Demonstration of a symmetric dark hole with a stroke-minimizing correction algorithm

Jason D. Kay[a], Laurent A. Pueyo[b], and N. Jeremy Kasdin[a]

[a]Princeton University, Dept. of Mechanical and Aerospace Engineering, Princeton, NJ, USA;

[b]Jet Propulsion Laboratory, Pasadena, CA, USA

ABSTRACT

The past decade has seen a significant growth in research targeted at space based observatories for imaging exosolar planets. The challenge is in designing an imaging system for high-contrast. Even with a perfect coronagraph that modifies the point spread function to achieve high-contrast, wavefront sensing and control is needed to correct the errors in the optics and generate a "dark hole". The high-contrast imaging laboratory at Princeton University is equipped with two Boston Micromachines Kilo-DMs. We review here an algorithm designed to achieve high-contrast on both sides of the image plane while minimizing the stroke necessary from each deformable mirror (DM). This algorithm uses the first DM to correct for amplitude aberrations and the second DM to create a flat wavefront in the pupil plane. We then show the first results obtained at Princeton with this correction algorithm, and we demonstrate a symmetric dark hole in monochromatic light.

Keywords: Adaptive optics, high-contrast imaging, wavefront correction

1. INTRODUCTION

The search for extrasolar planets has led to a great increase in coronagraphic research for high-contrast imaging. While indirect methods such as stellar wobble, orbital transit, and gravitational lensing are capable of detecting increasingly smaller planets farther from their parent stars, such methods have not been proven sensitive enough for the detection of Earth-like planets. It is believed that only direct imaging will detect such objects. Additionally, directly detecting such planets is necessary for spectroscopic studies. Direct detection of an Earth-like planet in visible light requires a contrast level of 10^{-10} between planet light and the star light in the image plane. More than a dozen types of coronagraphs have been developed to reach this necessary contrast level, and many have been tested in laboratories.[1–3] In the high-contrast imaging laboratory at Princeton University, we have pioneered the shaped pupil coronagraph (SPC). This coronagraph uses a binary mask to change the entrance pupil of the system to one that achieves a high-contrast point spread function (PSF) at the image plane.[4] Ref 5 shows the best broadband results achieved with the shaped pupil coronagraph, at the Jet Propulsion Laboratory (JPL), using a band-limited Lyot coronagraph, a contrast of 6.4×10^{-10} has been obtained in 10% broadband light at an angle of 4 λ/D.

While the SPC solves the diffraction in the image plane by changing the system's PSF, there still exists the far more complicated problem of wavefront aberrations. Errors in the surface shapes of the system optics and non-uniform reflectivity across the mirror surfaces lead to phase and amplitude aberrations in the wavefront. These aberrations cause a degradation of contrast in the image plane. The presence of these aberrations prevents the ability to detect a planet because the light thrown about by the aberrations leaks into the dark regions of the image plane. For planet detection to be successful, a space-based mission must be equipped with a high-quality wavefront correction system. Figure 1 shows an example of a perfect PSF for the SPC and the image plane with the presence of system aberrations.

To overcome the contrast degradation caused by phase and amplitude aberrations, a telescope must be equipped with a high-quality adaptive optics system. Such a system will most likely include at least two DMs

Further author information:
Send correspondence to Jason Kay
E-mail: jkay@princeton.edu

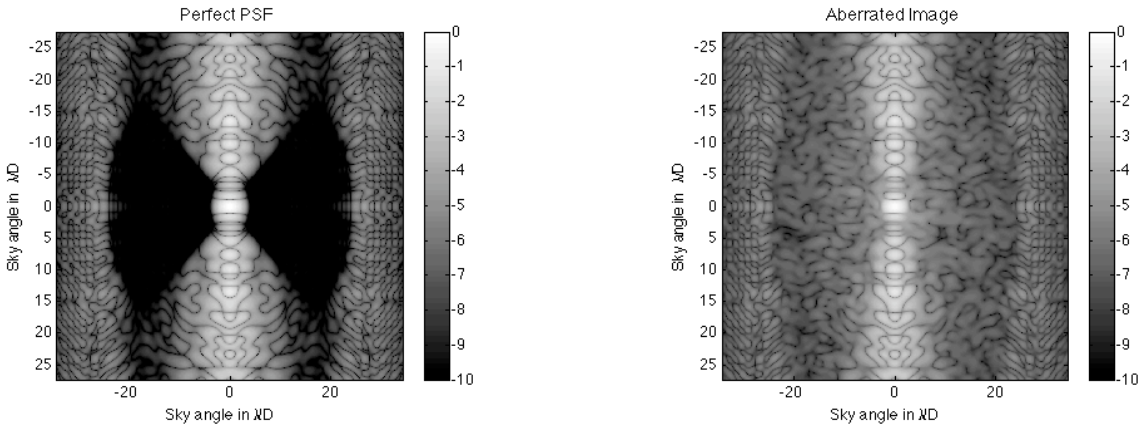

Figure 1. (left) The ideal PSF from a system using a shaped-pupil coronagraph. (right) The PSF from the same system with simulated aberrations. The phase aberration was generated to have an amplitude of $\lambda/100$ and a $-3/2$ power law, and the amplitude aberration has an amplitude of $1/20$ with a -2 power law. The figures are in a log scale with the log of contrast is shown in the colorbars.

operating in series as two DMs are required for correction of both phase and amplitude aberrations on all sides of the image plane. The most mature wavefront correction algorithms employed at both the High Contrast Imaging Testbed (HCIT) at JPL and on the testbed at Princeton are estimate-based algorithms. These correction methods require an estimate of the field in the region of interest in the image plane, the dark hole, to calculate the commands to send to the DMs. In § 2, we briefly review the electric field estimation algorithm used on the Princeton testbed for the experiments presented throughout this paper.

In § 3.1, we review the energy minimization correction algorithm presented in Ref. 6. This algorithm is designed to minimize the intensity of light in the dark hole region of the image plane. In § 3.2, we present the stroke minimization correction algorithm designed to minimize the stroke required on the DMs while achieving a desired contrast in the dark hole. We show experimental results, including a two-sided symmetric dark hole, in § 4.

2. WAVEFRONT ESTIMATION

Knowledge of the electric field is necessary in most wavefront correction algorithms. The difficulty is that cameras measure only the magnitude squared of the field, and as a result, all phase information is lost. To get an estimate of the field in the dark hole, we take multiple images in the image plane with diversity on the surface of the DM. Electric field estimation by DM diversity was originally proposed in Ref. 6 via three different images. This form of DM diversity is extremely sensitive to a priori knowledge of DM behavior. The specific variation of this algorithm that we use is called *pairwise estimation* as it uses a number of surfaces on the DM and their negatives.[7]

The pairwise estimation algorithm uses multiple DM settings and the corresponding intensity patterns in the image plane to reconstruct the complex electric field.[7] A pair of settings, ψ and $-\psi$, is put on the surface of the DM, and an image is taken for each setting. Given at least two pairs of images, the intensity patterns generated contain enough information to determine the full complex electric field.

Let $A(x, y)$ be the amplitude mask in the pupil plane of a coronagraph, and let the aberrations in the field be written as $1 + g(x, y)$. With $\psi(x, y)$ being the phase on the surface of the DM in radians, we can write the electric field in the pupil plane,

$$E_p = A(1 + g)e^{i\psi} \approx A + Ag + iA\psi, \tag{1}$$

where we have dropped the dependence on x and y, assumed that the surface on the DM is small, and ignored the cross term between the aberration and the DM surface. In addition, for purposes of this paper, we will assume that the DM and the pupil are in the same plane. With C being the linear operator that takes the electric field

in the pupil/DM plane to the plane of interest (often the combination of a series of Fourier Transforms and mask multiplications), the field in the image plane can be written as

$$E_{im} = \mathcal{C}\{A\} + \mathcal{C}\{Ag\} + i\mathcal{C}\{A\psi\} \approx \mathcal{C}\{Ag\} + i\mathcal{C}\{A\psi\}, \tag{2}$$

where we have assumed that the PSF of the system is negligibly small in the dark hole.

Pairs of images are taken with the DM having surfaces of ψ_j and $-\psi_j$. Let I_j^+ be the intensity measurement when the surface of the DM is ψ_j, and I_j^- be the intensity measurement when the surface is $-\psi_j$. The complex electric field in the plane of interest can be determined by taking two or more pairs of images. The only term in Eq. 2 that is not known a priori is $\mathcal{C}\{Ag\}$. This term can be found by taking the differences between the j pairs of images.

$$\frac{1}{4}\begin{bmatrix} I_1^+ - I_1^- \\ \vdots \\ I_j^+ - I_j^- \end{bmatrix} = \begin{bmatrix} \Re\{\mathcal{C}\{iA\psi_1\}\} & \Im\{\mathcal{C}\{iA\psi_1\}\} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ \Re\{\mathcal{C}\{iA\psi_j\}\} & \Im\{\mathcal{C}\{iA\psi_j\}\} \end{bmatrix} \begin{bmatrix} \Re\{\mathcal{C}\{Ag\}\} \\ \Im\{\mathcal{C}\{Ag\}\} \end{bmatrix} \tag{3}$$

The vector on the left of the equality in Eq. 3 is found by direct measurement of the intensities. The matrix to the right of the equality can be pre-computed with knowledge of the DM shapes. By inverting this matrix with a least-squares fit and multiplying the pseudo-inverse by the vector of intensities, the field due to the aberrations can be determined on a pixel-by-pixel basis, and based on this estimate, the proper DM settings can be applied based on the correction algorithm. This correction is the subject of the next section.

3. WAVEFRONT CORRECTION

The estimate of the complex electric field obtained with the method described in § 2 is inserted into the wavefront correction algorithm in order to create a dark hole in the image plane in which a planet can be found. We write the surface of the DM as a sum of N weighted basis functions,

$$\psi_{DM}(x,y) = \frac{2\pi}{\lambda}\sum_{k=1}^{N} a_k f_k(x,y), \tag{4}$$

and we use a correction scheme that solves for the optimal a_k's to give a dark hole where desired. λ is the wavelength.

The electric field in the dark hole of the image plane is written as in Eq. 2. The intensity is the magnitude squared of this field, so the total intensity in the dark hole can be written as

$$I_{DZ} = \int_{DH} \left| \mathcal{C}\{A\}(u,v) + \mathcal{C}\{Ag\}(u,v) + i\frac{2\pi}{\lambda}\sum_{k=1}^{N} a_k \mathcal{C}\{Af_k\}(u,v) \right|^2 dudv, \tag{5}$$

where the integral is taken over the dark hole.

We write a row vector of the function strengths, $X = [a_1, a_2, \ldots, a_{N-1}, a_N]$. The total intensity of Eq. 5 can be written as a quadratic form in the vector X.

$$I_{DH} = \left(\frac{2\pi}{\lambda}\right)^2 XMX^{\mathrm{T}} + \frac{4\pi}{\lambda}X\Im\left(b^{\mathrm{T}}\right) + d \tag{6}$$

We can define the above matrices as follows, where $\langle\cdot\rangle$ is the normal L_2 inner product.

$$\begin{aligned} M_{p,q} &= \left\langle \mathcal{C}\left\{Af_{k(p)}\right\}, \mathcal{C}\left\{Af_{k(q)}\right\} \right\rangle \\ b_p &= \left\langle \mathcal{C}\left\{Af_{k(p)}\right\}, \mathcal{C}\{A\} + \mathcal{C}\{Ag\} \right\rangle \\ d &= \langle \mathcal{C}\{Ag\}, \mathcal{C}\{Ag\} \rangle \end{aligned}$$

In words, M is a matrix that describes the effect of the DM on the intensity in the dark hole, b is a vector that shows the image plane effect of the DM with the aberrated field (or estimated field), and d is a scalar that represents the intensity due to the uncorrected, aberrated field. In the absence of a DM, the total intensity in the dark hole would be d. This correction scheme requires only a measurement of the aberrated field, $\mathcal{C}\{Ag\}$, to determine a DM setting, and this is obtained with the estimation algorithm as in § 2. Note that in deriving Eq. 6, we have assumed that $\langle \hat{A}, \hat{A} \rangle$ is insignificant as by design, the coronagraph, in this case the shaped pupil, gives an intensity in the dark zone below the necessary contrast. In addition, we have ignored the term $\langle \hat{A}, \widehat{Ag} \rangle$ and assumed that M is a symmetric matrix. Equation 6 is general for multiple sequential DMs. With one DM, M is an $N \times N$ matrix and b is $N \times 1$. With two sequential DMs, M is $2N \times 2N$ and b is $2N \times 1$.

3.1 Energy Minimization

The energy minimization algorithm is introduced in Ref. 6. The optimal vector X to minimize the energy in a specified region of the image plane can be found by taking the partial derivative of I_{DH} in Eq. 6 with respect to X.

$$X_{opt} M = -\frac{\lambda}{2\pi} \Re(b) \tag{7}$$

This algorithm suffers two major drakbacks. First of all, in most coronagraphs, including the Shaped Pupil Coronagraph, M is not invertible, so solving for X_{opt} requires the use of singular value decomposition to find a pseudo-inverse. Such a process utilizes a threshold identifying which singular values are kept and which are dropped. In this algorithm, the process of choosing such a threshold is ad hoc and is not done with any real understanding. Furthermore, the use of this algorithm does not guarantee the desired contrast. The stroke minimization presented in the next section solves both of these issues.

3.2 Stroke Minimization

In stroke minimization, we eliminate the invertibility problem by minimizing with respect to the finite number of coefficients in the DM surface expansion rather than the field itself. This has the effect of minimizing the average stroke of the DM actuators, an important criteria for the small devices being used. We include in the minimization the constraint that the field meet the contrast requirement in the dark hole.

The optimization problem is thus written,

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2} \sum_{k=1}^{N} a_k^2 \\ \text{subject to} \quad & I_{DH} \leq C \end{aligned}$$

where C is the contrast desired. The cost function to minimize is therefore,

$$J = X \left(I + \mu \left(\frac{2\pi}{\lambda} \right)^2 M \right) X^{\mathrm{T}} + 2\mu \frac{2\pi}{\lambda} X \Re(b^{\mathrm{T}}) + \mu C, \tag{8}$$

where μ is a Lagrange multiplier to ensure the constraint in contrast is satisfied. As above, we find the optimal value of X by taking the partial derivative of Eq. 8 in X. We find that

$$X_{opt} = -\left(\frac{\lambda}{2\pi} I + \mu \frac{2\pi}{\lambda} M \right)^{-1} \Re(b), \tag{9}$$

where this matrix inverse is found directly and no decomposition is needed. The optimal value for the Lagrange multiplier, μ, is determined by finding the lowest value of μ for which the contrast remains at the targeted level.

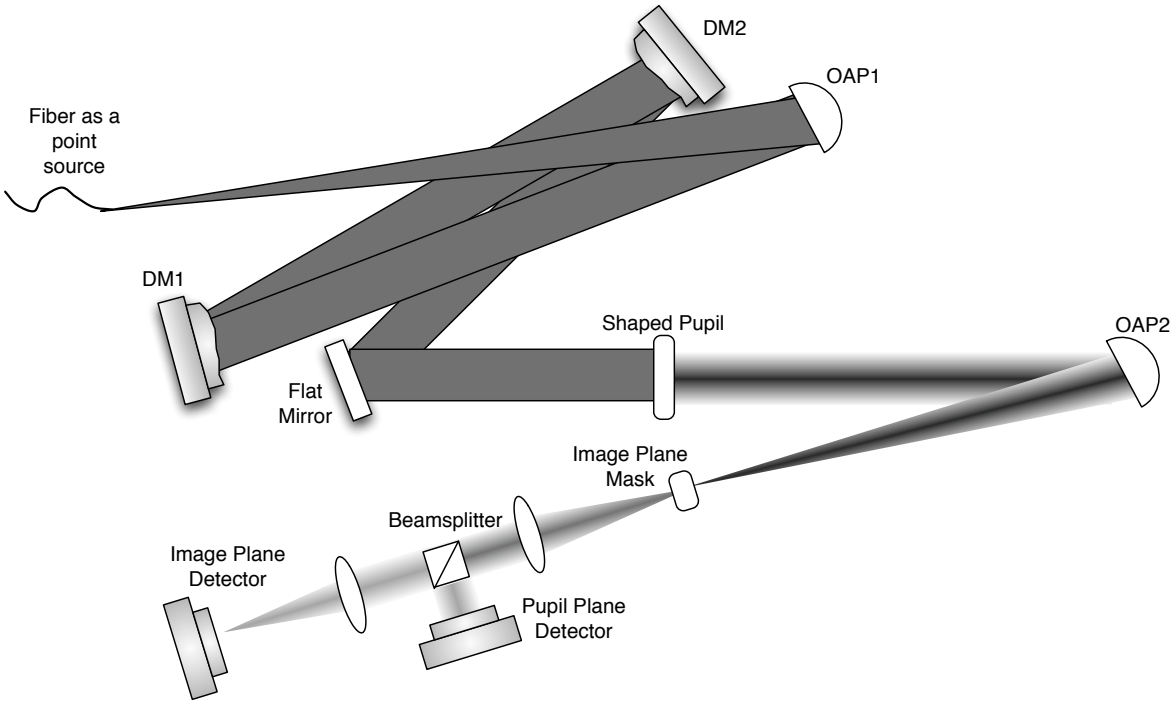

Figure 2. Optical layout of the princeton high contrast imaging testbed. For the experiment presented in § 4.1, only one DM is used for wavefront control and only the image plane camera is used for wavefront sensing.

4. LABORATORY RESULTS

The optical layout of the Princeton High Contrast Imaging Laboratory is shown in Fig. 2. The DM(s) size is 1 cm, and neither of them is located in a plane conjugate to the pupil, however, this does not impact the correction algorithms as this propagation can be taken into account in our knowledge of the $\mathcal{C}$ operator. For the experiments presented in this paper, only the image plane camera was used. All results were obtained in monochromatic 632 nm laser light, and the wavefront is flattened so that the image plane exhibits a dark hole from 7-10 λ/D in x and -2.5-2.5 λ/D in y.

The basis functions used for the DM model in these experiments (Eq. 4) are the influence functions. These are the shapes that the DM surface takes when voltage is sent to one actuator. The influence function is assumed to be the same at each actuator, and a Gaussian shape is fit to give the same intensity in the pupil plane as is seen on the optical table. While not the optimal basis, the use of influence functions leads to the simplest mapping between voltage and DM surface shape. The following two sections present the experimental results for both one and two sided dark hole creation using one and two DMs respectively.

4.1 Dark Hole Creation

To obtain these results, we use one DM to create a dark hole on one side of the image plane. With an influence function basis, there are 1,024 basis functions in our one DM correction system. Looking at the cost function, Eq. 8, M is 1024 $\times$ 1024, and b is 1024 $\times$ 1. The PSF obtained using the Stroke Minimization algorithm is shown in Fig. 3 and exhibits a dark hole of 2 orders of magnitude deeper than the non-corrected one. In 100 iterations, the contrast in the dark hole increases from 2×10^{-4} to 3×10^{-6}.

4.2 Symmetric Dark Hole

In this section, we present the first results of a symmetric dark hole using two DMs in sequence to create a dark hole on both sides of the image plane. With two DMs, M is now 2048 $\times$ 2048, and b is 2048 $\times$ 1. As in the case

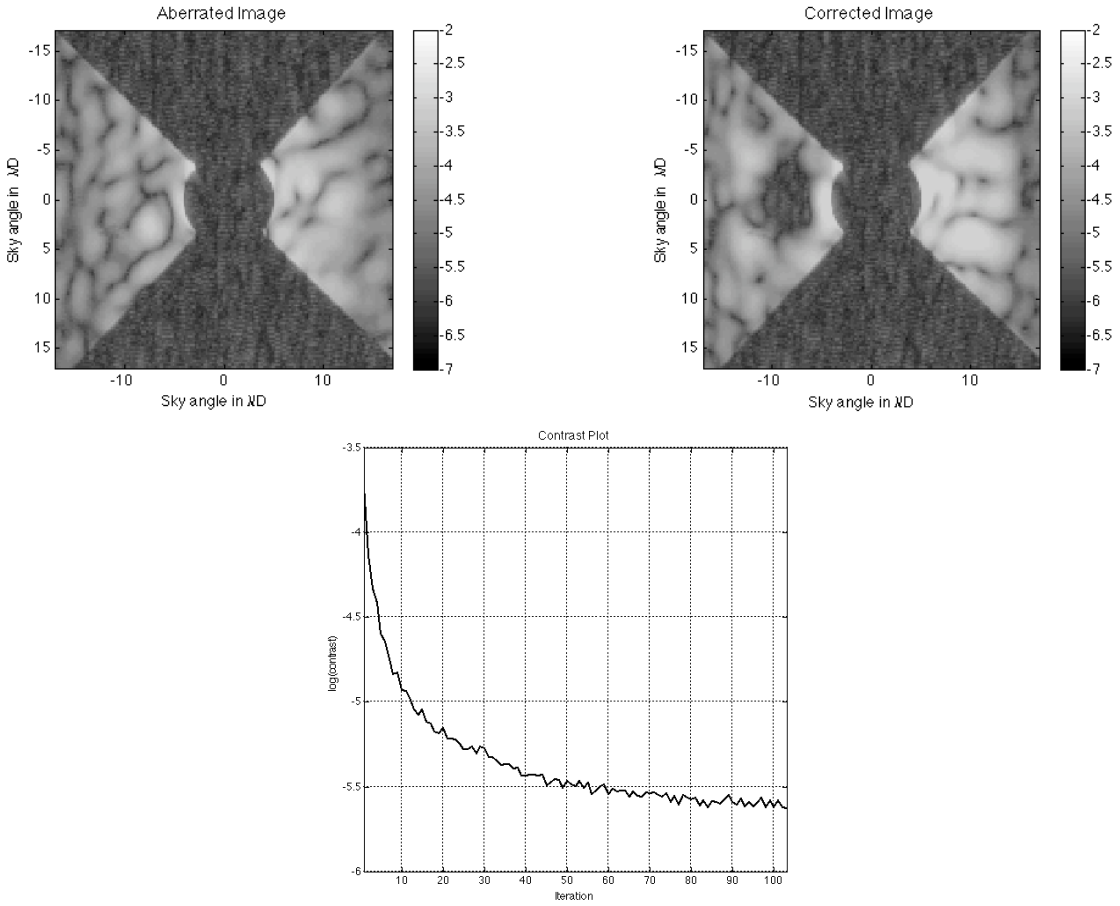

Figure 3. Aberrated (Left) and corrected (Right) PSF on the Princeton testbed in log(contrast). The wavefront is flattened so that half of the image plane exhibits a dark hole over a specified region, in this case X = 7-10 λ/D and Y = -2.5-2.5 λ/D.

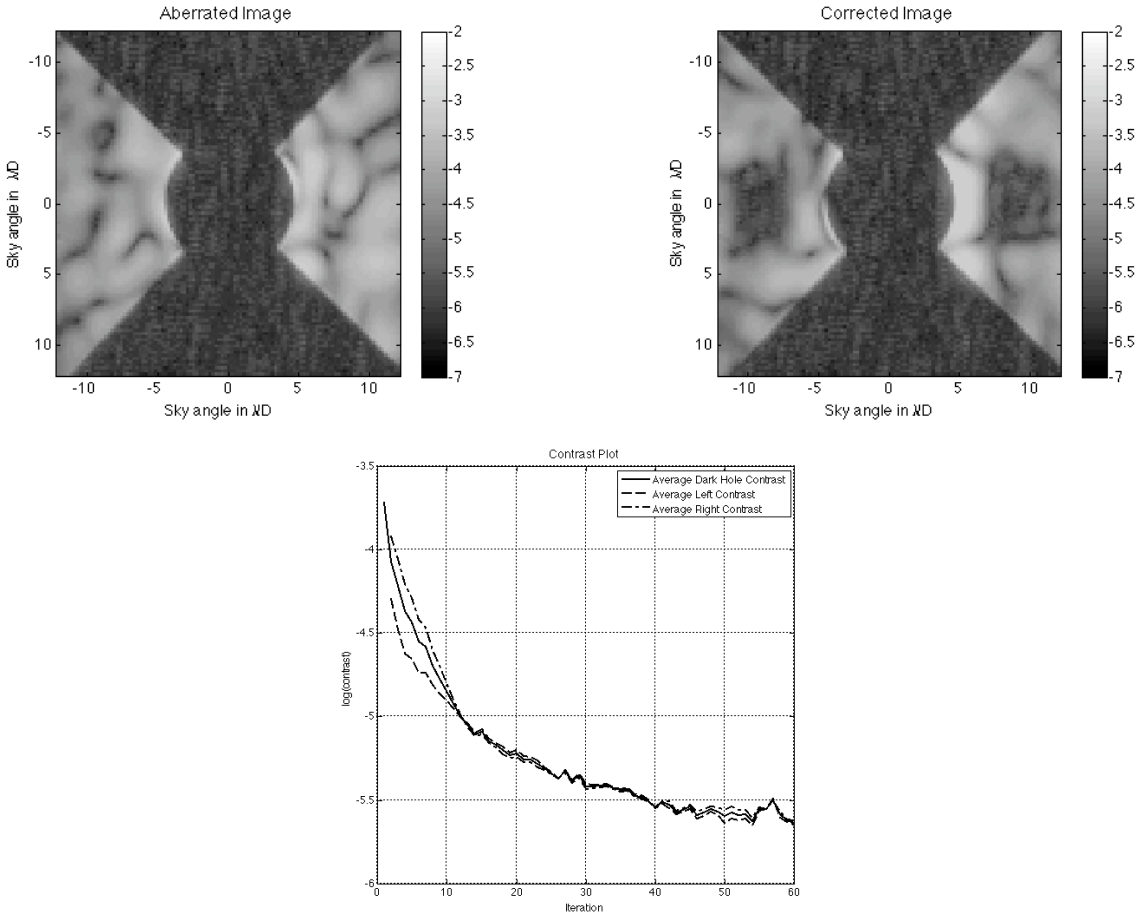

Figure 4. Aberrated image (left) and corrected image(right) of the 2 DM stroke minimization symmetric dark hole experiment. Bottom: A plot of contrast vs. iteration in each of the two dark holes and in the combination of the two.

of one DM, the corrected image displays a contrast of two orders of magnitude better than the initial, aberrated image. In 60 iterations, the contrast increases from 2.5×10^{-4} to 3×10^{-6}.

Figure 4 shows the aberrated image prior to correction as well as the image after 60 iterations of the stroke minimization correction algorithm. In addition, the figure shows a contrast plot as a function of iteration. We show a correction of just under two orders of magnitude on both sides of the image plane.

5. CONCLUSIONS AND FUTURE WORK

We have presented, here, a new algorithm for wavefront correction, the Stroke Minimization algorithm. Given an estimate of the electric field in a region of the image plane and a target contrast, the stroke minimization algorithm determines the shape to apply to the surface of one or multiple DMs that achieves this contrast in the image plane while minimizing the required stroke on the DMs. In creating one dark hole using one DM and in creating symmetric dark holes using two sequential DMs, we have demonstrated an improvement in contrast of two orders of magnitude.

Currently, the limiting factor in these experiments is in the wavefront estimation and the model of the DMs. As discussed in § 2, proper estimation of the aberrated wavefront requires proper knowledge of the DM surface. This requires good understanding of the mapping between voltages to DM surface shape, and such a mapping is not very well understood. Non-linearities in the voltage to height map, actuator coupling, and more accurate basis functions must be taken into account for improved results, and we are working on adaptive schemes to determine these effects.

REFERENCES

1. M. Kuchner, J. Crepp, and J. Ge, "Eighth-order image masks for terrestrial planet finding," *The Astrophysical Journal* **628**, pp. 466–473, 2005. (astro-ph/0411077).
2. O. Guyon, E. Pluzhnik, R. Galicher, R. Martinache, S. Ridgway, and R. Woodruff, "Exoplanets imaging with a phase-induced amplitude apodization coronagraph–i. principle," *Astrophysical Journal* **622**, p. 744, 2005.
3. B. Levine, M. Shao, D. Liu, J. Wallace, and B. Lane, "Planet detection in visible light with a single aperture telescope and nulling coronagraph," in *Proceedings of SPIE Conference on Astronomical Telescopes and Instrumentation*, *5170*, pp. 200–208, 2003.
4. N. Kasdin, R. Vanderbei, M. Littman, and D. Spergel, "Optimal one-dimensional apodizations and shaped pupils for planet finding coronagraphy," *Applied Optics* **44**, pp. 1117–1128, 2005.
5. R. Belikov, A. Give'on, B. Kern, E. Cady, M. Carr, S. Shaklan, K. Balasubramanian, V. White, P. Echternach, M. Dickie, J. Trauger, A. Kuhnert, and N. J. Kasdin, "Demonstration of high contrast in 10% broadband light with the shaped pupil coronagraph," in *Techniques and Instrumentation for Detection of Exoplanets III. Edited by Coulter, Daniel R. Proceedings of the SPIE, Volume 6693, pp. 66930Y-66930Y-11 (2007).*, *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference* **6693**, Sept. 2007.
6. P. Borde and W. Traub, "High-contrast imaging from space: Speckle nulling in a low aberration regime.," *Applied Physics Journal* **638**, February 2006.
7. A. Give'on, B. Kern, S. Shaklan, D. C. Moody, and L. Pueyo, "Electric Field Conjugation - A Broadband Wavefront Correction Algorithm For High-contrast Imaging Systems," in *Bulletin of the American Astronomical Society*, *Bulletin of the American Astronomical Society* **38**, pp. 975–+, Dec. 2007.