Application of the Hotelling and ideal observers to detection and localization of exoplanets

Luca Caucci,[1,*] Harrison H. Barrett,[1,2] Nicholas Devaney,[3] and Jeffrey J. Rodríguez[4]

[1]*College of Optical Sciences, University of Arizona, Tucson, Arizona 85721, USA*
[2]*Department of Radiology, University of Arizona, Tucson, Arizona 85724, USA*
[3]*Department of Physics, National University of Ireland, Galway, Ireland*
[4]*Department of Electrical and Computer Engineering, University of Arizona, Tucson, Arizona 85721, USA*
[*]*Corresponding author: caucci@email.arizona.edu*

Received April 12, 2007; accepted June 1, 2007;
posted June 22, 2007 (Doc. ID 82008); published September 25, 2007

The ideal linear discriminant or Hotelling observer is widely used for detection tasks and image-quality assessment in medical imaging, but it has had little application in other imaging fields. We apply it to detection of planets outside of our solar system with long-exposure images obtained from ground-based or space-based telescopes. The statistical limitations in this problem include Poisson noise arising mainly from the host star, electronic noise in the image detector, randomness or uncertainty in the point-spread function (PSF) of the telescope, and possibly a random background. PSF randomness is reduced but not eliminated by the use of adaptive optics. We concentrate here on the effects of Poisson and electronic noise, but we also show how to extend the calculation to include a random PSF. For the case where the PSF is known exactly, we compare the Hotelling observer to other observers commonly used for planet detection; comparison is based on receiver operating characteristic (ROC) and localization ROC (LROC) curves.

OCIS codes: 010.1080, 110.3000, 350.1260.

1. INTRODUCTION

Science images are ultimately acquired to fulfill a particular purpose or task. How well the task can be performed depends on the task itself (which could be a classification or an estimation task), the device that produced the image(s), and, finally, the way in which the task is performed. A classification task can be carried out by human observers (for example, a radiologist might look at radiographs and classify them according to the presence of a tumor), or we may consider computer algorithms or mathematical models to make a decision. In the latter case, it is possible to define an ideal observer that, given the images produced by the imaging system and the knowledge of their statistics, achieves the "best" performance on the task of interest. Thus, assessment of task performance will provide a figure of merit for the imaging system itself and also a way to assess the quality of the images it produces for the particular task of interest. This task-based interpretation of image quality has recently become very popular in the medical field for the evaluation of medical devices. Tasks of interest include the detection of anomalies (such as tumors) and the estimation of clinical parameters. Objective assessment of image quality plays a crucial role in many nonmedical applications; examples can be found in [1]. However, little has been done so far in astronomy in which, many times, the task of interest is the detection of a weak signal buried in noise or the estimation of some numerical parameters. In this paper, we show how the concepts of image quality and signal detection can be applied to the problem of detection of extrasolar planets (exoplanets).

To date, more that 240 exoplanets have been discovered [2], most of them by indirect detection methods, such as radial velocity or Doppler shift. However, indirect methods are not able to provide very much information about a newly discovered exoplanet. Direct exoplanet imaging is more attractive and far more could be learned from it. For example, a direct imaging method could be coupled with a spectroscopy analysis for the estimation of methane content. The presence on the planet (or in its atmosphere) of organic substances (such as carbon dioxide), oxygen, and liquid water or vapor will provide strong evidence of life—as we know it—or, at least, that such a planet is amenable to life. Many indirect methods for exoplanet detection are not even able to find planets that may host life. Indeed, some indirect methods work well only when the planet is considerably more massive than the Earth (for example, ten times the mass of Jupiter), making these methods unsuited for Earth-like exoplanet hunting.

Adaptive optics (AO) [3,4] has recently made it possible to perform ground-based imaging of exoplanets [5,6] but, up to now, only a few planets have been discovered by direct imaging. Consequently, algorithms for direct-imaging exoplanet detection have mushroomed. Many of these algorithms share a common characteristic: No analytical way to assess their performances (with respect to a meaningful and objective metric) has been investigated; therefore, they do not lend themselves to an image-quality study. In this paper, we propose a strategy for planet detection for which the proof of optimality among all detection methods can be provided, leading to a way to assess image quality in AO [7]. We use the optimal-linear Hotelling detector to solve two problems: simple detection at a known location and detection with location uncertainty.

This paper is organized as follows. Section 2 provides the motivation for this research and emphasizes the need for a meaningful and well-defined definition of task performance. Section 3 describes the components of an AO system and justifies the investigation of task performance for long-exposure AO images. Section 4 reviews some key concepts from statistical decision theory, describes the kind of problems with which we are concerned in this paper, and presents the derivations of the optimal observers for such problems. Section 5 presents the simulation and comparison results we obtained for the detection tasks. Finally, Section 6 summarizes our work, presents our conclusions, and suggests some possible further studies.

2. WHY DO WE NEED ANOTHER ALGORITHM?

We begin this section by giving a brief overview of the standard practice for the detection of point objects on a random background. After a preprocessing step that usually makes use of dark and flat-field images to obtain the observed image from the CCD data, the mean and variance of the sky (residual) background in the image are estimated. Such estimation may be carried out by first smoothing the image with a low-pass filter. The smoothed image is then divided into tiles, and the mean and variance for each tile of the smoothed image are computed. The sky background is then subtracted from the observed image, and objects are detected by searching for pixels whose intensity exceeds the sky background intensity by a given threshold τ. Usually, the threshold is of the form $\tau=k\sigma$, where σ^2 is the sky background variance and k is an integer between 3 and 10. Many variants of this basic idea have been used in the past.

Although the algorithms for exoplanet detection found in the literature are very interesting, and the papers in which they are presented support the design decisions with strong arguments, most of their authors do not address an important point: Among all of the possible detection algorithms, do the ones described in those papers maximize the probability of a correct detection for any value of the probability of a false alarm? If yes, is there a way to analytically assess their performance (with respect to this criterion) and prove their optimality? Questions like these are usually overlooked by the astronomy exoplanet detection community: A formal analytical study—such as the one in [8]—is the exception, not the rule. Indeed, a common way to show that a new detection algorithm outperforms a previous one is to create one or more test images containing the object(s) that we want to detect, run both algorithms on such data, and claim that the new algorithm performs better because it is able to correctly find more objects than the previous one. This approach takes no account of false alarms.

Bayesian approaches to object detection (and/or parameters estimation) have been proposed as well. For example, Hobson and McLachlan [9] presented an approach that uses Bayesian inference for detecting and characterizing the signal from discrete objects embedded in a diffuse background. This is accomplished by considering each possible alternative H_i and computing the likelihood $\mathrm{pr}(\mathbf{g}|H_i)$, where $\mathbf{g}$ is the observed noisy data. The alternative H_i that, for a fixed $\mathbf{g}$, corresponds to the highest value for $\mathrm{pr}(\mathbf{g}|H_i)$ is the output of the algorithm. This approach is optimal in the sense that it returns the alternative that most likely has produced the observed data. A similar approach can be used for parameter estimation: If the parameter vector θ is to be estimated, the likelihood $\mathrm{pr}(\mathbf{g}|\theta)$ for different θ is considered and the estimate $\hat{\theta}$ of θ maximizes $\mathrm{pr}(\mathbf{g}|\theta)$ as a function of θ. Maximum likelihood approaches have been applied in astronomy for the detection of cluster galaxies and the estimation of numerical parameters, such as redshifts [10,11]. Algorithms like these, however, are nonlinear and require the knowledge of the probability density functions of the alternatives/parameters and of the data vector $\mathbf{g}$ given the alternative/parameter. Such densities might be unknown or difficult to estimate. In [9], the authors use Markov-chain Monte Carlo (MCMC) sampling [12,13] to evaluate integrals and compute the values of the likelihood functions. The computational requirements can, however, be very high. A Bayesian approach has been proposed in [14] as well. In this paper, the authors analyze the measurement and noise statistics and devise a second method, based on the estimation of observation time. Their observation time estimates are optimal in the sense that, for a fixed value of the probability of a false positive detection, the probability of a true detection is maximized. Detection occurs when the signal estimate irradiance exceeds the background standard deviation by some amount. The probability of such an event is equated to a desired confidence value and, from this condition, the optimal integration time is computed. This can be seen as a first attempt to design a system/algorithm that is optimized (in a given sense) for the task of interest. The remarkable result of [14] is that, for high background intensity, the point-spread function (PSF) matching filter is the optimal linear detector.

The problem of exoplanet detection via direct imaging shares a great deal with the problem of tumor detection. In both cases, the objects that we want to detect are very small and dim. In the case of radiology, the smaller the tumor that can be accurately detected and localized, the more likely it is that the patient can recover if well treated. In astronomy, because exoplanets radiate much less energy than their parent stars, finding them in an image is not easy. In addition, the images we work with in both fields are usually noisy and contain other structures (different organs and tissues in the case of radiology, nearby stars or other objects in the case of astronomy), hence it seems reasonable to take advantage of techniques from medical radiology to study direct exoplanet detection. Even though such methods and techniques for assessment of image quality and task performance in radiology are general enough that they could be applied in other areas of imaging as well, they are just beginning to be applied in astronomy [7].

The goal of this paper is to discuss the use of receiver operating characteristic (ROC) and localization ROC (LROC) curves for the study of objective assessment of image quality and task performance in AO. In particular, ROC/LROC analysis is applied to the problem of exoplanet detection. Adopting the area under the ROC/LROC curve as a figure of merit, we devise optimal detec-

tors and show what improvements over current detection algorithms can be attained with them.

3. SIGNAL DETECTION WITH ADAPTIVE OPTICS

Figure 1 shows the typical configuration of an AO telescope. A distorted wavefront coming from a distant source is reflected by the telescope's primary and secondary mirrors to a deformable mirror, which attempts to reduce the wavefront distortion due to atmospheric turbulence by applying an appropriate phase correction. The light reflected by the deformable mirror is then split into two different pathways: Some light reaches the imaging camera, while the remaining light is sent to a wavefront sensor. The wavefront sensor estimates the instantaneous aberrations; these estimates are then sent to the wavefront reconstructor. The wavefront reconstructor uses one or more data sets coming from the wavefront sensor to compute the appropriate control signals to be sent to the deformable mirror in order to compensate for atmospheric turbulence. The deformable mirror then applies the desired wavefront correction (specified by the control signals) to the incoming wavefront. This process repeats in real-time in a closed-loop fashion. Sequences of science images—usually in a time scale of the order of a few minutes—are recorded. As the wavefront reconstructor accumulates more and more wavefront estimates for distortion compensation, the signal-to-noise ratio (SNR) in the science images increases. However, this process is not perfect: Wavefront estimation is affected by noise in the wavefront sensor, and the computation and application of the phase correction takes some time (at least some milliseconds). Atmospheric turbulence is rapidly changing and, because of the computational delay mentioned above, the applied correction will not be able to perfectly compensate and cancel the effect of the turbulence. A more subtle problem is that atmospheric phase aberrations can be mathematically described as elements of an infinite-dimensional vector space, whereas the finiteness of the number of control signals supplied to the deformable mirror implies that the vector space of phase corrections is finite-dimensional. This dimension mismatch makes it impossible for the deformable mirror to perfectly match and correct every possible phase distortion. The overall effect is that the PSF of the science images will have a residual halo.

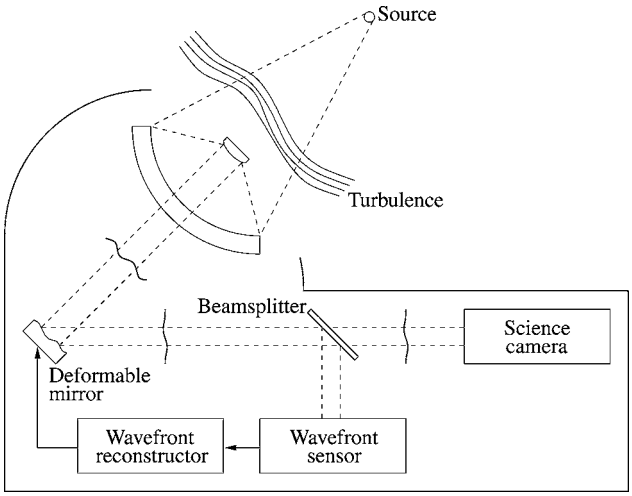


Fig. 1. Configuration of an adaptive optics telescope.

In [7], fourth in a series [15–17] of papers on objective assessment of image quality, the authors derived a very general formulation for the statistical properties of AO images, including the effects of measurement noise, randomness in the astronomical scene, and the turbulent atmosphere. It presented a fairly complete basic theory, but it stopped short of being a practical tool that one could use to assess and optimize the design of systems. A first attempt to provide a practical tool and, at the same time, overcome the heavy computational requirements for a full implementation of the spatiotemporal detector presented in [7] is to estimate the AO long-exposure PSF and apply such detection techniques to the long-exposure AO image. Considerable progress has been made in reconstructing the long-exposure PSF from recorded control-loop data [18], and this justifies this research on optimal-linear exoplanet detection when the long-exposure PSF is known. In addition, the quasi-static component of the PSF can be suppressed with a technique known as angular differential imaging [19]. In angular differential imaging, sequences of short-exposure images with an altitude/azimuth telescope are taken while the field derotator is turned off. Therefore, only the field of view (not the PSF) will rotate with time. The short-exposure images are then rotated to compensate for field of view rotation and added together to obtain a long-exposure image: Dim peaks due to companion objects get smeared, and the resulting long-exposure image represents the quasi-static component of the PSF. This component is then subtracted from the observed short-exposure images. It has been reported [19] that this technique allows quasi-static PSF distortion to be reduced by a factor of about 5.

4. HOTELLING AND IDEAL OBSERVERS IN ADAPTIVE OPTICS

The goal of a classification task is to assign the object that produced an image to one of two or more classes or hypotheses. A signal-detection task is a particular case of a classification task: In a signal-detection task, we have only two hypotheses that we call here signal absent and signal present. We will refer to these two hypotheses as H_0 and H_1, respectively. Given a data vector $\mathbf{g}$, the classification task performed by an observer (i.e., a binary classifier [1]) is made by computing a numerical value $t(\mathbf{g})$ and comparing it to a threshold τ. If $t(\mathbf{g}) \leqslant \tau$, then the observer will assign the object that generated $\mathbf{g}$ to the class H_0; if $t(\mathbf{g}) > \tau$, the observer will opt for H_1 [1].

The threshold τ controls the trade-offs between the probability of correctly classifying the signal as present and the probability of classifying the signal as present when it is in fact absent. In signal-detection applications, the first probability is usually called true-positive fraction (TPF), while the second is called false-positive fraction (FPF) [1]. The threshold τ parameterizes a curve on the TPF–FPF plane; such a curve is called the ROC curve. Meaningful figures of merit for a detection task include

the area under the ROC curve (AUC) and some detectability indexes such as d_A [1].

A detection task may also include a localization step: If $t(\mathbf{g}) > \tau$ and the signal location is not known, the observer will return an additional parameter that represents the estimated location of the signal. In this case, the TPF can be defined as the probability of correct classification of the image as signal present and localization of the signal within a given tolerance ε from the true location. A TPF–FPF plot for a detection and localization problem is called a LROC curve [1,20]. LROC curves can be seen as generalizations of estimation ROC (EROC) curves [21].

Statistical decision theory is concerned with how a decision-maker makes decisions and how such decisions can be made optimal in a well-defined sense. To our knowledge, statistical decision theory and methodology of objective assessment of image quality are just being applied to AO imagery. In a recent theoretical study, Barrett *et al.* [7] rigorously extended the theory of image-quality assessment to temporal sequences of AO images with random PSFs and derived general formulations and computational methods for detection and estimation tasks. The authors analyzed the structure of the data covariance matrix and showed how it can be rigorously decomposed into three terms representing the effect of the measurement noise (Gaussian for electronic readout noise and Poisson in the case of photon-counting statistics), the random PSF, and the random nature of the astronomical scene. Because AO systems deliver sequences of images, we will denote with $\mathbf{G}=\{\mathbf{g}^{(1)},\ldots,\mathbf{g}^{(J)}\}$ the complete data set from the science camera, with $\mathbf{F}=\{f^{(1)},\ldots,f^{(J)}\}$ the sequence of continuous-space objects, and, finally, with $\mathbf{P}=\{p^{(1)},\ldots,p^{(J)}\}$ the sequence of all continuous-space random PSFs. Formally, the mean image, given the object sequence $\mathbf{F}$ and the sequence $\mathbf{P}$ of PSFs, can be written as $\bar{\mathbf{G}}=\langle\mathbf{G}\rangle_{\mathbf{G}|\mathbf{P},\mathbf{F}}$, where the notation $\langle\ldots\rangle_{\mathbf{G}|\mathbf{P},\mathbf{F}}$ is understood as the average over noise for fixed $\mathbf{P}$ and $\mathbf{F}$. Likewise, $\bar{\bar{\mathbf{G}}}=\langle\bar{\mathbf{G}}\rangle_{\mathbf{P}|\mathbf{F}}$ is the mean image (averaged over noise and PSF randomness) given the object sequence $\mathbf{F}$ and, finally, $\bar{\bar{\bar{\mathbf{G}}}}=\langle\bar{\bar{\mathbf{G}}}\rangle_{\mathbf{F}}$ is the overall mean image averaged over noise, PSFs, and object variability. The data covariance matrix $\mathbf{K}_\mathbf{G}$ can then be written as [7]

$$\mathbf{K}_\mathbf{G}=\bar{\bar{\mathbf{K}}}_\mathbf{G}^{\text{noise}}+\bar{\mathbf{K}}_{\bar{\mathbf{G}}}^{\text{PSF}}+\mathbf{K}_{\bar{\bar{\mathbf{G}}}}^{\text{obj}},\tag{1}$$

where

$$\bar{\bar{\mathbf{K}}}_\mathbf{G}^{\text{noise}}=\langle\langle\langle[\mathbf{G}-\bar{\mathbf{G}}][\mathbf{G}-\bar{\mathbf{G}}]^{\mathrm{T}}\rangle_{\mathbf{G}|\mathbf{P},\mathbf{F}}\rangle_{\mathbf{P}|\mathbf{F}}\rangle_{\mathbf{F}},\tag{2}$$

$$\bar{\mathbf{K}}_{\bar{\mathbf{G}}}^{\text{PSF}}=\langle\langle[\bar{\mathbf{G}}-\bar{\bar{\mathbf{G}}}][\bar{\mathbf{G}}-\bar{\bar{\mathbf{G}}}]^{\mathrm{T}}\rangle_{\mathbf{P}|\mathbf{F}}\rangle_{\mathbf{F}},\tag{3}$$

$$\mathbf{K}_{\bar{\bar{\mathbf{G}}}}^{\text{obj}}=\langle[\bar{\bar{\mathbf{G}}}-\bar{\bar{\bar{\mathbf{G}}}}][\bar{\bar{\mathbf{G}}}-\bar{\bar{\bar{\mathbf{G}}}}]^{\mathrm{T}}\rangle_{\mathbf{F}},\tag{4}$$

in which $\bar{\bar{\mathbf{K}}}_\mathbf{G}^{\text{noise}}$ is the contribution to $\mathbf{K}_\mathbf{G}$ due to the noise, $\bar{\mathbf{K}}_{\bar{\mathbf{G}}}^{\text{PSF}}$ is the contribution due to the PSFs and, finally, $\mathbf{K}_{\bar{\bar{\mathbf{G}}}}^{\text{obj}}$ is the contribution due to the object being imaged. The decomposition in Eq. (1) is rigorous in the sense that it does not require any hypothesis of independence [7]. We remark, however, that in typical AO applications where the images can be large, say, 512×512 pixels, and a sequence $\mathbf{G}$ of noisy images can contain hundreds of such images, the size of $\mathbf{K}_\mathbf{G}$ can be of the order $10^7\times10^7$ or more. Handling such huge matrices is clearly a challenge [7]. In this paper, we avoid such difficulties by considering a preliminary simplified case. We will assume no object randomness (and so $\mathbf{K}_{\bar{\bar{\mathbf{G}}}}^{\text{obj}}$ is the zero matrix), and we will assume that $\mathbf{G}$ reduces to just a single long-exposure image $\mathbf{g}$ rather than a sequence $\{\mathbf{g}^{(1)},\ldots,\mathbf{g}^{(J)}\}$ of J short-exposure images (we will therefore denote the overall covariance matrix $\mathbf{K}_\mathbf{G}$ as $\mathbf{K}_\mathbf{g}$). Finally, the assumptions from the second part of this section will imply that $\bar{\mathbf{K}}_{\bar{\mathbf{G}}}^{\text{PSF}}$ is diagonal (or, at least, approximable with a diagonal matrix).

The problem we are concerned with in this paper is the detection of a planet located at $\mathbf{r}_\text{p}$ in focal plane coordinates. Generally speaking, a digital image is a 2-D array of M pixel values. Such data can be arranged in a column vector of size $M\times1$ by reading (for example, in a raster-scan fashion) all of the M values in the array and storing them in a column vector. This rearrangement makes it possible to work with images using no more than the usual matrix operations. If we denote with $\bar{\mathbf{g}}_0$ the mean (noise-free) image vector when the planet is not present, and with $\bar{\mathbf{g}}_{1,\mathbf{r}_\text{p}}$ the mean (noise-free) image vector when the planet is present and at location $\mathbf{r}_\text{p}$, we can express the signal we want to detect as $\mathbf{s}_{\mathbf{r}_\text{p}}=\bar{\mathbf{g}}_{1,\mathbf{r}_\text{p}}-\bar{\mathbf{g}}_0$. Furthermore, recalling the meaning of the hypotheses H_0 and H_1, we can model the noisy images under these hypotheses as

$$H_0:\mathbf{g}=\bar{\mathbf{g}}_0+\mathbf{n},\tag{5}$$

$$H_1,\mathbf{r}_\text{p}:\mathbf{g}=\underbrace{\bar{\mathbf{g}}_0+\mathbf{s}_{\mathbf{r}_\text{p}}}_{\bar{\mathbf{g}}_{1,\mathbf{r}_\text{p}}}+\mathbf{n},\tag{6}$$

where $\mathbf{n}$ denotes the randomly distributed zero-mean noise vector. The mean image vectors $\bar{\mathbf{g}}_0$ and $\bar{\mathbf{g}}_{1,\mathbf{r}_\text{p}}$ are the images without noise; such mean vectors can be obtained by averaging a large number of realizations of $\mathbf{g}$ under the hypotheses H_0 and H_1, respectively.

We will denote with $h_m(\mathbf{r})$ the mth pixel of the raster-scan ordering of the discretized (pixel-averaged) version of the long-exposure PSF centered at the focal plane location $\mathbf{r}$ (at which the long-exposure PSF assumes the value 1); A the postdetection intensity of the star (in electron units, e^- per pixel); a the postdetection intensity of the planet in e^-; b_m the background intensity in e^- per pixel; and, finally, σ_m^2 the variance of the readout noise at the mth detector pixel. If we denote $\mathbf{r}_\star$ the location of the star and $\mathbf{r}_\text{p}$ the location of the planet, then the mth pixel of the mean images under the two hypotheses can be written as

$$\bar{g}_{0,m}=Ah_m(\mathbf{r}_\star)+b_m,\tag{7}$$

$$\bar{g}_{1,m}=Ah_m(\mathbf{r}_\star)+ah_m(\mathbf{r}_\text{p})+b_m,\tag{8}$$

where, again, the $\bar{g}_{i,m}$ are measured in electron units. The noise in the image $\mathbf{g}$ is assumed to be readout Gaussian noise from the detector and Poisson photon noise from the star (plus, if present, the planet) and background.

Many authors have studied the statistical properties of ground-based AO-corrected images [22–26], and it was noted that ground-based AO observations are limited by

speckle noise, which is particularly problematic in exoplanet detection [22]. Indeed, one component of speckle noise, the lowest-order term leading to "speckle pinning" [23,24,27], can create bright speckles at specific locations that could look like exoplanets, leading to false positives. Thus, speckle noise, if not accounted for, reduces the performance of exoplanet detection at or near the diffraction limit. Many techniques are available to reduce speckle noise and increase the contrast. Simultaneous differential imaging (SDI) is a technique that can be used to suppress or reduce speckle noise. In SDI, pairs of images are taken simultaneously in two different and adjacent narrow bands for which the companion planet is expected to be very dim or absent in one and only one of such narrow bands. The speckle pattern will be the same in both images, but the contrast between a star and its companion planet (if present) will not be the same for the two bands. In such a case, subtracting one (scaled) image from the other will reduce the speckle while almost preserving the intensity of the science object (for example, a planet) present in the original images [28–32].

In this paper, we are concerned with single long-exposure images for which additional assumptions make it possible to derive an analytical expression for $\mathbf{K_g}$. In some applications, it could be reasonable to assume that the PSF is nonrandom and known by the observer. This is the case of space telescopes (such as the Hubble Space Telescope) or when the long-exposure PSF is reconstructed from control loop data [18]. For these cases, we can assume that the actual PSF is known. Another important case is the one in which speckle noise has been removed or reduced by applying one of the techniques described above.

Let $\bar{g}_{i,m}$ denote the mth component of the vector $\bar{\mathbf{g}}_i$. We note that the $\bar{g}_{i,m}$ are measured in electrons (e^-), and we may thus assume $\bar{g}_{i,m} \gg 1$ for all $m \in \{1, \ldots, M\}$. Therefore, by dint of the central limit theorem, the probability density functions under the two hypotheses [33] are

$$\mathrm{pr}(\mathbf{g}|H_0) = \left[\frac{1}{(2\pi)^M \det(\mathbf{K_g})}\right]^{1/2} \times \exp\left[-\frac{1}{2}(\mathbf{g}-\bar{\mathbf{g}}_0)^{\mathrm{T}}\mathbf{K_g}^{-1}(\mathbf{g}-\bar{\mathbf{g}}_0)\right], \tag{9}$$

$$\mathrm{pr}(\mathbf{g}|H_1,\mathbf{r}_\mathrm{p}) = \left[\frac{1}{(2\pi)^M \det(\mathbf{K_g})}\right]^{1/2} \times \exp\left[-\frac{1}{2}(\mathbf{g}-\bar{\mathbf{g}}_0-\mathbf{s}_{\mathbf{r}_\mathrm{p}})^{\mathrm{T}}\mathbf{K_g}^{-1}(\mathbf{g}-\bar{\mathbf{g}}_0-\mathbf{s}_{\mathbf{r}_\mathrm{p}})\right], \tag{10}$$

where $\mathbf{K_g}$ is the covariance matrix of the data vector $\mathbf{g}$. The matrix $\mathbf{K_g}$ is assumed to be the same under both hypotheses. This assumption is acceptable provided that the signal $\mathbf{s}_{\mathbf{r}_\mathrm{p}}$ is very weak compared to $\bar{\mathbf{g}}_0$. In a general case and if a sequence of short-exposure images is used, the covariance matrix $\mathbf{K_g}$ cannot be written analytically [7], and simulation of the whole AO system may be a possible way to estimate $\mathbf{K_g}$.

For the case of astronomy, the sky background can be estimated either from an image (or sequence of images) of a source-free field or by assuming that it is spatially constant and by applying a median filter to the actual image of the object. We will refer to this case as background known exactly (BKE). The signal can be assumed known in brightness and location for the pure detection problem (signal known exactly, or SKE) and only known in brightness for the detection and localization problem. If the brightness of the signal is not known, we will still talk about SKE because, as evident throughout the discussion, the optimality of the observers analyzed here does not depend on the value of the brightness of the signal. As an alternative, we can consider a hybrid detection and estimation task [1] (the brightness of the signal is the value being estimated) for which we will obtain the same optimal observers described here. Finally, as we noted, the PSF can be assumed known [18] for the long-exposure image, and we call this case PSF known exactly (PKE).

With these assumptions and the exact knowledge of the PSF, the matrix $\mathbf{K_g}$ is diagonal, and its (m,m')-th element can be written as

$$[\mathbf{K_g}]_{m,m'} = [Ah_m(\mathbf{r}_\star) + b_m + \sigma_m^2]\delta_{m,m'}, \tag{11}$$

where $\delta_{m,m'}$ is the Kronecker delta function. This expression for $\mathbf{K_g}$ ignores the small contribution to the covariance matrix due to the Poisson photon noise from the planet. We remark that the presence of the uncorrelated noise makes the matrix $\mathbf{K_g}$ invertible (because $\mathbf{K_g}$ is diagonal and $[\mathbf{K_g}]_{m,m} > 0$ for all m). If we substitute Eqs. (9) and (10) into the general expression of the likelihood ratio [1] and take the logarithm (a strictly increasing function) of $\Lambda(\mathbf{g}|\mathbf{r}_\mathrm{p})$, we obtain the log-likelihood ratio [1]

$$\lambda(\mathbf{g}|\mathbf{r}_\mathrm{p}) = \ln[\Lambda(\mathbf{g}|\mathbf{r}_\mathrm{p})] \tag{12}$$

$$= \frac{1}{2}(\mathbf{g}-\bar{\mathbf{g}}_0)^{\mathrm{T}}\mathbf{K_g}^{-1}(\mathbf{g}-\bar{\mathbf{g}}_0) - \frac{1}{2}(\mathbf{g}-\bar{\mathbf{g}}_0-\mathbf{s}_{\mathbf{r}_\mathrm{p}})^{\mathrm{T}}\mathbf{K_g}^{-1}(\mathbf{g}-\bar{\mathbf{g}}_0-\mathbf{s}_{\mathbf{r}_\mathrm{p}}). \tag{13}$$

Note that the matrix $\mathbf{K_g}$ is symmetric (because it is a covariance matrix) and so is its inverse $\mathbf{K_g}^{-1}$. Expanding the matrix products in Eq. (13) and simplifying, we can rewrite the log-likelihood ratio $\lambda(\mathbf{g}|\mathbf{r}_\mathrm{p})$ as

$$\lambda(\mathbf{g}|\mathbf{r}_\mathrm{p}) = -\frac{1}{2}[-\mathbf{g}^{\mathrm{T}}\mathbf{K_g}^{-1}\mathbf{s}_{\mathbf{r}_\mathrm{p}} + \bar{\mathbf{g}}_0{}^{\mathrm{T}}\mathbf{K_g}^{-1}\mathbf{s}_{\mathbf{r}_\mathrm{p}} - \mathbf{s}_{\mathbf{r}_\mathrm{p}}^{\mathrm{T}}\mathbf{K_g}^{-1}\mathbf{g} + \mathbf{s}_{\mathbf{r}_\mathrm{p}}^{\mathrm{T}}\mathbf{K_g}^{-1}\bar{\mathbf{g}}_0 + \mathbf{s}_{\mathbf{r}_\mathrm{p}}^{\mathrm{T}}\mathbf{K_g}^{-1}\mathbf{s}_{\mathbf{r}_\mathrm{p}}^{\mathrm{T}}] \tag{14}$$

$$= -\frac{1}{2}[2\mathbf{s}_{\mathbf{r}_\mathrm{p}}^{\mathrm{T}}\mathbf{K_g}^{-1}\bar{\mathbf{g}}_0 - 2\mathbf{s}_{\mathbf{r}_\mathrm{p}}^{\mathrm{T}}\mathbf{K_g}^{-1}\mathbf{g} + \mathbf{s}_{\mathbf{r}_\mathrm{p}}^{\mathrm{T}}\mathbf{K_g}^{-1}\mathbf{s}_{\mathbf{r}_\mathrm{p}}^{\mathrm{T}}]. \tag{15}$$

For a fixed $\mathbf{r}_\mathrm{p}$ and $\bar{\mathbf{g}}_0$ (as in the SKE case), the first and third terms that appear in Eq. (15) between brackets do not vary with $\mathbf{g}$; hence, they can be included in the constant against which $\lambda(\mathbf{g}|\mathbf{r}_\mathrm{p})$ will be compared. Therefore, the expression for the logarithm of the likelihood ratio when the densities are as in Eqs. (9) and (10) can be rewritten as

$$\lambda(\mathbf{g}|\mathbf{r}_p) = t_{\mathrm{Hot}}(\mathbf{g}|\mathbf{r}_p) = \mathbf{w}^T\mathbf{g}, \tag{16}$$

where

$$\mathbf{w} = \mathbf{K}_\mathbf{g}^{-1}\mathbf{s}_{\mathbf{r}_p} \tag{17}$$

is called a template vector [1]. The expression for $t_{\mathrm{Hot}}(\mathbf{g}|\mathbf{r}_p)$ in Eq. (16), which requires the knowledge of the mean vectors and covariance matrix under the two hypotheses, is the so-called Hotelling observer [1,7,15,34,35] and is a linear function of $\mathbf{g}$. As pointed out in [36], the Hotelling observer is still analytically tractable in realistic cases (for example, when the background can be described by a stationary random process), whereas computing the likelihood ratio in practical cases is, in general, a difficult problem. The derivation above shows that, if the data are normally distributed, the Hotelling observer is equivalent to the likelihood ratio, in the sense that they differ by an additive or positive multiplicative constant. Taking into account the particular structure of the matrix $\mathbf{K}_\mathbf{g}$ [see Eq. (11)] and observing that the components of $\mathbf{s}_{\mathbf{r}_p}$ are of the form $ah_m(\mathbf{r}_p)$, we obtain

$$t_{\mathrm{Hot}}(\mathbf{g}|\mathbf{r}_p) = \sum_{m=1}^{M} w_m g_m = \sum_{m=1}^{M} \frac{ah_m(\mathbf{r}_p)}{Ah_m(\mathbf{r}_\star) + b_m + \sigma_m^2} g_m. \tag{18}$$

The expression in Eq. (18) is ideal in the SKE/BKE/PKE cases when the location $\mathbf{r}_p$ of the planet is known. This expression shows that the knowledge of the exact value of a is not required because such a scaling factor can be absorbed into the threshold against which $t_{\mathrm{Hot}}(\mathbf{g}|\mathbf{r}_p)$ will be compared.

If the location of the planet is unknown, we can define a set T of test locations and introduce scanning observers. An observer can then be defined by introducing the costs $C_{i,j}$ of opting for the ith hypothesis when in fact the jth hypothesis is true. Actually, for detection with localization, the costs depend—in general—on the values of the true and estimated planet locations. An optimal detector can be defined as the one that minimizes the expected cost associated to the decision. If the expected cost is to be minimized, the optimal observer takes the form [21,37]

$$\Lambda(\mathbf{g}) = \max_{\mathbf{r}_p \in T} \Lambda(\mathbf{g}|\mathbf{r}_p) = \max_{\mathbf{r}_p \in T} \frac{\mathrm{pr}(\mathbf{g}|H_1, \mathbf{r}_p)}{\mathrm{pr}(\mathbf{g}|H_0)}, \tag{19}$$

which is called the generalized likelihood ratio [1]. If the observer concludes that a planet is present in the image, its estimated location $\hat{\mathbf{r}}_p$ is computed as

$$\hat{\mathbf{r}}_p = \arg\max_{\mathbf{r}_p \in T} \Lambda(\mathbf{g}|\mathbf{r}_p). \tag{20}$$

We still assume that the densities $\mathrm{pr}(\mathbf{g}|H_0)$ and $\mathrm{pr}(\mathbf{g}|H_1, \mathbf{r}_p)$ are as defined in Eqs. (9) and (10), respectively. Repeating the same process that led to Eq. (18), we can obtain expressions for the optimal observer and the estimated planet location when $\mathbf{r}_p$ is unknown. Substituting Eqs. (9) and (10) into Eq. (19) and taking the logarithm, we obtain

$$\lambda(\mathbf{g}) = \ln[\Lambda(\mathbf{g})] \tag{21}$$

$$= \max_{\mathbf{r}_p \in T} \left[\frac{1}{2}(\mathbf{g} - \bar{\mathbf{g}}_0)^T \mathbf{K}_\mathbf{g}^{-1}(\mathbf{g} - \bar{\mathbf{g}}_0) - \frac{1}{2}(\mathbf{g} - \bar{\mathbf{g}}_0 - \mathbf{s}_{\mathbf{r}_p})^T \mathbf{K}_\mathbf{g}^{-1}(\mathbf{g} - \bar{\mathbf{g}}_0 - \mathbf{s}_{\mathbf{r}_p}) \right]. \tag{22}$$

As before, the fact that $\mathbf{K}_\mathbf{g}$ is an invertible covariance matrix allows us to rewrite $\lambda(\mathbf{g})$ as

$$\lambda(\mathbf{g})\bar{g} = \max_{\mathbf{r}_p \in T} \left[-\frac{1}{2}(-\mathbf{g}^T\mathbf{K}_\mathbf{g}^{-1}\mathbf{s}_{\mathbf{r}_p} + \bar{\mathbf{g}}_0{}^T\mathbf{K}_\mathbf{g}^{-1}\mathbf{s}_{\mathbf{r}_p} - \mathbf{s}_{\mathbf{r}_p}^T\mathbf{K}_\mathbf{g}^{-1}\mathbf{g} + \mathbf{s}_{\mathbf{r}_p}^T\mathbf{K}_\mathbf{g}^{-1}\bar{\mathbf{g}}_0 + \mathbf{s}_{\mathbf{r}_p}^T\mathbf{K}_\mathbf{g}^{-1}\mathbf{s}_{\mathbf{r}_p}^T) \right] \tag{23}$$

$$= \max_{\mathbf{r}_p \in T} \left[-\frac{1}{2}(2\mathbf{s}_{\mathbf{r}_p}^T\mathbf{K}_\mathbf{g}^{-1}\bar{\mathbf{g}}_0 - 2\mathbf{s}_{\mathbf{r}_p}^T\mathbf{K}_\mathbf{g}^{-1}\mathbf{g} + \mathbf{s}_{\mathbf{r}_p}^T\mathbf{K}_\mathbf{g}^{-1}\mathbf{s}_{\mathbf{r}_p}^T) \right]. \tag{24}$$

This time, however, the quantities $\mathbf{s}_{\mathbf{r}_p}^T\mathbf{K}_\mathbf{g}^{-1}\bar{\mathbf{g}}_0$ and $\mathbf{s}_{\mathbf{r}_p}^T\mathbf{K}_\mathbf{g}^{-1}\mathbf{s}_{\mathbf{r}_p}^T$ vary as $\mathbf{r}_p$ is varied over T; they cannot be ignored because that will change the location $\hat{\mathbf{r}}_p$ of the maximum. The quantity that appears in Eq. (24) in square brackets is linear (or, more appropriately, affine) in $\mathbf{g}$; therefore, $\lambda(\mathbf{g})$ is the expression of the scanning Hotelling observer we were seeking. Again, the particular structure of the covariance matrix $\mathbf{K}_\mathbf{g}$ makes it possible to obtain an expression for the scanning Hotelling observer $t_{\mathrm{Hot}}(\mathbf{g})$ similar to Eq. (18):

$$t_{\mathrm{Hot}}(\mathbf{g}) = \max_{\mathbf{r}_p \in T} \sum_{m=1}^{M} \frac{ah_m(\mathbf{r}_p)}{Ah_m(\mathbf{r}_\star) + b_m + \sigma_m^2} \times \left[g_m - \bar{g}_{0,m} - \frac{1}{2}ah_m(\mathbf{r}_p) \right]; \tag{25}$$

also,

$$\hat{\mathbf{r}}_p = \arg\max_{\mathbf{r}_p \in T} \sum_{m=1}^{M} \frac{ah_m(\mathbf{r}_p)}{Ah_m(\mathbf{r}_\star) + b_m + \sigma_m^2} \times \left[g_m - \bar{g}_{0,m} - \frac{1}{2}ah_m(\mathbf{r}_p) \right]. \tag{26}$$

5. SIMULATION RESULTS

In this section, we report some simulation results for the detection at a known location and detection with location uncertainty. We adopt the area under the ROC/LROC curve as a figure of merit for performance assessment. Tests involve both simulated and real data.

A. ROC Curves for Simulated Data

In this test, we consider a detection problem for which the location of the planet (if present) is known. We use simulated data and assume that the atmospheric turbulence has been completely and perfectly corrected. This means that the mean image of the star (before sampling by the detector array) is an Airy disk. The location of the first zero of the Airy disk depends on the parameters (such as

wavelength, size of the aperture, focal length of the telescope, etc.) of the simulation. The images we simulated were of size 255×255 pixels. The Airy disk generated by the star was located at the center of the images, and its first zero was about 7.78 pixels away from the center of the image [see Fig. 2(a)]. The intensity (in electrons e^- per detector pixel) of the star was set in such a way that the brightness of the central pixel in Fig. 2(a) is $A=10^7\,e^-$. The constant sky background $b_m=10\,e^-$ was then added to the image. We simulated a mean image containing a planet by adding to the image $\bar{\mathbf{g}}_0$ of the star plus background a copy of the same image but scaled and shifted. In doing so, we assumed that the intensity of the pixels outside the region of support of the image was zero. The amount we shifted the image depended upon where we wanted the planet to be located. For this test, we assumed that the planet was 30 pixels to the right and 30 pixels down from the star. The shifted image was then scaled in such a way that the intensity of its brightest pixel was equal to the desired planet intensity $a=10\,e^-$. The image of the planet was then added to the image $\bar{\mathbf{g}}_0$; the resulting image $\bar{\mathbf{g}}_1$ is shown in Fig. 2(b).

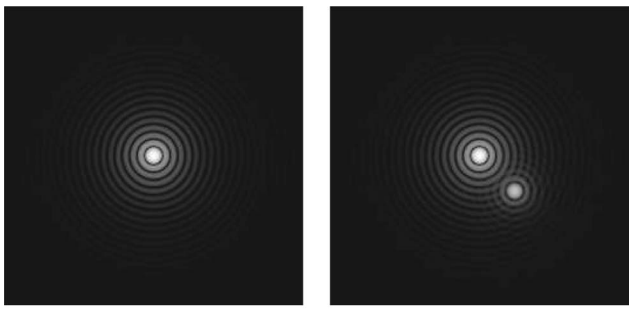
The detector readout noise variance was $\sigma_m^2=100$ for all $m\in\{1,\ldots,M\}$. We simulated 10,000 noisy images of the star without the planet and 10,000 noisy images with the planet present. In this test—and in those that follow—the photon noise for any mth pixel of either $\mathbf{g}_0$ or $\mathbf{g}_1$ followed a Poisson distribution with parameter $\bar{g}_{0,m}$ or $\bar{g}_{1,m}$, respectively. For each noisy image $\mathbf{g}$, we computed $t_{\mathrm{Hot}}(\mathbf{g})$ and stored the values in the $10{,}000\times1$ vector $\mathbf{t}_0$ (for the images where the planet was absent) and in $\mathbf{t}_1$ (for the images where the planet was present). By comparing the elements $t_{0,i}$ and $t_{1,i}$ to a threshold τ, the FPF and TPF as functions of τ were computed as

$$\mathrm{FPF}(\tau)=\frac{|\{i \text{ such that } t_{0,i}>\tau,\ i=1,\ldots,10{,}000\}|}{10{,}000}, \tag{27}$$

$$\mathrm{TPF}(\tau)=\frac{|\{i \text{ such that } t_{1,i}>\tau,\ i=1,\ldots,10{,}000\}|}{10{,}000}, \tag{28}$$

where the notation $|S|$ denotes the cardinality (number of members) of set S. The $\mathrm{FPF}(\tau)$ and $\mathrm{TPF}(\tau)$ were then plotted as an ROC curve.

For this first test, we decided to compare the Hotelling observer with another linear observer. Because we are assuming that the PSF, the background, and the location $\mathbf{r}_\star$ of the star are known, we may consider an observer that simply subtracts off the background and the image of the star and does pixel-based thresholding on the resulting image. More formally, if k is an odd integer number and the test location is $\mathbf{r}_\mathrm{p}$ (in focal plane coordinates), we can define the set $\mathcal{N}(\mathbf{r}_\mathrm{p},k)\subset\{1,\ldots,M\}$ that contains the pixel indices m of the $k\times k$ square neighborhood of the pixel that corresponds to $\mathbf{r}_\mathrm{p}$ in the image $\mathbf{g}$. Having defined $\mathcal{N}(\mathbf{r}_\mathrm{p},k)$, we can introduce the background-and-PSF-subtraction observer as follows:



(a) (b)

Fig. 2. Log-scale simulation of (a) image $\bar{\mathbf{g}}_0$ of the star and (b) image $\bar{\mathbf{g}}_1$ of the star and the planet. The intensity of the planet was set to $a=10^6\,e^-$ only in this figure and to make the planet location visible.

$$t_{\mathrm{sub}}^{(k)}(\mathbf{g}|\mathbf{r}_\mathrm{p})=\sum_{m\in\mathcal{N}(\mathbf{r}_\mathrm{p},k)}[g_m-b_m-Ah_m(\mathbf{r}_\star)]. \tag{29}$$

As we commented in Section 2, this way to operate—which gives rise to an affine observer—is one of the first steps for many state-of-the-art algorithms for point-source detection. Such algorithms then apply heuristic strategies which, to the authors' knowledge, have not been analyzed in ROC terms. We decided to compare the Hotelling observer against the background-and-PSF-subtraction observer defined in Eq. (29) because the latter embodies the general idea on which many detection methods used today are based. The notation used in Eq. (29) underlines the fact that the background-and-PSF-subtraction observer requires the additional parameter k and, as we shall see later, the performance of such an observer is greatly influenced by the value of k (we note, however, that in a practical case, an appropriate value of k can be estimated from assumptions on the relative size of the planet to the star, the apparent separation at which we are looking, and so on). More elaborate state-of-the-art algorithms for point-source detection may require dozens of parameters, which represents a major practical drawback of such algorithms: Detection performance depends on such parameters, and it may not always be easy to select the most appropriate values for a given image.

The Hotelling observer defined in Eq. (16) and the observer defined in Eq. (29) were run on the noisy images generated from $\bar{\mathbf{g}}_0$ [shown in Fig. 2(a)] and from $\bar{\mathbf{g}}_1$ [shown in Fig. 2(b)], and the corresponding values $t_{0,i}$ and $t_{1,i}$ for both observers were collected. From these values, the detectability d_A can be estimated as follows [1]:

$$d_A^{(\mathbf{t}_0,\mathbf{t}_1)}=\frac{\langle t_{1,i}\rangle_i-\langle t_{0,i}\rangle_i}{\sqrt{\frac{1}{2}\sigma_{t_{0,i}}^2+\frac{1}{2}\sigma_{t_{1,i}}^2}}, \tag{30}$$

where $\langle\ldots\rangle_i$ denotes ensemble average over the index i, and $\sigma_{t_{0,i}}^2$ and $\sigma_{t_{1,i}}^2$ are the variances of the $t_{0,i}$s and the $t_{1,i}$s, respectively. The detectability can also be expressed in terms of the AUC as [1]

$$d_A^{(\mathrm{AUC})}=2\,\mathrm{erf}^{-1}(2\,\mathrm{AUC}-1), \tag{31}$$

where erf^{-1} is the inverse of the error function $\mathrm{erf}(x)=2\pi^{-1/2}\int_0^x e^{-t^2}\,\mathrm{d}t$.

Table 1. Comparison between $t_{\mathrm{Hot}}(\mathbf{g}|\mathbf{r}_p)$ and $t^{(k)}_{\mathrm{sub}}(\mathbf{g}|\mathbf{r}_p)$ for Simulated Data and Known $\mathbf{r}_p$

Test Statistic		AUC	$d_A^{(\mathbf{t}_0,\mathbf{t}_1)}$	$d_A^{(\mathrm{AUC})}$
$t_{\mathrm{Hot}}(\mathbf{g}\|\mathbf{r}_p)$		0.894310	1.765742	1.767459
$t^{(k)}_{\mathrm{sub}}(\mathbf{g}\|\mathbf{r}_p)$	$k=1$	0.557286	0.203139	0.203775
	$k=3$	0.657921	0.574617	0.575298
	$k=5$	0.739889	0.913945	0.909342
	$k=7$	0.781343	1.099416	1.098470
	$k=9$	0.774432	1.070906	1.065644
	$k=11$	0.747979	0.945965	0.944895
	$k=13$	0.719856	0.824326	0.823657
	$k=15$	0.697693	0.732458	0.732248

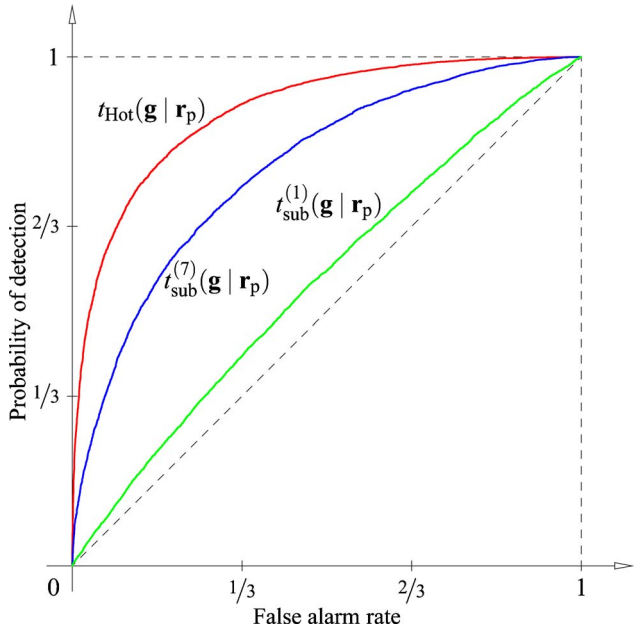


Fig. 3. (Color online) ROC curves for the simulated data.

Table 1 shows the values of the AUC, $d_A^{(\mathbf{t}_0,\mathbf{t}_1)}$, and $d_A^{(\mathrm{AUC})}$ for the test statistics $t_{\mathrm{Hot}}(\mathbf{g}|\mathbf{r}_p)$ and $t^{(k)}_{\mathrm{sub}}(\mathbf{g}|\mathbf{r}_p)$ as k is changed. Based on the values reported in Table 1, we selected the values $k=7$ and $k=1$ for the ROC curves reported in Fig. 3. The case $k=1$ is the single-pixel version of the background-and-PSF-subtraction observer. In this test, the ratio A/a was 10^6; if this ratio increases, we would obtain smaller values for the AUCs for the different methods, but their rank ordering should be the same. We also note that, for any test statistic, the values of $d_A^{(\mathbf{t}_0,\mathbf{t}_1)}$ and $d_A^{(\mathrm{AUC})}$ are very close to each other. That was expected, as they would be the same for normally distributed data [1].

B. LROC Curves for Simulated Data

In this test, we consider the performance of the Hotelling observer for detection with localization. We used again the image shown in Fig. 2(a) as the image of the star, and we set $A=10^7$ e$^-$. We simulated 100 planets located as shown in Fig. 4(a). The intensity of each planet was $a=20$ e$^-$. The background intensity and readout noise standard deviation were set to $b_m=100$ e$^-$ and $\sigma_m=10$ e$^-$, respectively.

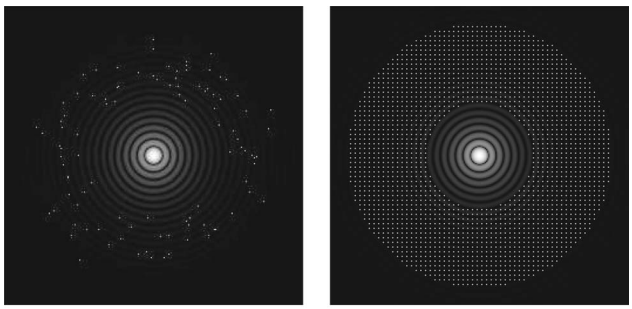


Fig. 4. Images used for the simulated data case and unknown $\mathbf{r}_p$: (a) locations of the planets for the test with the simulated image (Airy disk) and (b) test locations.

For each location, we simulated 100 noisy images where the planet was present at the desired random location and another 100 noisy images in which the planet was absent. The noisy images were generated in the same way as before. We used both the scanning Hotelling observer [see Eqs. (25) and (26)] and a scanning version of the background-and-PSF-subtraction observer

$$t^{(k)}_{\mathrm{sub}}(\mathbf{g}) = \max_{\mathbf{r}_p\in T} \sum_{m\in\mathcal{N}(\mathbf{r}_p,k)} [g_m - b_m - Ah_m(\mathbf{r}_\star)], \tag{32}$$

$$\hat{\mathbf{r}}_p = \arg\max_{\mathbf{r}_p\in T} \sum_{m\in\mathcal{N}(\mathbf{r}_p,k)} [g_m - b_m - Ah_m(\mathbf{r}_\star)]. \tag{33}$$

We set $k=7$ and $k=1$ for the background-and-PSF-subtraction scanning observer defined above. The set T of test locations is shown in Fig. 4(b): For each row or column, every other pixel location within a ring was in T. This means that the tolerance ε for the planet location we must use when we compute the TPF must be at least $\sqrt{2}$ in pixel units. In our simulation, we set $\varepsilon=1.45$ pixel units to avoid biased results due to numerical approximation.

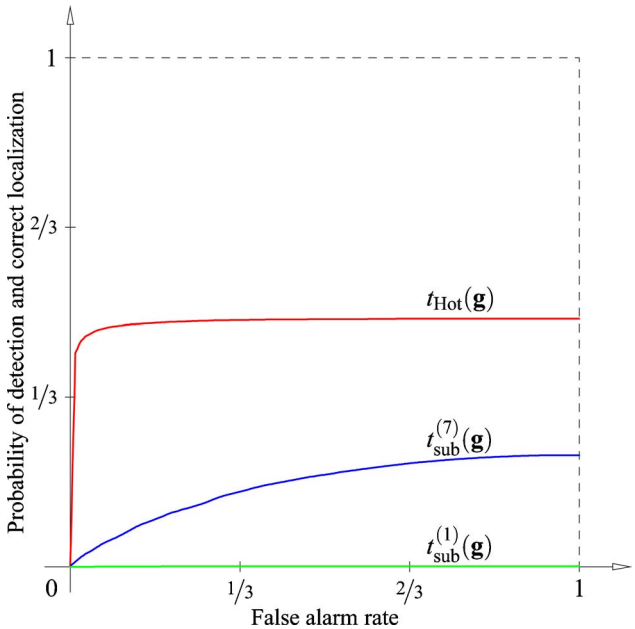


Fig. 5. (Color online) LROC curves for simulated data ($A=10^7$ e$^-$, $a=20$ e$^-$, $b_m=100$ e$^-$, $\sigma_m=10$ e$^-$).

Table 2. Comparison between $t_{\mathrm{Hot}}(\mathbf{g})$ and $t^{(k)}_{\mathrm{sub}}(\mathbf{g})$ for Simulated Data and Unknown $\mathbf{r}_p$

Test Statistic		ALROC
$t_{\mathrm{Hot}}(\mathbf{g})$		0.480074
$t^{(k)}_{\mathrm{sub}}(\mathbf{g})$	$k=1$	0.000667
	$k=7$	0.160270

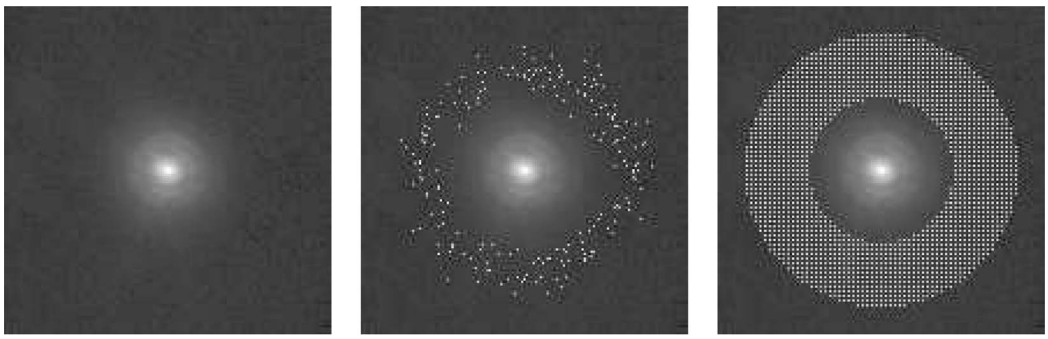
For each observer, we obtained 100 LROC curves—one for each true planet location shown in Fig. 4(a). These LROC curves were then averaged to obtain the mean LROC curve for any of the observers considered here. The mean LROC curves for the Hotelling scanning observer and the background-and-PSF-subtraction scanning observer are shown in Fig. 5, and the corresponding areas under the LROC curve (ALROC) are listed in Table 2. The LROC curve for the background-and-PSF-subtraction scanning observer when $k=1$ shows that, for this observer, the probability of detection and correct localization is very close to 0 at any false alarm rate (indeed, the LROC curve for $t^{(1)}_{\mathrm{sub}}(\mathbf{g})$ almost always overlaps the horizontal axis).

C. LROC Curves for Real Data

For this test, we used a real image of the star GJ 450. The image was generated by averaging 380 0.3596 s short-exposure images of GJ 450 taken in the L′ band (central wavelength $\lambda=3.809\ \mu$m; bandwidth $\Delta\lambda=0.623\ \mu$m). In order to reduce the background and subtract glints and other artifacts due to the telescope and/or the dome, the short-exposure images were obtained in nodded pairs. Moving the telescope periodically (nodding) and subtracting the images in each pair reduced the sky background in the short-exposure images at the expense of doubling the variance of the pixel intensities. The images were taken on the night of April 11, 2006 with the Clio 3–5 μm planet-finding AO camera [38] installed on the 6.5 m Multiple Mirror Telescope (MMT) [39] at the University of Arizona. It was reported [40] that the Strehl ratio for the MMT AO system can be as high as 80% in the L′ band. The pixel scale for these images was 0.048574 ± 0.000090 arcsec/pixel. The apparent magnitude (in the L′ band) of GJ 450 was determined from the raw counts using photometric calibration parameters and ranged from 5.27 to 5.29. Important parameters of the detector are as follows: approximate well depth 3.3×10^6 e$^-$, readout noise standard deviation 700 e$^-$, gain 87.6 e$^-$/DN (where DN stands for digital number), and linearity 1%. The background intensity was estimated from a set of 254 images of the same star with an exposure time of 2.0596 s. This longer exposure time resulted in the detector elements at the star location being always saturated. The long-exposure background was 16390 DN, which corresponds to about 250680 e$^-$ for the averaged image. This averaged image is the long-exposure image from which we obtained the image used in this test.

The central 141×141 pixel area [see Fig. 6(a)] of the original averaged 500×500 pixel image was used in this study; this was the image $\bar{\mathbf{g}}_0$ for the tests. We decided to use the central 141×141 pixel area of the original image to leave out of the field of view the artifacts due to nodding. The residual background was estimated by averaging the intensities of pixels more than 40 pixels away from the center of the star. This residual background was subtracted from the image before further processing: The resulting image is assumed to be the long-exposure residual halo. The image of each planet was simulated by scaling and shifting a copy of the image $\bar{\mathbf{g}}_0$ (padded with zeros outside its support region) and adding it to $\bar{\mathbf{g}}_0$ to obtain $\bar{\mathbf{g}}_1$. The intensity of the star was $A=3066000$ e$^-$, while the intensity of the planet was $a=1533$ e$^-$. The magnitude in the L′ band of the planet was about 13.53, while the magnitude of the star was about 5.28.

We simulated 200 planet locations distributed as shown in Fig. 6(b). Each of these planet locations had integer pixel coordinates. For each location, we simulated 100 noisy images where the planet was present at the desired random location and another 100 noisy images in which the planet was absent. We ran the scanning Hotelling observer [see Eqs. (25) and (26)] and the scanning version of the background-and-PSF-subtraction observer [see Eqs. (32) and (33)] on these images. This time, we decided to include in our tests one of the most commonly used algorithms for point-source detection. We opted for SExtractor [41,42], version 2.5.0. To the authors' knowledge, an ROC performance analysis of SExtractor has not been performed yet. We decided to use SExtractor because of its wide use for point-source detection, astrometry, and photometry. Implementations of this algorithm on different architectures have been developed, which has



(a) (b) (c)

Fig. 6. Images used for the real data case and unknown $\mathbf{r}_p$: (a) real image $\bar{\mathbf{g}}_0$ of the star (plotted on logarithmic scale), (b) locations of the planets, (c) test locations.

Table 3. SExtractor Parameters

Parameter	Value
CATALOG_TYPE	ASCII
DETECT_MINAREA	1
THRESH_TYPE	ABSOLUTE
FILTER_NAME	gauss_5.0_9× 9.conv
DEBLEND_NTHRESH	64
DEBLEND_MINCONT	0.00001
CLEAN	N
SATUR_LEVEL	40000
BACK_FILTERSIZE	10
VERBOSE_TYPE	QUIET

contributed to its popularity (see, for example, [43–49]). SExtractor is able to perform deblending of overlapping extended objects and can use a neural network to perform classification. Therefore, if correctly configured, SExtractor should perform well on separating and classifying one or more planets orbiting a star. SExtractor returns a list of objects it found in the image along with some information for each object. Every time SExtractor found exactly one object in the image, we assumed that no planet was found in the image. This is because the star is very bright, and SExtractor was always able to find and locate it. On the other hand, if SExtractor returned a list of two or more objects, we assumed that a planet was found, and its location was that of the second brightest object SExtractor found. Configuration parameters (stored in a file) are required for running SExtractor. We started with the default parameters for version 2.5.0 and changed them to try to obtain the best performance with respect to the area under the LROC curve. The values of the parameters that differ from the default are reported in Table 3. The parameters "DETECT_THRESH" and "ANALYSIS_THRESH" were changed to obtain the LROC curves. Despite our thorough parameter space search, SExtractor seemed to perform poorly on the task of interest.

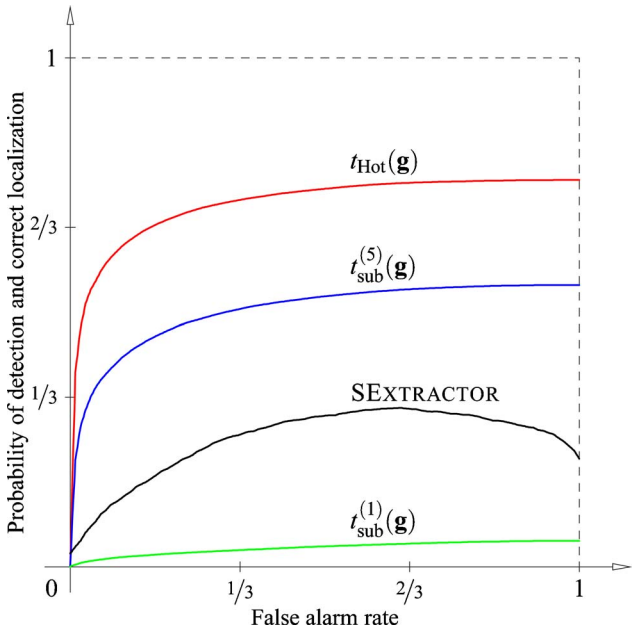


Fig. 7. (Color online) LROC curves for real data (A = 3,066,000 e$^-$, a = 1533 e$^-$, b_m = 250,680 e$^-$, σ_m = 700 e$^-$).

Table 4. Comparison between $t_{\mathrm{Hot}}(\mathbf{g})$, $t^{(k)}_{\mathrm{sub}}(\mathbf{g})$, and SExtractor for Real Data and Unknown $\mathbf{r_p}$

Test Statistic/Algorithm	ALROC
$t_{\mathrm{Hot}}(\mathbf{g})$	0.711847
$t^{(k)}_{\mathrm{sub}}(\mathbf{g})$ $k=1$	0.037027
$t^{(k)}_{\mathrm{sub}}(\mathbf{g})$ $k=5$	0.501799
SExtractor	0.248950

We set $k=5$ and $k=1$ for the background-and-PSF-subtraction scanning observer. The set T contained 1984 test locations distributed as shown in Fig. 6(c). For this simulation, we set the localization tolerance to $\varepsilon=1.45$ pixel units. For each observer, we obtained 200 LROC curves—one for each true planet location shown in Fig. 6(b). These LROC curves were then averaged to obtain the mean LROC curve for any of the observers considered here. The mean LROC curves for the Hotelling scanning observer, the background-and-PSF-subtraction scanning observer, and SExtractor are shown in Fig. 7. The corresponding values of the ALROC are listed in Table 4. We emphasize the fact that, for this image, the Hotelling observer markedly outperformed the other methods; in particular, SExtractor did not perform well, even if the planets we simulated had an apparent magnitude of about 13.53. If the apparent magnitude of the planet increases (i.e., the planet gets dimmer), the values of the ALROC decrease, but their rank ordering would still be the same, showing that the Hotelling observer still outperforms the other methods.

6. SUMMARY AND CONCLUSIONS

In this paper, we addressed the problem of objective assessment of image quality in AO, and we applied statistical decision theory to the problem of point detection in astronomy. Starting from the optimal (with respect to the area under the ROC/LROC curves) observers, we rigorously derived the Hotelling observers for both detection at a known location and detection with location uncertainty. For Gaussian noise, these observers are also optimal and are linear (or affine) functions of the data. We applied the Hotelling observers to long-exposure AO images for which we assumed that the atmospheric distortions were completely corrected by the AO system. Our derivations of the Hotelling observers from the optimal ones used the fact that, for large parameters, Poisson distributions can be approximated with Gaussian distributions. For this reason, the Hotelling observers derived here are optimal. A rigorous approach for assessment of image quality in AO stems from the application of these optimal observers to the problem of exoplanet detection.

We compared the Hotelling observers with other observers and algorithms, including one widely used in astronomy for point detection and photometry. Our tests, which we carried out on both simulated and real data, corroborate the theoretical superiority (with respect to the area under the ROC/LROC curves) of the Hotelling observers over the other methods. The promising results we obtained suggest that a similar approach could be

used to derive optimal estimators, such as the generalized Wiener estimator, for the estimation of flux in different wavelength ranges, or physical properties such as temperature, age, or mass. In turn, this will lead to a method for the assessment of image quality and task performance for estimation tasks.

The Hotelling observer can be applied to sequences of short-exposure AO images as well. Spatiotemporal information about the statistics of the random PSFs and random objects can be incorporated in the Hotelling observer as additional terms in the expression for the data covariance matrix. We commented that, for a sequence of AO images, the data covariance matrix can be written rigorously as the sum of three terms, referred to as the noise, PSF, and object terms. The data covariance matrix will no longer be diagonal, and the number of elements in it will increase by the square of the number of short-exposure images in each sequence. This will require efficient algorithms for the computation of the Hotelling test statistics. How well the Hotelling observer would perform in the case of short-exposure AO images is the topic of a future study.

Finally, we note that these observers are not limited to the case of AO systems but can also be applied to diffraction-limited images obtained from space.

ACKNOWLEDGMENTS

The authors would like to thank Michael Lloyd-Hart, Ari Heinze, Philip Hinz, Eric Clarkson, Christopher Dainty, Kyle Myers, Keith Hege, and Olivier Guyon for stimulating discussions and two reviewers for helpful comments on an earlier draft of the manuscript. The data for this research were supplied by Ari Heinze, Michael Lloyd-Hart, and Philip Hinz. This research was supported by Science Foundation Ireland under grant 01/PI.2/B039C and by SFI Walton Fellowship (03/W3/M420) for H. H. Barrett. Development of the basic methodology for objective assessment of image quality was also supported in part by the National Institutes of Health under grants R37 EB000803 and P41 EB002035.

REFERENCES

1. H. H. Barrett and K. J. Myers, *Foundations of Image Science* (Wiley-Interscience, 2004).
2. "The extrasolar planets encyclopaedia," http://exoplanet.eu/.
3. H. W. Babcock, "The possibility of compensating astronomical seeing," Publ. Astron. Soc. Pac. **65**, 229–236 (1953).
4. R. K. Tyson, *Principles of Adaptive Optics* (Academic, 1998).
5. B. A. Macintosh, J. Graham, L. Poyneer, G. Sommargren, J. Wilhelmsen, D. Gavel, S. Jones, P. Kalas, J. P. Lloyd, R. Makidon, S. Olivier, D. Palmer, J. Patience, M. Perrin, S. Severson, A. Sheinis, A. Sivaramakrishnan, M. Troy, and J. K. Wallace, "Extreme adaptive optics planet imager: XAOPI," Proc. SPIE **5170**, 272–282 (2003).
6. J. R. P. Angel, "Ground-based imaging of extrasolar planets using adaptive optics," Nature **368**, 203–207 (1994).
7. H. H. Barrett, K. J. Myers, N. Devaney, and J. C. Dainty, "Objective assessment of image quality. IV. Application to adaptive optics," J. Opt. Soc. Am. A **23**, 3080–3105 (2006).
8. E. Thiébaut and L. Mugnier, "Maximum *a posteriori* planet detection and characterization with a nulling interferometer," in *Direct Imaging of Exoplanets: Science and Techniques*, C. Aime and F. Vakili, eds. (Cambridge U. Press, 2006), pp. 547–552.
9. M. P. Hobson and C. McLachlan, "A Bayesian approach to discrete object detection in astronomical data sets," Mon. Not. R. Astron. Soc. **338**, 765–784 (2003).
10. M. Postman, L. M. Lubin, J. E. Gunn, J. B. Oke, J. G. Hoessel, D. P. Schneider, and J. A. Christensen, "The Palomar distant clusters survey. I. The cluster catalog," Astron. J. **111**, 615–641 (1996).
11. P. Schuecker and H. Böhringer, "Likelihood filter for cluster detection," Astron. Astrophys. **339**, 315–326 (1998).
12. C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. (Springer-Verlag, 2004).
13. W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice* (Chapman & Hall, 1998).
14. N. J. Kasdin and I. Braems, "Linear and Bayesian planet detection algorithms for the Terrestrial Planet Finder," Astrophys. J. **646**, 1260–1274 (2006).
15. H. H. Barrett, "Objective assessment of image quality: effects of quantum noise and object variability," J. Opt. Soc. Am. A **7**, 1266–1278 (1990).
16. H. H. Barrett, J. L. Denny, R. F. Wagner, and K. J. Myers, "Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance," J. Opt. Soc. Am. A **12**, 834–852 (1995).
17. H. H. Barrett, C. K. Abbey, and E. Clarkson, "Objective assessment of image quality. III. ROC metrics, ideal observers, and likelihood-generating functions," J. Opt. Soc. Am. A **15**, 1520–1535 (1998).
18. J.-P. Véran, F. Rigaut, H. Maître, and D. Rouan, "Estimation of the adaptive optics long-exposure point-spread function using control loop data," J. Opt. Soc. Am. A **14**, 3057–3069 (1997).
19. C. Marois, D. Lafreniére, R. Doyon, B. Macintosh, and D. Nadeau, "Angular differential imaging: a powerful high-contrast imaging technique," Astrophys. J. **641**, 556–564 (2006).
20. R. G. Swensson, "Unified measurement of observer performance in detecting and localizing target objects on images," Med. Phys. **23**, 1709–1725 (1996).
21. E. W. Clarkson, "Estimation ROC curves and their corresponding ideal observers," Proc. SPIE **6515**, 651504 (2007).
22. M. P. Fitzgerald and J. R. Graham, "Speckle statistics in adaptively corrected images," Astrophys. J. **637**, 541–547 (2006).
23. E. E. Bloemhof, R. G. Dekany, M. Troy, and B. R. Oppenheimer, "Behavior of remnant speckles in an adaptively corrected imaging system," Astrophys. J. Lett. **558**, 71–74 (2001).
24. E. E. Bloemhof, "Suppression of speckle noise by speckle pinning in adaptive optics," Astrophys. J. Lett. **582**, 59–62 (2003).
25. A. Sivaramakrishnan, J. P. Lloyd, P. E. Hodge, and B. A. Macintosh, "Speckle decorrelation and dynamic range in speckle noise-limited imaging," Astrophys. J. Lett. **581**, 59–62 (2002).
26. M. D. Perrin, A. Sivaramakrishnan, R. B. Makidon, B. R. Oppenheimer, and J. R. Graham, "The structure of high Strehl ratio point-spread functions," Astrophys. J. **596**, 702–712 (2003).
27. C. Aime and R. Soummer, "The usefulness and limits of coronagraphy in the presence of pinned speckles," Astrophys. J. Lett. **612**, 85–88 (2004).
28. R. Racine, G. A. H. Walker, D. Nadeau, R. Doyon, and C. Marois, "Speckle noise and the detection of faint companions," Publ. Astron. Soc. Pac. **111**, 587–594 (1999).
29. C. Marois, R. Doyon, R. Racine, and D. Nadeau, "Efficient speckle noise attenuation in faint companion imaging," Publ. Astron. Soc. Pac. **112**, 91–96 (2000).
30. C. Marois, R. Racine, R. Doyon, D. Lafrenière, and D. Nadeau, "Differential imaging with a multicolor detector assembly: a new exoplanet finder concept," Astrophys. J. Lett. **615**, 61–64 (2004).

31. O. Guyon, "Imaging faint sources within a speckle halo with synchronous interferometric speckle subtraction," Astrophys. J. **615**, 562–572 (2004).
32. R. Lenzen, L. Close, W. Brandner, B. Biller, and M. Hartung, "A novel simultaneous differential imager for the direct imaging of giant planets," Proc. SPIE **5492**, 970–977 (2004).
33. A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. (McGraw-Hill, 2002).
34. R. D. Fiete, H. H. Barrett, W. E. Smith, and K. J. Myers, "Hotelling trace criterion and its correlation with human-observer performance," J. Opt. Soc. Am. A **4**, 945–953 (1987).
35. W. E. Smith and H. H. Barrett, "Hotelling trace criterion as a figure of merit for the optimization of imaging systems," J. Opt. Soc. Am. A **3**, 717–725 (1986).
36. J. Yao and H. H. Barrett, "Predicting human performance by a channelized Hotelling observer model," Proc. SPIE **1768**, 161–168 (1992).
37. P. Khurd and G. Gindi, "Decision strategies that maximize the area under the LROC curve," IEEE Trans. Med. Imaging **24**, 1626–1636 (2005).
38. M. Freed, P. M. Hinz, M. R. Meyer, N. M. Milton, and M. Lloyd-Hart, "Clio: a 5-μm camera for the detection of giant exoplanets," Proc. SPIE **5492**, 1561–1571 (2004).
39. M. Lloyd-Hart, F. P. Wildi, B. Martin, P. C. McGuire, M. A. Kenworthy, R. L. Johnson, B. C. Fitz-Patrick, G. Z. Angeli, S. M. Miller, III, and J. R. P. Angel, "Adaptive optics for the 6.5-m MMT," Proc. SPIE **4007**, 167–174 (2000).
40. M. A. Kenworthy, P. M. Hinz, J. R. P. Angel, A. N. Heinze, and S. Sivanandam, "Whack-a-speckle: focal plane wavefront sensing in theory and practice with a deformable secondary mirror and 5-micron camera," Proc. SPIE **6272**, 1074–1080 (2006).
41. E. Bertin and S. Arnouts, "SExtractor: Software for source extraction," Astron. Astrophys., Suppl. Ser. **117**, 393–404 (1996).
42. E. Bertin, "Terapix–SExtractor," http://terapix.iap.fr/soft/sextractor.
43. K. Bundy, R. S. Ellis, C. J. Conselice, J. E. Taylor, M. C. Cooper, C. N. A. Willmer, B. J. Weiner, A. L. Coil, K. G. Noeske, and P. R. M. Eisenhardt, "The mass assembly history of field galaxies: detection of an evolving mass limit for star-forming galaxies," Astrophys. J. **651**, 120–141 (2006).
44. N. Caldwell, "Color-magnitude diagrams of resolved stars in Virgo cluster dwarf galaxies," Astrophys. J. **651**, 822–834 (2006).
45. B. E. Cobb, C. D. Bailyn, P. G. van Dokkum, and P. Natarajan, "Could GRB 060614 and its presumed host galaxy be a chance superposition?" Astrophys. J. Lett. **651**, L85–L88 (2006).
46. S. J. Schmidt, A. J. Connolly, and A. M. Hopkins, "The DRaGONS survey: a search for high-redshift radio galaxies and heavily obscured active galactic nuclei," Astrophys. J. **1**, 63–78 (2006).
47. T. Saito, K. Shimasaku, S. Okamura, M. Ouchi, M. Akiyama, and M. Yoshida, "Systematic survey of extended Lyα sources over $z \sim 3-5$," Astrophys. J. **648**, 54–66 (2006).
48. R. J. Foley, D. A. Perley, D. Pooley, J. X. Prochaska, J. S. Bloom, W. Li, B. Cobb, H.-W. Chen, G. Aldering, C. Bailyn, C. H. Blake, E. E. Falco, P. J. Green, M. P. Kowalski, S. Perlmutter, K. Roth, and K. Volk, "GRB 050408: a bright gamma-ray burst probing an atypical galactic environment," Astrophys. J. **645**, 450–463 (2006).
49. K. N. Allers, J. E. Kessler-Silacci, L. A. Cieza, and D. T. Jaffe, "Young, low-mass brown dwarfs with mid-infrared excesses," Astrophys. J. **644**, 364–377 (2006).