Reference-less Detection, Astrometry, and Photometry of Faint Companions with Adaptive Optics at 1, 2 and 5 μm

Szymon Gladysz  
Department of Experimental Physics, National University of Ireland, Galway

Julian C. Christou  
Division of Astronomical Sciences, National Science Foundation, Arlington, VA, USA

Matthew A. Kenworthy  
Steward Observatory, Tucson, AZ, USA

Nicholas M. Law, Richard G. Dekany  
Caltech Optical Observatories, California Institute of Technology, Pasadena, CA, USA

ABSTRACT

We propose a complete framework for detection, differential astrometry and photometry of faint companions, e.g. planets orbiting stars other than the Sun. The algorithms exploit the difference in statistics between the on-axis and off-axis intensity and in principle do not require the signal to be above the noise level. To test the accuracy of the proposed methods we used adaptive-optics short-exposure data from three observing campaigns: I-band observations with the 5m telescope at the Palomar Observatory, K-band data from the Lick Observatory's 3m telescope, and M-band images from the MMT Observatory's 6.5m telescope. We show that our algorithms outperform the classic procedures of PSF-subtraction for detection and PSF-fitting for photometry. All three proposed techniques are self-calibrating, i.e. they do not require observation of a calibration star after the science target thus improving the observing efficiency.

1. INTRODUCTION

In recent years the problem of imaging faint astronomical sources located close to bright objects has gained significant momentum after the discovery of the first planets outside the Solar System [1]. Most of these objects have been discovered by the Doppler technique which is inherently biased towards massive and close-in planets. Another important caveat is that radial velocity carries only partial information about the mass of an extrasolar planet (inclination angle ambiguity). Direct imaging is the natural next step in the search for extrasolar planets. The new challenge is to overcome formidable contrasts between planets and their parent stars, which can be anything from $10^{-4}$ in the infrared down to $10^{-10}$ and fainter in the visible regime [2,3]. The technology that can overcome such contrast ratios is being developed for both ground- and space-based observations [4-6]. In this paper we focus on ground-based adaptive-optics (AO) observations.

It has been shown through simulations and experiments that the limiting factors in high-contrast imaging from ground and space are static and quasi-static speckles [7-9]. While residual atmospheric aberrations after AO correction are random and will average out over time, these persistent speckles will stand out against the AO-corrected stellar halo and masquerade as faint sources even after long integrations. This is why future efforts to directly image planets will also entail post-processing schemes based on the concept of point spread function (PSF) subtraction assuming stable PSFs for the target and calibration objects. Proposed approaches utilize PSF estimates provided by on-sky rotation [10], as well as spectral [11] and polarization-based [12] discrimination between the light coming from the parent star and the companion. The problem with some of these techniques is that the PSF-subtracted images still contain static speckles (at a lower brightness level than in the direct images) due to errors in PSF estimation. These errors arise due to the inherent sensitivity of PSF subtraction to changes in “seeing”, mechanical flexure or the introduction of extra imaging channels. With a set of many object-like features in the PSF-subtracted image the question remains: What is a speckle and what is a planet?

We believe that the capability of current AO systems can be greatly enhanced by exploiting the statistical information present in corrected short exposures. Recently we showed that the statistical distribution of intensity at
the centre of the AO-corrected PSF has a different shape from the speckle distribution just outside the image core [13]. We propose to use this difference in the detection and estimation process. The advantage of the “stochastic speckle discrimination” methods [14] is that they are sensitive to any objects located close to bright stars because these techniques only depend on the properties of the image formation process. Their application is therefore not limited to cases of exoplanets with specific atmospheric composition, as opposed to spectral differential imaging [11] which is currently the most-widely used method for the discussed task.

2. THEORY OF STATISTICAL DETECTION

This section is only a brief introduction to the subject of speckle statistics. For a complete review the reader should consult earlier papers [15,14].

The AO speckle intensity distribution (probability density function – PDF) is described by the modified Rician function:

\[ p(I) = \frac{1}{I_s} \exp \left( -\frac{I + I_c}{I_s} \right) I_0 \left( \frac{2\sqrt{I} \sqrt{I_c}}{I_s} \right) \]

where \( I_c \) corresponds to the intensity produced by the deterministic (constant) part of the wavefront, and \( I_s \) corresponds to the halo produced by random intensity variations. \( I_0 \) is the zero-order modified Bessel function of the first kind.

The parameters \( I_c \) and \( I_s \) are related to the expected value \( E(I) \) and variance \( \sigma_I^2 \) of intensity through the following equations:

\[ E(I) = I_c + I_s \]
\[ \sigma_I^2 = I_s^2 + 2I_s I_c \]

The PDF given by Equation (1) is positively-skewed, i.e. it is asymmetric and its left tail is longer (Figure 1). We studied the statistics of AO speckle and we observed that histograms of on-axis intensity for AO-corrected PSFs cannot be described by the Rician PDF [13]. In our system characterization campaign at the Lick Observatory with moderate AO compensation we encountered only negatively-skewed histograms. We proposed a model for the PDF of the normalized on-axis intensity, i.e. the Strehl ratio (SR), which is equal to the measured on-axis intensity divided by the diffraction-limited peak intensity. This new PDF can be obtained by relating SR to wavefront phase variance \( \sigma_\phi^2 \) using the Maréchal approximation:

\[ S = e^{-\sigma_\phi^2} \]

The PDF expression is then

\[ p_{SR}(sr; k, \theta, \mu) = \frac{p_{\sigma_\phi^2}(-\ln sr; k, \theta, \mu)}{sr} \]

where \( SR \) denotes the random variable with possible values \( sr \), and \( p_{\sigma_\phi^2}(\cdot) \) is the distribution of the phase variance described by the gamma model:

\[ p(\sigma_\phi^2; k, \theta, \mu) = \frac{(\sigma_\phi^2 - \mu)^{k-1}}{\Gamma(k) \theta^k} \exp \left( \frac{-\sigma_\phi^2}{\theta} \right) \quad \text{for } \sigma_\phi^2 \geq \mu \]
where $k > 0$ is the shape parameter, $\theta > 0$ is the scale parameter, $\mu$ is the location parameter, and $\Gamma(x)$ denotes the gamma function. In Figure 1 the theoretical distribution of SR is plotted for three levels of AO compensation, as quantified by the total number of actuators on the deformable mirror of a 3m telescope. It should be noted that Equation (3) is only valid in the moderate and high SR regime.

The shape of the SR PDF is significantly different from that of speckle PDF in the moderate to high SR range, where the skewness of on-axis distribution is negative. This difference was studied by directly comparing estimates of the distributions [14]. The task required the computation of the parameters $I_c$ and $I_s$ for the time-series at the investigated location in a peak-shifted data cube. Subsequently only time-series with similar $I_c$ and $I_s$ were compared via distribution testing to avoid false alarms. The estimation of $I_c$ and $I_s$ was done using the method of moments, i.e. Equation (11) was inverted to produce estimates of the Rician parameters based on the sample estimators of the mean value and variance:

$$I_s = E(I) - (E(I)^2 - \sigma_I^2)^{1/2}$$

$$I_c = E(I) - I_s$$  \hspace{1cm} (6)$$

It was observed that at the locations of the artificial companions, inserted into the data-cubes estimation of parameter $I_c$ yielded values higher than for speckles of similar magnitude. Conversely, the value of $I_s$ was significantly lower for the companions’ peak pixels than for the speckles. This is easy to explain heuristically. On-axis intensity is the squared Fourier coefficient corresponding to the DC wavefront component, or the coherency of the wavefront. The goal of adaptive optics is to stabilize the wavefront (and therefore the SR). $I_c$, which captures the amount of coherence in a wavefront, will be large at the centre of the PSF. On the other hand, the parameter $I_s$ is a measure of the local intensity variations. For two pixels in the focal plane with similar mean intensity, one containing the companion and the other containing speckle, the intensity variations will be larger for the latter.

With this in mind we can design image post-processing techniques which transform spatial intensity variations (images) into maps of local statistics. One such transformation is

$$I_{SAA}(x,y) \rightarrow \frac{I_c(x,y)}{I_s(x,y)}$$  \hspace{1cm} (7)$$

where SAA stands for shift-and-add, i.e. the collapsed peak-centred data-cube which is traditionally used by detection algorithms. The transformation given by Equation (7) leads to a new representation of an object where centres of the PSFs have high values compared to the speckles which were equally bright in the SAA images. We illustrate this effect in the next section, but first we describe the observations which were used to test the new detection approach.
3. OBSERVATIONS

Data were obtained from three different observing programs representing poor, medium, and good AO correction corresponding to three different performance criteria as determined by the SR metric.

PALOMAR OBSERVATORY

At the 5m Hale telescope at Palomar observatory we used the LAMP Lucky Imaging camera [16] behind the PALMAO AO system. The observations used here were of a single star (HD 192849) imaged in 10nm bandpass centred around 950nm (I-band), with pixel size corresponding to Nyquist sampling. We obtained $10^4$ frames at the frame rate of 50Hz. The individual short exposures were registered with sub-pixel accuracy to produce SAA images (sub-pixel registering was applied to all observations described in this paper). Strehl ratio of the SAA image was 0.25 for this I-band data.

LICK OBSERVATORY

In our campaign at the Lick Observatory we observed over 50 single stars using the natural-guide-star AO system on the 3m Shane Telescope. Closed loop images of bright sources were obtained using the high-speed sub-array mode with a size of 64×64 pixels of the 256×256 pixel IRCAL camera. This corresponds to field size of 4.864×4.864 arcseconds. In the observations presented here the integration time was set to 22ms (followed by the readout time of 30ms) and we collected $\sim 10^4$ exposures per target. All data were obtained in the two-micron (K) band where the diffraction-limit is 151mas so that the PSF was effectively Nyquist-sampled. Measured SR in these SAA images were in the range 0.25 – 0.55. Details of the data reduction and computation of SR ratio are presented in the previous paper [13].

MMT OBSERVATORY

We collected 1000 short exposures with the optimized 3-5µm imager named Clio [3]. Clio can produce diffraction limited images in H, L' and M bands. Our observations were carried out in M-band (4.75µm). The important feature of the MMT dataset is that it is broadband (10% filter’s bandpass), and the successful application of speckle discrimination shows that the new methods can be used for polychromatic data. The exposure time was 64ms, and the PSFs were oversampled (3 pixels per $\lambda/D$). Measured mean SR was 0.8.

Figure 2 shows histograms of the instantaneous SR encountered in the observations described above. The reader should be warned that there is no direct correspondence between these histograms and the models shown in Figure 1. This is because the theoretical PDFs were plotted assuming constant parameters pertaining to atmospheric turbulence, same wavelength, same target brightness, etc. Nevertheless the change in shape of the on-axis PDF around 0.5 SR is easily noticeable.

![Fig. 2. Histograms of the instantaneous SR encountered in our observations.](image-url)
4. TESTS OF THE STOCHASTIC SPECKLE DISCRIMINATION APPROACH

Using real data we created artificial companions to stars by scaling and shifting single-star datasets. We varied the brightness and position of companions so that they would be hard to detect visually in the SAA images. For the relatively low-SR Palomar data the faintest companion in Figure 3 is 100 times less luminous than the star. For Lick data the furthest companion in Figure 3 is 400 times fainter than the central stars. The Lick dataset presented here corresponds to the star HD 216756 which is a binary of separation 1.4". We chose this dataset because of brightness of the central star and its relatively high SR (~0.5). Moreover this dataset allows for tests on real and simulated sources simultaneously. For the MMT dataset which consisted of only 1000 frames we were restricted to smaller contrast ratios than suggested by the relatively high SR of this data. This is because the estimation of $I_c/I_s$ via Equation (6) is noisier for thousand frames compared to ten thousand images in the other observations. The faintest simulated companion for the MMT test in Figure 3 is 400 less luminous than the central star.

Fig. 3. Top to bottom: Palomar, Lick and MMT datasets. Left: SAA images. Right: “maps” of $I_c/I_s$. The images are displayed on the same scale (square root), and using the same false-colour scheme.

The appearance of faint companions in the $I_c/I_s$ “maps” is striking. In many cases sources which were hidden in the SAA images are now very clearly seen. We will analyze these results separately.

In the case of the Palomar dataset the AO compensation was relatively low compared to the Lick and MMT. The target was also fainter. This resulted in the individual frames highly contaminated by Gaussian noise (background; readout noise was negligible). Speckles were barely visible in the short exposures. Because of the increased noise variance Equation (6) could not be computed for a large set of locations (no real value for the square root of a negative number). These locations were set to 0 in the $I_c/I_s$ map. With the new method we can detect sources 50 times fainter only 0.13" away from the stars.
The Lick observations contain many static speckles – this is what makes them challenging and interesting from the detection point of view. The visual ambiguity of speckles and companions in the SAA images disappears in the $I_c/I_s$ map. All five simulated companions are clearly visible. We can now detect sources 300 and 800 times fainter 0.7" and 1" away from the star, respectively.

In the case of the MMT data $I_c/I_s$ clearly suppresses the diffraction structure revealing five companions. We could detect sources 100 times and 400 times fainter 0.6" and 1" away from the star, respectively.

We also compared PSF subtraction to stochastic speckle discrimination performed on the PSF-subtracted data-cube. Two Lick datasets corresponding to the same star (HD 153832) were used. These observations were carried out one after the other and SR of the SAA images were very similar: 0.43, and 0.42. Again, five simulated companions were inserted in the data-cube and the SAA image. As can be seen, even a very-closely matched PSF leaves a significant structure in the PSF-subtracted image. In the $I_c/I_s$ map computed on the PSF-subtracted data-cube only peaks of the companions are non-zero (Figure 4).

We do not suggest that the new approach supersedes the standard technique of summing all frames. Rather, we see it as a complement to the classical detection. In studying exoplanets and sub-stellar companions to bright stars it is almost always necessary to use short exposures to avoid saturation of the detector. We suggest utilizing the statistical information which is present in these frames in the detection process.

5. DIFFERENTIAL ASTROMETRY

It is straightforward to extend the detection algorithm to find the relative position of a faint companion. One simply needs to interpolate the data-cube at the location of a suspected source, and find the maximum of the interpolated map. For interpolation we used the cubic convolution method because this algorithm closely approximates the theoretically optimum sinc interpolation. The $5 \times 5$ neighbourhood of the companion’s peak was sub-sampled with the spatial interval of 0.25 pixel. The ratio $I_c/I_s$ was then computed for each interpolated time-series. Subsequently two-dimensional Butterworth function was fitted to the data (for small separations we used Gaussian function).

We suggest utilizing the statistical information which is present in these frames in the detection process.

5. DIFFERENTIAL ASTROMETRY

It is straightforward to extend the detection algorithm to find the relative position of a faint companion. One simply needs to interpolate the data-cube at the location of a suspected source, and find the maximum of the interpolated map. For interpolation we used the cubic convolution method because this algorithm closely approximates the theoretically optimum sinc interpolation. The $5 \times 5$ neighbourhood of the companion’s peak was sub-sampled with the spatial interval of 0.25 pixel. The ratio $I_c/I_s$ was then computed for each interpolated time-series. Subsequently two-dimensional Butterworth function was fitted to the data (for small separations we used Gaussian function).

We suggest utilizing the statistical information which is present in these frames in the detection process.

5. DIFFERENTIAL ASTROMETRY

It is straightforward to extend the detection algorithm to find the relative position of a faint companion. One simply needs to interpolate the data-cube at the location of a suspected source, and find the maximum of the interpolated map. For interpolation we used the cubic convolution method because this algorithm closely approximates the theoretically optimum sinc interpolation. The $5 \times 5$ neighbourhood of the companion’s peak was sub-sampled with the spatial interval of 0.25 pixel. The ratio $I_c/I_s$ was then computed for each interpolated time-series. Subsequently two-dimensional Butterworth function was fitted to the data (for small separations we used Gaussian function).

We suggest utilizing the statistical information which is present in these frames in the detection process.

5. DIFFERENTIAL ASTROMETRY

It is straightforward to extend the detection algorithm to find the relative position of a faint companion. One simply needs to interpolate the data-cube at the location of a suspected source, and find the maximum of the interpolated map. For interpolation we used the cubic convolution method because this algorithm closely approximates the theoretically optimum sinc interpolation. The $5 \times 5$ neighbourhood of the companion’s peak was sub-sampled with the spatial interval of 0.25 pixel. The ratio $I_c/I_s$ was then computed for each interpolated time-series. Subsequently two-dimensional Butterworth function was fitted to the data (for small separations we used Gaussian function).

We suggest utilizing the statistical information which is present in these frames in the detection process.

5. DIFFERENTIAL ASTROMETRY

It is straightforward to extend the detection algorithm to find the relative position of a faint companion. One simply needs to interpolate the data-cube at the location of a suspected source, and find the maximum of the interpolated map. For interpolation we used the cubic convolution method because this algorithm closely approximates the theoretically optimum sinc interpolation. The $5 \times 5$ neighbourhood of the companion’s peak was sub-sampled with the spatial interval of 0.25 pixel. The ratio $I_c/I_s$ was then computed for each interpolated time-series. Subsequently two-dimensional Butterworth function was fitted to the data (for small separations we used Gaussian function).

We suggest utilizing the statistical information which is present in these frames in the detection process.
The accuracy of this approach to differential astrometry was tested on the Lick data for three locations shown in Figure 5 (left panel), approximately 0.45'', 0.8'' and 1.2'' away from the central star. The intensity ratio between the star and the companion was varied between 10 and 250 and expressed in stellar magnitude difference.

For companions located less than 0.5'' away from the bright star and being 10-250 times fainter, differential astrometry is accurate to within 0.05''. For separations larger than 1'' astrometric error is less than 0.01''. Errors never exceeded one pixel.

6. DIFFERENTIAL PHOTOMETRY

In Section 2 in Equations (4) and (5) we gave the expression for the SR PDF. SR is related to original measured intensity by a single numerical factor, i.e. peak of the diffraction-limited image $I^*$ normalized to have the same total power as the investigated PSF:

$$SR = \frac{I_{\text{peak}}}{I^*}$$  \hspace{1cm} (8)

We assume that on-axis intensity for each short exposure in a dataset is related to that frame’s instantaneous SR by a single number which is constant for the entire dataset. With this in mind one can find the distribution of on-axis intensity by transforming the SR PDF to obtain:

$$p(I; k, \theta, \mu, I^*_{\text{peak}}) = \frac{\left(-\ln \left(\frac{I}{I^*_{\text{peak}}}\right) - \mu \right)^{k-1} \exp \left(-\frac{-\ln \left(\frac{I}{I^*_{\text{peak}}}\right) - \mu}{\theta}\right)}{\Gamma(k)\theta}$$  \hspace{1cm} (9)

It is intuitively obvious that multiplication of a random variable by a scalar does not change the shape of its distribution and this is precisely what Equation (9) shows. The distribution of AO on-axis intensity has the same shape as the SR PDF, given by Equation (4), and this shape is significantly different from that of speckle PDF in the moderate to high SR range, where the on-axis distribution’s skewness is negative. Again, this difference will be used to separate speckle noise and companion’s signal. This time the task at hand is to estimate the relative brightness of a companion “sitting” on the speckle halo.

Two random processes take place at the location of the companion’s peak (ignoring Gaussian and Poissonian noise). The speckle and signal intensities add, and the PDFs are convolved. The analytical form of these PDFs is known. This offers an incentive to develop a “PDF deconvolution” technique. The distribution of the signal (peak intensity) described by Equation (9) is “blurred” by the speckle kernel given by equation (1). Between them, these distributions have six parameters. Fortunately, the first three parameters of the signal’s PDF are common for the bright star and the companion. These parameters could be estimated for the bright star, so that the deconvolution problem is reduced to finding only three parameters: $I_c$, $I_s$ and $I^*_{\text{peak}}$. We show how this could be done by minimizing the difference between an estimate of the composite PDF and the observed histogram.

We applied the method to the MMT dataset. The first step in our procedure is the fit of Equation (9) to the bright-star histogram. This is done using least-squares, one-dimensional fitting routine (Figure 6). The parameters we want to estimate are: $k$, which is related to the number of independent phase patches in the wavefront, $\theta$, which is related to the mean phase variance, and $\mu$, which corresponds to the static aberrations in the wavefront (shifting the phase variance PDF towards higher values). The term to be minimized is

$$E = \sum_{j=0}^{n-1} \left[h_j - \hat{p}_{on}(I_j; k, \theta, \mu, I^*_{\text{peak}})\right]^2$$  \hspace{1cm} (10)

where $h$ is the histogram with $n$ bins, and $\hat{p}_{on}(I_j; k, \theta, \mu, I^*_{\text{peak}})$ is the estimate of the on-axis PDF described by Equation (9) evaluated for intensities corresponding to the histogram bins $I_j$. 

The value of $I_{\text{peak}}^*$ obtained in the previous step is subsequently discarded, and the values for $k$, $\theta$ and $\mu$ are kept for the estimation of the simulated companion’s $\Delta m$, i.e. the magnitude difference. In the fitting routine these values are now kept fixed, and the algorithm only searches for $I_c, I_s$ and $I_{\text{peak}}^*$ which provide the best fit to the companion’s histogram. The term to be minimized is now

$$E = \sum_{j=0}^{n-1} \left[ I_j - \hat{p}_{\text{off}}(I_j; k, \theta, \mu, I_{\text{peak}}^*) \otimes \hat{p}_{\text{off}}(I_j; I_c, I_s) \right]^2$$

(11)

where $\hat{p}_{\text{off}}(I_j; I_c, I_s)$ is the estimate of the off-axis PDF given by Equation (1), and $k$, $\theta$ and $\mu$ are now kept fixed. $\otimes$ denotes convolution.

In the right panel of Figure 7 we plot the intensity histogram at the location of a simulated companion (left panel), together with the constrained fit of convolution of Equations (9) and (1) to this histogram. After the fitting routine converges the mean flux of the companion can be obtained by calculating the mean value of the signal’s random process. The photometric error is then computed as:

$$\text{photometric error} = \Delta m - \Delta^*$$

(12)

where $\Delta m$ is the true magnitude difference, and $\Delta^*$ is its estimate.

Fig. 6. Least-squares fit of Equation (9) to the intensity histogram at the central point in the MMT dataset.

Fig. 7. Left: MMT dataset; SAA image of a simulated faint companion. Right: Result of "PDF deconvolution" carried out on the corresponding data-cube.
In Figure 8 we present the comparison between our method and Starfinder which is a PSF-fitting package designed for the analysis of crowded fields imaged with AO [17]. In this paper we provide StarFinder with a PSF estimate – a single star matched as closely as possible to the science object. For the MMT dataset we did not have a separate calibration PSF therefore we used the SAA image of the first 500 frames as a very closely-matching PSF estimate.

It is clear from Figure 8 the “PDF deconvolution” outperforms PSF-fitting. It is easy to understand why. This is a one-dimensional approach and the complicated speckle structure present in the SAA image is irrelevant. It is also a reference-less approach, therefore it is not influenced by errors in the PSF estimation. It should be remembered that these errors were kept at a minimum in our comparison – we used almost the same data for the simulated target and the calibration PSF.

7. DISCUSSION

The point that has to be addressed is the number of frames and their maximum exposure times that can be used within the proposed short-exposure framework. We predict that the exposure time can be quite long (perhaps one second in K-band) because of the correlations observed for on-axis intensity [13]. These correlations mean that the central limit theorem (CLT) will smear the statistics slower than expected for independent samples, so the proposed algorithm could be classified as a “multiple-frame”, rather than a “short-exposure” technique. An initial analysis of this effect is presented in Figure 9.

Fig. 9. Left: $I_c/I_s$ map for the Lick dataset HD 216756 with one real companion and five simulated companions, $10^4$ frames with exposure time 22ms. Centre: $I_c/I_s$ map for the same dataset binned into $10^3$ frames with exposure time 0.2s. Right: $I_c/I_s$ map for the data-cube binned into $10^2$ frames with exposure time 2s.
The impact of Poisson noise for very faint sources and short exposures would be to transform all statistics, so that the speckle and companion PDFs will look very similar, invalidating our photometric method. The observer would have to balance that fact against the impact of the CLT which destroys the PDF asymmetry for longer exposures.

For coronagraphic imaging, where the estimation of the on-axis PDF is harder, we propose to use the wavefronts reconstructed from the wavefront sensor to obtain the parameters $k$, $\theta$ and $\mu$. We plan to carry out an experiment to test this approach.

This research was supported by Science Foundation Ireland under Grants 02/PI.2/039C and 07/IN.1/I906, as well as the National Science Foundation Science and Technology Center for Adaptive Optics, which is managed by the University of California at Santa Cruz under cooperative agreement AST 98-76783. SG was also funded by the European Office of Aerospace Research and Development under the Window on Science Cooperative Agreement Award No. FA8655-08-1-2070. We are indebted to Bill Bradford for endorsing our research.

8. REFERENCES