PROCEEDINGS OF SPIE

Phase-retrieval-based wavefront metrology for high-contrast coronography: 2. Reconstructions through a shaped pupil apodizer

Gregory R. Brady, Peter Petrone III, Iva Laginja, Keira Brooks, Manxuan (Rebecca) Zhang, et al.

SPIE.

Phase-retrieval-based wavefront metrology for high contrast coronography: 2. Reconstructions through a shaped pupil apodizer

Gregory R. Brady[a], Peter Petrone[a], Iva Laginja[a], Keira Brooks[a], Manxuan (Rebecca) Zhang[ab], Mamadou N'Diaye[c], Christopher Moriarty[d], John Hagopian[e], Rémi Soummer[a]

[a] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[b] California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA
[c] Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice Cedex 4, France
[d] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[e] Advanced NanoPhotonics, Inc., 4437 Windsor Farm Rd. Harwood, MD, 20776, USA

ABSTRACT

We discuss the use of parametric phase-diverse phase retrieval to characterize and optimize the transmitted wavefront of a high-contrast apodized pupil coronagraph with and without an apodizer. We apply our method to correct the transmitted wavefront of the HiCAT (High contrast imager for Complex Aperture Telescopes) coronagraphic testbed. This correction requires a series of calibration steps, which we describe. The correction improves the system wavefront from 16 nm RMS to 3.0 nm RMS for the case where a uniform circular aperture is in place. We further measure the wavefront with the apodizer in place to be 11.7 nm RMS. Improvement to the apodized pupil phase retrieval process is necessary before a correction based on this measurement can be applied.

Keywords: Phase retrieval, coronagraph, wavefront-sensing, deformable mirror, high-contrast imaging

1. INTRODUCTION

Coronagraph instruments on future space telescopes will have contrasts on the order of 10^{-10} to enable imaging and spectroscopy of exoplanets. This is enabled by precision wavefront sensing and control in which a dark zone is created iteratively by manipulating one or more deformable mirrors (DMs) using feedback from the science imagery. There are several methodologies for achieving this,[1–4] but generally they assume that the phase of the optical field is small enough that the approximation

$$e^{i\theta} \approx 1 + i\theta \tag{1}$$

is valid. In other words, the end-to-end wavefront error of the coronagraph must already be small before the dark-zone procedure is executed.

To achieve this level of wavefront correction, careful design, tolerancing and alignment of the coronagraph is of paramount importance so that the initial wavefront error is small. However this is only effective to a point, and the wavefront error is ultimately limited by fabrication errors in the optics, deformations of the optics due to mounting and thermal effects, and limitations on positioning accuracy and feedback during alignment. To reduce the residual errors further, before the dark zone procedure, it is possible to employ the DM to apply a correction, provided that we can measure the wavefront error that is to be corrected. This pre-correction ensures that the coronagraph possesses a wavefront for the linear approximation of Equation 1 to be valid.

Further author information: send correspondence to Gregory Brady: E-mail: gbrady@stsci.edu, Telephone: +1 410 338 6504

Here we describe a procedure for measuring the coronagraph's wavefront error using phase retrieval. Phase retrieval does not make any assumptions about the optical field except for those required for scalar diffraction theory and consequently does not rely on the linear approximation described above. The particular algorithm used here is of a parametric type, where parameters of interest describing the phase and other pertinent system characteristics are determined using an optimization algorithm. We describe calibration steps necessary to apply the measured wavefront to the DM. We have successfully applied these corrections to the DM in a testbed system and demonstrated dramatic improvement in the wavefront in the case where the system had a uniform circular pupil. We also have measured the system wavefront after the introduction of a complicated apodizer in the pupil plane. Finally, we describe the further application of phase retrieval in our system to calibrate and cross-check the results of a future low-order wavefront sensor (LOWFS) that will maintain wavefront stability of the system while the dark zone is created.

2. PARAMETRIC PHASE RETRIEVAL ALGORITHM

Phase retrieval refers to a procedure in which the phase of an optical field in the exit pupil of a system is determined from measurements of the intensity produced by that field in a plane (or planes, in the case of a phase-diverse phase retrieval algorithm) near its paraxial focus.[5] The relationship of the fields in the pupil and the measurement planes are assumed to be governed by scalar (Fourier) optics principles. One common type of phase retrieval algorithm can be referred to as iterative-transform-type algorithms, and includes the well-known Gerchberg-Saxton-Misell algorithm and variants.[6–9] These iterative algorithms propagate the field back and forth between the pupil and measurement planes, enforcing known constraints from the data and system characteristics, ultimately converging on an estimate of the phase in the pupil.

An alternate approach, used in this work, describes the system in terms of a set of model parameters and calculates an estimate of the intensity distribution in the measurement planes from them. These estimated intensities are compared to the measured intensities using an objective function. An optimization algorithm is used to vary the parameters in such a way that the minimum value of the objective function is found. Here the phase parameters are of the primary interest. The coefficients of a polynomial expansion (e.g. Zernike polynomials) or phase values on a sampled grid can be used as parameters.[8,10] Other system parameters can also be introduced to make the algorithm more robust to errors in the assumed values of system parameters, such as the distance from the exit pupil to the measurement planes,[11,12] the lateral positioning of the measured intensity patterns,[12,13] the pupil amplitude distribution,[14,15] or the pixel sampling of the image data.[16] It is important to note that this class of algorithm can straightforwardly make use of multiple intensity measurements, where there is a known phase relationship between the measurements. This is a technique known as phase-diverse phase retrieval and typically enhances algorithm robustness. This ability, combined with the ability to optimize for unknown system parameters, makes the parametric approach superior to the iterative-transform-type algorithms.

2.1 Phase Model

In a parametric phase retrieval algorithm the phase, $\widehat{\phi}(m,n;\widehat{\alpha}_j)$ is often represented as the coefficients $\widehat{\alpha}_j$ of a set of J basis functions $Z_j(m,n)$, so that

$$\widehat{\phi}(m,n;\widehat{\alpha}_j) = \sum_j \widehat{\alpha}_j Z_j(m,n), \qquad (2)$$

where the indices (m,n) correspond to one of the $N \times N$ samples in the pupil plane. For simplicity, we assume that the data can be cropped to a square array. Throughout this paper we use the convention that a variable or function topped with a circumflex is a quantity that could be estimated by the phase retrieval algorithm, while those with no circumflex are fixed values defined by measurement or choice of parameterization. Common basis sets that might be used include the Zernike polynomials or alternately a set of Kronecker delta functions $Z_j(m,n) = \delta_{m-m_j,n-n_j}$ centered at each pixel (usually termed a "point-by-point" estimate.) In the point-by-point case, $J = N^2$. Often it is desirable to change the representation of the phase used as the algorithm nears the minima. For example, we typically begin the phase optimization using a small number of low-order Zernike terms, which tends to smooth the optimization space and make the algorithm less prone to getting stuck in local minima. When the values of the Zernikes coefficients have converged we then switch to the point-by-point

representation to allow for reconstruction of higher spatial frequency features, since we are now more likely to be within the smaller capture range of the point-by-point optimization.

This parameterized phase is combined with a representation of the pupil amplitude $\widehat{A}(m,n)$ to calculate an estimate of the field in the pupil,

$$\widehat{E}_p(m,n;\widehat{\alpha}_j) = \widehat{A}(m,n)\exp\left[i\widehat{\phi}(m,n;\widehat{\alpha}_j)\right]. \tag{3}$$

Note that we have written the amplitude here as an estimable parameter. It is most common to fix the pupil amplitude as a distribution known by design or by an independent pupil camera measurement. However, with enough data or other constraints it is possible to estimate these amplitudes along with the phases, although this can adversely affect algorithm robustness and convergence.[12,14]

2.2 Propagation Model

We employ a two step propagation to calculate the estimated field in each of the measurement planes from the field in the pupil plane. First, Fresnel diffraction is used to calculate the field $\widehat{E}_f(p,q)$ in the paraxial focal plane,

$$\widehat{E}_f(p,q;\widehat{\alpha}_j) = \frac{1}{N}\exp\left[\frac{i\pi}{\lambda f}\left(p^2\Delta_p^2 + q^2\Delta_q^2\right)\right]\sum_{m,n}^{N}\widehat{E}_p(m,n;\widehat{\alpha}_j)\exp\left[\frac{-i2\pi}{N}(mp+nq)\right], \tag{4}$$

where λ is the wavelength, f is the distance between the exit pupil and paraxial focal plane, p and q are sample indices, and (Δ_p,Δ_q) are the spacing between samples (in units of length) in the paraxial focal plane. This is generally equivalent to the pixel spacing of the detector used to record the image data. The second propagation step uses the angular spectrum method to traverse the small distance between the paraxial focal plane and each of the slightly defocused measurement planes. The angular spectrum of plane waves $\widehat{U}_f(m,n;\widehat{\alpha}_j)$ in the paraxial focal plane is computed using a Fourier transform,

$$\widehat{U}_f(m,n;\widehat{\alpha}_j) = \frac{1}{N}\sum_{p,q}^{N}\widehat{E}_f(p,q;\widehat{\alpha}_j)\exp\left[-i\frac{2\pi}{N}(pm+nq)\right]. \tag{5}$$

The angular spectrum is then propagated to each of the K measurement planes,

$$\widehat{U}_k(m,n;\widehat{\alpha}_j) = \widehat{U}_f(m,n;\widehat{\alpha}_j)\exp\left(i2\pi\widehat{z}_k\sqrt{\frac{1}{\lambda^2} - m^2\Delta_m^2 - n^2\Delta_n^2}\right), \tag{6}$$

where (Δ_m,Δ_n) are frequency domain sample spacings (in units of inverse length) given by

$$\Delta_{m,n} = \frac{1}{N\Delta_{p,q}} \tag{7}$$

and $\widehat{z}_k$ is the distance to the kth measurement plane. This is typically used as a free parameter for at least a portion of the optimization to allow for correction of possible positioning errors. The propagated field in the spatial domain is computed from the angular spectrum using an inverse Fourier transform,

$$\widehat{E}_k(p,q;\widehat{\alpha}_j) = \frac{1}{N}\sum_{m,n}^{N}\widehat{U}_k(m,n;\widehat{\alpha}_j)\exp\left[i\frac{2\pi}{N}(mp+nq)\right]. \tag{8}$$

The intensity in each measurement plane is then calculated,

$$\widehat{I}_k(p,q;\widehat{\alpha}_j) = \left|\widehat{E}_k(p,q;\widehat{\alpha}_j)\right|^2. \tag{9}$$

The final step of our propagation model employs the Fourier shift theorem to laterally shift the computed intensity distribution so that it can be registered to the measured intensity as part of the phase retrieval optimization. First, we perform the forward transform

$$\widehat{f}_k(m, n; \widehat{\alpha}_j) = \frac{1}{N} \sum_{p,q}^{N} \widehat{I}_k(p, q; \widehat{\alpha}_j) \exp\left[-i\frac{2\pi}{N}(mp + nq)\right]. \tag{10}$$

Then, we multiply by the linear phase Fourier filter

$$\widehat{g}_k(m, n; \widehat{\alpha}_j) = \exp\left[\frac{-i2\pi}{N}(m\widehat{x}_k + n\widehat{y}_k)\right] \widehat{f}_k(m, n; \widehat{\alpha}_j), \tag{11}$$

where $(\widehat{x}_k, \widehat{y}_k)$ is the amount by which the image is shifted in pixel units. We then take the inverse Fourier transform to arrive at our final modeled intensity distribution,

$$\widehat{G}_k(p, q; \widehat{\alpha}_j) = \frac{1}{N} \sum_{m,n}^{N} \widehat{g}_k(m, n; \widehat{\alpha}_j) \exp\left[i\frac{2\pi}{N}(mp + nq)\right]. \tag{12}$$

2.3 Objective Function

The objective function calculates a value that is indicative of the agreement between the modeled and measured intensity distributions and is minimized by the optimization algorithm to arrive at our estimate of the phase. In this work we use a bias- and gain-insensitive objective function developed by Thurman et al.,[12, 17]

$$\Phi = 1 - \frac{1}{K} \sum_{k}^{K} \frac{\left[\sum_{p,q} W_k(p, q) G_k(p, q) \widehat{G}_k(p, q; \widehat{\alpha}_j)\right]^2}{\left[\sum_{p,q} W_k(p, q) G_k^2(p, q)\right]\left[\sum_{p,q} W_k(p, q) \widehat{G}_k^2(p, q; \widehat{\alpha}_j)\right]}, \tag{13}$$

where $G_k(p, q)$ is one of K measured intensity patterns and $W_k(p, q)$ is a weighting function used to ignore the contribution of bad pixels or to favorably weight certain portions of the intensity pattern with, for example, better Signal-to-Noise ratio (SNR). This metric can be thought of as a normalized overlap integral between $G_k(p, q)$ and $\widehat{G}_k(p, q; \widehat{\alpha}_j)$. If the two distributions are identical the value of the term to the right of the minus sign is unity, hence the objective function is zero.

2.4 Optimization

To minimize the objective function we employ widely-used optimization algorithms such as the conjugate gradient[18] or limited-memory BFGS[19–21] algorithms. These algorithms require a calculation of the gradient of the objective function with respect to the optimization parameters, i.e. for the parameters describing the phase we need to calculate

$$\frac{\partial \Phi}{\partial \widehat{\alpha}_j} \tag{14}$$

for each of the J phase parameters. While a finite difference approximate value of the gradient can be used, this is very computationally expensive when a large number of parameters are used. This is especially true when a point-by-point parameterization is used, where $N \times N$ derivatives must be calculated and N is typically a few hundred to a few thousand. It is thus typically dramatically more efficient to use an analytic expression for these derivatives. Efficient expressions for the parameters and propagation model discussed above have been derived and are compiled in the appendix of Reference 12, which we have employed here. Other researchers have demonstrated the use of algorithmic differentiation to automatically compute analytic derivatives without having to perform tedious longhand calculations.[22]

To summarize, our phase retrieval code has the capability to optimize over:

- Lateral spot position in each phase-diverse image, $\widehat{x}_k$, $\widehat{y}_k$

- Axial location of each phase-diverse image, $\widehat{z}_k$

- Pupil amplitude distribution, $\widehat{A}(m,n)$

- Pupil phase expressed as point-by-point phase values, $\widehat{\phi}(m,n)$

- Pupil phase expressed as coefficients of Zernike polynomials (or other basis), $\widehat{\alpha}_j$.

These can be used jointly or individually, with the output of a preceding optimization serving as the input for subsequent optimizations as necessary for the problem at hand. In the code used in this work, the following sequence is typically used:

1. Pre-process image data to calculate data weight masks, $W_k(p,q)$, and initial guesses for lateral positions, $\widehat{x}_k$, $\widehat{y}_k$, from image centroids.

2. Optimize over lateral positions alone, using the value from the image centroids of step 1 as a starting guess.

3. Optimize jointly over lateral position (using result from step 2 as a starting guess), axial position and Zernike coefficients.

4. Fix values for lateral and axial positions at values determined in step 3, optimize over point-by-point phase values using Zernike-derived phase from step 3 as starting guess.

5. Optimize over lateral and axial position and point-by-point phase values using results from step 4 as starting guess.

6. Optimize over lateral and axial position, point-by-point phase using results from step 5 as starting guess, and pupil amplitude using assumed pupil as starting guess.

This multiple-step process allows us to determine the unknown parameters of the system robustly. If we were to proceed directly to step 6 it is more likely that the optimization would stagnate in a local minima. Steps 4 to 6 are necessary only when point-by-point phase values are desired and may be omitted when Zernikes are all that are desired.

3. IMPLEMENTATION ON THE HICAT TESTBED

We use the phase retrieval methodology described above to measure and correct the transmitted wavefront of the HiCAT (High Contrast Imager for Complex Aperture Telescopes) coronagraphic testbed from the illuminating point source to the occulting aperture (also termed Focal Plane Mask) of the system. A schematic diagram of HiCAT is shown in Figure 1.

At its core HiCAT is an Apodized Pupil Lyot Coronagraph (APLC). The system simulates a star using a single-mode optical fiber collimated by an Off-Axis Parabolic (OAP) mirror, illuminating the pupil of a telescope simulator. The telescope simulator consists of a pupil amplitude aperture stop that is imaged onto an Iris AO hexagonal-segment deformable mirror by an OAP and a parabolic mirror. This DM allows us to simulate the piston, tip and tilt motions of the segments of a segmented aperture telescope. The DM pupil plane is again reimaged by the parabola and an OAP onto a reflective apodizer. The apodizer has two purposes. The first is to control diffraction and to begin establishing the dark zone of the coronagraph. The second purpose of the apodizer is to act as a tip-tilt mirror to stabilize the beam under closed-loop servo control with feedback from the tip-tilt sensor. The apodizer pupil plane is imaged by an OAP and a toroidal mirror onto one of a pair of Boston Micromachines continuous face-sheet Kilo DMs. The pupil plane DM directs the light onto the second DM which is located some distance from the pupil plane. This dual DM arrangements allows for control of both amplitude and phase to create a high-contrast dark zone.

From the second DM a second toroidal mirror and a spherical mirror focus the light to the plane of the FPM. The FPM is responsible for blocking the majority of the star's light. Between the toric and spherical mirrors a

high-quality mirror can be inserted using a flip mount to intercept the converging beam and direct it to a camera on a motorized stage. This camera is used to collect phase-diverse phase retrieval data sets, the results of which are our topic here. An extremely high-quality point-source image is formed at the FPM, allowing it to block star light most effectively. The FPM is a high quality mirror with a hole in it, with the desired coronagraphic light reflecting off of the FPM and the suppressed star light passing through it. After reflection off of the FPM a spherical mirror is used to re-image the pupil to the Lyot stop of the coronagraph. The Lyot stop blocks light diffracted by the hard edge of the FPM that falls outside of, and at the edge of, the system pupil. Finally, fold mirrors and lenses are used to form a coronagraphic image with the desired dark zone or a pupil image on the science camera and pupil cameras, respectively.

The portion of the light passing through the FPM hole is directed to two additional wavefront sensing subsystems. The first is a Low-Order WaveFront Sensor (LOWFS), which images the point source on a phase mask and subsequently images the pupil on a camera to capture a Zernike phase contrast image of the pupil. This image can be analyzed to arrive at a quantitative estimate of the phase in the pupil of the system while the coronagraph is in operation. This allows us to measure and correct wavefront errors that may vary while the dark zone is created or during coronagraphic imaging, such as those due to instabilities in the telescope. However the FPM constitutes a Fourier low-pass filter on the pupil image, so only low spatial frequency phase terms can be sensed with this arrangement. The other wavefront sensing subsystem is the tip-tilt sensor used to drive the tip-tilt actuators on the apodizer. The tip-tilt sensor is composed of a wide-field coarse target acquisition camera that drives the apodizer so that the point image falls on the smaller quadrant photodiode array, which gives the precision positioning information.
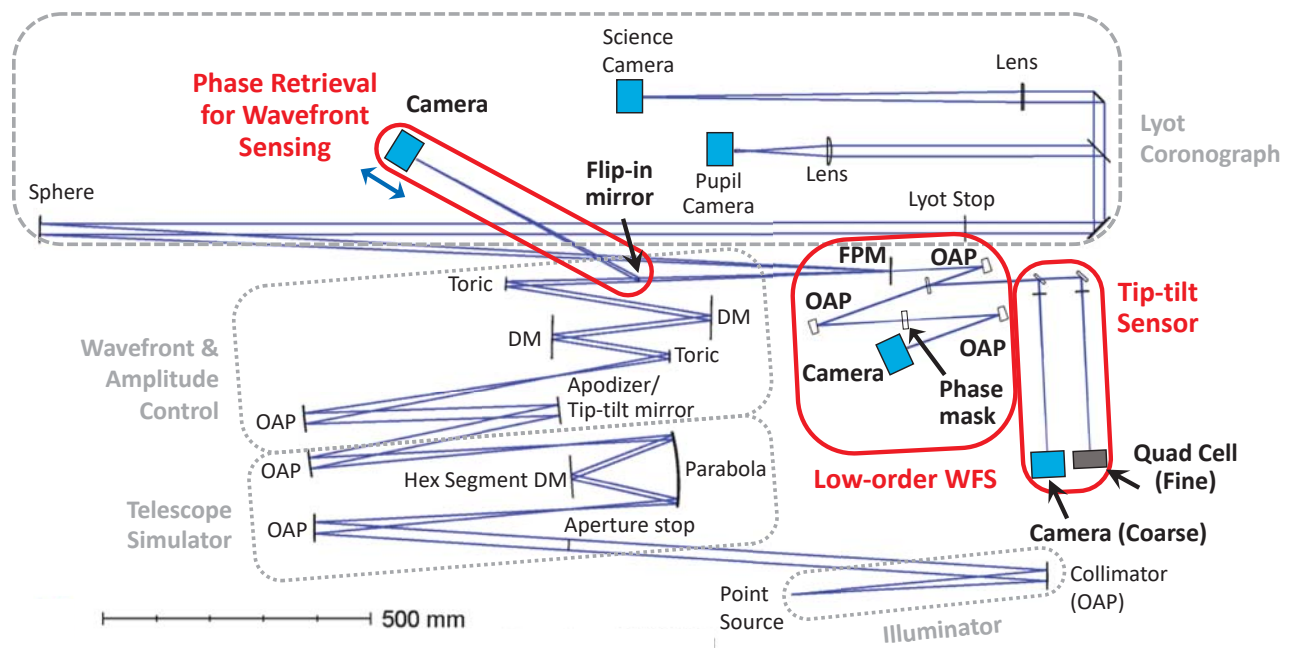

Figure 1. HiCAT testbed layout. The testbed includes a telescope simulator (IRIS-AO 37 hexagonal segment deformable mirror (DM) combined with a pupil mask/aperture stop to produce the central obstruction and support structures); a wavefront and amplitude control subsystem incorporating an apodizer and two Boston Micromachines Kilo DMs; a classical Lyot coronagraph with a hard-edge focal plane mask (FPM) and a Lyot stop as well as cameras in the coronagraphic focal plane and pupil plane; and a set of wavefront sensing subsystems including a phase retrieval arm, a Zernike-phase-contrast-based low-order wavefront sensor and a tip-tilt sensor.

3.1 Measurements Over a Uniform Circular Aperture

For our initial experiments we simplified the nominal HiCAT arrangement that was described above. First, the aperture stop corresponding to the segmented aperture was replaced with a 22 mm diameter circular aperture. This allowed us to sample more of the area of the Boston Micromachine DMs and thus better correct aberrations that are near the edge of the designed segmented system pupil, which has a circumscribed diameter of 19.725 mm. We also removed the apodizer and segmented DM and replaced them with high-quality flat mirrors to create a uniform pupil. After measuring the initial system wavefront over the 22 mm aperture, we applied a correction derived from it to the Boston Micromachines DM conjugate to the pupil. Then the wavefront was measured again over the 22 mm aperture and also over an 18 mm aperture, which is more representative of the diameter of the segmented, apodized pupil.

3.2 Circular Pupil Calibration

To apply the wavefront corrections derived from our measured phase distributions to the DM it was necessary to perform a geometric calibration. This calibration was done in three steps and was achieved by applying a series of calibration phase patterns to the DM and recovering them using our phase retrieval algorithm with high spatial resolution. First we applied a letter "F" pattern to the DM, which allowed us to determine how the DM and phase retrieval wavefront are flipped (left-right and up-down) and rotated with respect to each other. In this step we also confirmed the polarity of the retrieved phase as compared to the voltages applied to the DM. The DM command and retrieved phase are shown in Figure 2, with the determined flips and rotations applied.

(a)   (b)
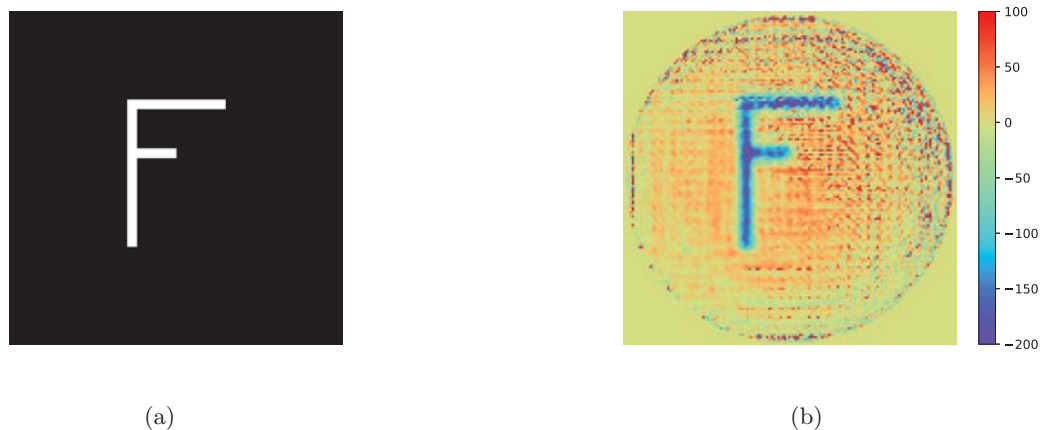
Figure 2. Letter F pattern (a) DM command and (b) retrieved phase. Color scale is in nm.

The second calibration step determined the location of the center of the DM in the wavefront data retrieved by the phase retrieval algorithm. Four actuators arranged in a square around the center of the DM were poked so that the intersection of the diagonals located the center of the DM aperture. This pattern and the corresponding retrieved phase is illustrated in Figure 3.

The third calibration step determined the location of individual DM actuators in the recovered wavefront data. This was done by putting a "checkerboard" pattern on the DM in which every fourth actuator is poked, illustrated in Figure 4. Sixteen data sets were recorded where each actuator in a 4 × 4 cell is poked, locating every actuator. The corresponding phase retrieval results are shown in Figure 5. Mapping the pupil in this way calibrates out any pupil imaging distortion between the DM plane and the exit pupil of the system.

3.3 Uncorrected System Wavefront Measurement over Circular Pupil

In order to measure the system wavefront, we first applied our best-known flat maps to the two Boston Micromachines DMs, determined before their installation using a 4D Technologies Fizeau interferometer. This, combined
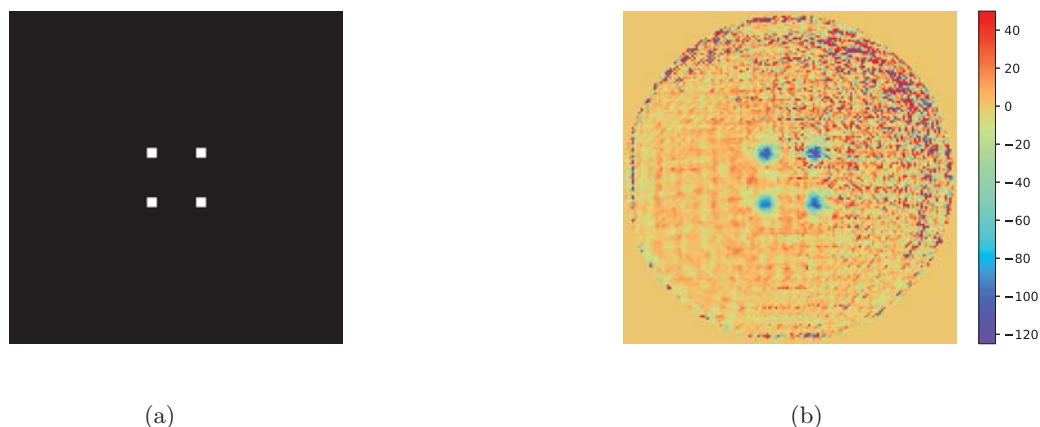
(a)

(b)

Figure 3. (a) DM command and (b) retrieved phase for the case where four actuators are poked to determine the location of the center of the DM aperture with respect to the phase retrieval phase map. Color scale is in nm.
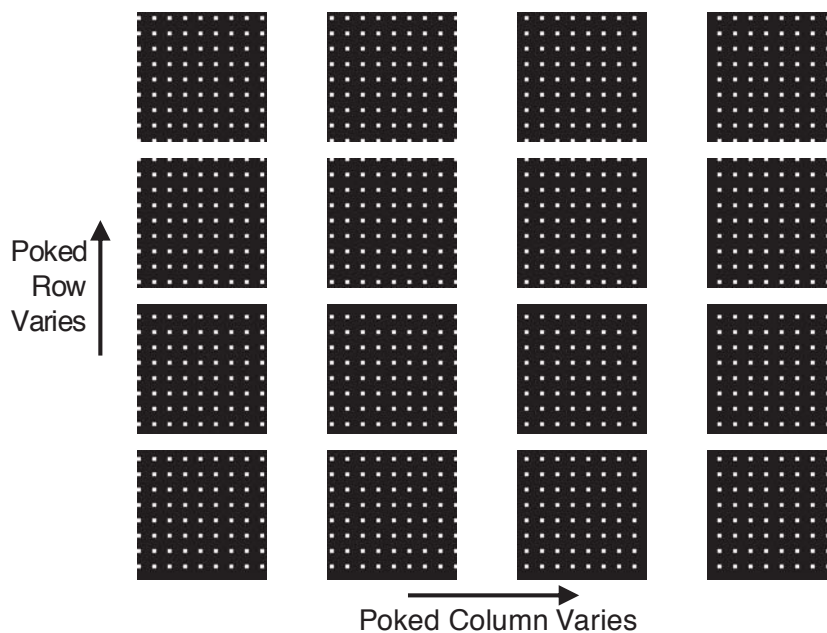

Figure 4. Sixteen checkerboard DM commands which poke, in turn, all of the DM actuators. The actuators on the DM are grouped into cells of 4×4 actuators, one of which is poked in each command. Note the shift of which actuator row and column is active in each cell between each of the 16 commands.

with careful specification and alignment of the optics, ensured that the system was as well-corrected as possible before any additional corrections were applied using our phase retrieval measurements. The through-focus intensity data used as the input to the phase retrieval algorithm is shown in the top row of Figure 6. Each of these images is the sum of 200 registered images. This summing process results in an image with high dynamic range and excellent SNR. The phase recovered from this data using the first 45 Zernike polynomials is shown in Figure 7 and the corresponding modeled intensity distributions are shown in the bottom row of Figure 6. The Peak-to-Valley (PV) wavefront error is 120 nm and the Root-Mean-Squared (RMS) wavefront error is 16 nm.
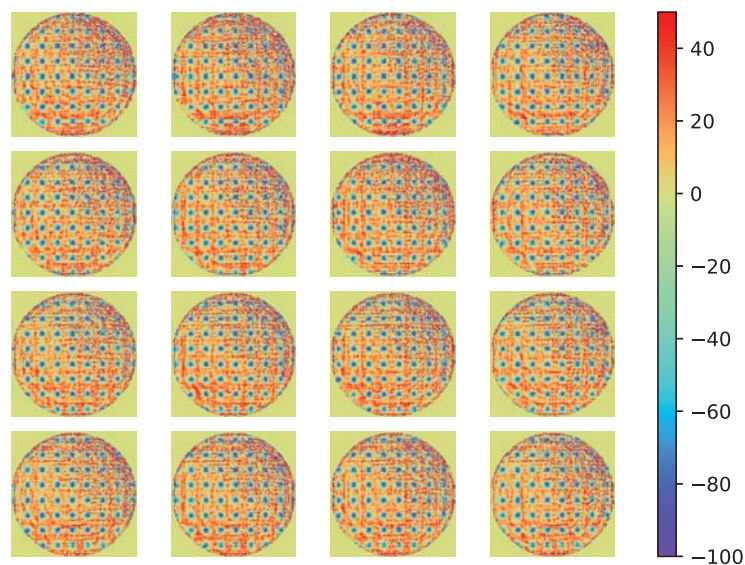
Figure 5. Retrieved phase maps for the 16 checkerboard patterns applied to the DM. Color scale is in nm.
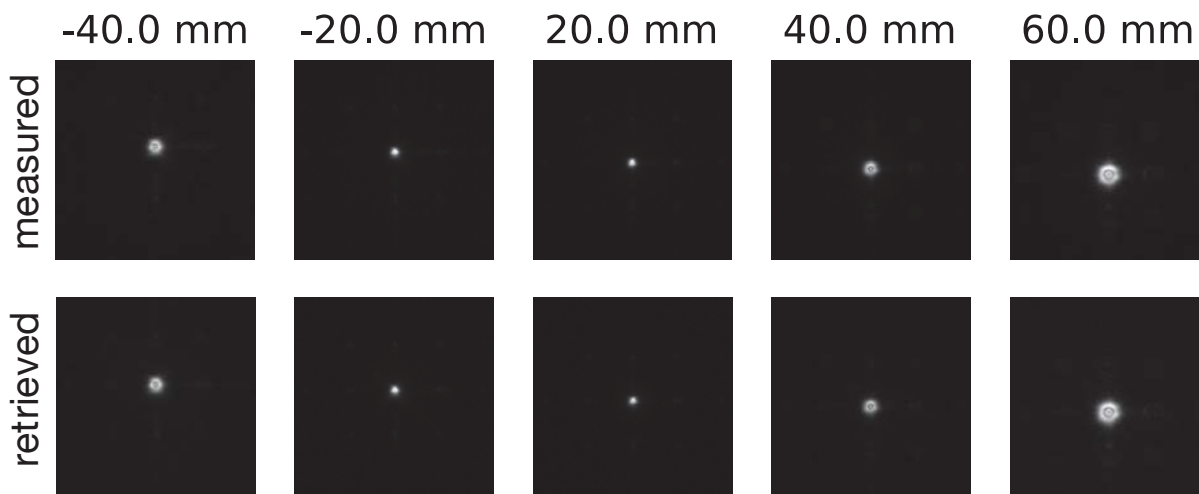

Figure 6. Intensity patterns measured in the system before any wavefront correction was performed using phase retrieval. The images in the top row are the minimally processed intensity measurements from the phase retrieval camera. The images in the bottom row are the intensity distributions modeled by our phase retrieval algorithm using the recovered wavefront. The distances labeled across the top of the figure indicate the axial position of the measurement plane with respect to the nominal focus plane.

Figure 7. Reconstruction of the initial wavefront in the system exit pupil before applying corrections. The PV wavefront error is 120 nm and the RMS wavefront error is 16 nm. The color scale is in units of nm.

3.4 Corrected Wavefront Over a 22 mm Circular Aperture

The conjugate of the wavefront in Figure 7 was applied to the pupil DM using the calibration and mapping determined previously. The recorded intensity data is shown in the top row of Figure 8. The wavefront recovered from this data using the first 45 Zernike polynomials (following the ordering of Born & Wolf[23]) is shown in Figure 9. The PV wavefront error is 35.2 nm and the RMS wavefront error is 5.5 nm. This is a significant improvement over the uncorrected case. This is a strong indication of the precision and accuracy of our phase retrieval results and demonstrates the validity of our methodology for applying the corrections based on phase retrieval to the DM.
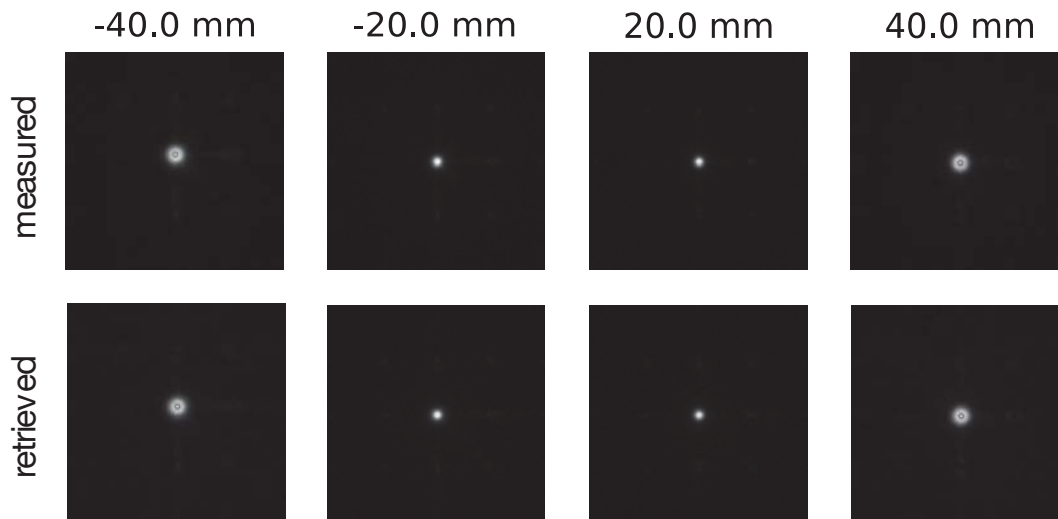

Figure 8. Intensity patterns measured in the system after a correction based on the wavefront shown in Figure 7 was applied over a 22 mm diameter circular pupil. The images in the top row are the minimally processed intensity measurements from the phase retrieval camera. The images in the bottom row are the intensity distributions modeled by our phase retrieval algorithm using the recovered wavefront. The distances labeled across the top of the figure indicate the axial position of the measurement plane with respect to the nominal focus plane.
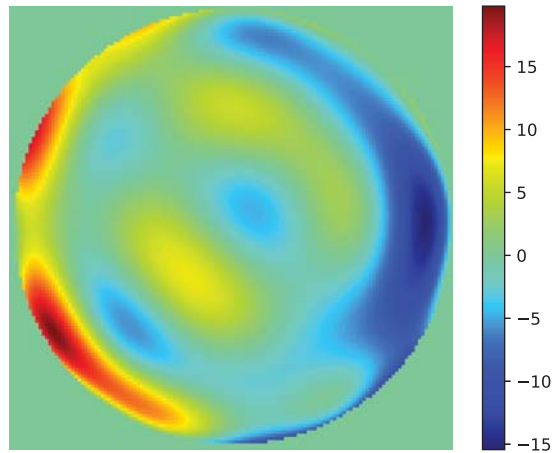
Figure 9. Corrected system wavefront using correction derived from phase retrieval data in Figure 7. This is the wavefront in the exit pupil of the system in the case where an oversized, 22 mm diameter aperture was used. The PV wavefront error is 35.2 nm and the RMS wavefront error is 5.5 nm. The color scale is in units of nm.

3.5 Corrected Wavefront Over an 18 mm Circular Aperture

After the first wavefront correction was applied the oversized 22 mm aperture was removed and the 18 mm pupil aperture, which is more representative of the final system aperture, was inserted. A subsequent data set was taken. Example intensity images are shown in the top row of Figure 10. The retrieved wavefront is shown in Figure 11. The PV wavefront error is 23.9 nm and the RMS wavefront error is 3.0 nm. An image of the in-focus PSF of the system, resulting from the average of 1000 images is shown in Figure 12. It is difficult to distinguish this from a plot of a perfect theoretical result.
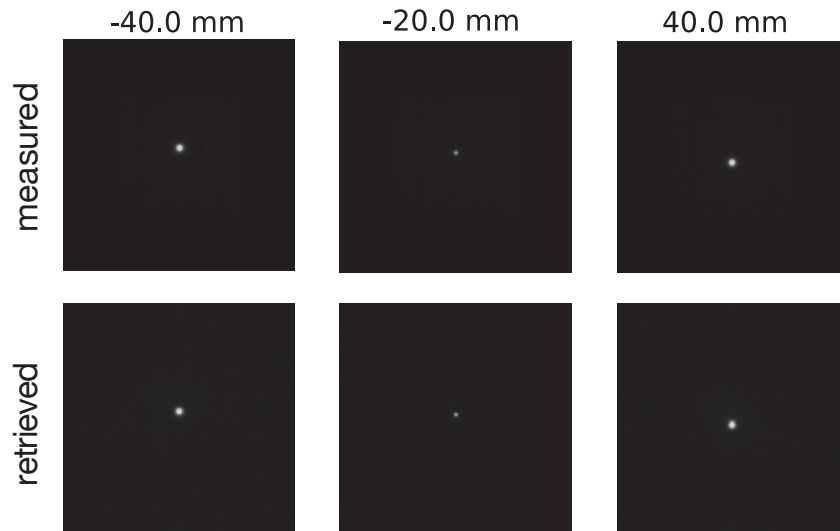

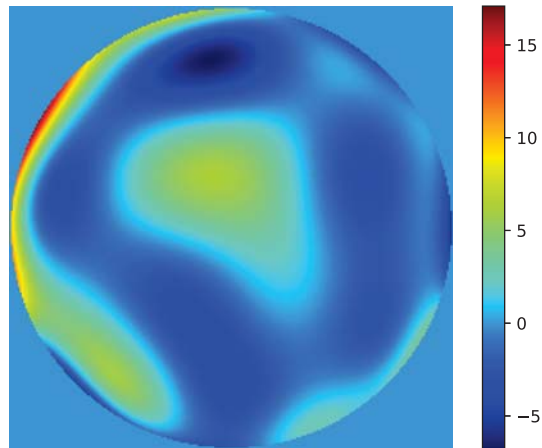Figure 10. Corrected PSFs, measured and retrieved over 18 mm aperture

Figure 11. Reconstruction of corrected system wavefront, using correction derived from phase retrieval reconstruction shown in Figure 9, over the nominal 18 mm aperture. The PV wavefront error is 23.9 nm and the RMS wavefront error is 3.0 nm. The color scale is in units of nm.
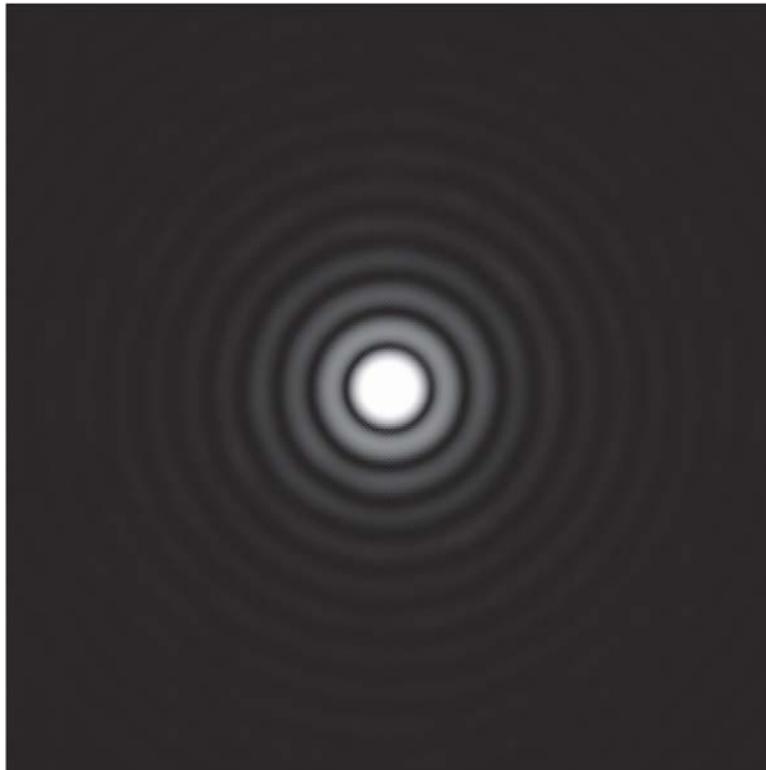

Figure 12. In-focus PSF over 18 mm aperture. The data here is the sum of 1000 registered camera frames, resulting in a very high SNR image with high dynamic range. The data is displayed on a logarithmic scale with a compressed dynamic range so that very dim outer rings are visible. The PSF is almost indistinguishable from a perfect theoretical result for a circular pupil with no aberrations and, qualitatively, is consistent with our measured wavefront error of 3 nm RMS.

3.6 Measurements Over an Apodized Aperture

As described above, HiCAT is an apodized pupil Lyot coronagraph for segmented apertures. As such, we need to retrieve the phase over a pupil of this type. It is not adequate to correct the phase using the circular aperture result since the apodizer introduces its own aberrations to the system. The type of apodizer we use employs a high spatial frequency binary pattern, where light is either reflected or strongly absorbed, to create a dark zone in the coronagraph image plane through diffraction from its amplitude distribution. A photograph of such an optical element is shown in Figure 13.

Figure 13. Photograph showing an example of a carbon nanotube apodizer.

The apodizers we use are precision flat silicon substrates on which carbon nanotubes are grown on a patterned catalyst layer. While the substrate at the the start of processing is very flat, there are concerns that high-temperature fabrication steps could slightly warp the substrate, despite efforts to anneal out residual stresses. Thus it is important to characterize the apodizer in the context of HiCAT and correct any errors due to it.

The complicated amplitude distribution of the apodizer increases the difficulty of the phase retrieval problem. Although the designed amplitude distribution is well-known, there are some practical considerations that must be taken into account in order to translate this into a suitable amplitude distribution for phase retrieval. First, the apodizer will be rotated at some angle with respect to the phase retrieval camera. Second, the amplitude pattern imprinted on the beam after reflection is typically narrower in one cardinal direction than the other due to the non-normal incidence of the beam on the apodizer. Third, the apodizer is typically designed on a much finer sampling grid than is used for phase retrieval. As a result of these factors it is necessary to downsample and interpolate the designed apodizer pattern to arrive at an amplitude distribution suitable for use in the phase retrieval algorithm. This processing constitutes a source of error for the phase retrieval estimate. An example of a processed apodizer amplitude distribution is shown in Figure 14.

In this work estimates of the apodizer rotation and foreshortening were arrived at by manual analysis of the phase retrieval intensity patterns in an image editing software package. It would be desirable to implement this estimation as part of the phase retrieval process and this may be a topic of future research.
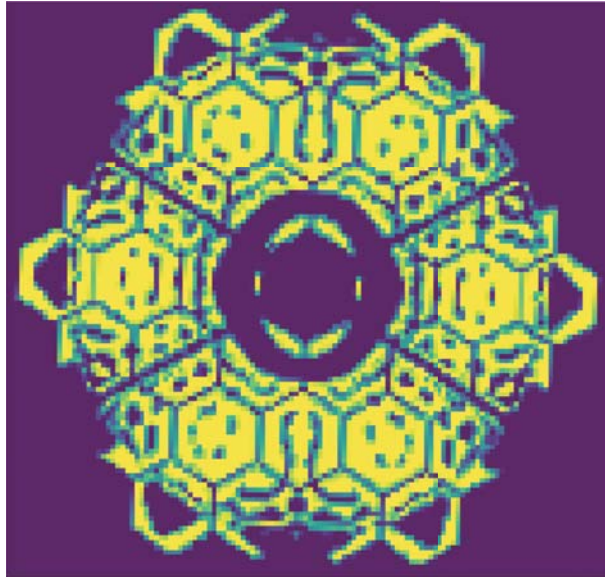
Figure 14. Processed apodizer data used as pupil amplitude distribution for phase retrieval algorithm. The pattern is rotated, foreshortened and downsampled relative to the designed apodizer. On this color scale dark blue represents an amplitude transmission of 0 and yellow represents an amplitude transmission of 1.

We collected a series of images with the apodizer in place, which appear as the top row of Figure 15. We then retrieved the phase in the pupil using 21 Zernike terms. The resulting modeled intensity distributions from our phase retrieval algorithm appear in the lower row of Figure 15. The corresponding wavefront is shown in Figure 16. The wavefront has a PV error of 86.2 nm and an RMS error of 11.2 nm. These values are roughly in-line with typical values measured using other surface metrology techniques on processed apodizer substrates.
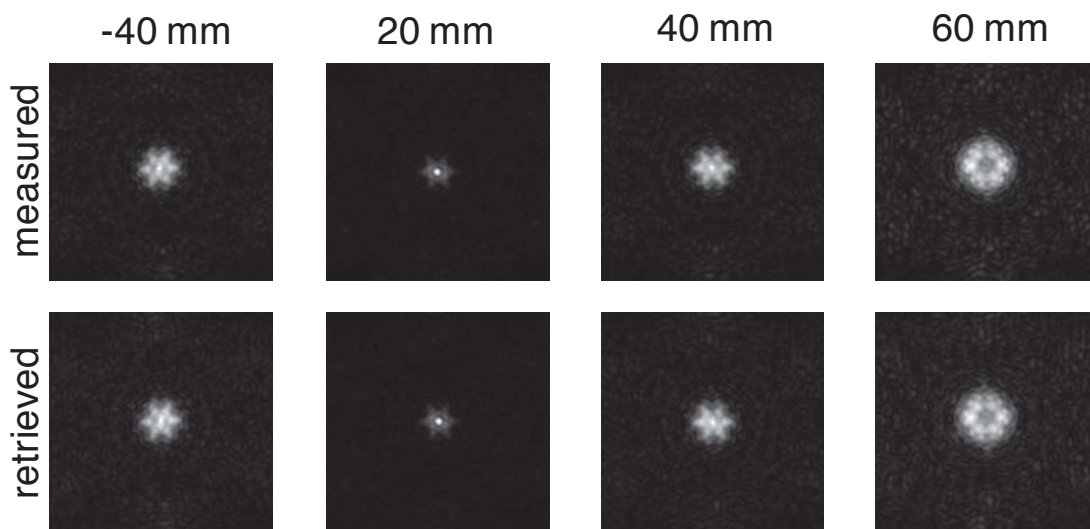

Figure 15. Measured (top row) and retrieved (bottom row) intensity patterns from the system with an apodizer in place. The distances labeled across the top of the figure indicates the axial position of the measurement plane with respect to the nominal focus plane.
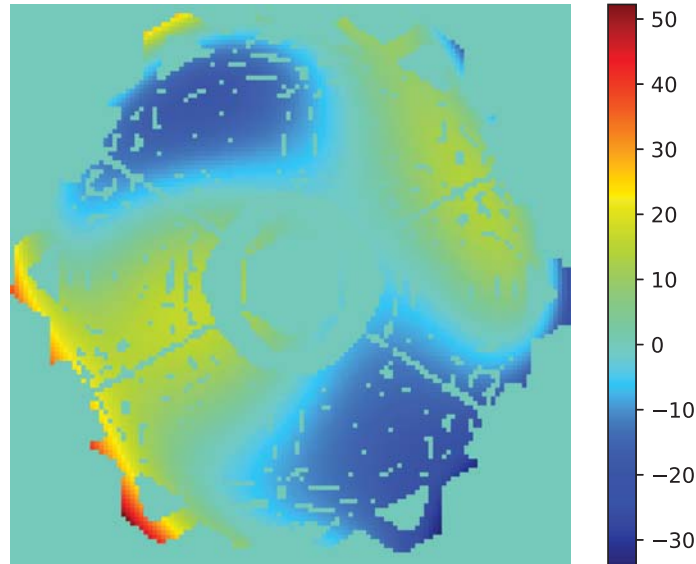
Figure 16. Measured wavefront with an apodizer in the system. The PV wavefront error is 86.2 nm and the RMS wavefront error is 11.7 nm. The color scale is in units of nm.

4. FUTURE WORK

4.1 Apply Correction with Apodizer in Place

We have not yet successfully used phase retrieval measurements to apply a wavefront correction to the DM with the apodizer in place. Phase retrieval with high spatial resolution has proven difficult to achieve with the high spatial frequency content of the apodizer and uncertainties in the apodizer's rotation angle, foreshortening, downsampling and interpolation. As a consequence we have not yet been able to complete a pupil calibration as we described in the circular aperture case above. We are currently working on solutions to this problem.

4.2 Further Application of Phase Retrieval in HiCAT to LOWFS Calibration

An additional key wavefront sensing subsystem of HiCAT is a low-order wavefront sensor that is currently in development. Figure 17 shows the optical arrangement of our Zernike sensor LOWFS. A pair of OAPs forms an image of the intensity in the FPM plane, truncated by the FPM aperture, on a Zernike contrast phase mask. The pupil is then reimaged, with phase encoded as amplitude, on the LOWFS camera by an additional OAP. This image is analyzed to arrive at a low-pass filtered version of the wavefront in the system pupil. Unlike with the phase retrieval approach, this estimate is available concurrently with the dark-zone digging procedure and corongrgaphic science imaging and requires less processing. As a result it can be used as feedback to maintain the systems dynamic stability, at least for low-order contributions.

The non-trivial optics in the LOWFS impart their own errors to the measured wavefront. However, if the wavefront error of the beam forming the FPM focus can be measured both using phase retrieval and the LOWFS, the errors of the LOWFS can be determined and subtracted from subsequent measurements,

$$\phi_{LOWFS\ Calibrated\ Measurement} = \phi_{LOWFS\ Raw\ Measurement} - \phi_{HiCAT\ PR\ Measurement}. \tag{15}$$

Consequently, in the future, we plan to also use phase retrieval as a cross-check and sanity check for LOWFS measurements during the test and integration of the subsystem.
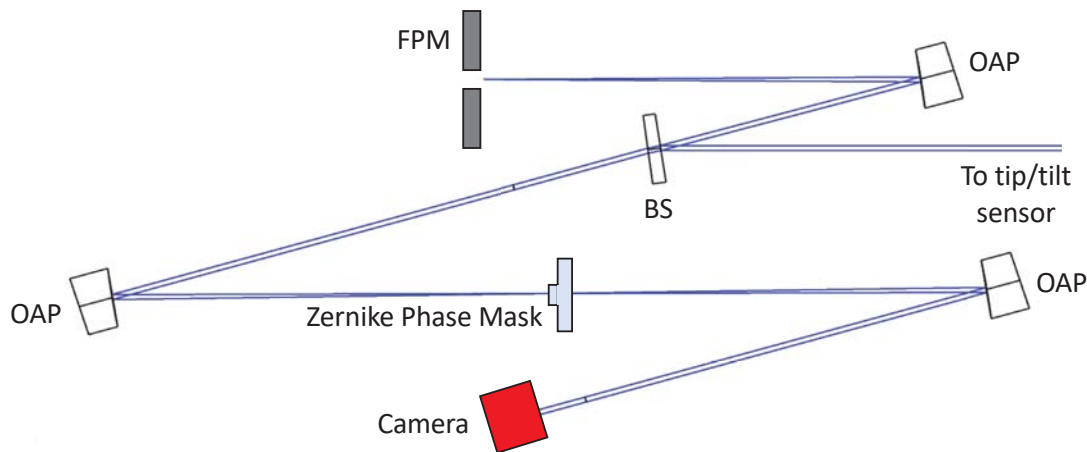
Figure 17. Optical layout of the low-order wavefront sensor in development for HiCAT.

5. SUMMARY

We described the parametric approach to phase retrieval, including details of our propagation model and the parameters and corresponding derivatives used as inputs to an optimization algorithm. We described its application to the measurement of the transmitted wavefront error of the HiCAT coronagraphic testbed system. We were successful in performing this measurement when the pupil of the system was a uniformly illuminated circle and when the pupil was apodized by a complex binary pattern. In the case of the circular pupil, we also successfully performed the calibrations necessary to register the measured wavefront maps to the DM actuators, including confirming the orientation and polarity of the data, determining the center point of the DM and calibrating for distortions. This allowed us to improve the measured wavefront error from 16 nm RMS before correction to 3.0 nm RMS after correction for the case of a circular, uniform pupil. We described the process of rotating, foreshortening and downsampling the apodized pupil pattern to determine the amplitude distribution required to perform phase retrieval with the apodized aperture. We were able to measure the transmitted wavefront error of the system with the apodizer in place and the circular pupil correction applied to be 11.7 nm RMS. However we have not yet been able to derive a further correction with the apodizer in place due to difficulties in high spatial resolution phase retrieval in the presence of the apodizer. Our results are summarized in Table 1. Our excellent results with a circular pupil demonstrate the utility and efficacy of phase retrieval for characterizing and correcting coronagraphs for high-contrast imaging.

Table 1. Summary of wavefront errors measured

Case	RMS Wavefront Error (nm)
Initial Aligned System (DMs flattened)	16
Correction Applied over 22 mm Aperture	5.5
Correction Applied over 18 mm Aperture	3.0
Apodizer in Place (circular aperture correction applied)	11.7

ACKNOWLEDGMENTS

REFERENCES

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

[2] Bottom, M., Femenia, B., Huby, E., Mawet, D., Dekany, R., Milburn, J., and Serabyn, E., "Speckle nulling wavefront control for Palomar and Keck," *Proc.SPIE* **9909**, 990955–1 – 990955–16 (2016).

[3] Miller, K., Guyon, O., and Males, J., "Spatial linear dark field control: stabilizing deep contrast for exoplanet imaging using bright speckles," *J. Astron. Telesc. Instrum. Syst.* **3**, 049002 (Oct. 2017).

[4] Pueyo, L., Kay, J., Kasdin, N. J., Groff, T., McElwain, M., Give'on, A., and Belikov, R., "Optimal dark hole generation via two deformable mirrors with stroke minimization," *Appl. Opt.* **48**, 6296–6312 (Nov 2009).

[5] Fienup, J. R., "Reconstruction of an object from the modulus of its Fourier transform," *Opt. Lett.* **3**, 27–29 (Jul 1978).

[6] Gerchberg, R. W. and Saxton, W. O., "Practical agorithm for determination of phase from image and diffraction plane pictures," *OPTIK* **35**(2), 237–& (1972).

[7] Misell, D. L., "A method for the solution of the phase problem in electron microscopy," *J. Phys. D* **6**(1), L6 (1973).

[8] Fienup, J. R., "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**, 2758–2769 (Aug 1982).

[9] Dean, B. H., Aronstein, D. L., Smith, J. S., Shiri, R., and Acton, D. S., "Phase retrieval algorithm for JWST flight and testbed telescope," *Proc. SPIE* **6265**, 626511–1 – 626511–17 (June 2006).

[10] Fienup, J. R., "Phase retrieval for undersampled broadband images," *J. Opt. Soc. Am. A* **16**, 1831–1837 (Jul 1999).

[11] Brady, G. R., *Application of Phase Retrieval to the Measurement of Optical Surfaces and Wavefronts*, dissertation, University of Rochester (2008).

[12] Thurman, S. T., DeRosa, R. T., and Fienup, J. R., "Amplitude metrics for field retrieval with hard-edged and uniformly illuminated apertures," *J. Opt. Soc. Am. A* **26**, 700–709 (Mar 2009).

[13] Brady, G. R., Guizar-Sicairos, M., and Fienup, J. R., "Optical wavefront measurement using phase retrieval with transverse translation diversity," *Opt. Express* **17**, 624–639 (Jan 2009).

[14] Brady, G. R. and Fienup, J. R., "Nonlinear optimization algorithm for retrieving the full complex pupil function," *Opt. Express* **14**, 474–486 (Jan 2006).

[15] Thurman, S. T. and Fienup, J. R., "Complex pupil retrieval with undersampled data," *J. Opt. Soc. Am. A* **26**, 2640–2647 (Dec 2009).

[16] Jurling, A. S. and Fienup, J. R., "Phase retrieval with unknown sampling factors via the two-dimensional chirp z-transform," *J. Opt. Soc. Am. A* **31**, 1904–1911 (Sep 2014).

[17] Thurman, S. T. and Fienup, J. R., "Phase retrieval with signal bias," *J. Opt. Soc. Am. A* **26**, 1008–1014 (Apr 2009).

[18] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., [*Numerical Recipes 3rd Edition: The Art of Scientific Computing*], Cambridge University Press, New York, NY, USA, 3 ed. (2007).

[19] Broyden, C. G., "Quasi-newton methods and their application to function minimisation," *Math. Comput.* **21**(99), 368–381 (1967).

[20] Byrd, R., Lu, P., Nocedal, J., and Zhu, C., "A limited memory algorithm for bound constrained optimization," *SIAM J. Sci. Comput.* **16**(5), 1190–1208 (1995).

[21] Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J., "Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization," *ACM Trans. Math. Softw.* **23**, 550–560 (Dec. 1997).

[22] Jurling, A. S. and Fienup, J. R., "Applications of algorithmic differentiation to phase retrieval algorithms," *J. Opt. Soc. Am. A* **31**, 1348–1359 (Jul 2014).

[23] Born, M. and Wolf, E., [*Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light (7th Edition)*], Cambridge University Press, 7th ed. (1999).