

# Journal of Astronomical Telescopes, Instruments, and Systems

AstronomicalTelescopes.SPIEDigitalLibrary.org

## **Shaped pupil Lyot coronagraphs: high-contrast solutions for restricted focal planes**

Neil T. Zimmerman  
A. J. Eldorado Riggs  
N. Jeremy Kasdin  
Alexis Carlotti  
Robert J. Vanderbei

# Shaped pupil Lyot coronagraphs: high-contrast solutions for restricted focal planes

Neil T. Zimmerman,<sup>a,\*</sup> A. J. Eldorado Riggs,<sup>a</sup> N. Jeremy Kasdin,<sup>a</sup> Alexis Carlotti,<sup>b</sup> and Robert J. Vanderbei<sup>c</sup>

<sup>a</sup>Princeton University, Department of Mechanical and Aerospace Engineering, Engineering Quadrangle, Olden Street, Princeton, New Jersey 08544, United States

<sup>b</sup>Centre National de la Recherche Scientifique, Institut de Planétologie et d'Astrophysique de Grenoble, BP 53, F-38041, Grenoble, Cédex 9, France

<sup>c</sup>Princeton University, Department of Operations Research and Financial Engineering, Sherrerd Hall, Charlton Street, Princeton, New Jersey 08544, United States

**Abstract.** Coronagraphs of the apodized pupil and shaped pupil varieties use the Fraunhofer diffraction properties of amplitude masks to create regions of high contrast in the vicinity of a target star. Here we present a hybrid coronagraph architecture in which a binary, hard-edged shaped pupil mask replaces the gray, smooth apodizer of the apodized pupil Lyot coronagraph (APLC). For any contrast and bandwidth goal in this configuration, as long as the prescribed region of contrast is restricted to a finite area in the image, a shaped pupil is the apodizer with the highest transmission. We relate the starlight cancellation mechanism to that of the conventional APLC. We introduce a new class of solutions in which the amplitude profile of the Lyot stop, instead of being fixed as a padded replica of the telescope aperture, is jointly optimized with the apodizer. Finally, we describe shaped pupil Lyot coronagraph (SPLC) designs for the baseline architecture of the Wide-Field Infrared Survey Telescope–Astrophysics Focused Telescope Assets (WFIRST-AFTA) coronagraph. These SPLCs help to enable two scientific objectives of the WFIRST-AFTA mission: (1) broadband spectroscopy to characterize exoplanet atmospheres in reflected starlight and (2) debris disk imaging. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.JATIS.2.1.011012]

Keywords: coronagraph; high-contrast imaging; space telescope; exoplanets.

Paper 15053SS received Jun. 16, 2015; accepted for publication Nov. 23, 2015; published online Jan. 14, 2016.

## 1 Introduction

The last two decades have witnessed tremendous advances in high-contrast imaging technology in tandem with the emergence of exoplanet research. There is now a mature and growing assortment of instrument concepts devised to isolate the light of an exoplanet from its host star and acquire its spectrum. Stellar coronagraphs descended from Bernard Lyot's invention represent a major component of this effort, complementing and intersecting the innovations in interferometry, adaptive optics, wavefront control, and data processing. With these tools in place, several exoplanet imaging programs at large ground-based observatories are underway.<sup>1–3</sup> Their observations have led to discoveries and astrophysical measurements that are steering theories of planet formation, planetary system evolution, and planetary atmospheres.<sup>4–7</sup> Meanwhile, laboratory testbeds are setting the stage for yet more ambitious instruments on new space telescopes.<sup>8–11</sup>

A coronagraph alters the point spread function (PSF) of a telescope so that a region of the image normally dominated by starlight is darkened by destructive interference, enabling observations of faint surrounding structures and companions. Starlight cancellation is accomplished with a group of optical elements that operate on the complex field of the propagating beam. The classical Lyot coronagraph functions with a pair of simple masks: one an opaque occulting spot at the focus and second a Lyot stop to block the outer edge of the recollimated

on-axis beam before it is reimaged.<sup>12</sup> To take advantage of the diffraction-limited imaging capabilities of high-order adaptive optics (AO) systems, beginning in the 1990s, classical Lyot designs were revised for high-contrast stellar coronagraphy.<sup>13–18</sup> Through Fourier optical analysis and modeling, researchers soon discovered the remarkable performance benefits of apodizing the entrance pupil of a coronagraph.<sup>19–21</sup> Since then, the transmission profile of this apodizer mask has been a topic of vigorous study.<sup>22–27</sup> One resulting family of designs, the apodized pupil Lyot coronagraph (APLC), has been successfully integrated with several AO-fed cameras to facilitate deep observations of young exoplanetary systems at near-infrared wavelengths.

As an alternative to a coronagraph with two or more mask planes, pupil apodization by itself is perhaps the simplest and oldest way to reject unwanted starlight from a telescope image.<sup>28,29</sup> Fraunhofer diffraction theory dictates how any change in the shape or transmission profile of the entrance pupil redistributes a star's energy in the image plane. This relationship can be used to design an apodizer whose PSF has a zone of high contrast near the star without additional coronagraph masks. This is the shaped pupil approach developed by Kasdin and collaborators, who pioneered the optimization of apodizers with binary-valued transmission.<sup>30–33</sup> In recent years, shaped pupil solutions have evolved to work around arbitrary two-dimensional telescope apertures, in parallel with similar breakthroughs in APLC design.<sup>34–39</sup> The relative simplicity of a single mask, however, comes with a sacrifice in how close the dark search region can be pushed toward the

\*Address all correspondence to: Neil T. Zimmerman, E-mail: [ntz@stsci.edu](mailto:ntz@stsci.edu)

star. At the contrast levels relevant to exoplanet imaging, the smallest feasible shaped pupil inner working angle (IWA) is between 3 and  $4\lambda/D$ .<sup>33</sup>

Shaped pupil coronagraphs (SPCs) and Lyot coronagraphs (both classical and APLC) both rely on masks that operate strictly on the transmitted amplitude of the propagating beam. Numerous coronagraph designs have been introduced that incorporate phase masks<sup>40–42</sup> and pupil remapping via aspheric mirrors and/or static deformations.<sup>43,44</sup> In general, coronagraphs that manipulate phase in addition to amplitude can achieve higher performance in terms of IWA and throughput than SPCs and APLCs. For example, the vector vortex coronagraph has a theoretical inner working limit of  $\sim\lambda/D$  separation from a star.<sup>45</sup> Recent theoretical innovations have improved the compatibility of phase-mask and pupil-remapping coronagraph concepts with segmented and obstructed telescope apertures.<sup>46–49</sup> For broad comparisons between coronagraph design families, see Refs. 9, 50, 51, 52.

Until the past two years, all SPC testbed experiments with wavefront control used freestanding shaped pupil designs with connected obstruction patterns. In particular, experiments in Princeton’s High Contrast Imaging Laboratory<sup>53</sup> and the High Contrast Imaging Testbed (HCIT)<sup>54</sup> at the Jet Propulsion Laboratory (JPL) have used ripple-style SPC masks along with two deformable mirrors in series.<sup>55–57</sup> In this issue, Cady et al.<sup>58</sup> report the first experimental results with a nonfreestanding SPC design. This mask was fabricated on a silicon wafer substrate with aluminized reflective regions and highly absorptive black silicon regions; the fabrication process is described in detail in this issue by Balasubramanian et al.<sup>59</sup> Cady et al. used a single deformable mirror in their experiments to create a single-sided dark hole from 4.4 to  $11\lambda/D$  in a 52-deg wedge.<sup>58</sup> They achieved  $5.9 \times 10^{-9}$  contrast in a 2% bandwidth about 550 nm and  $9.1 \times 10^{-8}$  contrast in a 10% bandwidth about 550 nm.

The Science Definition Team of NASA’s Wide-Field Infrared Survey Telescope–Astrophysics Focused Telescope Assets (WFIRST-AFTA) mission has proposed including a coronagraph instrument (CGI) to observe super-Earth and gas-giant exoplanets in reflected starlight at visible wavelengths.<sup>60</sup> Coronagraph designs for WFIRST-AFTA must be compatible with its heavily obscured telescope aperture, broad filter bands, and rapid development timeline. The SPC, recognized to match these demands, was selected as one of the two baseline coronagraph technologies to undergo extensive testing at JPL in advance of the mission formulation.<sup>61</sup> The method under development in parallel with the SPC is the hybrid Lyot coronagraph (HLC).<sup>62,63</sup> The HLC departs from the classical Lyot approach by using a focal plane mask (FPM) with a complex transmission profile.<sup>64</sup> The SPC and HLC can share the same optical path and wavefront control system. A third coronagraph type, the phase-induced amplitude apodization complex mask coronagraph, is being pursued in parallel as a backup option.<sup>48,65–67</sup>

In the course of our efforts to improve the SPC designs already meeting the minimum performance goals, we investigate a hybrid coronagraph architecture in which a binary shaped pupil functions as the apodizer mask in an APLC-like configuration.<sup>68,69</sup> In effect, this expands on the idea first put forward by Cady et al.,<sup>70</sup> who designed a hard-edged, star-shaped apodizer for the Gemini Planet Imager’s APLC. We identify this design category as the shaped pupil Lyot coronagraph (SPLC). The SPLC offers a persuasive union of the virtues of SPC and APLC: a binary apodizer with achromatic transmission properties

and promising fabrication avenues,<sup>59,71,72</sup> and the relatively small IWA and robustness to aberrations of an APLC.<sup>73</sup>

## 2 Lyot Coronagraphy with an Unobscured Circular Aperture

Although coronagraph designs for obscured apertures are the ones of highest practical interest and relevance to WFIRST-AFTA and the general community, a clear circular telescope aperture offers a natural starting point to understand how the SPLC relates functionally to the conventional APLC. For example, circular symmetry simplifies the analytical formulation, the numerical optimization problem, and the interpretation. The same qualitative relationships that occur for a simple aperture will reappear for more complicated cases [e.g., SP apodizer feature size and outer working angle (OWA)]. Furthermore, the clear circular aperture allows us to probe the limitations of pure amplitude Lyot coronagraphy, offering useful insights for exoplanet imaging mission design studies such as the recent Exo-C.<sup>74</sup>

For each of our numerical SPLC experiments, we consider two forms of FPM, illustrated in Fig. 1: first, the occulting spot of the conventional APLC, with radius  $\rho_0$ ; and second, an annular diaphragm with inner radius  $\rho_0$  and outer radius  $\rho_1$ . In our descriptions of the on-axis field propagation, we will make use of the complement of the FPM transmission function. For the spot and diaphragm FPM cases, we label these as  $M_a$  and  $M_b$ , respectively. In terms of the radial spatial coordinate,  $\rho$ , they are defined as

$$M_a(\rho) = \Pi(\rho/(2\rho_0)),$$

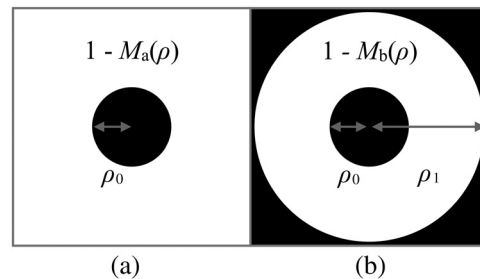
$$M_b(\rho) = 1 - \Pi(\rho/(2\rho_1)) + \Pi(\rho/(2\rho_0)). \quad (1)$$

Here,  $\Pi(x)$  is the rectangle function, equal to unity inside  $|x| < 1/2$  and zero elsewhere.

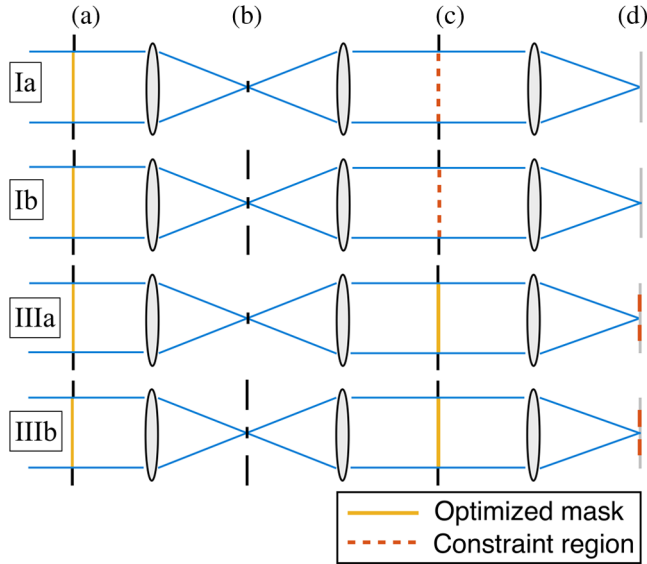
To find apodizer solutions for the circular SPLC, we use the same numerical optimization tools previously applied to shaped pupil mask designs. In addition to the two types of FPM above, we consider two different planes of field cancellation constraints, as diagrammed in Fig. 2. The results of all the circular SPLC trials are later summarized in Table 1. Details about our optimization method, including discrete algebraic models for the on-axis field propagation and definitions of the linear program objectives and constraints, are given in Appendix A1.

### 2.1 Focal Occulting Spot

Loosely following the nomenclature that Soummer et al. formulated in Ref. 21, we represent the scalar electric field in the



**Fig. 1** The two types of focal plane mask (FPM) considered for the circular aperture Lyot coronagraph: (a) occulting spot and (b) annular diaphragm.



**Fig. 2** A representative subset of the optimization schemes we apply to the circular aperture Lyot coronagraph. From left to right, the critical mask planes are (a) the apodizer, (b) the FPM, (c) the Lyot stop, and (d) the final image. The red (dashed) regions are where the on-axis field is constrained. The gold (solid) segments mark the masks that are optimized as free variables.

entrance pupil, focal plane, and Lyot plane, respectively, by  $\Psi_A$ ,  $\Psi_B$ , and  $\Psi_C$ . In a slight departure, we define the focal plane radial coordinate  $\rho$  in units of image resolution elements ( $f\lambda/D$ ) and the radial coordinate in the two conjugate pupil planes as  $r$ , normalized to the aperture diameter  $D$ . For brevity, we implicitly apply the pupil cutoff function in all instances of  $\Psi_A$  and  $\Psi_C$ , and we set  $D$  to 1. These provisions allow us to succinctly

express the on-axis scalar electric field in the Lyot plane after the occulting spot, in accordance with the Babinet principle:

$$\begin{aligned}\Psi_C(r) &= \Psi_A(r) - [\widehat{\Psi}_A(\rho)\widehat{M}_a(\rho)] \\ &= \Psi_A(r) - (2\pi)^2 \int_{\rho=0}^{\rho_0} \int_{r'=0}^{1/2} \Psi_A(r') J_0(2\pi\rho r') r' dr' J_0 \\ &\quad \times (2\pi r\rho) \rho d\rho.\end{aligned}\quad (2)$$

Hatted variables denote the application of the Hankel transform, and  $J_0$  is the zero-order Bessel function of the first kind.

Setting the condition for total field cancellation,  $\Psi_C(r) = 0$ , leads to an integral equation of the variable function  $\Psi_A(r)$ . In Ref. 21, Soummer et al. showed that the approximate solutions are a subset of Slepian's circular prolate spheroidal wave functions, originally published four decades prior.<sup>75</sup> The zero-order prolate spheroidal wave functions possess two exceptional apodization properties. First, they are by definition invariant to the finite Fourier transform, so the scalar field in the focal plane after the apodizer is equal to the unrestricted prolate function itself to within a scale factor. Second, the prolate apodizer maximizes the concentration of energy in the focal plane, within a radius set by the eigenvalue  $0 < \Lambda < 1$  of the integral equation.<sup>76</sup> Therefore, once a focal plane spot radius  $\rho_0$  has been chosen, the apodizer for optimum monochromatic extinction  $\Psi_A(r) = \Phi_\Lambda(r)$  is fully determined. Invoking the finite Hankel invariance property of  $\Phi_\Lambda(r)$ , one can start from Eq. (2) and arrive at a simple expression for the residual on-axis Lyot plane electric field:

$$\Psi_C(r) = \Phi_\Lambda(r)(1 - \Lambda).\quad (3)$$

Unlike an alternate configuration in which the prolate apodizer is combined with a Roddier  $\pi$  phase-shifter,<sup>19</sup> for the opaque

**Table 1** Summary of circular aperture Lyot coronagraph trials for various mask and optimization configurations. In the Lyot stop descriptions, I.D. and O.D. stand for inner and outer diameter, respectively. In the column of notes describing each solution, SP and LS stand for shaped pupil and Lyot stop, respectively.

Config.	Focal plane mask (FPM)	Lyot stop	Field constraint	Bandwidth	Throughput	PSF area	Notes
Ia	Spot $\rho_0 = 1.87$	Fixed O.D. = 1	$ \Psi_C  \leq 10^{-3}$	Mono	0.134	1.84	APLC for $\Lambda = 0.999$
Ib	Ann $[\rho_0 = 1.87, \rho_1 = 12]$	Fixed O.D. = 1	$ \Psi_C  \leq 10^{-3}$	Mono	0.148	1.73	8-ring SP
IIa	Spot $\rho_0 = 3$	Fixed O.D. = 1	$10^{-9}$ contrast, $3 < \rho < 12$	10%	0.095	2.42	9-ring SP
IIb	Ann $[\rho_0 = 3, \rho_1 = 12]$	Fixed O.D. = 1	$10^{-9}$ contrast, $3 < \rho < 12$	10%	0.108	2.18	8-ring SP
IIIa	Spot $\rho_0 = 3$	Free mask	$10^{-9}$ contrast, $3 < \rho < 12$	10%	0.334	1.25	10-ring SP, 3-ring LS
IIIb	Ann $[\rho_0 = 3, \rho_1 = 12]$	Free mask	$10^{-9}$ contrast, $3 < \rho < 12$	10%	0.144	1.39	7-ring SP, 7-ring LS
IVa	Spot $\rho_0 = 3$	Fixed [I.D. = 0.1, O.D. = 0.9]	$10^{-9}$ contrast, $3 < \rho < 12$	10%	0.317	1.32	8-ring SP
IVb	Ann $[\rho_0 = 3, \rho_1 = 12]$	Fixed [I.D. = 0.1, O.D. = 0.9]	$10^{-9}$ contrast, $3 < \rho < 12$	10%	0.121	1.93	8-ring SP

Note: PSF, point spread function; APLC, apodized pupil Lyot coronagraph; SP, shaped pupil.



occluding spot, the monochromatic on-axis cancellation is never complete because no prolate solution corresponding to  $\Lambda = 1$  exists.<sup>21</sup> But  $\Lambda$  is already 0.999 at  $\rho_0 = 1.87\lambda_0/D$ , for example, and it can be made arbitrarily close to unity by further widening the occulting spot at the expense of the IWA.

As a first experiment, we start with the coronagraph model portrayed at the top of Fig. 2, which we label config. Ia. For the clear circular aperture, an occulting focal plane spot of radius  $\rho_0$ , and a Lyot stop with the same diameter as the aperture, we ask what entrance apodizer results in a monochromatic Lyot field cancellation factor  $1 - \Lambda$  while maximizing the overall field transmission. Our aim is to independently recover one of the canonical circular prolate apodizers presented in Ref. 21, corresponding to  $\Lambda = 0.999$  and  $\rho = 1.87$ . We form a linear program relating the discretized apodizer vector to the resulting Lyot field, using a Riemann sum representation of the Hankel transforms on the right-hand side of Eq. (2). For details about this procedure, see Appendix A1.

The resulting discrete apodizer solution array is  $\mathcal{A}_{\text{SP}}(r_i)$ , where  $r_i$  is the normalized pupil radius. We check our result by evaluating the integrated energy transmission metric originally tabulated by Soummer et al.:  $2\pi \sum_{i=1}^{N_r} r_i \mathcal{A}_{\text{SP}}^2(r_i)$ . Our integrated energy transmission is 0.193, which is in close agreement with the corresponding value of the analytical prolate solution, 0.190.<sup>21</sup> A gray-scale map of the apodizer transmission is plotted on the left-hand side of Fig. 3.

We decompose the two algebraic components of the Lyot plane field to learn how the design constraints are fulfilled. In the upper right plot of Fig. 3; the apodizer curve is drawn in blue, followed by the Hankel transform of the field inside the occulting spot in gold. Recall that the latter curve is the function subtracted in Eq. (2) to compute the resulting Lyot field. The invariance of the apodizer to the finite Hankel transform is evident by the fact that within the aperture  $r < D/2$ , the subtrahend curve is indistinguishable from the apodizer transmission. Outside  $r = D/2$ , the subtrahend remains continuous since it recovers the unrestricted prolate function. The difference of the two functions reveals a slight deviation from the analytical solution. The residual Lyot field is not shaped like the circular prolate function, as prescribed by Eq. (3). Recall, however, that

we did not specify a point-wise constraint in the Lyot plane, instead imposing a less stringent requirement that  $|\Psi_C(r)| < 1 - \Lambda$  for  $r < D/2$ .

## 2.2 Focal Diaphragm

We express the on-axis Lyot field in terms of the apodizer transmission and the FPM profile, for both FPM types. To offer a slightly more intuitive description than Eq. (2), instead of explicitly writing out the Hankel transform integrals, this time we express the Lyot plane field components in terms of pupil-plane convolutions:

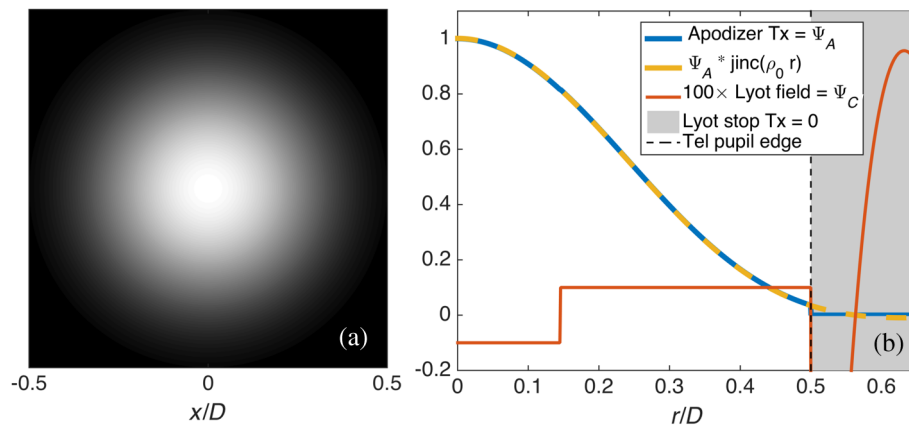
$$\begin{aligned} \Psi_{C_a}(r) &= \Psi_{A_a}(r) - [\widehat{\Psi_{A_a}(\rho)} M_a(\rho)] \\ &= \Psi_{A_a}(r) - \Psi_{A_a}(r) * \frac{\rho_0 J_1(2\pi\rho_0 r)}{r}. \end{aligned} \quad (4)$$

$$\begin{aligned} \Psi_{C_b}(r) &= \Psi_{A_b}(r) - [\widehat{\Psi_{A_b}(\rho)} M_b(\rho)] \\ &= \Psi_{A_b}(r) * \frac{\rho_1 J_1(2\pi\rho_1 r)}{r} - \Psi_{A_b}(r) * \frac{\rho_0 J_1(2\pi\rho_0 r)}{r}. \end{aligned} \quad (5)$$

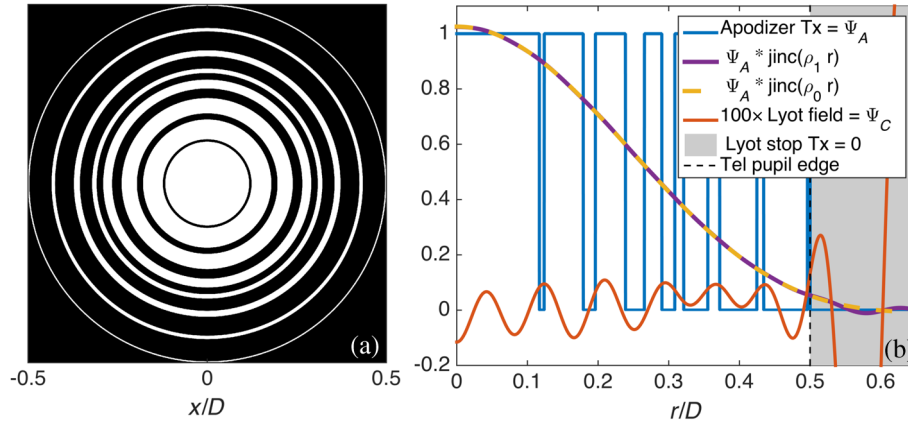
The circular analog of the sinc function,  $\text{jinc}(ar) = aJ_1(2\pi ar)/r$ , appears here ( $J_1$  is the first-order Bessel function of the first kind).<sup>77</sup> In both configurations, it serves as a low-pass filter kernel on the entrance pupil field.

Equation (5) does not yield the same form of integral equation as before, since the apodizer function appears inside an integral in both terms. Therefore, the original framework for the approximate analytical solution no longer applies. In spite of this, we show now that a similar cancellation of the on-axis Lyot field is easily achievable with the annular FPM.

For config. Ib (Fig. 4), we repeat the same problem as Ia, except that now we replace the occulting spot with the annular diaphragm [as defined in the second part of Eq. (1)]. The inner and outer radii are  $1.87\lambda_0/D$  and  $12\lambda_0/D$ , respectively. The resulting solution no longer resembles a circular prolate function, but instead a concentric ring shaped pupil mask of the kind previously described by Vanderbei et al.<sup>32</sup> By blocking the outer region of the focal plane, the solver is able to take



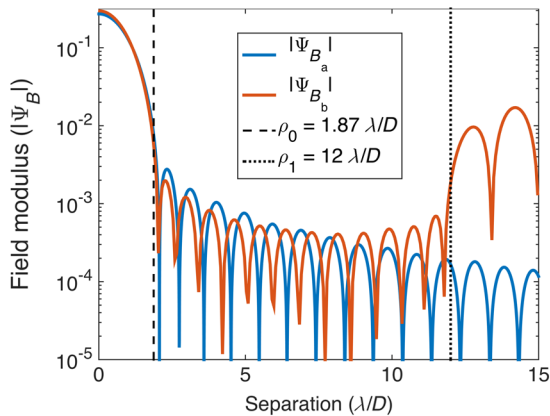
**Fig. 3** Circular Lyot coronagraph, config. Ia. The (a) apodizer is optimized for maximum transmission while achieving a monochromatic Lyot plane cancellation factor of  $10^{-3}$ , with an occulting focal plane spot of radius  $1.87\lambda_0/D$ . The Lyot stop is fixed to the diameter of the reimaged telescope pupil. In the right-hand plot (b), the algebraic components of the Lyot plane field ( $\Psi_C$ ) are compared, namely, Eq. (4). Due to the Fourier transform invariance property of the prolate-apodized field ( $\Psi_A$ ), the filtered subtrahend component [ $\Psi_A * \text{jinc}(\rho_0 r)$ ] traces the same profile within the radius of the stop, constraining the difference ( $\Psi_C$ ) near zero.



**Fig. 4** Circular Lyot coronagraph, config. Ib. The (a) apodizer is optimized for maximum transmission while achieving a monochromatic Lyot plane cancellation factor of  $10^{-3}$ , with an annular diaphragm FPM of inner radius  $1.87\lambda_0/D$  and outer radius  $12\lambda_0/D$ . The Lyot stop is fixed to the diameter of the reimaged telescope pupil. In the right-hand plot (b), the algebraic components of the Lyot plane field ( $\Psi_C$ ) are compared, namely Eq. (5). In the image domain,  $\Psi_A$  concentrates most energy within  $\rho_0$  (as shown in Fig. 5). Therefore, both low-pass filtered instances of the ring-apodized field [ $\Psi_A \times \text{jinc}(\rho_0 r)$  and  $\Psi_A \times \text{jinc}(\rho_1 r)$ ] are approximately equal, and the residual difference meets the design constraints inside the Lyot stop.

advantage of a mask whose Bessel harmonics lie outside  $\rho_1$ , because energy distributed there can no longer propagate on to the Lyot plane. The spatial frequencies of the strong Bessel harmonics of the concentric ring mask depend on the ring spacing and thickness. As a consequence, when we repeat the trial for larger  $\rho_1$ , the number of rings increases, while the overall open area decreases slightly. Conversely, when  $\rho_1$  is reduced, the apodizer solution has fewer, thicker rings and higher transmission.

The plot of the Lyot field decomposition on the right-hand side of Fig. 4 reveals another interesting result. The two components of the Lyot plane field, which we expressed before in Eq. (5) in terms of convolutions between the apodizer transmission and jinc functions, bear a striking resemblance to the original circular prolate function that appeared in config. Ia. Therefore, even though the apodizer is binary, the low-pass filter effect of the jinc convolution recovers a rough approximation of the circular prolate function for both the inner and outer



**Fig. 5** Comparison of the scalar focal plane fields for apodizer solutions Ia and Ib. The blue curve ( $\Psi_{B_a}$ ) is the on-axis field in the first focal plane after the circular prolate solution shown in Fig. 3. The red curve ( $\Psi_{B_b}$ ) is the focal plane field after the concentric ring mask apodizer of config. Ib (Fig. 4). Notably, the amplitude of  $\Psi_{B_b}$  rises immediately outside the outer radius of the annular FPM ( $\rho_1 = 12$ ), while the ripple envelope of  $\Psi_{B_a}$  continues to fall monotonically.

components. The residual ripple shows the two components are equal to within the  $10^{-3}$  field constraint specified in the design. This could only be the case for an apodizer that concentrates a great fraction of its energy within the inner edge of the annulus. We verify this characteristic in Fig. 5, where the field distributions produced in the first focal plane by the apodizers of Ia and Ib are compared. The ring apodizer (red curve) has a higher overall throughput and, therefore, a higher peak. Right outside the outer FPM edge  $\rho_1$ , the rejected high-frequency Bessel harmonics of the ring apodizer emerge and continue to oscillate beyond the plotted radius. By comparison, the ripple envelope of the circular prolate focal plane field (blue curve) decreases monotonically out to infinity.

### 2.3 Polychromatic Focal Plane Field Cancellation

While the previous trials offer a useful conceptual perspective on how binary apodizers can function in a Lyot coronagraph, we are ultimately concerned with image plane performance metrics and solutions that suppress starlight over a finite bandwidth. Therefore, the remaining trials carry the propagation to the final focal plane and constrain the contrast there to  $10^{-9}$  in a restricted region (details in Appendix A1). We define contrast here as the ratio of the intensity in the final image to the peak of the off-axis coronagraph PSF. Bandwidth is achieved by repeating the field constraints at three wavelength samples spanning a 10% fractional bandwidth.

For all the configurations, we compute the throughput and area of the coronagraph PSF, and assemble the results in Table 1. Following the convention of Krist et al.,<sup>78</sup> throughput takes into account the overall proportion of energy from an off-axis (planet-like) point source that reaches the final image, as well as the proportion of that energy concentrated in the main lobe of the corresponding PSF. We make the assumption that only in the main lobe of the off-axis PSF is the intensity high enough to generate a useful signal. We compute the throughput by propagating an off-axis plane wave through the coronagraph model, masking off the full-width half-maximum (FWHM) region of the resulting PSF and summing the intensity there. Then we repeat the same calculation when the off-axis source

is directly imaged by the telescope without a coronagraph. The ratio of these intensity sums gives a normalized metric indicating how efficiently off-axis point sources are preserved by the coronagraph. For a Lyot coronagraph (including classical, APLC, and SPLC), throughput is approximately constant over the field of view (FoV), as long as the off-axis PSF core clears line-of-sight obstruction by the FPM.

Independent of throughput, we also assess how tightly the energy is concentrated in the central lobe of the off-axis PSF based on the area of the FWHM region. A small PSF area is desirable, because for a given throughput value, a smaller area results in a higher peak signal on the detector. We again normalize this to the reference case of a PSF without a coronagraph. Since these designs have an unobstructed circular pupil, the reference telescope PSF is an Airy disk.

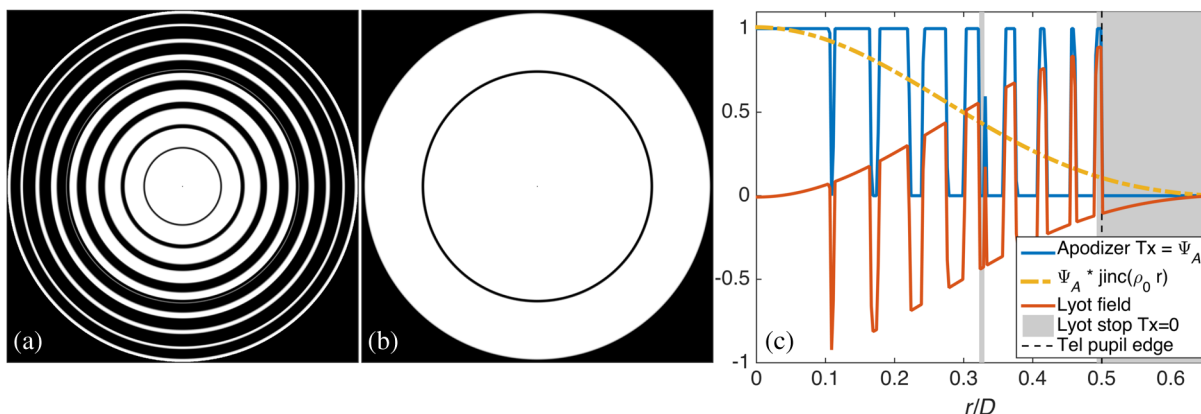
To enable the most meaningful comparison across the various configurations in the table, we relaxed the IWA from  $1.87\lambda_0/D$  to  $3\lambda_0/D$ , where  $\lambda_0$  is the center wavelength of the passband. This increase in the inner edge is needed because for some configurations, we failed to find any polychromatic solutions for  $\rho_0$  of  $2\lambda_0/D$  or below. The outer edge of the high-contrast region is arbitrarily fixed at  $12\lambda_0/D$  for all designs.

The first set of trials with image plane constraints are configs. IIa and IIb, with a fixed Lyot stop again matched to the telescope aperture. Following the previous nomenclature, type ‘‘a’’ designs use the occulting spot FPM, and type ‘‘b’’ designs use the annular diaphragm FPM. For both types of FPM, the apodizer with the highest throughput is a concentric ring shaped pupil (Table 1). Even for the spot FPM, the hard outer edge of our specified dark region means that the strong Bessel harmonics of the ring apodizer are tolerated outside  $\rho_1 = 12\lambda_0/D$ . If we had constrained the contrast out to an infinite radius from the star, we would instead expect the solution to revert to a smooth apodizer with a continuous derivative. More practically, we could have included derivative constraints in the optimization program.<sup>33,79</sup> In a recent APLC design study, for example, the contrast constraints were also imposed over a restricted area of the final image.<sup>27</sup> The authors, requiring a smooth apodizer transmission profile, added constraints on the spatial derivative of the apodizer in order to avoid binary solutions.

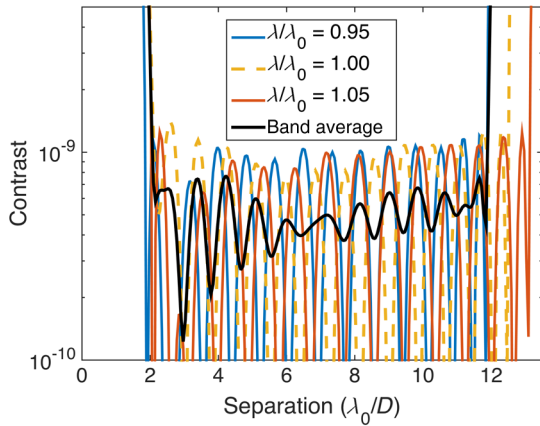
## 2.4 Joint Optimization of the Apodizer and Lyot Stop

For the conventional monochromatic APLC, the optimal Lyot stop is one exactly matched to the telescope aperture (after a 180 deg rotation, for telescope apertures lacking circular symmetry).<sup>25</sup> The Lyot stop is padded only for the purpose of alignment tolerance.<sup>26</sup> However, recent investigations have shown that APLC optimizations incorporating bandwidth and image constraints yield better results when the Lyot stop’s central obstruction replica is significantly oversized.<sup>27</sup> For example, in the course of optimizing an APLC for an aperture with central obstruction of diameter  $0.14D$ , aiming for contrast  $10^{-8}$  over a 10% bandwidth, N’Diaye et al. found that increasing the inner diameter of the Lyot stop to  $\sim 0.35D$  enabled the IWA to be reduced from  $3.7\lambda_0/D$  to  $2.4\lambda_0/D$  at the same throughput.<sup>27</sup> Evidently, the transmission profile of the Lyot stop offers an important parameter space to survey in addition to the apodizer.

Building on this notion, for configs. IIIa and IIIb, we recast the Lyot coronagraph optimization as a nonlinear program in which both the apodizer and the Lyot stop transmission profile are free vectors, optimized simultaneously. See Appendix A1 for further description of this procedure. The program seeks to maximize the sum of the transmission of both apodizer and Lyot stop, given the same contrast and 10% bandwidth goal as before. The results are illustrated in Figs. 6 and 8 and listed in Table 1. As in the case of config. IIa (not plotted), the mismatch between the apodizer and the Babinet subtrahend profiles leads to high amplitude, sharp residual features in the Lyot plane. This time, however, a subtle rearrangement of Lyot stop obstructions is enough to enable an apodizer with far more open area. Most of the sharp residual Lyot plane features are not obstructed by the freely varying stop, contrary to what one might expect. Apparently, not even a modest level of field cancellation in the Lyot plane is required to create deep, broadband destructive interference in the image plane. The FWHM throughput of this solution is 0.33, more than triple that of the comparable clear Lyot stop configuration (IIa). The coronagraph PSF also sharpens, giving an FWHM area only 25% larger than the Airy disk. We also tested the effect of decreasing the focal occulting spot radius from  $3\lambda_0/D$  to  $2\lambda_0/D$  and arrived at a similar design



**Fig. 6** Circular aperture Lyot coronagraph, config. IIIa. The (a) apodizer and (b) Lyot stop are jointly optimized for maximum transmission while achieving a contrast of  $10^{-9}$  in the image plane over a 10% passband over a working angle range of  $2$  to  $12\lambda_0/D$ . The FPM is an occulting spot of radius  $2\lambda_0/D$ . The right-hand plot (c) shows that for this configuration, the solution depends on high-amplitude discontinuities in the Lyot plane field.



**Fig. 7** On-axis intensity pattern for a circular shaped pupil Lyot coronagraph (SPLC) with a jointly optimized apodizer and Lyot stop and a working angle range of 2 to  $12\lambda_0/D$ . The nonlinear optimization program constrains the contrast at three wavelengths, resulting in a pseudo-achromatized dark search region fixed in sky coordinates. The black curve shows the average intensity across five wavelength samples spanning the 10% passband.

with a throughput of 17%. The contrast curve of this design is plotted in Fig. 7, showing the intensity pattern at three wavelengths, as well as the average over five wavelength samples spanning the 10% passband.

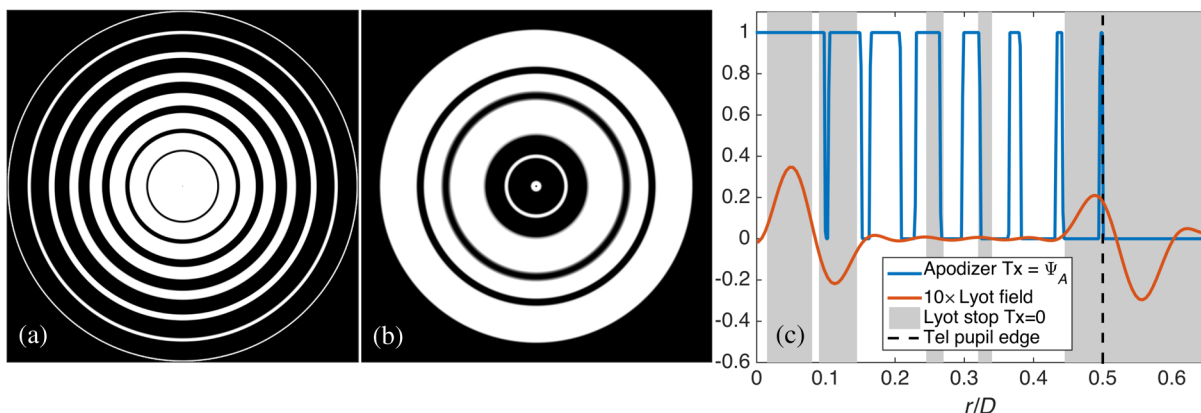
With the annular diaphragm FPM, allowing the Lyot stop transmission profile to vary also results in increased throughput, although the improvement here is less dramatic, climbing from 0.108 to 0.144 (Table 1). The apodizer and Lyot stop both have less open area than the spot FPM variant. Notably, however, the PSF is almost as sharp as the occulting spot variant, with FWHM area 39% larger than the Airy disk. The plot in Fig. 8 of the Lyot plane field alongside the Lyot stop transmission profile shows that the performance of this configuration benefits from notching out the radial peaks, which was not the case for the spot FPM. The resulting Lyot stop has five prominent opaque rings and one small dark spot at the center. Another important aspect of the Lyot plane behavior for the

diaphragm FPM is the relative smoothness of the field structure as compared to the spot FPM case. This is a direct outcome of the mathematical description of the Lyot field in Eq. (5), where both instances of the apodizer transmission function are convolved with a jinc function. This quality of the diaphragm FPM variant of the SPLC hints at a more generous tolerance to manufacturing and alignment. We revisit this point in Sec. 4.3, in the context of our WFIRST-AFTA designs.

From further experiments, we found that the nonlinear, non-convex program used to derive joint shaped pupil and Lyot stop solutions only converges for one-dimensional coronagraph models. In our circular aperture case, this two-plane optimization program operates near the limit of the interior point solver’s capability, and reliable outcomes require tuning. Even for low spatial resolution versions of obstructed two-dimensional apertures, there are too many variables to extend the tactic. This obstacle is algorithmic in nature rather than one that can be surmounted by expanding the computing hardware capacity. This difficulty, combined with the practical attractions of a simpler Lyot stop, suggest one might settle for an intermediate performance level by surveying an annular Lyot stop described by only two parameters (inner and outer radius). We have not yet explored the full range of inner and outer diameter Lyot stop combinations for the circular aperture. However, we found that for an arbitrary test design with a  $0.1D$  inner diameter and  $0.9D$  outer diameter (configs. IVa and IVb), performance is not far from the optimized Lyot stop: for the case of the spot FPM, throughput decreases only from 0.334 to 0.317 (Table 1). For the diaphragm FPM variant, the throughput loss resulting from the switch to the annular Lyot stop is also small. However, the coronagraph PSF deteriorates significantly, jumping in area from 1.39 to 1.93 times that of the Airy core.

## 2.5 Distinction between the Shaped Pupil Lyot Coronagraph and Microdot Realizations of the Apodized Pupil Lyot Coronagraph

Microdot lithography can be used to stochastically approximate the continuous prolate apodizer solutions derived from



**Fig. 8** Circular aperture Lyot coronagraph, config. IIIb. The (a) apodizer and (b) Lyot stop are jointly optimized for maximum transmission while achieving a contrast of  $10^{-9}$  in the image plane over a 10% passband over a working angle range of 2 to  $12\lambda_0/D$ . The FPM is an annular diaphragm with an inner radius of  $2\lambda_0/D$  and outer radius of  $12\lambda_0/D$ . The Lyot plane field illustrated in the right-hand plot (c) differs from the case of config. IIIa in two ways: (1) it varies smoothly, due to the filtering effect of the annular FPM, and (2) field nulls coincide with gaps in the Lyot stop.



Eqs. (2) and (3), as well as their analogs for more complicated telescope apertures.<sup>80–82</sup> The technique stems from long-established printing processes, in which an array of black pixels with varying spatial density imitates the halftones of a gray-scale image. A microdot apodizer for a Lyot coronagraph can be manufactured with an opaque metal layer deposited on a glass substrate at the locations of black pixels.<sup>83,84</sup> In testbed experiments, APLC designs with microdot apodizers have reached contrasts as low as  $5 \times 10^{-7}$ .<sup>85</sup> Microdot APLC apodizers are core components in several on-sky, AO-fed coronagraphs.<sup>1–3</sup>

Although the halftone microdot process results in a binary-valued transmission pattern, there is nonetheless a categorical distinction from the SPLC. A shaped pupil, rather than approximating a continuous mask solution in the apodizer plane, instead matches the desired destructive interference properties in the image plane. Consequently, on a macroscopic scale, the ring apodizer shown in Fig. 4 is qualitatively dissimilar to a halftone APLC approximation, despite solving a similar field cancellation problem. Instead, the image domain is where the strongest resemblance appears between the SPLC and APLC solutions. This is made evident by comparing their on-axis field distributions at the first focal plane within the bounded search region, shown in Fig. 5 for the most elementary design case (monochromatic cancellation in the Lyot plane).

Because the SPLC design process directly optimizes the performance, the fabrication instruction set for the apodizer realization is a one-to-one replica of the linear program solution. A microdot APLC apodizer, on the other hand, is one step removed from an underlying numerical solution. In this sense, the shaped pupil technique has a clear advantage for meeting the high precision required for the most demanding applications.

### 3 Shaped Pupil Lyot Coronagraph Designs for WFIRST-AFTA

#### 3.1 WFIRST-AFTA Coronagraph Instrument Concept

CGI proposed by the WFIRST-AFTA Science Definition Team aims to image and measure the spectra of mature, long-period gas giants in the solar neighborhood. This planet population, which at present can only be studied indirectly through radial velocity (RV) surveys, is out of reach of transit spectroscopy methods due to their strong bias toward highly irradiated planets on short orbital periods. Depending on orbital configuration and albedo characteristics, the planet-to-star contrast of an ex-Jupiter seen in reflected starlight is of the order of  $10^{-8}$  or below. Due to AO performance limitations, this contrast ratio may prove too extreme for ground-based imaging, regardless of telescope aperture or coronagraph design.<sup>86</sup>

The reflected spectra of gas giants are sculpted by a series of methane absorption bands in the range of 600 to 970 nm. Acquiring these fingerprints for an ensemble of planets, in conjunction with mass constraints from radial velocities and astrometry, will provide a wealth of insights into the structure, composition, and evolution of gas giants.<sup>87,88</sup> The Princeton team was tasked with providing shaped pupil designs for this characterization mode, covering the stated wavelength range in three 18% passbands, each corresponding to one filter setting of the integral field spectrograph (IFS).<sup>89–91</sup>

The optical path of the proposed CGI is shared between SPC/SPLC and JPL's HLC.<sup>62</sup> The HLC uses an FPM with a phase- and

amplitude-modulating transmission profile.<sup>64</sup> In the baseline configuration of the CGI, the HLC mode operates with two imaging filters, nominally 10% bandpasses centered at 465 and 565 nm. The HLC mode is optimized for detection and color measurements of the scattered continuum, rather than spectroscopic characterization with the wider bandpass of the IFS.<sup>91</sup>

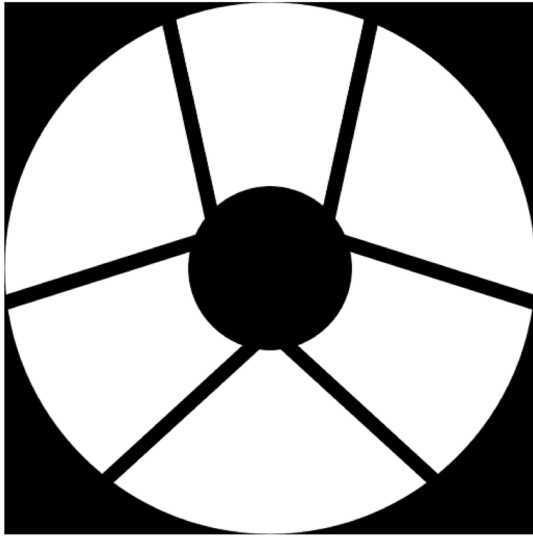
In addition to exoplanets, a closely related category of scientific opportunity for WFIRST-AFTA is circumstellar debris structure. One of the goals of the CGI will be to image scattered light from low-density, solar-system-like zodiacal disks that are below the noise floor of existing instruments. In addition, thick debris disks of the kind already studied with the Hubble Space Telescope will be probed at smaller angular separations than before. This will unveil the dynamic evolution of circumstellar debris and its interaction with planets in the habitable zones of exoplanetary systems.<sup>92</sup> Small angular separation observations of debris disks can be carried out with the HLC mode. However, some of the foreseen disk imaging programs require larger OWAs ( $> \sim 0.5$  arc second) than those relevant to reflected starlight exoplanet detection. Therefore, we explored separate SPLC mask solutions for a dedicated, wide-field disk science coronagraph mode.

Throughout our design process, we concentrate on three essential performance metrics: contrast, IWA, and throughput. The scientific goals require all WFIRST-AFTA designs to achieve a raw contrast of  $10^{-8}$ , defined at a given image position as the ratio of diffracted starlight intensity to the peak of the off-axis coronagraph PSF shifted to that location. We make the assumption that data postprocessing will further reduce the intensity floor, so that planets several times below this nominal contrast can be detected.<sup>78,89,91,93</sup>

IWA is defined as the minimum angular separation from the star at which the coronagraph's off-axis (planet) PSF core throughput reaches half-maximum.<sup>78</sup> For a Lyot coronagraph, planet throughput rises steadily with increasing angular separation from the edge of the FPM, leveling off when the core of the PSF clears the line-of-sight FPM occultation. Having a small IWA is especially important for a coronagraph aiming to detect starlight reflections, because the irradiance falls off with the square of the planet–star distance. The consequence—when considered along with the distances to nearby FGK stars and their expected distributions of planet semimajor axes—is two-fold: (1) the number of accessible planets rises steeply with reduced IWA and (2) those giant exoplanets at smaller angular separations tend to be the brightest targets.<sup>91,94</sup>

As in Sec. 2, we define throughput as the ratio of energy contained within the FWHM contour of the PSF core to that of the telescope PSF with no coronagraph. Planet signal-to-noise ratios will generally be low, and the number of targets the instrument can acquire over the mission lifespan will be limited by the cumulative integration times.<sup>91</sup> Detection times will depend on the total amount of planet light that survives propagation losses through the optical train and how tightly that remaining energy is concentrated on the detector. In characterization mode, the spectrograph will disperse the planet's light over many detector pixels. Therefore, the majority of the instrument's operational time budget will be consumed by integrations totaling one day or more per target.<sup>78</sup>

The 2.4 m-diameter WFIRST-AFTA telescope aperture is illustrated in Fig. 9. Its large central obstruction ( $0.31D$ ) and six off-center support struts, each oriented at a unique angle, pose a challenge for any coronagraph design. The SPC and



**Fig. 9** The WFIRST-AFTA telescope aperture from the most recent mission design cycle, as seen on-axis.

SPLC use the apodization pattern of the shaped pupil to confine the diffraction effects of these obstructions outside of the optimized dark region.<sup>37,38</sup> The penalties of this strategy are lower throughput and higher IWA than would be the case for a clear circular aperture. Alternatively, it is possible for a coronagraph to counteract these obstructions with static phase excursions applied to deformable mirrors<sup>44,62</sup> or custom aspheric optics.<sup>43,65,66,95</sup> However, amplitude-mask-based apodization places less demanding requirements on mirror surface manufacturing tolerances, alignment tolerances, and deformable mirror reliability, thereby mitigating the overall technological risk of our design.

The SPLC designs presented here build directly on the efforts of Carlotti et al.,<sup>38</sup> who led the first shaped pupil designs for WFIRST-AFTA. Those first-generation shaped pupil coronagraphs fulfill the basic mission requirements. They were described in further detail by Riggs et al.,<sup>68</sup> and in this issue, Cady et al. describe successful laboratory demonstrations of the first-generation characterization SPC design.<sup>58</sup> The SPLC designs here form part of a reference design case adopted by the Science Definition Team for the purpose of technology demonstrations, mission simulations, and cost assessment.<sup>60</sup> The flight design moving forward may differ significantly.

### 3.2 Characterization Mode Shaped Pupil Lyot Coronagraph

For the characterization design, the challenge is to achieve a small IWA while maintaining acceptable throughput. Although it is always desirable to create a full 360-deg dark search region around the star, we know from previous work that it is impossible for a shaped pupil alone to produce an annular FoV with IWA of  $4\lambda/D$  or below with the obscurations of the WFIRST-AFTA aperture.<sup>38</sup> However, knowing that the SPLC configuration should be able to reach a smaller IWA at the same contrast and throughput as the first-generation SPC, we now examine again how close we can push a 360-deg dark region in toward the star. At the same time, for effective broadband characterization, we strongly prefer a quasi-achromatic dark region,<sup>26</sup> so that

a target located near the IWA is detected across the full spectrograph passband. Therefore, for our parameter exploration, we always apply polychromatic image constraints, with inner and outer image radii defined in terms of central wavelength diffraction elements ( $\lambda_0/D$ ), as we did before in Sec. 2.3, and similar to previous APLC optimizations described by N'Diaye et al.<sup>27</sup> In Appendix A2, we describe the practical details of the optimization procedure used to test a given set of design parameters.

Informed by the results of our circular SPLC trials (Secs. 2.3 and 2.4), we use an occulting spot FPM, with radius either  $\rho_0 = 2.5$  or  $3.0\lambda_0/D$ . We surveyed the Lyot stop parameter space by repeating optimizations with different padding levels on the inner and outer edges of the telescope aperture replica, ranging from 2 to 12% of the diameter. We also varied the outer radius of the dark region between 8, 9, and  $10\lambda_0/D$ . Finally, we repeated these multiparameter trials for two bandwidths: 18% (the target characterization design) and 10%. A subset of the results is summarized in Table 2. Several conclusions can be drawn. First, for the smaller inner FPM radius  $\rho_0 = 2.5\lambda_0/D$ , there are no acceptable 360-deg SPLC solutions for the full 18% characterization bandwidth. At  $\rho_0 = 3.0\lambda_0/D$ , however, some weak solutions begin to appear. Performance here is sensitive to the Lyot stop padding level, and for the 18% bandwidth case, the best padding levels are in the range of 8 to 10% of the pupil diameter. When the bandwidth is reduced to 10%, the improvement in throughput is dramatic. In particular, we highlight a design with a throughput of 0.14 and FPM radius of  $3.0\lambda_0/D$ . We did not find a strong dependence on outer dark region radius  $\rho_1$  over the values we surveyed.

To reach an IWA smaller than  $3\lambda_0/D$ , and do so over the full characterization bandwidth, we need to restrict the azimuthal span of the constrained dark region. This strategy was originally developed to design the first-generation WFIRST-AFTA SPCs, resulting in a design with  $\rho_0 = 4\lambda_0/D$ . In a survey aiming to discover exoplanets, bowtie-shaped dark zones have the disadvantage of requiring repeat integrations for two or more mask orientations. However, the overhead for characterizing a planet with a known position is minor. Furthermore, restricting the azimuthal and radial FoV of the image plane search area alleviates the demands placed on the wavefront correction. When the wavefront control system aims to suppress light only in a small region, there are more degrees of freedom available than when

**Table 2** Throughput of 360-deg field-of-view WFIRST-AFTA shaped pupil Lyot coronagraph (SPLC) solutions for different combinations of inner working angle (IWA), bandwidth, and Lyot stop padding. The outer edge of the dark image constraint region is fixed at  $\rho_1 = 8\lambda_0/D$ .

Lyot stop padding (% diam.)	18% bandwidth $\rho_0 (\lambda_0/D)$		10% bandwidth $\rho_0 (\lambda_0/D)$	
	2.5	3.0	2.5	3.0
6	$<10^{-3}$	0.042	0.040	0.13
8	$<10^{-3}$	0.057	0.060	0.14
10	$<10^{-3}$	0.066	0.062	0.12
12	$<10^{-3}$	0.056	0.043	0.07

trying to suppress the full correctable region. Restricting the dark hole problem thus leaves greater tolerance for unknown aberrations in the propagation model.

For the SPLC configuration, we surveyed a range of bowtie-shaped focal plane geometries with inner radii between  $2.4\lambda_0/D$  and  $3.0\lambda_0/D$ , and opening angles between 30 and 90 deg. The trials are repeated for 18 and 10% bandwidths. All optimization attempts at smaller inner radii, such as  $2.2\lambda_0/D$ , failed to give results with reasonable throughput (above 0.01). Here, instead of an occulting spot, the FPM is a bowtie-shaped aperture matched to the optimized focal plane region in the final image, as illustrated in Fig. 10. The presence of the outer edge in the first focal plane makes this design most analogous to config. IVb, among the circular SPLCs described in Sec. 2 and Table 1.

The Lyot stop we use for the bowtie characterization design is a simple clear annulus rather than a padded replica of the telescope aperture (Fig. 10). That is because the low-pass filter effect of the bowtie FPM smears the support strut field features in the Lyot plane. We verified through separate optimization tests that there is no advantage to including matched support struts in the Lyot stop in this configuration. To survey the dependence of throughput on focal plane geometry, we fix the inner diameter of the Lyot stop annulus at  $0.3D$  and the outer diameter at  $0.9D$ . Later, we tune the inner and outer diameters of the Lyot stop for a specific characterization design.

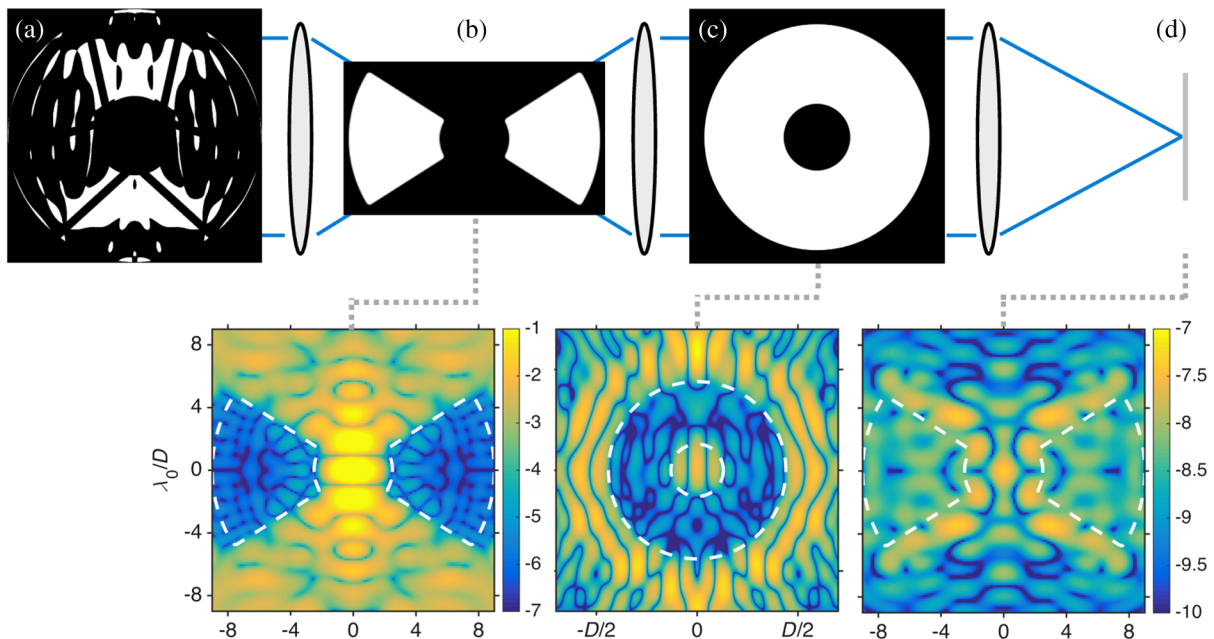
Some results from the focal plane geometry trials are collected in Table 3. Since we found that throughput depends relatively weakly on outer dark region radius, here we only tabulate

**Table 3** Throughput of bowtie characterization WFIRST-AFTA SPLC solutions for different combinations of inner radius and opening angle. The optimization passband is 18%, the outer radius is fixed at  $\rho_1 = 9\lambda_0/D$ , and the Lyot stop is a clear annulus with [I.D., O.D.] = [0.3, 0.9]. The approximate trade-off in inner radius, opening angle, and throughput we chose to pursue is highlighted in bold font.

	Opening angle (deg)				
$\rho_0 (\lambda_0/D)$	30	45	60	75	90
2.4	0.14	0.089	0.048	0.026	0.010
2.6	0.16	0.17	<b>0.11</b>	0.049	0.022
2.8	0.16	0.18	0.13	0.067	0.030
3.0	0.17	0.18	0.15	0.088	0.042

the throughput values for  $\rho_1 = 9\lambda_0/D$ . Throughput varies steeply with opening angle. In particular, from 60 to 90 deg, the throughput decreases by a factor of  $\sim 4$  to 5. At an inner radius of  $2.4\lambda_0/D$ , the only opening angle with throughput above 0.1 is the 30-deg bowtie.

We find the most compelling trade-off at  $\rho_0 = 2.6\lambda_0/D$ , which has throughput a of 0.11 at an opening angle of 60 deg. Similar to the first-generation SPC, this enables the full FoV to be covered with three pairs of shaped pupils and



**Fig. 10** Diagram of the characterization-mode SPLC mask scheme, along with the plots of the intensity of the on-axis field at each critical plane. The shaped pupil (a) forms a bowtie-shaped region of the destructive interference in the first focal plane (b), which is then occulted by a diaphragm with a matched opening. The on-axis field is further rejected by an annular stop in the subsequent Lyot plane (c) before it is reimaged at the entrance of the integral field spectrograph (d). The propagation is shown at the central wavelength of the design for the case of a perfectly flat wavefront with no planet or disk present. The flux scale bars indicate the intensity on a  $\log_{10}$  scale. In the first focal plane (b), the flux scale is normalized to the point spread function (PSF) peak, whereas in the final focal plane (d), the scale is normalized to the peak of the unseen off-axis PSF, in order to map the contrast ratio in a way that accounts for the Lyot stop attenuation. The mean contrast in the dark bowtie region (averaged over azimuth angle, then averaged over wavelength, and then averaged over radial separation) is  $6 \times 10^{-9}$ .

FPMs oriented at 120-deg offsets. Due to their limited utility, the 10% bandwidth trials are not tabulated here, but we can summarize them by pointing out that throughput increases by a factor of 1.1 to 3 over the 18% bandwidth case, with the largest changes occurring for small  $\rho_0$  and wide opening angle.

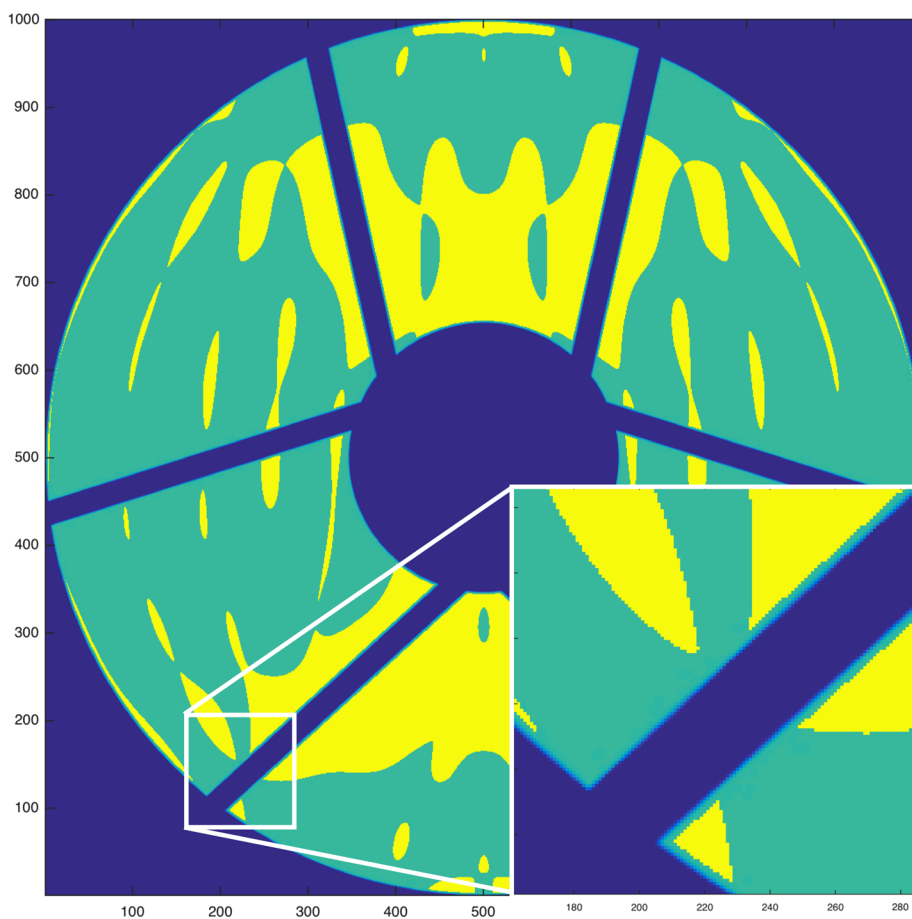
A full set of mask designs for a bowtie characterization SPLC have been delivered to JPL for fabrication and experiments on the HCIT. The detailed structure of the  $1000 \times 1000$  pixel shaped pupil apodizer array is illustrated in Fig. 11.

For this version, we raised the opening angle above 60 deg to provide a margin of FoV overlap among the three mask orientations needed to cover an annulus around the star. The overlap slightly reduces the likelihood of a scenario where the location of an exoplanet coincides with the edge of a bowtie mask, cropping the PSF core and requiring extra integration time to compensate. After the opening angle was fixed at 65 deg, we decremented  $\rho_0$  from  $2.6\lambda_0/D$  to the smallest radius that maintains the throughput above an arbitrary goal of 0.10, thereby revising  $\rho_0$  to  $2.5\lambda_0/D$ . The Lyot stop is an annulus with inner diameter of  $0.26D$  and outer diameter of  $0.88D$ . We stress that these design choices are provisional and that

maximizing the scientific yield would require integrating the parameter survey with end-to-end observatory and data simulations.

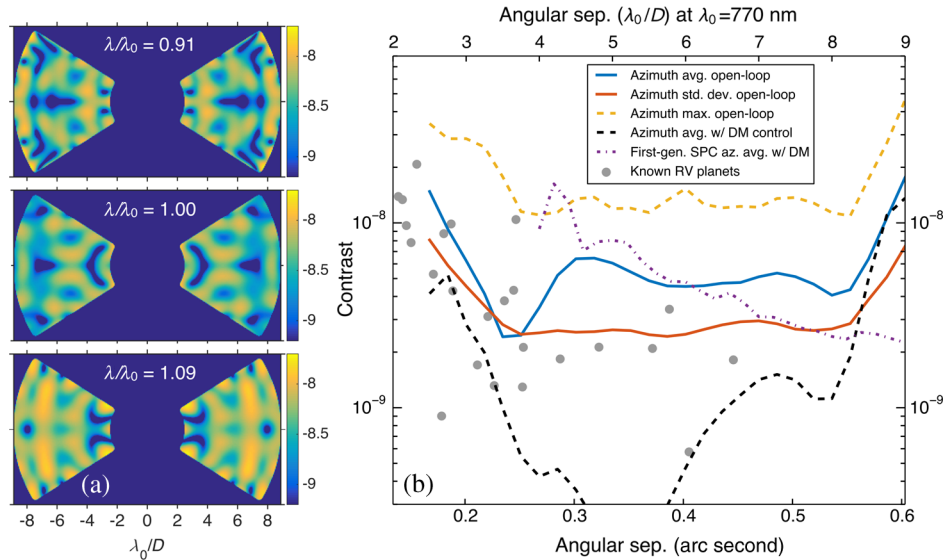
The ideal model PSF and contrast curves with zero wavefront error are shown in Fig. 12. At center wavelength, the FWHM PSF area is 1.6 times that of the WFIRST-AFTA PSF. The contrast constraint is slightly relaxed relative to the previous parameter trials:  $2 \times 10^{-8}$  for separations below  $3.5\lambda_0/D$  and  $1.5 \times 10^{-8}$  in the rest of the bowtie region. Still, the average contrast curve (here averaged over azimuth and then over wavelength) is well below the worst-case intensity values, as plotted in the right-hand side of Fig. 12, below  $7 \times 10^{-9}$  at all angular separations in the range of 3 to  $8\lambda_0/D$ .

The two deformable mirrors integrated with the WFIRST-AFTA CGI are expected to improve the nominal SPLC performance. Since the SPLC design optimization only makes use of amplitude operations, the extra degrees of freedom from deformable mirror (DM) phase control can yield higher contrast. To demonstrate this, we simulate the effect of DM control with an unaberrated wavefront. We simulated wavefront control on a layout similar to the actual WFIRST-AFTA CGI with two



**Fig. 11** Detail of the  $1000 \times 1000$  point shaped pupil mask solution for the WFIRST-AFTA characterization mode, corresponding to the design exhibited in Figs. 10 and 12. The obscurations of the WFIRST-AFTA telescope aperture are colored blue, the regions of the pupil masked by the shaped pupil apodizer in addition to the telescope pupil are colored green, and the regions of the pupil transmitted by the apodizer are colored yellow. The magnified inset shows the granular quality of the square, binary elements of the shaped pupil array. The inset also shows the gap between the edge of the telescope aperture features and the open regions of the apodizer, which is reserved in order to ease the alignment tolerance between the shaped pupil apodizer and the telescope pupil.





**Fig. 12** (a) Ideal on-axis PSF of the characterization SPLC design, with no aberrations, shown at three wavelengths: the short-wavelength extreme of the passband (top;  $\lambda/\lambda_0 = 0.91$ ), the center wavelength (middle;  $\lambda/\lambda_0 = 1.00$ ), and the long-wavelength extreme (bottom;  $\lambda/\lambda_0 = 1.09$ ). (b) Raw contrast plotted alongside the estimated contrasts and separations of the known long-period radial velocity exoplanet population. The contrast curves are shown for the SPLC design centered on  $\lambda_0 = 770$  nm, the middle of the three spectrograph filters. For reference, we also include the contrast curve of the first-generation shaped pupil coronagraph (purple dashed curve).

$48 \times 48$  actuator DMs upstream of the SPLC. We divided the 18% passband into nine wavelength samples and weighted each equally to control the dark hole with a stroke minimization algorithm, originally described by Pueyo et al.<sup>96</sup> The inner region of the bowtie is most critical since more exoplanets are expected to be observed at small angular separations, so we weighted the intensity from  $2.5$  to  $4.5\lambda_0/D$ , three times higher for a slight improvement. The resulting contrast curve is plotted in the right-hand-side plot of Fig. 12. With DM control, the average intensity in the separation range of  $2.5$  to  $3.5\lambda_0/D$  is reduced by a factor of 2, and at separations of  $4$  to  $8\lambda_0/D$ , by a factor of 4 or more. In addition to the azimuthally averaged contrast, in the same figure, we also plot the standard deviation of the intensity pattern as a function of separation from the star, measured in concentric annuli.

We use the distribution of known RV exoplanets to indicate where the performance of this coronagraph sits relative to plausible characterization targets. We take the same assumptions made by Traub et al.<sup>91</sup> in their science yield calculations, resulting in a representative target population plotted on the right-hand side of Fig. 12. Planets are assumed to be on circular orbits inclined by  $60$  deg and observed at a favorable mean anomaly of  $70$  deg. The longest period RV exoplanet detections are generally on the upper end of the mass distribution, so in the absence of other constraints, they are assigned a size equal to Jupiter and a geometric albedo of  $0.4$ . Finally, to compare the SPLC contrast curve on an angular scale, we set the central wavelength of the characterization design to  $770$  nm, the middle of the three nominal IFS filters.<sup>90</sup> It can be seen that 12 planets outside of the  $2.8\lambda_0/D$  IWA have contrasts and separations placing them above the band-averaged contrast floor obtained from the wavefront control simulation.

The first-generation characterization SPC design for WFIRST-AFTA is overplotted as the dashed purple contrast

curve in Fig. 12. With an IWA of  $4\lambda/D$ , that coronagraph could only access half of the exoplanets in the mock target sample. There is an additional disadvantage of the first-generation design that is not apparent in the contrast plot. The IWA of a shaped pupil PSF scales directly with wavelength. Without a Lyot stop, there is no possibility to anchor the inner radius of the dark bowtie region across the spectrograph filter bandpass, as we do to optimize the SPLC. Therefore, an exoplanet falling near the inner edge of the first-generation contrast curve would be undetected at the long-wavelength end of the filter.

We acknowledge that the raw contrast prediction for the characterization SPLC is optimistic since it does not include aberrations, tip-tilt jitter, alignment errors, and so forth. But the comparison verifies that our design functions in the regime of contrast and angular separation needed to meet the top-level mission requirement of acquiring the reflected spectra of six or more gas giants.<sup>91</sup> In this issue, Krist et al. analyze the sensitivity of the SPLC performance to realistic aberrations and incorporate the coronagraph in an end-to-end simulation of the WFIRST-AFTA telescope for a complete observing scenario.<sup>78</sup>

### 3.3 Debris Disk Mode–Shaped Pupil Lyot Coronagraph

To design an SPLC for debris disk imaging over a much wider FoV, we carry out a parameter survey similar to that of the  $360$ -deg characterization design trials. Assuming a deformable mirror with a  $48 \times 48$  actuator array, the maximum correctable aberration spatial frequency corresponds to an angular separation of  $21.8\lambda_0/D$  at the short wavelength end of an 18% passband. To approximately match this, we fix the outer radius of the polychromatic dark region at  $20\lambda_0/D$ . Within that dark annulus, we constrain the contrast to  $\leq 10^{-8}$  over an 18% bandwidth. We test FPM spot radii of  $6.0$ ,  $6.5$ , and  $7.0\lambda_0/D$ , and Lyot stop padding levels between  $2$  and  $8\%$ . The throughput results are tabulated in

**Table 4** Throughput of debris disk science WFIRST-AFTA solutions for different combinations of inner radius and Lyot stop padding. The trade-off in inner radius and throughput corresponding to the design exhibited in Figs. 13 and 14 is highlighted in bold font.

Lyot stop padding (% diam.)	$\rho_0$ ( $\lambda_0/D$ )		
	6.0	6.5	7.0
2	0.17	0.21	0.22
4	0.18	<b>0.23</b>	0.24
6	0.17	0.20	0.22
8	0.14	0.15	0.17

Table 4. We highlight the solution with  $\rho_0 = 6.5\lambda_0/D$  and padding level of 4%, since it gives a throughput of 0.23, almost as high as the best design at  $\rho_0 = 7\lambda_0/D$ . With the smaller focal occulting spot radius of  $\rho_0 = 6\lambda_0/D$ , on the other hand, there is a significant throughput drop for all the Lyot stops.

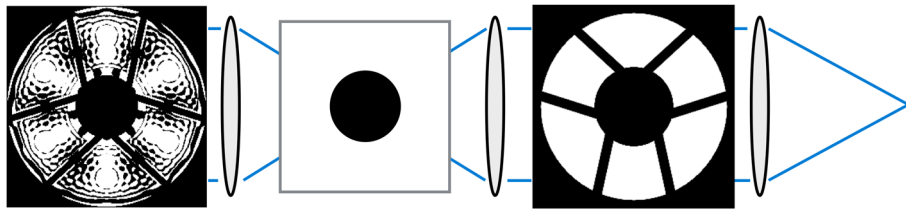
In Fig. 13, we illustrate the SPLC mask scheme for the highlighted disk science design. The apodizer maintains over 59% of the available open area around the WFIRST-AFTA

pupil obscurations, and the FWHM PSF area is only 1.11 times that of the WFIRST-AFTA telescope. The on-axis PSF of the coronagraph at the center wavelength is plotted on the left-hand side of Fig. 14, along with the ideal contrast curves. Like the characterization design, the mean contrast is significantly deeper than the constraint value, due to the lumpy structure of the diffraction pattern.

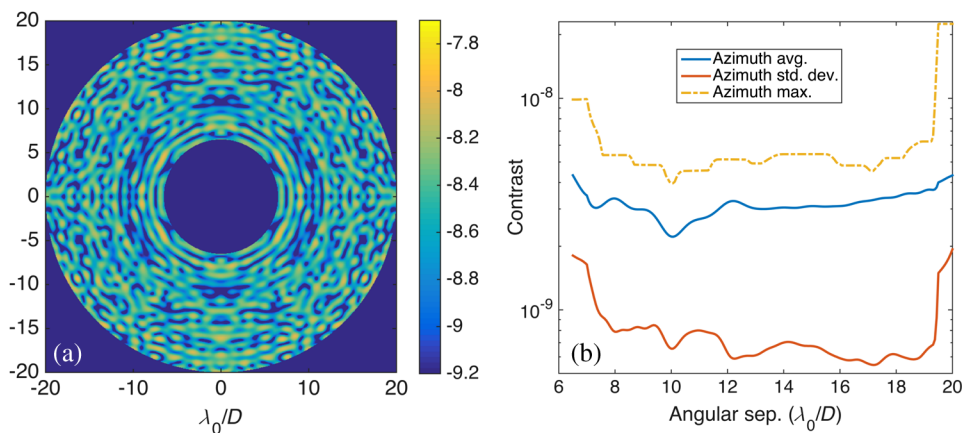
We have also optimized debris disk designs that use an annular diaphragm FPM instead of an occulting spot. The throughput in this configuration is approximately half that of the spot FPM case for the same contrast, FoV, and bandwidth parameters. However, a general advantage of SPLC designs that use the annular diaphragm FPM is their greater tolerance to Lyot stop mask misalignment, an issue we discuss in Sec. 4. Therefore, despite the lower theoretical performance, for the initial testbed implementation of the debris disk SPLC, we will use an annular FPM variant.<sup>97</sup>

### 3.4 Summary of WFIRST-AFTA Shaped Pupil Lyot Coronagraph Designs

We assemble in Table 5 the parameters and performance metrics of our candidate SPLC designs for WFIRST-AFTA. To be consistent with other coronagraph descriptions, the IWA and OWA are measured by the half-maximum crossings of the throughput



**Fig. 13** Mask scheme for the debris disk mode SPLC for WFIRST-AFTA, from left to right: shaped pupil apodizer, FPM, and Lyot stop. The relative sizes of the masks in conjugate planes are not shown to scale. This design produces a broadband, annular dark region with  $\leq 10^{-8}$  contrast over an 18% bandwidth from  $6.5\lambda_0/D$  to  $20\lambda_0/D$ . The FPM is an occulting spot of radius  $6.5\lambda_0/D$ . The Lyot stop is a replica of the telescope aperture, with the inner and outer edges padded at 4% of the pupil diameter.



**Fig. 14** (a) The ideal, on-axis, center wavelength PSF of a broadband disk science SPLC design. On the right-hand side (b), the ideal contrast is averaged over azimuth samples and then averaged over seven wavelength samples spanning the 18% passband. Also plotted are the maximum contrasts at each separation over all azimuth and wavelength samples.

**Table 5** Summary of WFIRST-AFTA SPLC designs.

Configuration	IWA ( $\lambda_0/D$ )	OWA ( $\lambda_0/D$ )	Field of view	Contrast	Bandwidth	Throughput	PSF area
Characterization mode	2.8	8.7	$2 \times 65$ deg	$6.0 \times 10^{-9}$	18%	0.10	1.6
Disk mode / spot FPM	6.6	19.9	360 deg	$3.2 \times 10^{-9}$	18%	0.22	1.1
Disk mode / ann. FPM	6.8	19.7	360 deg	$3.1 \times 10^{-9}$	10%	0.15	1.2

Note: OWA, outer working angle.

curve rather than the dimensions of the FPM and optimization constraints.<sup>50</sup> We use the same definitions for throughput and PSF area first given in Sec. 2.3. The PSF area is the FWHM region of the PSF for an off-axis (planet-like) source, normalized to the FWHM area of the PSF of the WFIRST-AFTA telescope without a coronagraph. We note that two designs are listed for the debris disk mode. One is the occulting spot configuration described in Sec. 3.3 and Figs. 13 and 14; the other is the annular diaphragm variant that is undergoing fabrication for testbed evaluation at HCIT, described in more detail in Ref. 97.

## 4 Error Tolerances

### 4.1 Shaped Pupil Apodizer

For the first-generation SPC designs for WFIRST-AFTA, Riggs et al. found that the contrast performance was sufficiently robust to etching errors.<sup>68</sup> The current testbeds in the JPL HCIT are using shaped pupil masks with diameters between 14 and 22 mm. With 1000 optimized transmission points across the mask diameter, the binary array is therefore composed of square pixels of width 14 to 22 microns. The standard etching tolerance in JPL's Microdevices Lab is below 1 micron, so the etching error on each SP pixel is less than  $\sim 5\%$ . For a 5% uniform over/underetching error, the first-generation characterization SPC had an open-loop contrast degradation of 2 $\times$ , from 1 to  $2 \times 10^{-8}$ . For the mission payload, even tighter etching tolerances can be achieved, so uniform etching errors are not a major concern for the SP apodizer mask.

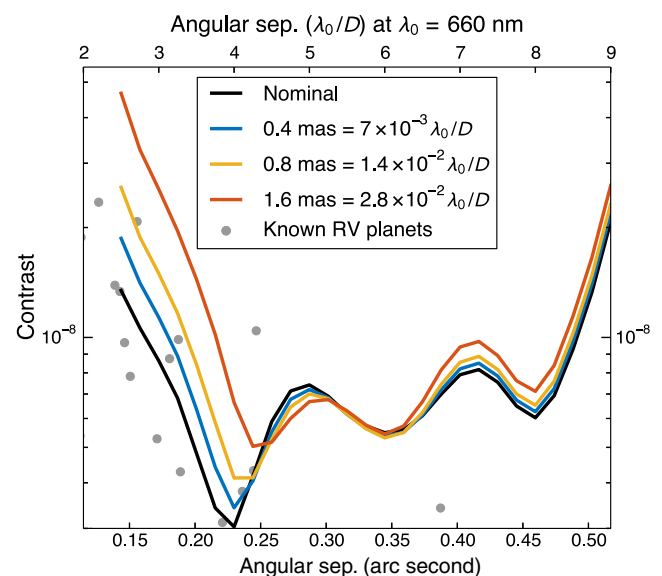
### 4.2 Focal Plane Mask and Pointing Sensitivity

For the bowtie characterization SPLC described in Sec. 3.2, we modeled the sensitivity to systematic errors in the FPM fabrication and alignment, as well as line-of-sight pointing errors originating from spacecraft jitter. The width of the stellar PSF core in relation to the inner radius of the bowtie mask is highest at the long-wavelength end of the design passband. Any disparity from the transmission profile assumed in the on-axis optimization therefore causes unwanted starlight to leak into the inner part of the dark region. Conversely, at the short-wavelength end of the passband, the outer perimeter of the dark bowtie region is sensitive to transmission profile disparities along the outer edge of the bowtie mask.

We plot the effect of pointing errors and, in particular, their effect on the inner part of the image in Fig. 15. In our Fourier propagation model, we apply phase ramps in the apodizer plane corresponding to a set of tilt errors along the long axis of the

bowtie for the characterization design with central wavelength of 660 nm. This wavelength corresponds to the bluest of the three characterization filters, chosen here because when IWA is fixed in resolution elements ( $\lambda_0/D$ ), the shortest-wavelength coronagraph realization—as defined by the physical scale of the FPM—is the one most sensitive to a given telescope pointing error. We apply pointing errors of 0.4, 0.8, and 1.6 milliarc second. These are on the same angular scale as the residual jitter levels that may be present on the WFIRST spacecraft.<sup>78,89,91</sup> At the 660 nm center wavelength of the characterization passband, these tilts translate to focal plane offsets of  $7 \times 10^{-3} \lambda_0/D$ ,  $1.4 \times 10^{-2} \lambda_0/D$ , and  $2.8 \times 10^{-2} \lambda_0/D$ , respectively. To show the upper bound of the impact on the polychromatic coronagraph PSF, in Fig. 15, we plot the contrast only at the long-wavelength end of the passband, 719 nm, where the effect is worst. At each separation, we plot the azimuthal average of the contrast in the degraded half of the bowtie region.

The results plotted in Fig. 15 show that contrast outside of  $4\lambda_0/D$  is not degraded significantly for pointing errors up to 1.6 milliarc second, roughly equivalent to an inner FPM radius error of  $\sim 3 \times 10^{-2} \lambda_0/D$ . At the interior, the long-wavelength intensity increments for 0.4, 0.8, and 1.6 milliarc second are, respectively,  $6 \times 10^{-9}$ ,  $1.4 \times 10^{-8}$ , and  $3.6 \times 10^{-8}$ , as expressed in units of contrast. We emphasize that these values only indicate



**Fig. 15** Contrast degradation of WFIRST-AFTA characterization design for three pointing errors, shown at  $\lambda = 719$  nm, the red end of the design centered on the  $\lambda_0 = 660$  nm spectrograph filter.

**Table 6** Contrast degradation of the characterization SPLC for a set of horizontal and vertical translation offsets of the FPM.

	Mean $\Delta$ Contrast; all spatial and wavelength samples	Mean $\Delta$ Contrast; sep. $<3.8\lambda_0/D$ , red edge of 18% band, worst quadrant
Horz. $5 \times 10^{-3} \lambda_0/D$	$3.0 \times 10^{-11}$	$1.5 \times 10^{-9}$
Horz. $1 \times 10^{-2} \lambda_0/D$	$1.2 \times 10^{-10}$	$3.4 \times 10^{-9}$
Horz. $2 \times 10^{-2} \lambda_0/D$	$4.8 \times 10^{-10}$	$8.5 \times 10^{-9}$
Horz. $4 \times 10^{-2} \lambda_0/D$	$1.9 \times 10^{-9}$	$2.4 \times 10^{-8}$
Vert. $5 \times 10^{-3} \lambda_0/D$	$5.8 \times 10^{-11}$	$3.2 \times 10^{-9}$
Vert. $1 \times 10^{-2} \lambda_0/D$	$2.3 \times 10^{-10}$	$8.0 \times 10^{-10}$
Vert. $2 \times 10^{-2} \lambda_0/D$	$9.3 \times 10^{-10}$	$2.3 \times 10^{-9}$
Vert. $4 \times 10^{-2} \lambda_0/D$	$3.8 \times 10^{-9}$	$7.6 \times 10^{-9}$

the contrast degradation at the red edge of the characterization filter. The impact over the rest of the band is much smaller: with a 1.6 milliarc second pointing error, for example, the intensity increment at the interior of the bowtie averaged over seven wavelength samples is  $3.6 \times 10^{-9}$ , a factor of 10 below the increment at the red extreme. It remains to be seen through integrated modeling on the full scope of the observatory, including low-order wavefront sensing and control and data postprocessing, how tip-tilt error affects the science yield of the characterization SPLC. For example, there may be an advantage in redesigning the apodizer for a passband extended slightly beyond the actual spectrograph filter in order to provide a buffer against pointing or alignment errors, at some cost in throughput.

Errors in the lateral alignment between the apodizer and the FPM have a comparable impact on the on-axis intensity as line-of-sight pointing errors. We summarize the results of trial offsets in Table 6. We quantify this in two ways. First, we compute the mean change in contrast over all the spatial samples and wavelength samples in the dark bowtie region. Then we isolate the region of the image that is most severely affected: the inner part of the bowtie (within  $3.8\lambda_0/D$ ), at the red edge of the bandpass. Offsets of up to  $4 \times 10^{-2}\lambda_0/D$  cause a mean contrast degradation of only 2 to  $4 \times 10^{-9}$ . However, in the most sensitive part of the image, for a horizontal offset of  $2 \times 10^{-2}\lambda_0/D$ , the contrast increment approaches  $10^{-8}$ .

**Table 7** Contrast degradation of WFIRST-AFTA characterization SPLC for horizontal Lyot stop alignment errors.

Lyot stop translation (% pupil diam.)	Mean $\Delta$ Contrast	Max. $\Delta$ Contrast
0.5	$2.1 \times 10^{-10}$	$1.3 \times 10^{-9}$
1.0	$9.2 \times 10^{-10}$	$5.5 \times 10^{-9}$
2.0	$4.7 \times 10^{-9}$	$2.5 \times 10^{-8}$

We have also tested the effect of bowtie FPM clocking errors on the characterization SPLC performance. The clocking angle of the bowtie FPM in the first focal plane needs to be accurate to within 0.5 deg to keep the worst contrast degradation (over all spatial samples and wavelengths) below  $5 \times 10^{-9}$ .

### 4.3 Lyot Stop

We test the Lyot stop alignment tolerance of the WFIRST-AFTA characterization SPLC by modeling the propagation of flat, on-axis wavefronts when the Lyot stop is translated off-center. Over 7 wavelength samples and 27 angular separations spaced at  $(\lambda_0/D)/4$ , we compute the azimuthally averaged contrast. Then we compute the mean and maximum increment relative to the nominal contrast values over all those wavelengths and separations. We find that the coronagraph performance is more sensitive to horizontal than vertical Lyot stop translations (here horizontal and vertical orientations are used in the same sense as the diagram in Fig. 10). In Table 7, we summarize the effect for horizontal translations of 0.5, 1.0, and 2.0% of the pupil diameter.

The FPM of the WFIRST-AFTA characterization SPLC is opaque outside the optimized bowtie region (Fig. 10). Any such diaphragm in the first focal plane has a low-pass filter effect on the morphology of the field in the Lyot plane, as we showed in Eq. (5) for the simple circular case. The effect can also be examined visually in the evaluation plots of the type-b SPLC configurations in Sec. 2. The lack of sharp field transitions near the Lyot stop edges helps to keep the alignment tolerance reasonable, in spite of the very high contrast goal in the final image.

This smooth Lyot field characteristic is not the case for the WFIRST-AFTA debris disk design we presented in Sec. 3.3, which operates with an occulting spot FPM. Initial calculations indicate that its Lyot stop alignment tolerance is at least an order of magnitude tighter than that of the diaphragm FPM variants. It is possible that an expanded optimization procedure can counteract this sensitivity. For example, if the optimizer model propagates the field not only through perfectly aligned masks but also through a set of cases with translated Lyot stops, then the final field could be constrained simultaneously for misaligned mask scenarios. However, if that approach is not feasible, then there is a substantial practical advantage for SPLC designs optimized for a diaphragm-type FPM, despite the fact that their throughput is in most cases lower for the same image constraints (e.g., comparing the metrics of configs. IVa and IVb in Table 1, and the debris disk designs in Table 5).

## 5 Conclusion

We have described a hybrid coronagraph configuration that uses a shaped pupil as the apodizing mask in a Lyot-style architecture. An optimized SPLC reaches the contrast and IWA of the well-established APLC design family, while benefitting from a precise, achromatic transmission characteristic that is most feasibly implemented with a binary apodizer.<sup>59</sup>

Our numerical optimization experiments have revealed a rich parameter space in the Lyot stop transmission profile. The apodizer and Lyot stop can be optimized simultaneously, leading to solutions with higher throughput and a sharper PSF for a given contrast and bandwidth. For example, we noted one design (Fig. 7) that surpasses  $10^{-9}$  contrast starting from an angular separation of  $2\lambda_0/D$  while maintaining an FWHM throughput of 17% over a 10% bandpass. At present, however, due to



optimizer limitations, the approach is only feasible for telescope apertures with pure circular symmetry.

The SPLC is compatible with two types of FPM: a conventional occulting spot and an annular/bowtie diaphragm. Once the Lyot stop is tuned, the throughput is generally higher for the occulting spot solutions. However, alignment and manufacturing tolerances may hinder their practicality, due to sharp field features in the Lyot plane originating from the binary profile of the apodizer. By distinction, the low-pass filter effect of the diaphragm FPM dramatically relaxes the tolerance on the Lyot stop profile accuracy. Future efforts will determine if an expanded optimization procedure can produce occulting spot solutions that are less sensitive to this effect.

By applying the same design principles tested for the circular case, we explored the parameter space of SPLC solutions for WFIRST-AFTA. We arrived at a mask scheme optimized for the spectroscopy mode of the coronagraph. This design produces a bowtie-shaped ( $2 \times 65$  deg), quasi-achromatic dark region of  $<10^{-8}$  contrast over an 18% bandwidth, with an IWA of  $2.8\lambda_0/D$  (0.19 arc second at  $\lambda_0 = 770$  nm). Experiments at JPL HCIT are underway to test the ability of this coronagraph to meet the exoplanet characterization goals of the mission.<sup>58</sup> We will also evaluate promising designs for a wider-angle disk imaging mode, operating from  $6.5\lambda_0/D$  to  $20\lambda_0/D$  (angular separations of 0.3 to 1.0 arc second at  $\lambda_0 = 565$  nm).

We limited the practical aspects of this study to the WFIRST-AFTA mission concept, but SPLC designs have broad applicability to high-contrast imaging problems with obscured telescope apertures. Upcoming work by N'Diaye et al. will survey SPLC solutions for segmented space telescope apertures.

## Appendix: Optimization Scheme

### A1 Circular aperture shaped pupil Lyot coronagraph

For each circular shaped pupil Lyot coronagraph (SPLC) configuration, a discrete, algebraic propagation model enables us to exactly define the optimization objectives and constraints we explored in Sec. 2. We mimic the notation used in past descriptions of conventional (non-Lyot) shaped pupil coronagraph optimizations.<sup>35,36</sup> As in those cases, we code the algebraic model and design goals as a linear program in the AMPL programming language. For each of the circular SPLC experiments, we used the LOQO interior point solver<sup>98</sup> to solve the AMPL program and obtain the mask solution.

Due to the circular symmetry of the telescope pupil, the on-axis (stellar) scalar field in each coronagraph plane is expressed as a purely real, one-dimensional, radial function. The numerical Fourier propagation between the coronagraph planes is computed via the discrete Hankel transform. We use spatial coordinate  $r_i$  in the reimaged telescope pupil and  $\xi_j$  in the image plane. The image coordinate maps to a true physical radius and does not scale with wavelength. A unitless wavelength ratio  $\gamma_k = \lambda_k/\lambda_0$ , where  $\lambda_0$  is the center wavelength, captures the chromatic dependence of the field.

We use the variable  $\mathcal{A}_{\text{SP}}$  to represent the radial transmission function of the shaped pupil apodizer and  $\mathcal{A}_{\text{LS}}$  to represent the Lyot stop. The variables  $\Psi_B$ ,  $\Psi_C$ , and  $\Psi_D$  represent the scalar fields in the first focal plane, Lyot plane, and final focal plane, respectively. For pupil plane variables (namely,  $\mathcal{A}_{\text{SP}}$ ,  $\mathcal{A}_{\text{LS}}$ , and  $\Psi_C$ ), the radial coordinate  $r_i$  is normalized to the reimaged

telescope pupil diameter. Therefore, if there are  $N_R$  points across the pupil radius, spaced at intervals of  $\Delta r = \frac{1}{2}/N_R$ , then the radial samples occur at  $r_i = (i - 1/2)\Delta r$ , for integers  $i = 1, 2, \dots, N_R$ .

#### A1.1 Focal occulting spot

In the case where the FPM is an occulting spot (configs. Ia, IIa, IIIa, and IVa in Sec. 2), we use the semianalytical APLC modeling approach of Soummer et al.<sup>99</sup> In the first focal plane, we compute the field only within the occulting spot rather than the (ideally) unbounded transmitted region. Then Babinet's superposition principle can be applied to determine the Lyot plane field, as expressed before in Eq. (2). We sample the field at  $N_{\rho_0}$  points within the spot radius  $\rho_0$ , at spacing  $\Delta\xi = \rho_0/N_{\rho_0}$ . Then for integers  $j = 1, 2, \dots, N_{\rho_0}$ , we compute the interior field at image radii  $\xi_j = (j - 1/2)\rho_0/N_{\rho_0}$ , in units of center wavelength resolution elements. The expressions for the field in the first focal plane and the Lyot plane are as follows:

$$\begin{aligned} \Psi_B(\xi_j, \gamma_k) &= 2\pi/\gamma_k \sum_{i=1}^{N_R} r_i \mathcal{A}_{\text{SP}}(r_i) J_0(2\pi\xi_j r_i/\gamma_k) \Delta r, \\ \Psi_C(r_i, \gamma_k) &= \mathcal{A}_{\text{SP}}(r_i) - 2\pi/\gamma_k \sum_{j=1}^{N_{\rho_0}} \xi_j J_0(2\pi r_i \xi_j/\gamma_k) \\ &\quad \times \Psi_B(\xi_j, \gamma_k) \Delta\xi. \end{aligned} \quad (6)$$

In our trials in Sec. 2, the sampling interval is as fine as  $\Delta\xi = (1/16)$  in the first focal plane and  $\Delta r = \frac{1}{2}/2000$  in the pupil planes. For config. Ia, the Lyot plane  $\Psi_C$  is the last stage of propagation computed by the optimizer. As in the case of the conventional APLC, the on-axis field is constrained here.<sup>21</sup> Our goal is to maximize the sum of the apodizer mask field transmission over the pupil area while meeting some level of on-axis field cancellation. Since the design is monochromatic,  $\Psi_C$  is only computed and constrained at  $\gamma_k = 1$ . Now we have the elements needed to declare the optimization objective, along with the design constraints:

$$\begin{aligned} \text{Maximize } \mathcal{O}_I &= 2\pi \sum_{i=1}^{N_R} r_i \mathcal{A}_{\text{SP}}(r_i) \Delta r, \\ \text{subject to: } &0 \leq \mathcal{A}_{\text{SP}}(r_i) \leq 1, \quad \text{and} \\ &-10^{-s} \leq \Psi_C(r_i) \leq 10^{-s}, \quad \text{for } 0 \leq r_i \leq \frac{1}{2}. \end{aligned} \quad (7)$$

The parameter in the field cancellation exponent,  $s$ , is set to 3.0 in the case illustrated in Sec. 2.1.

In order to more directly prescribe the performance, as we do for configs. IIa to IVa, the optimization model must propagate the field from the Lyot plane to the final focal plane of the coronagraph. Here, we switch the spatial coordinate variable from  $\xi$  to  $\zeta$  to indicate a change in the radial sampling. The new sampling interval,  $\Delta\zeta$ , must be no larger than 1/2 of a center wavelength resolution element (to meet the Nyquist–Shannon sampling criterion) and preferably close to 1/4. The expression for the scalar electric field in the final plane is

$$\Psi_D(\zeta_j, \gamma_k) = 2\pi/\gamma_k \sum_{i=1}^{N_R} r_i \mathcal{A}_{LS}(r_i) \Psi_C(r_i, \gamma_k) J_0(2\pi\zeta_j r_i / \gamma_k) \Delta r, \quad (8)$$

computed at radii  $\zeta_j = (j - \frac{1}{2})\Delta\zeta$  for integer indices  $j$  satisfying  $\rho_0 \leq \zeta_j \leq \rho_1$ . For configs. IIa/IIb, the Lyot stop is a replica of the telescope pupil; therefore,  $\mathcal{A}_{LS}$  is equal to unity for all radii in the summation bounds. However, for configs. IVa/IVb, the annular Lyot stop is equal to unity for  $0.1 \leq r_i/\frac{1}{2} \leq 0.9$  and zero elsewhere.

Our goal again is to maximize the integrated field transmission of the apodizer mask  $\mathcal{A}_{SP}$ . This time, however, the on-axis field is constrained in an annular region of the final image. For each wavelength  $\gamma_k$  sampling the operating bandwidth, we compute the peak field in the first focal plane. This value is a proxy for the star's peak intensity and serves as a reference for the contrast constraints:

$$\Psi_{B_{\text{Peak}}}(\gamma_k) = 2\pi/\gamma_k \sum_{i=1}^{N_R} r_i \mathcal{A}_{SP}(r_i) \Delta r. \quad (9)$$

Now we can declare the optimization objective and the design constraints:

$$\begin{aligned} \text{Maximize } \mathcal{O}_I &= 2\pi \sum_{i=1}^{N_R} r_i \mathcal{A}_{SP}(r_i) \Delta r, \\ \text{subject to: } &0 \leq \mathcal{A}_{SP}(r_i) \leq 1, \quad \text{for } 0 \leq r_i \leq \frac{1}{2}, \quad \text{and} \\ &-10^{-c/2} \leq \frac{\Psi_D(\zeta_j, \gamma_k)}{\Psi_{B_{\text{Peak}}}(\gamma_k)} \leq 10^{-c/2} \quad \text{for } \rho_0 \leq \zeta_j \leq \rho_1 \quad \text{and} \\ &1 - w/2 \leq \gamma_k \leq 1 + w/2. \end{aligned} \quad (10)$$

The  $c$  parameter in the exponent is the base 10 logarithm of the desired contrast in intensity. In practice,  $c$  must be increased slightly above this specification to compensate for the off-axis field attenuation caused by the Lyot stop. The parameter  $w$  defining the wavelength bounds is the fractional operating bandwidth (equal to 0.1 for most trials in Sec. 2). By repeating identical constraints at multiple wavelength samples, the true spatial dimensions of the dark search region (and equivalently, its angular projection on the sky) are fixed across the operating bandwidth. Similar achromatization procedures have been applied to APLC designs.<sup>26,27</sup> At the 10% bandwidth we investigated for the circular aperture, three wavelength samples suffice to maintain a broadband null at  $10^{-9}$  contrast.

If, as in config. IIIa, we define the Lyot stop as a variable rather than a fixed parameter, then the optimization objective must take into account the transmission of two masks. In our trials described in Sec. 2.4, we weight them equally:

$$\begin{aligned} \text{Maximize } \mathcal{O}_{II} &= 2\pi \sum_{i=1}^{N_R} r_i [\mathcal{A}_{SP}(r_i) + \mathcal{A}_{LS}(r_i)] \Delta r, \\ \text{subject to: } &0 \leq \mathcal{A}_{SP}(r_i) \leq 1, \quad \text{for } 0 \leq r_i \leq \frac{1}{2}, \quad \text{and} \\ &0 \leq \mathcal{A}_{LS}(r_i) \leq 1, \quad \text{for } 0 \leq r_i \leq \frac{1}{2}. \end{aligned} \quad (11)$$

Note that in the case where the Lyot stop is a free variable array, the function being constrained by the optimizer ( $\Psi_D$ ) is no longer a linear function of the free variables. That is because each point in the final field is now determined by products of transmission values in the apodizer and Lyot stop. Although some solvers, such as LOQO, are flexible in accepting non-linear, nonconvex programs, convergence on a solution is not guaranteed.

## A1.2 Focal diaphragm

For the configurations where the focal plane mask (FPM) is a diaphragm rather than a spot, our computational approach is distinct. Now, the transmitted region between radii  $\rho_0$  and  $\rho_1$  is the part of the field computed in the first focal plane, which is in turn directly propagated to the Lyot plane:

$$\begin{aligned} \Psi_B(\xi_j, \gamma_k) &= 2\pi/\gamma_k \sum_{i=1}^{N_R} r_i \mathcal{A}_{SP}(r_i) J_0(2\pi\xi_j r_i / \gamma_k) \Delta r, \\ \Psi_C(r_i, \gamma_k) &= 2\pi/\gamma_k \sum_{j=M_{\rho_0}}^{M_{\rho_1}} \xi_j J_0(2\pi r_i \xi_j / \gamma_k) \Psi_B(\xi_j, \gamma_k) \Delta \xi, \\ \Psi_D(\zeta_j, \gamma_k) &= 2\pi/\gamma_k \sum_{i=1}^{N_R} r_i \mathcal{A}_{LS}(r_i) \Psi_C(r_i, \gamma_k) J_0(2\pi\zeta_j r_i / \gamma_k) \Delta r. \end{aligned} \quad (12)$$

$M_{\rho_0}$  and  $M_{\rho_1}$  correspond, respectively, to the lowest and highest integers  $j$  satisfying  $\rho_0 \leq \xi_j \leq \rho_1$ , where  $\xi_j = (j - 1/2)\Delta\xi$ . The definitions for the optimization objective and constraints given for the spot FPM configuration, Eqs. (7), (10), and (11), remain valid for the corresponding configs. Ib, IIb, and IIIb, respectively.

## A2 WFIRST-AFTA shaped pupil Lyot coronagraph

The same approach we used to define a discrete, algebraic propagation model for the clear circular aperture shaped pupil Lyot coronagraph (SPLC) can be applied to an arbitrary telescope aperture. However, the propagation now relies on two-dimensional Fourier transforms rather than Hankel transforms, resulting in a combination of real and imaginary scalar field components. We again code the linear program in AMPL. However, we use the Gurobi<sup>100</sup> package to implement the solver algorithm instead of LOQO, since it better accommodates the much larger size of the two-dimensional problem.

At each stage, we expand the complex exponential of the discrete Fourier transform into cosine and sine terms; doing so reveals simplifications arising from the geometric symmetry of the telescope pupil, thereby reducing the computational complexity and speeding up the optimization. We align the telescope pupil (Fig. 9) so that one of its three symmetry axes coincides with the vertical axis ( $y$ ) in our Cartesian representation. In the first stage of the propagation, this enables us to restrict the bounds of the horizontal Riemann sum to one half of the pupil plane and also to drop sine terms with a horizontal dependence. The field in the first focal plane is then

$$\begin{aligned}\Psi_{B_{\text{Re}}}(\xi_u, \eta_v, \gamma_k) &= 2/\gamma_k \sum_{j=-N_y}^{N_y} \cos(2\pi\eta_v y_j / \gamma_k) \sum_{i=1}^{N_x} \mathcal{A}_{\text{SP}}(x_i, y_j) \\ &\quad \times \cos(2\pi\xi_u x_i / \gamma_k) \Delta x \Delta y, \\ \Psi_{B_{\text{Im}}}(\xi_u, \eta_v, \gamma_k) &= 2/\gamma_k \sum_{j=-N_y}^{N_y} \sin(2\pi\eta_v y_j / \gamma_k) \sum_{i=1}^{N_x} \mathcal{A}_{\text{SP}}(x_i, y_j) \\ &\quad \times \cos(2\pi\xi_u x_i / \gamma_k) \Delta x \Delta y.\end{aligned}\quad (13)$$

The real and imaginary components of the field are distinguished with subscripts Re and Im. Combining the facts that  $\mathcal{A}_{\text{SP}}$  is real and symmetric about the vertical axis, it can be shown that (1)  $\Psi_{B_{\text{Re}}}(\xi_u, \eta_v, \gamma_k)$  has even symmetry over  $\xi$  and  $\eta$  and (2)  $\Psi_{B_{\text{Im}}}(\xi_u, \eta_v, \gamma_k)$  has even symmetry over  $\xi$  and odd symmetry over  $\eta$ . Using these symmetry properties, we need only evaluate the Riemann sums for  $\Psi_{B_{\text{Re}}}$  and  $\Psi_{B_{\text{Im}}}$  in one quadrant of the focal plane to determine the full field.

As before, the horizontal and vertical coordinates are normalized to the reimaged telescope pupil diameter. The fabrication process for the WFIRST-AFTA shaped pupils assumes a binary mask array 1000 pixels in diameter.<sup>59</sup> Therefore, in order to optimize testbed-ready designs, as in the case of the characterization design presented in Sec. 3.2, we set  $N_x$  and  $N_y$  to 500,

$$\begin{aligned}\Psi_C(x_i, y_j, \gamma_k) &= \mathcal{A}_{\text{SP}}(x_i, y_j) - 4/\gamma_k \sum_{v=1}^{N_{\xi_0}} \cos(2\pi\eta_v y_i / \gamma_k) \sum_{u=1}^{N_{\xi_0}} M(\xi_u, \eta_v) \Psi_{B_{\text{Re}}}(\xi_u, \eta_v, \gamma_k) \cos(2\pi\xi_u x_i / \gamma_k) \Delta\xi \Delta\eta \\ &\quad + 4/\gamma_k \sum_{v=1}^{N_{\eta_0}} \sin(2\pi\eta_v y_i / \gamma_k) \sum_{u=1}^{N_{\xi_0}} M(\xi_u, \eta_v) \Psi_{B_{\text{Im}}}(\xi_u, \eta_v, \gamma_k) \cos(2\pi\xi_u x_i / \gamma_k) \Delta\xi \Delta\eta, \\ \Psi_{D_{\text{Re}}}(\zeta_u, \mu_v, \gamma_k) &= 2/\gamma_k \sum_{j=-N_y}^{N_y} \cos(2\pi\mu_v y_j / \gamma_k) \sum_{i=1}^{N_x} \mathcal{A}_{\text{LS}}(x_i, y_j) \Psi_C(x_i, y_j, \gamma_k) \cos(2\pi\zeta_u x_i / \gamma_k) \Delta x \Delta y \\ \Psi_{D_{\text{Im}}}(\zeta_u, \mu_v, \gamma_k) &= 2/\gamma_k \sum_{j=-N_y}^{N_y} \sin(2\pi\mu_v y_j / \gamma_k) \sum_{i=1}^{N_x} \mathcal{A}_{\text{LS}}(x_i, y_j) \Psi_C(x_i, y_j, \gamma_k) \cos(2\pi\zeta_u x_i / \gamma_k) \Delta x \Delta y.\end{aligned}\quad (14)$$

The field in the Lyot plane,  $\Psi_C$ , is real and symmetric about the vertical axis. The final focal plane field,  $\Psi_D$ , retains the same symmetry properties as  $\Psi_B$ , so again it is most efficient to only evaluate one quadrant.

In our investigations of WFIRST-AFTA solutions, we found that the spatial sampling in the first focal plane is especially critical for maintaining the accuracy of designs with a small inner working angle. When the FPM has an inner radius below  $3\lambda_0/D$ , a resolution of  $\Delta\xi = \Delta\eta = \frac{1}{8}$  is needed to ensure agreement with high-resolution evaluations of the solution.

Like the circular SPLC, we use the central peak in the first focal plane as a proxy for the star's flux:

$$\Psi_{B_{\text{Peak}}}(\gamma_k) = 2/\gamma_k \sum_{j=-N_y}^{N_y} \sum_{i=1}^{N_x} \mathcal{A}_{\text{SP}}(x_i, y_j) \Delta x \Delta y.\quad (15)$$

Then the optimization objective and constraints are defined as follows:

and  $\Delta x$  and  $\Delta y$  to  $1/1000$  in Eq. (13). However, for efficient parameter surveys, the spatial resolution can be much coarser, for example,  $\Delta x = 1/256$ .

### A2.1 Focal occulting spot

Similar to the circular SPLC, the region within the quadrant where we evaluate  $\Psi_B$  depends on the FPM configuration. For the occulting spot FPM, the field is evaluated only in the interior of the occulting spot, since Babinet's principle applies conveniently again when propagating to the Lyot plane.

We represent the discretized profile of the FPM explicitly by the variable array  $M(\xi_u, \eta_v)$ . Consistent with the convention used in Eqs. (1) and (2),  $M(\xi_u, \eta_v)$  is the complement of the mask transmission: zero-valued in the transmitted region and unity in the occulted region. As is necessary in order to approximate round and diagonal features on a Cartesian grid,  $M(\xi_u, \eta_v)$  takes on gray values between 0 and 1 at the edges of features in the mask profile in proportion to the fraction of area occulted on the mask array pixel.

If we sample the interior of the occulting spot of radius  $\rho_0$  with  $N_{\xi_0}$  horizontal samples at interval  $\Delta\xi$  and  $N_{\eta_0}$  vertical samples at interval  $\Delta\eta$ , then the on-axis field propagations to the Lyot plane and final focal plane are modeled as follows:

$$\begin{aligned}\text{Maximize } \mathcal{O}_I &= 2 \sum_{j=-N_y}^{N_y} \sum_{i=1}^{N_x} \mathcal{A}_{\text{SP}}(x_i, y_j) \Delta x \Delta y, \\ \text{subject to: } &0 \leq \mathcal{A}_{\text{SP}}(x_i, y_j) \leq \mathcal{A}_{\text{TEL}}, \text{ for } 0 \leq x_i \leq \frac{1}{2}, -\frac{1}{2} \leq y_j \leq \frac{1}{2}, \\ &\text{and} \\ &-10^{-c/2}/\sqrt{2} \leq \Psi_{D_{\text{Re}}}(\zeta_u, \mu_v, \gamma_k) / \Psi_{B_{\text{Peak}}}(\gamma_k) \leq 10^{-c/2}/\sqrt{2}, \\ &-10^{-c/2}/\sqrt{2} \leq \Psi_{D_{\text{Im}}}(\zeta_u, \mu_v, \gamma_k) / \Psi_{B_{\text{Peak}}}(\gamma_k) \leq 10^{-c/2}/\sqrt{2}, \\ &\text{for } \rho_0 \leq \sqrt{\zeta_u^2 + \mu_v^2} \leq \rho_1, \text{ and } 1 - w/2 \leq \gamma_k \leq 1 + w/2.\end{aligned}\quad (16)$$

The variable array  $\mathcal{A}_{\text{TEL}}$  represents the transmission of the telescope pupil, including its central obstruction and support struts, as illustrated in Fig. 9. This condition forces all points in the pupil already obstructed by the telescope to remain opaque in the shaped pupil apodizer solution. When defining  $\mathcal{A}_{\text{TEL}}$ , we pad the telescope obstruction features by 0.25% of the pupil diameter in order to allow for some alignment error between the shaped pupil apodizer and the relay optics.

## A2.2 Focal diaphragm

In the propagation model for the diaphragm FPM configuration, the region of the first focal plane quadrant with nonzero transmission is the only one we compute. In the case of the characterization SPLC design for WFIRST-AFTA described in Sec. 3.2, this region is bowtie-shaped; for other designs, it can be annular. For convenience, we define an FPM variable that is the complement of  $M$ :  $\tilde{M}(\xi_u, \eta_v) = 1 - M(\xi_u, \eta_v)$ . Therefore,  $\tilde{M}$  is equal to unity in the transmitted region and zero-valued in the occulted region. Starting from Eq. (13), the on-axis field propagations to the Lyot plane and final focal plane are modeled as follows:

$$\begin{aligned} \Psi_C(x_i, y_j, \gamma_k) &= 4/\gamma_k \sum_{v=1}^{N_{\eta_1}} \cos(2\pi\eta_v y_j / \gamma_k) \sum_{u=1}^{N_{\xi_1}} \tilde{M}(\xi_u, \eta_v) \\ &\quad \times \Psi_{B_{\text{Re}}}(\xi_u, \eta_v, \gamma_k) \cos(2\pi\xi_u x_i / \gamma_k) \Delta\xi \Delta\eta \\ &\quad - 4/\gamma_k \sum_{v=1}^{N_{\eta_1}} \sin(2\pi\eta_v y_j / \gamma_k) \sum_{u=1}^{N_{\xi_1}} \tilde{M}(\xi_u, \eta_v) \\ &\quad \times \Psi_{B_{\text{Im}}}(\xi_u, \eta_v, \gamma_k) \cos(2\pi\xi_u x_i / \gamma_k) \Delta\xi \Delta\eta, \\ \Psi_{D_{\text{Re}}}(\zeta_u, \mu_v, \gamma_k) &= 2/\gamma_k \sum_{j=-N_y}^{N_y} \cos(2\pi\mu_v y_j / \gamma_k) \sum_{i=1}^{N_x} \mathcal{A}_{\text{LS}}(x_i, y_j) \\ &\quad \times \Psi_C(x_i, y_j, \gamma_k) \cos(2\pi\zeta_u x_i / \gamma_k) \Delta x \Delta y, \\ \Psi_{D_{\text{Im}}}(\zeta_u, \mu_v, \gamma_k) &= 2/\gamma_k \sum_{j=-N_y}^{N_y} \sin(2\pi\mu_v y_j / \gamma_k) \sum_{i=1}^{N_x} \mathcal{A}_{\text{LS}}(x_i, y_j) \\ &\quad \times \Psi_C(x_i, y_j, \gamma_k) \cos(2\pi\zeta_u x_i / \gamma_k) \Delta x \Delta y. \end{aligned} \quad (17)$$

The optimization constraints are defined in a manner identical to the previous configuration, except that the points where the contrast is constrained in the image need to be matched to the profile of the FPM rather than an assumed annular region:

$$\begin{aligned} -10^{-c/2}/\sqrt{2} &\leq \Psi_{D_{\text{Re}}}(\zeta_u, \mu_v, \gamma_k) / \Psi_{B_{\text{Peak}}}(\gamma_k) \leq 10^{-c/2}/\sqrt{2}, \\ 10^{-c/2}/\sqrt{2} &\leq \Psi_{D_{\text{Im}}}(\zeta_u, \mu_v, \gamma_k) / \Psi_{B_{\text{Peak}}}(\gamma_k) \leq 10^{-c/2}/\sqrt{2}, \\ \text{for } (\zeta_u, \mu_v) &\text{ such that } \tilde{M}(\zeta_u, \mu_v) > 0, \quad \text{and} \\ 1 - w/2 &\leq \gamma_k \leq 1 + w/2. \end{aligned} \quad (18)$$

For the WFIRST-AFTA designs presented with a 10% bandwidth ( $w = 0.10$ ), we constrained the contrast at five wavelengths. For the 18% bandwidth designs ( $w = 0.18$ ), we constrained the contrast at seven wavelengths.

## Acknowledgments

We are grateful to Mamadou N'Diaye and Rémi Soummer for helpful suggestions and conversations about coronagraph design. We also thank Wesley A. Traub for providing us with his program to estimate the separation–contrast distribution of radial velocity planets. This work was partially funded by NASA Technology Development for Exoplanet Missions Grant Number NNX09AB96G and by the Jet Propulsion Laboratory of the California Institute of Technology. A. J. E. Riggs is

supported by NASA Grant Number NNX14AM06H, a NASA Space Technology Research Fellowship. Robert J. Vanderbei acknowledges research support from the Office of Naval Research.

## References

1. S. Hinkley et al., “A new high contrast imaging program at Palomar observatory,” *Publ. Astron. Soc. Pac.* **123**, 74–86 (2011).
2. B. Macintosh et al., “First light of the Gemini Planet Imager,” *Proc. Natl. Acad. Sci.* **111**, 12661–12666 (2014).
3. J. L. Beuzit et al., “SPHERE: a ‘planet finder’ instrument for the VLT,” *Proc. SPIE* **7014**, 701418 (2008).
4. B. R. Oppenheimer et al., “Reconnaissance of the HR 8799 exosolar system. I. Near-infrared spectroscopy,” *Astrophys. J.* **768**, 24 (2013).
5. B. Macintosh et al., “Discovery and spectroscopy of the young jovian planet 51 Eri b with the Gemini Planet Imager,” *Science* **350**, 64–67 (2015).
6. M. A. Millar-Blanchaer et al., “Beta Pictoris’ inner disk in polarized light and new orbital parameters for Beta Pictoris b,” *Astrophys. J.* **811**, 18 (2015).
7. S. P. Quanz et al., “Confirmation and characterization of the protoplanet HD 100546 b—direct evidence for gas giant planet formation at 50 au,” *Astrophys. J.* **807**(1), 64 (2015).
8. J. T. Trauger and W. A. Traub, “A laboratory demonstration of the capability to image an Earth-like extrasolar planet,” *Nature* **446**, 771–773 (2007).
9. P. R. Lawson et al., “Survey of experimental results in high-contrast imaging for future exoplanet missions,” *Proc. SPIE* **8864**, 88641F (2013).
10. J. Mazoyer et al., “High-contrast imaging in polychromatic light with the self-coherent camera,” *Astron. Astrophys.* **564**, L1 (2014).
11. M. N’Diaye et al., “High-contrast imager for complex aperture telescopes (HiCAT): 3. First lab results with wavefront control,” *Proc. SPIE* **9605**, 96050I (2015).
12. B. Lyot, “The study of the solar corona and prominences without eclipses (George Darwin Lecture, 1939),” *Mon. Not. R. Astron. Soc.* **99**, 580 (1939).
13. D. A. Golimowski et al., “High-resolution ground-based coronagraphy using image-motion compensation,” *Appl. Opt.* **31**, 4405–4416 (1992).
14. T. Nakajima, “Planet detectability by an adaptive optics stellar coronagraph,” *Astrophys. J.* **425**, 348–357 (1994).
15. F. Malbet, “High angular resolution coronagraphy for adaptive optics,” *Astron. Astrophys. Suppl. Ser.* **115**, 161 (1996).
16. D. Mouillet et al., “A stellar coronagraph for the COME-ON-PLUS adaptive optics system. II. First astronomical results,” *Astron. Astrophys.* **324**, 1083–1090 (1997).
17. A. Sivaramakrishnan et al., “Ground-based coronagraphy with high-order adaptive optics,” *Astrophys. J.* **552**, 397–408 (2001).
18. B. R. Oppenheimer et al., “The Lyot project: toward exoplanet imaging and spectroscopy,” *Proc. SPIE* **5490**, 433–442 (2004).
19. O. Guyon and F. J. Roddier, “Direct exoplanet imaging possibilities of the nulling stellar coronagraph,” *Proc. SPIE* **4006**, 377–387 (2000).
20. C. Aime, R. Soummer, and A. Ferrari, “Total coronagraphic extinction of rectangular apertures using linear prolate apodizations,” *Astron. Astrophys.* **389**, 334–344 (2002).
21. R. Soummer, C. Aime, and P. E. Falloon, “Stellar coronagraphy with prolate apodized circular apertures,” *Astron. Astrophys.* **397**, 1161–1172 (2003).
22. R. Soummer, “Apodized pupil Lyot coronagraphs for arbitrary telescope apertures,” *Astrophys. J.* **618**, L161–L164 (2005).
23. C. Aime, “Principle of an achromatic prolate apodized Lyot coronagraph,” *Publ. Astron. Soc. Pac.* **117**, 1012 (2005).
24. L. Abe et al., “Multi-stage apodized pupil Lyot coronagraph experimental results,” *Proc. SPIE* **7014**, 701467 (2008).
25. R. Soummer et al., “Apodized pupil Lyot coronagraphs for arbitrary apertures. II. Theoretical properties and application to extremely large telescopes,” *Astrophys. J.* **695**, 695–706 (2009).
26. R. Soummer et al., “Apodized pupil Lyot coronagraphs for arbitrary apertures. III. Quasi-achromatic solutions,” *Astrophys. J.* **729**(2), 144 (2011).



27. M. N'Diaye, L. Pueyo, and R. Soummer, "Apodized pupil Lyot coronagraphs for arbitrary apertures. IV. Reduced inner working angle and increased robustness to low-order aberrations," *Astrophys. J.* **799**, 225 (2015).
28. J. F. W. Herschel, *Results of astronomical observations made during the years 1834, 5, 6, 7, 8, at the Cape of Good Hope; being the completion of a telescopic survey of the whole surface of the visible heavens, commenced in 1825*, Smith, Elder and Co., London (1847).
29. E. E. Barnard, "Observations of the companions of Procyon," *Astron. Nachr.* **182**, 13 (1909).
30. D. N. Spergel, "A new pupil for detecting extrasolar planets," ArXiv Astrophysics e-prints astro-ph/0101142, <http://arxiv.org/abs/astro-ph/0101142> (2001).
31. N. J. Kasdin et al., "Extrasolar planet finding via optimal apodized-pupil and shaped-pupil coronagraphs," *Astrophys. J.* **582**, 1147–1161 (2003).
32. R. J. Vanderbei, D. N. Spergel, and N. J. Kasdin, "Spiderweb masks for high-contrast imaging," *Astrophys. J.* **590**, 593–603 (2003).
33. N. J. Kasdin et al., "Optimal one-dimensional apodizations and shaped pupils for planet finding coronagraphy," *Appl. Opt.* **44**, 1117–1128 (2005).
34. K. Enya and L. Abe, "A binary shaped mask coronagraph for a segmented pupil," *Publ. Astron. Soc. Jpn.* **62**, 1407 (2010).
35. A. Carlotti, R. Vanderbei, and N. J. Kasdin, "Optimal pupil apodizations of arbitrary apertures for high-contrast imaging," *Opt. Express* **19**, 26796–26809 (2011).
36. R. J. Vanderbei, "Fast Fourier optimization: sparsity matters," *Math. Program. Comput.* **4**, 53–69 (2012).
37. A. Carlotti et al., "Optimized shaped pupil masks for pupil with obscuration," *Proc. SPIE* **8442**, 844254 (2012).
38. A. Carlotti et al., "Shaped pupil coronagraphy with WFIRST-AFTA," *Proc. SPIE* **8864**, 886410 (2013).
39. K. Haze et al., "Experimental demonstration of binary shaped pupil mask coronagraphs for telescopes with obscured pupils," *Publ. Astron. Soc. Jpn.* **67**, 28 (2015).
40. F. Roddier and C. Roddier, "Stellar coronagraph with phase mask," *Publ. Astron. Soc. Pac.* **109**, 815–820 (1997).
41. D. Rouan et al., "The four-quadrant phase-mask coronagraph. I. Principle," *Publ. Astron. Soc. Pac.* **112**, 1479–1486 (2000).
42. D. Mawet et al., "Optical vectorial vortex coronagraphs using liquid crystal polymers: theory, manufacturing and laboratory demonstration," *Opt. Express* **17**, 1902–1918 (2009).
43. O. Guyon et al., "Exoplanet imaging with a phase-induced amplitude apodization coronagraph. I. Principle," *Astrophys. J.* **622**, 744–758 (2005).
44. L. Pueyo and C. Norman, "High-contrast imaging with an arbitrary aperture: active compensation of aperture discontinuities," *Astrophys. J.* **769**, 102 (2013).
45. D. Mawet et al., "Annular groove phase mask coronagraph," *Astrophys. J.* **633**, 1191–1200 (2005).
46. D. Mawet et al., "Ring-apodized vortex coronagraphs for obscured telescopes. I. Transmissive ring apodizers," *Astrophys. J. Suppl. Ser.* **209**, 7 (2013).
47. A. Carlotti, L. Pueyo, and D. Mawet, "Apodized phase mask coronagraphs for arbitrary apertures. II. Comprehensive review of solutions for the vortex coronagraph," *Astron. Astrophys.* **566**, A31 (2014).
48. O. Guyon et al., "High performance Lyot and PIAA coronagraphy for arbitrarily shaped telescope apertures," *Astrophys. J.* **780**, 171 (2014).
49. G. J. Ruane et al., "Optimized focal and pupil plane masks for vortex coronagraphs on telescopes with obstructed apertures," *Proc. SPIE* **9605**, 96051I (2015).
50. O. Guyon et al., "Theoretical limits on extrasolar terrestrial planet detection with coronagraphs," *Astrophys. J. Suppl. Ser.* **167**, 81–99 (2006).
51. O. Guyon, "Coronagraphic techniques for direct imaging of exoplanets and disks," in *American Institute of Physics Conference Series*, T. Usuda, M. Tamura, and M. Ishii, Eds., American Institute of Physics Conference Series 1158, pp. 285–292 (2009).
52. M. Levine et al., "Overview of technologies for direct optical imaging of exoplanets," in *astro2010: The Astronomy and Astrophysics Decadal Survey*, *Astronomy 2010* 37 (2009).
53. N. J. Kasdin et al., "Shaped pupil coronagraphs for planet finding: optimization, manufacturing, and experimental results," *Proc. SPIE* **5905**, 59050G (2005).
54. A. E. Lowman et al., "High-contrast imaging testbed for the Terrestrial Planet Finder coronagraph," *Proc. SPIE* **5487**, 1246–1254 (2004).
55. T. D. Groff and N. J. Kasdin, "Kalman filtering techniques for focal plane electric field estimation," *J. Opt. Soc. Am. A* **30**(1), 128–139 (2013).
56. A. J. E. Riggs et al., "Demonstration of symmetric dark holes using two deformable mirrors at the high-contrast imaging testbed," *Proc. SPIE* **8864**, 88640T (2013).
57. A. J. E. Riggs, N. J. Kasdin, and T. D. Groff, "Recursive starlight and bias estimation for high-contrast imaging with an extended Kalman filter," submitted to *J. Astron. Telesc. Instrum. Syst.* (2015).
58. E. J. Cady et al., "Demonstration of high contrast with an obscured aperture with the WFIRST-AFTA shaped pupil coronagraph," *J. Astron. Telesc. Instrum. Syst.* **2**(1), 011004 (2015).
59. K. Balasubramanian et al., "WFIRST-AFTA coronagraph shaped pupil masks: design, fabrication and characterization," *J. Astron. Telesc. Instrum. Syst.* **2**(1), 011005 (2015).
60. D. Spergel et al., "Wide-field infrared survey telescope-astrophysics focused telescope assets (WFIRST-AFTA) 2015 report," [http://wfirst.gsfc.nasa.gov/science/sdt\\_public/WFIRST-AFTA\\_SDT\\_Report\\_150310\\_Final.pdf](http://wfirst.gsfc.nasa.gov/science/sdt_public/WFIRST-AFTA_SDT_Report_150310_Final.pdf) (2015).
61. I. Poberezhskiy et al., "Technology development towards WFIRST-AFTA coronagraph," *Proc. SPIE* **9143**, 91430P (2014).
62. J. Trauger et al., "Hybrid Lyot coronagraph for WFIRST-AFTA: coronagraph design and performance metrics," *J. Astron. Telesc. Instrum. Syst.* **2**(1), 011016 (2015).
63. B. Seo et al., "Hybrid Lyot coronagraph for WFIRST-AFTA: occulter fabrication and high contrast narrowband testbed demonstration," submitted to *J. Astron. Telesc. Instrum. Syst.* (2015).
64. J. Trauger et al., "Complex apodization Lyot coronagraphy for the direct imaging of exoplanet systems: design, fabrication, and laboratory demonstration," *Proc. SPIE* **8442**, 84424Q (2012).
65. O. Guyon et al., "Phase induced amplitude apodization (PIAA) coronagraphy: recent results and future prospects," *Proc. SPIE* **8442**, 84424V (2012).
66. E. A. Pluzhnik et al., "Design of off-axis PIAACMC mirrors," submitted to *J. Astron. Telesc. Instrum. Syst.* (2015).
67. B. Kern et al., "PIAACMC mask fabrication, characterization, and modeling for WFIRST-AFTA," *J. Astron. Telesc. Instrum. Syst.* **2**(1), 011013 (2015).
68. A. J. E. Riggs et al., "Shaped pupil design for future space telescopes," *Proc. SPIE* **9143**, 914325 (2014).
69. M. N'Diaye et al., "High-contrast imager for complex aperture telescopes (HiCAT): APLC/shaped-pupil hybrid coronagraph designs," presented at *American Astronomical Society Meeting 225*, 4–8 January 2015, Seattle, Washington.
70. E. Cady et al., "Shaped pupil design for the Gemini Planet Imager," *Astrophys. J.* **698**, 938–943 (2009).
71. K. Balasubramanian et al., "Fabrication and characteristics of free-standing shaped pupil masks for TPF-coronagraph," *Proc. SPIE* **6265**, 62653N (2006).
72. K. Enya et al., "Comparative study of manufacturing techniques for coronagraphic binary pupil masks: masks on substrates and free-standing masks," *Publ. Astron. Soc. Jpn.* **64**, 123 (2012).
73. A. Sivaramakrishnan et al., "Sensing phase aberrations behind Lyot coronagraphs," *Astrophys. J.* **688**, 701–708 (2008).
74. K. Stapelfeldt et al., "Exo-C STDT final report," Technical Report, NASA, [http://exep.jpl.nasa.gov/stdt/Exo-C\\_Final\\_Report\\_for\\_Unlimited\\_Release\\_150323.pdf](http://exep.jpl.nasa.gov/stdt/Exo-C_Final_Report_for_Unlimited_Release_150323.pdf) (2015).
75. D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty? IV: extensions to many dimensions; generalized prolate spheroidal functions," *Bell Syst. Tech. J.* **43**, 3009–3057 (1964).
76. D. Slepian, "Analytic solution of two apodization problems," *J. Opt. Soc. Am.* **55**, 1110–1114 (1965).
77. R. Bracewell, *Fourier Analysis and Imaging*, Kluwer Academic/Plenum Publishers, New York (2003).
78. J. Krist, B. Nemati, and B. Mennesson, "Numerical modeling of the proposed WFIRST-AFTA coronagraphs and their predicted performances," *J. Astron. Telesc. Instrum. Syst.* **2**(1), 011003 (2015).
79. R. J. Vanderbei, D. N. Spergel, and N. J. Kasdin, "Circularly symmetric apodization via star-shaped masks," *Astrophys. J.* **599**, 686–694 (2003).

80. P. Martinez et al., “Design, analysis, and testing of a microdot apodizer for the apodized pupil Lyot coronagraph,” *Astron. Astrophys.* **495**, 363–370 (2009).
81. P. Martinez et al., “Design, analysis, and testing of a microdot apodizer for the apodized pupil Lyot coronagraph. II. Impact of the dot size,” *Astron. Astrophys.* **500**, 1281–1285 (2009).
82. P. Martinez et al., “Design, analysis, and testing of a microdot apodizer for the apodized pupil Lyot coronagraph. III. Application to extremely large telescopes,” *Astron. Astrophys.* **520**, A110 (2010).
83. P. Martinez et al., “Half-toning for high-contrast imaging: developments for the SPHERE and EPICS instruments,” September 2009, <http://www.eso.org/sci/facilities/develop/ao/sys/messengerCORO.pdf> (9 December 2015).
84. A. Sivaramakrishnan et al., “Gemini Planet Imager coronagraph testbed results,” *Proc. SPIE* **7735**, 773586 (2010).
85. S. J. Thomas et al., “Testing the apodized pupil Lyot coronagraph on the laboratory for adaptive optics extreme adaptive optics testbed,” *Astrophys. J.* **142**, 119 (2011).
86. O. Guyon, “Limits of adaptive optics for high-contrast imaging,” *Astrophys. J.* **629**, 592–614 (2005).
87. M. Marley et al., “A quick study of the characterization of radial velocity giant planets in reflected light by forward and inverse modeling,” ArXiv Astrophysics e-prints astro-ph/1412.8440, <http://arxiv.org/abs/1412.8440> (2014).
88. A. Burrows, “Scientific return of coronagraphic exoplanet imaging and spectroscopy using WFIRST,” ArXiv Astrophysics e-prints astro-ph/1412.6097, <http://arxiv.org/abs/1412.6097> (2014).
89. C. Noecker, F. Zhao, and R. Demers, “Coronagraph instrument for WFIRST-AFTA,” (2015).
90. M. McElwain et al., “WFIRST-AFTA coronagraph IFS,” submitted to *J. Astron. Telesc. Instrum. Syst.* (2015).
91. W. A. Traub et al., “Science yield estimate with the WFIRST-AFTA coronagraph,” submitted to *J. Astron. Telesc. Instrum. Syst.* (2015).
92. G. Schneider, “A quick study of science return from direct imaging exoplanet missions: detection and characterization of circumstellar material with an AFTA or EXO-C/S CGI,” ArXiv Astrophysics e-prints astro-ph/1412.8421, <http://arxiv.org/abs/1412.8421> (2014).
93. M. Ygouf et al., “Data processing and algorithm development for the WFIRST-AFTA coronagraph: reduction of noise free simulated images, analysis and spectrum extraction with reference star differential imaging,” *Proc. SPIE* **9605**, 96050S (2015).
94. W. A. Traub et al., “Science yield estimation for AFTA coronagraphs,” *Proc. SPIE* **9143**, 91430N (2014).
95. L. Abe et al., “Removal of central obscuration and spider arm effects with beam-shaping coronagraphy,” *Astron. Astrophys.* **451**, 363–373 (2006).
96. L. Pueyo et al., “Optimal dark hole generation via two deformable mirrors with stroke minimization,” *Appl. Opt.* **48**(32), 6296–6312 (2009).
97. N. T. Zimmerman et al., “Shaped pupil Lyot coronagraph designs for WFIRST/AFTA,” *Proc. SPIE* **9605**, 96050A (2015).
98. R. J. Vanderbei, “LOQO: an interior point code for quadratic programming,” *Optim. Methods Softw.* **11**(1–4), 451–484 (1999).
99. R. Soummer et al., “Fast computation of Lyot-style coronagraph propagation,” *Opt. Express* **15**, 15935 (2007).
100. Gurobi Optimization Inc., “Gurobi optimizer reference manual,” 2015, <http://www.gurobi.com/documentation/6.5/refman/index.html> (10 December 2015).

**Neil T. Zimmerman** recently joined the staff of the Space Telescope Science Institute after postdoctoral research appointments at the Max Planck Institute for Astronomy and at Princeton University. He earned his PhD in astronomy from Columbia University in 2011. He is a member of SPIE and the American Astronomical Society, and he specializes in exoplanet imaging instrumentation and techniques.

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

**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 a member of the American Astronomical Society and SPIE.

**Alexis Carlotti** is an assistant astronomer at the Institute of Planetology and Astrophysics in Grenoble, France. He specializes in coronagraph design and adaptive optics instrumentation.

**Robert J. Vanderbei** is a professor in the Department of Operations Research and Financial Engineering at Princeton University. He has degrees in chemistry (BS), operations research and statistics (MS), and applied mathematics (MS and PhD). He is a fellow of the American Mathematic Society, the Society for Applied and Industrial Mathematics, and the Institute for Operations Research and the Management Sciences.