Electric Field Conjugation in the Presence of Model Uncertainty

David Marx*, Byoung-Joon Seo, Brian Kern, Erkin Sidick, Bijan Nemati, and Ilya Poberezhskiy
Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109

ABSTRACT

The Wide-Field Infrared Survey Telescope (WFIRST) is a 2.4m diameter space telescope NASA program. The payload will include a coronagraph instrument (CGI). The CGI designs under development use deformable mirrors (DM) to create a point spread function (PSF) with a dark region around the obscured star object. Electric field conjugation (EFC) is an iterative nonlinear optimization procedure that uses measurements of the electric field at the image to determine the DM displacements to modify the PSF to create the region of high contrast in the image. EFC requires a numerical model of the coronagraph to calculate the Jacobian of the system, which is used, along with regularization, to solve for the DM displacements for each iteration of the nonlinear optimization. Ideally, the coronagraph is aligned and calibrated, and the calibration data are used in the numerical model for calculating the Jacobian. However, calibration and alignment measurements always contain uncertainty resulting in calibration error. Therefore, the Jacobian calculated from the numerical model is not an exact representation of the physical coronagraph, and the resulting DM solution for an EFC iteration does not have the exact impact on the electric field of the coronagraph as predicted by the EFC. The result is slow convergence, and, as will be shown, drives the use of more restrictive regularization. Using Monte Carlo trials, we investigate the effect of calibration error on EFC convergence and regulatization. Comparison to results from the High Contrast Imaging Testbed Hybrid Lyot Coronagraph are also presented.

Keywords: Coronagraphy, adaptive optics, deformable mirrors, space telescopes, exoplanets, nonlinear optimization

1. INTRODUCTION

The Occulting Mask Coronagraph (OMC) testbed was assembled, aligned, and placed in a vacuum chamber at JPL within the past year. The OMC testbed was designed to test and demonstrate coronagraph performance for the OMC flight instrument to be on the Wide-Field InfraRed Survey Telescope (WFIRST). The common architecture accommodates both a hybrid Lyot coronagraph (HLC) and a shaped pupil coronagraph. The pupil of the testbed is also representative of the WFIRST pupil. For a more complete description of the OMC testbed, including experimental results the for HLC and SPC coronagraphs, see Seo [1] and Cady [2]. All of the modeling results and discussion presented below are for the HLC.

Once in the chamber and stabilized, the coronagraph is calibrated by measuring optical alignments, DM registration, and wavefront error using a variety of techniques. A wavefront flattening procedure consists of iteratively measuring the wavefront error and correcting the wavefront error with the DMs, until no further improvement is possible. The results of all the calibrations, including the measured, residual wavefront error are used to build a numerical model of the coronagraph. The numerical model is an integral part of the EFC control as it is used for calculating the Jacobian and estimating coronagraph performance for various parameters such as regularization

The question we address here is the effect of measurement errors, or other mismatch between the coronagraph and the numerical model used in EFC on the performance of EFC and the coronagraph. We define model error, or model uncertainty, as any mismatch between the numerical model used in the EFC calculations and the physical coronagraph in the vacuum chamber. To study the effect of calibration errors on EFC, we constructed a simulation where the physical coronagraph was replaced by another numerical model. The EFC control, including probing, DM actuator gains, and DM actuator constraints was the identical software and model used by the testbed. The actuator constraints include maximum stroke and neighbor rules. By replacing the physical coronagraph with a numerical model, we could introduce controlled calibration errors and run the EFC convergence for many different cases calibration errors. The results presented below are derived from Monte Carlo random trials of calibration error, and the types and ranges of calibration error were derived from realistic estimates of measurement uncertainties from the testbed.

*David.S.Marx@jpl.nasa.gov; phone 1 818 354-2837

In the next section, we review EFC in the current context and describe the architecture of the numerical simulations. Details of the calibration errors realistic for the HLC testbed, and their incorporation into the Monte Carlo simulation, are described in Section 3. Section 4 discusses the results of the simulations, and metrics used to describe the effects of calibration errors on EFC. We also discuss the interaction of calibration error and the choice of regularization. Finally, we describe a regularization schedule that improves the performance of EFC in the presence of calibration errors.

2. REVIEW OF EFC

Electric field conjugation (EFC) is an iterative nonlinear optimization method to minimize the diffracted intensity of a source star in a coronagraph field of view [Give'on 2007]. Viewed as a system, the free parameters of the optimization (state of the system) are the DM actuator displacements, and the metric (output) is the average normalized intensity in the coronagraph field of view, or "dark hole". The optimization algorithm uses a Jacobian matrix, calculated for the current state, to solve for the change in state to improve contrast.

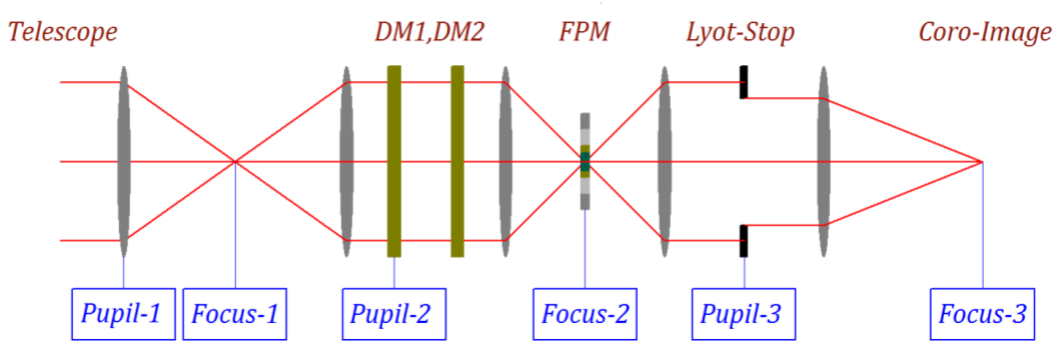

Figure 1. The key elements of the OMC testbed layout. The light source ("starlight") is a collimated beam delivered by a telescope to input pupil (Pupil-1), and a CCD science camera is located at the coronagraphic image (Coro-Image, Focus-3) plane for detecting the image of the "starlight".

As shown in Give'on [3], for small actuator movements, the change in the electric field at the coronagraph image is approximately linearly dependent on the change in the DM actuators,

$$J\underline{\delta DM} = \underline{\delta E}, \tag{1}$$

where J is the Jacobian matrix of the approximately linear coronagraph system, $\underline{\delta DM}$ is a vector of actuator movements, and $\underline{\delta E}$ is the vector of change electric field real and imaginary values in the coronagraph image field of view due. Because the system is in fact nonlinear, the Jacobian is dependent on the current actuator state.

Given the current state of the coronagraph, and the electric field in the image field of view, $\underline{E}$, we desire to solve for the DM actuator movement that reduces the image field to zero. That is, we want to solve,

$$J\underline{\delta DM} = -\underline{E}, \tag{2}$$

for $\underline{\delta DM}$, which is, in general, an overdetermined system. In practice, we add a constraint on the total DM movement, and solve for $\underline{\delta DM}$ that gives the minimum squared error [5],

$$\left[J^T J + \alpha^2 10^{-\beta} I\right]\underline{\delta DM} = -J^T \underline{E}, \tag{3}$$

where $\alpha^2 = \max\left(\mathrm{diag}(J^T J)\right)$, β is the regularization parameter, and I is the identity matrix. Equation (3) is numerically solved using standard linear algebra routines, such as LU decomposition. As the underlying system is in fact nonlinear, solutions to equation (3) are iterated. Each iteration, the electric field, $\underline{E}$, is measured using the probe method [6][4], a new Jacobian is calculated from the new DM state, and equation (3) is solved to produce a new vector of DM actuator movements. The new DM actuator state is sent to the coronagraph, and the process repeats.

The elements of the Jacobian relate the change in the electric field at a pixel in the field of view to a change in a DM actuator displacement, $J_{m,n} = \frac{\partial E_m}{\partial DM_n}$, where E_m is the electric field at pixel m, and DM_n is the displacement of DM actuator n. The Jacobian is calculated from a numerical model of the coronagraph because estimating the Jacobian from the real coronagraph is prohibitively time consuming. Therefore, a numerical model of the coronagraph system is an integral part of the EFC iterative procedure. The error between the numerical model used in EFC and the real coronagraph instrument is the main consideration of this paper.

The regularization parameter, β, is typically considered a free parameter to constrain the solution and the nonlinear error. Smaller β values are less restrictive, and allow the DM actuator to make a greater mean-squared movement, while a larger β value places a greater constraint, or cost, on DM movement. Thus, the solution produced with a smaller β value is more likely to contain nonlinear error, and at some β value, the nonlinear error overwhelms the solution, and the contrast degrades rather than improves for the next iteration. We define the "optimal β" as that β where the EFC model solution predicts the best improvement in contrast for the following iteration. If the EFC model also includes calibration error, then the actual improvement of the testbed contrast, for any chosen β, is worse than predicted by the EFC model.

After a new DM state is determined, the coronagraph DM's are modified to the new state, and the coronagraph is probed to estimate the electric field speckles and contrast. The new electric field estimate is then used in the next iteration. As shown in Figure 2, each iteration cycles between a model of the coronagraph and the real coronagraph.

So that the model is an accurate representation of the coronagraph, the coronagraph is characterized and calibrated. Examples include measuring the alignment offsets of the DMs, alignment of the Lyot stop, alignment of the focal plane mask, gain values of the DM actuators, and measurement of the pupil wavefront error. The model is not a perfect representation of the physical coronagraph because the characterizations have measurement uncertainty, not all possible alignments and other non-idealities are characterized, and many of the non-idealities are lumped into the pupil plane wavefront error. If the model is not a perfect representation of the coronagraph, then the Jacobian calculated from the model is not exact for the coronagraph. Then the estimation of the change in the DM state and regularization contains error.

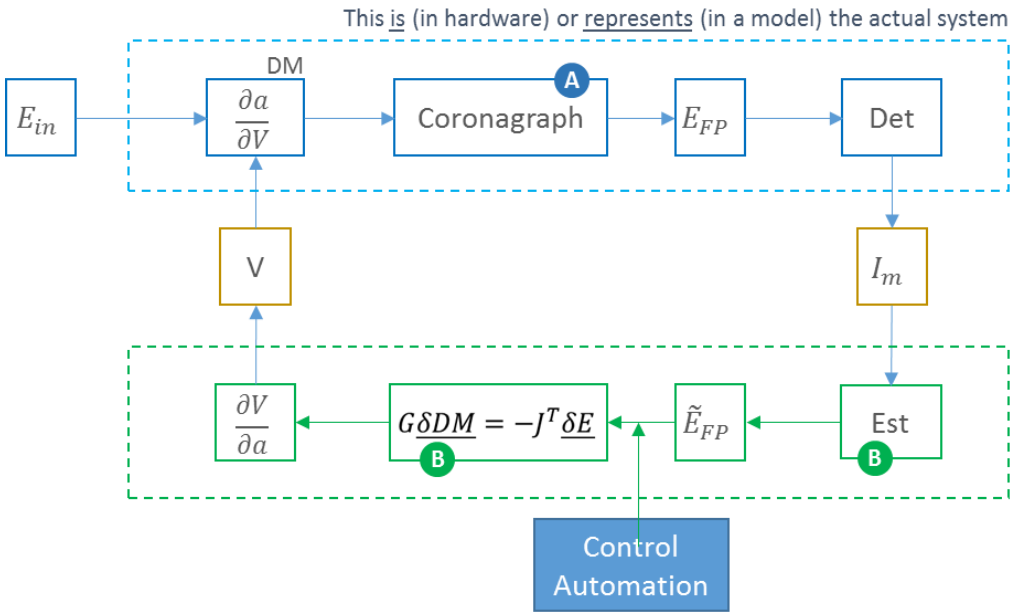

Figure 2. Flow chart of coronagraph operation, including the EFC control loop. The dashed blue box encloses the coronagraph instrument, which could be the real instrument in the testbed, or a numerical model. $E_{in}$is the initial pupil wavefront; $\frac{\partial a}{\partial V}$ represents the actuator gains in displacement per Volt; E_{FP} is the electric field at the focal plane. The dashed

green box encloses the EFC control that takes the electric field at the focal plane as estimated from probing, calculates the Jacobian, solves the linearized equation for DM actuator displacement at a chosen regularization, and outputs actuator voltages to the coronagraph. The only communication between the EFC control and the coronagraph are intensity images and DM actuator voltages. "Control Automation" represents the automated control parameters such as regularization.

3. SIMULATION OF MODEL ERROR

Model error is the difference between the electric at the coronagraph field of view and the change in the electric field predicted by the EFC model,

$$\epsilon_{\mathrm{M}} = E_{\mathrm{testbed}}(DM + \delta DM) - E_{\mathrm{model}}(DM + \delta DM). \tag{5}$$

Contributions to model error include calibration error, testbed drift, and uncorrelated noise such as shot noise, camera noise, and source intensity noise. In our simulation, we consider only calibration error and ignore the others. Model error does not include nonlinear error, which is the difference, $\epsilon_{\mathrm{NL}} = J\underline{\delta DM} - \underline{\delta E}$. Nonlinear error is the difference between the solution of (3) and the electric field as calculated from the diffraction model used in EFC to evaluate solutions, calculate the Jacobian, and estimate E_{model}. Nonlinear error restricts the magnitude $|\underline{\delta DM}|$ in each iteration because of the original assumption of small actuator deviations to linearize the diffraction model. However, calibration error causes error in the direction of $\underline{\delta DM}$.

To model and understand the effects of calibration error on EFC, we simulated testbed operation, as shown in Figure 2, but with the physical testbed replaced with a numerical diffraction model of the coronagraph. The only information exchanged between the EFC control and the simulated testbed are camera images and DM actuator instructions, just as in the physical testbed. Thus, the same probing for electric field estimation, control strategy, and diffraction model used in the physical testbed was also used in the model simulation. The control strategy governs EFC parameters such as regularization β, probe strength, control region, actuator gain, etc. Testbed operation has long known that EFC using regularization β predicted by the EFC model leads to poor results for the testbed coronagraph. As we will show in the results, calibration error must play a role in this divergence between model and testbed. Because β predicted by the EFC model cannot be used, a control strategy to automate the β value as the coronagraph contrast converges must be chosen. Thus, the control strategy is an integral part of the simulation of the testbed with calibration errors.

Calibration errors were added to the numerical testbed using Monte Carlo method. Table 1 shows the list of calibration errors used in the Monte Carlo simulations. For each parameter listed in the table, the "nominal value" is the measured alignment on the OMC testbed. Thus, the nominal value is a known error between the coronagraph design and its alignment. This type of error might cause the coronagraph to perform below the design expectation, but it is not an error that causes the EFC diffraction model to be in error to the testbed, because the nominal alignment error is common to both the testbed and the EFC model. The last column of the table lists the standard deviations of the calibration errors added to the numerical testbed in each Monte Carlo trial. Some errors that were ignored in our simulation are any wavelength dependent errors, DM actuator gain errors, and jitter.

Pupil phase and amplitude errors require special consideration because they are not easily described by a single parameter. The baseline pupil amplitude and phase field were the measured amplitude and phase in the OMC testbed. Normal testbed practice is to measure the pupil wavefront error and amplitude through the use of phase retrieval [1][8]. DM1 is actuated to correct the measured pupil wavefront error, but a residual wavefront error remains. The residual is the nominal wavefront error common to both the testbed and the EFC model. To create pupil calibration errors for Monte Carlo trials, we added noise representative of camera noise to the images used for the phase retrieval. The resulting pupil wavefronts contained a calibration error of 0.05 rad rms.

4. SIMULATION RESULTS

All of the numerical simulations presented here are for the HLC coronagraph in the OMC testbed, scored in the dark region of the field of view between $3\lambda/D$ and $9\lambda/D$, with a 10% bandwidth centered at 550nm, and the AFTA pupil. The reported metric of normalized intensity is defined as the ratio of the average intensity in the scored field of view with the FPM in to the peak intensity of the centered star with the FPM removed.

Table 1. Alignment parameters measured on the coronagraph testbed to calibrate the EFC model. The last column, "MC Error σ" is the standard deviation of calibration error added to the nominal value for the Monte Carlo testing of the effect of calibration errors on EFC.

Name	Parameter	Unit	Nominal Value	MC Error σ
Pupil	Pupil Diameter	pix	305.3	0
DM1	X-Decenter	um	-94.1	50
	Y-Decenter	um	-475.1	50
	X-Rotation (Tip)	deg	8.25	0
	Y-Rotation (Tilt)	deg	-0.07	0
	Z-Rotation (Clocking)	deg	-0.86	0.05
DM2	X-Decenter	um	799	50
	Y-Decenter	um	-83.5	50
	X-Rotation (Tip)	deg	2.37	0
	Y-Rotation (Tilt)	deg	0.004	0
	Z-Rotation (Clocking)	deg	0.45	0.05
Lyot	X-Decenter	um	0	25
	Y-Decenter	um	0	25
FPM	X-Decenter	um	0	0.5
	Y-Decenter	um	0	0.5
Pupil	Amplitude Variation	rel. int.	Measured	Intensity Noise
	Phase Variation	rad (rms)	Measured	0.05

Figure (3a) compares the EFC convergence of normalized intensity (NI) for the Monte Carlo trials with calibration error to the baseline case of no calibration error. The few iterations for the baesline case where the NI spikes are due to DM actuators violating actuator constrain limits. Those actuators are then constrained, and EFC recovers the following iteration. Except for those few iterations, the NI for the baseline case quickly converges to a value of about 1.0e-9 in about 58 iterations.

Convergence for the Monte Carlo trials is slower compared to the baseline case, and appears to approach a contrast floor roughly ten times worse. This result is qualitatively similar to initial OMC testbed results before various improvements and calibrations. We can conclude that for the later iterations, the calibration error prevents EFC from improving the NI. In all cases, β was chosen each iteration to maximize improvement, the "optimal β," as discussed above. We can gain some understanding by graphing β for the baseline and Monte Carlo cases (Figure 3b). For the baseline case, β continues to trend to smaller values (more negative) as the NI converges. Thus, the constraint on actuator movement relaxes, and the actuator solutions are free to address finer details of the residual electric field. However, β for the Monte Carlo cases oscillates around a floor and does not reach the similar low values as for the baseline case. Then the actuator constraint for the Monte Carlo cases does not relax to the same degree as the baseline, and so the Monte Carlo cases are not able to address fine details of the residual electric field, thereby limited to the NI floor. The result is that calibration error prevents EFC from correcting higher modes of the electric field, thereby limiting NI convergence. For further discussion of the relationship between β and electric field modes, see Sidick [7].

Another indication of the mismatch between model and coronagraph is the difference in response to the regularization parameter. Figure (4) compares, as a function of β, the EFC model predicted NI to the NI realized by the simulated coronagraph with calibration error. This illustrates why the optimal β strategy works well for the model but in practice when calibration error is present. For this specific example, EFC will choose $\beta = -4.40$, and send the corresponding $\underline{\delta DM}$ solution to the testbed. The EFC model (blue curve) expects an improvement in NI from approximately 1e-7 to approximately 3e-9, but the testbed with calibration error only improves to approximately 6e-8. Furthermore, the solution that would make the most improvement for the testbed NI is from a significantly larger $\beta = -1.25$. The large slope of the

testbed curve in the region of the "optimal β" implies that the testbed is result includes nonlinear error as well as the calibration error.

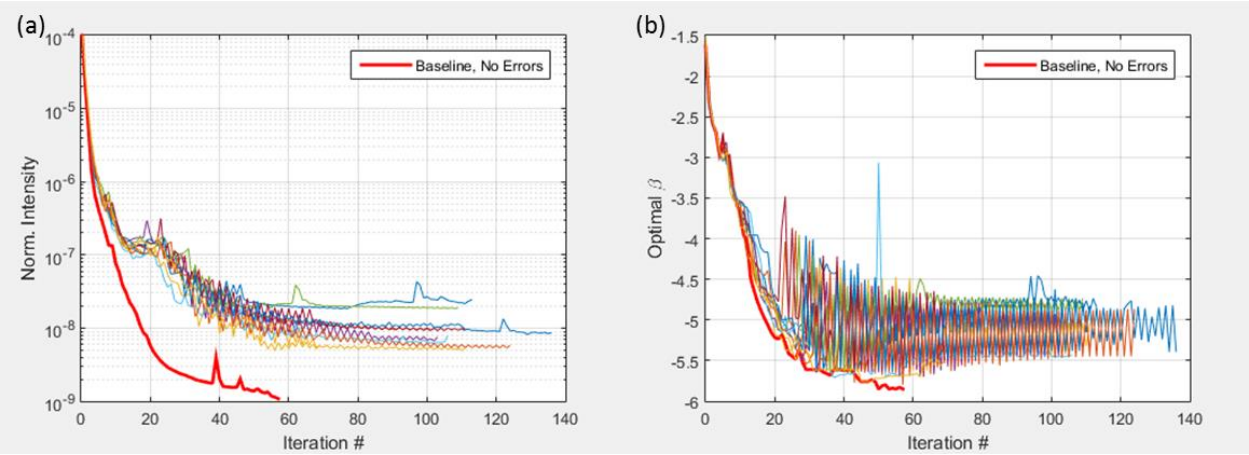

Figure 3. (a) The addition of calibration errors to the numerical model of the testbed prevents EFC convergence to the normalized intensity that would otherwise be achieved, plotted as the "Baseline". (b) Regularization parameter β plotted for the same iterations. While β relaxes for the Baseline case with no calibration errors, it oscillates and remains constraining on the DM solution when calibration errors are introduced.

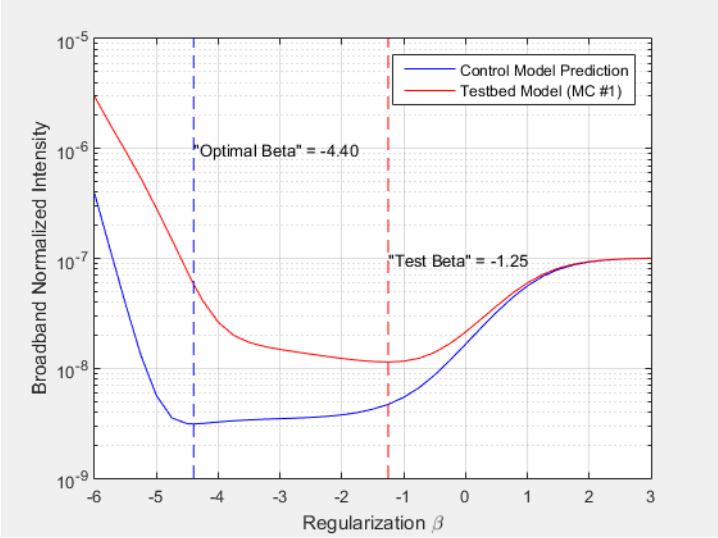

Figure 4. Comparison of the normalized intensity that would result for the next iteration as a function of the regularization parameter β used in EFC. The blue curve is what the EFC model predicts for the next iteration, but because of calibration error added to the numerical model of the testbed, the actual testbed result is as shown in red.

Actual testbed operation rarely uses the optimal β. The reason is that optimal β as estimated by the EFC model typically causes the NI to increase the following iteration. In that case, the model error, Equation (5), is very large, a direct result of calibration error. As a result, real coronagraph testbed operation typically uses β values with less magnitude (more restrictive) than the optimal β estimated by the EFC model. Reasoning that without relaxing β, more complicated fields can never be corrected [7], we experimented with schemes using the optimal β for some iterations, and more restrictive β for other iterations. Iterations where the optimal β was used caused the NI to increase, but then following iterations with a more restrictive β recovered the NI, generally with an overall improvement relative to the initial iteration. Thus, we

empirically developed the control strategy with β values alternating between the optimal β, and a value of -2.5. Results using this new control strategy for the same models with calibration error are plotted in Figure (5). Compared to the initial results using optimal β, Figure (3a), the new strategy achieves better NI with the same calibration errors. The results clearly show that for the iterations where optimal $\beta$is used, NI is larger and not converging. But only these iterations can address higher order modes in the image electric field. The following iterations with $\beta = -2.5$ then achieve a lower NI, but quickly converge and cannot make further improvements because the restrictive β does not allow solutions addressing high orders. The first iteration with $\beta = -2.5$ after iterations with optimal β is similar to the order of magnitude improvement for the testbed model shown in Figure (4) (red curve).

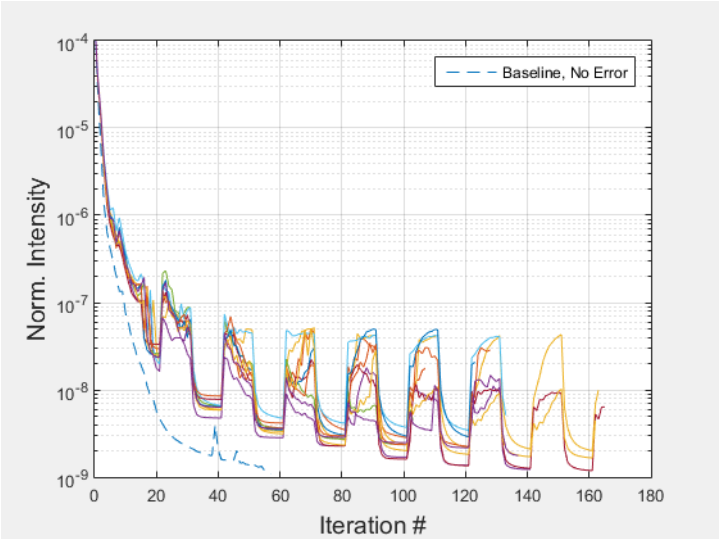

Figure 5. Convergence curves for the same Monte Carlo trials of calibration error as Figure 3, except a new regularization strategy of periodically using the optimal β or a fixed β of -2.5. Those iterations where the normalized intensity jumps high are when optimal β are used, and those iterations where the normalized intensity drops to a stable level are when β = -2.5. Figure shows the EFC convergence result for the Monte Carlo calibration error trials compared to the ideal baseline. Each cycle of β produces an overall improvement in NI, but only after allowing iterations with a relaxed β and a temporary loss of performance. Compared to the strictly optimal β Monte Carlo trials in Figure , however, the overall convergence is significantly improved.

To test our hypothesis that NI increases for the optimal β estimated by the EFC model because of calibration error, we ran another simulation using the new β control strategy, but with no calibration error. As shown in Figure (6), the NI does not degrade for any iteration, except for a few iterations where actuator constraints come in to play. For iterations where β = -2.5, there is very little improvement because the less complex field modes have already been addressed, and the lack of nonlinear or model error in the iterations where optimal β is used does not increase the error for these lower order modes. For iterations using the optimal β estimated by the EFC model, NI improves at a rate similar to the ideal case. Since the only difference between this simulation and the Monte Carlo trials is the calibration error, the calibration error must be what causes the field of the testbed to degrade when the EFC model predicts an improvement. In summary, two simulations with the identical system, initial condition, and control strategy responded quite differently for less constrained regularization because of calibration error.

5. CONCLUSIONS

We define calibration error as the difference between the real coronagraph system and the model representation of the coronagraph used by the EFC control. The EFC model is used to create the Jacobian, as well as to evaluate estimates of coronagraph performance, such as for various regularization values. Sources of calibration error include measurement uncertainty of DM alignments, Lyot stop alignments, and system wavefront error. Since measurement uncertainty always exists, some degree of calibration error must exist. Error in the calculated Jacobian causes error in the calculation of the

$\underline{\delta DM}$ solution vector for the next iteration. This error in the solution vector is magnified when regularization is relaxed. However, if regularization is never relaxed, the fine detail of the residual electric field in the image can never be corrected.

We constructed a simulation where the coronagraph EFC control exactly replicates the control used for the testbed, including probing and measured alignments and wavefront error, but the physical coronagraph testbed is replaced by a numerical model containing Monte Carlo generated alignment and wavefront errors. The Monte Carlo generated errors represent the unknown calibration errors of the real coronagraph. We then demonstrated with the simulation that calibration errors slow or prevent EFC convergence. We then implemented an empirical EFC control strategy that modulates the regularization constraint to allow for relaxed regularization followed by strongly constrained regularization. The result of the new control strategy was a stair-step reduction in NI to a level below what was achieved otherwise. The new control strategy was also demonstrated on the HLC testbed in the JPL HCIT lab with positive results [1].

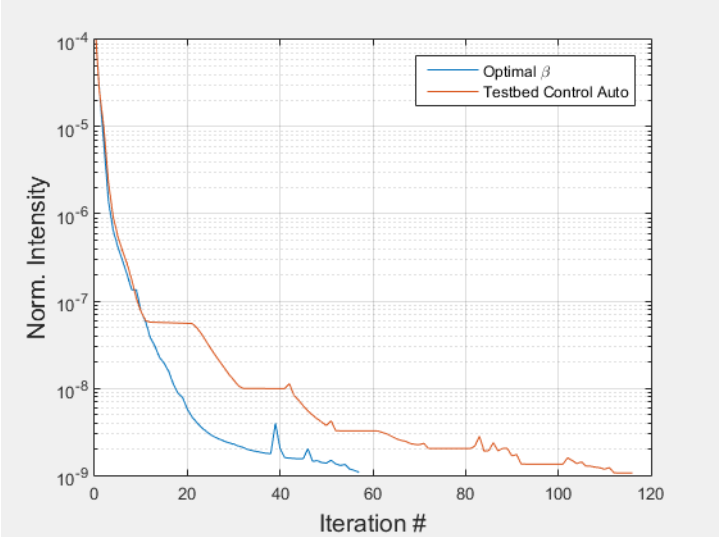

Figure 6. When the new regularization strategy is used for the case of no calibration error, the normalized intensity does not get worse for the iterations of optimal β. Furthermore, the iterations of optimal β are those where normalized intensity improves, and the iterations using β = -2.5 are those where normalized intensity does not change.

ACKNOWLEDGEMENT

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

REFERENCES

[1] Seo, B., Cady, E. J., Gordon, B., Kern, B. D., Marx, D. S., Moody, D., Muller, R., Patterson, K., Poberezhskiy, I., Shi, F., Sidick, E., Trauger, J. and Wilson, D., "Hybrid Lyot coronagraph for WFIRST: high-contrast broadband testbed demonstration," Proc. SPIE 10400, 10400-15 (2017).

[2] Cady, E. J., Kunjithapatham, B., Gersh-Range, J., Kasdin, J., Kern, B. D., Lam, R., Prada, C. M., Moody, D., Patterson, K., Poberezhskiy, I., Riggs, A. J. E., Seo, B., Shi, F., Tang, H., Trauger, J., Zhou, H., Zimmerman, N. T., "Shaped pupil coronagraphy for WFIRST: high-contrast broadband testbed demonstration," Proc. SPIE 10400-14 (2017).

[3] Give'on, A., Kern, B., Shaklan, S., Moody, D. C., Pueyo, L., "Broadband wavefront correction algorithm for high-contrast imaging systems," Proc. SPIE 6691, 66910A-1-11 (2007).

[4] Give'on, A., Kern, B. K., and Shaklan, S., "Pair-wise, deformable mirror, image plane-based diversity electric field estimation for high contrast coronagraphy, Proc. SPIE 8151, 815110-01-10 (2011).

[5] Fletcher, R., *Practical Methods of Optimization, Second Edition*, John Wiley & Sons, West Sussex, England, Ch. 6 (1987).

[6] Zhou, H., Nemati, B., Krist, J., Cady, E., Prada, C. M., Kern, B. and Poberezhskiy, I., "Closing the Contrast Gap Between Testbed and Model Prediction with WFIRST-CGI Shaped Pupil Coronagraph," Proc. SPIE 9904, 990419-1-11 (2016).

[7] Sidick, E., Seo, B., Kern, B. D., Marx, D. S., Poberezhisky, I. and Nemati, B., "Optimizing the regularization in broadband wavefront control algorithm for WFIRST coronagraph," Proc. SPIE 10400, 10400-74 (2017).

[8] Marx, D., and Kern, B., "Phase Retrieval Implementation for the WFIRST Coronagraph Development Testbed," OSA Imaging and Applied Optics Congress, CT4C.4, 2016.