

Data Post-Processing and Algorithm Development for the WFIRST Coronagraph: FY16 Interim Report

Marie Ygouf, Neil T. Zimmerman,
Laurent Pueyo, Marshall D. Perrin, Rémi Soummer^(PI)

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



Abstract

Advanced post-processing techniques have the potential to broaden the landscape of scientific targets accessible to the WFIRST Coronagraph Instrument (CGI). The core objective is to bridge the order-of-magnitude contrast gap between the baseline CGI raw starlight suppression ($< 10^{-8}$) and the $< 10^{-9}$ level needed for WFIRST to characterize a diverse sample of exoplanets in reflected starlight. In continuation of our previous efforts to advance the maturity of post-processing algorithms for the WFIRST CGI, to date in FY16 we have completed the following tasks:

- (i) We analyzed a new set of Hybrid Lyot Coronagraph (HLC) data simulations, and compared the post-processed contrast gain for various raw data assumptions (photon noise, low-order

wavefront control, and DM stroke quantization) and observing scenarios (Reference Differential Imaging versus Angular Differential Imaging / roll subtraction). We found that ADI is more robust to both speckle and photon noise than RDI, in some cases offering a threefold gain in sensitivity.

- (ii) We investigated the impact of spatial sampling on post-processed sensitivity, in the context of design trade studies for the Integral Field Spectrograph (IFS) component of the instrument. Our preliminary results suggest that the spatial sampling can be halved from the baseline sampling rate (~ 4 lenslets per λ/D) without any degradation in final contrast, thereby reducing the integration time required for spectroscopic characterization.
- (iii) We analyzed new Shaped Pupil Coronagraph (SPC) data acquired at JPL. With this data set—likely the most flight-like yet acquired on the SPC testbed in terms of wavefront stability—our mock RDI scenario demonstrates a broadband contrast gain factor of $14\times$ to reach a 1σ contrast floor of 10^{-9} . In the narrowband version of the same data, the exceptional aberration stability enables a contrast gain of $35\times$.

Contents

1	Introduction	4
2	Simulated data	6
2.1	Description of simulations	6
2.2	Wavefront aberration stability assessment	6
2.3	Comparison of classical PSF subtraction and KLIP for the RDI scenario	8
2.4	Comparison of RDI and ADI observing strategies	13
3	Impact of spatial sampling on the post-processing performance	17
4	Laboratory data	22
4.1	Data description	22
4.2	Results of mock RDI scenario	24
4.3	Residual noise analysis	26
4.4	RDI with simulated planets	29
5	Conclusions	30
6	Acknowledgments	32
7	Acronym list	32

1 Introduction

After a successful mission concept review in December 2015 and the start of mission phase A in February 2016, the WFIRST [1] coronagraph instrument (CGI) [2] is under development to be the first actively controlled starlight suppression system in space [3]. The combination of an Occulting Mask Coronagraph (OMC) with deformable mirrors (DM) and low-order wavefront sensing and control (LOWFSC) [4] will be used to achieve raw contrasts below 10^{-8} at angular separations $3\lambda/D$ [5]. On top of numerous hardware innovations, advanced data post-processing techniques will be needed to detect exoplanets in reflected starlight, extract their astrometry and spectra, and image faint circumstellar debris [6].

Our group has previously conducted data processing and algorithm development using simulated and laboratory WFIRST coronagraph data [7, 8, 9, 10]. Those efforts produced the first systematic studies of coronagraph post-processing techniques in the performance regime of the WFIRST CGI. We investigated both coronagraph designs comprising the baseline OMC, the Hybrid Lyot Coronagraph (HLC) [11] and the Shaped Pupil Coronagraph (SPC) [12], under various assumptions of wavefront stability, spacecraft+telescope+instrument thermal models [13], and observing scenario.

In Section 2, we build on the prior investigation of HLC data simulation to compare two observing strategies, reference star differential imaging (RDI) and angular differential imaging (ADI) [14]. The RDI strategy consists of observing one reference star to calibrate the coronagraphic PSF and subtracting it from the image of the target star that contains the astrophysical signal. RDI is the current baseline for the WFIRST coronagraph [1]. The ADI strategy consists of repeated observations of the target, introducing diversity between instrumental artifacts and astrophysical signal by rolling the spacecraft at several different orientations. Even though ADI has been an effective strategy for science results on both coronagraphic and non-coronagraphic Hubble Space Telescope (HST) images [15, 16, 17, 18], there is not yet a clear plan in place to implement ADI for the WFIRST coronagraph.

The Integral Field Spectrograph (IFS) mode of the WFIRST CGI is designed to detect molecular absorption features in the reflected spectra of giant exoplanet atmospheres. A key tradeoff parameter in the IFS design is the spatial sampling defined by the lenslet array. Increasing the lenslet size reduces the integration time needed to reach a given photon count in the spectrally dispersed signal corresponding to an individual spatial sample, at the expense of spatial resolution. In the baseline design of the WFIRST IFS, the lenslet pitch is $0.25 \lambda_0/D$ in the bluest bandpass [1]. In view of potentially relaxing the spatial sampling of the current design in order to reduce the integration times required for spectroscopic characterization, we used simulated HLC images to examine what effect the spatial sampling has on the post-processed contrast floor.

In our most recent report [10], we noted that the post-processing performance on SPC labo-

ratory data was limited by the changing deformable mirror (DM) actuator settings between each exposure (a byproduct of the iterative “dark hole” control loop that does not represent a flight-like scientific observing scenario) and the high level of uncorrelated noise relative to the speckle brightness, partly due to short integration times. These effects limited our ability to test the functionality of speckle subtraction on the SPC data, despite the fact that raw contrast were near the baseline 10^{-8} . In a new SPC data set taken by JPL in December 2015, the experiment parameters were changed so that the DM settings remained fixed throughout a sequence of many exposures, meanwhile allowing the testbed to drift due to ambient environment effects. We show in Section 4 that with this new lab data we can now demonstrate order of magnitude RDI contrast gains reaching down to the 10^{-9} regime.

2 Simulated data

2.1 Description of simulations

We examined data sets simulated by the Jet Propulsion Laboratory (JPL) of CGI Observing Scenario 5 (OS5), with and without photon counting noise, using a diffractive model of the coronagraph [19, 13]. The temporal variability of the wavefront at the entrance of the CGI was derived from a comprehensive thermal model of the spacecraft generated by the Goddard Space Flight Center (GSFC). Throughout this section, we label the data versions with and without photon noise *noisy* and *noiseless*, respectively. On top of the inclusion/exclusion of photon noise, these simulations include data versions with and without low-order wavefront control (LOWFC). Further, of the two data sets with LOWFC, one was simulated without digital to analog converter (DAC) stroke digitization (unlimited actuator stroke resolution); the second assumed a DAC with 16-bit stroke resolution and randomization of the least significant bit to avoid spatially-correlated residual wavefront patterns. Successive observations of two stars were simulated in OS5: (1) first a bright star is targeted for dark hole generation (β UMa; A1IV, $V=2.37$); (2) this is followed by observations of the science target (47 UMa; G1V, $V=5.04$) at two different spacecraft orientations ($+13^\circ$ and -13° off nominal roll). There are 30 short exposure images of 1000 sec each for β UMa and 50 for 47 UMa in each roll. Table 1 summarizes this observing scenario. These simulated data were provided for one of the baseline WFIRST coronagraphs, the Hybrid Lyot Coronagraph (HLC), in the broadband imaging channel. Synthetic planets were added to the 47 UMa images in all data sets. The nominal spatial sampling of images is $0.35\lambda/D$ per pixel. Figure 1 shows the combined long exposure images for each simulated data sets. Table 2 summarizes the different files of the OS5 data sets¹.

Star	Exposition Time
β UMa at roll $+13^\circ$	30 short exposures of 1000 sec each (30000 sec total)
47 UMa at roll $+13^\circ$	50 short exposures of 1000 sec each (50000 sec total)
47 UMa at roll -13°	50 short exposures of 1000 sec each (50000 sec total)

Table 1: **OS5 observing scenario.**

2.2 Wavefront aberration stability assessment

The overall performance of post-processing depends strongly on the stability of the wavefront errors [7]. The temporal profile of correlation coefficients between exposures is a simple way to assess the stability of the aberrations [8]. We computed the correlation coefficients of the first or last (according to the case) exposure of the 47 UMa (target) data cube at $+13^\circ$ with each exposure

¹These files as well as their complete description can be downloaded on the WFIRST pages of the Infrared Processing and Analysis Center (IPAC) website: https://wfirst.ipac.caltech.edu/sims/Coronagraph_public_images.html

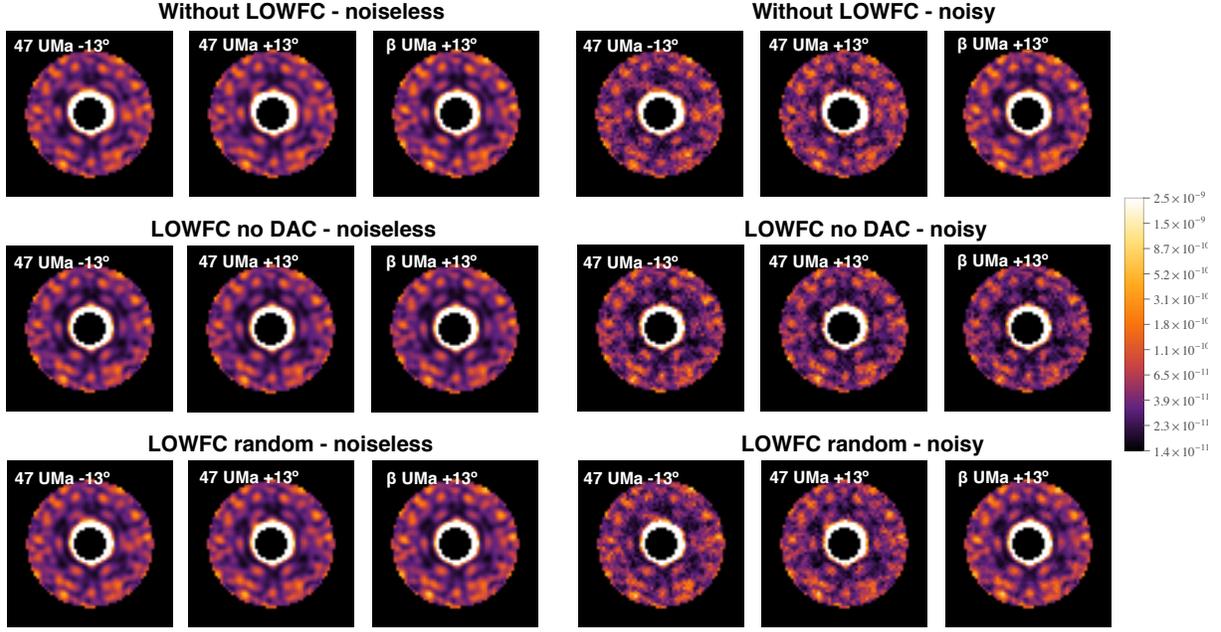


Figure 1: **HLC OS5 simulated data sets provided by JPL.** The simulated observation scenario 5 (OS5) corresponds to an observatory thermal model for the cycle 6 design. For each test case, without LOWFC, with LOWFC no DAC and with LOWFC random, without and with shot noise, we present the raw long exposure images at the nominal sampling ($0.35\lambda/D$) of the target star 47 UMa at two spacecraft orientations ($+13^\circ$ and -13° off nominal) and the reference star β UMa, normalized to the total contrast in the raw 47 UMa image at $+13^\circ$. Each image was multiplied by a binary mask to enhance the details in the dark hole.

of the reference data cube of the OS5 time series:

$$C^i = \frac{1}{N} \sum_{x,y} \frac{[I_1^i(x,y) - \bar{I}_1^i] [I_2(x,y) - \bar{I}_2]}{\sigma_{I_1^i} \sigma_{I_2}}, \quad (1)$$

where I_1^i is the i -th of the reference data cube, I_2 is the first image of the target data cube, N is the total number of pixels in the images, $\sigma_{I_1^i}$ and σ_{I_2} are respectively the standard deviations of the images I_1^i and I_2 . Two different reference cubes were considered: β UMa and 47 UMa at -13° for the RDI and ADI strategies respectively. We plotted the evolution of correlation with respect to the number of the reference short exposure (30 for β UMa and 50 for 47 UMa at -13°) in Fig. 2. Because each short exposure of the reference cube corresponds to the successive times at which the images were taken, the evolution of correlation with the number i of the reference cube short exposure can also be seen as an evolution of correlation with time. The level of correlation, superior to 99%, is representative of a very stable set of aberrations. In such a case, a simple reduction technique like classical point-spread-function (PSF) subtraction, also called coronagraphic PSF template subtraction (CPSFTS) [20], leads to similar levels of contrasts [7] as those obtained with more sophisticated techniques based on a principal component analysis (PCA). It is also interesting to note that, in the cases without LOWFC, the correlations are slightly better with ADI than with

<p>Reference star: β UMa ($V = 2.37$, A1IV) at $+13^\circ$ off nominal <i>30000 sec total integration (30 exposures \times 1000 sec)</i> OS5_adi_*_polx_**_beta_Uma_HLC_sequence.fits OS5_adi_*_polx_**_beta_Uma_HLC_sequence_noisy.fits OS5_adi_*_polx_**_beta_Uma_HLC_unocculted_psf.fits OS5_adi_*_polx_**_beta_Uma_HLC_zernikes.txt Science Target: 47 UMa ($V = 5.04$, G1V, has planets) at $+13^\circ$ off nominal <i>50000 sec total integration (50 exposures \times 1000 sec)</i> OS5_adi_*_polx_**_47_Uma_roll_p13deg_HLC_sequence_with_planets.fits OS5_adi_*_polx_**_47_Uma_roll_p13deg_HLC_sequence_noisy_with_planets.fits OS5_adi_*_polx_**_47_Uma_roll_p13deg_HLC_unocculted_psf.fits OS5_adi_*_polx_**_47_Uma_roll_p13deg_HLC_zernikes.txt Science Target: 47 UMa ($V = 5.04$, G1V, has planets) at -13° off nominal <i>50000 sec total integration (50 exposures \times 1000 sec)</i> OS5_adi_*_polx_**_47_Uma_roll_m13deg_HLC_sequence_with_planets.fits OS5_adi_*_polx_**_47_Uma_roll_m13deg_HLC_sequence_noisy_with_planets.fits OS5_adi_*_polx_**_47_Uma_roll_m13deg_HLC_unocculted_psf.fits OS5_adi_*_polx_**_47_Uma_roll_m13deg_HLC_zernikes.txt</p>

where * can be replaced by either "2" or "3_highres" and ** can be replaced by either "lowfc_no_dac", "lowfc_random" or "no_lowfc" as the case may be

Table 2: OS5 data sequence file structure.

RDI. Thus, the aberrations are generally more stable in the ADI case.

2.3 Comparison of classical PSF subtraction and KLIP for the RDI scenario

In this section, we make sure that our prediction about the similar efficiency of classical PSF subtraction and PCA-based techniques is correct for the OS5 data sets. We applied the classical PSF subtraction and the PCA-based KLIP algorithm [21] techniques to the OS5 data sets. Classical PSF subtraction consists of a simple subtraction of the collapsed and normalized reference data cube from the normalized long exposure raw target image. In the RDI strategy, the 47 UMa data set at $+13^\circ$ is the target and the reference is β UMa. Before reducing the data with KLIP, we subtract the average value of the science and reference data sequences so that they have zero spatial mean. We collapse the science data sequence (47 UMa) to obtain a long exposure raw image. We then apply the KLIP algorithm, which 1) decomposes the reference data into principal components or KLIP modes; 2) projects the long exposure raw image into these modes to create a synthetic reference; and 3) subtracts the synthetic reference from the target.

We compared the results of our reductions for different OS5 data sets (no LOWFC, LOWFC no DAC, LOWFC random, without and with noise) by plotting on the same dynamic range: the classical PSF subtraction and KLIP reductions using 30 modes (cf. Figure 3). Table 3 summarizes the 5σ contrast and contrast gains for the different cases. Apart from the case without LOWFC,

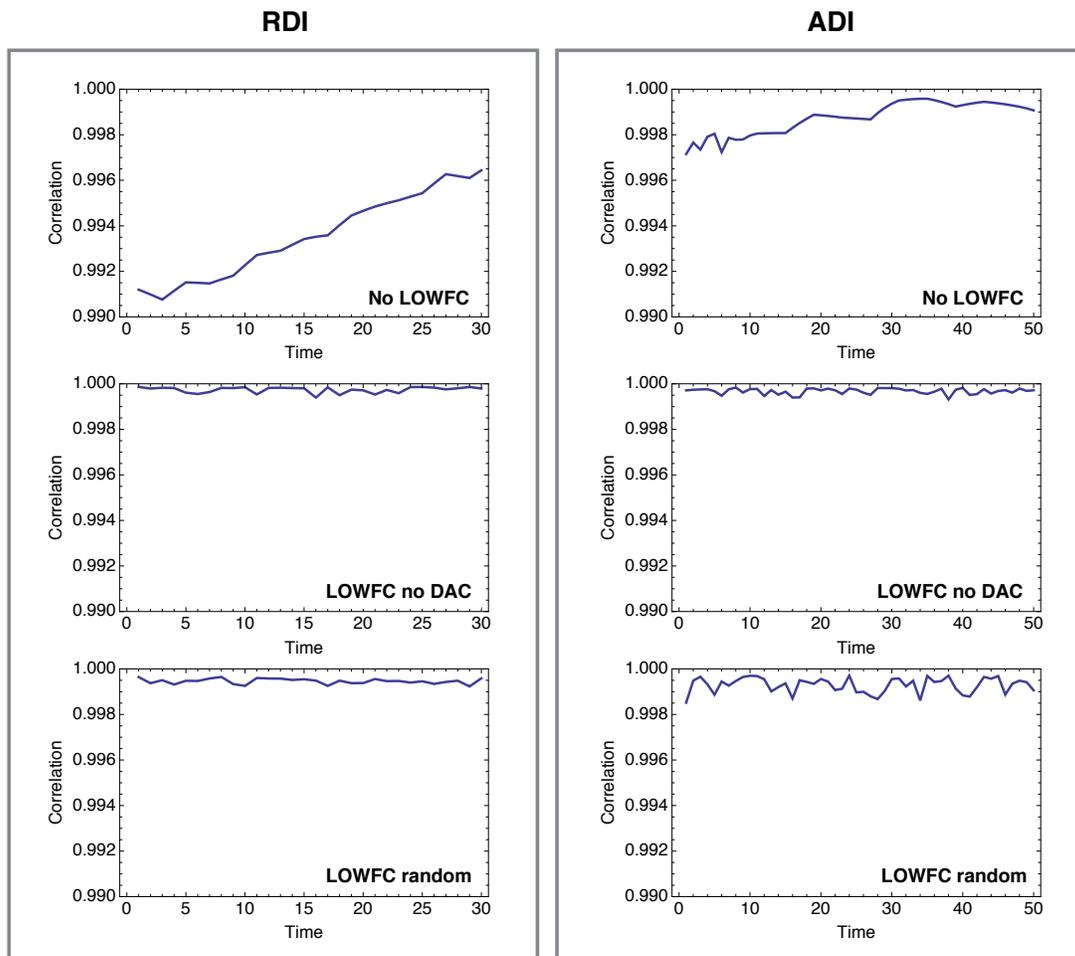


Figure 2: **Temporal stability of the aberrations.** [Left] Correlation of the last short exposure of the target 47 UMa +13° with each short exposure of beta UMa in the cases without LOWFC, with LOWFC no DAC and with LOWFC random. [Right] Correlation of the first short exposure of the target 47 UMa at +13° with each short exposure of the target 47 UMa at -13° in the cases without LOWFC, with LOWFC no DAC and with LOWFC random. With a level of correlation higher than 99% in all cases, the classical PSF subtraction is as efficient [7] as modern PCA-based reduction techniques.

the contrast gain over the field of view is quite similar in the two reductions. This demonstrates the positive impact of the LOWFC on post-processing. The stabilization of aberrations from the LOWFC improves the performance of the classical PSF subtraction, and decreases the differences between the two reduction techniques. In the two cases with the LOWFC, the classical subtraction performs nearly as well as KLIP, with both methods providing enough improvement compared to the raw data to detect all three simulated planets in the noiseless data. These results are consistent with previous results obtained in the reduction of OS3 data sets [8].

The contrast curves as a function of separation in Fig. 4 demonstrate that increasing the number of KLIP modes only improves the contrast level in the case without LOWFC. In the cases with LOWFC the level of contrast is similar regardless of the number of KLIP modes used. This is another demonstration of the extreme stability of aberrations in the OS5 data sets. However, an important limitation of these simulations is that there is not any jitter in these data sets. The classical PSF subtraction may be less effective in presence of jitter but any PCA-based technique such as KLIP [21] should be able to mitigate this effect.

It is also interesting to note that no striking difference in the performance of these two post-processing techniques is observed between the two cases with LOWFC (LOWFC no DAC and LOWFC random). For the sake of simplicity, the discussion which follows is restricted to data sets without and with LOWFC random reduced with the classical PSF subtraction.

We identified two different regimes in the reduced images according to the data set: a regime without photon noise (noiseless images) and a regime where speckle noise is negligible with respect to photon noise (noisy images). Table 4 shows the RMS of noiseless reduced image (speckle noise only), noisy minus noiseless reduced images (photon noise only) and noisy reduced image (speckle noise + photon noise). In noisy images, photon noise is dominant over speckle noise for all cases. In the following noiseless and noisy data are respectively used to study the cases of a speckle-noise and a photon-noise dominated regime.

	Noiseless				Noisy			
	Classical PSF Sub		KLIP 30 modes		Classical PSF Sub		KLIP 30 modes	
	Contrast	Gain	Contrast	Gain	Contrast	Gain	Contrast	Gain
Without LOWFC	2.3×10^{-10}	5.0	1.5×10^{-10}	7.7	8.0×10^{-10}	1.7	8.1×10^{-10}	1.6
LOWFC no DAC	1.5×10^{-10}	7.6	1.5×10^{-10}	7.6	7.3×10^{-10}	1.8	7.2×10^{-10}	1.8
LOWFC random	1.7×10^{-10}	6.9	1.4×10^{-10}	7.9	7.3×10^{-10}	1.8	7.4×10^{-10}	1.7

Table 3: **Final 5σ contrast and contrast gain factors for the RDI strategy.** 5σ contrast and contrast gains over the raw data, averaged over the field of view, for all cases reduced with classical PSF subtraction and 30 KLIP modes: without LOWFC, with LOWFC no DAC, with LOWFC random, with and without photon noise.

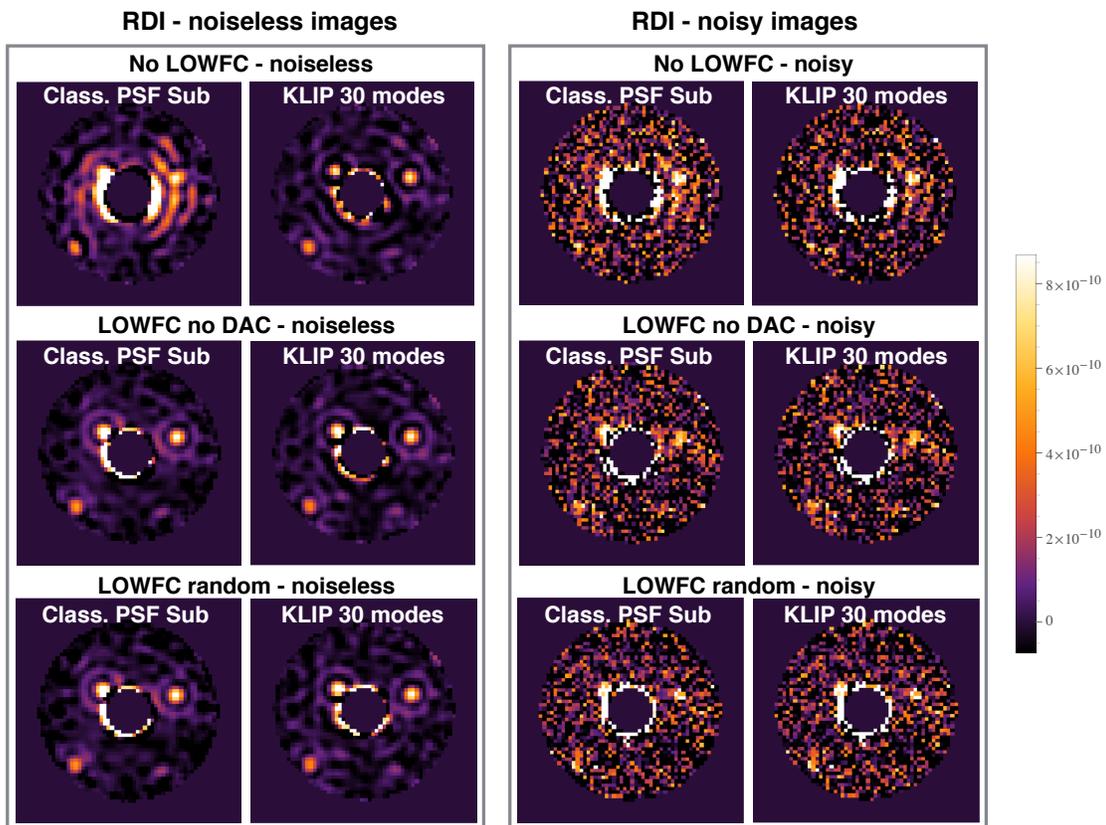


Figure 3: **Reductions of HLC OS5 data sets simulated with the RDI strategy.** Comparison of reductions of the long exposure image of the target 47 UMa for the data sets without LOWFC, with LOWFC no DAC and with LOWFC random, without and with noise, with same dynamic range. For each case, we compare classical subtraction and KLIP with 30 modes (maximum number of short exposures of the reference β UMa). Detection is more difficult in the noisy images than in the noiseless images. Apart from the case without LOWFC, the contrast gains over the field of view are pretty similar for both techniques (see Table 3).

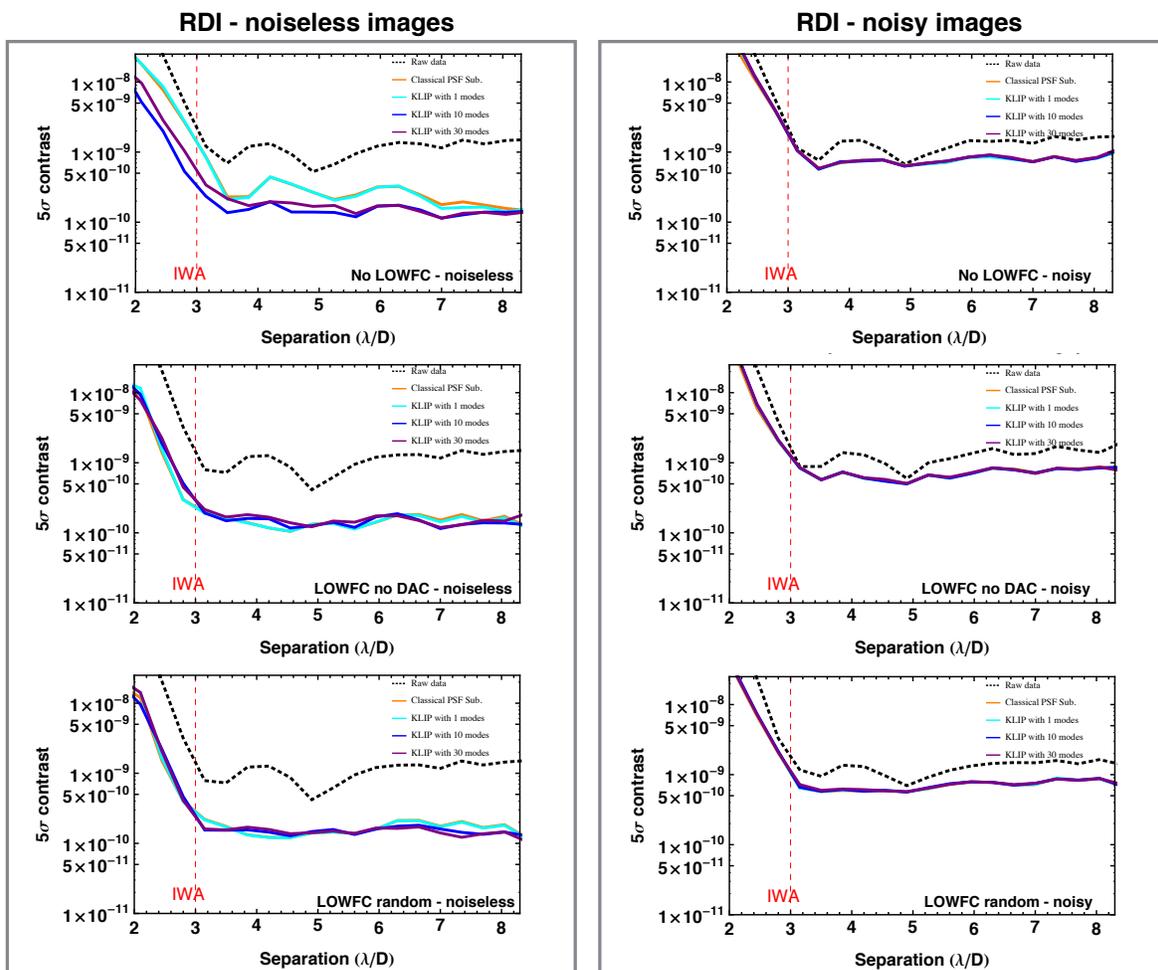


Figure 4: Contrast curves for reduced HLC OS₅ data sets for the RDI strategy. Comparison of 5σ contrast as a function of separation for the data sets without LOWFC, with LOWFC no DAC and with LOWFC random, without and with noise, with same dynamic range.

	Noiseless image (Speckle noise only)	Noisy image - Noiseless image (Photon noise only)	Noisy image (Speckle noise + Photon noise)
Without LOWFC	4.1×10^{-10}	7.7×10^{-10}	8.2×10^{-10}
LOWFC no DAC	1.8×10^{-10}	7.0×10^{-10}	7.3×10^{-10}
LOWFC random	2.0×10^{-10}	7.1×10^{-10}	7.3×10^{-10}

Table 4: **Identification of speckle-noise and photon-noise dominated regimes in reduced images.** 5σ contrast averaged over the field of view of noiseless reduced image (speckle noise only), noisy minus noiseless reduced images (photon noise only) and noisy reduced image (speckle noise + photon noise). In noisy images, photon noise is dominant over speckle noise for all cases. We thus identify two different regimes in the reduced images according to the data set: speckle-noise dominated regime for noiseless images (no photon noise at all) and photon-noise dominated regime for noisy images (speckle noise is negligible with respect to photon noise).

2.4 Comparison of RDI and ADI observing strategies

We applied the classical PSF subtraction to reduce the data sets without and with LOWFC random. Recall that this technique consists of a simple subtraction of the collapsed and normalized reference data cube from the normalized long exposure raw target image. In both RDI and ADI strategies, the 47 UMa data set at $+13^\circ$ is the target. The reference are β UMa and 47 UMa at -13° for the RDI and ADI strategies respectively. Table 5 summarize these strategies. In Fig. 5, we compare the results of the RDI and ADI reductions for the different OS5 data sets, without and with LOWFC. Two different regimes are considered: noiseless and noisy data are respectively used to study the cases of a speckle-noise and a photon-noise dominated regimes.

In the speckle-noise dominated regime, three planets are detected in all data sets with both strategies. However, ADI enables up to a twofold and a threefold gains in speckle subtraction with respect to RDI respectively without and with LOWFC (see Table 6), helping to better discriminate between planets and residual speckles. This is the case for both low-order (see noiseless data set without LOWFC) and high-order (see noiseless data set with LOWFC) aberration residuals. The contrast, defined as the azimuthal standard deviation, and the contrast gain (over the raw contrast) as a function of separation for noiseless data are represented respectively on the left of Fig. 6 and Fig. 7. The classical PSF subtraction benefits more from ADI than RDI in the case without LOWFC. At small separations, there is no significant difference between the two observing strategies. However, at large separations, ADI enables significant contrast improvements over RDI. The contrast gains are multiplied by factors of up to 2 with LOWFC and up to 3 without LOWFC. This is consistent to the results of the correlation study performed in 2.2 that demonstrated a slightly more stable set of aberrations in the ADI strategy than in the RDI strategy. The ADI works this well presumably because the spacecraft thermal and structural design is good enough that there is little thermal change between the balanced symmetric cases at $+13^\circ$ and -13° roll. There must be less thermal and optical change between these states than between the observing attitude for β UMa and either of them. This is consistent with the correlation study that we have done in 2.2 (see comparison of RDI and ADI cases without LOWFC in Fig. 2). In presence of LOWFC, the lower

spatial frequencies are stabilized so the difference between RDI and ADI is smaller at short separations. The LOWFC improves the contrast at less than $7\lambda/D$ but makes no difference outside of that (see Fig. 7, left, plain lines). Seems like the LOWFC has the biggest impact at shorter separations. Without LOWFC there is an improvement of ADI over RDI for all separations (see Fig. 7, left, dashed lines).

The reduction of noisy data gives insight into the performance of RDI and ADI strategies in a photon-noise dominated regime. Even though all the reductions show an improvement in contrast compared to the raw images of Fig. 1, the signal-to-noise ratio for each planet is degraded by the presence of photon-noise. The contrast and the contrast gain as a function of separation for noisy data are represented respectively on the right of Fig. 6 and Fig. 7. With a total exposure time on the target multiplied by two in the ADI strategy (see table 5 for exposure times), contrast gains are slightly better than for the RDI strategy, i.e. up to a factor of 1.3 at $4\lambda/D$ when looking at the gain ratio between the RDI and ADI cases with LOWFC.

We studied the impact of two different observing strategies on the contrast gain after post-processing on WFIRST simulated images. We considered the cases of noiseless and noisy data to study the impact in two different regimes: speckle-noise and photon-noise dominated. We found that the ADI observing strategy is always as effective as or up to 3 times better than RDI for this particular data set. In the speckle-noise dominated regime, ADI is better than RDI at subtracting both low-order and high-order aberration residuals. The ADI gain increases dramatically for separations higher than $5\lambda/D$ to reach about 30 in the case without LOWFC. In the photon-noise dominated regime, we observe slightly better gains for all separations due to the gain in signal-to-noise ratio offered by ADI over RDI. We thus recommend that the telescope be able to roll for coronagraphy.

As a conclusion, either in the speckle-noise dominated and photon-noise dominated regimes, it is worth to roll the telescope. As demonstrated in the OS5 data, two spacecraft orientations at $\pm 13^\circ$ off nominal show a significant improvement in all tested cases. In the speckle-noise dominated regime, the gain over RDI is up to 3 at large separations where the LOWFC seems to be less efficient at correcting speckles. ADI also increases contrast gains at short separations with respect to RDI in the case of a "deficient" LOWFC.

Strategy	Observing sequence	Integration time per star	Total integration time
RDI	β UMa at roll $+13^\circ$	30000 sec	80000 sec
	47 UMa at roll $+13^\circ$	50000 sec	
ADI	47 UMa at roll $+13^\circ$	50000 sec	100000 sec
	47 UMa at roll -13°	50000 sec	

Table 5: **RDI and ADI observing strategies.**

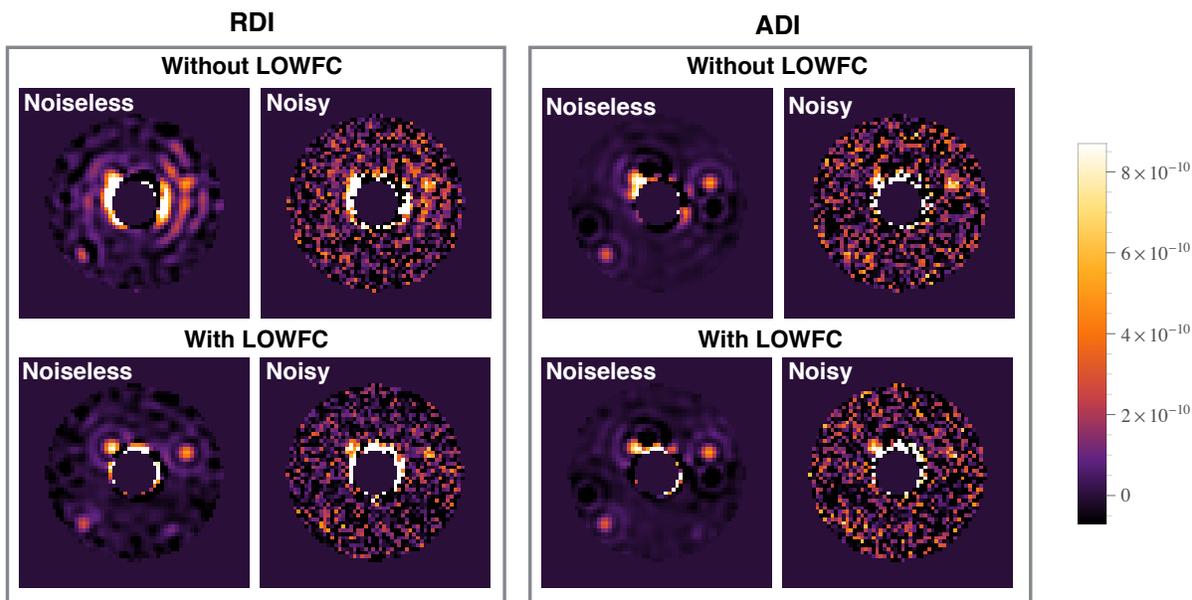


Figure 5: **Post-processed data for both RDI and ADI observing scenarios.** Comparison of [left] RDI and [right] ADI reductions of the long exposure image of 47 UMa for the noiseless and noisy data sets, without and with LOWFC. Reductions have been performed with the classical PSF subtraction technique. The ADI strategy enables a slightly better speckle subtraction than the RDI strategy, helping to better discriminate between planets and residual speckles. This is particularly obvious in the case without LOWFC.

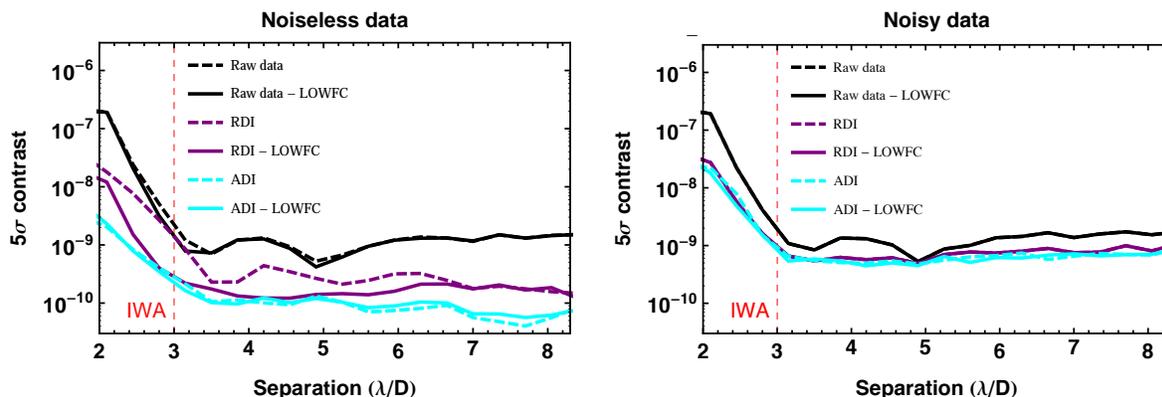


Figure 6: **Contrast curves in the speckle-noise dominated regime for RDI and ADI strategies.** Comparison of 5σ contrast as a function of separation of [left] RDI and [right] ADI reductions for data without and with the LOWFC.

	RDI				ADI			
	Noiseless		Noisy		Noiseless		Noisy	
	Contrast	Gain	Contrast	Gain	Contrast	Gain	Contrast	Gain
Without LOWFC	2.3×10^{-10}	5.0	8.0×10^{-10}	1.7	7.8×10^{-11}	14.8	6.6×10^{-10}	2.0
LOWFC random	1.7×10^{-10}	6.9	7.3×10^{-10}	1.8	8.6×10^{-11}	13.1	6.4×10^{-10}	2.1

Table 6: **5σ contrast and contrast gains.** 5σ contrast and contrast gains over the raw data, averaged over the field of view, for all 8 cases: RDI without LOWFC, RDI with LOWFC, ADI with LOWFC, ADI without LOWFC – with and without photon noise.

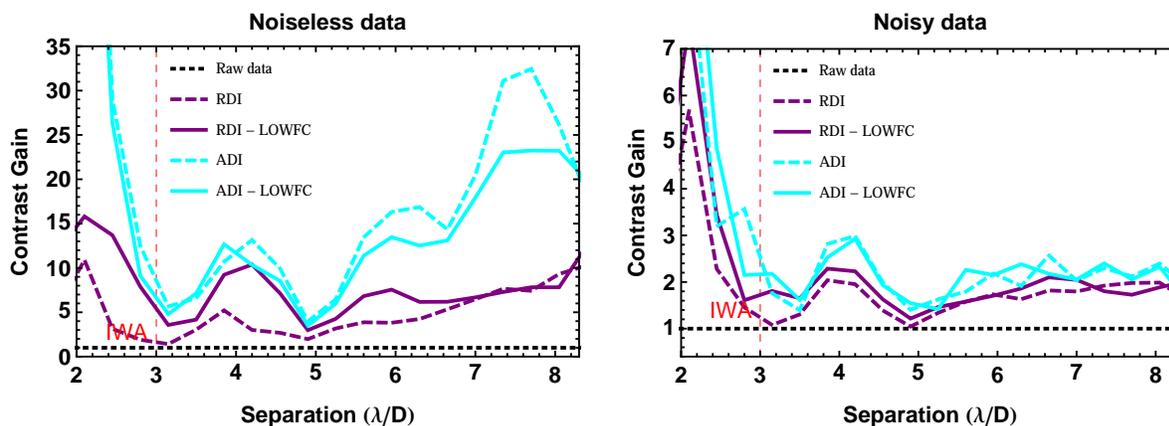


Figure 7: **Contrast gain curves for RDI and ADI strategies.** Comparison of contrast gain as a function of separation for [left] noiseless and [right] noisy data without and with the LOWFC. In the speckle-dominated regime, the classical PSF subtraction benefits more from ADI than RDI in the case without LOWFC. At large separations, ADI enables to multiply the gains by a factor up to 2 with the LOWFC and up to 3 without the LOWFC. In the photon-noise dominated case, ADI benefits from doubled exposure time on the target, with gains multiplied by a factors up to 1.3 at $4\lambda/D$ when looking at the gain ratio between the RDI and ADI cases with LOWFC.

3 Impact of spatial sampling on the post-processing performance

The long integration times required for spectroscopic characterization will consume a large portion of the CGI mission time [22]. The current Cycle 6 IFS design oversamples the coronagraph image: at the center wavelengths of the three characterization bandpasses the sampling rates are 3.3, 3.9, and 4.6 lenslets across each λ_0/D resolution element². Since the SNR of the dispersed spectrum produced on the detector for an individual spatial element is directly proportional to the lenslet collecting area, the lenslet scale is a direct lever on the time required to characterize an exoplanet. Based on this consideration, Project and Science Investigation Team personnel requested our group to conduct a preliminary study on the impact of coarser spatial sampling on post-processing performance.

We used the already available OS5 simulations as a test data set, though now starting from a highly oversampled version (10 pixels per λ/D) specially prepared by J. Krist for this test. We manipulated these oversampled images to derive various downsample/shift cases. Notably, there was no IFS model included, so effectively we assumed a perfect reconstruction of the coronagraph image from the IFS focal plane. We restrict the observing scenario to RDI with the LOWFC-corrected version of the images. We tested instrument sampling rates of 3, 2, and 1.5 pixels per λ/D . We repeated all tests on both the noiseless and noisy data simulation, thereby probing the speckle-noise and photon-noise dominated regimes respectively.

We assumed that the star in the reference image sequence can appear at a fixed, arbitrary sub-pixel offset from the target star, so that it must be counter-shifted to align with the science target before PSF subtraction. This mimics a scenario where the registration of the coronagraph image on the lenslet array jumps between the two observations. The high-resolution reference exposures (sampled at 10 pixels per λ/D) were shifted by the trial offset, binned to the test instrument spatial sampling, and finally counter-shifted back to align with the science target PSF. Unlike the reference images, we did not shift the science target images but only binned them to the instrument sampling of the given test case. We assumed perfect knowledge of the reference-target registration offset.

Because the images are signals of finite support and therefore not strictly band-limited, some interpolation noise occurs after the counter-shift, and it grows as the sampling rate is reduced to near or below the Nyquist rate. We used the bicubic spline interpolation routine from the Python SciPy [23] library to apply the shifts, with a pre-filter step to suppress the noise introduced by high spatial frequency content.³ We illustrate in two figures the noise introduced by the counter-shift step. First in Figure 8, we show the distribution of errors that appear for one sampling case (2.0 pixels per λ/D) when an image is counter-shifted by 0.6 pixels, taking the difference from the same image simulated with no registration offset. In Figure 9, we plot distribution of interpolation

²https://wfirst.ipac.caltech.edu/sims/Param_db.html

³Of the various options available in the SciPy interpolation and shift routine, we found the pre-filtered bicubic spline offered the best performance, in terms of the noise level introduced at coarse (sub-Nyquist) sampling rates.

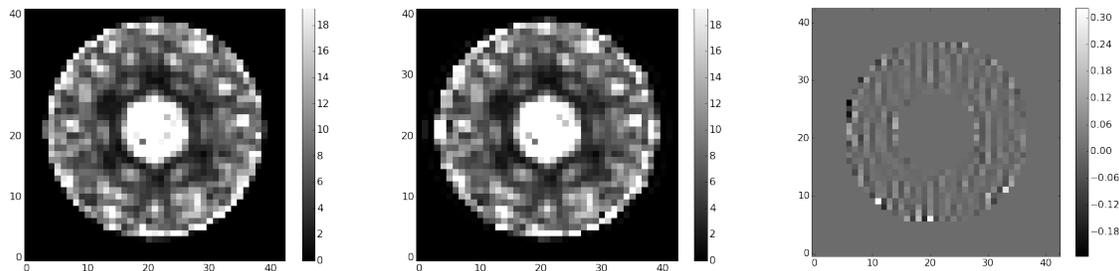


Figure 8: **Effect of interpolation noise on aligning sub-pixel offset images.** [Left] Simulated HLC coronagraph image down-sampled to 2 pixels per λ/D without any offset; [Center] the same image after a horizontal offset of $+0.3 \lambda/D$ in the high-resolution frame followed by a counter-shift of $-0.3 \lambda/D$ in the down-sampled frame (-0.6 pixels); [Right] the difference between the two down-sampled image realizations, indicating residual errors around 10% of the raw intensity values.

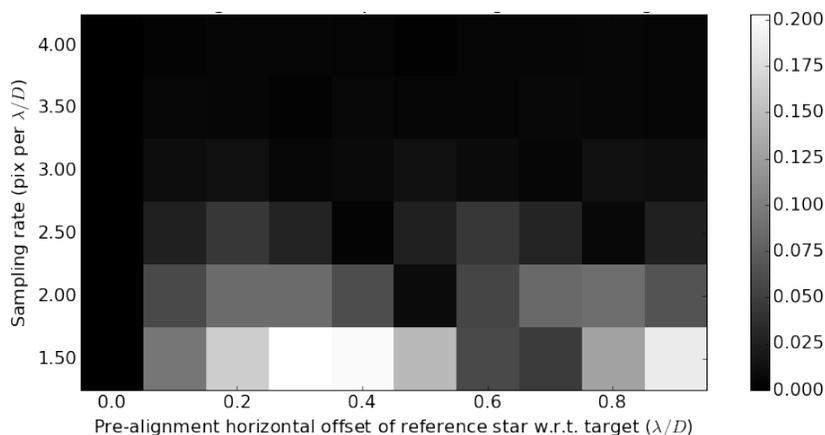


Figure 9: **Quantifying the noise introduced by shifting an image: the relative RMS flux error (normalized with respect to the raw speckle RMS) that appears in a reference image, after interpolation and alignment to the target star frame.** When the target-reference registration offset is zero (far left column), no shift is required and hence the added noise is zero. In general, the introduced noise is a periodic function of the offset which peaks at 0.5 pixels. The shifting noise is near or below 10% of the raw speckle level when the image is Nyquist sampled (2.0 pixels per λ/D).

noise over all the tested instrument sampling rates and registration offsets. There we see that the shifting noise can be significant (20% of raw speckle noise) at sub-Nyquist spatial sampling (1.5 pixels per λ/D), and that this noise generally peaks when the counter-shift is close to 0.5 pixels in the down-sampled frame.

We carried out RDI post-processing, first for the case when there is zero registration offset between reference and target images (no shifting) and second for the case when the reference images are offset by $0.3 \lambda/D$ (near the worst-case shift). In Figure 10 we show the final PSF-subtracted image as a function of spatial sampling and offset for the data version with planets inserted, and no photon noise. In Fig. 11 we show the corresponding contrast curves. The curves show that down-sampling alone does not degrade the post-processed contrast floor. At Nyquist sampling (2.0 pixels per λ/D) or higher, reference-to-target offsets have negligible impact on the post-processed con-

trast floor. However, at the coarse spatial sampling (1.5 pixels per λ/D), counter-shifting a reference image introduces noise that can significantly degrade the post-processed contrast floor.

We performed the same analysis for the photon noise-dominated version of the data. The reduced images are displayed in Fig. 12. In this case, the planets become more obvious with a coarser spatial sampling, as the uncorrelated photon noise averages out within the larger pixels. This is particularly true for the faintest planet in the bottom left of the image, which is not apparent in presence of photon noise with the finest sampling. As before for the noiseless data, we plot in Fig. 13 the 5σ contrast curves for the different considered test cases. As the spatial sampling rate is decreased, the smoothing effect on the uncorrelated noise floor has the effect of decreasing the contrast curve. However, this should not be interpreted as a preference for the coarsest, sub-Nyquist (1.5) instrument sampling. Rather, this trend is an artifact of our flawed measure of the contrast floor in the presence of uncorrelated noise: we could get the same result or better by starting from the intermediate (2.0) sampling case, and convolving the image with the PSF model to act as a matched filter.

We arrive at two main conclusions from our spatial sampling study:

1. Within the limited noise components included in the OS5 simulations (no IFS image reconstruction noise, no systematic detector noise), relaxing the spatial sampling does not degrade the post-processed contrast floor.
2. At a spatial sampling equal to or above the Nyquist rate (2 pixels per λ/D) we can apply arbitrary alignment offsets to reference library images without adding significant noise.

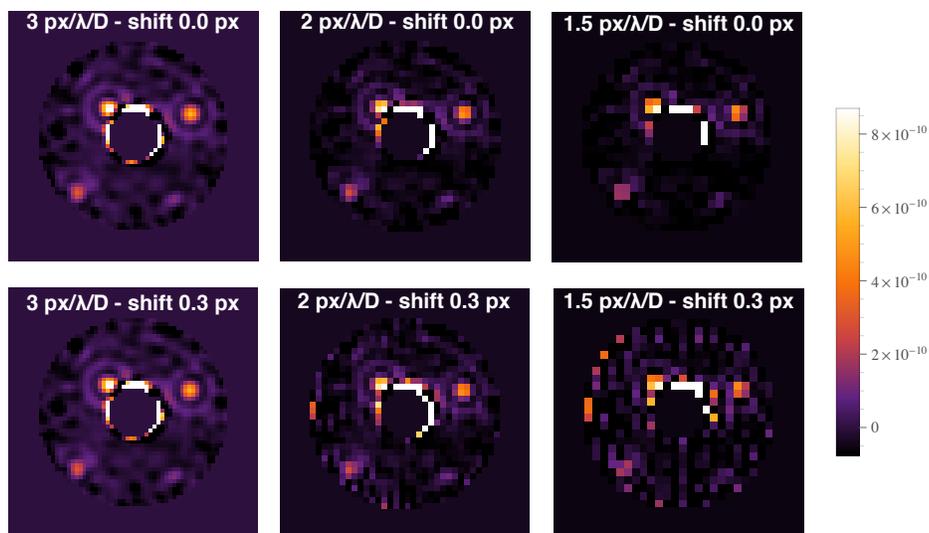


Figure 10: **Post-processed OS5 images, with fake planets and no photon noise, for different instrument spatial samplings and reference shifts.** From left to right, reduced images are represented for samplings of 3, 2 and 1.5 pixels per λ/D . [Top] No shift was introduced to the reference. [Bottom] a shift of 0.3 pixels was introduced to the reference.

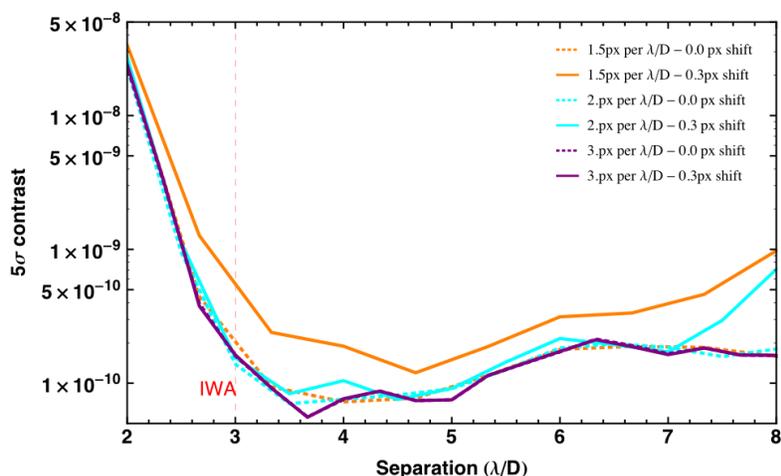


Figure 11: **5σ contrast as a function of separation after post-processing for different samplings and shifts.** The contrast of post-processed noiseless data is not impacted by a coarser sampling of the data. However, a coarser sampling is less robust to a re-alignment of the reference frame as we can see in the case sampling case (1.5 pixels per λ/D) with a $0.3 \lambda/D$ reference-target offset.

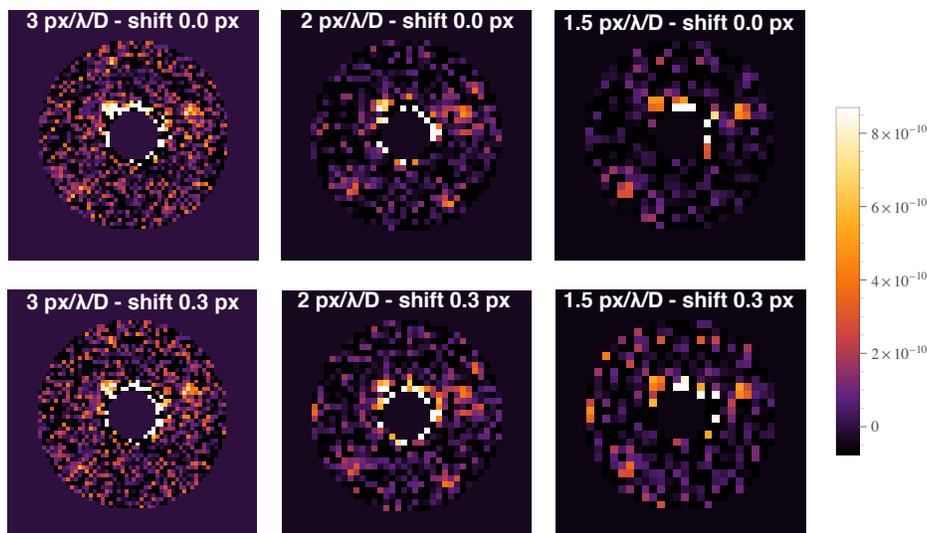


Figure 12: **OS5 residual images with photon noise, post-processed for different spatial samplings and reference-target offset.** From left to right, reduced images are represented for spatial samplings of 3, 2 and 1.5 pixels per λ/D . [Top] No offset between the reference and target images. [Bottom] A simulated offset of $0.3 \lambda/D$ was applied to the reference sequence, which was then countered in the down-sampled frame before PSF subtraction. The planets become more visible with a coarser sampling as the photon noise averages out. This is particularly true for the faintest planet in the bottom left of the image, which is less obvious at the finest sampling.

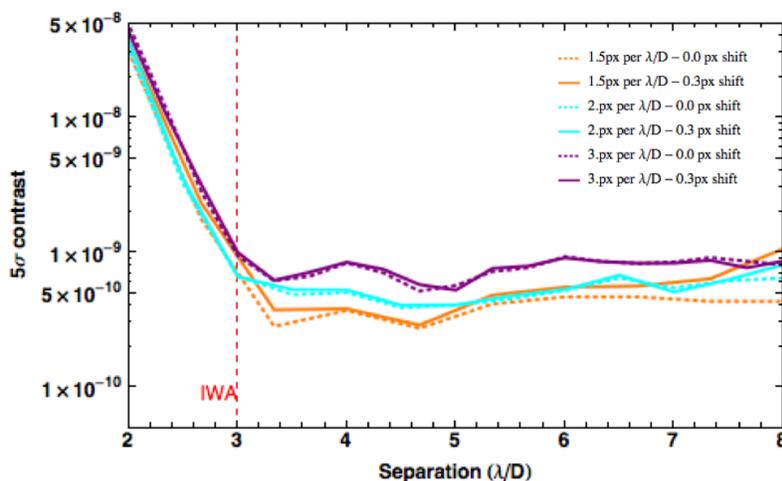


Figure 13: **HLC OS5 simulated noisy data sets with different samplings and shifts - Contrast curves.** Comparison of 5σ contrast as a function of separation for data with LOWFC with different samplings and shifts. RDI classical PSF subtraction reductions. As the residual image is dominated by spatially uncorrelated photon noise, the contrast curve decreases with coarser sampling. Shifting the reference image only slightly deteriorates contrast at the coarse sampling of 1.5 pixel per λ/D .

4 Laboratory data

4.1 Data description

JPL acquired new laboratory data sets on the SPC testbed at HCIT in December 2015. This new “drift” coronagraph data addressed two of the limitations we identified in the SPC testbed data from 2015 [10]: (i) changing DM actuator settings between each exposure, a byproduct of the dark hole generation loop where at each control iteration there is a sweep of a regulation parameter on the DM solution, and (ii) short integration times which imposed a detector noise floor on the RDI result. The new “drift” data includes only frames taken after the dark hole was acquired, and the DM actuator settings frozen, resulting in a speckle pattern that evolves solely due to the thermal environment, mechanical stresses on the testbed optics, and possible small drifts in the DM actuators. A second major feature of the “drift” data are integration times long enough such that the ratio of raw speckle-to-background noise ratio in an individual frame increased by roughly on order of magnitude from the ratio of ~ 3 measured in the 2015 data [10]. This enabled us to test post-processing in an opposite noise regime, where correlated speckle noise always dominates the subtraction residual. Furthermore, the “drift” data were acquired with the 2nd generation SPC design that uses a Lyot mask configuration. This allowed us to measure detection limits at working angles down to $2.5 \lambda/D$ instead of the $4 \lambda/D$ inner working angle limit of the 1st generation design [12].

The basic properties of the SPC drift data are summarized in Table 7. There were two acquisition runs, each recorded at a different light source bandpass: The *narrowband* data were acquired in a 2% bandpass centered at 572 nm, while the *broadband* data were acquired at a 10% bandpass centered at 550 nm. To subtract detector bias and dark current, we were provided with a set of dark frames acquired at the same integration time as the coronagraph data (100.0 s per frame). The raw detector images highly oversample the PSF, at 11.9 pixels per λ/D . The off-axis light source calibration information for each bandpass was provided directly by JPL at a 2×2 pixel binning from the raw detector image (spatial sampling 5.95 pixels per λ/D). During our post-processing analysis we further down-sampled the image using 4×4 pixel bins, resulting in a spatial sampling of 2.98 pixels per λ/D (close to the nominal OS5 data simulation spatial sampling at 2.86 pixels per λ/D).

Example raw images from each data set, displayed in contrast units after 4×4 spatial binning, and masking off pixels outside of the bowtie-shaped FoV, are shown in Figure 14. Notably, there is a significant gap in raw contrast between the narrowband and broadband data sets; in contrast units the brightest intensity value in narrowband FoV is over $5 \times$ higher than in the broadband case. We further examine the raw image statistics in Figure 15. There we see that the speckle noise is increasing with time in each data set, as the wavefront degrades with respect to the optimized DM solution. The disparity between the maximum intensity values appears again for the spatial mean and standard deviation values. The fact that these raw contrast metrics are higher in the narrowband data than in the broadband data went against our expectation, since in general it is

	Narrowband	Broadband
Bandpass	566–578 nm (2%)	523–578 nm (10%)
Integration sequence	108 frames \times 100 s	108 frames \times 100 s
Light source peak count rate	1.87×10^{10} ADU/s	5.32×10^{10} ADU/s
Spatial mean count rate in coronagraph FoV	1.00×10^3 ADU/s	1.10×10^3 ADU/s
Flight equivalent light source V -band mag	−4.78	−4.16
Flight equivalent T_{int} per frame, β UMa ($V = 2.37$)	20.1 h	11.4 h
Flight equivalent T_{int} per frame, 47 UMa ($V = 5.04$)	235 h	133 h

Table 7: Summary of the “drift” SPC data set acquired at JPL HCIT in December 2015. The laboratory light source count rates were measured by JPL during a separate calibration procedure, and represent the peak value of the off-axis, unocculted PSF after 2×2 -binning the CCD array pixels. We compute the flight-equivalent source brightness based on the Cycle 5 WFIRST CGI model which includes losses in the telescope and instrument optical train, and losses due to the SPC masks assuming no polarizing filter. This model assumes a flat bandpass response and a source spectral type GV5. The corresponding flight-equivalent, per-frame integration times for the canonical simulation reference-target star pair, β UMa and 47 UMa, are included here for reference.

more difficult to control the coronagraph wavefront over a broader bandpass.

We checked the light source calibration data and our interpretation of those values with JPL staff member Eric Cady, in order to eliminate a possible bookkeeping error. We also independently checked the light source calibration in a relative sense, by measuring the background scattered light level in a box adjacent to the coronagraph PSF. The broadband/narrowband ratio of this scattered light level in raw detector counts is 2.67. This ratio is within 6% of the ratio of JPL’s calibration source values, 2.84 (comparing the values in the third row of Table 7). The broadband/narrowband performance disparity is partly explained by the fact that the narrowband data used the channel at the red extreme (572 nm) of the 10% bandpass, where the accuracy of the wavefront correction is generally worse. The root cause of the disparity has not been pursued any further at the time of this report. For the time being, we place a higher emphasis on the broadband data results, partly because the deeper raw contrast achieved in this data sequence are much closer to the baseline 10^{-8} starlight suppression spec of the CGI design. In addition, the wider bandpass (10%) matches that of the baseline imaging channel design and the bandpass of the OS5 HLC data, so it is operationally the more relevant test case.

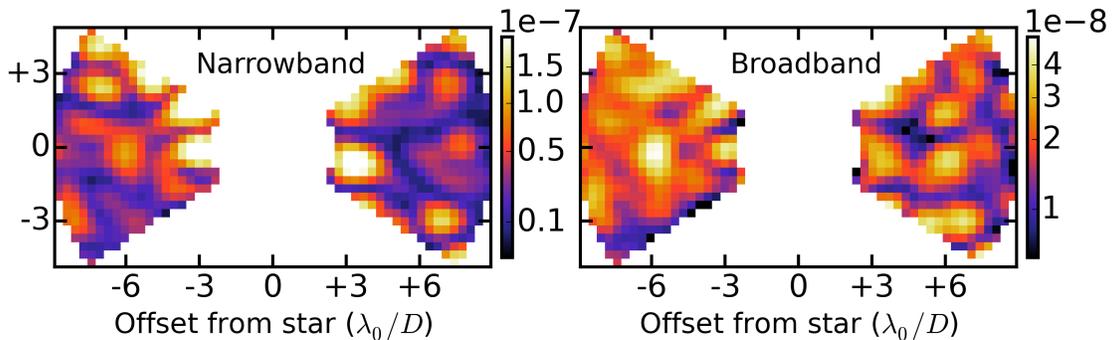


Figure 14: Individual raw images from the “drift” SPC data sets, displayed in contrast units on a square-root color stretch.

Also apparent in Figure 15, there is a “knee” in the evolution of the raw broadband statistics that

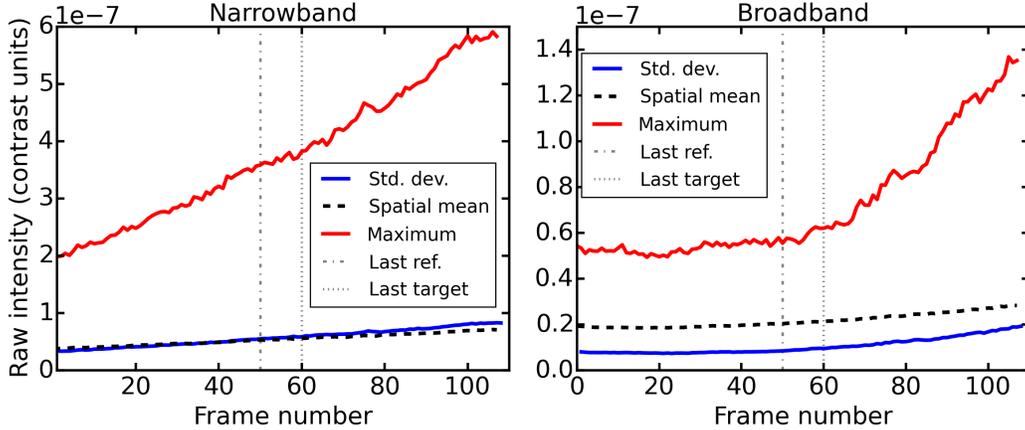


Figure 15: Time evolution of raw image statistics in the “drift” SPC data sets. The gray vertical lines mark the partition of frames between the reference and target sets: the first 50 serve as the reference library and the 10 succeeding frames serve as the target frames.

occurs near the half-way mark in the 108-frame data sequence. This appears in the raw images as a sudden brightening of some of the speckles. For our mock RDI scenario, therefore, we chose to cut off the data set at 60 frames. In view of the relatively high per-frame speckle SNR in the drift SPC data (and consequently long flight-equivalent integration time on the scientific target), we selected only the last 10 of these 60 frames as the target subset, and assigned the first 50 frames in each data set to the reference image library. In some of our analysis, we further restricted the target data subset to a single frame.

4.2 Results of mock RDI scenario

Table 8 summarizes the results of applying KLIP subtraction to the RDI scenario with 50 reference images and 10 combined target images. The 1σ contrast floor levels defined here are the standard deviations measured in the image, scaled up by two penalty factors: first the off-axis PSF throughput of the coronagraph, which varies with field point, and is mainly determined by the hard truncation of the planet PSF near the FoV edge defined by the bowtie-shaped focal plane mask. The second correction factor, accounted for in the final contrast but not the raw contrast, is the KLIP algorithm’s oversubtraction of the planet signal. We calibrated these correction factors by modeling the off-axis PSF at 14 radial separations between $2.5 \lambda/D$ and $9.0 \lambda/D$ in increments of $0.5 \lambda/D$, and 7 position angles on each side of the 65-degree opening angle bowtie FoV (at offset angles -30 , -20 , -10 , 0 , $+10$, $+20$, $+30$ degrees with respect to the bowtie axis).

We modeled the off-axis SPC PSF with the WebbPSF Python package, using its new WFIRST CGI module currently under development [24, 25]. The projection coefficient of the off-axis planet PSF model onto the K-L basis, computed as a function of the K-L mode truncation (K_{KLIP}), determines the oversubtraction correction factor to apply at a given image position. In total, we cali-

brated the contrast floor at 196 positions over the coronagraph FoV, for K_{KLIP} values ranging from 1 to 30 modes. For Table 8 we averaged the resulting contrast values in two sub-regions: first over the full bowtie FoV, and then the inner sub-region that spans radial separations 2.5–4.5 λ/D . The inner sub-region is of particular interest due to the greater abundance of potential exoplanet targets at small angular separations [22]. When considering the full field of view, we found the optimal K_{KLIP} for the narrowband and broadband data sets were 13 and 23 K-L modes, respectively. When restricting the measurement to the inner sub-region (2.5–4.5 λ/D), the optimal K_{KLIP} is 13 and 10 K-L modes, respectively. In all cases, the improvement in contrast floor beyond a K_{KLIP} of ~ 10 K-L modes is marginal.

Data set	Bowtie FoV measurement zone	Raw contrast floor (1σ)	Final contrast floor (1σ)	Gain factor
Narrowband	Full (2.5–9.0 λ_0/D)	8.45×10^{-8}	2.39×10^{-9}	35.3
	Inner (2.5–4.5 λ_0/D)	1.53×10^{-7}	4.00×10^{-9}	38.2
Broadband	Full (2.5–9.0 λ_0/D)	1.32×10^{-8}	9.66×10^{-10}	13.7
	Inner (2.5–4.5 λ_0/D)	1.80×10^{-8}	1.29×10^{-9}	13.9

Table 8: Summary of KLIP subtraction results for the RDI scenario where 10 target images are combined, using the preceding 50 images as a reference library. The stated contrast floors are standard deviation measurements scaled up by correction factors to account for field-dependent coronagraph throughput and KLIP oversubtraction. The throughput-corrected contrast values were spatially averaged over the indicated measurement zone.

One of the most striking difference between the narrowband and broadband results shown in Table 8 is the disparity in contrast gain. This was caused by the higher wavefront stability during the narrowband data sequence, despite the fact that the overall quality of wavefront correction (and hence raw contrast) was worse than the broadband version. To test this hypothesis, within each data set we computed the correlation coefficients (as defined in Equation 1) between the first target frame (number 51) and the preceding 50 reference images. The correlation between the narrowband frames grows in a roughly linear fashion from 0.974 at the first reference frame to 0.999 at the last reference frame. The same correlation metric in the broadband data follows a similar trajectory, but ranges from a significantly lower value of 0.772 in the first reference frame up to 0.996 in the best reference-target pair. Based on previous experience with both laboratory and simulated WFIRST data, we expect the contrast gain to depend strongly on these correlation figures, which are a proxy for the wavefront stability. Table 8 also shows that the larger contrast gain seen in the narrowband data does not make up for the poorer raw contrast, so the final contrast floor remains a factor of 2–3 \times higher than in the broadband result.

In Figure 16 we plot the same RDI KLIP subtraction results discussed above, but now showing how the contrast floor varies as a function of radial separation, for three different K_{KLIP} values: 1, 5, and 10 K-L modes. The $K_{\text{KLIP}} = 1$ case (akin to classical PSF subtraction) is surprisingly far from optimal, compared to previous experiments. This might be due to the strong inhomogeneities in the speckle distribution observed in the raw data (see above Figure 14). A few extremely bright

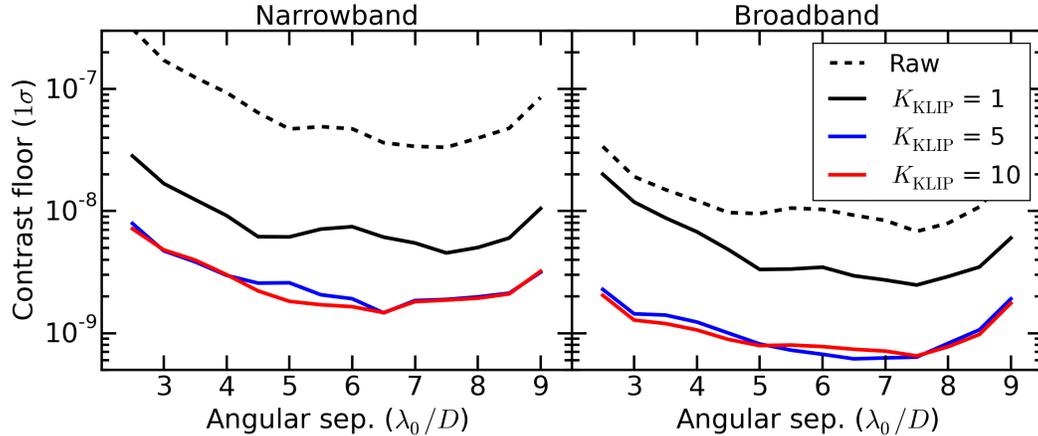


Figure 16: Contrast floor measured as a function of radial separation, before and after three KLIP RDI cases: 1 K-L mode (classical subtraction), 5 K-L modes, and 10 K-L modes (near optimal). In all cases, 10 reduced target images were averaged, and the first 50 frames in the sequence acted as the reference library.

speckles in the field of view can “throw off” the step in KLIP where the spatial mean is subtracted from each image to form a zero-centered vector. With just a few additional K-L modes included, however, the subtraction improves dramatically. Also notice the bowl-shaped profile of all the contrast curves, as they curl up near the inner and outer limits of angular separation. This is a consequence of the correction factor we applied for coronagraph PSF throughput, as discussed above: along the edge of the FoV, the planet signal is truncated by the coronagraph, and the detection limit there is penalized accordingly based on our off-axis PSF model.

4.3 Residual noise analysis

As we noted before, the broadband speckle pattern exhibits a relatively rapid decorrelation with time (as measured by the correlation coefficients, which are 0.77 between broadband frames 1 and 51, versus 0.97 between narrowband frames 1 and 51). This rapid degradation in wavefront motivated to examine the useful timespan of the reference image library. To explore this, we restricted the target data subset to a single exposure (frame number 51) and carried out the KLIP subtraction with varying reference library sizes. We started with only the nearest reference image (frame 50) and sequentially added more, looking further back in time. In Figure 17 we show how the resulting inner sub-region ($2.5\text{--}4.5 \lambda/D$) contrast gain varies as a function of reference library look-back duration. Interestingly, the contrast gain improves all the way to the most distant reference images, despite their lower correlations. With 10 K-L modes the gain factor increases from $11\times$ with a single neighboring reference frame up to $14\times$ with the full 50-frame reference library. Additional analysis is needed to understand why this is the case. One possible explanation for why the KLIP subtraction benefits from the history captured in the larger reference library is that the temporal variation in the wavefront is quasi-periodic.

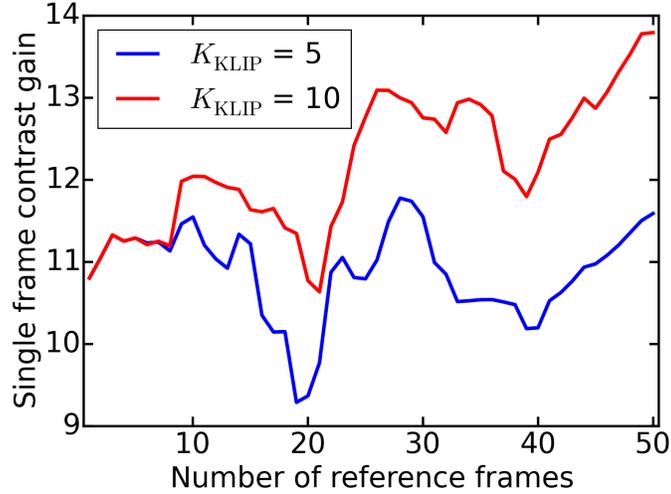


Figure 17: Contrast gain in the inner measurement zone ($2.5\text{--}4.5 \lambda_0/D$) of a single broadband target frame, as a function of the reference library lookback span. The contrast gain is highest when the full 50-frame reference library is used, and relatively poor when the reference library is restricted to a subset of frames immediately preceding the acquisition of the target frame.

	Narrowband	Broaband
Detector (uncorrelated read noise and dark/bias errors)	1.82×10^{-10}	6.38×10^{-11}
Photon counting (due to raw speckles)	8.49×10^{-10}	3.13×10^{-10}
Overall noise floor (1σ)	2.14×10^{-9}	9.58×10^{-10}

Table 9: Comparison of noise contributions in an individual KLIP-subtracted target frame, listed in contrast units. In an individual KLIP-subtracted target frame, the effects of detector noise and photon counting noise are small compared to residual speckles, which dominate the overall measured noise floor. Note the latter quantity is not corrected for planet throughput, to enable a more direct statistical comparison.

A closely related question is how the detection limits are affected by the number of target images we combine. Or equivalently, whether or not we can place a bound on the useful integration time spent on a scientific target. One factor to consider here is the relative contributions from the various uncorrelated noise sources that impact the detection limits on top of residual speckles. In this regard, we isolated and quantified two effects in Table 9: the contrast floor set by raw speckle photon counting noise, and separately, the uncorrelated detector noise due to read noise and errors in dark/bias subtraction. We measured the residual detector noise by examining the spatial standard deviation in median-subtracted dark frames. Forming a one-dimensional, empirical metric of the photon counting noise floor is slightly more nuanced. In this case, we looked at the raw counts in the 10 original target frames, at the 4×4 pixel binning used in our post-processing analysis. Each of those values defines the expectation count of an independent random variable with a Poisson distribution. We then computed the spatial variance that would result from the corresponding ensemble of random variables. We found that this raw speckle photon noise is the predominant

uncorrelated noise source in an individual frame. In both data sets, the raw speckle photon noise sets a contrast floor almost a factor of $5\times$ above the detector noise floor. Importantly, this uncorrelated noise floor is still a factor of ~ 3 below the overall one we measure in an individual KLIP-subtracted frame. Therefore, the detection limit of an individual exposure was largely determined by correlated, residual speckles rather than any uncorrelated, random process we could identify.

With this assessment of the relative proportions of the noise contributions in mind, we examined how the contrast floor varied for each of the 10 individual target frames, and as a function of the number of combined target frames. In Figure 18 we plot the 1σ contrast floor measured in the inner sub-region ($2.5\text{--}4.5 \lambda/D$) with K_{KLIP} at 5 and 10 K-L modes. We see that the contrast floor of the KLIP-subtracted “running average” does not improve beyond the third combined frame, and in fact deteriorates as we approach 10 frames. This can be explained by the predominance of residual speckle noise that gets stronger in each successive frame (due to the gradual decorrelation with respect to the reference library), and which furthermore does not average out over time like an uncorrelated noise process would.

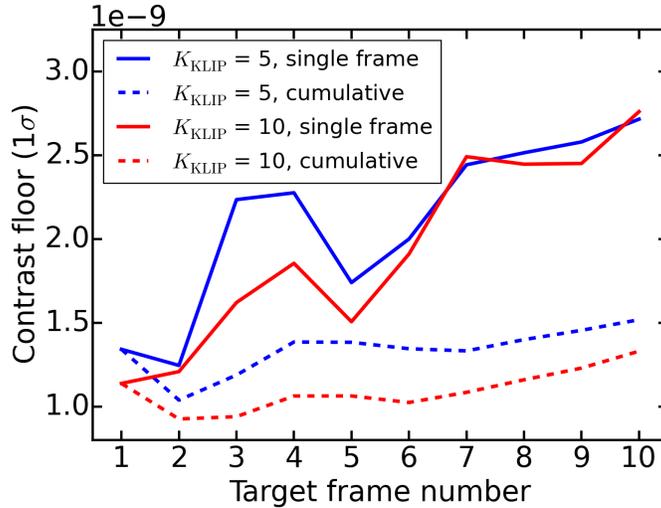


Figure 18: Broadband contrast floor measured in the inner field-of-view region ($2.5\text{--}4.5 \lambda_0/D$) of individual KLIP-subtracted target frames. We compare the single-frame contrast floor (solid lines) to the same measurement repeated on a cumulative “running average” of KLIP-subtracted target frames (dashed lines). Interestingly, the final contrast does not improve beyond the combination of 2 to 3 target frames. This is a consequence of two effects: (i) the relative dominance of residual speckles over uncorrelated noise sources (Table 9); and (ii) the gradual decorrelation of target frames with respect to the reference library. As a consequence, for this particular RDI scenario and SNR regime, we find that combining multiple target frames offers little or no benefit to detection sensitivity.

At the beginning of this section, in Table 7, we translated the light source intensity observed in the laboratory to stellar magnitude, given our best knowledge of the combined telescope and instrument throughput at 550 nm. We also listed the effective flight integration times per frame to measure the same speckle intensity if we were observing the reference star (β UMa) or scientific target (47 UMa) during our canonical RDI observing scenario. For 47 UMa, the corresponding

per-frame integration times are well over 100 h. Therefore, for the purpose of comparing this data set to a realistic flight observing scenario, it is encouraging that we can demonstrate a 10^{-9} contrast floor with a single target frame. Furthermore, based on the noise floors in Table 9, we can speculate that the laboratory integration time could have been reduced by a factor of 10, so that the photon noise level would be roughly equal to the correlated residuals, and we would still achieve nearly the same contrast floor in $\frac{1}{10}$ of the current integration time.

4.4 RDI with simulated planets

As a cross-check on the post-processed contrast gains and detection floors stated above, for one final illustration we inserted fake planets at two positions in a single raw target image (frame number 51). Figure 19 shows the residual images after KLIP subtraction over the full field of view with 20 K-L modes. The planet PSFs were computed using our WebbPSF/WFIRST off-axis PSF model, which in turn defined a Poisson random variable distribution to draw a photon counting realization from. We modeled the planets at a contrast of 3×10^{-9} , corresponding to $\text{SNR} \simeq 3\sigma$. The planets are placed at angular separation $3.5 \lambda/D$ from the star, first in the center of the left-hand side of the FoV, and second in the center of the right-hand side of the FoV. The expectation count in the peak pixel of the planet PSF is 574 ADU before KLIP subtraction. After the random photon count realizations and KLIP subtractions, the peak pixel counts in the left and right residual planet PSFs are 455 ADU and 590 ADU, respectively.

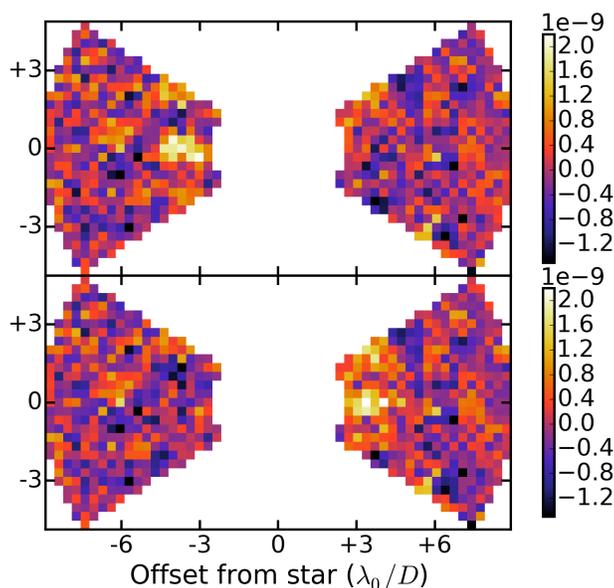


Figure 19: An individual, broadband, KLIP-subtracted target frame (exposure number 51) with a fake planet inserted at contrast 3×10^{-9} ($\text{SNR} \simeq 3$) and angular separation $3.5 \lambda_0/D$ from the star. In the top image, the planet was placed in the left half of the bowtie field of view; in the bottom image the planet was placed in the right half. The planet PSF model includes photon counting noise, and was added to the target image before applying KLIP subtraction with 20 K-L modes over the full field of view using the 50-frame reference library.

5 Conclusions

We applied speckle subtraction algorithms, including classical PSF subtraction and the PCA-based KLIP algorithm, to analyze simulated and laboratory WFIRST coronagraph images. Building on our previous work [7, 8, 9, 10], in this report we covered the following topics:

- (i) RDI and ADI post-processing applied to a new simulated HLC data set for various noise and wavefront control scenarios;
- (ii) evaluation of the impact of spatial sampling on post-processing, motivated by the possibility of relaxing the IFS lenslet spatial sampling;
- (iii) reduction and analysis of new SPC testbed data.

Below, we list the conclusions from each part of our investigation.

(i) RDI and ADI post-processing with simulated HLC data. We analyzed the OS5 simulation under various assumptions: speckle-noise and photon noise dominated images, and with and without LOWFC. After analyzing the aberration stability and testing both classical PSF subtraction and KLIP for both the RDI and ADI strategies, we concluded that:

- owing to the relatively deep raw contrast ($\sim 2 \times 10^{-9}$) observed in these HLC simulations, in all cases our 5σ post-processed detection limits are well below 10^{-9} , and in the most idealized cases reach to near 10^{-10} .
- in presence of LOWFC, the level of correlation between the reference and target images is high enough (> 0.99) such that classical PSF subtraction leads to similar levels of contrasts as those obtained with KLIP;
- an important omission from these simulations is spacecraft jitter. We expect classical subtraction to be less effective in the case of less stable pointing, whereas KLIP should be able to mitigate the contrast degradation;
- two opposing regimes of post-processing behavior can be identified in the various data versions: without photon noise (“noiseless” images) the contrast gains are a factor of 3 to 10 times higher than the case where speckle noise is below the photon noise level (“noisy” images).

We compared separately the impact of the RDI and ADI strategies on the contrast gain after post-processing with classical PSF subtraction. We repeated these tests in the speckle noise- and photon noise-dominated image versions. We found that:

- the ADI observing strategy is always as effective as RDI and up to 3 times better than RDI for some OS5 simulation cases;

- in the speckle-noise dominated regime, ADI is better than RDI at subtracting both low-order and high-order aberration residuals. The ADI contrast gain increases dramatically for separations higher than $5\lambda/D$ to reach about 30× in the case without LOWFC;
- in the photon-noise dominated regime, we observe slightly better gains with ADI at all separations due to the gain in signal-to-noise ratio.

We thus recommend that WFIRST retain the operational capability to carry out rolled coronagraph observations by at least $\pm 13^\circ$ off nominal.

(ii) Impact of spatial sampling on post-processing performance. Our analysis showed that:

- within the limited noise components included in the OS5 simulations (no IFS image reconstruction noise, no systematic detector noise), relaxing the spatial sampling does not degrade the post-processed contrast floor;
- at a spatial sampling equal to or above the Nyquist rate (2 pixels per λ/D) we can apply arbitrary alignment offsets to reference library images without adding significant noise.

(iii) SPC testbed data. Our analysis of “drift” laboratory data acquired on the SPC testbed at HCIT achieved an encouraging RDI/KLIP contrast gain of 14× in broadband light (10%) to reach a 1σ contrast floor of 1.3×10^{-9} at angular separations 2.5–4.5 λ/D and 9.7×10^{-10} averaged over the full field of view. In the course of our analysis we also established that:

- in the speckle-dominated noise regime seen in this data, we reach the same contrast floor in a single frame treated as the entirety of the science observation, versus a combined set of many frames;
- KLIP subtraction benefits from the history of wavefront diversity captured in a large reference image library, even when images with relatively poor correlation coefficients (< 0.8) are included.
- the narrowband version of the “drift” data demonstrated a high level of wavefront stability (frame-to-frame correlation coefficients ≥ 0.97) and the corresponding RDI/KLIP test scenario achieved a contrast gain factor of 35× averaged over the field of view.

6 Acknowledgments

This WFIRST study was carried out at Space Telescope Science Institute (STScI) under subcontract No. 1506553, funded by the Jet Propulsion Laboratory (JPL), which is managed for NASA by the California Institute of Technology. The authors are grateful to John Krist for producing the data simulations used in this study, to Gautam Vasisht and Eric Cady for assistance with the SPC testbed data, and to Bertrand Mennesson and Bijan Nemati for helpful discussions.

7 Acronym list

ADI - Angular Differential Imaging
 ADU - Analog to Digital Units
 AFTA - Astrophysics Focused Telescope Assets
 CGI - CoronaGraph Instrument
 CPSFTS - Coronagraphic PSF Template Subtraction
 DAC - Digital to Analog Converter
 DM - Deformable Mirror
 FoV - Field of View
 FPF - False Positive Fraction
 FY - Fiscal Year
 GSFC - Goddard Space Flight Center
 HCIT - High-Contrast Imaging Testbed
 HLC - Hybrid Lyot Coronagraph
 HST - Hubble Space Telescope
 IFS - Integral Field Spectrograph
 JPL - Jet Propulsion Laboratory
 JWST - James Webb Space Telescope
 KLIP - Karhunen-Loève Image Projection Algorithm
 LOCI - Locally Optimized Combination of Images
 LOWFC - Low-Order Wavefront Control
 LOWFSC - Low-Order Wavefront Sensing and Control
 OMC - Occulting Mask Coronagraph
 OS - Observing Scenario
 PCA - Principal Component Analysis
 PDF - Power Density Function
 PSF - Point Spread Function
 RDI - Reference Star Differential Imaging

RMS - Root Mean Square
 ROC - Receiver Operating Characteristic
 SNR - Signal-to-Noise Ratio
 SPC - Shaped Pupil Coronagraph
 STScI - Space Telescope Science Institute
 TPR - True Positive Fraction
 WFIRST - Wide-Field InfraRed Survey Telescope

References

- [1] D. Spergel et al. “Wide-Field Infrared Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report”. In: *ArXiv e-prints* (Mar. 2015). arXiv: [1503.03757](https://arxiv.org/abs/1503.03757) [astro-ph.IM].
- [2] S. Shaklan et al. “The AFTA coronagraph instrument”. In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 8864. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Sept. 2013, p. 15. DOI: [10.1117/12.2024560](https://doi.org/10.1117/12.2024560).
- [3] I. Y. Poberezhskiy et al. “WFIRST-AFTA coronagraph starlight suppression technology progress”. In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 9904. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. 2016.
- [4] F. Shi et al. “Low-order wavefront sensing and control for WFIRST coronagraph”. In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 9904. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. 2016.
- [5] R. T. Demers. “The WFIRST coronagraph instrument: from technology advancement to flight design”. In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 9904. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. 2016.
- [6] J. T. Trauger, D. C. Moody, and J. E. Krist. “Managing the optical wavefront for high contrast exoplanet imaging with the WFIRST-AFTA coronagraph”. In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 9904. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. 2016.
- [7] M. Ygouf et al. “Data processing and algorithm development for the WFIRST-AFTA coronagraph: reduction of noise free simulated images, analysis and spectrum extraction with reference star differential imaging”. In: *Techniques and Instrumentation for Detection of Exoplanets VII*. Vol. 9605. Sept. 2015, 96050S. DOI: [10.1117/12.2188669](https://doi.org/10.1117/12.2188669).
- [8] M. Ygouf et al. *Data Post Processing and Algorithm Development for the WFIRST-AFTA Coronagraph: First Progress Report*. Tech. rep. WFIRST-STScI-TR1503. 2015. URL: <http://www.stsci.edu/wfirst/technicalreports/WFIRST-STScI-TR1503A.pdf>.
- [9] B. Mennesson et al. “WFIRST-AFTA coronagraph performance: feedback from post-processing studies to overall design”. In: *Techniques and Instrumentation for Detection of Exoplanets VII*. Vol. 9605. Proc. SPIE. Sept. 2015, p. 96050D. DOI: [10.1117/12.2189403](https://doi.org/10.1117/12.2189403).

- [10] M. Ygouf et al. *Data Post-Processing and Algorithm Development for the WFIRST-AFTA Coronagraph: FY2015 Report*. Tech. rep. WFIRST-STScI-TR1601. 2016. URL: <http://www.stsci.edu/wfirst/technicalreports/WFIRST-STScI-TR1601A.pdf>.
- [11] John Trauger et al. “Hybrid Lyot coronagraph for WFIRST-AFTA: coronagraph design and performance metrics”. In: *Journal of Astronomical Telescopes, Instruments, and Systems* 2.1 (2016), p. 011013. DOI: [10.1117/1.JATIS.2.1.011013](https://doi.org/10.1117/1.JATIS.2.1.011013). URL: <http://dx.doi.org/10.1117/1.JATIS.2.1.011013>.
- [12] N. T. Zimmerman et al. “Shaped pupil Lyot coronagraphs: high-contrast solutions for restricted focal planes”. In: *Journal of Astronomical Telescopes, Instruments, and Systems* 2.1, 011012 (Jan. 2016), p. 011012. DOI: [10.1117/1.JATIS.2.1.011012](https://doi.org/10.1117/1.JATIS.2.1.011012). arXiv: [1601.05121](https://arxiv.org/abs/1601.05121) [astro-ph.IM].
- [13] J. E. Krist et al. “An overview of WFIRST/AFTA coronagraph optical modeling”. In: *Techniques and Instrumentation for Detection of Exoplanets VII*. Vol. 9605. Proc. SPIE. Sept. 2015, p. 960505. DOI: [10.1117/12.2188361](https://doi.org/10.1117/12.2188361).
- [14] C. Marois et al. “Angular Differential Imaging: A Powerful High-Contrast Imaging Technique”. In: *ApJ* 641 (Apr. 2006), pp. 556–564. DOI: [10.1086/500401](https://doi.org/10.1086/500401). eprint: [arXiv:astro-ph/0512335](https://arxiv.org/abs/astro-ph/0512335).
- [15] R. Soummer et al. “Orbital Motion of HR 8799 b, c, d Using Hubble Space Telescope Data from 1998: Constraints on Inclination, Eccentricity, and Stability”. In: 741, 55 (Nov. 2011), p. 55. DOI: [10.1088/0004-637X/741/1/55](https://doi.org/10.1088/0004-637X/741/1/55). arXiv: [1110.1382](https://arxiv.org/abs/1110.1382) [astro-ph.EP].
- [16] A. Rajan et al. “Detection and characterization of the atmospheres of the HR 8799 b and c planets with high contrast HST/WFC3 imaging”. In: *American Astronomical Society Meeting Abstracts*. Vol. 225. American Astronomical Society Meeting Abstracts. Jan. 2015, #323.07.
- [17] E. Choquet et al. “Archival Legacy Investigation of Circumstellar Environments using KLIP algorithm on HST NICMOS coronagraphic data”. In: *IAU Symposium*. Ed. by M. Booth, B. C. Matthews, and J. R. Graham. Vol. 299. IAU Symposium. Jan. 2014, pp. 30–31. DOI: [10.1017/S1743921313007722](https://doi.org/10.1017/S1743921313007722).
- [18] J. H. Debes et al. “Wide-Field Infrared Survey Telescope-Astrophysics Focused Telescope Assets coronagraphic operations: lessons learned from the Hubble Space Telescope and the James Webb Space Telescope”. In: *Journal of Astronomical Telescopes, Instruments, and Systems* 2.1, 011010 (Jan. 2016), p. 011010. DOI: [10.1117/1.JATIS.2.1.011010](https://doi.org/10.1117/1.JATIS.2.1.011010). arXiv: [1511.06277](https://arxiv.org/abs/1511.06277) [astro-ph.IM].
- [19] J. E. Krist. “End-to-end numerical modeling of AFTA coronagraphs”. In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 9143. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Aug. 2014. DOI: [10.1117/12.2056759](https://doi.org/10.1117/12.2056759).
- [20] G. Schneider et al. “NICMOS Coronagraphy: Recalibration and the NICMOS Legacy Archive PSF Library”. In: *Hubble after SM4. Preparing JWST*. July 2010, p. 15.
- [21] R. Soummer, L. Pueyo, and J. Larkin. “Detection and Characterization of Exoplanets and Disks Using Projections on Karhunen-Loève Eigenimages”. In: 755, L28 (Aug. 2012), p. L28. DOI: [10.1088/2041-8205/755/2/L28](https://doi.org/10.1088/2041-8205/755/2/L28). arXiv: [1207.4197](https://arxiv.org/abs/1207.4197) [astro-ph.IM].
- [22] Wesley A. Traub et al. “Science yield estimate with the Wide-Field Infrared Survey Telescope coronagraph”. In: *Journal of Astronomical Telescopes, Instruments, and Systems* 2.1 (2016), p. 011020. DOI: [10.1117/1.JATIS.2.1.011020](https://doi.org/10.1117/1.JATIS.2.1.011020). URL: <http://dx.doi.org/10.1117/1.JATIS.2.1.011020>.

- [23] Eric Jones, Travis Oliphant, Pearu Peterson, et al. *SciPy: Open source scientific tools for Python*. [Online; accessed 2016-07-15]. 2001-. URL: <http://www.scipy.org/>.
- [24] M. D. Perrin et al. "Simulating point spread functions for the James Webb Space Telescope with WebbPSF". In: *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*. Vol. 8442. Proc. SPIE. Sept. 2012, p. 84423D. DOI: [10.1117/12.925230](https://doi.org/10.1117/12.925230).
- [25] M. D. Perrin et al. "An Update on Simulating Imaging, Spectroscopic, and Coronagraphic PSFs for JWST (and WFIRST too!)" In: *American Astronomical Society Meeting Abstracts*. Vol. 228. American Astronomical Society Meeting Abstracts. June 2016, #216.20.