A Possible Advantage of Telescopes with a Noncircular Pupil

Guy Nir1, Barak Zackay2, and Eran O. Ofek1

1 Benoziyo Center for Astrophysics, Weizmann Institute of Science, 76100 Rehovot, Israel
2 Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, USA

Received 2018 September 26; revised 2019 June 5; accepted 2019 June 5; published 2019 July 18

Abstract

Most telescope designs have a circular aperture. We quantify the advantages that telescopes with an elongated pupil have over circular-pupil telescopes in terms of contrast at small separations between a bright central star and a faint companion. We simulate images for an elongated-pupil telescope and for a circular-pupil telescope of equal aperture area and integration time, and we specifically investigate the maximal contrast when finding faint companions around bright stars as a function of angular separation. We show that this design gives better contrast at lower separation from a bright star. This is shown for diffraction-limited (for perfect and imperfect optics) and seeing-limited speckle images assuming an equal aperture area and observing time. We also show that the results are robust to errors in measurement of the point-spread function. To compensate for the wider point-spread function of the short axis, images should be taken at different rotation angles, either by rotating the telescope around the optical axis or by allowing a stationary mirror array to scan different parallactic angles with time. Images taken at different rotation angles are added using the proper image coaddition algorithms developed by Zackay & Ofek. The final image has the same contrast at all angles, rather than in specific areas of diffraction nulls. We obtained speckle observations with a small ground-based elongated-aperture telescope and show that the results are consistent with simulations.

Key words: methods: data analysis – methods: observational – techniques: high angular resolution – techniques: image processing – telescopes

1 Introduction

Telescope design is usually restricted to a circular pupil, which also has an approximately circular-symmetric point-spread function (PSF). In the past, telescopes with a noncircular pupil design have been considered as a means to improve resolution in one axis or to generate strong diffraction nulls that help detect faint companions or planets around bright stars. For example, the Large Binocular Telescope (Byard & Bonaccini 1994) uses a noncircular pupil to achieve high resolution; the proposed design for DARWIN (Kaltenegger & Fridlund 2006) suggests that a constellation of satellites is used and their light is combined on long baselines, which allows the detections of faint planets at very small angular separation; the proposed Terrestrial Planet Finder mission (Brown et al. 2003; Coulter 2003) had two designs, one with multiple satellites and one with a coronagraph. The use of a nonapodized coronagraph (Spergel & Kasdin 2001) and an aperture that is specifically apodized to create strong nulling of the central object have been suggested for this mission (Spergel 2001). Similar ideas have also been suggested by Rafanelli & Rehfield (1993) and more recently by Rafanelli et al. (2017).

For any telescope design, resolution will be hindered by imperfect optics, and for ground-based telescopes, by the turbulent atmosphere. We show that under these conditions, a rotating, elongated-pupil telescope will outperform a symmetric design in terms of resolution for equal aperture area and exposure time. The improved resolution is not restricted to one axis where the elongated pupil is larger than the circular one, but is improved at all angles when proper coaddition methods are used. We describe the image-processing algorithms for adding images at multiple angles while conserving all the constant-in-time information in the individual images. The main goal of this work is to test whether telescopes with an asymmetric pupil design have an advantage over circular pupil telescopes in terms of average angular resolution and contrast.

A few applications that may benefit from this approach are segmented mirror telescopes, where the same number of mirrors can be aligned in a row instead of a disk to achieve higher resolution; or space-based missions, where a folding telescope would have less mass and fewer moving parts if it only needs to open up in one axis.

Some engineering constraints may complicate the construction and use of elongated pupil telescopes that also need to be rotated. These problems should be addressed for each individual application and are beyond the scope of this work. Others have recently addressed some of these issues, e.g., Monreal et al. (2018) have proposed a specific design implementing a telescope with a long and narrow pupil, while Green et al. (2018) proposed using a long and narrow aperture to facilitate the long baselines required for detection of exoplanets in the mid-infrared.
In Section 4 we present ground-based observations using a of the elongated-pupil telescope and relevant image processing. In Section 4 we present simulations of the elongated-pupil telescope and relevant image processing.

2. Review of the Image Coaddition Algorithm

In this section, we outline the methods used to coadd images in the subsequent simulations and data analysis. In the background-dominated case, Zackay & Ofek (2017a, 2017b) derived from first principles a coaddition method that is numerically stable, produces images with uncorrelated noise, and preserves information at all spatial frequencies. We refer to such an image as a “proper image.” The algorithm gives more weight to frequency bins that have more information. The highest frequencies in the PSF of this coadded image are similar to the highest frequencies in the individual PSFs.

The resulting image in Fourier space is given by

\[ \hat{R} = \frac{\sum_j f_j \hat{P}_j M_j}{\sqrt{\sum_j \sigma_j^2 |\hat{P}_j|^2}} \equiv T \hat{P}_R + \epsilon_R, \]  

Here \( M_j \) are the measured images with image index \( j \), and \( P_j \) are the unity-normalized PSFs for each image. The notation \( \hat{\square} \) represents a two-dimensional Fourier transform, while \( \square \) represents complex conjugation. The overall flux of each image is \( f_j \) and the background-noise standard deviation of each image is \( \sigma_j \). The resulting image \( R \) can be represented as the true image \( T \) convolved with an effective PSF \( P_R \), and some uncorrelated noise \( \epsilon_R \). The effective PSF of the proper image is

\[ \hat{P}_R = \left[ \sum_j \frac{f_j^2}{\sigma_j^2 |\hat{P}_j|^2} \right]^{1/2}. \]  

For ground-based long-exposure images (more than a few seconds), the PSF of each image can be measured by observing a bright star (or multiple stars) in the frame. For exposures on timescales of \( \lesssim 10 \) ms, the PSF varies from image to image and over small angular distances within the same image (e.g., Title et al. 1975; Chassat et al. 1989). In this case, it may still be possible to estimate the PSF using a wavefront sensor (Primot et al. 1990; Fried 1987), phase retrieval techniques (Knox & Thompson 1974; Lohmann et al. 1983), or using the image as its own PSF (Zackay & Ofek 2017b).

If the PSF is known and the images are background-noise dominated, the resulting image given by Equations (1) and (2) is both optimal and a sufficient statistic; e.g., the coadded final image can be used to find faint companions or as a reference image for transient searches (Zackay et al. 2016).

2.1. Speckle Coaddition: The Square-root Method

In certain applications, e.g., when obtaining speckle images without any nearby reference star, it is difficult to measure the PSF directly. In some cases, when the image is dominated by a point source, we can approximate the PSF of each image by the image itself, as discussed in Zackay & Ofek (2017b):

\[ f_j P_j \approx M_j. \]  

Here we use the approximation that the noise \( \sigma_j \) is constant for all images and absorb the flux term \( f_j \) into the overall normalization of the PSF. Note that for extended object such as galaxies, this approximation is no longer valid, and a different method should be used to estimate the PSF. The resulting proper coadded image is

\[ \hat{Q} \approx \frac{\sum_j M_j \hat{M}_j}{\sqrt{\sum_j |M_j|^2}} = \sqrt{\sum_j |M_j|^2} \approx \hat{P}_R. \]  

The result is similar to the correlation map of the image, as was used, e.g., by Labeyrie (1970), and more recently by Tokovinin (2018). The main difference being in the normalization, in Fourier space, by the standard deviation of each frequency, which ensures that if the noise in the original images is independent and identically distributed (i.i.d), then the noise in the resulting image is also i.i.d.

For images taken at short exposure times (\( \lesssim 10 \) ms), the speckle pattern of each star in every exposure is a good approximation for the PSF. Using Equation (4) without any additional information will result in an image with dramatically improved resolution compared to simply summing the images (Zackay & Ofek 2017b). In this approximation, we do not recover the true image, \( T \), but an estimate of the power spectrum of the true image, \( |\hat{T}| \), so that the resulting image, \( Q \), is no longer a sufficient statistic for general hypothesis testing or measurement. However, it is still possible to answer specific statistical questions, e.g., looking for faint companions around bright stars, differentiating point sources from extended sources, or for measuring the flux of multiple adjacent stars.

To calculate this statistic, we zero-pad each image to twice its original size, Fourier transform each image, and take the absolute value squared of each pixel. The resulting power spectrum for each image is summed, and finally, the square root of the sum of power spectra is obtained. When searching for point sources, the coadded image is filtered by its own PSF:

\[ \hat{Q} \hat{P}_R \approx \sum_j |M_j|^2. \]  

Zero-padding of the input images is done so that the multiplication in Fourier space is equivalent to convolution without cyclical boundaries (the FFT itself uses cyclical boundary conditions). When the inverse FFT is performed to obtain the coadded image, we also crop back to the original image size.

2.2. Treatment of Correlated and Uncorrelated Noise

The image \( Q \) represents a correlation map, therefore the central pixel of the image in position space contains information on the observed object but also on the sum of all the noise contributions from all spatial frequencies. To remove the “zero-point” correlation of the noise, we subtract a constant term, the minimum of the image in Fourier plane:

\[ \hat{Q}_{\text{adj}} = \hat{Q} - \text{min}(\hat{Q}). \]  

---

3 A statistic is sufficient with respect to a statistical model and its associated unknown parameter if no other statistic that can be calculated from the same sample provides any additional information as to the value of the parameter (Fisher 1922).
If the image suffers from strong correlated noise (e.g., line noise sometimes seen in sCMOS devices), we instead subtract the minimum of each row and then the minimum of each column (the order of operations is based on the strength of the line noise in the horizontal and vertical directions):
\[
\hat{Q}_\text{adj} = \hat{Q} - \min_u (\hat{Q}_u) - \min_v (\hat{Q} - \min_u (\hat{Q}_u)), \quad \text{or}
\]
\[
\hat{Q}_\text{adj} = \hat{Q} - \min_v (\hat{Q}_v) - \min_u (\hat{Q} - \min_v (\hat{Q}_v)).
\]
These treatments improve the overall image and the maximum achievable contrast. Throughout this work, the algorithm presented above is used for simulations and measurements where no PSF data are available. When PSF data are also available, a different algorithm is used, as described in Section 2.3.

### 2.3. High-contrast Imaging with Known PSF

When looking for dim companions around bright stars, there is an optimal\(^4\) coaddition technique that is not limited to the case of background-dominated noise. If the PSF of each image is known (e.g., in some space-based missions or when the wavefront aberrations are measured independently), an optimal statistic can be used specifically for the detection of faint companions (hereafter referred to as the binary coaddition statistic). This method is developed from first principles in Appendix A, and will be discussed further in B. Zackay et al. (2019, in preparation). This method is used for all simulations where we assume that the PSF is known. The resulting statistic is given by

\[
S = \sum_j f_j \left( \overrightarrow{P}_j \otimes \frac{M_j - f_j P_j}{f_j P_j + \sigma_j^2} - \sum_{x,y} P_{x,y} \frac{M_j - f_j P_j}{f_j P_j + \sigma_j^2} \right),
\]

and the variance for each point is given by

\[
V_s = \sum_j f_j^2 \left( \overrightarrow{P}_j^2 \otimes \frac{1}{f_j P_j + \sigma_j^2} - 2 \overrightarrow{P}_j \otimes \frac{P_j}{f_j P_j + \sigma_j^2} + \sum_{x,y} \frac{P_{x,y}^2}{f_j P_j + \sigma_j^2} \right).
\]

Here we use \(\overrightarrow{\cdot}\) to denote coordinate reversal \((x, y \rightarrow -x, -y)\) and \(\otimes\) to denote two-dimensional convolution. Using this statistic, which is optimal for the detection of high-contrast companions, the detection limit for each point in the image is given by

\[
C = \frac{\sqrt{V_s}}{S/N},
\]

where \(S/N\) is the signal-to-noise ratio (in units of standard deviation) required for detection. In this work, we adopt \(S/N = 5\) as a threshold for detection.

### 2.4. Coaddition of Elongated-pupil Telescope Images

In the case of the elongated-pupil telescope, the images have a wide PSF in one direction, corresponding to the narrow axis of the pupil. By imaging the source at different rotation angles (i.e., different position angles of the pupil’s long axis as projected on the sky), and using the algorithm of Zackay & Ofek (2017b) that we presented in Section 2.3, information is recovered from frequency bins in all directions. Thus the elongated pupil can recover details on separations as small as the diffraction limit of a circular pupil with the same diameter as the long axis of the elongated pupil, as shown in Figure 3.

In order to recover a final image that has a symmetric, round PSF, the pupil needs to be rotated around the optical axis by a full 180°, and the position angle, \(\theta\), should be sampled in intervals smaller than the ratio between the long and narrow sides of the PSF. In our simulation (with a width-length ratio of 10), we found that \(\Delta \theta \sim 5^\circ\) is sufficient (see Section 3.5).

### 3. Simulations

We conducted several sets of simulations to compare the theoretical ability of circular- versus elongated-pupil telescopes to detect close and faint companions around bright stars. We simulated PSFs from a circular and an elongated pupil. A 180 cm diameter was chosen for the circular pupil, while the elongated-pupil telescope had a 500 by 50 cm rectangular aperture. This results in both telescopes having nearly the same aperture area.

The goal of these simulations was to test the relative performance of the circular- and elongated-pupil telescopes, not the coaddition algorithms themselves. The same methods were used for both telescopes, so that they could be compared under equivalent conditions.

We ran three sets of simulations, the first assuming perfect optics (Section 3.1); the second assuming imperfect optics (Section 3.2), where the wavefront aberrations have so-called red-noise properties\(^5\); and the last set using atmospheric aberrations of the wavefront (Section 3.3).

The code we used is available as part of the MATLAB astronomy & astrophysics toolbox (Ofek 2014).

#### 3.1. Simulation of Diffraction-limited Images

In the first set of simulations, we assume that the optics are perfect and the PSF is known. This regime represents the absolute maximum contrast attainable for either circular- or elongated-pupil telescopes.

We simulated a 10th mag star observed in the \(V\) band for 36 exposures of 125 s, equivalent to a total of approximately \(10^{10}\) photons. A circular aperture was employed to simulate a symmetric PSF, which is used in two simulations: the unrotated simulation, where a single orientation is used for a long exposure of \(36 \times 125\) s; and the rotated simulation, where the PSF was rotated by 5° between each of 36 consecutive orientations, exposing each for 125 s. The elongated-pupil aperture was used to produce an elongated PSF, which was used in 36 consecutive 125 s exposures with a 5° rotation angle step. To minimize interpolation errors, we used three-shear rotation, using Fourier interpolation for the different skews (Larkin et al. 1997). For each exposure, the pixel scale was set to 32 pixels per \(\lambda/D\), where \(D\) was chosen to be the long axis of each telescope. Oversampling by a factor of 16 over the Nyquist sampling was used so that features (such as the contrast curves) of the resulting coadded images could be

---

\(^{4}\) Optimal in the sense that for a given detection rate, it has a minimum false-alarm rate.

\(^{5}\) In the case of optics, the aberrations typically have a higher amplitude at lower spatial frequencies in the pupil plane.
measured more accurately. In real measurements, a sampling of 2 pixels per $\lambda/D$ is sufficient. Source noise and an additional noise of 1 electron e.g., read noise were added to all images. The noise for the image of the unrotated circular-pupil simulation was set to $\sqrt{36}$ to compensate for having one exposure instead of many. The pupils and resulting PSFs for the circular- and elongated-pupil simulated telescopes are shown in Figure 1.

The images were summed using the generalized proper coaddition algorithm given in Equation (1). In Figure 2 we show for both apertures a 1D cut through the resulting PSF, $P_R$ from Equation (2). In Figure 3 we show the approximated 1D modulation transfer function (MTF), which is the absolute value of the Fourier transform of the coadded PSF, $|P_R|$. Above $\lambda/D$, the information content in the MTFs is zero, but residuals in the approximated MTFs remain. These noise residuals are caused by clipping of the PSF in the simulation, which saves some calculation time in the coaddition. It should be noted that the values beyond the diffraction limit of each aperture, on the order of $10^{-5}$, can be effectively treated as going to zero.

The high-contrast statistic given by Equations (9) and (10) was used to calculate the maximum contrast for detection of a faint companion around the central star. We used the map of the square root of the variance as in Equation (11) and calculated the median of all the pixels at a certain radius from the center, i.e., within an annulus with a two-pixel width. The circular-pupil simulations resulted in a variance map with many features, similar to the rings in an Airy disk, so that each annulus contained pixels with widely varying contrast values; a median statistic for each annulus was found to give stable contrast values. To find the contrast, we assumed a $5\sigma$ detection threshold. The results for perfect optics in the known-PSF case are shown in Figure 4.

The simulations show that the elongated-pupil telescope has a similar contrast to that of a circular telescope with the same area, as long as the separation is larger than the diffraction limit of the circular telescope. Inspecting the contrast curves below $0.06/\lambda$ (300 cm of the circular telescope) or the MTF in Figure 3, we can see that the elongated pupil preserves information on smaller angular scales than the circular pupil. The simulations suggest that the theoretical resolution of the elongated pupil is the diffraction limit of the long axis. For a 1:10 axis ratio, the diffraction limit is 2.78 times smaller than for a circular pupil of the same area.

Note that the contrast curves of the rotated and nonrotated circular telescope are nearly identical. While this is expected because the circular-pupil PSF has circular symmetry, it also suggests that the theoretical resolution of the elongated pupil is the diffraction limit of the long axis. For a 1:10 axis ratio, the diffraction limit is 2.78 times smaller than for a circular pupil of the same area.

Figure 1. Simulated pupils and PSFs for the circular- and elongated-pupil telescopes. (a) The aperture shape of the circular-pupil telescope has a shorter diameter than the long edge of the elongated-pupil telescope. (b) The aperture shape of the elongated-pupil telescope has the same area as the circular pupil. (c) The PSF from the circular pupil is an Airy disk. (d) The PSF from the elongated pupil is wider on one axis but narrower on the other. The two PSFs are shown on the same angular scale, with the diffraction limit shown in each corner for reference.

Figure 2. A 1D profile (cut through) of the proper coaddition PSF, $P_R$, as given by Equation (2). The PSF of the elongated pupil (black line) is narrower than that of the circular pupil (gray line).

Figure 3. A 1D profile (integrated over circles of increasing radii) of the approximate MTF of the two apertures for the case of perfect optics. The approximate MTF is given as the absolute value of the coaddition PSF, in Fourier space, $|P_R|$. The elongated-pupil MTF has information content at higher frequencies than the circular-pupil MTF. The theoretical diffraction limit of the circular- and elongated-pupils is shown as gray and black dashed lines, respectively. The radius of the inner circle of the Airy disk (1.22 $\lambda/D$) for each pupil is given by the gray and black dotted lines. For the elongated pupil we see that there is information in the MTF up to the diffraction limit calculated based on the long edge of the pupil. Note that the residual $10^{-5}$ beyond $\lambda/D$ is caused by clipping of the PSF in the simulation and does not contain any information.

Figure 4. The highest contrast between a primary star and a companion that can be detected at $S/N = 5$ for different angular separations. This simulation of circular- and elongated-pupil telescopes used perfect optics and no atmospheric aberrations. The bottom x-axis uses an angular scale to compare different pupil sizes, but can be scaled linearly if the sizes of all pupils are changed by the same proportion. The upper x-axis shows the angular distance in units of the diffraction limit of the circular pupil. All simulations used the same total exposure time and aperture area. The thin black curve describes the result for a single orientation of a circular telescope with an aperture of 180 cm. The thick gray curve describes the results for 36 images taken at different angles with a circular telescope. The thick black curve describes the results for 36 images taken at different angles using a 50 × 500 cm elongated-pupil telescope. As expected for perfect optics, the rotated and nonrotated circular telescopes give nearly identical results. The contrast is similar for the circular- and elongated-pupil telescopes until the diffraction limit is reached, at which point the contrast drops sharply. The elongated pupil has a smaller diffraction limit, consistent with the long axis of the pupil, and thus preserves contrast at lower angular separations.
shows that proper rotation (using FFT-skew transformations) does not bias the results.

All angular scales are shown for specific aperture sizes that were chosen arbitrarily, and can be scaled linearly to any telescope size. For example, the contrast at 0\(^\prime\)1 for the 180 cm circular pupil (or 500 x 50 cm elongated pupil) is equivalent to the same contrast at 0\(^\prime\)05 for a 360 cm (or 1000 x 100 cm) telescope as long as the total photon count is preserved.

### 3.2. Simulation of Diffraction-limited Images with Imperfect Optics

Real-life telescopes have imperfect optics. Simulations were made for apertures with imperfections of the optical surfaces. The amplitude of the deviation of the aberrated wavefront from the planar wavefront was randomly chosen from a normal distribution with a variance following a power law in the spatial frequency:

\[
\sigma_{k}^{2} = Af^{-2},
\]

where \(A = 0.2\) rad\(^2\) is the variance of the base frequency, and \(f\) is the spatial frequency in units of pixel\(^{-1}\). The base frequency is set by the size of each of the pupils (the diameter or the long edge). The aberrated PSFs are then rotated and used to produce images (as was done in Section 3.1). The images were coadded using proper coaddition, as in Equation (1). The resulting PSF and MTF showed no substantial difference from the perfect-optics case because the inputs to the coaddition algorithm are the aberrated PSFs, so that the imperfections are accounted for in the weighing of the coadded image.

The images were also coadded using Equations (9) and (10) (i.e., assuming the PSF is known). The resulting contrast curves for the rotated circular-pupil telescope and the elongated-pupil telescope are shown in Figure 5. As each realization of the optical aberrations had a very different contrast curve, we avoided strong biases that may arise from any specific realization by simulating 30 complete sets of 36 images of each pupil shape. The resulting contrast curves were averaged, and the variation between each simulation was used to derive 1\(\sigma\) scatter intervals.

As expected, the addition of optical imperfections reduces the contrast achieved in all the simulations. This causes both telescopes to start losing contrast at higher separations than the diffraction limit, but the elongated telescope is more robust to the aberrations. This is likely due to the fact that the elongated pupil samples only a narrow band out of the space of possible spatial frequencies of the aberrations.

### 3.3. Simulation of Speckle Images

To simulate ground-based telescope observations, we generated a random phase-screen using the prescription in Jia et al. (2015). We gave each spatial frequency a random value from a normal distribution following the power law

\[
\sigma^{2}(f) = 0.033\left(\frac{r_{0}}{D}\right)^{-5/3}\left(f\sigma_{i}\right)^{-11/6}\exp(-f^{2}/f_{p}^{2}),
\]

where \(f = |f|\) is the size of the two-dimensional spatial frequency vector \(f\), in units of pixel\(^{-1}\), while \(f_{i} = 5 \times 10^{3}\) pixel\(^{-1}\) and \(f_{p} = 0.5\) pixel\(^{-1}\) are the frequencies corresponding to the inner and outer scale of the atmosphere, respectively.\(^6\) The Fried length is given by \(r_{0} = 10\) cm, corresponding to a seeing of \(\sim 1^\prime\), and the telescope aperture is given by \(D\). For each image, we produced one phase screen that was wide enough to encompass both the circular- and elongated-pupil mask arrays, so that both telescopes were simulated to view a different cut of the exact same sky in each of the images. The resulting PSFs were used for three simulations: the circular PSF with source noise and an additional noise of 1 count/pixel; the elongated PSF with source noise and an additional noise of 1 count/pixel; and the same elongated PSF with noise scaled by the angular size of each pixel (e.g., scaling the noise as background rather than readout noise).

The number of photons in the simulations was set to be equivalent to that produced by a magnitude 4 star in the V band. Batches of 50 images were simulated at a 10 ms exposure time for each image. For each batch, the image was rotated by 5\(^\circ\), with 36 batches covering the range of 180\(^\circ\). The high magnitude was chosen so that the total photons in all exposures is \(\sim 10^{10}\), while requiring only a few hundred simulated images. Simulation of fainter stars, while maintaining the total number of photons, gave similar results, but required more speckles to be generated and thus a longer runtime.

The pupils for the two telescopes, overlaid with an example phase screen, are presented in Figure 6, along with the resulting speckle PSFs.

We used the same simulation to compare the two telescopes using three coaddition methods: the proper coaddition method given by Equation (1), the approximation given by Equation (4) for situations where the PSF is not known, and the specialized high-contrast imaging method given by Equation (9), where we assume that the PSF is known.

The coadded PSF, \(PR\), from Equation (2) is shown in Figure 7, while the approximated MTF, \(|\tilde{PR}|\), is shown in Figure 8. As before, the elongated pupil gives a narrower PSF.

---

\(^6\) The outer scale is the scale at which energy enters the atmosphere, from which it cascades in turbulent eddies down to smaller scales until the inner scale is reached, where dissipation overtakes turbulence. Here we assume that the outer scale is 10 m, while the inner scale is 1 mm.
Fried length of approximated MTF of the two apertures, under atmospheric conditions with a Fried length of \( r_0 = 10 \) cm, assuming knowledge of the PSF. The PSF of the elongated pupil is narrower than that of the circular pupil.

The contrast curves for the circular- and elongated-pupil telescopes for the case of unknown PSF are presented in Figure 9. To find the contrast in this method, we compared the peak intensity to the noise in an annulus with a two-pixel width in the filtered coadded image, as given by Equation (5). Because there is no phase information in this method, the intensity of the companion would be split symmetrically into two peaks on either side of the center of the image. So we divide the contrast we find by a factor of two and by another factor of \( N_p = 5 \) to find the contrast at which we expect to make a detection with an \( S/N = 5 \).

The simulations show that the elongated-pupil telescope, for the same area, has similar contrast to that of the circular telescope down to the diffraction limit of the circular-pupil telescope. The elongated-pupil telescope preserves contrast at smaller separation angles, down to the diffraction limit set by the long axis of the pupil. Compared to the previous results for diffraction-limited images, the contrast for both apertures is lower because of the atmospheric conditions and lack of knowledge of the PSF.

In the second case, we assume knowledge of the PSF for each speckle image, e.g., by using a wavefront sensor. In this case we use the method for detecting high-contrast companions given by Equation (9). The results are shown in Figure 10. The contrast is dramatically improved in comparison to the previous method, but the atmosphere still reduces the attainable maximum contrast as compared to the diffraction-limited case. For this coaddition method, we show two simulations of the elongated pupil, one with noise of equal magnitude as the circular-pupil telescope, represented by the dotted black line, and one with noise scaled by the angular size of the pixels, represented by the solid black line. The elongated pupil has smaller pixels to match the smaller diffraction limit of the telescope. If the noise is constant per pixel (e.g., read noise), the contrast from the elongated pupil is lower than from the circular pupil. If the noise is constant per angular area (e.g., sky background), the contrast for the two pupil shapes is very similar for an angular separation greater than the diffraction limit of the circular telescope. Below that limit, the circular-
telescope contrast is lower than the elongated-telescope contrast. Because the diffraction limit of the elongated telescope is set by the long edge of the pupil, its contrast begins to diminish at an angular separation that is about three times smaller than the limit for the circular telescope.

3.4. The Effect of Uncertainty in the PSF

The binary coaddition method (Equation (9)) depends on knowledge of the PSF. We tested the ability of this method to recover a faint companion around a bright star in situations where the PSF is not perfectly measured. Contrast curves are calculated analytically using the input PSFs, and not the images, as seen in Equation (10). Therefore, we cannot use them to test the effects of the difference between the aberrated PSFs and the images, which are simulated using the original PSFs. Instead, we planted a faint companion with a contrast ratio of 1:10⁴ at an angular separation of 0″5 from a simulated 4th mag star. We simulated atmospheric conditions and used perfect optics. We generated 18,000 images in batches of 100 images separated by an angle of 1″ between each batch. We used the same atmospheric phase screen for both the circular- and elongated-pupil telescope simulations, and coadded the resulting images. The PSFs produced by the atmospheric phase screen were used directly to create the images by multiplying by the flux of the two stars, with the correct offset for the faint companion, and then adding noise proportional to the angular scale of each pixel, i.e., 1 count per pixel for the circular-pupil telescope and 0.13 counts per pixel for the elongated-pupil telescope. The PSFs were given to the coaddition algorithm only after adding Gaussian noise to each pixel of the PSF, with rms proportional to the intensity in that pixel. We repeated this simulation five times and averaged the score maps of each coaddition result.

The statistical score maps allow the detection of the companion up to a fractional PSF noise of 0.1. For the case of perfect knowledge of the PSF, the bright star at the center of the score map should be completely removed by the coaddition method. However, when PSF errors are added, it leaks back into the score map and contaminates the companion’s signal. Even though the companion’s signal shows a peak at the correct location, the signal is washed out by the false signal from the main star, as seen in the example in Figure 11. To estimate the recovered signal, we measured the intensity of the peak, subtracted from it the average of two points at a distance of 5 pixels to the left and right of the companion peak, and then divided by the standard deviation of all pixels at a distance between 4 and 10 pixels from the companion peak. The resulting loss of signal is evident in Figure 12.

Further work on testing the effects of other kinds of PSF inaccuracies will be detailed in B. Zackay et al. (2019, in preparation). Initial tests show that PSFs with red-noise optical aberrations suffer even more leakage from the central star. It may be possible to mitigate this effect by aligning the PSF to the image.

3.5. The Effect of the Rotation Angle Sampling

We measured the maximum rotation angle sampling step size that can be used without losing contrast. We simulated images using perfect optics, perfect knowledge of the PSF, and no atmosphere and coadded the images as described in
and 90° companion with a contrast of 5. The minimum contrast to detect a companion in each annulus is plotted to highlight the angle step where some parts of the image lose contrast. We find that for a 1:10 pupil axes ratio, the contrast is the same up to steps of $\Delta \theta = 45^\circ$. However, by inspection of the coaddition results, the overall coadded image quality starts to degrade at around $\Delta \theta = 5^\circ$.

The simulations were conducted several times using different rotation angle sampling. We used $\Delta \theta = 1^\circ$, $5^\circ$, $15^\circ$, $45^\circ$, $60^\circ$, and $90^\circ$ sampling steps. In each case, we simulated 180 images. For $\Delta \theta = 1^\circ$, each image was simulated at a different rotation angle, and for the other simulations, the same angles were repeated several times so that the total number of images was the same. Source noise and an additional 1 electron per pixel were added as before. The contrast curves remained constant for all rotation angle steps, but at higher step values, the 2D contrast surface is less smooth, and polygon modulations dominate the image. Because the contrast image is no longer uniform, we do not expect the median on each annulus to be a good representation of the contrast. Instead, we display the minimum contrast in each annulus, which shows the largest loss of contrast when large rotation angle steps are used. The results are shown in Figure 13.

We see that for a 1:10 ratio elongated-pupil telescope, contrast is fairly well maintained up to $\sim 45^\circ$. We also planted a companion with a contrast of $5 \times 10^6$ at a separation of $0^\prime 3$, which should be detectable at that distance. For all simulations up to and including $\Delta \theta = 60^\circ$, the companion is recovered with the expected S/N. To calculate a proper image, rather than look for point sources with high-contrast ratios, we coadded the simulated images using the proper coaddition technique given by Equation (1). In the coadded image, we see diffraction spikes around the PSF even at $\Delta \theta = 5^\circ$, with increasing intensity of spikes as $\Delta \theta$ increases. It is hard to quantify the effect on image quality from these results. However, we can estimate that the PSF will be sufficiently sharp for angle steps smaller than $\Delta \theta = 5^\circ$.

4. Proof of Concept Using Real Measurements

We performed some observations under atmospheric conditions using a small telescope with an elongated pupil. We demonstrate the coaddition method in the case of an unknown PSF and show that the basic reconstruction methodology works.

Observations were taken at the Weizmann Institute during one night on 2018 January 8, using the Kraar observatory 40 cm telescope at f/25. A fast sCMOS camera (Andor Zyla 5.5) was used to capture speckle images at 100 Hz with 0.5 ms exposure times and $\approx 2.5 \, e^-$/pix read noise. The telescope was outfitted with a cardboard mask to simulate a long and narrow aperture. The width of the mask was 2.2 cm, while its length was 40 cm, excluding an obscuration of $\approx 10$ cm in the center of the aperture, where the secondary mirror blocks the light. Observations were made in batches of 3000 images with the mask set at the same orientation during each batch. Between each batch, the mask was rotated by $\approx 12$ deg. We observed Sirius ($M \sim -1.4$) using a V filter.

We also produced simulations using the same aperture and assuming a Fried length of $r_0 = 4$ cm. We matched the total photon count and background noise of the simulations to the observed images. We used the same ratio of long-to-narrow aperture, and excluded the central obscuration. The PSF was calculated by generating a random phase screen and then shifting it three times and averaging the resulting PSFs to simulate a $10 \, m \, s^{-1}$ wind. To match the telescope results, we added specific optical aberrations to the wavefront in the form of defocus and astigmatism terms. The two terms are given by

$$\Delta \varphi_{\text{def}}(\rho, \theta) = \sqrt{3} (2 \rho^2 - 1), \quad \Delta \varphi_{\text{ast}}(\rho, \theta) = \sqrt{6} \rho^2 \cos 2\theta,$$

where $\rho$ and $\theta$ are the pupil position parameters, and $\Delta \varphi_{\text{def}}$ and $\Delta \varphi_{\text{ast}}$ are the defocus and astigmatism added phase term, respectively, given in radians. The amplitude of each of these aberrations was chosen empirically to be one, by comparing simulations to the measured results. In addition to these errors, we also added red-noise optical aberrations with an amplitude of $A = 0.2$, as in Equation (12). Unlike in previous simulations, the optical aberrations were stationary when the elongated pupil mask was rotated during the simulation because in the actual measurements we did not rotate the optics of the telescope but only the cardboard mask.

Some sample images from the simulation and from the actual measurements are shown in Figure 14. The effects of the seeing and optical aberrations is visible, although for these specific images the simulation appears to have somewhat sharper speckles. The coaddition result, using Equation (4), also shows that the simulation and data are fairly similar, even though the simulation is slightly sharper. A profile through the coaddition result is shown in Figure 15 for slices through the 2D map, in a vertical direction and in a $45^\circ$ angle.
We show the MTFs for the simulation and the data in Figure 16. We see that the optical aberrations cause a loss of information at scales lower than the diffraction limit. The coaddition can also be used to find the maximal contrast for a companion to be detectable at a given angular separation. The contrast curve results for the simulation and the data are shown in Figure 17.

The drop in contrast below the seeing limit, which is more pronounced in the data and the simulations discussed in this section relative to the simulations shown in Section 3.3, is most likely due to the astigmatism of the telescope. The asymmetry is hard to notice in individual speckle images, but is apparent in the coaddition result. It is clearly visible in Figure 18, where the image has been subtracted from its own transpose. We calculated the contrast as the ratio of peak intensity to the noise standard deviation in annuli around the peak. The astigmatism asymmetry has a strong effect on the measured noise in each annulus, which reduces the contrast dramatically. It is likely that this effect would be smaller if the astigmatism in the telescope were lower; if the mirror were rotated and the asymmetry averaged out over all angles; if the PSF were measured and used in the coaddition; or, ideally, all of the above.

5. Summary

In this work, we quantified the advantages of an elongated-pupil telescope and showed that it is able to detect faint companions around bright stars at a higher contrast ratio and closer separation than a circular-pupil telescope of equal area. We conducted simulations comparing the two designs using two coaddition algorithms, which were used when the PSF is known and when it is unknown. We tested them on images simulated assuming no atmosphere and under the assumptions of perfect, and then imperfect, red-noise optics, while also assuming perfect knowledge of the PSF. We performed tests on different rotation angle steps and found the minimum angle needed to maintain contrast for a 1:10 axes ratio.

We conducted simulations assuming atmospheric conditions in the two cases of known and unknown PSF. In the latter, we applied the coaddition algorithm discussed in Section 2.1. We also presented observations with a small telescope fitted with a mask limiting the aperture to a long slit, and compared the results to simulations.
The simulations comparing circular- and elongated-pupil telescopes all show that the elongated-pupil telescope has an advantage when high-contrast companions at close separations are searched for. We demonstrated that using image coaddition techniques allows the use of new asymmetric designs where higher resolution can be recovered while maintaining the total area of the telescope. These results also highlight the importance of using proper image coaddition when images with very different PSFs are added.

G.N. would like to thank N. Eliyahu and H. Azzam for their participation in observations for this work. B.Z. acknowledges the support from the infosys fund. E.O.O. is grateful for the support by grants from the Israel Science Foundation, Minerva, Israeli Ministry of Science, Weizmann-UK, the US-Israel Binational Science Foundation, and the I-CORE Program of the Planning and Budgeting Committee and The Israel Science Foundation.

Appendix A
Derivation of the Binary Coaddition Method

Source detection under the assumptions of Gaussian independently and identically distributed noise has an optimal solution: the matched-filter (Turin 1960). This well-known result can be derived from first principles, i.e., from the lemma of Neyman & Pearson (1933). In some cases, when a faint object is situated next to a very bright object (i.e., high-contrast imaging), the noise from the bright object is no longer identically distributed; it depends on the proximity to the star. Furthermore, methods for coadding images that are optimal under the assumption of background-dominated noise (Zackay & Ofek 2017a, 2017b) are no longer valid because the source noise from the bright object is dominant.

We show that using the lemma of Neyman & Pearson (1933), we can derive an optimal statistical test for finding a faint companion around a bright star. The null hypothesis is that of a single star with flux \( f_j \), while the alternative is that another star (e.g., a companion) is at position \( q \) with flux \( c f \). The data models, under these two hypotheses, are given by

\[
D_j(x) = \begin{cases} 
 f_j P_j(x) + \epsilon & \text{if } H_0 \\
 f_j P_j(x') \otimes [(1-c)\delta(x') + c\delta(x' - q)] + \epsilon & \text{if } H_1.
\end{cases}
\]  

(15)

The likelihood ratio statistic is the difference of these terms

\[
S(q) = \sum_{j,k} \frac{-2 cf_j P_j(x - q) [D_j(x) - f_j P_j(x)] + 2 c f_j P_j(x) [D_j(x) - f_j P_j(x)] + c^2 [f_j P_j(x) - f_j P_j(x - q)]^2}{f_j P_j(x) + B_j(x)}.
\]  

(20)

Here \( j \) is the image index, where \( D_j(x) \) are images of the same object, \( x \) is the 2D position, \( P_j(x) \) is the PSF of each frame, \( f_j \) is the flux of the primary in each frame (taking into account, e.g., changing transmission of the atmosphere), \( \epsilon \) is the noise (assumed to be normally distributed), and \( \otimes \) is the convolution operator over the dummy coordinates \( x' \).

The most powerful test to tell these hypotheses apart would be the likelihood ratio test (Neyman & Pearson 1933). The likelihood for the data given the null hypothesis is

\[
L(D|H_0) = \exp \left[ -\frac{1}{2} \sum_{j,x} \frac{[D_j(x) - f_j P_j(x)]^2}{\sigma^2_j(x)} \right].
\]  

(16)

The variance of each pixel, \( \sigma^2_j(x) \), is a combination of source noise and background (e.g., read noise, sky brightness, dark current):

\[
\sigma^2_j(x) = \sigma^2_{sx}(x) + \sigma^2_{bg}(x) = f_j P_j(x) + B_j(x),
\]  

(17)

where \( B_j(x) \) is the total background for each pixel and each frame, and the source noise variance is equal to the star’s mean brightness, where all quantities are in units of electrons. Note that we are assuming that the noise contribution of the companion is negligible compared to the background and primary source noise.

We can define the log likelihood \( LL \equiv -2 \log(L) \):

\[
LL(D|H_0) = \sum_{j,x} \frac{[D_j(x) - f_j P_j(x)]^2}{f_j P_j(x) + B_j(x)}.
\]  

(18)

On the other hand, the log likelihood for the alternative is

\[
LL(D|H_1) = \frac{\sum_{j,x} [D_j(x) - f_j P_j(x')] \otimes [(1-c)\delta(x') + c\delta(x' - q)]^2}{f_j P_j(x) + B_j(x)}.
\]  

(19)

The likelihood ratio statistic is the difference of these two terms

\[
S(q) = \sum_{j,x} \frac{f_j P_j(x - q) [D_j(x) - f_j P_j(x)] - f_j P_j(x) [D_j(x) - f_j P_j(x)]}{f_j P_j(x) + B_j(x)}.
\]  

(21)

Because \( c \ll 1 \), we can remove \( c^2 \) terms. After rescaling by \(-2c\), we obtain

\[
S(q) = \sum_{j,x} \frac{f_j P_j(x - q) [D_j(x) - f_j P_j(x)] - f_j P_j(x) [D_j(x) - f_j P_j(x)]}{f_j P_j(x) + B_j(x)}.
\]  

(20)

\[
S(q) = \sum_{j,x} \frac{f_j P_j(x - q) [D_j(x) - f_j P_j(x)] - f_j P_j(x) [D_j(x) - f_j P_j(x)]}{f_j P_j(x) + B_j(x)}.
\]  

(21)
In the first term in the second line we have used the definition of convolution: \( \sum_j F(x - q)G(x) = (F \otimes G)(q) \).

To make any decisions based on this statistic, we must calculate its variance:

\[
V[S] = \sum_{j,x} \left[ \frac{f_j (p_j(x) - P_j(x)(D_j(x) - f_j P_j(x)))}{f_j P_j(x) + B_j(x)} \right]
\]

\[
= \sum_{j,x} \left[ \frac{(p_j(x) - P_j(x))(D_j(x) - f_j P_j(x))}{f_j P_j(x) + B_j(x)} \right]
\]

\[
= \sum_{j,x} \left[ \frac{(p_j(x) - P_j(x))^2}{f_j P_j(x) + B_j(x)} \right] V[D_j(x) - f_j P_j(x)]
\]

\[
= \sum_{j,x} \left[ \frac{(p_j(x) - P_j(x))^2}{f_j P_j(x) + B_j(x)} \right] V[D_j(x)]
\]

\[
= \sum_{j,x} \left[ \frac{(p_j(x) - P_j(x))^2}{f_j P_j(x) + B_j(x)} \right] [f_j P_j(x) + B_j(x)]
\]

\[
= \sum_{j,x} \left[ \frac{(p_j(x) - P_j(x))^2}{f_j P_j(x) + B_j(x)} \right] [f_j P_j(x) + B_j(x)]
\]

\[
= \sum_{j,x} \left[ \frac{(p_j(x) - P_j(x))^2}{f_j P_j(x) + B_j(x)} \right]
\]

\[
- 2 \left( \frac{p_j(x)}{f_j P_j(x) + B_j(x)} \right) \sum_x \frac{P_j^2(x)}{f_j P_j(x) + B_j(x)}
\]

\[
+ \sum_x \frac{P_j^2(x)}{f_j P_j(x) + B_j(x)} \right].
\]

Finally, the resulting statistic \( S / \sqrt{V(S)} \) would give the detection \( S/N \) for any possible detection of a companion. The results can be used to estimate the probability for a false-alarm detection in the usual way.

If we wish to determine the maximum contrast for detection, we calculate the expectation value \( \langle S \rangle \):

\[
\langle S \rangle = \sum_{j,x} \left[ \frac{(p_j(x) - P_j(x))(D_j(x) - f_j P_j(x))}{f_j P_j(x) + B_j(x)} \right]
\]

\[
= \sum_{j,x} \left[ \frac{(p_j(x) - P_j(x))(f_j P_j(x) + [(1 - c) \delta(x) + c \delta(x - q)] - f_j P_j(x))}{f_j P_j(x) + B_j(x)} \right]
\]

\[
= \sum_{j,x} \left[ \frac{f_j (p_j(x) - P_j(x))(f_j P_j(x) + (1 - c) f_j P_j(x) + c f_j P_j(x) - f_j P_j(x))}{f_j P_j(x) + B_j(x)} \right]
\]

\[
= \sum_{j,x} \left[ \frac{c f_j (p_j(x) - P_j(x))^2}{f_j P_j(x) + B_j(x)} \right] = c V[S],
\]

where we have used the expectation value \( \langle D \rangle = f_j P_j(x') \otimes [(1 - c) \delta(x') + c \delta(x' - q)] \). Note that the last equality is evident from the fifth line of Equation (22). We can calculate the \( S/N \) for a given value of \( c \):

\[
S/N = \frac{\langle S \rangle}{\sqrt{V(S)}} = c \sqrt{V(S)},
\]

from which we can determine the maximum contrast \( C \) for detection at a given \( S/N \):

\[
C = c^{-1} = \frac{\sqrt{V(S)}}{S/N}.
\]

Appendix B

Comparison with Wiener Filtering of Multiple Images

In some cases, images with very different PSFs (e.g., from different rotation angles) can be used to try to reconstruct the original image. One way to produce this deconvolution image is to use a Wiener filter, which uses the mean \( S/N \) of all the different images (Yaroslavsky & Caulfield 1994, hereafter YC):

\[
\hat{T}_{\text{reconstructed}} = \sum_j \frac{1}{P_j} \frac{\hat{T}^2 D_j / \sigma_j^2}{1 + \sum_j \hat{T}^2 D_j / \sigma_j^2},
\]

where all sizes are calculated in Fourier space, and we have assumed that the images are background dominated with noise variance \( \sigma_j^2 \) that can be position dependent or uniform. To produce this estimate of the true image \( \hat{T} \), we require some knowledge of the power spectrum of the true image, \( \hat{T}^2 \), as well as knowledge of the PSF. It is not guaranteed that a good estimate of the power spectrum is available, e.g., when trying to detect a faint companion around a bright star, we wish to discern between the spectrum of a single star and the spectrum of a binary, and cannot a priori assume it is one or the other.

While for some applications it may be useful to find the reconstruction of the true image, it is not necessarily the best statistical result for performing measurements. The coaddition result presented in Equation (1) can be used to perform any kind of measurements that can be done on the constant-in-time content of all the images (see the discussion in Zackay & Ofek 2017a, 2017b). Thus it is possible to recover the deconvolution results of YC from the proper coadded image, and its PSF, without having to save the individual images. If we are given an estimate of the power spectrum of the true image, we can apply a Wiener filter directly to the single coadded image, replacing \( M_j \rightarrow R \) and \( P_j \rightarrow P_R \):

\[
\hat{T}_{\text{reconstructed}} = \frac{1}{P_R} \frac{\hat{T}^2 R / \sigma_R^2}{1 + \sum_j \hat{T}^2 R / \sigma_j^2},
\]
where we have replaced the sums with the single image, $\hat{R}$, and its PSF, $\hat{P}_R$, and because the proper image is normalized by the background of the input images, the background noise of the proper image is, by construction, $\sigma_R = 1$. From this we can expand using Equations (1) and (2):

$$\hat{T}_{\text{reconstructed}} = \frac{1}{\sqrt{\sum_j \left| \hat{P}_j \right|^2 / \sigma_j^2}} \times \frac{\sum_j \hat{P}_j \hat{M}_j / \sigma_j^2}{\left( 1 + \sum_j \left| \hat{P}_j \right|^2 / \sigma_j^2 \right)^{1/2}}. \quad (28)$$

Finally, if we multiply and divide the $\hat{P}_j$ term with $\hat{P}_j$, the resulting expression becomes equivalent to the YC result:

$$\hat{T}_{\text{reconstructed}} = \sum_j \frac{\left| \hat{P}_j \right|^2 \hat{M}_j / \sigma_j^2}{\left( 1 + \sum_j \left| \hat{P}_j \right|^2 / \sigma_j^2 \right)^{1/2}}. \quad (29)$$

**ORCID iDs**

Guy Nir @ https://orcid.org/0000-0002-7501-5579
Eran O. Ofek @ https://orcid.org/0000-0002-6786-8774