Methods and limitations of focal plane sensing, estimation, and control in high-contrast imaging

Tyler D. Groff
A. J. Eldorado Riggs
Brian Kern
N. Jeremy Kasdin
Methods and limitations of focal plane sensing, estimation, and control in high-contrast imaging

Tyler D. Groff,a,* A. J. Eldorado Riggs,a Brian Kern,b and N. Jeremy Kasdinb

aPrinceton University, Department of Mechanical & Aerospace Engineering, Equal Olden Street, Princeton, New Jersey 08544, United States
bJet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California 91109, United States

Abstract. Coronagraphy is a very promising method for directly imaging exoplanets, but the performance of a coronagraph is highly sensitive to quasi-static aberrations within the telescope. The resultant speckles are suppressed in the final focal plane using a wavefront control system that estimates the field at the final focal plane to avoid any noncommon path error. This requires a set of probe images that modulate the field so that it may be estimated. With an estimate of the focal plane electric field, a control law is defined to suppress the speckle field so that the planet can be imaged. Characterizing the planet requires that the speckle field be suppressed simultaneously over the bandpass of interest. The choice of control law, bandpass, estimator, and probing methodology has implications in the control solutions and contrast performance. Here, we compare wavefront probing, estimation, and control algorithms, and describe their practical implementation.

Keywords: aberrations; wavefront compensation; wavefront sensors; wavefronts; signal detection; Fourier optics.

1 Introduction

Directly imaging exoplanets, particularly Earth-like exoplanets, has been a goal of the scientific community for some time and was identified as a key technological priority in the last decadal review.1 Starshades and coronagraphs are the key enabling technologies for direct imaging missions in space. Both are experiencing great successes in their technology development, with a few modern examples of these designs being the Exo-S,2 Exo-C,3 and WFIRST-AFTA4 studies. The starshade option requires formation flight of the telescope with a second satellite that occults the star to a diffraction-limited level so that only the planet light enters the telescope. A coronagraph relies on a series of apodizers, masks, and stops in the instrument to suppress the starlight so that the off-axis planets can be imaged. The trade for a single satellite means that the starlight enters the telescope, resulting in aberrations that leak starlight into the coronagraphic image. Even small aberrations affect contrast at a 10−10 contrast level, the relative intensity at which we expect to see Earth-like exoplanets. Furthermore, these aberrations are dynamic in even the most stable telescope and result in a quasi-static speckle field in the image plane. The solution is to include deformable mirrors in the instrument so that the resultant speckles can be actively suppressed, allowing us to recover a dark hole of high contrast in the image plane. The precision required to create a dark hole at extremely high contrast drives wavefront-sensing methods away from sensors that introduce noncommon path optics. Thus, the control loop is closed around the final science image where we define model-based estimators and control laws where the space on the image plane electric field so that we can ultimately form an image from the telescope. What follows is the simplest possible focal plane wavefront control instrument. A schematic of the Princeton HCIL layout is given in Fig. 1 to illustrate the concept. One can consider the fiber launch in Fig. 1 to be the input image from the telescope. What follows is the simplest possible control system with two deformable mirrors (DMs). The beam is collimated by an off-axis parabola (OAP), reflects off two DMs in series, and propagates through the apodizer. In the case of the HCIL, the apodizer is a shaped pupil coronagraph,5,6 but the choice of apodizer only affects the transformation required to reach the image plane (which is generalized in this paper). With the shaped pupil located in the first pupil plane, we then form the image with a second OAP, block the core of the point spread function (PSF) with a focal plane mask, and reimagine the PSF onto the science camera. Note that there is no wavefront sensor, and the purpose of this section is to derive the transformations that describe the effect of perturbations in collimated space on the image plane electric field so that we can ultimately define model-based estimators and control laws where the sensing is done in the image plane. Augmenting the controller to multiple DMs, and mathematically correcting for moving DM surfaces to planes nonconjugate to the pupil is fully described by

*Address all correspondence to: Tyler D. Groff, E-mail: tgroff@princeton.edu

2 Wavefront Control

We begin by deriving the underlying mathematics for a generalized focal plane wavefront control instrument. A schematic of the Princeton HCIL layout is given in Fig. 1 to illustrate the concept. One can consider the fiber launch in Fig. 1 to be the input image from the telescope. What follows is the simplest possible control system with two deformable mirrors (DMs). The beam is collimated by an off-axis parabola (OAP), reflects off two DMs in series, and propagates through the apodizer. In the case of the HCIL, the apodizer is a shaped pupil coronagraph, but the choice of apodizer only affects the transformation required to reach the image plane (which is generalized in this paper). With the shaped pupil located in the first pupil plane, we then form the image with a second OAP, block the core of the point spread function (PSF) with a focal plane mask, and reimagine the PSF onto the science camera. Note that there is no wavefront sensor, and the purpose of this section is to derive the transformations that describe the effect of perturbations in collimated space on the image plane electric field so that we can ultimately define model-based estimators and control laws where the sensing is done in the image plane. Augmenting the controller to multiple DMs, and mathematically correcting for moving DM surfaces to planes nonconjugate to the pupil is fully described by

2329-4124/2015/$25.00 © 2015 SPIE

Downloaded From: https://nanophotonics.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 25 Jan 2020
Terms of Use: https://nanophotonics.spiedigitallibrary.org/terms-of-use
Pueyo et al.,7 with various details of controllability addressed by Shaklan and Green8 and Groff.9 Once again, these will have small effects on the form of the transformations that we have generalized here.

Treating the designed pupil field as \( A(u, v) \), the DM surface as a phase perturbation at the pupil plane, \( \phi(u, v) \), and including a distribution of complex aberrations, \( g(u, v) \), the electric field at the pupil plane at any control step \( k \) is given by

\[
E_{\text{pup},k}(u, v) = A(u, v)[1 + g(u, v)]e^{i\phi_k(u,v)}. \tag{1}
\]

For controller design, we write the controlled phase at step \( k \) as a sum of the previous DM surface and a small perturbation, \( \delta \phi \), about which we will ultimately linearize the phase induced by the DM. It is important to note that we should account for the aberrated field as a separate additive complex error, \( g(u, v) \), but this makes it less obvious that the initial field about which one linearizes should be the entire phase distribution in the pupil. In other words, \( \phi_0(u,v) \) would include an estimate of the entire aberrated field. At HCIT, this comes from the initial phase diversity estimates used to flatten the pupil field and determine its starting complex field.

Assuming that the phase perturbation from the DM, \( \delta \phi_k \), is adequately small, we let \( \phi_k = \phi_{k-1} + \delta \phi_k \) and take a first-order approximation of the exponential, resulting in the linearized form of the pupil field

\[
E_{\text{pup},k}(u, v) \cong A(u, v)[1 + g(u, v)]e^{i\delta \phi_k(u,v)}[1 + i\delta \phi_k(u,v)]. \tag{2}
\]

Additionally, we assume that the product \( g\delta \phi_k \) is negligible compared to the other terms since they will both be small and are of approximately the same magnitude, letting us rewrite the expanded field as

\[
E_{\text{pup},k}(u, v) \approx A(u, v)e^{i\delta \phi_k(u,v)}[1 + g(u, v) + i\delta \phi_k(u,v)]. \tag{3}
\]

Treating the propagation through the coronagraph and imaging system as a linear transformation between the image and pupil planes, \( \mathcal{C}[\cdot] \), the linearized form of the image plane electric field is written as

\[
E_{\text{in},k}(x, y) = \mathcal{C}[A(u, v)e^{i\phi_k(u,v)}[1 + g(u, v)]]+i\mathcal{C}[A(u, v)e^{i\phi_k(u,v)}g(u, v)]. \tag{4}
\]

Wavefront control is accomplished by commanding DM actuators, not setting the phase across a mirror. Thus, we apply a physical model of the DM in terms of actuator commands. Normalizing the surface height of the mirror, \( H(u, v) \), by the wavelength, \( \lambda \), the phase shift, \( \phi(u, v) \), induced by the DM is given by

\[
\phi(u, v) = \frac{4\pi}{\lambda}H(u, v). \tag{5}
\]

We describe this surface profile as a combination of the two-dimensional height maps imposed by each actuator. Since we are using a DM with a continuous face sheet, the contribution of any actuator will be highly localized but will still deform the entire DM surface. Thus, we define the contribution of any \( q \)th actuator as a characteristic two-dimensional phase map over the entire plane of the DM surface with unitary amplitude, \( f_q(u, v) \). We can then describe \( H(u, v) \) as a superposition of these characteristic shapes, commonly referred to as influence functions, over all \( N_{\text{act}} \) actuators. Defining the amplitude of each influence function (represented as a phase shift in meters to eliminate a factor of two) as \( a_q \), the phase perturbation induced by the DM becomes

\[
\phi(u, v) = \frac{2\pi}{\lambda} \sum_{q=1}^{N_{\text{act}}} a_q f_q(u, v). \tag{6}
\]

Since Eq. (6) is linear in actuator amplitudes, we can substitute Eq. (6) for \( \delta \phi_k \) in Eq. (4). The image plane response to a change in actuator amplitudes at the \( k \)th step is given by

\[
E_{\text{in},k}(u, v) = \mathcal{C}[A(u, v)e^{i\phi_k(u,v)}[1 + g(u, v)]]+\frac{2\pi i}{\lambda} \sum_{q=1}^{N_{\text{act}}} \delta a_q \mathcal{C}[Ae^{i\phi_k(u,v)}f_q(u, v)]. \tag{7}
\]

Since we only measure the field intensity at a finite number of pixels, we convert the continuous functions in Eq. (7) to values at each pixel and rewrite in matrix form

\[
E_{\text{in},k} = \mathcal{C}[Ae^{i\phi_k(1 + g_k)}] + G_k u_k, \tag{8}
\]

where \( E_{\text{in},k} \) is a matrix of dimension \( N_{\text{pix}} \times 1 \) containing the electric field at each pixel.

\[
G_k = i\mathcal{C}[Ae^{i\phi_k}] \tag{9}
\]

is the \( N_{\text{pix}} \times N_{\text{act}} \) control effect matrix accounting for the propagation of each influence function, each of which is embedded in a matrix \( f \), and \( u_k \) is a column matrix of changes in actuator amplitudes at step \( k \), \( u_k = [\delta a_1, \ldots, \delta a_{N_{\text{act}}}]^T \). The discretization of the field is naturally subject to numerical effects. One must take care to correctly sample each plane, and the choice of transform method can also become important. Numerical transforms are widely used in high-contrast imaging experiments, and their effect has been well studied. One notable study was done in the development of the PROPER optical propagation library by Krist.10

---

**Fig. 1** Optical layout of the Princeton High Contrast Imaging Laboratory (HCIL). Collimated light is incident on two deformable mirrors (DMs) in series, which propagates through a shaped pupil, the core of the point spread function is removed with an image plane mask, and the 90 deg search areas are reimaged on the final camera when using the Ripple3 shaped pupil coronagraph.
Finally, the measured intensity in the image plane is just the square of the field, \( I_{\text{im},k}(x,y) = |E_{\text{im},k}(x,y)|^2 \). We often refer to the area in the image plane where sufficient contrast has been achieved as the dark hole and use the notation \( I_{\text{DH}} \) for the high-contrast intensity (either designed or controlled) there. The objective of the wavefront controller is to find the control matrix \( u_k \) that will recover the desired contrast in the dark hole in the presence of the aberrations.

## 3 Speckle Nulling

Speckle nulling, whose fundamental equations are described fully in Ref.\(^1\) along with a speckle energy minimization and sensing scheme, is the simplest form of suppression in the final image plane. It does not require an estimate of the electric field and, therefore, only closes the loop on intensity images in an attempt to slowly drive the unobservable field states to zero. All that is needed is a mapping of spatial frequencies at the DM plane to image plane locations. Speckles are then singled out in the image and suppressed based on their intensity and phase. The amplitude of the command for any spatial frequency is directly proportional to the intensity of the speckle. Specifically, the DM amplitude should scale as \( \lambda \sqrt{I_{\text{speckle}}} \). The correct phase shift for each spatial frequency command on the DM is determined by minimizing the energy of the speckle in the image plane. Since the amplitude and phase of any spatial frequency on the DM are computed directly from intensity measurements at the detector, speckle nulling tends to be highly robust to model error and does not require as much computational overhead. However, it is highly inefficient compared to other control methods because speckles are dealt with one at a time and there is no guarantee that canceling a speckle will not degrade the contrast of other speckles. This has been mitigated in some sense by simultaneously solving for the multiple speckles at a time, which has been reported at the Ames Coronagraph Experiment Laboratory, where the algorithm is typically used in conjunction with electric field conjugation (EFC).\(^2\) Despite this, speckle nulling tends to be slow and is largely relied upon to clean up stray artifacts. The main advantage of speckle nulling is its robustness to model error, which is thoroughly demonstrated on sky by Martinache et al. without the aid of extreme adaptive optics.\(^3\)

### 4 Monochromatic Model–Based Control Algorithms

The primary control laws used in focal plane wavefront control for the past decade at Princeton and JPL are stroke minimization\(^4\) and EFC.\(^5\) Though other variants, such as the coronagraphic focal plane wavefront estimation for exoplanet detection (COFFEE)\(^6\) algorithm, have also been studied, stroke minimization and EFC have achieved the highest contrast of any controllers to date and are the primary control laws under consideration for the WFIRST-AFTA coronagraph system. In this section, we describe each, highlighting their differences and similarities.

#### 4.1 Stroke Minimization—Lagrange Multipliers

The stroke minimization algorithm\(^7\) seeks to regularize the control solution by directly minimizing stroke under the constraint that the solution for the desired DM perturbation achieves a specified contrast level.

Applying the matrix form of the control amplitudes in Eq. (9) and taking the inner product (represented as \( \langle ., . \rangle \)) of Eq. (4), we find the intensity in the dark hole at the operation wavelength, \( \lambda_0 \), to be

\[
I_{\text{DH}}(k) = \frac{4\pi^2}{\lambda_0^2} u_k^T M_k u_k + \frac{4\pi}{\lambda_0} u_k^T \mathcal{J} \{ b_k \} + d_k,
\]

where \( \mathcal{J} \{ \cdot \} \) is the imaginary component of the operand, and

\[
M_k = \langle C(Ae^{i\phi_{k-1} f_k}), C(Ae^{i\phi_{k-1} f_k}) \rangle = G_k^* G_k,
\]

\[
b_k = \langle E_{\text{ab},k}, C(Ae^{i\phi_{k-1} f_k}) \rangle = G_k^* E_{\text{ab},k},
\]

\[
E_{\text{ab},k} = C[Ae^{i\phi_{k-1}} (1 + g_k)].
\]

Conceptually, \( d_k \) is the column matrix of the intensity contribution from the aberrated field, \( E_{\text{ab},k} \). The measured intensity can be used to construct \( d_k \) if there is not a large bias in the signal. The matrix \( b_k \) represents the interaction of the DM electric field with the aberrated field, and \( M_k \) describes the additive contribution of the DM to the image plane intensity for a given DM shape, \( f_k \), at control step \( k \). Having represented \( I_{\text{DH}} \) in a quadratic form in terms of the control signal, we can use Eq. (10) to produce an optimal control strategy. The optimization problem in monochromatic light is functionally stated as

\[
\text{minimize} \quad \sum_{q=1}^{N} a_{q,k}^2 = u_k^T u_k,
\]

subject to \( I_{\text{DH}}(k) \leq C_k \)

where \( C_k \) is the contrast constraint at each iteration, which changes over time to steadily decrease contrast.

To solve the optimization problem, Pueyo et al.\(^7\) realized a control law by modifying the inequality constraint in Eq. (15) to be represented as the equality constraint \( (I_{\text{DH}} - C_k) = 0 \). Incorporating the equality constraint for the central wavelength into the minimization on stroke via a Lagrange multiplier, \( \mu_k \), yields a quadratic cost function given by

\[
J_k = u_k^T u_k + \mu_k [I_{\text{DH}}(k) - C_k]
\]

\[
= u_k^T u_k + \mu_k \left( \frac{4\pi^2}{\lambda_0^2} u_k^T M_k u_k + \frac{4\pi}{\lambda_0} u_k^T \mathcal{J} \{ b_k \} + d_k - C_k \right)
\]

\[
= u_k^T \left( I + \mu_k M_k \right) u_k + \mu_k \frac{4\pi^2}{\lambda_0^2} u_k^T \mathcal{J} \{ b_k \} + \mu_k (d_k - C_k),
\]

where \( I \) represents the identity matrix. The quadratic form guarantees a single minimum to the linearized equation. This is identical to a linear least squares solution. Recognizing that \( M_k \) is guaranteed to be symmetric, and hence \( M_k = M_k^T \), we take the partial derivative to find actuation that minimizes the cost function. Evaluating

\[
\frac{\partial J}{\partial u_k} \bigg|_{u_k} = 2 \left( I + \mu_k \frac{4\pi^2}{\lambda_0^2} M_k \right) u_k + \mu_k \frac{4\pi}{\lambda_0} \mathcal{J} \{ b_k \} = 0
\]

and solving for the optimal control input, we find

\[
\mu_k \frac{4\pi}{\lambda_0} \mathcal{J} \{ b_k \} = -2 \left( I + \mu_k \frac{4\pi^2}{\lambda_0^2} M_k \right) u_k
\]

\[
\mu_k \frac{4\pi}{\lambda_0} \mathcal{J} \{ b_k \} = -2 \frac{4\pi}{\lambda_0} \mathcal{J} \{ b_k \}
\]

\[
\mu_k = \frac{2}{\lambda_0} \mathcal{J} \{ b_k \}
\]
\[ u_{j,k} = -\left( \frac{1}{\mu_0} \mathcal{I} + \frac{2\pi}{\lambda_0} M_k \right)^{-1} \Delta \{b_k\}. \]  

(18)

To find the value of \( u_{j,k} \), all that is left is to find the value of \( \mu_k \) that minimizes the cost function, Eq. (16). This has historically been done via a line search on \( \mu_k \), continuously increasing \( \mu_k \), evaluating \( u_k \) with Eq. (18) for each value of \( \mu_k \) until the quadratic intensity given by Eq. (10) goes below the targeted contrast value. Since this is a linearized subprogram, we must relinearize about the new DM shape at each iteration of the correction algorithm (or every time we apply a new DM command). This is computationally expensive, so we tend not to do this in the experiment. However, the nature of the controller is such that the solution will not deviate dramatically from the optimal one found by relinearizing at each iteration. Since the controller is trying to minimize stroke, the control signal tends to remain in the regime of a particular linearization for as long as possible. Statistics on the control history for a representative two-DM experiment at HCIL is shown in Fig. 2. We observe that the vast majority of stroke is used to eliminate strong aberrations in the first five iterations of the control history. After this, the extrema and mode reduce dramatically, frequently a factor of 10 less stroke than the first iteration. Throughout the entire control history, the median never deviates far from zero, meaning that a dc drift never develops and the shape about which we are linearizing is nearly identical as we approach our ultimate achievable contrast.

4.2 Electric Field Conjugation—Tikhonov Regularization

EFC is an alternate controller approach first presented by Give’on et al.\textsuperscript{14} It was originally modeled on the traditional, pupil-based adaptive optics approach, where the DM is set to conjugate the phase in the pupil to recover as flat an entering field as possible. In focal plane wavefront control, the analogous field conjugation would be done in the image. That is, given an estimate of the electric field at step \( k-1 \) from Eq. (8), \( \hat{E}_{im,k} = C(Ae^{i\phi_{k-1}} + C(Ae^{i\phi_{k-1}}) \), the EFC controller seeks to drive to a desired field, \( E_D \), allowing us to define the targeted field perturbation as

\[ \delta E_k = \hat{E}_{im,k} - E_D. \]  

(19)

Substituting for the control effect in Eq. (8), the goal is to use the DM to create a field perturbation, \( \delta E_k \), such that the desired field becomes

\[ E_D = C\{A(u,v)e^{i\phi_{k-1}}[1 + g(u,v)]\} + \delta E_k = E_{im,k}(x,y). \]  

(20)

In its original form, EFC seeks to find a control \( G_k u_k \) that conjugates the residual aberration terms in Eq. (20), to create the desired final field in the image plane. The effect of different choices of \( E_D \) will become clearer once the controller has been fully defined.

Unfortunately, given the limitations of the influence function form of the DMs, the image plane field perturbation in Eq. (20) is not always reachable. In other words, one cannot simply invert \( G_k \) to solve for \( u_k \) given the desired field perturbation defined in Eq. (19). Instead, EFC seeks to find a control as close to the desired field perturbation as possible. We do so by defining an error, \( \epsilon_k \), between the controlled field and the desired field as

\[ \epsilon_k = G_k u_k - \delta E_k. \]  

(21)

EFC seeks to minimize this error rather than minimize the stroke as in Sec. 4.1. Doing so via a simple least-squares solution, however, often produces an ill-posed problem. Thus, EFC uses a regularized cost function to guarantee inversion. The regularization parameter, \( \Gamma_k \), is introduced via a Tikhonov regularization, making the quadratic cost function

\[ W_k = (G_k u_k - \delta E_k)^* (G_k u_k - \delta E_k) + u_k^T \Gamma_k u_k. \]  

(22)

Evaluating the partial derivative,

\[ \frac{\partial W_k}{\partial u_k} = 2(G_k^* G_k + \Gamma_k) u_k - 2G_k^* \delta E_k = 0, \]  

(23)

the control vector that minimizes Eq. (22), \( u_{w,k} \), is given by

![Fig. 2 Time history of actuation statistics for each DM. We plot the mode, median, and peak-to-valley actuation levels for (a) DM1 and (b) DM2 in a representative control run at the Princeton HCIL. The starting contrast is \( 1 \times 10^{-4} \) and final contrast is \( 2.3 \times 10^{-7} \) in symmetric dark holes.](image-url)
\[ u_{w,k} = (G_k^*G_k + \Gamma_k^T \Gamma_k)^{-1} G_k^* \delta E_k. \]  

(24)

With no particular reason to weight any specific field position or actuator, the regularization matrix, \( \Gamma_k \), is defined as

\[ \Gamma_k = \alpha_k T_k. \]  

(25)

The regularized solution and its associated cost are then given by

\[ W_k = (G_k u_k - \delta E_k)^T (G_k u_k - \delta E_k) + \alpha_k^2 u_k^T u_k, \]  

(26)

\[ u_{w,k} = (G_k^*G_k + \alpha_k^2 T_k)^{-1} G_k^* \delta E_k. \]  

(27)

As with stroke minimization, the task is now to search through \( \alpha_k \) to evaluate \( u_{w,k} \) in Eq. (27) and minimize Eq. (26). What is interesting is that the solution directly weights the minimization of stroke against the residual from the targeted electric field. As the regularization parameter, \( \alpha_k \), is increased, the solution will seek smaller solutions of \( u \), but the targeted electric field may become unreachable.

What remains is to chose \( \delta E_k \) for the field perturbation at each step, or equivalently, the desired field \( E_D \). If we attempt to drive to the nominal electric field that the coronagraph would achieve in a perfect optical system, making \( E_D = C[A] \), the target field will be given by

\[ \delta E_k = \hat{E}_k - C[A] \]  

(28)

\[ = C[A e^{i \phi_{kd}} (1 + g)] - C[A]. \]  

(29)

Under the optimistic assumption that the estimated component of the nominal field matches the expected nominal field, the target becomes

\[ \delta E_k = C[Ag]. \]  

(30)

If the intent is to drive the field to zero, \( E_D = 0 \), then the controller should attempt to conjugate the current field estimate, making

\[ \delta E_k = \hat{E}_k. \]  

(31)

Referring back to Eq. (12), we see that in this second scenario, \( G_k^* \delta E_k \approx b_0 \), and the form of the monochromatic solution is identical when using EFC or stroke minimization.

### 4.3 Comparing Control Laws

Where stroke minimization is attempting to achieve a targeted average contrast in the image plane via an equality constraint, EFC is attempting to minimize the desired field error and stroke is minimized via the regularization parameter. Thus, the true mathematical difference between stroke minimization and EFC is only in the inherent priority lent in each cost function. Where the primary metric of minimization in stroke minimization is the total DM stroke, leaving the resultant field as a free variable, the primary metric of minimization in EFC is the field residual, leaving the required stroke as a free variable. In monochromatic light, this simply means the Lagrangian for our central wavelength, \( \mu_k \), and Tikhonov regularization parameter, \( \alpha_k^2 \), are inverses of one another. This is entirely due to the fact that the inequality constraint of the intended algorithm in Eq. (15) has been implemented as an equality constraint in a quadratic cost function.

For monochromatic control, the real difference is not from the choice of control law but in its algorithmic application. Details in the choice of minimization parameters, \( (\mu_k, C_k) \) for stroke minimization and \( (\alpha_k^2, \delta E_k) \) for EFC, will yield different (sometimes dramatically different) solutions. The logic used in these choices seems to be a function of the user/laboratory in question and can have the greatest impact on varying control solutions to achieve a specific contrast level in the image plane.

In the case of the regularization parameter, the principal difference we see in their algorithmic applications is how this value is tuned, or if it is tuned at all. Here we point out that the proper mathematical application of these algorithms is to relinearize after every iteration and to perform a high-resolution line search on the Lagrange multiplier (or equivalently the regularization parameter) to find the true minimum for that time step.

Choosing the targeted contrast (or desired field) tends to be more of an ad hoc decision. Experimentally, this is generally driven by practically achievable step sizes in early iterations. At some point, one can pick a target that is unreachable by the controller at that iteration. There are essentially two scenarios for this value. The first is at the end of control when we have reached our final achievable contrast. This tends to be limited by noise, be it sensor or process noise. In this case, the contrast target is insignificant because we are ultimately limited by the estimator, which is discussed in Sec. 6. More interesting for the controller is the second scenario, where we observe unreachable states in early iterations. This would seem curious because these states are not below the instrument’s ultimate achievable contrast and we are not yet sensing at an unfavorable signal-to-noise ratio (SNR). The simplest explanation is that the actuator strokes are highest during early iterations and the contrast target will be limited by the accuracy of the current Jacobian. In this case, the next higher-order terms that were neglected in Eq. (8) become significant for the probe and control amplitudes in question. Thus, while theoretically reachable, the linearization defining the controller’s Jacobian, \( G_k \), will effectively have lower accuracy in early iterations than in later iterations because higher-order terms in the linear expansion are a function of the DM amplitude. This is true even if the same Jacobian is used in later iterations and is a separate issue from whether one chooses to relinearize at each iteration of the algorithm. As a functional example, the Princeton HCIL has demonstrated dark holes that suppress the aberrated field from \( \approx 1 \times 10^{-4} \) to \( \approx 2 \times 10^{-7} \) without relinearizing the Jacobian, \( b_0 \), but we cannot achieve \( 2 \times 10^{-7} \) in a single time step when we start at \( 1 \times 10^{-4} \). The effect of probe amplitude does play a role in accuracy, and its error propagation is discussed in Sec. 8. Here we are only making the case that the controller amplitude limits the ability of the Jacobian to accurately predict the shape required to achieve a particular contrast value or field target. Thus, we tend to throttle the target so that we are not left with undue DM residuals that will have to be compensated anyway.

## 5 Broadband Model–Based Control Algorithms

Monochromatic control of the image plane field represents a significant step toward detection, but the controller must ultimately
be capable of creating a dark hole in broadband light so that we can take spectra of the planets. We are typically only concerned with \(\approx 10\%\) bandpasses, but it would be useful to achieve a bandpass as large as 20\%. Some of the first broadband results, presented in Ref. 6, used the monochromatic controller at a central wavelength and accepted the contrast degradation in the rest of the bandpass. However, both EFC and stroke minimization have been developed into broadband control laws by augmenting the cost function with multiple monochromatic estimates that span the bandpass of concern.

5.1 Windowed Stroke Minimization

To define the bandpass of the controller, we begin with the monochromatic constraint in Eq. (15). The center wavelength constraint at \(\lambda_0\) is now augmented with two more contrast constraints \(C_{i,k}\) at bounding wavelengths \(\lambda_1, \lambda_2\), to define a window over which the correction will be made. With three separate constraints, the optimization now becomes

\[
\text{minimize} \quad \sum_{i=1}^{N_k} a_{i,k}^2 = u_k^T u_k \\
\text{subject to} \quad I_{DH}(\lambda) \leq C_{i,k}, \\
I_{DH}(\lambda_1) \leq C_{i,k}, \\
I_{DH}(\lambda_2) \leq C_{i,k},
\]

where \(\lambda_1 = \gamma_1 \lambda_0\)
\(\lambda_2 = \gamma_2 \lambda_0\).

Acknowledging that the variables in Eq. (10) are a function of wavelength, the variables defined in Eqs. (11) to (13) are written here as \(M_{i,k}, b_{i,k}\), and \(d_{i,k}\), respectively. Using these variables, the wavelength-dependent form of the dark hole intensity defined in Eq. (10) is given by

\[
I_{DH}(\lambda) = w(\lambda) \frac{4\pi^2}{\lambda^2} u_k^T M_{i,k} u_k + w(\lambda) \frac{4\pi}{\lambda} u_k^T 3 \{b_{i,k}\} + w(\lambda_2) d_{i,k},
\]

where the normalization value, \(w(\lambda)\), accounts for the fact that the relative intensity of each wavelength will vary as a function of wavelength. This additional normalization step prevents any wavelength from being artificially weighted in the cost function. Defining the wavelength as a scalar multiple of the central wavelength, \(\lambda_i = \gamma_i \lambda_0\), we rewrite Eq. (33) as

\[
I_{DH}(\lambda_i) = w(\lambda_i) \frac{4\pi^2}{\gamma_i^2 \lambda_0^2} u_k^T M_{i,k} u_k + w(\lambda_i) \frac{4\pi}{\gamma_i \lambda_0} u_k^T 3 \{b_{i,k}\} + w(\lambda_2) d_{i,k}.
\]

Using Eq. (34), we define a cost function from the problem described in Eq. (32) by applying three separate constraints, one per wavelength. Using a separate Lagrange multiplier for each wavelength, \(\mu_{i,0}, \mu_{i,1}\), and \(\mu_{i,2}\), the most general form of the broadband cost function is given by

\[
J_k = u_k^T u_k + \mu_{i,0} [I_{DH}(\lambda_i) - C_{i,k}] + \mu_{i,1} [I_{DH}(\lambda_i) - C_{i,k}] \\
+ \mu_{i,2} [I_{DH}(\lambda_i) - C_{i,k}]
\]

\[
= u_k^T \left[ I + \frac{4\pi^2}{\gamma_i^2 \lambda_0^2} \left\{ \mu_{i,0} M_{i,k} + \mu_{i,1} \frac{w(\lambda_i)}{\gamma_i^2} M_{i,k} \\
+ \mu_{i,2} \frac{w(\lambda_2)}{\gamma_i^2} M_{i,k} \right\} u_k + \frac{4\pi}{\gamma_i \lambda_0} u_k^T \left\{ \mu_{i,0} 3 \{b_{i,k}\} \\
+ \mu_{i,1} \frac{w(\lambda_i)}{\gamma_i} 3 \{b_{i,k}\} + \mu_{i,2} \frac{w(\lambda_2)}{\gamma_i} 3 \{b_{i,k}\} \right\} \\
+ \mu_{i,0} (d_{i,k} - C_{i,k}) + \mu_{i,1} w(\lambda_i) (d_{i,k} - C_{i,k}) \\
+ \mu_{i,2} w(\lambda_2) (d_{i,k} - C_{i,k}) \right].
\]

With multiple constraints, we now have the problem that the optimal solution in the three-dimensional space \((\mu_{i,0}, \mu_{i,1}, \mu_{i,2})\) does not necessarily guarantee the targeted suppression at all three wavelengths. However, with three Lagrange multipliers, we can guarantee suppression of all three wavelengths by restricting the search to a single dimension. We write the Lagrange multipliers of the two bounding wavelengths as weighted values of the first so that \(\mu_{i,1} = \delta_i \mu_{i,0}\) and \(\mu_{i,2} = \delta_2 \mu_{i,0}\). Applying this relationship to Eq. (35), the windowed stroke minimization cost function becomes

\[
J_k = u_k^T \left[ I + \frac{4\pi^2}{\gamma_i^2 \lambda_0^2} \left\{ \mu_{i,0} M_{i,k} + \delta_i \frac{w(\lambda_i)}{\gamma_i^2} M_{i,k} + \delta_2 \frac{w(\lambda_2)}{\gamma_i^2} M_{i,k} \right\} u_k \\
+ \mu_{i,0} \frac{4\pi}{\gamma_i \lambda_0} \left\{ 3 \{b_{i,k}\} + \delta_i \frac{w(\lambda_i)}{\gamma_i} 3 \{b_{i,k}\} + \delta_2 \frac{w(\lambda_2)}{\gamma_i} 3 \{b_{i,k}\} \right\} \\
+ \mu_{i,0} (d_{i,k} - C_{i,k}) + \delta_i w(\lambda_i) (d_{i,k} - C_{i,k}) \\
+ \delta_2 w(\lambda_2) (d_{i,k} - C_{i,k}) \right].
\]

We are left with a weighted optimization where we can effectively tune the contrast requirement on the bounding wavelengths, \((\lambda_1, \lambda_2)\), against the central wavelength, \(\lambda_0\), with our choice of \(\delta_1, \delta_2\). Following the same procedure as in Sec. 4.1, we take the partial derivative of Eq. (36) with respect to the control vector. In this case, the broadband control analog to Eq. (18) is given by

\[
u_{j,k} = \frac{1}{\mu_{i,0} 2\pi} \left\{ M_{i,k} + \frac{2\pi}{\lambda_0} \frac{w(\lambda_i)}{\gamma_i^2 M_{i,k}} \right\} \left( \frac{1}{\mu_{i,0} 2\pi} \left\{ M_{i,k} + \frac{2\pi}{\lambda_0} \frac{w(\lambda_i)}{\gamma_i^2 M_{i,k}} \right\} \right)^{-1} \\
\cdot \frac{3 \{b_{i,k}\}}{\gamma_i} + \frac{w(\lambda_i)}{\gamma_i} 3 \{b_{i,k}\} + \frac{w(\lambda_2)}{\gamma_2} 3 \{b_{i,k}\}.
\]

With only one Lagrange multiplier left, we apply the same algorithmic procedure as in Sec. 4.1, using Eq. (37) to determine the value of \(u_{j,k}\) as we search for value of \(\mu_{i,0}\) that minimizes the cost function, Eq. (36).

With the controller defined, we now evaluate its ability to produce a uniform contrast in wavelength space. The most recent broadband results from the Princeton HCIL producing symmetric dark holes in a targeted 10\% band around \(\lambda_0 = 0.533\) nm are shown in Fig. 3. In all cases, the controller corrects two rectangular areas spanning 6 to 11 \(\times\) 3 to \(3\lambda_{\text{iso}}/D\) on both sides of the image plane. The contrast level is evaluated over two rectangular areas spanning 7 to 10 \(\times\) 2 to \(2\lambda_{\text{iso}}/D\) on both sides of the image plane. Note that this area is defined by
the central wavelength for all measurements so that contrast performance correlates to a fixed sky angle, $\beta$, defined as $\beta = \tan^{-1}(n\lambda_0/D)$. Since both DMs are nonconjugate to the pupil and have very large wavefront error, we anticipate there to be a great deal of chromaticity in the Princeton HCIL. We will briefly address these limitations in the context of the experimental results, but for now, we chose to slightly underweight the bounding wavelengths in the optimization, choosing $\delta_1 = \delta_2 = 0.75$ at the bounding wavelengths, $\lambda_1 = 600$ and $\lambda_2 = 650$ nm. In this particular experiment, we estimate each wavelength separately using the batch process estimator discussed in Sec. 6.1. The final dark holes, shown in Fig. 3(d), reached a contrast of $5.67 \times 10^{-6}$ in a $\sim 10\%$ band and $1.364 \times 10^{-5}$ over the full band of the laser, shown in Fig. 3(b).

Individual monochromatic images of the final dark hole are given in Fig. 4. With the bounding wavelengths only slightly underweighted in the optimization, we see in Fig. 3(a) that the contrast is relatively constant in wavelength over the 10% band. We also see in Figs. 4(a) and 4(c) that there is very little evolution of the dark hole at the bounding wavelengths of the 10% bandpass.

The question that now arises is what limits contrast as a function of our choice in bandpass. Correcting over a 10% band automatically costs us an order of magnitude in our final achievable contrast. The most obvious answer is the inherent chromaticity of our pupil from highly aberrated, nonconjugate planes. Shaklan et al. show that the ultimate achievable contrast is dependent on the propagation of phase-induced amplitude errors from optics that are nonconjugate to the pupil plane. They also point out that since the control system has a finite set of controllable spatial frequencies, the ultimate achievable contrast will become a function of the bandwidth we choose in our dark hole. Using the results of Shaklan et al., we have previously estimated the lower limit of the Princeton HCIL to be $\approx 1 \times 10^{-6}$ for a 20% bandwidth. This order of magnitude bound only accounts for the quilting of the two DM surfaces, both of which are shown in Fig. 1 to be at planes nonconjugate to the pupil. This bounding estimate both highlights the importance of bandwidth in the optical design and indicates that our current performance is still worse than the fundamental limitations of this optical system.

Given that we do not believe we have reached a fundamental contrast floor in a 10% band, we now consider other limitations that could affect broadband control. At the time of these measurements, the contrast of speckle measurements in the Princeton HCIL was only stable to within $\approx 0.5$ to $1 \times 10^{-6}$ over the period of tens of minutes to an hour. It is possible
that we reached a stability limit in the experimental setup in this particular broadband experiment. The broadband source is of much lower power than the monochromatic source, driving up the integration time for each image. The broadband solution also requires three individual estimates to achieve the results shown in Fig. 3. Thus, we are estimating over a time scale that will affect our measurements by \( \approx 1 \times 10^{-6} \), and it is not surprising that we were only able to correct down to the \( 1 \times 10^{-6} \) level at any wavelength. This highlights the importance of setting stability requirements, which must scale with the minimum allowable integration time required to obtain an adequate SNR. HCIL, being an air experiment with no active temperature control directly on the optical bench, has very poor stability compared to what is expected from a space environment. However, we can control under equivalent time scales as the observatory environment by taking shorter exposures (between 100s of milliseconds to a minute depending on the experiment) with a brighter source. This makes yet another point; to accurately predict the performance of a flight instrument from an experimental testbed, one must extrapolate the system stability to time scales relevant to the expected integration time. Broadband estimation, taking inherently longer to accomplish than monochromatic estimation, should be the time scale associated with the control vector that minimizes Eq. (22), \( u_{w,k} \), is

\[
W_k = \sum_{i=0}^{N} (G_{j,i,k}u_k - \delta E_{j,i,k})^T (G_{j,i,k}u_k - \delta E_{j,i,k}) + u_k^T \Gamma_k u_k
\]

(39)

Evaluating the partial derivative,

\[
\frac{\partial W_k}{\partial u_k} = 2 \left( \sum_{i=0}^{N} G_{j,i,k}^* G_{j,i,k} + \Gamma_k^{-1} \right) u_{w,k} - 2 \sum_{i=0}^{N} G_{j,i,k}^* \delta E_{j,i,k} = 0
\]

(40)

5.2 Broadband Electric Field Conjugation

As with stroke minimization, we create a broadband form of EFC by augmenting the cost function with multiple wavelengths. We redefine the error function in Eq. (21) for the monochromatic case as

\[
e_{j,i,k} = G_{j,i,k}u_k - \delta E_{j,i,k},
\]

(38)

where \( G_{j,i,k} \) and \( \delta E_{j,i,k} \) are the wavelength-dependent Jacobian and field target, respectively. Summing a discrete set of these error metrics in wavelength, the broadband EFC cost function is given by

\[
W_k = \sum_{i=0}^{N} (G_{j,i,k}u_k - \delta E_{j,i,k})^T (G_{j,i,k}u_k - \delta E_{j,i,k}) + u_k^T \Gamma_k u_k,
\]

(42)

If we once again simplify the regularization matrix using Eq. (25), the cost function and associated control solution are given by

\[
W_k = \sum_{i=0}^{N} (G_{j,i,k}u_k - \delta E_{j,i,k})^T (G_{j,i,k}u_k - \delta E_{j,i,k}) + u_k^T \Gamma_k u_k,
\]

(42)

\[
u_{w,k} = \left[ \sum_{i=0}^{N} G_{j,i,k}^* G_{j,i,k} + \Gamma_k^{-1} \right]^{-1} \sum_{i=0}^{N} G_{j,i,k}^* \delta E_{j,i,k}.
\]

(43)

As Give’on et al.\(^1\) have pointed out, it is also possible to avoid the summation by augmenting the Jacobian as a larger diagonal matrix. The mathematics are identical to the summation presented here, and the approach is largely limited by the maximum
practical (or allowable) matrix size in the computation. In either case, the task is once again the same. We must search through \( \alpha_k \) to evaluate \( u_{w,k} \) in Eq. (43) and minimize Eq. (42).

The resemblance between the broadband EFC algorithm and windowed stroke minimization is not as strong as it was in the monochromatic case. Their reciprocal treatment of the regularization parameter and Lagrange multipliers becomes more significant in the presence of multiple parameters. Since the constraints are applied to the contrast target in windowed stroke minimization, each wavelength has its own Lagrange multiplier. With the modification made in Eq. (36) to further constrain the solution, the two become much more similar in that there is only a single regularization parameter we search through in the optimization. If arbitrary weights were included in Eq. (39), then Eqs. (42) and (43) would be nearly identical in form to the optimization defined in Eqs. (36) and (37). As in the monochromatic case, their exact functional similarity will be associated with the choice of \( \delta E_{b,k} \) in the EFC control law, and the algorithmic approach to their implementation remains unchanged save for the choice in weights.

6 Wavefront Estimation

The principle of focal plane wavefront control is that high contrast is best achieved by avoiding all possible noncommon errors that arise from path, temporal, or wavelength differences. Thus, the controller is provided with an estimate of the electric field at the final image plane, rather than from a measurement taken by a separate wavefront sensor. The three principal methods used to sense the image plane electric field in high-contrast imaging are Gerchberg-Saxton based multipane solutions,19-21 differential fringe imaging, such as the self-coherent camera22 or pinhole masks,23,24 and DM probing. Multipane solutions are typically used as an initialization step where the pupil field is flattened in an attempt to initialize control with the cleanest beam possible. Creating fringe patterns in the image is a very clever way to sense the field and has demonstrated great success in high-contrast imaging.25 As a flight instrument, this means adding a mechanism and is not necessarily compatible with all coronagraphs. This is particularly true if multiple classes of coronagraphs are used in the same instrument, as is the case with the WFIRST-AFTA concept. With or without a fringe-based technique, DM probing is always possible and is attractive because it only relies on the DM already needed for wavefront control. No additional sensors or mechanisms are required to probe the electric field in the image plane. There are, however, limitations with such a method. Here we derive the estimation scheme associated with DM-based probing where the image plane is modulated with a set of diversity, or probe, images.

6.1 Batch Process Method

Pair-wise difference imaging, as first developed by Give’on et al.,14 allows the estimation of the focal plane electric field from several intensity images, each with a different DM setting. As originally conceived, at least two pairs of probed images are required to invert the measurements and obtain the real and imaginary parts of the electric field. The choice of these probe shapes is important and is the principal topic for the remainder of this paper. Here, in Sec. 6, we first derive the basic least-squares estimator, which will subsequently be compared to a Kalman filter estimator.

Letting \( u_{j,k} \) be the signal for the \( j \)th positive probe shape in the \( k \)th iteration, we refer back to Eq. (10) to define the change in the focal plane electric field for a given pair of probe shapes, \( p_{j,k} = \pm G_{u,j,k} \). Recalling Eq. (14), we apply these shapes to Eq. (10) to define their image plane intensities as

\[
I_{j,k} = |E_{ab,k}|^2 + |p_{j,k}|^2 + 2\Re\{E_{ab,k}p_{j,k}\},
\]

where \( \Re\{\cdot\} \) takes the real part of the field. The difference of the positive and negative probed images is just the cross-term.

\[
\Delta I_{j,k} = I_{j+k} - I_{j-k} = 4\Re\{E_{ab,k}, p_{j,k}\}.
\]

Recognizing that \( \Re\{A,B\} = \Re\{A\} \Re\{B\} + \Im\{A\} \Im\{B\} \)

\( \Re\{A^*B\} = \Re\{A\} \Re\{B\} + \Im\{A\} \Im\{B\} \)
takes the imaginary part of the field, we can write the observation as

\[
\begin{pmatrix}
\Delta I_{1,k} \\
\vdots \\
\Delta I_{n_p,k}
\end{pmatrix} = 4
\begin{pmatrix}
\Re\{p_{1,k}\} & \Im\{p_{1,k}\} \\
\vdots & \vdots \\
\Re\{p_{n_p,k}\} & \Im\{p_{n_p,k}\}
\end{pmatrix}
\begin{pmatrix}
\Re\{E_{ab,k}\} \\
\Im\{E_{ab,k}\}
\end{pmatrix},
\]

(46)

where \( n_p \) is the number of probe pairs. The relationship between the observation and state in Eq. (46) has now taken the linear form

\[
z_k = H_k x_k,
\]

(47)

where the observation \( z_k \), observation matrix \( H_k \), and image plane electric field state \( x_k \) are defined as

\[
z_k = \begin{bmatrix}
\Delta I_{1,k} \\
\vdots \\
\Delta I_{n_p,k}
\end{bmatrix},
\]

(48)

\[
H_k = 4
\begin{pmatrix}
\Re\{p_{1,k}\} & \Im\{p_{1,k}\} \\
\vdots & \vdots \\
\Re\{p_{n_p,k}\} & \Im\{p_{n_p,k}\}
\end{pmatrix},
\]

(49)

\[
x_k = \begin{bmatrix}
\Re\{E_{ab,k}\} \\
\Im\{E_{ab,k}\}
\end{bmatrix}.
\]

(50)

For any given image plane pixel, we can now estimate the electric field state at the \( k \)th iteration by taking the left pseudo-inverse of \( H \).

\[
\hat{x}_k = (H_k^T H_k)^{-1} H_k^T z_k,
\]

(51)

where \( \hat{x}_k \) is the least-squares estimate of the electric field state. The batch process estimation method has proven to be reliable in high-contrast imaging, but suffers from losing information over the control history. There is no choice but to take a full set of probes at each iteration.

6.2 Filtering Method

A more recent form of estimation employs a Kalman filter to estimate the field.16 In its initial formulation, the state estimate, \( \hat{x} \), and the observation matrix, \( H_k \), are identical to those constructed in Sec. 6.1. The construction of \( \hat{x}_k \) and \( H_k \) define the form of the controller update, \( \Gamma_{k-1} \), used to extrapolate the field estimate given the control applied in the last iteration,
perturbations, $\phi_j$, that must be chosen to modulate the intended control area well to maximize the signal in the image plane and to guarantee a well-conditioned observation matrix, $H_k$, associated with that set of probes. As was discussed in Sec. 6.2, this is the only factor that can result in measurement updates making the estimate worse over time.

These shapes are traditionally based on analytical functions for which we know the Fourier transform. Following Give'on et al., we will simplify the problem of coverage/shape of the dark hole by choosing symmetric rectangular regions that span the region we wish to estimate and, for the sake of choosing the shape only, assume that the operator is a Fourier transform (i.e., $C\{\cdot\} = F\{\cdot\}$). We define an offset pair of rectangles of width $w_x$ and height $w_y$ by convolving the rectangle function with a set of Dirac delta functions, $\delta\{\cdot\}$. Shifting the rectangles off axis in the image plane by a distance $a$ in the $x$ dimension and a distance $b$ in the $y$ dimension, the intended probe field in the image plane from a pupil plane perturbation, $A(u, v)$, $\phi(u, v)$, is defined as

$$F\{A\phi\} = F\{A\} \ast c \text{rect}(w_x x \text{rect}(w_y y)$$

$$\ast \delta(x-a) + \delta(x+a) \ast \delta(y-b) + \delta(y+b)$$

$$= F\{c A \text{sinc}(w_x u) \text{sinc}(w_y v) \cos(au) \cos(bv)\},$$

(57)

where $c$ is an arbitrary amplitude applied to the shape. Inverse transforming, the shape we would like the DM to approximate is

$$\phi = c \text{sinc}(w_x u) \text{sinc}(w_y v) \cos(au) \cos(bv).$$

(58)

This particular probe function is very simple to use during control because the size and location of the probed area in the image plane in $\lambda/D$ units are exactly equal to the spatial frequency applied to the DM in units of cycles per aperture. Thus, to produce a probe pair, we simply specify the desired spatial frequencies in Eq. (58) and apply $+\phi_j(u, v)$ and $-\phi_j(u, v)$ to get $I_j^+$ and $I_j^-$, respectively. To get multiple probe pairs, one simply applies phase shifts to one of the cos functions in Eq. (58).

It is important to note that Eq. (57) should not be used to construct the observation matrix. This should be done by solving for the DM commands from Eq. (58) and applying them to the true system Jacobian in Eq. (9). The PSF will alter the field so that the field from a DM shape given by Eq. (58) will not have exact unit amplitude and the edges of the rectangle will extend by one radius of the PSF. The distribution will still be relatively uniform, most likely modulating each pixel in the dark hole.

## 8 Error Analysis for Pair-Wise Probing

With DM probe shapes defined in Sec. 7, we analyze the variance of the electric field estimate given bounding cases of measurement noise. We begin with the effect of stellar shot noise on the estimate, demonstrate how zodi and dark current drive us to brighter probes, and determine the order of magnitude limit on how bright we might make the probes. We then determine what strategies can be used to obtain the most precise and accurate electric field estimates possible. In all cases, we propagate the uncertainty in the probe through the batch process estimator to more easily see the effect of noise in a single iteration. Since this amounts to a measurement error, the noise value will be identical in the Kalman filter and the optimal gain will more heavily weight information from the prior iteration. Thus, the
batch process estimator represents the worst-case contribution of probe errors to our achievable contrast.

8.1 Stellar Shot Noise

Here, we will calculate the variance in the electric field estimate at a given pixel in the presence of only stellar shot noise. By utilizing probe shapes that actuate the entire focal plane evenly in field (such as the products of sinc and sinc functions mentioned in Sec. 7), we can safely apply these calculations to all pixels in the correction region.

We choose probes \( p_{j,k} \) in the \( k \)th iteration such that they have the same amplitude \( p = |p_{j,k}| \forall j \).

\[
p_{j,k} = p e^{i\theta_{j,k}} = p \cos(\theta_{j,k}) + ip \sin(\theta_{j,k}),
\]

(59)

where \( \theta_{j,k} \) is an arbitrary angle. Uniformly distributing the phase of the probes, \( \theta_{j,k} \), within \([0, \pi]\) to get good coverage of the real-imaginary plane, we substitute Eq. (59) into Eq. (46) to rewrite the observation as

\[
\begin{bmatrix}
\Delta I_1 \\
\vdots \\
\Delta I_n_p
\end{bmatrix} = 4p \begin{bmatrix}
\cos(\alpha + \pi \frac{n_p n_x}{n_p}) & \sin(\alpha + \pi \frac{n_p n_x}{n_p}) \\
\vdots \\
\cos(\alpha + \pi \frac{n_p n_x - 1}{n_p}) & \sin(\alpha + \pi \frac{n_p n_x - 1}{n_p})
\end{bmatrix} \begin{bmatrix}
E_R \\
E_I
\end{bmatrix}
\]

(60)

where \( \alpha \) is some arbitrary phase angle offset, and \( E_R \) and \( E_I \) are shorthand for \( \Re \{ab_{j,k}(x,y)\} \) and \( \Im \{ab_{j,k}(x,y)\} \), respectively. We define \( H_\theta \) as above (such that \( H = 4pH_\theta \)), to obtain our new state estimate equation.

\[
\hat{x} = \frac{1}{4p}(H_\theta^T H_\theta)^{-1}H_\theta^T z.
\]

(61)

With some algebra, it can be shown that

\[
H_\theta^T H_\theta = \begin{bmatrix}
\sum_{j=0}^{n_p-1} \cos^2 \left( \alpha + \pi \frac{n_p n_x - 1}{n_p} \right) & \sum_{j=0}^{n_p-1} \sin \left( \alpha + \pi \frac{n_p n_x - 1}{n_p} \right) \\
\sum_{j=0}^{n_p-1} \sin \left( \alpha + \pi \frac{n_p n_x - 1}{n_p} \right) & \sum_{j=0}^{n_p-1} \cos^2 \left( \alpha + \pi \frac{n_p n_x - 1}{n_p} \right)
\end{bmatrix}
\]

\[
= \frac{n_p}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\]

(62)

We can then write Eq. (61) as

\[
\begin{bmatrix} E_R \\ E_I \end{bmatrix} = \frac{1}{4p n_p} \begin{bmatrix}
\sum_{j=0}^{n_p-1} \Delta I_j \cos \left( \alpha + \pi \frac{n_p n_x - 1}{n_p} \right) \\
\sum_{j=0}^{n_p-1} \Delta I_j \sin \left( \alpha + \pi \frac{n_p n_x - 1}{n_p} \right)
\end{bmatrix}.
\]

(63)

We now want to calculate the variance \( \sigma^2 \{ \cdot \} \) of the real and imaginary parts of the electric field, which is dependent on the variance of the \( \Delta I_j \)'s and any uncertainty in \( p, \sigma^2 \{ p \} \).

Assuming that \( \sigma^2 \{ \Delta I \} = \sigma^2 \{ \Delta I \} \forall j \), we can write

\[
\begin{align*}
\sigma^2 \{ E_R \} &= \frac{1}{4p^2} \left( \sum_{j=0}^{n_p-1} \Delta I_j \cos \left( \alpha + \pi \frac{n_p n_x - 1}{n_p} \right) \right)^2 \\
\sigma^2 \{ E_I \} &= \frac{1}{4p^2} \left( \sum_{j=0}^{n_p-1} \Delta I_j \sin \left( \alpha + \pi \frac{n_p n_x - 1}{n_p} \right) \right)^2
\end{align*}
\]

(64)

The second term in Eq. (64) is a fractional error in \( E_R \) and \( E_I \), equal to the fractional error (or uncertainty) in \( p, \sigma \{ p \} / p \). The most likely sources of this error are DM calibration errors and DM motions not being produced as they were commanded (e.g., hysteresis, electronic drift, etc.). It is likely that these systematic issues can be better understood by consistent use of the same probes and empirical calibration. The rest of this derivation will focus on the fundamental Poisson limit, implicitly assuming that \( \sigma \{ p \} \) is significantly smaller than \( p \) such that the first term in Eq. (64) dominates.

The fundamental limit to measurement of \( E_R \) and \( E_I \) comes from Poisson statistics on the intensity measurements in Eq. (64), which is rewritten (without the \( \sigma \{ p \} \) term) as

\[
\sigma^2 \{ E_R \} = \sigma^2 \{ E_I \} = \left( \frac{1}{4p} \right)^2 \left( \frac{2}{n_p} \right)^2 \sigma^2 \{ \Delta I \}.
\]

(65)

Referencing Eq. (44), we evaluate Eq. (65) in the two limiting cases of large, \( p \gg E_0 \), and small, \( p \ll E_0 \). In the case of large probes of equal amplitude, Eq. (44) can be approximated as

\[
I_{pk} \sim p^2,
\]

(66)

where we drop the \((x,y)\) notation and write \(|p_j| = p\) since we assume the probes are of equal amplitude. The Poisson statistics of any of the intensity measurements come from the number of detected photoelectrons in that measurement (photoelectrons being assumed for most Si detectors), so it is necessary to relate physical units to \( I \). In high-contrast coronagraphy, it is common to relate the intensities measured throughout the image plane to the peak of an image with no occulter, \( I_{pk} \), which is determined by integration time per intensity measurement \( t \) and a flux \( F_{pk} \) (expressed in detected photoelectrons/s) such that

\[
I_{pk} = F_{pk} t.
\]

(67)

The intensities at each image plane location where the probes are being applied can be assigned a contrast value \( C \), which can be used very generally to relate to any intensity measured, such that

\[
I = C I_{pk} = C F_{pk} t.
\]

(68)

where the units of \( I \) are explicitly photoelectrons. In general, we assume that \( t \) is the same for every intensity measurement in the probe sequence and that \( I_{pk} \) is unchanged by the probe DM settings (which is accurate on the order of \( p^2 / I_{pk} \)), so that only \( C \) changes through the probing sequence.

To determine the variances relevant to Eq. (65), we define \( C_p \) such that
\[ p^2 = C_p F_{pk} t, \quad (69) \]

where \( C_p \) has a contrast interpretation, e.g., \( C_p = 10^{-8} \) corresponds to probes that add \( 10^{-8} \) intensity in the image plane. With this formulation, the units of \( p^2 \) are photoelectrons. For an individual intensity measurement, Poisson statistics dictate that the variance in photoelectrons is equal to the expected number of photoelectrons. For large probes,

\[ \sigma^2 \{ I \} \sim p^2, \quad (70) \]

remembering that the units of both sides of this equation are photoelectrons. The photoelectron normalization implicitly comprises an expectation value \( \langle \rangle \) corresponding to probes that add \( p^2 \) carries units of photots perform correctly because photots, not to the accu-

The photoelectron unit is difficult for the notation to carry suc-

It is instructive to define a new term, a dimensionless noise-

Another observation about \( N_c \) is worth revisiting. While the \( N_c \) limit from Poisson statistics is a fundamental limit, the fractional error \( \sigma(p)/p \) from Eq. (64) also affects the maximum improvement from a single probe estimate. For example, con-

### 8.2 Stellar Shot Noise Using Small Probes

Section 8.1 dealt with large probes, where \( p^2 \gg |E_0|^2 \). While the intermediate case where \( p^2 \sim |E_0|^2 \) does not have much algebraic simplicity, the case of small probes, \( p^2 \ll |E_0|^2 \), is simple to express. For small probes, the individual measured intensities, and therefore their Poisson noise terms, are dominated by \( |E_0|^2 \); based on Eq. (44) and analogous to Eq. (66), the probe intensity can be approximated as

\[ I_{pk} \sim |E_0|^2. \quad (77) \]

Following the same logic as with the large probes, the noise equivalent contrast \( N_c \) for small probes becomes

\[ N_c = \frac{1}{F_{pk} t_{tot}} |E_0|^2. \quad (78) \]

Since, by construction, small probes have \( |E_0|^2/p^2 \gg 1 \), this noise is far larger than the corresponding large probe \( N_c \).
8.3 Other Noise Sources

We can expect there are other sources of measurement noise that will degrade our estimate quality. Here we will include the effects of zodi or exozodi light $Z$ (in units of photoelectrons, integrated over time $i$), detector dark current $D$ (also in units of photoelectrons), and readout noise variance $\sigma_{\text{ran}}^2$ (in units of photoelectrons$^2$). Large probes are assumed, as they will have the least noise in the electric field estimates. A measurement $\Delta I$ comprising $n_{\text{exp}}$ stacked exposures per probe image, where $n_{\text{exp}}$ exposures together accumulate $t$ integration time, will have a variance

$$\sigma^2\{\Delta I\} = 2\sigma^2\{I_{j,k}\} + 2\sigma^2\{Z\} + 2\sigma^2\{D\} + 2n_{\text{exp}}\sigma_{\text{ran}}^2 = 2p^2 + 2Z + 2D + 2n_{\text{exp}}\sigma_{\text{ran}}^2.$$  

(79)

The factors of 2 in the variance components all come from the two intensity measurements, each containing $n_{\text{exp}}$ individual exposures, which form a $\Delta I$ measurement. Normalizing by $I_{pk}$, we find

$$N_e = \frac{1}{F_{pk}i_{\text{tot}}} \left(1 + \frac{Z + D + n_{\text{exp}}\sigma_{\text{ran}}^2}{p^2}\right).$$  

(80)

We can now see some new properties of the variance of the electric field estimate. We see that increasing the probe brightness $p$ reduces the variance contribution from incoherent background (zodi, dark current) and readout noise. If probes can be made arbitrarily large, $N_e$ will approach the fundamental limit from stellar shot noise alone.

8.4 Maximum Allowable Probe Brightness

Earlier we found that brighter probes (larger $p$) reduce the variance of the electric field estimate. However, if we make $p$ too large, higher-order terms of the Taylor series neglected in Eq. (3) will become the dominant source of error in Eq. (44). Here we will find the maximum allowable intensity of a probe before nonlinear terms become significant relative to the nominal contrast level, $|E_{ab}|^2$. For simplicity, we will once again evaluate a single pixel in the image plane. Rather than assuming we already know the probe $p_{jk}$, we will treat the additive electric field change of the $j'$th probe at the $k$'th iteration as a general value $\Delta E_{jk}$. In this case, the least-squares solution in Eq. (46) is rewritten as

$$[\Delta I_{1,j,k}] = 4\left[\begin{array}{c} \Re\{\Delta E_{1,j,k}\} \\ \vdots \\ \Re\{\Delta E_{n_{jk},j,k}\} \\ \Im\{\Delta E_{1,j,k}\} \\ \vdots \\ \Im\{\Delta E_{n_{jk},j,k}\} \end{array}\right] = 4\left[\begin{array}{c} \Re\{E_{ab,j,k}\} \\ \vdots \\ \Re\{E_{ab,j,k}\} \end{array}\right].$$  

(81)

We will now evaluate the effect of higher-order terms appearing in the observations, $z_{j,k}$, and whether we can compensate for those nonlinear terms by constructing the observation matrix, $H_{jk}$, with the full nonlinear model. Expanding about the DM surface as we did in Eq. (3), a higher-order expansion at the DM plane is given by

$$E_{\text{pup},j,k}e^{i\phi_i} \approx \bar{E} \left[1 + i\delta \phi_i - \frac{1}{2} \delta \phi_i^2 - \frac{i}{3!} \delta \phi_i^3 - \frac{1}{4!} \delta \phi_i^4 \right].$$  

(82)

where $\bar{E} = E_{\text{pup},j,k}e^{i\phi_i}$. Applying a specific positive or negative probe to the expanded field in Eq. (82), the focal plane electric becomes

$$E_{ab,j,k}(\pm \delta \phi_i) = C\{E_{\text{pup},j,k}e^{i\phi_i}\} \approx C\{\bar{E}\} \pm \frac{1}{2} C\{\bar{E}\delta \phi_i^2\} + \frac{1}{3!} C\{\bar{E}\delta \phi_i^3\} + \frac{1}{4!} C\{\bar{E}\delta \phi_i^4\}.  

(83)

We expect the average change in electric field $\Delta E$ between the positive and negative probe to be

$$\Delta E = \frac{E_{ab,j,k}(+\delta \phi_i) - E_{ab,j,k}(-\delta \phi_i)}{2} \approx \frac{G_k u_{j,k} - i}{3!} C\{\bar{E}\delta \phi_i^3\}.  

(84)

The higher-order terms in Eq. (84) can become significant compared to $G u$ for sufficiently large probe amplitudes, $|\delta \phi_i|$. Since these terms will be observed in the measurement, we evaluate their contribution to our measurement, $\Delta I_{jk}$, for a single probe. Calculating $\Delta I_{jk}$ up to fourth order in $\phi_i$ and $\delta \phi_i$, we find

$$\Delta I_{j,k} = |E_{j,k} + \Delta E_{j,k}|^2 - |E_{j,k}|^2 = 4\Re\{\langle E_{ab,j,k}, G_k u_{j,k} \rangle\} - 2\Re\{\langle G_k u_{j,k}, C\{\bar{E}\delta \phi_i^2\} \rangle\} + \frac{1}{3} \Re\{\langle C\{\bar{E}\delta \phi_i^3\}, C\{\bar{E}\delta \phi_i^2\} \rangle\} + \frac{1}{6} \Re\{\langle C\{\bar{E}\delta \phi_i^4\}, C\{\bar{E}\delta \phi_i^3\} \rangle\}.  

(85)

where $\langle \cdot, \cdot \rangle$ still denotes the complex inner product. With the observation nonlinearities in place, we now determine what nonlinear terms exist on the right-hand side of Eq. (81) by expanding one row of the observer. Recalling that $\Re\{A\} \Re\{B\} + \Im\{A\} \Im\{B\} = \Re\{\langle A, B \rangle\}$, the nonlinear terms from a probe that appear in the observer are given by

$$4\Re\{\langle \Delta E_{j,k}, E_{ab,j,k} \rangle\} = 4\Re\{\langle \Delta E_{j,k}, E_{ab,j,k} \rangle\} - \frac{2}{3} \Re\{\langle E_{ab,j,k}, C\{i\bar{E}\delta \phi_i^3\} \rangle\}.  

(86)

Subtracting Eq. (86) from Eq. (85), the unmodeled residuals in the measurement that are not eliminated by constructing the observation matrix with a nonlinear model are given by

$$\Delta I_{\text{err}} = -2\Re\{\langle G_k u_{j,k}, C\{\bar{E}\delta \phi_i^2\} \rangle\} + \frac{1}{3} \Re\{\langle \bar{E}\delta \phi_i^3, C\{i\bar{E}\delta \phi_i^2\} \rangle\} + \frac{1}{6} \Re\{\langle G_k u_{j,k}, C\{\bar{E}\delta \phi_i^4\} \rangle\}.  

(87)

By constructing the observation matrix, $H_{jk}$, from the full nonlinear model, we do indeed eliminate some of the dominant nonlinear residuals in our probe images. Had we simply used $\Delta E_{j,k} = G_k u_{j,k}$ to construct $H_{jk}$, the estimate would have been subjected to all of the nonlinear residuals in Eq. (85).
However, Eq. (87) shows us that not all higher-order terms are canceled in $\Delta I$ by using a nonlinear model to construct $H_k$. This is because formulating the observer fundamentally requires that we define a linear system. Thus, we have mitigated errors in our estimate from unmodeled measurement nonlinearities, but we have not eliminated them entirely. It is the nonlinear residuals of Eq. (87) that define the maximum allowable probe amplitude. Recalling from our development of Eq. (83) that $G_k u_{j,k} = C\{\tilde{E}\delta_{j,k}\}$, the first term in Eq. (87) will be of order $|\delta_{j,k}|^2$ and the other two are both of order $|\delta_{j,k}|^3$. Making the simplifying assumption that the fifth-order terms will be small compared to the third-order term, Eq. (87) can be approximated as

$$\Delta I_{\text{err}} \approx -2\Re\{G_k u_{j,k}, C\{\tilde{E}\delta_{j,k}\}\}. \quad (88)$$

We now wish to determine the contrast level at which the unmodeled nonlinear residuals of Eq. (88) begin to dominate our estimate for a given probe amplitude, $p = Gu$. Although $C\{\tilde{E}\delta_{j,k}\}$ does not simplify in terms of $G$ and $u$, we approximate it as $C\{\tilde{E}\delta_{j,k}\} \approx (G_k u_{j,k})^2$ in Eq. (88) to get an order of magnitude estimate of this floor. Thus, the dominant nonlinear error in our estimate will scale as

$$\Delta I_{\text{err}} \approx |G_k u_{j,k}|^3 = |C\{\tilde{E}\delta_{j,k}\}|. \quad (89)$$

If the peak contrast of the probes (assuming all $n_p$ probes have equal amplitude) in the $k$th iteration is $10^{-4}$, $|G_k u_{j,k}| \approx 10^{-2}$ and our unmodeled nonlinear errors are at the $10^{-6}$ contrast level. We therefore cannot expect to control below that. Solving in the other direction, if we are probing a dark hole at a contrast of $10^{-3}$, the maximum probe contrast is $10^{-6}$. It is worth mentioning that uncertainties in the probe shape will result in imperfect cancelation of nonlinearities in the observation. The effect of probe uncertainty is minimized by measuring the probe amplitude directly as explained in Ref. 29, but its phase must still be obtained from the model. One can always use Eq. (89) as a starting point for wavefront correction simulations to hone in on the maximum allowable brightness of the probe.

9 Conclusions

We have outlined what we consider to be the current best practices of focal plane wavefront probing, estimation, and control. Both the EFC and stroke minimization control laws are model-based, and their performance relies heavily on the accuracy of the Jacobian used to define the transfer function between DM actuation and image plane field. When stroke minimization algorithm is implemented with an equality constraint in a quadratic cost function, it is identical to EFC when its desired field is equal amplitude) in the probe. However, the EFC algorithm, but a multiparameter optimization is inherent to the formulation of windowed stroke minimization. However, the biggest difference in the control solution for a specific dark hole comes in the user’s choice of contrast/field target, how often the Jacobian is relinearized, and how the regularization parameters are chosen at each iteration. The choice in size and location of the dark hole also plays a major role in performance, but the outcome is highly dependent on the initial state realized in any individual optical system.

Control will only be as accurate as the wavefront estimation and probing steps. Here we have described the original batch process method as a segue to more advanced Kalman filtering algorithms. So long as the Kalman filter’s observation matrix is well conditioned, we are guaranteed that the estimate will not get worse over time, a very important fact as the dark hole reaches higher contrast levels and noise plays a more significant role. Since the estimator will be limited by the probes we apply, we propagated measurement errors through the estimator to quantify how they limit our contrast performance. Various noise sources and uncertainties bound the brightness of the probes, suggesting that brighter probes are better until nonlinear errors in the estimate become significant.

Acknowledgments

Tyler D. Groff, A. J. Eldorado Riggs, and N. Jeremy Kasdin are funded for this work by the NASA Nancy Grace Roman Technology Fellowship (NNX15AF28G), the NASA Space Technology Research Fellowship (NNX14AM06H), and Jet Propulsion Laboratory (JPL) subaward 1499194 under contract from National Aeronautics and Space Administration (NASA). For Brian Kern, this research was carried out at JPL, California Institute of Technology, under a contract with NASA.

References

Groff et al.: Methods and limitations of focal plane sensing, estimation, and control in high-contrast imaging


Tyler D. Groff is an associate research scholar at Princeton University. He received his BS in mechanical engineering and astrophysics from Tufts University in 2007 and his PhD from Princeton University in 2012. His current research interests include exoplanet imaging, optomechanics, optical design, wavefront estimation and control, infrared spectroscopy, integral field spectrographs, deformable mirrors, and ferrofluidics. He is a member of the American Astronomical Society and SPIE.

A. J. Eldorado Riggs is a PhD candidate in the Department of Mechanical and Aerospace Engineering at Princeton University. He received his BS in physics and mechanical engineering from Yale University in 2011. His research interests include coronagraph design and wavefront estimation and control for exoplanet imaging.

Brian Kern is an optical engineer at NASA Jet Propulsion Laboratory (JPL). He received his PhD from Caltech in 2002. His research interests are in high-contrast imaging. He works on the JPL High Contrast Imaging Testbed.

N. Jeremy Kasdin is a professor of mechanical and aerospace engineering and vice dean of the School of Engineering and Applied Science at Princeton University. He is the principal investigator of Princeton’s High Contrast Imaging Laboratory. He received his PhD from Stanford University in 1991. His research interests include space systems design, space optics and exoplanet imaging, orbital mechanics, guidance and control of space vehicles, optimal estimation, and stochastic process modeling. He is an associate fellow of the American Institute of Aeronautics and Astronautics and member of the American Astronomical Society and SPIE.