Diffraction Modeling of Finite Subband EFC Probing on Dark Hole Contrast with WFIRST-CGI Shaped Pupil Coronagraph

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

ABSTRACT

The current NASA WFIRST (Wide-Field InfraRed Survey Telescope) coronagraph instrument (CGI) design allocates two subband filters per each full science band in order to contain system complexity and cost. We present our detailed investigation results on the adequacy of such a limited number of finite subband filters in achieving broadband dark hole contrast with a shaped pupil coronagraph (SPC). The study is based on diffraction propagation modeling with realistic WFIRST optics, where each subband's image plane electric field is estimated from pairwise pupil plane deformable mirror (DM) probing and image plane intensity averaging of the resulting fields of multiple (subband) wavelengths. Multiple subband choices and probing and control strategies are explored, including standard subband probing; mixed wavelength and/or weighted Jacobian matrix; subband probing with intensity subtraction; and extended subband probing with intensity subtraction. Overall, the investigation shows that the achievable contrast with a limited number of finite subbands is about 2~2.5x worse than the designed contrast for the current SPC. The result suggests that future shaped pupil designs should be optimized for slightly broader bandwidths than the intended science bandpasses if limited subbands are used for wavefront sensing via probing.

Key words: diffraction modeling, coronagraph, shaped pupil coronagraph, Electric Field Conjugation algorithm, wavefront sensing and control, EFC probing, WFIRST

1. INTRODUCTION

The shaped pupil coronagraph (SPC) and hybrid Lyot coronagraph (HLC) are the baseline techniques in the current WFIRST coronagraph instrument design [1-4]. The SPC operates in two modes: disk imaging mode, which has a 360° dark hole region around the star but with relatively large inner (IWA) and outer (OWA) working angles, and characterization mode for integral field spectroscopy (IFS), which has a bow-tie-shaped dark hole region but with smaller IWAs and OWAs. The current design's science bandpasses for the imager are 10% wide (the IFS bandwidth is 18%), too wide to sufficiently characterize the highly chromatic speckles as required to create a good dark hole. Multiple narrower bandpass (subband) filters are needed. However, to contain system complexity and cost, it is desirable to limit the number of components, so the current CGI design allocates only two 5% subband filters per full band.

The purpose of this study is to evaluate, through diffraction modeling of realistic WFIRST optics, the adequacy of such limited subbands in achieving the desired contrasts, and to investigate the tradeoff between the numbers of subbands, the best bandpasses for the filters, and the wavefront sensing strategy. This study is part of a chain of investigations designed to answer the question of how long it would take to produce a dark hole on-orbit [5].

While the HLC is more likely to ultimately define the number and types of subbands, we started our investigation with the less complicated SPC imaging configuration. The shaped pupil coronagraph uses a binary pupil mask, together with a focal plane mask and Lyot stop, to alter the star's diffraction pattern and achieve a desired dark hole at image plane. Its simple design renders it relatively insensitive to jitter, chromaticity, misalignment, etc.

Our study began with standard subband wavefront probing. We then explored the possibility of using mixed wavelengths and/or weighted wavefront control matrices. Next we investigated subband wavefront probing with intensity subtraction of the subband images from the broad one to create a virtual 3rd subband. And finally we looked into the possibility of using extended subband probing with intensity subtraction.

2. GENERAL CONFIGURATION AND EFC PROBING PROCEDURE

The modeling in this study used the WFIRST generation 2 shaped pupil coronagraph design (SPC 20141007) [6], together with the Cycle 5b CGI optics layout, which has about 31 surfaces (details can be found in [7]). This SPC design is intended for circumstellar disk imaging with a 360° dark hole field and working angles 6.5 - 20 λ/D over a 10% bandpass, centered at λ_o = 550 nm. The pupil mask and its companion Lyot stop are shown in Figure 1. The intermediate focal plane mask (not shown) is a simple annulus.

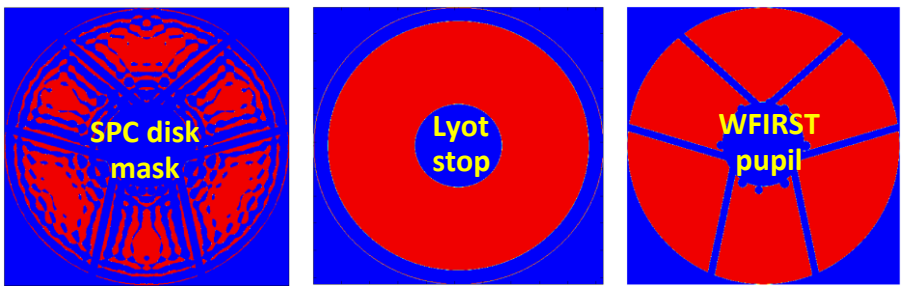


Figure 1. SPC disk pupil mask, Lyot Stop, and WFIRST pupil mask used

We used a full diffraction-based PROPER model [8, 9] to compute the SPC image plane speckle pattern. PROPER uses angular spectrum and Fresnel algorithms to propagate the beam from surface to surface according to the layout's physical distances, with the system represented as an unfolded linear sequence of optics. Realistic WFIRST optical aberrations, in the form of phase and amplitude errors, were included for each surface. No other types of aberrations were included. A broadband speckle pattern was generated by adding multiple monochromatic speckle patterns spanning the passband. Contrast (or more accurately, the normalized intensity) was evaluated over a 10% bandwidth with 5 wavelengths (centered at λ_o = 550nm), regardless of the probing subband choices used.

For wavefront sensing and control, the baseline method (used in simulation, on a testbed, or in flight) is to use pairs of positive and negative (pupil plane) deformable mirror (DM) patterns to sense (probe) the image plane electrical field at multiple wavelengths spanning the science bandpass. These are inputs to the Electric Field Conjugation (EFC) algorithm that determines the DM settings that minimize the image plane speckle pattern [10, 11]. The control matrix, typically termed *G* matrix or the Jacobian, is a linear approximation of image plane electric field response, *E*, to unit-strength single DM actuator pokes. It is commonly arranged by real and imaginary parts of the dark hole electric field, pixel by pixel, wavelength by wavelength for each actuator. Weighting (*W*) can be applied, either spatially or spectrally, or both [10, 11]:

$$G = \begin{bmatrix} \Re\{\mathrm{E}(\lambda_1, DH_1, DM_1)\cdot W(\lambda_1, DH_1)\} & \ldots & \Re\{\mathrm{E}(\lambda_1, DH_1, DM_m)\cdot W(\lambda_1, DH_1)\} \\ \Im\{\mathrm{E}(\lambda_1, DH_1, DM_1)\cdot W(\lambda_1, DH_1)\} & & \Im\{\mathrm{E}(\lambda_1, DH_1, DM_m)\cdot W(\lambda_1, DH_1)\} \\ \vdots & & \vdots \\ \Re\{\mathrm{E}(\lambda_k, DH_n, DM_1)\cdot W(\lambda_k, DH_n)\} & \ldots & \Re\{\mathrm{E}(\lambda_k, DH_n, DM_m)\cdot W(\lambda_k, DH_n)\} \\ \Im\{\mathrm{E}(\lambda_k, DH_n, DM_1)\cdot W(\lambda_k, DH_n)\} & & \Im\{\mathrm{E}(\lambda_k, DH_n, DM_m)\cdot W(\lambda_k, DH_n)\} \end{bmatrix} \quad (1)$$

A Tikhonov regularization factor, β, is added to damp the imperfect correction due to the linear approximation being used for an intrinsically nonlinear problem, imperfect sensing, and imperfect system calibration. One way of doing this can be:

$$G_{inv} = \left[G^T G + \beta \cdot \max\left(diag\left(G^T G \right) \right) \cdot I \right]^{-1} G^T \quad (2)$$

Pair-wise DM probing is used to sense the dark hole electric field [11]:

$$I_j^{+/-} = \left|E_{DH} \pm \Delta P_j\right|^2 + I_{inc} = \left|E_{DH}\right|^2 + \left|\Delta P_j\right|^2 \pm 2\Re\left\{E_{DH}^* \Delta P_j\right\} + I_{inc} \tag{3}$$

where ΔP_j is the probing field added to the existing to-be-measured DH field, E_{DH}, and I_j are the intensities measured as a result of jth probe, either positive (+) or negative (-). From Eq. (3):

$$\Delta I_j \equiv I_j^+ - I_j^- = 4\Re\left\{E_{DH}\Delta P_j\right\} \tag{4}$$

or in the form of a matrix for N pairs of probes:

$$\begin{bmatrix} \Delta I_1 \\ \vdots \\ \Delta I_N \end{bmatrix} = 2\begin{bmatrix} -\Im\{\Delta P_1\} & \Re\{\Delta P_1\} \\ \vdots & \vdots \\ -\Im\{\Delta P_N\} & \Re\{\Delta P_N\} \end{bmatrix}\begin{bmatrix} \Re\{E_{DH}\} \\ \Im\{E_{DH}\} \end{bmatrix} \tag{5}$$

Here ΔP_i are obtained by propagating the DM probe patterns through the coronagraph model. One can then find the speckle pattern, E_{DH}, for each pixel with N measured intensity pairs based on Eq. (5).

Once G is established and the dark hole field E_{DH} is sensed, the DM correction needed at each iteration is:

$$a = G_{inv} \cdot E_{DH} \tag{6}$$

The quality of the speckle field correction at each step depends on the accuracy of the control Jacobian matrix, G, and the accuracy of the sensed electric field, E_{DH}. In this paper our focus is related to the later (our other paper [12] focuses on the former).

The DM probing pattern is typically sinc-like; it adds a near-uniform phase (imaginary) offset on top of the existing field, or a real-value offset of opposite signs on opposing sides of the field. In our simulations the patterns were created by using the Fourier transform of the desired complex image plane offset pattern defined over a region slightly larger than the designed dark hole area. The center of the DM pattern location was shifted to avoid the telescope's obscurations. Additionally, the real-part probe pattern (the one that introduces opposite sign offsets on opposing sides of the image field) and location were rotated between even and odd iterations, and median filtering was applied for the estimated field; both of which were intended for reducing large errors at center seam which was caused by unmodulated field values along the line of opposing half fields. As was done for the full broadband image, the subband probe fields were simulated using fields generated at multiple wavelengths spanning each subband and averaged in intensity.

Detector noise was not added for probing on the basis that the dark hole acquisition star was assumed to be much brighter than the science target star. Two pairs of DM probes were used for each subband at each iteration. Heterodyning behavior of the complex field allows the probe modulation contrasts to be much brighter than the unprobed field contrast (e.g., 10^{-7} contrast probes can be used to accurately measure a 10^{-9} contrast field). The probe strength was chosen such that the resulting probe contrast was roughly 0.15~50x times the to-be-estimated field contrast of 3×10^{-6}~3×10^{-9} ,based on a simple, empirically-derived formula to scale DM probe strength (from a normalized 1nm max probe pattern):

$$dm_scale_factor = -50 * \log(field_contrast) \tag{7}$$

or loosely in terms of probe contrast:

$$probe_contrast \approx 10^{-7}\log\left(field_contrast\right) + 10^{-6} \tag{8}$$

In all of the cases described below, we started with a flat DM. We then computed the WFE at the exit pupil (equivalent to measuring the phase using phase retrieval) and set the DM to compensate (often termed as wavefront flattening). We then iteratively ran probing and EFC to create the dark hole.

The SPC design mean contrast over the dark hole region is ~3 x 10^{-9} with no aberrations or any wavefront control. Contrast was evaluated over a 10% bandwidth sampled using 5 wavelengths (λ_o = 550 nm), regardless the probing subband choices used.

3. FINITE SUBBAND PROBING

Our investigation began with standard (baseline) subband probing. Although 2 subbands are currently allocated, we included 3 and 4 subbands cases for establishing a reference for later comparison.

3A. Baseline subband probing

Figure 2 illustrates the wavelengths at which monochromatic fields were generated to comprise finite-width subband images for three cases of probing using 2, 3, and 4 subbands. Figure 3 shows the corresponding post-EFC dark hole contrasts along with a post-EFC field in which perfect knowledge was assumed (i.e., the computed, rather than probing-derived, E-fields were used).

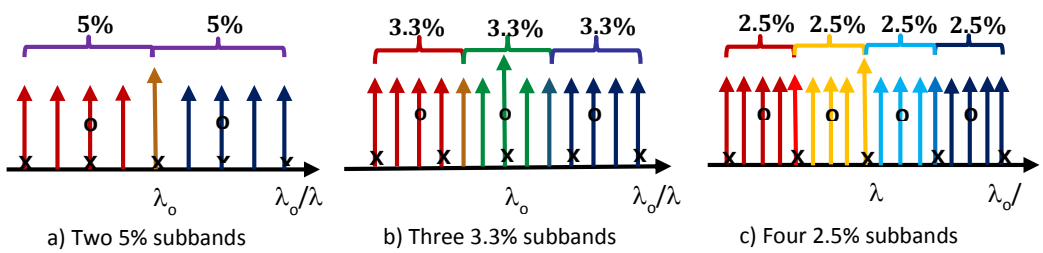


Figure 2. Wavelengths for 2, 3, and 4 subband probing. "o": subband central lambda (which are also G matrix lambda); "x": the contrast evaluation lambda

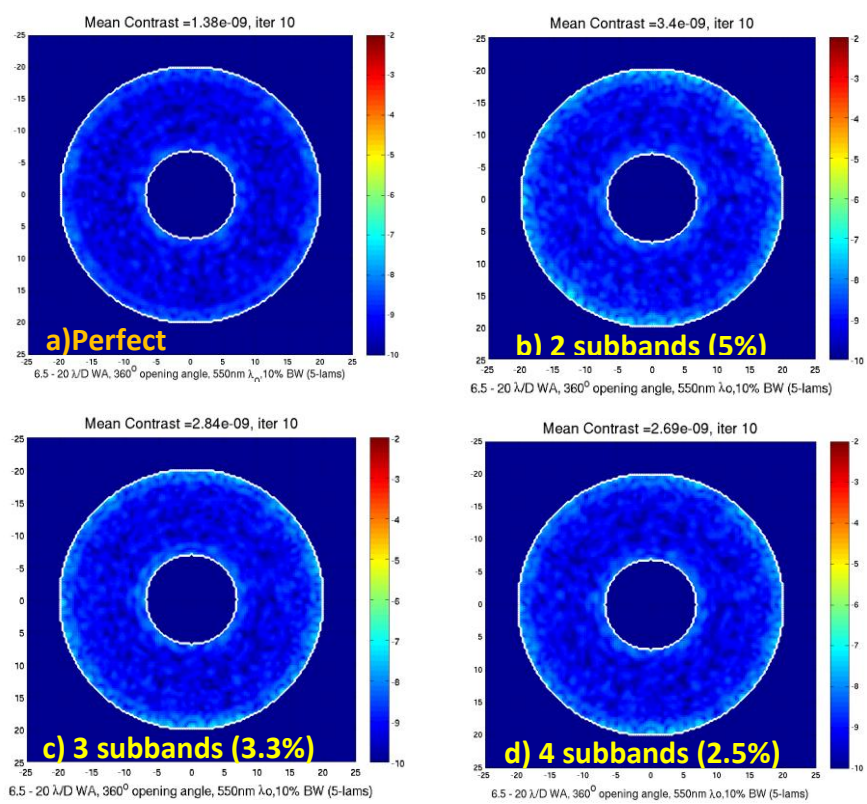


Figure 3. Contrast for perfect knowledge EFC and 2, 3, and 4 subband probing (b~d)

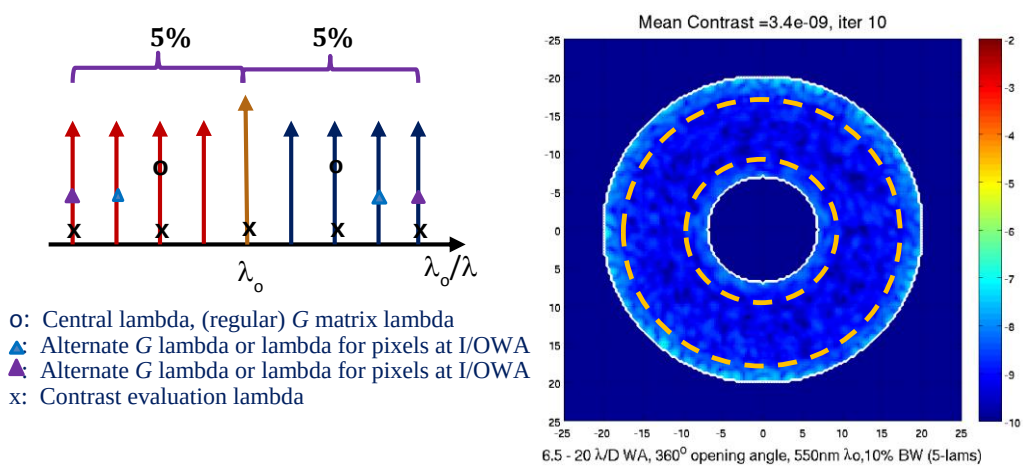
The results shown in Figure 3 indicate that there is only a small contrast gain from 2 to 3 or 4 subbands probing, whereas in all cases there is roughly a 2~2.5x degradation relative to using perfect knowledge. Notice that the degradations are mostly are at the outer and inner field edges.

3B. Subband probing with mixed wavelengths and/or weighted G matrix

The limited improvement in contrast from 2 to 4 subbands is not entirely surprising since the SPC has a fairly low chromaticity. It was assumed that the dark hole residuals along the field edges were due to field leakage by short wavelengths at OWA and by longer ones at IWA.

The control matrix, G, is composed of model-generated monochromatic field responses to individual DM actuator pokes for N_λ wavelengths. It is used in a linear equation (6) along with the measured E-fields that are derived from probing in N_λ finite-width subbands, and so represent the mean E-field over all wavelengths encompassed by each subband. An assertion must be made relating the wavelengths used to construct the G matrix to the representative wavelengths of the subbands (i.e., are the G matrix wavelengths equal to the central wavelengths of the subbands, or the shortest or longest wavelengths?). The chromaticity of the speckles over even a 5% bandpass may be sufficient to influence how the subband wavelength is related to the G matrix and the eventual ability to dig a broadband dark hole.

The first experiment we tried was to use mixed or alternative wavelengths in the *G* matrix and the E-field estimations. Specifically, for the alternate wavelength case, we generated two sets of G matrices: one with the central subband wavelength and one with an off-center wavelength, as illustrated in Figure 4), and alternated the direction of the offset between even and odd EFC iterations. The E-field estimation also alternated similarly between even and odd iterations. For the mixed wavelength case, we generated a G matrix using the central subband wavelength for most pixels, but used the short wavelength of the subband for the outer edge pixels in the short subband, and the long subband wavelength for inner edge pixels of the long subband.



a) b)

Figure 4. Wavelengths for alternate or mixed wavelength G matrix & E-field estimation for 2 subbands.

These modifications did not bring the desired improvement. The best result was an insignificant contrast improvement from the baseline probing's 3.36 x 10^{-9} to 3.1 x 10^{-9} with a mixed wavelength *G* of the central and end wavelengths. The reason was most likely due to the subband off-central wavelengths producing significantly larger field estimation errors than the central wavelengths, as illustrated in Figure 5. This was especially so at poor contrasts. Contrast converges slowly over EFC iterations and stalls using an off-center wavelength *G* in either alternate or mixed wavelength *G* Matrix from beginning. But even if one starts with a center-wavelength *G* matrix and then switches to a mixed wavelength *G* when contrast has improved (e.g., after reaching <10^{-8}), improvement is not guaranteed. We believe the error in the estimated off-central subband wavelength partially nullifies the potential benefit of knowledge at a more desired wavelength.

The second experiment was to weight the OWA/IWA pixels. We tried various combinations of 1.5x, 2x, and 0.5x weighting at the OWA and/or IWA edge pixels. The best case, though trivially so, used 0.5x at the OWA and IWA for the short subband, with a contrast of 2.81 x 10^{-9} achieved vs the previous 2.84 x 10^{-9} for 3 subbands. The reason was likely due to attenuation of the estimation error at the IWA and OWA edges, which tended to be larger than in the rest of the DH area. However the improvement was so minor, it is not worth adopting this complicated scheme.

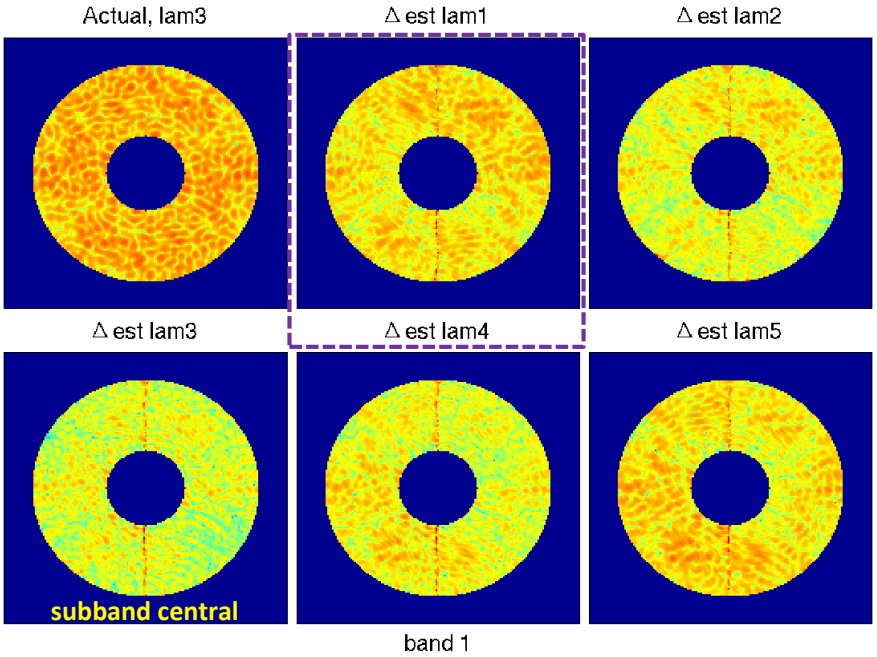


Figure 5. E field estimation error over wavelength (short subband, unfiltered)

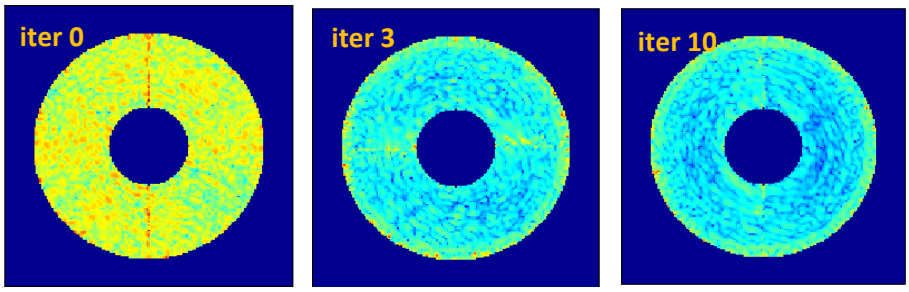


Figure 6. Field estimatin errors are larger at IWA/OWA edges, especially at later iteration

3C. Subband probing with intensity subtraction

While the limited contrast improvement gained by using more than two subbands does not seem to warrant the addition of more filters (at least in the case of the SPC evaluated here), it is possible to achieve the equivalent of 3 subbands using only 2 subband filters and the 10% science filter. The idea works like this: apply the DM probe and then measure the field intensity in two fairly narrow (2.5-3.3%) subbands located at the ends of the science filter's bandpass *and* in the full 10% science bandpass; subtract the two subband images from the full bandpass to get the intensity in a virtual third, central bandpass.

The results of doing this are shown in Figure 7 with almost the same contrasts as their baseline counterparts (Figure 3c and 3d). This means that it is feasible to use two subband filters plus one full band filter (2+1) to perform 3 subband probing (assuming perfect intensity subtraction).

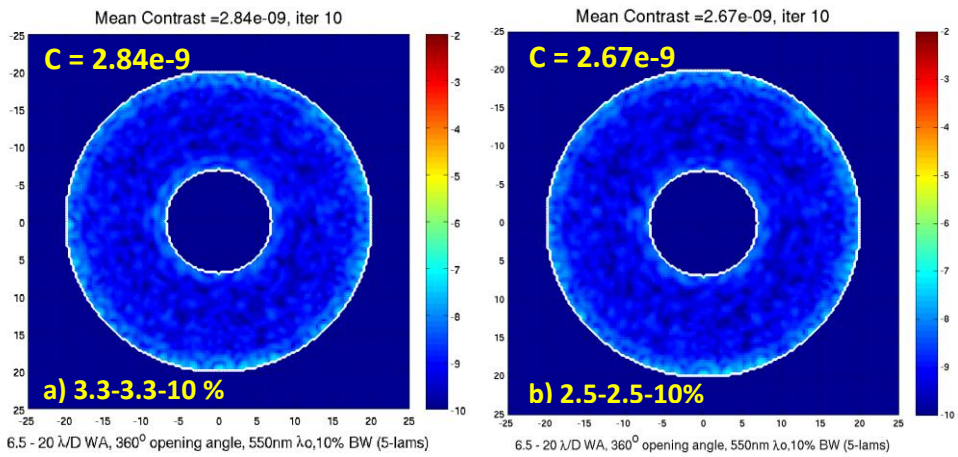


Figure 7. Contrasts of subband probing with intensity subtraction

3D. Extended subband probing with intensity subtraction

To see if the current achievable contrast was limited by probing and WFE estimation, we performed perfect knowledge EFC (no probing, computed E-fields used) at two sets of equivalent 3-subband wavelengths as illustrated in Figure 8a. In Figure 8b, their contrasts vs EFC iteration were plotted. Perfect electric field knowledge at the central wavelength of the 2+1 subbands , whether 3.3% (blue line) or 2.5% (green line), did not improve contrast - it actually resulted in slightly worse contrast than the probing version. Only the perfect electric field knowledge at the edge wavelengths plus the central one (solid triangle marked wavelengths choice in Fig.8a) provided the near-designed contrast (black line in Fig. 8b). This indicates that wavelength choice, rather than the probing itself, might be a limiting factor in achievable subband probing DH.

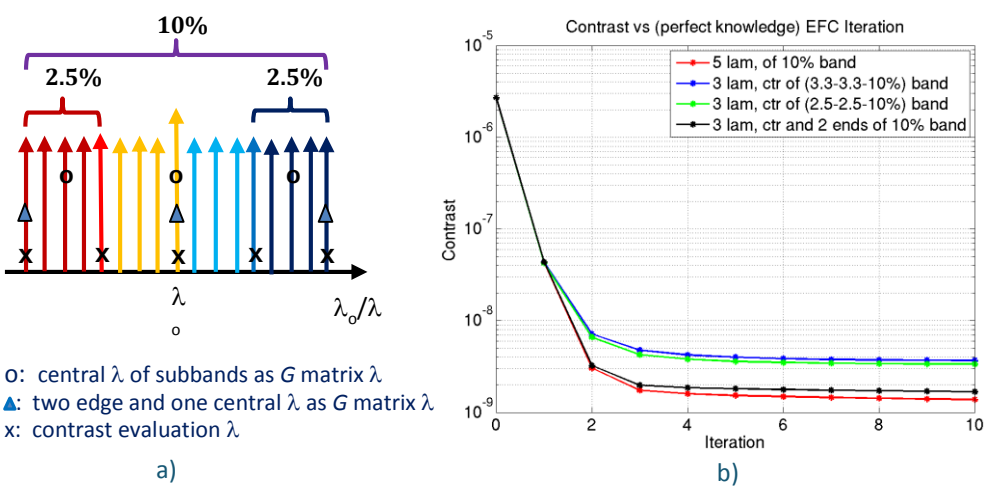


Figure 8. Contrasts of perfect knowledge EFC (no probing) with equivalent 3-subbnad wavelengths

Based on the above observations, our next study used extended subband probing, as illustrated in top of Figure 9. The procedure is similar to the regular subband probing with intensity subtraction as in sect 3C:

the middle band intensity is acquired using the 10% full band probing pattern intensity image, but subtract out half of the probing intensity in each end subband:

$$I_{mid} = I_{full} - 0.5*(I_{short} + I_{long}) \qquad (9)$$

The result, however, yielded poorer contrast compared with its standard counterpart. After careful examination, we believe the reason was poor field estimation due to the imbalanced intensity ratios among the subband wavelengths as shown in bottom part of Figure 9. The contrast at wavelengths beyond the designed bandwidth was significantly poorer than those within, especially at the OWA edges in the short extended subband.

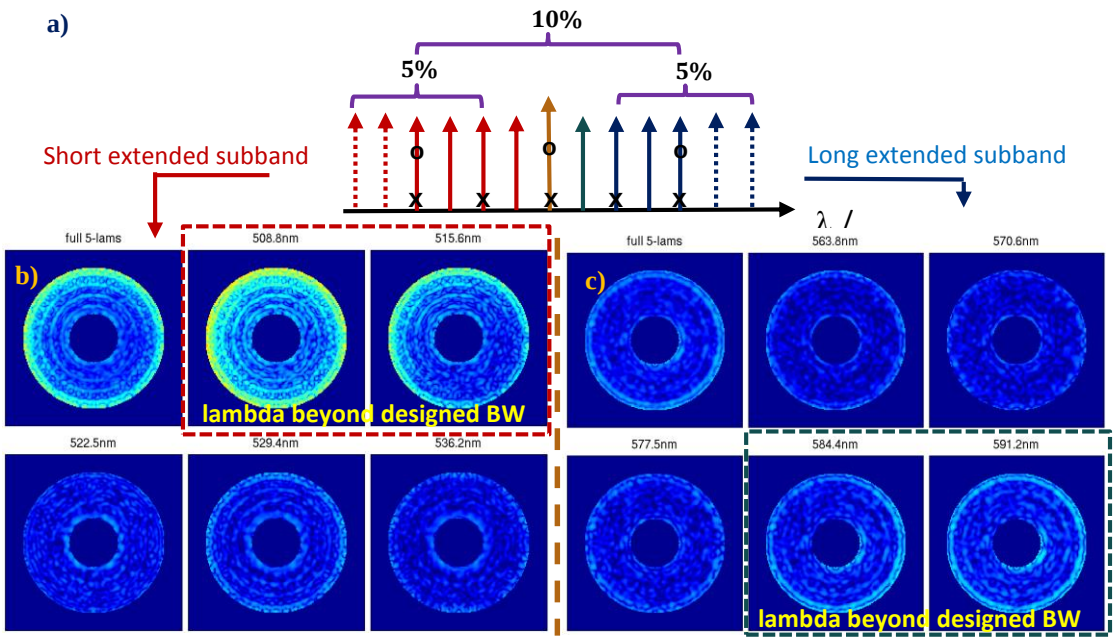


Figure 9. Top (a): Wavelengths in extended subband probing with intensity subtraction. Bottom (b and c): each subband's full speckle pattern and its wavelength components show the intensity imbalance between the inside and extended wavelength components.

4. DISCUSSION AND SUMMARY

We set out to evaluate the adequacy of WFE estimation using DM probing with a limited number of finite bandwidth filters to achieve the designed broadband coronagraphic contrast and to investigate the tradeoffs between the number of subbands, the best bandpasses for the filters, and the probing strategy. Beginning with standard subband probing, we explored the possibility of using a mixed wavelength and/or weighted G matrix, subband probing with intensity subtraction, and the possibility of using extended subband probing with intensity subtraction.

The results so far show that: 1) variations of subband probing strategy have mostly limited or insignificant improvement in contrast; 2) the mean contrast degradation (as compared with perfect electrical field knowledge EFC) is about 2.5x worse for 2 subband probing and about 2x worse for 3 or 2+1 (subtraction) subband probing; 3) the dark hole residuals are mostly at the OWA and IWA at the shortest wavelength. Overall, probing with two 3.3% subbands plus one 10% science band with intensity subtraction appears to be a good tradeoff, though probing with two 5% subbands is not much different. We summarize our results in Figures 10 and 11.

It is worth noting that the main reason for the limited improvement from the various probing strategies appears to be that the current SPM design has sharp contrast degradations immediately outside its designed bandwidth (especially at short wavelengths). If we narrow the evaluation bandwidth (i.e., from 10% bandwidth to 8%), the mean contrast would be much closer to the perfect knowledge case, as shown in Figure 12. Therefore, to limit contrast degradation when using only 2 subbands or 2+1 subbands with intensity subtraction, it would be desirable for the shaped pupil to be designed for a slightly broader bandwidth than its intended science bandpass, e.g., by 1% on each side for a 10% science bandwidth.

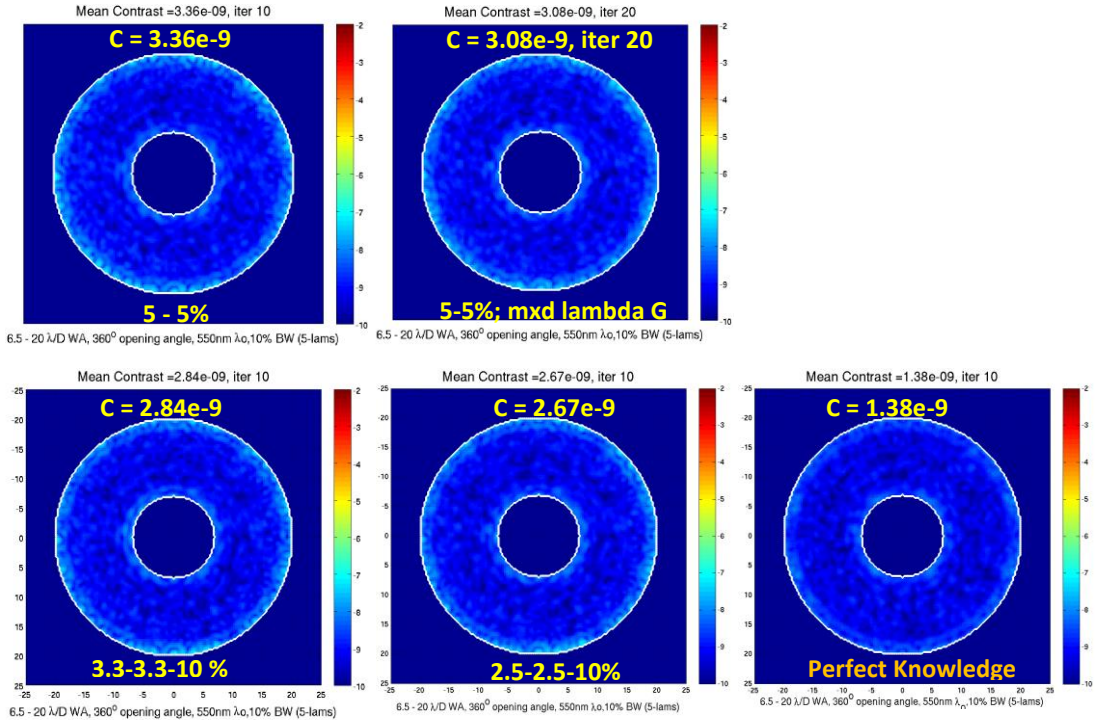


Figure 10. Recap of achieved final contrasts with various subband probing schemes.

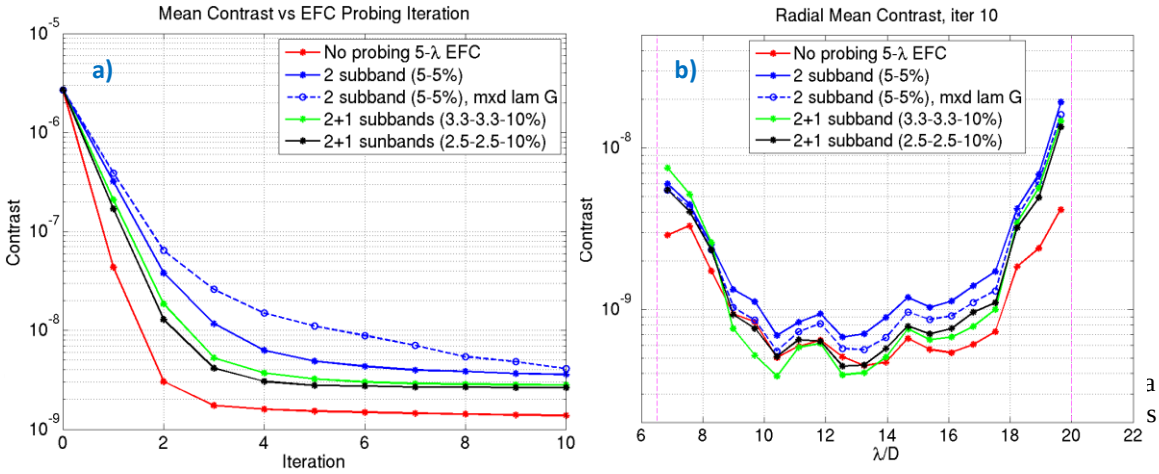


Figure 11. (a) Contrast convergence versus EFC iteration for various subband probing schemes. (b) Corresponding radial mean contrast at the end of convergence

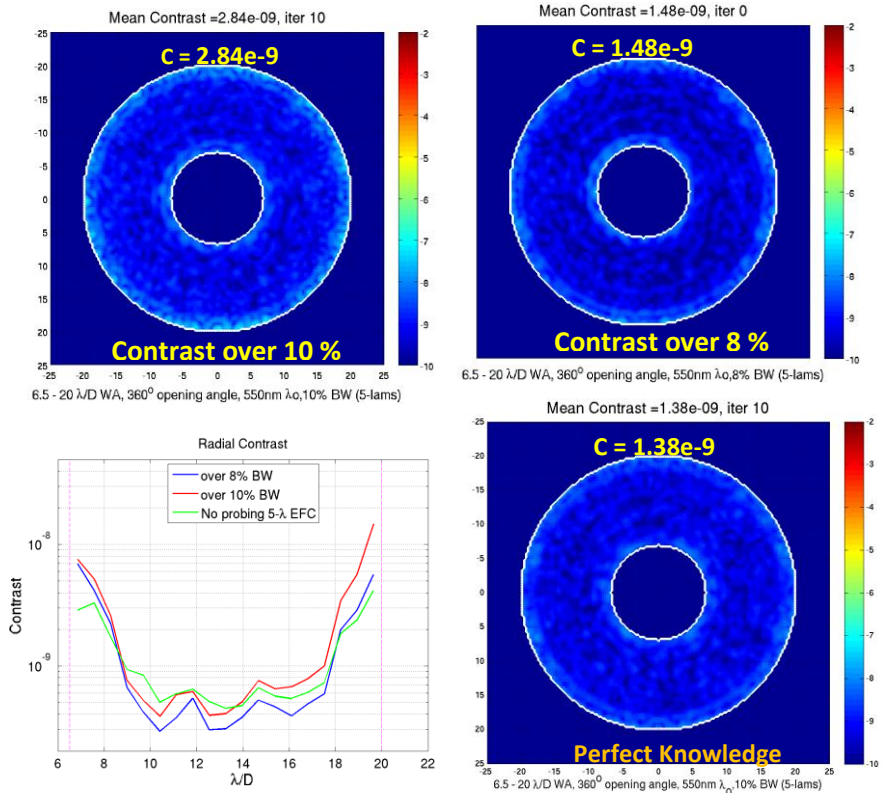


Figure 12. (a) Contrast over full science bandwidth (10%), and (b) over narrower (8%) bandwidth. Both with 3.3%-3.3. %-10% subband probing scheme.

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. © 2016. California Institute of Technology. Government sponsorship acknowledged.

REFERENCES

[1] Poberezhsky, I. et al., "WFIRST-AFTA coronagraph starlight suppression technology progress," Proc. SPIE. 9904, (2016)

[2] Cady, E., et al., "Demonstration of broadband wavefront control with a shaped pupil Lyot coronagraph for the WFIRST-AFTA Observatory," Proc. SPIE. 9904, (2016)

[3] Seo, B., et al, "Hybrid Lyot Coronagraph Testbed Demonstration for WFIRST-AFTA," Proc. SPIE. 9904, (2016)

[4] D. Spergel, et al., "Wide-Field InfraRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA. Final Report," https://arxiv.org/abs/1503.03757 (2015).

[5] J. Krist, et al., "An overview of WFIRST-AFTA coronagraph modeling," Proc. SPIE, 9605, 960505-1 (2015).

[6] Zimmerman, N. T., et al., "Shaped Pupil Lyot coronagraph Designs for WFIRST-AFTA," Proc. SPIE, 9605, 96050A-1 (2015)

[7] Tang, H., et al, "The WFIRST/AFTA Coronagraph Instrument Optical Design," Proc. SPIE. 9605, 960504-1 (2015)

[8] Krist, J., et al., “Numerical modeling of the proposed WFIRST-AFTA coronagraphs and their predicted performance,” JATIS 2(1), (2016)
[9] Krist, J., “PROPER: An optical modeling program for IDL,” Proc. SPIE. 6675, 66750P (2007).
[10] Give’on, A., et al., “Broadband wavefront correction algorithm for high contrast imaging systems,” Proc. SPIE. 6691, 66910A (2007).
[11] Give’on, A., et al., “Pair-wise, deformable mirror, image plane-based diversity electric field estimation for high contrast coronagraph,” Proc. SPIE. 8151, 815110 (2011).
[12] Zhou, H., et al, “Closing the Contrast Gap between Testbed and Model Prediction with WFIRST-CGI Shaped Pupil Coronagraph,” Proc. SPIE. 9904, (2016)