Closed-Loop Adaptive Suppression of Residual Coronagraph Halos using a Focal Plane Interferometer and Anti-Halo Apodization

Johanan L. Codona

***Abstract*—High-contrast imaging systems for studying extra-solar planets require contrasts of 8 – 10 decades at angles of a few λ/D, requiring unprecedented precision and control. Of the methods proposed to accomplish this, many involve some form of coronagraph, where a focal plane mask is used to block the bright star, simulating an artificial eclipse. However, wavefront imperfections of as little as 8 pm, or transmission variations of as little as 0.01% cause planet-like speckles to appear at 10^{-8} of the star, causing confusion.**

We have developed an interferometric focal-plane wavefront sensor using the starlight normally discarded by the coronagraphic focal plane mask to form an interferometric reference beam. The interferometer permits us to estimate the phase and amplitude of the residual halo relative to the bright center of the star's diffraction pattern. The measurement is efficient and has the significant benefit of being able to overcome the effects of incoherent background noise such as zodiacal light, even when the halo is much fainter. We have demonstrated its use in halo suppression by driving a pupil deformable mirror (DM) in closed loop.

We have also developed a new halo suppression method, Anti-Halo Apodization (AHA), which uses the coronagraph's discarded starlight as the raw material to construct a coherent negative copy of the halo directly in the focal plane. A major advantage of the AHA approach is that it uses an attenuating filter to set the overall level of the antihalo, rather than the ultra-precise control needed with a pupil DM. This allows us to achieve a great deal of additional halo suppression with relatively low-precision phase modulators: two additional decades of suppression can be achieved with 0.1 radian of phase control. The AHA benefits can be used to ease the tolerances throughout the system.

***Index Terms*—High-Contrast Imaging, Coronagraphs, Interferometers, Extreme Adaptive Optics.**

I. Introduction

Direct imaging of extrasolar planets, especially non-self-luminous terrestrial extrasolar planets, is an incredibly daunting challenge. Depending on the observed wavelength and the size and characteristics of the terrestrial exoplanet, the expected contrast is of the order of ten decades. Giant exoplanets and young, self-luminous exoplanets will be detectable with lower contrast, but finding and studying an analog of the Earth will involve contrasts of 10 decades. Detecting such an exoplanet in the habitable zone of a star makes the problem even more challenging, by restricting the search to an angular region very near the star. A consequence of geometry is that the habitable zones of stars with the same apparent magnitude appear the same size in the sky. Therefore, the habitable zone of a fifth magnitude star (e.g. the Sun at 10 pc) is about 100 mas radius. For specificity, we consider a 4 m space telescope operating in the visible with a diffraction-limited resolution of about 30 mas, so we are talking about meeting the optical challenge of 10 decades contrast at 3–4 λ/D. The TPF-C strategy is to use a very high-performance space-based coronagraph with a deformable mirror (DM) to correct residual wavefront errors and suppress any remaining diffraction and scattered light(Malbet et al., 1995). This has now been demonstrated at the JPL High-Contrast Imaging Testbed (HCIT) to levels that are appropriate to terrestrial exoplanet detection (Trauger and Traub, 2007).

In this paper we describe two extensions of the base coronagraph design, now demonstrated at the University of Arizona, that can supplement a high-contrast coronagraphic imaging system. Both modifications use the starlight that is normally blocked and discarded by the coronagraphic focal plane mask to measure and suppress the residual coronagraphic halo. The modifications can be used to bring a less capable coronagraph to higher and more robust levels of performance, or reduce performance risk in a high-performance design. The proposed coronagraph modifications can be introduced in such a way as to preserve the unmodified performance, while providing both greater robustness and more reliable halo suppression performance.

II. Coronagraphs and residual halos

A stellar coronagraph is an optical telescope designed to create an artificial eclipse by blocking the bright starlight in the focal plane with an occulting mask. Especially when this mask is only a few diffraction resolutions of the telescope (λ/D_{pupil}), diffraction around the mask is significant, leaving a bright residual halo. For a simple mask this residual halo light is concentrated near the edges of the reimaged pupil, but is present everywhere in the pupil. By blocking the normally dark regions of the reimaged pupil and trimming the regions of the pupil where the halo is brightest, the residual starlight halo in a second focal plane is greatly reduced. This reimaged pupil "Lyot stop" can be made more aggressive by blocking more of the halo region, but at the cost of angular resolution. A clever refinement of the Lyot coronagraph is to include a spatial modulation on the focal plane mask, limiting where in the reimaged pupil plane the diffracted halo goes. This makes it possible to get very great halo suppression with a less aggressive Lyot stop (Kuchner and Traub, 2002). Depending on the size and design of the focal plane mask and the the corresponding design of the Lyot stop, it is possible to attenuate the halo at a few λ/D to 6 or more decades below the un-

blocked diffraction-limited core of the star image.

Novel alternatives to conventional coronagraphs have also been proposed and may have very different architectures. A leading concept is the Phase-Induced Amplitude Apodization (PIAA) system(Guyon, 2003; Guyon et al., 2006). Since PIAA involves a mask blocking the bright starlight, the techniques described here are still completely applicable in spite of the significant differences in the main halo suppression architecture.

Even a perfect coronagraph will leave a halo. This problem becomes even more serious as we try to image nearer the star, often requiring a smaller focal plane mask. However, incomplete diffraction suppression is not expected to be the limit to detection, rather optical aberrations, reflection and transmission variations across the pupil, and mechanical imperfections. A sinusoidal ripple in the optical wavefront with an amplitude z, acting like a weak diffraction grating, will create a diffraction speckle with a contrast $(2\pi z/\lambda)^2$ below the star. An optical element with a surface accuracy of even $\lambda/1000$ will leave focal plane speckles that are typically brighter than 5 decades below the star. Imperfections in the transmission and reflection of the optics have a similar effect, with a 1% transmission variation causing speckles of about the same order. Even with highly precise manufacturing, such imperfections are inevitable, and unwanted speckles will be the norm at levels of at least 5 or 6 decades below the star. This when the target contrast is 10 decades.

Fortunately, these unwanted halo speckles can be suppressed using a deformable mirror (DM) in an upstream pupil plane (Malbet et al., 1995). Since the halo–speckles included–is temporally coherent with itself across the focal plane, we can use the DM to induce an "anti-speckle" with the same amplitude but opposite phase to any given speckle within the DM control radius. Since the speckle and the created antispeckle are coherent, they cancel, leaving a dark spot in the halo. By simultaneously steering many such anti-speckles, an extended search region in the halo can be suppressed, including not only the unwanted random speckles, but the residual coronagraph diffraction. A general limitation arises from the fact that DM phase patterns result in focal plane waves which are antisymmetric about the star (the same symmetry as the aberration speckles), while the transmission flaws and general diffraction residual leave symmetric halo speckles waves. In narrowband light this means that while phase speckles can be suppressed on both sides of the star, suppressing transmission speckles and residual diffraction on one side of the star coherently reinforces the halo on the other side, making it brighter. In general, there will be a combination of both speckle symmetries, and the halo can only be deeply suppressed on one side of the star at a time. This has been dramatically demonstrated in the JPL High Contrast Imaging Testbed (HCIT), with laser light Trauger and Traub (2007). The DM phase adjustments are wavelength-dependent, and the correction is chromatic, making this approach less effective over large bandwidths such as the 10–20% bands that would be desired for a TPF instrument. For the purposes of this paper, we will limit consideration to a 2% band, which is small enough to get acceptable suppression performance, and large enough to imagine channelized replication to achieve an instrument with the full desired bandwidth (e.g. Guyon et al. (2006)).

To achieve generalized halo suppression to a given level, the DM must be able to generate and maintain a phase ripple across the pupil with an rms amplitude accuracy of $\sigma_z \leq \sqrt{C/2}\lambda/4\pi$, where C is the desired contrast. For visible light ($\lambda \approx 600$nm) and a desired contrast of 10 decades, the DM surface ripple must have an accuracy better than 0.47 pm. This is the combined effort of all of the illuminated actuators in the DM, which for a 1024-actuator DM are of order 30^2 actuators. The accuracy requirement on a single actuator is therefore 30×larger, or about 10 pm rms.

Estimating the halo to be suppressed

The pupil DM can steer an antispeckle to any point within the control radius of the DM ($n\lambda/2D$, where n is the number of actuators across the pupil). Furthermore, the antispeckle can be given any phase relative to the PSF core by changing the phase of the DM ripple, and its amplitude is proportional to the ripple amplitude. We can therefore expect to be able to cancel any given point or region (subject to the symmetry and control radius restrictions) in the halo, so long as we have: (1) adequate knowledge of the phase and amplitude of the halo to be suppressed, and (2) the ability to adequately control the DM ripple's phase and amplitude and thus the phase and amplitude of the corresponding anti-speckle.

Consider first the problem of estimating the phase and amplitude (or the *complex amplitude* $ae^{i\varphi}$) of a halo speckle. The JPL speckle-nulling technique is too complicated to be properly summarized here (Bordé and Traub, 2006), but can be roughly analyzed with the intent of understanding its performance level and implications for systems engineering. The approach is to apply a superposition of small ripples to the DM creating a speckle at a set of points in the focal plane where the halo is to be estimated. By measuring the resulting *halo+speckle* image with at least three ripple-speckle phases allows an estimate of the halo complex amplitude to be computed. The greatest contrast for this measurement is achieved when the speckle has the same amplitude as the underlying halo. Ignoring detailed numerical factors for this discussion, the complex halo estimate accuracy becomes sufficient to make a servo correction (e.g. halo phase known to a radian and amplitude to within a factor of 2) when the ripple phase images have intensity sigmas of the same order as the halo intensity. This means that according to this rough estimate, we would have to observe at least 2 photons per speckle per phase image in order to measure the halo to an adequate level. In other words, to measure and suppress a point in the halo to the 10^{-10} level, the phase-shifted ripple intensity images must have an rms measurement accuracy to that level. Using a laser in the lab meets this requirement with relative ease. Making the same measurement

with a real star in a 2% band is much more challenging.

Using our model case of an sun-like star at 10 pc, its apparent magnitude will be $\sim$ 5 mag, or a photon flux of $\sim 10^7$ photons/s/m^{2}in a 2% band. Using a detection throughput of 1/3, a 4 m diameter collection area of 12.5 m^2, and a reasonable 75% encircled energy in the diffraction-limited core, our 5th magnitude star should yield $\sim 3 \times 10^7$detected photons/s in the PSF core. We should therefore expect to detect one photon every five minutes from a speckle suppressed to the desired contrast level of 10 decades fainter than the PSF core. Assuming two photons per speckle and a minimum of 3 phase images, we would estimate being able to make a servo update after 30 minutes. This would be enough to keep the halo suppressed to an interesting level. However, the presence of background noise makes this estimate much worse.

This calculation does not include background noise, most notably from zodiacal and exo-zodiacal light, which can be much brighter than our target planet. Assuming that the halo remains stable long enough to make a meaningful measurement, we need to integrate with each ripple phase until the sigma of the photon noise is comparable with the speckle intensity. That is, $\sigma_{background} = \sqrt{N_{background}T} \approx N_{speckle}T$, or $T \gtrsim N_{zodi}/N^2_{speckle}$, where T is the integration time and $N_{speckle}$ and $N_{background}$ are the speckle and background photon fluxes measured over a speckle-sized area of the focal plane. For us to be able to measure the minimum of three phases in 90 minutes, the background-to-speckle flux ratio has to be less than the total number of speckle photons detected–about 18 for our assumptions. This corresponds with a 1 zodi background (22 mag/arcsec^2 at 550 nm and our assumed diffraction scale and throughput). When the background is much brighter than this minimum level, an adequate measurement of the speckle's complex amplitude will be proportionately longer, resulting in a diminished ability to suppress the halo. Also, since the background is independent of the star's apparent magnitude, background noise will inevitably be a limiting factor in some cases. As we shall see in the next section, this problem can be completely side-stepped by using a modified coronagraph to measure the halo.

III. Coronagraphic Focal Plane Interferometer

The conventional design for a Lyot coronagraph simply discards the starlight blocked by the focal plane mask. Because this starlight is coherent with the halo, it is a valuable resource in that it can be used to both measure and cancel the halo. Roger Angel (Angel, 2002) proposed using a modified coronagraph where the starlight blocked by the focal plane mask is diverted and spread out to form an interferometric reference beam. The halo path is processed as usual, including the reimaged pupil plane with the Lyot stop, but is folded to be coherently mixed with the reference beam using a beamsplitter (figure 1). The resulting configuration is a form of Mach-Zehnder interferometer allowing the measurement of the halo phase and amplitude relative to the PSF core. It is possible to derive the complex amplitude of the halo by taking at least three images with different phase shifts (path length offsets) introduced by moving the reference beam fold mirror. If we capture both the reflected and transmitted outputs from the beamsplitter, we can measure the ϕ and $\phi + \pi$ phases simultaneously, with no loss of halo light. By steering ripple-speckles with known phases using the pupil DM, we can fully calibrate the interferometer, including all relevant parameters in the optical system, such as plate scale, DM alignment relative to the optical axis, placement of the pupil image on the DM, and the difference between the reference beam wavefront and the halo in the focal plane. As was the case with the ripple phase-shifting method, a halo suppression servo will converge so long as the measured phase sigma is less than approximately 1 radian. Therefore, even though the calibration can be performed quite accurately, the actual requirements on the calibration are rather loose.

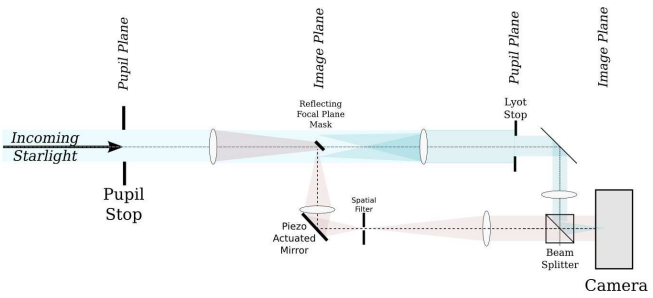

Figure 1. The coronagraph folded to mix the halo with a reference beam formed from the bright starlight diverted from the diffraction-limited core of the star's image. The reference beam is phase-shifted relative to the halo by moving the reference beam fold mirror.

Since we calibrate the interferometer using DM ripple speckles, we can easily determine the phase and amplitude of a ripple to suppress a given speckle directly from the interferometer measurement. Just as in the case of halo measurements using phase-shifted ripples, an update to the DM settings will diminish the halo intensity as long as the rms phase measurement sigma is less than ~ 1 radian and the amplitude is known to within a factor of two. If we take four images with the interferometer, phase shifted at 90^0 intervals, yielding four images I_0, I_{90}, I_{180}, an I_{270}, we can compute the complex interference term between the halo and reference waves as

$$4a_{ref}a_{halo}e^{i(\phi_{halo}-\phi_{ref})} \propto I_0 - I_{180} + (I_{270} - I_{90})i, \quad (1)$$

where $a_{halo}e^{i\phi_{halo}}$ is the complex halo field, $a_{ref}e^{i\phi_{ref}}$is the complex reference field, and the proportionality constant and the reference beam phase are determined from the ripple-induced speckle calibration. The measurement accuracy is determined by the four images' photon noise. The normal use of an interferometer suggests that we maximize fringe visibility by matching the reference beam and halo intensities. However, an examination of equation (1) shows that while the fringe contrast becomes lower as we increase the reference beam intensity above the halo, we have more photons with which to make the measurement in the same

amount of time. Since $I \propto a^2$ and $I \propto \sigma_I^2$, the accuracy of the complex halo measurement becomes independent of the reference beam intensity(Angel, 2002). The resulting complex halo measurement becomes sufficiently accurate to make a DM update when a few ($\gtrsim 3$) halo photons have been detected per speckle. Of course, we will never actually know which photons were halo photons since they must be indistinguishable from the overwhelming number of reference beam photons. For our reference case of a 5th magnitude star, with a 10-decade down speckle photon being detected at a rate of once in 5 minutes, the focal plane interferometer can make an adequate measurement of the speckle phase in 15 minutes. Depending on the strategy the time between halo wavefront sensing (reference beam on) and science image integration (reference beam off), we should still be well within the target timescale of 90 minutes to make a servo update.

The real benefit here is that the halo measurement can be made insensitive to zodiacal and other background noise so long as the photon noise from the reference beam dominates. That is, if $I_{ref} \gg I_{zodi}$, the sigma in the halo measurement will be independent of the background. To estimate the reference beam intensity, assume that we take the PSF core with size λ/D and spread it out to $10\lambda/D$. The mean surface brightness will drop by a factor of 100, allowing a reference beam that is as much as 8 decades as bright as the speckles being measured. This is, of course, far too bright to be practical, but it shows that the photons needed to overwhelm the background are available. A more reasonable limit to how bright a useful reference beam would be is 2 or 3 decades brighter than the halo. This is 5 or 6 decades fainter than the brightest reference beam, meaning that there is plenty of core starlight left over for other purposes. (We make use of this extra core starlight in the active halo suppression method described in the next section.) If we consider the difficult case where the background is 100× the speckle intensity, we would have to integrate long enough to detect more than 100 halo photons per speckle using the phased-speckle method, while the same measurement accuracy can be achieved in $\sim 1/30$ the time using the interferometer. With even brighter backgrounds, the advantage of the coronagraphic interferometer becomes even greater.

As part of our NASA-funded research, we have prototyped the focal plane interferometer and successfully used it in a closed-loop to drive a Boston Micromachines MEMS DM in the pupil, suppressing the halo to more than 5 decades below the PSF core. While this result is not in the same league as the JPL HCIT work, it is close to the best that can be achieved using a MEMS DM in this configuration. Our work points the way to a relatively simple coronagraph modification that can take full advantage of higher quality components, with a new robustness to incoherent background noise.

IV. Anti-Halo Apodization (AHA)

AHA is an active halo suppression technique that uses a portion of the same diverted coronagraph starlight used in the focal plane interferometer (Codona and Angel, 2004). The core starlight is spread out over the focal plane region that is to be suppressed and is explicitly imprinted with a negative copy of the measured complex field using a fully-complex spatial light modulator. The constructed antihalo is then mixed with the coronagraph halo using an asymmetric beamsplitter, keeping as much of the coronagraph halo and exoplanet light as possible. When the antihalo is properly modulated, it coherently cancels the halo, leaving the incoherent light from the exoplanet unaffected. Unlike the pupil DM method of creating antispeckles, the AHA method creates its field in a separate optical path, starting with a highly-attenuated piece of the core starlight, meaning that the modulator has fewer decades to drop the antihalo to match the halo. The incoming starlight used to create the antihalo is controlled by a simple ND filter rather than by the incredibly fine control of a pupil DM. We have

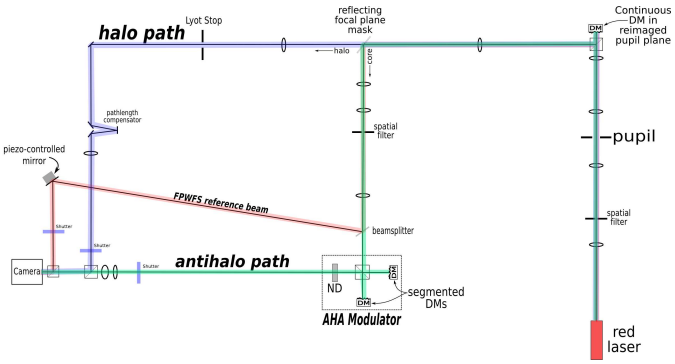

Figure 2. AHA lab diagram. This shows both the focal plane interferometer and the AHA antihalo subsystems.

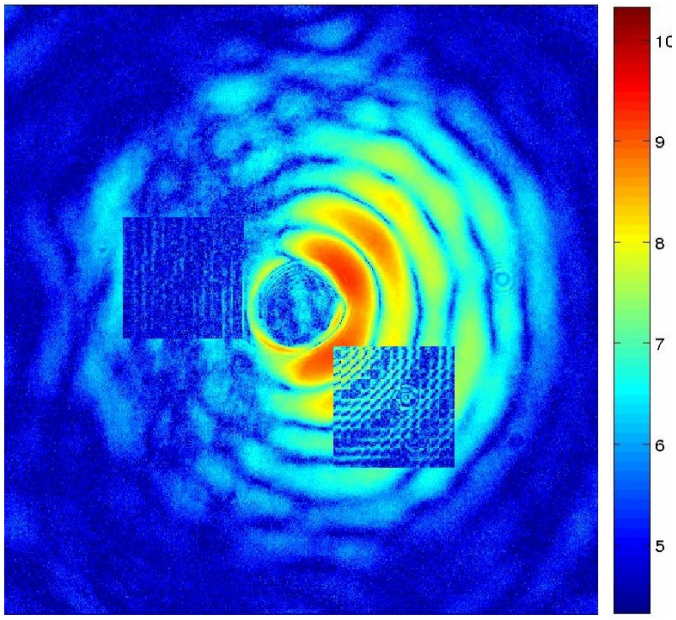

Figure 4. A small 12x12 AHA modulator would not cover much of the focal plane, but could be placed wherever needed. The halo+antihalo in the AHA window would be interferometrically measured and adjusted to suppress the combination to more than 2 decades below the halo alone.

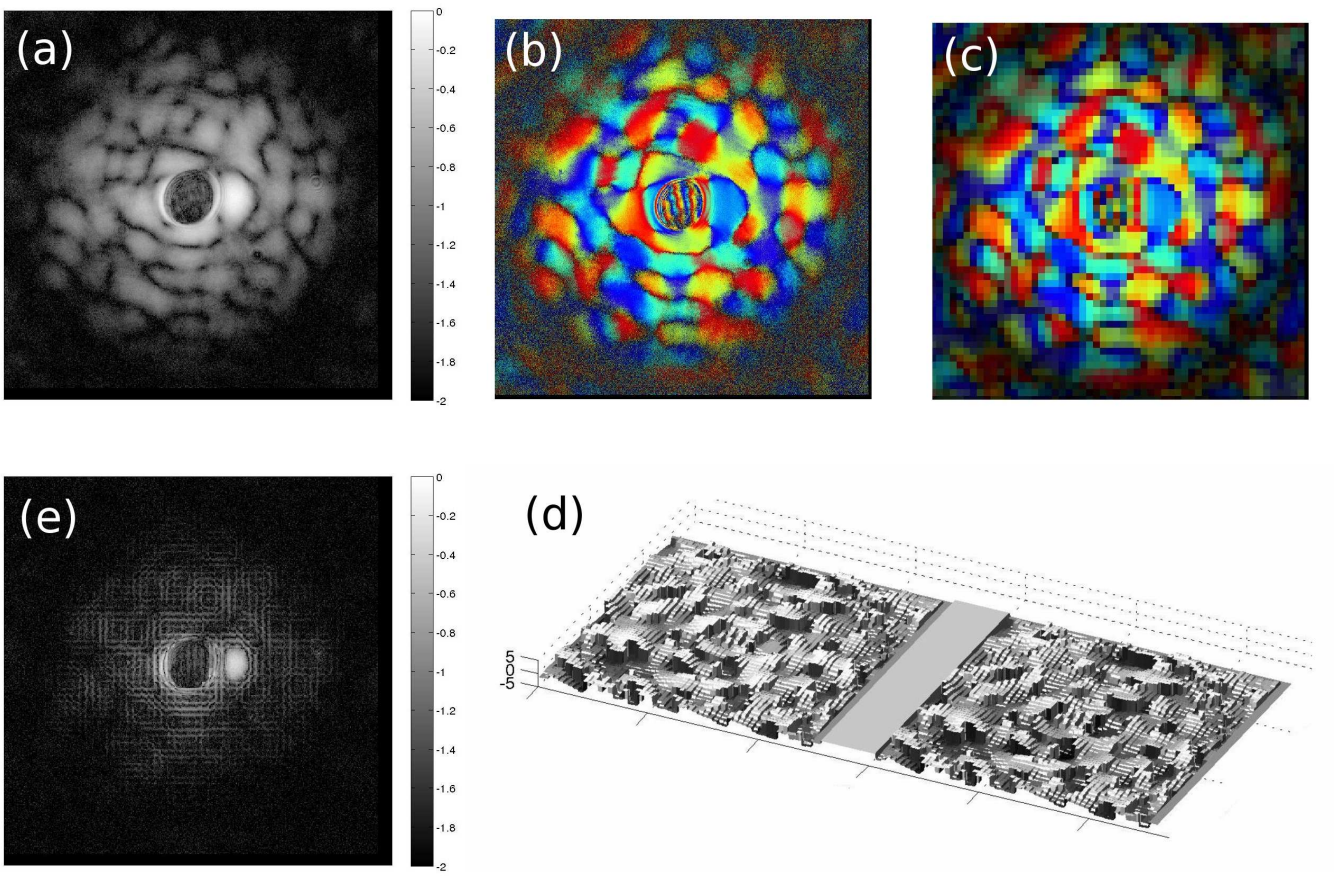

Figure 3. Conceptual flow of the AHA method. (a) A measured intensity image of a speckled halo (high-gamma stretch). (b) The complex halo as measured with the focal plane interferometer (phase is shown as color). (c) Pixelized version implemented by actuator displacements in the AHA modulator. (d) The computed displacements for the two DMs in the AHA Michelson to create (c). (e) The intensity resulting from the difference between (b) and (c). Note the bright speckle at the 3:00 position that was too bright to be reached by the addressable complex values with the AHA modulator illumination. Increasing the modulator illumination would enable this speckle to also be suppressed.

implemented an AHA system in the lab, and demonstrated its ability to suppress the coronagraphic halo. Our AHA modulator is constructed using two MEMS DMs. We have recently devised a second method for creating the antihalo that uses only a single DM, but we have yet to implement it in the lab.

The AHA modulator is a complex spatial light modulator (SLM) built using a Michelson interferometer with a DM on each arm. The core starlight is spatially filtered and spread out in a relayed focal plane over what will become the suppressed region of the halo. For optimal contrast and chromatic performance, the two arms of the modulator are adjusted to be as close as possible to each other, and are, in turn, matched to the length of the coronagraph halo path. The DMs were adjusted such that the images of each actuator overlaps its counterpart to within a small fraction of an actuator. By moving the DM actuators, the path length along each arm is altered, causing the phase shift along each arm to become $2\pi(2z_{arm})/\lambda$, where z_{arm} is the path length to an actuator along one of the arms. After remixing the two light paths in a beam splitter cube, the resulting complex amplitude applied to the featureless field derived from the diffraction core is proportional to $\exp(4\pi i z_1/\lambda) - \exp(4\pi i z_2/\lambda)$. This gives the antihalo any complex value within an amplitude controlled by the ND filter. The created antihalo has only one settable value for each DM actuator pair, making it necessary to have at least 3 or 4 actuators (i.e. "AHA pixels") per λ/D. Since the AHA modulator can be placed anywhere in the focal plane, a pair of 1024-element DMs can be used to create an $8 \times 8$to $10 \times 10\lambda/D$ dark window in the halo wherever desired in the search region.

The operational concept is illustrated in figure 3. The halo used in this figure was actually measured using our laboratory focal plane interferometer, but the AHA portion was simulated. In figure (3), we have the speckled halo of a star imaged through a weak phase screen. The Strehl ratio is still high enough to retain a substantial diffraction-limited core, which was reflected for use as the reference beam in the FPI, and available for use in creating the AHA antihalo. There is a bright speckle at the 3-o'clock position that we will intentionally miss with the AHA modulator illumination, just to show the effect. Figure (3b) shows the complex halo as measured with the FPI, and (3c) shows the pixelized version of the halo that can be created using the AHA modulator for some illumination strength.

This is accomplished by adjusting the two DM surfaces to appropriate values, shown in figure (3d) for a hypothetical huge DM. The two surfaces are very similar when the desired antihalo amplitude is much fainter than the light illuminating the modulator, and differ more where the antihalo is brightest. Careful examination of the two surfaces near the positions of the brightest speckles shows this effect. Finally, figure (3e) shows the result of the complex sum of the halo and antihalo, which takes place in the final beam splitter. Since the light directly entering the camera is halo-antihalo, while the light exiting the beam splitter in the other direction is antihalo-halo, capturing both outputs and summing the detected photons means that no light need be lost in the mixing stage.

V. Discussion

Both the coronagraphic focal plane interferometer and the AHA halo suppression system make use of normally discarded starlight to improve performance and reduce risk. The focal plane interferometer allows us to overwhelm background noise with a bright reference beam and measure the complex amplitude of the halo in a time dependent only on the halo photon flux. The AHA halo suppression system can be used to supplement the coronagraph halo suppression by an additional 2 or 3 decades with no negative impact on the science light path or the ability to operate the system without AHA. The 2–3 decades of AHA halo suppression can be used to improve the performance of a less-capable coronagraph, or reduce the risk of a more capable coronagraph. Since both techniques use discarded starlight and relatively low-cost optical components, and are compatible with the leading coronagraph design concepts, it makes sense to incorporate them into future system designs.

Acknowledgment

This work was supported by NASA grant NNG05GD28G.

References

Angel, R. (2002). Imaging Exoplanets from the ground. In S. Seager and D. Deming, editors, *Scientific Frontiers in Research on Extrasolar Planets*, Washington D.C. ASP, ASP Conference Series.

Bordé, P. J. and Traub, W. A. (2006). High-Contrast Imaging from Space: Speckle Nulling in a Low-Aberration Regime. *ApJ*, 638:488–498.

Codona, J. L. and Angel, R. (2004). Imaging Extrasolar Planets by Stellar Halo Suppression in Separately Corrected Color Bands. *ApJ*, 604:L117–L120.

Guyon, O. (2003). Phase-induced amplitude apodization of telescope pupils for extrasolar terrestrial planet imaging. *A&A*, 404:379–387.

Guyon, O., Angel, J. R., Bowers, C., Burge, J., Burrows, A., Codona, J., Greene, T., Iye, M., Kasting, J., Martin, H., McCarthy, D. W., Meadows, V., Meyer, M., Pluzhnik, E. A., Sleep, N., Spears, T., Tamura, M., Tenerelli, D., Vanderbei, R., Woodgate, B., Woodruff, R. A., and Woolf, N. J. (2006). Telescope to Observe Planetary Systems (TOPS): A High Efficiency Coronagraphic 1.2-m Visible Telescope. In *American Astronomical Society Meeting Abstracts*, volume 209.

Kuchner, M. J. and Traub, W. A. (2002). A Coronagraph with a Band-limited Mask for Finding Terrestrial Planets. *ApJ*, 570:900–908.

Malbet, F., Yu, J. W., and Shao, M. (1995). High-Dynamic-Range Imaging Using a Deformable Mirror for Space Coronography. *PASP*, 107:386–+.

Trauger, J. T. and Traub, W. A. (2007). A laboratory demonstration of the capability to image an Earth-like extrasolar planet. *Nature*, 446:771–773.

Johanan L. Codona, Ph.D. is a Senior Research Scientist with the University of Arizona's Center for Astronomical Adaptive Optics in Tucson, AZ.