SENSING PHASE ABERRATIONS BEHIND LYOT CORONAGRAPHS

ANAND SIVARAMAKRISHNAN[1,2] AND RÉMI SOUMMER[2,3,4]

Department of Astrophysics, American Museum of Natural History, 79th Street at Central Park West, New York, NY 10024

LAURENT PUEYO

Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544

AND

J. KENT WALLACE[2] AND MICHAEL SHAO

Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109

Received 2006 November 8; accepted 2008 July 16

ABSTRACT

Direct detection of young extrasolar planets orbiting nearby stars can be accomplished from the ground with extreme adaptive optics and coronagraphy in the near-infrared, as long as this combination can provide an image with a dynamic range of 10^7 after the data are processed. Slowly varying speckles due to residual phase aberrations that are not measured by the primary wave-front sensor are the primary obstacle to achieving such a dynamic range. In particular, non–common optical path aberrations occurring between the wave-front sensor and the coronagraphic occulting spot degrade performance the most. We analyze the passage of both low and high spatial frequency phase ripples, as well as low-order Zernike aberrations, through an apodized pupil Lyot coronagraph in order to demonstrate the way coronagraphic filtering affects various aberrations. We derive the coronagraphically induced cutoff frequency of the filtering and estimate coronagraphic contrast losses due to low-order Zernike aberrations: tilt, astigmatism, defocus, coma, and spherical aberration. Such slowly varying path errors can be measured behind a coronagraph and corrected by a slowly updated optical path delay precompensation or offset asserted on the wave front by the adaptive optics (AO) system. We suggest ways of measuring and correcting all but the lowest spatial frequency aberrations using Lyot plane wave-front data, in spite of the complex interaction between the coronagraph and those mid–spatial frequency aberrations that cause image plane speckles near the coronagraphic focal plane mask occulter's edge. This investigation provides guidance for next-generation coronagraphic instruments currently under construction.

Subject headings: instrumentation: adaptive optics — instrumentation: high angular resolution — planetary systems — space vehicles — techniques: high angular resolution

1. INTRODUCTION

Imaging and low-resolution spectroscopy with a coronagraph (Lyot 1939) on next-generation "extreme adaptive optics" (ExAO) systems on 8–10 m ground-based telescopes (Sivaramakrishnan et al. 2001) promises to open up a search space that is complementary to current radial velocity planet detection techniques (Oppenheimer et al. 2003; Macintosh et al. 2004; Mouillet et al. 2004; Beuzit et al. 2006; Graham et al. 2007; Hinkley et al. 2007; Macintosh et al. 2007), as long as the ExAO system, coronagraph, observing strategy, and postacquisition data analysis methods mesh well enough to deliver an *H*-band photometric dynamic range of $\sim 10^7$. With residual coronagraphic light leak suppressed to well below this required contrast, advanced speckle detection and removal techniques can be brought to bear on residual slowly varying speckles that the coronagraph cannot eliminate (Sivaramakrishnan et al. 2005). These static and quasi-static speckles set the real contrast floor of such instruments, as extensive simulations suggest (Macintosh et al. 2006). The cited work predicts that the required contrast is achievable after speckle subtraction in the data analysis as long as the coronagraph by itself can deliver a raw contrast of 6×10^{-8} in the absence of all aberrations and, furthermore, mid–spatial frequency aberrations

in the optics are kept below ~ 1–2 nm. An apodized pupil Lyot coronagraph can provide such suppression at a separation of $0.17''$ from an on-axis AO guide star. For scale, this angular distance is approximately five resolution elements at the 1.65 μm *H*-band central wavelength on an 8 m telescope.

Multiwavelength speckle suppression techniques used on low spectral resolution (~ 40) integral field unit spectrographic data will enable a factor of ~ 30 to be achieved if chromatic wave-front errors are kept below ~ 1 nm rms (Macintosh et al. 2006). This is a challenging goal, as phase errors introduced by optics that are not in, or close to, pupil planes can complicate the speckle reduction data processing (e.g., Marois et al. 2004; Marois 2004; Marois et al. 2005). Furthermore, aberrations prior to any image slicing or spectrographic focal plane must not differ by more than about 1 nm for straightforward speckle suppression (Marois et al. 2000; Sparks & Ford 2002) to be used.

The *H* band is a favored ground-based observing bandpass, since Jovian gas giant planets and brown dwarfs exhibit a methane absorption feature across half of the band, combined with the fact that ambient temperature optics do not contribute excessive thermal radiation to the beam at *H*. Thermal background at the longer wavelength *K* band starts to be problematic (e.g., Robberto et al. 2000; Sosey & Sivaramakrishnan 2004). Thus, a noncryogenic AO system can be used at *H* with little impact from the thermal background. Phase aberrations scale approximately inversely with wavelength, so the longer the wavelength, the better the image quality. Furthermore, contrast between AO target stars and their faint companions—young gas giant planets or brown

---

[1] Department of Physics and Astronomy, State University of New York at Stony Brook, Stony Brook, NY 11794.
[2] Center for Adaptive Optics, University of California, Santa Cruz, CA 95064.
[3] Michelson Postdoctoral Fellow.
[4] Kalbfleisch Postdoctoral Fellow.

dwarfs—is lower at H than in the optical bandpasses. With an integral field spectrograph behind the coronagraph, not only can sophisticated speckle subtraction be performed on the images to enhance the photometric dynamic range by factors of 10–30 (Racine et al. 1999; Marois et al. 2000, 2004, 2005, 2006; Sparks & Ford 2002; Lafrenière et al. 2007; Hinkley et al. 2007), but spectra of extrasolar planets can also be acquired. Such spectra will enable a study of planet formation and evolution to be made by providing information on the physical state of planetary atmospheres. In addition, low-resolution spectra in J and H can help spectral classification in the late M, L, and T dwarf spectral classes (Kirkpatrick 2005). ExAO coronagraphs can also be used to study protoplanetary disk structure, brown dwarfs, and stellar binarity and multiplicity (Metchev 2006; McElwain et al. 2007; Kalas et al. 2007; Oppenheimer et al. 2008).

The manner in which wave-front aberrations interact with coronagraphic optics can complicate postprocessing techniques (Lloyd & Sivaramakrishnan 2005; Sivaramakrishnan et al. 2005; Digby et al. 2006; Marois et al. 2006; Hinkley et al. 2007). Theoretical work on speckle morphology and statistics (Bloemhof et al. 2001; Sivaramakrishnan et al. 2002; Perrin et al. 2003; Marois 2004; Aime & Soummer 2004; Soummer & Ferrari 2007) suggests that the presence of static or quasi-static speckles in the final image can reduce the efficacy of speckle subtraction techniques that utilize simple predictions of the way speckles behave chromatically. This makes the calibration of the ExAO system's residual wave-front error against the wave-front error seen by the coronagraph critical to the scientific output of the entire instrument. ExAO coronagraphs currently under development plan to perform non–common-path aberration sensing and correction, either in a closed loop at a rate of a few Hz (Macintosh et al. 2006; Wallace et al. 2006) or, without the coronagraph, in between science exposures (Fusco et al. 2006; Sauvage et al. 2007), so as to remove as much of the residual static and quasi-static speckle pattern as possible. In space applications such as the *Terrestrial Planet Finder*, these aberrations can be controlled with image-based wave-front sensing using data from the coronagraphic image (Give'on et al. 2006; Bordé & Traub 2006; Trauger & Traub 2007; Pueyo & Kasdin 2007). It is therefore timely to estimate how wave-front errors are changed by passage through a coronagraph, since the coronagraph modifies the wave front in ways that are sensitive to the spatial frequency content of the wave front's aberrations.

Lloyd & Sivaramakrishnan (2005) and Sivaramakrishnan et al. (2005) studied the passage of selected individual low-order aberrations through a coronagraph. Aime & Soummer (2004) and Soummer et al. (2007a) analyzed the effects of the coronagraph on the statistics of the residual speckle noise. Numerical tolerancing studies (Shaklan & Green 2005) and laboratory experiments (Crepp et al. 2006) on the propagation of particular phase errors through coronagraphs have also been carried out. Here we present some guidelines for sensing phase aberrations behind a Lyot-style coronagraph. Any implementation of this idea will need to be refined with studies using the details of the instrument, the detectors used, control loop rates, and the character of noise in the system.

2. WAVE-FRONT SENSING IN THE CORONAGRAPH'S LYOT PUPIL

Wallace et al. (2004), Shao et al. (2004), Macintosh et al. (2006), Wallace et al. (2006), and Shao (2007) describe components of an ExAO coronagraph with a slow postcoronagraphic calibration system dedicated to reducing non–common-path phase errors between a kilohertz-rate AO wave-front sensor and the optical path difference (OPD) at the focal plane, where the coronagraph's focal

plane mask occults the core of the on-axis star. This is accomplished by feeding back offsets to the fast AO control loop, thereby driving the deformable mirror to deliver as good an image as possible to the coronagraphic occulting spot, rather than correcting the wave front as well as is possible at the fast wave-front sensor (Wallace et al. 2006). Here we study the interaction between the coronagraph and such a calibration system. We adopt a general approach to the discussion to this topic, since we utilize a straightforward Fourier-based description of aberrations and a well-understood control-theoretic construct, the Bode diagram. We use a first-order expansion to estimate the transfer function of the coronagraph, and we present ways of performing a full correction of the wave front at the science wavelength in spite of the action of the coronagraph.

Scintillation is the phenomenon of phase irregularities in a wave front inducing intensity variations after the wave travels some distance. The magnitude of atmospheric scintillation at H is smaller at optical wavelengths than in the IR (e.g., Angel 1994), and, from rough measurements at the Advanced Electro-Optical System (AEOS) telescope, appears to be tolerable, given the required contrast limits of the coronagraph (Sivaramakrishnan et al. 2006). We ignore the effects of both atmospheric and internal scintillation here.

The theory of the apodized pupil Lyot coronagraph (APLC) for various aperture geometries (including arbitrary aperture shapes and segmented extremely large telescope configurations) was developed by Aime et al. (2002), Soummer et al. (2003), Soummer (2005), and Soummer et al. (2007b). The representative APLC design we use in this work is taken from Soummer (2005). This specific design possesses a focal plane mask (FPM) occulting spot with a diameter of ∼5 resolution elements, matched to an entrance pupil geometry corresponding to the 8 m Gemini telescope. The Gemini secondary obstruction diameter is 14% of the primary mirror's diameter. The theoretical throughput of the coronagraph's apodizer is ∼50%, and the coronagraph design delivers contrast ratios of 6×10^{-8} and 2×10^{-8} at separations of $0.23''$ and $0.45''$, respectively, over a 10% bandpass centered at 1.65 μm. Our study provides a theoretical basis for sensing high- and mid-frequency non–common-path aberrations behind such coronagraphs, using information acquired in the coronagraph's Lyot pupil, immediately after the coronagraphic focal plane mask.

3. REVIEW OF APLC CORONAGRAPHIC THEORY

The basic ideas underlying Lyot coronagraphy behind adaptive optics systems are discussed by Malbet (1996) and Sivaramakrishnan et al. (2001). Diverging from the notation used in the original APLC design papers, we use the notation of Lloyd & Sivaramakrishnan (2005), Sivaramakrishnan & Lloyd (2005), and Sivaramakrishnan et al. (2005), where an approach to tolerancing Lyot coronagraphs based on an expansion of the point-spread function (PSF) in terms of the Fourier transform of the phase aberrations in the pupil (Bloemhof et al. 2001; Sivaramakrishnan et al. 2002; Perrin et al. 2003; Marois 2004) is developed.

To recapitulate this notation, we use a telescope aperture transmission function $A(\boldsymbol{x})$, where $\boldsymbol{x} = (x, y)$ is a location in the aperture plane in units of the wavelength of the light (see Fig. 1). The first image is attenuated by a focal plane mask (FPM) whose transmission function is $m(\boldsymbol{k})$ [where $\boldsymbol{k} = (k_x, k_y)$ is the field angle, in units of radians]. We assume that the Fraunhofer approximation of Fourier optics holds, so that the aperture plane and image plane fields are Fourier transform pairs (e.g., Born & Wolf 1993). We change the case of a function to denote its Fourier transform. The Lyot pupil field is $E_{\mathrm{Lyot}} = M(\boldsymbol{x}) * A(\boldsymbol{x})$,
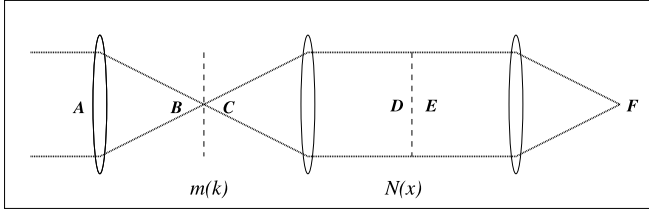
FIG. 1.—Essential planes and stops in a Lyot-style coronagraph. The entrance aperture is A; the direct image at B falls on a FPM whose transmission function is $m(k)$. The reimaged pupil plane D, after being modified by passage through a Lyot stop with a transmission function $N(x)$, is sent to the coronagraphic image at F. Here A, D, and E are pupil planes and B, C, and F are image planes. In the apodized-pupil Lyot coronagraph, the entrance pupil is apodized and the Lyot pupil stop has the same geometry as the entrance pupil (a classical Lyot coronagraph undersizes the entrance aperture at the Lyot plane).

FIG. 2.—Numerically optimized apodized Gemini pupil transmission function $T(r)$ for a $5\lambda/D$ FPM diameter occulter. Transmission at the edge of this apodizer is 16%, and its throughput is 49%.

where $*$ denotes the convolution operator, and $M(\boldsymbol{x})$ is the Fourier transform of the FPM occulter transmission profile $m(\boldsymbol{k})$.

The APLC design is characterized by a nonuniform transmission in the entrance aperture,

$$A(\boldsymbol{x}) = T(\boldsymbol{x})A_{\text{clear}}(\boldsymbol{x}), \tag{1}$$

combined with a hard-edged FPM profile

$$m(\boldsymbol{k}) = 1 - \Pi(k/s), \tag{2}$$

where $k = |\boldsymbol{k}|$. The function $\Pi(k)$ is 1 for $|k| < 1/2$ and 0 elsewhere, s is the diameter of the occulting spot in units of λ/D, and T is the apodizing function (see Fig. 2). Here D is the telescope diameter and λ is the observing wavelength. The quantity A_{clear} is 1 over the aperture and 0 elsewhere. The key feature of the APLC design is that the apodizing function T and the diameter of the FPM, s, are related to each other by an eigenvalue equation involving a functional of T, notably $E_{\text{Lyot}} = (1 - \Lambda)T(\boldsymbol{x})$. Here Λ is the eigenvalue that is close to but less than unity for FPM occulter diameters larger than a few resolution elements. After a stop in the Lyot pupil plane sized to admit just the light from within the footprint of the clear aperture, the fraction of on-axis energy passing through the coronagraph is $(1 - \Lambda)^2$. Soummer et al. (2003) describe the relevant mathematical properties of this eigenvalue problem.

4. PROPAGATION OF WAVE-FRONT ERRORS THROUGH AN APLC

4.1. *First-Order Phase Expansion*

The aperture illumination function with a phase aberration ϕ is $A_{\text{aber}} = A(\boldsymbol{x})e^{i\phi(\boldsymbol{x})}$. We expand the exponential in this equation in powers of ϕ and restrict the largest excursion of ϕ from its mean value over the clear aperture to be considerably less than 1 rad, thus staying within the linear range of the Taylor expansion $e^{i\phi(\boldsymbol{x})} = 1 + i\phi(\boldsymbol{x}) - i\phi(\boldsymbol{x})^2/2! + \dots$. For scale, in the H band on an 8 m telescope, a one resolution element tilt corresponds to a wavefront slope of $0.042''$. The tilted wave front has a peak-to-valley (P-V) difference of 2.0×10^{-7} rad across the telescope diameter. On a 1 cm diameter laboratory testbed aperture, a one resolution element tilt at the same wavelength yields a P-V difference of 1.6×10^{-4} rad across the pupil. Certainly ExAO systems can control tilt to far better than a resolution element. As long as all phase aberrations stay below ~0.3 rad, the first-order Taylor expansion will be a good guide to the behavior of the optical system (Sivaramakrishnan et al. 2002; Perrin et al. 2003). Barring phase errors with pathological functional forms, such as aberrations due to immobile actuators on deformable mirrors (Makidon
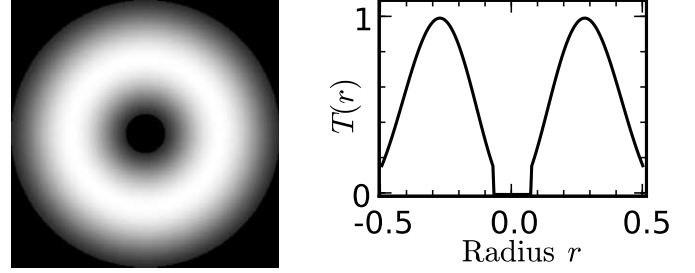
et al. 2005; Sivaramakrishnan et al. 2006), the first-order Taylor expansion will capture the behavior of the optical system quantitatively at Strehl ratios exceeding $\sim90\%$.

4.2. *The Fourier Approach*

We combine analytical methods (Sivaramakrishnan et al. 2002; Perrin et al. 2003; Lloyd & Sivaramakrishnan 2005; Sivaramakrishnan et al. 2005; Wallace et al. 2004, 2006) with a numerical study of the way Fourier modes of phase aberrations leak through a coronagraph. If we use the first-order expansion and treat the phase aberration ϕ by means of its Fourier transform Φ, the entrance pupil field is

$$e^{i\phi(\boldsymbol{x})} \simeq 1 + i\phi(\boldsymbol{x}) = 1 + i\int\int \Phi(\boldsymbol{k})e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\,d\boldsymbol{k}. \tag{3}$$

We find that a Fourier approach to the speckle calibration problem engenders insight and suggests practical ways of improving instrument performance. Such Fourier decomposition might seem artificial at first glance, since in reality wave-front aberrations are broadband (in the spatial frequency domain), but, as others have pointed out, some recondite problems in mathematical physics are very amenable to treatment by Fourier theory. Fourier-based thinking presented by Sivaramakrishnan et al. (2002) resulted in an analytical proof that the theory of speckle decorrelation (Angel 1994) is untenable. An extension of this Fourier approach (Perrin et al. 2003) provided a springboard for the development and analysis of spatially filtered wave-front sensing (Poyneer & Macintosh 2004). Poyneer & Véran (2005) and Poyneer et al. (2006) use the fact that different Fourier components of aberrations are uncorrelated in Kolmogorov-spectrum turbulence to show that a Fourier-based decomposition of wave-front aberrations enables one to make wave-front estimation and predictive control schemes that are computationally tractable.

4.3. *High and Break Spatial Frequency Aberrations*

Whereas Sivaramakrishnan et al. (2005) considered effects of higher order terms in the expansion of $e^{i\phi}$, we focus here on the first-order term in order to understand the propagation of small residual phase errors through the APLC coronagraph. Since $E_{\text{Lyot}} = M(\boldsymbol{x}) * A(\boldsymbol{x})$ and $M(\boldsymbol{x}) = \delta(\boldsymbol{x}) - s\,\text{jinc}(s\boldsymbol{x})$, at high Strehl ratios a small phase aberration $\phi(\boldsymbol{x})$ in the entrance pupil results in a Lyot pupil field of

$$E_{\text{Lyot,aber}} = E_{\text{Lyot}} + i\phi(\boldsymbol{x})A(\boldsymbol{x}) - is[\phi(\boldsymbol{x})A(\boldsymbol{x})] * \text{jinc}(s\boldsymbol{x}) + \dots \tag{4}$$

to first order in ϕ, where E_{Lyot} denotes the field in the Lyot plane in the absence of aberrations and $\delta(\boldsymbol{x})$ is the two-dimensional
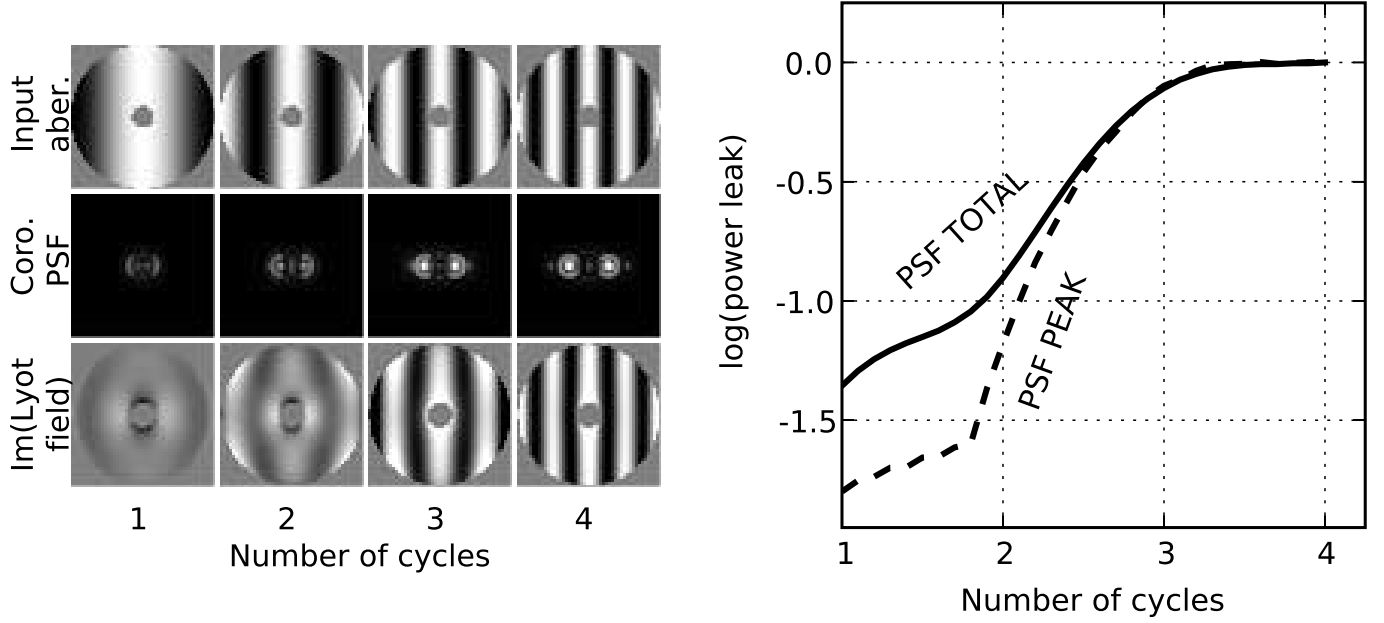
Fig. 3.—Coronagraphic filtering of input aberrations. *Left, top row*: Sinusoidal 0.316 rad peak-to-valley input phase aberrations with different spatial frequencies (i.e., cycles per aperture). *Left, middle row*: Coronagraphic PSFs with a $2.5\lambda/D$ radius occulting spot (*not shown*). As the spatial frequency of the phase ripple increases, first-order speckles approach the edge of the FPM occulter and their intensity increases. Their intensities become asymptotic when they move a distance of $\sim\lambda/D$ past the FPM's edge. PSFs are displayed on a logarithmic stretch between 1 and 10^{-3} of their asymptotic peak brightness. *Left, bottom row*: Imaginary component of the Lyot plane field (after dividing by the pupil apodization). As first-order speckles move further than λ/D past the FPM edge, the imaginary component of the Lyot field complex amplitude tracks the input phase aberration. Higher spatial frequency phase ripples pass through to the Lyot pupil with little attenuation or distortion. *Right*: Total coronagraphic PSF power and peak PSF intensity due to small sinusoidal phase ripple aberrations. Both curves are normalized to unity at a spatial frequency of four cycles per aperture.

Dirac δ-function. Since $\mathrm{jinc}(x)$ is the Fourier transform of the two-dimensional function $\Pi(k)$, this equation can be rewritten as

$$E_{\mathrm{Lyot,aber}} = E_{\mathrm{Lyot}} + i\phi(\boldsymbol{x})A(\boldsymbol{x}) - i\mathcal{F}([a(\boldsymbol{k}) * \Phi(\boldsymbol{k})]\Pi(k/s)) + \ldots,$$
(5)

where $\mathcal{F}(...)$ denotes a Fourier transform. In the case of a pure phase ripple, e.g., $\phi(x) = \cos 2\pi n_0 x$, where n_0 is the number of cycles across the aperture, the convolution in equation (5) simplifies to

$$a(\boldsymbol{k}) * \Phi(\boldsymbol{k}) = \frac{1}{2}[a(\boldsymbol{k} - \boldsymbol{k}_0) + a(\boldsymbol{k} + \boldsymbol{k}_0)],$$
(6)

where $|\boldsymbol{k}_0| = n_0\lambda/D$. The first-order aberration leak into the Lyot pupil is therefore

$$i\phi(\boldsymbol{x})A(\boldsymbol{x}) - \frac{i}{2}\mathcal{F}([a(\boldsymbol{k} - \boldsymbol{k}_0) + a(\boldsymbol{k} + \boldsymbol{k}_0)]\Pi(k/s)). \quad (7)$$

The behavior described here is generic in Lyot coronagraphs with hard-edged FPMs. With increasing spatial frequency of the pupil plane phase ripple, the resultant image plane speckles migrate out from behind the occulting FPM. In doing so, they mimic the way in which pure tilt causes complex PSF behavior in Lyot coronagraphs (Lloyd & Sivaramakrishnan 2005). The full expansion of the PSF in terms of the Fourier transform of the phase aberration described by Perrin et al. (2003) enables one to come to a deeper understanding of this: a phase ripple's Fourier transform is just a pair of Dirac δ-functions in the frequency domain. The PSF expansion describes how these δ-functions combine nonlinearly to create second- and higher order δ-function

components. Every one of these δ-functions are convolved with a, the transform of the unaberrated aperture function A. When a phase ripple places a δ-function at a particular location $\boldsymbol{k}_0$ in the image plane, this creates an offset copy, $a(\boldsymbol{k} - \boldsymbol{k}_0)$, of a (eq. [6]). This is mathematically indistinguishable from a tilted PSF located at $\boldsymbol{k}_0$. Thus, the way in which coronagraphs respond to tilt contains the kernel of the description of the way in which small phase ripples propagate through the coronagraph.

A hard-edged focal plane occulter $s\lambda/D$ in diameter introduces a frequency scale at $s/2$ cycles across the aperture. This characteristic spatial frequency provides a separator or "break frequency" between high and low spatial frequency behavior, since the mask function $m(\boldsymbol{k})$ filters the two offset copies of a (see eq. [7]). The propagation of phase ripples through the coronagraph is illustrated in the left panel of Figure 3, where we show the input aberration, the coronagraphic image plane intensity (to illustrate the evolution of two [first order] offset copies of a with various spatial frequencies), and the imaginary part of the field in the Lyot plane. The right panel of Figure 3 shows the way power from input sinusoidal phase aberrations leaks through the coronagraph and is essentially a transfer function for aberrations passing through the coronagraph.

The top row of the left panel of Figure 3 shows harmonic input aberrations with the same peak-to-valley amplitude (0.316 rad of phase). In the absence of a coronagraphic occulting spot, a small phase ripple with n cycles across the aperture produces two first-order speckles located at $\pm n\lambda/D$ in the image plane. The magnitude of the phase ripple we chose shows first- and second-order speckles, with first-order speckles being the dominant ones. Four spatial frequencies—the number of cycles across the aperture— are shown, in order to elucidate coronagraphic behavior as first-order speckles caused by the aberration migrate past the edge of the coronagraphic occulting spot. The occulter diameter is $5\lambda/D$,

so its radius, $2.5\lambda/D$, corresponds to a spatial frequency of 2.5 cycles across the aperture, a number that lies within the range of the spatial frequencies shown in the figure. The middle row of this panel shows the final coronagraphic image plane intensities when these aberrations are present. As is evident from this figure, speckles that fall behind the occulter are reduced in intensity, but the power that does leak through them is modified in form by passage through the coronagraph. When we compare the morphology of the speckles visible in the middle row, it is apparent that it is only when the speckles get more than a resolution element further away from the occulter's edge that their form is simple and undistorted. At spatial frequencies of two and three cycles across the aperture the speckles are distorted, whereas at four cycles across the aperture these speckles start to resemble the PSF of the unocculted system. This behavior is reflected in the complexity of the coronagraph's Lyot pupil field, especially in its imaginary component. As described above, the Lyot pupil field's imaginary component embodies the first-order characteristics of the coronagraphic response to aberrations (Sivaramakrishnan et al. 2005). As the bottom row of this panel shows, the Lyot field imaginary component morphology is distorted (relative to the input aberration morphology) until the first-order speckles move sufficiently far outside the occulter's edge. At one, two, and three cycles across the aperture, the Lyot field structure shows both attenuation and warping relative to the input aberration structure seen in the top row.

The right panel of Figure 3 shows the magnitude of the coronagraphic effect on the input aberration, measured in two ways. The total energy in the coronagraphic PSF is one measure of light leak. Another measure is the peak intensity of the brightest speckle in the image. Normalization is performed by dividing the appropriate quantity by that of the highest spatial frequency ripple's coronagraphic PSF. At this spatial frequency ($n = 4$), the effect of the coronagraph is negligible. As the figure shows, the light leak of the coronagraph becomes asymptotic once the first-order speckles move a resolution element or so outside the occulter's edge. Second-order speckles—faint spots at $2n$ resolution elements away from the center of the PSFs—are just visible in some of these images. A fourth-order effect, on-axis faint spots, are also sometimes apparent in the PSFs. In this figure, we identify different regimes of the propagation of harmonic ripples through a coronagraph.

1. When $n_0 > s/2 + 1$, the phase ripple propagates through the coronagraph with barely any attenuation. The two offset copies of a are virtually untouched by the mask function $m(\boldsymbol{k})$, so the aberration leak is simply the input aberration itself, attenuated in magnitude by the pupil apodization T. The phase aberration manifests itself as a component $i\phi A(\boldsymbol{x})$ of the Lyot field, in phase quadrature to the unperturbed Lyot pupil interior field E_{Lyot}. This is a mathematical representation of the fact that the first-order image plane speckles caused by a high spatial frequency phase ripple miss the FPM occulting spot by at least one resolution element. The left panel of Figure 3 shows this behavior at spatial frequencies of $n_0 > s/2 + 1 = 3.5$ cycles across the aperture. These high spatial frequency aberrations can be retrieved from a measurement of the imaginary part of the Lyot stop complex amplitude and can be corrected by applying a wave-front phase correction of $-\phi(\boldsymbol{x})$ with a deformable mirror at a pupil in front of the coronagraph.

2. When $s/2 + 1 > n_0 > s/2 - 1$, the phase ripple is altered by the image plane occulting stop. This "break" spatial frequency regime separates the high and low spatial frequency regimes. It is not easy to deduce the coronagraphic effect on these spatial frequencies from equation (4), where the convolution term $is[\phi(\boldsymbol{x})A(\boldsymbol{x})] * \mathrm{jinc}(sx)$ is nonnegligible and affects the aberration leak significantly. The left panel of Figure 3 shows this for spatial frequencies of $1.5 < n_0 < 3.5$.

4.4. Low Spatial Frequency Aberrations

At low spatial frequencies, when $n_0 < s/2 - 1$, the ripple is essentially damped out by the image plane occulting stop. Little information about low spatial frequency input aberrations reaches the Lyot plane, so Lyot pupil data cannot be used directly for low spatial frequency wave-front sensing.

On the other hand, if $n_0 \ll s/2$, convolution with the jinc function resembles convolution with a δ-function [since the FWHM of $\mathrm{jinc}(sx)$ is far smaller than the spatial period of an oscillation of the ripple]. In this case, the contribution of equation (7) to the Lyot field is much reduced because of the almost complete cancellation of the two terms. The left panel of Figure 3 shows this at spatial frequencies of $n_0 < s/2 - 1 = 1.5$ cycles across the aperture. The signal from a low spatial frequency phase ripple is highly attenuated by passage through the coronagraph.

In order to provide an alternative understanding of the propagation of aberrations in this low spatial frequency regime, we modeled coronagraphic sensitivity to low-order aberrations, expressed using Zernike polynomials. Our pupil is represented with 352 samples across its diameter, and the first image plane is sampled at $\lambda/6D$. The pupil and the focal plane occulter are grayscaled at their edges to mitigate pupil boundary pixelation effects. Final coronagraphic images are computed on a $\lambda/3D$ pitch. To describe the tilt, as well as the other low-order aberrations, in a uniform manner, we describe aberration strengths by their rms phase error (OPD). This quantity is a measure of the Strehl ratio decrease of a direct (i.e., noncoronagraphic) PSF by the Maréchal approximation at Strehl ratios above 0.3 (Maréchal 1947). Since the rms phase error is not a natural measure of the tilt, we note that a wave-front slope of 1 mas on an 8 m circular aperture results in a 10 nm rms tilt of the wave front over the aperture. Contrast at a point in the coronagraphic image plane is defined as the ratio of the intensity at that point when compared to the peak intensity of a PSF with the same pupil apodization and Lyot stop, but without the FPM present.

Figure 4 shows coronagraphic contrast at the design wavelength of the coronagraph with low-order Zernike phase aberrations applied at its entrance pupil. Since the aberration size is well within the linear theory of PSF formation, cross talk between different aberrations is negligible. We note that the way astigmatism behaves is in good agreement with the first-order coronagraphic aberration theory described in Sivaramakrishnan et al. (2005), which also predicts the kind of sensitivity to the other aberrations shown in this figure. Lloyd & Sivaramakrishnan (2005) discuss the curious increase in contrast with increasing tilt that is evident in this figure.

5. IMPLICATIONS FOR WAVE-FRONT SENSING POSTCORONAGRAPH AND SPECKLE CONTROL

5.1. Wave-Front Sensing in the Lyot Plane

Aberrations with spatial frequencies that are not filtered out by the FPM can be sensed directly in the Lyot plane, as illustrated in equation (7), if a measurement of the field strength (i.e., the wave's complex amplitude) in this plane is obtained. Wallace et al. (2004, 2006) proposed using an interferometer to obtain such data. As a result of a coronagraphically imposed cutoff
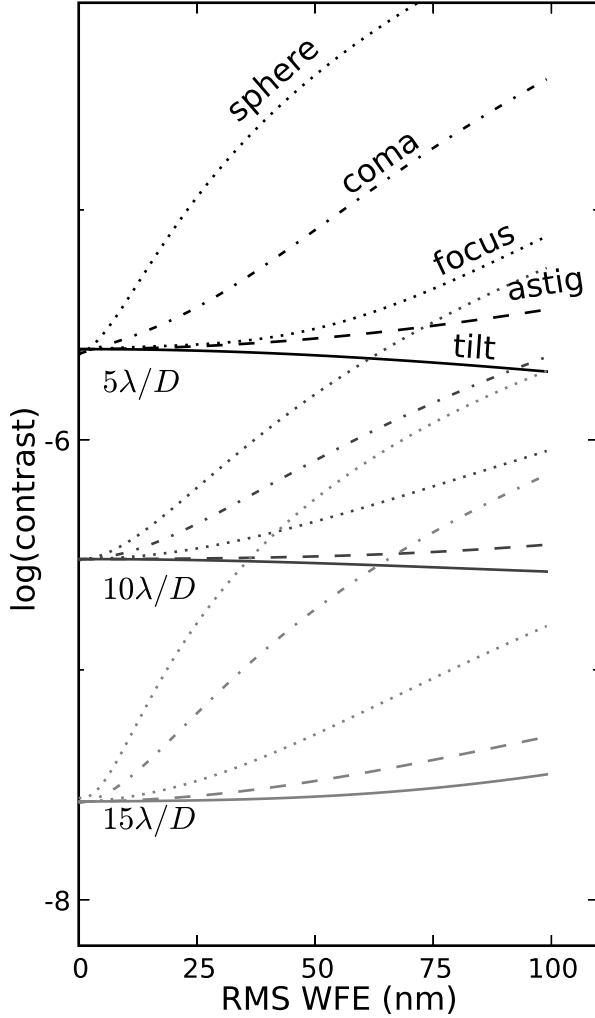
Fig. 4.—Azimuthally averaged contrast at separations of $5\lambda/D$, $10\lambda/D$, and $15\lambda/D$ from a star behind the APLC design optimized for the Gemini pupil geometry. The coordinate system is centered at the middle of the $5\lambda/D$ diameter focal plane mask, which is equivalent to a 100 mas radius occulter at the H band on an 8 m telescope. A tilt of 1 mas corresponds to an rms phase error of 10 nm on an 8 m aperture.

frequency, non–common-path errors above $s/2 + 1$ cycles across the aperture are measured by the imaginary component of the field in the Lyot plane and could be corrected with a deformable mirror. While low-frequency aberrations cannot be retrieved from the imaginary part of the Lyot field, they can be measured in the light occulted by the FPM. In particular, if the FPM is implemented as a reflective mirror with a hole (as is done on the Lyot Project coronagraph; Oppenheimer et al. 2004; Sivaramakrishnan et al. 2006), light passing through the hole can be captured and analyzed for low-frequency aberrations (Wallace et al. 2006). This is the approach that has been selected for the Gemini Planet Imager ExAO coronagraph (Macintosh et al. 2007).

Although this wave-front sensing approach is based on a first-order expansion, second-order speckles are clearly visible in the coronagraphic PSF formed from a phase ripple with a single spatial frequency (Fig. 3). As Sivaramakrishnan et al. (2002) and Perrin et al. (2003) discuss, these second- (and higher) order speckles are typically small at Strehl ratios above $\sim 80\%$–90%, so the sensing described in this paper is relevant when image quality is more or less diffraction-limited. This image quality can be obtained on large ground-based telescopes possessing ExAO systems operating in the near-IR.

Wave-front sensing often uses a shorter wavelength, λ_S, than the science wavelength λ_L. A physical occulter of $s\lambda_L/D$ in angular diameter is $s' = s\lambda_L/\lambda_S$ resolution elements across at λ_S. If postcoronagraphic sensing is done at a sufficiently short wavelength

$$\lambda_S = s/(s + 4)\lambda_L, \tag{8}$$

the break spatial frequency phase errors at the science wavelength can be "shifted down" to be sensed in light that passes through the occulting hole in the FPM. This can be accomplished using a hole in a flat mirror for an occulter, or an optical quality reflective spot. Sensing this occulted light's wave front (with a low-order Shack-Hartmann sensor, for example) will measure low-order aberrations effectively. At this sensing wavelength λ_S, a speckle that falls one resolution element outside the occulter's edge at the longer wavelength λ_L will lie one resolution element inside the occulter edge, thus passing through the hole without being affected much by the presence of the occulter's edge.

As long as the chromaticity of the optics and the atmospheric chromaticity do not change across the wavelength range, exploiting this wavelength-dependent behavior of break frequency aberrations could enhance post-coronagraphic wave-front sensing if the sensing is performed at a sufficiently short wavelength.

Similar arguments are made by Poyneer & Macintosh (2004) regarding their innovative wave-front sensing arrangement that utilizes a spatial filter implemented as a hole in the image plane. In this case the hole size is tuned to the highest spatial frequency that the wave-front sensor is capable of measuring.

In a practical application of this calibration system, the imaginary component of the Lyot plane complex amplitude could be Fourier-transformed, sent through a high-pass numerical filter, and then retransformed to pupil plane space. The high spatial frequency components of this signal can be applied by the adaptive optics system as Shack-Hartmann wave-front sensor centroid offsets by a deformable mirror. Sensing low-frequency aberrations using light that is occulted requires a separate low spatial frequency sensing system. The details of matching the signals from the two sensors are beyond the scope of this work, but certainly a graded transition between the two is an obvious strategy. The form of this transition will depend on the relative wavelengths of the two sensors and the size of the coronagraphic occulter. Further analysis and simulation is required to address the issue of matching the two spatial frequency regimes, and the entire approach will need to be verified with experiments.

5.2. Controlling Non–Common-Path Aberration

Details of a system that corrects aberrations on the basis of Lyot plane measurements are beyond the scope of this paper. We simply illustrate the possibility of such a correction scheme based on the first-order expansion model. We discuss the validity of a wave-front sensing and correction scheme based on a first-order approximation, together with perfect sensing of the wave front in the Lyot plane interior. This can be accomplished using a phase-shifting interferometric arrangement in a Mach-Zehnder arrangement (Wallace et al. 2006). This implementation relies on using a single high-quality beam splitter to extract a fraction of the science light that misses the coronagraphic occulting spot, in order to create an image of the Lyot pupil plane on a detector. Light passing through the coronagraphic occulting spot (implemented as a hole in a reflective mirror) is spatially filtered using a sub–resolution-element pinhole to provide an unaberrated reference beam. This reference beam is made to interfere with the Lyot pupil complex amplitude on the detector, enabling
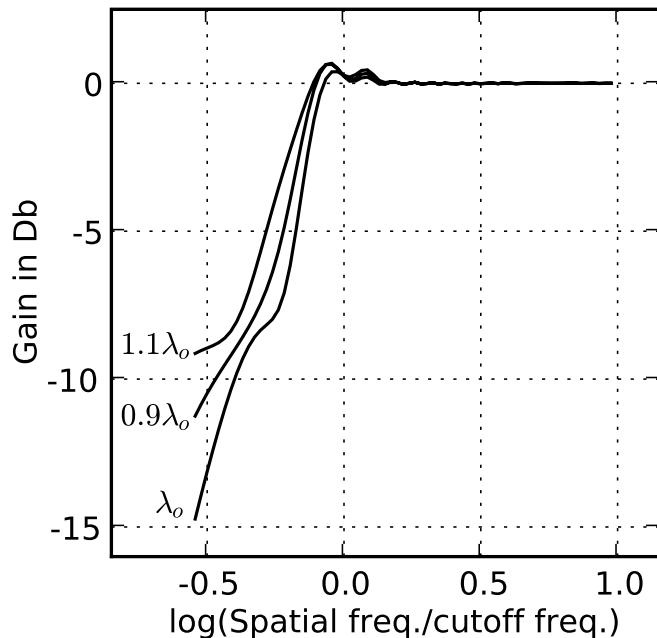
FIG. 5.—Coronagraph as a high-pass filter: the Bode diagram of a gray apodizer APLC at the design wavelength λ_0 and at two nearby wavelengths. The amplitude of a phase ripple aberration in the Lyot plane relative to the amplitude of the same ripple in the entrance pupil is plotted (a ratio r is 20 log r dB). For spatial frequencies above the cutoff frequency (which is $s/2 + 1$ cycles across the aperture), the response is unity.
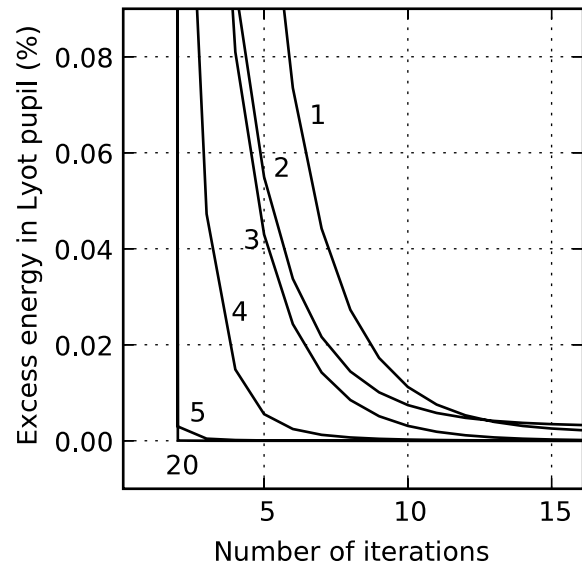

FIG. 6.—Coronagraphic energy leak vs. number of iterations of the calibration loop. Phase ripples of 1, 2, 3, 4, 5, and 20 cycles per aperture in the entrance pupil are measured by an ideal interferometric measurement of the Lyot plane complex amplitude, and removed in an iterative loop in order to correct the entrance pupil aberration. At spatial frequencies above the coronagraphic cutoff frequency (3.5 cycles per aperture), rapid convergence to a flat wave front is seen. At lower spatial frequency aberrations, convergence is slower or fails entirely. Complex behavior of low-frequency aberrations can degrade the overall coronagraphic performance. The leak is expressed as a percentage of the energy within the entrance pupil's projection to the Lyot pupil plane in the absence of all aberrations.

phase-shifting interferometric measurements of complex amplitude in the Lyot plane.

Figure 5 shows the ratio of the amplitude of an input phase ripple to that of the same ripple measured in the Lyot plane after its passage through the coronagraph. This ratio is plotted as a Bode diagram, a standard control-theoretic construct. Although this construct here is not a rigorous Bode diagram (which assumes a linear process where the output signal is modified in phase and amplitude, without any cross talk between frequencies), it captures the essence of the effect of the coronagraph on aberrated wave fronts. At high frequencies, the ratio of the aberrations is unity and the phase can be retrieved easily from the Lyot plane measurement. The coronagraph behaves as a wavelength-dependent high-pass filter in the spatial frequency domain. In the transition regime this loop construct is not rigorously valid, but some signal leaks through the Lyot plane. In this regime we note the presence of an overshoot near the break frequency. In the low-frequency region, the phase is not only attenuated, but also distorted (as illustrated in Fig. 3).

The effect of a control loop based on first-order theory on the wave front in the low spatial frequency region is demonstrated here. We assume perfect wave-front sensing and perfect application of wave-front corrections, together with small wave-front aberrations, in order to test the applicability of our approach. We do not consider additional noise sources such as photon noise or detector noise. Figure 6 shows the behavior of a slow calibration loop that utilizes the imaginary part of the Lyot plane interior complex amplitude to correct for wave-front aberrations in the entrance pupil. As expected, the loop performs well for spatial frequencies higher than those corresponding to the FPM diameter. Good correction is obtained after a single pass through the control loop, since high spatial frequency aberrations are measured perfectly in the Lyot plane. At lower spatial frequencies, the loop behavior is complicated, but nevertheless convergent (although at a slower rate). Figure 3 shows that the diffraction of low spatial

frequency ripples by the FPM yields higher spatial frequency content of the propagated aberration in the Lyot plane. In consequence, in addition to correcting high spatial frequencies, each iteration of this speckle calibration loop transfers energy from lower spatial frequencies to higher ones, making them correctable in the next iteration. This ensures the convergence of the wavefront calibration algorithm as long as second- and higher order terms in the PSF expansion are small compared to the first-order term.

The Bode diagram presented here can be used to estimate a gain as a function of spatial frequency, in order to accelerate the removal of low spatial frequency aberrations from a wave front. However, such amplification will increase sensitivity to measurement noise and nonlinearities. As shown in Figure 5, optimal rejection of low spatial frequency aberrations is only obtained for the central wavelength; viz., the wavelength at which the apodization is optimized. Indeed, chromatic effects reduce the efficiency of the coronagraph, and more aberrations leak through the coronagraph. Any increased loop gain used to accelerate loop convergence for low spatial frequency aberrations is limited by coronagraphic behavior at the wavelength that has the worst coronagraphic response (i.e., the highest throughput of on-axis incident light).

6. CONCLUSION

In this paper we study the propagation of aberrations through an apodized pupil Lyot coronagraph (Soummer 2005) by calculating a first-order approximation to the coronagraph's response to an aberrated wave front. Our treatment is valid in the regime of diffraction-limited imaging, such as might be expected from the next generation of extreme adaptive optics systems currently under construction (Sivaramakrishnan et al. 2001). We analyze a wave-front sensing approach described by Wallace et al. (2004, 2006; notably, that phase aberrations at the input pupil of the coronagraph can be measured in the Lyot pupil plane, in phase

quadrature to the incoming light. We study the effect of an APLC on an aberrated wave front using a Fourier decomposition of the wave front's phase errors, using the approach of Sivaramakrishnan et al. (2002). We summarize the throughput of aberrations of various spatial frequencies using a Bode diagram, thereby refining our understanding of control loop behavior. We find that high spatial frequency aberrations that correspond to a scale larger than the focal plane occulting mask size (in the pupil's Fourier transform domain) can be measured and corrected by our approach. These aberrations are responsible for residual speckles in the science field of view after the coronagraph. Break spatial frequencies that cause speckles near the edge of the occulting mask are filtered by the occulting spot. They are not measured perfectly in the Lyot plane, but seem to be correctable with some effort. If the control loop gain for these frequencies is increased, they are amenable to correction using postcoronagraphic Lyot pupil plane interferometric data. Alternatively, if a short enough sensing wavelength is used, information on these frequencies can be downshifted to below the break frequency and then sensed in the light that passes through a hole that serves as the coronagraphic occulter. The lowest spatial frequency aberrations must also be sensed in such occulted light, since it appears to be impractical to measure these aberrations after the starlight has been modified by a focal plane occulting mask. Any implementation of the ideas described in this paper will require detailed modeling using the actual instrument parameters, following the precepts we outline here.

REFERENCES

Aime, C., & Soummer, R. 2004, ApJ, 612, L85
Aime, C., Soummer, R., & Ferrari, A. 2002, A&A, 389, 334
Angel, J. R. P. 1994, Nature, 368, 203
Beuzit, J.-L., et al. 2006, in IAU Colloq. 200, Direct Imaging of Exoplanets: Science and Techniques, ed. C. Aime & F. Vakili (Cambridge: Cambridge Univ. Press), 317
Bloemhof, E. E., Dekany, R. G., Troy, M., & Oppenheimer, B. R. 2001, ApJ, 558, L71
Bordé, P. J., & Traub, W. A. 2006, ApJ, 638, 488
Born, M., & Wolf, E. 1993, Principles of Optics (6th ed.; Oxford: Pergamon)
Crepp, J. R., Ge, J., Vanden Heuvel, A. D., Miller, S. P., & Kuchner, M. J. 2006, ApJ, 646, 1252
Digby, A. P., et al. 2006, ApJ, 650, 484
Fusco, T., et al. 2006, Proc. SPIE, 6272, 62720K
Give'on, A., Kasdin, N. J., Vanderbei, R. J., & Avitzour, Y. 2006, J. Opt. Soc. Am. A, 23, 1063
Graham, J. R., et al. 2007, preprint (arXiv:0704.1454)
Hinkley, S., et al. 2007, ApJ, 654, 633
Kalas, P., Duchene, G., Fitzgerald, M. P., & Graham, J. R. 2007, ApJ, 671, L161
Kirkpatrick, J. D. 2005, ARA&A, 43, 195
Lafrenière, D., Doyon, R., Nadeau, D., Artigau, É., Marois, C., & Beaulieu, M. 2007, ApJ, 661, 1208
Lloyd, J. P., & Sivaramakrishnan, A. 2005, ApJ, 621, 1153
Lyot, B. 1939, MNRAS, 99, 580
Macintosh, B., et al. 2006, Proc. SPIE, 6272, 62720L
———. 2007, Comptes Rendus Phys., 8, 365
Macintosh, B. A., et al. 2004, Proc. SPIE, 5490, 359
Makidon, R. B., Sivaramakrishnan, A., Perrin, M. D., Roberts, L. C., Jr., Oppenheimer, B. R., Soummer, R., & Graham, J. R. 2005, PASP, 117, 831
Malbet, F. 1996, A&AS, 115, 161
Maréchal, A. 1947, Rev. d'Opt., 26, 257
Marois, C. 2004, Ph.D. thesis, Univ. Montréal
Marois, C., Doyon, R., Nadeau, D., Racine, R., Riopel, M., Vallée, P., & Lafrenière, D. 2005, PASP, 117, 745
Marois, C., Doyon, R., Racine, R., & Nadeau, D. 2000, PASP, 112, 91
Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., & Nadeau, D. 2006, ApJ, 641, 556
Marois, C., Racine, R., Doyon, R., Lafrenière, D., & Nadeau, D. 2004, ApJ, 615, L61
McElwain, M. W., et al. 2007, ApJ, 656, 505
Metchev, S. A. 2006, Ph.D. thesis, Caltech
Mouillet, D., Lagrange, A. M., Beuzit, J.-L., Moutou, C., Saisse, M., Ferrari, M., Fusco, T., & Boccaletti, A. 2004, in ASP Conf. Ser. 321, Extrasolar Planets: Today and Tomorrow, ed. J.-P. Beaulieu, A. Lecavelier des Etangs, & C. Terquem (San Francisco: ASP), 39

Oppenheimer, B. R., Sivaramakrishnan, A., & Makidon, R. B. 2003, in The Future of Small Telescopes in the New Millennium, Vol. 3, ed. T. D. Oswalt (Dordrecht: Kluwer), 155
Oppenheimer, B. R., et al. 2004, Proc. SPIE, 5490, 433
———.. 2008, ApJ, 679, 1574
Perrin, M. D., Sivaramakrishnan, A., Makidon, R. B., Oppenheimer, B. R., & Graham, J. R. 2003, ApJ, 596, 702
Poyneer, L. A., & Macintosh, B. 2004, J. Opt. Soc. Am. A, 21, 810
Poyneer, L. A., & Véran, J.-P. 2005, J. Opt. Soc. Am. A, 22, 1515
Poyneer, L. A., Véran, J.-P., Dillon, D., Severson, S., & Macintosh, B. A. 2006, Proc. SPIE, 6272, 62721N
Pueyo, L., & Kasdin, N. J. 2007, ApJ, 666, 609
Racine, R., Walker, G. A. H., Nadeau, D., Doyon, R., & Marois, C. 1999, PASP, 111, 587
Robberto, M., Sivaramakrishnan, A., Bacinski, J. J., Calzetti, D., Krist, J. E., MacKenty, J. W., Piquero, J., & Stiavelli, M. 2000, Proc. SPIE, 4013, 386
Sauvage, J.-F., Fusco, T., Rousset, G., & Petit, C. 2007, J. Opt. Soc. Am. A, 24, 2334
Shaklan, S. B., & Green, J. J. 2005, ApJ, 628, 474
Shao, M. 2007, Comptes Rendus Phys., 8, 340
Shao, M., Wallace, J. K., Levine, B. M., & Liu, D. T. 2004, Proc. SPIE, 5487, 1296
Sivaramakrishnan, A., Koresko, C. D., Makidon, R. B., Berkefeld, T., & Kuchner, M. J. 2001, ApJ, 552, 397
Sivaramakrishnan, A., & Lloyd, J. P. 2005, ApJ, 633, 528
Sivaramakrishnan, A., Lloyd, J. P., Hodge, P. E., & Macintosh, B. A. 2002, ApJ, 581, L59
Sivaramakrishnan, A., Soummer, R., Sivaramakrishnan, A. V., Lloyd, J. P., Oppenheimer, B. R., & Makidon, R. B. 2005, ApJ, 634, 1416
Sivaramakrishnan, A., et al. 2006, in IAU Colloq. 200, Direct Imaging of Exoplanets: Science and Techniques, ed. C. Aime & F. Vakili (Cambridge: Cambridge Univ. Press), 613
Sosey, M. L., & Sivaramakrishnan, A. 2004, Proc. SPIE, 5487, 299
Soummer, R. 2005, ApJ, 618, L161
Soummer, R., Aime, C., & Falloon, P. E. 2003, A&A, 397, 1161
Soummer, R., & Ferrari, A. 2007, ApJ, 663, L49
Soummer, R., Ferrari, A., Aime, C., & Jolissaint, L. 2007a, ApJ, 669, 642
Soummer, R., Pueyo, L., Sivaramakrishnan, A., & Vanderbei, R. J. 2007b, Opt. Express, 15, 15935
Sparks, W. B., & Ford, H. C. 2002, ApJ, 578, 543
Trauger, J. T., & Traub, W. A. 2007, Nature, 446, 771
Wallace, J. K., Bartos, R., Rao, S., Samuele, R., & Schmidtlin, E. 2006, Proc. SPIE, 6272, 62722L
Wallace, J. K., Green, J. J., Shao, M., Troy, M., Lloyd, J. P., & Macintosh, B. 2004, Proc. SPIE, 5490, 370