

PROCEEDINGS OF SPIE

SPIDigitalLibrary.org/conference-proceedings-of-spie

Numerically optimized coronagraph designs for the Habitable Exoplanet Imaging Mission (HabEx) concept

A J Eldorado Riggs, Garreth Ruane, Kevin Fogarty, Laurent Pueyo, Kunjithapatham Balasubramanian

A J Eldorado Riggs, Garreth Ruane, Kevin Fogarty, Laurent Pueyo, Kunjithapatham Balasubramanian, "Numerically optimized coronagraph designs for the Habitable Exoplanet Imaging Mission (HabEx) concept," Proc. SPIE 10698, Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave, 106980V (9 August 2018); doi: 10.1117/12.2313843

SPIE.

Event: SPIE Astronomical Telescopes + Instrumentation, 2018, Austin, Texas, United States

Numerically Optimized Coronagraph Designs for the Habitable Exoplanet Imaging Mission (HabEx) Concept

A J Eldorado Riggs^a, Garreth Ruane^b, Kevin Fogarty^c, Laurent Pueyo^c, Kunjithapatham Balasubramanian^a

^aJet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109

^bCalifornia Institute of Technology, Pasadena, CA 91125

^cSpace Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

ABSTRACT

The primary science goal of the Habitable Exoplanet Imaging Mission (HabEx), one of four candidate flagship missions under investigation, is to image and spectrally characterize Earth-like exoplanets. It is well known that pupil obscurations degrade coronagraphic performance and complicate coronagraph design, so HabEx is planned to have an off-axis, unobscured primary mirror. We utilize the circular symmetry of the aperture to investigate 1D-radial coronagraph optimization methods that are prohibitively time-consuming or intractable in 2D, such as diffractive pupil remapping and concurrent, multi-plane optimization. We also directly constrain sensitivities to dynamic, low-order Zernike aberrations, which are separable in polar coordinates and can thus be propagated as 1D-radial integrals. The mask technologies in our designs claim heritage from the extensive modeling and testbed experiments performed by the Wide-Field Infrared Survey Telescope (WFIRST) Coronagraph Instrument (CGI) project. In this paper, we detail our optimization methods and outline future work to complete our design survey.

Keywords: HabEx, coronagraph, exoplanet, wavefront control, optimization

1. INTRODUCTION

The primary science goal of the Habitable Exoplanet Imaging Mission (HabEx), one of four concept flagship missions under investigation for the 2020 Decadal Survey, is to image and spectrally characterize Earth-like exoplanets.¹ The proposed mission plan is to image exoplanets in visible light with a coronagraph instrument and to obtain spectra out to longer wavelengths with a starshade. Because coronagraphs for unobscured, circular apertures have been found to be much more efficient and robust to disturbances, the HabEx concept is baselining a monolithic, off-axis primary mirror. Cost and launch vehicle size set the baselined primary mirror diameter at 4.0 meters. In order to image exo-earths with its 4m aperture, HabEx requires a high-throughput, broadband coronagraph with an inner working angle (IWA) less than about $2.5\lambda_c/D$, where λ_c is the central wavelength of the bandpass and D is the telescope diameter. The spectral bandwidth ($\Delta\lambda/\lambda_c$) of each imaging channel is 20% and is limited by the expected polychromatic controllability with two deformable mirrors (DMs).²

The baselined coronagraph architecture for HabEx has a charge 6 vector vortex coronagraph (VVC) and two 64×64 -actuator Boston Micromachines Corporation (BMC) DMs.³ As described in detail by Ruane et al.,⁴ the VVC offers the best known combinations of high throughput and low-order aberration (LOA) robustness at small working angles. The only tuning parameter for the VVC is the charge of the vortex; a higher charge improves LOA robustness at the expense of a larger IWA.

In this paper, we investigate designs for HabEx's backup option, the hybrid Lyot coronagraph (HLC).^{5,6} The primary justification for this alternate path is that the HLC has extensive testbed and modeling heritage from the Wide-Field Infrared Survey Telescope (WFIRST) Coronagraph Instrument (CGI) project.⁷⁻¹¹ In addition, the HLC may be able to provide a comparable scientific yield if the VVC continued to be operable in only

Send correspondence to A.J. Riggs: aj.riggs@jpl.nasa.gov

©2018 California Institute of Technology. Government sponsorship acknowledged.

Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave, edited by Makenzie Lystrup,
Howard A. MacEwen, Giovanni G. Fazio, Proc. of SPIE Vol. 10698, 106980V
© 2018 SPIE · CCC code: 0277-786X/18/\$18 · doi: 10.1117/12.2313843

one polarization state at a time. Whereas VVC design has a single tuning parameter, HLC optimization has thousands that define the DMs, hybrid focal plane mask (FPM), and Lyot stop. The enormous size of this nonlinear, non-convex optimization problem makes the HLC difficult and time-consuming to design. Our main goal in this paper is to develop a tractable, reproducible HLC design strategy to meet a given set of performance requirements. To accomplish this goal, we break the optimization into several stages and perform much of our analysis in 1D-radial coordinates.

The sections of this paper are as follows. Section 2 derives the mathematics used for the 1D-radial optimization problem. Section 3 presents the design steps and results of our 1D-radial HLC optimization procedure. Next, Section 4 describes the steps necessary to accommodate and use a 1D-radial solution in an actual 2D system. Finally, Section 5 lists the remaining work to complete our HLC design and analysis pipeline.

2. FORMULATION OF THE PROBLEM IN 1D-RADIAL COORDINATES

The focus of this paper is coronagraph optimization in 1D. The immense time savings of optimizing in 1D-radial instead of 2D allows us to pursue more design strategies and perform more comprehensive surveys. With 1D-radial coordinates, it is also simple to include direct constraints on LOA robustness for many common types of coronagraphs. For observatories with circularly symmetric apertures such as HabEx, 1D-radial optimization also produces designs that are usable either directly or with minor modifications.

2.1 1D-Radial Optical Propagation

All propagations in our idealized optical systems for design are either Fourier transforms between a pupil and focus or angular spectrum (AS) Fresnel propagations between DMs. The 1D-radial equivalent of a Fourier transform, the zeroth-order Hankel transform, converts a pupil plane E-field $E(r)$ to a focal plane E-field $E(\rho)$,

$$E(\rho) = 2\pi \int_0^\infty E(r) J_0(2\pi\rho r) r dr, \quad (1)$$

where J_0 is a zeroth-order Bessel function of the first kind. The order of a Hankel transform matches the order of the Bessel function within it.

The 1D-radial AS propagation, however, is substantially more complicated. If the AS propagation equation is directly translated from Cartesian to polar coordinates, energy is not conserved during propagation and erroneous, high-frequency ringing appears in the E-field no matter what resolutions are used in the input and output planes. Guizar-Sicairos and Gutiérrez-Vega¹² provide the solution, which is to use non-uniform coordinate sampling chosen to match the zeros of the relevant Bessel function order.

2.2 Low-Order Aberrations

Zernike polynomials are straightforward to propagate through a coronagraph in 1D-radial coordinates. Most observatory dynamics, against which a coronagraph needs to be robust, are well represented by Zernike polynomials with Noll indices Z2-Z11. These correspond to tip/tilt, defocus, astigmatism, coma, trefoil, and spherical. In order to 1D-radially propagate Zernike modes with azimuthal components, we add together orthogonal mode pairs with one mode being real and the other being imaginary. This creates a single, complex-valued mode with a circularly symmetric amplitude component, allowing us to separate variables and move the mode's azimuthal component outside the radial integral.

As described by Ruane et al.,⁴ two changes occur to the Hankel transform when propagating a Zernike mode. First, the order of the Hankel transform changes to the radial order n of the Zernike mode. Second, the radial polynomial of that Zernike is included as another factor with the integral. Table 1 lists the radial polynomials $R_n^{|m|}$ for the first few Zernike modes, where m is the known as the degree. The equation for the Hankel transform including a Zernike mode is thus

$$E(\rho, \phi) = 2\pi(-i)^m e^{im\phi} \int_0^\infty R_n^{|m|} E(r) J_n(2\pi\rho r) r dr, \quad (2)$$

from which separation of variables yields

$$E(\rho) = 2\pi \int_0^\infty R_n^{|m|} E(r) J_n(2\pi\rho r) r dr \quad (3)$$

$$E(\phi) = (-i)^m e^{im\phi}. \quad (4)$$

An important point is that the order of the Hankel transform stays the same in subsequent propagations due to how Eq. 4 is derived.^{4,13}

Zernike Mode	Radial Polynomial
Piston: $R_0^0(r)$	= 1
Tip & Tilt: $R_1^1(r)$	= $(\frac{r}{a})$
Defocus: $R_2^0(r)$	= $2(\frac{r}{a})^2 - 1$
Astigmatism: $R_2^2(r)$	= $(\frac{r}{a})^2$
Coma: $R_3^1(r)$	= $3(\frac{r}{a})^3 - 2(\frac{r}{a})$
Trefoil: $R_3^3(r)$	= $(\frac{r}{a})^3$
Spherical: $R_4^0(r)$	= $6(\frac{r}{a})^4 - 6(\frac{r}{a})^2 + 1$

Table 1. The first few Zernike modes and their radial polynomials. The radial coordinate r is normalized by the circle's radius a . Values are from Appendix A of Ruane et al.⁴

Most coronagraph types can be modeled and optimized in 1D-radial coordinates. The previous equations are valid as long as all the coronagraph optics are circularly symmetric and the Fresnel approximation still holds. One notable exception is the vortex coronagraph, whose focal plane mask has an azimuthal phase ramp. Fortunately, similar to how the Hankel transform was modified for Zernike modes with azimuthal components, the vortex mask can also be modeled in 1D-radial coordinates by using a different radial polynomial and incrementing the order of the Hankel transform by the charge of the vortex. These modifications are listed in Appendix C of the paper by Ruane et al.⁴

2.3 1D-Radial Wavefront Control

As in our 2D optimization code, the Fast Linearized Coronagraph Optimizer (FALCO),¹⁴ we use linearized control in our 1D model to find solutions for the larger nonlinear optimization problem. Linearized control is a simple-to-implement yet powerful technique, so we use it exclusively in our initial 1D work presented in this paper. The two well known and closely related controllers for coronagraphic wavefront control, electric field conjugation (EFC)¹⁵ and stroke minimization,¹⁶ were originally derived in 2D-Cartesian coordinates. By default, they assume even spatial weighting for all pixels in the image plane correction region, called the dark hole. Because 1D radial pixels represent rings with different areas in the actual 2D image plane, the linearized controllers 1D must be adjusted to counteract this implicit weighting.

2.3.1 Unconstrained Linearized Control

We first address unconstrained focal plane wavefront control, which is commonly used with DMs. The cost function, J_k , to minimize at control iteration k is composed of the average intensity in the dark hole, \bar{I}_k plus a scalar weight γ times a quadratic penalty on the actuation command vector, Δu_k . The average intensity term is expanded as a sum weighted by the radial coordinate ρ with pixel index j ,

$$\begin{aligned} J_k &= \bar{I}_k + \gamma \Delta u_k^T \Delta u_k \\ &= (2\pi \Delta \rho \sum_j \rho_j I_{j,k}) / (2\pi \Delta \rho \sum_j \rho_j) + \gamma \Delta u_k^T \Delta u_k \\ &= (\sum_j \rho_j I_{j,k}) / (\sum_j \rho_j) + \gamma \Delta u_k^T \Delta u_k. \end{aligned} \quad (6)$$

The intensity I is just the squared E-field E at that iteration, and the E-field at the current iteration is linearly approximated as the previous E-field plus the control Jacobian G times the latest command, giving

$$\begin{aligned} I_{j,k} &= E_{j,k}^* E_{j,k} \\ &= (E_{j,k-1} + G_{j,k} \Delta u_k)^* (E_{j,k-1} + G_{j,k} \Delta u_k) \\ &= E_{j,k-1}^* E_{j,k-1} + \Delta u_k^T G_{j,k}^* G_{j,k} \Delta u_k + 2 \Delta u_k^T \mathcal{R}\{G_{j,k}^* E_{j,k-1}\}, \end{aligned} \quad (7)$$

where $\mathcal{R}\{\cdot\}$ returns the real part of a complex value and $*$ is the complex conjugate. After condensing the regularization coefficient γ and the sum $\sum_j \rho_j$ into the scalar term ν , we further expand the cost function and convert it to matrix form,

$$\begin{aligned} J_k &= \sum_j \rho_j I_{j,k} + \nu \Delta u_k^T \Delta u_k \\ &= \sum_j (\rho_j E_{j,k-1}^* E_{j,k-1} + \rho_j \Delta u_k^T G_{j,k}^* G_{j,k} \Delta u_k + 2 \rho_j \Delta u_k^T \mathcal{R}\{G_{j,k}^* E_{j,k-1}\}) + \nu \Delta u_k^T \Delta u_k \\ &= E_{k-1}^* \mathbb{P} E_{k-1} + \Delta u_k^T G_k^* \mathbb{P} G_k \Delta u_k + 2 \Delta u_k^T \mathcal{R}\{G_k^* \mathbb{P} E_{k-1}\} + \nu \Delta u_k^T \mathbb{I} \Delta u_k, \end{aligned} \quad (8)$$

where \mathbb{P} is a square matrix with the coordinates ρ_j of the dark hole pixels along the diagonal, and \mathbb{I} is the identity matrix.

To find the optimal unconstrained controller command, we take the derivative of the cost function with respect to the command and set it to zero,

$$\begin{aligned} \frac{\delta J_k}{\delta \Delta u_k} = 0 &= \frac{\delta}{\delta \Delta u_k} \left(\Delta u_k^T (G_k^* \mathbb{P} G_k + \nu \mathbb{I}) \Delta u_k + 2 \Delta u_k^T \mathcal{R}\{G_k^* \mathbb{P} E_{k-1}\} \right) \\ &= 2(G_k^* \mathbb{P} G_k + \nu \mathbb{I}) \Delta u_k + 2 \mathcal{R}\{G_k^* \mathbb{P} E_{k-1}\} \end{aligned} \quad (9)$$

and then re-arrange to solve for the command,

$$\Delta u_k = (G_k^* \mathbb{P} G_k + \nu \mathbb{I})^{-1} \mathcal{R}\{G_k^* \mathbb{P} E_{k-1}\}. \quad (10)$$

The 1D-radial control solution is the same as that for EFC in Cartesian coordinates except for some differences given by the matrix \mathbb{P} for pixel weighting and by the extra factor of $\sum_j \rho_j$ in the regularization matrix coefficient.

2.3.2 Constrained Linearized Control

When optimizing mask properties with the controller, constraints are often needed to enforce physical limitations. For example, the amplitude transmission of a mask needs to stay in the range $[0, 1]$, and the thicknesses of mask materials must be nonnegative. It is sub-optimal to modify the unconstrained solution from Eq. 10 with after-the-fact adjustments such as railing particular values or re-scaling the entire command. (Those strategies work well when using only DMs because the actuators typically start in the middle of their stroke ranges and few actuators are affected.) For the concurrent DM and mask optimizations in this paper, we have observed better results with direct constraints applied during the optimization.

For this study, we used the optimization package CVX^{17, 18} called from MATLAB. The convex solver used by CVX minimizes the constrained cost function directly, so we do not have to take the derivative of Eq. 8. Instead, we directly pass to CVX the quadratic cost function along with constraints and bounds,

$$\begin{aligned} &\text{minimize} && \Delta u_k^T \left((G_k^* \mathbb{P} G_k + \nu \mathbb{I}) \Delta u_k + 2 \mathcal{R}\{G_k^* \mathbb{P} E_{k-1}\} \right) \\ &\text{subject to} && -\Delta u_{max} \leq \Delta u_k \leq \Delta u_{max} \\ &\text{subject to} && u_{min} \leq u_{k-1} + \Delta u_k \leq u_{max}, \end{aligned} \quad (11)$$

where Δu_{max} is the max positive or negative step size at each step, and u_{min} and u_{max} are the absolute bounds on the cumulative controller commands.

2.4 1D-Radial Model of Deformable Mirrors

Deformable mirrors can be integral parts of coronagraph designs via diffractive pupil remapping, as demonstrated with the HLC⁹ and Active Compensation of Aperture Discontinuities-Optimized Stroke Minimization (ACAD-OSM).^{19,20} One of the main purposes of the FALCO¹⁴ software package is to perform such diffractive pupil remapping quickly in 2D.

There are two issues with implementing diffractive pupil remapping in 1D-radial coordinates. The first, as discussed in Section 2.1, is that the AS propagation between DMs requires a special formulation with non-uniform coordinate sampling. The second is that the DMs used in high-contrast imaging are not circularly symmetric; their actuator grids are either square or hexagonal. We nevertheless have to use a circularly symmetric DM representation to optimize in 1D. As shown in Fig. 1 with every fourth actuator poked, we defined our circularly symmetric DM actuators as rings with a Gaussian cross section. The actuator spacing was chosen to be the same as it would be along the actuators of a true DM on a square grid. That is, there are 32 ring actuators across the radius of the 1D DM model to represent a real 64×64 -actuator DM. This actually means that the 1D model has significantly fewer radial modes than the 64×64 DM, but we leave finding the most realistic basis set of 1D-radial actuators for future work.

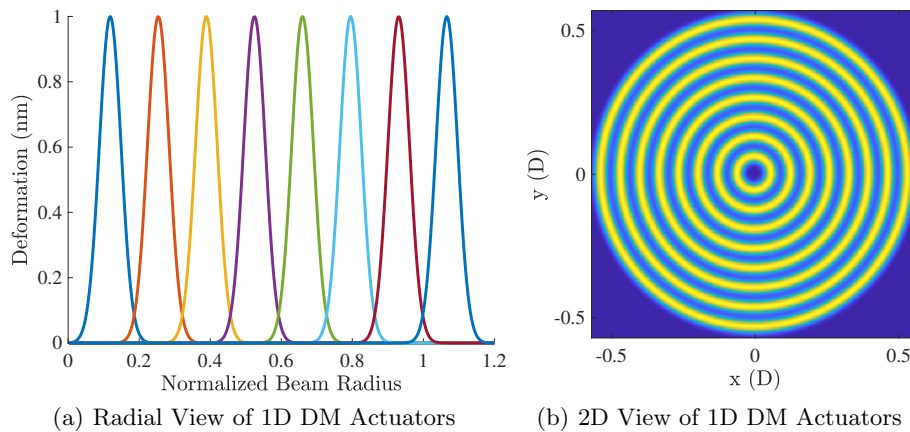


Figure 1. Basis set of DM “actuators” in the 1D-radial model. Every fourth actuator is poked in (a) a 1D plot and (b) a 2D map to show the contents of the basis set. The radial cross-section of the ring actuators is a Gaussian.

3. 1D-RADIAL HYBRID LYOT CORONAGRAPH DESIGN

We found that the most effective way to optimize the HLC is in discrete steps or stages. When we attempted to optimize the DMs and FPM properties concurrently from the start, the optimization converged poorly. Based on our observations, we believe that that initial strategy failed because the FPM was trying to counter the shape of the central incident PSF at the same time that the DMs were apodizing the PSF. In other words, the DMs were changing the inner PSF heavily while the FPM wanted that part of the PSF to stay the same. Therefore, we split the optimization into two main parts: (1) DM-only optimization and (2) concurrent DM and FPM optimization. We call the result of the first step a deformable mirror Lyot coronagraph (DMLC) because the FPM is still a simple opaque spot at that stage.

All the designs in this section have 10% spectral bandwidth instead of 20% in order to use fewer wavelength samples while honing a design strategy. We have begun repeating our analysis with 20% bandwidth designs and have not observed any significant differences or degradations in performance.

3.1 Deformable Mirror Lyot Coronagraph (DMLC)

As with the design strategy for 1D shaped pupil Lyot coronagraphs (SPLCs) described by Coker et al.,²¹ we performed an extensive set of DMLC optimizations for different combinations of inter-DM Fresnel number, FPM spot radius, Lyot stop inner radius, and Lyot stop outer radius. One slice of the results are shown in Fig. 2. Note that we exclude the two columns for Lyot stop inner diameter (ID) of 2% and 4% of D because they are

too small to implement in practice. As expected from the opening of the Lyot stop, the core throughput in Fig. 2(a) is higher for smaller Lyot stop ID and larger Lyot stop outer diameter (OD). However, we also would like small peak-to-valley (P-V) DM stroke and a smaller normalized intensity, which favors the opposite direction for larger Lyot stop ID and smaller Lyot stop OD. In later work, we will be able to continue all of these survey points into the second optimization stage (with concurrent FPM optimization) and score each design based on its estimated scientific yield. For exploratory purposes right now, we simply choose a conservative point in the parameter space (inter-DM Fresnel number = 620, FPM inner radius = $2.2 \lambda_c/D$, Lyot stop ID = $0.14D$, Lyot stop OD = $0.78D$) and carry that configuration into the next optimization stage.

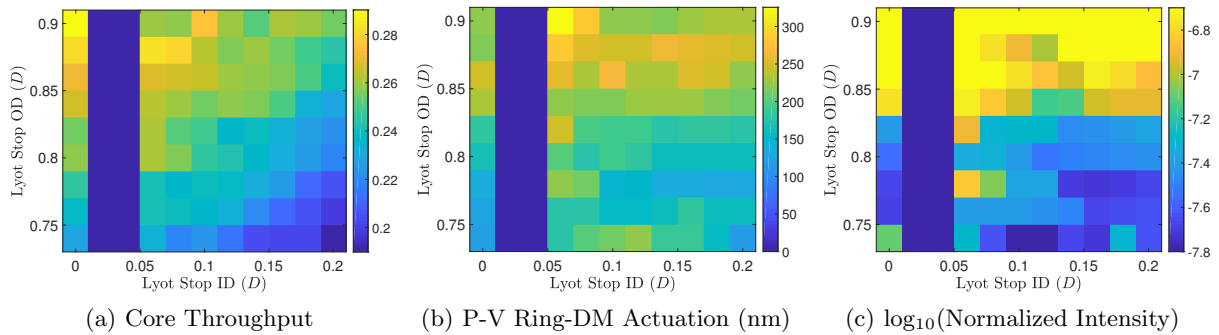


Figure 2. 1D-radial DMLC survey results for an inter-DM Fresnel number of 620 and an FPM inner radius of $2.2 \lambda_c/D$. We exclude the two columns for Lyot stop inner diameter (ID) of 2% and 4% of D because they are too small to implement in practice.

3.1.1 Comparison of 1D and 2D DMLC Results

Before switching to the hybrid part of the HLC optimization, we compare the the DMLC results computed with our 1D-radial code and the 2D-Cartesian FALCO software. As shown in Fig. 3, the DM shapes are vastly different but the intensity profiles in the final image mostly overlap. Other than differing in morphology, the 1D-radial, ring-DM shapes in Fig. 3(a) have almost twice the P-V surface deformation of the azimuthally-averaged 2D DM shapes. (The DM surfaces in 2D-Cartesian were overwhelmingly circularly symmetric.) We believe this supports our hypothesis—that our current 1D-radial ring actuator basis set has fewer independent radial modes than the actual 2D DM on a square actuator grid—because the 1D-radial DM has to work harder to achieve the same result. In the end, the important aspect to keep in mind is that the FPM is the hardest part to optimize, so it will be reasonable for us to optimize the FPM with the incorrect DM shapes in 1D as long as there exist real, 2D DM shapes that provide the same performance (or better if the P-V DM stroke is lower). For now we operate on this assumption and move forward with the HLC optimization.

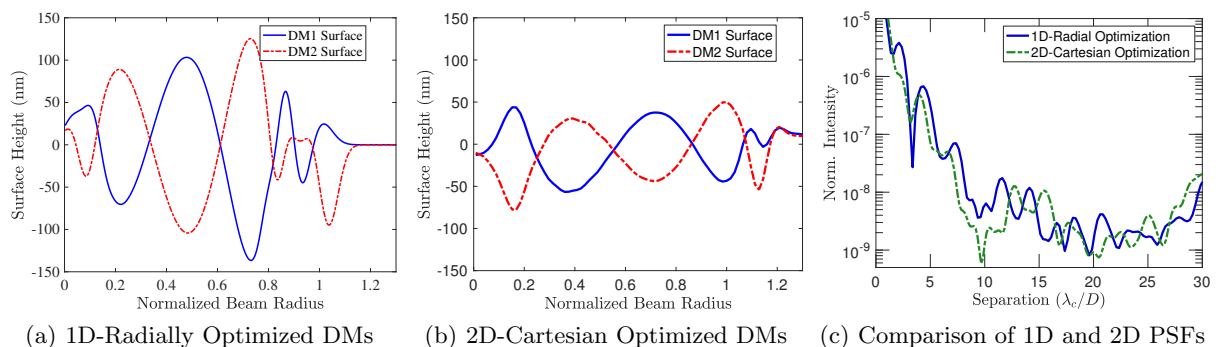


Figure 3. Comparison of DMLC results in 1D-radial and 2D-Cartesian for the point design with inter-DM Fresnel number = 620, FPM inner radius = $2.2 \lambda_c/D$, Lyot stop ID = $0.14D$, Lyot stop OD = $0.78D$. The 2D result was computed with FALCO. The 2D DM shapes have different morphology and approximately half the P-V stroke, but the intensity curves in the final focal plane are nearly the same.

3.2 Hybrid Lyot Coronagraph (HLC)

3.2.1 Material Properties of FPM

Starting from the DMFC design result, the HLC optimization concurrently optimizes the FPM and the DMs. The HLC FPMs developed for the WFIRST CGI have one or more layers of nickel for the central spot and a layer of profiled dielectric, usually PMGI, that covers the nickel and can extend far past it.⁶ For our 2D-Cartesian optimization of the simplest HLCs with FALCO, described by Sidick et al.,²² we found only a minor degradation in performance when moving from a first-order, simplified model of nickel and PMGI to using the actual physics with thin-film effects. Therefore, for these initial 1D-radial HLC investigations, we start with the same simplified model in which the nickel provides achromatic amplitude-only control and the PMGI is purely phase-shifting and provides the same phase delay (in meters) at all wavelengths. As shown in Fig. 4, these assumptions are quite accurate when looking varying only one material without the other. Small—but not ignorable—errors arise when nickel and PMGI layers are stacked because of thin film interference effects.

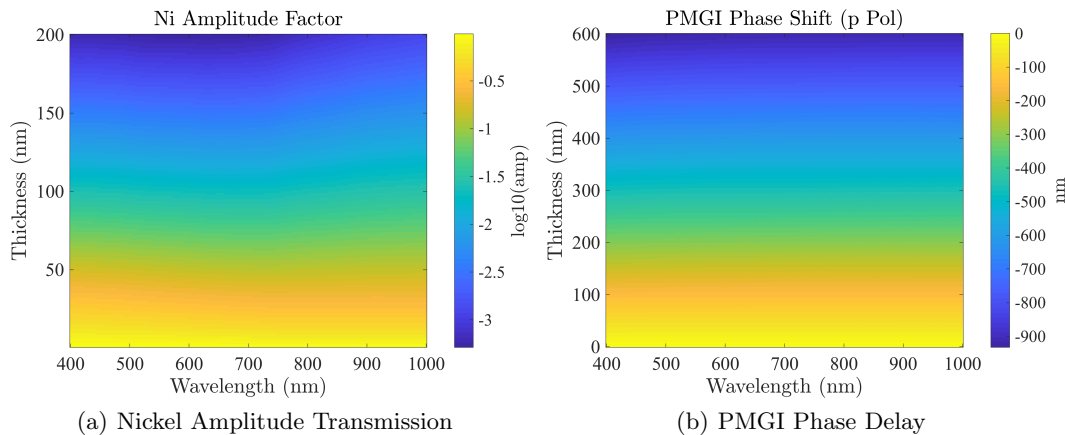


Figure 4. Material properties of the HLC's focal plane mask. The amplitude decreases linearly with the log of the nickel thickness and the phase delay increases linearly with PMGI thickness in the visible and near-infrared. When the nickel and PMGI are stacked, the linear model assumptions degrade somewhat from interference effects between the layers.

3.2.2 Initial 1D-Radial HLC Results

Fig. 5 shows the results of the first 1D-radial HLC optimization. For this case, there were no low-order aberrations included. The average normalized intensity in the dark hole is shown for each of the four design steps in Fig. 5(a). In Step 1, only the DMs were used to dig a dark hole out to the maximum correctable separation, about $29 \lambda_c/D$. In the following stages, the dark hole was reduced to have an outer working angle of $15 \lambda_c/D$, which explains the blip at the 11th control iteration. In the final three stages, we eased into letting the FPM change. In Step 2 the DMs and FPM amplitude could change for two iterations, and in Step 3 the DMs and FPM phase could change for three iterations. Finally, in Step 4 the DMs and both the FPM's amplitude and phase could vary for another fifteen control iterations. The radial profile of the final 10% bandwidth PSF in Fig. 5(b) is below the 10^{-10} targeted ceiling except near the IWA. That part of the intensity profile can easily be brought below the intensity target by applying spatial weighting to that region in the controller.

The FPM amplitude and phase profiles underwent significant changes as shown in Figs. 5(c) and 5(d). The FPM amplitude softened from a hard-edge disk to a decaying, ringing profile. The FPM phase, starting from flat, clearly changed to apply slightly opposite phases to each of the rings in the amplitude profile.

The amplitude and phase control of the starlight between the IWA and OWA on the FPM are critical for obtaining LOA robustness because low-order Zernike aberrations still have non-negligible amounts of energy in the mid-spatial frequencies. However, there is a danger of reducing coronagraphic throughput if the FPM amplitude drops too low or the the phase oscillations are too large. The phase oscillations within the dark hole in Fig. 5(d) are not worrying, though, because they are smaller than $\lambda_c/10$ P-V over the width of a PSF core. In any case, it is important in these initial studies not to over-constrain the problem with artificial bounds on FPM amplitude and phase in order to observe and interpret the naturally occurring tradeoffs in performance.

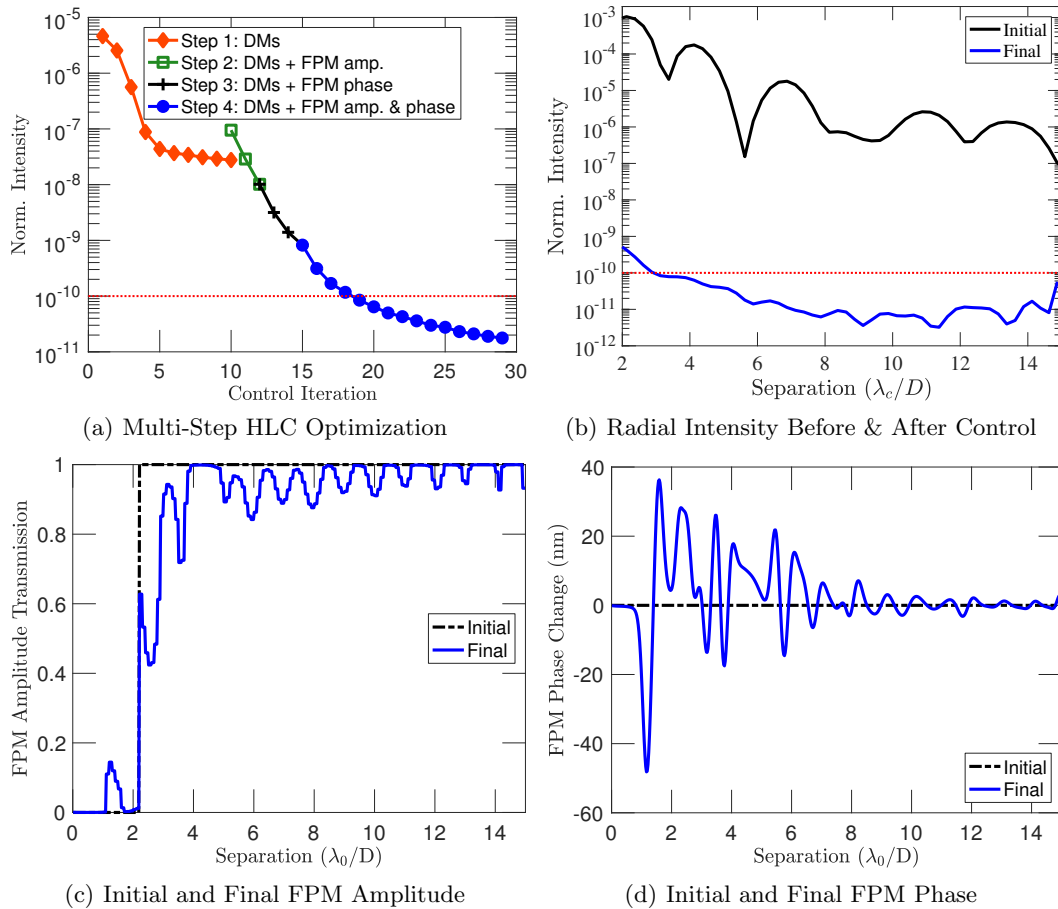


Figure 5. Results of the first 1D-radial HLC design run. Step 1 of the optimization process is the same as that performed in Section 3.1. The series of design steps enable the HLC to reach its targeted contrast goal. The final FPM wants to have a ringing amplitude profile with a slightly opposite phase applied to neighboring amplitude oscillations.

3.2.3 1D-Radial HLC Optimization with Low-Order Aberrations

The goal with including Zernike modes in the control Jacobian is to obtain directly the desired levels of LOA robustness. To start with, we choose values similar to those that the VVC with charge 4 or 6 is resistant against. Ruane et al.⁴ list these values in their Table 1. For our test, we used 10nm for tip/tilt, 1.0nm for defocus, and 1.0nm for astigmatism. Fig. 6 shows the resulting FPM amplitudes and average intensity curves as each mode is added one at a time. When the controller has difficulty suppressing the specified amount of the next Zernike mode, the FPM amplitude features become less sharp and the achievable average intensity worsens. In future work, we will need to characterize these tradeoffs and consider the intensity vs. separation curves as well.

4. 2D HLC DESIGN

For HabEx, the majority of the HLC design work can and should be performed in 1D-radial coordinates. The extreme time-savings in 1D and the directly usable mask optimizations make the decision clear. The place for 2D-Cartesian work with HabEx is in analysis and verification of 1D results. For example, DM shapes for a 1D-optimized FPM should be re-computed in 2D to see if performance value do not degrade, and the performance metrics (e.g., LOA robustness and throughput) should all be verified in 2D.

4.1 Segmented Apertures

Although the HabEx concept is currently unobscured and monolithic, it is important for us to consider how to generalize our HLC design process to mitigate pupil obscurations. A circular central obscuration would pose

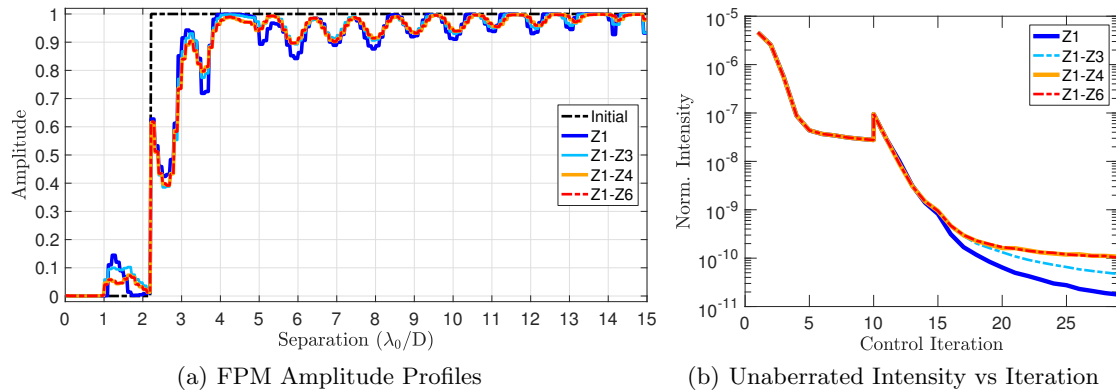


Figure 6. Changes to the HLC as the controller is asked to suppress additional Zernike aberrations up to Z6. The included LOA amounts were 10nm of tip/tilt, 1.0nm of defocus, and 1.0nm of astigmatism. When the controller has difficulty suppressing the specified amount of the next Zernike mode, the FPM amplitude features become less sharp and the achievable average intensity worsens.

no additional difficulty to our 1D-radial code. Diffraction from segment gaps and/or struts can be handled suppressed with diffractive pupil remapping solutions computed by FALCO. Ruane et al.²³ investigate the maximum segment gap size for which diffractive pupil remapping is a viable option.

Before the “HabEx Lite”²⁴ concept settled on a segmented, on-axis telescope paired with only a starshade, we analyzed another option having a segmented, off-axis architecture. As shown in Fig. 7, we used FALCO to compute small ($\approx 30\text{nm}$ P-V) DM surface shapes to apodize the segment gaps along with the baselined charge 6 VVC. The maximum achievable OWA was about two-thirds what it would be with a monolithic aperture, but the throughput was identical.

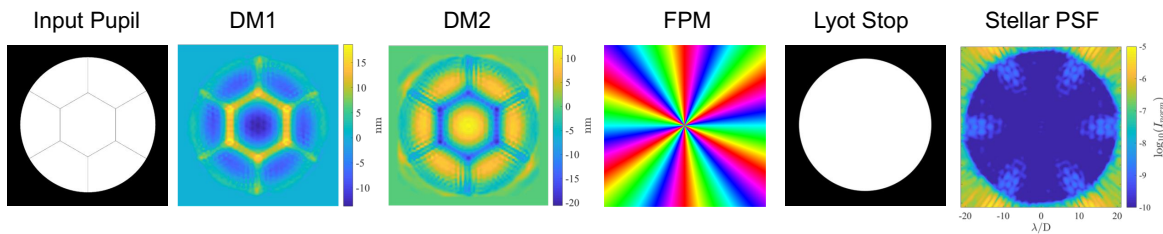


Figure 7. DM-augmented VVC computed with FALCO for a segmented, circular aperture. The outer working angle had to be reduced, but the core throughput was the same as for the monolithic primary mirror case.

5. FUTURE WORK

There are a number of remaining steps to complete our design pipeline. The three main areas of development are the software for coronagraph optimization, design evaluation, and science yield modeling.

Within the optimization, one area of concern is the 1D-radial DM model. We should determine a different radial basis set that better matches the circularly symmetric modes produced by a square grid of DM actuators. Until that is performed, we will need to verify the LOA robustness and contrast performance in 2D to be sure that a 1D-optimized design is viable.

The thin film effects of the HLC’s FPM materials need to be added in the 1D-radial optimization software. That code already exists and was used by Sidick et al.²² for 2D HLC design, so it will be straightforward to incorporate into the 1D-radial HLC design code.

Another unfinished part of the design work is the tuning of controller settings. We can experiment with adjusting the relative weights, such as for the DMs versus the FPM amplitude and phase or for the pixels in the

dark hole. A different option is to use CVX to apply more explicit constraints to ensure that critical performance values are met. For example, we could directly constrain the contrast at the IWA and the contrast degradation from LOAs.

Evaluation software in 2D is the next priority in order to analyze the performance of any given design. After the 1D optimization, there needs to be standardized software to convert 1D-radial masks in to 2D-Cartesian maps usable by FALCO. Riggs et al.¹⁴ are adding scripts to FALCO for evaluating the LOA robustness and other common performance metrics.

Finally, the science yield calculation methods performed for the VVC need to be repeated with the HLC. We will ideally use the science yield results from hundreds or thousands of 1D-radial designs to characterize performance tradeoffs and indicate which part of parameter space to explore next.

6. SUMMARY

In this paper, we presented the start of an HLC design and analysis pipeline for the HabEx mission concept. The HLC is the backup coronagraph architecture for HabEx because of its technology development heritage for the WFIRST CGI. Making the most of the circular, monolithic, unobscured aperture of HabEx, our optimization code largely uses 1D-radial coordinates to sharply reduce the computational time and complexity of the problem. A fast, 1D-radial design survey may be the only tractable method of exploring the vast HLC parameter space. The 1D-radial formulation of the coronagraph optimization also allows us to constrain the HLC to work with expected levels of low-order Zernikes from observatory dynamics and polarization aberrations. Further improvements to our 1D-radial optimization code and the FALCO package will complete our software pipeline and enable us to provide fine-tuned HLC designs for the HabEx concept and other observatories.

Acknowledgements

The HabEx Concept design is pre-decisional information presented for planning and discussion purposes only. This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

- [1] Kuan, G. M., Warfield, K., Mennesson, B., Kiessling, A., Martin, S., Nissen, J. A., Shaklan, S. B., Alvarez-Salazar, O. S., Mandic, M., Webb, D. R., Stahl, H. P., and Warwick, S., “Overview of the 4m baseline architecture concept of the habitable exoplanet imaging mission (habex) study,” (2018).
- [2] Pueyo, L. and Kasdin, N. J., “Polychromatic compensation of propagated aberrations for high-contrast imaging,” *ApJ* **666**, 609–625 (Sept. 2007).
- [3] Morgan, R., Warfield, K., Kuan, G., Stahl, H. P., Mennesson, B., Balasubramanian, K., Mawet, D., Nikzad, S., Nissen, J., Serabyn, E., Shaklan, S., Stapelfeldt, K., and Warwick, S., “Technology maturity for the habitable-zone exoplanet imaging observatory (habex) concept,” (2018).
- [4] Ruane, G., Mawet, D., Mennesson, B., Jewell, J. B., and Shaklan, S. B., “Vortex coronagraphs for the Habitable Exoplanet Imaging Mission concept: theoretical performance and telescope requirements,” *Journal of Astronomical Telescopes, Instruments, and Systems* **4**, 4 – 4 – 14 (2018).
- [5] Trauger, J., Moody, D., Gordon, B., Krist, J., and Mawet, D., “A hybrid lyot coronagraph for the direct imaging and spectroscopy of exoplanet systems: recent results and prospects,” in [*Proceedings of SPIE*], **8151**, 81510G, SPIE (2011).
- [6] Trauger, J., Moody, D., Krist, J., and Gordon, B., “Hybrid Lyot coronagraph for WFIRST-AFTA: coronagraph design and performance metrics,” *Journal of Astronomical Telescopes, Instruments, and Systems* **2**(1), 011013 (2016).
- [7] Seo, B.-J., Gordon, B., Kern, B., Kuhnert, A., Moody, D., Muller, R., Poberezhskiy, I., Trauger, J., and Wilson, D., “Hybrid Lyot coronagraph for wide-field infrared survey telescope-astrophysics focused telescope assets: occulter fabrication and high contrast narrowband testbed demonstration,” *Journal of Astronomical Telescopes, Instruments, and Systems* **2**(1), 011019 (2016).

- [8] Krist, J., Nemati, B., and Mennesson, B., “Numerical modeling of the proposed WFIRST-AFTA coronagraphs and their predicted performances,” *Journal of Astronomical Telescopes, Instruments, and Systems* **2**(1), 011003 (2015).
- [9] Seo, B.-J., Cady, E., et al., “Hybrid Lyot coronagraph for WFIRST: high-contrast broadband testbed demonstration,” in [*Proc. of SPIE*], **10400**, 15 (2017).
- [10] Krist, J., Effinger, R., Kern, B., Mandic, M., McGuire, J., Moody, D., Morrissey, P., Poberezhskiy, I., Riggs, A. J., Saini, N., Sidick, E., Tang, H., and Trauger, J., “Wfirst coronagraph flight performance modeling,” (2018).
- [11] Shi, F. et al., “Milestone #9 report: Occulting mask coronagraph dynamic broadband demonstration,” tech. rep., Jet Propulsion Laboratory (November 2016).
- [12] Guizar-Sicairos, M. and Gutiérrez-Vega, J. C., “Computation of quasi-discrete hankel transforms of integer order for propagating optical wave fields,” *J. Opt. Soc. Am. A* **21**, 53–58 (Jan 2004).
- [13] Watson, G. N., [*A Treatise on the Theory of Bessel Functions*], Cambridge University Press, 2nd ed. (1966).
- [14] Riggs, A. J. E., Ruane, G. J., Sidick, E., Coker, C. T., and Kern, B. D., “Fast Linearized Coronagraph Optimizer (FALCO) I. a software toolbox for rapid coronagraphic design and wavefront correction,” in [*Proc. of SPIE*], **10698** (2018).
- [15] Give’on, A., Kern, B., Shaklan, S., Moody, D., and Pueyo, L., “Broadband wavefront correction algorithm for high-contrast imaging systems,” *Proceedings of SPIE* **6691**, 66910A–1 – 66910A–11 (2007).
- [16] Pueyo, L., Kay, J., Kasdin, N., Groff, T., McElwain, M., Give’on, A., and Belikov, R., “Optimal dark hole generation via two deformable mirrors with stroke minimization,” *Applied Optics* **48**(32), 6296–6312 (2009).
- [17] Grant, M. and Boyd, S., “CVX: Matlab software for disciplined convex programming, version 2.1,” (Mar. 2014).
- [18] Grant, M. and Boyd, S., “Graph implementations for nonsmooth convex programs,” in [*Recent Advances in Learning and Control*], Blondel, V., Boyd, S., and Kimura, H., eds., *Lecture Notes in Control and Information Sciences*, 95–110, Springer-Verlag Limited (2008).
- [19] Mazoyer, J., Pueyo, L., N’Diaye, M., Fogarty, K., Zimmerman, N., Leboulleux, L., Laurent, K. E. S., Soummer, R., Shaklan, S., and Norman, C., “Active Correction of Aperture Discontinuities-Optimized Stroke Minimization. I. a new adaptive interaction matrix algorithm,” *ApJ* **155**(1), 7 (2018).
- [20] Mazoyer, J., Pueyo, L., N’Diaye, M., Fogarty, K., Zimmerman, N., Soummer, R., Shaklan, S., and Norman, C., “Active Correction of Aperture Discontinuities-Optimized Stroke Minimization. II. optimization for future missions,” *ApJ* **155**(1), 8 (2018).
- [21] Coker, C. T., Ruane, G. J., Riggs, A. J. E., et al., “Fast Linearized Coronagraph Optimizer (FALCO) III. optimization of key coronagraph design parameters,” in [*Proc. of SPIE*], **10698** (2018).
- [22] Sidick, E., Riggs, A. J. E., Ruane, G. J., Krist, J. E., Moody, D. C., and Coker, C. T., “Fast linearized coronagraph optimizer (FALCO) II: optical model validation and time savings over other methods,” (2018).
- [23] Ruane, G., Riggs, A. J. E., Coker, C. T., et al., “Fast Linearized Coronagraph Optimizer (FALCO) IV. coronagraph design survey for obstructed and segmented apertures,” in [*Proc. of SPIE*], **10698** (2018).
- [24] Redding, D., Coste, K., Polanco, O., Pineda, C., Hurd, K., et al., “HabEx Lite: a starshade-only habitable exoplanet imager alternative,” (2018).