Jet Propulsion Laboratory
California Institute of Technology

JPL Document D- 74425

TECHNOLOGY DEVELOPMENT FOR EXOPLANET MISSIONS

Assessing the performance limits of internal coronagraphs through end-to-end modeling

Technology Milestone #1 Results Report

John Krist, PI
Ruslan Belikov
Dimitri Mawet
Dwight Moody
Laurent Pueyo
Stuart Shaklan
John Trauger

7 December 2012, corrected 25 June 2013

Approved for public distribution

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.

© 2012 All rights reserved.

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

Approvals

Released by

John Krist
Principal Investigator

26 Oct 2012
Date

Approved by

Peter Lawson
Exoplanet Exploration Program Chief Technologist, JPL

31 Oct. 2012
Date

Marie Levine
Exoplanet Exploration Program Technology Manager, JPL

10/31/12
Date

Peter Lawson
Exoplanet Exploration Program Manager (Acting), JPL

31 Oct. 2012
Date

Douglas Hudgins
Exoplanet Exploration Program Scientist, NASA HQ

8 Nov. 2012
Date

Michael Moore
Exoplanet Exploration Program Executive, NASA HQ

8/11/12
Date

Assessing the performance limits of internal coronagraphs

Table of Contents

1. Objective **1**
2. Introduction **1**
Exoplanet imaging **1**
Coronagraphs **2**
2.1.1 Phase induced amplitude apodization (PIAA) coronagraph 2
2.1.2 Vector vortex coronagraph (VVC) 4
2.1.3 Hybrid band-limited coronagraph (HBLC) 5
Coronagraph numerical models **6**
Goals of this study **8**
Application to future NASA missions **8**
Modeling framework **9**
Caveats **9**
3. Milestone #1 Description **10**
Milestone Prerequisites **10**
3.1.1 Coronagraph contrast and image plane field dimensions 10
3.1.2 Optical system layout 10
3.1.3 PIAA 10
3.1.4 VVC 11
3.1.5 HBLC 11
Milestone Requirements **11**
3.1.6 Algorithm efficiency 11
3.1.7 Algorithm accuracy 12
Milestone Metrics **13**
3.1.8 Algorithm efficiency 13
3.1.9 Algorithm accuracy 14
4. Success Criteria **15**
5. Milestone #1 Results Summary **16**
6. Milestone Certification Data Package **18**
7. References **19**
Appendix A: Optical system layout **22**
Appendix B: Verification of PROPER **26**
Introduction to PROPER **26**
Verifying PROPER accuracy **27**
Appendix C: Phase-Induced Amplitude Apodization (PIAA) **29**
Introduction **29**
PIAA Design **30**
PIAA Optical Modeling **31**
S-Huygens Propagation **32**
Verification of S-Huygens Against Rayleigh-Sommerfeld **34**
PIAA Angular Spectrum Propagator (PASP) **37**
Post-Apodizer design and implementation **42**
Effect of edge weighting **50**
PIAA Milestone #1: Accuracy of the PASP algorithm **52**
Accuracy of pupil remapping without explicit propagation **63**
PIAA Milestone #1: Efficiency of the PASP algorithm **63**
Impact of not meeting the combined PIAA milestones **67**

Appendix D: Vector Vortex Coronagraph (VVC) 69
The Vector Vortex Coronagraph Unique Characteristics 69
Accurate propagator for the Vector Vortex Coronagraph 69
Sensitivity to sampling at the focal plane mask 73
VVC Milestone #1: Efficiency and Accuracy of the Vortex Representation 75
Evaluation of the leakage term 78
Appendix E: Hybrid Band-Limited Coronagraph (HBLC) 83
HBLC overview 83
Optimizing the HBLC components 83
Redefining the HBLC design solution: DM map vs. actuator settings 85
Re-optimizing the HBLC design 87
More actuators, smaller dark hole = darker dark hole 88
Implementation of the HBLC 90
HBLC Milestone #1: Efficiency and accuracy 90
Appendix F. Acronyms 94

TDEM Milestone #1 White Paper: Assessing the Performance Limits of Internal Coronagraphs Through End-to-End Modeling

1. Objective

In support of NASA's Exoplanet Exploration Program and the ROSES Technology Development for Exoplanet Missions (TDEM) study program, this report describes the results from the TDEM Milestone #1 for developing and verifying modeling algorithms for selected internal coronagraphs.

This study sets the foundation for Milestone #2, which is an assessment of the theoretical performance limits of the selected coronagraphs in a system with realistic optical aberrations as derived through end-to-end modeling. That milestone uses the algorithms developed and verified in Milestone #1 to characterize the wavefront control behavior of each coronagraph and identify the limiting factors for achieving 10^{-10} contrast over a broad bandpass for future coronagraphic missions. Milestone #2 is described in a separate report.

2. Introduction

The technology milestone described here serves to gauge the progress of optical modeling for a space-based coronagraphic mission such as ACCESS (Trauger et al. 2008) or the Terrestrial Planet Finder Coronagraph (TPF-C; Traub et al. 2006) that would detect and characterize exoplanets.

Exoplanet imaging

A mission like ACCESS or TPF-C must be able to detect a planet whose apparent brightness relative to the star is on the order of 10^{-9} (Jupiter-like) to 10^{-10} (Earth-like) at visible wavelengths. For even a nearby system such a planet would be seen at a small angle from the star. Diffracted and scattered light from best space telescope would overwhelm the feeble signal from the planet at these small separations. The diffracted light can be significantly reduced using a coronagraph, which is an optical component that specifically filters the wavefront to remove the light from the central source (star). However, unavoidable optical fabrication errors (e.g. imperfect polishing and non-uniform coatings) will scatter light that the coronagraph cannot suppress. Wavefront control using one or more deformable mirrors can largely compensate for these errors.

To enable observations of exoplanets the coronagraph and the wavefront control system operate together to create a ***dark hole*** centered on the image of the star. The dark hole usually begins at an inner radius where, by convention, the intensity transmission of the coronagraphic occulter is 50%. This is the ***inner working angle*** (IWA). Depending on the coronagraph design, the practical limit to the IWA is between 2 λ/D to 4 λ/D radians (D is the telescope diameter; hereafter, angles expressed as multiples of λ/D will be in radians). The ***outer working angle*** (OWA) is the outer radius of the dark hole and is limited by the number of deformable mirror (DM) actuators, *N*, that map across the pupil diameter. The maximum dark hole radius is (*N*/2) *λ*/D at the shortest wavelength in the passband (though best contrast is achieved with an OWA smaller than this).

The light level in the dark hole is defined in terms of its ***contrast***. The broadband contrast is the ratio of the average (across the bandpass) scattered starlight level in the dark hole to the average (again over the bandpass) peak light level of an image of the star when the coronagraph mask (which blocks the starlight at the image plane) is removed. In other words, it is the brightness of a field point source, relative to the star, whose peak is equal in intensity to the mean dark hole brightness. NOTE: The peak light level of a source varies near the inner working angle of the coronagraph due to modulation by the mask or occulter. In Milestone #1, this variation is not included in the calculations; it increases the contrast by 2× at the IWA but quickly levels beyond that as the mask transmission increases to its maximum. The small effected region does not significantly impact the reported full-dark-hole contrast values. In Milestone #2 this effect is taken into account.

Coronagraphs

There are a wide variety of coronagraphs. Three of the most studied of those proposed for space missions are the hybrid band-limited coronagraph (HBLC; Moody et al. 2008), the vector vortex coronagraph (VVC; Mawet et al. 2009a), and the phase-induced amplitude apodization (PIAA; Guyon et al. 2005) coronagraph. To suppress diffraction, each modifies the wavefront in a different manner: in the focal plane, the VVC primarily affects phase (via polarization manipulation) and the HBLC alters amplitude and phase, while PIAA geometrically remaps the wavefront at the pupil to create an apodized beam. All three have been tested in the High Contrast Imaging Testbed (HCIT; Trauger et al. 2007) at the Jet Propulsion Laboratory down to contrast levels (monochromatic light, one-half dark hole, one deformable mirror) of 5×10^{-10} (HBLC with a linear occulter), 3.5×10^{-9} (VVC; Mawet 2011a, 2011b), and 4×10^{-7} (PIAA, with the deformable mirror after the beam is remapped; Kern et al. 2009, Belikov et al. 2009). Given that these methods alter the wavefront in different ways, one may expect that they have different sensitivities to the aberrations present in any optical system.

2.1.1 Phase induced amplitude apodization (PIAA) coronagraph

Apodization of a telescope pupil will produce a point spread function with significantly reduced wing intensity, making it a technique of interest for high contrast imaging. Conventional apodizers using transmission-altering masks reduce throughput too much to be acceptable for exoplanet imaging, and they would be difficult to accurately fabricate

and operate over broad bandpasses. As an alternative method of apodization, PIAA utilizes two optics (lenses or mirrors) that geometrically distort the wavefront (Figure 1). The first optic, M1, is located at a pupil and remaps the beam onto the second optic, M2, which corrects for the phase errors from the remapping to create a "flat" wavefront (Figure 2).

Pure PIAA systems that provide 10^{-10} contrast are impractical due to diffraction effects and stringent surface tolerances. These can be reduced in a hybrid system using a weak apodizer (*post-apodizer*) placed near or at a subsequent image of M2 (Pluzhnik et al. 2006a, 2006b). In practice, post-apodizers are implemented using binary transmission masks composed of a series of narrow, opaque rings spaced and sized to provide the desired diffraction reduction. These have been fabricated using lithographic techniques (Kern et al. 2009).

After the post-apodizer the beam is brought to a focus where an occulting spot masks the central lobe of the apodized stellar point spread function (PSF). At this point any field sources are highly distorted and blurred. To restore the original wavefront mapping and thus image quality, the beam is fed through another set of PIAA optics, but in reverse order (no apodizer is required for the reverse system).

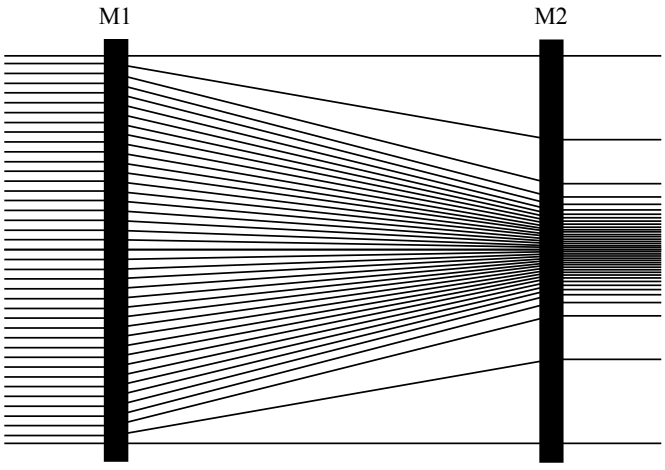

Figure 1. Schematic ray diagram showing how the PIAA M1 optic creates an apodized beam. The beam remains collimated at the very edge during propagation, but towards the center it becomes compressed. The PIAA M2 optic corrects for phase distortions introduced by the remapping.

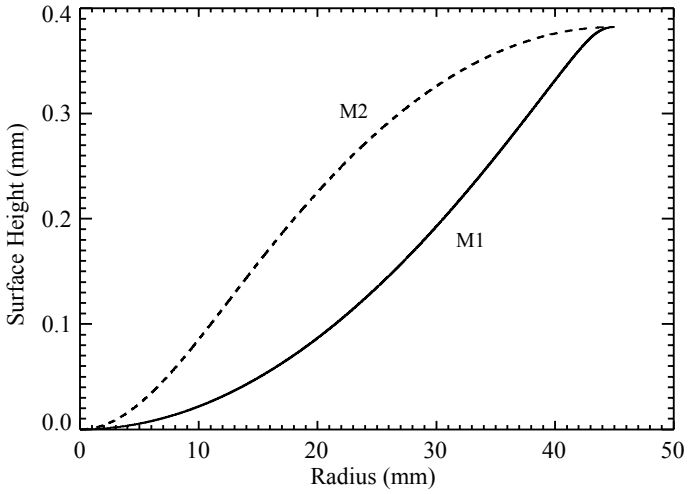

Figure 2. Surface height profiles of a set of PIAA optics (90 mm diameter, 900 mm separation). Note that the Y axis is magnified compared to the X axis.

2.1.2 Vector vortex coronagraph (VVC)

An optical vortex is created using a mask that introduces an azimuthally-varying phase shift to the wavefront, creating a "phase screw" with a singularity at the center. This results in self-interference as the wavefront propagates, forming a dark central hole in the beam at some distance from the mask. An optical vortex can be used in a Lyot coronagraph by placing a vortex-generating mask at an intermediate image plane that results in a dark central hole at a subsequent pupil plane, where a Lyot stop is placed. The advantages of an optical vortex coronagraph are that it can provide imaging very close to the star (< 2 λ/D) and has high throughput (>90%) due to a relatively wide-open Lyot stop.

Previous vortex masks were transmissive, stepped spiral patterns etched into substrates. They suffered from manufacturing defects due to the imperfect singularity at the center of the spiral and the small, nanometer-scale steps required to produce an approximation to a smooth phase ramp. They were also inherently chromatic.

A new technique (Mawet et al. 2009a) has been developed that creates a "geometrical" phase spiral by manipulating the polarization of incoming light with novel coatings made of hardened birefringent liquid crystal polymers (LCPs). This *vector vortex* is nothing more than a rotationally symmetric halfwave plate (HWP) providing a geometrical phase shift that applies opposite phase screws to the two orthogonal circular polarization states (Figure 3). In the vector vortex, for a linearly polarized input field (or for natural light projected onto a linear basis), the rotationally symmetric HWP rotates the polarization vector as in Figure 3a. The definition of circular polarization is a linear polarization rotating at the angular frequency ω (equal to that of the electromagnetic field); a rotation $\varphi = 2\theta$ of the polarization vector is strictly equivalent to a phase delay (Figure 3b). If, at any given point in space, the polarization vector is rotated such as in Figure 3a, it implies that the given circular polarization (Figure 3b) has acquired a geometric phase ramp $e^{i\varphi} = e^{i2\theta}$ such as that represented in Figure 3c. The term φ thus represents both an angle and a phase – hence the term "geometrical" phase.

The details of this technique are given by Mawet et al. (2009a). This form, called the vector vortex coronagraph (VVC), has many advantages over the physical vortex mask. Very smooth and accurate phase ramps can be created, and with multilayer coatings it can be made fairly achromatic. It does, however, require a small circular mask ($r < 0.5$ λ/D) to cover manufacturing errors at the singularity. Depending on the *charge* of the vortex (the number of waves the phase ramp goes through - 4 to 6 is practical), the inner working angle is 1.8 λ/D – 2.4 λ/D.

At high contrasts, it is not possible to create a dark hole in a VVC system simultaneously for both orthogonal polarizations. After passing through the vortex mask, the beam is circularly polarized in one direction. Aberrations on optics downstream of the mask will pickup the same polarization direction, and thus can be corrected only for that polarization. As reported in section 8.7.2.4.1 of the ACCESS report (Trauger et al. 2008), the dark hole contrast for one polarization can be ~10^{-10} while using the same deformable mirror solution for the other one the contrast is ~10^{-8}.

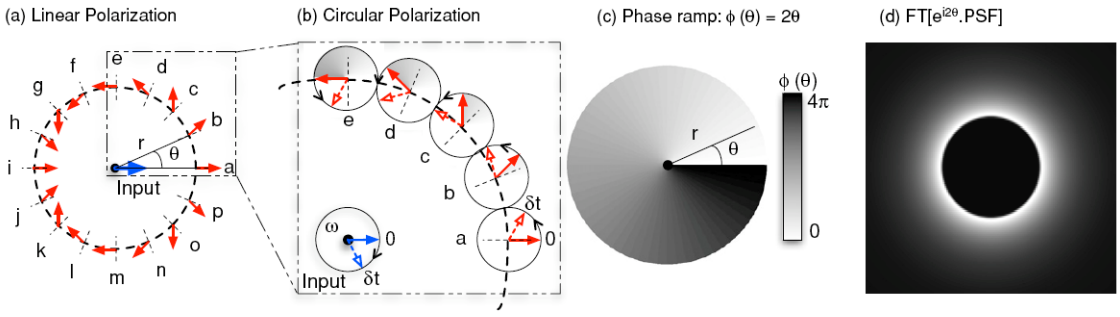

Figure 3. The VVC azimuthal phase ramp. Panel A: rotationally symmetric HWP with an optical axis orientation that rotates about the center (dashed lines perpendicular to the circumference). The net effect of a HWP on a linear impinging polarization is to rotate it by $-2 \times \alpha$ where α is the angle between the incoming polarization direction and the fast optical axis. An incoming horizontal polarization (blue arrow) is transformed by the vector vortex so that it spins around its center twice as fast as the azimuthal coordinate θ (red arrows). Panel B: for circular polarization, the output field rotation is strictly equivalent to a phase delay (the starting angle 0 is rotated; therefore phase shifted). The angle of local rotation of the polarization vector corresponds to a "geometrical" phase: upon a complete rotation about the center of the rotationally symmetric HWP, it has undergone a total $2\times2\pi$ phase ramp, which corresponds to the definition of an optical vortex of topological charge 2 (panel C). Upon propagation from the focal plane to the subsequent pupil plane, the Fourier transform (FT) of the product of the PSF by the azimuthal phase ramp sends the light outside the original pupil area (Panel D).

2.1.3 Hybrid band-limited coronagraph (HBLC)

The HBLC is a variation of the classical Lyot coronagraph having an occulting mask at a focus that blocks the central portion of the stellar point spread function. A graded transmission mask is used for very high contrast imaging as it provides for better diffraction suppression than a solid occulter (Kuchner & Traub 2002). At a subsequent

image of the entrance pupil a simple aperture mask, the Lyot stop, blocks light along the outer edge of the beam. The stop diameter is sized specifically to the occulter; the smaller the occulter is, the smaller the Lyot stop clear aperture is.

A graded-transmission occulter can be created using a variable-thickness layer of a metal, and such occulters have been used successfully in the HCIT (Trauger et al. 2007). This coating introduces wavelength-dependent transmission and phase shifts that are related to the layer thickness and material properties. Careful selection of materials can reduce these effects; nickel, for instance, offers a favorable relationship between refractive indices and wavelength (Balasubramanian 2005), providing useful and nearly constant optical depths across the visible spectrum. To partly compensate for the phase dispersion over a broad wavelength range, multilayer variable-thickness dielectric coatings are deposited on top of the transmission pattern or on the back side of the substrate (Figure 4).

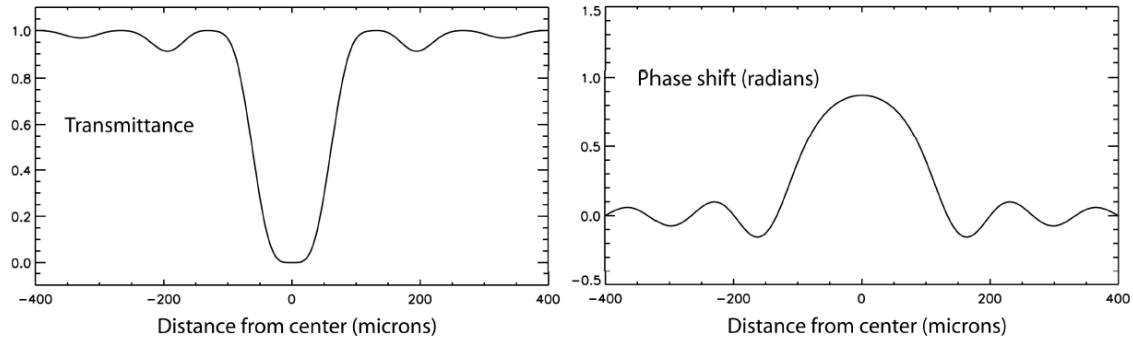

Figure 4. Attenuation and phase shift profiles for a 4th order linear occulter HBLC mask with an inner working angle of 3 λ/D. From Moody et al. (2008).

With realistic materials and number of coating layers, the HBLC mask and Lyot stop cannot achieve a band-limited, 20% broadband contrast of 10^{-10} without additional wavefront modification provided by deformable mirrors. The DMs, focal plane mask, and Lyot stop are all considered as part of the coronagraph. The DM settings required to achieve the required performance are derived using an iterative wavefront control algorithm in an unaberrated, simulated layout.

Coronagraph numerical models

Accurate wavefront propagation methods are needed to derive the behavior and performance of a coronagraph in a realistic system (Krist et al. 2006, 2009). They can also be used to investigate data analysis techniques prior to obtaining real data (Krist et al. 2008). Purely analytical models may sometimes be useful for initial performance estimates in the absence of aberrations or other system defects, but they are usually impractical for predicting what happens for realistic systems. For those cases, accurate numerical propagation modeling techniques are required to predict real-world performance.

In addition to performance predictions, propagation models are needed to build the deformable mirror response matrix that is used for wavefront control (Give'on et al. 2007). This contains the predicted changes in the electric field in the image plane produced by motion of each DM actuator. The wavefront control algorithm uses it to determine the DM settings that minimize the starlight inside the dark hole. It is generated by applying a small piston to each actuator in a system model and then computing the resulting field. Using two 48×48 actuator DMs with five sensing wavelengths (to provide a broadband solution), the number of image plane fields that must be computed is over 18,000. Even when the problem is reduced in number by invoking symmetry (if possible), the number of full system propagations is over 2,000. Rigorous diffraction algorithms such as Rayleigh-Sommerfeld take days to propagate just one wavefront, and even if they took just an hour, it would still take over three months to generate the DM response matrix. Every time the system design is altered, the matrix must be recomputed. Thus, an *accurate* and *efficient (fast)* propagator is needed for operation of the coronagraph, whether in simulation or the real world. A mission waiting for an updated response matrix would be wasting science time.

The accuracy of a propagation algorithm must be verified. This can be done either by comparing the models against results from actual hardware in a testbed or against more rigorous, physics-based algorithms or analytical representations. Of the three coronagraphs discussed here, only the band-limited coronagraph (with a linear rather than circular occulter) has achieved dark hole contrasts of 10^{-9} or better in a testbed. The VVC and PIAA have to date reached, monochromatically, only the 10^{-7} - 10^{-8} level. Thus, current testbed results cannot be used to verify new modeling techniques to the 10^{-10}-10^{-12} contrast levels required for extrasolar Earth-type planet observations.

Current modeling techniques for these coronagraphs vary in terms of efficiency, predicted accuracy, and verified accuracy. The HBLC is a modification of the amplitude-only bandlimited masks that have been evaluated in the HCIT over the past few years, including for TPF-C Milestones 1 and 2 (and the current TPF-C Milestone 3 being done under the TDEM program). Most of the experiments, and indeed the design of the HBLC itself, have been driven by numerical simulation (Trauger et al. 2007; Moody et al. 2008). The proprietary modeling software used for those, written by Dwight Moody at JPL, implements the angular spectrum and Fresnel algorithms. The model predictions agree well with the testbed results at the 10^{-9} – 10^{-10} levels achieved on HCIT. PROPER (Krist 2007), a library of optical propagation routines that is used in this study (Section 2.6), implements the same algorithms (some results have been successfully compared between the two packages; note that Moody's software is not documented and not publicly available; it and PROPER have the same capabilities, with no intrinsic support for any of the coronagraphs, but PROPER is used in this study because it is publicly available and documented). Prior to this study, verification against rigorous methods had not been done to the 10^{-10} level. The VVC is a relatively recent development and has achieved ~10^{-9} contrast on HCIT. The VVC was developed and its performance predicted to that level using PROPER (Mawet et al. 2009b). However, the models have not been verified in any manner (either using testbed or reference algorithms) down to 10^{-10} contrast levels, nor has their speed been established. Accurate models for PIAA have been developed using the S-Huygens algorithm (Vanderbei 2006; Belikov et al. 2006), but that method is too slow to be practical for repeated end-to-end modeling or DM response matrix generation

(which would take over a week on a current workstation for the matrix size proposed here). In summary, there is a need for efficient algorithms whose accuracies have been verified and documented against more rigorous methods.

This TDEM study verifies the efficient model accuracies against physics-based algorithms. Validation of the models (be they the efficient or rigorous ones used here) against results from actual coronagraphic hardware is required to ensure that all important properties of the system have been included. Such validation is beyond the scope of this study. The NASA Exoplanet Exploration program has a long-term goal to validate coronagraphic models against testbed results. For instance, Exoplanet Exploration Technology Milestone #3A (Shaklan 2009) has been defined to validate bandlimited coronagraph models against results in the HCIT for 10^{-9} contrast fields over a broad bandpass. This is the subject of another TDEM study.

Until such time that testbed validation can be done to the contrast levels of concern for a particular coronagraph, verifications of practical, efficient algorithms against more rigorous ones are the only means of establishing their reliability. In the end, these new algorithms are required to implement wavefront control on the testbeds that will be used for validation; the models will define the expected performance against which the testbed results will be judged (i.e., a testbed experiment may not achieve the performance goal set by the simulations due to misalignments or fabrication errors not included in the models). In past experience, only when a modeling algorithm has not been verified against a more thorough method will it poorly match the performance of the real system. For example, initial calculations of the performance of a pure PIAA system were done using geometrical optics, with contrast performances of $10^{-9} – 10^{-10}$ predicted; however, using more detailed modeling Vanderbei (2006) showed that diffraction effects would limit contrasts to $10^{-5} – 10^{-7}$.

Goals of this study

In this study we develop accurate and efficient numerical propagation algorithms for the three selected coronagraphs that are verified against rigorous analytical or numerical models. These algorithms may be used to evaluate the performance of these systems and to generate system response matrices used for wavefront control of simulated and real coronagraphs. The accuracies determined in this study can be used to assess the numerical errors on the predicted tolerances derived from future modeling efforts. These goals take the form of milestones with defined merit criteria, as described in Sections 3 and 4. The algorithms developed for each coronagraph meet the efficiency and accuracy requirements in those milestones.

Application to future NASA missions

Any future mission that uses one of these coronagraphs will require the fast and accurate models developed in this study to

- Determine the performance of the coronagraph in real-world conditions
- Plan testbeds used to evaluate prototypes and perhaps test flight coronagraphs

- Generate the DM response matrix that is used on-orbit for determining the DM settings that produce a dark hole in the image plane around the star allowing for high contrast imaging

Modeling framework

We use IDL (Interactive Data Language) and the PROPER optical propagation library (Krist 2007) to model the coronagraphic systems. PROPER includes routines to propagate a wavefront using Fourier-based angular spectrum and Fresnel algorithms, as well as functions to create apertures, deformable mirrors, and a variety of wavefront aberrations (both amplitude and phase). The package includes extensive documentation. New routines that are integrated with the PROPER functions are needed to represent some of the coronagraphs (VVC, PIAA). These are written in IDL and C.

PROPER has been used for modeling of proposed space-based exoplanet missions (Krist et al. 2006, 2008, 2009). It was also used in the design and performance predictions of the NIRCam coronagraph for the James Webb Space Telescope (Krist et al. 2007), the VVC masks made for HCIT and ground-based telescopes (Mawet et al. 2009b), the Gemini Planet Imager (Marois et al. 2008), the EPICS coronagraph on the European Extremely Large Telescope (Vérinaud et al. 2010), and the FOROS optical modeling system for SPHERE (Yaitskova et al. 2010).

The accuracy of the PROPER algorithms is derived in this study by comparison to a more rigorous algorithm (S-Huygens) for a simple Lyot coronagraph system. As detailed in Appendix B, the root-mean-square difference between the fields computed by these two methods is <0.6% (for a $>10^{-10}$ mean contrast field). This is within the milestone goal of 1% error.

These codes are being made publicly available as an add-on package to PROPER. The system layouts and aberration maps used in Milestone #2 will also be made available. These allow the community to both use these codes for additional study of these coronagraphs and to test their own coronagraphs in the same layouts to compare results.

PROPER is freely available at www.openchannelsoftware.com/projects/PROPER.

Caveats

The modeling undertaken in this study assumes scalar propagation of the wavefront. Vector propagation, which includes the physical effects of electric field interactions with conductive and non-conductive materials at small scales, is not used. Vector propagation becomes important when small apertures may act as waveguides and the electrical properties of the aperture substrate are significant (Lieber et al. 2005). In this study the impact of any vectorial effects would be mainly in the PIAA binary post-apodizer, the small occulting spot at the center of the VVC mask, and the amplitude-modifying structure of the HBLC. Based on previous studies for the Terrestrial Planet Finder Coronagraph, we expect such effects to be small as these structures are thin (i.e. we are not using thick apertures such as those used for early shaped pupil experiments). The realm of vector propagation is also well beyond the time and financial limits of this study.

3. Milestone #1 Description

We identify, implement in code, and verify efficient numerical methods for representing wavefront modification by the Hybrid Band-Limited Coronagraph (HBLC), the Vector Vortex Coronagraph (VVC), and the Phase-Induced Amplitude Apodization (PIAA) coronagraph that are accurate to 1% or better relative to the mean field contrast for contrasts down to 10^{-10}.

Milestone Prerequisites

3.1.1 Coronagraph contrast and image plane field dimensions

All of the coronagraphic designs evaluated in this study must be capable of providing, in an aberration-free system, a mean contrast of $\leq 10^{-10}$ within an annulus centered on the star extending between r = 2.5 λ_c/D – 18 λ_c/D in the image plane across a λ = 500 – 600 nm bandpass (λ_c = 550 nm). The inner radius is set by the occulter size and the outer radius by the expected number of deformable mirror actuators across the pupil (which will be 46 in Milestone #2) and the shortest wavelength (18 λ_c/D ≈ 20 λ/D at λ = 500 nm). The coronagraphs in Milestone #1 are modeled *as designed*, excluding manufacturing defects. In Milestone #2 we evaluate systems with realistic fabrication errors and include end-to-end modeling with wavefront control.

3.1.2 Optical system layout

The modeling algorithm for each coronagraph is evaluated with a representation of a realistic telescope+coronagraph optical system layout derived from the ACCESS mission concept study. There are two layouts: one for PIAA and one shared by VVC and HBLC (see Appendix A). Both systems include two deformable mirrors. The DMs are only used for the HBLC in Milestone #1, as that coronagraph requires them.

The layouts are represented as unfolded (linear layout) systems and implemented using the PROPER software with custom routines specific for the coronagraphs. In reality, off axis optics and inclined fold mirrors can generate polarization-dependent aberrations. These can be minimized by balancing (similar pairing of) the off-axis mirrors, keeping incident angles close to normal, and using tailored coatings. In our simulations we assume that a single polarization channel is being modeled.

3.1.3 PIAA

The PIAA design is described in Appendix C. A pupil apodization profile, A, was chosen that produces a diffraction PSF having a contrast of 10^{-10} or better at $r_{image} \geq 2.5\ \lambda_c$/D over λ=500 - 600 nm. The specification is equivalent to the PIAA 2 design, which has already been implemented and tested on the High Contrast Imaging Testbed at JPL. A binary post-apodizer (a transmissive substrate with opaque rings) was derived as part of the design. The surface profile of M1 is set by the desired apodization profile, the distance

from M1 to M2, and the diameter of the beam. The profile of M2 is then the surface that flattens the phase variations introduced by remapping and collimates the beam.

The modeling and verification process for PIAA is detailed in Appendix C.

3.1.4 VVC

The VVC design has a charge of 4, resulting in 4^{th} order aberration rejection. It assumes a 5-layer achromatic design optimized for the λ = 500 - 600 nm passband. The opaque spot at the center of the mask has a diameter of 0.5 λ_c/D. This provides an inner working angle of 1.8 λ_c/D (the coronagraph is evaluated at 2.5 – 18 λ_c/D like the others). A simple circular Lyot stop with a clear aperture of 90% of the diameter of the beam is used. The results are computed for only one polarization channel; in a dual channel system, each would have its own DMs and detector, so the performance of one channel would be the same as the other.

The modeling and verification process for the VVC is detailed in Appendix D.

3.1.5 HBLC

The HBLC used in this study has a circular focal plane occulter that provides an approximately 4^{th} order aberration rejection response. It has an inner working angle of 2.5 λ_c/D. A variable-thickness layer of nickel was chosen to provide amplitude modulation, while cryolite was specified for the phase modulation layer. The resulting wavelength-dependent phase shifts were derived from thin-film calculations using the properties of these materials. The Lyot stop is a circular aperture mask whose diameter is 60% of the beam's. It also includes a central opaque spot with a diameter of 10%. DM settings that provide the required broadband performance are derived.

The modeling and verification process for the HBLC is detailed in Appendix E.

Milestone Requirements

3.1.6 Algorithm efficiency

Milestone #1 Efficiency Requirement: *The efficient algorithm for each coronagraph will allow for the generation of a 2077 component (number of actuators per DM × number of DMs × number of sensing wavelengths) DM response matrix within 48 hours on a single modern workstation.*

Rationale:

The efficiency of the propagator is needed early on in the mission, when the dark hole is being created using an iterative control process prior to the first science observations. The initial DM matrix is typically generated using a system model with no aberrations. In some cases (testbeds and simulations) it has been seen that during the first few iterations the contrast level in the dark hole improves but then stagnates, typically when the non-linearities in the system that are not in the model begin to become apparent (e.g. modification of the DM image-plane influence functions due to unmodeled system aberrations). It is often possible to get things converging again by regenerating the DM

matrix with updated system information (e.g. phase retrieved maps of the system aberrations from on-orbit, defocused star images). If contrast is stuck at 10^{-8} and we need to get down to 10^{-10} and it takes a week to generate a new matrix, then no useful science will likely be done for that week. If it can be done in a few hours, then little of the mission lifetime is lost. Given the need for a fast turn-around to maintain observational efficiency, the numerical modeling algorithms must be fast.

As described previously, thousands of separate wavefront propagations will be required to construct the DM response matrix for either a model or real coronagraphic system having wavefront control. Given the Milestone #2 proposed convention of 46 DM actuators across the pupil, five sensing wavelengths, and invoking an 8-fold symmetrical system, we need ~2077 propagations through the entire system[1]. The rigorous, physics-based, accurate propagation algorithms that are currently available (S-Huygens, Rayleigh-Sommerfeld) would take many days or even weeks to compute 2077 end-to-end simulations, even on multiprocessor systems. During all that time the in-space coronagraph would not be performing at its required contrast level, essentially wasting time.

To determine the 48 hour limit, the PIAA layout was implemented in PROPER, with the PIAA optics replaced with flat mirrors. Using 2048×2048 element wavefront grids, it took 43 seconds to perform one end-to-end simulation. Multiplied by 2077, it would require 24.8 hours to generate the DM response matrix. This ignores the additional time that would be required to account for propagation between actual PIAA optics. Thus, 48 hours on a single computer seems reasonable. The matrix generation process is easily parallelized, so we assume that the computation time on a single workstation can be scaled to a multi-node distributed computing system to provide an update time of a few hours or less, which agrees well with spacecraft communication intervals.

3.1.7 Algorithm accuracy

Milestone #1 Accuracy Requirement: *The efficient algorithm for each coronagraphic technique must provide an accuracy, as defined in Section 3.1.9, that is equal to or less than 1% of the mean contrast in the annulus defined in Section 3.1. The accuracy will be measured in separate simulations with aberrations chosen that produce 10^{-5} and 10^{-10} mean contrast fields (accuracies of $\leq 10^{-12}$ and $\leq 10^{-7}$ for 10^{-10} and 10^{-5} mean contrast fields, respectively).*

[1] In our hypothetical system, we use 48 × 48 actuator DMs with 46 actuators spaced across the projected pupil diameter. This provides $\pi(46/2)^2 = 1662$ useful actuators within the projected area of the pupil. With two DMs, the total number of actuators is 3324. We need to compute the response function for each actuator on each DM at each sensing wavelength. We assume 5 wavelengths, so a total of 3324 × 5 = 16620 response functions are needed. In an unaberrated, circularly-symmetric system, the response function of a DM actuator is the same as the one for an actuator on the opposite side of, and the same distance from, the DM center, with a rotation and/or transposition. We can divide the DM into 8 pie-slice sectors and compute the response functions only for the actuators in one sector. Then, we can rotate and/or transpose those functions as appropriate to create the responses for the other actuators on the DM. So, we can divide the total number of system propagations required to generate the response function matrix by 8: 16620 / 8 ≈ 2077. The charge=4 VVC violates this assumption and allows only a 4, rather than 8, fold symmetry.

Rationale:

The algorithm chosen to model a coronagraph must have sufficient accuracy to properly reflect the system's behavior in the presence of wavefront aberrations. Because scattering from individual component wavefront errors corresponding to contrasts of 10^{-11}-10^{-12} add up to create the ensemble of light within the dark hole, the algorithms must be accurate to equivalent contrasts of 10^{-12} for fields with mean contrasts of 10^{-10}. Accuracy is also important when generating the deformable mirror response matrix used for wavefront control, which requires that the model system match the real one as closely as possible - large differences will result in either a poor solution (low contrast dark hole) or divergence without any improvement. Smaller but still significant errors might allow convergence to a solution but only after many iterations.

The optical system aberrations, which correspond to power spectral density curves similar to those of actual optics, were chosen to produce fields with these contrasts within the defined annulus (Appendix A).

Milestone Metrics

3.1.8 Algorithm efficiency

Milestone #1 Efficiency Metric:

The efficiency of the coronagraphic modeling algorithm is the elapsed time required to propagate one arbitrary, monochromatic wavefront from the first deformable mirror of the specified optical layout, through the optical system (including coronagraph), and then to the final image plane using a current workstation (default is a dual quad-core Xeon workstation) multiplied by the number of system propagations required to generate a DM response matrix.

Rationale:

For each coronagraph, a single wavefront is propagated from the telescope primary mirror, through the defined optical system, and then to the final image plane using a combination of PROPER and the efficient coronagraph modeling algorithms. The elapsed time is multiplied by 2077, the number of full system propagations required to generate the DM response matrix (Section 3.1.6) for an 8-fold, perfectly symmetric system.

3.1.9 Algorithm accuracy

Milestone #1 Accuracy Metric:

The root-mean-square of the differences of the E_{eff} and E_{ref} electric fields generated by, respectively, the efficient and reference (assumed perfectly accurate) methods measured within the annulus specified in Section 3.1.1 will represent the accuracy, in terms of contrast, of the efficient algorithm:

$$accuracy = RMS\left(\left|E_{eff} - E_{ref}\right|^2\right) \Big/ \max\left(PSF\right)$$

Rationale:

For each coronagraph the accuracy of the algorithm is established relative to a reference algorithm that is assumed to be perfect, as detailed further in the appendices. How this is done depends on the coronagraph and is detailed further in the following sections. When verifying an efficient method against a more rigorous one, the computed complex-valued electric fields are compared within the final image plane annulus described above for the same input wavefront. One field is subtracted from the other, the modulus-squared taken, and the resulting intensity converted to contrast by dividing it by the peak unocculted stellar PSF value (the stellar PSF is the modulus square of the unocculted electric field). In essence, this is correcting the efficient method's field (E_{eff}) using the negative of the reference field ($-E_{ref}$) and then measuring the resulting contrast.

4. Success Criteria

The following items summarize the requirements and metrics detailed in Sections 2 and 3.

4.1 The measurements to be evaluated are comparisons between the image plane monochromatic electric fields computed by the efficient models and the reference algorithms.

4.2 The fields are computed for a wavelength of 550 nm.

4.3 The fields are measured and compared within an annulus of $2.5\ \lambda_c/D \leq r \leq 18.0\ \lambda_c/D$ radians centered on the star (λ_c = 550 nm, D = diameter of the entrance pupil).

4.4 Two fields are separately evaluated, one with a mean contrast of approximately 10^{-5} and another of approximately 10^{-10} with the input aberrations scaled to provide those levels.

4.5 The RMS difference between the modeled and reference fields, expressed in terms of contrast, must be less than or equal to 1% of the mean field contrast specified in item 4.4.

4.6 The execution time required to separately propagate 2077 single monochromatic wavefronts through the chosen optical layout using the PROPER library routines and the efficient coronagraph models must be less than 48 hours on a current workstation (dual quad-core Xeon system) and are evaluated based on the execution time of one wavefront propagation through the system.

4.7 Items 4.1 – 4.6 are satisfied with simulations representing three different coronagraphs: Hybrid Band-Limited Coronagraph (HBLC), Vector Vortex Coronagraph (VVC), and Phase-Induced Amplitude Apodization (PIAA) coronagraph. The system optical layouts are defined in Appendix A and the coronagraphs and their respective algorithms are defined in Appendices C, D, and E.

5. Milestone #1 Results Summary

The requirements for accuracy and efficiency defined in the Milestone #1 whitepaper for the HBLC and vector vortex coronagraphs have been met (Table 1). In most cases, the limiting accuracy is caused by numerical errors in the PROPER algorithms, as derived in Appendix B, Table 4. For PIAA, the efficiency metric at 10^{-10} contrast is met for 2048 × 2048 arrays, but not the accuracy metric (5% realized versus 1% requirement). The accuracy metric is met using 4096 × 4096 arrays (1%), but not the efficiency metric. This is a limit set by the filtered representation of the binary post-apodizer. The accuracy of the PASP propagation method when compared to the reference method (S-Huygens), however, is well within the milestone accuracy requirement for 2048 × 2048 arrays (0.03% at 10^{-10} contrast). If a grayscale rather than binary apodizer were used for PIAA, then this would be the relevant accuracy.

Table 1. Summary of Milestone #1 Achieved Metrics

	~10^{-5} Contrast[(1)] Accuracy	~10^{-10} Contrast[(1)] Accuracy	Execution Time[(2)]	For Details, Appendix
Requirement	≤1%	≤1%	≤48 hrs	
PIAA	0.3%[(3)]	5%[(3)]	33.8 hrs	C, Tables 5 & 6
VVC	0.02%[(4)]	0.6%[(4)]	15.7 hrs	D, Table 7
HBLC	0.02%[(4)]	0.6%[(4)]	8.3 hrs	E, Table 9

[(1)] Approximate contrast. See corresponding appendices for exact values.
[(2)] Time to compute 2077 (HBLC), 2080 (PIAA), or 4154 (VVC) actuator influence functions on a dual quad-core Xeon system using 2048 × 2048 arrays.
[(3)] The stated accuracy for PIAA is set by the comparison of a high-resolution, one-dimensional propagation of an aberrated field at PIAA M2 to the same field converted to a circularly-symmetric, two-dimensional field (2048 x 2048) and then propagated (see Figure 26 and related text).
[(4)] This value is set by the PROPER accuracy determined in Appendix B because it is worse than that determined from the comparison to the reference for this coronagraph.

Other accomplishments in this study

- Determination of the accuracies of the PROPER algorithms compared to a more exact method.

- Successful merging of the PASP and S-Huygens algorithms to provide an efficient and accurate means of propagating between PIAA optics.
- Development of an accurate and efficient means to compute the VVC, including the small opaque spot.
- Development of techniques for the optimization of circular HBLC focal plane masks and the design of the 1st realistic circular mask.
- Development of code that can be used by the community for PIAA, HBLC, and VVC propagations.

6. Milestone Certification Data Package

The results of this study have been reviewed by NASA Headquarters. The documentation, code, and data products that provide evidence that the requirements of this milestone have been met are:

a. Documentation detailing the modeled optical system layout and coronagraph optical parameters (Appendices A, C, D, & E).

b. Documentation describing the efficient and rigorous algorithms used for each coronagraph (Appendices C, D, & E).

c. Documented computer codes (IDL, C) that implement the efficient and rigorous algorithms (provided as an add-on package to PROPER).

d. Documentation detailing the accuracies achieved for the efficient algorithm for each coronagraph and the execution times necessary to create the DM response matrices (Appendices C, D, & E; summaries in Section 6)

e. The complex-valued fields at the final image planes generated by the efficient and rigorous algorithms for each coronagraph and at each required contrast level that were used to demonstrate achievement of the milestone requirements. These will be distributed as two FITS files each (real and imaginary field components) and shown as color-coded images in this document.

f. FITS files containing the input phase and amplitude error maps used to generate the results in item (e) and shown as color-coded images in this document (Appendix A).

g. Contrast maps at each required contrast level for each coronagraph as generated by the efficient algorithms. These are distributed as FITS files and as color-coded maps in this document (Appendices C, D, & E)

7. References

Abramowitz, M., and Stegun, I., "Handbook of Mathematical Functions", New York:Dover (1972).

Balasubramanian, K. et al, "Polarization compensating protective coatings for TPF Coronagraph optics to control contrast degrading cross polarization leakage," *Proc. SPIE,* 5905 (2005).

Belikov R., Kasdin N.J., Vanderbei R.J., "Diffraction-based Sensitivity Analysis of Apodized Pupil-Mapping Systems," *Astrophysical Journal,* 652, 833 (2006).

Belikov, R., Pluzhnik, E., Connelley, M. S., Lynch, D. H., Witteborn, F. C., Cahoy, K., L., Guyon, O., Greene, T. P., McKelvey, M. E., "First results on a new PIAA coronagraph testbed at NASA Ames," *Proc SPIE* 7440, 74400J (2009).

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

Guyon, O., Pluzhnik, E., Galicher, R., Martinache, F., Ridgeway, S., Woodruff, R., "Exoplanet Imaging with a Phase-Induced Amplitude Apodization Coronagraph. I. Principle," *Astrophysical Journal*, 622, 744 (2005).

Kern, B. D., Belikov, R., Give'on, A., Guyon, O., Kuhnert, A. C., Levine, M. B., Moody, D. C., Jr., Niessner, A. F., Pueyo, L. A., Shaklan, S. B., Traub, W. A., Trauger, J. T., "Phase-Induced Amplitude Apodization (PIAA) coronagraph testing at the High Contrast Imaging Testbed," *Proc. SPIE,* 7440, 7440H (2009).

Krist, J., Trauger, J., Moody, D., "Studying a simple TPF-C", *Proc. SPIE*, 6265 (2006).

Krist, J., et al., "Hunting planets and observing disks with the JWST NIRCam coronagraph," *Proc. SPIE*, 6693, 66930H (2007).

Krist, J., "PROPER: An optical modeling program for IDL," *Proc. SPIE*, 6675, 66750P (2007).

Krist, J., Shaklan, S., Levine, M., "Extraction of extrasolar planet spectra from realistically simulated wavefront corrected coronagraphic fields," *Proc. SPIE*, 7010, 701044 (2008).

Krist, J., Moody, D., Mawet, D., Trauger, J., Belikov, R., Shaklan, S., Guyon, O., Vanderbei, R., "End-to-end simulations of different coronagraphic techniques," *Proc. SPIE*, 7440 (2009).

Kuchner, M., Traub, W., "A Coronagraph with a Band-limited Mask for Finding Terrestrial Planets," *Astrophysical Journal*, 750, 900 (2002).

Lieber, M., Neureuther, A., Ceperley, D., Kasdin, J., Hoppe, D., Eisenman, A., "Evaluating the end-to-end performance of TPF-C with vector propagation models: Part I – pupil mask effects," *Proc. SPIE*, 5905, 59050K (2005).

Marois, C., Macintosh, B., Soummer, R., Poyneer, L, Mayman, B., "An end-to-end polychromatic Fresnel propagation model of GPI," *Proc. SPIE*, 7015, 70151T (2008).

Mawet, D., et al., "Infrared achromatic phase shifters using modulated total internal reflection", *Proc. SPIE*, 6268, 79 (2006).

Mawet, D., "Annular Groove Phase Mask:An Achromatic Vortex Coronagraph Intended at Differential Polarimetric Imaging", *Proceedings of In the Spirit of Bernard Lyot: The Direct Detection of Planets and Circumstellar Disks in the 21st Century* (2007).

Mawet, D., et al., "Optical Vectorial Vortex Coronagraphs using Liquid Crystal Polymers: theory, manufacturing and laboratory demonstration," *Optics Express*, 17, 1902 (2009a).

Mawet, D., Trauger, J., Serabyn, E., Moody, D., Liewer, K., Krist, J., Shemo, D., O'Brien, N., "Optical vector vortex coronagraph: first results in the visible," *Proc. SPIE*, 7440, 74400X (2009b).

Mawet, D., Pueyo, L., Moody, D., Krist, J., and Serabyn, E., "The Vector Vortex Coronagraph: sensitivity to central obscuration, low-order aberrations, chromaticism, and polarization", *Proc. SPIE,* 7739, 773914 (2010).

Mawet, D., et al., "Taking the vector vortex coronagraph to the next level for ground- and space-based exoplanets imaging instruments: review of technology developments in the USA, Japan, and Europe", *Proc. SPIE*, 8151, 815108 (2011a).

Mawet, D., et al., "Recent results of the second generation of vector vortex coronagraphs on the high-contrast imaging testbed at JPL", *Proc. SPIE*, 8151, 81511D (2011b).

Moody, D., Gordon, B., Trauger, J., "Design and demonstration of hybrid Lyot coronagraph masks for improved spectral bandwidth and throughput," *Proc. SPIE*, 7010, 70103P (2008).

Murakami, N., Baba, N., Ise, A., Sakamoto, M., and Oka, K., "Laboratory demonstration of an optical vortex mask using photonic crystal", Proceedings of the 2nd In the Spirit of Lyot Conference (2010).

Pluzhnik, E., Guyon, O., Ridgway, S., Martinache, F., Woodruff, R., Blain, C., Galicher, R., "Exoplanet Imaging with a Phase-Induced Amplitude Apodization Coronagraph. III. Diffraction Effects and Coronagraph Design," *Astrophysical Journal*, 644, 1246 (2006a).

Pluzhnik, E., Guyon, O., Warren, M., Ridgway, S., Woodruff, R., "PIAA coronagraph design: system optimization and first optics testing," *Proc. SPIE*, 6265, 62653S (2006b).

Pueyo, L, Shaklan, S., Krist, J., "Numerical propagator through PIAA optics," *Proc. SPIE*, 7440, 74400E (2009).

Shaklan, S., "Technology Milestone #3A White Paper: Coronagraph Starlight Suppression Model Validation," JPL Document D-64582 (2009).

Soummer, R., Pueyo, L., Sivaramakrishnan, A., Vanderbei, R., "Fast computation of Lyot-style coronagraph propagation," *Optics Express*, 15, 15935 (2007).

Traub, W., et al., "TPF-C: status and recent progress," Proc. SPIE, 6268, 62680T (2006).

Trauger, J., Stapelfeldt, K., Traub, W., Henry, C., Krist, J., et al. "ACCESS: a NASA mission concept study of an actively corrected coronagraph for exoplanet system studies," *Proc. SPIE*, 7010, 701029 (2008).

Trauger, J., Give'on, A., Gordon, B., Kern, B., Kuhnert, A., Moody, D., et al., "Laboratory demonstrations of high-contrast imaging for space coronagraphy," *Proc. SPIE*, 6693, 66930X (2007).

Vanderbei, R., Spergel, D., and Kasdin, N., "Spiderweb Masks for High-Contrast Imaging", *Astrophysical Journal*, 590, 593 (2003).

Vanderbei, R., "Diffraction Analysis of Two-dimensional Pupil Mapping for High-Contrast Imaging," *Astrophysical Journal*, 636, 528 (2006).

Vanderbei, R., and Traub, W., "Pupil Mapping in Two Dimensions for High-Contrast Imaging", *Astrophysical Journal*, 626, 1079 (2005).

Vérinaud, C., Kasper, M., Beuzit, J-L., "System study of EPICS: the exoplanets imager for the EELT," *Proc. SPIE*, 7736 (2010).

Yaitskova, N., et al., "FOROS: Fresnel optical propagation code for SPHERE," *Proc. SPIE*, 7735 (2010).

Appendix A: Optical system layout

We assume a telescope layout very similar to that used for the ACCESS mission concept study that represents a realistic system. The ACCESS telescope has a 1.5 m diameter primary mirror, which we will also assume. However, the size of the telescope will not impact the results our study, and for a larger system the components may be scaled and/or the beam feeding the coronagraph may be resized by preceding optics to keep the same instrument dimensions. The results are therefore applicable to the same coronagraphs on larger telescopes. The telescope has an off-axis Gregorian configuration that avoids the obscurations caused by a secondary mirror and its supports. The beam from the secondary is diverted behind the primary and into the coronagraph by a fold mirror. An off-axis parabola (OAP) forms a collimated beam and an image of the primary mirror on a deformable mirror (DM1). Some additional distance down the optical path a second DM (DM2) is placed to provide leverage for controlling wavefront amplitude errors. Another OAP then focuses the beam. At this point the system has two configurations, one for the HBLC /VVC (Figure 5) and another for PIAA (Figure 6):

HBLC & VVC: The occulting/OVV focal plane mask is located at the intermediate focus. An OAP then forms an image of the primary mirror on the Lyot stop, which is a simple aperture mask. Yet another OAP or lens focuses the beam onto the detector. The prescription is listed in Table 2.

PIAA: The beam passes through the intermediate focus without alteration and to an OAP that collimates it and forms an image of the primary mirror on the PIAA M1 mirror. The beam is remapped onto PIAA M2 and then an OAP collimates the beam and forms an image of M2 on the post-apodizer. Another OAP focuses the beam onto the occulter. The beam is then fed directly to the first reverse PIAA mirror, M1R. This reverses the apodization produced by the forward PIAA optics, forming a geometrically undistorted wavefront on the second mirror, M2R. An OAP or lens then focuses the beam onto the detector. The prescription is listed in Table 3.

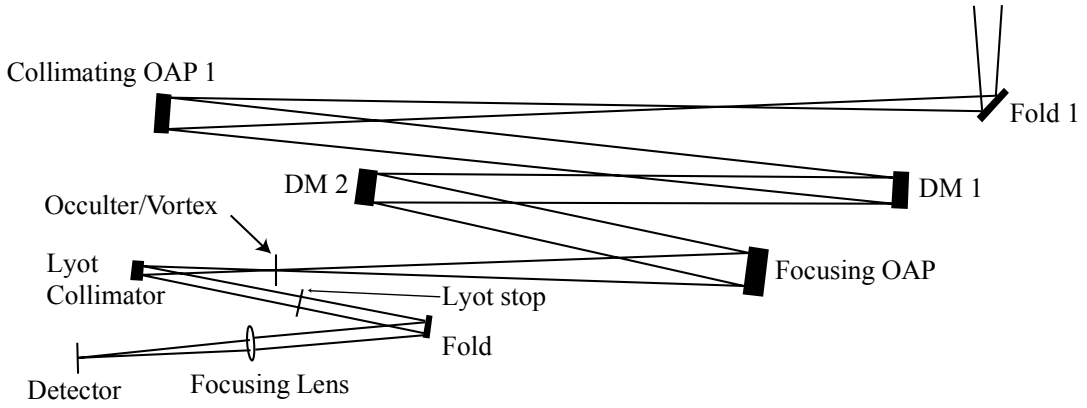

Figure 5. Schematic optical layout for the HBLC/VVC. Not shown are the telescope primary and secondary mirrors that feed Fold 1 in the upper right.

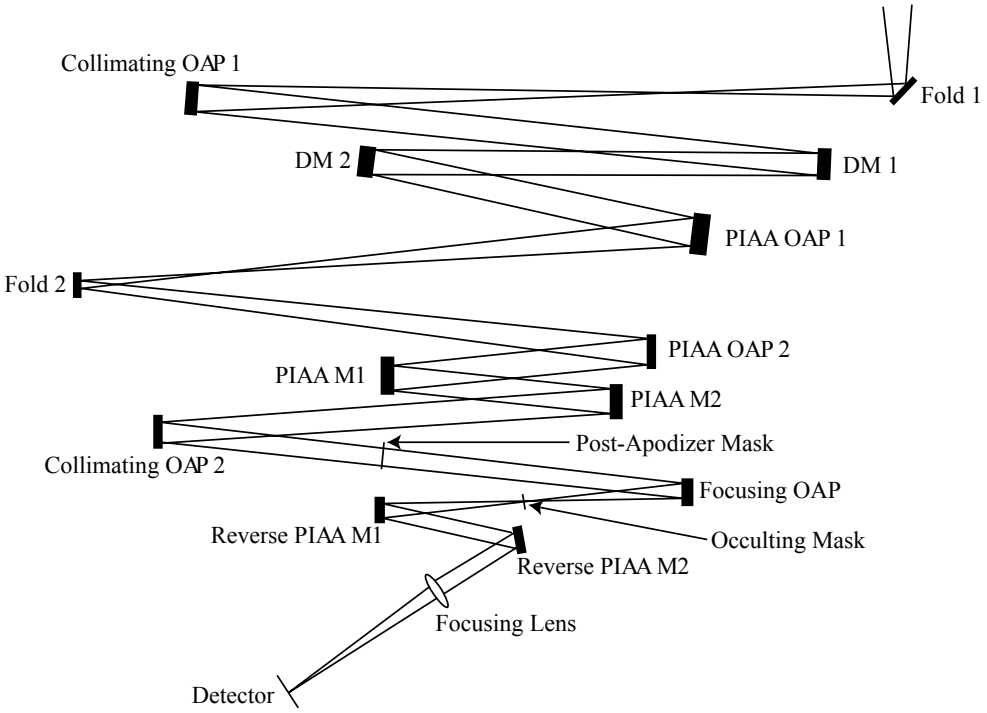

Figure 6. Schematic optical layout for the PIAA coronagraph. Not shown are the telescope primary and secondary mirrors that feed Fold 1 in the upper right.

Table 2. Optical prescription for the HBLC/VVC Layout

Optic	Beam Diameter* (m)	Distance from Previous Surface (m)	Focal Length (m)
Primary	1.500		3.798
Secondary	0.237	4.398	0.542
Fold	0.017	5.250	
DM collimator	0.046	1.512	1.100
DM 1	0.046	1.340	
DM 2	0.046	1.000	
Focuser	0.046	0.750	1.000
Occulter		1.000	
Lyot collimator	0.023	0.500	0.500
Lyot stop	0.023	0.312	
Fold	0.023	0.050	
Focuser	0.023	0.100	1.773
Fold	0.012	0.883	
Image		0.890	

*The beam diameter ignores limiting apertures such as stops, except for the primary.

Table 3. Optical prescription for the PIAA Layout

Optic	Beam Diameter* (m)	Distance from Previous Surface (m)	Focal Length (m)
Primary	1.500		3.798
Secondary	0.237	4.398	0.542
Fold	0.017	5.250	
DM collimator	0.046	1.512	1.100
DM 1	0.046	1.340	
DM 2	0.046	1.000	
PIAA collimator 1	0.046	0.750	1.225
PIAA fold	0.022	1.811	
PIAA collimator 2	0.090	1.811	2.397
PIAA M1	0.090	0.387	
PIAA M2	0.108	0.900	0.667
PIAA apodizer collimator	0.054	1.000	0.333
PIAA apodizer	0.054	0.500	
PIAA focuser	0.054	0.667	0.333
PIAA occulter		0.333	
PIAA reverse M2	0.054	0.333	0.333
PIAA reverse M1	0.045	0.450	
PIAA lens	0.045	0.350	1.425
PIAA fold	0.020	0.800	
Image		0.625	

*The beam diameter ignores limiting apertures such as stops, except for the primary.

Wavefront Error Definition

The Milestone #1 accuracy requirement entails propagating an aberrated wavefront through the optical system using both an assumed-perfect reference algorithm and an efficient one. The fields at the final image plane are then compared. The phase error map used in these propagations was derived from a power-spectral-density (PSD) curve. The PSD specifies how much power each spatial frequency introduces. The PSD was defined to be similar to current high-quality optics. Spatial frequencies above 48 cycles/diameter were rejected, for reasons described in Appendix C. The square root of the PSD was taken to convert it from power (intensity) to amplitude and then interpolated to produce a 2-D, complex-valued symmetric function. Random phases were generated for each element, then the result was Fourier transformed and the real value taken to produce the phase error map, shown in Figure 7. By trial and error, normalizations of this map were chosen to produce $\sim 10^{-5}$ and $\sim 10^{-10}$ contrast fields in the final image plane of a simple bandlimited coronagraph. The measured mean PSD of this map is shown in Figure 8.

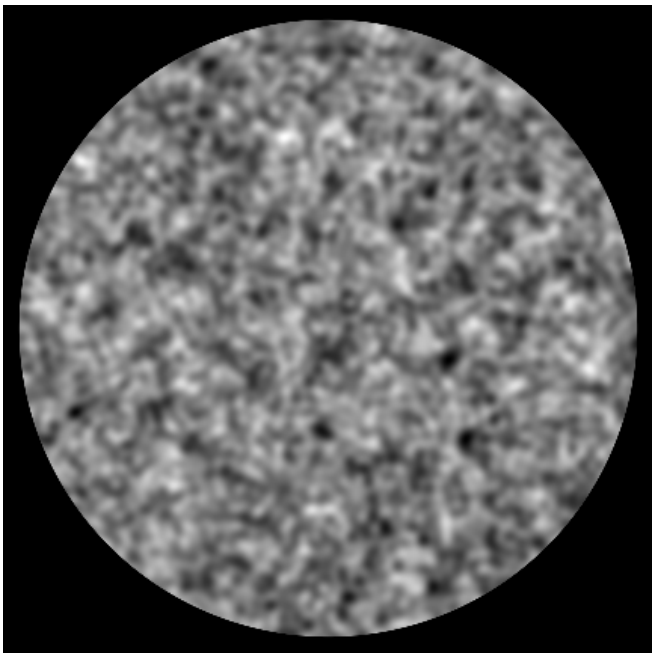

Figure 7. Phase error map used for the coronagraph propagator accuracy tests.

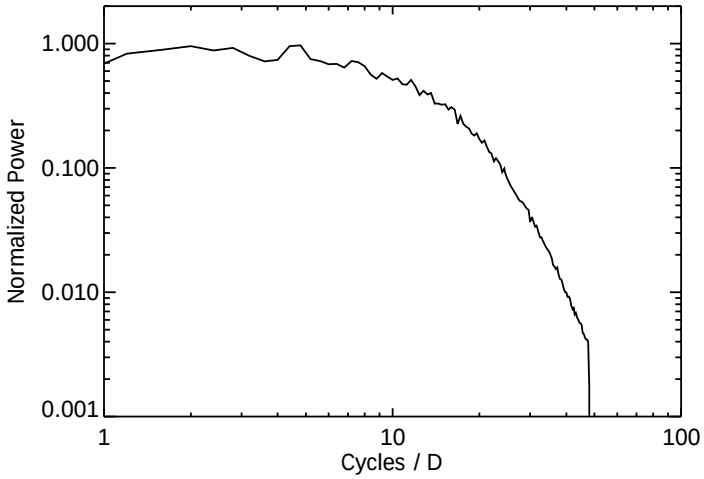

Figure 8. Normalized power spectral density (2-D) curve for the map shown in Figure 7. Note that there are no aberrations with spatial frequencies above 48 cycles/D.

Appendix B: Verification of PROPER

Introduction to PROPER

The goal of Milestone #1 is to develop efficient and accurate procedures for propagating an aberrated wavefront through a coronagraphic system. Certain coronagraphs (PIAA, vector vortex) contain components that differ significantly from the standard optics found in most systems (parabolic, spherical, or flat mirrors and lenses, for instance). Calculations of wavefront propagation through these "special" optics require non-standard algorithms, unlike the Fourier-based Fresnel or Angular Spectrum propagators typically used for "normal" systems. In later sections we will describe techniques for modeling the PIAA and vector vortex coronagraphic elements.

The optics in the system before and after these special elements are usually rather conventional (e.g., a telescope, deformable mirrors, collimators, and focusing optics), and the propagation between them can be modeled using efficient Fourier-based algorithms. In this study, we use the PROPER library for IDL (Interactive Data Language) to propagate the wavefronts between the optics, except when alternative algorithms must be used. PROPER was developed at JPL for the Terrestrial Planet Finder Coronagraph program to provide a common set of tools for the community for simulating optical systems. It can be freely downloaded from www.openchannelsoftware.com and is well documented. The PI of this study, John Krist, is its developer. It consists of a set of routines that uses standard Fourier-based Angular Spectrum and Fresnel propagation algorithms, with automatic selection of the appropriate method. The library includes routines for the creation of complex apertures and of phase and amplitude aberrations (specified by Zernike coefficients or power spectrum curves). It includes a deformable mirror model that includes measured actuator influence functions from DMs used in the HCIT. To simulate a polychromatic image, multiple monochromatic ones generated at wavelengths sampling the passband will be added together.

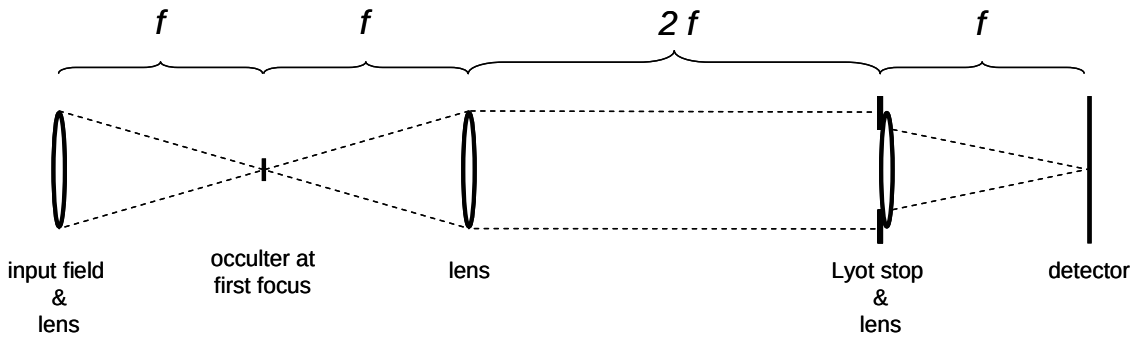

Figure 9. Schematic of the simple Lyot coronagraph used to verify PROPER.

Verifying PROPER accuracy

To establish the accuracy of the PROPER routines, we define a very simple, unobscured Lyot coronagraph (Figure 9). The occulter is a bandlimited, amplitude transmission mask:

$$amp(r) = 1 - \mathrm{sinc}(r)^2$$

The mask is scaled to provide 50% intensity transmission at 2.5 λ/D radians. The Lyot stop is a circular aperture with a 53% diameter clear opening. The entrance aperture is 2.5 cm in diameter, and every lens is parabolic with a focal length of 70 cm. The uniformly-illuminated input wavefront has phase errors composed of concentric ripples with a variety of spatial frequencies and amplitudes, and it is normalized to produce a $\sim 10^{-10}$ or 10^{-5} mean contrast over an $r = 2.5 - 18$ λ/D field in the final image plane.

The input field is propagated through the system with PROPER at λ = 500 nm. The entrance pupil is mapped across 307.2 pixels and embedded in a 2048 × 2048 array. This produces a sampling of 0.15 λ/D in the image plane.

The S-Huygens algorithm is used to produce the reference field against which the PROPER results are compared for accuracy. As described in the PIAA section (Appendix C), S-Huygens involves explicit computation (no Fourier transforms) of the propagation of a highly-sampled, one-dimensional radial slice of a two-dimensional wavefront. In that section, the accuracy of S-Huygens is verified against the Rayleigh-Sommerfeld diffraction integral for propagation between PIAA optics, so we use it as our reference algorithm here. Because we use a symmetrical, circularly-aberrated wavefront, only a single 1-D wavefront slice needed to be propagated between each optic. The lenses are represented by parabolic mirror surfaces, rather than the thin-lens phase terms used by PROPER. At the entrance, 80,000 points are used to sample the radius (0.156 μm/element). The occulter plane image is computed out to 300 λ/D radians with a sampling of 0.15 λ/D per element. The field at the 2nd lens extends 60% further than the entrance pupil radius, sampled by 128,000 points (0.156 μm/element). The field at the Lyot stop is at the same sampling but only over 43,200 points along the radius due to the stop. The final image plane is sampled at 0.15 λ/D.

The wavefront is propagated through the system using both methods at λ = 500 nm. In the final image plane, the field is normalized by the peak of the PSF (which is computed by propagating through the system without the occulter), so that the field intensity is in units of contrast. A global phase offset between the two methods is also removed. Within the 2.5 – 18 λ/D image plane annulus the mean contrast of the S-Huygens high-contrast field is 9.9×10^{-11} and PROPER's is 9.7×10^{-11}. In this case, the accuracy of PROPER, as the metric is defined in the Milestone requirement, is 5.6×10^{-13} (0.6%), which is within the goal of 1%. The contrast plots for both fields are shown in Figure 10. In the lower contrast case ($\sim 10^{-5}$), the accuracy is 0.02%. These results are summarized in Table 4.

Table 4. Milestone #1 Metrics for PROPER

	Reference	Efficient
Algorithm	S-Huygens	PROPER (Fourier)
Array size at pupil ***Beam radius***	80000 pix (1-D) 80000 pix	2K × 2K pix 128 pix
Sampling at focal plane mask	0.15 λ/D	0.125 λ/D
LC mean contrast in DH ***LC ACCURACY*** ***LC Requirement***	9.9×10^{-6}	9.7×10^{-6} ***1.7×10^{-9} (0.02%)*** *$<9.7 \times 10^{-8}$ (<1%)*
HC mean contrast in DH ***HC ACCURACY*** ***HC Requirement***	9.9×10^{-11}	9.7×10^{-11} ***5.6×10^{-13} (0.6%)*** *$<9.7 \times 10^{-13}$ (<1%)*

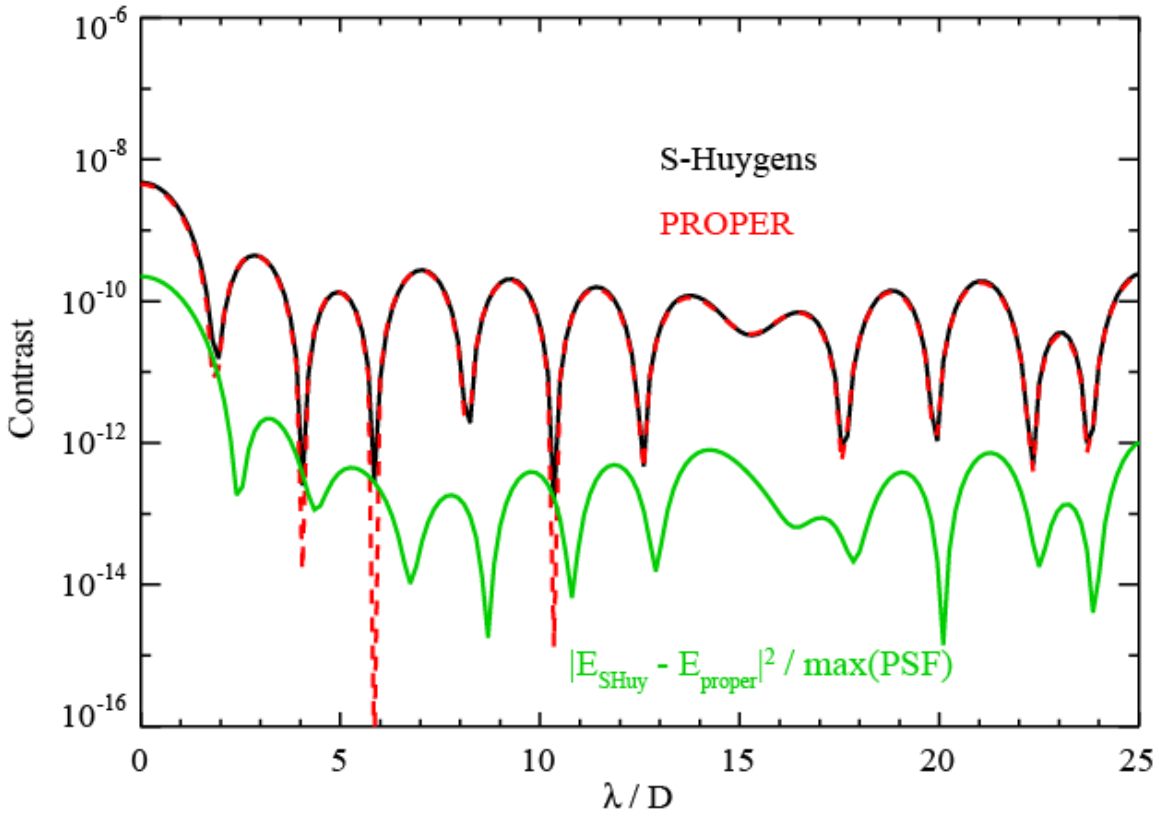

Figure 10. Plots of the ~10^{-10} contrast coronagraph image fields calculated by PROPER and S-Huygens. The modulus square of the difference between the electric fields is also plotted.

Appendix C: Phase-Induced Amplitude Apodization (PIAA)

Introduction

Apodization in a telescope will produce a point spread function with significantly reduced wing intensity, making it a technique of interest for high contrast imaging (Figure 11 & Figure 12). Conventional apodizers using transmission-altering masks reduce throughput too much to be acceptable for exoplanet imaging, and they would be difficult to accurately fabricate and operate over broad bandpasses.

As an alternative method of apodization, Phase-Induced Amplitude Apodization (PIAA) utilizes two optics (lenses or mirrors) that geometrically distort the wavefront. The first optic, M1, is located at a pupil and remaps the beam onto the second optic, M2, which corrects for the phase errors from the remapping to create a "flat" wavefront (Figure 1).

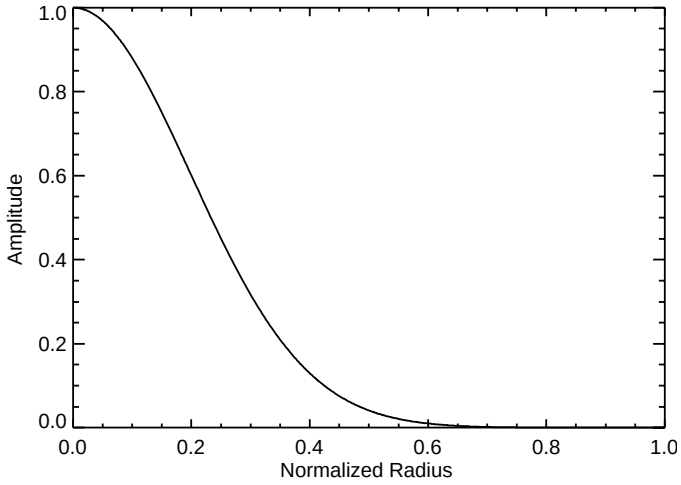

Figure 11. Ideal radial amplitude apodization profile.

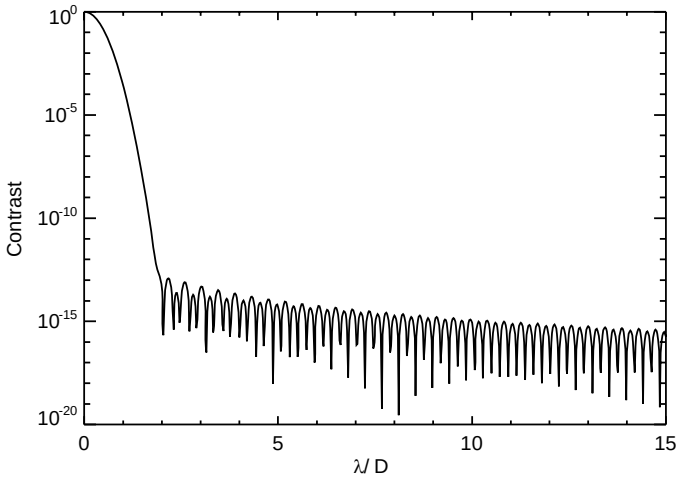

Figure 12. Radial point spread function produced by the ideal apodization profile

Pure PIAA systems that provide 10^{-10} contrast are impractical due to diffraction effects and stringent surface tolerances. These can be reduced in a hybrid system by using a weak apodizer (post-apodizer) placed near or at a subsequent image of M2 (Pluzhnik et al. 2006a, 2006b). In practice, post-apodizers are implemented using binary transmission masks composed of a series of narrow, opaque rings spaced and sized to provide the desired diffraction reduction. These have been fabricated using lithographic techniques (Kern et al. 2009).

After the post-apodizer, the beam is brought to a focus where an occulting spot masks the central lobe of the apodized stellar point spread function (PSF). At this point any field sources are highly distorted and blurred. To restore the original wavefront mapping and thus image quality, the beam is fed through another set of PIAA optics, but in reverse order (no apodizer is required for the reverse system). Figure 13 shows a PIAA coronagraphic system with forward and reverse mapping optics.

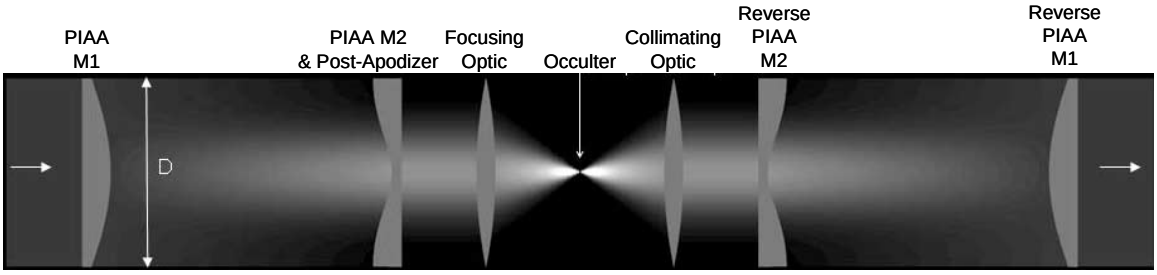

Figure 13. Schematic layout of a PIAA coronagraph with forward and reverse pupil remapping optics. PIAA M1 is at a pupil. In a more realistic system, additional optics would form an image of PIAA M2 onto the apodizer. Adapted from Belikov et al. (2006).

PIAA Design

We use a PIAA optical design that, when combined with a grayscale occulter, would create an apodization profile, $A_{ideal}(r)$, shown in Figure 11. This profile has the form:

$$A_{ideal}(r_{M2}) = \exp\left[-14.414531\left(\frac{r_{M2}}{R_{M1}}\right)^{2.00395} + 1.660601\left(\frac{r_{M2}}{R_{M1}}\right)^{2.036393}\right]$$

where r_{M2} is the distance from the center of M2 and R_{M1} is the radius of M1 (45 mm). Generating this apodization profile with only PIAA optics would introduce high levels of chromatic diffraction due to the high curvature on the edge of M1. The diffraction can be reduced by offloading some of the apodization onto a separate apodizer, with the PIAA

optics producing only a portion of the total apodization. To ease fabrication at the edge, the curvature can be lessened by extending the PIAA apodization profile to have a uniform level beyond a specified radius ("overrun"). In the PIAA system we use in this study, the PIAA-induced apodization extends out to a radius of 54 mm on M2 (Figure 14):

$$A_{PIAA}(r_{M2}) = A_{ideal}(r_{M2}),\ A_{ideal}(r_{M2}) \geq 0.411239$$

$$A_{PIAA}(r_{M2}) = A_{ideal}(r_{M2}) + 0.070506\left(\frac{1}{2}+\frac{1}{2}\cos\left[\frac{\pi A_{ideal}(r_{M2})}{0.411239}\right]\right)^{1.642964},\ A_{ideal}(r_{M2}) < 0.411239$$

$$A_{PIAA}(r_{M2}>R_{M1}) = A_{PIAA}(R_{M1}),\ r_{M2} \leq 54 \text{ mm}$$

The apodization profile is renormalized to conserve intensity. It is possible to define A_{PIAA} so that the M2 radius is equal to or smaller than M1's. We chose the above prescription to create a beam with approximately the same radius *after applying the post-apodizer*. This prescription produces a rolled edge at M1 (Figure 14). We describe the post-apodizer design in a later section.

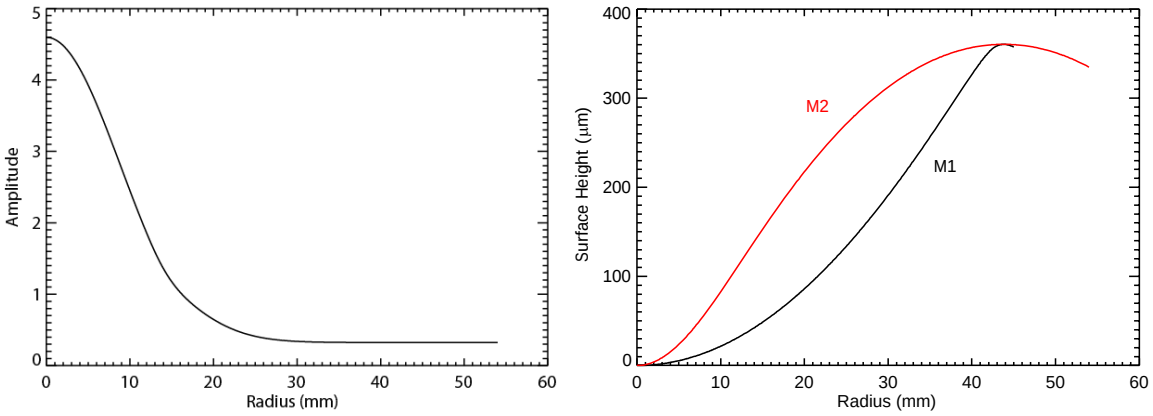

Figure 14. (Left) PIAA-induced apodization profile, including "overrun"; (Right) Surface height profiles for PIAA M1 (R=45 mm) and M2 (R=54 mm).

The forward (apodizing) PIAA system was specified to have a separation between M1 and M2 of 0.9 m. The surface heights and M1-to-M2 radial mapping functions were obtained using A_{PIAA} as defined above and the equations in Vanderbei & Traub (2005). The reverse PIAA (distortion-correction) system uses a 50% scale version of the forward PIAA (27 mm diameter reverse M2, 45 mm diameter reverse M1, 0.45 m separation).

PIAA Optical Modeling

The optical modeling of PIAA is much more difficult than for nearly any other coronagraphic technique because the 1st PIAA optic that creates the apodized beam has a

surface with a high rate of curvature change near its outer edge. With conventional Fourier-based propagation methods (*e.g.,* Fresnel, angular spectrum) this change in the phase of the wavefront is too great to be accurately represented by a uniformly spaced grid without aliasing unless the array is unreasonably large (tens of thousands of points in each dimension). These techniques also require a nearly-parabolic or flat wavefront phase and uniform spacing between rays during propagation, which is not the case for PIAA. One can use brute force methods, such as integration of the Rayleigh-Sommerfeld diffraction equation, but the time required to propagate just one wavefront would be prohibitive (weeks or, even on a large parallel processing cluster, days).

A new, accurate means of propagation has been devised called S-Huygens (Vanderbei 2006; Belikov et al. 2006). This is an approximation to the integration of the Huygens wavelet diffraction integral, but with fewer simplifications than used by the Fourier-based algorithms. It is too slow for practical modeling because any arbitrarily-aberrated input wavefront must be decomposed into numerous (hundreds) of orthogonal terms that are then separately propagated, and the propagation of a single term can take tens of seconds or more. It is, however, useful as a means of verifying other, more efficient methods without resorting to the even more computationally intensive methods like Rayleigh-Sommerfeld.

One efficient method, used in this study, is geometric remapping with propagation of individual component harmonic aberrations, which we call PIAA Angular Spectrum Propagation (PASP). In this case the aberrated wavefront is decomposed into pure harmonic components (ripples in X and Y) that are each analytically propagated followed by wavefront geometric distortion mapping via interpolation.

S-Huygens Propagation

Algorithm

The S-Huygens algorithm propagates a 1-D radial slice of the 2-D wavefront from one surface to another (Vanderbei 2006; Belikov et al. 2006). The diffraction integral must be explicitly computed – there is no quick, Fourier-based implementation. In a circularly-symmetric system, either unaberrated or with circularly-symmetric aberrations, only one radial integral need be computed. To propagate an arbitrarily-aberrated 2-D wavefront, the input field must be decomposed into one-dimensional vectors representing azimuthal harmonics (ripples in the azimuthal direction) with varying radial weights. The results are added together at the end to create a new 2-D wavefront, as shown in Belikov et al. (2006).

The 2-D input wavefront at the first surface (r,θ,h) (h is the surface height) is described by a radially-and-azimuthally-varying electric field coefficient $C(k,r)$ multiplied by an azimuthally-varying harmonic term of k cycles:

$$E_{in}(r,\theta,k) = C(k,r)e^{ik\theta}$$

For a perfectly circularly-symmetric wavefront, only the k=0 terms need be computed. $C(k,r)$ is the azimuthal frequency spectrum of the wavefront, derived independently for each r:

$$C(k,r) = \sum_{\theta=0}^{2\pi} E_{in}(r,\theta) e^{-ik\theta} \text{ , } k\text{=-}n_k \text{ to } n_k$$

Interpolation is used to extract a Cartesian-sampled wavefront at each radius r over all angles θ.

The corresponding output wavefront at the second surface $(\tilde{r},\theta,\tilde{h})$ separated from the first by distance Z is:

$$E_{out}(\tilde{r},\theta,k) = E_o(\tilde{r},k) e^{ik\theta}$$

where

$$E_o(\tilde{r},k) = \frac{2\pi}{i^{k+1}\lambda Z} \int_0^\infty K(r,\tilde{r}) E_i(r) J_k\left(\frac{2\pi r\tilde{r}}{\lambda S(\tilde{r},r)}\right) r\, dr$$

$$K(r,\tilde{r}) = A_{out}(\tilde{r}) e^{2\pi i \left/ \lambda \right. \left[\tilde{r} / S(\tilde{r},r) + S(\tilde{r},r) + \left[\tilde{h}(\tilde{r}) - h(r)\right] \right]} A_{in}(r)$$

$$S(\tilde{r},r) = \sqrt{r^2 + \tilde{r}^2 + \left[h(r) - \tilde{h}(\tilde{r})\right]^2}$$

Note that E_o is a one-dimensional vector for each k, representing the field at θ = 0. The mapping of $\tilde{r}$ from r is derived from the prescribed PIAA apodization, which also defines the surface heights (Vanderbei & Traub, 2005). In our case the apodization at PIAA M1, A_{in}, is 1 inside the beam and 0 outside. A_{out} is the defined apodization produced by M1 at M2.

The choice of n_k and number of points along the radial direction, n_r, are chosen to sufficiently sample the wavefront and encompass all of the spatial frequencies of interest.

Implementation

We have implemented the S-Huygen algorithm as a C routine that is called from IDL. This code is multithreaded to improve performance on multi-core CPUs.

Sampling

The number of samples along the radial direction required to produce an accurate result is dependent on the wavelength, the diameter and curvature of the optics, and the

propagation distance. The sampling for the k=0 component needs to be high because it contains most of energy (all of the energy in an unaberrated or circularly-aberrated system) and dominates the edge diffraction of the propagated wavefront. For other values of k, which sample non-circularly symmetric aberrations, the sampling can be eased. The maximum value of k ($-k$ to $+k$, actually) depends on the highest azimuthal spatial frequency that needs to be sufficiently sampled at the edge of the input beam.

When verifying PASP against S-Huygens, as described later, in the forward PIAA (M1 D=90 mm) we use 100,000 points radially on M1 and M2 for the k=0 terms and for the others 25,000 on M1 and 15,000 on M2. Because the k=0 term has been largely reduced by the time the wavefront enters the reverse PIAA, we use 15,000 points at reverse M2 and 25,000 at reverse M1.

When running the full telescope+coronagraph layout, we use 100,000 points/radius at M1 and 60,000 at M2 to generate the field from the tapered edge.

Verification of S-Huygens Against Rayleigh-Sommerfeld

S-Huygens is used to verify the results from the PASP method, so first S-Huygens itself must be verified. One of the three PIAA-specific requirements defined in the Milestone #1 whitepaper includes verification of the S-Huygens algorithm against the Rayleigh-Sommerfeld diffraction integral:

Milestone statement:

We will verify the accuracy of the S-Huygens algorithm by propagating a simple wavefront between the PIAA M1 and M2 optics using both it and direct integration of the Rayleigh-Sommerfeld diffraction equation (which we consider exact). A grayscale post-apodizer will be applied at M2 and the wavefronts propagated via Fourier transform to a focus where the field will be compared.

To verify the accuracy of S-Huygens, the same simple wavefront is propagated between the two PIAA optics using it and Rayleigh-Sommerfeld (RS) diffraction. The RS code is implemented in C to propagate a circularly-symmetric symmetric wavefront between two surfaces. Because of the symmetry, only the field along one radial direction on the second surface needs to be computed, and the two-dimensional RS wavefront integration only needs to be explicitly performed over one-half of the first surface and then doubled to constitute the full wavefront. The code is designed to calculate a specified subset of points along the radial direction of the 2nd surface. This allows the program to be easily ported to the JPL parallel-processing supercomputers, each having 256 processors. Separate instances of the RS program are simultaneously executed on 256 processors (128 on each on two systems), with each one dealing with one chunk of the M2 radial wavefront vector.

The test configuration used is the PIAA design previously described (*D*=90 mm M1, 0.9 m separation) working at λ = 1 μm. A simple, symmetric, aberrated wavefront (i.e., circular ripples of varying heights and spatial frequencies) is experimentally derived using repeated S-Huygens runs to produce a ~10^{-10} mean contrast level over the 2.5 – 20 λ/D field of interest. The aberrations are added to the surface profile of the 1st PIAA optic. The same optical surface profiles are used by the S-Huygens and RS programs.

The sampling used for RS on M1 is λ (Δ=1 μm; 90,000 points across the diameter, 6.4 billion points over the full surface) and the wavefront at M2 is sampled at 35,000 points along a radial vector. This configuration requires an elapsed time of ~31 hours (~332 CPU days) to compute using the RS algorithm on the 256 processors. The same radial sampling is used for S-Huygens, and it takes 70 seconds on an 8-CPU workstation (note that this is only for a single propagation given the circularly-symmetric wavefront; asymmetric wavefronts would take much longer because each azimuthal harmonic term would have to be propagated). The amplitude of the propagated wavefront at M2 is shown in Figure 15. It has high-spatial-frequency components that are produced in both the RS and S-Huygens propagations.

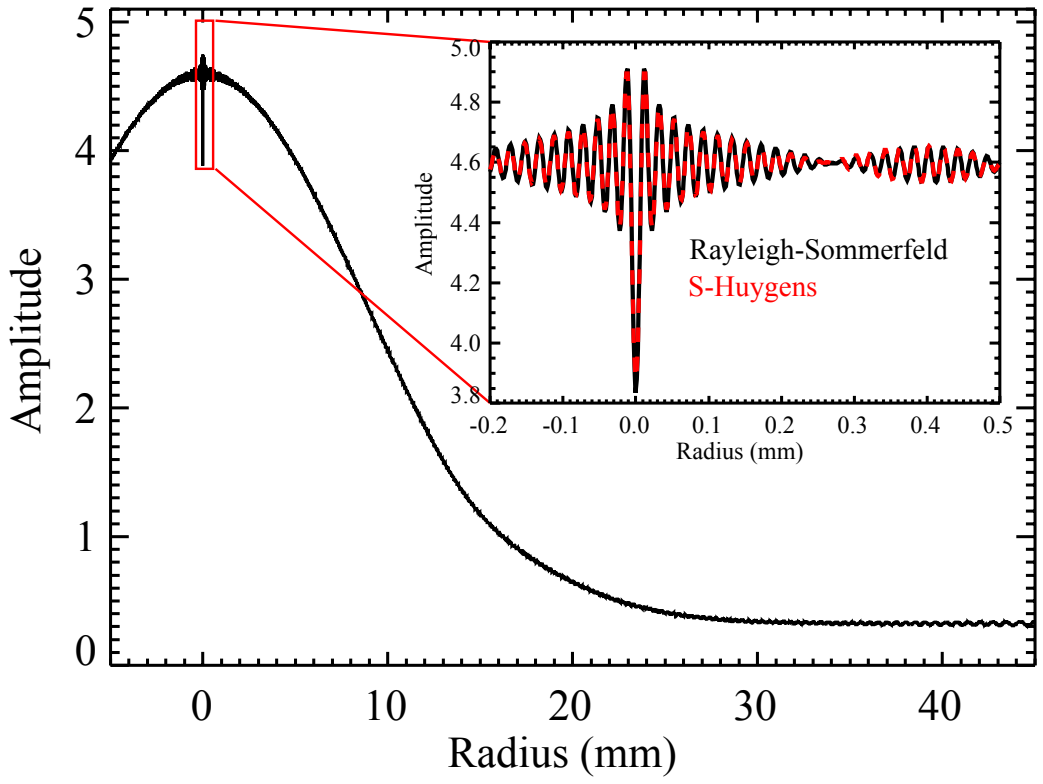

Figure 15. Amplitude at PIAA M2 of the wavefront propagated from M1 using Rayleigh-Sommerfeld and S-Huygens diffraction algorithms.

The one-dimensional RS and S-Huygens wavefronts at M2 are multiplied by a grayscale post-apodizer and then propagated to the PIAA focus using Hankel transforms (a Hankel transform computes the one-dimensional radial slice of a symmetrical, two-dimensional output wavefront at focus given a one-dimensional radial slice of an input wavefront at a pupil).

As defined in Section 3.1.9, the root-mean-square of the differences of the electric fields generated by the efficient (here, S-Huygens) and reference (RS, assumed perfectly accurate) methods measured within a 2.5 – 20 λ/D (on the sky) annulus will represent the accuracy, in terms of contrast, of the efficient algorithm.

Milestone result:

Between 2.5 – 20 λ/D, the mean contrast of the aberrated wavefront field is 1.4 × 10^{-10}, and the accuracy of the S-Huygens algorithm is 5.1 × 10^{-14}, well below the requirement of 10^{-12}. The contrast curves are shown in Figure 16.

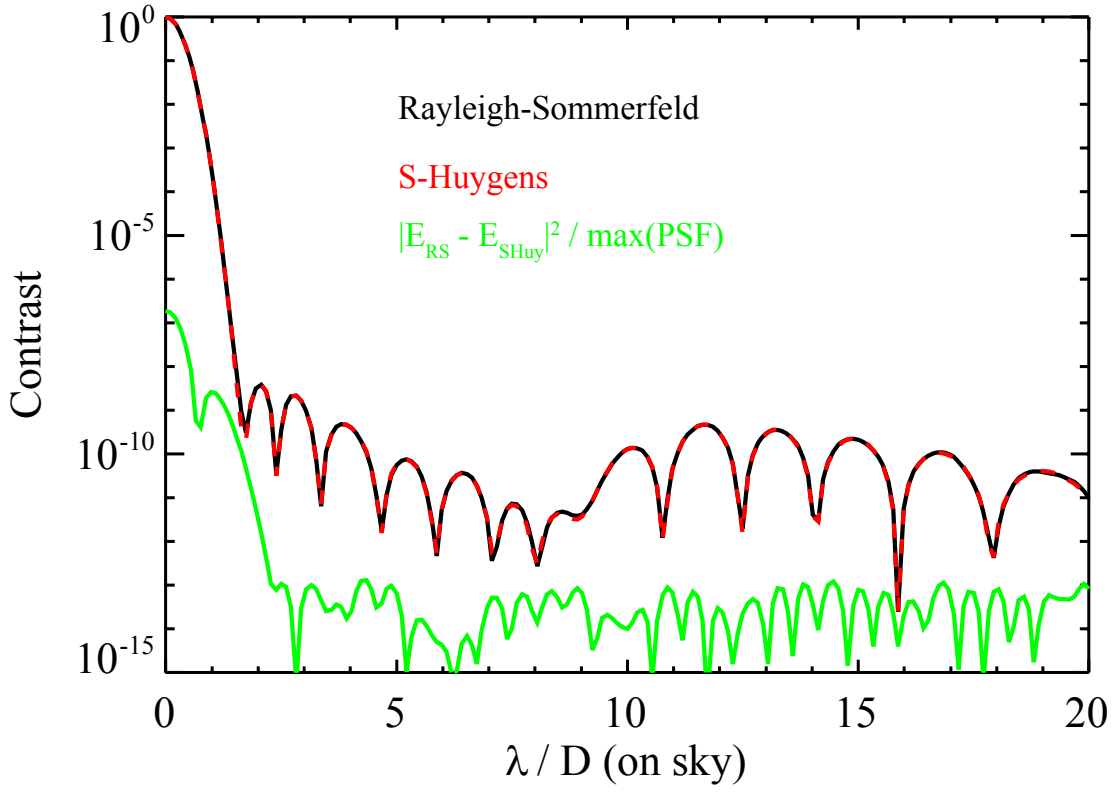

Figure 16. Contrast profiles at the PIAA focus for an aberrated wavefront propagated between PIAA M1 and M2 using S-Huygens and Rayleigh-Sommerfeld algorithms, multiplied by a grayscale apodizer, and then propagated to focus using Hankel transforms.

PIAA Angular Spectrum Propagator (PASP)

Algorithm

Pueyo et al. derived a modification of the angular spectrum propagation algorithm that can be used for PIAA systems, which we call here PIAA Angular Spectrum Propagation (PASP). It can also be applied to other optical systems where the conventional method may not work. In PASP, like the conventional algorithm, the input wavefront is decomposed into X and Y oriented, complex-valued spatial harmonics (ripples) using a Fourier transform, producing the angular spectrum. Each component is separately propagated, and the results are added together at the end. The conventional algorithm, by virtue of requiring uniform input and output wavefront sampling, can very quickly perform the propagation of all of the components at once by multiplying the spectrum by a simple two-dimensional phase function and then taking the inverse Fourier transform of the result. PASP, on the other hand, must explicitly propagate the two-dimensional representation of each ripple explicitly because of the distortion introduced by PIAA. It both accounts for distortion (remapping) of the ripple as well as the changes that occur as the ripple propagates between the two PIAA optics, including phase-to-amplitude transformations and vice-versa.

We present here the algorithm in a form that should be reasonably easy to implement in code. It propagates the field at PIAA M1 (E_{M1}) to produce the field at PIAA M2 (E_{M2}). Note that $R(\tilde{r})$ is the radius of a point on PIAA M1 that is remapped to radius $\tilde{r}$ on PIAA M2 (we assume a circularly symmetric, on-axis system), and similarly for $X(\tilde{x})$ and $Y(\tilde{y})$. The parameters are:

Z = propagation distance (separation of PIAA M1 and M2)

λ = wavelength

X, Y = location in input wavefront (at PIAA M1); $R = \left(X^2 + Y^2\right)^{1/2}$

$\tilde{x}, \tilde{y}$ = location in output wavefront (at PIAA M2); $\tilde{r} = \left(\tilde{x}^2 + \tilde{y}^2\right)^{1/2}$

$c\left(\nu_X, \nu_Y\right)$ = complex amplitude of spatial harmonic; $c(\nu_X, \nu_Y) = \text{FFT}(E_{M1})$

ν = spatial harmonic (ripple) frequency (*e.g.* cycles/meter); $\nu = (\nu_X^2 + \nu_Y^2)^{1/2}$

φ = rotation angle of the ripple in input wavefront (at PIAA M1); $\varphi = \tan^{-1}\left(\nu_Y / \nu_X\right)$

θ = angle in the output wavefront; $\theta = \tan^{-1}\left(\tilde{y} / \tilde{x}\right)$

$a(\tilde{r})$ = amplitude apodization profile produced at PIAA M2 by PIAA M1

$remapping\left(\tilde{x}, \tilde{y}, \nu\right) = 2\pi i \nu\left[X(\tilde{x}, \tilde{y})\cos\varphi + Y(\tilde{x}, \tilde{y})\sin\varphi\right]$

$$propagation(\tilde{x},\tilde{y},\nu) = -i\pi\lambda Z\nu^2\left[\frac{\tilde{r}\,a(\tilde{r})^2}{R(\tilde{r})}\cos(\theta-\varphi)^2 + \frac{R(\tilde{r})}{\tilde{r}}\sin(\theta-\varphi)^2\right]$$

$$E_{M2}(\tilde{x},\tilde{y}) = \sum_{\nu_X,\nu_Y} c(\nu_X,\nu_Y)\exp\left[remapping(\tilde{x},\tilde{y},\nu) + propagation(\tilde{x},\tilde{y},\nu)\right]$$

For a reverse PIAA system, replace M1 above with reverse M2, and M2 with reverse M1.

Note that the exponential components in the final equation are constant for a given system and wavelength, so they only need to be computed once and stored in a matrix that can be used for repeated propagations (such as computing the response function for each DM actuator). Because we assume circular symmetry in our simulated PIAA system, we can also speed things up by a factor of eight by computing a matrix corresponding to just one 45° sector of coefficient space and then flipping and transposing the matrix (or more efficiently, the coefficients and the resulting fields) as necessary to match the orientation of the other sectors.

After the wavefront at M2 has been computed it can be propagated through the system using conventional Fourier-based algorithms, from M2 to the post-apodizer to the occulter and up to the reverse PIAA. There, the PASP method must be used again but with the remapping and apodization relations appropriately modified to undistort and de-apodize the beam (the effect of the post-apodizer cannot be removed).

Using a PASP + S-Huygens hybrid to capture edge diffraction

For practical reasons we must limit the number of wavefront components that are propagated by PASP to within some specified maximum spatial frequency. Doing so means that we cannot capture the diffraction effects from the edge of the PIAA M1 optic that require very high frequency terms. A practical solution is to use S-Huygens to compute the diffraction just from the edge region of M1, ignoring aberrations. The result is rotated in azimuth to produce a two-dimensional field, assuming a symmetric, circular system. PASP is then used to propagate the aberrated two-dimensional wavefront components, and the two fields are added together. Because it is computed for an unaberrated wavefront, the S-Huygens edge diffraction field only needs to be generated once for each wavelength.

A tapered weighting function that is zero at the edge of M1 and smoothly increases to one towards the center is used to gradually reduce the amplitude of the wavefront to provide a transition between the PASP and S-Huygens regions. This "virtual apodizer", a_{virt}, is applied prior to taking the Fourier transform of the wavefront at M1 to compute the angular spectrum whose terms are propagated by PASP. The inverse function, 1-a_{virt}, is used as the amplitude along one radial direction of an unaberrated wavefront at M1 that is propagated by S-Huygens. This apodizer is also of use in reducing the spread in the angular spectrum caused by ringing from the bounded transform that would introduce anomalous spatial frequencies that would impact accuracy. The low weight along the

edge also serves to reduce the impact of computational errors at the edge of the pupil. In this study we used a cosine-tapered virtual apodizer along the outer 5% of M1 (RM1 is the M1 outer radius):

$$a_{virt}\left(r < 0.95R_{M1}\right) = 1 \quad \text{and} \quad a_{virt}\left(0.95R_{M1} \le r \le R_{M1}\right) = \frac{1}{2} + \frac{1}{2}\cos\left(\frac{\pi(r - R_{M1})}{0.05R_{M1}}\right)$$

While the transition region can be any size, it should be kept to the minimum practical thickness to allow as much of the aberrated wavefront to be propagated as possible, especially if there are significant aberrations or deformable mirror actuators near the edge of the pupil. Our experiments show that a 5% radius annulus is generally suitable, as described in a following section.

When propagating through the reverse PIAA optics only PASP is needed. At this stage the star has been suppressed so there is no edge diffraction to worry about. The remapping relations must be reversed along with the apodization function, of course, and a separate propagation matrix generated.

Wavefront sampling considerations

The wavefront is represented by a complex-valued, two-dimensional array. The sampling interval and size of the array are critical to the accuracy of the propagation. These are determined by the highest spatial frequency aberration of interest and the characteristics of the system. If one is using a deformable mirror with N_{act} actuators projected across the entrance pupil, then the highest spatial frequency error that can be corrected is (N_{act}/2) cycles/D (D is the telescope diameter). This corresponds to an outer control radius of (N_{act}/2) λ/D radians in the image plane (hereafter, multiples of λ/D will be in radians). However, errors with higher spatial frequencies than N_{act}/2 can combine to create harmonics that produce scattered light inside the control region through an effect called frequency folding. These errors must also be included in order to judge their impact on the wavefront control performance of the system. For this study, we include aberrations up to N_{act} cycles/D.

To meet the Nyquist criterion and prevent aliasing of the wavefront errors, the sampling in the wavefront array must be no coarser than ½ cycle per pixel for the highest spatial frequency, ω, and finer sampling is preferred. This means the number of samples across the beam must be at least 2ω (we used at least 2.5ω). The PIAA M1 optic compresses the beam (Figure 17), turning spatial frequencies of N cycles/D at M1 to ω=N×m cycles/D near the center of M2, where m is the compression factor (typically 3 – 5). Thus, in a PIAA system the highest spatial frequency we consider is ω=N_{act}×m. Any aberrations of >ω cycles/D must be filtered out, otherwise they will be aliased and produce erroneous results when the wavefront is compressed by M1.

With the sampling of the wavefront array now set, the size of the array is determined by the required sampling in the final focal plane. To produce a Nyquist-sampled focal plane, the diameter of the beam at the entrance pupil, D, as represented in the wavefront array

must be ½ the diameter of the array, N. Finer sampling is better so we used an array width of at least N=2.5×D.

If we assume a 48×48 actuator DM with 46 actuators across the pupil, then the highest spatial frequency that we consider important is 46 cycles/D. If our PIAA system has a compression of $m \approx 4.7$, then the beam must have at least 2.5×46×4.7 ≈ 541 samples across it in the wavefront array at the entrance pupil. This implies that the wavefront array itself must be at least 2.5×541≈1353 elements wide. For efficiency when using fast Fourier transforms, we used arrays 2048 elements wide with a beam diameter of 2048/2.5=819.2 elements.

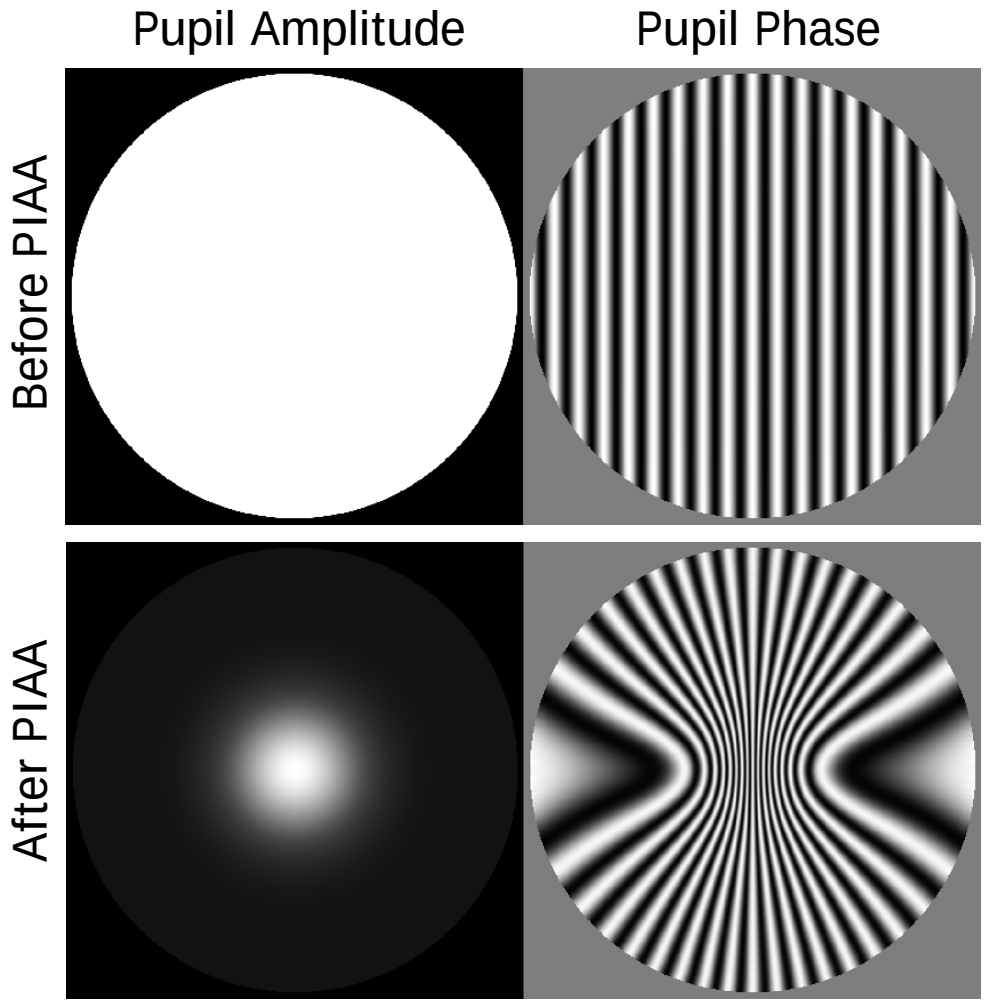

Figure 17. Demonstration of the pupil remapping effects of the PIAA optics. Aberrations of a given spatial frequency are compressed to higher spatial frequencies near the center of the wavefront.

Using Symmetry to Increase Speed

In this study we are modeling a circular, symmetric optical system. We make use of this symmetry to achieve a factor of eight improvement in efficiency in two situations:

1) Propagating harmonic terms in PASP: The wavefront at the entrance to the PIAA system is decomposed into its spatial frequency spectrum via a Fourier transform, and the two-dimensional spatial harmonic (ripple) corresponding to each X,Y spatial frequency term is generated and propagated.
2) Creating the deformable mirror response matrix: To create the deformable mirror actuator response matrix used for wavefront control (part of the Milestone #1 efficiency test), each DM actuator is poked and the corresponding change in the field at the final image plane is recorded.

Symmetry allows us to divide either the frequency spectrum or the DM into eight 45° sectors. As illustrated in Figure 18, the DM actuators in sector 5 are equivalent to those in sector 1 with a 180° rotation. For example, one can obtain the wavefront change caused by poking an actuator in sector 5 by poking the actuator in sector 1 located directly opposite of it, computing the resulting wavefront, and then rotating the result about the optical axis by 180°. With a combination of rotations and flips, one can compute the response of every DM actuator by just poking those in sector 1, reducing the number of actual full wavefront computations by a factor of 8. The same symmetry can be used when computing the harmonic terms in PASP for the frequency spectrum coefficients.

Note that the use of symmetry is only valid as long as the system is perfectly centered and there are no distortions. While these conditions can be enforced in a numerical simulation, they would not likely exist in a real optical system.

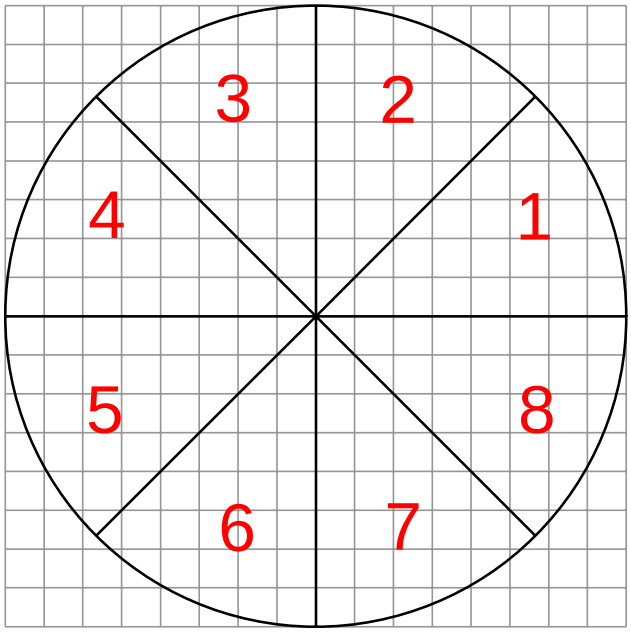

Figure 18. Schematic demonstration of the eight-fold symmetry available in a perfectly symmetrical system with a Cartesian grid (DM actuators or 2-D spatial frequency spectrum values). This schematic, for example, could represent the pupil mapped onto square DM actuators.

Post-Apodizer design and implementation

The role of the post-apodizer

The high curvature of the PIAA mirrors creates chromatically-dependent diffraction effects that prevent a pure PIAA system from reaching 10^{-10} contrast (Vanderbei et. al. 2006). High curvature of the optics is needed when the apodization that is to be created takes on small amplitudes, and high contrast apodizations are usually taper to zero at the aperture edge. These chromatic diffraction effects can be mitigated by combining the PIAA optics with a separate pupil apodizer (Pluzhnik et. al. 2006a, Belikov et. al. 2006). The PIAA optics create an intermediate apodization that is close to the desired one, except it does not have very small amplitudes. A "post-apodizer" changes this intermediate apodization into the desired high contrast one. The theoretical throughput of this hybrid PIAA-apodizer system is still high: 80%-90%, depending on the design.

Ideally, one would use smooth, grayscale transmission post-apodizers, but in reality fabrication errors and wavelength-dependent transmission effects prevent their use. The alternative is to use a binary patterned post-apodizer with regions that either fully transmit or are completely opaque. Such masks can be produced by etching of metal, either free standing or on a glass substrate. A binary apodizer can be designed to approximate the transmission profile of a grayscale one, but not with sufficient similarity to produce an equivalently high contrast field. Instead, the binary patterns are specified to interfere with the incoming light to produce a suitably dark field with a given outer radius (a perfect grayscale apodizer would produce a darker field with an unlimited outer radius).

Designing the post-apodizer

The binary apodizer we adopt here is of the "concentric ring" type. It is designed using a method similar to Vanderbei et. al. (2003) but with some differences to incorporate a minimum feature size (ring thickness) constraint. The method is essentially a linear optimization on the discrete samples $A_{bin}[n]$ of the radial apodizer profile. These values are constrained between 0 and 1 but are allowed to vary freely between those bounds as is required to preserve the linearity of the optimization. However, as it turns out, most of these values converge to either 0 or 1 naturally, creating an (almost) binary sequence of $A_{bin}[n]$ which in two dimensions corresponds to a series of alternating transparent and opaque concentric rings of different widths. The sample of the apodizer that is closest to the edge of each ring actually does not converge to 0 or 1. Its value is used to determine the exact width of the ring. A summary of the algorithm is as follows:

1. Set the apodizer sampling spacing equal to minimal feature size.
2. Perform the linear optimization as described in Vanderbei et al. (2003), producing the sampled apodizer $A_{bin}[n]$.

3. For each ring edge, find the (non-binary) sample closest to that edge. Let these samples be at indexes $S = \{n_1, n_2, ..., n_N\}$.
4. (This step is not always necessary and in fact was not necessary for the apodizer we created for this work.). If $n_{k+1} - n_k = 1$ for any n_k, n_{k+1} in S, remove these samples from S (and round $A_{bin}[n_k]$ and $A_{bin}[n_{k+1}]$ to 0 or 1). This prevents rings thinner than the minimum feature size. However, it usually degrades contrast, which is repaired by repeating the linear optimization, but only on $A_{bin}[n]$ for n in S. Now we have a sampled apodizer satisfying the minimum feature size constraint.
5. Compute ring radii analytically from the resulting $A_{bin}[n]$ by equating the integrated area of the rings in the two apodizers. This generates an error in the resulting PSF, but in practice this error is usually below 10^{-10} contrast.

Iterating the design

Some iterating is required to arrive at the final apodizer design used in this study. There are certain requirements: the field created by the unaberrated PIAA optics and the binary post-apodizer must have a mean contrast of $<10^{-10}$ inside of 20 λ/D radians over a wavelength range of λ=500-600 nm, and the minimum post-apodizer ring thickness must be >4 μm for a 108 mm diameter apodizer, to ease fabrication.

One of the early attempts is shown in Figure 19. Note that it includes rings of transmission at the edge of the beam at PIAA M2 (which has a radius of 54 mm). A one-dimensional, highly-sampled (<1 μm/element) representation of the PIAA-apodized field at M2 was multiplied by the discretely-sampled post-apodizer. This was Hankel-transformed to the PIAA focus (the Hankel transform takes a 1-D slice of a 2-D input field and produces the 1-D slice of the 2-D Fourier transform output field, for a circularly-symmetric system). The field at M2 was generated either by the smooth, ideal, geometric PIAA apodization function (no diffraction) or by propagating an unaberrated, one-dimensional field from M1 to M2 using S-Huygens (diffraction included).

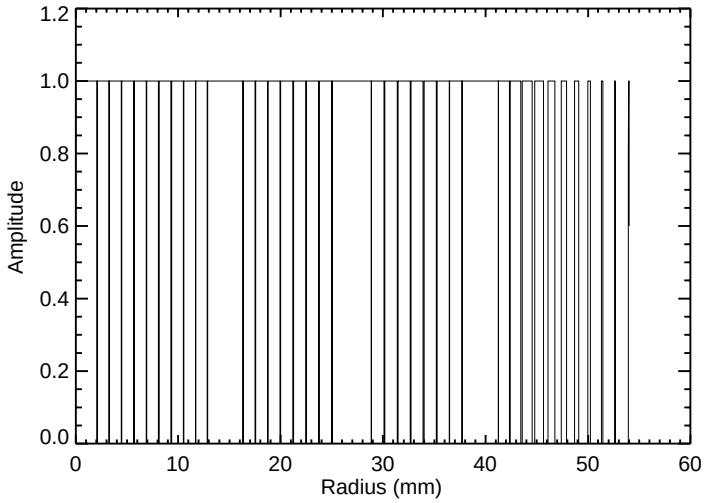

Figure 19. Radial plot of the version 2 binary post-apodizer (M2 beam radius = 54 mm).

There are considerable differences between the two fields when compared at the PIAA focus (Figure 20). The non-diffractive M2 field produces considerably better contrast than the diffractive S-Huygens ones. This suggests that this post-apodizer is poorly optimized for broadband performance and very sensitive to the diffraction effects, which is demonstrated by the λ=500 nm and 1000 nm profiles. Figure 21 shows the wavelength-dependent diffraction at the edge of the beam at M2, where there are narrow rings in this particular post-apodizer. The outer region of the field must therefore be fully masked to reduce chromatic effects. Besides having an unsatisfactory chromatic sensitivity, this post-apodizer also does not provide a sufficient contrast and outer radius.

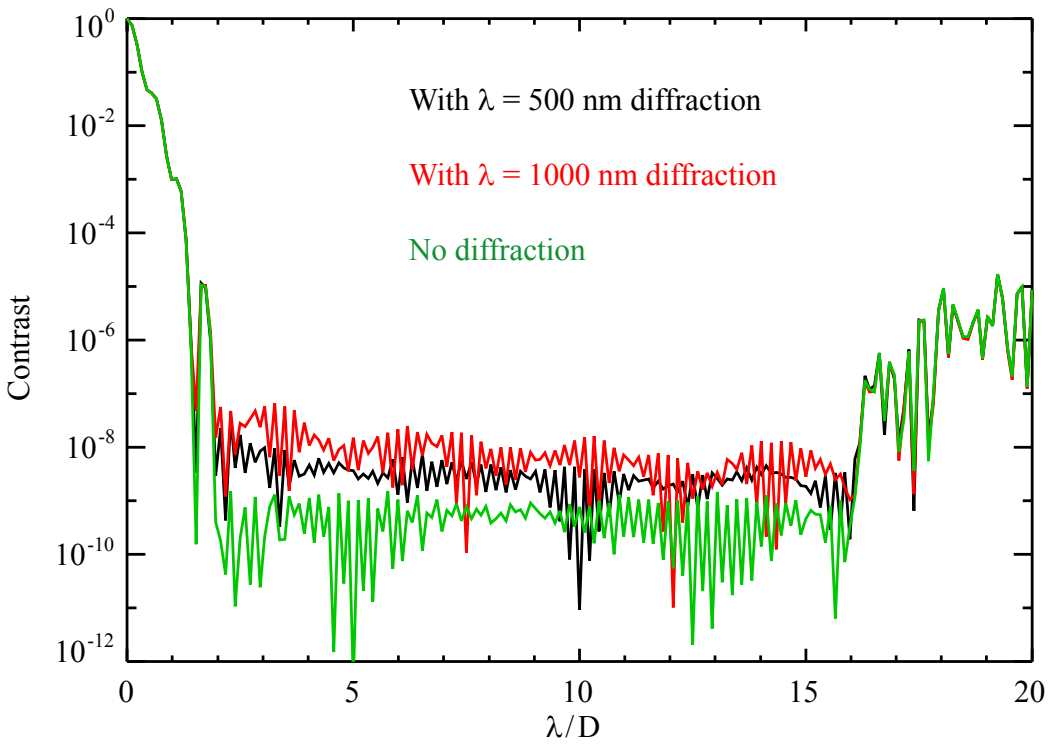

Figure 20. Contrast at the intermediate focus of PIAA (occulter plane) using the binary post-apodizer version 2. The green line was computed using the ideal (geometric, non-diffractive) field at M2 multiplied by the post-apodizer and then Hankel-transformed to the PIAA focus. The other lines are the same except that diffractive fields at λ=500 nm (black) and λ=1000 nm (red) at M2 computed by S-Huygens were used. The differences show that the version 2 post-apodizer is not optimized for broadband performance (besides not offering a sufficiently dark or wide field).

Figure 22 shows the final post-apodizer design (version 7). This mask has no transmission beyond a radius of 43 mm (the beam radius at M2 is 54 mm). As demonstrated in Figure 23, the agreement between the diffractive and non-diffractive results for this mask are much closer and the broadband performance is better. This mask meets the requirements for contrast, outer working angle, and minimum ring thickness (4

μm). The contrast in the final coronagraph image plane (after the reverse PIAA) is even better (about a factor of 5) because the field point sources are much sharper after the wavefront distortion is removed.

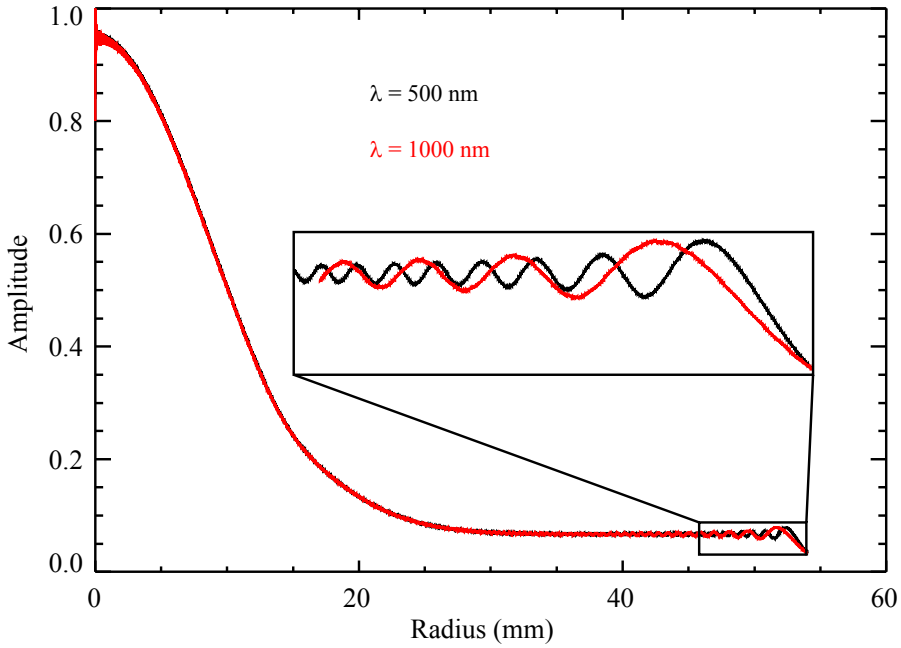

Figure 21. The amplitude at PIAA M2 at λ=500 nm (black) and λ=1000 nm (red) as computed by S-Huygens. As the inset demonstrates, there are significant differences near the edge of the beam.

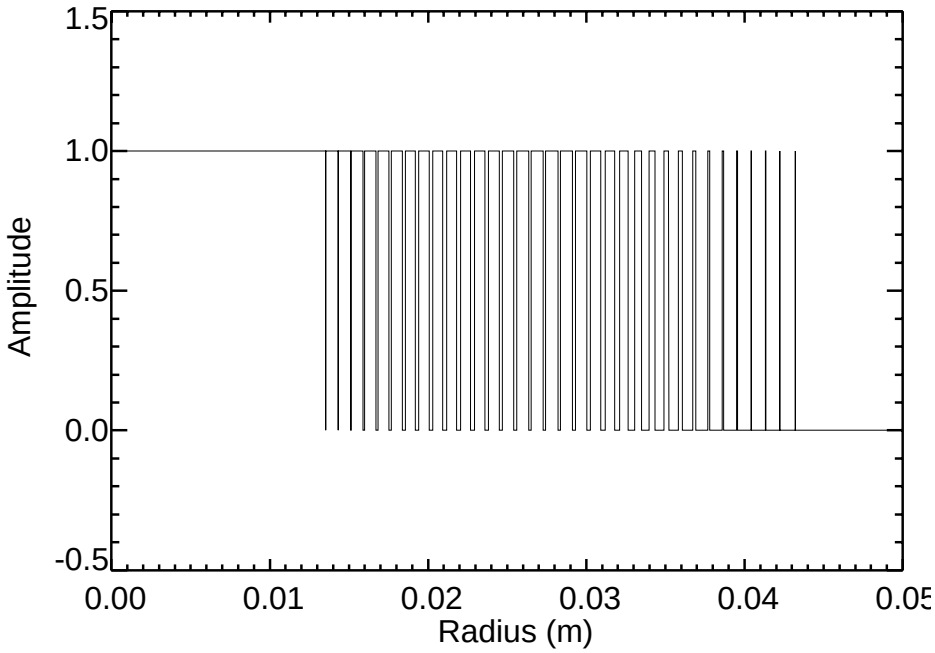

Figure 22. Radial amplitude transmission of the final (v. 7) post-apodizer design.

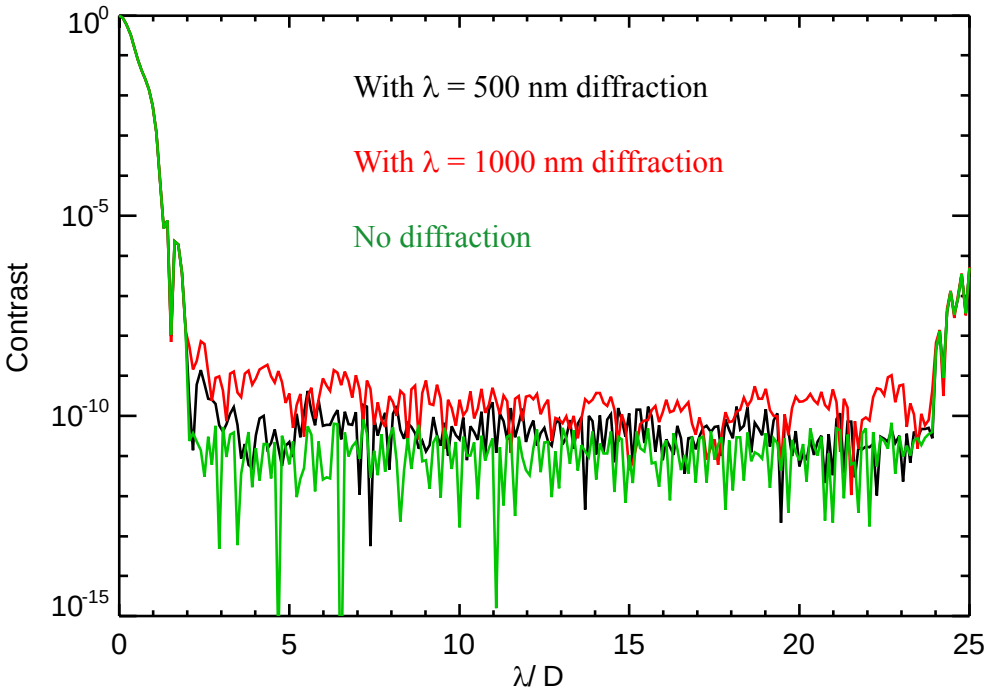

Figure 23. Same as Figure 20, but using the final post-apodizer design (version 7).

Representing the binary post-apodizer at lower sampling

Another PIAA-specific requirement in the Milestone #1 Whitepaper states:

> ***Milestone statement:***
>
> ***We will verify that the binary post-apodizer approximation and discrete binary representation are equivalent down to a 10^{-10} contrast level by propagating a simple wavefront though both types using S-Huygens with fine sampling.***

The thin rings that comprise a binary apodizer cannot be directly represented with the sampling typically used in 2-D wavefront propagation (e.g. ~100-1000 samples across the pupil). Representing the rings using antialiased patterns, where transmission in each pixel is proportional to the pixel area covered by a ring, has been shown to provide insufficient accuracy.

One way to implement the binary apodizer with limited sampling is to compute a spatially-filtered version of a high resolution mask, creating a non-physical representation

(one with negative values) that produces the same net effect in the image region of concern. To construct it, a very high resolution (tens of thousands of points), one-dimensional array representing the uniform transmission across the wavefront is created, and elements are set to zero where masked by the apodizer rings. This array is then Fourier transformed to produce the frequency spectrum of the apodizer (alternatively, the spectrum can be analytically computed). The spectrum is multiplied by an apodized windowing function to limit the spatial frequencies to those that can be sufficiently represented with the lower resolution of the 2-D wavefront array. The resulting 1-D spectrum array is inverse Fourier transformed, creating the filtered apodizer profile that is then interpolated onto a 2-D grid.

To create a filtered version of the post-apodizer, a 1-D array with a sampling of 0.225 μm/pixel (240,000 pixels across the radius of M2; the thinnest ring is 4 μm) is used to represent the binary apodizer and is then Fourier transformed. A tapered windowing function is then applied to this spectrum with an amplitude of 1.0 out to a spatial frequency of 48 cycles/diameter with a cosine-squared taper from 48 – 54 cycles/diameter. After transforming back to apodizer space, the resulting field contains oscillations and excursions with negative real values. This filtered apodizer is shown in Figure 24.

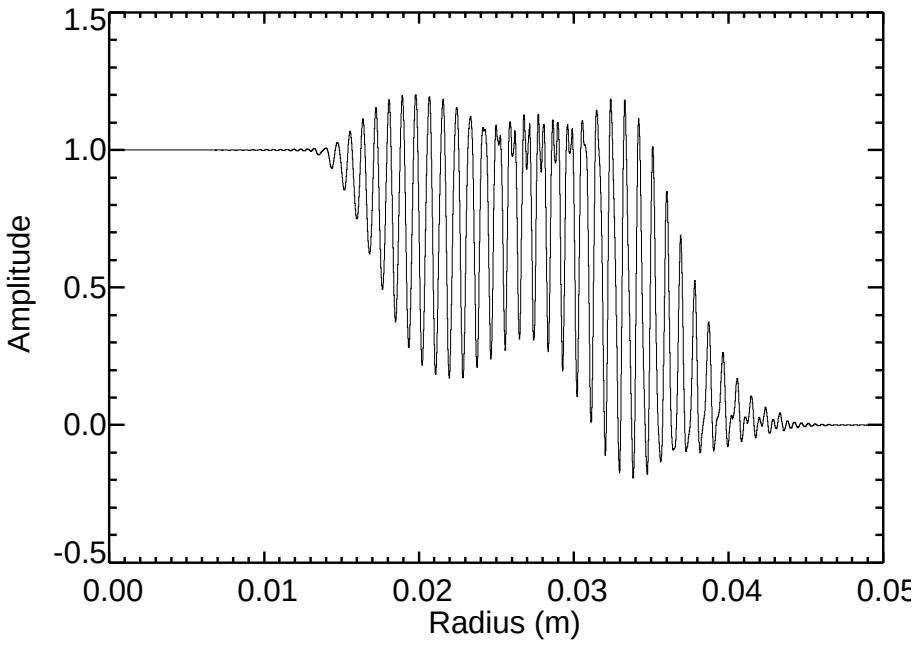

Figure 24. Radial amplitude transmission profile of the filtered version 7 binary post-apodizer. Note that the profile contains negative values.

To determine the validity of the filtered apodizer, a 240,000 element, 1-D radial wavefront at M2 is created using the ideal PIAA-induced apodization profile (no diffraction, no errors). This is multiplied by the finite binary apodizer and Hankel transformed to the PIAA focus. The same ideal wavefront is separately multiplied by the filtered version of the apodizer and transformed to focus. The contrast profiles for these two versions agree well (Figure 25). Note that a perfect grayscale apodizer applied to the same wavefront produces better contrast with an unlimited outer radius.

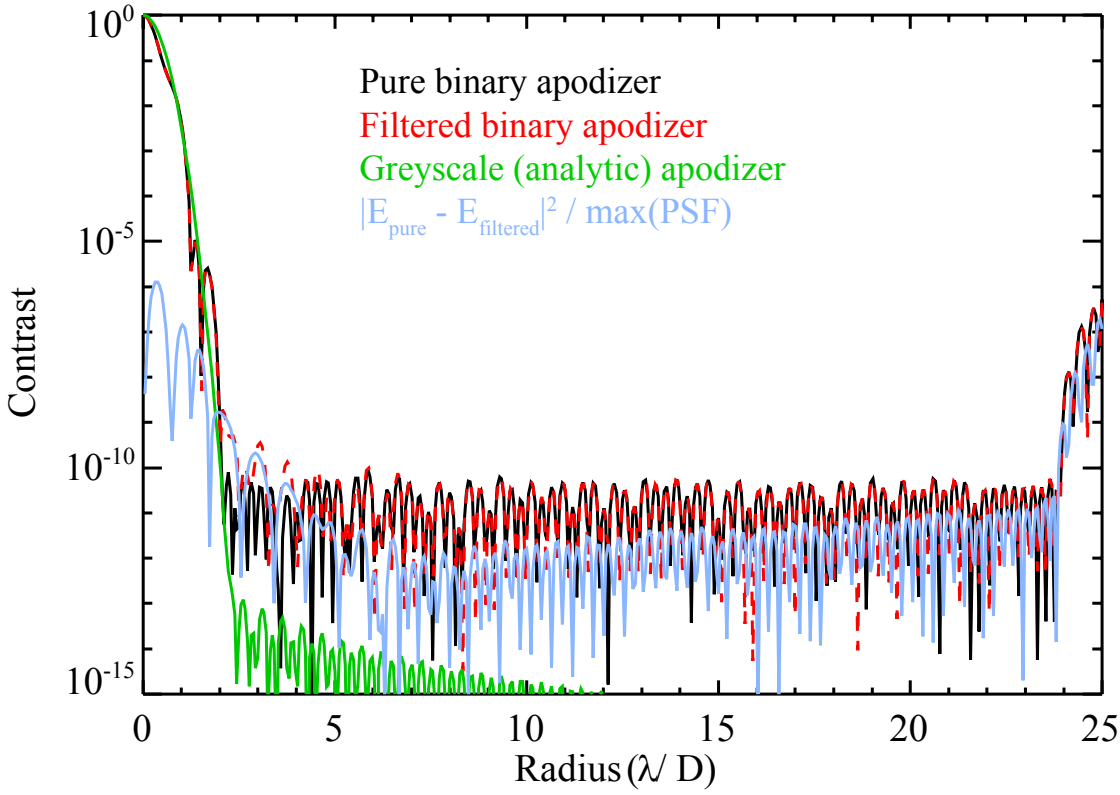

Figure 25. Contrast at the PIAA focus obtained by multiplying the highly-sampled, geometric (non-diffractive) PIAA apodization profile by the discrete binary post-apodizer, the spatially filtered binary post-apodizer, and the perfect grayscale one. Except at small angles (<3.3 λ/D), the finite and filtered apodizer fields agree to a contrast of below 10^{-10}.

The next step is to verify the two-dimensional representation of the filtered apodizer. First, a 1-D, 90,000 element (0.5 μm/element), aberrated wavefront at M1 is created. The aberrations (a pattern of radial ripples on M1) are iteratively derived to produce a ~10^{-10} contrast field at the PIAA focus at λ=500 nm. This is propagated via S-Huygens to produce a 60,000 element (0.9 μm/element) aberrated field at M2. One copy of this field is multiplied by the finite binary apodizer and Hankel transformed to the PIAA focus. The second copy is converted into a 2-D field by downsampling via binning and

interpolation. Two instances of this were tested: the first with 1434 samples across M2 and embedded in a 2048 × 2048 array, and the second with 2867 samples across a 4096 × 4096 array. These fields were multiplied by a similarly sampled version of the filtered apodizer and then Fourier transformed to the PIAA focus.

To determine the contrasts in the final image plane, we continue propagation beyond the PIAA focus. First, the fields at the PIAA focus are multiplied by a central occulter of radius 2.1 λ/D and an outer radius mask of 20 λ/D. They are then transformed (Hankel or Fourier) to the reverse PIAA optics, where a mask with a 96% opening diameter is applied (to mask residual diffraction effects from the binary post-apodizer). Then, rather than explicitly propagating between the reverse optics, interpolation is used to reverse the wavefront distortion and PIAA-induced apodization. The fields are then propagated to the final focus and the contrasts computed (a 1-D slice is taken from the 2-D array for comparison).

The fields are shown in Figure 26. The 1-D field has a contrast of 9.6×10^{-11}. The 2-D contrast is 1.0×10^{-10} for the 2K array and 9.9×10^{-11} for the 4K one. The RMS of the differences between the 1-D and 2-D electric fields, converted to contrast, are 5.0×10^{-12} (2K array) and 8.7×10^{-13} (4K array), corresponding to accuracies of 5.0% and 0.9%, respectively. The same wavefront error pattern, adjusted to provide $\sim10^{-5}$ contrast (9.9×10^{-6}, specifically), produces accuracies of 0.3% for both 2K and 4K arrays.

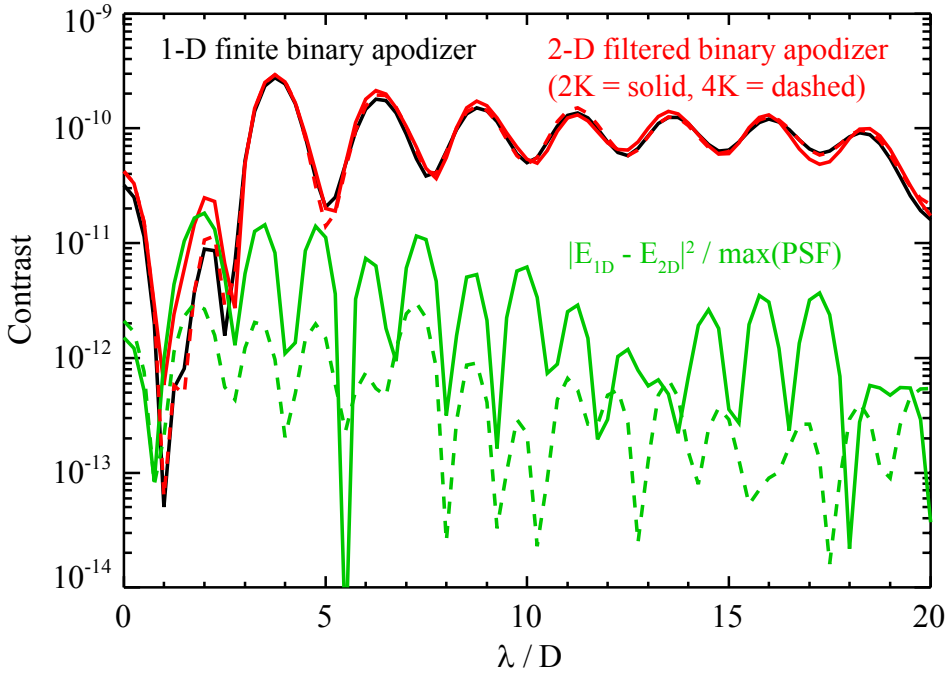

Figure 26. Contrast at the final image plane of a 10^{-10} equivalent-contrast aberrated field (circularly-symmetric) propagated using S-Huygens through the forward+reverse PIAA coronagraph system. The black line is the highly-sampled 1-D propagation using a finite binary apodizer. The red lines are 2-D propagations using a filtered binary apodizer. The green lines are the differences between the 1-D and 2-D electric fields converted to contrast. For the 2-D profiles, the solid lines are for 2K arrays, dashed for 4K.

Increasing the spatial frequency cutoff in the filtered apodizer from 48 cycles/D to 68 does not significantly improve the accuracy. Increasing to 88 cycles/D actually degrades the accuracy (as does reducing the cutoff to 38 cycles/D).

Instead of using the binary post-apodizer in the 1-D and 2-D propagations, we do another set of experiments with the graded transmission (smooth transmission profile) post-apodizer (the combination of the geometric PIAA-induced apodization and the graded transmission profile produces the ideal apodization). Using the same aberrations, the 1-D result has a mean contrast of 9.4×10^{-11} and the 2-D is 9.5×10^{-11}. The RMS of the difference of the electric fields is 2.6×10^{-13}, an accuracy of 0.3%. This demonstrates that the 2-D, filtered representation of the binary post-apodizer sets the accuracy of the PIAA simulations for 10^{-10} contrast fields.

Effect of edge weighting

As detailed previously, propagation of the field at the PIAA M1 optic is split between S-Huygens, which computes the edge diffraction absent any aberrations, and PASP, which handles the interior wavefront including aberrations. The allocation is weighted by a cosine taper over the outer 5% of the radius (i.e., S-Huygens is used only for the outer 5%). Aberrations in this outer 5% are thus effectively reduced, approaching zero at the field edge.

To demonstrate the impact of not including these edge errors in full, two M1 fields were propagated at high resolution using S-Huygens from M1 to M2 at high resolution (90000 points along the radius of M1 at $\lambda = 500$ nm, 240000 points along M2). The fields were identical except that in one the phase errors were downweighted with the same 5% outer radius cosine taper used in the milestones. The errors (Figure 27) were derived from the azimuthal average surface error profile from a measured PIAA M1 optic, which has high deviations along the edge. The profile was normalized to produce a $\sim 10^{-10}$ mean contrast field. The fields at M2 were multiplied by the finite binary apodizer derived in the previous section and then Hankel transformed to the focus of the PIAA system. An occulting mask was applied and the fields propagated via Hankel transform to the reverse M2, where the wavefront was multiplied by an aperture stop and remapped via interpolation. The fields were then finally propagated to final focus.

The mean field contrast (Figure 28) is 1.0×10^{-10}. The RMS of the difference of the electric fields converted to contrast is 6.1×10^{-16}. This demonstrates that there is no significant impact from downweighting the outer 5% of the field errors. This is a consequence of the PIAA design. The outermost region of the M1 optic is intentionally curved to disperse light over a wider region along the outer region of M2 to reduce the effect of edge diffraction. The post-apodizer masks the large majority of this light.

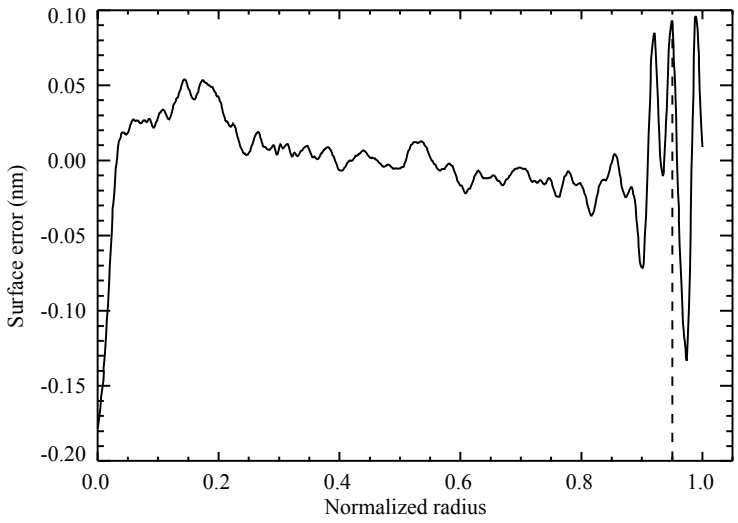

Figure 27. Plot of the radial surface error profile used for the edge weighting test. The errors have been normalized to produce a 10-10 mean contrast field. The dasked line at 0.95 indicates where the edge weighting would begin.

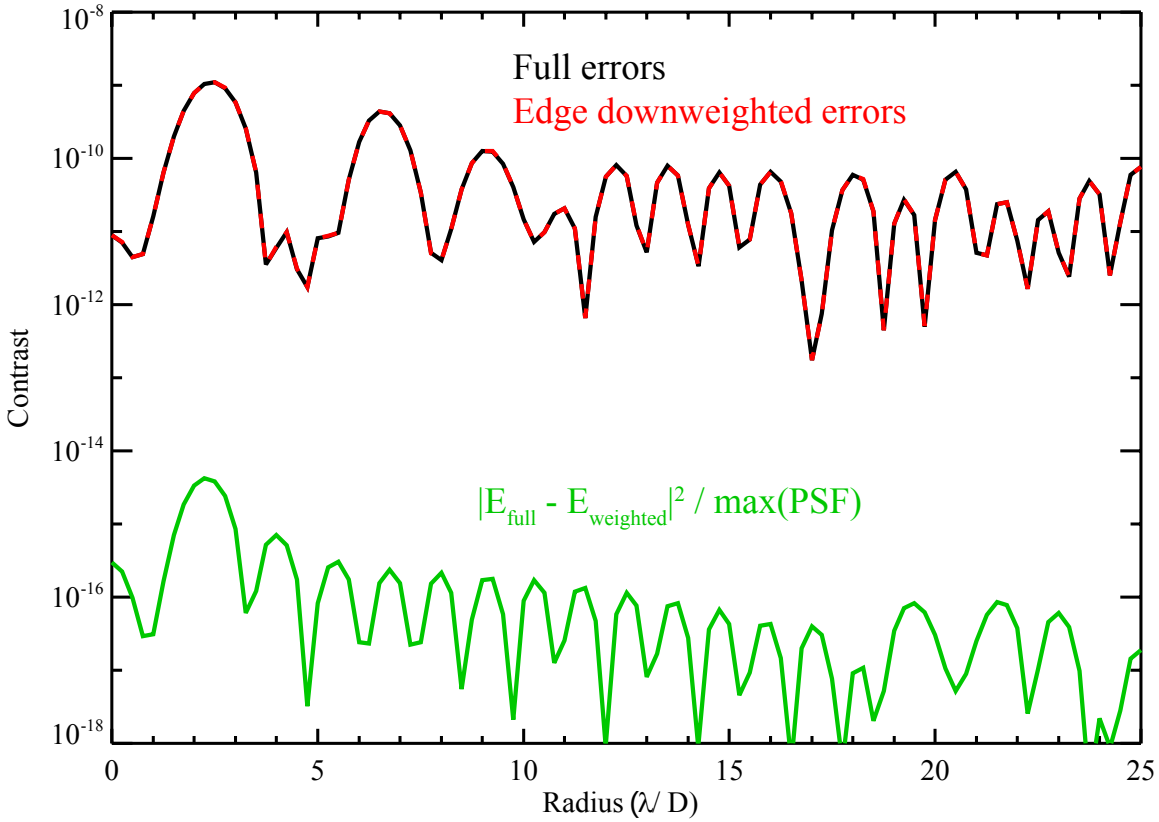

Figure 28. Comparison at the final image plane of aberrated fields at M1 propagated by S-Huygens (λ = 500 nm). The black line is the result of using the full error term at M1, and the red is the same field with the outer 5% radius downweighted with a cosine taper. The green line is the difference of the electric fields converted to contrast.

PIAA Milestone #1: Accuracy of the PASP algorithm

The purpose of Milestone #1 is to show that the efficient algorithm is sufficiently accurate to predict how the coronagraph behaves to wavefront aberrations and their control using deformable mirrors. This is done by propagating an aberrated wavefront through the coronagraph using both efficient and known-to-be-accurate algorithms. The fields at the final focus produced by both are differenced and converted to intensity described in terms of contrast. For PIAA, PASP is the efficient method and S-Huygens is the reference.

Wavefront Error Map

The phase aberration map, shown in Figure 7 was created to produce a generally uniform-intensity speckle field in the coronagraphic image plane. As discussed, the PIAA beam remapping compresses the wavefront towards the center of the M2 optic, increasing the spatial frequency of an aberration. Therefore, the error map is filtered to reject spatial frequencies above 48 cycles/diameter to prevent aliasing during beam compression. This is sufficient to include aberrations that directly create speckles inside the dark hole region (those of ~20 cycles/diameter or less) as well as higher spatial frequency ones that create speckles in the dark hole due to frequency folding. The wavefront errors are scaled to produce the desired contrasts inside the dark hole area.

Coronagraph Layout and Propagation Sequence

To evaluate the accuracy of the PASP algorithm, the simple PIAA coronagraphic system shown in Figure 13 is used, with a lens at reverse PIAA M1 to focus at the final image plane. The reverse PIAA uses the same optical setup as the forward version but in reverse order. The entrance pupil in this system is at the PIAA M1 optic, with a uniformly-illuminated input wavefront that includes the phase aberration map. The PASP algorithm is used to propagate from PIAA M1 to M2. At M2 the wavefront is multiplied by the filtered version of the binary apodizer. PROPER is used to propagate to a focusing optic, the occulter, the collimating optic, and then on to the reverse PIAA M2 optic. PASP is used to propagate from reverse PIAA M2 to reverse PIAA M1, where a focusing lens is applied. PROPER then propagates the wavefront to the final focus.

PASP Initialization (edge propagation, correction function)

Prior to using PASP for the forward PIAA, the outer unaberrated edge of M1 (weighted by the inverse of the taper used for the interior wavefront in PASP) is propagated to M2 with S-Huygens (Figure 29, right) and stored for later use (it only needs to be computed once per wavelength). Because of the limited number of spatial frequencies propagated by PASP for efficiency reasons, a PASP+S-Huygens propagation of an unaberrated wavefront is not equal to that produced by a pure S-Huygens propagation of the entire

wavefront (Figure 30). To compensate for this, the interior portion (taper weighted at the edge) of an unaberrated M1 wavefront is propagated to M2 using PASP and added to the S-Huygens-generated M1 edge wavefront at M2 to create a complete M2 wavefront. Separately, the full unaberrated M1 wavefront is propagated to M2 using S-Huygens (Figure 30, left). The difference between these fields produces a correction function (Figure 31) that compensates for the PASP limitations for the perfect, unaberrated term of the wavefront (which contains most of the energy in low-aberration systems like those we study here). This function needs to be computed only once at each wavelength and stored.

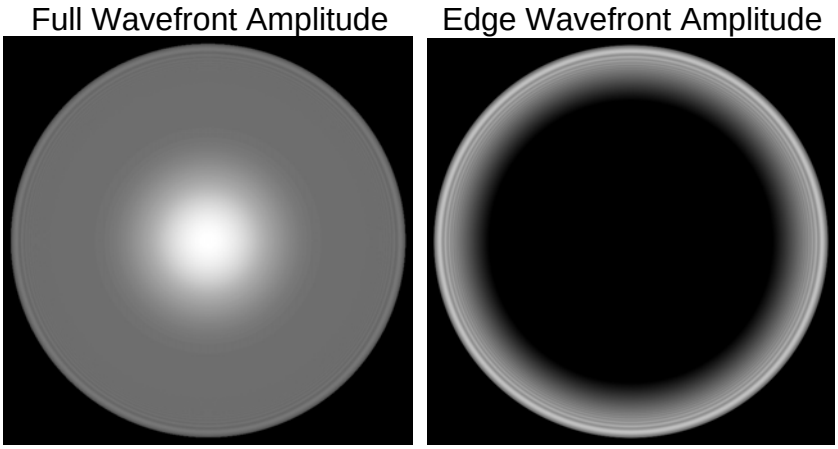

Figure 29. Unaberrated wavefront amplitude at M2 computed by S-Huygens for (left) the full M1 wavefront, and (right) the outer 5% of M1 weighted by the inverse taper.

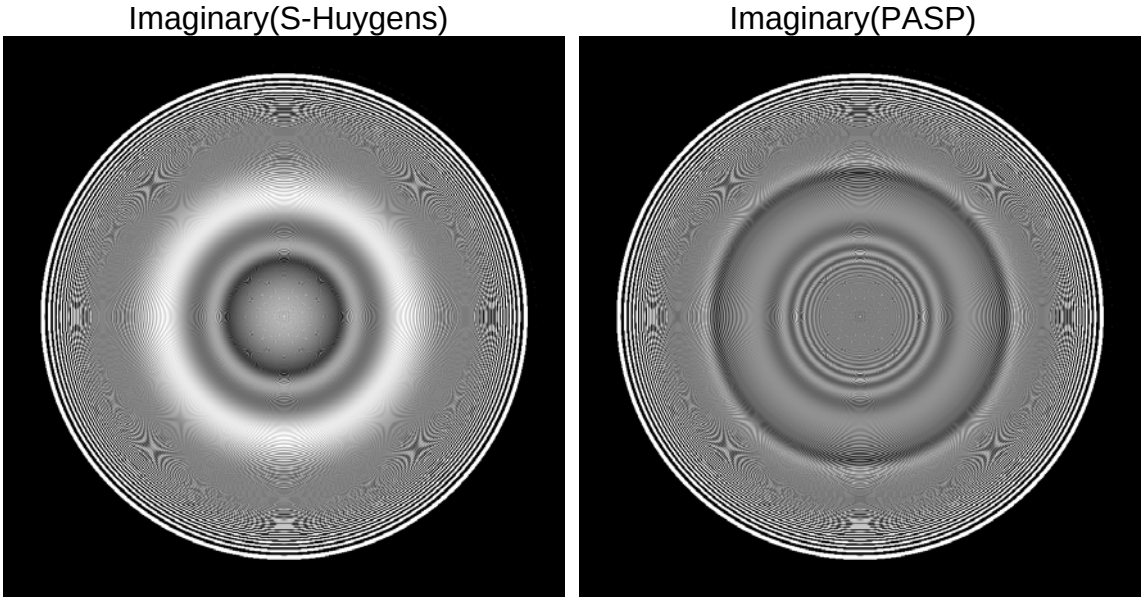

Figure 30. Imaginary (~phase) component of the wavefront at M2 for an unaberrated wavefront at M1 propagated using (left) only S-Huygens and (right) PASP+S-Huygens.

Figure 31. Phase difference at M2 between the PASP(interior)+S-Huygens(edge) and S-Huygens(full) propagations for an unaberrated wavefront.

Stage 1. Propagation from forward PIAA M1 to M2+post-apodizer

Propagation using PASP+S-Huygens

The complex-value, phase-aberrated wavefront is defined at M1 with a pupil 819.2 pixels in diameter (0.1099 mm/pixel) embedded in a 2048 × 2048 array (i.e., the pupil occupies 40% of the diameter of the grid). It is then multiplied by the 5% edge tapering function. A Fourier transform converts the wavefront into its frequency spectrum with a sampling of 0.4 cycles/diameter, and aberrations with frequencies beyond r = 60 cycles/D are zeroed. The spectrum is then segmented into 45° sectors, and each sector is flipped/rotated/transposed as needed to match the orientation of the propagation matrix, which is computed for only one sector.

The propagation matrix is composed of one 460 × 460 element (458 element beam diameter) complex-valued wavefront array for each non-zero spatial frequency in the spectrum sector. It contains the corresponding harmonic (ripple) that has been propagated using PASP. As there are approximately 8800 terms in a sector, this amounts to an array size (double-precision complex values) of 460 × 460 × 8800 × 16 bytes/element ≈ 30 Gigabytes.

For each spectrum sector, the propagation matrix wavefront are multiplied by the corresponding spectrum terms using a custom C routine called from IDL. This code is multithreaded and performs the multiplication twice as fast as can be done with IDL's operators (which are themselves multithreaded but not optimized for this particular data layout). The sum of the wavefronts for each sector is rotated/transposed/flipped as

necessary to return the wavefront to the nominal orientation. The wavefronts from all the sectors are then added together to create the field at M2. The 460 × 460 pixel wavefront is interpolated to create a 1433.6 pixel diameter wavefront embedded in a 2048 × 2048 array (the higher pupil sampling, chosen by experimentation, is required to adequately handle the higher frequency terms during propagation introduced by the filtered binary apodizer). The similarly-sampled, unaberrated edge wavefront previously generated by S-Huygens is read in and added to create the complete wavefront at M2. The correction function is also added.

Propagation using S-Huygens Only

The same 2048 × 2048 element (819.2 element pupil diameter) M1 aberrated wavefront array used above for the PASP+S-Huygens propagation is used for the pure S-Huygens propagation. For each pixel-increment radius in the array, the one-dimensional wavefront at all azimuths at that radius is extracted via interpolation. The *C(k,r)* terms for that radius are derived as described earlier in the description of the S-Huygens algorithm. The terms are computed for k = -175 to +175 cycles/360°, which contain the highest spatial frequencies of concern at the edge of the pupil. The frequency spectrum is shown in Figure 32. It has the dimensions $r \times k = 820 \times 351$ elements.

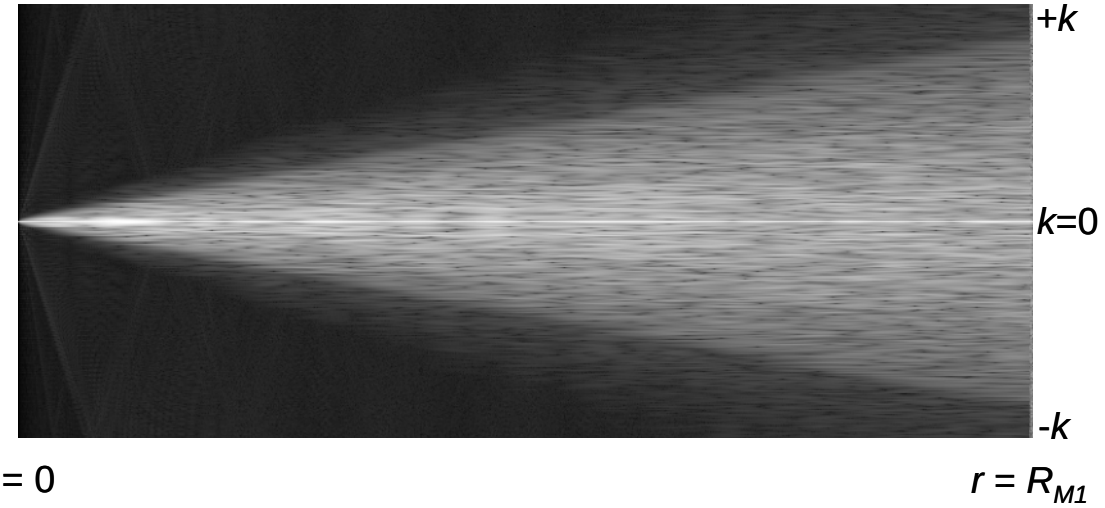

Figure 32. Azimuthal spatial frequency spectrum $C(r,k)$ of the aberrated field at M1. The dominant k=0 term has been truncated in intensity and the contrast stretched to show the other components.

The *C* radial vector for each k is interpolated to higher sampling and propagated using S-Huygens. To improve speed, only terms greater than 5× the median value of the entire *C* array are used. For the k = 0 term, which contains most of the energy and has the large edge discontinuity, 100,000 points are used along the radius of M1 and M2. For other k's, reduced sampling can be used, so 25,000 points are used along the radius of M1 and 15,000 along the radius of M2. The results are downsampled to the same resolution used in the PASP propagation and converted to two-dimensional wavefronts by multiplying by

the appropriate azimuthal harmonic function. These are added together to form the M2 wavefront (Figure 33 and Figure 34).

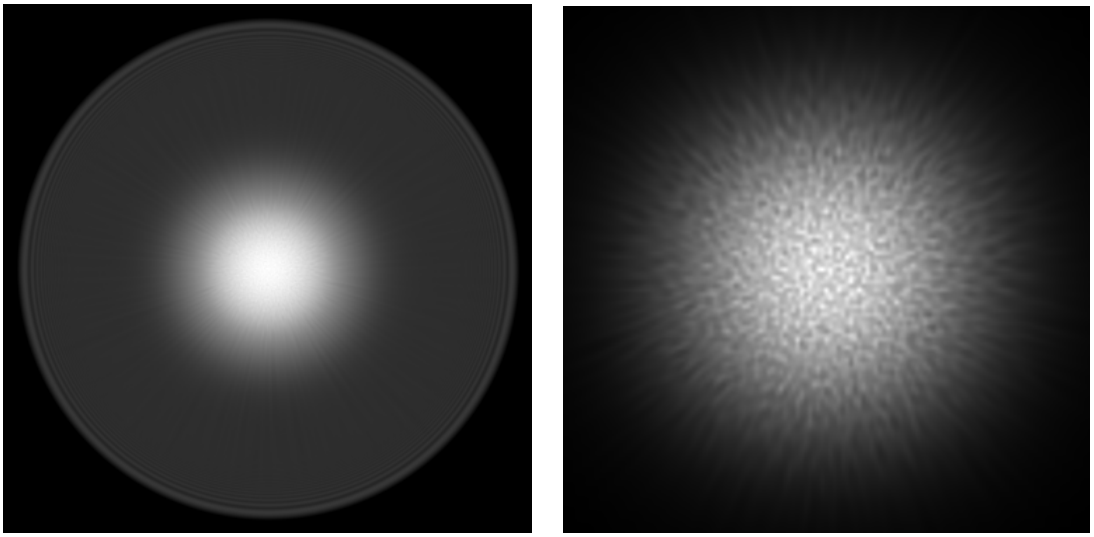

Figure 33. (Left) The amplitude of the aberrated field at M2 produced by S-Huygens; (Right) Magnified center portion of the field. The amplitude errors are generated by phase errors at M1.

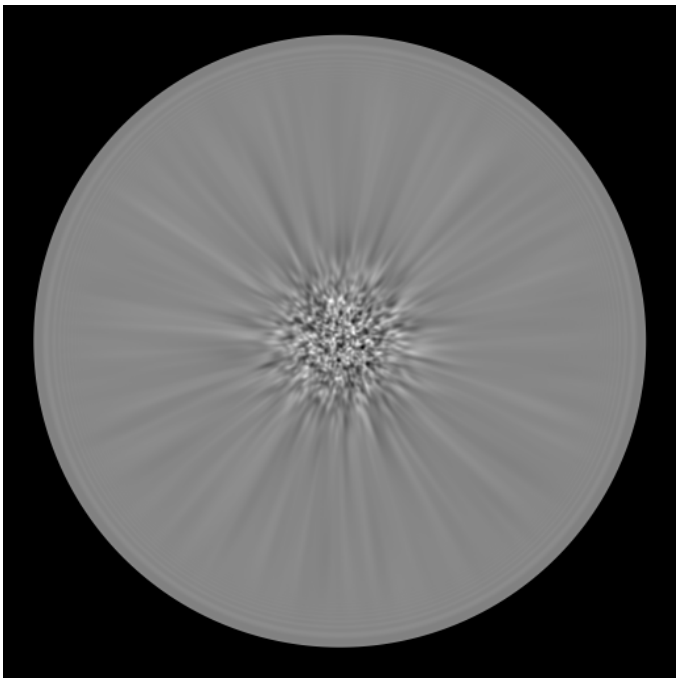

Figure 34. The PASP-computed imaginary term (phase) of the aberrated wavefront at PIAA M2.

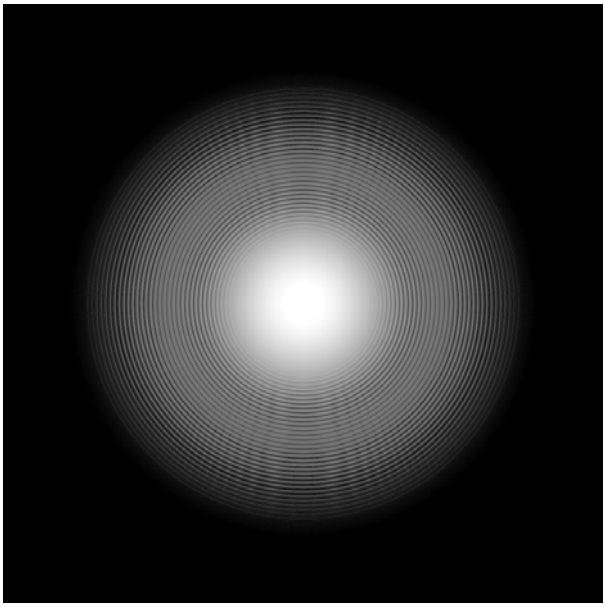

Figure 35. Amplitude at the post-apodizer.

Stage 2. Propagation from M2+post-apodizer to the occulter to the reverse PIAA

In the second stage, the PASP and S-Huygens fields at M2 are propagated using PROPER to the occulter and then on to the reverse PIAA system.

First, the wavefront is multiplied by the filtered post-apodizer (Figure 35), which is at M2. A lens is placed at M2, and the field is propagated to the PIAA focus (Figure 36). The post-apodizer creates a diffraction pattern at the focus with bright rings outside of a certain radius. The field is truncated beyond r = 22 λ/D to prevent this light from scattering at subsequent surfaces. The central core is also masked inside of r = 2.4 λ/D. The field is propagated to the reverse PIAA M2 optic (Figure 37), where a lens collimates the beam. An aperture stop with a 75% clear diameter is applied to mask additional diffracted light created by the post-apodizer. Because this is applied to the already heavily apodized beam, it has a minor impact on throughput.

Stage 3: Propagation from reverse PIAA M2 to reverse PIAA M1

Propagation using PASP

The field at the reverse PIAA M2 optic produced by PROPER using the PASP+S-Huygens generated forward PIAA M2 wavefront is propagated to reverse PIAA M1 using only PASP. Because the 0^{th} order component (uniform pupil illumination) of the initial wavefront has been removed by the occulter, there is no need for a separate propagation of the edge region using S-Huygens. As with the forward PIAA propagation, the reverse M2 wavefront is converted into it spatial frequency spectrum via a Fourier

transform. Because the field at the occulter was masked beyond 22 λ/D, the spectrum contains spatial frequencies below 22 cycles/D. The propagated harmonic component wavefronts (460 × 460, as before) are added together and interpolated to form the reverse M1 wavefront with 819.2 pixels across the beam embedded in a 2048 × 2048 array. Note that in the reverse PIAA, the apodization $a(\tilde{r})$ in the PASP algorithm is actually the *deapodization* function that reverses the forward-PIAA-induced apodization.

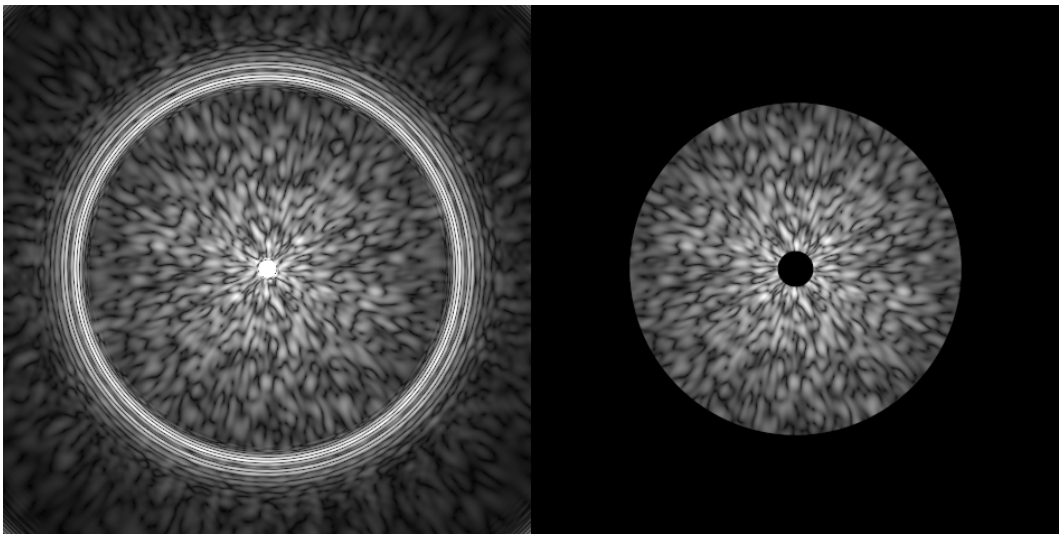

Figure 36. (Left) Amplitude at the PIAA focus (using the PASP-generated wavefront). The bright rings on the outside are from diffraction by the binary post-apodizer. (Right) The field after applying the occulter and outer radius mask.

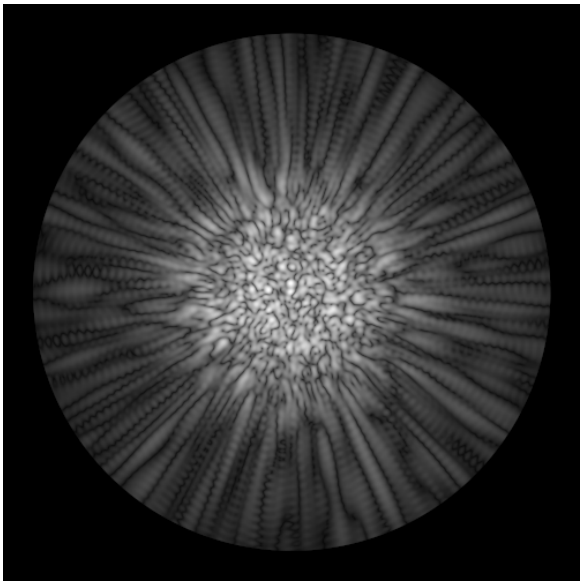

Figure 37. Amplitude of the aberrated wavefront at reverse M2.

Propagation using S-Huygens

The propagation from reverse PIAA M2 to M1 with S-Huygens generally follows the same procedure as the forward propagation. The wavefront at reverse M2 is decomposed into the azimuthal spatial frequency spectrum, $C_{reverse}$ (Figure 38). The spectral coefficients are again sampled from k = -175 to +175 cycles/360°. Because there is no significant 0^{th} order term, the sampling can be relaxed – the radial $C_{reverse}$ vectors are resampled to 25,000 points along the radius of reverse M2 and propagated to 15,000 points along the radius of reverse M1. The two-dimensional wavefront is then constituted as before, creating a 819.2 pixel diameter beam embedded in a 2048 × 2048 array. The PASP and S-Huygens reverse M1 wavefronts are shown in Figure 39.

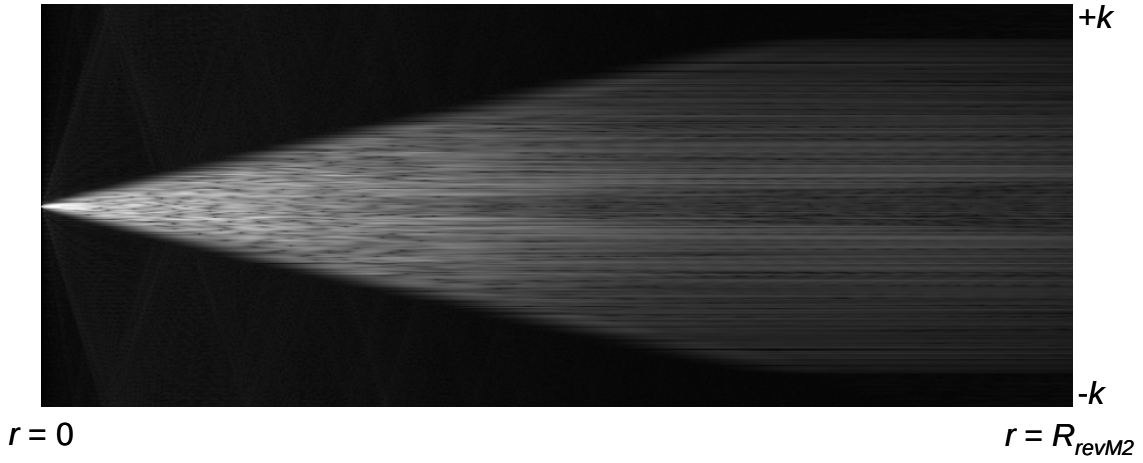

Figure 38. Azimuthal spatial frequency spectrum $C(r,k)$ of the aberrated field at reverse M2.

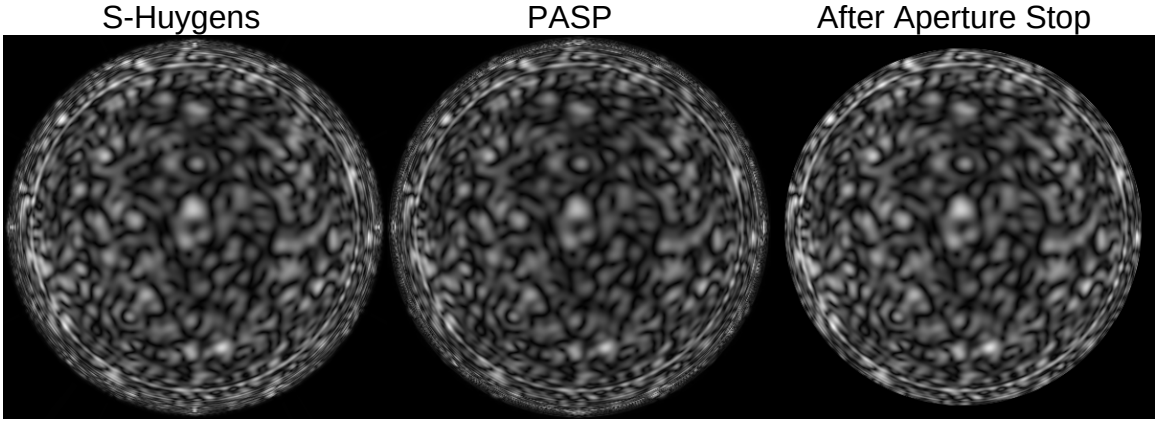

Figure 39. Amplitude of aberrated field (10^{-10} contrast) at the reverse M1 optic. The S-Huygens propagated field is on the left, with the PASP one in the middle. On the right is the field after the 91% pupil radius aperture stop is applied.

Stage 4: Propagation from reverse M1 to final focus

PROPER is used to propagate the wavefront from reverse M1 to the final focus. First, an aperture stop with a 91% diameter clear opening is applied to mask high spatial frequency artifacts at the very edge of the beam, likely introduced by the post-apodizer. This does not reduce throughput as much as might be expected because the outer ~5% of the (undistorted) beam was already blocked by the post-apodizer. A lens is inserted and the wavefront is propagated to focus (Figure 40), with a sampling of 0.4 λ/D per pixel.

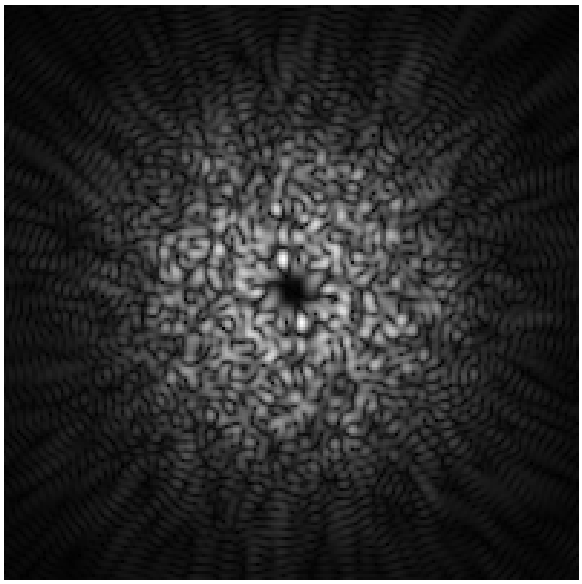

Figure 40. PASP-computed amplitude of the aberrated wavefront at the final image plane. The shadow of the occulter can be seen at the center. The speckles disappear towards the edge due to the filtering of the aberration map to <48 cycles/D.

Computing the Milestone #1 Accuracy Metric

Milestone #1 involves measuring the accuracy of the PASP algorithm at aberration levels of 10^{-5} and 10^{-10} equivalent contrasts within a 2.5 – 18 λ/D annulus. The stages described above are executed in separate runs using the same error map scaled to provide these mean contrasts. For each contrast level, the image plane complex-valued field generated using S-Huygens for the PIAA propagations is subtracted from that using PASP. The modulus-square of the difference field is taken to convert the result to intensity. The root-mean-square of the pixels inside the annulus is computed and then divided by the peak of the unocculted stellar PSF (generated by running a simulation without the occulter) to convert it to units of contrast.

The final image plane fields and the difference fields are shown in Figure 41 to Figure 44. The measured mean contrasts and accuracies are given in Table 5. The accuracies easily meet the 1% requirement by nearly two orders of magnitude, which is satisfyingly surprising given that the algorithms and codes are completely different.

However, it was demonstrated earlier that the accuracy of a two-dimensional representation of a circularly-symmetric, aberrated field using a filtered version of the

binary post-apodizer, when compared to a one-dimensional, high resolution field with a finite binary post-apodizer, is limited to ~5% at contrasts of ~10^{-10} when 2048 × 2048 arrays are used. Using 4096 × 4096 arrays improves the accuracy to 1%.

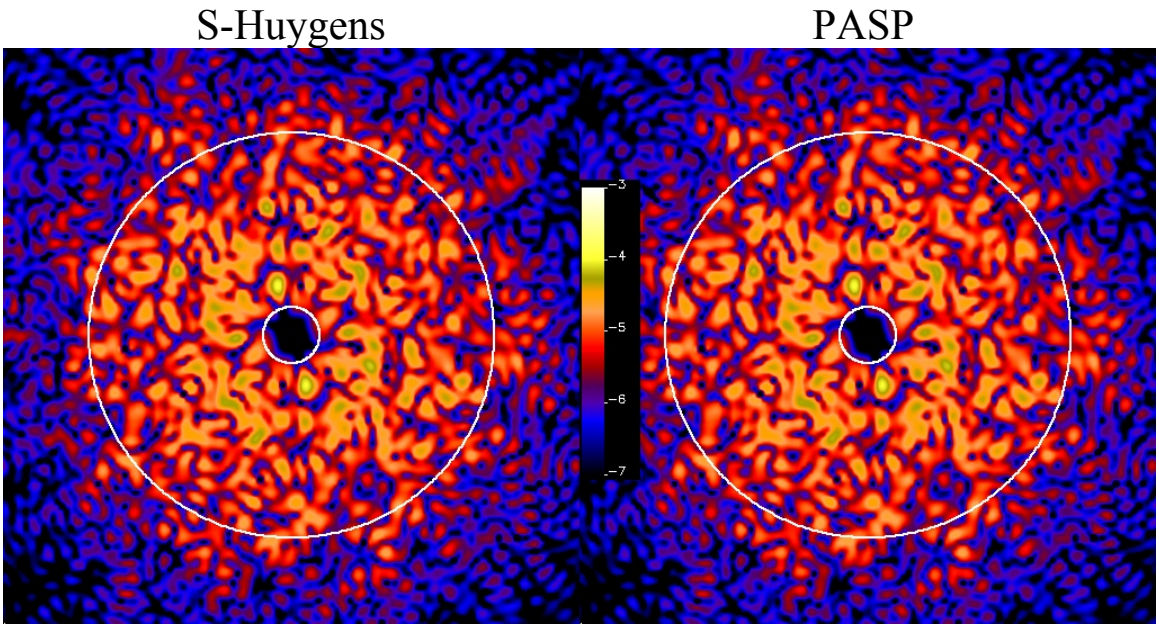

Figure 41. Contrast maps at the final focus for a field with ~10^{-5} mean contrast. The inscribed circles mark the inner (2.5 λ/D) and outer (18 λ/D) radii. The color bar indicates $\log_{10}$(contrast).

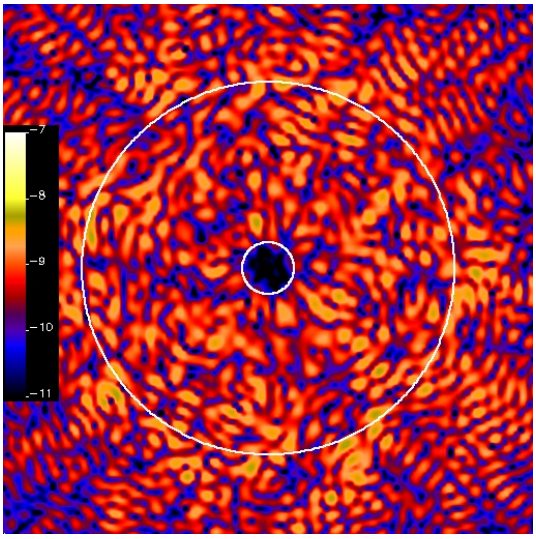

Figure 42. Contrast of the difference of the PASP and S-Huygens wavefronts for the ~10^{-5} mean contrast case. Note that the color scale is different than for the previous figure.

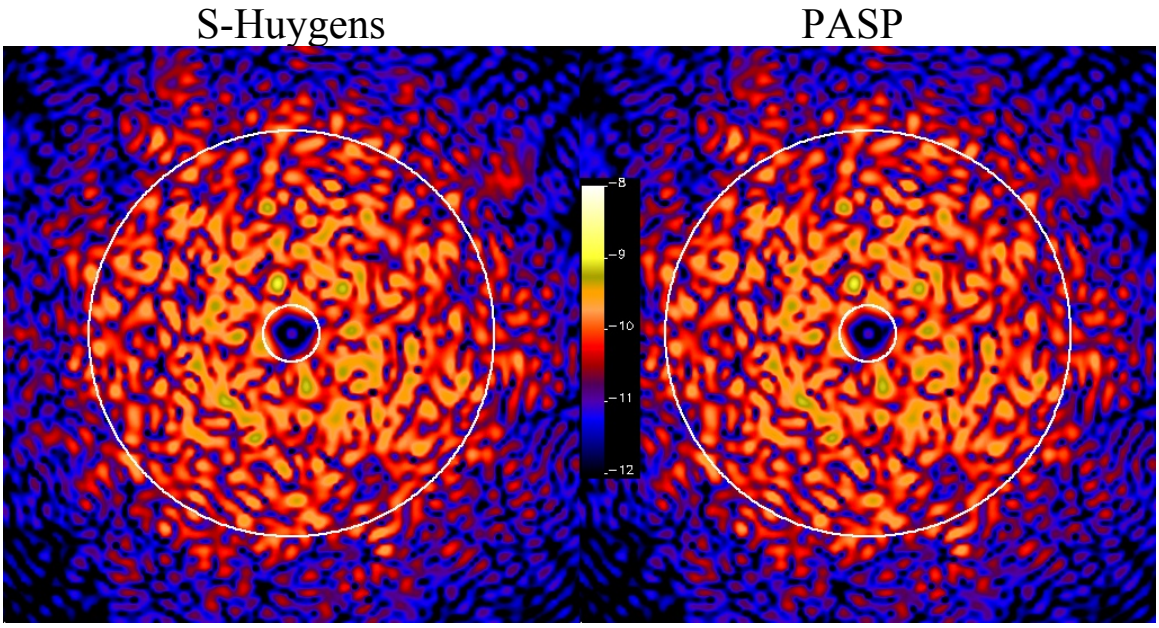

Figure 43. Contrast maps at the final focus for a field with ~10^{-10} mean contrast. The inscribed circles mark the inner (2.5 λ/D) and outer (18 λ/D) radii. The color bar indicates $\log_{10}$(contrast).

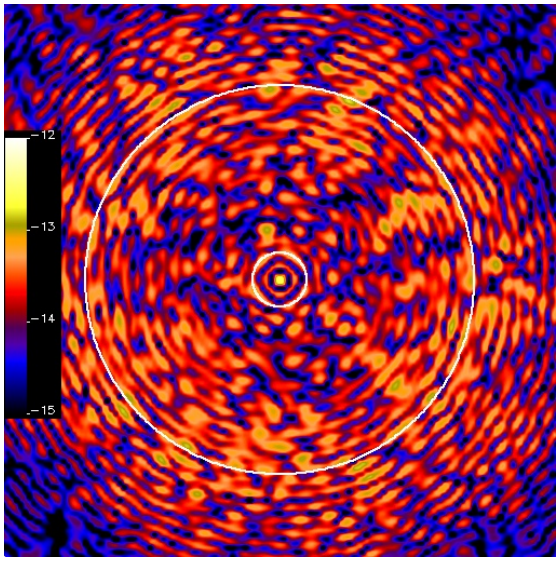

Figure 44. Contrast of the difference of the PASP and S-Huygens wavefronts for the ~10^{-10} mean contrast case. Note that the color scale is different than for the previous figure

.

Table 5. Milestone #1 Accuracy Metrics for the PIAA Coronagraph

	Reference (S-Huygens)	Efficient (PASP)
Low Contrast field mean	9.5×10^{-6}	9.5×10^{-6}
LC Accuracy(1)		1.9×10^{-9} (0.02%)
LC Milestone #1 Accuracy(2)		**0.03%**
LC Requirement		<1%
High Contrast field mean	9.2×10^{-11}	9.0×10^{-11}
HC Accuracy(1)		3.0×10^{-14} (0.03%)
HC Milestone #1 Accuracy(2)		**5%**
HC Requirement		<1%

(1) The accuracy derived from comparing the S-Huygens and PASP results.
(2) The official accuracy is set by the comparison of the 1-D, high resolution propagation using a finite binary post-apodizer to a 2-D (2048 x 2048) one using a filtered binary post-apodizer. Note that the HC accuracy requirement is met when 4096 x 4096 arrays are used (but not the efficiency requirement).

Accuracy of pupil remapping without explicit propagation

A "quick and dirty" means of representing PIAA in a simulated system is to simply remap the wavefront at M1 via interpolation onto M2 and apply the analytic PIAA-induced amplitude apodization. The wavefront errors are thus not actually propagated – phase errors at M1 remain phase errors at M2, amplitude errors remain amplitude errors, only their spatial distribution and amplitude change. This ignores the phase-to-amplitude transitions that occur in a real PIAA system.

While much faster (the interpolation is nearly instantaneous), accuracy suffers. The same accuracy aberrated wavefront was propagated through the test system using remapping at the forward and reverse PIAA optics. For a 10^{-5} contrast field, the accuracy vs. S-Huygens (as defined previously) was 6.9×10^{-7} (7.2%) and for a 10^{-10} field it was 1.3×10^{-11} (14.6%).

PIAA Milestone #1: Efficiency of the PASP algorithm

Having shown the PASP algorithm accuracy, we now demonstrate whether it is fast enough to create the deformable mirror response matrix (sometimes called the *G* matrix) that is used for wavefront control (Give'on et al. 2007). This matrix is used to solve for the DM actuator settings required to minimize energy within the dark hole region of the image, given a measurement of the field there.

The *G* matrix is constructed by poking each DM actuator in a simulated system by some amount and calculating the change in the field in the final image plane. This has to be done at each sensing wavelength at which contrast is being optimized (3 – 5 wavelengths for a 20% bandpass). In our layout, two deformable mirrors with 46 actuators across the beam are used with 5 sensing wavelengths, amounting to over 16600 separate system

propagations required to determine the response of all actuators at all wavelengths. Because we are enforcing a perfectly symmetrical system, we can use the eight-fold symmetry described earlier to reduce the number of propagations by a factor of 8, down to about 2077. Still, if it took an hour to accurately propagate the poke of one DM through a system containing forward and reverse PIAA optics, it would take 87 days to complete the entire run of 2077, or without shortcuts in a real, non-symmetric system, almost two years for all the actuators.

Milestone #1 requires that the PIAA propagator be fast enough to allow the propagation of 2077 separate actuator pokes to be computed through the entire optical system within 48 hours on a modern workstation. To demonstrate this, we will propagate 416 actuators (1/5 the total) in under 48/5 = 9.6 hours.

Optical System Definition

The optical layout used in this test (Figure 6) is a realistic and complete telescope with DMs and a PIAA coronagraph. It is based on the PIAA coronagraph layout developed for the ACCESS ASMCS07 study (Trauger et al. 2008, Krist et al. 2009). This same system will be used for end-to-end modeling to be undertaken in Milestone #2 to demonstrate the wavefront control behavior of the PIAA system, part of which will involve generating and using the *G* matrix.

An unobscured (off-axis), circular Gregorian telescope is used. The image of the primary mirror is formed on the first DM, which has 46 actuators across the pupil. A second DM with the same actuator spacing is located 1.0 m downstream. An image of the 1st DM is created on the forward PIAA M1 optic, which apodizes the beam at M2. The image of M2 is created at the post-apodizer, after which the beam is focused onto the occulter/field stop. From there the beam goes to the reverse PIAA system and then is focused at the final image plane. The optical system has no aberrations in this instance.

Method

A flowchart of the propagation stages is shown in Figure 45, along with the execution times for each stage.

Stage 1. Propagation from primary mirror to deformable mirror

The optical system between the primary mirror and the deformable mirror does not change with modification of the DM. Therefore, the propagation from the primary to DM1 needs to be computed only once per wavelength and saved. Likewise for the primary to DM2. These propagations are done using PROPER with a pupil diameter of 819.2 pixels embedded in a 2048 × 2048 array.

Stage 2. Propagation from deformable mirror to forward PIAA M1

For each actuator & wavelength being processed, the field at DM1 or DM2 computed in Stage 1 is read in, an actuator is poked, the wavefront propagated to the PIAA M1 mirror. The Fourier transform of the wavefront is then computed to produce the spatial frequency spectrum, and terms with spatial frequencies below 60 cycles/diameter are kept (the same as in the Milestone #1 accuracy test). These terms are written to disk (this is more efficient than saving the entire wavefront). This is done using PROPER.

To improve efficiency, eight separate instances of IDL are started, each simultaneously running the PROPER DM-to-PIAA prescription. Each instance processes 52 actuators. On a modern workstation (dual Xeon processors with 4 cores per processor), this improves speed by a factor of ~5, even though the intrinsic IDL operators and the FFTW routines used by PROPER are already multithreaded.

Stage 3. Propagation from PIAA M1 to PIAA M2

The PASP+S-Huygens hybrid algorithm is used to propagate between the forward PIAA optics (the S-Huygens edge wavefront component is computed beforehand). As with the accuracy test, the harmonic component wavefronts are created on 460 × 460 grids that form the propagation matrix. The propagation matrix is first created and stored in memory. Tests show that it is actually faster to compute the matrix from scratch than to read it in from disk. In a serial order, the M1 wavefront spatial frequency spectrum for each actuator produced in Stage 2 is read in and multiplied by the propagation matrix, and the resulting 460 × 460 wavefront is saved to disk.

Stage 4. Propagation from PIAA M2 to the occulter and then to reverse PIAA M2

Using the eight-simultaneous-IDL-processes method applied in Stage 2, the 460 × 460 wavefront for each actuator is read in and resampled onto a 1433.6 pixel diameter pupil embedded in 2048 x 2048 array. The field is the propagated using PROPER to the post-apodizer, where the filtered apodizer is applied, and then to focus. The occulter and field stop are applied (as in the accuracy test), and the field propagated to the reverse PIAA M2 optic, where a 75% beam diameter clear aperture stop is applied. As in Stage 2, a Fourier transform creates the spatial frequency spectrum, the significant components of which are extracted and written to disk.

Stage 5. Propagation from reverse PIAA M2 to reverse PIAA M1

The PASP propagation matrix for the reverse PIAA system is generated and stored in memory. Serially, the spatial frequency spectrum for an actuator produce in Stage 4 is read in, multiplied by the matrix, and the resulting 460 × 460 element array written to disk.

Stage 6. Propagation from reverse PIAA M1 to final focus

Using the eight simultaneous IDL processes method again, the wavefronts from Stage 5 are read in, sampled to 819.2 pixels across the beam embedded in a 2048 × 2048 array, multiplied by a 91% clear diameter stop, and propagated using PROPER to the final image plane (where sampling is 0.4 λ/D per pixel).

Computing the Milestone #1 Efficiency Metric

The Milestone #1 efficiency test is executed on a Linux workstation with dual 3.0 GHz Xeon processors (4 cores/processor) and 48 GB of RAM and 2048 × 2048 wavefront grids. A total of 416 actuators are processed using the method described above, equal to the number of actuators to be propagated at one wavelength.

The execution times for each stage are listed in Table 6. The total time required for 416 actuators was 6.75 hours. Extrapolating that to a full five wavelengths of actuators (416×5=2080), the execution time is 33.75 hours, below the 48 hour time requirement.

Table 6. Milestone #1 Efficiency Metrics for PIAA

Primary to DM to PIAA M1	23.7 m
Forward PIAA M1 to M2	171.6 m
PIAA M2 to occulter to reverse PIAA M2	30.0 m
Reverse PIAA M2 to reverse PIAA M1	165.3 m
Reverse PIAA M1 to final focus	14.6 m
Time for 416 actuators (measured)	6.75 hr
Time for 416×5=2080 actuators	33.75 hr
Required time for 2077 actuators	<48 hr

Note: Elapsed time is given for a dual quad-core Xeon system.

While the efficiency milestone is met using 2048 × 2048 arrays during the wavefront propagation, the accuracy milestone is not met for 10^{-10} contrast fields. The accuracy is met with 4096 × 4096 arrays, but using those would increase the computation time by at least a factor of 3, exceeding the efficiency milestone limit. It is possible to reduce the execution time for the larger arrays by perhaps ~10% by avoiding propagations to surfaces that have no aberrations (an thus do not affect the resulting DM response functions). The bulk of the execution time, however, is in the PASP computations, and those cannot be reduced further. As always, one can throw additional computational power at the problem, distributing the load across a number of computers (including graphical processor units).

Impact of not meeting the combined PIAA milestones

The inability to achieve the 1% accuracy at 10^{-10} contrast milestone with 2K arrays, while disappointing, is not a critical failure in terms of representing the wavefront control behavior of PIAA. A 5% error still allows DM control matrices to be generated that are within the linear response regime of the system (i.e., a 5% error in the DM response matrix will not cause the wavefront control solution to become nonlinear and diverge). The larger error may require one or two more iterations to reach the same contrast goal as a more accurate matrix, but will not prevent it (the sensitivity is something that is worth exploring in future studies of the various control algorithms). The fundamental characteristics of the PIAA system are still represented by the models.

Of course, there is always the option of using larger arrays. One only needs to wait a couple of years for systems to gain the factor of 3 or 4 in speed necessary to match today's 2K computation times. Also, given that the DM response matrix is so easily parallelized, simply adding more processors can solve the problem today. Also, the PASP propagation through the reverse PIAA system may be unnecessary and could be replaced with a simple geometric remapping given that no significant scattering occurs there due to the already-suppressed PSF.

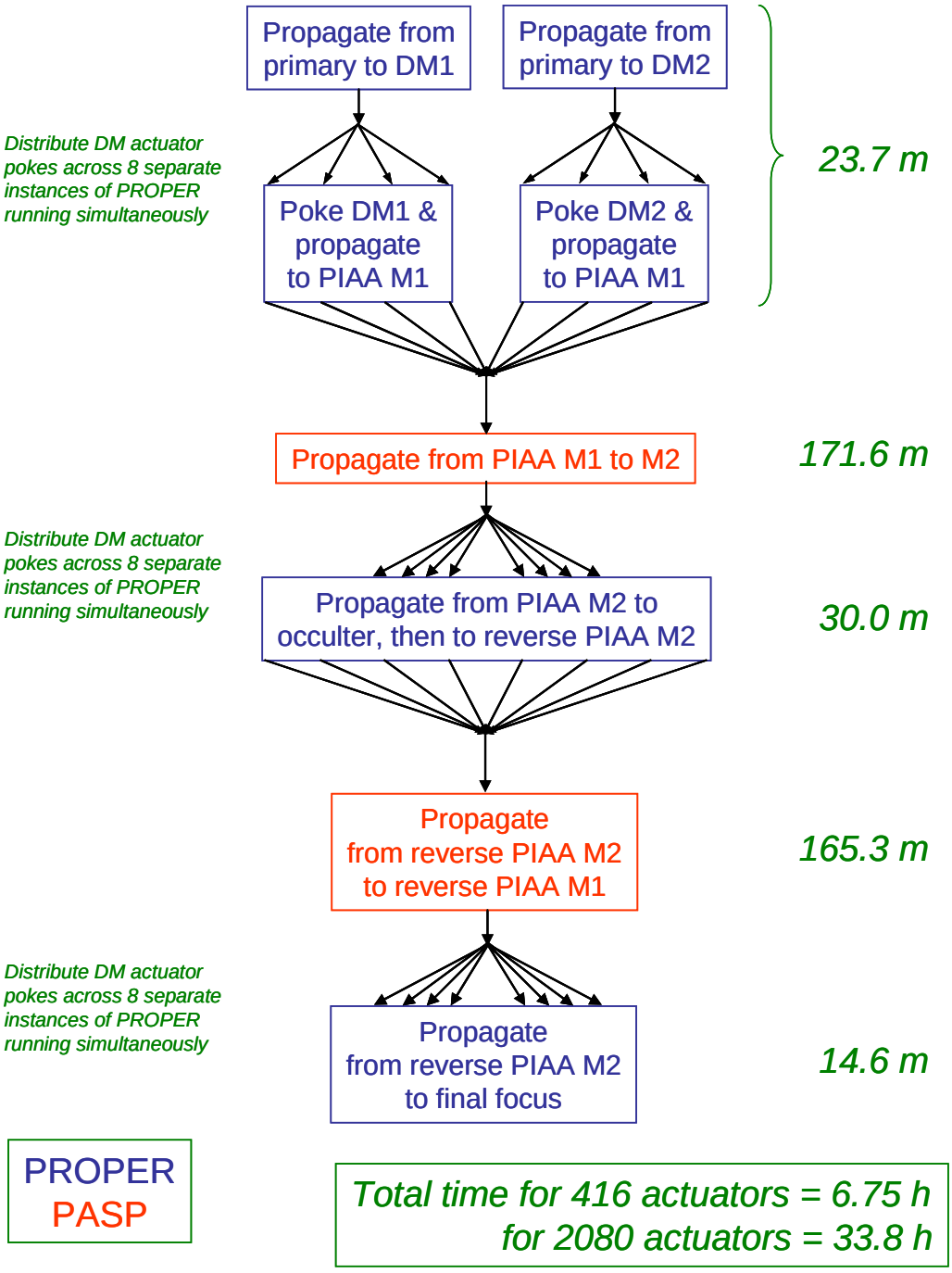

Figure 45. Flowchart illustrating the stages required to propagate DM actuators to the image plane to create the DM response matrix. Stages that use PROPER are in blue, those that use PASP are in red. The execution time required for each stage is listed on the right.

Appendix D: Vector Vortex Coronagraph (VVC)

The Vector Vortex Coronagraph Unique Characteristics

The concept of an optical vortex and its vectorial form used in this study are discussed in Section 2.1.2. Recent work has shown that wavefront errors passing through the VVC can only be corrected in one polarization direction. This would require a polarizer to block light with the perpendicular polarization from reaching the detector, effectively halving the throughput. There are two options to get around this problem. The first is to have two separate polarization channels, each with their own deformable mirrors and coronagraphs. The second is to make use of wide-band half wave plates to produce a single polarization channel coronagraph capable of operation over a 50% bandwidth. In the second case, however, the coronagraph may be able to suppress diffraction over that bandwidth, but wavefront aberrations are unlikely to be correctable to the same level.

To date, the VVC has achieved, with DM wavefront control, contrasts of 3.5×10^{-9} (monochromatic, 785 nm), 2.6×10^{-8} (10% broadband, λ_c = 800 nm), and 4×10^{-8} (20% broadband, λ_c = 800 nm) on JPL's HCIT (Mawet et al. 2011a, 2011b). The demonstrated performance indicates that the VVC can be rated at TRL=3.

VVC modeling techniques utilize subsampling to accurately reproduce the phase ramp in the region near the singularity.

Accurate propagator for the Vector Vortex Coronagraph

Breaking down the Jones matrix

In the Milestone #1 whitepaper, we showed that the action of the vector vortex coronagraph (VVC) on the input vectorial wavefront W^i is rigorously represented by the following Jones matrix product:

$$\begin{bmatrix} W_L^o(r,\theta,\lambda) \\ W_R^o(r,\theta,\lambda) \end{bmatrix} = \begin{bmatrix} L(r,\theta,\lambda) & V(r,\theta,\lambda)e^{il\theta} \\ V(r,\theta,\lambda)e^{-il\theta} & L(r,\theta,\lambda) \end{bmatrix} \times \begin{bmatrix} W_L^i(r,\theta,\lambda) \\ W_R^i(r,\theta,\lambda) \end{bmatrix}$$

As stated in the white paper, a single polarization at a time will be analyzed, which simplifies the equation as follows:

$$\begin{bmatrix} W_L^o(r,\theta,\lambda) \\ W_R^o(r,\theta,\lambda) \end{bmatrix} = \begin{bmatrix} L(r,\theta,\lambda) & V(r,\theta,\lambda)e^{il\theta} \\ V(r,\theta,\lambda)e^{-il\theta} & L(r,\theta,\lambda) \end{bmatrix} \times \begin{bmatrix} W_L^i(r,\theta,\lambda) \\ 0 \end{bmatrix}$$
$$= \begin{bmatrix} L(r,\theta,\lambda) \times W_L^i(r,\theta,\lambda) \\ V(r,\theta,\lambda)e^{-il\theta} \times W_L^i(r,\theta,\lambda) \end{bmatrix}$$
$$= \begin{bmatrix} L \times W \\ VVC \times W \end{bmatrix}$$

In conclusion, the propagation through the VVC requires two main propagations, adding up incoherently: (1) the *VVC* term that will itself be broken down (see below) to represent the action of the opaque spot covering the central region of disorientation and (2) the leakage term. The chromatic leakage term *L* transmits a fraction of the incoming light without phase modification but with the amplitude modification induced by the opaque spot.

Action of the VVC in the perfect case

Before starting, here is a quick reminder of the perfect VVC ideal case (circular filled uniform pupil without aberration) and its mathematical property (Mawet et al. 2005). The field in the pupil after the vortex focal plane mask is:

$$E_{pup} = FT[E \times VVC(l)] = HT_i[J_1] = \begin{cases} 0 & r < R_{pup}; l = 2,4,... \\ f(r) & r > R_{pup} \end{cases}$$

The action of the vortex on a perfect field E_{perf} at focus corresponds to a Hankel transform ($HT[]$) of the l_{th} order of the Bessel J_1 function. This transform has analytical solutions, known as the Weber-Schafheitlin integral, which reduces to the Sonine integral in the *l=2* case (Abramowitz & Stegun 1972, p. 487). The final rigorous analytical result of these integral solutions is that the energy inside the pupil downstream from the vortex is 0 for non-zero even values of the *topological charge l*.

Superposition principle applied to the VVC

Now in the real world, nothing is perfect, and both the wavefront and the VVC are affected by systematic errors. The superposition principle allows rigorous decomposition of these errors linearly. The aberrated system field at the vortex plane, *E,* can be decomposed into the sum of a perfect field, E_{perf}, and an aberrated one, E_{ab}:

$$E = E_{perf} + E_{ab}.$$

Both of which correspond to the Fourier transform of a perfect wavefront at a pupil and the aberrated one:

$$W = W_{perf} + W_{ab}.$$

The action of a real world VVC_{ab} (VVC_{perf} + *central opaque spot S*) can be decomposed into the difference of a perfect VVC_{perf}, whose mathematical properties with a perfect wavefront are well known (see here above), and a finite one limited to the spot area:

$$VVC = (VVC_{perf} - VVC_{spot})$$

The superposition principle allows us to break down the action of an imperfect vortex on an imperfect wavefront, as follows

$$\begin{aligned} E \times VVC &= (VVC_{perf} - VVC_{spot}) \times (E_{perf} + E_{ab}) \\ &= (VVC_{perf} \times E_{perf}) + (VVC_{perf} \times E_{ab}) - (VVC_{spot} \times E_{perf}) - (VVC_{spot} \times E_{ab}) \end{aligned}$$

Evaluation of the first term

The first term (Figure 46), bearing a large fraction of the incoming beam energy (in the high Strehl ratio regime), *cancels out inside the pupil geometric area* (cf. the mathematical demonstration in the perfect case), which allows us to avoid computing it uselessly and be confronted to the well-know aliasing error of the vortex coronagraph representation at the very center (where most of the energy is). Outside of the pupil the light is masked by the Lyot stop.

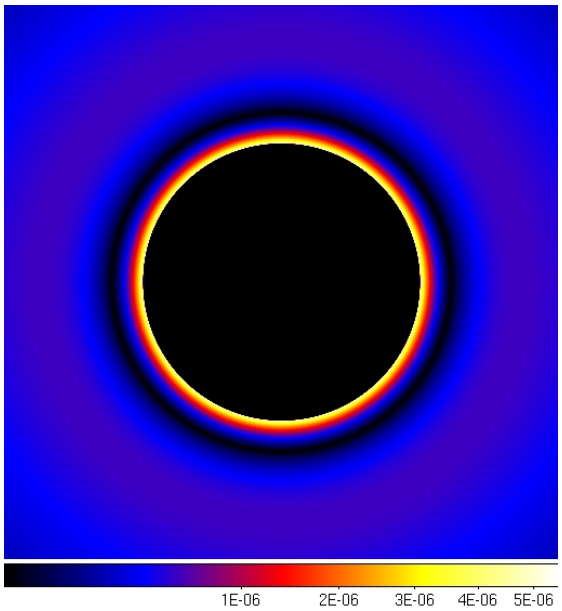

Figure 46. First term in the pupil downstream of the coronagraph (FT[perfect vortex x perfect field]) showing the expected diffraction pattern of the 4th order VVC. The energy inside the pupil is 0.

For this term, the efficient method simply consists of considering this term as being 0. The reference method is analytical and gives 0 as well. The match is therefore perfect, *by definition*.

IMPORTANT NOTE:

Our representation of the vortex boils down to propagating differential terms (2, 3, & 4) only, using the perfect vortex solution for the dominant (1st) one. However, leaving out this main term, which carries most of the energy, does not capture the effect of downstream aberrations. Therefore, to represent the low-level interactions between the 1st vortex term and errors on the OAP between the mask and the Lyot stop, one must carry out the propagation of the "vortexified field" ($E_{perf} \times VVC_{perf}$) from the mask, through the relay optic, and on to the Lyot stop.

The numerical representation of the vortex mask is generated by computing a phase ramp on a grid 5× larger on each side than the wavefront array (e.g., a 10K × 10K array for a 2K × 2K grid). This subsampled grid is then block averaged (real and imaginary parts separately) down to the default sampling. This method produces a good, but not perfect, representation of the vortex mask. When it is multiplied by the perfect field at focus and propagated to a pupil, aliasing and Fourier transform artifacts introduce energy inside the pupil interior, when it should be zero.

To avoid these numerical aliasing problems specific to the 1st term (and affecting low spatial frequencies), one can propagate the $E_{perf} \times VVC_{perf}$ field to the Lyot stop, zero-out the pupil, and then propagate back to the focal plane, using the result for future propagations. Because Fresnel algorithms are used, the field at the focus has a phase term introduced by the focusing optic, so a simple Fourier transform cannot be used to go back and forth between the focal and pupil planes. Instead, one must do a full propagation from focus to the collimator, apply a positive lens, propagate to the Lyot stop and zero out the interior, then propagate backwards to the collimator, apply a negative lens, and then back to focus. This only has to be done once per wavelength and the result stored, since the 1st term does not change.

Generally speaking an optical system could be designed in such a way that the number of relay optics is reduced to the only required OAP. This optic would be sufficiently far from the mask, hence close enough to the pupil so that the dark hole is nearly fully formed, making the effect of surface errors nearly irrelevant after filtering by the Lyot stop.

Evaluation of the second term

The second term (Figure 47, left) can then be propagated using the conventional FFT-based technqiues, aliasing (sampling artifacts) errors being negligible where the aberrations are (further away from the center of the vortex).

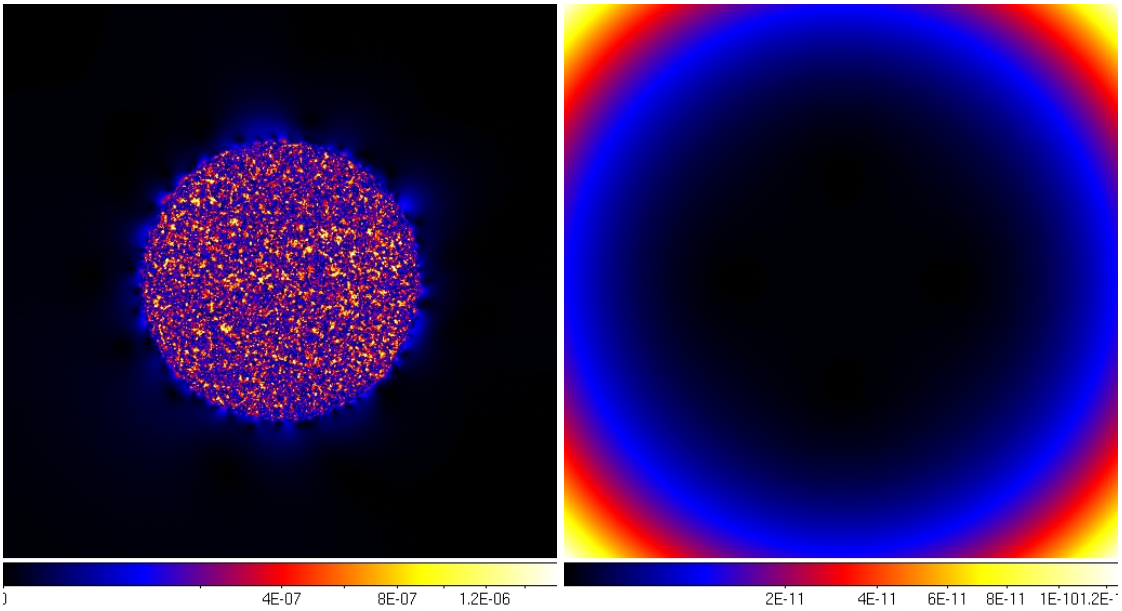

Figure 47. Visualization of the second term including aberrations, (left), FT[perfect vortex × aberrated part of the field] and the third term (right), FT[vortex limited to the spot area × perfect field]. The scales have different stretches and are not linear.

Evaluation of the third and fourth term

The third term can be very accurately represented using partial Fourier transforms, e.g. the matrix Fourier transform (MFT, Soummer et al. 2007); we only need to evaluate the effect of the mask in a very limited area (the spot). This computation can be done once per wavelength, and stored for subsequent use. The perfect field at focus is Fourier transformed to a virtual pupil plane, then it is transformed back to focus using the MFT to create a very highly sampled field only within the region of the spot. This is multiplied by the vortex phase ramp and then transformed back to the virtual pupil using another MFT. The result is then Fourier transformed back to focus with the original field sampling and stored for subsequent use. We have verified that the spot, given its small size (0.5 λ/D) has no influence on the aberrated field (verified down to about 10^{-19} for a 10^{-10} nominal contrast), so the fourth term can be neglected (see Figure 47, right).

Sensitivity to sampling at the focal plane mask

The critical part of the vortex mask is the small central region where the phase changes rapidly and there is an opaque occulter. Poor sampling of the field here may result in numerical anomalies that will erroneously degrade the computed contrast. To determine the sampling sensitivity, we propagate an unaberrated wavefront through our default VVC system using different beam/grid diameter ratios (this beam ratio determines the sampling at the focal plane; if the beam is 512 pixels across and the wavefront grid is 2048 × 2048 pixels, the beam ratio is 512/2048 = 0.25 and the sampling at the focal plane will be 0.25 λ/D radians per pixel).

Figure 48 shows radial contrast plots for simulations with different focal plane samplings (0.1, 0.2, 0.3, and 0.4 λ/D with 2048 × 2048 arrays and 0.1 and 0.2 λ/D for 4096 × 4096 arrays). These show that with our method of oversampling the spot, there is no significant difference between any of the samplings used. In all cases, the contrast profile is that expected for an occulted VVC field. Omitting the occulting spot produces a contrast that is practically zero (specifically, $<10^{-32}$ everywhere); this is expected since the pure vortex × perfect PSF field is set to zero inside the pupil by our procedure and the field outside the pupil is masked by the Lyot stop.

For all of our VVC simulations to be done in the Milestone #2, we will use a beam diameter of 322 pixels (each DM actuator is 7.0 pixels across) in a 2048 × 2048 pixel wavefront grid, resulting in a focal plane sampling of 0.157 λ/D.

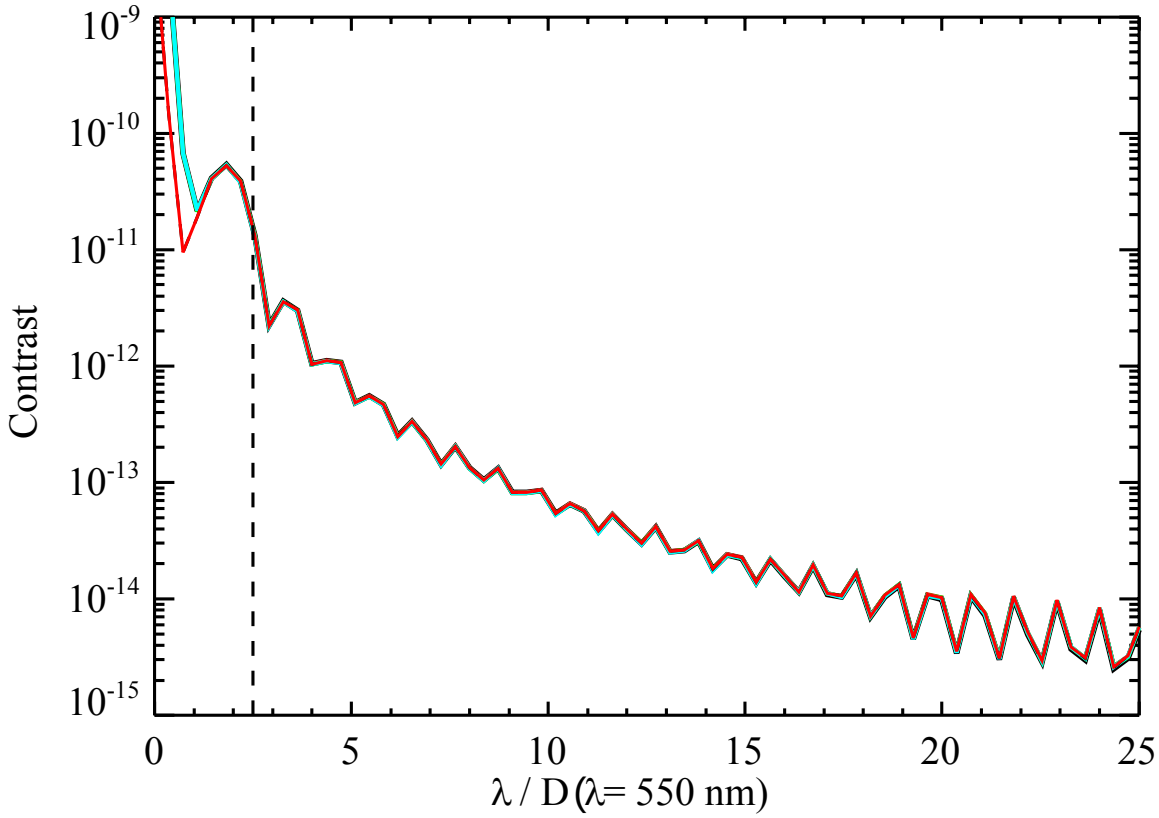

Figure 48. Monochromatic radial contrast profiles for an unaberrated VVC system modeled using different sampling densities at the focal plane mask. The sampling is given in radians per pixel. The dashed vertical line represents the inner working angle of 2.5 λ/D. The only significant difference is at $r < 1.5$ λ/D, where the profiles from the 4K × 4K array propagations (for samplings of 0.1 and 0.2 λ/D, both of which are plotted and overlay each other) are a little lower than those from the 2K × 2K ones (0.1, 0.2, 0.3, and 0.4 λ/D samplings are plotted and overlay each other).

VVC Milestone #1: Efficiency and Accuracy of the Vortex Representation

There is no known analytical method for propagation of an aberrated field through a vector vortex to serve as a reference algorithm to determine the accuracy of the methods used here. Instead, to estimate the accuracy we have run separate simulations using different array sizes (4096 × 4096 & 2048 × 2048) and corresponding beam/grid diameter ratios (0.0625 & 0.125, respectively; the pupil size is 256 pixels across in each case). We also include results for 10^{-10} contrast for the beam ratio to be used in Milestone #2. The VVC chosen for this simulation has a central opaque spot 0.5 λ/D in diameter. We used the phase map defined earlier to provide a contrast floor of 10^{-5} (low contrast case) and 10^{-10} (high contrast case). The simulation is run through the system layout shown in Figure 5. This is the same layout that will be used for Milestone #2.

The parameters and results of the simulations are presented in Table 7. Because the accuracies derived from the array size comparison are better than those established for PROPER in Appendix B, we use the PROPER accuracies as the Milestone #1 metric for the VVC. Figure 49 to Figure 52 show the contrast maps and differences for the 10^{-5} and 10^{-10} contrast fields. The efficient VVC propagation algorithms meet the Milestone #1 efficiency and accuracy requirements .

Table 7. Milestone #1 Metrics for the Vector Vortex Coronagraph

	Reference	Efficient 2K	Milestone #2 2K
Array size	4K × 4K	2K × 2K	2K × 2K
***Beam ratio*(1)**	0.0625	0.125	0.157
***Spot region sampling*(2)**	8K pix per λ/D	4K pix per λ/D	4K pix per λ/D
***Execution time*(3)**	54.5 sec	13.6 sec	13.6 sec
***" for 2077×2 actuators*(4)**	62.9 hrs	15.7 hrs	15.7 hrs
Requirement		<48 hrs	<48 hrs
Low Contrast field mean	1.1×10^{-5}	1.1×10^{-5}	
***LC Accuracy*(5)**		7.2×10^{-12} (~0%)	
***LC Milestone #1 Accuracy*(6)**		**0.02%**	
LC Requirement		$<1.1\times10^{-7}$ (<1%)	
High Contrast field mean	8.7×10^{-11}	8.7×10^{-11}	8.7×10^{-11}
***HC Accuracy*(5)**		2.6×10^{-15} (~0%)	2.8×10^{-15} (~0%)
***HC Milestone #1 Accuracy*(6)**		**0.6%**	**0.6%**
HC Requirement		$<8.8\times10^{-13}$ (<1%)	$<8.8\times10^{-13}$ (<1%)

(1) The diameter of the beam to the wavefront grid diameter. Also equivalent to the sampling at the focal plane in units of λ/D radians.
(2) The sampling in the magnified spot region field produced using the MFT.
(3) Propagation time from DM1 to final image on a dual quad-core Xeon system.
(4) Total time to propagate 2077 × 2 actuators (the factor of 2 is due to the 4-fold rather than 8-fold symmetry of the VVC+deformable mirror system).
(5) The accuracy derived from comparing the 2K × 2K result to the 4K × 4K one.
(6) The accuracy of PROPER when compared to S-Huygens for the given contrast level for a Lyot coronagraph. This is considered to be the Milestone #1 accuracy for the VVC because it is worse than that derived from the 2K vs. 4K comparison.

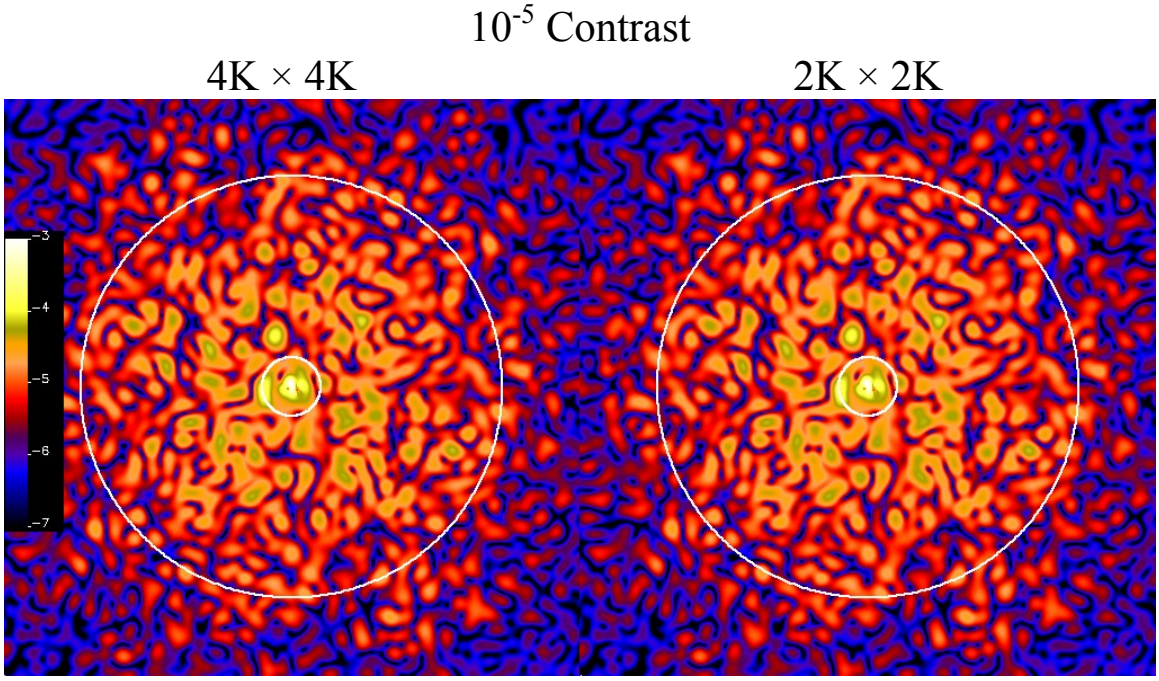

Figure 49. Monochromatic (550 nm) contrast maps for VVC Milestone 1 simulation runs (1.1×10^{-5} mean contrast) using 4K × 4K arrays (left) and 2K × 2K arrays (right). The contrast evaluation region of r = 2.5 - 18 λ/D is marked with circles. The colorbar indicates $\log_{10}$(contrast) and ranges from 10^{-7} up to 10^{-3}.

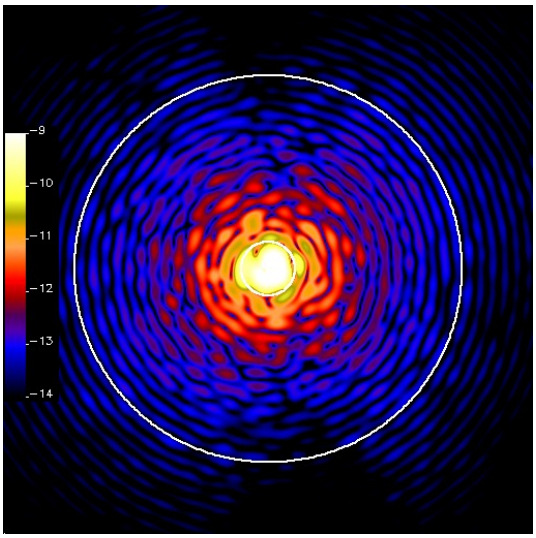

Figure 50. Difference between the 4K × 4K and 2K × 2K electric fields, displayed as contrast. The contrast evaluation region of *r* = 2.5 - 18 λ/D is marked with circles. The colorbar indicates $\log_{10}$(contrast) and ranges from 10^{-14} up to 10^{-9} (different from contrast maps above).

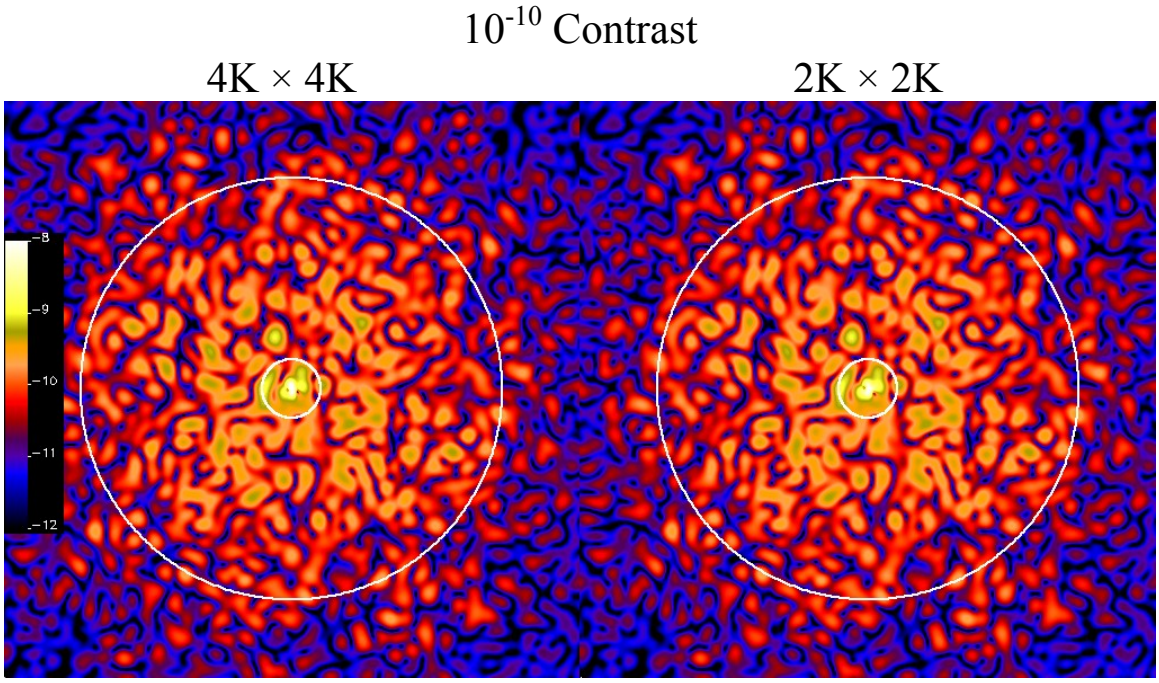

Figure 51. Monochromatic (550 nm) contrast maps for VVC Milestone 1 simulation runs (8.8×10^{-11} mean contrast) using 4K × 4K arrays (left) and 2K × 2K arrays (right). The contrast evaluation region of r = 2.5 - 18 λ/D is marked with circles. The colorbar indicates $\log_{10}$(contrast) and ranges from 10^{-12} up to 10^{-8}.

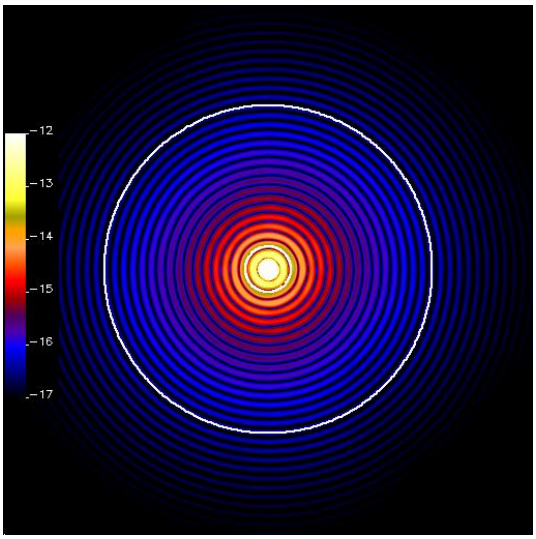

Figure 52. Difference between the 4K × 4K and 2K × 2K electric fields, displayed as contrast. The contrast evaluation region of r = 2.5 - 18 λ/D is marked with circles. The colorbar indicates $\log_{10}$(contrast) and ranges from 10^{-17} up to 10^{-12} (different from contrast maps above). The asymmetry in this residual map is due to the action of the opaque spot, which injects diffracted starlight enhancing the anti-symmetric cross-terms in the vortex.

Evaluation of the leakage term

By design, the chromatic leakage term $(L \times PSF_{abb})$ is below the noise floor at the nominal working wavelength (see here below). Its representation is straightforward and does not imply any phase modification. It can be evaluated with MFT to capture the effect of the spot (residual light levels are usually so small that classical FFT will do). See Figure 53.

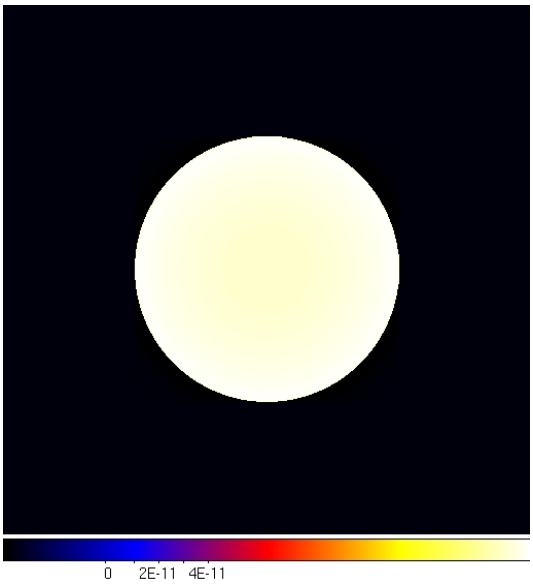

Figure 53. Pupil plane pattern of the fourth term, which is the constant leakage term modulated by the diffraction due to the finite spot size.

Chromaticity of the leakage term

Chromatic effects are inherent to phase-mask coronagraphs, but their importance depends on the technological approach chosen to render the phase distribution. Well-known techniques to make phase changes achromatic have been successfully used (achromats, Mawet et al. 2006). For the vector vortex, chromatic effects appear in the wavelength dependence of the local polarization characteristics of the birefringent medium, i.e. the retardance and the ordinary and extraordinary transmittances. The consequence is that the weighting coefficients of the leakage term L is wavelength dependent. For the current generation of devices, which is based on a single layer of LCP tuned to a target wavelength, usually picked at the center of the working astronomical filter, the natural dispersion of the LCP material rapidly degrades contrast by contributing quadratically to the leakage on each side of the central wavelength (see top red curve on Figure 55).

To improve the achromaticity of the vectorial vortex made out of LCP, the first obvious solution is to make its local polarization characteristics immune to wavelength changes. This first technique consists of using a stack of three or five LCP layers. Another solution (that can be used jointly with the first one) is to filter the leakage term. This solution, which could potentially bring tremendous improvements with off-the-shelf polarization

components, has already been investigated theoretically elsewhere (Mawet 2007, 2010). Experimental validation of this filtering method with the VVC will be the subject of future work, but has already been validated for the VVC by Murakami et al. 2010.

3-layer and 5-layer achromatic VVC

The 3-layer achromatic design is currently being investigated for manufacturing. This design again makes use of the geometrical phase, but in a vertical implementation rather than a lateral one (Figure 54). This principle has been known for half a century and is even commercially available from waveplate vendors. The basic three-layer design consists of identical retarder layers (1, 2, 3) having nominal retardance values $L_1 = L_2 = L_3 = L_{nom} = \pi$ radians at a given unique wavelength λ. The respective optical axes α_1, α_2, α_3 of those layers are nominally defined as follows (relative to some arbitrary angle reference): $\alpha_1 = -30^o$, $\alpha_2 = 30^o + \delta$, $\alpha_3 = -30^o$ (Figure 54), where δ is a small modifier angle (i.e. on the order of 0 to a few degrees) that is adjusted to produce the desired bandwidth, at the expense of retardance variation within the band. In practice, the six parameters at each spatial location (L_1, L_2, L_3, α_1, α_2, α_3) are subject to certain fabrication errors which result in some amount of deviation from the nominal design behavior. The retarder material has also a certain known dispersion of birefringence with wavelength that has to be taken into account in the design.

The 3-layer design can be extended to additional layers. A 5-layer design would satisfy all the requirements comfortably without the need to filter the residual chromatic leakage (Figure 55). All 5 layers are identical and also provide nominal retardance values $L_1 = L_2 = L_3 = L_4 = L_5 = L_{nom} = \pi$ radians at a given unique wavelength λ. The respective optical axes α_1, α_2, α_3, α_4, α_5 of those layers are nominally defined as follows (relative to some arbitrary angle reference): $\alpha_1 = 0^o$, $\alpha_2 = 70^o + \delta_1$, $\alpha_3 = -140^o + \delta_2$, $\alpha_4 = \alpha_2$, $\alpha_5 = \alpha_1$.

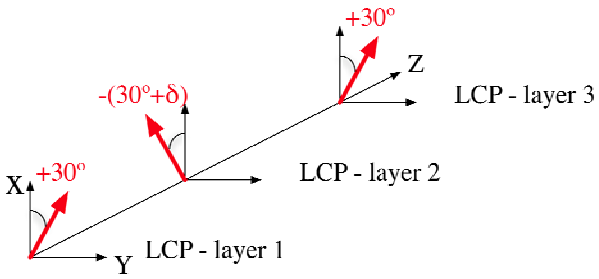

Figure 54. Three-layer LCP design. Optical axis offset between the three identical halfwave layers. Note that each individual layer is only a couple of microns thick (depending on the actual wavelength range of the device). The total thickness of the stack is always very small, on the order of 10 µm.

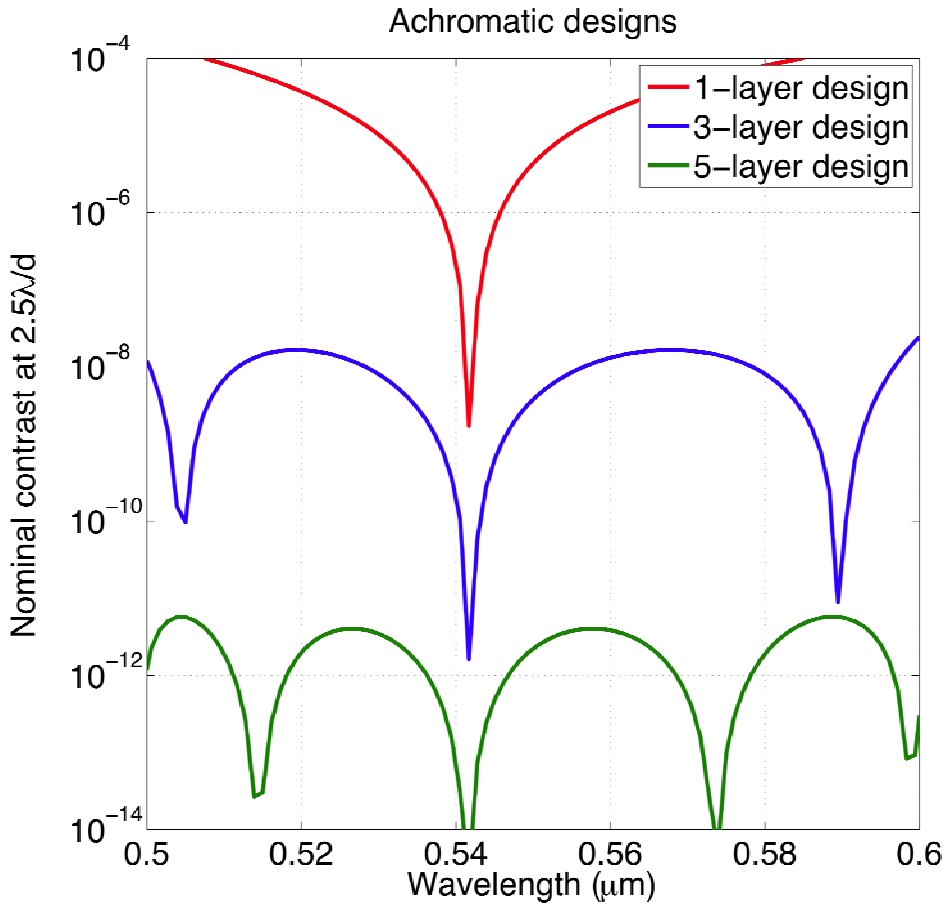

Figure 55. Contrast versus wavelength for different VVC focal plane mask coatings.

A 3-layer LCP is currently fabricated under NASA SBIR funding. See Table 8 for a quantitative estimation of the chromatic leakage contribution to contrast at 2.5 λ/d.

Chromatic leakage filtering

The VVC Jones matrix presented here above, which represents an imperfect chromatic VVC, consists of 2 terms, as already mentioned. A pure geometrical vortex, and a uniform leakage term bearing the chromatic defects. The polarization structure of both terms is orthogonal, which suggests that one could very easily get rid of the leakage term. Indeed, an input left (resp. right) circular polarization state yields a right (resp. left) - circularly polarized vortex term and a left (resp. right) - circularly polarized leakage. An output right (resp. left) - circular analyzer would let the pure vortex term out only, blocking the left (resp. right) - circularly polarized leakage.

To assess the performance gain of such a filtering setup, we use the Jones formalism and realistic assumptions concerning the performance of the circular polarizer/analyser. Indeed, the latter consists of a polarizer with a finite extinction ratio (ER), and a

quarterwave plate (QWP) itself affected by the chromaticism of its own retardance $\Delta\varphi_{QWP}(\lambda)$. Doing the matrix multiplication leads to a very simple rule of thumb for the filtering attenuation gain G of the orthogonal chromatic leakage:

$$G \approx \left[\frac{1}{ER} + \Delta\phi_{QWP}^2 \right]^{-1}$$

Assuming off-the-shelf polarizers and waveplate performances, for instance an ER of 10^4:1 and a QWP retardance accuracy $\Delta\varphi_{QWP}$ of 0.025 radian rms (over the working passband), the relative gain in contrast amounts to ~1400. This unique property of the VVC would allow us to significantly improve the broadband contrast floor and reduce the requirements on the mask itself. Note that the planetary signal is not affected by the filtering process. A dedicated experiment demonstrating this setup is currently being integrated at JPL (Mawet et al. 2010, in preparation ; see also Murakami et al. 2010).

Filtering the chromatic leakage imposes the use of polarizing elements upstream of the coronagraph. Note that the polarizing elements introduce their own optical aberrations. These must be taken into account in the model the same way as any other optical aberrations. While adding complexity to the system, it also represents an opportunity to perform an instantaneous polarization measurement *and* a real-time calibration of residual speckles. Indeed, polarization can be used both as a tool to facilitate or increase the dynamic range of instruments, and as a scientific gauge to probe physical phenomenon, but this topic is out of the scope of the current report.

Table 8. LCP designs and performances vs bandwidth in terms of chromatic leakage contribution to the resulting contrast (at ~2.5 λ/d), with and without chromatic leakage filtering. These theoretical values are applicable to the entire working wavelength range of the LCP material (from the optical to the near-infrared). Note that as the bandwidth increases, the definition of the IWA becomes flawed because of the PSF size variation with wavelength. Here we used the size of the PSF at the center of the passband.

Design	Bandwidth	Intrinsic Leakage	Filtered Leakage
1-layer LCP VVC	522-577 nm (10%)	10^{-5}	10^{-8}
	500-600 nm (20%)	6×10^{-5}	6×10^{-8}
	510-690 nm (30%)	1.4×10^{-4}	1.4×10^{-7}
3-layer LCP VVC	522-577 nm (10%)	10^{-10}	10^{-13}
	500-600 nm (20%)	10^{-8}	10^{-11}
	510-690 nm (30%)	10^{-7}	10^{-10}
5-layer LCP VVC	522-577 nm (10%)	10^{-15}	10^{-18}
	500-600 nm (20%)	10^{-12}	10^{-15}
	510-690 nm (30%)	10^{-10}	10^{-13}

Appendix E: Hybrid Band-Limited Coronagraph (HBLC)

HBLC overview

The HBLC is a variation on the classical Lyot coronagraph. The classical version uses a small, opaque occulting spot in an intermediate focal plane to mask the core of the stellar PSF and an aperture (Lyot) stop in a subsequent pupil plane to block most of the remaining diffracted light. The HBLC replaces the spot with specifically-tailored amplitude-and-phase-modulating, variable-thickness, patterned coatings. The graded amplitude transmission coating is usually a metal, such as nickel, on top of which a patterned dielectric coating (e.g., cryolite) is applied. By modifying both the amplitude and phase in a controlled manner, and using the known material properties (indices of refraction, etc.) the HBLC focal plane mask can control wavelength-dependent transmission variations inherent in any real material as well as increase the throughput by allowing for a slightly more open Lyot stop. Even greater improvement can be achieved by modifying the wavefront with one or more deformable mirrors.

The coatings and DM settings can be simultaneously optimized to produce a high-contrast field with specified inner and outer radii over a given bandpass. The procedure is somewhat similar to that used for wavefront control, except with the additional mask parameters. Our system was optimized to meet the milestone requirements of having $\leq 10^{-10}$ mean contrast in an aberration-free system in an $r = 2.5 - 18\ \lambda_c/D$ field over a $\lambda =$ 500 – 600 nm bandpass with a DM with 46 actuators across the pupil.

We must emphasize that the HBLC described here and used in our simulations includes the effects of coating-induced, wavelength-dependent, phase and amplitude variations based on real material properties as calculated with standard thin film equations. This is a completely realistic design that can be used in an actual system and whose performance should closely match the predictions (excluding any mask deposition errors).

Optimizing the HBLC components

The starting focal plane mask profile consists of real-and-imaginary-valued, circular Bessel functions implemented with variable-thickness nickel and cryolite coatings. These functions are band-limited with the half-transmittance point set to $2.5\ \lambda_c/D$, and they extend out to about $80\ \lambda/D$, beyond which the transmission was 1.0. Exact thin film equations are used to find the optical density and phase profiles at 17 different wavelengths spanning the passband. The circular Lyot stop pattern is clear between 0.1 and 0.6 of the pupil radius (its configuration is fixed). The spot at its center provides an additional degree of design freedom without significantly reducing the throughput. Only DM #1, located at a pupil, is used in the solution.

A DM response matrix is computed that described the effect in the focal plane of small strokes (probes) of each actuator at each of the 17 wavelengths, and this serves as the basis for an iterative, minimum least squares solution. This procedure imposes penalties

for large changes in actuator pistons and coating thicknesses in order to maintain the validity of the linear approximation necessary to describe the nonlinear system. At the beginning of each iteration the DM response matrix is recalculated using the current coating patterns. The DM settings and mask profiles are simultaneously optimized over nearly 200 iterations. The outputs of this process are a map of the wavefront correction introduced by the DM (Figure 56) and the focal plane mask real and imaginary profiles for each wavelength (Figure 57 and Figure 58).

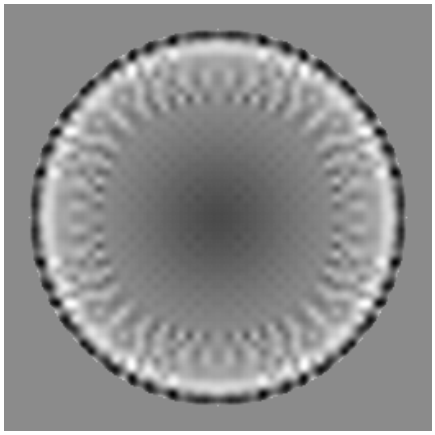

Figure 56. Deformable mirror #1 (located at a pupil) surface map, including actuator influence functions, for the original HBLC solution. The surface amplitude range is -17.5 nm to +14.6 nm.

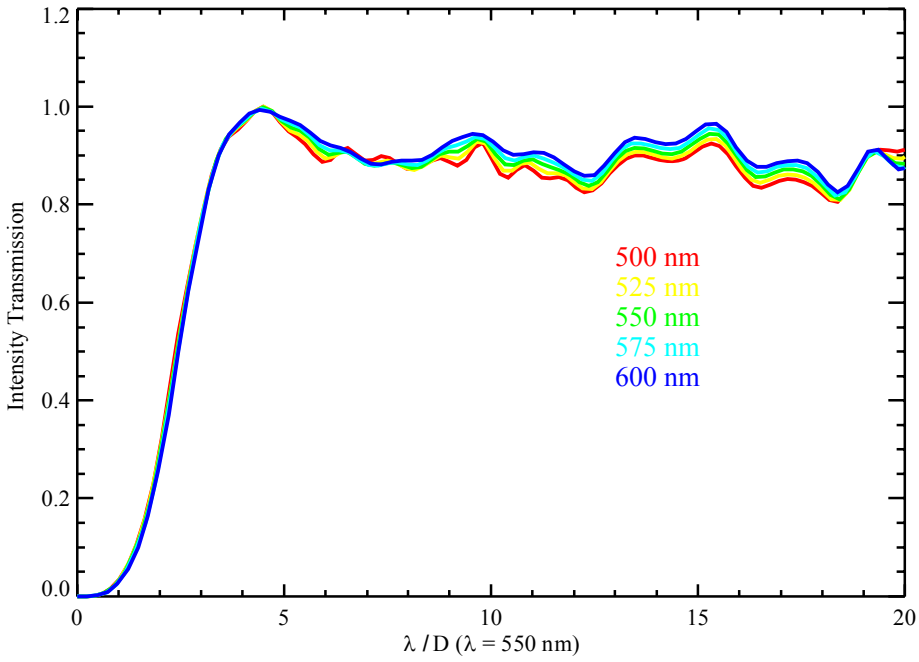

Figure 57. HBLC focal plane mask intensity transmission (square of amplitude modulation) for various wavelengths. The chromaticity is due to the material properties of the coatings as computed using thin film equations.

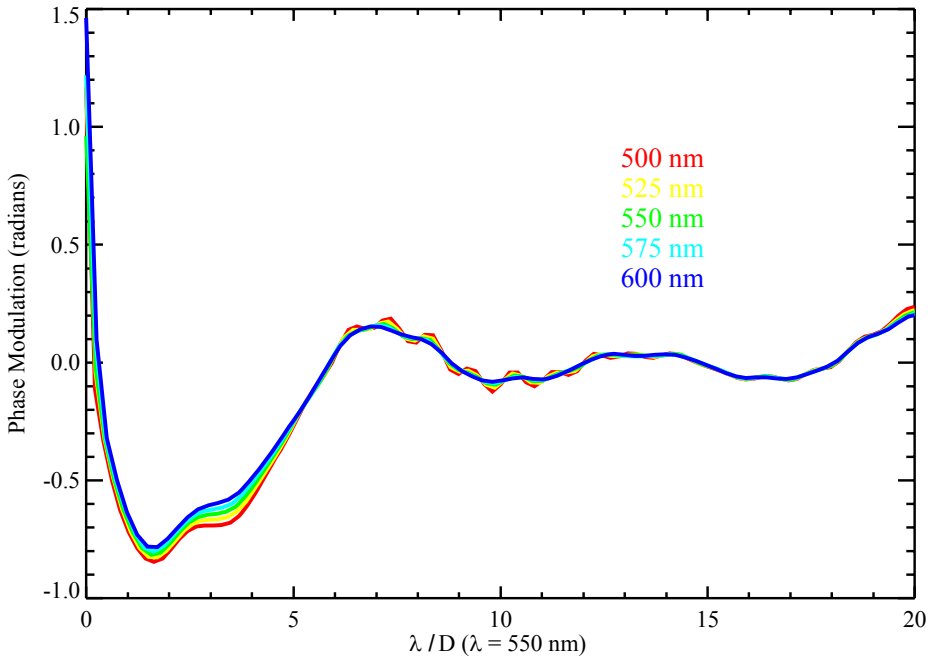

Figure 58. HBLC focal plane mask phase modulation for various wavelengths. The chromaticity is due to the material properties of the coatings as computed using thin film equations.

The optimization procedure (developed by team member Dwight Moody) uses Fourier-based propagation methods like those used by PROPER (but they are independent codes). For establishment of the Milestone metrics, the DM map and focal plane profiles are inserted into the PROPER prescription and an unaberrated wavefront is propagated through the system at multiple wavelengths to produce a broadband field. This result matches that produced by the originating code. As shown in Figure 59, this implementation provides 4×10^{-11} mean contrast in both $r = 2.5 – 3.5$ and $2.5 – 18$ λ/D annuli. About 10% of the dark hole area has a contrast $>10^{-10}$, mostly in a ring between $r = 5 – 7$ λ/D where some points are above 2×10^{-10}, but the statistic meets the Milestone 1 design requirements.

Redefining the HBLC design solution: DM map vs. actuator settings

When the DM wavefront map (rather than the actuator settings) provided by the optimization software is used in another program (i.e., PROPER), the resulting field produced by the HBLC is as expected. If the actuator pistons from the optimizer are used in PROPER instead of the DM wavefront map, the results do not agree because the two programs use slightly different actuator surface influence functions (the influence function describes how the surface of the DM changes when a single actuator is poked; the surface will deform over an area of several actuators). As an experiment, we fit the DM map using PROPER's influence functions to derive the actuator settings, and these

are used instead of the original DM map to generate the field for an unaberrated input wavefront. The result (Figure 60) is a bit different from the original shown in Figure 59.

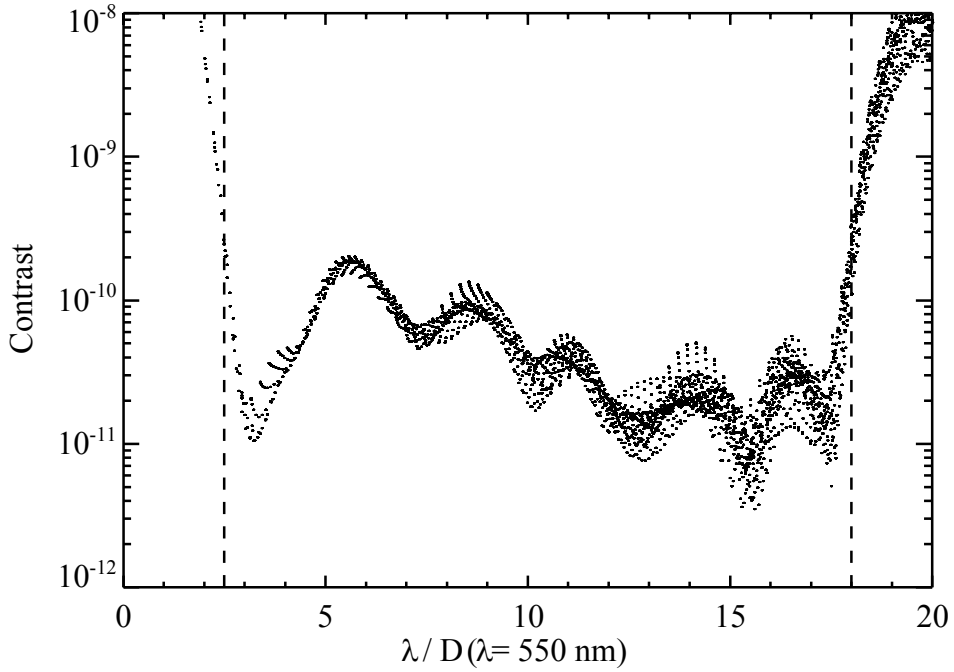

Figure 59. Contrast at each point in the broadband (λ = 500 – 600 nm) field using the original HBLC optimized design.

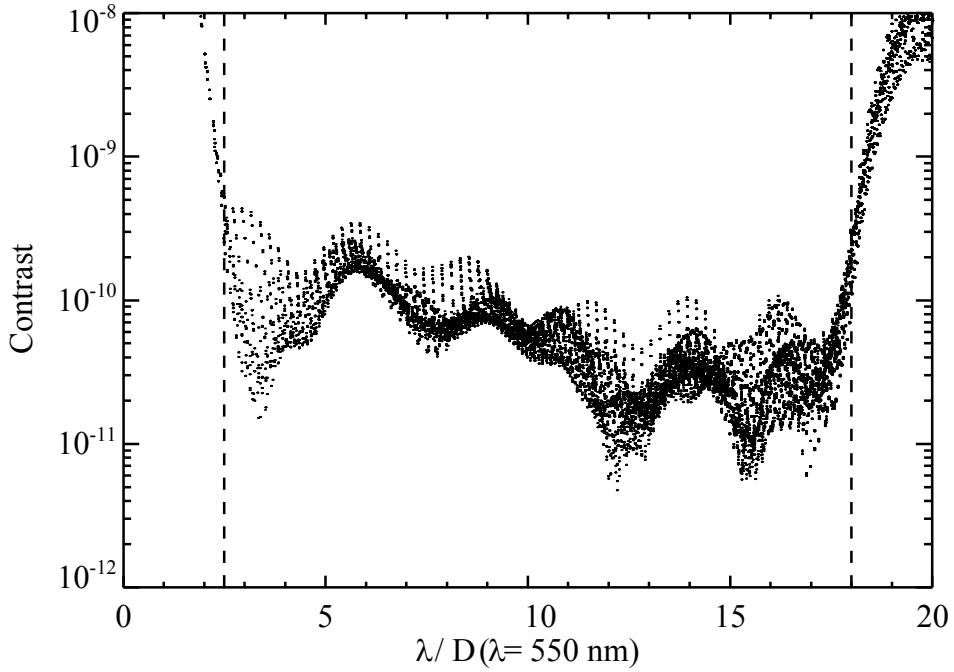

Figure 60. Contrast at each point in the broadband (λ = 500 – 600 nm) field using the derived DM actuator settings instead of the original DM surface map.

The mean contrast in this "actuator settings" version is 1.4×10^{-10} over $r = 2.5 – 3.5$ λ/D and 5.8×10^{-11} over $r = 2.5 – 18$ λ/D, and the maximum value is 4.5×10^{-10}. These differences highlight the distinction between the accuracy of the propagation algorithm and the accuracy of the match between two systems (be they model versus model or model versus actual). The only significant difference between the two runs is the choice of DM influence function. The same problem would occur if the function assumed in the model does not match the real one used in a testbed or in an actual telescope. The latter case is often the more difficult, as it involves both accurately measuring and then representing the system.

Re-optimizing the HBLC design

While the "actuator settings" version technically meets the Milestone 1 requirement by having a mean contrast of $<10^{-10}$ over 2.5 – 18 λ/D, the result is not appealing due to the large fraction of pixels (14%) above 10^{-10} and the degraded contrast at the inner radius. To restore the result to better match the original solution and without changing the focal plane mask profiles, we undertake another round of optimization, this time adjusting both the DM #1 and DM #2 actuators (and nothing else). Rather than using the original optimization software, we use a combination of PROPER and the iterative Electric Field Conjugation (EFC) wavefront optimization algorithm. EFC operates very much like the original software, using the predicted image-plane field changes for each DM actuator to derive settings that minimize the energy inside the dark hole. The details of EFC are discussed elsewhere (Give'on et al. 2007) and in the TDEM Milestone #2 report, where it is used in our simulations to create dark holes in aberrated systems.

Using the map-fitted DM #1 actuator settings as a starting point, the PROPER+EFC routines are applied to derive updated DM #1 and DM #2 actuator pistons (the routines are run until convergence). The resulting broadband field agrees closely with that produced by the original optimizer. However, given that the original field has a number of values above 10^{-10} contrast, we adjust the weighting function to force EFC to put more effort into pushing those regions down. The result (Figure 61) is a significant improvement over the previous one (Figure 60) with a mean contrast of 3.7×10^{-11} over 2.5 – 18 λ/D. Only 7% of the pixels are above 10^{-10}, which is a little better than the original solution. However, the contrast between 2.5 – 3.5 λ/D is 1.2×10^{-10}, a bit worse than the original.

To see if we can improve the field at the inner radius, we adjust the weights accordingly and ran the optimizer with the new solution. However, the result is simply close to the original: the inner radius values go down but those elsewhere in the field (especially at $r = 5 – 7$ λ/D) go up. Further experiments confirm this result – pushing down in one region inevitably pushes up another. In the end, we choose to use the solution described above and shown in Figure 61 as our default, as it provides the largest area below 10^{-10} contrast. The DM settings are shown in Figure 62. As a reminder, the original solution's focal plane mask profiles are not changed, nor is the Lyot stop configuration.

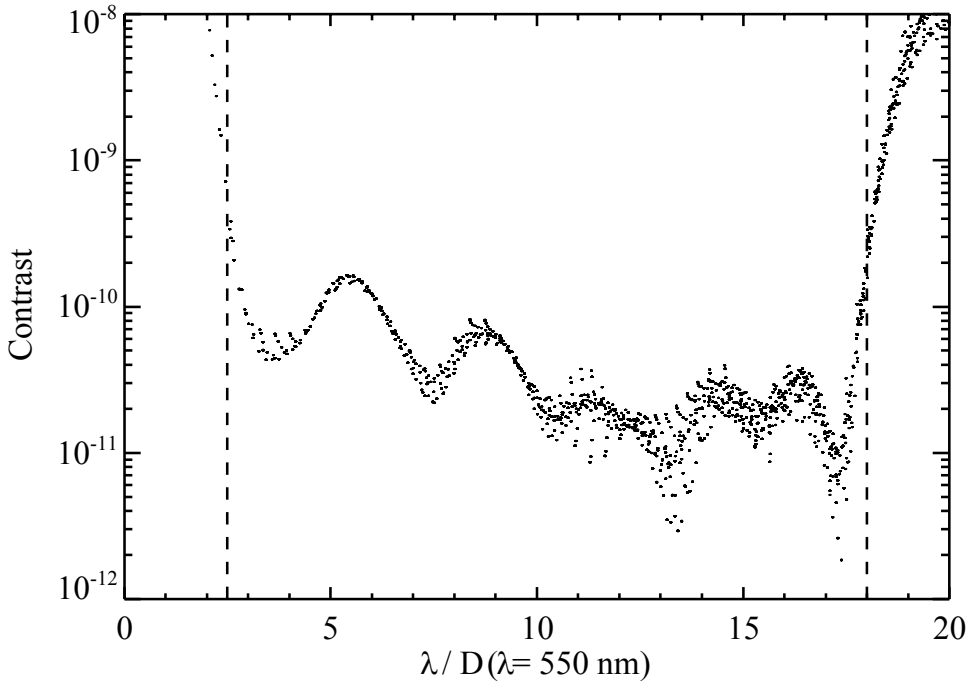

Figure 61. Contrast at each point in the broadband (λ = 500 – 600 nm) field using the final HBLC solution.

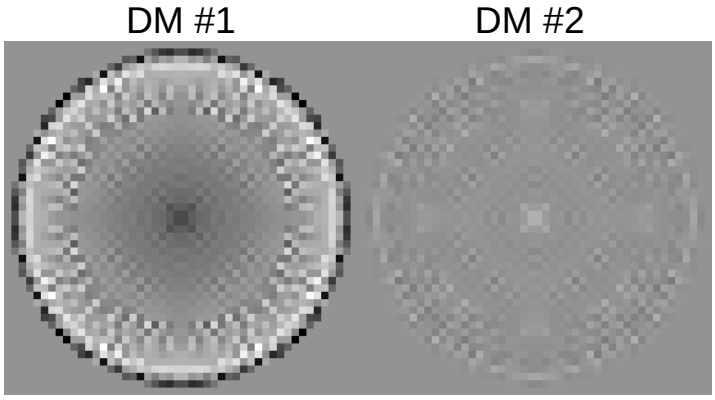

Figure 62. DM actuator piston settings for the final HBLC solution. Both DMs are shown scaled between -10.2 nm to +6.4 nm in piston. DM #1 is located at a pupil image and DM #2 is 1.0 meters downstream.

More actuators, smaller dark hole = darker dark hole

The HBLC design described above is tailored to our pre-defined Milestone #1 requirements of a dark hole size of 2.5 – 18 λ/D, DMs with 46 actuators across the pupil, and a bandpass of 500 – 600 nm. What if instead we use the full 48 actuators available on a DM to span the pupil and shrink the dark hole size down bit? We run the optimizer to

produce a design including 48 actuators across the pupil and a dark hole size of 2.5 – 15 λ/D (same bandpass). The resulting field, plotted in Figure 63, provides a substantially darker result than our Milestone #1 design, with a mean contrast of 5.3×10^{-12} over both r = 2.5 – 3.5 λ/D and 2.5 – 15 λ/D. Only DM #1 is used. This, like the previous designs, utilize real materials and their corresponding optical properties as computed with thin film equations. This focal plane mask could actually be fabricated and used in a real system.

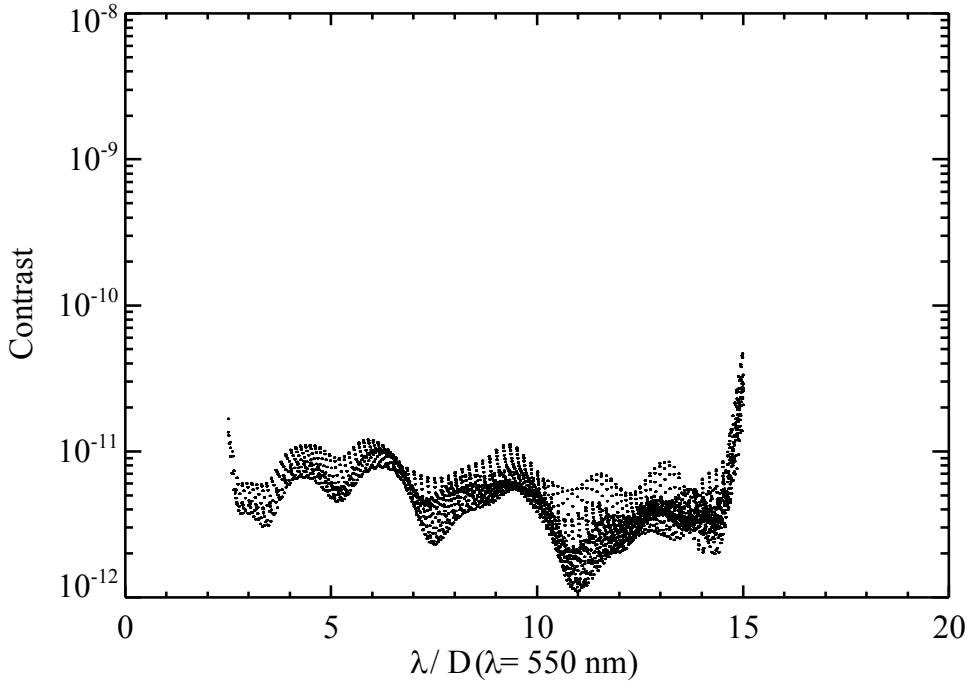

Figure 63. Contrast at each point in the broadband (λ = 500 – 600 nm) field for an experimental HBLC solution optimized for r = 2.5 – 15 λ/D and using 48 DM actuators across the pupil.

The additional actuators are partially responsible for the improvement, but most of it is due to shrinking the dark hole outer radius. In a "pure" system where DM actuators are represented by delta functions, the DM can control wavefront errors with spatial frequencies of $\leq N_{act}/2$ cycles/beam diameter, by Nyquist theory. This corresponds to a control radius of $N_{act}/2$ λ/D radians in the image plane. However, in reality the DM, which is not a set of delta functions, has somewhat less spatial frequency control. More importantly, however, the edge of the dark hole is adjacent to the outer region of the field that is several orders of magnitude brighter. This edge transition cannot be instantaneous because it is convolved with the point spread function (the PSF itself is broadened by the reduction in the beam diameter due to the Lyot stop). Therefore, it is necessary to "back off" from the edge a bit. The decrease in the dark hole outer radius allows greater freedom to achieve a deeper, albeit smaller, hole.

This deeper, smaller dark hole solution is ***not*** used to meet the TDEM Milestones.

Implementation of the HBLC

The HBLC is the easiest of the three coronagraphs in our study to represent numerically in the simulations. Unlike PIAA, which has the issue of wavefront remapping, and the VVC, which has sampling issues near the vortex singularity and requires accurate calculation of the effect of the small opaque spot, the HBLC can be accurately represented by simple phase and amplitude modulations of the wavefront and settings on the DMs.

The optimization program produces a well-sampled, complex-valued focal plane mask profile for each specified wavelength spanning the bandpass. These can be interpolated to match the sampling used in the simulations. The DM settings (provided by the 2nd optimization program) are used as-is and fed to PROPER's DM model.

HBLC Milestone #1: Efficiency and accuracy

The HBLC does not require any special propagation algorithms or "tricks" to represent it in the numerical simulations since it simply modifies phase and amplitude with relatively smooth patterns (i.e., not hard-edged and thus easily interpolated to different samplings). The wavelength-dependent phase and amplitude modulations are computed using analytic thin film equations, assumed to be exact, and the known material properties. There is, in essence, no reference algorithm result to which a faster algorithm can be compared. For the HBLC Milestone #1 accuracy requirement, we mostly rely on the demonstrated accuracy of the PROPER propagation algorithms detailed in Appendix B.

As with the VVC, we also compare results from different computational grid sizes (4096 × 4096 & 2048 × 2048) and corresponding samplings at the focal plane mask (0.079 & 0.157 λ/D, respectively, with 322 pixels across the pupil in both cases). We use the phase map defined earlier to provide a contrast floor of 10^{-5} (low contrast case) and 10^{-10} (high contrast case). The simulation is run through the system layout shown in Figure 5. This is the same layout that is used for Milestone #2.

In the initial runs, the default aberration map that produced a 10^{-10} contrast field in PIAA and the VVC actually generate a ~2.5×10^{-10} contrast field in the HBLC. This is likely due to the greater masking of the pupil by the Lyot stop, which both blocks a portion of the aberration map and makes the PSF broader (degrading contrast). The 10^{-5} contrast map has the same problem. To compensate, these maps are multiplied by a constant in the prescription to bring the resulting fields to the desired contrast levels.

The parameters and results of the simulations are presented in Table 9. Figure 64 to Figure 67 show the raw and difference contrast maps for both aberration maps. Surprisingly, unlike for PIAA and the VVC, the accuracy for both contrast levels is about the same, ~7×10^{-14}. The residuals in the 10^{-5} contrast case (Figure 65) are a little more irregular, but are generally the same as those for 10^{-10} (Figure 67).

The accuracies mentioned above are based on the differences between the 4K × 4K and 2K × 2K wavefront grid fields, and it is likely that some numerical artifacts cancel out. They are actually better than the PROPER accuracies specified in Appendix B.

Therefore, we define the Milestone #1 accuracy metric for the HBLC to be equal to PROPER's. The HBLC PROPER propagation algorithms meet the Milestone #1 efficiency and accuracy requirements .

Table 9. Milestone #1 Metrics for the Hybrid Bandlimited Coronagraph

	Reference 4K	**Efficient 2K**
Array size	4K × 4K	2K × 2K
***Beam ratio*(1)**	0.079	0.157
***Execution time*(2)**	53.3 sec	14.3 sec
***" for 2077 actuators*(3)**	30.8 hrs	8.3 hrs
Requirement		<48 hrs
Low Contrast field mean	1.0×10^{-5}	1.0×10^{-5}
***LC Accuracy*(4)**		6.9×10^{-14} (~0%)
***LC Milestone #1 Accuracy*(5)**		**0.02%**
LC Requirement		$<1.0 \times 10^{-7}$ (<1%)
High Contrast field mean	1.1×10^{-10}	1.1×10^{-10}
***HC Accuracy*(4)**		6.9×10^{-14} (0.07%)
***HC Milestone #1 Accuracy*(5)**		**0.6%**
HC Requirement		$<1.1 \times 10^{-12}$ (<1%)

(1) The diameter of the beam to the wavefront grid diameter. Also equivalent to the sampling at the focal plane in units of λ/D radians.
(2) Propagation time from DM1 to final image on a dual quad-core Xeon system.
(3) Total time to propagate a poke of each of 2077 actuators (based on the time for a single propagation).
(4) The accuracy derived from comparing the 2K × 2K result to the 4K × 4K one.
(5) The accuracy of PROPER when compared to S-Huygens for the given contrast level for a Lyot coronagraph. This is considered to be the Milestone #1 accuracy for the HBLC because it is worse than that derived from the 2K vs. 4K comparison..

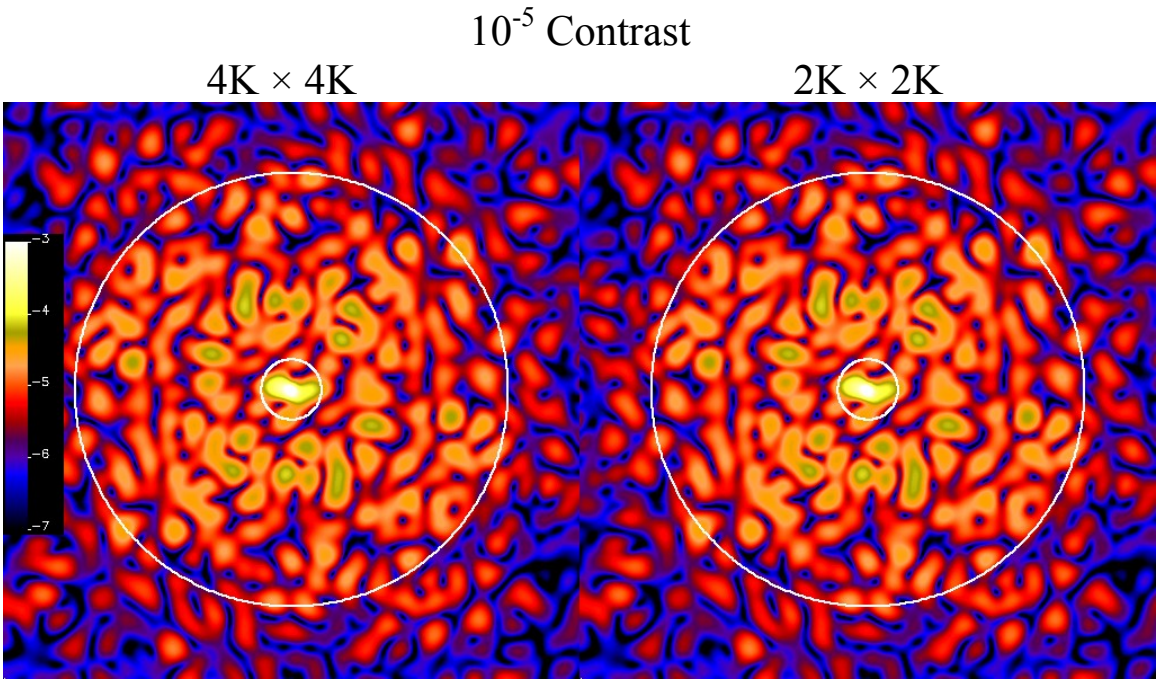

Figure 64. Monochromatic (550 nm) contrast maps for HBLC Milestone 1 simulation runs (1.0×10^{-5} mean contrast) using 4K × 4K arrays (left) and 2K × 2K arrays (right). The contrast evaluation region of r = 2.5 - 18 λ/D is marked with circles. The colorbar indicates $\log_{10}$(contrast) and ranges from 10^{-7} up to 10^{-3}.

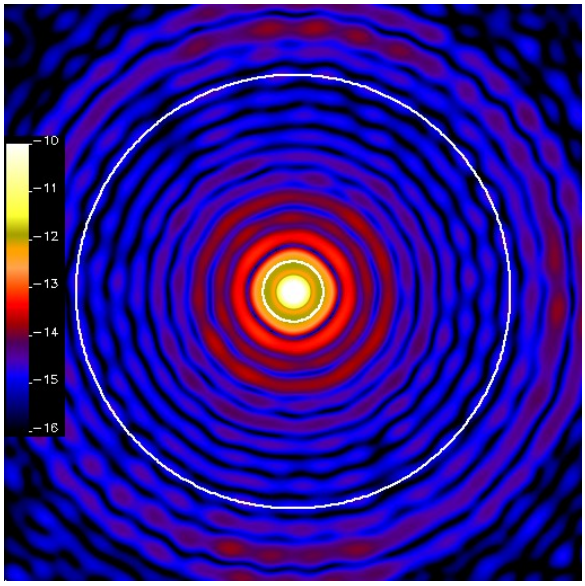

Figure 65. Difference between the 4K × 4K and 2K × 2K electric fields (10^{-5} contrast), displayed as contrast. The contrast evaluation region of r = 2.5 - 18 λ/D is marked with circles. The colorbar indicates $\log_{10}$(contrast) and ranges from 10^{-16} up to 10^{-10} (different scaling from contrast maps above).

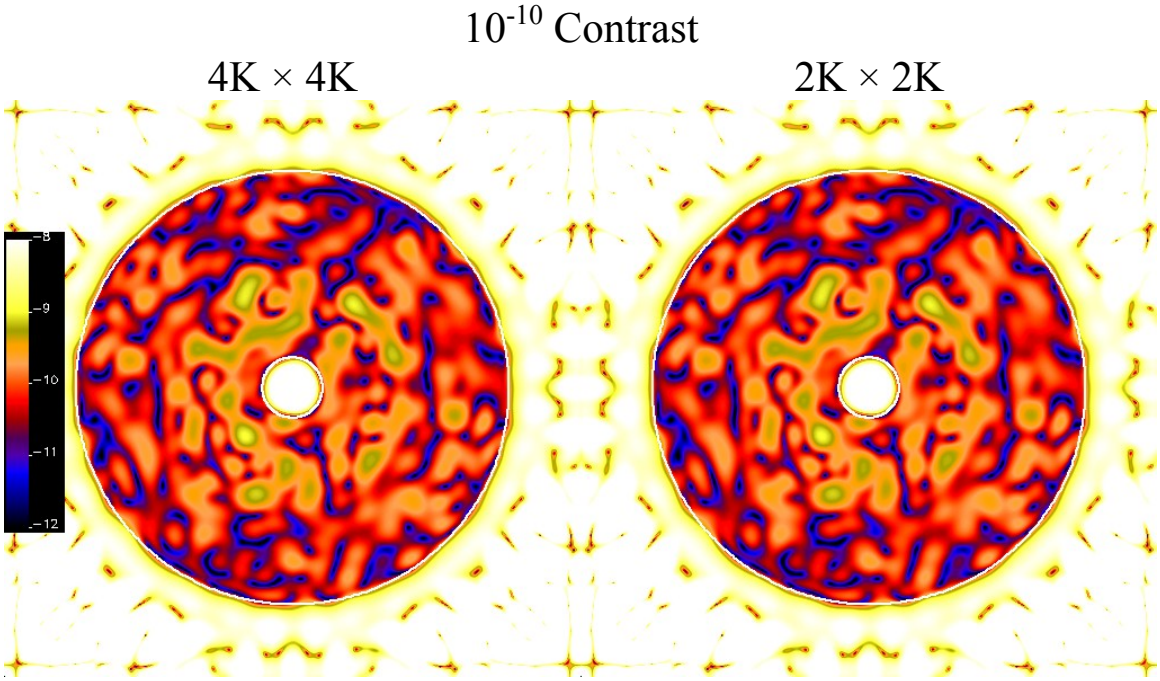

Figure 66. Monochromatic (550 nm) contrast maps for HBLC Milestone 1 simulation runs (1.1×10^{-10} mean contrast) using 4K × 4K arrays (left) and 2K × 2K arrays (right). The contrast evaluation region of r = 2.5 - 18 λ/D is marked with circles. The colorbar indicates $\log_{10}$(contrast) and ranges from 10^{-12} up to 10^{-8}.

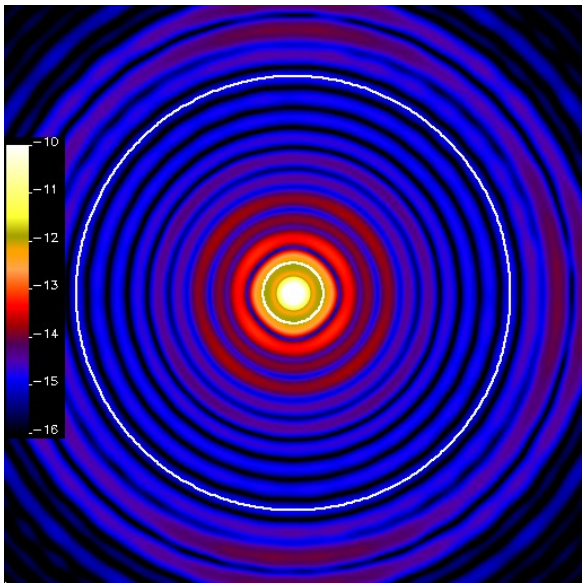

Figure 67. Difference between the 4K × 4K and 2K × 2K electric fields (10^{-10} contrast), displayed as contrast. The contrast evaluation region of *r* = 2.5 - 18 λ/D is marked with circles. The colorbar indicates $\log_{10}$(contrast) and ranges from 10^{-16} up to 10^{-10} (different scaling from contrast maps above).

Appendix F. Acronyms

ASMCS	Advanced Space Mission Concept Study
D	Diameter of the entrance pupil (typically)
DM	Deformable mirror
EFC	Electric field conjugation
FFT	Fast Fourier transform
FITS	Flexible Image Transport Specification
FT	Fourier transform
GB	Gigabyte
HBLC	Hybrid band-limited coronagraph
HCIT	High Contrast Imaging Testbed
HWP	Half-wave plate
IDL	Interactive Data Language
IWA	Inner working angle
JPL	Jet Propulsion Laboratory
LCP	Liquid crystal polymer
MFT	Matrix Fourier transform
OAP	Off axis parabola
OWA	Outer working angle
PASP	PIAA angular spectrum propagation
PIAA	Phase induced amplitude apodization
PSD	Power spectral density
PSF	Point spread function
QWP	Quarter wave plate
RAM	Random access memory
RMS	Root mean square
RS	Rayleigh Sommerfeld (propagation algorithm)
TDEM	Technology development for exoplanets missions
TPF-C	Terrestrial Planet Finder - Coronagraph
VVC	Vector vortex coronagraph