

Wavefront Control Performance Modeling with WFIRST Shaped Pupil Coronagraph Testbed

Hanying Zhou*, Bijan Nemati, John Krist,
Eric Cady, Brian Kern, Ilya Poberezhskiy
Jet Propulsion Laboratory/California Institute of Technology,
4800 Oak Grove Drive, Pasadena, CA 91109 USA
*hanying.zhou@jpl.nasa.gov

ABSTRACT

NASA's WFIRST mission includes a coronagraph instrument (CGI) for direct imaging of exoplanets. Significant improvement in CGI model fidelity has been made recently, alongside a testbed high contrast demonstration in a simulated dynamic environment at JPL. We present our modeling method and results of comparisons to testbed's high order wavefront correction performance for the shaped pupil coronagraph. Agreement between model prediction and testbed result at better than a factor of 2 has been consistently achieved in raw contrast (contrast floor, chromaticity, and convergence), and with that comes good agreement in contrast sensitivity to wavefront perturbations and mask lateral shear.

Keywords: High contrast imaging, coronagraph, wavefront sensing and control, model validation, exoplanets, WFIRST-CGI, shaped pupil coronagraph

1. INTRODUCTION

NASA's Wide-Field InfraRed Survey Telescope (WFIRST) mission includes a coronagraph instrument (CGI) for direct imaging of exoplanets and their spectral characterization [1-3]. To achieve and maintain high contrast for the desired science goal, CGI employs a dual wavefront sensing and control (WFSC) system: a low-order WFSC subsystem using a Zernike wavefront sensor, and an electric field conjugation (EFC) based high order subsystem. The system's ability to suppress star light and maintain a high contrast dark hole ($<1e-8$) in a simulated dynamic environment was recently demonstrated on Occulter Mask Coronagraph (OMC) testbed at JPL [4-6]. Parallel to this technology development is the effort to advance the CGI's WFSC model fidelity, as it will be crucial for error budgeting, flight system design and flight system performance verification and validation.

Coronagraph modeling has been mostly based on an assumed system with realistic but often synthetic imperfections and limitations of optic components [7~9]. It has been extensively used to study throughput, bandwidth, and sensitivity to perturbation, of various coronagraph architecture and design, and is the basis of many error budget estimations. As such, modeling has been mostly focused on determining the requirement on the most sensitive parameters. Performance modeling with as-built real system has been relatively uncommon and of limited success, due to various difficulties in testbed environment. For example, poorly understood incoherent light on the testbed has often plagued the WFSC effort in reaching the designed contrast floor on testbed. Accurate calibration can also be challenging; e.g., it is difficult to measure the shaped pupil mask wavefront error with phase retrieval methods due to limited clear area. In [10], for example, the mean *contrast change* (the quadratic coefficient of the contrast with occulter mask lateral translation) was predicted to $\sim 67\%$ of testbed result, while mean *raw contrast* differs from testbed result by a much larger factor (up to a few orders of magnitude in some cases). In a more recent study [11], Jacobian error in testbed control was shown to account for the difference between testbed result and model prediction on raw contrast floor. As remarkable as it is, the fact that it used time-consuming post-WFSC measurement (Jacobian) data makes it less practical for a model as a tool for the purpose of error budgeting or flight system design and verification.

During the past year, our effort continued in pushing the limit of coronagraph model fidelity, with an emphasis on improving the model prediction of testbed raw contrast. This emphasis on raw contrast partly reflects the fact that there is

relative success on contrast sensitivity prediction in the past but much less so on contrast floor. But as importantly, it also reflects our thinking that key performance metrics of coronagraph operation to CGI science goal include contrast floor, efficiency in reaching the contrast floor (at low flux), and robustness to contrast instability in presence of some common disturbances [12]. Hence, a good model would require accurate predictions on all these aspects. While the stability to WFE perturbations more directly relates to false positive probability in science images (of differential detection), we note that the sensitivity to various perturbation is often proportional to (the pre-disturbed) contrast itself. Thus, confidence and accuracy in predicting raw contrast should improve those of the contrast sensitivity.

In the following, we present our SPC modeling method and compare our model predictions to that of testbed results, using available pre-WFSC calibrations from WFIRST CGI Milestone 9 (MS9) and subsequent testbed runs. Significant improvement in model fidelity comes from: 1) better understanding and inclusion of testbed imperfections in the model (e.g., chromatic pupil aberrations), and 2) incorporation of testbed-like WFSC operation features such as probing, deformable mirror voltage constraints, regular updates of the control model, etc. Testbed error budget analysis and Monte Carlo modeling incorporating estimated uncertainties help to identify the most critical (or missing) testbed knowledge that impacts contrast floor and/or convergence. As a result, a better than a factor of 2 agreement between model prediction and testbed result has been consistently achieved in raw contrast floor, chromaticity, and convergence. In terms of sensitivity, which we managed to do a few important ones (even though not the original focus of the study), the agreement is even better in general, validating the notion that the better agreement in raw contrast the better in contrast sensitivity also. Testbed contrast performance enhancement from the modeling feedback further confirms our model's good fidelity.

A few notes before we go to details. Since an EFC-based WFSC does not respond to unmodulated (incoherent) light, all of our comparison to testbed results are limited to modulated (coherent) part of testbed result. Also, even though both nulling and contrast maintenance under dynamic conditional were successfully demonstrated, the focus of this study is limited to the static high-order WFSC part. No dynamic low-order WFSC contrast stability performance modeling beyond Zernike sensitivity study is in our scope. Lastly, throughout the paper, we mostly use "contrast" in its strict definition, which is the speckle intensity normalized by a PSF peak map with coronagraphic mask (the peak of the PSF changes with radial offset due to occulter transmission). Occasionally, we use "normalized intensity" in place of contrast, which is the speckle intensity normalized by the PSF peak at origin without including coronagraphic mask throughput. Contrast is generally worse than normalized intensity near the inner working angle by this definition.

2. TESTBED SPC MODEL SETUP AND MODELING APPROACH

2.1 OMC testbed layout and SPC coronagraph mask

The test data used in this study were collected from the OMC testbed which is commissioned for WFIRST-CGI technology development at JPL. Figure 1 shows its layout schematic. The OMC combines both Hybrid Lyot coronagraph (HLC) and Shaped Pupil coronagraph (SPC) in one optics layout, and comes with an optical telescope assembly (OTA) simulator for simulating dynamic wavefront disturbances. It also includes a Zernike wavefront sensor based LOWFS/C, which uses rejected star light (from occulter mask) to sense disturbances and control through fast-steering mirror (FSM) and one of the two deformable mirrors (DMs) to correct line-of-sight and low-order WFE respectively. Although the OTA simulator was later replaced with an OAP to reduce large incoherent light that plagued WFSC during much of the early times of OMC operation, the testbed's ability to inject and correct focus, jitter or line-of-sight retained. More detailed description of the testbed and testbed operation can be found in [4-6].

The SPC mask installed on OMC testbed for this study, shown in Figure 2, is for the characterization mode for eventual use with the integral field spectrograph (IFS). It is similar to previous designs [13].

2.2 SPC Baseline Model

Rather than a full propagation diffraction model that is often used for coronagraph modeling at JPL, a compact SPC model is used for this purpose, driven largely by the computation need as well as the fact that it is the choice used on OMC testbed. As shown in Fig.3, the compact model starts from DM1 (a pupil plane), Fresnel propagates from DM1 to DM2

(and back to DM1, equivalent to next pupil plane), and uses Fourier transforms for all of rest of propagations between pupil plane to focal plane. Except for the shaped pupil mask WFE, which is put at shaped pupil plane, all the rest of system aberrations are compressed into a single entrance pupil aberrations. All aberrations are phase retrieval (PR) measured as they exist in the system. Typical testbed pupil WFE (before DM flattening) is on the order of $\sim 100\text{nm}$ RMS, pupil amplitude $\sim 6.5\%$ RMS (shown in Figure 3 for *Config.1*), and SPM WFE $\sim 40\text{nm}$ RMS (mostly astigmatism), all of them larger than what expected in a flight system. Besides aberrations, DM registrations, and masks alignments, and DM gains are the only other calibrations that go into the model.

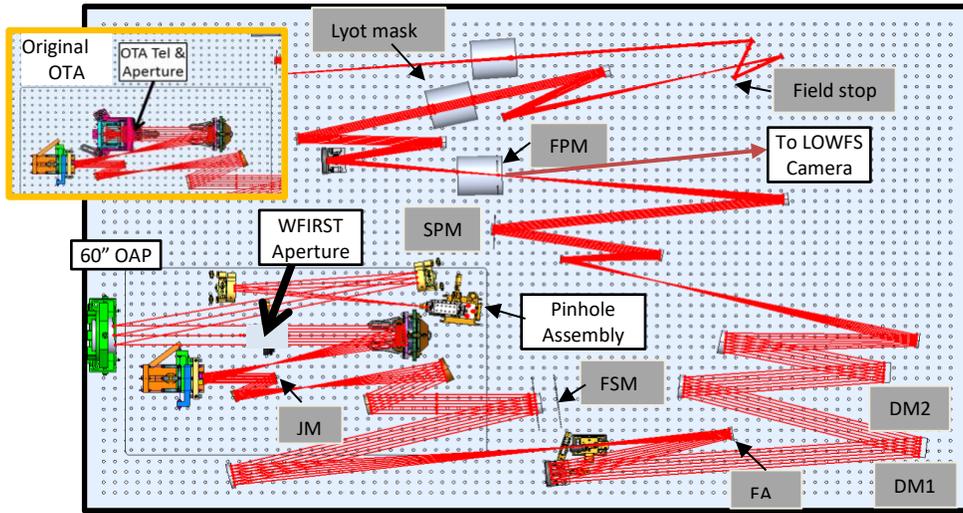


Figure 1. OMC testbed layout. The original OTA (in the center of yellow rectangular insert) was used in *Config.1* but was replaced by 60" OAP in *Config.2*. FPM: focal plane mirror; SPM: shaped pupil mask; FSM: fast steering mirror; FA: focus adjustor; JM: jitter mirror; OAP: off-axis parabola; DM: deformable mirror

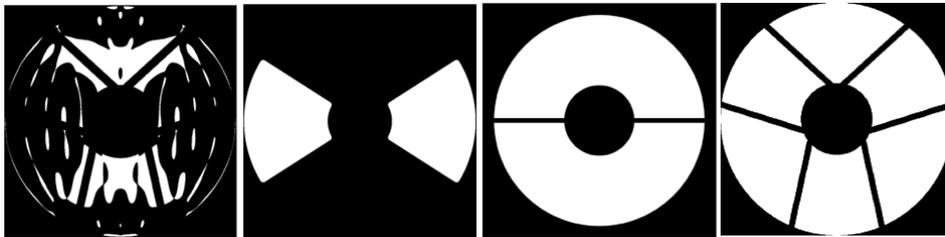


Figure 2: Shaped pupil Lyot coronagraph masks used in this study. From left: shaped pupil mask, occulter mask, Lyot stop mask, and testbed entrance pupil

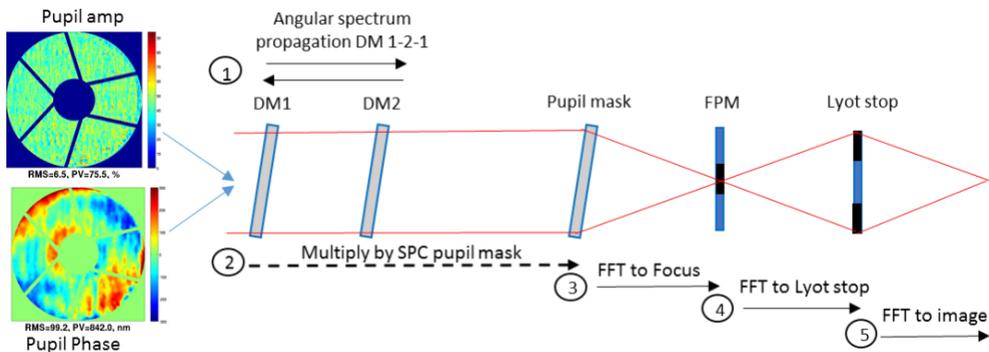


Figure 3: Compact CGI model is used in this study

The WFSC method used is the standard image-plane based EFC algorithm [14, 15] on both testbed model and prediction model. In this approach, the field is *sensed* through a *model-based* pair-wise DM probing measurement:

$$I_j^{+/-} = |E \pm \Delta P_j|^2 + I_{inc} = |E|^2 + |\Delta P_j|^2 \pm 2\Re\{E^* \Delta P_j\} + I_{inc} \quad (1)$$

$$\Delta I_j \equiv I_j^+ - I_j^- = 4\Re\{E \Delta P_j\}; \quad (2)$$

where ΔP_j are image plane (unaberrated) fields due to probe application and are obtained by propagating the coronagraph model with the probing field. Combined with measured ΔI_j , the aberrated field E can then be estimated.

During *control*, actuator adjustment at each iteration is found to be:

$$\Delta h = \left[J^T J + 10^\beta \cdot \max(\text{diag}(J^T J)) I \right]^{-1} J^T E \quad (3)$$

where J is the control matrix (aka Jacobian), which is a collection of (linear approximation of) image plane electric field response to each actuator's unit strength poke (pupil plane) through model. J is calculated by propagation through the coronagraph model. β is a regularization parameter to damp the imperfect control due to nonlinearity approximation, imperfect sensing, as well as imperfect system calibration.

Note that both sensing and control involve the model with known calibration inputs.

While there is simplification in terms of using compact model, in other aspects, more testbed WFSC operation features and constraints are included in our SPC model than in a typical coronagraph modeling practice. These includes:

1. **Initial low order pupil WFE flattening** (aka ‘DM flattening’). On testbed, it is done through a few iterations of phase retrieval measurement and application of DM voltage. To simulate this in the model, we use measured collective WFE input and then fit it to the DM surface and smooth the fit for a few iterations (This part is normally implemented in typical coronagraph modeling if not starting with flattened pupil aberration input).
2. **Electric field (E-field) sensing / estimation**. In the past coronagraph modeling often skipped the sensing part, practically assuming perfect knowledge of image plane E-field that is used to compute the DM voltage adjustment for next iteration. The testbed, of course, senses the E-field imperfectly through a pair-wise probing procedure. In our model, we also implemented the pair-wised probing with finite subband probing to get the estimated E-field at each iteration.
3. **Regular update of control matrix** (aka ‘Jacobian’, or G matrix). Both testbed and model uses image plane based WFSC algorithm, with control matrix calculated based on coronagraph model. Due to imperfect or sometime missing knowledge (as well as the large aberration /misalignment themselves) of the as-built testbed optics, testbed found regular control matrix update a necessity for effective nulling. In our model, Jacobian is also recalculated on a regular basis during the course of an EFC run.
4. **DM voltage constraints and neighborhood rule**. Again, typically coronagraph modeling ignores the fact that DM has limited stroke range (0~100 volt in our case, with bias at 50V for DM1, and 30V for DM2), and that there is safety rule put in place that says any neighboring DM's voltage difference cannot be greater than 30V. This is to prevent potential DM facesheet damage. These DM constraints are similarly implemented in our model.
5. **Regularization strategy**. Regularization strategy on testbed evolved during the course of MS9 effort, from early ‘standard’ regularization (fixed at a nominal β , e.g., $\beta = -3$ in Eq. (3), with occasional manual adjustment now and then), to a new strategy that alternates between nominal and aggressive regularizations over iterations. The aggressive beta is found as ‘optimal’ regularization in nonlinearity sense. This latter strategy comes from in-depth SVD analysis of Jacobian and E-field [6, 16]. Essentially, it was found that with small nominal $|\beta|$ value, one corrects few ‘easy’ SVD modes with small DM strokes (cheap), while an aggressive $|\beta|$ values one is able to correct more ‘hard’ SVD modes with large strokes (more expensive) but often causes worsening contrast. Better contrast is achieved when interleaving iterations at nominal β with aggressive β than by using a fixed strategy at either regularization alone.

Other parameters / strategies the testbed sometimes used during WFSC include dark hole weighting (e.g., more weight near IWA), dark hole control region (e.g., larger bow-tie region than the scoring/design region), control bands (more and/or incremental application of subbands), control gains (to further damp the imperfect control). These are not used (or not consistently used) for the SPC nulling runs used in this study, and we opted to skip these operation features in modeling. Additionally, detector noise was ignored in our model as testbed was operated with bright pseudo-stars in a high photon flux condition.

2.3 SPC “Testbed” model and knowledge uncertainty

The above baseline model uses known calibration input to give a baseline prediction of testbed performance. As calibrations are always less than perfect, a Monte Carlo (MC) technique was used to enhance the robustness of the prediction. Note that a ‘model’ is used on testbed in both sensing part (for probe induced unaberrated E-field calculation, ΔP_i in Eq. (2), during estimation of aberrated E field) and control part (for Jacobian calculation J in Eq. (3)). In MC runs, we use known imperfection parameters only in these parts; calibration error is then added for all contrast evaluation and calculation of “measured” probing intensity I during E field sensing. Therefore, there are effectively two models used together: one for control with our best understanding of the testbed, and a second with knowledge error (additional unmeasured aberrations, misalignments, or miscalibrations) to represent a real testbed where we don’t have complete system knowledge for control. We sometime refer the model that has added knowledge error as “Testbed” model.

The type and amount of knowledge uncertainties that were used in our model are listed in Table 1. They can be broadly grouped in two categories: 1) alignment error such as DMs, SPM, occulter, Lyot stop, pseudo star source, etc., and 2) aberration errors such as pupil phase and amplitude (achromatic in early models as well chromatic in later models), and SPM reflective surface phase. Estimations of uncertainty are not particularly rigorous for certain items but were estimated as best we could at the time. Below are some short notes on a few important calibration errors (more details can be found in [6, 16]).

1. **DM registration and gain.** These calibrations and error estimations are based on phase retrieval (PR) measurements. First DM is poked (often in group of actuators), typically a few volts from a flat WFE (or a decent dark hole) state. Each actuator is then 5 DoF matched, using a measured influence function, in actuator pitch, orientation, and offset. This data is also used to estimate individual actuator gains. The procedure produces the (mis)registration information of DMs relative to pupil, their gains, as well as uncertainty of them (the standard deviation of the collection of individual actuator’s registrations and gains). The mean and standard deviation gives the upper bound of the uncertainty. In many of the calibration sets taken, DM gain errors are typically found to be between 5 ~ 8%. A conservative 10% is used as upper bound for DM gain error.
2. **Pupil phase and amplitude.** These are again obtained through PR measurement. Difference of measured WFEs between two different times are evaluated and served as uncertainty. For example, difference between 1 hour apart WFEs is typically ~1nm rms, a 5 days apart have ~20nm rms, and 43nm for a month apart. The mean and standard deviation of these gives the upper bound of the uncertainty. Typically testbed measures WFE at the beginning of EFC run, which may on and off last a day or two. In our modeling, we put uncertainty error at 5nm rms, mostly to account for the testbed drift.
3. **SPM WFE.** This is usually obtained as the difference between measurement with SPM in and SPM out. Shaped pupil mask aberration is often difficult to measure accurately due to its large opening area. During MS9, we have found inaccuracy in PR processing of SPM from time to time. We put uncertainty error at 5nm rms, mostly to account for the PR accuracy.
4. **Chromatic pupil WFE and amplitude.** In early model, no measurement was taken for chromatic phase and amplitude across the entrance pupil. The phase error was estimated based on polarization modeling of pinhole in source head and of OMC testbed mirrors. It was concluded that small, long tunnel-like commercial pin holes used have significantly more wavelength dependence, likely 0.03rad or up, due to dispersion of the tunnel’s waveguide modes over the bandwidth (10%). Additionally two OTA mirrors (PM and SM) have non-negligible polarization effect with about 3mrad rms wavelength dependent WFE (and a weaker chromatic amplitude as well). We put a combined chromatic pupil WFE of 0.08rad rms and pupil amplitude of 2% as upper bound for *Config.1* (see Sect.

3.1). In *Config.2* where both original OTA and source head pinhole were replaced, the amounts were reduced by a factor of 100.

Calibration errors are assumed to be simple Gaussian distribution. Each error instances are generated as a truncated Gaussian, with an upper limit as listed in Table 1 and truncation done at $\pm 2\sigma$. We generated 20 sets of error instances, but in reality, due to the time-consuming execution of EFC iterations, only a little over 10+ sets were run.

	Parameter Name	Estimated Error Upper Limit	Note
DM gain & alignment	DM (lateral) decenter	75 um	
	DM rotation	[0 0 0.3] deg	[tip tilt clocking]
	DM gain	10 pct	
SPM alignment & manufacture error	SPM (lateral) decenter	32 um	
	SPM rotation	[0 0 0.25] deg	[tip tilt clocking]
	SPM undercut	1 um, all sides	Fixed in all MC instances, <i>Config.2</i> only
Occulter alignment	Occulter (lateral) decenter	1 um	<i>Config.2</i> only
	Occulter defocus	100 um	
	Occulter rotation	[0 0 0.5] deg	[tip tilt clocking]
Lyot Stop alignment	Lyot stop lateral decenter		
	Lyot stop rotation	[0 0 0.5] deg	[tip tilt clocking]
Source alignment	Source lateral shift	0.125 λ/D	<i>Config.2</i> only
Pupil amp error (Achromatic)	pupil amp zernike term	2:3	slope-like amplitude droop
	Pupil amp zernike rms	2 pct	
	Pupil amp high order		From difference of repeat flat DM pairs mears
	Pupil amp PSD rms	2 pct	
Pupil WFE (Achromatic)	Pupil WFE Zernike term	5:15	
	Pupil WFE Zernike rms	0.05rad @550nm	drift/change since last measured
	Pupil WFE PSD		From difference of repeat flat DM pairs mears
	Pupil WFE PSD rms	1.5 nm	
SPM WFE	SPM zernike term	15	
	SPM zernike rms	0.05rad @550nm	less accurate PR due to thin edges of mask
	SPM PSD params		= [SPdiam/pupil_diam spm_psd_rms 4 3]
	SPM PSD rms	1.5 nm	
Chromatic pupil WFE & amp	Pol amp	2 pct	+/- slope-like Z2Z3 at end bands, λ dependent
	Pol WFE	0.08rad @550nm	+/- Ast (Z5Z6) at end bands, λ dependent

Table 1: List of knowledge error terms and uncertainty bounds

3. CLOSE LOOP RAW CONTRAST

Throughout the course of CGI MS9 course, many real-world problems on OMC testbed, particularly the large unexplained incoherent (unmodulated) light among others, took up substantial testbed time to diagnose and minimize it, making frequent hardware changes along the way. This posed an additional challenge to our model validation effort, as only few rounds of nulling runs performed along with needed complete testbed calibrations were taken. On the other hand, it also

provided an opportunity to test our model under different hardware /software configurations. In Table 2, we list two testbed configurations for which we had calibration data to construct our SPC model and compare against testbed result.

	Src generation	Src pinhole	Testbed EFC Ctrl strategy	Pupil WFE in model(s)	note
<i>Config.1</i>	OTA-s	COTS 1um, 10 um thick	Mostly fixed regularization	Achromatic	Incomplete model
				Chromatic, estimated	Imprecise data
<i>Config.2</i>	Long F# OAP	MDL 3um, thin	Alternating regularizations	Chromatic; measured	Better model and data

Table 2: Testbed configurations used in this study

3.1 Performance comparison of *Config.1*

We started out with testbed *Config.1*, which has the original OTA simulator part of the layout. The pinhole in the source head is a laser-burnt 1um diameter, 10um-thick, and commercially made. This configuration turned out having significant polarization effect that contributes to incoherent light which the EFC is not designed to suppress (though not in flight relevant way, as the pinhole was part of the simulated star upstream of the coronagraph). However, this fact was not initially apparent to testbed and modeling teams. As a result, pupil WFE was measured at a single central wavelength only (achromatic) that fed to both testbed control model as well as in (baseline and early MC runs of) prediction model.

Figure 4 shows raw contrast performance of *Config.1* (note we used normalized intensity in this plot). Clearly, the base model prediction (solid blue line) has large discrepancy from testbed result (solid red line). So the results of early MC runs (#1~5, dotted lines in the plot, just above solid blue line), even though they have included many knowledge uncertainty as best as we can estimate, they did not come close to testbed result.

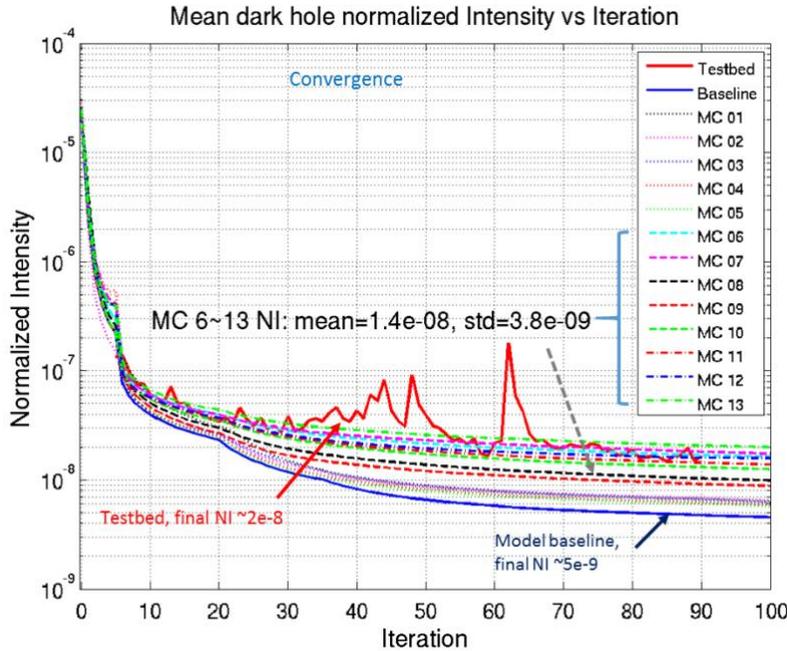


Figure 4: Normalized intensity vs iteration, testbed data (solid red line) vs model predictions, *Config.1*

After investigation indicating polarization effect being the likely source of the observed incoherent light, we added chromatic pupil aberrations due to polarization to our testbed error budget analysis (the incoherent part was ignored since

EFC won't affect it; the polarization cross term was also not included as testbed used a polarizer before imaging camera). We modeled chromatic WFE as astigmatism Z5 Z6, opposite sign at two end bands and wavelength proportional, and chromatic pupil amp as tip tilt Z2 Z3, again opposite sign at two ends of bands and wavelength proportional. The result shows they would be top contrast floor contributors among other error sources. We then put chromatic aberrations with estimated uncertainty in MC runs # 6 ~11 (dashed lines in Figure 4). The result shows a much better agreement with the testbed result. MC #12, the most matching one, has 0.05rad (4.4nm rms) in end bands chromatic WFE, within 2X the polarization modeling estimates. Both of its contrast profile in terms of field position (Figure 5a) and chromaticity (Figure 5b) matches well with testbed result. Aside from a few bumps in testbed iteration curve, its (envelope of) convergence is also remarkably comparable to the model prediction. The bumps in testbed curve were caused by manual control strategy adjustment (regularization, number of subbands) intended to speed up the convergence. Note that in both simulation and on testbed, nulling often makes up minor "bad" iterations once it gets back to right regularization/Jacobian by quickly returning to its nominal course. Figure 6 compares the predicted speckle pattern with that of the testbed result. Again, we see good agreement in spatial content statistics.

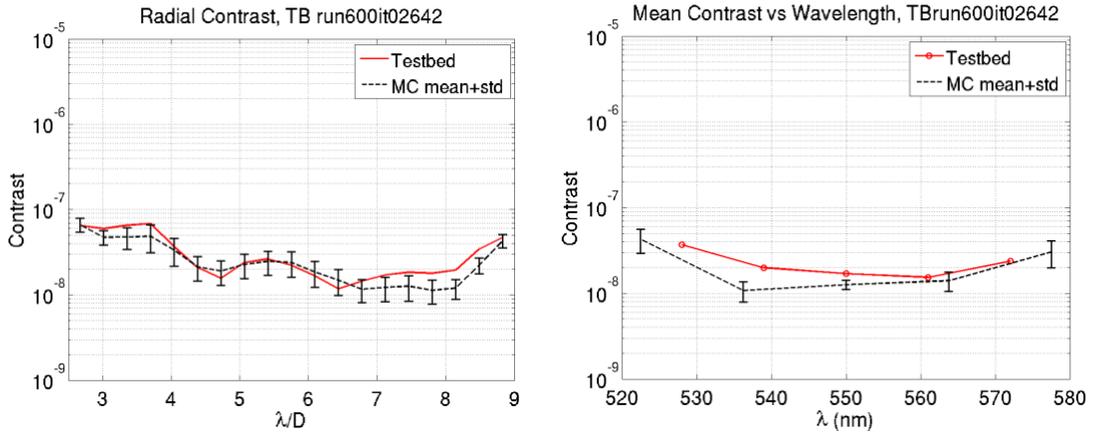


Figure 5: Radial contrast (left) and chromaticity (right) at final of EFC nulling

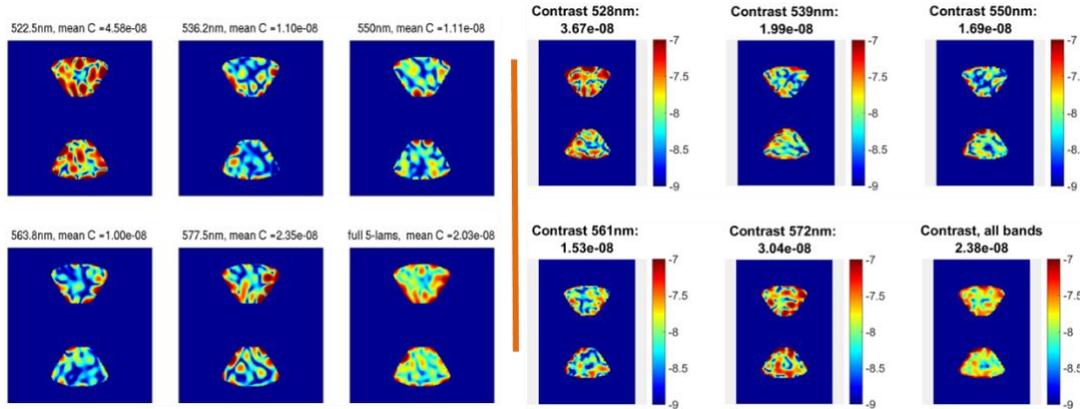


Figure 6: Speckle pattern of model prediction (left 6 subplots, per wavelength and then full bandwidth) and testbed run result (right 6 subplots, per wavelength and then full bandwidth)

3.2 Performance comparison of Config.2

In a second configuration, the OTA was removed, in its place is a 60" OAP of F/33. The pinhole in source head was also replaced by a thin 3um one made by the Microdevices Lab (MDL) at JPL, etched in silicon with clean edges. By this time, we have measured chromatic pupil aberrations in both testbed control model and our prediction model. Another important

change in model(s) is the use of a new regularization strategy (alternating nominal and aggressive regularization). Shown in Figure 7, the tall spikes in iteration curves are when regularization switched to more aggressive ones. Since the new regularization strategy on testbed was not automatic for the SPC data taken, we did not attempt in modeling to time the testbed control cycle precisely.

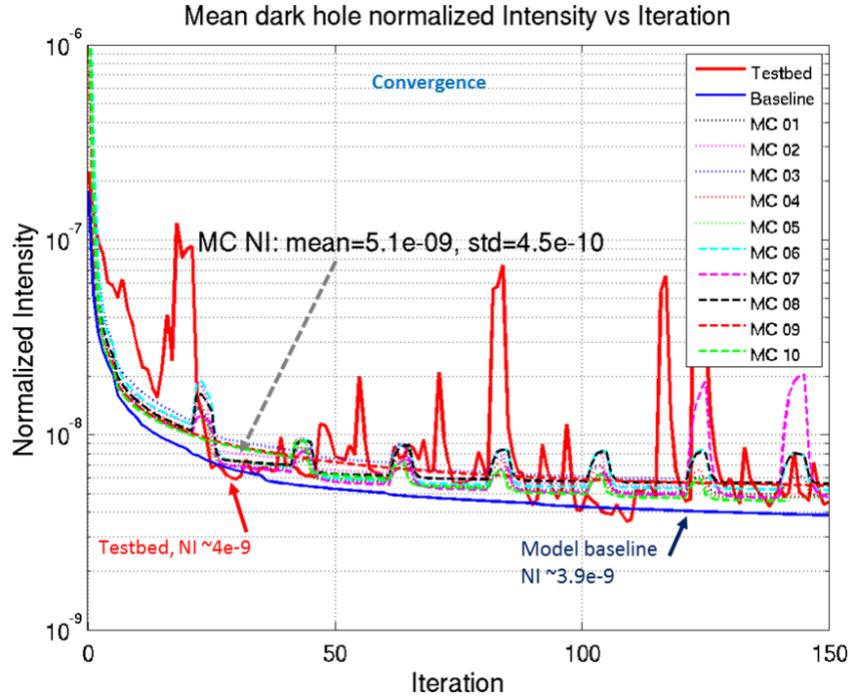


Figure 7: Normalized intensity vs iteration, testbed data vs model prediction, *Config.2*

Compared to Fig.4 of *Config.1*, one noticeable difference is now we have a much closer baseline prediction to the testbed result in both final contrast as well as in contrast convergence. The envelope of testbed iteration curve (solid red line) is much more tightly bounded between the base model prediction (blue line) and MC runs predictions (dashed lines), indicating both calibration and calibration error are adequately captured. In *Config.1*, initially only achromatic pupil aberration was included in both baseline model as well as MC runs #1~5, and results failed to explain the testbed raw contrast floor. Capturing and including key knowledge and knowledge error in model (in this case, the system chromatic WFE caused by polarization) is essential to an as-built system performance modeling.

Figure 8 shows final contrast vs field position (left) and that vs wavelength (right). Figure 9 compares the predicted speckle pattern with that of the testbed result. Except radial contrast at IWA $4 \lambda/D$, we see generally good agreement between model prediction and testbed result.

A note on the alternating regularization strategy: our later simulation shows that it is possible to avoid alternating regularization procedure by using a vector regularization to achieve same effective regularization strategy.

4. OPEN LOOP CONTRAST SENSITIVITY

Speckle stability is what ultimately matters during coronagraph science imaging, since the high order WFSC loop will not be operating during science observation (although low-order control will). Some of the top stability concerns include 1) telescope thermal instability (low order WFE, pupil shear), 2) telescope pointing and jitter, and 3) coronagraph mask deployment (mask offset). While contrast floor was the focus of the current model validation effort, we did collect a few open loop testbed contrast sensitivity data sets and compared them to our model predictions.

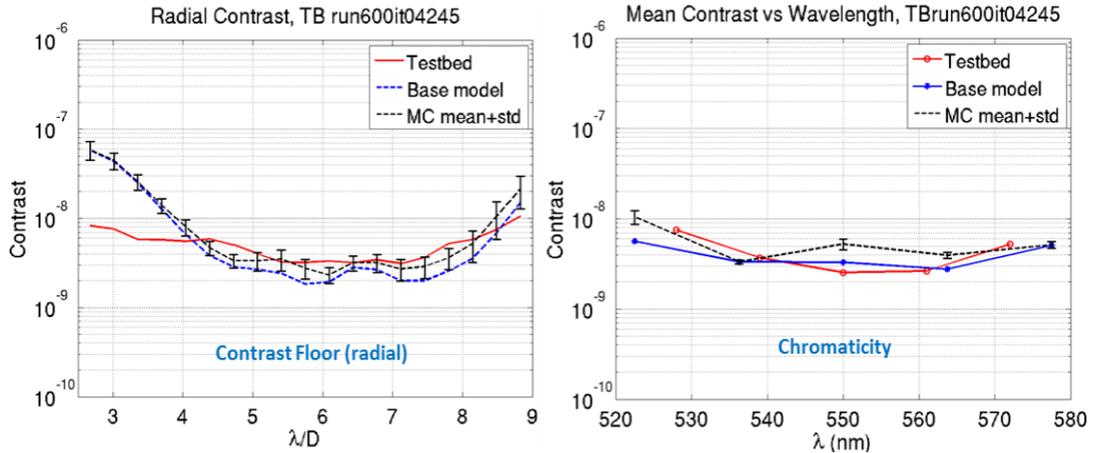


Figure 8: Radial contrast (left) and chromaticity (right) at final nulling sequence, *Config.2*

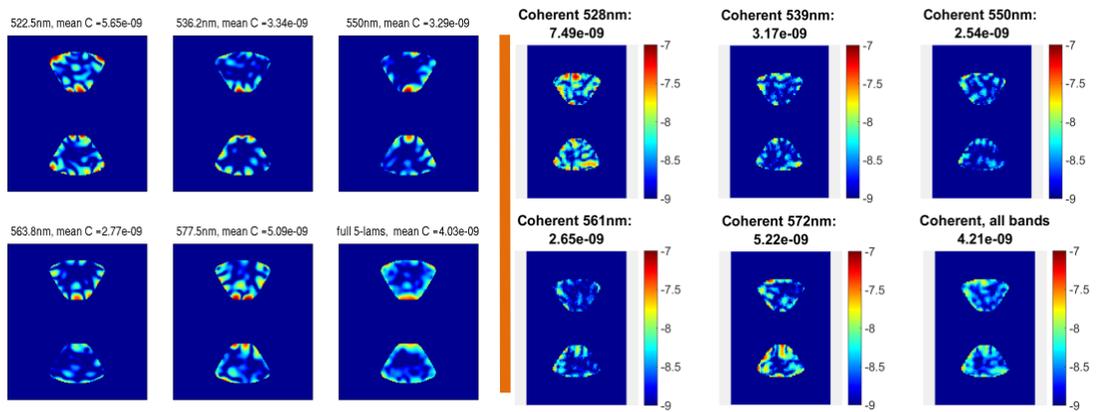


Figure 9: Speckle pattern (final nulling sequence) comparison: model prediction (left 6 subplots, per wavelength and then full bandwidth) vs testbed result (right 6 subplots, per wavelength and then full bandwidth), *Config.2*

4.1. Low order Zernike WFEs

Despite the success of the LOWFS/C in maintaining dark hole better than $1e-8$ in a simulated dynamic environment [4 - 6], better stability is eventually desired for science, perhaps $1e-9$ or better in contrast change. As part of the investigation along this line, sensitivities to low order Zernike WFE perturbation were collected from testbed and compared with model predictions.

Since the original OTA was removed from OMC, the main method on testbed to inject Zernike perturbations is to apply fitted DM voltages (to Zernike terms) converted from DM gain map. Typically, this is done in equivalent 0.2nm rms step size over ± 2 nm scan range around nominal DM setting, one Zernike at a time, and record the corresponding dark hole image series. For tip tilt, testbed also used two alternative methods: 1) through controlled injection at jitter mirror, and 2) through lateral offset of image plane occulter mask (image plane offset is equivalent to pupil plane tilt). On testbed, a conversion factor 0.13 μ m occulter offset equals to 1nm rms tip tilt was determined. For focus term, one alternative method is used: by source Z (axial) motion, with a conversion scale = 1nm rms focus / 32 μ m linear motor motion.

On model prediction, we also used two different ways to apply Zernike perturbations:

- Method #1:* Apply directly at the pupil plane as thermal perturbation would have caused
- Method #2:* Apply as fitted DM voltage through DM gain map conversion, same as testbed implemented.

In both cases, sensitivity is evaluated with *Config.2* model post WFSC (i.e., the starting contrast is $\sim 4e-9$).

Since the metric of interest in general is speckle field sensitivity, but the testbed collects intensity rather than E-field images (which would require probing that is not practical), we fit the testbed delta intensity image series (relative to the nominal setting one) pixel by pixel to a 2nd order polynomial. The quadratic coefficient of the fit gives us field sensitivity for that pixel per nm rms. Averaging over all dark hole region pixels, we obtain the testbed's field sensitivity for each Zernike perturbation term. The results are largely similar to that from raw intensity based fitting (which itself is mostly well behaved), indicating that a simple contrast based sensitivity should be largely adequate in general [5]. Figure 10 illustrates a typical well-agreed fitting result case (left, Z4 shown) and one slightly off fitting case (right, Z6).

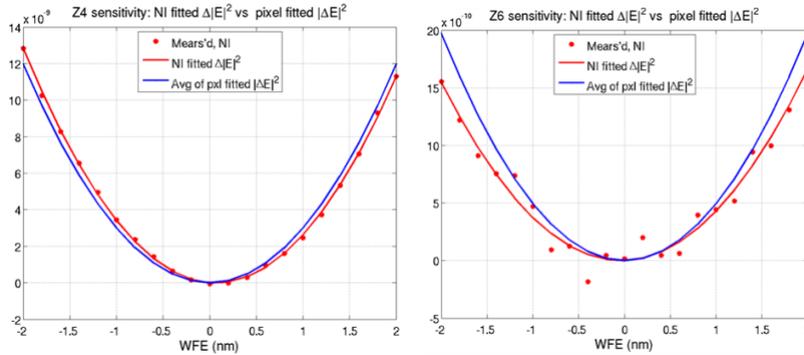


Figure 10. Examples of pixel based fitting for field sensitivity vs NI based fitting for contrast sensitivity.

In Fig.11 we plot the comparison of Zernike field sensitivity per 1nm rms from testbed measurement and that of model predictions. The measured data include measurement via DM voltage, tip tilt via jitter mirror, (equivalent) tip tilt via occulter offset, and defocus via source Z motion. The model used here is based on *Method #1* (perturbation applied at pupil phase). We see mostly good agreement between model prediction and testbed result. The only notable exception is testbed tip tilt via DM voltage (the solid blue segment), which are off by several times in magnitude. We postulated it was caused by imperfect DM gain calibration and tried the second modeling method mentioned above.

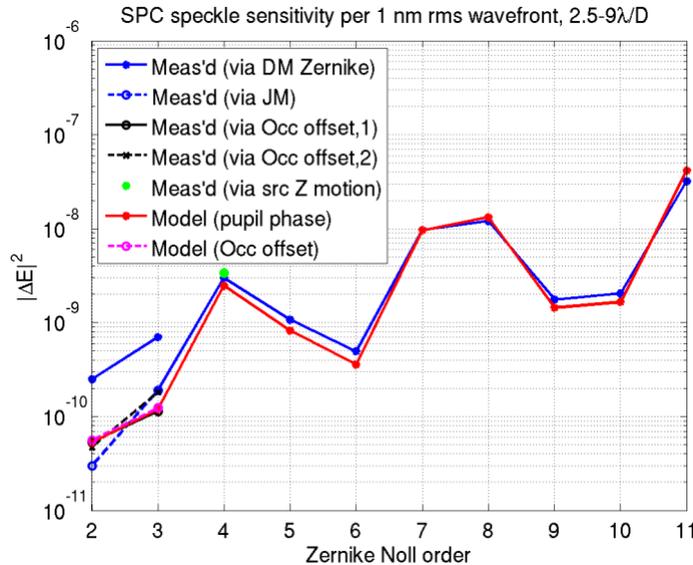


Figure 11: Zernike sensitivities *Method #1* pupil phase model (red) vs testbed (blue, black, or green, lines or dot)

In second method, we apply each Zernike term as fitted DM voltage through DM gain map, just as testbed actually implemented. Recall that the testbed DM gain typically has calibration error between 5~8%. Empirically, we found that by adding a random gain error of 6% to fitted DM voltage, the result matches the observed tip tilt sensitivity very well, as shown in Figure 12, where the dashed red line and error bar is the result of 10 instances of random DM gain error of 6%. The gain error has no or little impact on higher Zernike terms, as they have higher contrast sensitivity to start with (compared to the gain error effect). The fractional error between the testbed result (via DM voltage application, the solid blue) and model prediction (the dashed red) are all within +/-25%, with a mean (of absolute fraction error) ~ 9%

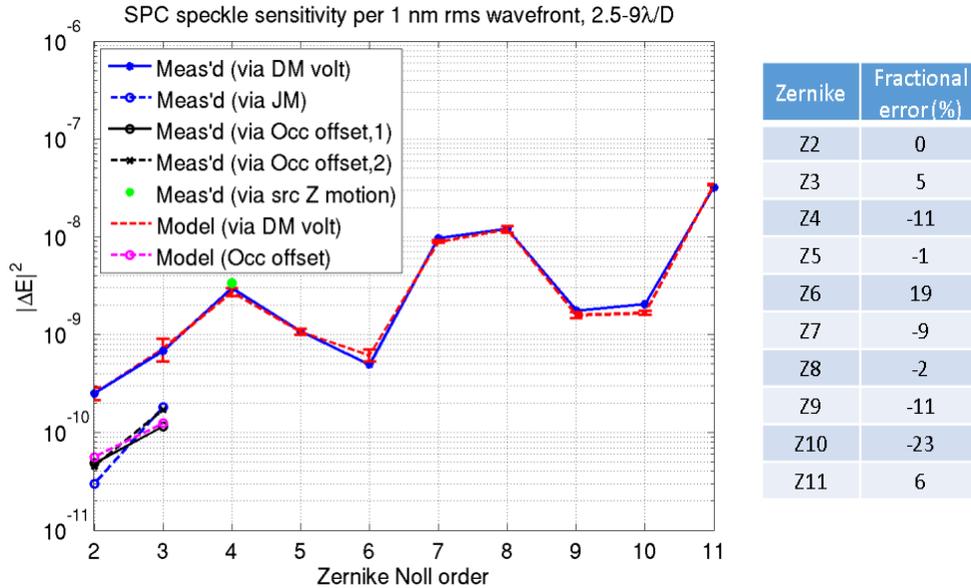


Figure 12: Zernike sensitivities at 1nm rms; Method #2 DM gain model (dashed red line) vs testbed (blue, black, or green, lines or dot). Fractional error is calculated as: $(model - measured) / model * 100$

The excellent match in Figure 12 is a strong evidence that the apparent differences in tip tilt sensitivity via DM voltage application from pupil phase model prediction (and from those via occult offset or jitter mirror) are almost certainly due to the DM gain error. The slightly less great match in pupil phase model prediction (Figure 11) can be explained as the phase applied on testbed is not precisely what it intended to be due to DM gain error. When factoring the DM gain error of a typical amount found on testbed, the model prediction error of individual Zernike term is within 25% of its true value for all Zernike's Z2~11 induced via DM voltage application. The difference in tip tilt sensitivity of other testbed methods (JM, occulter offset) from model prediction (via pupil phase application) have a max fractional error 60%. This consistency among multiple methods (on testbed and in model) demonstrates our model's good fidelity (and supports the notion that the good agreement in raw contrast brings better agreement in contrast sensitivity).

4.2. Mask lateral shift

Stability (as well as accuracy and repeatability) of various mask deployment can have significant impact on contrast performance during science operation. Previous filter wheel mask error budget analysis identified occulter mask lateral shift of shaped pupil coronagraph as the driving requirement (for stability), and shaped pupil mask lateral shift as the next. In this section, we describe the comparison of our model predictions to testbed results and verification of these two error budget terms.

The testbed data were collected by scanning the focal-plane occulter in x and y directions, in coarse steps of 1um and finer steps of 0.1 um over +/-10 um span. Figure 13 shows the measured and fitted sensitivity. Although not a perfect agreement in the x axis scan at large shear, it is a reasonably good match between measured data and model prediction out to the

limits of the relevant range (of error budget). The measured data are not very accurate below 1um, but we still can get reasonable number from the fitted testbed data. The sensitivity at $1e-10 \Delta C$ is about 0.35um and 0.2um, for x- axis and y- axis, respectively. The fractional error between model and testbed result is 9% and 20% for x and y, respectively. Extrapolated these to $1e-11 \Delta C$, the sensitivity is within 2X the previous filter wheel error budget allocation (based on full Proper design model). This serves as one empirical verification of filter wheel mechanism error budget.

This occulter shift data was also used as equivalent tip tilt sensitivity in Sect. 4.1, with a conversion factor of 0.13um occulter offset per 1nm rms tip tilt.

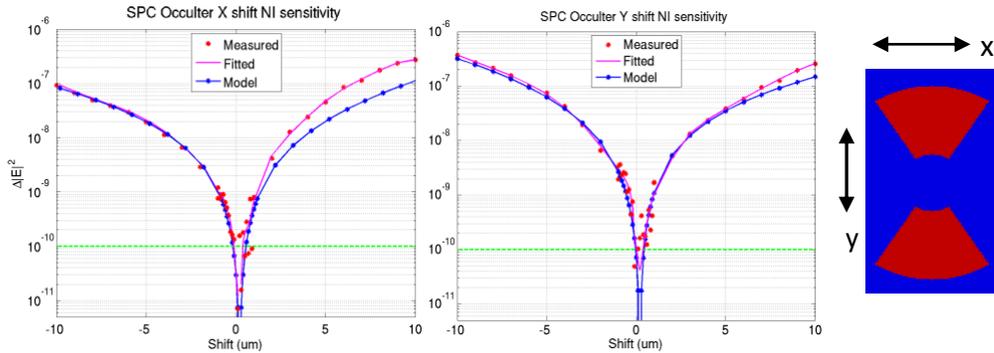


Figure 13: Occulter lateral shear sensitivity, model prediction (blue) vs testbed data (red dot and line)

For shaped pupil mask lateral shift, only y direction coarse scan data (in steps of 4um) was collected due to limitation in hardware motion. Figure 14 shows the result. Again, we see good agreement between testbed measured and model prediction, with a fractional error at $1e-8 \Delta C$ about 33%. The extrapolated sensitivity at $1e-11 \Delta C$ is about 2.4X the previous analysis; slightly larger than desired due to coarse scan.

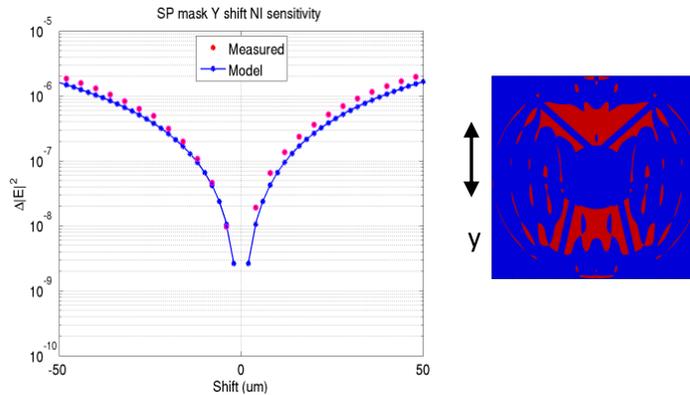


Figure 14: Shaped pupil mask shear sensitivity, model prediction (blue) vs testbed data (red dot and line)

5. SUMMARY AND FUTURE WORK

We have constructed a testbed SPC model using the OMC testbed results and available characterization data and compared its WFSC performance predictions to that of testbed under different hardware and software (control model) configurations and in a number of metrics. In the process, we identified the missing chromatic pupil aberration as the most plausible cause in early model's contrast gap from testbed, and improved contrast agreement in later model (and contrast on testbed as well in later runs) after the inclusion of it.

We achieved good agreement between testbed results and model predictions in all three aspects of contrast performance metrics: raw contrast floor, contrast convergence rate, and some key contrast sensitivities such as low order zernike wavefront and occulter mask and shaped pupil mask lateral shears:

- Contrast floor: mean <30% ; chromaticity ~34% (average)
- Contrast convergence rate: comparable (envelope)
- Contrast sensitivity:
 - Low order Zernike WFE (Z2 ~ Z11): all <25%, average ~9%, at 1nm rms
 - Occulter mask lateral shear: <20% , at 1e-10ΔC
 - Shaped pupil mask lateral shear: ~33%), at 1e-8 ΔC

Our results illustrate the impact of an incomplete model in the wavefront control loop on the predicted performance of an as-built system. While flight system will not suffer the specific polarization problem as on testbed (due to imperfect hardware used for simulating star), there might be other phenomenon that we have not have a good understanding.

In terms of calibration, the existing calibration as carried on testbed was reasonably adequate in the two cases studied here, in the sense that their collective uncertainty has small impact on contrast floor in MC runs; the close base model prediction to testbed result of *Config.2* can also be interpreted that calibration errors are small. The missing part of model (chromatic pupil WFE in this case) contributes to a much bigger contrast gap in *Config.1*; while in *Config.2* a better prediction (and better testbed contrast) comes with a more complete model. Of course, desire of deeper contrast likely will demand better calibration of some items, a part of flight model error budget analysis we are currently working on.

More works are still needed in refining our model to have even better fidelity and accommodating for new test conditions. For example, the radial contrast profile near IWA in *config.2* prediction has a large discrepancy from that of testbed result. Future testbed tests such as low flux nulling may require modified sensing and/or control method.

ACKNOWLEDGMENTS

This work was performed at the Jet Propulsion Laboratory of the California Institute of Technology, under contract with the National Aeronautics and Space Administration. © 2017 California Institute of Technology. Government sponsorship acknowledged.

REFERENCES

- [1] D. Spergel, et al., "Wide-Field InfraRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report." <https://arxiv.org/abs/1503.03757>, (2015)
- [2] Macintosh, B., et al., "Science capabilities of the WFIRST coronagraph," Proc. SPIE, 10400 (this conference), (2017)
- [3] Poberezhskiy, I., et al. "Technology development towards WFIRST-AFTA coronagraph." *SPIE Astronomical Telescopes+ Instrumentation*. International Society for Optics and Photonics, 2014
- [4] Shi, Fang, et al., "Testbed demonstration of low-order wavefront sensing and control for WFIRST coronagraph", Proc. SPIE, 10400 (this conference), (2017)
- [5] Cady, E., et al., " Shaped pupil coronagraph for the WFIRST: high-contrast broadband testbed demonstration," Proc. SPIE. 10400, (this conference) (2017)
- [6] Seo, J., et al. "Hybrid Lyot coronagraph for WFIRST: high-contrast broadband testbed demonstration," Proc. SPIE, 10400 (this conference), (2017)
- [7] Krist, J. et al., "Numerical modeling of the proposed WFIRST-AFTA coronagraphs and their predicted performances," J. Astron. Telesc. Instrum. Syst. 2(1), 011003 (2016).
- [8] Krist, J., "End-to-end numerical modeling of AFTA coronagraphs", *Proc. SPIE*, 9143, 91430V (2014).
- [9] Krist, J., et al., "WFIRST coronagraph optical modeling," Proc. SPIE, 10400 (this conference), (2017)

- [10] Shaklan, S., et al., “Coronagraph starlight suppression model validation,” https://exep.jpl.nasa.gov/files/exep/COMBINEDv5_Milestone%203A%20Final%20Report%20072915.pdf ExEp Technology Milestone #3A Final Report (2015)
- [11] Zhou, H., et al, “Closing the Contrast Gap between Testbed and Model Prediction with WFIRST-CGI Shaped Pupil Coronagraph,” *Proc. SPIE.* 9904, (2016)
- [12] Nemati, B., et al., “Sensitivity of the WFIRST coronagraph performance to key instrument parameter,” *Proc. SPIE.* 10400, (this conference) (2017)
- [13] Zimmerman, N. T., et al., “Shaped Pupil Lyot coronagraphs: high-contrast solutions for restricted focal planes,” *Journal of Astronomical Telescopes, Instruments, and Systems*, 2(1):0110912, (2016)
- [14] Give'on A. *et al*, “Broadband wavefront correction algorithm for high-contrast imaging system,” *Proc. SPIE.* 6691, 66910A (2007).
- [15] Give'on A. *et al*, “Pair-wise, deformable mirror, image plane-based diversity electric field estimation for high contrast coronagraph,” *Proc. SPIE.* 8151, 815110-2, (2011)
- [16] Marx, D., “Electric field conjugation in the presence of model uncertainty,” *Proc. SPIE.* 10400, (this conference) (2017)