

Roman Coronagraph Instrument Post Processing Report - OS9 SPC Distribution

March 5, 2022

Marie Ygouf¹, Neil Zimmerman², and Vanessa Bailey¹

1. Jet Propulsion Laboratory, California Institute of Technology
2. Goddard Flight Space Center

Contents

1	Introduction	1
1.1	Key Definitions	1
2	Data Description	1
2.1	OS9 Simulation	1
2.2	Important note on the OS9 SPC data set	2
2.3	Differences with previous observing scenarios	4
3	Simulation of Spectroscopic Data	4
3.1	Scene generation	5
3.1.1	Target simulation	5
3.1.2	Planet injection	5
3.1.3	Normalization	6
3.1.4	Slit spectroscopy	6
3.2	Detector	10
4	Calibration and Data Reduction	10
4.1	Spectrum Calibration	10
4.2	Data Reduction	11
4.2.1	Procedure	11
4.2.2	Factor Above Classical Evaluation	12
4.2.3	Spectrum extraction	13
5	Conclusions and Next Steps	15

6 Acknowledgment	16
7 References	16

1 Introduction

This document reports on findings from applying post-processing techniques to the Observing Scenario 9 - Shaped Pupil Coronagraph (SPC) simulated data for the Roman Space Telescope Coronagraph Instrument. In particular, this analysis focuses on the prism-dispersed spectroscopy observations. This is an on-going work and thus the information contained in this document is subject to future updates. In the meantime, please send questions and/or comments to marie.ygouf@jpl.nasa.gov.

1.1 Key Definitions

ADI: Angular Differential Imaging

CGI: Coronagraphic Instrument

DM: Deformable Mirror

EFC: Electric Field Conjugaison

EMCCD: Electron Multiplying Charge Coupled Device

FAC: Factor Above Classical (also called "post-processing factor (fpp)" in some documents)

FPM: Focal Plane Mask

FSAM: Field Stop Alignment Mechanism

GSFC: Goddard Space Flight Center

HLC: Hybrid Lyot Coronagraph

JPL: Jet Propulsion Laboratory

LOWFS/C: Low Order Wavefront Sensing and Control

MUFs: Model Uncertainty Factors

OS: Observing Scenario

PCA: Principal Component Analysis

PSF: Point Spread Function

QE: Quantum efficiency

RDI: Reference Star Differential Imaging

SPC: Shaped Pupil Coronagraph

STOP: Structural Thermal Optical Performance

WFI: Wide Field Instrument

2 Data Description

2.1 OS9 Simulation

The Observing Scenario (OS) image time sequences are generated by John Krist and the integrated modeling team at JPL to create simulated CGI data for the Shaped Pupil Coronagraph (SPC) which includes the most updated concept of operations observing strategy. For this work, we used the most recent OS dataset made public on May 6th 2020, called OS9 ¹, with the broadband SPC-20190130 (Phase B flight design) in a 15% bandpass filter centered at 730 nm (Band 3: 675 - 785 nm). The target star used for these simulations is 47 UMa

¹https://roman.ipac.caltech.edu/sims/Coronagraph_public_images.html#CGI_OS9

($V=5.0$ mag, G1V) and the reference star is ζ Pup ($V=2.25$ mag, O4I). This is the first time this star is used as a reference in one of the Roman observing scenarios. Although being located 90 degrees away from 47 UMa, this new reference star was selected because it is only 3.5 degrees different in solar pitch from 47 UMa (see paragraph ?? for more information). The stellar spectra of the two stars (G1V and O4I), and stellar diameter are included.

The observing sequence begins with a slew from a Wide Field Instrument (WFI) high latitude survey target to the reference star, followed by an observatory settling time of 30 hours. Assuming that the dark hole was previously dug at some earlier time, the observation cycle starts with 4 hours of dark hole touchup or Electric Field Conjugaison (EFC) maintenance on ζ Pup. One EFC iteration is performed every 2 hours for a total of 2 iterations over each 4 hours touch-up span. Then, one hour is spent on the imaging of ζ Pup before slewing to the target star 47 UMa. The target is observed for about 2 hours at rolls angles of -11° , $+11^\circ$, -11° , $+11^\circ$ relative to the solar-normal roll before slewing back to ζ Pup for about one hour of imaging. This observation cycle is repeated 14 times. Figure 1 shows a visualization of the observing time series is presented in OS9 for the SPC. Note that the data corresponding to the EFC maintenance at the beginning of each cycle is not included in the OS9 distribution.

The STOP model was run with 20 min timesteps to produce aberration timeseries and results were interpolated to 5 min timesteps. The STOP model assumed instantaneous slews, but the jitter model included slew/roll times. Slew times were taken out of STOP on-target times, resulting in an irregular number of timesteps on each star/roll. These issues will be fixed in future scenarios.

These simulations include wavefront error changes from thermal drift, errors from Low Order Wavefront Sensing and Control (LOWFS/C) (Z4 – Z11 LOWFS sensed & corrected with measurement & DM errors), errors from beam shears, Deformable Mirror (DM) thermal drift (from mini-STOP CGI-only model) and pointing jitter (LOWFS-corrected), all optical aberrations, including polarization and Focal Plane Mask (FPM) fabrication errors.

Those simulations were produced with optical model uncertainties. Of the previous versions of OS data, only OS6 did include MUFs. The optical MUFs are: Polarization aberrations = $1.5\times$, contrast sensitivity to low-order aberrations = $2.0\times$ (includes sensitivity to polarization and wavefront jitter), starting contrast = $2.0\times$, structural deformation = $2.0\times$ (includes beam shear increases and wavefront error drift), and frequency dependent jitter = $3.0\times$ (<20 Hz), $4.27\times$ (40-100 Hz) and $8.0\times$ (>100 Hz) (in Phase A, OS6: $2.07\times$ (<20 Hz), $4.27\times$ (>40 Hz)). Table 1 summarizes the fits files that contains the OS9 SPC dataset images.

Note that those data does not include noise and they were not fed into a EMCCD model like the OS9 Hybrid Lyot Coronagraph (HLC) data. No planets have been injected in the SPC data.

2.2 Important note on the OS9 SPC data set

In the most recent (OS9) SPC time series, the residual starlight pattern between reference and target star is stable to better than 3%, in terms of RMS flux bias on an extracted spectrum at $4\lambda/D$ separation. After classical subtraction this translates to about 1.6×10^{-9} flux ratio noise floor, ignoring all other noise sources. If the residual starlight is this

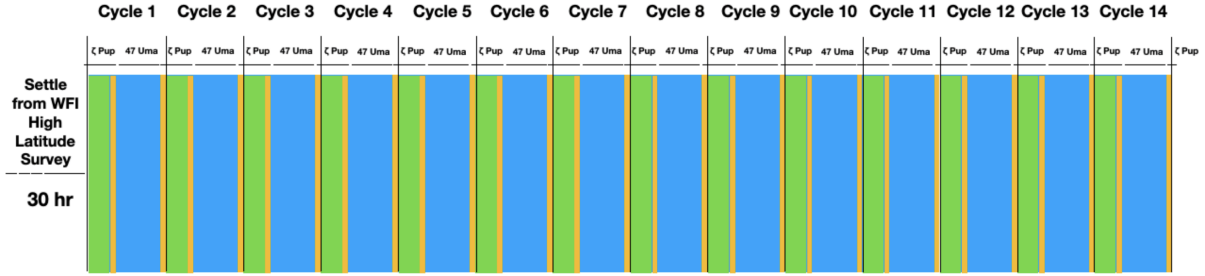


Figure 1: Visualization of the observing sequence simulated in the OS9 SPC PSF time series. The sequence begins with a slew from a WFI high latitude survey followed by an observatory settling time of 30 hours. Then Cycle 1 begins with 4 hours of EFC maintenance on the reference ζ Pup (green), about 1 hour observing ζ Pup (yellow) and about 8 hours observing the target 47 UMa (blue) at two different roll angles before going back to the reference for an additional hour (yellow). This cycle is repeated 14 times in total. Data corresponding to the EFC maintenance are not included in the OS9 distribution.

File name	MUFs	Noise	Planets	Flux Units
os9_images_cycle1.fits	✓	-	-	average flux units
os9_images_cycle2.fits	✓	-	-	average flux units
os9_images_cycle3.fits	✓	-	-	average flux units
os9_images_cycle4.fits	✓	-	-	average flux units
os9_images_cycle5.fits	✓	-	-	average flux units
os9_images_cycle6.fits	✓	-	-	average flux units
os9_images_cycle7.fits	✓	-	-	average flux units
os9_images_cycle8.fits	✓	-	-	average flux units
os9_images_cycle9.fits	✓	-	-	average flux units
os9_images_cycle10.fits	✓	-	-	average flux units
os9_images_cycle11.fits	✓	-	-	average flux units
os9_images_cycle12.fits	✓	-	-	average flux units
os9_images_cycle13.fits	✓	-	-	average flux units
os9_images_cycle14.fits	✓	-	-	average flux units

Table 1: Description of OS9 SPC Datasets. Each file contains a $300\text{px} \times 300\text{px} \times 29$ wavelengths \times ntimes array of monochromatic speckle field images. Only simulations with the sensitivity MUFs are included in this distribution. No noise or planets are included in those data.

stable, it will require very long integration times, greater than 400 hours (for a nominal $V=5$ target), before post-processing algorithm performance limits the sensitivity. Therefore, we can already predict that the numbers we extract from SPC spectroscopy KLIP trials are not likely to impact time-to-SNR calculations for tech demo targets.

The wavefront errors may behave very differently during the on-sky operations from what the integrated model PSF time series have shown us so far. If we have brighter speckles with more time variability, then we will need KLIP or similar processing to maximize planet SNR.

Given all that, this analysis shows a good preliminary investigation, but the input simulation is probably not in the right performance regime to get meaningful results.

2.3 Differences with previous observing scenarios

For differences with previous observing scenarios, we invite the reader to consult the Roman Coronagraph Instrument Post Processing Report for the OS9 HLC Distribution².

3 Simulation of Spectroscopic Data

The OS9 SPC distribution contains noiseless data with SPC bowtie, as they would appear if imaged by a detector. Because no data dispersion is performed in OS9 we need to perform it independently. In addition, the OS9 SPC data does not include planets or detector model. This section describes how those features were added to the data.

The simulation of spectroscopic data is performed with tools developed at the Goddard Space Flight Center by Neil Zimmerman and Hari Subedi (GSFC). The `zodprism_sim` script takes as input the Roman Coronagraph SPC speckle time series simulated by John Krist (JPL), and creates a star plus exoplanet point source scene based on a specified target star apparent mag, spectral type, exoplanet source position, and albedo profile. A slit mask is then applied to the image model, to approximate the slit in the field stop plane (FSAM). The mask is a long vertical slit with aperture width approximately $2\lambda/D$, matched to the half-max cross-section of the PSF. The script then applies the specified prism dispersion profile to the star plus exoplanet scene, computes the detector count rate map, and applies an approximate EMCCD detector model to simulate these 4 co-added data products:

- Occulted science target star with exoplanet
- Unocculted science target star
- Occulted reference star
- Unocculted reference star

²https://roman.ipac.caltech.edu/sims/Coronagraph_public_images.html#CGI_OS9_report

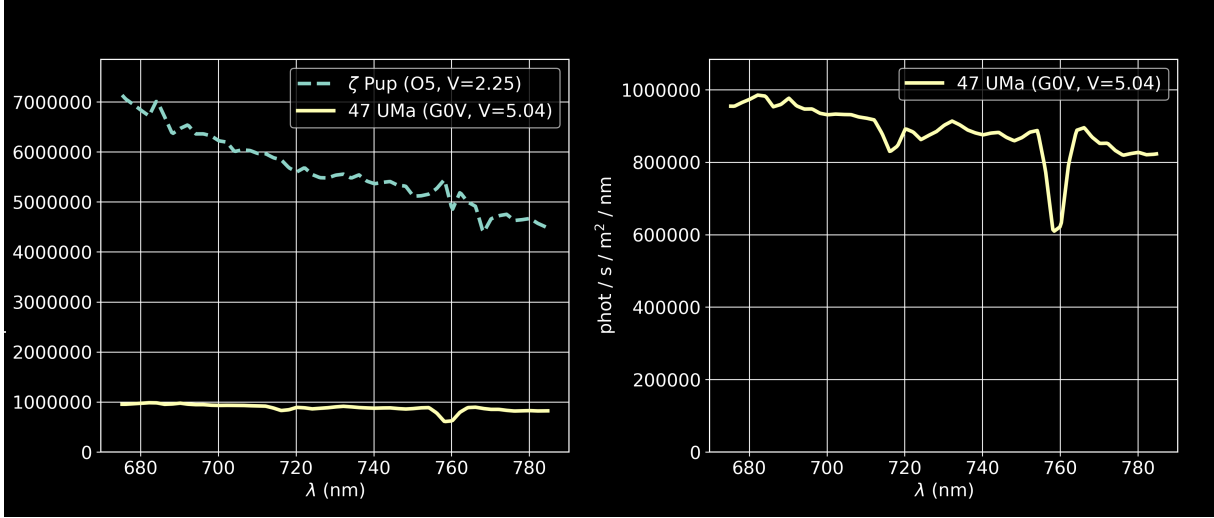


Figure 2: Stellar spectra for 47 UMa and ζ Pup. ζ Pup spectral type is O4I but O5 was used for those simulations because it is the closest available spectral type the in BPGS library of templates. Note the spectral feature at 760 nm in the spectrum of 47 UMa.

3.1 Scene generation

3.1.1 Target simulation

For these simulations we kept the target and reference simulated in OS9 but the SNR in the continuum spectral bin determines the integration time, which was set to 28.3 and 23.1 hours for the target and reference stars respectively.

The script does not have all spectral types and looks for the closest match. For example, ζ Pup spectral type is O4I but O5 is the closest available spectral type the in BPGS library of templates Bruzual-Persson-Gunn-Stryker Atlas³.

We first set the integration time and then add noise in the data an other assumptions. Only half the data set was used (either roll minus or roll plus). No Angular Differential Imaging (ADI) was simulated. The entire stack of speckle patterns was summed, integration time being a parameter. The script does not make use of the time steps described in the OS9 scenario and compute its own timesteps. Integration time was set so that we obtain a SNR of 20. The detector model includes dark current and CIC noise, and a uniform QE factor. Figure 2 represents the stellar spectra for 47 UMa and ζ Pup.

3.1.2 Planet injection

Planet injection was performed using the available unocculted PSFs that are included in the OS9 SPC distribution. Note that those unocculted PSFs were offset by various amounts in radius and angle from the center of the FPM over one quadrant of the dark hole. In order to simulate a planet at a specific location in the instrument field of view, we used the closest available PSF and shift it to the chosen planet location. Therefore, it is worth noting that

³<https://stsynphot.readthedocs.io/en/latest/stsynphot/appendixa.html#stsynphot-appendixa-bpgs>

the shape of the PSF might be slightly different from what it would be at this location but this effect is negligible and does not impact the results of this study.

For the planet spectrum, we chose a planet that matches the old requirement of being able to detect a planet at a contrast of 5×10^{-8} with a SNR of 10. This planet was injected at $4\lambda/D$. We used a solar system-like Jupiter with methane features model (Albedo simulations on the ipac website cahoy 2010⁴).

Data in the OS9 SPC distribution are provided at 29 wavelengths evenly sampled within the Band 3 bandpass. However the used spectrum model is provided at 301 wavelengths from 675 to 785 nm. An interpolation of the data cubes is needed before injecting the planet spectrum.

3.1.3 Normalization

It is important to compute photon rates and throughput correction factor for each wavelength.

Data that provided in the OS9 distribution does not account for the loss factor due to mirror reflections. Those optical losses do not depend on wavelength and are modeled by a simple scale factor. The factor capturing those reflection losses is computed using the Roman CGI ETC.

On the other hand, the throughput correction depends on the wavelength. The energy in the PSF model at each wavelength will not be the same and depends on the shape of the PSF and varies depending on the offset we have as well.

3.1.4 Slit spectroscopy

The SPC band 3 mode of the Roman Coronagraph uses slit spectroscopy. The slit was simulated as a simple binary mask with a size equal to the half max diameter of the PSF to cover exactly the PSF cross section. The current size might change in future simulations to match the final design. The simulation slit width has granularity set by the $0.1\lambda/D$ focal plane sampling, which results in a 5% disagreement from the 127 mas width of the baseline slit. In future versions of the code we may apply a "gray" partial transmission map to approximate finer steps in slit width.

The dispersion profile of the prism is then applied. Figure 4 represents the wavelength dispersion scale on the detector. The dispersion depends on the wavelength. This effect needs to be taken into account to properly disperse the flux while simulating the prism dispersion.

Figures 5 and 6 represent the noiseless dispersed noiseless countrate maps for the occulted and unocculted target and reference star, as well as for the injected planet before and after applying the slit mask and prism dispersion.

Figure 7 represents the noisy dispersed photoelectron count rates for the target and planet of flux ratio 5×10^{-8} at a separation of $4\lambda/D$, as well as the sum of the target plus planet before and after post-processing.

⁴https://wfirst.ipac.caltech.edu/sims/Exoplanet_Albedos.html

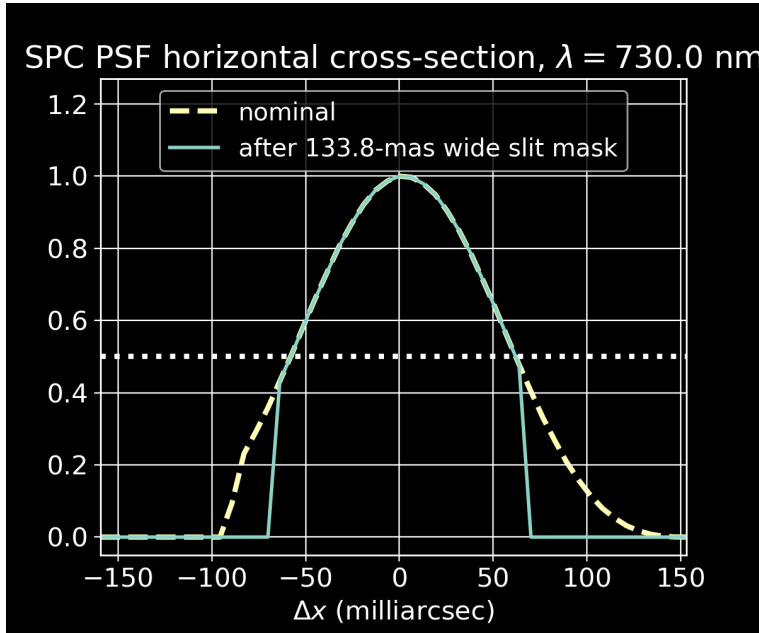


Figure 3: SPC PSF horizontal cross-section at $\lambda = 730\text{nm}$. The yellow curve is the nominal cross-section. The green curve is the cross-section after applying a 133.8-mas wide slit mask.

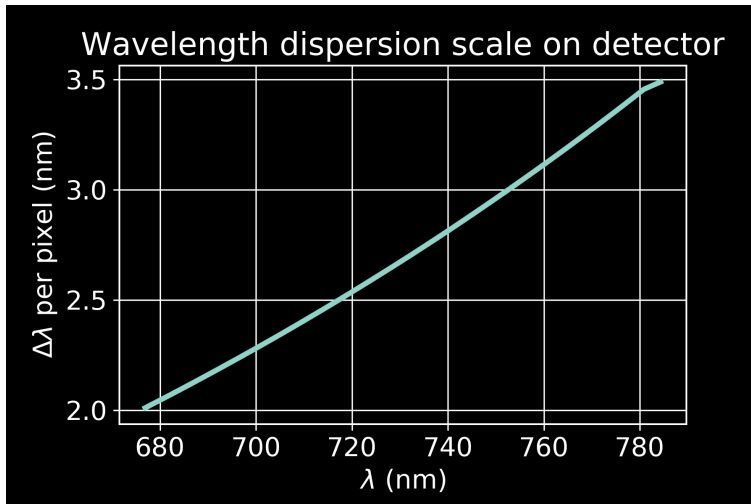


Figure 4: Wavelength dispersion scale on detector.

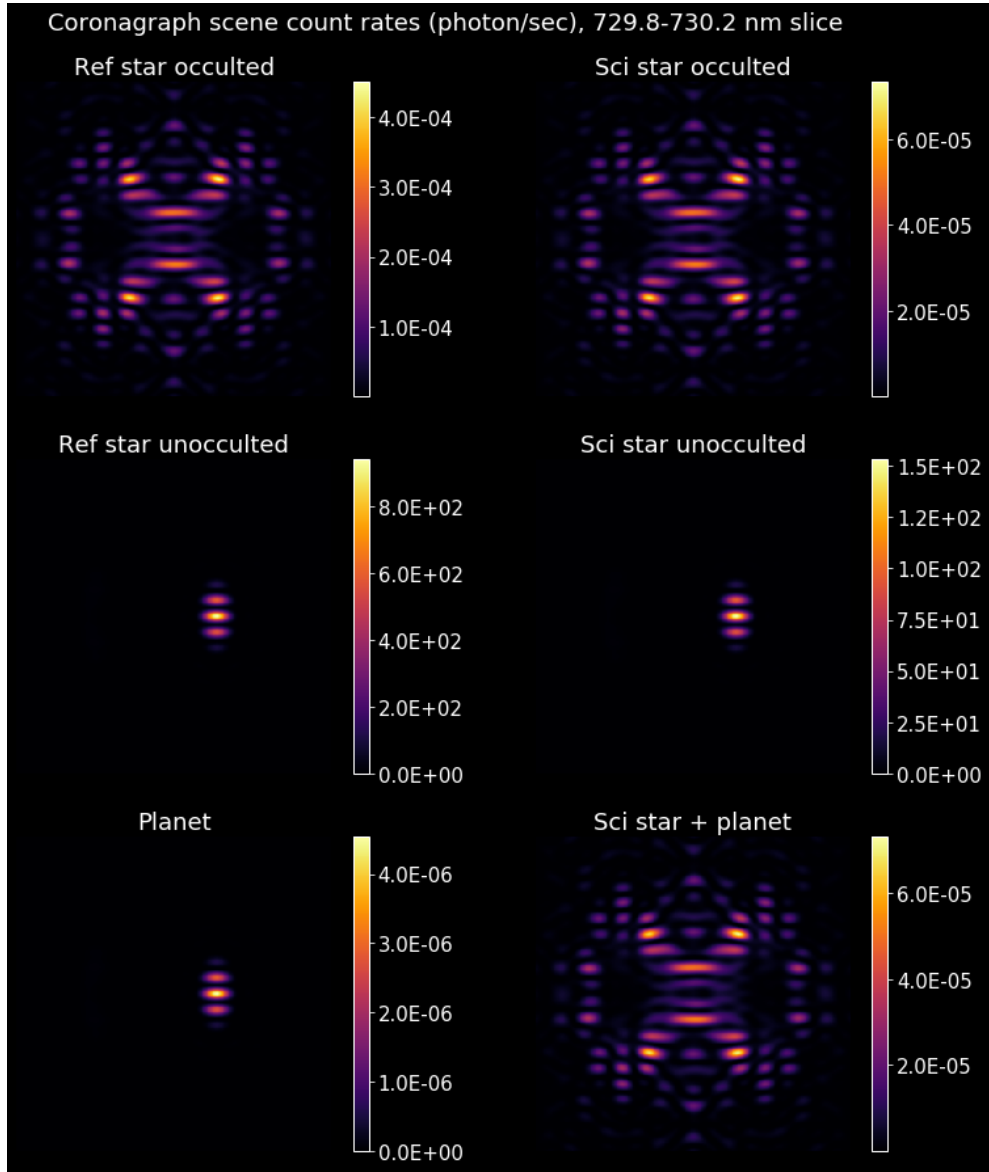


Figure 5: Countrate maps before dispersion.

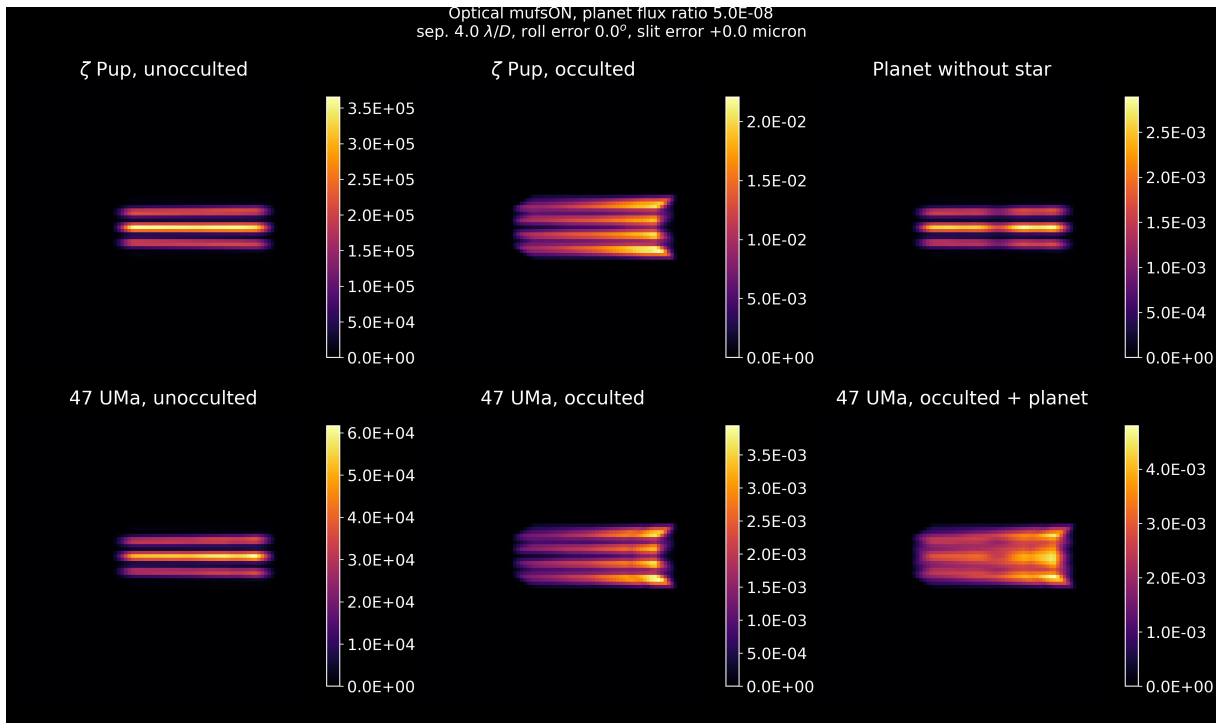


Figure 6: Dispersed noiseless countrate maps in the MUFs ON case. The planet contrast is 5×10^{-8} . Top panel, from left to right: unocculted reference star, occulted reference star and planet without star. Bottom panel from left to right: unocculted target star, occulted target star and occulted target + planet.

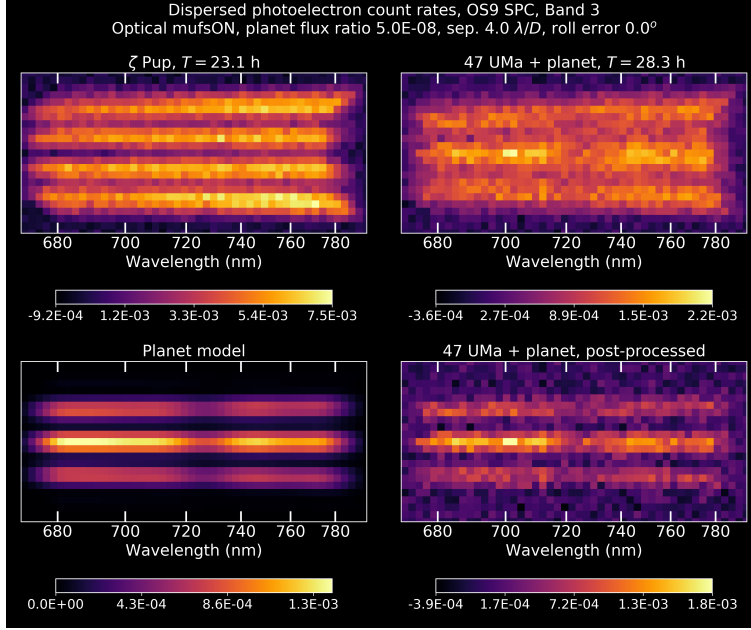


Figure 7: Dispersed photoelectron count rates after adding noise in the MUFs ON case. The planet contrast is 5×10^{-8} . Top left panel: reference star, top right panel: target star + planet, bottom left panel: planet model, bottom right panel: target star + planet, post-processed.

3.2 Detector

Data are simulated in pure analog mode (no photon counting mode). Although the photon counting processing is not simulated, a combined QE factor (borrowed from the ETC spreadsheet) was intended to account for the approximate losses for photon counting, in addition to the CCD’s estimated QE. The values for that are in the zodprism config file, along with the dark current and CIC noise parameters.

4 Calibration and Data Reduction

The zodprism_sim code models the calibration and reference differential imaging procedures, as well as the extraction of the flux ratio spectrum of the exoplanet point source. This part describes those procedures and present some preliminary results of data reduction and spectrum extraction.

4.1 Spectrum Calibration

Converting the spectrum measurement in count rate to flux ratio requires to measure the science target spectrum when it is unocculted. This measurement consists in recording how the PSF scales with wavelength is impacted by the slit. This calibration should be performed at the same location as the planet to avoid any flat field or detector issues. The calibrated

spectrum is then obtained by extracting the flux column by column through the slit and binning by a certain number of columns to obtain a calibrated spectrum.

`zodprism_sim` simulates this calibration with the caveat that the on-sky setup will be slightly different from the current concept for operations setup in `zodprism_sim`. During on-sky calibrations, the target will need to be observed through a ND filter to avoid the detector saturation. However, the ND filter is in the same plane as the slit mask, it will then not possible to apply the ND filter and the slit at the same time. When measuring the unocculted star, the slit will be removed. The profile of the PSF that gets dispersed will then be different without the slit being applied. The `zodprism_sim` simulations assume that this has only marginal effects on the dispersed pattern. However, this effect had not been quantified yet at the time of this study and will likely have an impact on the flux ratio calibration.

4.2 Data Reduction

We performed two different types of data reductions with different goals:

- Goal 1: Evaluate the Factor Above Classical (FAC) for Roman OS9 SPC Band 3 Spectroscopy data
- Goal 2: Demonstrate our ability to perform spectrum extraction

The Factor Above Classical (FAC), also called post-processing factor (fpp) in some other documents, is defined as the improvement in sensitivity over that achieved with classical RDI (or cRDI) PSF subtraction with a single observatory roll angle. This factor is defined for the noiseless case, because it is used in the CGI error budget as an assumed attenuation factor on the speckles only. The factor above classical is computed with respect to the single roll processed with classical subtraction. Before this work, this factor was assumed equal to 2.0 on the basis of the studies that had been done using Roman Coronagraph data in the imaging mode. This is the first time this factor is actually measured from SPC spectroscopy data.

Demonstrating our ability to perform spectrum extraction is an important goal because our ability to extract information such as the presence of methane or other chemical components in exoplanet atmospheres will rest on the quality of this spectrum extraction. In particular we want to understand what is the impact of different techniques of post-processing on this extraction. If the comparison of different techniques is out of the scope of this paper, we address the first part, which is to demonstrate our ability to extract a planet spectrum and how the noise added in the data impacts the extraction.

4.2.1 Procedure

We applied two different techniques of post-processing that take advantage of the Reference Star Differential Imaging (RDI) strategy. This strategy consists in observing a reference star to calibrate the stellar residuals. Note that the other strategy that is part of OS9, the Angular Differential Imaging (ADI) strategy, consists in calibrating the stellar residuals by observing the target at different rolls angles (e.g., +/- 11 degrees). Note that the `zodprism_sim` setup

used for those simulations does not incorporate the roll angles, ignoring an entire half the OS9 sequence. Thus, studying ADI and the corresponding post-processing techniques is out of the scope of this report.

Two families of techniques were explored:

- *Classical PSF subtraction RDI (cRDI)*. Our implementation of classical PSF subtraction uses linear regressions to minimize the least square error in the final image;
- *Principal Component Analysis (PCA)*. We used KLIP (Soummer et al. 2012), a Principal Component Analysis (PCA) technique that uses the reference cube as a library of reference point spread functions.

Figure ?? shows an example of subtracted data using cRDI without and with a cross-dispersion mask that limits the detector area to a small amount of rows (5 pixels wide) to minimize the impact of speckles.

4.2.2 Factor Above Classical Evaluation

In a previous version of `zodprism_sim`, only cRDI was applied, without splitting up time series from all cycles in different chunks. A new version of the code was then developed to independently process the 14 cycles of the OS9 SPC data and thus generating a library of 14 reference PSFs with the goal of implementing KLIP RDI. cRDI and KLIP RDI were implemented on noiseless dispersed data without planet.

A first implementation of KLIP was performed but without correcting from the algorithm throughput. When processing the data with Principal Component Analysis, the library of references is decomposed into principal components. A synthetic PSF is then generated by projecting the target into those principal components. The synthetic PSF is then subtracted from the data to obtain the final reduced image. However, when projecting the target into the basis of principal components (PC), hidden planets will project as well and may be partly subtracted. In fact, the more PC used for the decomposition, the planet signal on the reduced data. In order to account for this loss, a dispersed model of offset planet is propagated through the algorithm for different separations and position angles to evaluate the loss and apply it to the computed contrast gains.

Without taking into account the KLIP throughput, we find that KLIP RDI performs better than cRDI, however even in this favorable case the FAC is 1.22, which is well below the previously adopted factor of 2.

The KLIP throughput was computed by propagating a dispersed model of offset PSF at the location of the planet through the KLIP algorithm. With the KLIP throughput taken into account, the FAC is 0.76 and thus cRDI performs better than KLIP RDI. Note that results was obtained using 6 PC in total and that using less PC would give slightly better results (we obtained 0.80 using only one PC).

Discussion. We applied classical PSF subtraction (cRDI) and KLIP RDI on our simulated noiseless data in order to compute the factor above classical (FAC) OS9 SPC spectroscopy.

The KLIP throughput is computed by propagating a dispersed offset PSF at the location of the planet through the KLIP algorithm. With KLIP throughput taken into account, the

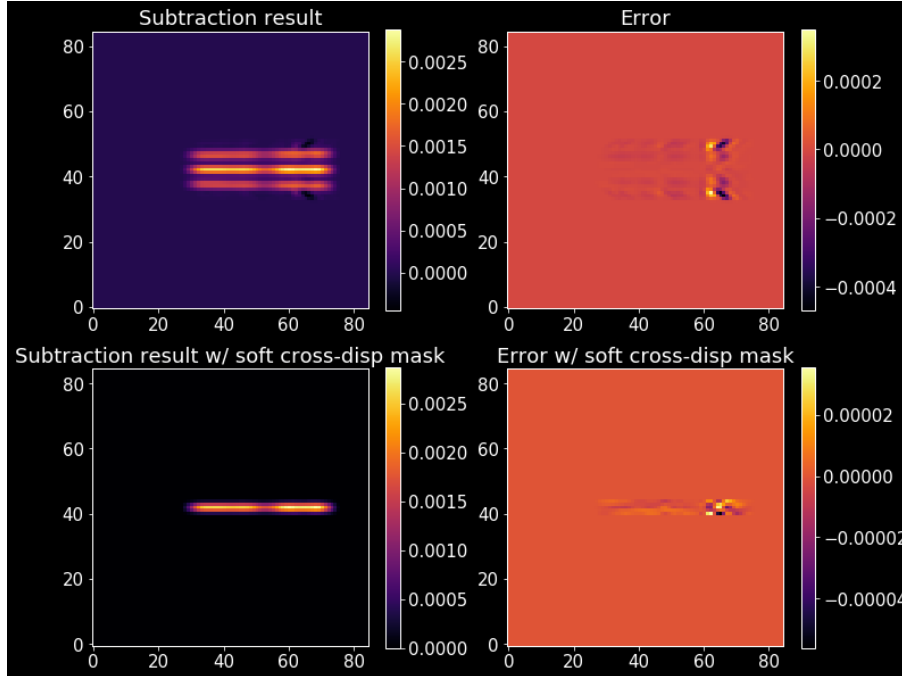


Figure 8: Noiseless data processed with classical PSF subtraction RDI. The top panels represent the subtraction result and error without the cross-dispersion mask while the bottom panels represent the subtraction result and error with the cross-dispersion mask. This mask is the same mask that was used for spectrum extraction.

FAC is 0.8, which is lower than 1. In consequence cRDI performs better than KLIP RDI in that case. Also, the FAC is much lower than 2 which is the value computed in the case of OS9 HLC data. This is because of the presence of the slit in spectroscopic mode that masks most of the field of view and the pixels or region used for post-processing are/is contaminated by the presence of the planet. It is thus not possible to select regions that are not impacted by the presence of the planet to optimize the post-processing and minimize over-subtraction such as in imaging mode. Here, the region used to optimize KLIP is overlapping with the planet, which leads to over subtraction.

Those results may be improved by only selecting less heavily contaminated regions of the spectrum that are just above or just below the central lobes of the planet line spread function or by taking into account the presence of the planet while performing the post-processing. However, taking into account the presence of the planet while performing the post-processing might be trickier than for imaging data because of the absence of a template of the dispersed planet line spread function.

Note that the estimated FAC value may slightly vary depending on the speckle realization and location of the planet (and thus of the slit) in the field of view.

4.2.3 Spectrum extraction

Spectrum extraction was done using a mask that limits the detector area to a small amount of rows (5 pixels wide), summing up the count levels for each column and binning those columns

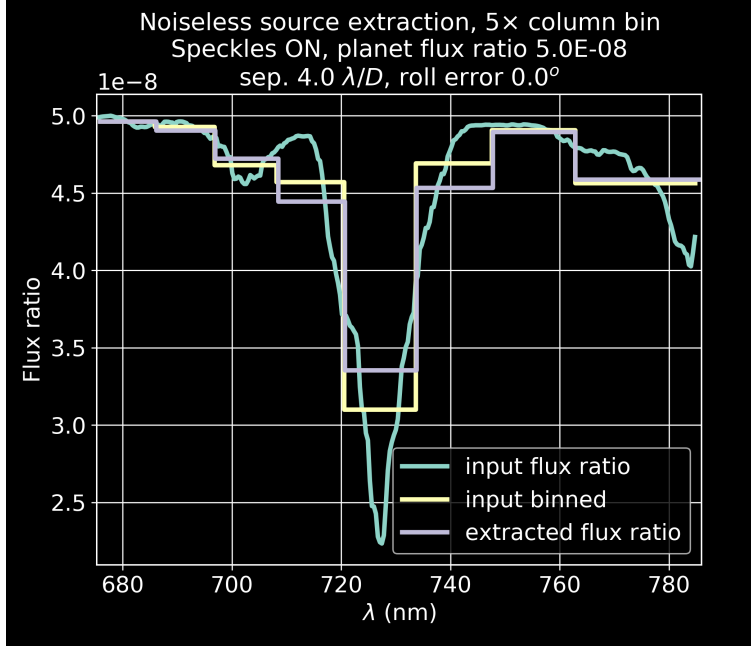


Figure 9: Extracted planet spectrum binned to the detector sampling in the noiseless MUFs ON case. The planet contrast is 5×10^{-8} and separation is $4\lambda/D$. The green curve represents the input planet spectrum in high resolution. The yellow curve represents the input planet spectrum binned to the CGI spectral resolution. The purple curve is the extracted planet spectrum binned to the CGI spectral resolution.

up to get the native spectral resolution. Figure 9 represents the noiseless source extraction binned to the spectral resolution of the Roman CGI for a planet flux ratio of 5×10^{-8} at a separation of $4\lambda/D$. The input flux ratio spectrum, which is the spectrum of the planet divided by the spectrum of the star, is represented in green. This input spectrum is binned to the spectral resolution of the Roman CGI and represented in yellow for comparison with the extracted flux ratio binned to the spectral resolution of the Roman CGI in purple. The difference observed between the input and extracted binned flux ratios seems proportional to the inverse of the flux ratio. Indeed, the signal of the planet decreases with the spectral feature and speckle noise largely dominates. Thus, it is no surprising that we observe a higher difference at the wavelength where the spectral feature is located.

Figure 10 represents the source extraction after post-processing in the noisy case. The extracted flux for each bin is represented by purple dots and the error bar correspond to the uncertainty due to the noise (including shot noise, read noise and QE). Those error bars do not include the uncertainty due to the presence of speckle, although it exists and it is not negligible (as discussed earlier, the impact of the speckles is all the more important as the speckles as bright and/or the signal of the planet is faint such as the spectral feature wavelength for example or at the location of the residual speckles).

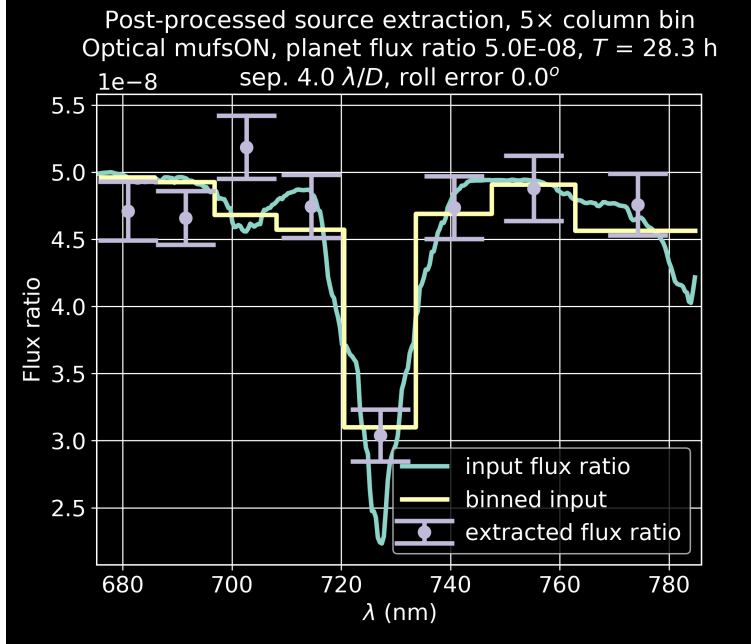


Figure 10: Post-processed source extraction in the case of noisy MUFs ON data. The planet contrast is 5×10^{-8} and separation is $4\lambda/D$. Green curve: input spectrum at high resolution. Yellow curve: input spectrum binned to CGI spectral resolution. Purple curve: extracted spectrum binned to CGI spectral resolution.

5 Conclusions and Next Steps

We simulated Roman Coronagraphic SPC Band 3 spectroscopic data and implemented the classical PSF subtraction cRDI and KLIP RDI post-processing techniques in order to estimate the Factor Above Classical for this mode. Our preliminary results and ideas for possible mitigation are:

- We found no improvement using KLIP RDI relative to classical RDI;
- It is more difficult to isolate the planet when applying KLIP to SPC data than it is in the case of HLC data. This is because the proportion of the planet signal relative to the speckle light is much higher in the slit aperture than for an image that contains the full field of view. As a result, the projection of the target data vector onto the reference basis is not well posed;
- Those results may be improved by only selecting regions of the spectrum that are just above or just below the central lobes of the planet LSF or by taking into account the presence of the planet while performing the post-processing;
- In addition, the relatively high stability of this PSF time series is an important factor when comparing the performance of classical PSF subtraction vs KLIP. If the correlations are close to unity then the FAC approaches 1.0;

- Thus, it is possible that the FAC slightly improve when processing less stable data sets;
- Taking into account the presence of the planet while performing the post-processing might be trickier than for imaging data because we will not have a template of the dispersed planet PSF;
- Might need to consider cross-correlations techniques with library of templates to optimally extract the planet spectrum.

In addition, we added noise to our simulated data, injected a planet in the noisy data and performed an extraction of the planet spectrum. Those tests show that even though the dispersion is supposed to mitigate the issue of speckles, the presence of a speckle at the location of the planet is a possibility and may strongly impact the spectrum extraction.

Next steps. As possible next steps, we may want to answer the following questions:

- Is the Factor Above Classical a pertinent metric to use for the spectroscopic mode of the Roman Coronagraph? Should we rather define a metric based on the spectrum extraction or reconstruction? How would cRDI and KLIP RDI perform in that case?
- Would implementing angular differential imaging (ADI) work for spectroscopy? If yes, what gains can we expect?
- What would the spectrum extraction would look like for a different planet and in particular for a self-luminous one? Would that be easier than for a reflected-light planet? What can we learn from this process? Which self-luminous target would be a good candidate for the Roman Coronagraph Tech Demo?
- What is the sensitivity of planetary atmosphere models to residual speckles at the location of the planet and how can we mitigate this issue to improve the spectrum extraction?
- What is the impact of removing the slit during the spectrum calibration?

6 Acknowledgment

The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).

7 References

1. Soummer et al., “Detection and Characterization of Exoplanets and Disks Using Projections on Karhunen-Loève Eigenimages”, *ApJ*, 2012.

2. Zimmerman et al., “WFIRST Coronagraph Instrument post-processing algorithms for advanced PSF subtraction”, 2018.
3. Ygouf et al., “Roman Coronagraph Instrument Post Processing Report - OS9 HLC Distribution”, 2021.
4. Ygouf et al., “Data processing and algorithm development for the WFIRST-AFTA coronagraph”, Proc. of SPIE, 2016.
5. 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”, Proc. of SPIE, 2015.
6. Ygouf et al., “Data Post Processing and Algorithm Development for the WFIRST-AFTA Coronagraph”, FY15 final report, Space Telescope Science Institute, 2016.
7. Ygouf et al., “AFTA Coronagraphic Technology – Data Post-Processing and Algorithm Development”, FY14 final report, Space Telescope Science Institute, 2014.
8. Ygouf et al., “Data Post-Processing and Algorithm Development for the WFIRST-AFTA Coronagraph”, First Progress Report, Space Telescope Science Institute, 2015.