Residual Wave Front Phase Estimation in the Reimaged Lyot Plane for the Eclipse Coronagraphic Telescope

Stuart Shaklan*, Dwight Moody, Joseph Green
Jet Propulsion Laboratory, California Institute of Technology

ABSTRACT

Coronagraphs for extra-solar planet detection remove diffracted stellar light through the combination of a coronagraphic mask and a Lyot stop. When the entrance pupil contains a nearly perfect wave front, most of the stellar light is absorbed at the mask. Light scattered around the spot due to mid- and high-spatial frequency phase errors in the pupil appears at the Lyot plane as speckles whose amplitudes are proportional to the local wave front phase residuals. The speckles scale with optical wavelength but are not radially smeared. The Eclipse deformable mirror (DM) can be used to modify the Lyot amplitude distribution, providing a simple means of estimating the residual phase content and controlling the wave front. To reduce the detrimental noise carried by uncontrollable high-spatial frequency wave front components, the Lyot plane signal is filtered at the science plane to pass only the controllable spatial frequencies that contribute to the dark hole. The Lyot stop is then reimaged onto a detector. We demonstrate through simulations that this approach significantly improves the signal-to-noise ratio of the planet measurement.

Keywords: Coronagraph, wave front sensing, wave front control

1. INTRODUCTION

Advanced coronagraphs on space-based telescopes are a promising technology for the direct detection of extra-solar planets. The Eclipse mission is a proposed 1.8 m visible/near-IR telescope whose goal is to detect Jupiter-mass planets around a sample of hundreds of nearby stars [1]. The Eclipse coronagraph is designed to remove diffraction to levels < $10^{-9}$ of the light in the Airy core. The coronagraph incorporates a high-density deformable mirror capable of correcting the scattered wave front to sub-Angstrom precision and creating a so-called “dark hole” [2] extending to 40-50 Airy rings.

The ability to sense wave fronts at the Angstrom level using starlight is a critical technology for Eclipse and other coronagraphs. Wave fronts can be sensed in several ways. First, a ‘front-end’ sensor – one that collects light in front of the coronagraph mask – can be used. A proposed method of front-end sensing is the application of a Gerchberg-Saxton (GS) algorithm to defocused images and a pupil image, as discussed in Redding et al [3] and Green et al [4]. This approach uses all the available starlight but places challenging requirements on detector calibration (e.g. MTF calibration and flat-fielding). Chromatic smearing of speckles in the (defocused) image plane may limit the useful optical bandwidth of the algorithm. We speculate that the GS algorithm can be applied at the ‘back end’ images of the Lyot stop and the final focal plane. Again, chromatic smearing is a limiting factor but it may be less important as the dynamic range for wave front sensing is orders of magnitude smaller at the back-end where virtually all of the starlight has been removed. Traub has proposed a technique that uses the deformable mirror (DM) to iteratively remove speckles from the center of the image out toward the edges [5]. It is similar in principal to the Clean-like algorithm proposed by Burrows [6].

In this paper we propose a back-end wave front sensing approach that does not make use of image plane data. Instead, we measure the wave front through two images of the Lyot plane; the first image uses the nominal wave front, while the second calls on the DM actuators to move by a known amount. Being pupil images, there is no radial chromatic smearing. Thus the technique works in broadband light providing excellent sensitivity.

* Contact information: Address: 4800 Oak Grove Dr., MS 301-486, Pasadena, CA 91109. e-mail stuart.shaklan@jpl.nasa.gov
The following section derives the simple theory and algorithm behind this approach. We also calculate the shot-noise limited signal-to-noise ratio. In Section 3 we show several examples and introduce the science-plane spatial filter that removes uncontrolable high-spatial frequency noise from the wave front measurement.

2. REIMAGED LYOT PLANE WAVE FRONT SENSING

When a small, opaque mask is placed on-axis at the image plane of a telescope, the image of an unresolved on-axis source is largely obscured. However, mid- and high-spatial frequency light is scattered off-axis, avoiding the mask and continuing on through the system. Several authors have proposed different masks, including general tapering [7], band-limited [8], central phase-shifting [9] and 4-quadrant phase shifting [10]. The various masks can be optimized to improve throughput, maximize search space, decrease sensitivity to pointing errors, allow close-in searches near the central lobe, etc. The wave front sensing technique described below can be used with any of the masks.

Assume that the wave front phase deviations \( \phi(x) \) are small compared to a radian, \( \phi(x) \ll 1 \), and likewise that amplitude variations \( a(x) \) across the pupil are small compared to the average amplitude, \( a(x) \ll 1 \). Then given the pupil function \( P(x) \) we can express the wave front at the pupil by

\[
E = P(x)[1 + a(x)]e^{i\phi(x)} = P(x)[1 + a(x)][1 + i\phi(x)]
\]

Expressing Fourier-Transformed quantities using the tilde symbol, e.g. \( \tilde{E} = FT(E) \), and convolution using the \( \otimes \) symbol, the field at the image plane is given by

\[
\tilde{E} = \tilde{P}(r) \otimes [\delta(r) + \tilde{a}(r) + i\tilde{\phi}(r) + i\tilde{a}(r) \otimes \tilde{\phi}(r)] .
\]

We will now ignore the second order cross-term \( \tilde{a} \otimes \tilde{\phi} \) as it is much smaller than the first order terms under the stated conditions.

The image plane mask is a function \( M(r) \) that multiplies the image plane field to yield the transmitted field

\[
\tilde{E}_T = M(r)[\tilde{P}(r) \otimes [\delta(r) + \tilde{a}(r) + i\tilde{\phi}(r) + i\tilde{a}(r) \otimes \tilde{\phi}(r)]] .
\]

Mask \( M(r) \) mainly absorbs the delta-function in eq. 3 while passing off-axis \( (r \neq 0) \) light. For purposes of simplicity, we will ignore the Airy function \( \tilde{P}(r) \) in eq. (2) and consider an ideal coronagraph mask that eliminates only the delta-function, so that the light passing the mask is given by

\[
\tilde{E} = \tilde{a}(r) + i\tilde{\phi}(r) .
\]

The approximation ignores blurring of the speckles in the image plane due to \( P \) as well as blurring of speckles in the reimaged pupil plane due to \( M \). This leads to small errors in the simplistic wave front control algorithm (Sect. 4). This can be addressed with a more rigorous mathematical approach. It is also correctable by iterating the solution, as the approximation errors are a small fraction of the recovered wave front.

Reimaging of the pupil (normally to the plane where the Lyot stop is applied) then yields the field amplitude

\[
E = a(x) + i\phi(x)
\]

and intensity
\[ I = |E|^2 = a(x)^2 + \phi(x)^2 \]  

This is the observed intensity at the reimaged Lyot plane (RLP).

Let's now change the wave front phase by a known amount \( \Delta(x) \) to form a new wave front \( \phi_2(x) = \phi(x) + \Delta(x) \). The observed RLP intensity is then

\[ I_2 = a(x)^2 + \phi_2(x)^2 = a(x)^2 + \phi(x)^2 + \Delta(x)^2 + 2\phi(x)\Delta(x) \]  

(7)

We can now estimate the wave front phase \( \phi(x) \) by forming

\[ \hat{\phi} = (I_2 - I - \Delta(x)^2)/2\Delta(x) \]  

(8)

Eq. 8 is the fundamental equation for RLP wave front sensing. It includes two measured quantities, RLP images \( I_2 \) and \( I \), and a known phase change \( \Delta(x) \). In practice, \( \Delta \) is the change in phase at the RLP for a given phase change at the DM. This is calculated using an accurate model of the system that propagates the DM influence function to the RLP. If knowledge of \( \Delta \) is in error, a situation caused by DM calibration errors or DM drift, then \( \hat{\phi} \) will also be in error, but the process should converge as \( \hat{\phi} \) improves. Given accurate DM calibration and adequate signal-to-noise, this approach requires just two images \( I \) and \( I_2 \).

We note that eq. 8 senses only the imaginary part of the wave front, that is the part related to pupil phase \( \phi \) and not amplitude errors \( a \). Probably the best means of measuring \( a \) is to make a direct image of the pupil at the front end. A signal-to-noise of \( \sim 500 \) is required to estimate \( a \) to 0.001, which is equivalent to estimating phase to 0.001 radians according to eq. 5. The broadband nature of the overall complex field solution will be impacted because amplitude errors propagate with different wavelength dependence than optical path errors. Given front-end measurements of \( a \) and RLP measurements of \( \phi \), we have a wave front sensing solution capable of describing the complex field in the pupil.

An important advantage of our purely pupil-plane approach is that the RLP signal is unaffected by the presence of a planet, as long as the integrated planet signal is small compared to the integrated light in the dark hole. This is in contrast to image plane wave front sensing, where the planet signal, if it is near the inner portion of the dark hole, is indistinguishable from a speckle.

### 3. SIGNAL-TO-NOISE IN THE SHOT-NOISE LIMIT

Consider the signal at the RLP when the coronagraph mask is removed. All of the light reaches the RLP and the intensity is \( I = 1 \) when amplitude errors are ignored. As amplitude errors fluctuate much more slowly in time than phase errors, they can be considered static during wave front sensing so they don’t impact the signal-to-noise ratio (SNR).

We define \( n \) as the mean number of photons per coherence cell (or per DM actuator) detected without the mask. When the mask is in place, the intensity at the RLP is \( I = \phi^2 \) and the mean number of photons detected per coherence cell is

\[ \langle n_1 \rangle = \bar{n}\phi^2 \]
\[ \langle n_2 \rangle = \bar{n}(\phi + \Delta)^2 \]  

(9)

for the nominal wave front and modified wave front, respectively. As the photon counts are sampled from a Poisson distribution, their second moments are given by
\[
\langle n_1^2 \rangle = \bar{n}^2 \phi^2 + \bar{n} \phi^2
\]
\[
\langle n_2^2 \rangle = \bar{n}^2 (\phi + \Delta)^2 + \bar{n} (\phi + \Delta)^2
\]

We now redefine the wave front estimation equation in terms of detected and mean photon counts:
\[
\hat{\phi} = (n_2 - n_1 - \bar{n} \Delta^2)/2\bar{n} \Delta,
\]
and note that \( \langle \hat{\phi} \rangle = \phi \) when eq. 9 is substituted into eq. 11.

Using eqs. 9 -11, the variance of the phase estimator is determined to be
\[
\sigma_{\hat{\phi}}^2 = \langle \hat{\phi}^2 \rangle - \langle \hat{\phi} \rangle^2 = (2\phi + \Delta^2 + 2\phi \Delta)/4\bar{n} \Delta^2.
\]

The SNR is then given by
\[
\text{SNR}_{\phi} = \frac{1}{\sigma_{\hat{\phi}}^2} = 2\sqrt{\bar{n}} |\phi| |\Delta| / \sqrt{2\phi^2 + \Delta^2 + 2\phi \Delta}.
\]

Eq. 13 shows that for a given phase value \( \phi \), the SNR is maximized at large \( \Delta \), \( \lim_{\Delta \to \infty} \text{SNR} = 2\sqrt{\bar{n}} |\phi| \).

Averaging over the pupil, and assuming that \( \phi \) has a Gaussian probability distribution with variance \( \sigma_{\phi}^2 \), it can be shown that the average signal-to-noise ratio is \( \langle \text{SNR} \rangle = 1.6\sqrt{\bar{n}} \sigma_{\phi} \) for \( \Delta \gg \sigma_{\phi} \). There is, however, a practical limitation to the accuracy of \( \hat{\phi} \) at large \( \Delta \). As noted earlier, DM drift and calibration errors, and errors in the optical model of the system lead to errors in the 'known' value \( \Delta \). These are exacerbated as \( \Delta \) grows to values >> the nominal phase error \( \sigma_{\phi} \). In the following section, we have chosen somewhat arbitrarily \( |\Delta| = 2\sigma_{\phi} \), leading to a reduction of the SNR to \( \text{SNR}_{\phi} \sim 2\sqrt{\bar{n}} \sigma_{\phi} \), or \( \sigma_{\phi}^2 = 1/\bar{n} \). The optimal choice of \( \Delta \) will ultimately depend on the systematic noise floor.

4. SIMULATIONS

We have simulated the RLP wave front sensing process for several cases, namely noise-free sensing, sensing in the presence of uncontrollable spatial frequencies \( f \geq f_{\text{Nyquist}} \) (the DM cutoff frequency, sensing in broad-band light, and the shot-noise-limited case.

At the core of our simulation engine is a physical optics propagation code developed by D. Moody at JPL. The code handles high-precision wave front propagation both in the Fraunhofer regime and in near-field cases. It also generates DM wave fronts by convolving the commanded actuator pistons with a measured influence function. It is also capable of generating shaped wave front noise on the DM to simulate uncontrolled (\( \phi \)-Nyquist) spatial frequencies. While the code is set up to simulate plane-to-plane propagation for the Eclipse Testbed, it is a general-purpose tool and for simplicity we have chosen to utilize only the Fraunhofer propagations for the present work.

The model parameters for our coronagraph are given in Table I. The Gaussian mask and Lyot stop are paired to reduce diffraction to levels below \( 10^{-6} \) of the Airy peak. The Gaussian mask FWHM of 425 microns reaches half-power at the fourth Airy ring (after accounting for the Lyot diameter of 0.75 pupil). The edge of the Lyot stop is smoothed with a half-width of 3\% of the pupil diameter. This helps reduce ringing in the image plane and decreases the bleeding of high-spatial frequency light into the central darkhole. The DM is a large format version of the \( 42 \times 42 \) device currently under test at JPL. The Xinetics DM has a 1 um piston capability with 1 mm pitch. Trauger et al [11] provides a detailed description of the DM.
Figure 1. Simulation process.

The simulation process is shown in Figure 1. First, we use a normal distribution to place a random wave front on the DM. We choose a surface r.m.s. of 2 Angstrom (for a wave front r.m.s. of 4 Å). We also select a wave front change \( \Delta \) with an amplitude of +/- 4 Å and a random sign. That is, each actuator has a phase change randomly selected from +/- 4A. Another possible choice is a checkerboard pattern, but we found that the regularity of the pattern introduced a large spike at \( f_{\text{Nyquist}} \) along the axes. We chose an amplitude of 4 Å because it is small enough that DM calibration errors should be negligible, it does not require large dynamic range on the detector, and it carries a moderate noise penalty compared to the limiting case \( \Delta \rightarrow \infty \) (eq. 13). While the constant 4 Å amplitude ensures some wave front sensing ability at all points on the pupil, the sensitivity is not uniform as congruent blocks of like-sign (e.g. large areas where all actuators have +4 Å) will be more or less filtered by the coronagraph. Figure 2 shows the initial, second, and difference wave fronts.

The back-focus image is formed after Fourier propagations from the pupil to the mask, and the
Figure 2. DM actuator settings. (A) is a 21 x 21 subset of the DM with a 2 A r.m.s. surface. (B) is the same set of actuators with +/- 4A added to each actuator. (C) is the random pattern of +/- 4 A. The figures depict the commanded actuator piston: the DM wave front is the convolution of this pattern with the actuator influence function (fig. 3).

mask to the Lyot plane. At the back-focus, the dark hole forms with a dimension determined roughly by the DM Nyquist frequency $f_{\text{Nyquist}}$. Around the dark hole, a large ‘wall’ of light exists for which there is no means of control in the system. The shaped noise (described in Example 2 below) determines the distribution and amplitude of this light, mostly concentrated near the edge of the hole. The light carries much energy and therefore much noise. We remove the light by placing a square filter around the dark hole, passing only the ‘good’ light that we wish to control. The filter is a few speckles narrower than the formal Nyquist limit because the large wall of light partially diffracts into the dark hole. This reduces resolution on the DM but it does not affect the ability to measure and correct the wave front within the filtered region.

A final Fourier Transform takes us to the reimaged Lyot plane (RLP) where we measure $I_1$ and $I_2$ and add shot noise. We also form an auxiliary image (in field, not intensity) using the difference wave front (starting with the DM amplitudes in fig. 2C) that becomes $\Delta$ in eq. 11.

Example 1: Noise-free imaging

In our first example, we assume that the DM is a perfect mirror whose surface is limited only by the actuator piston – there are no spatial frequencies present beyond the Nyquist frequency (0.5 cycles/mm). The actuators are driven to a random position resulting in a 2 A r.m.s. surface (4 A wave front). There are no other optical errors in the system. The influence function for an actuator is shown in fig. 3.
We follow the procedure outlined in figure 1 to estimate the wave front. Once the wave front phase is estimated, it is filtered to estimate the correct phase to add to the DM to improve the wave front. At the present time we use a simple algorithm to perform the filtering—it results in some performance degradation compared to more complex approaches, but it allows us to demonstrate sub-Angstrom wave front control nonetheless. The algorithm takes the estimate of $\phi(x)$, which is sampled at 250 microns, and bins it to a 1 mm scale through simple Fourier plane filtering. We divide by the single-actuator amplitude response (FT of the back-end influence function, shown in fig. 3) to obtain a more uniform response across the dark hole and to reduce high-frequency noise contributions. The result is a 96x96 estimate of the phase at each DM actuator. This simple process does not properly treat the convolution of the influence function with the
Figure 5. Wave front sensing in the presence of shaped high-frequency noise. (A) shows the 'wall' due to the 4 nm r.m.s. wave front at $> f_{ny}$. The light level is ~100x larger than the light level inside the dark hole. (B) shows the region inside the dark hole: only this light is passed to the RLP for sensing. Wave front sensing performance is comparable to the noise-free case (Fig. 4).

actuator amplitudes, as the detected signal is the square of the convolved field. This limits us to factors of ~6-7 in wave front improvement and reduces the useful size of the dark hole by ~20%.

The estimate of wave front phase is only valid over the portion of the pupil passed by the Lyot stop, 72 actuators wide. We form a new DM wave front by subtracting from the DM our estimated values for the 72-actuator region, and leave the other actuators unaffected. We propagate the wave front through the system and form a new dark hole image, as shown in Fig. 4. Both the image mean and standard deviation are reduced by a factor of 40 compared to the one obtained with a 4A wave front. The new image is consistent with a 0.6 A r.m.s. wave front.

**Example 2: Sensing with $f \geq f_{nyquist}$ wave front errors**

Shaped noise at Nyquist and above results from polishing and manufacturing limitations. The DM cannot reduce the amplitude of the resultant speckles outside the dark hole. However, bleeding of their tails into the dark hole can be corrected using a half-dark hole algorithm at the expense of increasing light levels in one-half of the dark hole. This will be demonstrated in a future paper.

We add wave front errors with a power spectral density that follows $f^{-3}$. There is no power at frequencies below Nyquist other than the DM actuator piston. The r.m.s. value of the high-frequency surface errors is chosen to be 2 nm - ten times larger than the 2 A DM errors, carrying 100 times more energy. Fig. 5a shows the 'wall' of light around the dark hole.

Our wave front estimation technique filters most of this light at the image plane. Only the light within the dark hole passes through the mask, allowing one to obtain wave front reconstructions with performance comparable to Example 1.

**Example 3: Broad-band light**

An important advantage of our approach is the ability to sense wave fronts over a broad optical band. Fig. 6 shows the RLP signal (proportional to $\phi^2$) for a single wavelength 850 nm and a broad band, spanning 700 – 800 nm. The two are very similar and allow us to perform wave front estimation (albeit with reduced performance) over 25% and larger bands. In these simulations, the shaped high-frequency light is on and as in example 2 it is filtered at the image plane by the square mask. The difference between the
monochromatic and composite cases is due to two effects; first, the back-end influence function changes as the coronagraph mask has a slightly different effect at each wavelength. Second, the image plane mask, which has a fixed physical size, is matched to the dark hole at the shortest wavelength and is therefore undersized for the longer wavelengths. This widens the influence function at longer wavelengths. Fig 7 shows the improved image at 850 nm after wave front sensing over the range 700-800 nm. The image plane speckle improvement, a factor of 6.7 in intensity (2.6 in wave front) is mostly limited by our simple algorithm which is used unmodified from its monochromatic implementation. We also performed the simulation over the 27% bandpass spanning 650-850 nm. The results were still good: a factor of 4 improvement in speckle intensity (factor of 2 improvement in wave front). One obvious way to improve the result is to form an average $\Delta$ for the band pass (we used a $\Delta$ calculated for 750 nm).

Even if the broadband performance is partially limited by the differences in speckle amplitude at the RLP (fig. 6), the wave front estimate can be improved with additional iterations. This is because the speckle intensity is proportional to the square of the wave front deviations, and we have just shown that the wave front deviations are improved by a factor of two in a single iteration. We thus conclude that the optical bandwidth is not a limiting factor; the technique demonstrates high-accuracy wave front retrieval even over very large optical bandwidths.

Example 4: Shot-noise limit
We added noise to the RLP images of example 1 by considering an integration that gave $10^7$ photons per coherence cell, or $7 \times 10^{10}$ photons total. For Eclipse with a 100 nm band pass in the near infrared, this is obtainable on a V=5 source in about 400 seconds. Given $\sigma_\phi = 4 \lambda$, or 0.003 radians at $\lambda$ 850 nm, we expect $SNR_\phi = 9$.

Our simulation result, shown in Fig. 8, obtains an image improvement of 25 over the image formed with a 4 A wave front. This is equivalent to a 0.8 A r.m.s. result for the spatial frequencies of interest. There are two independent contributions to the 0.8 A result: the 0.63 A algorithm-limited portion of Ex. 1, and the shot noise. Assuming the two add in quadrature, we derive a shot-noise contribution of 0.5 A r.m.s. This compares well with the expected result of $1/9$ of 4 A, or 0.44 A.
5. CONCLUSION

The RLP approach is an effective means of estimating the front-end pupil phase through a back-end pupil measurement. It works in broadband light, it obtains the phase estimate independent of small amplitude errors, it is not affected by the presence of a planet, and it obtains a signal-to-noise ratio given by the square root of the number of photons reaching the back end.

The last point is important because it shows that the technique efficiently utilizes the light in the dark hole to estimate the wave front. For example, let’s look at how the SNR improves when operating at $Q=1$, that is, when the planet flux is equal to the mean scattered light flux. In a given time $3T$, we collect, say $3\bar{n}\sigma_p^2$ photons per speckle and the same number for the planet signal. The planet detection SNR is $\sqrt{n}/2$ when the background speckles are stable, and unity when background speckle fluctuations are the limiting factor. Let’s divide the available time into $3$ equal periods of duration $T$, two of which are used for planet wave front retrieval (to estimate $I_1$ and $I_2$) and the third for integration on the planet. The noise on the phase estimate is $\sigma_{\phi} = 1/\sqrt{n}$ and we will form a new phase estimate $\phi' = \hat{\phi} - \phi = 1/\sqrt{n}$ on average. The back end signal, originally $\bar{n}\sigma_p^2$, is improved to $\bar{n}\sigma_p^{2'} = n/n = 1$ photon! This is borne out by our simulation in example 4: we started with 65 photons per coherence cell and a 4 A wave front. We achieved a 0.5 A shot-noise-limited wave front reconstruction, improving the wave front by a factor of 8 and reducing the intensity at the science plane by $8^2=64$. Thus the reconstructed wave front, had it not been limited to 0.6 A r.m.s. by our simple wave front control algorithm, would have achieved $\sim 1$ photon per coherence cell after RLP wave front measurement.

For our planet imaging example, the $Q$ value is improved to $\bar{n}\phi^2$ (the number of detected planet photons in time $T$), while the planet detection SNR is improved by $\sqrt{2}$ when background shot-noise is important, and to $\sqrt{\bar{n}}\sigma_{\phi}$ when background speckle fluctuations are the limiting factor. For 100 detected photons, this is a factor of 10 improvement in SNR even though 2/3 of the available time is used for wave front sensing.

We noted in Sect. 2 that the while the RLP technique measures only the phase component of the complex pupil function, the amplitude component can be measured with a direct front-end pupil image. With knowledge of both the phase and amplitude, the DM can be commanded to generate a phase profile that cancels the light in one-half of the image plane, while doubling the energy in the other half of the plane. This is similar to the local dark-hole concept proposed by Malbet, Yu, and Shao [2], but the algorithm for determining the actuator setting is straightforward since the entire pupil function is known. This will be the subject of future work as we learn to improve the RLP approach.

Figure 8. Shot-noise limited example. There are $1x10^7$ photons per coherence cell at the front-end, and 65 photons per coherence cell detected at the back-end given the 4 A r.m.s. wave front at lambda = 850 nm. The shot noise contribution is $\sim 0.4$ A r.m.s. This adds in quadrature to the algorithm-limited performance of 0.6 A r.m.s. The algorithm limit can be improved through proper treatment of the complex back-end actuator influence function.
ACKNOWLEDGMENTS

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

REFERENCES