Speckle noise in highly corrected coronagraphs

E. E. Bloemhof
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA USA 91109

ABSTRACT

Speckles in a highly corrected adaptive optic imaging system have been studied through numerical simulations and through analytic and algebraic investigations of the Fourier-optical expressions connecting pupil plane and focal plane, which simplify at high Strehl ratio. Significant insights into the behavior of speckles, and the speckle noise caused when they vary over time, have thus been gained. Such speckle noise is expected to set key limits on the sensitivity of searches for companions around other stars, including extrasolar planets. In most cases, it is advantageous to use a coronagraph of some kind to suppress the bright primary star and so enhance the dynamic range of companion searches. In the current paper, I investigate speckle behavior and its impact on speckle noise in some common coronagraphic architectures, including the classical Lyot coronagraph and the new four quadrant phase mask (FQPM) concept.

Keywords: adaptive optics, speckles, atmospheric turbulence

1. INTRODUCTION

Spectacular recent advances in the development of adaptive optics (AO) systems for astronomy have delivered rather highly corrected images in the near-infrared. As one example, the PALAO system on the Palomar 5 meter telescope, with 241 controlled deformable-mirror (DM) actuators and a 16x16 Shack-Hartmann wavefront sensor, achieves Strehl ratios $S$ greater than ~0.6 at K band (2.2 μm). At such high correction, regularities begin to emerge in the residual halo of focal-plane speckles that account for the fraction of image power, $(1-S)$, not contained in the diffraction-limited core. A sample short-exposure image is shown in Figure 1.

Figure 1 – Highly corrected (Strehl ratio $S$~0.6; DM actuator density $D/a$~16, where $a$ is the DM actuator spacing) short-exposure image at 2.2 μm obtained by the Palomar adaptive optics system (PALAO). (Left) Surface plot shows the diffraction-limited core with symmetric structures on the first “Airy” ring, surrounded by the speckle halo. (Right) Logarithmically stretched two-dimensional intensity plot brings out the fainter features in the outer halo, including a few instrumental ghosts and “waffle mode” artifacts. Regularities in speckle structure should become significant at levels of correction only slightly higher than this.
The image in Figure 1 was obtained with an exposure time of about 0.5 s, comparable to \( \tau_0 \), the characteristic timescale for the atmospheric turbulence above the telescope pupil to rearrange itself. Hence it represents the focal-plane intensity distribution that results from a single phase screen, frozen in time, above the telescope. It was also obtained in a relatively narrow spectral band, minimizing chromatic smearing of speckles.

Images such as these have inspired detailed study of the properties of the speckle halo. One highly suggestive effect in particular is clearly visible in the short exposure image of Figure 1: slowly varying non-common-path aberrations internal to the AO system cause temporally persistent knots of 3- or 4-fold symmetry on the inner Airy ring. As the instrument experiences changing flexure at Cassegrain focus with telescope pointing, these aberrations creep into the difference between the wavefront propagated through the AO system's science and wavefront sensing paths, which is calibrated only periodically. (The science path directs near-infrared light into a sophisticated science camera/spectrometer; the wavefront sensing path directs visible light into a very high frame rate CCD fed by a Shack-Hartmann.) Figure 1 suggests an interaction in some sense between the non-common-path aberrations and the diffraction-limited point-spread function (PSF), particularly the Airy rings that surround the image core. This connection motivates an interpretation of related effects that will be studied in more detail in the next section, but may be thought of loosely as an interaction between the diffraction-limited PSF and the instantaneous phase screen presented by aberrations impressed across the telescope pupil by atmospheric phase fluctuations.

2. SPECKLES AT HIGH CORRECTION IN THE DIRECT IMAGING CASE (NO CORONAGRAPH)

Image formation in a telescope may be modeled very simply by a Fourier optical picture in which the optical fields in the focal plane are related to those in the pupil plane by Fourier transformation. The input pupil is the telescope aperture, having two spatial coordinates \((\xi, \eta)\), and described by an aperture function \( A(\xi, \eta) \) and a phase function \( \Phi(\xi, \eta) \), each a real function. (Scintillation is neglected.) The effect of the adaptive optics system is to reduce the phase function to a very small "remnant" level that cannot be corrected, having mean-square value \( (1+\delta) \). The image-plane intensity in a short exposure due to a single remnant phase screen is

\[
\text{Intensity } I(x, y) = \left| \text{F.T.}[A(\xi, \eta)e^{i\Phi(\xi, \eta)}] \right|^2 = \left| Ae^{i\Phi} \right|^2 \tag{1}
\]

Here \( \text{F.T.}[...] \) denotes two-dimensional spatial Fourier transformation into the \((x, y)\) coordinate system of the focal plane. (The Fourier transform may also be denoted, more compactly, by an overbar.) The exponential in Equation (1) may sensibly be expanded in a truncated Taylor series when the remnant phase \( \Phi \) is small:

\[
\text{Intensity } I(x, y) = \left| Ae^{i\Phi} \right|^2 \approx \left| A[1 + i\Phi] \right|^2 \tag{2}
\]

Carrying out the algebra for this first-order expansion, one finds three terms, one that is the diffraction-limited PSF corresponding to aperture \( A \), one that is linear in the Fourier transform of \( \Phi \), and one that is quadratic:

\[
\text{Intensity } (x, y) \approx \left| A \right|^2 + 2 \text{Re}\left[ i \left( \overline{A} \otimes \overline{\Phi} \right) \overline{A^*} \right] + \left| A \otimes \overline{\Phi} \right|^2 \tag{3}
\]

The symbol \( \otimes \) denotes two-dimensional convolution, and the asterisk denotes complex conjugation. "Re[J]" denotes the real component of a complex-valued expression. The first term in Equation (3) is by far the largest in amplitude, since the remnant phase function is small in the context of high-correction imaging. The two speckle terms, the last two in Equation (3), represent two distinct kinds of patterns of light that has been distributed in the spatially extended speckle halo. The relative brightness of speckles from these two terms will vary as a function of correction, both Strehl ratio \( S \) and deformable mirror (DM) actuator density \( D/a \), but speckles from the second term must eventually dominate as the degree of adaptive correction gets very large. The second term of Equation (3) may be called the "linear" term, being roughly linear in \( \overline{\Phi} \), and the speckles it describes called "linear-term" or simply linear speckles. The third term of Equation (3) and its speckles may similarly be called "quadratic".
If the aperture $A$ is an unobstructed and unapodized (or "clear") circle of diameter $D$, the first term in Equation (3) takes on the familiar form of an Airy pattern, the diffraction-limited PSF for this pupil geometry. The PSF is simply an image of an unresolved star at perfect correction, i.e. $\Phi = 0$ or $S = 1$:

$$PSF(x, y) = |\overline{A}|^2 = Airy = \left(\frac{2J_1(\pi D \theta / \lambda)}{\pi D \theta / \lambda}\right)^2$$

The unapodized case allows some simplification, since then $A\Phi = \Phi$; a symmetric aperture, such as a circle, has a real Fourier transform. Then Equation (3) becomes the following expression for PSF plus two kinds of speckles:

$$Intensity (x, y) = |\overline{A}|^2 - 2\text{Im}[\Phi] \overline{A} + |\Phi|^2$$

An unusual property of the second term is that it inherits the nulls of the PSF from the multiplicative factor $\overline{A}$, and so linear speckles are "pinned" to (spatially located on) Airy rings.

The symmetry properties of the speckle terms are very important, and follow from the fact that the Fourier transform of $\Phi$, a real function, is Hermitian. It follows that the third, or quadratic, term has spatial (centro)symmetry, while the second, or linear, term has spatial antisymmetry. One consequence is that the linear-term speckles have zero net image power when integrated over the focal plane in a single short exposure (or single realization of the phase screen), though individual linear-term speckles may be bright. Linear-term speckles are in fact dominant in single-speckle intensity at sufficiently high Strehl ratio $S$ and DM actuator density $D/a$, particularly on the inner Airy rings.

It is formally possible to retain additional terms in the Taylor series expansion for the phase exponential in Equation (2). However, some recent work on properties of higher order terms has confirmed that the new terms that result are less interesting physically than the two terms in Equation (3) derived from the first-order expansion. The brightest term not included in Equation (3) is brightest near the origin, but this is primarily a correction to the PSF, as may be seen from the following decomposition (here $\langle \ldots \rangle$ denotes an average over the pupil):

$$-A \text{Re}\left\{\Phi^2\right\} = -A \text{Re}\left\{\Phi^2 - \langle \Phi^2 \rangle + \langle \Phi^2 \rangle\right\}$$

$$= \text{speckles} - A \text{Re}\left\{(1-S)\overline{A}\right\}$$

Adding this PSF correction to the first term of Equation (3) gives a net intensity equal to $S \times PSF$, which is what would be expected for an AO system operating at Strehl ratio $S$: it indicates that a fraction $S$ of the total light is put into the diffraction-limited PSF. This PSF correction is nominally static, and should not be counted as a speckle term. The true speckles from this term, the first term in Equation (6), are much fainter. They are pinned, spatially symmetric, and they have zero temporal mean at any point in the focal plane.

3. A CORONAGRAPH AT HIGH ADAPTIVE CORRECTION

A strength of adaptive imaging is the ability to search for companions close to bright stars, which indeed provide guide stars for the AO wavefront sensor. These searches are of great scientific interest, as companions may include planets, or even earth-like planets, which are otherwise detected only by somewhat indirect means. It is natural to couple an AO system to a coronagraph, for even more powerful companion detection capability, because a coronagraph is designed to suppress the bright on-axis star and permit the regions around it to be searched to a much deeper level. A coronagraph will not only suppress the core of the image of the on-axis star, but also all of the secondary diffraction ("Airy") rings extending out into the focal plane, where faint companions will be sought.
A schematic depiction of an astronomical coronagraph is shown in Figure 2. The imaging properties are modified by the insertion of two masks, one in a focal plane and one in the pupil plane. The first of these is an opaque spot a few times the diffraction limit in diameter, in the classical Lyot coronagraph, but some innovative approaches have emerged recently, including the four quadrant phase mask (FQPM) shown in Figure 2. The FQPM applies a phase shift of 0 or π alternately in four quadrants of the focal plane, and this can in principle achieve very high rejection of the central star if the aperture is unobscured and there is very little remnant tip/tilt error. Both focal-plane masks work by diverting light from the bright central star to the periphery of the pupil plane, and so both employ a similar “Lyot” mask in the reimaged pupil, which blocks light from the outer edges and slightly reduces the spatial resolution of the final image.

![Figure 2 - Schematic of a coronagraph on an astronomical telescope. The telescope primary, represented by a lens, defines the entrance pupil P1. Focal-plane optical fields are related to those in the pupil plane by Fourier transformation. A focal-plane mask “f” is inserted in the first focal plane FP1, and the light at the periphery of the reimaged pupil P2 is blocked by a Lyot mask before the final image is formed.](image)

In the coronagraph case, optical fields in the first focal plane and the reimaged pupil plane are intentionally modified by the masks “f” and “p” placed there. As shown in detail in Figure 2, these multiply the incident optical fields so that the final image intensity becomes, in analogy to Equation (1),

$$\text{Intensity } I(x, y) = \left| p(\xi, \eta) f(x, y) A(\xi, \eta) e^{i\Phi(\xi, \eta)} \right|^2 = \left| f A e^{i\Phi} \right|^2 \tag{7}$$

One may now carry out the speckle expansion for a coronagraph in analogy with the direct-imaging case of Section 2. Assuming the correction is high so that the remnant phase Φ is small, the phase exponential may be expanded to retain only the lowest non-trivial order:
\[ I(x, y) \approx |pfA[1 + i\Phi]|^2 = |pfA|^2 + 2\text{Re} \left\{ i \left( pfA \right)^* \left( pf\Phi \right) \right\} + |pf\Phi|^2 \]  

(8)

As in the direct-imaging case, three terms are obtained to this order. The first is the diffraction-limited PSF of the bright on-axis star, modified (presumably, strongly attenuated) by the action of the coronagraph. A convenient phenomenological expression for this on-axis attenuation is

\[ \frac{|pfA|^2}{R_{\text{cor}}} \]

(9)

where the coronagraphic suppression factor \( R_{\text{cor}} \) is greater than unity and possibly large. It is a function of position within the focal plane, and as defined here applies only to an on-axis source, so is not in any strict sense a PSF. It may be quoted as an azimuthal average over the image, for a given radial offset from the central star; it may also be sensible to average over some radial region, since the coronagraph will in general shift the positions of Airy rings.

4. PROPERTIES OF THE CORONAGRAPHICALLY-FILTERED SPECKLE TERMS AT HIGH CORRECTION

An extremely important property of speckles in the direct imaging case at high correction is that they appear in patterns of definite spatial symmetry. These symmetries occur in the case of speckles in a coronagraph as well, by fairly similar arguments, and can probably be exploited to help suppress the noise level due to speckles. Properties such as temporal and spatial means and spatial pinning to Airy rings hold for coronagraphic speckles in fairly close analogy with the direct imaging case. An additional interesting property, evident from the form of Equation (8) or from numerical simulations, is that the coronagraph has little effect on the intensity of quadratic speckles, but will sharply attenuate the intensity of linear-term speckles. In the following sections the properties of the two dominant speckle types shown in Equation (8), the ones arising from a first-order expansion of the phase exponential in Equation (7), are presented; as with direct imaging, terms from higher orders in the expansion are less important physically to imaging at high correction.

a. Quadratic-Term Speckles

The third term in Equation (8) is roughly quadratic in the coronographically-filtered speckle amplitude, \( pfA \), and is the natural analog of the quadratic speckles seen in the direct-imaging case, which arise from the term quadratic in the (unfiltered) speckle amplitude \( \Phi \). In fact, as is shown in Figure 3, quadratic-term speckles are little affected by the coronagraph, as might be expected if one thinks of them as sources in the first focal plane of the telescope that are reimaged by subsequent coronagraph optics: a coronagraph is designed to pass off-axis sources with little attenuation. As the Figure shows, both the Lyot and FQPM coronagraphs modeled here preserve the basic off-axis quadratic speckle pattern, though the Lyot blocks speckles near the center of the field behind its occulting spot. More subtly, since each coronagraph reduces the pupil diameter with a Lyot mask in the reimaged pupil, each slightly degrades the spatial resolution and peak intensity of quadratic speckles.

Also apparent in the Figure is the spatial (centro)symmetry of quadratic speckles. This results from the fact that the phase screen \( \Phi \) is real, so its Fourier transform is Hermitian. For the real filters \( f, p \) of either the Lyot or FQPM coronagraphs, the coronographically-filtered speckle amplitude is Hermitian, so its squared modulus is spatially symmetric.

Not surprisingly, the symmetric speckles in a coronagraph are distributed over a halo of diameter \( \sim \lambda/a \), where \( a \) is the typical transverse coherence scale of the phase screen (equal to \( r_\nu \) for ground-based observations). They may be found anywhere in this region: they experience no pinning to Airy rings, as the linear-term speckles do. The quadratic speckles, when integrated over the focal plane, fully account for the fractional light contained in speckles, \( I-S \).
b. Linear-Term Speckles

The second term in Equation (8) is roughly linear in the coronagraphically-filtered speckle amplitude, and also contains a multiplicative factor proportional to the amplitude of the diffraction pattern from the bright on-axis star. As in the direct imaging case, this leads to the spatial pinning of linear speckles onto the secondary maxima (Airy rings) of that pattern: i.e., these speckles vanish on the nulls of the on-axis diffraction pattern. Because the on-axis diffraction pattern is strongly attenuated by the action of the coronagraph, it is expected that the linear-term speckles will be attenuated as well; this can be seen to occur in the numerical simulations presented in Figure 4.

Coronagraphic linear-term speckles occur in spatially antisymmetric patterns, in close analogy with linear speckles in the direct imaging case. This is apparent in Figure 4, and can also be seen by considering the algebraic form in Equation (8). The coronagraphically filtered speckle amplitude and PSF amplitude both derive from the Fourier transform of a real function, and passing through the real filters f, p result in both terms being Hermitian; the imaginary part of their product is then spatially antisymmetric.
Figure 4 – Numerical simulations of linear-term speckles showing they are substantially attenuated by passage through a coronagraph. (Top left) the PSF for a telescope with no coronagraph sets the spatial scale. (Top right) the linear speckle pattern for a clear circular aperture with no coronagraph. (Bottom left) for the same phase screen, the linear speckle pattern after passage through a Lyot coronagraph. (Bottom right) again for the same phase screen, the linear speckle pattern after passage through a four quadrant phase mask (FQPM) coronagraph. Intensity scales are arbitrary, but the same in all three speckle panels.

From the form of the second term in Equation (8), the distribution of linear-term speckles is peaked towards the center of the field, where the Airy rings are brightest, and this is apparent in Figure 4. The spatial antisymmetry means that linear speckles contribute nothing to the image power integrated over the focal plane, although individual speckles may be bright. (As with direct imaging, linear speckles can have negative intensity, but this simply implies partial reduction of an Airy ring, in such a way that the net intensity remains non-negative.)

5. ESTIMATED INTENSITIES OF THE CORONAGRAPHICALLY-FILTERED SPECKLE TERMS AT HIGH CORRECTION

The properties of the remnant speckle halo will depend on which of the two brightest speckle types presented in the previous section is dominant under given conditions. To determine that, it is useful to derive simple heuristic estimates for the typical intensities of speckles of the two types. This derivation assumes that speckles are diffraction-limited objects of diameter \( \sim \lambda/D \), randomly distributed within a halo of diameter \( \sim \lambda/a \), where \( a \) is a characteristic transverse coherence scale of the remnant phase screen after correction. On the ground, the Fried scale \( r_0 \) of the atmosphere is a simple parametrization of the spatial frequency content of turbulence, and the spacing of deformable mirror actuators is typically chosen to roughly equal \( r_0 \). In space, the deformable mirror actuator spacing \( a \) is taken as a reasonable measure, but there could be other characteristic scales in the optics.
The speckle halo diameter $\lambda/a$ will obviously be populated by $-(D/a)^2$ diffraction-limited speckles; Roddier has calculated the prefactor to find the following expression for the number of speckles in the halo:

$$N_S \approx (0.342) \left(\frac{D}{a}\right)^2$$

(10)

The total light in the halo is $(1-S)$ for correction characterized by Strehl ratio $S$, and the typical intensity of a single speckle is found by dividing this light among the number of speckles found in Equation (10). (The spatial smoothness of the corrected wavefront is thus playing a key role, along with the Strehl ratio, by determining the size of the speckle halo and hence the number of speckles among which the uncorrected light is divided.) A refinement not present in the direct imaging case is a slight reduction in peak intensity of any individual speckle with the spatial broadening implied by a Lyot mask of diameter $D_p$, typically 90% of $D$. Including this effect, the intensity of a typical speckle is

$$\text{quadratic} = \left| \frac{p f \Phi}{p f A} \right|^2 \approx \frac{(1-S)}{S} \frac{1}{0.342} \left( \frac{a}{D} \right)^2 \left( \frac{D_p}{D} \right)^2$$

(11)

Here the peak height is referred to the height of the corrected PSF (of relative intensity and peak height $S$), rather than to that of an ideal ($S=1$) PSF; the difference is negligible at high Strehl ratio. The typical speckle intensity of linear-term speckles may be found to be, using Equations (11) and (9):

$$\frac{\text{quadratic}}{\text{linear}} \approx \sqrt{2} \left| \frac{p f \Phi}{p f A} \right| \approx \sqrt{\frac{R_{cor}}{2}} \frac{1}{\sqrt{2}} \sqrt{\frac{(1-S)}{S} \frac{1}{0.342} \left( \frac{a}{D} \right)^2 \left( \frac{D_p}{D} \right)^2}$$

(12)

These equations may be used to estimate speckle noise from the two types, using the following equation derived by Racine et al.:

$$\text{variance(speckles)} \approx \frac{1}{16} F_S^2 \left( \frac{t}{\tau_0} \right)$$

(13)

where $F_S$ is the intensity of a typical speckle, $t$ is the integration time, and $\tau_0$ is the timescale for realization of a new, statistically independent phase screen: milliseconds on the ground, but possibly a rather long time in space. In the latter case, a more appropriate picture may be the potential for false companion detections, where an unusually bright speckle mimics an off-axis source. These effects are potentially important in space coronagraphy being carried out now: for example, the HST/NICMOS coronagraph operates under conditions equivalent to $S=0.98$ and $D/a=21$. (There is no adaptive correction, but the wavefront is flat because the optics are of high quality and situated above the earth’s atmosphere; an equivalent to the DM actuator density is the inverse of the observed scattered-light halo diameter.)

6. CONCLUSIONS

The Fourier-optical algebra for a coronagraph coupled to an adaptive optics system operating at high correction has been presented. Two speckle types, from the first-order expansion of the phase exponential, are found to be most important in determining image properties, in close analogy with the direct-imaging case. These terms, roughly linear and quadratic in the Fourier transform of the remnant phase screen, are spatially antisymmetric and symmetric, respectively. Quadratic-term speckles are little affected by passage through a coronagraph, while linear-term speckles are substantially attenuated, along with the diffraction pattern from the on-axis source.

As in the direct-imaging case, linear-term speckles are pinned to Airy rings, and have zero temporal mean at any point in the focal plane. Their spatial antisymmetry means they contribute nothing to the net image power contained in speckles, a fraction $(1-S)$ of the total from the central star, but the intensity of individual speckles may be great. Estimates of
The intensities of typical speckles for different conditions are given. It is found that linear-term speckles are brightest on the inner Airy rings, and become dominant at high Strehl ratio $S$ and high deformable-mirror actuator density $D/a$.

The speckle properties presented here suggest the following tactics for suppressing the speckle noise or false companion detections that could result. First, antisymmetric linear-term speckles might be suppressed by use of a coronagraph or by observations at the position of Airy nulls. (Use of broad spectral bands will tend to cancel and suppress linear speckles, too.) Then each individual short exposure could be antisymmetrized by subtracting a spatially-inverted copy of itself, suppressing symmetric speckles from the quadratic term (this will in general delete half of the image power from any bona fide companion, which may be resolved into equal parts symmetric and antisymmetric image signature.) If these processed short exposures are co-added, the antisymmetric part of the true companion image will persist, while the linear-term speckles, having zero mean at any point in the image, should eventually average away. Particularly in the case of space-based observations, the ideality of successive phase screen realizations (randomness, statistical independence, timescale over which they are refreshed) should be examined. Relative speckle intensities from Equations (11) and (12) might be used to fine-tune image processing strategies depending on which type of speckle is dominant for given conditions, particular $S$ and $D/a$. These effects become prominent at somewhat higher correction than is now routinely available on the ground, but may be important to space-borne instruments such as HST.

ACKNOWLEDGMENTS

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

REFERENCES