Piquéras et al. (2010). Fig. 5 provides a map of combined flux losses due to the telescope PSF being truncated and diffracted by the slit, for the case of grating G235H. The typical level of intensity loss is $\sim$ 5%, with peak-to-valley variations of 0.03% across the grid.

The pointing stability of JWST is required to be better than 7 mas (1σ) over 10,000 s, this however include random jitter over time scale of seconds and longer term pointing drift that would appear as systematic signal drift, but would not directly contribute to the noise, in a series of integrations of a fraction of a second each. Thus, to estimate the noise contribution in a multi-integration exposures of exoplanets, we adopt the pointing requirements for NIRCam for time scales of a few seconds, which is 3.5 mas in line of sight and roll stability, corresponding to a jitter radius (1σ) of 5 mas. The relative noise in one exposure introduced by the telescope movement in random direction is then the RMS of the throughput of S1600A on a circle of 5 mas-radius relative to the value at its center, at the different wavelengths. The mean relative noise in an exposure of a point source over the wavelength range of the different dispersers are summarized in Table 1, where one can see that this noise contribution is expected be below 25 ppm for all modes.

Of higher concern could be the interplay between the telescope pointing instability and intrapixel sensitivity variations. Data on characterizing the intrapixel sensitivity for these pixels are very limited. Hardy et al. (2008)

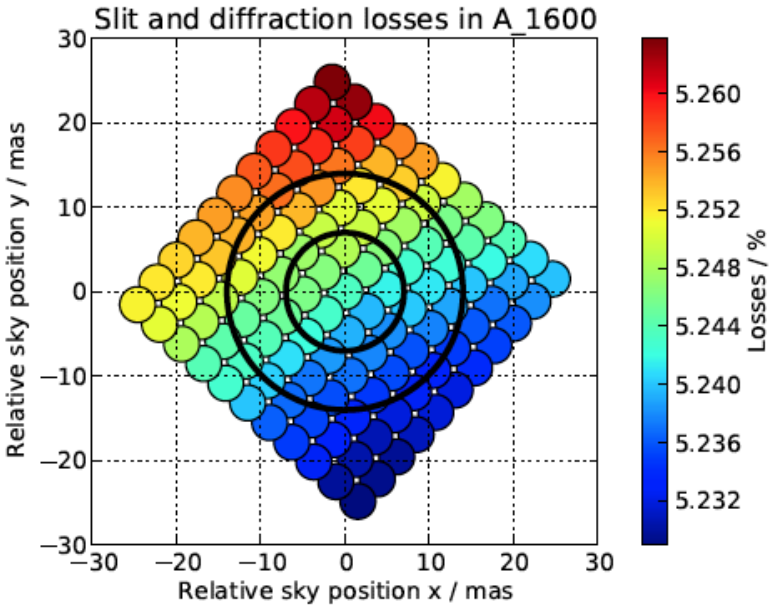

Figure 5: Map of combined slit- and diffraction-losses in the central part of S1600A for the grating G235H at 2.45 μm. The black circles have radius of 7 and 14 mas, corresponding to 1 and 2σ of the allowed pointing drift over 10,000 s, as specified in JWST requirements.

Table 1: Mean relative noise in an exposure of a point source in S1600A due to slit- and diffraction losses, calculated for a random jitter of 5 mas, for the different bands.

	Relative Noise $[10^{-5}]$		
Resolution	band I	band II	band III
R100		2.2 (over 0.6−5μm)	
R1000	1.7	2.0	2.5
R2700	1.6	1.8	2.3

have characterized the intra-pixel sensitivity variations of an older-generation H2RG detector with a resolution of approximately 4 μm (0.22 pixel) in the wavelength range between 0.65 and 2.2 μm, for a sample of 64 pixels. The intra-pixel response is overall smooth with variations across the pixels at the level of 2-3% RMS, which is down to the measurement accuracy level, however there appear to be some localized sensitivity dips and long linear defects affecting roughly 10% of the pixels examined. Hardy et al. (2014) completed a similar set of measurements on the new design H2RG detector similar to the two detectors installed in NIRSpec. They find intra-pixel sensitivity variations with similar or somewhat smaller amplitude than those from the earlier set of measurements.

Using the intrapixel sensitivity maps from Hardy et al. (2008) convolved with the NIRSpec PSF at 2 m, we evaluated the interplay between intra-pixel sensitivity and telescope pointing variations (RMS $<$ 7 mas), by drawing 1000 random samples around each sub-pixels within an area with 1σ radius of 6.5 mas and then computing the STD over the 1000 samples. We found that the typical STD of this noise component could be as high as 400 ppm of the signal amplitude, if uncorrected for (and if the small sample of pixels available is reasonably representative). In this respect, it is very encouraging to see from our analysis of the CV3 time-series that a simple bi-linear fit of the source spatial jitters and long term (large) drift that corrects for the PSF under-sampling (and therefore, potentially, also correct for intrapixel sensitivity variations) allows us to reach N/S

below 200 ppm. Note that the 5-color light curve division applied to our data set would not be able to correct for the effect of intrapixel sensitivity variations as this noise component acts strictly at the level of individual pixels.

An other instrumental systematic behavior that could affect NIRSpec observations of bright targets and that cannot be probed with the CV3 data (due to the light source instability) is detector persistence. Persistence refers to the detector's memory of previous exposures. It is one manifestation of charge trapping. Charge traps are localized defects in the detector that present unintentional potential wells that can temporarily capture mobile charges. If a moving charge falls into a trap before being integrated, it will become stuck and not appear as part of the signal in that exposure. When the trap decays and releases the charge at a later time, it becomes mobile again. Most of the time, the newly released charge will integrate along with the photocurrent from the current exposure to appear as persistence, or memory from the previous exposure.

The effect of persistence when observing bright host stars with *Spitzer* or HST has been reported by several authors – e.g. Charbonneau et al. (2008); Agol et al. (2010); Berta et al. (2012); Wakeford et al. (2016) – and is referred to as the "ramp" or "hook" effect, because of the shape of the light curve over a set of contiguous exposures: sharply rising and quickly saturating ramp-like features, as, for example, can be clearly seen in the observation with WFC3 in stare mode of GJ1214 by Berta et al. (2012). The "charge trapping" model leads to exponential ramps when observing bright sources as the excess dark current increases sharply at first but slows its increase as the population of charge traps begins to approach steady state.

NIRSpec detectors are also affected by persistence, and as exoplanet observation will be conducted without dithers within slit S1600A, we can expect a detector behavior qualitatively similar to that of WFC3 when used in stare mode, that is an initially rising light curve during the first integrations of a bright targets followed by an equilibrium state in which flux levels are highly repeatable. Quantitatively speaking, however, the effect is expected to be significantly smaller for the much newer NIRSpec detectors.

Rauscher et al. (2014) have characterized the persistence behavior of the two NIRSpec detectors. They measured persistence in dark exposures acquired soon after illuminating the detector with flat-field light at a fluence of 10 $\times$ the full-well level and found that the persistence of NIRSpec flight SCAs was well approximated by the power law, $f(t) = f_0(t/1\ \mathrm{s})^{-1}$ with f_0 equal to 29 and 42 $\mathrm{e^{-}s^{-1}pix^{-1}}$, respectively for SCA491 and SCA492. In comparison, McCullough & Deustua (2010) used data from persistence tests of WFC3 IR flight array acquired at GSFC DCL by Hilbert (2008) using 100-times full well overexposed flat field illumination to show that the persistence decay in WFC3 detector follows a power law of the form 0.156 $(t/1\ \mathrm{hour})^{-1}$. This corresponds to a latent signal of $\sim 5.6\ \mathrm{e^{-}s^{-1}pix^{-1}}$ after 100 s since the end of the over-exposure, compared to a corresponding signal from NIRSpec's detectors of less than 0.5 $\mathrm{e^{-}s^{-1}pix^{-1}}$. Note that in the case of the NIRSpec detector arrays, using 10-times or 100-times fluence level in the over-exposure does not impact the latency level as persistent images do not get stronger over $\sim 2-3\times$ saturation. Also, persistence does not depend on the wavelength of the illumination, but there appears to be a dependency with illumination time, that is the longer the over-illumination time, the stronger the persistence image (Rauscher et al. 2014).

We also note that in the case of NIRSpec up to 65,535 integrations can be taken during one (hours-)long exposure without breaks, because, unlike for HST's orbit, JWST's position at L2 enables its instruments to perform continuous observations of a target for many hours or even days. This will allow observers to minimize the time the detectors will be in a non-equilibrium state with respect to the effect of charge trapping for transit, occultation and phase curve observations.

4 REFERENCES

Agol, E., Cowan, N. B., Knutson, H. A., Deming, D., et al. 2010, ApJ, 721, 1861

Berta, Z. K., Charbonneau, D., Désert, J.-M., et al. 2012, ApJ, 747, 35

Charbonneau, D., Knutson, H. A., Barman, T., et al. 2008, ApJ, 686, 1341

Hardy, T., Baril, M. R., Pazder, J., & Stilburn, J. S. 2008, in Proc. SPIE, Vol. 7021, High Energy, Optical, and Infrared Detectors for Astronomy III, 70212B

Hardy, T., Willot, C., & Pazder, J. 2014, in Proc. SPIE, Vol. 9154, High Energy, Optical, and Infrared Detectors for Astronomy VI, 91542D

Hilbert, B. 2008, WFC3 TV3 Testing: IR Channel Thermal Background Signal, Tech. rep.

Kimble, R. A., Vila, M. B., Van Campen, J. M., et al. 2016, in Proc. SPIE, Vol. 9904, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 990408

McCullough, P. & Deustua, S. 2010, WFC3 TV3 Testing: IR Persistence, Tech. rep.

Moseley, S. H., Arendt, R. G., Fixsen, D. J., et al. 2010, in Proc. SPIE, Vol. 7742, High Energy, Optical, and Infrared Detectors for Astronomy IV, 77421B

Piqueras, L., Legay, P.-J., Legros, E., et al. 2008, in Proc. SPIE, Vol. 7017, Modeling, Systems Engineering, and Project Management for Astronomy III, 70170Z

Piquéras, L., Legros, E., Pons, A., et al. 2010, in Proc. SPIE, Vol. 7738, Modeling, Systems Engineering, and Project Management for Astronomy IV, 773812

Rauscher, B. J., Arendt, R. G., Fixsen, D. J., et al. 2012, in Proc. SPIE, Vol. 8453, High Energy, Optical, and Infrared Detectors for Astronomy V, 84531F

Rauscher, B. J., Boehm, N., Cagiano, S., et al. 2014, PASP, 126, 739

Rauscher, B. J., Fox, O., Ferruit, P., et al. 2007, PASP, 119, 768

Rauscher, B. J., Fox, O., Ferruit, P., et al. 2010, PASP, 122, 1254 (Erratum)

Wakeford, H. R., Sing, D. K., Evans, T., Deming, D., et al. 2016, ApJ, 819, 10