WFIRST coronagraph flight performance modeling

John Krist, Robert Effinger, Brian Kern, Milan Mandic, Dwight Moody,
Patrick Morrissey, Ilya Poberezhskiy, A.J. Riggs, Navtej Saini, Erkin Sidick, John Trauger
Jet Propulsion Laboratory, California Institute of Technology,
4800 Oak Grove Drive, Pasadena, CA 91109

ABSTRACT

As it has for the past few years, numerical modeling is being used to predict the on-orbit, high-contrast imaging performance of the WFIRST coronagraph, which was recently defined to be solely a technology demonstrator. A consequence has been a realignment of modeling priorities and revised applications of modeling uncertainty factors and margins, which apply to multiple factors such as pointing and wavefront jitter, thermally-induced deformations, polarization, and aberration sensitivities. At the same time, the models have increased in fidelity as additional parameters have been added, such as time-dependent pupil shear and mid-spatial-frequency deformations of the primary and secondary mirrors, detector effects, and reaction-wheel-speed-dependent pointing and wavefront jitter.

Keywords: WFIRST, coronagraph, modeling

1. INTRODUCTION TO THE WFIRST CORONAGRAPH

The Wide-Field Infrared Survey Telescope (WFIRST)[1] is a space observatory planned for launch in the mid-2020s. Its primary instrument is a wide-field infrared imaging camera for dark energy and exoplanet microlensing studies. A secondary camera is a visible-light coronagraph instrument (CGI)[2] that is designed to be capable of high-contrast imaging of large exoplanets and circumstellar dust disks. Following a NASA-directed review in 2017 that evaluated cost and schedule the coronagraph was redesignated a technology demonstrator with no science requirements. Its task is to show the maturity of particular technologies such as deformable mirrors (DMs), picometers-level low-order and high-order wavefront sensing and control (LOWFS/C, HOWFS/C), extremely-high-contrast integral field spectroscopy, and photon counting detectors that may be used by future exoplanet science missions such as Habex or LUVOIR.

> *Note:* We will use here the term *contrast* in two different ways: 1) when referring to the brightness of a planet, contrast is the ratio of the planet's observed flux to that of its star; 2) when referring to instrumental performance, contrast will be the per-pixel instrumental background (speckle) brightness equated to the contrast of a planet whose peak pixel brightness is the same.

The WFIRST coronagraph instrument (CGI) will be the first space coronagraph with DMs and LOWFS/C. The coronagraphs on the Hubble Space Telescope and the upcoming James Webb Space Telescope lack either, limiting them to raw (unprocessed) contrasts of $\sim 10^{-5}$ and post-processed exoplanet detection limit contrasts of $\sim 10^{-7}$ near the star. With active wavefront control, the WFIRST CGI is designed to provide instrumental contrasts of 10^{-8} or better and post-processed exoplanet detection limits of $\sim 10^{-9}$. This can be achieved despite the WFIRST telescope providing a non-optimal obscuration pattern for coronagraphy, with a large central obscuration and six relatively thick secondary support struts (Figure 1). These are compensated within the coronagraph by pupil masks (Figure 1) or wavefront modification by the DMs.

Since it will be impossible to test the telescope and coronagraph together prior to launch to these contrast levels, end-to-end numerical modeling of the system is critical for design evaluation, tolerancing, and performance estimation. Prior reports[3,4] have detailed many aspects of the numerical modeling used to simulate WFIRST CGI data. Here we present an updated overview of recent results, notably the inclusion of dynamic errors such as pointing and wavefront jitter, pupil shear, and DM thermal drift. We also discuss the incorporation of model uncertainty factors (MUFs).

1.1 Coronagraph layout and components

The telescope and coronagraph layouts are detailed elsewhere[5,6], so provided here is only a brief description. The telescope has a 2.36 m diameter primary that directs light to a secondary. It then enters the Tertiary Collimator Assembly (TCA), a series of flat and off-axis parabolic (OAP) mirrors that collimate the beam into the CGI and correct for off-axis aberrations (the CGI field is 0.4° from center of the telescope's focal plane). The first pupil image is at the Fast Steering Mirror (FSM) that corrects for telescope pointing errors based on measurements from the Zernike phase-contrast low-order wavefront sensor (LOWFS)[7,8]. It is followed by a pistoning fold mirror in a converging beam to provide fine focus control (Focus Offset Mirror, FOM). Subsequent optics form another pupil image on the first of two DMs, each with an array of 48 x 48 actuators spaced by 1 mm. The second DM is located 1.0 meter downstream from the first, and both together provide phase and amplitude control of wavefront errors to create a dark hole of suppressed starlight inside which planets and disks can be observed. The DMs are also used as an integral part of the Hybrid Lyot Coronagraph (HLC)[9,10] to compensate for the telescope's obscurations. After the DMs another pupil image is formed on flat mirrors located in a selection mechanism, where reflective pupil masks (Figure 1) for the Shaped Pupil Coronagraph (SPC)[11] are located; a simple flat is used for the HLC. The beam is then focused onto one of a selection of focal plane masks (FPMs). After that another pupil image is formed and a Lyot stop mask is inserted, one for each coronagraph. Another intermediate focus is then produced for a field stop that blocks bright light outside the dark hole field that would saturate the detector. After this is a filter wheel. At a subsequent location a selectable fold mirror directs the beam into an integral field spectrograph (IFS)[12,13] or into the direct imaging camera. The IFS contains an array of lenslets that partition the field and disperse each element through a prism and onto a detector, forming a series of spectra.

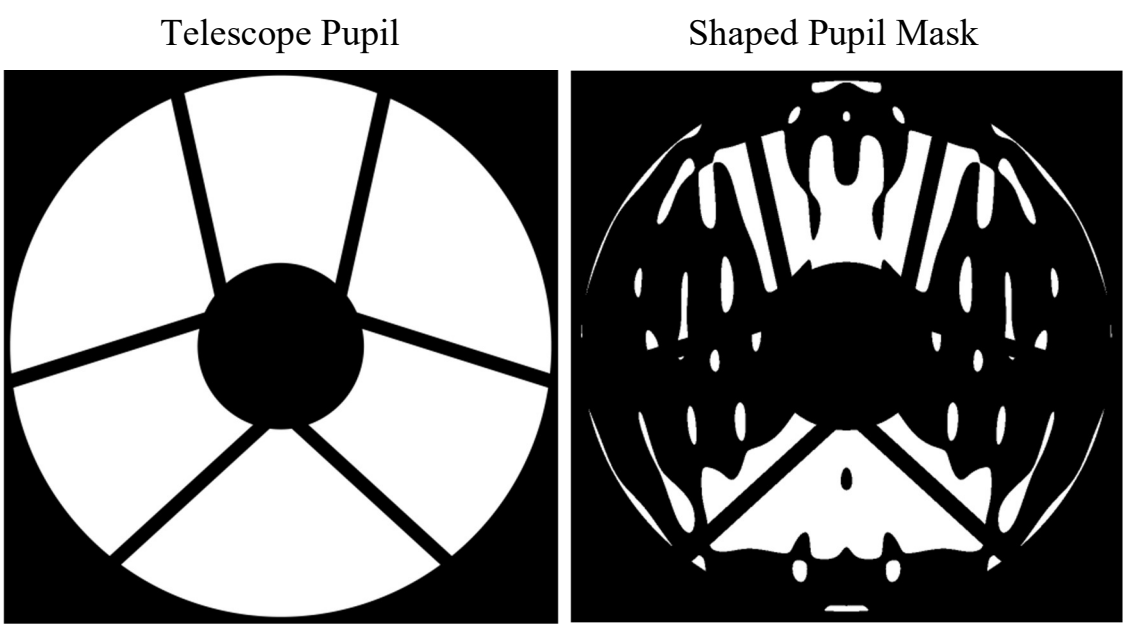

Figure 1. (Left) WFIRST telescope obscuration pattern. (Right) Shaped pupil coronagraph pupil mask.

The HLC will be used for direct imaging in a 360° annulus from $r = 3 – 9\ \lambda/D$ in a 10% width bandpass centered at 575 nm. It relies on the DMs to compensate for the telescope obscurations using relatively large strokes (a couple hundred of nanometers). In the HLC design used here (HLC20150818), the FPM is a partially-transmissive nickel spot of $r \approx 2.6\ \lambda/D$ overlaid with a patterned, transmissive dielectric coating, and together these provide both phase and amplitude modulation of the beam. A Lyot stop in the final pupil plane masks the spiders, central obscuration, and outer aperture edge.

The SPC uses a binary pupil mask to redefine the obscuration pattern in a manner that diffracts light outside of the dark hole region. The SPC IFS mode design (SPC20170714) operates over an 18% bandpass centered at 660 nm or 760 nm. It has a binary "bow tie" FPM with a field of view spanning $r = 3 – 9\ \lambda/D$ over two 65° azimuths (separate FPMs are provided

for each bandpass). The limited field assumes that a target planet was previously detected in HLC direct imaging mode and thus can be positioned to a known location in the IFS field. An annular Lyot stop blocks light along one axis.

A separate SPC wide field design provides imaging in a 10% bandpass centered at 825 nm of extended objects (e.g., circumstellar disks). It has a binary FPM with a 360° annulus from $r = 6.5 – 20\ \lambda/D$. Its Lyot stop is also an annulus. This design will not be discussed here.

1.2 Wavefront control

Critical to the performance of the CGI is wavefront sensing and control in low-order (low-numbered Zernikes) and high-order (zonal errors) regimes. The LOWFS/C maintains the stability of the system in the presence of pointing errors and thermally-induced long-term drifts of low-order aberrations. The HOWFS/C[14] is used to measure and control static higher spatial-frequency errors that would introduce light in the dark hole region. It is also used to derive the DM patterns necessary to compensate for the telescope's obscurations in the HLC mode.

The LOWFS/C uses the image of the stellar core reflected by the front surface of the FPM. A phase-modulating layer of patterned dielectric within a central spot on the FPM introduces interference with the exterior unmodulated beam as it propagates to the LOWFS detector, turning wavefront phase errors into intensity variations as seen at the reimaged pupil. Differences in the intensity pattern over time are decomposed into Zernike aberration changes by matching precomputed mode images. Commands are then sent to the FSM (for tip/tilt), the focus corrector (FOM), and the DMs (for other Zernikes) to compensate.

The HOWFS/C uses estimates of the wavefront as derived from speckle field images using the direct imaging or IFS detectors. The field is intentionally disturbed using specific patterns applied to the DMs (*DM probing*), and the measured changes in intensity are then compared to those predicted by the models for those patterns to estimate the complex electric field (E-field) at the image plane. An iterative wavefront control algorithm uses these E-fields in conjunction with a wavefront control matrix to determine the DM settings necessary to reduce the light within the dark hole. The matrix is generated using the model-predicted changes in E-field due to each DM actuator piston.

1.3 The CGI image

A telescope+CGI image consists of astronomical sources (planets, dust disks, background stars and galaxies) embedded in a field of speckles of residual light from the primary star. The diffraction pattern of the star is suppressed by the coronagraph. The speckles are the result of the imperfections in the optics and coronagraphic masks and are substantially reduced by wavefront control (WFC) using the DMs. The post-WFC speckles are of similar brightness to the astronomical sources of interest. The speckle level defines the background shot noise.

The WFIRST CGI will utilize electron multiplying charge-coupled device (EMCCD)[15,16] detectors for both the direct imaging, IFS, and LOWFS channels. The EMCCD provides practically read-noise-free photon counting, which is especially critical for the IFS due to the low signal rates of potential exoplanets. The detector does, however, suffer from other noise sources such as dark current, clock-induced-charge, charge transfer efficiency, and cosmic rays. These combine with the speckle shot noise to set the required integration time to reach a desired signal-to-noise ratio.

To distinguish the astronomical objects from the speckles requires some sort of post-processing of the images[17-19]. The most common method, reference differential imaging (RDI), involves observing a bright reference star and subtracting in the computer its speckles from those of the target star. An alternative method, angular differential imaging (ADI), observes the target star at two or more roll orientations of the telescope; since the speckles are generated in the optics, they will appear fixed on the detector while the actual astronomical sources will appear to rotate about the star. Various algorithms can be used to distinguish what is rotating in the image and what is fixed. Both methods have been used on the ground and with the Hubble Space Telescope for high contrast imaging. Both require that the speckle patterns in the images of both stars or at both orientations of the target star be largely similar. Any differences will result in residual noise against which the sources must be distinguished. The speckle pattern is sensitive to changes in the wavefront that may be caused by thermally-induced deformations, vibrations from the reaction wheels used to orient the spacecraft, and instabilities in the DMs.

1.4 Observing scenario

A realistic WFIRST CGI observing scenario (OS) is defined by timing and orbital parameters for a specific set of stars that then determine the observatory's orientation relative to the Sun. The most likely WFIRST CGI OS includes visits to two stars, the reference and target. The target star (generally V > 5) is likely to be a host to a radial-velocity-detected planet

or previously imaged circumstellar dust disk. It will be observed at two roll orientations of the telescope to allow ADI post-processing; the WFIRST roll range is limited to ±15° from the nominal orientation (the maximum solar illumination on the solar panels) and in practice is limited to ±13° in current plans. For CGI observations only two rolls are used to ensure that the incident total solar illumination is the same at each to minimize thermal disturbances. The reference star is brighter (V < 3) and is used to sense and optimize the wavefront using the DMs as well as to provide reference images for RDI post-processing.

There are a variety of ways to order the reference and target star observations. In the OS used in earlier modeling runs all of the images of the reference star were taken first, followed by the target at each of the two rolls. The problem with this is that it maximizes the effects of any long-term (relative to the OS duration) drifts in the system, leading to mismatches between the reference and target stars' speckles. These effects can be reduced by interleaving observations of the two stars, chopping back-and-forth between them multiple times. This comes at the expense of increased total slew time.

The OS used for the models presented here, OS6, incorporates chopping. The reference star is η UMa (V = 1.9; B3V) and the target is 47 UMa (V=5.0; G1V), a known multiple-planet host. Our assumption is that a week or two prior to the observations the wavefront control patterns previously derived for the given coronagraph and bandpass will be put on the DMs. This allows the DMs to settle and significantly reduces their surface drift when observations commence (the ceramic used in the DMs currently planned for the CGI relaxes over time after application of large actuator offsets). The patterns are then updated for the current wavefront state by spending 8 hours on η UMa while running the HOWFS/C algorithm. Starting with the old patterns rather than from scratch will likely save a significant amount of time to reacquire a good dark hole. After that the chopping sequence begins with 8 hours on 47 UMa observed back-and-forth at two orientations every two hours, followed by two hours on η UMa. This 47 UMa/ η UMa pattern then repeats for 13 cycles. Including the initial 8 hours on η UMa, the sequence lasts 138 hours (which also includes 10 minutes for each slew or roll of the telescope and to reacquire the target in the CGI).

2. ERROR SOURCES

A multitude of defects, both constant and time-varying, perturb the coronagraph from its ideal performance. Wavefront control can only reduce but not eliminated them. Residual static errors set the minimum possible speckle brightness while dynamic ones define the speckle stability and thus the post-processed exoplanet detection limits.

2.1 Static errors

Static errors are fabrication and alignment defects that are invariant over time. They define the minimum possible dark hole speckle brightness. Even though they are constant, they can amplify the effect of time-dependent errors through the heterodyned combination of the static and dynamic electric fields. Prior to wavefront control, the expected static wavefront errors will create a dark hole region background of ~10^{-4} contrast speckles in the WFIRST CGI.

Static wavefront errors are sensed using a combination of phase retrieval and DM probing. Phase retrieval, which in the CGI is enabled using a set of selectable defocusing lenses, provides the best estimate of large low-order errors (the main effects of which occur inside the region blocked by the FPM). Early in the on-orbit commissioning of the CGI it will be used without the coronagraphic masks in place to derive the DM pattern necessary to "flatten" the bulk of the static phase error; when the masks are reinserted the dark hole contrast should be ~10^{-6}. DM probing will then be used in conjunction with higher-order wavefront control to complete the dark hole digging process by correcting the remaining phase and amplitude errors, bring the mean speckle brightness down another 2-3 orders of magnitude.

Surface errors

For an extreme high-contrast coronagraph like WFIRST's, it is critical that the optics be ultrasmooth[20], with root-mean-square (RMS) surface deviations measuring in single digit nanometers. Though the coronagraph has DMs to control the resulting wavefront errors, there is a limit to how well they can be corrected. Surface errors on optics at non-pupil locations can generate chromatic phase and amplitude errors that cannot be fully negated by the DMs over a broad bandpass. Therefore, to produce accurate results the numerical models must include realistic aberrations on each optic. The optical fabrication errors of concern to our models can be divided into two regimes: (1) low-order aberrations such as focus, coma, astigmatism, trefoil, and spherical from imperfect figuring and (2) higher-spatial-frequency ones such as zonal aberrations and roughness from imperfect polishing.

The largest surface errors are on the primary and secondary mirrors. Both of these optics exists and have measured surface error maps that are incorporated into the models presented here. Though both will be refigured prior to launch to correct some lower-spatial-frequency errors and zonal defects, these maps are sufficiently representative of the on-orbit surfaces, especially in terms of mid-and-higher spatial frequency errors. The primary map, for example, shows the backing rib structure print-through and rolled-over edge common to large telescope primaries.

All other optics in the system must be represented by synthetic error maps, with each optic having defects appropriate for its type (flat or powered) and size. The CGI optics error budget defines the maximum aberration levels allowed on each surface, and the synthetic maps are generated to meet those. To establish a realistic surface error spatial frequency distribution, the measured error maps of multiple Gemini Planet Imager flats and off-axis-parabolic (OAP) mirrors were analyzed (maps provided by Bruce Macintosh, Stanford University). These were used to determine a likely spread in the allocation of low-order errors (defined by Zernike polynomials) and mid-spatial ones (defined by power spectral density, PSD, curves). These GPI optics have properties similar to or even better than the CGI requirements. Realizations matching these statistics were then generated for each CGI optic (Figure 2). The TCA optics have RMS surface errors of 14 nm (OAPs) and 7.5 nm (folds). The internal CGI optics have RMS surface errors of 1.9 nm (OAPs), 0.8 nm (folds), and 2.5 nm (DMs).

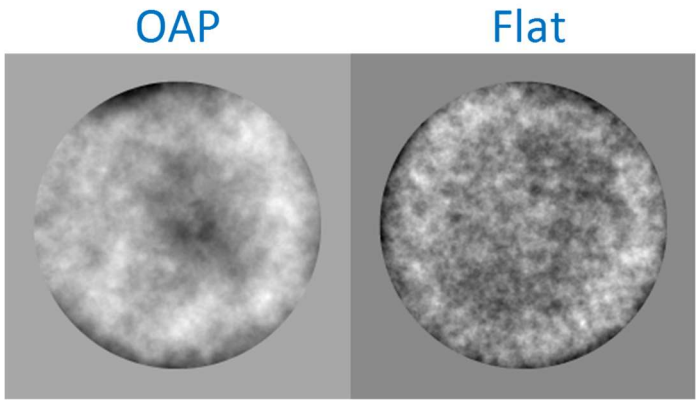

Figure 2. Synthetic surface error maps of a selected (left) off-axis parabola and (right) a flat mirror as used in the CGI models. The OAP map has an RMS wavefront error of 3.5 nm and the flat is 1.0 nm RMS.

Polarization

The fast (f/1.2) WFIRST primary mirror introduces significant polarization-dependent aberrations that vary strongly with wavelength, almost entirely astigmatism and tilt (because the coronagraph is located off-axis in the telescope focal plane). They can be described in terms of four wavefront components corresponding to orthogonal polarization input/output states: X_{in}/X_{out}, Y_{in}/Y_{out}, X_{in}/Y_{out}, and Y_{in}/X_{out}. The first two "direct" components contain the large majority of the light while the latter two "cross terms" are orders of magnitude smaller in intensity. The four fields are incoherent with each other. A simulation including these polarization errors must propagate each wavefront component through the system separately, and the final four image-plane fields are added in intensity. Were a polarizer to be used at the end of the CGI, the direct and cross term wavefronts for a given output polarization would be individually propagated and the resulting two fields added in intensity.

Figure 3 plots the difference in wavefront phase error between the X_{in}/X_{out} and Y_{in}/Y_{out} polarizations versus wavelength as predicted at the coronagraph entrance. These were computed with coatings matching those to be used in the telescope and CGI (the particular wavelength dependence behavior is a function of the coatings). The difference is minimal at 600 nm (near the center of the HLC bandpass) and increases rapidly towards shorter wavelengths and less so towards longer ones. There is also polarization-dependent amplitude error, but it is less important given the coronagraph's sensitivities (but it is still included in the models). The DMs can correct up to about half of the difference in X and Y polarization wavefront error at a single wavelength (excluding the cross terms). However, due to the wavelength dependence, the ability to correct these is limited by the bandpass width and central wavelength.

In practice we have chosen in our modeling to use wavefronts computed for $+45°_{in}/X_{out}$, $-45°_{in}/X_{out}$, $+45°_{in}/Y_{out}$, and $-45°_{in}/Y_{out}$. These can be readily converted to the four X-and-Y-only terms described above via subtraction or addition of pairs of wavefronts. The advantage of this method is that the wavefronts are low-order and continuous, and thus are

well described by Zernikes which allow accurate scaling of the errors to any grid sampling. In comparison, the X_{in}/Y_{out}, and Y_{in}/X_{out} cross terms may have sharp patterns or even singularities. Like before, each term is propagated separately through the system and the resulting fields are added in intensity.

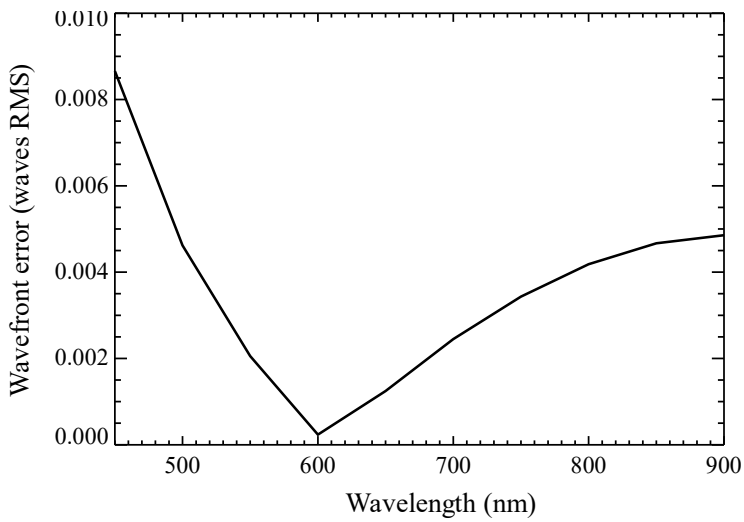

Figure 3. Predicted difference in phase error between the X and Y polarization-induced aberrations versus wavelength.

Misalignments

To represent the aberrations caused by static misalignments of the optics, low-order Zernike errors (focus, coma, and astigmatism) are added in amounts consistent with the values in the optical error budget (22 nm RMS at the TCA).

Mask defects and contamination

Errors in mask fabrication and any contamination (e.g., dust) are not currently included in the observing scenario models. Mask defects (e.g., sizing, registration, rotation) are included in the system error budget, which uses sensitivities to errors determined through modeling (note that a pupil magnification error due to inaccurate optic power or separation can be interpreted as a mask sizing error).

2.2 Dynamic errors

Dynamic wavefront errors, as discussed here, include all those that vary with time. They can be induced via thermal changes or vibrations. Some, but not all, of these can be sensed by the LOWFS and compensated with the FSM, focus corrector, and DMs.

Thermal

The main driver of thermal variations in the WFIRST telescope and instrument carrier is the orientation of the observatory relative to the Sun; the impact of heating from the Earth (the primary cause of Hubble Space Telescope thermal variations) is avoided by WFIRST's L2 orbit. The solar panel acts as a sunshield for most of the observatory, and since it is not actuated its normal must stay to within ~15° of the Sun to maintain sufficient power. The telescope (primary & secondary mirrors and their support structures), TCA, and the CGI are temperature controlled with multiple sensors and heaters, but the instrument carrier (IC) that is the structural interface between the telescope and CGI is not.

The changing radiative and conductive thermal conditions induce structural deformations that alter the alignments and surface figures of the optics and shift the beam (e.g., pupil shear) with respect to the optical surface errors. This is especially significant in regards to where the beam lands on the DMs. The DM patterns are established for a specific registration of to the optical wavefront errors, and if those shift relative to the pattern then the dark hole contrast will degrade. Unexpectedly large misalignments could also shift the telescope obscuration pattern outside of the SPC pupil mask area (which is built to tolerate 0.2% pupil shear), which would also quickly degrade performance.

Internal to the coronagraph deformations may occur due to external (solar) and internal (electronics) thermal inputs. The CGI optical bench and the DMs are temperature controlled, reducing these effects. The ceramic blocks in the DMs may also drift due to stresses introduced by putting large strokes on the actuators. The drift decays as the ceramic relaxes over time, so it can largely be avoided by putting any large-stroke patterns on the DMs a week or two prior to observations.

Vibrations

Six reaction wheels on WFIRST will be used to slew the telescope and maintain pointing to compensate for solar pressure torque, and they induce vibrations in the structure that cause rapid pointing and optical alignment errors (pointing and wavefront jitter). These are too fast for the LOWFS to fully sense and correct, and so they create an additional incoherent source of speckles. The telescope pointing jitter is expected to be no more than 12 mas RMS per axis, and this will be reduced by the CGI FSM down to 0.4 – 1.6 mas RMS.

Jitter varies over time as the wheel speeds change during the observations. To minimize vibrations during CGI observations, the wheels are limited in speed to approximately 5 revolutions per second (rps) in either spin direction. In OS6 it is possible to achieve this over a 20 hour timespan. The wheels start at a maximum rotational velocity of ~5 revolutions per second (rps), cross 0 rps after about 10 hours (depending on the wheel) and end up spinning in the opposite direction at 5 rps at 20 hours (this spans two iterations of the reference/target star sequence). After that the wheels' momentum is dumped using the spacecraft thrusters. This pattern is repeated every 20 hours during OS6.

3. NUMERICAL MODELING

The primary goal of the WFIRST CGI numerical modeling activity is to develop accurate simulations of the in-space coronagraph for a realistic observing scenario. Both the observatory and the CGI must be modeled together with wavefront control incorporated. The combined interaction of the static and dynamic errors may lead to unexpected degradations in performance, so it is best to find out well before the hardware designs are finalized and fabricated. There are four primary components to the observation scenario numerical modeling: combined structural/thermal/optical (STOP) modeling, vibration modeling, diffraction propagation (including the IFS model), and detector modeling.

3.1 STOP modeling

The STOP model, as it has been exercised in the analyses presented here, calculates the thermal variations in the telescope+CGI system due to changes in the solar incidence as the telescope is slewed to/from the stars and rolled about the target star (in upcoming analyses the impact of the heat from CGI electronics will be included). The structural elements, including the optical surfaces, deform in response. The resulting displacements and deformations are used to modify the nominal optical prescription, which is then ray traced to produce a map of the wavefront aberrations at a given plane for each timestep in the observing sequence.

The finite-element models used in STOP modeling have a multitude of nodes meshed to represent the various components, including the relevant properties of each material such as the coefficient of thermal expansion. Thermal sensing and heater control loops used to stabilize the system are implemented. The full STOP model is provided by the WFIRST Project Integrated Modeling Group at the NASA Goddard Space Flight Center (GSFC), who integrates separate models of the telescope (from Harris Corp.) and the CGI (from JPL) with theirs for the rest of the observatory. The version used for the analyses presented here is Phase A[21] Revision 1. It includes detailed representations of the telescope and instrument carrier, up through the latches mounting the CGI to the IC. The CGI in this version is rather simply represented, without detailed meshes, thermal control, or the thermal effects of the electronics. A much more complete model will be provided in Phase B (Summer 2018 – late 2019).

The STOP modeling sequence (Figure 4) begins with the solar illumination angles as defined by the OS time and orbital parameters. Thermal Desktop is used to determine the temperatures of each thermal node; this stage takes the large majority of the STOP modeling time, especially when short timesteps (e.g., 10 seconds) are used for the thermal solver. The results are fed to NASTRAN, which computes the corresponding structural deformations. Sigfit then converts these to perturbations of the optical prescription that is ray traced with Code V. The outputs are a series of wavefronts (Zernikes) at the entrance pupil to the CGI (located at the FSM) and critical offsets versus time (e.g., shear of the telescope pupil at the FSM). GSFC uses the same chain of programs, though they utilize pre-computed linear optical models for some stages that are explicitly computed for each timestep by JPL.

Both GSFC and JPL run the same OS through their STOP procedures to verify their results against each other, which may differ slightly due to the random number generator used to seed the thermal modeling. After confirmation of agreement, JPL can run the OS with different parameters (e.g., finer timesteps) or entirely different scenarios. To facilitate this, JPL has automated the Thermal Desktop-through-CodeV stages using the Luigi pipeline module for Python, creating the Integrated Modeling Pipeline[22]. One simply needs to define the OS and then these steps are executed without further intervention. This is particularly useful when fine timesteps are used, which can extend run times to nearly a week.

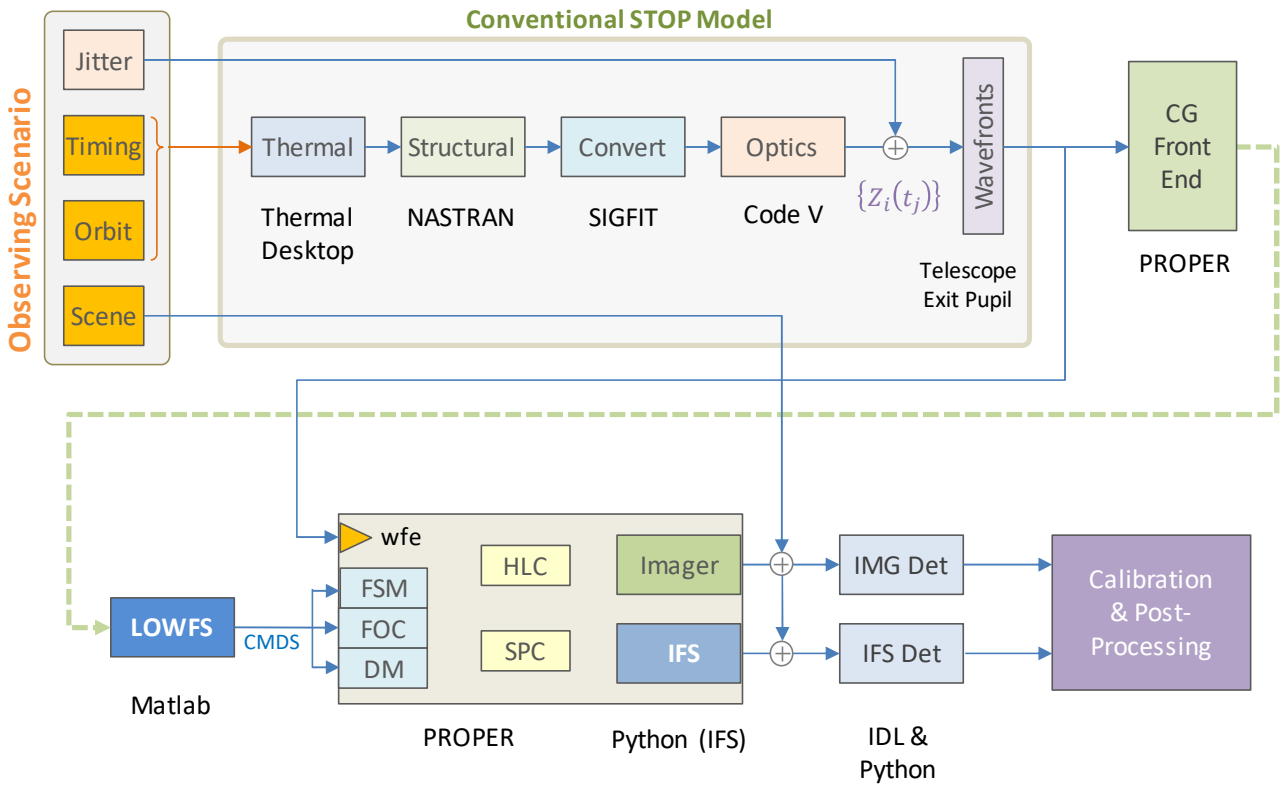

Figure 4. Schematic diagram of the flow of the combined STOP + diffraction propagation modeling process for on-orbit simulations. The associated software packages are identified. Modified from an original figure by Bijan Nemati.

3.2 Vibration (reaction wheel) modeling

Using the timing and orbital parameters provided by JPL for OS6, the Goddard modeling team determined the speed of each reaction wheel versus time. As noted before, the wheels vary in speed over a 20 hour span with a limit of ±5 rps, then the wheel momentum is dumped, and the cycle repeats. The vibrations induced in the telescope structure were computed and the corresponding induced aberrations (from misalignments) derived as power spectra for each aberration (amount of aberration per temporal harmonic). The pointing errors were filtered in temporal frequency to match the expect LOWFS/C pointing control.

At JPL the power spectra were turned into realizations of rapid pointing and aberration errors over time at fine resolution (1 kHz) over the 20 hour span. The post-FSM pointing offsets were integrated over one second intervals and converted to elliptical Gaussian distributions with major and minor axes RMS widths and the major axis orientation angles. The wavefront error jitters were converted per aberration to sinusoidal patterns with the computed RMS variation over one second for use by the optical models; the optical modeling routines require a symmetric oscillatory function for computationally-efficient wavefront error jitter representation.

3.3 Optical modeling

The optical model propagates the wavefront through the entire optical system, accounting for diffraction effects from each optical surface and coronagraph mask. It utilizes a representation of the static telescope+CGI system, including surface errors, polarization aberrations, coronagraphic masks, deformable mirrors, and (if modeling the IFS) lenslets. With the same wavefront sensing and control algorithms that are used on the testbeds and that will be used on-orbit, DM settings are determined to create a dark hole of suppressed scattered light around the star. The dynamic errors are then introduced as additional wavefront aberrations or component offsets.

PROPER

The wavefront must be propagated, including diffraction effects, from the telescope entrance to each optic, through the coronagraph masks, and to the final image plane. The commercial optical ray tracing software that offer such physical optics propagation, like Code V or Zemax, are rather inefficient for this (as well as too expensive and complicated for many modelers to use). Thus, more specialized code is needed. The results presented here used the PROPER[23] optical propagation software library, a freely available set of routines useful for modeling coronagraphs. PROPER is available for IDL (Interactive Data Language), Python (v2.7 & v3.x) and Matlab and can be downloaded from **`proper-library.sourceforge.net`**. It can do angular spectrum and Fresnel propagation with automated control, modeling of deformable mirrors using a measured actuator influence function, draw complex apertures, and create aberrations specified by Zernike coefficients, power spectral density (PSD) curves, or error maps. PROPER uses an unfolded layout (all optics positioned along a straight line) that can be derived from the system prescription. Powered optics are represented by pure parabolic phase functions. The static surface errors and computed polarization errors are included. PROPER has been verified for coronagraph modeling in a NASA Technology Demonstration for Exoplanet Missions (TDEM) study.

Broadband images are created by adding monochromatic images together in intensity. Nine wavelengths are used spanning the HLC 10% bandpass, 11 for the SPC 18% one. Each wavelength is propagated at each of the four polarizations, so a single broadband image requires 36 or 44 separate runs through the system.

The coronagraph masks are provided by their designers (Dwight Moody for the HLC and A.J. Riggs for the SPC, both at JPL). The HLC phase-and-amplitude modulating FPM is provided as a set of complex-valued masks at 9 wavelengths spanning the 10% bandpass with 0.159 λ/D spatial sampling. These include the thin-film effects of the mask materials. The Lyot stop is provided as a separate image file. The corresponding DM patterns used to compensate for the obscurations are provided as maps of the wavefront change produced by each DM.

The SPC pupil mask is a 1000-pixel-wide binary image, while the FPM and Lyot stop are specified by annular radii and opening angles. For best accuracy, it is necessary to highly sample the hard-edged SPC FPM. To implement this directly with PROPER would require gigantic arrays with sufficient padding to produce a wavefront with sufficient sampling. To avoid this, the wavefront is propagated to the FPM plane and then, outside the context of PROPER, Fourier transformed to a virtual pupil image. Then a Matrix Fourier Transform (MFT) is used to propagate the field back to the FPM at high sampling (0.025 λ/D) only over the clear region of the FPM, which is then applied as a mask. This field is then MFTed back to the virtual pupil, which is then Fourier transformed back to the FPM at the sampling used by PROPER, after which the wavefront is propagated as usual through the remainder of the optics.

Digging the dark hole

Prior to any wavefront control, the PROPER models produced broadband fields of $\sim 10^{-4}$ mean contrast with the coronagraphic masks inserted, limited by the uncorrected surface errors. The first stage to digging the dark hole is flattening the wavefront. While in reality this is accomplished by using phase retrieval without the masks to measure the phase error, in the models the computed phase error at pupil at the Lyot stop can be used (though the facility to duplicate phase retrieval is provided in the models with the inclusion of the defocusing lenses). DM1, located at a pupil conjugate to the primary, is fit to the opposite of this pattern. With the coronagraph masks reinserted the phase-flattened field contrast improved to $\sim 10^{-6}$, providing the starting point for the iterative high-order wavefront control (HOWFC).

Before one can control the high-order errors, they must be sensed. In the real world this is done by DM probing: putting multiple specific positive-and-negative pairs of pattern on DM1 (typically modified sinc patterns) and measuring the image plane intensity changes. Together with model-predicted intensity changes for those same patterns, one can solve for the complex-valued electric field (E-field). For a good broadband solution, the E-field must be derived for multiple sub-bands within the primary bandpass, which is done using multiple narrow filters or the IFS (for a 10% bandpass, at least 3 filters

are needed, each about 3.3% in width). When running models, one can instead use the actual computed fields at multiple wavelengths rather than derive them through probing, which saves considerable time. However, when incoherent effects that are dependent on the wavefront control are important, such as with polarization aberrations, probing is needed. The sub-band intensity images can be represented by the summation of the intensities at multiple wavelengths (3 for a 3.3% filter).

Electric Field Conjugation[14] (EFC) is the HOWFC algorithm used for the modeling shown here as well as on the JPL coronagraphic testbeds. It is based on a DM response matrix (the Jacobian) that relates the sensed image plane E-field to DM actuator settings that zero-out the energy inside the specified dark hole field. The Jacobian is built using the numerical model (whether for simulations or in the real world): each actuator on each DM is pistoned by some amount and the resulting change in E-field is recorded at each sensing wavelength. This large matrix is sized [*number of dark hole pixels* × *number of sensing wavelengths*] by [*number of useful DM actuators* × *number of DMs*]. Such complex-valued matrices can reach hundreds of megabytes in size for large dark holes with many pixels. A linear solver can be used to derive the zero-dark-hole-intensity DM settings. Due to nonlinear behavior of the coronagraph or errors in the model used to generate the Jacobian, it is necessary to apply regularization to damp the solution. Tikhonov regularization is used here, with the damping factors applied as diagonal matrices appended to the Jacobian, which is then pseudo-inverted using singular value decomposition prior to use by the solver. The best regularization parameters used here were derived by trial-and-error. Due to the linear approximation to a nonlinear system, the sense-then-correct procedure will not provide a good solution in a single step and must be iterated until convergence.

The implementation of the DM solution for the HLC as used here warrants some additional detail. The DMs are integral to the compensation of the telescope obscurations, primarily by creating wavefront modulations that, in conjunction with the FPM and Lyot stop, result in destructive interference inside the dark hole (without them and only using the masks, the field contrast would be ~10^{-5}). They are also critical to the reduction in the sensitivity of the coronagraph to pointing (tip/tilt) errors. The DM patterns and masks are jointly optimized for an unaberrated, obscured pupil, using a procedure similar to EFC. The tip/tilt sensitivity is included in the optimization by intentionally offsetting the star from the center of the FPM by multiple amounts. The mask design parameters along with the DM patterns are provided to the modelers, so rederiving the patterns from scratch are not necessary. In the real world, the situation is not so simple, as testbed experience has shown. The ideal DM patterns cannot be sufficiently matched by putting them on the real DMs due to differences in real and model DM parameters (actuator behavior, pupil distortions, etc.), and so the DM solution must be obtained from scratch by duplicating the optimization process. In the case of the models shown here, however, the ideal DM patterns were put on prior to running EFC.

Thermal drift & LOWFS/C

The Integrated Modeling Pipeline produces a table of Zernike aberrations and pupil shear at the FSM versus time. The aberrations are added in the PROPER model at the primary mirror. The pupil shear is assumed to be a shift of the entire coronagraph at the FSM perpendicular to the optical axis. This is done by propagating the wavefront to the FSM, taking the Fourier transform, adding an appropriate tilt phase term corresponding the to pupil offset, and Fourier transforming back.

The wavefront is propagated to the FPM and its on-reflection phase-modulating pattern. The appropriate phase pattern is added to the wavefront and the reflected core region is then propagated to the LOWFS detector (Figure *5*). Stellar brightness and realistic detector noise and effects, including the high-speed readout smearing seen in the LOWFS simulations, are included. The measured focus change is converted to a corrective FOM piston and the other low-order aberrations (astigmatism, coma, trefoil, spherical) are corrected by patterns sent to DM1, including actuator gain errors. The original wavefront is then run again through PROPER, this time all the way to the final image plane with the LOWFS/C commands included.

The response of the DMs to thermal drifts has not been well characterized, but it will be in Phase B. For now, a very notional model of its response, as a function of temperature change and actuator stroke, is included. The temperature of the DM over the OS is obtained from the Thermal Desktop output, based on the current crude representation of the internal CGI layout. The temperature variations are ~0.4 mK, but due to the high uncertainty in both these values and the DM response to them, we scale the temperature variation to 3 or 10 mK, depending on whether an optimistic or conservative case is being evaluated.

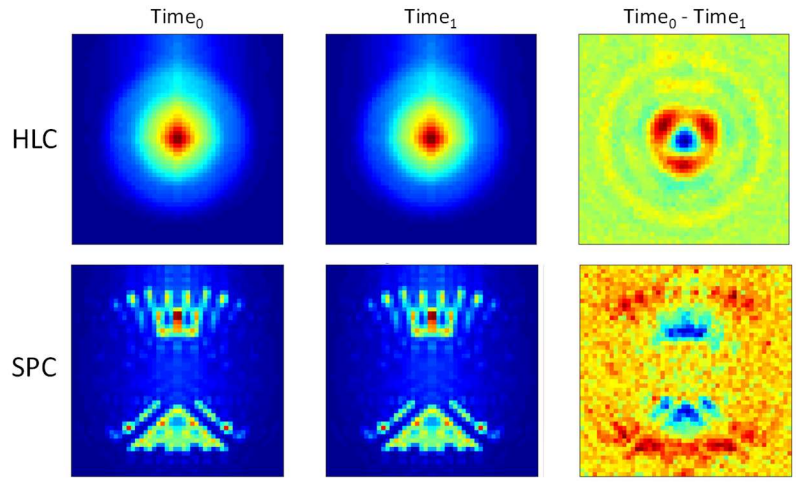

Figure 5. Simulated LOWFS images for the HLC and SPC shown at two different times. The right column shows the differences, which are fit to measure the Zernike aberrations. The corresponding correction commands are then sent to the FOM and DM. The LOWFS also measures and corrects pointing errors, but those are not modeled here.

At this stage, the wavefront for each STOP model timestep (10 minutes here) has been propagated through the system with PROPER at each wavelength (9 for HLC, 11 for SPC) and each of the four polarizations, and the complex-valued E-fields are stored at 0.1 λ/D spatial sampling. For the remaining stages, these are interpolated in time to finer steps (2 minutes).

Pointing jitter & stellar disk

The pointing errors are simulated by introducing tip and tilt at the primary mirror in the model. In order to represent elliptical Gaussian jitter and the circular finite disk of the star (assumed uniformly illuminated here), the fields for a large number (197) of irregularly-space X and Y source offsets are computed. The offsets range in spacing from very fine near the center (0.15 mas) to sample stellar disks of 1 mas diameter or larger, to 12.8 mas at the largest offset (45 mas) to sample the wings of the Gaussian distribution for very large jitters. These are propagated through the static system for each polarization component, and the E-field difference from the zero-offset field is stored in the "jitter table".

The elliptical Gaussian jitter distribution is computed for each timestep and convolved with a uniform circular distribution to account for the stellar disk, forming a weighting function (only weights ≥2% of the max weight are used). The zero-offset E-field for each timestep, generated as described in the previous section, is then added to each E-field in the jitter table, converted to intensity, and then multiplied by the corresponding weight. These images are then averaged together to represent the pointing-jittered image at that timestep. This method is much more efficient than generating fields for offsets at each timestep and has been experimentally shown to be accurate to 10^{-10} contrast.

Wavefront jitter

Wavefront jitter can be efficiently represented in the models if we assume that each component aberration (coma, astigmatism, focus, etc.) oscillates about a mean in equal positive and negative amounts over a short time span. It can be shown that for a sufficiently small amount of an aberration (e.g. picometers of coma) the common term, the unperturbed E-field, drops out when the two opposite aberration components are added in intensity, as they are to represent time-averaged wavefront jitter.

Let $E_0(t)$ represent the E-field at time t without any wavefront jitter aberrations included (t is some instant within the timespan $\bar{t}$ that covers one cycle of oscillatory wavefront jitter variation during which time no other changes in the system occur). $E_{+ab}(t)$ and $E_{-ab}(t)$ are the E-fields at time t with positive and negative wavefront jitter aberrations included. In this simple case where wavefront jitter is approximated by two timesteps:

$$I_{jit}(\bar{t}) = |E_{+a}\ (t)|^2 + |E_{-ab}(t)|^2 = [\ |E_{+ab}(0) - E_0(0)|^2 + |E_{-ab}(0) - E_0(0)|^2\ ]/2 + |E_0(t)|^2$$

where I_{jit} is the intensity image with wavefront jitter. This shows that the jitter aberrations do not have to be propagated through the system for every timestep. Rather, the differences between the initial perturbed and unperturbed E-fields can be stored and later added in intensity to the unperturbed image at any specified time. Of course, one must include enough steps within timespan $\bar{t}$ to adequately sample the jitter; we use 10 over one cycle of a sinusoidal oscillation. The amount of aberration will vary, so $E_{ab} - E_0$ is computed for multiple values ranging from -500 pm to +500 pm RMS for each aberration (focus, astigmatism, coma, trefoil, spherical). A 3rd order polynomial is then fit separately to the real and complex values and the coefficients stored and used to compute $E_{ab} - E_0$ for any value.

At each 2 minute timestep, the RMS variation over one second of each wavefront jitter aberration is determined, which defines the amplitude of the sinusoidal oscillation. The intensity images are computed at each of the 10 samplings of this cycle using the fitted coefficients and added to the pointing jittered image. The images at the four polarization components are then summed to represent an image observed without a polarizer. If the HLC is being modeled, the images at each wavelength are added together with weights appropriate for the system transmission and stellar spectrum, producing a data cube of dimensions [N_x, N_y, N_{times}], where N_{times} = 3630 here. If it is the SPC, then the separate wavelength images are kept to allow interpolation to finer spectral resolution in the IFS model, resulting in a data cube of dimensions [N_x, N_y, N_λ, N_{times}], where N_λ = 11 here.

IFS model

The IFS model[24] is developed and maintained by NASA Goddard. Written in Python, it takes as input the speckle field intensities versus time and wavelength. These are first projected onto the lenslet array, which has a sampling of λ/2D (Nyquist). The correspondence between the mean intensity over the lenslet and the resulting distribution in the spectrally-dispersed image on the detector is precomputed for each lenslet, allowing for a rapid mapping of input speckle field to output spectra. This lookup table is generated by propagating a uniform beam from the lenslet to its focus at a pinhole (each lenslet has a separate pinhole) and then through collimating optics using PROPER. Because PROPER does not explicitly handle propagation through a prism like that used in the IFS, the effective dispersion and induced aberrations are added using values computed with ray tracing software. The beam is then propagated through a couple more lenses and onto the detector plane, with distortions included (also computed via ray tracing). This rapid method was compared to a much more computationally complex propagation of the inter-lenslet, spatially-varying E-field through the IFS with good agreement and orders of magnitude speed improvement. The IFS and its model are described in more detail elsewhere.

Detector model

The EMCCD has traits that are notably different from normal CCDs, and thus it needs different modeling techniques. Operating in photon counting mode at high gain (3000 – 5000), there is effectively zero read noise. The resulting measured amplified signal is determined from a pulse height probability distribution function. Additional EMCCD-specific clock-induced charge (CIC) noise is added, along with dark current and hot pixels. These noise factors are dependent on the amount of radiation damage over time (mostly from solar flares), so an estimate is made of the mean radiation experienced over a specified time at the L2 orbit. Cosmic rays are also added, and because they saturate the serial register they produce a horizontal trail in the images. The details of the detector model is beyond the scope of this paper, so the reader is referred elsewhere.

4. MODELING UNCERTAINTY FACTORS

In a system as complex as WFIRST, where some of the technologies and components are unproven in space or have not been finely characterized, there is a good likelihood that things will not behave as expected. For example, the telescope may have a greater response to temperature changes than the STOP models show due to imperfectly known coefficients of thermal expansion, or the reaction wheels may vibrate the optics more than predicted due to imbalances. Due to the high sensitivity of the coronagraph, even tiny errors in the assumptions of the state of the system at any given time can have a large impact on performance estimates.

To avoid over-optimistic results, modeling uncertainty factors (MUFs) are applied to various critical parameters to account for unknowns in the assumptions. Some MUFs are based on prior experience with similar space-qualified components or

on Monte Carlo-type simulations encompassing various probability distributions. Since no coronagraph with wavefront control has flown in space, many of the MUFs for the WFIRST coronagraph models are based on testbed experiments. In this case, uncertainties in the testbed (e.g., limited calibrations of the DM actuator-to-actuator gains) can lead to MUFs that may be reduced in time as more experience and calibrations are obtained (notably sensitivities to low-order aberrations). Some WFIRST MUFs are imposed by the top-level project management (e.g., observatory STOP modeling), while others are introduced lower down the chain (e.g., low-order aberration sensitivity of the HLC).

Some MUFs multiply together, potentially driving the models to unexpectedly poor performance levels. For example, the coronagraphs in the testbeds show ~2× higher sensitivities to low-order aberrations in terms of contrast changes than expected from the physics of the designs, possibly due to the limitations in the calibration of the testbed rather than the behavior of the coronagraph. Thus, there is a low-order aberration contrast sensitivity MUF of 2. The STOP models also have a MUF of 2 on displacements caused by thermal changes, including optical offsets and deformations, resulting in a similar ~2× change in wavefront aberrations. Because intensity (contrast) is proportional to the square of the wavefront error, the corresponding speckles can increase in brightness by $2^2 \times 2 = 8\times$. As another example, a MUF of 1.5 is applied to the polarization-induced aberrations computed by ray tracing software like Code V to account for possible errors in the coating uniformities or indices of refraction (this MUF is an educated guess, given that such aberrations have not been measured to high accuracies on optics of the size of WFIRST's). The 1.5× higher aberrations lead to $1.5^2 = 2.25\times$ higher contrast speckles when wavefront is converted to contrast. The polarization errors are low-order, so these also get multiplied by the 2× sensitivity MUF for a total increase of 4.5× in speckle brightness (Figure *6*).

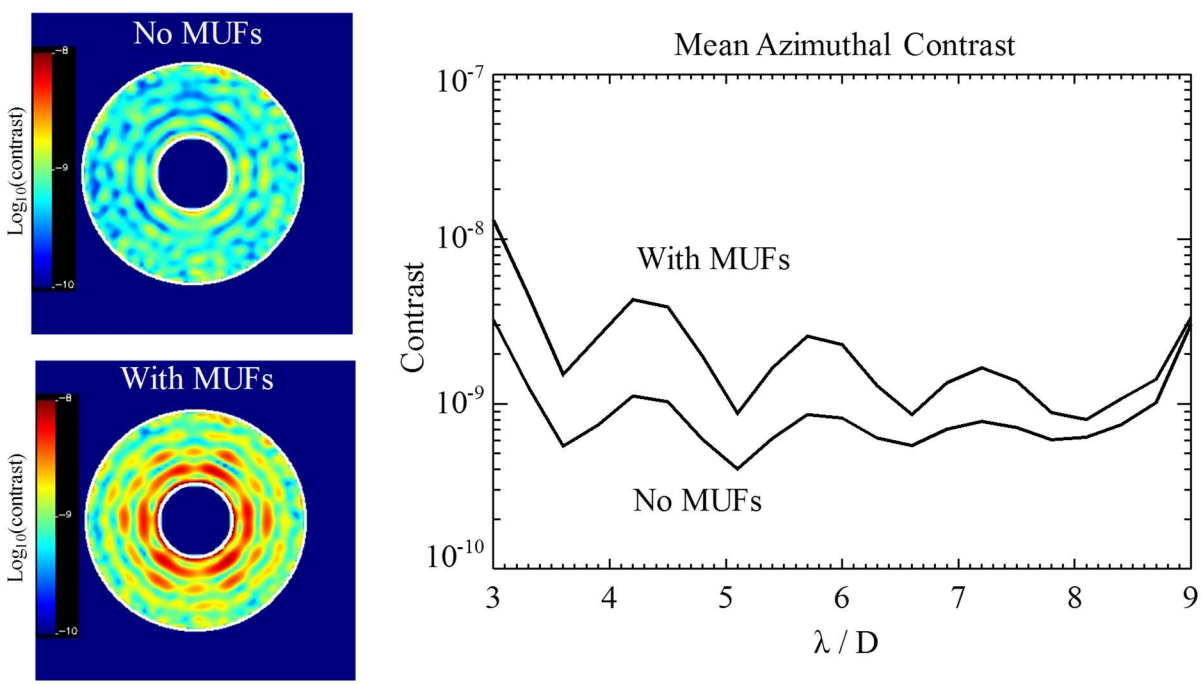

Figure 6. The impact of polarization-related optical MUFs on dark hole contrast. The HLC (λ_c = 575 nm bandpass) post-EFC dark hole fields on the left include the system static aberrations including polarization. On top is the result using the as-predicted polarization aberrations and low-order aberration contrast sensitivities. On the bottom is the result with the polarization aberrations increased by a factor of 1.5 and the low-order sensitivities increased by a factor of 2. On the right are azimuthal-mean plots of the fields.

The MUFs used in the simulations presented here are shown in Table 1. We distinguish "structural" MUFs (deformations, including the resulting wavefront errors, and jitter) from optical MUFs (aberration sensitivities and polarization). We apply the former to all of our models, while for our "optimistic" cases the optical MUFs are omitted.

Table 1. WFIRST CGI Modeling uncertainty factors (MUFs)

Error	MUF	Applied as	MUF source
Low-order aberration contrast sensitivity	2.0	$\sqrt{2} \times phase\ error$	Testbed
Polarization aberrations	1.5	$1.5 \times phase\ error$	Guess
Tip/tilt contrast sensitivity	2.0	$1.6 \times jitter$	Testbed
Thermally-induced structural displacements	2.0	$2 \times displacements$	JWST
Wavefront jitter & post-FSM pointing jitter due to reaction wheel vibrations	2.04 (f < 20 Hz) 2.04 – 4.13 linearly from f = 20 - 40 Hz 4.13 (f > 40 Hz)	$MUF \times jitter\ power\ spectrum$	JWST

5. MODELING RESULTS

5.1 STOP results

The OS6 STOP model for the full 138 hours was run by the modeling groups both at Goddard and JPL for 2 hour timesteps with a 60 second thermal solver resolution; the JPL results for defocus are shown in the top panel of Figure 7. This took about 1 day to execute. After verifying that the GSFC and JPL models agree, the scenario was run again in the JPL Integrated Modeling Pipeline for 10 minute timesteps and a 10 second thermal resolution. These results, shown in the lower panel of Figure 7 and in Figure 8, took a week to run. The improvement in smoothness of the oscillating pattern due to the higher thermal solver resolution is apparent, with the remaining high-frequency noise apparently caused by the non-optimized thermal control loops in the model.

Other low-order aberrations are shown in Figure 8. Defocus is by far the largest, due mostly to the change in spacing between the primary and secondary mirrors in response to the differences in solar illumination between pointings on the two stars. There does not appear to be any sensitivity to the two roll orientations on 47 UMa; this is as intended, with the rolls chosen to maintain the same observatory cross-section with respect to the Sun. The other aberrations are ±15 pm RMS or less, and since the LOWFS/C measurement and correction errors (e.g., DM gain errors) would introduce larger disturbances, only defocus was compensated, using the FOM, in the OS6 diffraction runs.

The pupil shear at the FSM versus time is shown in Figure 9, and it seems to be mostly from motion at the latches between the IC and CGI. The model-predicted shear sensitivity curve in Figure 10 shows that the ~0.9 µm of shear would introduce ~2×10^{-10} of dark hole contrast change over OS6. While this seems small compared to the >10^{-9} required contrast performance, it is just one component of the speckle stability budget.

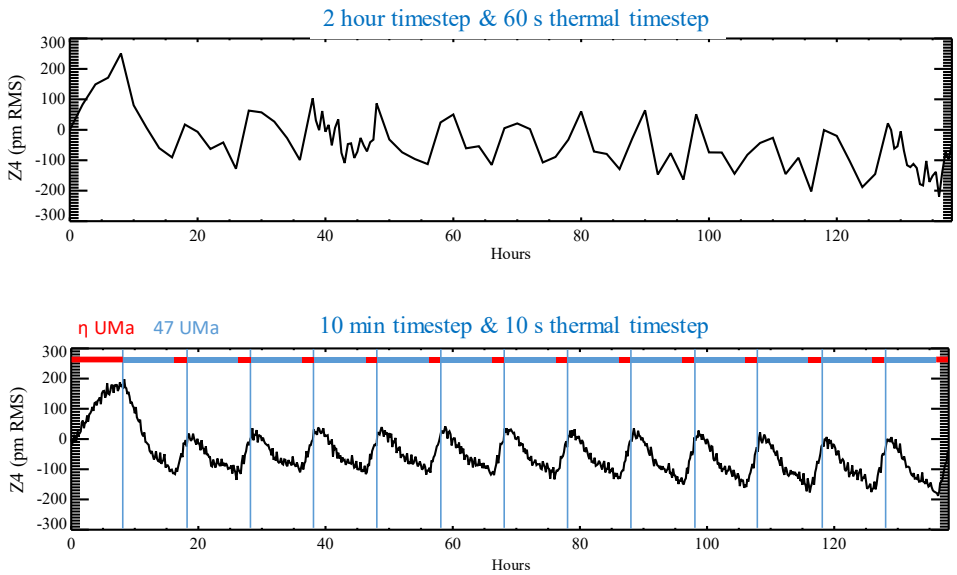

Figure 7. Plots of defocus (RMS picometers) at the pupil located at the FSM versus time over OS6. (Top) Results at 2 hour cadence (except for 2 spans at 30 minute cadence) with 60 second thermal solver resolution; (Bottom) Results at 10 minute cadence with 10 second thermal solver resolution. The timespans corresponding to pointings on the reference star η UMa are indicated in red and those on 47 UMa in blue.

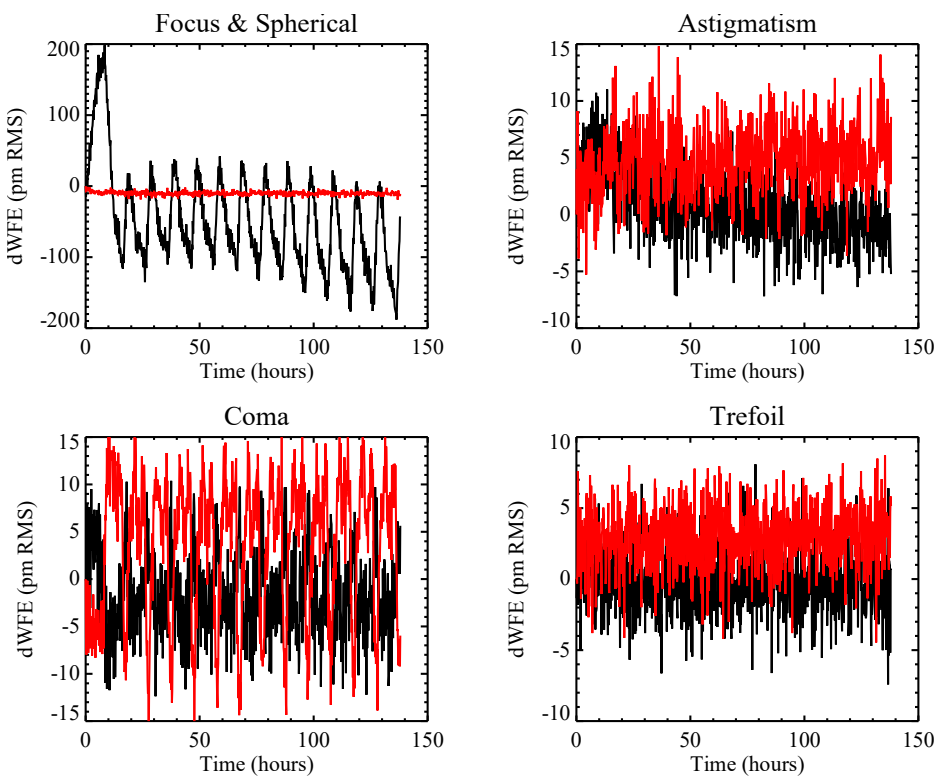

Figure 8. Plots of OS6 low-order aberrations (RMS picometers) at the pupil at the FSM versus time. The red and black lines for astigmatism, coma, and trefoil indicate the two orthogonal components of those aberrations.

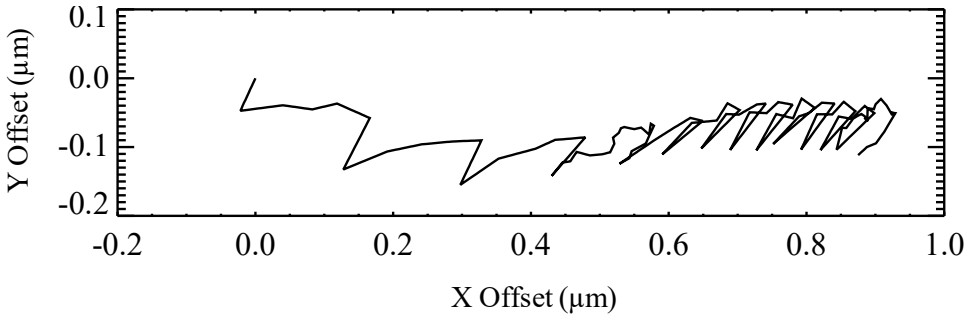

Figure 9. Offset of the FSM with respect to the chief ray versus time over OS6 (starting at 0,0), corresponding to pupil shear at the DMs. The beam diameter at the FSM is 40 mm. The loops are associated with one iteration of the chopping sequence.

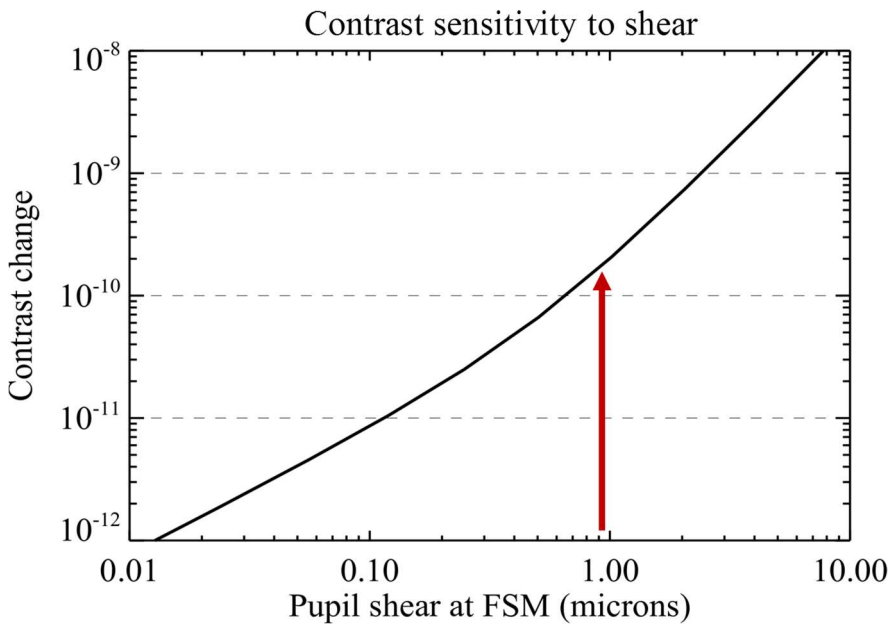

Figure 10. Model-predicted change in mean HLC dark hole contrast due to pupil shear at the FSM, primarily due to shear downstream at the DMs. The arrow indicates the amount of contrast change (~2×10^{-10}) over OS6 due to shear.

5.2 Jitter results

The Goddard modeling group computed the wheel speeds and corresponding pointing and wavefront jitters over a 20 hour span of OS6, beginning with 2 hours on η UMa. The telescope pointing jitter was reduced to the level expected post-FSM using a frequency-dependent filtering function specified by JPL. The RMS variation in post-FSM pointing jitter over 60 second intervals is shown in Figure 11. As expected, the jitter is highest when the wheel speeds are fastest, at the beginning and end of the 20 hour timespan. The jitter is below 0.5 mas RMS over ~85% of the span, which is favorable since both the HLC and SPC are largely insensitive to that amount (excluding application of the sensitivity MUF). As with pointing jitter, the wavefront jitter follows a similar trend with wheel speeds (Figure 12). The largest such jitter is in focus, which varies by ±30 pm RMS.

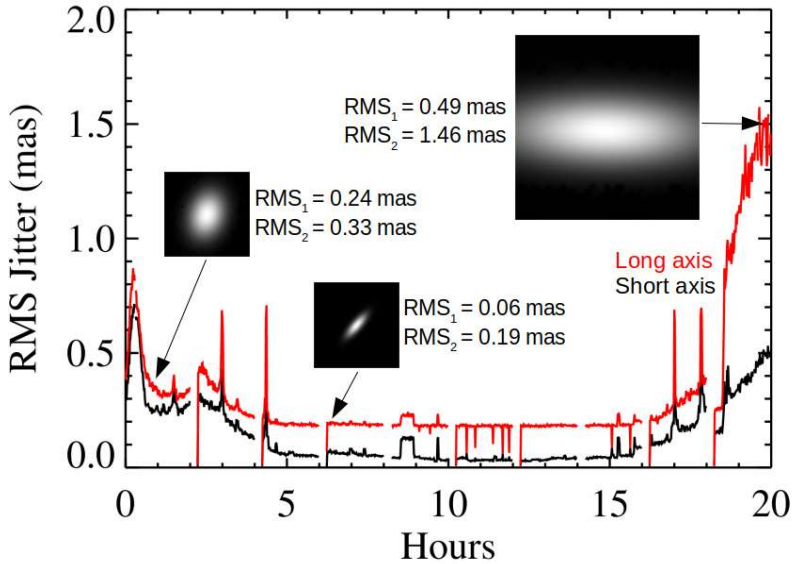

Figure 11. Predicted post-FSM pointing jitter (mas RMS per axis) over a 20 hour span of OS6, starting with 2 hours on η UMa followed by 8 hours on 47 UMa, and then a repeat of this sequence. The red line shows the major axis jitter and the black line is the minor axis (the orientation of the elliptical Gaussian jitter distribution varies with time as well). Realizations of the two-dimensional jitter distributions are shown at specific times. This sequence is repeated every 20 hours over OS6, with small random offsets introduced for each span.

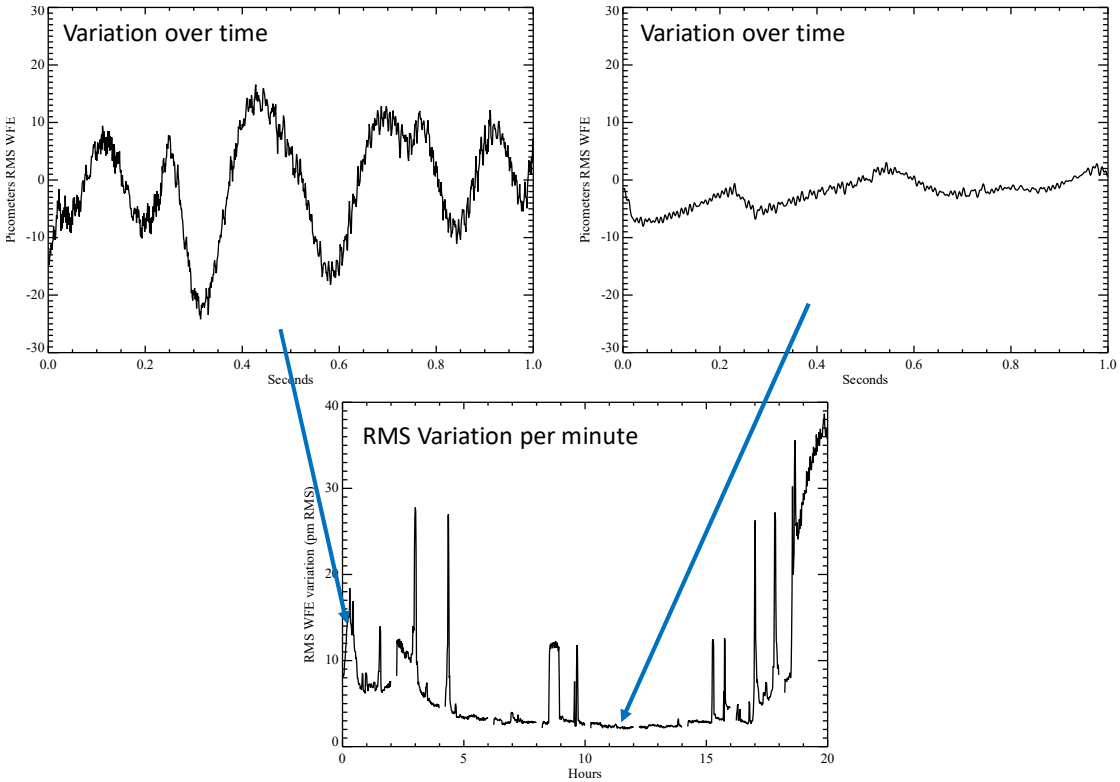

Figure 12. (Bottom) Focus wavefront jitter (RMS over 60 sec intervals) over the 20 hour timespan of OS6 corresponding to the predicted wheel speed variations. (Top) Realizations of instantaneous focus variation over a one second timespan computed by the JPL Pointing Group for two specific times.

5.3 Optical modeling results

The speckle fields incorporating the static errors, LOWFS/C corrected thermal drifts, DM drifts, and pointing and wavefront jitter are shown in Figure 13Figure 14. Figure 15Figure 16 illustrate the speckle brightness variation over time for the with-MUFs HLC OS6 fields. These time series are available from **wfirst.ipac.caltech.edu**.

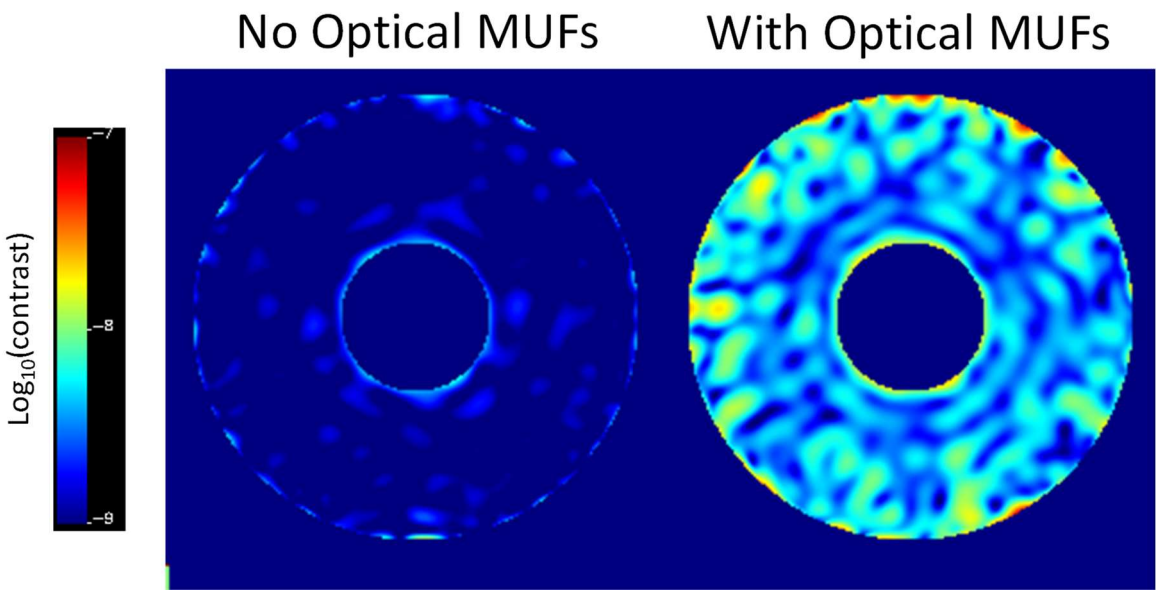

Figure 13. Contrast maps of the OS6 HLC 575 nm 10% bandpass speckle fields (Left) without optical MUFs and (Right) with all MUFs. The first image of the time series is shown for each. The inner and outer annulus radii are r = 3 & 9 λ/D.

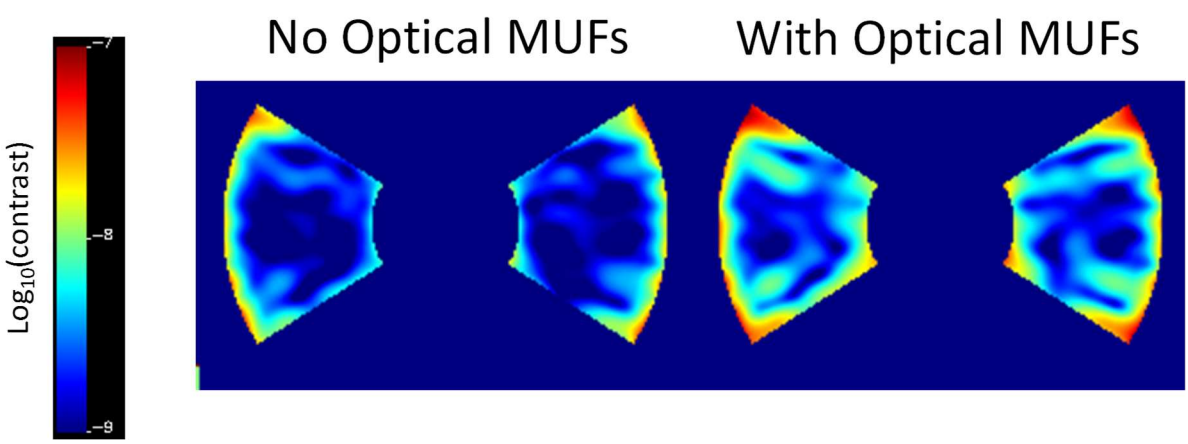

Figure 14. Contrast maps of the OS6 SPC 770 nm 18% bandpass speckle fields (Left) without optical MUFs and (Right) with all MUFs. The first image of the time series is shown for each. The inner and outer annulus radii are r = 3 & 9 λ/D.

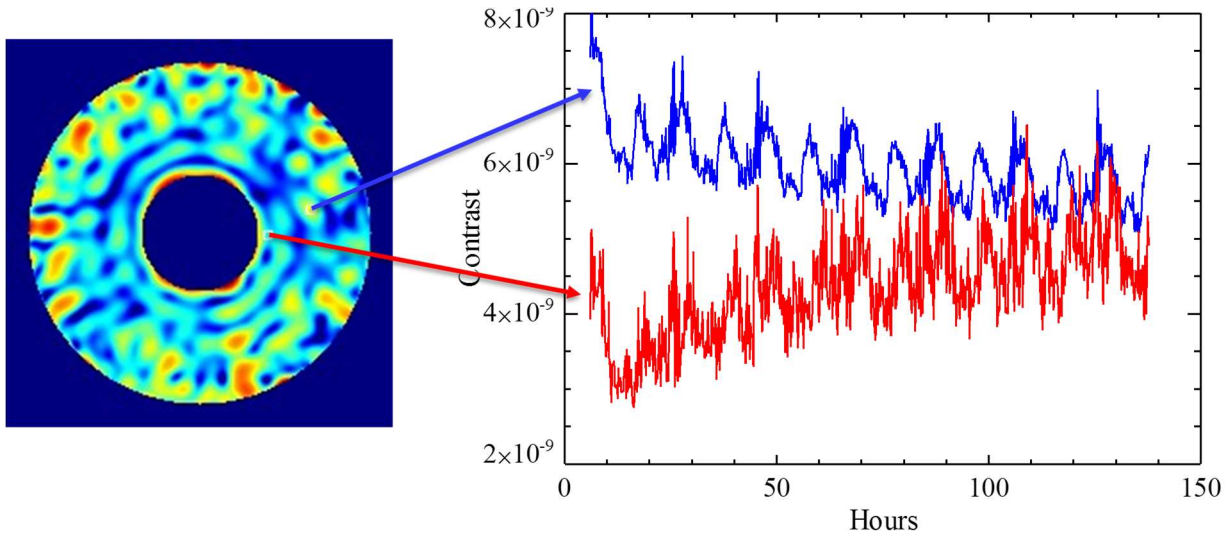

Figure 15. (Left) Intensity map of the time-averaged OS6 HLC 575 nm 10% bandpass dark hole speckles (all MUFs included). (Right) Plots of the contrast variations with time of selected 0.3 × 0.3 λ/D regions of the fields.

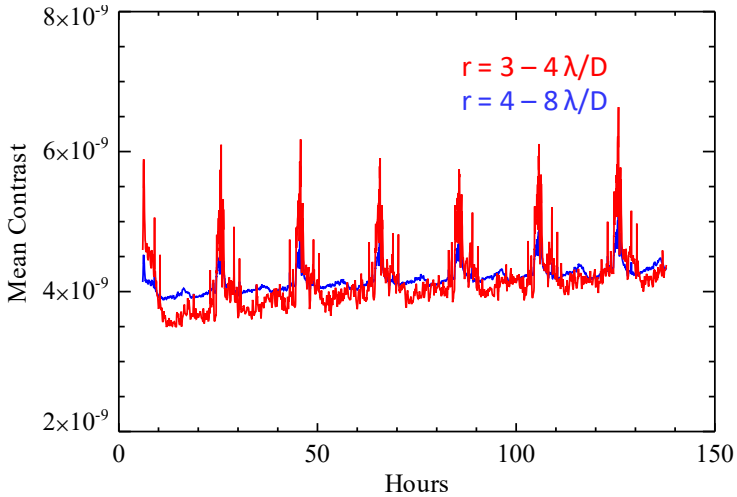

Figure 16. OS6 mean speckle contrast versus time in the HLC 575 nm 10% bandpass fields in annuli of r = 3 – 4 λ/D (red) and r = 4 – 8 λ/D (blue); all MUFs included. The increase in speckle brightness every 20 hours is due to large pointing jitter as a result of the relatively high reaction wheel speeds.

5.4 IFS modeling

The SPC speckle fields versus time and wavelength were provided to the NASA Goddard team for processing through their IFS model. See elsewhere in these proceedings[25] for these results.

5.5 Detector modeling

The OS6 HLC speckle fields were processed with the EMCCD detector model assuming 3 sec frame rates for η UMa and 30 sec for 47 UMa. Since the speckle fields themselves were generated at 2 min timesteps, each field was replicated 10 or 40 times, respectively, and then processed. The fields were normalized to the predicted stellar flux rates. For both stars the gain = 5000 e^-/photon, dark current = 0.5 e^-/pixel/hour, read noise = 99 e^-/readout, hot pixel rate = 3 e^-/hour, CIC noise = 0.01 e^-/pixel/readout, cosmic ray rate = 5.0 e^-/cm^2/sec, and the on-orbit lifetime = 1 year. The results for the no-optical-MUFs fields are shown in Figure *17*.

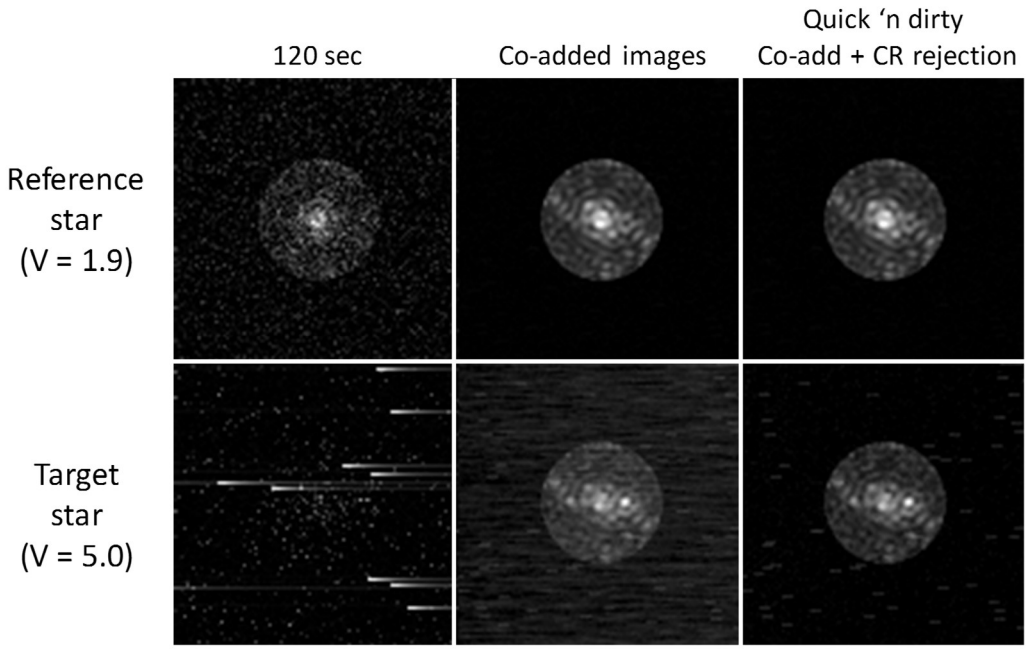

Figure 17. OS6 HLC speckle fields (no optical MUFs) processed with the EMCCD detector model. The top row shows the reference star (η UMa) and the bottom the target (47 UMa). The left column shows the summation of 120 sec worth of frames. Cosmic ray trails are especially obvious in the target star image. The middle column shows the summation of all frames for the reference and target (roll #1) star, without any cosmic-ray rejection. The left column shows the weighted summation after all rows in each image containing cosmic rays were set to zero. Two planets are visible in these images (the mean speckle contrast is 8×10^{-10}); see These planets would not be discernable from the background speckles in the with-MUFs case.

5.6 Simple post-processing

As discussed earlier, to robustly discern the exoplanets from the speckles requires additional image processing after the images have been obtained. One can either use the image of the reference star to subtract out speckles from the target (RDI), or use the images of the target at two orientations (ADI). Both methods were applied to the coadded, cosmic-ray-rejected images. Exoplanets with contrast ratios of 1×10^{-8} and 3×10^{-9} were added to the target images at distances of 3.5 and 4.5 λ/D, respectively, from the star prior to running the detector model. The RDI results shown in Figure *18* are simple target – reference subtractions at each target roll angle. The ADI results, which used only the target images at the two rolls, utilized an iterative solver, the same algorithm that has been used on Hubble Space Telescope coronagraph images. In both RDI and ADI cases the background speckles have been subtracted, even in the presence of detector noise and MUF-amplified instabilities.

Figure 19 shows plots of the azimuthal RMS variations in the pre- and post-processed speckle fields, including detector effects. These represent the background noise against which a planet is detected. ADI produces slightly better results, but it cannot always be used; if the target has a circumstellar disk, using ADI may cause the disk to subtract portions of itself during the processing. Subtraction seems to be more effective in the with-MUFs case, in terms of factors of improvement. However, given that the subtraction results for both MUF cases are similar, it is likely that detector noise, much of which is independent of image brightness, dominates much of the residuals.

6. SUMMARY AND FUTURE WORK

The simulation of a WFIRST CGI observing scenario requires a large amount of inputs and weeks of computer time. Static system errors such as surface polishing defects and polarization-induced aberrations define the minimum instrumental speckle level. Dynamic errors such as thermally-induced deformations and reaction-wheel-vibration-induced pointing and wavefront jitter alter the speckle patterns over time, setting a limit on the ability of post-processing to remove the speckles

to allow detection of faint planets and circumstellar disks. STOP modeling at NASA Goddard and JPL provides the thermal changes while Goddard produces estimates of jitter. These are incorporated into the diffraction models that create time series of speckle fields. These may be processed through a model of the IFS and/or into a EMCCD detector model. These datacubes provided to the community allow analysis of the speckle variations and testing of post-processing algorithms.

The results shown here use the Phase A STOP model. In Phase B a much higher fidelity model of the CGI will be incorporated. The current results do not include MUFs on the thermal variations, just the resulting structural deformations. It is likely that new MUFs will be applied in Phase B, some due to uncertainties in thermal sensing and control hardware. As the models of the reaction wheels gain fidelity, it is also possible that the jitter performance may degrade from what is presented here.

The coronagraph designs use here are from 2015 (HLC) and 2017 (SPC). There is a reluctance to incorporate newer designs using technologies that have not yet been verified on the testbeds, such as azimuthally-varying HLC FPM patterns. As new designs are tested, the improved performance they may offer will be incorporated into the models.

The detector model currently does not include CTE. A separate model does, but it is not currently efficient enough to process the large number of frames produced by an observing scenario. Work will continue on finding a faster but sufficiently accurate method.

As new observing scenarios and revised models are executed, the results will be made available to the public at the IPAC WFIRST web site: **wfirst.ipac.caltech.edu**.

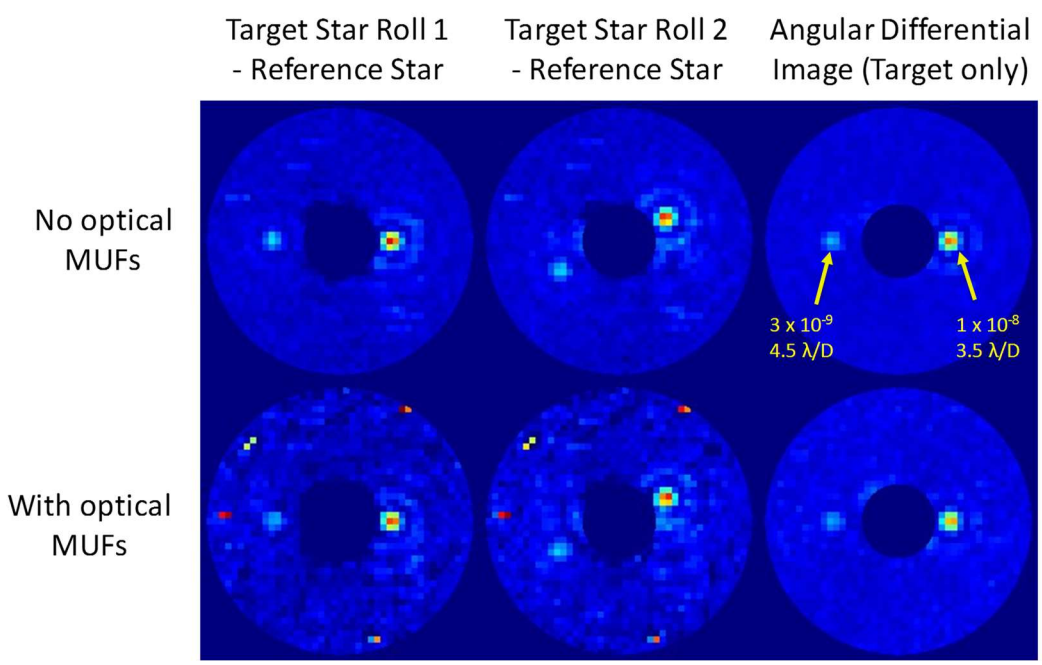

Figure 18. Simple post-processing results of the OS6 HLC simulated fields. The two indicated planets were added to the target star fields prior to running them through the EMCCD model. They were added at the correct orientations for the two rolls. All of the frames for each star (and for each roll) were coadded with cosmic ray rejection. The left and middle columns show RDI applied to each orientation of the target star. The right column shows the result of running an iterative ADI algorithm on the two rolls of the target star.

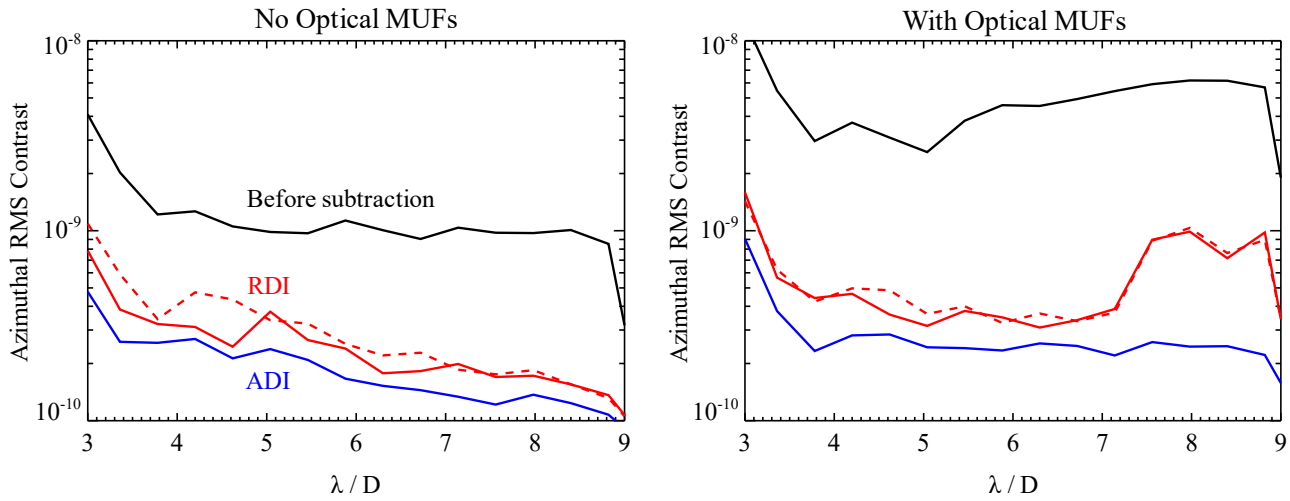

Figure 19. Plots of the azimuthal RMS variations (excluding planet regions) of the unsubtracted target speckle field (black), the RDI subtraction (red, solid and dashed for the two rolls), and ADI subtraction (blue). These represent the background noise against which a planet may be detected (note: these are not detection limits, only background noise).

ACKNOWLEDGEMENTS

The authors thank the NASA Goddard WFIRST Integrated Modeling Team for their work with the STOP and jitter models. This work was performed at the Jet Propulsion Laboratory/California Institute of Technology under contract to NASA. The decision to implement WFIRST will not be finalized until NASA's completion of the National Environmental Policy Act (NEPA) process. This document is being made available for information purposes only.

REFERENCES

[1] D. Spergel, et al., "Wide-Field InfraRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA Final Report", http://arxiv.org/abs/1305.5422 (2013).
[2] Kasdin, N. Jeremy, et al. "The WFIRST coronagraph instrument: technology demonstration and science potential," Proc. SPIE 10698-87 (2018).
[3] Krist, J., Nemati, B., Mennesson, B., "Numerical modeling of the proposed WFIRST-AFTA coronagraphs and their predicted performances," Journal of Astronomical Telescopes, Instruments, and Systems, 2, 011003 (2016).
[4] Krist, J., et al., "WFIRST coronagraph optical modeling," Proc. SPIE 10400, 10400-4 (2017).
[5] Tang, H., Demers, R., Krist, J., McGuire, J., Mayer, R., Zhao, F., "The WFIRST coronagraph instrument optical design update," Proc. SPIE 10400, 10400-2 (2017).
[6] Demers, R., "Review and update of WFIRST coronagraph instrument design and technology," Proc. SPIE 10698-98 (2018).
[7] Shi, F., et al., "Low-order wavefront sensing and control for WFIRST-AFTA coronagraph," Journal of Astronomical Telescopes, Instruments, and Systems, 2, 011021 (2016).
[8] Shi, F., et al., "WFIRST low order wavefront sensing and control performance under low photon flux," Proc. SPIE 10698-94 (2018).
[9] Trauger, J., Moody, D., Krist, J., Gordon, B., "Hybrid Lyot coronagraph for WFIRST-AFTA: coronagraph design and performance metrics," Journal of Astronomical Telescopes, Instruments, and Systems, 2, 011013 (2016).
[10] Trauger, J., Krist, J., Moody, D., Seo, B.-J., "Hybrid Lyot coronagraph instrument (CGI) design for high tolerance to WFIRST telescope aberrations," Proc. SPIE 10698-91 (2018).

[11] Eldorado Riggs, A.J., Zimmerman, N., Shaklan, S., Jewell, J., Gersh-Range, J., "Shaped pupil coronagraph design developments for the WFIRST coronagraph instrument," Proc. SPIE 10400, 10400-73 (2017).
[12] Mandell, A., et al., "Current science requirements and planned implementation for the WFIRST-CGI Integral Field Spectrograph (IFS)," Proc. SPIE 10400, 10400-9 (2017).
[13] Groff, T., et al., "WFIRST integral field spectrograph design and operations concept", Proc. SPIE 10698, 10698-84 (2018).
[14] Give'on, A., et al., "Broadband wavefront correction algorithm for high contrast imaging systems," Proc. SPIE 6691, 66910A (2007).
[15] Harding, L., et al., "Electron multiplication CCD detector technology advancement for the WFIRST-AFTA coronagraph," Proc. SPIE 9605, 96050F (2015).
[16] Harding, L., et al., "Technology advancement of the CCD201-20 EMCCD for the WFIRST coronagraph instrument: sensor characterization and radiation damage," Journal of Astronomical Telescopes, Instruments, and Systems, 2, 011007 (2016).
[17] Ygouf, M., et al., "Data processing and algorithm development for the WFIRST coronagraph: comparison of RDI and ADI strategies and impact of spatial sampling on post-processing," Proc. SPIE 9904, 99045M (2016).
[18] Ygouf, M., Pueyo, L., Soummer, R., Perrin, M., van der Marel, R., Macintosh, B., "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. SPIE 9605, 96050S (2015).
[19] Ygouf, M., Zimmerman, N., Pueyo, L., Perrin, M., Soummer, R., "Data Post-Processing and Algorithm Development for the WFIRST Coronagraph: FY16 Interim Report," Space Telescope Science Institute Technical Report WFIRST-STScI-TR1605 (2016).
[20] Roulet, M., et al., "Superpolished OAPs for WFIRST CGI," Proc. SPIE 10698-96 (2018).
[21] Casey, T., et al., "The wide field infrared survey telescope (WFIRST) observatory: design formulation (phase-A) overview," Proc. SPIE 10698, 10698-82 (2018).
[22] Saini, N., Anderson, K., Chang, Z., Gutt, G., Nemati, B., "IMPipeline: an integrated STOP modeling pipeline for the WFIRST coronagraph," Proc. SPIE 10400, 10400-8 (2017).
[23] Krist, J., "PROPER: an optical modeling program for IDL," Proc. SPIE 6675, 66750P (2007).
[24] Rizzo, M., et al, "Simulating the WFIRST coronagraph Integral Field Spectrograph," Proc. SPIE 10400, 10400-11 (2017).
[25] Rizzo, M., et al., "WFIRST CGI integral field spectrograph performance and post-processing in the OS6 observing scenario," Proc. SPIE 10698, 10698-246 (2018).