

# Optimizing coronagraph designs to minimize their contrast sensitivity to low-order optical aberrations

Joseph J. Green<sup>†</sup> and Stuart B. Shaklan  
Jet Propulsion Laboratory, California Institute of Technology

## ABSTRACT

The presence of optical aberrations in the entrance pupil of a coronagraph causes the stellar light to scatter about the occulting spot, reducing the effective contrast achievable. Even if these aberrations are sufficiently corrected with a deformable mirror to enable planet detection, small drifts in the optical alignment of the telescope introduce additional low-order aberrations. The design parameters of the coronagraph itself (e.g. occulting spot size, Lyot stop diameter, etc.) affect how these aberrations impact the contrast in the focal plane. In this study, we examine the sensitivity of contrast to low-order optical errors for several coronagraph concepts over their respective design parameters. By combining these sensitivities with the telescope throughput, we show that for each coronagraph concept there is an optimum selection of the design parameters that provides efficient, high-contrast imaging at the inner working distance in the presence of alignment errors.

Keywords: coronagraphic telescope, error modeling, optical aberrations, extrasolar planets

## 1. INTRODUCTION

The Terrestrial Planet Finder (TPF) mission seeks to detect and characterize extra-solar terrestrial planets around 150 nearby stars. The TPF project is currently in pre-phase A and is exploring two architectures: a structurally connected or free-flying infrared interferometer, and a visible coronagraph ([http://planetquest.jpl.nasa.gov/TPF/tpf\\_index.html](http://planetquest.jpl.nasa.gov/TPF/tpf_index.html)). The coronagraph detects starlight reflected from the planet; for an earth-like planet at 1 AU from its parent star, the planet brightness is  $\sim 1e-10$  of its parent star. For many targets the planet will appear within 100 milli-arcseconds of the star, just a few resolution elements away from the image core for a 5-10 m class telescope. Thus the visible coronagraph must achieve extraordinary dynamic range very close to the diffraction-limited core of the stellar image.

Achieving and maintaining the required sub-Angstrom level of wave front control is challenging to say the least. In particular, small drifts in the position and shape of the optics, especially the primary and secondary mirrors, will lead to low-order aberrations that scatter light in the neighborhood of the exo-planet image. In this paper, we address the sensitivity of several coronagraph designs to changes in the low-order aberrations. We show that there is trade between their efficiency (the fraction of exo-planet light that they transmit) and their aberration sensitivity. We then calculate optimal shapes that maximize the signal-to-noise ratio in the presence of time-variable aberrations.

Our study includes several image-plane masks, described below. To date we have not studied the efficiency and aberration sensitivity of pupil plane masks (e.g. shaped pupil<sup>1</sup> and apodized pupil<sup>2,3</sup>), nor have we analyzed the four-quadrant phase mask<sup>4</sup> or the newly-proposed pupil-reshaping apodization<sup>5</sup>. We plan to address these in future work.

## 2. CORONAGRAPH DESIGNS UNDER STUDY

In this paper we consider four Lyot-Coronagraph concepts: *radial-Gaussian (RG)*, *radial-cosine (RC)*, *linear-cosine (LC)*, and a mask with  $\sin(x)\sin(y)$  transmission. The RG represents the case of *band-unlimited* function, the RC, LC represents radial, and linear forms of *band-limited* occulting functions, which have been shown be preferable for improved diffraction rejection at high throughput<sup>6</sup>. The  $\sin(x)\sin(y)$  occulting spot is really a coronagraph representation

---

<sup>†</sup>Contact information: 4800 Oak Grove Dr., M/S 306-336, Pasadena CA 91109, E-mail: Joseph.J.Green@jpl.nasa.gov

for the visible-nuller (VN) concept<sup>7</sup>. Assuming a general functional form for occulting spot field transmission of  $o(x, y) = 1 - s(x, y)$ , the four occulting spot field absorption functions can be written as

$$s_{RG}(x, y) = \exp\left(-\pi \frac{(x^2 + y^2)}{\sigma^2}\right) \quad (1)$$

$$s_{RC}(x, y) = \frac{1}{2} + \frac{1}{2} \cos\left(\pi \frac{(x^2 + y^2)}{\sigma^2}\right) \quad (2)$$

$$s_{LC}(x, y) = \frac{1}{2} + \frac{1}{2} \cos\left(\pi \frac{x}{\sigma}\right) \quad (3)$$

$$s_{VN}(x, y) = 1 - \sin\left(\pi \frac{\sqrt{2}x}{\sigma}\right) \sin\left(\pi \frac{\sqrt{2}y}{\sigma}\right). \quad (4)$$

Kuchner pointed out that using sinusoidal occulting function of the form shown in equation (3) in a coronagraph is the analog to a nulling interferometer<sup>6</sup>. In this paper we explore the three possible ways of applying this analogy through the RC, LC and VN masks. These functions, of course, have the virtue of being band-limited. In a sense, these masks represent the limit to how band-limited a function can be made. The LC, RC and VN functions all contain energy along a fixed set of discrete points in their respective Fourier transforms (FT).

In consideration of the methods that may employed in fabricating the occulting masks<sup>8</sup> we wish to limit the spatial extent that these functions span. To limit the spatial extent of the periodic RC, LC and VN masks while maintaining that these functions remain band-limited, we multiply their respective spot functions with a positive band-limited function  $s_{apod}(x, y)$ . In these cases, the occulter field transmission becomes

$$o(x, y) = 1 - s(x, y) \cdot s_{apod}(x, y). \quad (5)$$

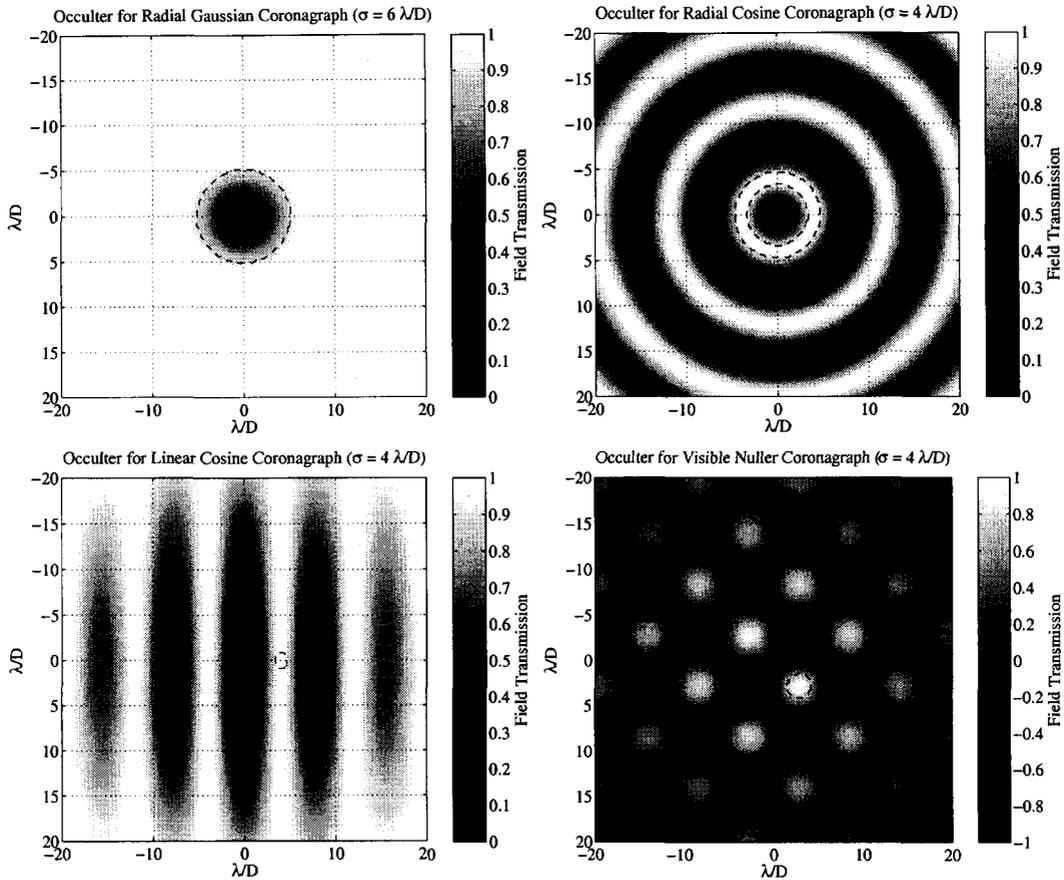
For our studies we employ a *radial-sinc*<sup>2</sup> tapering to the periodic occulters of the functional form

$$s_{apod}(x, y) = 1 - \left( \frac{\sin\left(\pi \sqrt{x^2 + y^2} / 4\sigma\right)}{\sqrt{x^2 + y^2} / 4\sigma} \right)^2, \quad (6)$$

where  $\sigma$  is the same as in eqs. (1)-(4). Figure 1 shows examples of the field transmissions that are provided by the four occulting masks under study. It should be noted that the VN equivalent mask has field locations with both positive and negative transmissions where the coronagraphs offers positive throughput. While we recognize the physical difficulties in fabricating an achromatic occulting mask of this type, we show this mathematical construction accomplishes the same pupil-shearing as the visible-nuller itself.

Aside from the occulting mask at a telescope focus, a Lyot coronagraph requires there to be a stop at a pupil downstream of the mask. To design an appropriate stop for a particular occulting mask, we first need to compute the pupil field after the occulter. We make the assumption that there exists an idealized, aberration-free exit-pupil,  $P_{exit}(u, v)$ , that is solely defined by the limiting aperture of the telescope (i.e. the diameter of the primary mirror,  $D$ ). As such, the field at the Lyot plane,  $P_{Lyot}(u', v')$ , at a particular wavelength,  $\lambda$ , is computed as

$$P_{Lyot}(u', v') = FT[FT[P_{exit}(u, v)] \cdot o(x, y)]. \quad (7)$$



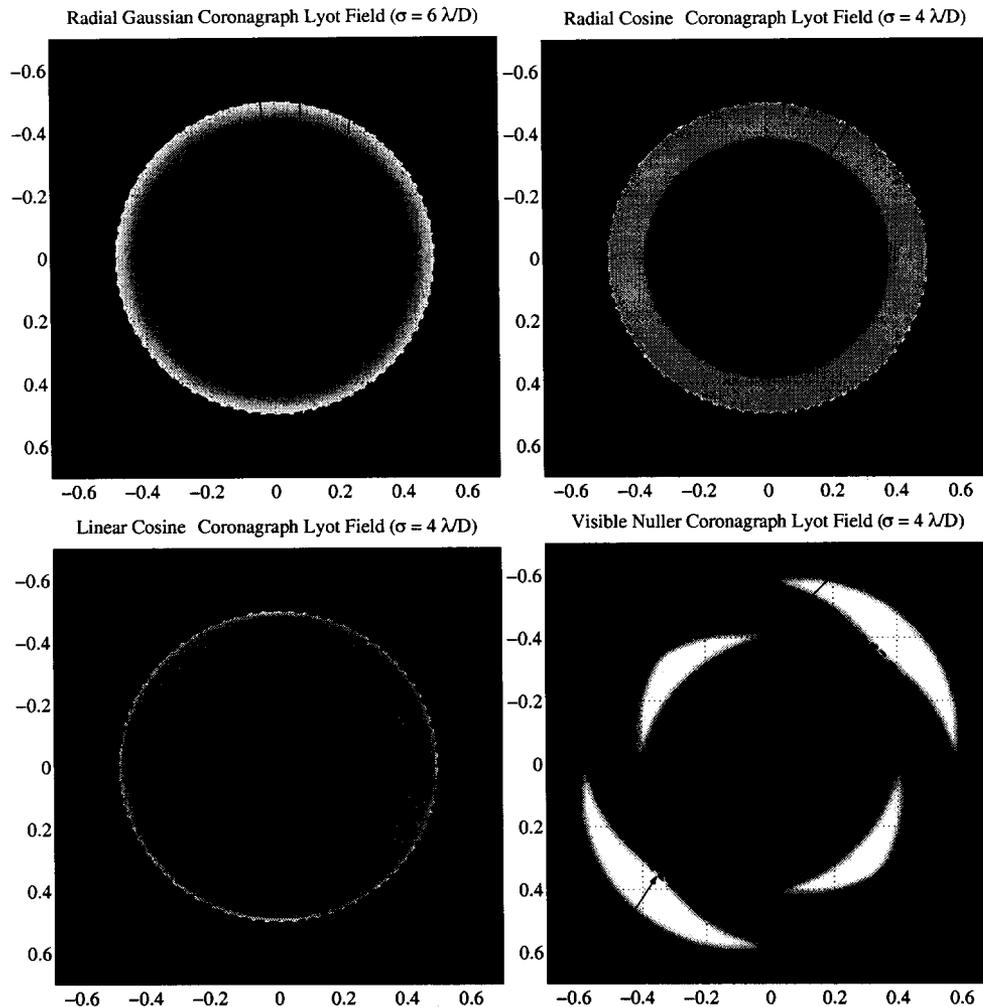
**Figure 1:** The occulter masks under study are the *radial-Gaussian* (top-left), *radial-cosine* (top-right), *linear-cosine* (bottom-left) and *visible-nulter* equivalent (bottom-right) masks. The dashed region is later used for evaluating efficiency (Section 3) and contrast (Section 4).

For the purposes of our study, we employed a fixed design rule for developing the appropriate hard-edged stop at the Lyot plane,  $S_{Lyot}(u', v')$ , given a particular occulting mask. This rule is expressed as

$$S_{Lyot}(u', v') = \begin{cases} 1 & |P_{Lyot}(u', v')|^2 > L_{tol} \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where  $L_{tol}$  is a fixed threshold level that we set at about  $10^{-9}$  of the peak of  $|P_{Lyot}(u', v')|^2$ .

Figure 2 shows the real-part of the Lyot plane that is resultant from each of the four occulting masks. These fields can be generally thought of as  $P_{exit}(u, v)$  minus a low-pass filtered version of  $P_{exit}(u, v)$  with the filter kernel being defined by the FT of  $o(x, y)$ . In the case of the RG occulter, this kernel is infinite in extent and thus there is substantial energy leakage into the interior of the field. The choice of  $L_{tol}$  thus has significant impact in the relative diameter  $D_{Lyot}$  of the designed stop. For the other three occulters, the low-pass filter kernel is finite in extent and results in essentially perfect diffraction cancellation in the interior of their respective Lyot planes. In the case of the VN occulting mask, the pupil-shearing in the Lyot plane demonstrates that this design maintains a strong analogy to the visible-nulter-interferometer imager<sup>7</sup>.

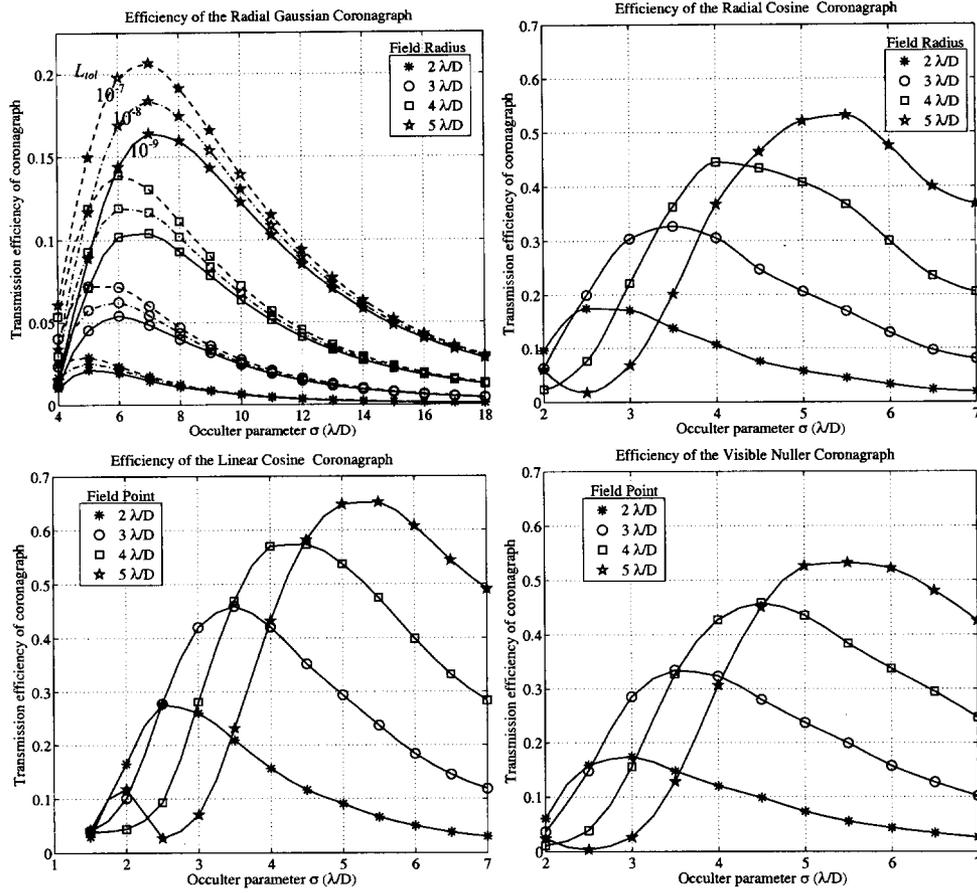


**Figure 2:** The real components of the Lyot pupil field of the four occulting masks under study are shown above. One approach to constructing a hard-edged stop for this pupil is to threshold the field amplitude at a small fraction of the peak field intensity. Because the radial Gaussian occulter is not a band-limited function, the Lyot stop diameter here has a relatively strong dependency upon the choice of threshold.

### 3. EFFICIENCY CONSIDERATIONS FOR CORONAGRAPHS

While it is essential that a coronagraph be designed to provide sufficient contrast for exo-planet detection and characterization, the transmission efficiency through the coronagraph has strong implications for the stability requirement of TPF. With our fixed design rule for specifying a hard-edged stop at the Lyot plane, we can now explore the throughput characteristics of coronagraphs at their inner working angle as function of the occulting spot width parameter,  $\sigma$ .

Figure 3 shows that at a fixed working angle there is some optimal choice of  $\sigma$  that maximizes the transmission efficiency of the coronagraph. It is interesting to note that the optimally efficient  $\sigma$ 's place the peak mask transmission



**Figure 3:** Using the threshold design rule for Lyot stops, the net intensity transmission at the field point is strongly coupled to the width parameter,  $\sigma$ , of the particular occulting spot. For a fixed field angle of interest, there is a particular choice of  $\sigma$  that maximizes the coronagraph efficiency. With the radial Gaussian coronagraph, the choice of Lyot thresholds plays a more significant role in the overall efficiency but this efficiency comes at the cost of increasing the contrast noise floor.

outside the field point. For instance,  $\sigma=3 \lambda/D$  places the peak mask transmission at  $3\lambda/D$  while the peak efficiency of the resulting coronagraph occurs at  $\sigma=3.5 \lambda/D$ . The increased size of Lyot stop that is allowed with an increased occulting spot size makes up for the added mask transmission loss at  $3\lambda/D$ . For two circularly symmetric occulting masks, RG and RC, the efficiency that is plotted in Figure 3 is the averaged intensity efficiency within a field ring (such as was shown Figure 1). The LC and VN occulting masks do have a symmetry that facilitates such averaging. For these cases, we simply averaged intensity efficiency within a circular region bounded by the full-width half-max of the PSF core. We present formal definitions for these regions in the next section.

#### 4. LOW-ORDER ABERRATION SENSITIVITY CONSIDERATIONS FOR CORONAGRAPHS

We now consider the case where the occulter exit-pupil is no longer aberration free. The aberrated pupil is written as

$$P_{exit}(u, v; t) = a(u, v; D) \cdot \exp\left(i \frac{2\pi}{\lambda} \phi(u, v; t)\right), \quad (9)$$

with the phase error function  $\phi(u, v; t)$

$$\phi(u, v; t) = \sum_m \phi_m(t) z_m(u, v; t) + \phi_{mid}(u, v; t) + \phi_{high}(u, v; t) + \phi_{wfc}(u, v; t_c) \quad (10)$$

In equation (10) we show that at any instantaneous time,  $t$ , the exit-pupil phase errors consist of a composition of low-order Zernike modes<sup>9</sup>,  $\{z_m\}$ , mid-spatial frequency errors (that are correctable over some temporal bandwidth), high-spatial frequency errors (that are uncorrectable) and the wavefront control (WFC) correction to the systems that was applied at time  $t_c$ . At any instantaneous moment, the errors are fixed and we can write the intensity at the far-field to the Lyot plane as

$$I(x', y') = \left| FT \left[ P_{Lyot}(u', v') \cdot S_{Lyot}(u', v') \right] \right|^2. \quad (11)$$

Equation (11) is the coronagraphic point-spread function (PSF). To compute the average contrast over some region of interest, we need to also compute the far-field intensity function,  $I_{open}$ , which is the case where the occulting mask is perfectly transparent (i.e.  $o(x, y) = I$ ) but with the Lyot stop present. With these two intensity calculations, we can approximate the average contrast over some set of field points,  $\chi_n$  as

$$C_{\chi_n} = \frac{\iint_{x, y \in \chi_n} I(x, y) dx dy}{|\chi_0| \iint_{x, y \in \chi_0} I_{open}(x, y) dx dy \cdot \iint_{x, y \in \chi_n} |o(x, y)|^2 dx dy} \quad (12)$$

As with our efficiency calculations, for the circularly symmetric occulting masks, we choose a region of interest consisting of a field-ring at a working angle,  $\rho_n$ . The set of all points contained within such a field-ring can be written as

$$\chi_n = \left\{ (x, y) : \left( \rho_n - \frac{\lambda}{2D_{Lyot}} \right) < \sqrt{x^2 + y^2} \leq \left( \rho_n + \frac{\lambda}{2D_{Lyot}} \right) \right\}. \quad (13)$$

For the other occulting masks, we select the region of interest as a field point located at  $(x_n, y_n)$ .

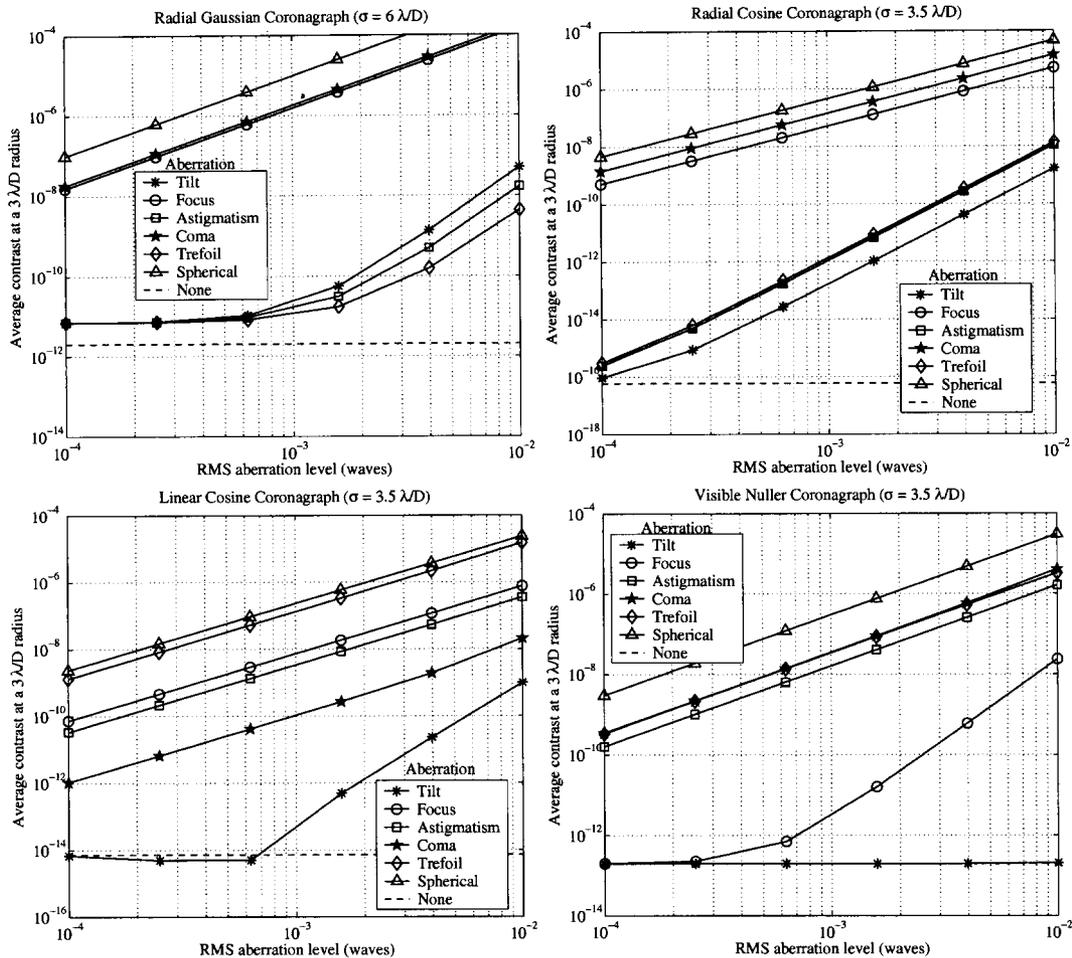
$$\chi_n = \left\{ (x, y) : \sqrt{(x - x_n)^2 + (y - y_n)^2} \leq \frac{\lambda}{2D_{Lyot}} \right\} \quad (14)$$

The set of field points contained in  $\chi_0$  are simply the same as in equation (14) for the case that  $x_n$  and  $y_n$  are both at zero (i.e. about the optical axis). The integral over the unocculted PSF core in equation (12) normalizes the integrated energy in the coronagraphic PSF into contrast, with the smaller numbers being interpretable as lower residual speckle energy. The normalization by the integrated intensity transmission of the occulter accounts for the average throughput losses over the region of interest.

As we mentioned earlier, small drifts in the alignment of the optic preceding the occulter along with deforming modes of the optics themselves include low-order optical errors in the exit-pupil. These errors will no doubt evolve during the integration time used for making planet detections. The growing aberration scatters starlight into the field, particularly at smaller inner working angles and, as we are about to show, the amount of aberration that is tolerable is dependent upon the coronagraph design.

Figure 4 shows that for increasing levels of particular modes, there is an increase in the scattered light level (in this example at  $3 \lambda/D$ ). For certain modes of aberration, this level increases linearly with the variance of the aberration. It so happens that these modes also have the strongest impact on the system performance. Consider the plot for the RC coronagraph. Focus, coma and spherical errors all have a significant impact on the contrast floor at low levels of aberration while tilt, astigmatism and trefoil have much weaker effect. These weaker modes, have a *fourth-order* impact

on contrast and do not affect performance until the aberration is many nanometers in amplitude. We derive this fourth-order relationship for the focus error in the VN coronagraph in the appendix.

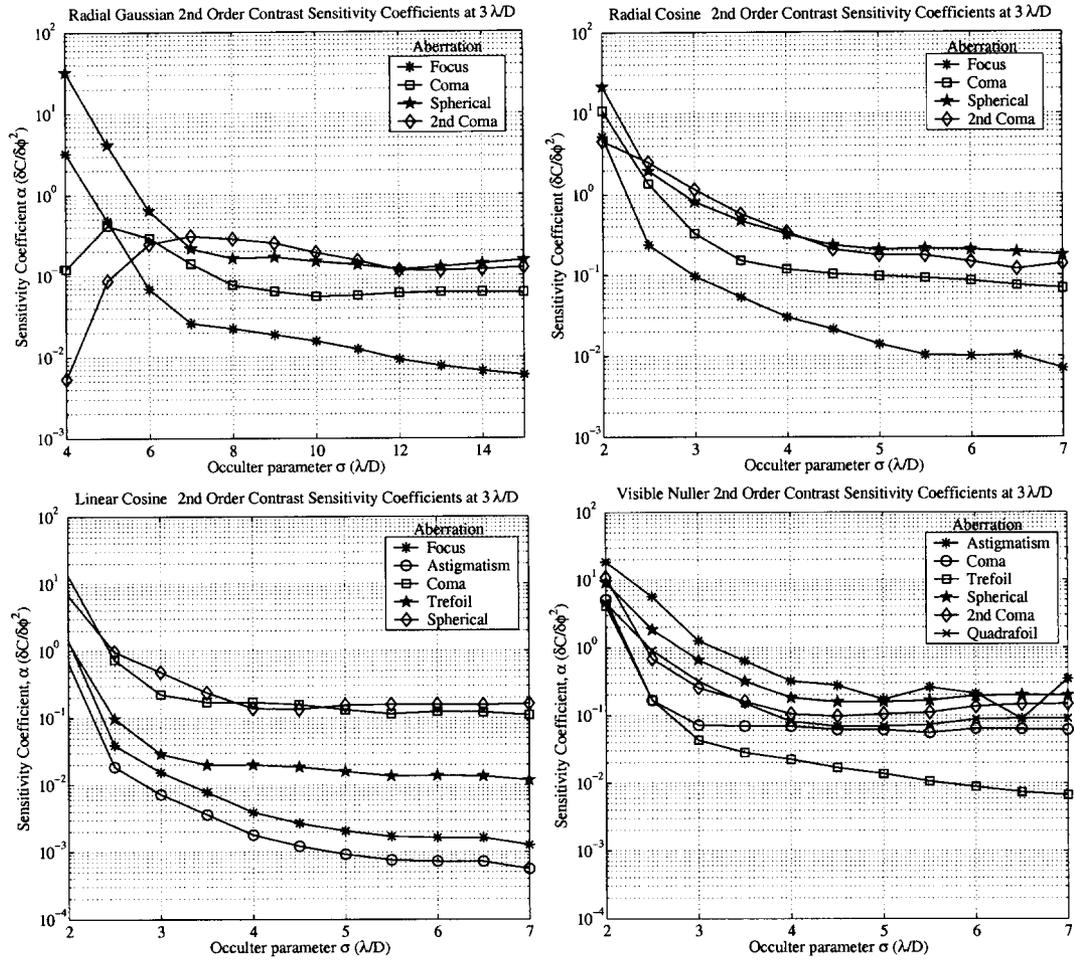


**Figure 4:** The increased contrast noise floor brought forth by the presence of low-order phase aberrations are shown for the four occulting masks. In each case, the particular mask has its  $\sigma$  tuned to provide the optimum coronagraph efficiency at a working distance of  $3 \lambda/D$ .

Using standard sensitivity analysis, we can write the contrast for the modes having second-order sensitivity as a combination of low-order errors as

$$C_{\chi_n} \approx \sum_m \alpha_m \phi_m^2, \quad (15)$$

where  $\alpha_m = \partial C_{\chi_n} / \partial \phi_m^2$ . The coefficient  $\alpha_m$  is the sensitivity of contrast to variance (in waves) of a particular Zernike mode  $m$ . In figure 5, we show the dominant low-order sensitivity coefficients for four coronagraph concepts at a working angle of  $3 \lambda/D$ . In every case, the sensitivity decreases as the occulting spot width increases. So while the efficiency of the RC coronagraph is maximized at  $\sigma=3.5 \lambda/D$ , the contrast floor increase per unit variance of a low-order mode is reduced at larger widths.



**Figure 5:** The contrast sensitivity coefficients for working at  $3 \lambda/D$  are shown above as a function of  $\sigma$ . The sensitivities that are plotted in each case were chosen because they represent the low-order modes that have the strongest impact upon the contrast floor.

### 5. SIGNAL-TO-NOISE RATIO OPTIMIZATION IN THE PRESENCE OF DYNAMIC ERRORS

We have shown that the choice of mask width  $\sigma$  affects both the optical efficiency  $\epsilon$  and the aberration sensitivity  $\alpha$ . The signal-to-noise ratio (SNR) increases with  $\epsilon$  but decreases with  $\alpha$  in the presence of time-variable aberrations. In this section, we derive the mask configuration for optimizing the SNR at a given point in the image plane.

Consider a coronagraph that begins to observe a star immediately after measuring and controlling the wavefront. Ideally, the wavefront does not contribute noise to the measurement – all the diffracted and scattered starlight is rejected by the coronagraph. The only sources of noise are the shot noise of the exoplanet light, the exozodiacal light, and solar system zodiacal light. The required SNR is high enough that shot noise dominates detector read noise.

Now consider what happens when the telescope and coronagraph optics inevitably drift from their initial state. Aberrations arise, and light leaks past the coronagraph into the image plane. Speckles of unknown intensity and position appear and form a background. The SNR is given by

$$SNR = \frac{\text{signal}}{\text{shotnoise} + \text{speckles}} \quad (16)$$

If a planet is present in the image plane, it provides a signal of  $\epsilon p$  photons/s, where  $p$  accounts for the source brightness, aperture diameter, optical reflections, and detector quantum efficiency, while  $\epsilon$  accounts for the fundamental efficiency of the coronagraph. Thus at a given time  $t_o$ , the integrated signal is  $s = \epsilon p t_o$ . The shot noise associated with both the planet photons  $p$  and background photons  $b$  is  $n_s = (\epsilon(p+b)t_o)^{0.5}$ . In the absence of speckles, the SNR increases as the sqrt( $t$ ).

The speckles arise quickly; if the wave front,  $\phi$ , changes linearly with time, e.g.  $\phi = \omega t$ , where  $\omega$  is a constant, then the scattered intensity is proportional to  $t^2$ . The coronagraph filters part of the energy so that the instantaneous speckle intensity in the image plane is  $n_c = \alpha \omega^2 t^2$  where  $\alpha$  is aberration sensitivity coefficient for some dominant mode that is evolving. After integrating for a period  $t_o$ , the total energy of the speckles is

$$n_c = \alpha \omega^2 t_o^3 / 3 \quad (17)$$

While the shot noise-limited SNR increases as sqrt( $t$ ), the speckle energy accumulates as  $t^3$ . Eventually the SNR decreases as  $t^{-2}$  and the speckles overwhelm the background. There exists an optimum uninterrupted integration time,  $t_{opt}$ , that maximizes the SNR for a given  $\epsilon$ ,  $\alpha$ .

The total integration time,  $t_{tot}$ , required to obtain a desired SNR is achieved by averaging together multiple ‘short’ exposures, each  $t_{opt}$  long. (Each one may be minutes to hours in length). After each exposure, the wave front is remeasured and controlled, a process that may take as long as the science exposure. Thus it is desirable to maximize the SNR of individual science exposures to minimize the wave front control overhead. This is where  $t_{opt}$  comes in – it determines the maximum SNR for an uninterrupted exposure. It does *not* determine the maximum rate of SNR return in an exposure – that only occurs at very short exposures where the aberrations are negligible (and assuming detector read noise is negligible). But at such short exposures, the wave front control overhead can be prohibitive. What we offer here is an approach to maximizing SNR while minimizing the overhead.

In what follows, we calculate  $t_{opt}$  and  $SNR_{opt}$ , and then show how the SNR varies as a function of the mask width  $\sigma$ . We will study two cases: first, as discussed above, systematic drift of the optics causes the aberration to increase linearly with time. Second, random drifts cause the aberration to follow a random walk,  $\phi = \omega \sqrt{t}$ . In both these cases we make the assumption that one particular mode dominates the evolving wavefront error.

Case 1: Aberration arises linearly with time

The SNR is readily optimized by minimizing its inverse, the Noise-to-Signal Ratio, NSR. From the discussion above, the NSR is given by

$$NSR = \frac{\sqrt{\epsilon(p+b)t} + \alpha \omega^2 t^3 / 3}{\epsilon p t} \quad (18)$$

We take the derivative of eq. (18) with respect to  $t$  and find that the minimum occurs at

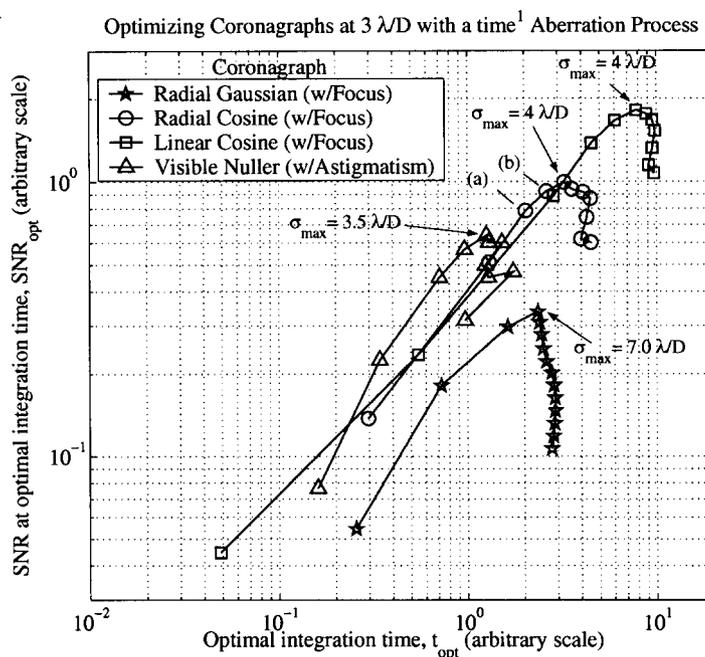
$$t_{opt} = \left( \frac{3\sqrt{(p+b)}}{4\omega^2} \right)^{2/5} \left( \frac{\epsilon^{1/5}}{\alpha^{2/5}} \right) \quad (19)$$

This is a very weak function of mask efficiency; the integration time favors masks that reduce sensitivity to aberrations at the expense of throughput. Substituting eq. (19) into eq. (16), the optimum SNR is determined to be

$$SNR_{opt} = \left( \frac{\epsilon^{3/5}}{\alpha^{1/5}} \right) \left( \frac{0.755p}{\omega^{2/5}(p+b)^{2/5}} \right) \quad (20)$$

We also point out that at time  $t_{opt}$ , the shot noise contribution  $\sqrt{\epsilon(b+p)t_{opt}}$  is four times larger than the speckle energy  $\alpha\omega^2 2t_{opt}^3 / 3$ . This confirms our earlier assumption that the speckle shot noise is well below the signal and background shot noise.

In figure 6, we plot the  $t_{opt}$  and  $SNR_{opt}$  for a signal evaluated at  $3\lambda/d$ , for the four coronagraphs under consideration. In the case of RC occulter, the SNR is maximized at  $\sigma = 4\lambda/D$ . Even though the transmission efficiency (for  $\sigma = 4\lambda/D$ ) of 30.6% is a little below the peak transmission of 32.7% ((see Figure 3), the weak dependence of  $SNR_{opt}$  on  $\alpha$  shifts the peak away from maximum transmission in favor of reduced aberration.



**Figure 6:** This is a plot of the optimal integration time for each coronagraph design point against the resulting optimal SNR under the assumptions that the dominant aberration increases linearly with time. Each line represents a particular coronagraph concept and the individual markers along the lines represent different choices of  $\sigma$ , with  $\sigma$  increasing from left to right. The points, which are demarked as (a) and (b), are referred to in the narrative below.

Ultimately, one wants to optimize the final SNR by choosing the best  $t_{opt}$  to maximize the SNR per unit time while minimizing wave front control overhead. By plotting  $SNR_{opt}$  vs.  $t_{opt}$  (Figure 6), we demonstrate how the SNR per unit time varies with mask width. Given a point on the curve (e.g. (a) on the RC curve), the rate of improvement of SNR with time has a slope (on the log-log scale) of 0.5 in the shot-noise limit. In practice the slope will be somewhat less than 0.5 due to the wavefront control overhead. Now moving from left to right along the curve (increasing  $t_{opt}$  and  $\sigma$  from point (a) to point (b)), the slope on the log-log scale is close to unity. This means that the larger masks have a higher SNR per unit time than the smaller masks used with multiple integrations. The slope at the point (b) on the RC curve,

where  $\sigma=3.5$ , is approximately 0.5. Beyond this point, the larger mask has lower SNR per unit time than a smaller mask used with multiple shorter exposures. The exact choice of the optimum mask depends on the wavefront control overhead.

Case 2: aberration evolves in a random walk.

For this case, the integrated speckle energy at time  $t_0$  is given by  $n_c = \alpha\omega^2 t^2/2$ , and the NSR is expressed by

$$NSR = \frac{\sqrt{\varepsilon(p+b)t + \alpha\omega^2 t^2/2}}{\varepsilon p t} \quad (21)$$

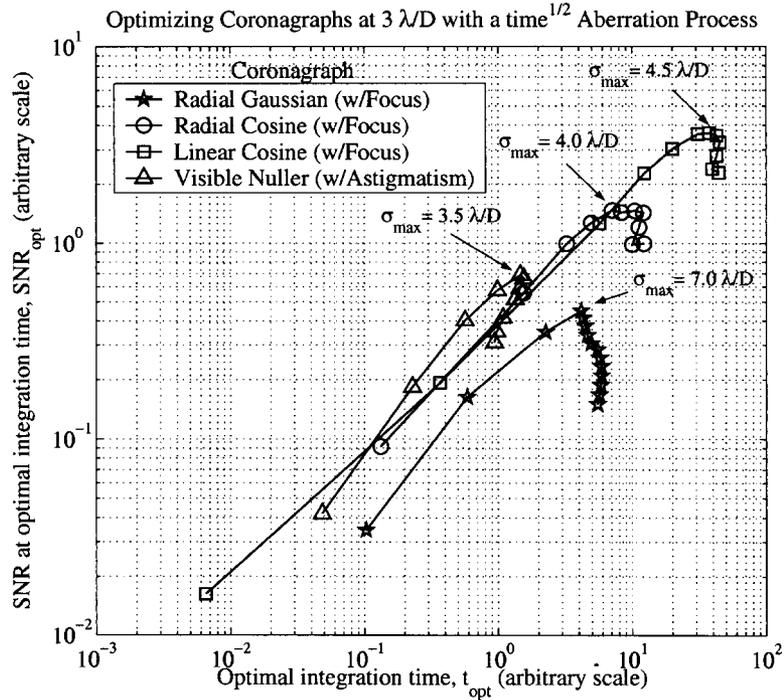
Minimizing NSR with respect to  $t$ , we find

$$t_{opt} = \left( \frac{\sqrt{(p+b)}}{\omega} \right)^{2/3} \left( \frac{\varepsilon^{1/3}}{\alpha^{2/3}} \right) \quad (22)$$

and the optimum SNR is

$$SNR_{opt} = \left( \frac{\varepsilon^{2/3}}{\alpha^{1/3}} \right) \left( \frac{1.5p}{\omega^{1/3}(p+b)^{1/3}} \right) \quad (23)$$

The solution leads to a speckle energy equal to one half the shot noise level, *twice* the value allowed in case 1.



**Figure 7:** With the assumption that the dominant aberration follows a random walk process, the range of achievable integration times and SNRs is greatly increased. Nevertheless, each coronagraph has essentially the same  $\sigma$  that maximizes the integration time allowable before requiring a wavefront correction.

In figure 7, we plot  $SNR_{opt}$  vs.  $t_{opt}$  for varying mask width  $\sigma$ . As with Case 1, the plot shows that the SNR per unit time is constant until the mask becomes so large ( $>4\lambda/D$ ) that the diminishing efficiency is not offset by reduced sensitivity to aberrations. The more slowly evolving aberration considered here not surprisingly allows longer integration times and higher SNR than the linearly evolving aberration.

## 6. CLOSING REMARKS

In this paper, we have shown that there are many trade-offs to be considered in the design of coronagraph masks. It is, of course, essential that a coronagraph provide the contrast necessary to achieve planet detections and their subsequent characterizations. The efficiency of the design along with its aberration sensitivity play a central role in determining the maximum uninterrupted integration time allowable that will still offer an improved contrast in the presence of evolving aberrations. The fixed costs of updating the wavefront correction coupled with its finite accuracy makes such an optimization of a coronagraph design that much more important. It remains to be seen how other occulting functions as well as pupil apodization and pupil shaping methods will compare to the simple hard-edged Lyot stop design approach employed in this paper.

To be successful, the wavefront controls methods employed by TPF must be operationally efficient and highly accurate so that a substantial portion of the mission can be used for science. Even so, TPF will have to be an extraordinarily stable telescope that employs coronagraphs optimized for the particular telescope optical sensitivities weighted by their excitation characteristics.

## ACKNOWLEDGEMENTS

The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

## APPENDIX: VISIBLE NULLER 4<sup>TH</sup> ORDER SENSITIVITY TO FOCUS

As with a band-limited mask, the Visible Nuller coronagraph aperture multiplies the image by a band-limited function, in this case two orthogonal sine waves. In pupil space, this is equivalent to convolving the pupil function with 4 delta functions,  $\delta(x-x_0)$ ,  $-\delta(x+x_0)$ ,  $\delta(y-y_0)$ , and  $-\delta(y+y_0)$ , where  $x_0$  and  $y_0$  depend on the spatial period of the sine waves. If the pupil function contains a focus term of amplitude  $a$ , we can express the pupil amplitude function as

$$P = \exp\left(ia\left(x^2 + y^2\right)\right) \quad \forall \sqrt{x^2 + y^2} \leq 1 \quad (24)$$

inside the aperture and  $P=0$  outside. The amplitude at the edge of the pupil,  $ax^2$ , is  $\ll 1$  radian. After convolving with the four delta functions, the field amplitude is given by

$$\begin{aligned} E \propto & \left( \exp\left(ia\left((x-x_0)^2 + y^2\right)\right) - \exp\left(ia\left((x+x_0)^2 + y^2\right)\right) \right) \\ & \times \left( \exp\left(ia\left(x^2 + (y-y_0)^2\right)\right) - \exp\left(ia\left(x^2 + (y+y_0)^2\right)\right) \right) \end{aligned} \quad (25)$$

Rearranging terms, we find that this is equivalent to

$$E \propto \exp\left(ia\left(x^2 + y^2 + x_o^2 + y_o^2\right)\right) \cdot 4 \sin(2axx_o) \sin(2ayy_o) \quad (26)$$

$$\cong a^2 x_o y_o xy$$

The field at the output pupil appears to be an astigmatism function whose amplitude is proportional to the square of the focus amplitude,  $a^2$ . The intensity then varies as

$$I \propto a^4 x_o^2 y_o^2 \quad (27)$$

That is, it is proportional to the fourth power of the aberration and it increases as the ‘shear’ terms  $x_o$  and  $y_o$  increase. The latter result shows that the leakage increases as the sine wave frequency increases (that is, the mask width gets narrower). Small masks – those designed to allow planet searches close to the target star - increase sensitivity to aberration (in this case, as the 4<sup>th</sup> power of the shear!!) while wide masks by virtue of their small pupil shear, decrease aberration sensitivity, as first pointed out to us by M. Shao (private communication).

Similar analysis reveals that the linear cosine mask behaves in a similar fashion with respect to one of the two Zernike astigmatism terms. We see (but have not performed the analytical derivation to verify) the same 4<sup>th</sup> order behavior for both astigmatism and both trefoil terms with the radial cosine mask.

## REFERENCES

1. N. J. Kasdin, R. Vanderbei, D. Spergle, and M. Glittman, “Optimal Shaped Pupil Coronagraphs for Extrasolar Planet Finding,,, Proc. SPIE, vol. 4860, pp. 240-250, (Waikoloa 2002).
2. P. Nisenson, C. Papaliolios, “Detection of Earth-like Planets Using Apodized Telescopes,,, 2001 ApJ vol 548, L201.
3. R. Lyon, J. Hollis, J. Dorband, “Comparative Optical Analysis of Extra-Solar Planetary Imaging Techniques,,, Proc. SPIE, vol. 4860, pp. 251-265, (Waikoloa 2002).
4. D. Rouan et al, “Estimated Performances of a Four-Quadrant Phase mask Coronagraph for Planet Detection,,, Proc. SPIE, vol. 4860, pp. 192-199, (Waikoloa 2002).
5. O. Guyon, “Phase Induced Amplitude Apodization of Telescope Pupils for Extrasolar Terrestrial Planet Imaging,,, 2003, Astron. And Astrophys, vol 404, pp. 379-387.
6. M. J. Kuchner and W. A. Traub, “A Coronagraph with a Band-Limited Mask for Finding Terrestrial Planets,,, ApJ vol 570, 2002.
7. B. Mennesson et al, "Optical Planet Discoverer: how to turn a 1.5 m telescope into a powerful exo-planetary systems imager," SPIE vol 4860 pp. 32-44, (Waikoloa, 2002).
8. D. Wilson, et al, “Eclipse apodization: realization of occulting spots and Lyot stops,,, Proc. SPIE, vol. 4860 (Waikoloa 2002).
9. R. J. Noll, “Zernike polynomials and atmospheric turbulence,,, J. Opt. Soc. Am., vol. 66(3), pp. 207-211, 1976.
10. F. Malbet, J. W. Yu and M. Shao, “High Dynamic-Range Imaging Using a Deformable Mirror for Space Coronagraphy,,, PASP, vol. 107, pp. 386-398, 1995.
11. J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics*, John Wiley & Sons, New York, 1978.
12. B. R. Frieden, *Probability, Statistical Optics, and Data Testing*, Springer-Verlag, Berlin, 1991.
13. D. J. Schroeder, *Astronomical Optics*, Academic Press, San Diego, 2000.