WFIRST coronagraph optical modeling

John Krist[a], A.J. Riggs[a], James McGuire[a], Hong Tang[a], Nikta Amiri[a], Gary Gutt[a], Luis Marchen[a], David Marx[a], Bijan Nemati[b], Navtej Saini[a], Erkin Sidick[a], Hanying Zhou[a]
[a]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA USA 91109; [b]Center for Applied Optics, University of Alabama, 301 Sparkman Dr., Huntsville, AL USA 35899

ABSTRACT

End-to-end numerical optical modeling of the WFIRST coronagraph incorporating wavefront sensing and control is used to determine the performance of the coronagraph with realistic errors, including pointing jitter and polarization. We present the performance estimates of the current flight designs as predicted by modeling. We also describe the release of a new version of the PROPER optical propagation library, our primary modeling tool, which is now available for Python and Matlab in addition to IDL.

Keywords: WFIRST, coronagraph

1. INTRODUCTION

The Wide-Field Infra-Red Survey Telescope (WFIRST) is a 2.4 m diameter space telescope that is currently planned to be launched into an L2 orbit in 2025. Its primary purpose is wide-field imaging to determine the dark matter distribution in the universe using a large format imaging detector array, as well as planet detection via microlensing. It will also have a separate stellar coronagraph for imaging and spectrally characterizing large extrasolar planets and circumstellar dust disks in reflected light. This will be the first highly optimized coronagraph in space (i.e., with active wavefront control) and the first one capable of imaging 10^{-9} contrast planets. The WFIRST telescope was not designed for astronomical observations, and certainly not coronagraphic ones. The on-axis telescope configuration creates an obscured pupil due to the secondary mirror and six relatively thick support struts (Figure 1). High contrast coronagraphs, however, favor clear, unobscured apertures that minimize the diffracted light.

We discuss here updates since our last report[1] to the design of the WFIRST coronagraphs and their predicted performances[2] based on numerical optical modeling. The reader is referred to there for details of the modeling procedures that are not repeated here.

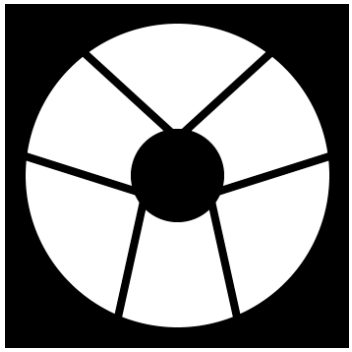

Figure 1. The WFIRST obscuration pattern.

© Copyright 2017 California Institute of Technology. All rights reserved.

1.1 Coronagraph requirements for high-contrast imaging

The WFIRST coronagraph is designed to image and characterize planets with flux contrast ratios relative to their stars of ~10^{-9}. It must first suppress the considerable diffraction due to the telescope's obscurations (the primary mirror's edge, the secondary mirror's shadow, and six thick secondary support struts). This is accomplished by either modifying the pupil pattern (as with the shaped pupil coronagraph, SPC) or via a combination of wavefront modulation using deformable mirrors, focal plane masks, and Lyot stops (as in the Hybrid Lyot coronagraph, HLC). Next, it must correct for the static wavefront errors caused by imperfections in the optics that create speckles of scattered light in the region around the star. Two deformable mirrors can be used to control both phase and amplitude errors to dig the dark hole region inside which the planets can be seen. The remaining speckles, which are as bright as the planets of interest, must be kept stable to about 2×10^{-10} in contrast in order to remove them via post-observation image processing methods such as reference star subtraction. The coronagraph must thus sense the time-dependent wavefront errors on the order of ten picometers caused by thermal, pointing, and vibration variations and compensate for them. This requires a low-order wavefront sensor (LOWFS) and wavefront phase modulators (e.g., fast steering mirror, focus adjustment mechanism, deformable mirrors) to keep the wavefront stable to within tens of picometers.

1.2 The need for numerical modeling

It is not possible on the ground to replicate with sufficient fidelity the operating conditions that the coronagraph will encounter on orbit. While testbed experiments are critical to demonstrate the readiness of various components, they are limited in the variety of investigations that can be undertaken due to constraints of time, labor, budget, and environment. Detailed computational modeling of the coronagraph is therefore necessary to predict its on-orbit performance. The simulations can combine predictions of time-dependent wavefront variations from finite element thermal and structural modeling with computed optical propagation of the wavefront from optic to optic and with wavefront control included, creating realistic images. Pointing and wavefront jitter (from vibrations of the optics by the reaction wheels), finite stellar diameter, stellar color, optic misalignments, and detector defects can all be included as well. Parameters derived from the models (throughput, contrast, aberration sensitivity, speckle stability, post-processing limits) are used to estimate the number of planets that can be imaged and characterized[2].

2. CORONAGRAPH DESCRIPTION

2.1 Coronagraph layout

The layout has been modified over the past years to accommodate changes in the location and allocated volume on the spacecraft and to improve the wavefront stability[3]. The foreoptics include the primary and secondary mirrors and three flats to direct the beam behind the primary. These multiple flats (the Pickoff Mirror Assembly) are used to avoid incidence angles that would generate significant polarization-induced aberrations. Two off-axis parabolas (OAPs) and a couple of flats (one of which is mechanized to provide pupil alignment), collectively called the Tertiary Optical Mirror Assembly (TOMA), are used to resize the beam and correct for the off-axis aberrations resulting from the 0.4° offset of the coronagraph. The beam is then sent to a fast steering mirror (FSM), located near a pupil image; starting there, the remainder of the layout is shown in Figure 2. There are a considerable number of optics necessary to provide internal focus adjustment, additional separate pupil images for the reflective shaped pupil masks, deformable mirror (DM) #1, and Lyot stops, and intermediate foci for the focal plane masks (FPMs) and field stops. The two 48 × 48 actuator deformable mirrors are separated by 1.0 meter to provide phase and amplitude wavefront control. In imaging mode the beam is sent directly to a detector, currently set to be an electron-multiplication CCD (EMCCD)[4,5]. A fold mirror in the pupil wheel can divert the beam instead into an integral field spectrograph[6] (IFS; not shown in Figure 2) with its own EMCCD that provides $R = 50$ spectral characterization over the dark hole field.

The focus adjustment mirror (FAM), in conjunction with the low order wavefront sensor[7] (LOWFS), allows for correction of slow focus changes during an observation. The FSM works with the LOWFS to correct pointing errors at ~1000 Hz rates, reducing the ~12 mas RMS body pointing jitter to about 0.4 - 1.6 mas RMS. The Zernike phase contrast LOWFS looks at a reflection of the starlight rejected at the focal plane mask to measure changes in the low order wavefront aberrations with an accuracy of tens of picometers.

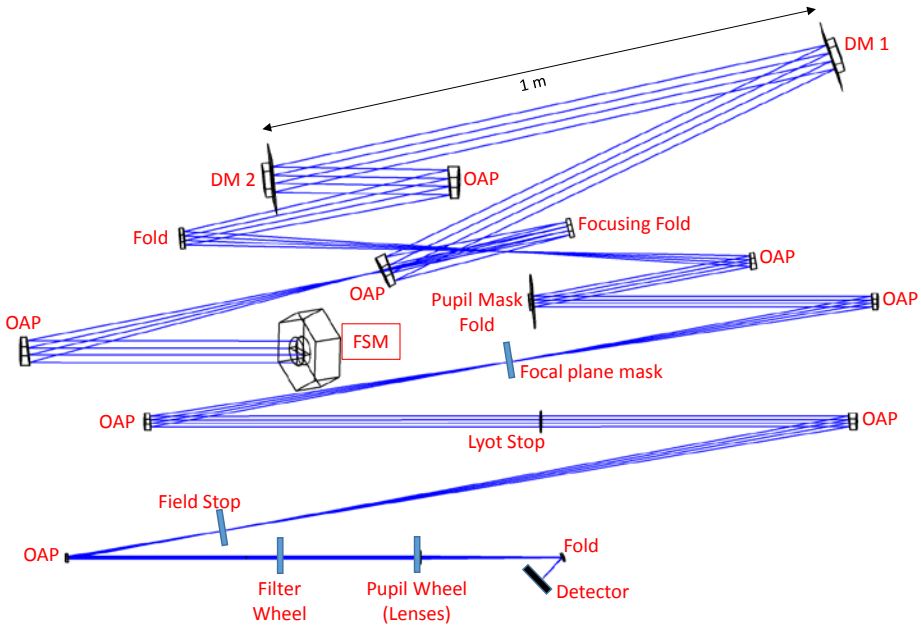

Figure 2. Current layout of the WFIRST coronagraph, starting at the Fast Steering Mirror (FSM). Prior to the FSM and not shown here are the telescope optics, a variety of fold mirrors, and some off-axis parabolas, with the telescope's optical axis oriented perpendicular to the page. Also not shown are the low order wavefront sensor and integral field spectrograph.

The coronagraph is designed to work over a ~2 arcsecond diameter field of view spanning λ = 450 - 900 nm with different filters. There are three 10% science imaging bandpasses (λ_c = 506, 575, & 661 nm) for detection and photometric color characterization (the 661 nm one is for CH_4 continuum measurement). There are also 5% - 6% bandpasses at λ_c = 721, 883, and 940 nm for imaging in and out of CH_4 and water absorption lines. Three 18% bands at λ_c = 660, 770, and 890 nm are available for the IFS.

2.2 Hybrid Lyot coronagraph

The Hybrid Lyot Coronagraph[8] (HLC) uses a phase-and-amplitude-modulating FPM at an intermediate focus followed by a Lyot stop at a reimaged pupil. Together with the masks, DM patterns are derived to compensate for the telescope's obscurations. The combination of the Lyot stop and the large strokes on the DMs (which scatter light to large angles) results in a low effective *core throughput*, which is the amount of a planet's light incident on the primary mirror that ends up in the core (≥50% of the peak) of its point spread function as seen at the detector (ignoring losses from reflections, filters, detector efficiency, etc.). Optimization of HLC designs over the past years has concentrated on maximizing throughput while minimizing sensitivities to polarization-induced aberrations and changes in low order aberrations from pointing jitter, optical vibrations, and thermally-induced structural deformations. Effort has also been made to reduce the DM actuator strokes.

The latest HLC design (HLC 20161228) developed by Dwight Moody at JPL provides a 360° field of view out to r = 9 λ_c/D (150 - 450 mas at λ_c = 550 nm) in a 10% bandpass. It uses an r = 2.8 λ_c/D partially-transmissive nickel disc overlaid with a radially-and-azimuthally varying dielectric coating (Figure 3). A matched Lyot stop (Figure 4) blocks the outer and inner obscurations and struts. DM patterns were derived that compensate for the obscurations in an unaberrated system and reduce pointing jitter sensitivity at the same time. The core throughput of this coronagraph is 4.5% (WFIRST without a coronagraph is 34%). Unlike a classic Lyot coronagraph, the complex phase modulation by the DMs and FPM results in a large fraction (24%) of the starlight falling within the clear aperture of the Lyot stop, and interference creates the dark hole zone around the star in the image plane. Because so much light makes it past the stop, a subsequent field stop must be used to block that which falls outside of the dark hole so that it is not scattered by back-end optics and does not saturate the detector.

In the simulations, the HLC is represented by (1) DM actuator pistons that initially are set to the as-designed, unaberrated-system values; (2) complex-valued representations of the FPM computed for each simulation wavelength, accounting for material properties and thin-film effects; and (3) the Lyot stop aperture. There is also a circular field stop mask placed at a focus between the Lyot stop and final image planes.

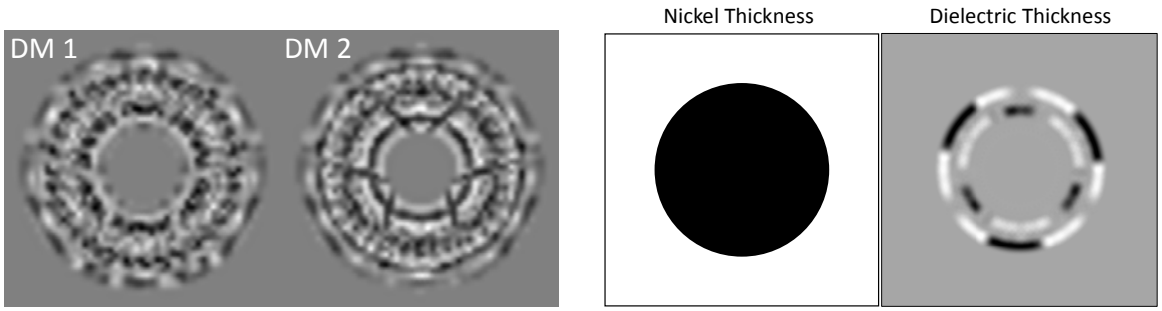

Figure 3. (Left) Deformable mirror surface variations used for compensation of the telescope obscurations in HLC 20161228; (Right) HLC focal plane mask coatings. The nickel coating is partially transmissive and 2.8 λ_c/D in radius.

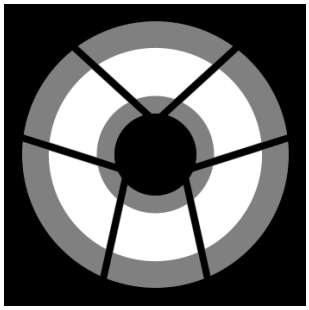

Figure 4. The HLC Lyot stop (grey) superposed on the WFIRST telescope obscuration pattern (black).

2.3 Shaped pupil coronagraph

The shaped pupil coronagraph[9,10] (SPC) redefines the obscuration pattern of the telescope using a binary mask at an intermediate pupil image plane. The openings in the mask are chosen to confine the diffracted light into a limited region, either in radius, azimuth, or both. The pupil mask blocks all of the telescope's obscurations and is slightly oversized to accommodate pupil alignment and magnification errors. A subsequent mask in an intermediate focal plane blocks the majority of the light that is diffracted outside of the field of interest. In a pure SPC, the FPM simply prevents scatter by optics after it or saturation of the detector by the bright areas outside of the dark hole region. In a Lyot coronagraph version of the SPC, like those being incorporated into WFIRST, a Lyot stop is added in a pupil plane after the FPM to aide diffraction suppression. This has advantages over a pure SPC in terms of contrast and throughput, but the additional elements increase the sensitivities to alignment errors and aberrations.

There are two SPC designs for WFIRST, one for characterizing planets with the IFS and another to provide a large field of view for imaging disks. The current characterization design (SPC 20170714; see Figure 5) is of the Lyot type and was optimized for throughput and aberration sensitivity in an 18% bandpass. The pupil mask consists of finite-width squares arranged on a 1000 by 1000 element grid to aide fabrication and avoid small apertures that could introduce anomalous diffraction effects. The mask can tolerate a translational displacement relative to the pupil of up to 0.2% of the pupil diameter before the telescope's obscurations become visible and contrast is degraded. The mask is located on a reflective flat substrate, with the mask defined by absorptive black silica; the substrate is on a wheel to allow a simple flat mirror to replace it when the HLC is being used. The FPM has two 65°, r = 2.6 - 9 λ_c/D (172 - 517 mas at λ_c = 660 nm) openings on opposite sides of the star. It is followed by a Lyot stop with 115° clear openings on opposite sides of the pupil between r = 0.25 and 0.82 of the pupil diameter. The core throughput is 4.3%. Because of its limited azimuthal field of

view, it is assumed that the exoplanet will have been detected previously in the 360° imaging channel with the HLC. Three sets of pupil masks, FPMs and matching Lyot stops will be provided to allow observation at any azimuth.

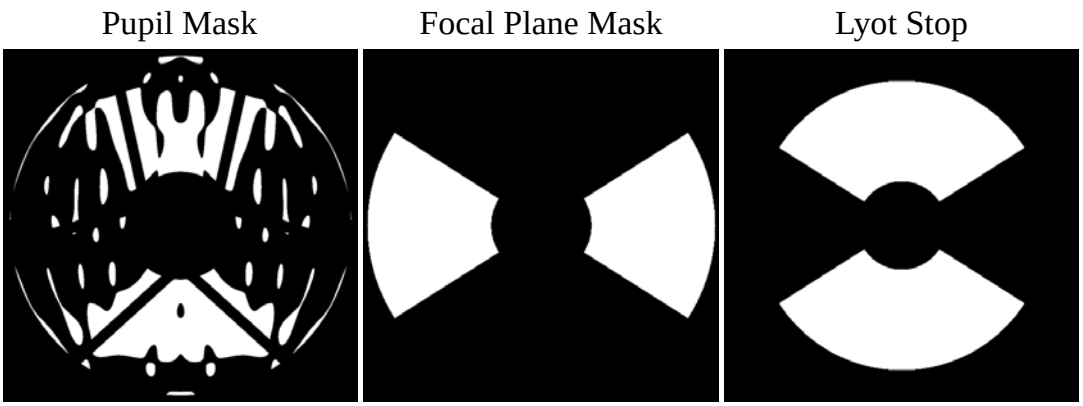

Figure 5. Shaped pupil coronagraph (SPC 20170714) pupil mask, focal plane mask, and Lyot stop.

The disk imaging SPC design offers a 360° field of view over r = 6.5 - 20 λ_c/D in a 10% bandpass in the imaging channel. A new version is being developed and thus will not be detailed further here.

In the models, the SPC pupil mask is represented at full 1000-pixel-diameter resolution. At the FPM the wavefront is upsampled and a high-resolution version of the mask (0.025 λ/D per pixel) is applied, after which the result downsampled. The Lyot stop is then applied conventionally.

2.4 Descope of the PIAACMC

In 2013 a review committee chose coronagraphs to continue development for WFIRST. HLC and SPC were designated as the primary methods and the PIAACMC (phase-induced amplitude apodization with a complex mask coronagraph[11]) as the backup. Though PIAACMC offered high throughput and a small inner working angle, modeling showed that its tolerance to low order aberrations, including polarization-induced ones, were too high. Work on it for WFIRST was therefore discontinued.

3. SYSTEM MODEL

3.1 Wavefront propagation using PROPER: new Python and Matlab versions

The performance of the WFIRST coronagraph is critically dependent on the characteristics of the wavefront errors in the system. Phase errors on an optic can morph into wavelength-dependent amplitude errors during propagation to the next optic (the Talbot effect). It is thus necessary to compute propagation from surface-to-surface using accurate but efficient algorithms. As with the prior coronagraph models, the results presented here used the PROPER[12] optical propagation software library, a freely available set of routines useful for modeling coronagraphs. 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. Aberrations are applied to each optic as described below. Broadband images are created by separately generating monochromatic images at a variety of wavelengths and adding them together in intensity.

PROPER was previously available for IDL (Interactive Data Language), and now there are Python (v2.7 & v3.x) and Matlab versions as well. PROPER is well documented with a complete manual and a number of examples. It can be downloaded from **`proper-library.sourceforge.net`**.

3.2 Wavefront control

The wavefront control procedures used to dig the dark hole, including Electric Field Conjugation[13] (EFC) and the simulation steps, were described in detail in the prior reports. Briefly, the complex-valued electric field at the final image plane is sensed by applying a series of particular patterns to the DMs, measuring the intensity changes on the detector, and then solving for the field using the model-predicted changes. The field is used in a linear equation along with model-predicted field changes due to each actuator piston (the Jacobian) to derive the DM settings necessary to minimize the energy in the dark hole. The model predictions are used for both the sensing and control stages, whether in simulation or in reality. Errors in the model, either in computational accuracy or (as is usually the case) incomplete knowledge of the system, can thus limit the achievable dark hole contrast.

3.3 Optical surface errors

All optics have surface errors from imperfect figuring (low spatial frequency), polishing (mid spatial frequency), and material roughness (high spatial frequency). Where in the system an optic resides can define the properties of the speckles in the final image plane created by these defects. Realistic estimates of the surface errors are thus important for accurate predictions of coronagraphic performance. Lacking actual maps of the optical errors in the WFIRST coronagraph, synthetic maps must be created. Both low order aberrations (described by Zernike polynomials) and mid-spatial frequency ones (described by power spectral density curves) are included; high spatial frequencies that contribute to large angle scattering are generally less important for coronagraphy since the field of interest around the star is small.

Interferometrically measured maps of the surface errors in the WFIRST primary and secondary mirrors exist. The errors can be well compensated for by the deformable mirrors, except at the outer edge of the primary where, like most large optics, the surface is overturned (the optics will be repolished and recoated, but the turnover cannot be corrected). While these maps can be included in the simulations, due to export regulations and manufacturer-imposed proprietary restrictions they cannot be accessed except by JPL or Goddard Space Flight Center employees. In the past studies synthetic primary and secondary maps were used instead, each having similar levels of low and mid spatial frequency aberrations as the actual optics. However, the synthetic primary map lacked the rollover seen in the real one. In early simulations comparing results using both types of maps there were no significant differences in the final results. The HLC is not sensitive to the rollover largely due to its Lyot stop. The early SPC design, which was a pure, non-Lyot version, was also not sensitive because there were no openings along the edge of the pupil, masking the rollover. However, recent SPC (Lyot) designs have thin apertures along the pupil edge, and the large wavefront errors (which are spatially too narrow to correct well with the DMs) severely limit the performance. To compensate for this the SPC 20170714 design[9] specifically masks out the outer edge.

At the moment, the only WFIRST optics in existence are the primary and secondary, so errors for all of the others in the system must be synthesized for inclusion in the models. Rather than simply making guesses as to what these aberrations may be, we instead utilized interferometric measurements of a variety of Gemini Planet Imager (GPI) off-axis parabolas and flats (kindly provided by Bruce Macintosh) to define the properties of our synthetic maps. Low order aberrations in each map were fit using Zernikes and then subtracted. The PSD was then computed and fit with a fitted power law as a measure of the mid-spatial frequency errors. For both GPI flats and OAPs, the PSD is a $\upsilon^{-3.5}$ power law (υ is spatial frequency) for $\upsilon > 10$ cycles/D (flats) and $\upsilon > 3$ cycles/D (OAPs), curving to a flat distribution below those limits. The low order aberrations in the synthetic maps were generated by randomly selecting values for each Zernike within the range of corresponding measured values. Inside the coronagraph (FSM and later), the OAPs have RMS low order wavefront errors of 1 - 6 nm with 1.8 nm RMS of higher orders while the flats are 1 - 5 nm RMS low order and 0.8 nm RMS higher. The larger optics between the secondary and coronagraph (five folds and two OAPs) are budgeted to have larger errors, twice the low order and about 5 times worse higher order errors than the coronagraph optics. These maps are filtered to ≤64 cycles/diameter to allow accurate interpolation to any wavefront sampling.

3.4 Polarization-induced aberrations

The WFIRST telescope optics create polarization-dependent wavefront errors that are significant at the contrast levels reached by the coronagraph[1]. The dominant aberrations are tilt (due to the coronagraph being located off axis in the telescope's focal plane) and astigmatism. Since different coronagraphs are sensitive to these low-order errors to varying degrees, their effect on dark hole contrast is coronagraph-dependent. The errors are functions of the incident angles of the reflections (worse in faster systems), the coating properties, and wavelength. The system wavefront can be decomposed into orthogonal polarization input/output direct components (*e.g.*, X_{in}/X_{out} & Y_{in}/Y_{out}) and cross-term (leakage) components (*e.g.*, X_{in}/Y_{out} & Y_{in}/X_{out}), all of which combine incoherently and thus cannot be fully corrected

via the DMs simultaneously. Without using a polarizer, the DMs can correct only the mean of these four wavefronts, and using a polarizer (e.g., X polarizer) they can correct the direct term (X_{in}/X_{out}) but not the cross-term (Y_{in}/X_{out}). The polarization variation with wavelength also imposes a limit to the ability to compensate for the induced aberrations over a finite bandpass. The polarization aberrations can be computed for any combination of input/output polarizations and wavelengths using common ray tracing software (*e.g.*, Code V or Zemax), include coating effects.

The WFIRST coronagraph layout and coatings have been designed to minimize polarization aberrations. While the fast telescope configuration cannot be changed, its enhanced silver coatings provide minimum polarization errors at $\lambda \approx 600$ nm, increasing rapidly to shorter wavelengths and less so toward longer ones. The coronagraph coatings are similar. To prevent the generation of large cross-terms, the coronagraph optics are oriented to provide as close to normal incidence as possible. In order to bend the beam behind the primary, for instance, three fold mirrors are used instead of a 45° fold.

Prior WFIRST coronagraph layouts included a beamsplitter (Wollaston prism) to produce two orthogonal polarization images on the same detector. The wavefront could then be optimized either for a single polarization (providing the best contrast) or both simultaneously (with degraded contrast). Recent analyses showed that the chromatic dispersion by the beamsplitter was unacceptable, so the layout was switched to a single channel with selectable polarizers in a wheel. To maximize throughput, most imaging is expected to be done without polarizers, and the wavefront control will be optimized for that. When polarization measurements are required, the appropriate polarizer can be inserted; in that case, the wavefront control will likely not be reoptimized for each polarization, so the speckles will have the same general contrast as those in the unpolarized image but with different morphologies.

For the simulations presented here, the aberration maps (both amplitude and phase) were computed using Code V for a range of wavelengths with the expected coatings with input polarizations of -45° and +45° and output polarizations of 0° (X) and +90° (Y); the ±45° inputs result in more readily-interpolated aberration maps and produce similar intensities in all four cases. The aberrations were computed up to the pupil at the FSM. In the no-polarizer simulations, during the wavefront control loop to dig the dark hole around the star, the wavefront could be sensed in two ways: (1) by directly computing the image plane electric field using the mean of the four polarization aberration maps, which is the fastest method; or (2) using probing with the DM to derive the mean wavefront by propagating each polarization aberration separately with the probe to create four image plane intensity fields that are added together. In either case, the derived DM solution was evaluated by separately computing each polarization wavefront and adding the resulting intensity fields together. For a single output polarization, the two appropriate ±45° input polarizations are used instead of all four. We have tested both methods (direct and probing), and they result in very similar dark hole levels.

We note that a prior study[14] by another group stated that WFIRST would be limited to ~10^{-7} contrast due to polarization aberrations. However, they used an incorrect combination of coatings, layout, and calculations, and their results did not agree with our predictions that showed orders of magnitude better contrast. ***They have since conducted new calculations[15] with more appropriate coatings and now agree with our results presented previously and here - polarization is generally a less than 10^{-8} contrast effect.***

3.5 Dynamic changes and the JPL integrated modeling pipeline

The WFIRST system will not be static. A multitude of time-dependent variations will impact the stability of the speckles in the dark hole, and thus the exoplanet detection limit. For a thorough estimate of the coronagraphic performance, all of these must be included in the full system simulations. Previously we defined an observing scenario (OS5) to model:

1. Slew to a bright star (β UMa; V=2.4), which is used for creating the dark hole and to provide reference images for post-processing, and image it for 8.3 hours.
2. Slew the exoplanet host star (47 UMa; V = 5.0) and take images over 11.1 hours.
3. Roll the telescope 26° about 47 UMa and take more images at a different sky orientation for 11.1 hours.

Using realistic spacecraft-to-Sun orientations, a time series of resulting wavefront variations were computed using a combination of thermal, structural, and optical ray trace modeling. These were then used in our representation of the optical system to first simulate the measurement of the wavefront changes with the LOWFS and determine DM settings to correct them, and then to compute the final image plane speckle fields incorporating those DM settings (together with the DM dark hole solution). These results have been used to test planet extraction and post-processing algorithms[16-19], such as reference differential imaging (subtracting the image of a reference star) and angular differential imaging (subtracting the image of the same star observed at a different orientation). These fields are publically available at

https://wfirst.ipac.caltech.edu/sims/Coronagraph_Image_sim.html. In general, the simulations show that the LOWFS/C can measure and correct the slow, low order wavefront changes caused by thermal effects to the stability levels required for coronagraphic imaging.

These simulations, however, are incomplete as there are other time-dependent system properties that have yet to be included. Besides thermal effects, the optical system will be perturbed by vibrations from the reaction wheels used to orient the telescope and keep it on target. These induce pointing errors that are largely corrected by the FSM, but variations in the residual pointing jitter can lead to speckle instability. There is also wavefront jitter, rapid variations in aberrations such as focus, coma, and astigmatism caused by vibrations of the optics, cannot be sensed or corrected. We will discuss the impact of pupil alignment instabilities later.

Another source of instability that was only recently identified is DM actuator drift. In the Xinetics DMs used in the High Contrast Imaging Testbed at JPL, which are the same as planned for use in the WFIRST coronagraph, some actuators slowly drift from their commanded positions, notably altering the speckle pattern. At the current time we must assume that these drifts will also be in the flight DMs, though work is underway to try to identify and fix their cause. It will be difficult to measure and correct them on-orbit, given the limited amount of photons that will reach the focal plane. It may be possible during post-processing deal with these using an optimized observation sequence that cycles more frequently between reference and science targets. For future modeling runs we have defined a new strategy (OS6):

1. Observe β UMa for a couple of hours.
2. Observe 47 UMa for a couple of hours.
3. Roll the telescope 26° about 47 UMa and image it again for a couple hours.
4. Repeat steps 1 - 3 multiple times.

By cycling more rapidly between the reference star and science target, as well as repeating the rolls, the speckle pattern will be better matched within an iteration. The final image would then be the summation of the post-processed images from these.

Up to now, the thermal/structural/ray modeling has been done at Goddard Space Flight Center. As they are also responsible for modeling the Wide Field Imager on WFIRST as well, they are unable to provide results rapidly. To rectify this situation and to verify Goddard's results, JPL has developed an integrated modeling pipeline[20] to replicate the modeling stages, combining Thermal Desktop, NASTRAN, SigFit, and Code V software under an automated execution management program. An observing sequence is provided, along with the thermal and finite element models of the telescope and coronagraph, and the output is a corresponding time series of wavefronts. These are fed to a separate optical modeling pipeline that includes LOWFS and produces the final speckle fields. A detector model[21] will provide realistic noise, including CTE and cosmic rays. The speckles can also be fed to an IFS simulator[22] developed at Goddard. At the time of writing, the pipeline is being used to diagnose differences in the Goddard and JPL models prior to running the new Phase A prescriptions and OS6.

4. INVESTIGATIONS

4.1 HLC performance

The HLC has been subject to nearly continuous reoptimization over the past years to improve contrast, reduce sensitivity to low order aberrations (for polarization and pointing errors), increase effective throughput, and reduce DM stroke. In general, one can optimize one or two of these factors, but at the expense of the others. The current flight design (HLC 20161228) provides better aberration tolerance (notably to tip/tilt) and contrast than the prior one (HLC 20140623) (see Figure 6). Core throughput is used in the WFIRST project as a comparative measure of the effective throughput of a coronagraph; the new and old designs have fairly similar core throughputs (0.045 new versus 0.042 old; WFIRST without a coronagraph is 0.34).

This HLC design was specifically developed for a 10% bandpass centered at 550 nm, but evaluation of the sensitivity to wavelength-dependent polarization in other bands was needed. Because it takes so long to optimize the HLC for other wavelengths, it was assumed that the sensitivity of this design to ***waves*** (rather than meters, for instance) of low order aberrations would be the same as for designs optimized for those bandpasses. For example, for a 10% bandpass from 447

- 494 nm we would use the polarization aberrations computed at those wavelengths *in waves* over the 523 - 578 nm bandpass at which the design was defined.

Including surface errors, EFC was used to produce the dark holes shown in Figure 7 and plotted in Figure 8. These clearly demonstrate how the polarization-induced aberrations can limit contrast, especially at shorter wavelengths. The prior coronagraph design specified 10% width imaging bandpasses centered at 470 and 550 nm to provide color information of planets and disks, but the bright speckles in the 470 nm bandpass due to polarization limited performance. By shifting the centers of these two bandpasses longword, the color can still be derived but the contrast improves by a factor of 2 - 3. The latest bandpasses are 10% centered at λ_c = 506 and 575 nm (the simulations shown here for slightly earlier version of the specification which had λ_c = 505 for the shorter wavelength bandpass). These new filter definitions have been accepted as the new imaging channel bandpasses.

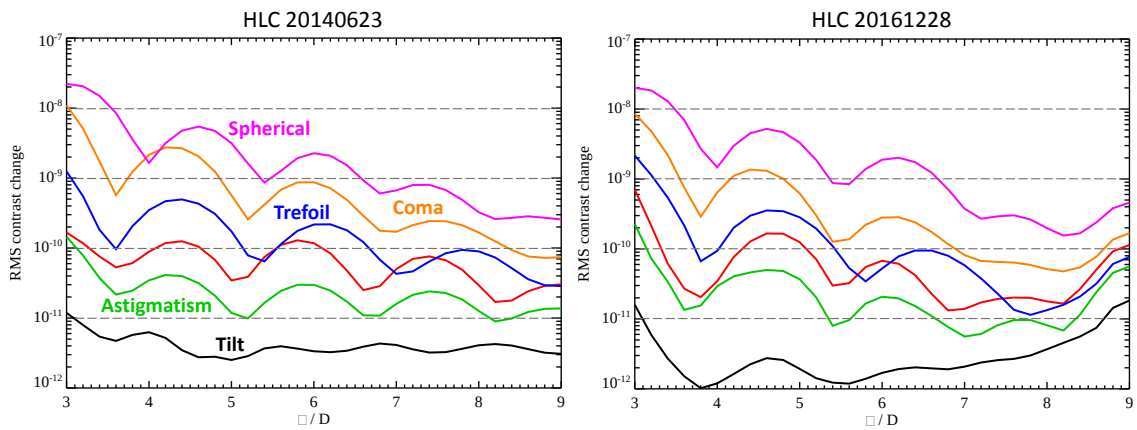

Figure 6. Plots of the contrast sensitivity to particular low order aberrations of the prior HLC flight design (left, HLC 20140623) and the new one (right, HLC 20161228) over a 10% bandpass centered at 550 nm. Each curve represents the difference between the unaberrated field and one perturbed by 100 picometers RMS of the specified wavefront aberration measured as the azimuthal mean versus field radius. Such plots are useful to gauge improvements in coronagraph designs and estimate tolerances to given aberrations.

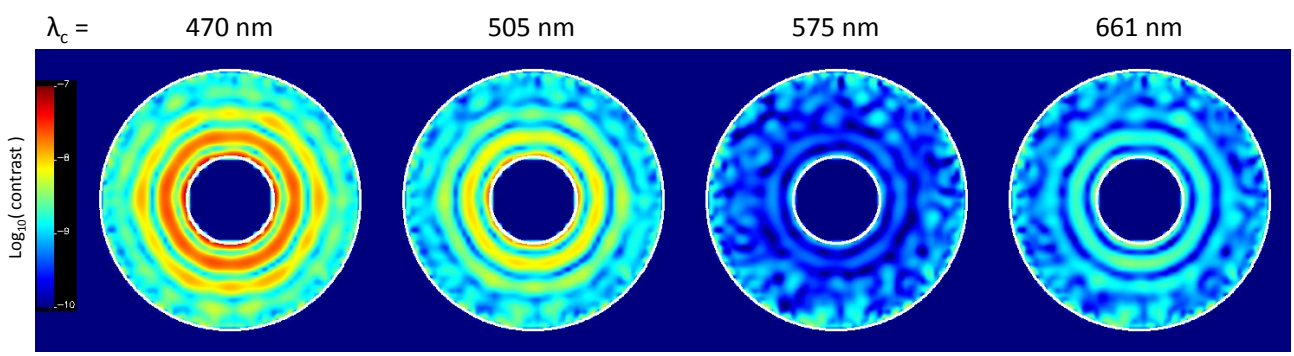

Figure 7. Post-wavefront-control contrast maps for HLC 20161228 in different 10% bandpasses without a polarizer. Surface errors and wavelength-dependent, polarization-induced aberrations were included. Note that due to the wavelength-dependent properties of the coatings, the polarization aberrations are lowest near 600 nm. The circles correspond to r = 3 and 9 λ_c/D.

Besides polarization, another limiting factor for performance is pointing jitter. The pointing stability of the spacecraft itself is ~12 mas RMS. The coronagraph's LOWFS and FSM will reduce the pointing error to 0.4 - 1.6 mas RMS. The impact of these jitter levels are shown in Figure 9. In terms of raw background contrast, jitter is not an issue for 0.8 mas

RMS or below. For faint targets (e.g., young stellar objects), the LOWFS will not be able to track the pointing changes, and the jitter will approach that of the spacecraft's. In that case, the contrast will be significantly degraded (Figure 10). One aspect that has yet to be evaluated via modeling is how variations in jitter over time impact the speckle stability and thus limit the effectiveness of post-processing methods such as PSF subtraction.

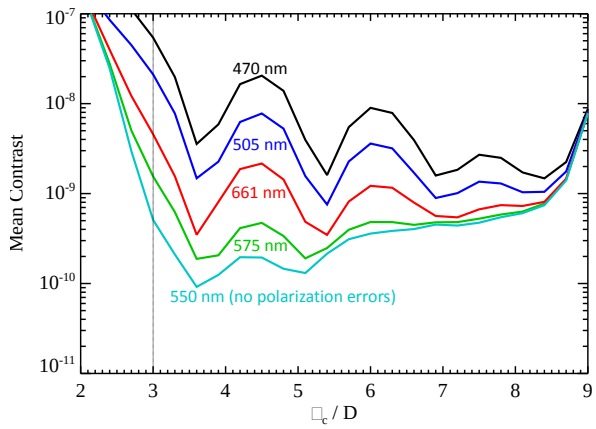

Figure 8. Mean contrast versus radius for HLC 20161228 in different 10% bandpasses without a polarizer. Wavelength-dependent polarization-induced aberrations (unless otherwise noted) and surface errors are included. The 3 λ/D inner working angle limit is shown.

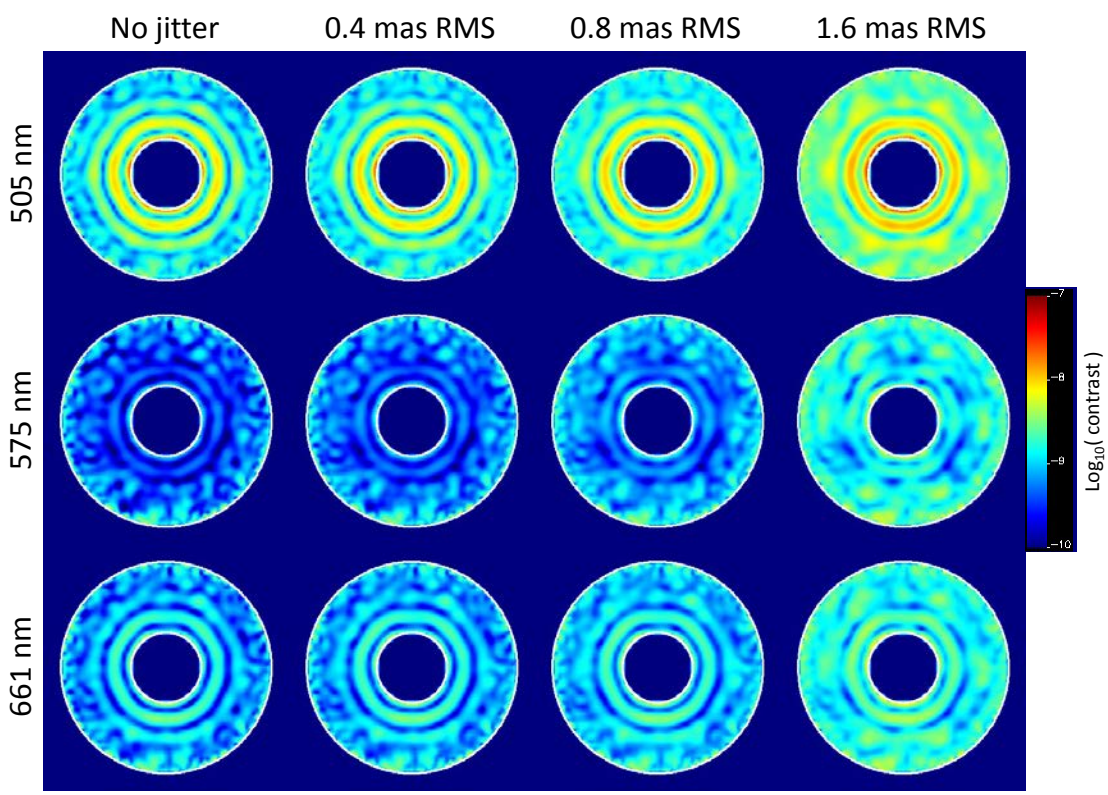

Figure 9. Contrast maps for HLC 20161228 in different 10% bandpasses for different amounts of pointing jitter and without a polarizer. Wavelength-dependent polarization-induced aberrations and surface errors are included. The circles correspond to r = 3 and 9 λ_c/D. In terms of raw contrast, jitter is not a significant effect below 0.8 mas RMS, but variations in jitter over time can be.

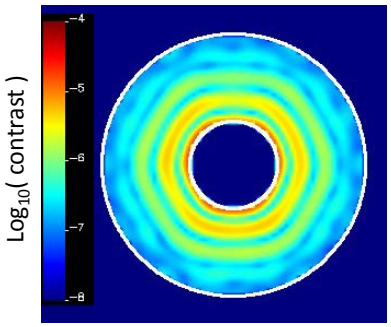

Figure 10. Contrast map for HLC 20161228 in a 10% bandpass centered at 550 nm with 12 mas RMS pointing jitter. The circles correspond to $r = 3$ and 9 λ_c/D. This is the raw background contrast that may be expected when observing faint (V > 9) targets and the LOWFS is unable to sense rapid pointing variations. This mode would be used for observing lower contrast features such as circumstellar disks.

4.2 SPC performance

Like the HLC, the SPC is continually being optimized for throughput, contrast, and aberration sensitivity. A recent redesign (SPC 20170714) was necessary to mask large wavefront errors along the overturned edge of the primary mirror (Figure 11). Although this required effectively reducing the pupil mask diameter, the performance improved. The core throughput of this new flight SPC is higher (4.3%) compared to the prior SPC 20140902 design (3.7%), while offering improved tolerance to pointing (tip/tilt) errors (Figure 12).

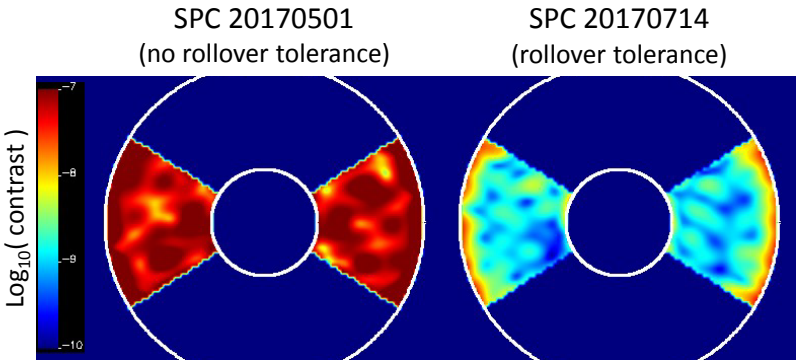

Figure 11. Post-wavefront-control Contrast maps in an aberrated system using the measured WFIRST primary and secondary error maps, including the primary's edge rollover. The SPC design on the left (SPC 20170501) does not mask the rollover region and wavefront control is unable to compensate for the error. The new design on the right (SPC 20170714) does, and achieves its nominal contrast. The circles correspond to $r = 3$ and 9 λ_c/D.

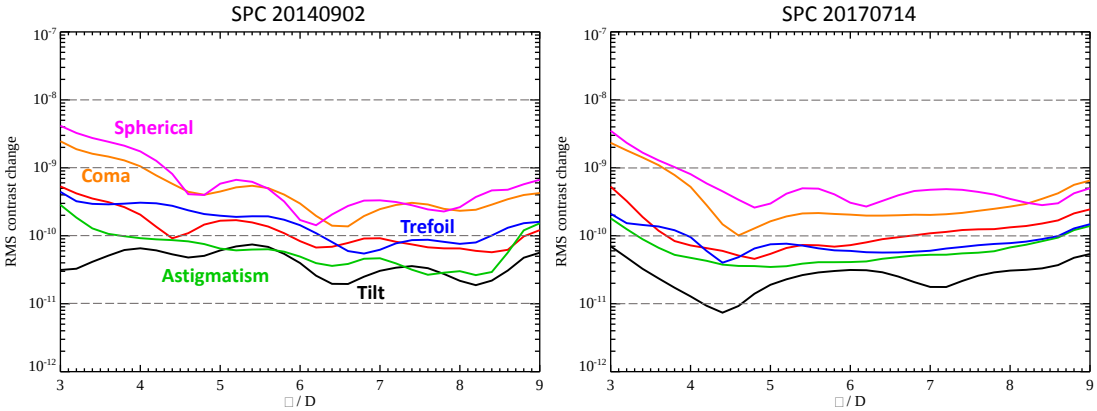

Figure 12. Plots of the contrast sensitivity to particular low order aberrations of the prior SPC flight design (left, SPC 20140902) and the new one (right, SPC 20170714) over an 18% bandpass centered at 660 nm. Each curve represents the difference between the unaberrated field and one perturbed by 100 picometers RMS of the specified wavefront aberration measured as the azimuthal mean versus field radius.

© Copyright 2017 California Institute of Technology. All rights reserved.

Unlike the HLC, whose FPM and DM settings are specific to each bandpass, the same SPC masks can work in any 18% bandpass. The SPC contrast performance is fairly constant over wavelength, jitter, and polarization, as shown in Figure 13. Note that the SPC contrast is not as good for the HLC.

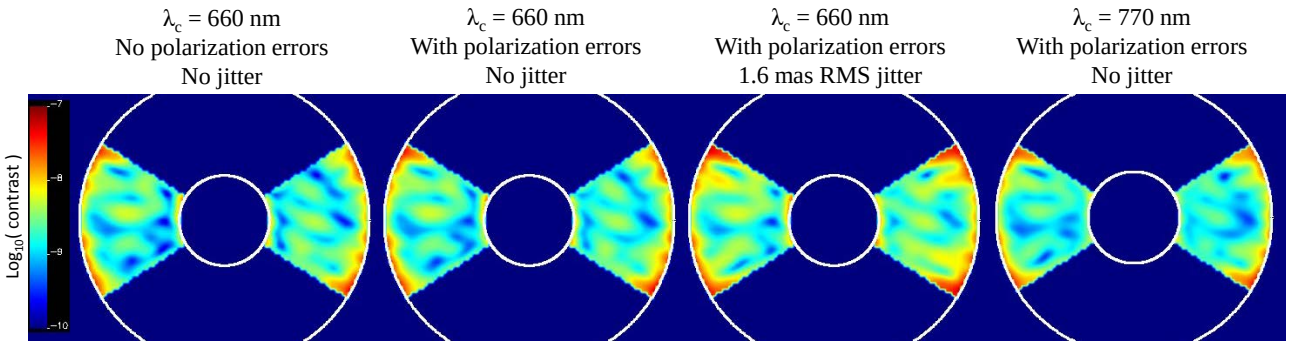

Figure 13. Contrast maps for SPC 20170714 in 18% bandpasses (λ_c = 660 and 770 nm) in aberrated systems after wavefront control. The SPC is tolerant of both polarization and pointing jitter errors. The circles correspond to r = 3 and 9 λ_c/D.

4.3 Number of filters and spatial sampling required for wavefront sensing & control

When simulating wavefront sensing and control using numerical models, it is common to generate purely monochromatic image plane electric fields at a limited number of wavelengths spanning the bandpass and use those directly as inputs to the dark hole optimizer. This corresponds to an unrealizable exact knowledge of the system. In reality, to create a broadband dark hole the wavefront must be derived from DM-probed images taken in multiple finite-width filters that sample the science bandpass. The number of these narrower sensing filters will be limited by the available slots in a filter wheel. However, if the sensing bandpasses are few and thus too wide, the wavelength-dependent variations in the wavefront become washed out and result in a poor broadband solution[23].

The field can also vary significantly over the area of a pixel if spatial sampling is too crude, leading to sensing errors as well. Signal-to-noise statistics favor fairly crude sampling in order to concentrate all of the light of a planet into a pixel and avoid the shot noise penalty that comes from distributing it over multiple pixels. However, post-processing algorithms favor finer sampling ($\leq 0.5\ \lambda$/D) to avoid spatial aliasing artifacts when shifting and/or rotating images.

To evaluate the sensitivity of wavefront sensing and control to the number of sub-band filters and sampling, a broad bandpass science image was represented by many (tens) of monochromatic fields computed at high resolution and binned to the final sampling to represent integration over the detector pixel's area. Each sub-band image was then computed as the mean of the field intensities within that band. DM probing was used to sense the wavefront and dig the dark hole in the aberrated (but no polarization errors) system.

For a 10% bandpass HLC the simulations (Figure 14) show that three 3.3% sub-band filters are sufficient to create a good dark hole with pixel sizes up to 0.5 λ_c/D. With careful definition of sub-band filter bandpasses and shapes, it may be possible to use only two sub-bands together with the 10% science filter to do wavefront sensing by subtracting the two 3.3% band images from the 10% one to create a virtual 3.3% filter. Note that these results are for an earlier version of the HLC (HLC 20160129), but the behavior should be the same for the current design.

The SPC will be used for the IFS, which has a spectral resolution of 50, so effectively many narrow sub-bands will be available simultaneously. However, to avoid long integration times to get good signal-to-noise in each spectral channel, it will be necessary to bin these into fewer, broader channels. Simulations with the 18% bandpass SPC show (Figure 15) that 5 channels with 3.6% widths and 0.5 λ_c/D sampling are sufficient to obtain a good dark hole. These results are for an earlier version of the SPC (SPC 20140902), but the behavior should be the same for the current design.

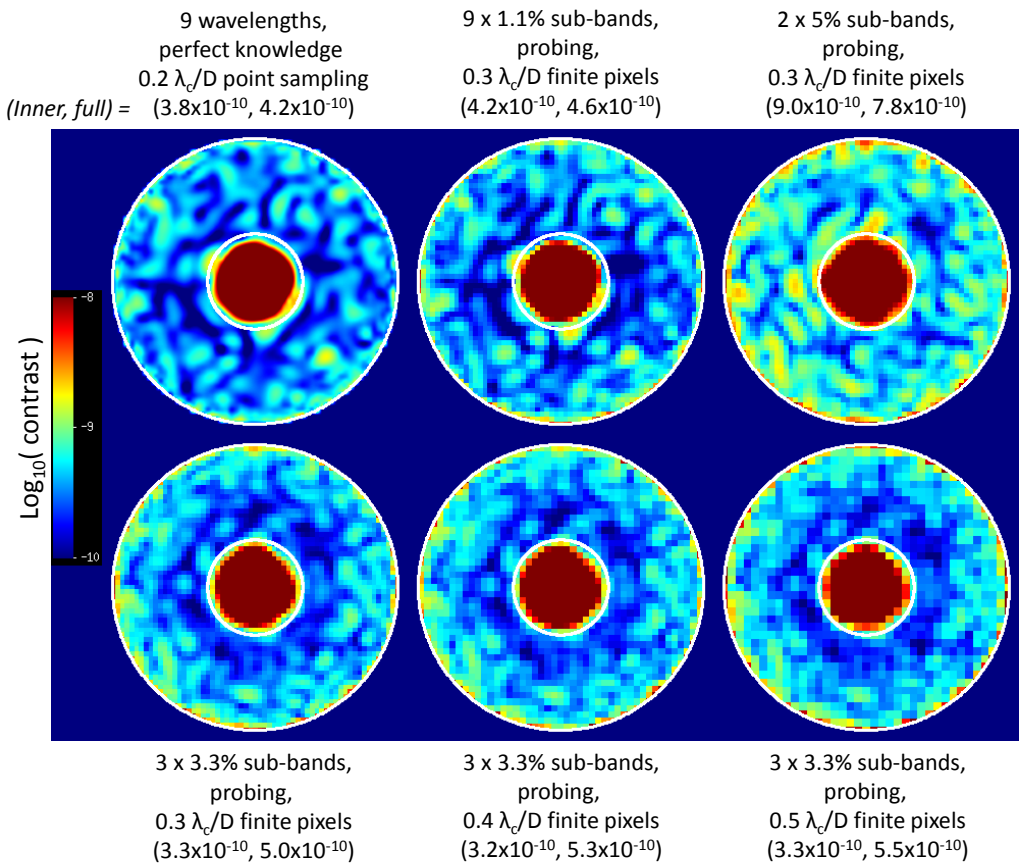

Figure 14. Post-wavefront control (EFC) contrast maps for the same aberrated HLC (10%, λ_c = 550 nm) showing the differences when using perfect knowledge of the image plane electric fields with point-sampling (upper left) and DM-probing-derived electric fields using finite-width sub-bands and finite area pixels. The contrast only degrades significantly with poor wavelength sampling using just two 5% filters (upper right). The mean contrasts between r = 3 - 4 λ_c/D and 3 - 9 λ_c/D are also indicted. The circles are r = 3 & 9 λ_c/D. Note that these are displayed on a narrower color scale that the other contrast maps in this document.

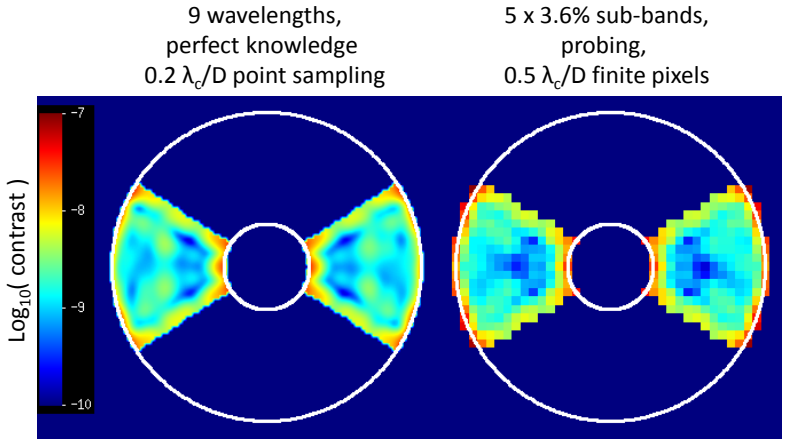

Figure 15. Like Figure 14, but for the SPC over an 18% bandpass centered at λ_c = 800 nm. The circles are r = 2.5 & 9 λ_c/D.

4.4 Pupil shear tolerance

Within the WFIRST coronagraph there are a number of places where an image of the pupil is formed and modified. The first pupil is at the FSM, followed by other pupil images at DM1, the SPC pupil mask, and the Lyot stop. Displacement of the pupil image relative to a surface (mask or DM), may occur due to a physical offset of the surface or by motion of a prior optic at a non-pupil location, offsetting the beam. A change in pupil magnification may also occur. Such variations will alter the wavefront and thus the speckles in the image plane. If the offset is static, it may likely be compensated by wavefront control, but if it is varying then it will alter the speckle stability and thus the achievable contrast after post-processing. Modeling can determine the allowable tolerances on pupil shear at each plane.

To shift the pupil by sub-pixel amounts the wavefront at that pupil is Fourier transformed to a virtual focus, an appropriate tilt term is applied, and it is then Fourier transformed back to the pupil. The pupil is then altered by the DM or mask. If the shift is a bulk displacement of the optical system after this location, then propagation continues as usual. If instead it is a displacement of only the optic/mask at that pupil, then the wavefront is then Fourier transformed, the original tilt removed, and then transformed back. At pupils prior to the focal plane mask, where the telescope obscurations are sharp, this method will introduce some numerical error because the obscurations are not bandlimited. However, we will show that this is not a problem at the levels of interest.

Figures 17 and 18 show the pupil shear sensitivities for the HLC and SPC. The "Coronagraph Shear" plots correspond to a bulk shear of the coronagraph at the FSM relative to the incoming pupil. This shifts the telescope obscuration and wavefront pattern not just on the FSM but also the DMs and Lyot stop. The "Pupil Mask Shear" and "Lyot Stop Shear" plots correspond to shifts of the masks in isolation relative to the rest of the system.

The coronagraph shear sensitivities for the aberrated HLC are shown in Figure 16 (the dark hole was generated first, then the shears applied open loop). The HLC has position-sensitive pupils at the DMs and the Lyot stop. The sensitivity at DM1 is to shifts of the wavefront aberrations relative to the DM solution intended to correct them (the same actually applies to DM2, which is a meter away from the pupil in collimated space). The 10^{-10} contrast change limit occurs with a coronagraph shear of 0.002% of the pupil diameter, which corresponds to ~0.8 μm of offset at the FSM. The Lyot stop sensitivity is about the same since so much starlight passes the stop.

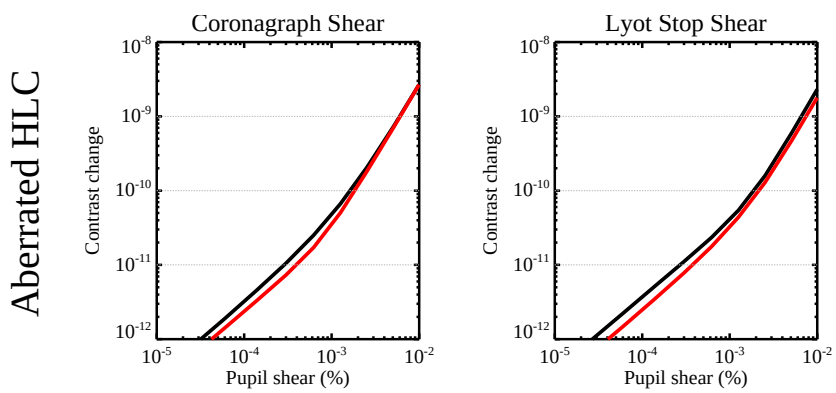

Figure 16. Mean contrast change (open loop) between r = 3.0 - 3.5 λ_c/D versus pupil shear for the aberrated (dark hole generated) HLC 20161228 in a 10% bandpass at λ_c = 575 nm. There is no unaberrated case because the HLC utilizes large stroke DM patterns as part of its design. Black lines show translation in X, red in Y. "Coronagraph Shear" is a transverse translation of the wavefront at the FSM, resulting in the wavefront being misaligned at all subsequent pupil planes, including the DM. "Lyot Stop Shear" is translation of the Lyot stop in isolation relative to the unperturbed wavefront.

In the SPC there are important pupils at the DMs, the pupil mask, and the Lyot stop. The sensitivity of the pupil mask to shear is dependent on blocking the obscurations, and the mask has a built-in tolerance of 0.2% shear due to oversizing of the mask. At the Lyot stop the primary sensitivity is to the position of the upstream pupil mask.

The SPC sensitivities are plotted in Figure 17 for both the unaberrated (hence flat DMs) and aberrated (with DM solution) cases. Because of the built-in pupil shear tolerance in the pupil mask, any contrast changes for shifts below 0.2% in the unaberrated case are due to numerical error, and these impact contrast by $<10^{-11}$ for shears below 0.005% of the pupil diameter. Note that the sensitivity to the Lyot stop shear is significantly less than for the other shears or for the HLC Lyot stop shear - after the focal plane mask, little light reaches the Lyot stop. The shear response is asymmetrical due to the orientation of the openings in the pupil mask.

The aberrated SPC case shows the impact of shearing an aberrated wavefront that was previously corrected via wavefront control. The bulk shear of the coronagraph displaces the aberrations relative to the correction pattern on the DMs, resulting in a notable contrast change (one that is significantly larger than the numerical errors featured in the unaberrated case). The 10^{-10} contrast change limit is at 0.001% coronagraph shear, corresponding to ~0.4 μm at the FSM.

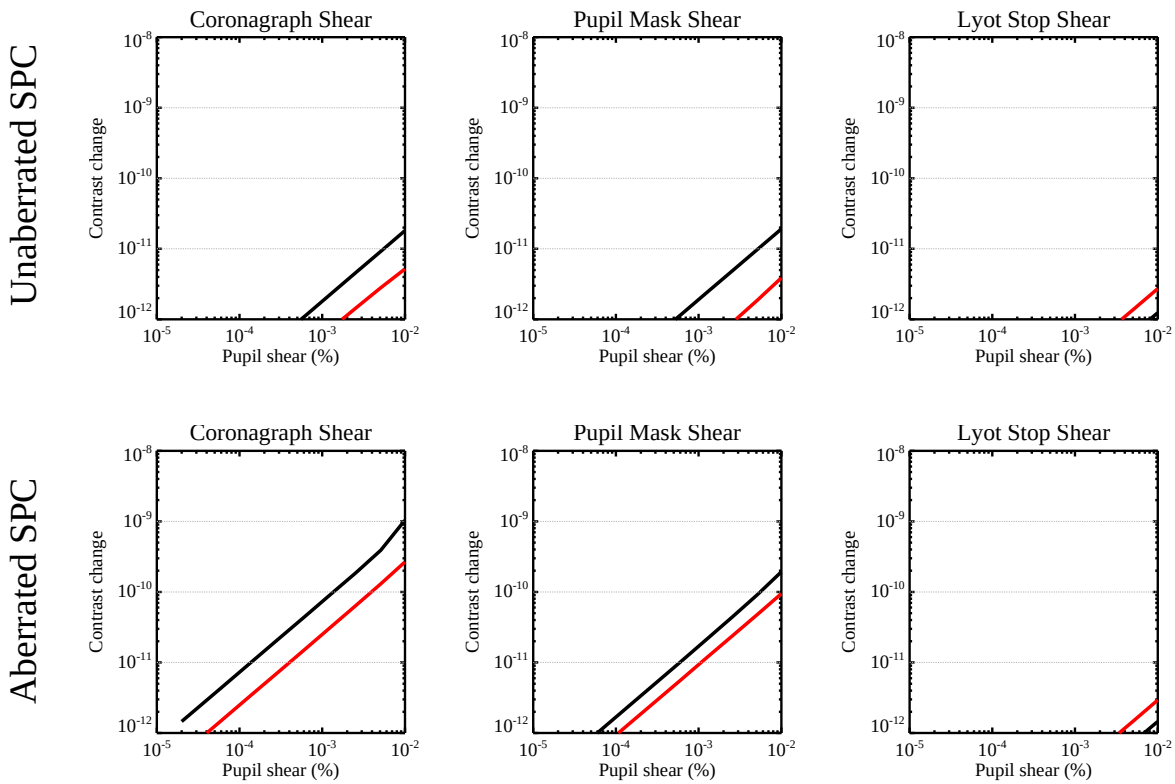

Figure 17. Contrast change (open loop) between r = 3.0 - 3.5 λ_c/D versus pupil shear for SPC 20170714 in an 18% bandpass at λ_c = 660 nm. On top is for an unaberrated system, on the bottom is for the post-wavefront-control aberrated system. Black lines show translation in X, red in Y. "Coronagraph Shear" is a transverse translation of the wavefront at the FSM, resulting in the wavefront being misaligned at all subsequent pupil planes, including the DM. "Pupil Mask Shear" and "Lyot Stop Shear" are translations of the shaped pupil mask or Lyot stop, respectively, relative to the unperturbed wavefront. The SPC has an as-designed pupil shear tolerance of 0.2%, so in the aberrated Coronagraph Shear case, any contrast change is due to numerical artifacts; in the aberrated case the contrast change is largely due to shearing of the wavefront relative to the DM solution.

4.5 The effect of knowledge errors in system models used for wavefront control

As stated before, the wavefront control matrix (Jacobian) is generated using a numerical model of the system, whether wavefront control is being simulated or exercised in reality. At some level differences between the model and actual system will limit how deep a dark hole can be achieved as the system will no longer respond as expected. Regularization is used to dampen the deformable mirror response to prevent both knowledge errors and nonlinear responses from creating a diverging solution. During recent testbed experiments where significant model/hardware mismatches existed,

new methods for dealing with such errors were developed, including using large swings in regularization with iteration. These are described elsewhere[24-29]. This work was used to validate the models during the Milestone 9 testbed experiments[30,31].

4.6 Fast computation of the wavefront control matrix (Jacobian)

In addition to the full model used for realistic Fresnel-based simulations and the compact model used for faster, mostly Fourier-based simulations, it can be quite useful to have a third model used solely to calculate the DM response matrix, also known as the control Jacobian. This so-called differential model can compute a control Jacobian in seconds to minutes (compared to hours to days with the compact model) by utilizing up to three time-saving strategies. Note that the physical equations used are still accurate for this differential model, and the only approximation is the one we already assume of the DM actuation being linearized. The first tactic is to propagate only the region of the poked influence function from each deformable mirror; by definition, the rest of the DM surface subtracts out when calculating the first-order differential term. The second gain comes from using matrix Fourier transforms (MFTs) as appropriate instead of only FFTs. For instance, for the shaped pupil Lyot coronagraph, the small region of the pupil containing the DM influence function is propagated via an MFT to the small region of the bow-tie focal plane mask where light is transmitted. With the HLC, Babinet's principle, as described in Soummer 2007 for the anodized pupil Lyot coronagraph, can instead be used to propagate efficiently from the DMs through the occulting part of the focal plane mask and then to the Lyot stop. Finally, just as with the other two types of model, lower resolution can be used at one or more planes to improve computational speed at the expense of accuracy.

Efficient calculation of the control Jacobian via the differential model can aid modeling and operations in several ways. It can enable faster modeling surveys or testbed experiments of the effects of various changes to an optical system. If the calculations are efficient enough, the control Jacobian for a space coronagraph may even be calculated on orbit and obviate the need for regular uploads of these large matrices (in the hundreds of megabytes for WFIRST) from ground stations. We are currently investigating this possibility for the WFIRST CGI.

ACKNOWLEDGEMENTS

This work was performed at the Jet Propulsion Laboratory/California Institute of Technology under contract to NASA.

REFERENCES

[1] 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).
[2] Nemati, B., Krist, J., Mennesson, B., "Sensitivity of the WFIRST coronagraph performance to key instrument parameters," Proc. SPIE 10400, 10400-7 (2017).
[3] 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).
[4] Harding, L., et al., "Electron multiplication CCD detector technology advancement for the WFIRST-AFTA coronagraph," Proc. SPIE 9605, 96050F (2015).
[5] 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).
[6] Mandell, A., et al., "Current science requirements and planned implementation for the WFIRST-CGI Integral Field Spectrograph (IFS)," Proc. SPIE 10400, 10400-9 (2017).
[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] 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).
[9] Zimmerman, N., Eldorado Riggs, A.J., Kasdin, J.N., Carlotti, A., Vanderbei, R., "Shaped pupil Lyot coronagraphs: high-contrast solutions for restricted focal planes," Journal of Astronomical Telescopes, Instruments, and Systems, 2, 011012 (2016).

[10] 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).
[11] Guyon, O., Hinz, P., Cady, E., Belikov, R., Martinache, F., " High Performance Lyot and PIAA Coronagraphy for Arbitrarily Shaped Telescope Apertures," Astrophysical J. 780, 171 (2014).
[12] Krist, J., "PROPER: an optical modeling program for IDL," Proc. SPIE 6675, 66750P (2007).
[13] Give'on, A., et al., "Broadband wavefront correction algorithm for high contrast imaging systems," Proc. SPIE 6691, 66910A (2007).
[14] Breckinridge, J.B., Chipman, R.A., "Telescope polarization and image quality: Lyot coronagraph performance," Proc. SPIE 9904, 99041C (2016).
[15] Chipman, R., Breckinridge, J., "Challenges in coronagraph optical design," Proc. SPIE 10374, 10374-2 (2017).
[16] 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).
[17] 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).
[18] 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).
[19] Mennesson, B., et al., "WFIRST-AFTA coronagraph performance: feedback from post-processing studies to overall design," Proc. SPIE 9605, 96050D (2015).
[20] 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).
[21] Nemati, B., et al., "The effect of radiation-induced traps on the WFIRST coronagraph detectors," Proc. SPIE 9915, 99150M (2016).
[22] Rizzo, M., et al, "Simulating the WFIRST coronagraph Integral Field Spectrograph," Proc. SPIE 10400, 10400-11 (2017).
[23] Zhou, H., Krist, J., Nemati, B., "Diffraction modeling of finite subband EFC probing on dark hole contrast with WFIRST-CGI shaped pupil coronagraph," Proc. SPIE 9911, 99111S (2016).
[24] Sidick, E., Seo, B.-J., Marx, D., Poberezhskiy, I., Nemati, B., "WFIRST/AFTA coronagraph contrast performance sensitivity studies: simulation versus experiment," Proc. SPIE 9912, 99126M (2016).
[25] Zhou, H., et al., "Closing the contrast gap between testbed and model prediction with WFIRST-CGI shaped pupil coronagraph," Proc. SPIE 9904, 990419 (2016).
[26] Zhou, H., et al., "Wavefront control performance modeling with WFIRST shaped pupil coronagraph testbed," Proc, SPIE 10400, 10400-5 (2017).
[27] Sidick, E., et al., "Sensitivity of WFIRST coronagraph broadband contrast performance to DM actuator errors," Proc. SPIE 10400, 10400-6 (2017).
[28] Marx, D., et al., "Electric field conjugation in the presence of model uncertainty," Proc. SPIE 10400, 10400-23 (2017).
[29] Sidick, E., Byoung-Joon, S., Kern, B., Marx, D., Poberezhskiy, I., Nemati, B., "Optimizing the regularization in broadband wavefront control algorithm for WFIRST coronagraph," Proc. SPIE 10400, 10400-74 (2017).
[30] Cady, E., et al., " Shaped pupil coronagraphy for WFIRST: high-contrast broadband testbed demonstration," Proc. SPIE 10400, 10400-14 (2017).
[31] Seo, B.-J., et al., " Hybrid Lyot coronagraph for WFIRST: high-contrast broadband testbed demonstration," Proc. SPIE 10400, 10400-15 (2017).