Advanced Speckle Sensing for Internal Coronagraphs
Charley Noecker, a Stuart B. Shaklan, b James K. Wallace, b Brian D. Kern, b Amir Give’on, b Jeremy Kasdin, c Ruslan Belikov d, Steve Kendrick a

aBall Aerospace & Technologies Corp., 1600 Commerce St., Boulder, CO, 80301, USA
bJet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, 91109 USA
cPrinceton University, Princeton, NJ 08544, USA
dNASA Ames Research Center, Moffett Field, CA 94035, USA

ABSTRACT
A 4-8m telescope carrying a coronagraph instrument is a leading candidate for an anticipated flagship mission to detect and characterize Earth-size exoplanets in the 2020s. 1 Many candidate coronagraph instruments have been proposed, and one has met many of the principal requirements for that mission. But the telescope and instrument will need exquisite stability and precise control of the incoming wavefront to enable detection of faint companions (10^-10 of the star) at an angular separation of 2-4 Airy radii. In particular, wavefront errors cause speckles in the image, and variations in those speckles can confound the exoplanet detection. This challenge is compounded by the background light from zodiacal dust around our Sun and the target star, which limits the speed with which we can estimate and correct the speckles. We are working on developing coherent speckle detection techniques that will allow rapid calibration of speckles on the science detector, allowing subtraction in post-processing or correction with deformable mirrors. The expected speed improvement allows a much quicker timeline for measurement & calibration, which reduces the required telescope stability requirement and eases both the flight system design and the challenge of ground testing. We will describe the experiments and summarize progress to date.

Keywords: Terrestrial Planet Finder, coronagraph, wavefront sensing, coherent detection

1 INTRODUCTION
One of the most compelling and exciting human endeavors is the study of our origins, from the Big Bang to galaxy formation to star formation to planetary system formation. But the most inspiring of these is the search for other planets and planetary systems like our own, especially those that harbor some form of life. This goal is tantalizingly close to our grasp, and could be achieved within our lifetime. Instruments have been designed and to some extent tested, 1,2,3,4,5,6 that are capable of direct detection—blocking the light from a nearby star and revealing the faint light of any planets orbiting around it. In this directly detected exoplanet light, we can observe atmospheric absorption lines, including biomarkers that provide evidence of life on that exoplanet.

The Kepler mission 7 also searches for exoplanets, by measuring the small decrease in the star’s brightness when each exoplanet passes in front of the star. Kepler has already found 1235 planet candidates, 8 most of which are still unconfirmed, and many more are expected. These exoplanet candidates range in size from 0.6-60 Earth radii (R_⊕) and their orbits are clustered at very small radii, with a modest number further out near the habitable zone (HZ). The habitable zone for any given star is the range of orbital semi-major axis at which liquid water can exist on the surface of an Earth-size rocky planet; typically this is agreed to be about 0.7-1.8 AU for a Sun-like star, 9 and it scales roughly as the square root of the stellar luminosity. Kepler has been very successful, and has produced the first large dataset of Earth-size planets; but its measurement technique and the brevity of the initial data record strongly both emphasize planets orbiting near their stars. As the mission data record lengthens, this bias will persist, but not so strongly.

But the Kepler planets are difficult targets for follow-up study, because they are typically very far away (few kiloparsec). Transit spectroscopy of Earth-size planets around these relatively faint stars is not practical, and direct detection and spectroscopy is essentially impossible because the planets are faint and at milliarcsec separation from their host stars. This is why the most fruitful direction for future exoplanet studies is direct detection of exoplanets around nearby stars.

* Email: charley.n@att.net, Phone: 303-475-5386
A stellar coronagraph in space is one way of doing that. It uses an ordinary-looking telescope with a special instrument that blocks the star and reveals faint planets at a small angle from it. The coronagraph instrument includes a mask at the image of the star, often followed by a mask later in the optical system. The image that results contains the exoplanets and a small amount of residual starlight, including starlight speckles arising from amplitude and phase errors on the optics. These speckles are a stippled pattern of starlight in the image plane, caused by diffraction from telescope optical errors like wavefront and amplitude variations. A high-performance wavefront sensing and control system incorporating a deformable mirror (DM) can reduce the speckles to be as faint as the planet within a modest bandwidth. In this way, the DM generates an optical electric field at the detector which cancels the residual fields from the telescope.

It is still a challenge to distinguish planet light from stellar speckles, because speckles are very similar to planets in angular size, brightness, and spectrum. After adjusting the DM to remove speckles from the image as best we can, we must calibrate and subtract the remaining speckles, to reveal any planets. A common method for calibrating speckles is roll subtraction: basically subtracting images taken at two different telescope roll angles around the line of sight axis. Because speckles are caused in the telescope, they stay fixed on the detector during a roll, while exoplanets rotate as the telescope rolls. But if the speckles change between these two images, those differences can mimic the signal of a planet; speckle stability sets an uncertainty floor on any negative or positive result for exoplanet detection. To control this uncertainty contribution, the speckle change must be significantly less than the desired detectable exoplanet signal. The stability timescale is approximately the speckle calibration or exoplanet measurement time. The integration time to detect an Earth-like exoplanet or the surrounding speckles in a zodiacal light background is several hours because of low photon rates. During this time, we rely on the stability of the optical system to ensure speckle stability well below the exoplanet signal. Thus, so far we find that our “ordinary-looking telescope” must be much more stable than any space telescope that’s ever been built. A typical wavefront stability tolerance is <0.5Å RMS spatially, for many hours.

The present study is aimed at reducing this stability time scale, by calibrating speckles instead via their optical coherence with the star. Building on the prior work of others, we will develop coherent speckle detection (CSD) methods for use in an internal coronagraph. By adding a known reference beam of starlight, we can measure the interference of the unknown speckle fields with the reference fields, and from that estimate the brightness of the speckles. By tuning the reference beam to be modestly brighter than the local zodi and exozodi light, we can increase the brightness of the speckle-related heterodyne signals and thus improve the SNR of the measurement in a given time. This allows a quicker estimate of the speckle brightness and optical phase, which can be used either as a calibration (for subtraction in post-processing of science images) or to produce a command for the DM (to reduce the speckle brightness). For this study we will concentrate on accurate measurement of speckles, for quick calibrations between science integrations, and not on control. Thus the mean speckle brightness is not a driving concern, but rather the uncertainty in the CSD-derived speckle measurements.

Our SNR model shows that coherent methods offer the possibility of quick measurements, limited in the ideal case mainly by shot noise in the speckle brightness, and almost independent of zodi and other backgrounds. For typical numbers, we see an integration time of order 1 ksec with CSD for the same uncertainty that takes 30-150 ksec with direct speckle measurement. Quick measurement allows quick alternation between calibration and science observations, which allows shorter time scales for “passive” wavefront stability; this enables relaxed optical stability requirements.

2 SPECKLE STABILITY AND CALIBRATION

The main challenge of speckles in the science measurement is that they mimic a planet signal. Speckles are caused by optical imperfections—wavefront phase and amplitude ripples that diffract and allow starlight to leak past the field stop into the final science image. By adjusting the DM we can reduce speckles, but technical limitations of speed, accuracy, and optical bandwidth set a limit on how faint they can be made. The residual speckles are a type of background signal in the science image, which must be estimated and subtracted in order to reveal a planet. But the appearance and spectrum of a speckle is generally similar to the reflected light of a planet; its typical size.

Coronagraphs suppress the speckles to a level roughly 10^{10} times fainter than the unblocked star image. This aggressive goal is not meant to allow the exoplanet to stand out by eye above the speckles (and it wouldn’t); it is to make it easier to control speckle variation. An Earth-like exoplanet’s brightness is about 10^{10} of the star’s, and it’s usually covered by a larger astronomical background flux from zodiacal dust in our local Solar System and around the exoplanet’s host star. So if the suppressed speckles match this exozodi, they are being driven well below the level that appreciably affects the
The speckles are a greater threat to the measurement than the zodiacal backgrounds, because they can vary in time and spatial distribution, whereas the zodiacal backgrounds are very uniform and stable in time. If \( E \) is the static speckle field and \( \delta E \) is its time variation, then the leading term in the intensity variation is \( 2E \cdot \delta E \). To keep this within about 1/5 of the exoplanet signal, \( 2E \cdot \delta E < 2 \times 10^{-11} \), then with \( E = 10^{-5} \) we need \( \delta E < 10^{-6} \) i.e. \( \delta E / E < 10\% \). But if we relax \( E \) to \( 3 \times 10^{-5} \) we need to tighten \( \delta E < 3.3 \times 10^{-7} \), requiring \( \delta E / E < 1.1\% \). Thus if we relax the accuracy of the DM control, we’ll pay a price in tighter stability; instead we’d rather make an engineering choice to keep the mean speckle brightness comparable to the exoplanet signal, so that a 10% variation is tolerable.

The challenge in this case is not so much speckle brightness as the uncertainty in speckle brightness. Thus the goal of wavefront sensing and control is managing this uncertainty. Frequent measurements and modeling are tools in that effort, i.e. estimation of the speckle intensity from various kinds of measurements over time (Figure 1). One possibility is measurements of thermal and other disturbances, combined with integrated modeling of the effects of those disturbances on the optics and supports, yielding a prediction of the changes in wavefront in the pupil and speckles in the focal plane. Another possibility is direct measurements of the speckle intensity, and estimation of the time evolution by extrapolation or interpolation in time. A third possibility is the topic of this study: measurements of speckle fields via CSD to predict the speckle intensity.

### 2.1 Speckle sensing methods and integration time

The earliest method of measuring and correcting speckles was speckle nulling, an iterative process in which the intensity pattern in the image plane is repeatedly measured while the DM is exercised to adjust the field to minimize the brightest speckles in the image.

\[
\text{Signal} = I_{LZ} + I_{EZ} + I_S \quad I_S \ll I_{LZ} + I_{EZ}
\]

where \( I_S \) is the speckle photoelectron rate, \( I_{LZ} \) is the local zodi rate, and \( I_{EZ} \) is the exozodi rate. The shot-noise limit for integration time for this “direct speckle measurement” method is

\[
\tau_0 = \text{SNR}^2 \frac{2(I_{LZ} + I_{EZ})}{I_S^2}
\]

Between corrections by this iterative process, there might be occasional measurements to feed the time-evolution model. Since the speckles are significantly fainter than the zodiacal background, it takes a long time to measure them directly for each iteration. A typical integration time for an Earth-like exoplanet or an equally bright speckle is several hours. This in turn requires very long timescales for the stability of the speckles. The extreme requirements for telescope stability led to elaborate thermo-mechanical telescope designs and heroic efforts in performance modeling to demonstrate feasibility.

Coherent techniques employ a substantial stellar “reference beam” field added at the location of the speckle in the image plane. This can be produced by a number of different ways; we will discuss two options. The speckles interfere with the reference beam in the science focal plane, yielding large interference cross-terms, which can rise above the background:

\[
\text{Signal} = I_{LZ} + I_{EZ} + I_S + I_R + 2\sqrt{I_SI_R \cos(\phi_S - \phi_R)} \quad I_R > (I_{LZ} + I_{EZ})
\]
Several measurements with differing $I_R$ and $\phi_R$ allow isolation of the last term, the interference cross-term, and thereby an identification of $I_S$ and $\phi$. This information can be used to calculate the speckle intensity as a calibration, or to develop an estimated correction to apply to the DM.

The shot-noise limit for integration time drops with increasing $E_R$:

$$\tau = \frac{SNR^2}{I_S \left(1 + \frac{I_{LZ} + I_{EZ}}{I_R} \right)}$$

(4)

In the limit of very large $I_R$, this integration time decreases to the shot-noise limit for zero background. (There is a maximum practical value of $I_R$ arising from the uncertainty in $I_R$ itself due to technical issues, as discussed below.) The ratio of this integration time to the direct-speckle-detection integration time in Eq. 2 is

$$\frac{\tau_R}{\tau_0} = \frac{1}{4 \left(1 + \frac{I_S}{I_{LZ} + I_{EZ} + I_R} \right)}$$

(5)

We see that this ratio decreases with increasing zodiacal light background or reference beam fluxes. Thus when the integration time is background-limited, CSD affords an improvement when $I_R > I_S$.

2.2 Speckle measurements with CSD

We have chosen two methods for generating a reference beam. One is to impose a pattern on the DM to create a deliberate stellar leakage field; we call this phase-diverse CSD (PD-CSD). The other is to use a pickoff to carry a sample of the starlight traveling by a separate path; we call this Mach-Zehnder CSD (MZ-CSD).

The PD-CSD method is in a family of methods proposed by Borde and Traub\textsuperscript{10} and developed by Give’on et. al.\textsuperscript{13} and colleagues.\textsuperscript{16,17} A virtue of this technique is that it requires no additional hardware on the testbed (the blue beams and yellow mirrors in Figure 2). The MZ-CSD approach uses a separate beam path to capture and craft the reference field. The originally concept (Figure 2 top) was a beam picked off from the coronagraph field occulter (CFO), transported beside the coronagraph beam, and recombined with the coronagraph channel after the Lyot stop, so they could both play on the science detector together. This had the virtue of arbitrary phase shifting, but required a more complex optical system to transport the reference beam. Instead, our chosen method of creating a MZ reference beam uses Fizeau recombination as in Bau- doz et. al.\textsuperscript{11} but substituting a small pickoff at the Lyot plane (Figure 2 bottom). This pickoff takes the form of a pinhole through the Lyot stop, creating a spherical wave which is recollimated to a plane wave at the final image. This method yields a stable and repeatable reference beam, with a white-light fringe guaranteed within the image, and requiring a very simple hardware modification.

The MZ-CSD approach requires at least two reference beams with different optical phase, so that both quadratures of the speckle fields can be probed by interference with these reference fields. this is done with multiple pinholes through the Lyot stop. The light which illuminates these pinholes is from the bright rings at the image of the pupil rim. These rings
are caused by diffraction of starlight from the CFO mask.18 By locating the pinholes laterally near these bright rings, we can get a reference beam of suitable spatial coherence and brightness from a pinhole of 300µm diameter. Because the Lyot plane is a Fourier conjugate to the science focal plane, the spherical wave emerging from the pinhole is recollimated to a plane wave incident on the science focal plane. The tilt of this plane wave is

\[
[\theta_{rx}, \theta_{ry}] = \left[ \tan^{-1}(x_p / f), \tan^{-1}(y_p / f) \right]
\]

(6)

where \((x_p, y_p)\) is the lateral position of the pinhole in the Lyot plane, and \(f\) is the effective focal length of the optics from there to the science focal plane. The tilts of the reference beam \((\theta_{rx}, \theta_{ry})\) determine the reference beam’s optical phase at each pixel in the focal plane. For the central (on-axis) pixel in the focal plane, the reference beam ideally is in phase with all other fields from the Lyot plane. At any other pixel, the phase is

\[
\phi(x_f, y_f) = \frac{2\pi}{\lambda} (\theta_{rx} x_f + \theta_{ry} y_f) = \frac{2\pi}{\lambda f} (x_p x_f + y_p y_f)
\]

(7)

where \((x_f, y_f)\) are positions in the focal plane. To solve for the speckle field amplitude and phase, we must probe each pixel with fields that have significantly different phases \(\phi(x_f, y_f)\), which we can ensure by providing enough reference beams with carefully chosen \((x_p, y_p)\).

The SNR and integration time continue to improve with increasing reference beam intensity \(I_R\), but there is a practical limitation to that improvement—when the errors or uncertainty in the reference field are comparable to the speckle sensitivity we hope to achieve. Suppose there is an unknown complex spatial fluctuation \(\delta E_R\) of the reference field \(E_R\) at the science focal plane due to errors in the optics and the calibration uncertainty of the \(I_R\) profile. The observed intensity is

\[
I_R = |E_R + \delta E_R|^2 = E_R^2 + 2E_R \delta E_R \cos(\phi_S - \phi_R) + 2E_R \delta E_R \cos(\phi_S - \phi_R) + 2E_R \delta E_R \cos(\phi_S - \phi_S) + E_S^2 + \delta E_R^2
\]

(8)

We can ignore the last 3 terms because they are doubly small. The second term must be smaller than the third term by a chosen factor \(C\) (of order 5-10⁻⁸) so that our measurement of \(E_S\) is sufficiently unbiased. Let us define \(\sigma_R = 2E_R \delta E_R\), the uncertainty in the reference beam intensity \(I_R\). Then

\[
C\sigma_R < 2\sqrt{I_S I_R} \quad \text{and} \quad I_R < I_S \left(\frac{C\sigma_R}{2I_R}\right)^{-2}
\]

(9)

If \(\sigma_R/I_R\) is a fixed 1%, and \(C=10\), then \(I_R < 400 I_S\). For a typical \(I_S \approx 10^{-10}\) of the star, that means \(I_R < 4 \times 10^{-4}\) of the star. If the calibration of \(I_R\) is less accurate (\(\sigma_R/I_R > 1\%)\), or if the error budget for speckle knowledge is tightened (\(C>10\)), then \(I_R\) must be reduced.

2.3 MZ-CSD implementation

We have chosen four pinhole locations \((x_p, y_p)\) that allow reasonably robust solutions for the speckle field—good distribution of phases at each pixel, leading to good condition numbers for the solutions. (Figure 3.) And we have designed a shutter mechanism for selecting which pinholes are open to illuminate the focal plane.

- With all pinholes blocked, we can take science data, and calibrate the speckle intensity with no reference beams.
- Using each pinhole in turn, we can capture interference images to support a solution for the speckle fields.
- Using all combinations of two pinholes at a time, we can also cross-calibrate the reference fields themselves, validating the reference field amplitude and phase estimates and improving the speckle solution.

With our CSD-derived estimate of the speckle fields, we may choose to correct the speckles or simply to estimate and subtract the speckle intensity from the science data.

Figure 3 shows the arrangement of pinholes around the outside of the Lyot stop opening, and the condition number for the optimal solution for speckle fields. In this scenario, 5 images are taken: one with the speckles and no reference beam, and four with speckles and each of the 4 reference beams in turn. (Each reference beam is engaged by opening the corresponding pinhole.) We derive an optimal solution for the speckle fields based on those five images; that optimal solution is characterized by a condition number, representing the degree to which the measurements are independent of each

*the value \(C=10\) corresponds to a 10% calibration of \(E_S\), but a 20% calibration of the speckle intensity.*
other and therefore constitute “rigid” constraints on the solution. If all 4 reference beams have similar optical phase at a particular pixel, there will be a high degeneracy between the 4 measurements in that pixel and thus a poor solution there.

Currently this requires 4 pinholes, and thus 5 images taken in succession. We expect that experience will allow us to use fewer pinholes and fewer images, and thus reduce the time needed to calibrate the speckles by CSD.

2.4 Shutter mechanism

We want to be able to open each pinhole in turn, or none of the pinholes, or combine them in pairs. With four holes to be switched on or off, this constitutes 11 states (Table 1). A sliding mask that moves on a 3×4 grid of positions (with two motorized translation stages) can handle these 11 states. In each of these positions, the pinholes in the Lyot stop line up with 0, 1, or 2 holes in the moving mask.

Figure 3. MZ-CSD pinholes and quality of solution. Left: Arrangement of pinholes around the outside of the central opening of the Lyot stop. The “Football” is the traditional opening in the Lyot mask, where the planet light and speckles pass through the Lyot stop. This one is tailored for a linear band-limited mask like those currently in use in HCIT. The four pinholes are arranged on an arc at right. Right: the condition number at each pixel for a best solution for CSD-derived speckle fields. XY axes are equivalent sky angles in units of λ/D. Within the range from 4-10λ/D, the minimum condition number is 1 (ideal), and the maximum is 3.17.

Figure 4 Sliding-mask shutter mechanism. With 3 vertical positions and 4 horizontal, the dark gray sliding mask offers 12 possible combinations of open/closed pinholes. Shown here are 4 of the 11 states we actually need: (1) Only pinhole C is open; (2) only pinhole A is open; (3) pinholes A and B are both open; and (4) none of the pinholes are open.
A separate shutter blocks the entire “Football,” as shown in Figure 3 left. This may ultimately prove unnecessary, but for now we want a means of isolating each pinhole alone, for reference beam calibration without any stray light or speckle interference via the Football.

This system is more complicated than some concepts we considered, but makes use of motion stages that have been used before and are available now. Thus very little new software will be needed to operate the 3 mechanisms to govern the two shutters. The only new hardware needed was a Lyot mask with four 300 µm holes, a sliding mask with 16 holes of 1 mm diameter, and a sliding mask to block the Football.

2.5 Phase-diverse CSD

Previous work using phase-diverse CSD measurements and analysis was focused on controlling the wavefront and speckles to the best possible level. In this study we plan to use PD-CSD techniques primarily to measure and calibrate the speckles without necessarily controlling them. The goals are to establish the accuracy of the method and develop its speed.

A key limitation for this family of techniques has been the calibration of the action of the DM. What we call the reference field depends on the detailed response of the wavefront to the applied DM voltages. Traditionally, researchers have simplified their analysis by assuming the DM response is a linear superposition of influence functions—i.e. that the change in wavefront from a millivolt applied to one actuator is the same for every actuator, and independent of the initial state of the neighboring actuators. However, the response is likely to encompass a number of complicated effects, including mechanical stiffnesses, hysteresis, and perhaps electrical properties of the actuator and drivers. As such, the uncertainty in DM response is expected to be the largest contribution to the uncertainty in PD-CSD.

Our goal is to build on prior work and advance the use of DM-generated reference beams for speckle field measurement and speckle intensity estimation, particularly for background-limited observations. As shown above, this requires a reference field calibration accuracy of a few percent to retain sufficient accuracy at high background fluxes and high reference beam intensities. This in turn calls for improved understanding of the DM calibration. We will investigate MZ-CSD as a possible diagnostic capable of supplying that calibration.

3 CONCLUSIONS

The study was conceived with the goal of proving techniques for calibration of speckle intensity with near-optimal speed. The quickest possible integration time is the zero-background case, in which we only face the shot noise of the speckles themselves. If CSD methods can correctly measure the speckle intensity pattern with sufficient accuracy and in the zero-background integration time, regardless of the actual background, then the calibration measurements can be done very frequently. Then the telescope stability requirements would apply to much shorter time scales, e.g. an hour or two vs. a day, and that would substantially relax the challenge of designing, building, and testing the telescope.

The hardware needed for MZ-CSD experiments has been installed and tested, and we have begun experimenting with it in vacuum. The testing procedures have been defined, and preliminary analysis algorithms have been developed and demonstrated.
Currently our concept requires 4 pinholes, and thus 5 images taken in quick succession. We expect that experience will enable us to use fewer pinholes and fewer images, and thus reduce the time needed to calibrate the speckles by CSD. In addition, our earlier concept for generating a reference beam (captured from the CFO and transported downstream) has several potential advantages, which warrant follow-up experiments. In particular, this method appears to be more time-efficient, provides more uniform sensitivity across the science focal plane, and is applicable to virtually every kind of internal coronagraph instrument.

This work is funded by a NASA TDEM proposal. The work was partially conducted at the Jet Propulsion Laboratory, California Institute of Technology, under contract to the NASA.

REFERENCES

15. C.J. Burrows, private communication.