



Stochastic Speckle Discrimination with Time-tagged Photon Lists: Digging below the Speckle Noise Floor

Alex B. Walter¹, Clinton Bockstiegel, Timothy D. Brandt, and Benjamin A. Mazin

Department of Physics, University of California, Santa Barbara, Santa Barbara, CA 93106, USA

Received 2019 June 7; accepted 2019 August 2; published 2019 October 7

Abstract

We present an algorithm that uses the distribution of photon arrival times to distinguish speckles from incoherent sources, like planets and disks, in high-contrast images. Using simulated data, we show that our approach can overcome the noise limit from fluctuating speckle intensity. The algorithm is likely to be most advantageous when a coronagraph limits the coherent diffraction pattern in the image plane but the intensity is still strongly modulated by fast-timescale uncorrected stellar light, for example from atmospheric turbulence. These conditions are common at small inner working angles of highly corrected adaptive optics images and will allow probing of exoplanet populations at smaller angular separations. The technique requires a fast science camera that can temporally resolve the speckle fluctuations, and the detection of many photons per speckle decorrelation time. Because the algorithm directly extracts the incoherent light, standard differential imaging postprocessing techniques can be performed afterwards to further boost the signal.

Key words: methods: data analysis – methods: statistical – planets and satellites: detection – techniques: high angular resolution – techniques: image processing

Online material: color figures

1. Introduction

Direct imaging is a challenging exoplanet discovery and characterization technique due to the extreme contrast ($<10^{-4}$ for ground-based targets) and small angular separations ($\lesssim 1''$) between the planetary companion and its stellar host. Despite this, adaptive optics (AO) and coronagraphy have enabled the discovery of planets up to $\sim 10^6$ times fainter than their host stars (Marois et al. 2008; Lagrange et al. 2010; Kuzuhara et al. 2013; Macintosh et al. 2015; Keppler et al. 2018). Imaging an exoplanet requires subtracting the light of its host star in the form of the point-spread function (PSF). If this background were static and could be subtracted perfectly, exoplanet imaging would be limited only by the photon shot noise of the bright host star. Instead, high-contrast imaging is limited by uncontrolled scattered and diffracted light, which produces a coherent speckle halo in the image plane (Guyon 2005).

Fast atmospheric speckles average down over an observation, while slower, quasistatic speckles must be removed using postprocessing techniques. Angular differential imaging (ADI, Marois et al. 2006) exploits the rotation of the Earth, and hence the field-of-view of an altitude-azimuth telescope, to distinguish diffraction speckles from astrophysical sources. Spectral differential imaging (SDI, Racine et al. 1999; Marois et al. 2000; Sparks & Ford 2002) uses the scaling of diffraction speckles with wavelength. Since the initial development of ADI and SDI, a variety of postprocessing algorithms have

refined their approaches to dig deeper into the stellar PSF (e.g., Lafrenière et al. 2007; Soummer et al. 2012; Marois et al. 2014).

The time variability and chromaticity of quasistatic speckles limit the performance of ADI and SDI (Gerard et al. 2019). Both techniques also suffer at small separations where exoplanets are more likely to hide. The speckle spectral dispersion used by SDI is proportional to the separation: close to the star, it becomes smaller than the planet's PSF. For ADI, the arclength traced by the companion's sky rotation is proportional to the separation. Furthermore, the precision of the background estimate for PSF subtraction is limited by low counting statistics at small separations (Mawet et al. 2014). Even without these issues, the variability induced by speckle fluctuations can dominate the photon noise and be well above the shot noise expected from the total number of photons.

Stochastic Speckle Discrimination (SSD) is a postprocessing technique for planet detection designed to reduce the additional noise caused by speckle fluctuations by temporally resolving the intensity variations in the focal plane. It relies on the difference in shape between the on-axis and off-axis intensity distributions (Soummer & Ferrari 2007; Gladysz et al. 2008) or the difference between various moments or combinations thereof (Gladysz & Christou 2009; Gladysz et al. 2010) to enhance the signal of faint companions hidden in stellar light. The first algorithm (Gladysz & Christou 2008) compared the

empirical intensity distribution at the location of a suspected companion to a theoretical probability density function of a speckle. The difference between distributions was quantified and translated to a confidence level associated with the detection decision. The moment-based approach compared only the parameters of measured distributions (Gladysz & Christou 2009; Gladysz et al. 2010) or their combinations. Expressions involving these parameters are chosen so that stellar light is suppressed while light from a companion is amplified. SSD-like approaches have been shown to reduce speckle noise from even fast atmospheric speckles and improve the contrast limit for the detection of substellar companions (Frazin 2016; Meeker et al. 2018; Stangalini et al. 2018).

In this paper, we present an improved version of SSD to exploit noise-free photon-counting cameras like MEC, the Microwave Kinetic Inductance Detector (MKID) Exoplanet Camera (Meeker et al. 2018; Walter et al. 2018) on Subaru Telescope’s SCExAO instrument (Lozi et al. 2018). Our approach statistically distinguishes a combination of constant and speckle intensity (both ultimately from the bright star) from incoherent light (from a planet or disk). At small separations, where ADI and SDI are least effective and the scientific questions are most pressing (Mawet et al. 2012), we expect SSD to offer strong improvements in the limiting detection contrast. This paper describes the new photon-counting SSD technique and demonstrates its performance on simulated data.

We organize the paper as follows. In Section 2 we detail the simulation of photon lists to emulate the data expected from photon-counting cameras like MEC. Section 3 describes a formal extension of previous SSD analysis techniques with millisecond images using a maximum likelihood algorithm. We simulate the performance for various atmospheric conditions, planet brightnesses, and effective exposure times. Section 4 presents the new photon-counting SSD algorithm that estimates the incoherent light from a companion or disk directly from individual photon arrival timestamps. We demonstrate the algorithm on a simulated telescope image. We discuss the main results in Section 5, and conclude with Section 6.

2. Simulating Photon Arrival Times

2.1. Modeling the Stellar Speckle Intensity

The statistics governing the off-axis intensity distribution of coherent light with a partially developed speckle pattern have been studied at length. Originally derived by Goodman (1975) and verified experimentally by Cagigal & Canales (2001) and Fitzgerald & Graham (2006), the probability of getting an instantaneous intensity I given I_c and I_s is governed by the Modified Rician (MR) distribution, defined as

$$\rho_{\text{MR}}[I|I_c, I_s] = \frac{1}{I_s} \exp\left[-\frac{I+I_c}{I_s}\right] I_0\left[\frac{2\sqrt{I I_c}}{I_s}\right], \quad (1)$$

where $I_0[x]$ denotes the zero-order modified Bessel function of the first kind. The parameter I_c represents the intensity of the “constant” part of the diffraction pattern, i.e., the PSF of a star without the atmosphere, while I_s is the intensity of the seeing halo which manifests as a “speckle” pattern (Canales & Cagigal 1999; Aime & Soummer 2004; Soummer et al. 2007).

Typically, the AO system attempts to confine all the speckles into the constant diffraction pattern. A coronagraph can remove or transform this coherent portion out of the image plane. For all of the calculations in this paper we assume that the Strehl ratio, and by extension I_c and I_s , remain constant.

Speckle intensity is correlated temporally. In the limit of Kolmogorov atmospheric turbulence and frozen flow, the speckle decorrelation time can be thought of as the wind crossing time across the telescope pupil. In reality, the turbulence is not Kolmogorov, the atmospheric turbulence evolves, and there are often multiple turbulence layers with different wind speeds and directions (Roddier et al. 1982; Macintosh et al. 2005). Other high speed processes such as dome seeing, telescope vibrations from wind buffeting, or the AO loop itself can further complicate the speckle temporal power spectrum density (PSD) (Stangalini et al. 2016). This may result in a faster speckle decorrelation time, a temporal PSD described by multiple exponential timescales, or possibly a dependence on the position in the focal plane.

For the purposes of creating the simulated data in this paper, we characterize the speckle PSD as a simple exponential decay with a characteristic speckle lifetime of $\tau_s = 0.1$ s to roughly match empirical data in the near-infrared (Fitzgerald & Graham 2006; Goebel et al. 2018; Meeker et al. 2018). The speckle lifetime can change drastically depending on the atmospheric conditions but our results here can be qualitatively understood in those cases by scaling all parameters in time. A core condition for the SSD technique is for many photons to arrive in a single speckle decorrelation time: if this is satisfied, individual speckle fluctuations can be probed without relying on prior knowledge of the true speckle temporal PSD.

2.2. Modeling an Incoherent Source

For high levels of AO correction that minimize the variations in instantaneous Strehl the light from a companion can be approximated as a delta function with a Poisson intensity distribution (Gladysz et al. 2008). This works well for faint sources where only the core of the PSF is visible. By ignoring variations in the Strehl we simplify our model; incoherent Poisson sources are parameterized by an intensity I_p and can be readily injected into images with the relevant photons uniquely identified in simulation. Any incoherent Poisson sources will be represented by this term including binary companions, planetary companions, extended sources, dark current, and read noise. For the coherent stellar light, in the case that $I_s = 0$, the

MR distribution reduces to a Poisson distribution with intensity I_c and will be indistinguishable from I_p . In the demonstrations that follow, we identify I_p as an injected planetary companion.

The intensity fluctuations associated with speckles elevate the noise floor above the typical shot noise, i.e., it is harder to measure the planet's intensity when it is embedded in a boiling speckle field. If the speckle temporal information is marginalized over, then we show in the [Appendix](#) that the total variance of temporally correlated intensities obeying MR statistics combined with photon shot noise is

$$\sigma_{I_{\text{tot}}}^2 \approx \frac{2\tau_s(I_s^2 + 2I_c I_s) + I_c + I_s + I_p}{T_{\text{tot}}}. \quad (2)$$

The variance of the measured intensity is inversely proportional to the total integration time, T_{tot} , as expected. This is the long exposure ($t_{\text{exp}} \gg \tau_s$) photon noise limit faced by all PSF subtraction techniques like ADI and SDI. In the limit that $2\tau_s I_s \ll 1$, this variance reduces to pure shot noise on the number of photons. For high-contrast imaging, typical parameters might be $\tau_s = 0.1$ s, $I_s = 50$ s⁻¹, and $2\tau_s I_s = 10$, in which case the noise from speckle fluctuations will dominate. In Section 4, we show that this noise can be overcome by temporally resolving individual fluctuations in the speckle background.

2.3. Generating Mock Photon Lists

We have developed code¹ for quickly generating mock photon lists with an optional injected planet, corresponding to the output of a single MKID-like pixel which is single photon counting with low noise. The photon lists obey the following rules:

1. The underlying intensities are MR distributed, but are correlated in time with $\langle I(t)I(t + \delta t) - \langle I \rangle^2 \rangle \propto \exp[-\delta t/\tau_s]$.
2. I_c , I_s , and I_p are independently specified by the user, such that I_c and I_s govern the MR statistics of the stellar intensity and I_p is the mean count rate of a Poisson source. The total intensity should have an expectation value of $\langle I \rangle = I_c + I_s + I_p$.
3. Due to the intrinsic dead time in an MKID, photons are removed from the list if they arrive within τ_0 of the previous valid photon's arrival time.

The procedure for generating photon timestamps begins with creating a correlated list of random numbers that follow a Gaussian distribution. The random numbers are transformed to a uniform distribution ranging from 0 to 1, and finally transformed again to a MR distribution. The correlated MR sequence defines the intensity in photon counts for a small

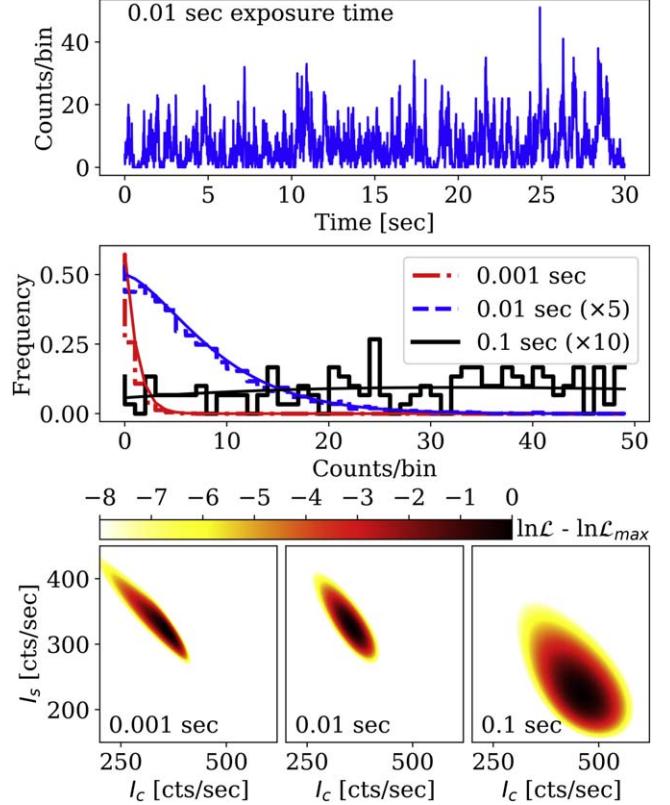


Figure 1. Top: a 30 s mock photon list simulated using the parameters $I_c = 300$, $I_s = 300$, $I_p = 0$ photons/s with $\tau_s = 0.1$ s, is binned into 0.01 s exposures to form a light curve. Middle: The same photon list is binned by three different exposure times and used to plot intensity histograms. The best-fit modified Rician functions are overplotted. The distribution changes shape as the bin size is varied; the corresponding fitted parameters also change. The histograms for 0.01 and 0.1 s exposure times have been scaled by factors of 5 and 10 respectively. Bottom: the likelihood is marginalized over I_p for the three different exposure times to illustrate how the best-fit parameters evolve with bin size. The darkest area in the plot represents the maximum likelihood and determines the best-fit values of I_c and I_s . The same photon list was used for all plots.

(A color version of this figure is available in the online journal.)

($\ll \tau_s$) time bin (we used 200 μ s), and a Poisson draw on that “instantaneous” intensity determines the number of photons that will finally be placed into that bin. The photons are distributed according to a uniform distribution in each bin. For simplicity, we assume neighboring pixels to have uncorrelated photon lists.

Since MKIDs do not have the same dark current or readout noise as conventional semiconductor detectors we do not add any additional noise. However, during high count rates photons can be lost due to a firmware triggering lockout that acts as a non-paralyzable dead time (van Eyken et al. 2015). We use $\tau_0 = 10 \mu$ s to match the latest firmware implemented on MEC. This dead time formulation could be used with quasi-photon-counting EMCCDs to account for photon pile-up.

¹ Part of the MKID Pipeline python package available at <https://github.com/MazinLab/MKIDPipeline>.

3. SSD with Short Exposure Images

Past efforts have been successful in detecting the position of (real or synthetic) faint companions in high-contrast images. Gladysz & Christou (2008) and Meeker et al. (2018) fit histograms of light curves (see top/middle of Figure 1) and discriminate planets by extracting parameters of hypothesized distributions using e.g., the method of moments and comparing those parameters, or their combinations, to the rest of the field. For example, a large I_c/I_s ratio can be used as a merit function because faint companions can masquerade as static speckles. Stangalini et al. (2018) use a discriminator from a direct recurrence quantification analysis of the light curves to search for companions. Gladysz & Christou (2009), Gladysz et al. (2010) also fit histograms of the light curves but use a deconvolution of the modeled probability distribution functions to further extract the companion’s photometry.

In this section we extend previous work by extracting the position and photometry of astrophysical sources (point or extended sources) using a maximum likelihood algorithm.

3.1. Maximum Likelihood Model for Discrete Light Curves

We model the intensity in the focal plane with three parameters I_c , I_s , and I_p and calculate their most likely values with an algorithm operating on the light curve. This allows for a direct detection of the non-stellar intensity, I_p , in the form of a point-source like a planet or an extended source like a protoplanetary disk. Furthermore, we can naturally introduce photon noise from a low-intensity photon-counting regime. By allowing the MR distribution in Equation (1) to suffer a Poisson-Mandel transformation (Cagigal & Canales 1999; Aime & Soummer 2004), the discrete stellar intensity distribution becomes

$$\begin{aligned} p_*[n|I_c, I_s] &= \int_0^\infty \frac{I^n}{n!} \exp[-I] \rho_{\text{MR}}[I] dI \\ &= \frac{1}{I_s + 1} \left(1 + \frac{1}{I_s}\right)^{-n} \exp\left[-\frac{I_c}{I_s}\right] \\ &\quad \times L_n\left[\frac{-I_c}{I_s^2 + I_s}\right] \exp\left[\frac{I_c}{I_s^2 + I_s}\right], \end{aligned} \quad (3)$$

where L_n is the n th Laguerre polynomial. The units of intensity for this section are number of photons per exposure time, t_{exp} . If a planet (or some other source that is incoherent with the star) exists in the field, the intensity distribution at that location will be the discrete convolution of p_* with the planet’s independent probability distribution p_p . To simplify, we assume the incoherent source has a Poisson probability distribution, $p_p[m|I_p] = \exp[-I_p] I_p^m / m!$ with average intensity I_p . The likelihood of the i th bin of a light curve containing k photons

given I_c , I_s , and I_p is

$$\mathcal{L}_i = \sum_{m=0}^{k_i} p_p[m|I_p] p_*[k_i - m|I_c, I_s], \quad (4)$$

which becomes

$$\begin{aligned} \mathcal{L}_i &= \frac{1}{I_s + 1} \exp\left[-I_p - \frac{I_c I_s}{I_s^2 + I_s}\right] \\ &\quad \times \sum_{m=0}^{k_i} \frac{I_p^m}{m!} \left(1 + \frac{1}{I_s}\right)^{-(k_i-m)} L_{k_i-m}\left[\frac{-I_c}{I_s^2 + I_s}\right]. \end{aligned} \quad (5)$$

The likelihood of the entire light curve is $\mathcal{L} = \prod_i \mathcal{L}_i$. We use a L-BFGS-B optimizer to find the most likely values of I_c , I_s , and I_p .

The maximum likelihood estimate of the intensity distribution is overplotted onto the light-curve histograms for three different exposure times in the middle panel of Figure 1. The bottom panel shows the likelihood functions marginalized over I_p for these three exposure times.

3.2. Performance of Millisecond Imaging SSD

To understand the performance of this millisecond imaging SSD algorithm and the extent to which it can improve exoplanet detections, we produce receiver operator characteristic (ROC) curves (Tanner & Swets 1954; DeLong et al. 1988; Krzanowski & Hand 2009; Jensen-Clem et al. 2017). We achieve this by generating an ensemble of mock photon lists with the same nominal values of I_c , I_s , and $I_p = 0$ and calculating maximum likelihood estimates (MLE) to build up a distribution of I_p corresponding to the signal-absent hypothesis. Next we inject a planet with $I_p > 0$ and calculate MLEs to build a distribution for I_p corresponding to the signal-present hypothesis. These distributions are shown in the left column of Figure 2. By choosing a detection threshold, we can calculate a true-positive rate and false-positive rate from the signal-present distribution and signal-absent distribution respectively. The ROC curves in the right column of Figure 2 are generated by varying the detection thresholds.

The right column of Figure 2 demonstrates the performance of the millisecond imaging SSD algorithm for different exposure times. The shape of the discrete intensity distribution can change depending on exposure time, which systematically affects the resulting MLEs for I_c , I_s , and I_p along with their uncertainties. In some cases there is a significant probability for the most likely value of I_p to be equal to zero. In the case that $I_c \gg I_s$, the MLE’s ensemble probability can be multimodal with a significant peak at MLE $I_p \approx \text{True } I_c + \text{True } I_p$. Both these behaviors appear for the same reason: when there is little modulation by I_s of photons associated with I_c , then I_c becomes difficult to distinguish from I_p . However, the total flux is still accurately recovered.

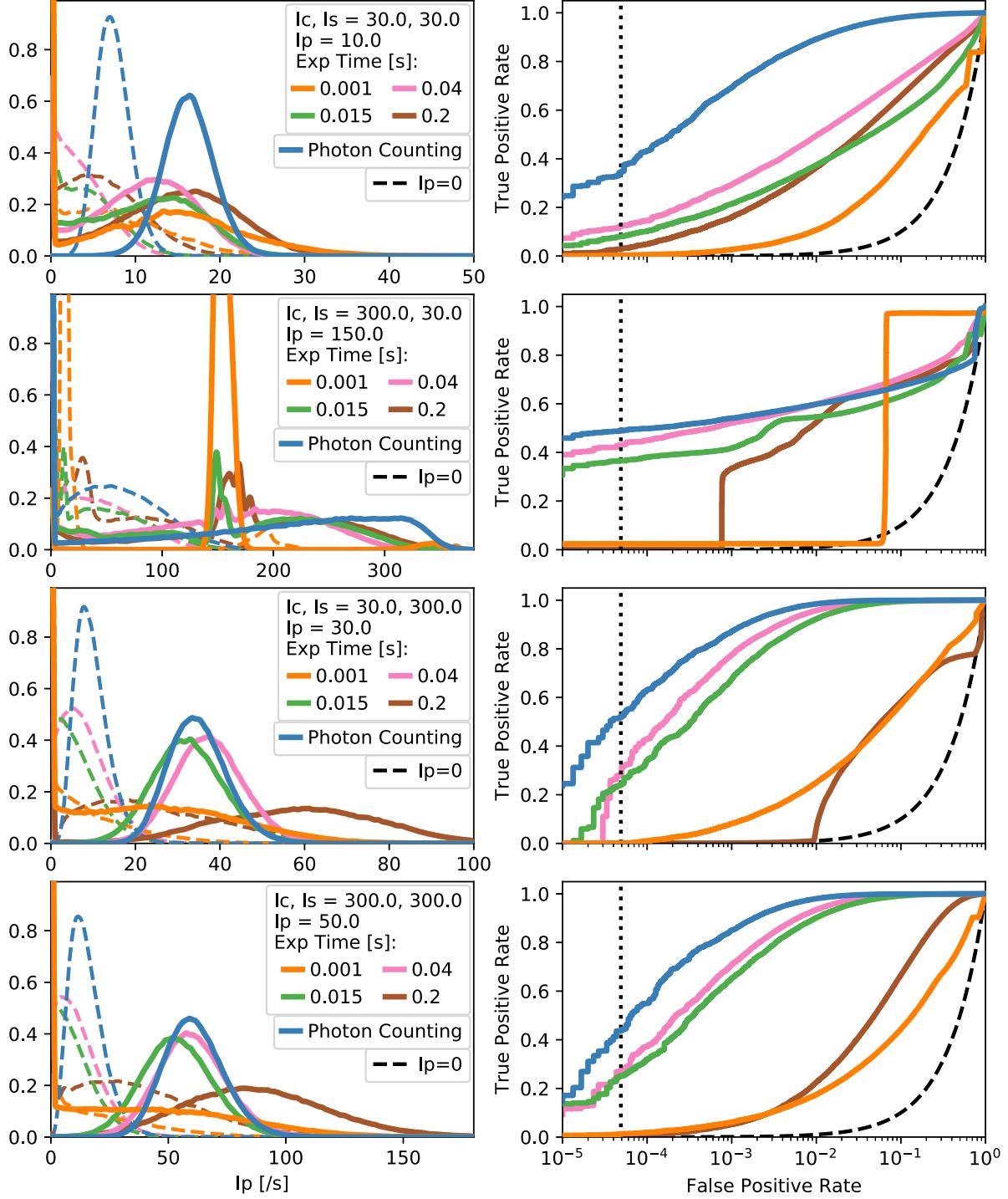


Figure 2. Performance of our millisecond imaging SSD algorithm (Section 3) compared to our photon-counting SSD algorithm (Section 4). Left panels: histograms of the maximum likelihood estimates of I_p , computed using $3 \cdot 10^5$ 30 s mock photon lists. Maximum likelihood estimates for I_p are calculated for various effective exposure times and for the case with (solid line) and without ($I_p = 0$, dashed line) an injected planet. The y-axis in the left column is arbitrary. The probability distributions are used to calculate the true/false-positive rates for the receiver operator characteristic (ROC) curve (right). The vertical dotted line at $1/20,000$ (for the 20,000 pixels in MEC) roughly indicates the maximum tolerable false-positive rate. The full photon-counting SSD algorithm (blue lines) described in Section 4 outperforms the cases with nonzero exposure times.

(A color version of this figure is available in the online journal.)

We find that there exists an optimal camera frame rate that leads to the most precise probability distribution for I_c , I_s , and I_p . For exposure times that are too short ($t_{\text{exp}} \ll \delta t$, the photon inter-arrival time) the correlation between subsequent photon arrivals is lost because frames are interpreted as an unordered set containing only 0 or 1 photons. For exposure times that are too long ($t_{\text{exp}} \gg \tau_s$) the speckle temporal information is averaged over. This tends to artificially decrease the I_s component of the speckle intensity directly in exchange for increasing the I_c component (see Figure 1 bottom). For those that use I_c/I_s as a merit function this means the part of the image without a planet will be polluted making it harder to discriminate the location of a companion.

4. SSD in the Photon-counting Regime

4.1. Maximum Likelihood Model for Photon Arrival Times

In the previous section the most precise extraction of the model parameters I_c , I_s , and I_p required the use of an optimal exposure time. Since this optimal exposure time is dependent on the speckle intensity as well as the speckle decorrelation time, the optimal exposure time will vary across the image. Conversely, if a single exposure time is chosen there may be a systematic loss of precision across some portions of the image. In order to avoid this pitfall, we develop the posterior probability for I_c , I_s , and I_p directly from the set of photon inter-arrival times.

We start by considering the (normalized) probability density for the next inter-photon arrival interval, δt , given a fixed intensity I :

$$p[\delta t|I] = I e^{-I\delta t}. \quad (6)$$

The speckle field intensity is not fixed but varies in time with the MR probability density described in Equation (1). At a fixed point in time, the probability density of the next photon arrival time given I_c and I_s becomes

$$p[\delta t|I_c, I_s] = \int_0^\infty p[\delta t|I] \rho_{\text{MR}}[I|I_c, I_s] dI, \quad (7)$$

where we have integrated the Poisson probability density from Equation (6) over all possible instantaneous stellar intensities, I . Intensities are considered to be in units of photons per second. We assume here that the speckle lifetime is much longer than the time it takes a photon to arrive, $\tau_s \gg \delta t$.

With a set of photon inter-arrival times, $\{\delta t_i\}$, we want the relative probability that a δt is realized. At moments that happen to have higher instantaneous intensities the number of short δt_i 's will be increased. Thus, the relative probability of a photon inter-arrival time being in our data becomes

$$p[\delta t|I_c, I_s] \propto \int_0^\infty p[\delta t|I] \rho_{\text{MR}}[I|I_c, I_s] dI. \quad (8)$$

Finally, we consider the case that light incoherent with the star, such as from a planet, is in the field. We consider only the simplest case in which this source has constant intensity I_p and is governed by Poisson statistics. In this case, the relative probability of δt given I_c , I_s , I_p is

$$p[\delta t|I_c, I_s, I_p] \propto \int_0^\infty p[\delta t|I + I_p] \rho_{\text{MR}}[I|I_c, I_s] (I + I_p) dI, \quad (9)$$

which can be evaluated analytically. We then find the normalization constant by setting

$$\int_{\tau_0}^\infty p[\tau|I_c, I_s, I_p] d\tau = 1. \quad (10)$$

Equation (10) accounts for the non-paralyzable detector dead time, τ_0 , intrinsic to MKIDs by replacing the lower limit of integration with τ_0 . In Equation (9), we set the likelihood equal to zero for $\delta t < \tau_0$. It is convenient to use the change of variables,

$$u_i = \frac{1}{1 + I_s \delta t_i}, \quad u_{\text{max}} = \frac{1}{1 + I_s \tau_0}. \quad (11)$$

The log likelihood, $\log(\mathcal{L})[I_c, I_s, I_p]$, then becomes

$$\begin{aligned} \log(\mathcal{L}) &= \sum_{i=1}^N p[\delta t_i|I_c, I_s, I_p] \\ &= \sum_{i=1}^N (u_i - 1) \left(\frac{I_p}{I_s u_i} + \frac{I_c}{I_s} \right) \\ &\quad + \sum_{i=1}^N \log [I_c^2 u_i^5 + 4I_c I_s u_i^4 + (2I_s^2 + 2I_p I_c) u_i^3 \\ &\quad + 2I_p I_s u_i^2 + I_p^2 u_i] \\ &\quad - N \frac{(u_{\text{max}} - 1)(I_p + I_c u_{\text{max}})}{I_s u_{\text{max}}} \\ &\quad - N \log [I_c u_{\text{max}}^3 + I_s u_{\text{max}}^2 + I_p u_{\text{max}}]. \end{aligned} \quad (12)$$

As in Section 3, we assume that AO performance remains stable so that I_c and I_s remain constant over the course of observations.

Equation (12) for the photon-counting SSD algorithm replaces Equation (5) from Section 3. We use a Newton conjugate-gradient search to find the maximum of the log-likelihood space and recover the best estimates for I_c , I_s , and I_p .

The photon-counting SSD algorithm consistently outperforms the millisecond imaging SSD algorithm from Section 3, which marginalizes over temporal information via the exposure time (see Figure 2). In Figure 3 we show the performance of the photon-counting SSD algorithm under conditions given by various combinations of I_c and I_s , and with various planet brightnesses I_p . The algorithm performs well with a high true positive detection rate even for the case when I_c and I_s are both large. However, the performance suffers in the case of $I_c \gg I_s$.

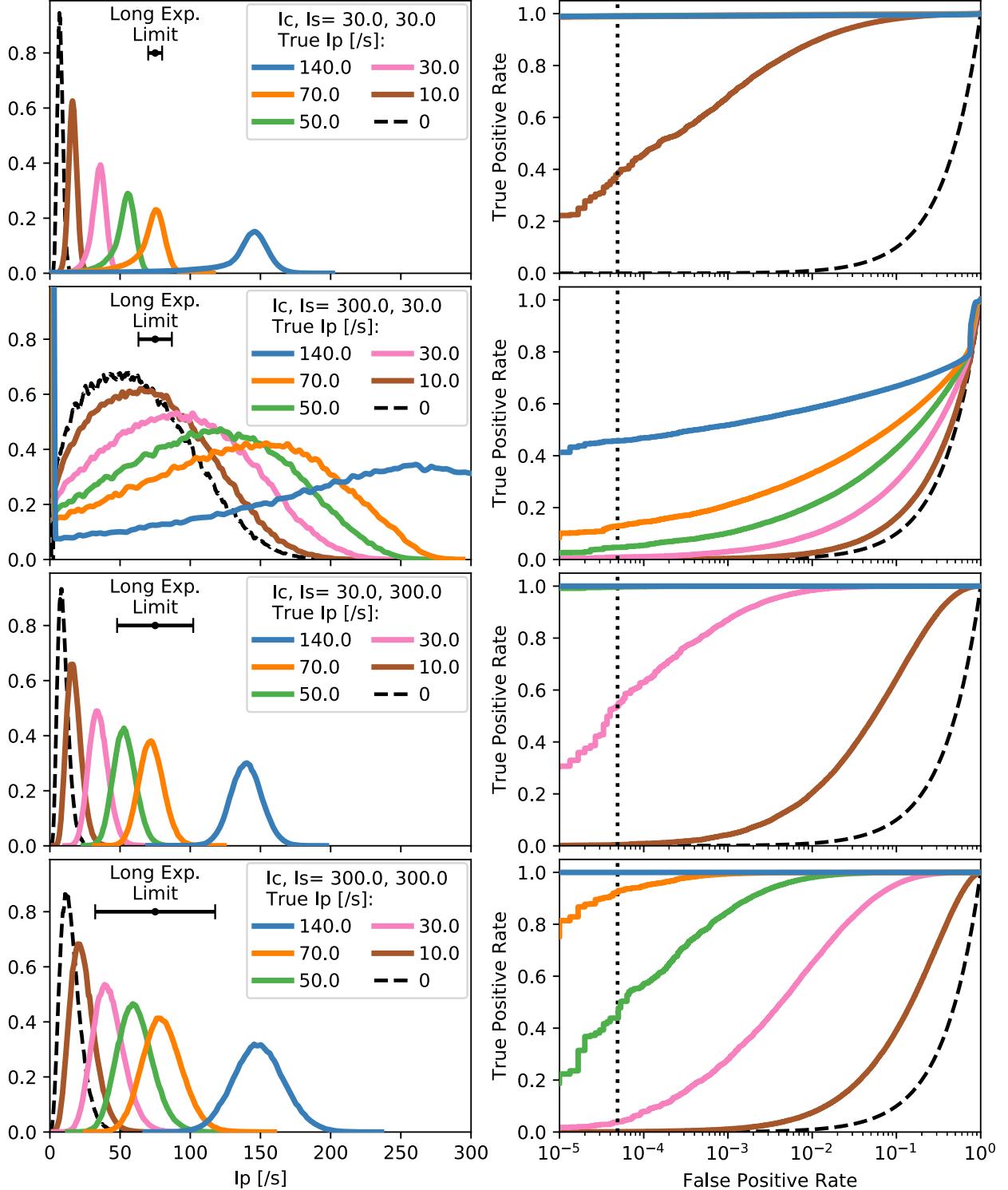


Figure 3. Performance of our photon-counting SSD algorithm. Left panels: histograms of the maximum likelihood estimates of I_p , computed using $3 \cdot 10^5$ 30 s mock photon lists for each set of parameters. The $\pm\sigma$ for the long exposure photon noise limit (Equation (2)) is shown with an error bar. The y-axis in the left column is arbitrary. The probability distributions are used to calculate the true/false-positive rates for the receiver operator characteristic (ROC) curve (right panels). The vertical dotted line at 1/20,000 (for the 20,000 pixels in MEC) roughly indicates the maximum tolerable false-positive rate. The algorithm performs well with a high true positive detection rate even for the case when I_c, I_s are both large. However, the performance suffers in the case of $I_c \gg I_s$ (row 2).

(A color version of this figure is available in the online journal.)

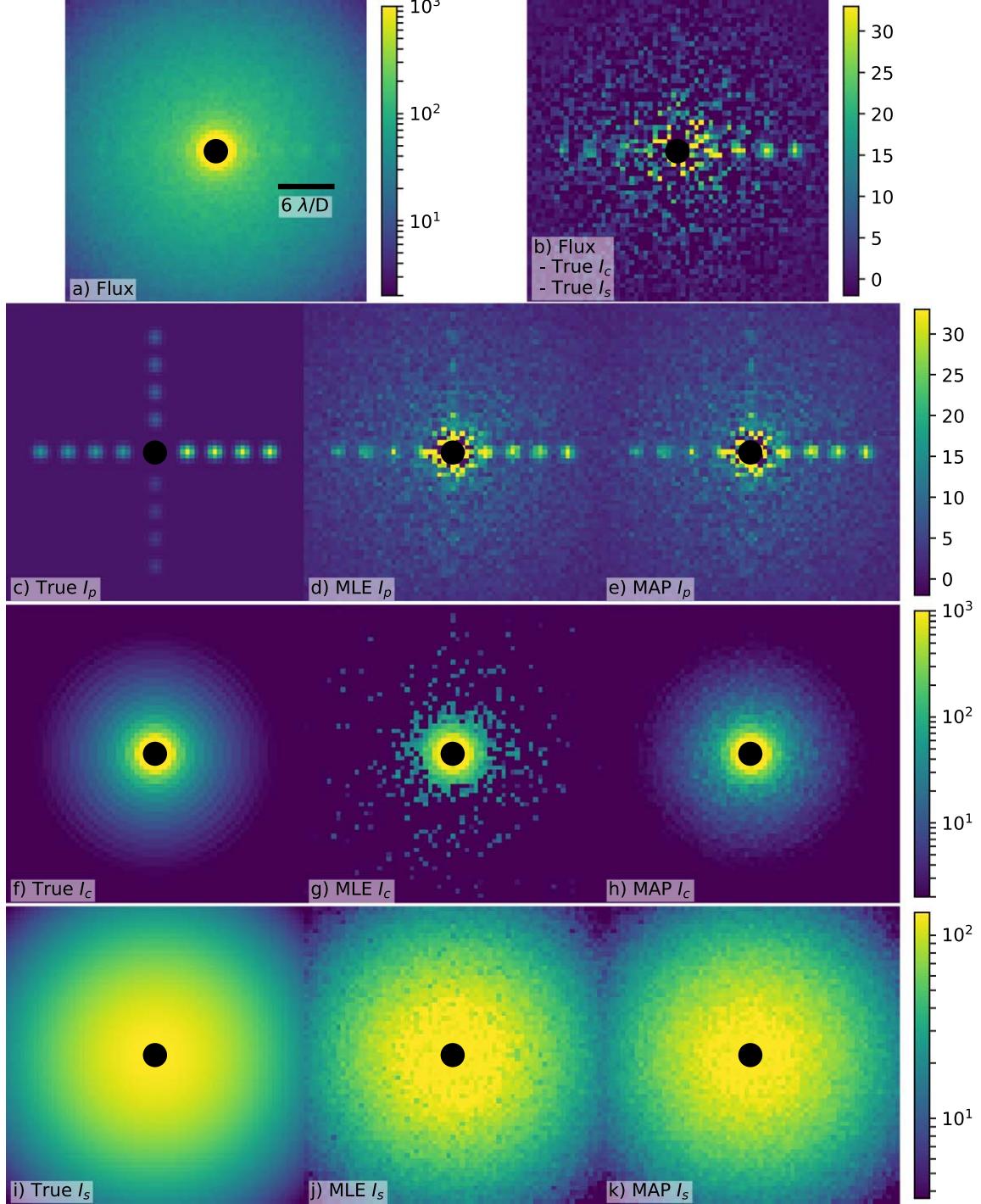


Figure 4. Performance of our photon-counting SSD algorithm on simulated telescope images. Panel (a) shows the average intensity of simulated photon lists in each pixel; panels (c), (f), and (i) show the parameters used. Panel (b) subtracts stellar flux from (f) and (i) from the average intensity, (a), illustrating the speckle variance in a long exposure image. Panel (b) represents the theoretical limit of perfect PSF subtraction in a single exposure subject to MR intensity fluctuations. The photon-counting SSD algorithm results in the maximum likelihood estimates shown in (d), (g), and (j). Using a priori knowledge of the I_c parameter (we used a Gaussian prior of $\text{True } I_c \pm 3 \cdot \sqrt{\text{True } I_c}$) we generate the maximum a posteriori (MAP) estimates shown in (e), (h), and (k). The planet signals extracted from the MAP I_p estimate in (e) are not significantly improved compared to the MLE I_p in (d). However, the SSD results in (d) and (e) both extract the planet better than the perfect PSF subtraction shown in (b). All images represent 30 s of data on a magnitude $J = 10$ star with an 8.2 m telescope; all units are photons/second. (A color version of this figure is available in the online journal.)

Table 1
Companion SSD Signal-to-noise Ratio

Contrast	Separation = $3.5 \lambda/D$			Separation = $6.5 \lambda/D$			Separation = $9.5 \lambda/D$			Separation = $12.5 \lambda/D$		
	MLE	MAP	Limit ^a	MLE	MAP	Limit ^a	MLE	MAP	Limit ^a	MLE	MAP	Limit ^a
$4 \cdot 10^{-4}$	4.8	5.3	2.3	6.8	7.0	3.8	7.6	8.3	5.7	7.0	9.4	8.8
$2 \cdot 10^{-4}$	2.8	3.0	1.2	4.0	4.0	1.9	4.9	4.9	2.9	5.6	5.9	4.5
$1 \cdot 10^{-4}$	1.6	1.7	0.6	2.3	2.3	1.0	2.8	2.8	1.4	3.5	3.5	2.3
$5 \cdot 10^{-5}$	0.9	0.9	0.3	1.3	1.2	0.5	1.6	1.5	0.7	2.0	2.0	1.1

Note.

^a Detection limit for a 30 s long exposure with perfect PSF subtraction of the stellar light.

As in millisecond imaging SSD, I_p and I_c become indistinguishable in this limit.

4.2. Maximum A Posteriori Estimation

Prior knowledge of a parameter can improve the estimates of I_c , I_s , and I_p . In our case, we commonly have information on the I_c parameter which corresponds to the static or quasistatic speckle PSF either from a telescope model or measured on a reference star. For such situations, the log likelihood function in Equation (12) can be modified with a Gaussian prior as

$$\log(\mathcal{L}) \rightarrow \log(\mathcal{L}) - \frac{1}{2}((I_c - \tilde{I}_c)/\sigma[\tilde{I}_c])^2 \quad (13)$$

where $\tilde{I}_c \pm \sigma[\tilde{I}_c]$ is the prior on I_c . The new estimates of I_c , I_s , and I_p become maximum a posteriori (MAP) estimates.

4.3. Performance on Simulated Telescope Image

We evaluate the performance of the photon-counting SSD algorithm on a simulated 30 s telescope image without a coronagraph. We identify I_c as the Airy-ring pattern of a diffraction limited telescope with a circular unobstructed aperture, I_s as the seeing halo from atmospheric speckles with a Strehl ratio of 0.7, and I_p as a series of injected planets at various brightnesses and separations. The faintest planet's total intensity is 40 photons/s, for a contrast of $5 \cdot 10^{-5}$ with the host star, and is separated by $3.5 \lambda/D$. Assuming a 5% end-to-end throughput on an 8.2 m telescope, the stellar magnitude in the near-infrared is approximately $J = 10$. We assume a speckle decorrelation time of $\tau_s = 0.1$ s and detector dead time of $\tau_0 = 10 \mu\text{s}$.

Each pixel has an independent 30 s photon list generated from the “True” I_c , I_s , and I_p shown in panels (c), (f), and (i) respectively of Figure 4. Spatial correlations in the photon lists are ignored for simplicity. The average intensity realized for each pixel is shown in Figures 4(a). (b) is the average intensity minus the expected light from the star, ($\text{True } I_c + \text{True } I_s$), which illustrates the best possible long exposure PSF subtraction (compare the background variance to Equation (2)). The MLE I_c , I_s , and I_p are shown in Figures 4(d), (g), and (j) respectively. In Figures 4(e), (h), and (k), we calculate the MAP estimates. We

used the True $I_c \pm 3 \cdot \sqrt{\text{True } I_c}$ as a Gaussian prior on I_c ; in practice one could use a telescope model or a reference PSF. The central black dot with radius $1.22 \lambda/D$ is not a coronagraph but simply obscures the on-axis light for convenience.

Comparing Figures 4(d) to (b) shows that the MLE I_p from the photon-counting SSD algorithm recovers the injected planets better than a perfect stellar PSF subtraction (i.e., subtraction of the true I_c and I_s). Figures 4(d) and (e) show that the MAP estimate for I_p is not significantly better than the MLE I_p (although the MAP estimate for I_c is more precise). This is surprising because the MAP estimate includes a prior on I_c that should help discriminate between I_c and I_p . That this is not the case indicates we can take full advantage of the photon-counting SSD algorithm without prior knowledge of the telescope PSF.

For the central pixel of each planet, we used 10^5 independent photon lists to calculate the signal-to-noise ratio $S/N = (\langle I_p \rangle - \langle \text{Background} \rangle)/(\text{std. dev. } \langle I_p \rangle)$ where the $\langle \text{Background} \rangle$ is estimated by not injecting a planet. These are recorded in Table 1. The long exposure photon noise limit is also recorded in Table 1 where the estimated I_p is equal to the total flux minus the light from the star, ($\text{True } I_c + \text{True } I_s$). While the results are from the simulated ensemble, they match the results from Equation (2). Table 1 is representative of 30 s of data, but the S/N will scale with $\sqrt{T_{\text{tot}}}$. For a 2 minute exposure, all values in Table 1 should be scaled up by a factor of 2. The S/N was calculated using only the central pixel for convenience but would be larger if the surrounding pixels were considered.

5. Discussion

While it remains impossible to beat the photon shot noise \sqrt{N} , Figure 4 shows that the photon-counting SSD algorithm can beat the long exposure ($t_{\text{exp}} \gg \tau_s$) photon noise limit described by Equation (2). Table 1 quantifies the improvement in the S/N as a factor of 3 in the case of faintest planet ($5 \cdot 10^{-5}$) at the nearest separation ($3.5 \lambda/D$). This is possible because the speckle fluctuations are temporally resolved ($t_{\text{exp}} \ll \tau_s$) and individual speckles are probed by multiple photons ($\delta t \ll \tau_s$). In the case that the fluctuations from stellar speckles dominate the variance of the total intensity

$(2I_s \tau_s \gg 1)$, fast, noiseless detectors like MKIDs or EMCCDs are needed to dig beneath the noise. This can occur in high-contrast imaging at separations between $1 \lambda/D > r <$ seeing radius but is especially important at $\lesssim 5 \lambda/D$ where ADI and SDI start to lose their effectiveness (depending on spectral coverage, sky rotation, and AO performance).

SSD will benefit ADI and SDI by reducing speckle noise from the data that is fed into those algorithms. ADI processing can be approached the same way as usual, but instead of using raw images one would use the I_p images produced with SSD. The modulation of the planet location will be unaffected by SSD. Similarly for SDI, the algorithm would be given I_p maps at various wavelengths. This approach would require wavelength information for each detected photon, which is an intrinsic feature of an MKID detector.

Our SSD algorithm does not perform well when $I_c \gg I_s$ as seen in Figure 3. In this regime there is little modulation of the static speckle intensity I_c by the atmospheric speckle field I_s , and as a result I_c starts to become indistinguishable from the Poisson distributed I_p . This can be greatly mitigated with a coronagraph and possibly active speckle nulling (Martinache et al. 2014) which directly reduce I_c in the image.

On a space-borne telescope atmospheric speckles are not a concern, implying that the image-plane intensity will have different temporal behavior from ground-based observatories. While our SSD algorithm relies on the intensity following a MR distribution, in general any distribution can be used so long as it is known. It may also be possible to modulate speckles in a controlled way using onboard AO.

6. Conclusions

In this paper we exploit photon arrival time statistics with a SSD algorithm to distinguish planets from speckles. We first extend previous work with a formalized maximum likelihood algorithm operating on light curves with fixed, albeit fast, exposure times. We find that the likelihood space can sometimes result in bimodal behavior in the case of $I_c \gg I_s$. Additionally, the choice in exposure time can systematically skew the MLE of I_p and inflate its variance. More generally, with a fixed exposure time, the performance will change as a function of parameters I_c , and I_s . This is a problem because I_c and I_s can change with observing conditions as well as with separation from the host star.

To overcome these difficulties, we have developed a new photon-counting SSD algorithm that calculates the maximum likelihood for I_p directly from the individual photon arrival times. With this approach the likelihood space becomes smooth and unimodal and the precision is maximized. The planet detection performance can be better by a factor of 2 than perfect stellar PSF subtraction of a long exposure. This requires fast noiseless detectors like MKIDs.

We have made several simplifying assumptions in our analysis. We take the speckle temporal PSD of our simulated data to be described by the single exponential timescale τ_s , we assume I_c , I_s , and I_p remain constant (ignoring variations to the instantaneous Strehl), and we assume that the MR distribution accurately describes the off-axis stellar intensity. Finally, we ignore chromaticity. These assumptions represent avenues of exploration for future work. The speckle temporal PSD can be measured and used to more accurately simulate photon lists. Since the instantaneous Strehl can be measured, a future implementation of this algorithm might use that information in a more sophisticated model for the incoherent light like in Gladysz et al. (2008). This would require an understanding of how Strehl variations effect I_c and I_s .

SSD algorithms for highly corrected AO images are ultimately constrained on two fronts. First, the photon arrival time δt must be much shorter than the speckle decorrelation time τ_s ; we need many photons to characterize the properties of a materialized speckle. Second, the performance degrades when I_c is large but I_s is small, because the incoherent planet light masquerades as the static speckles described by I_c . Fortunately, this can be improved with a coronagraph.

In return, SSD algorithms are most useful when I_s is large, which is often inescapable at small inner working angles. Furthermore, the results are not directly dependent on separation (although they are dependent on I_c and I_s , which are larger at small separations). This makes SSD a powerful postprocessing technique at small inner working angles where it can complement more established techniques like ADI and SDI. At high speckle intensities, SSD can even outperform the theoretical limits of a perfect implementation of ADI and SDI.

This work was supported by the National Science Foundation grant 1710385 and NASA ROSES grant NNX15AG23G. We would like to thank Michael Fitzgerald (University of California, Los Angeles) for helpful comments during the early stages of this work.

Appendix

We derive here the noise of a long ($t_{\text{exp}} \gg \tau_s$) exposure subject to speckle statistics. This is not simply photon shot noise, but arises due to the fluctuations of the MR itself. We consider a MR parametrized by I_c and I_s with an exponential decorrelation time τ_s . The basic statistics of the distribution are

$$\langle I \rangle = I_c + I_s \quad (14)$$

$$\sigma_I^2 = \langle I^2 \rangle - \langle I \rangle^2 = I_s^2 + 2I_c I_s \quad (15)$$

$$\langle (I_i - \langle I \rangle)(I_j - \langle I \rangle) \rangle = \sigma_I^2 \exp\left[-\frac{|t_i - t_j|}{\tau_s}\right]. \quad (16)$$

where brackets, $\langle \rangle$, denote an ensemble average.

We wish to compute the variance of the mean intensity measured over a finite time interval T_{tot} , which we divide into N subintervals, each of length δt :

$$\bar{I} = \frac{1}{N} \sum_i I_i. \quad (17)$$

Here, the overbar denotes a temporal average. The variance is

$$\langle \bar{I}^2 \rangle - \langle \bar{I} \rangle^2 = \frac{1}{N^2} \left\langle \sum_i \sum_j I_i I_j \right\rangle - \frac{1}{N^2} \left\langle \sum_i I_i \right\rangle^2 \quad (18)$$

$$= \frac{1}{N^2} \left\langle \sum_i \sum_j (I_i - \langle I \rangle)(I_j - \langle I \rangle) \right\rangle \quad (19)$$

$$= \frac{\sigma_I^2}{N^2} \sum_i \sum_j \exp \left[-\frac{|t_i - t_j|}{\tau_s} \right]. \quad (20)$$

Now we will set $t_i = i\delta t$. For the third step below, we use $\delta t/\tau_s \ll 1$. More generally, we take $\delta t \rightarrow 0$, $N \rightarrow \infty$:

$$\frac{1}{N^2} \sum_i \sum_j \exp \left[-\frac{|t_i - t_j|}{\tau_s} \right] = \frac{1}{N^2} \sum_i \sum_j \exp \left[-\frac{|i - j|\delta t}{\tau_s} \right] \quad (21)$$

$$= \frac{1}{N^2} \sum_{i=0}^{N-1} \left(\sum_{j=0}^{N-i-1} \exp \left[-\frac{j\delta t}{\tau_s} \right] + \sum_{j=0}^i \exp \left[-\frac{j\delta t}{\tau_s} \right] - 1 \right) \quad (22)$$

$$= \frac{1}{N^2} \sum_{i=0}^{N-1} \left(\frac{\tau_s}{\delta t} \left(2 - \exp \left[-\frac{(N-i)\delta t}{\tau_s} \right] \right) - \exp \left[-\frac{(i+1)\delta t}{\tau_s} \right] \right) - 1 \quad (23)$$

$$= \frac{2\tau_s}{N\delta t} - \frac{2\tau_s}{N\delta t} \frac{1}{N} \left(\sum_{i=1}^N \exp \left[-\frac{i\delta t}{\tau_s} \right] \right) - \frac{1}{N} \quad (24)$$

$$= \frac{2\tau_s}{N\delta t} - 2 \left(\frac{\tau_s}{N\delta t} \right)^2 \left(1 - \exp \left[-\frac{(N+1)\delta t}{\tau_s} \right] \right) + \frac{2\tau_s}{N^2\delta t} - \frac{1}{N}. \quad (25)$$

Taking $N \rightarrow \infty$, $\delta t \rightarrow 0$, and $N\delta t = T_{\text{tot}}$, we have

$$\langle \bar{I}^2 \rangle - \langle \bar{I} \rangle^2 = \left(\frac{2\tau_s}{T_{\text{tot}}} - 2 \left(\frac{\tau_s}{T_{\text{tot}}} \right)^2 \left(1 - \exp \left[-\frac{T_{\text{tot}}}{\tau_s} \right] \right) \right) \sigma_I^2 \quad (26)$$

$$= \left(\frac{2\tau_s}{T_{\text{tot}}} - 2 \left(\frac{\tau_s}{T_{\text{tot}}} \right)^2 \left(1 - \exp \left[-\frac{T_{\text{tot}}}{\tau_s} \right] \right) \right) (I_s^2 + 2I_c I_s). \quad (27)$$

In the limit of a very short integration, $T_{\text{tot}} \ll \tau_s$, the prefactor is unity (as expected). For an integration time much longer than the decorrelation time, $T_{\text{tot}} \gg \tau_s$ (as is more typical), Equation (27) simplifies to

$$\langle \bar{I}^2 \rangle - \langle \bar{I} \rangle^2 \approx \left(\frac{2\tau_s}{T_{\text{tot}}} \right) (I_s^2 + 2I_c I_s). \quad (28)$$

Assuming $2I_s\tau_s \gg 1$ (the inter-photon arrival time from the speckle field is much shorter than the decorrelation time τ_s), Equation (27) dominates over simple shot noise $\sigma^2 = (I_c + I_s)/T_{\text{tot}}$. Interestingly, setting the number of independent realizations of the MR equal to T_{tot}/τ_s would miss the factor of two in Equation (28). Adding shot noise back in, assuming $T_{\text{tot}} \gg \tau_s$, and including a component I_p incoherent with I_c and I_s , we have

$$\sigma_{I_{\text{tot}}}^2 \approx \frac{2\tau_s(I_s^2 + 2I_c I_s) + I_c + I_s + I_p}{T_{\text{tot}}}. \quad (29)$$

ORCID iDs

Alex B. Walter <https://orcid.org/0000-0002-4427-4551>

References

Aime, C., & Soummer, R. 2004, 12th European Signal Processing Conf. (Vienna, Austria: IEEE), 1071

Cagigal, M. P., & Canales, V. F. 1999, *Proc. SPIE*, 3866, 165

Cagigal, M. P., & Canales, V. F. 2001, *OptEn*, 40, 2690

Canales, V. F., & Cagigal, M. P. 1999, *ApOpt*, 38, 766

DeLong, E., DeLong, D., & Clarke-Pearson, D. 1988, *Biometrics*, 44, 837

Fitzgerald, M. P., & Graham, J. R. 2006, *ApJ*, 637, 541

Frazin, R. A. 2016, *Proc. SPIE*, 9909, 99093X

Gerard, B. L., Marois, C., Currie, T., et al. 2019, *ApJ*, 158, 36

Gladysz, S., & Christou, J. C. 2008, *ApJ*, 684, 1486

Gladysz, S., & Christou, J. C. 2009, *ApJ*, 698, 28

Gladysz, S., Christou, J. C., Bradford, L. W., & Roberts, L. C. 2008, *PASP*, 120, 1132

Gladysz, S., Yaitskova, N., & Christou, J. C. 2010, *JOSAA*, 27, A64

Goebel, S. B., Guyon, O., Hall, D. N. B., et al. 2018, *PASP*, 130, 104502

Goodman, J. W. 1975, in *Statistical Properties of Laser Speckle Patterns*, ed. J. C. Dainty (Berlin, Heidelberg: Springer), 9

Guyon, O. 2005, *ApJ*, 629, 592

Jensen-Clem, R., Mawet, D., Gonzalez, C. A. G., et al. 2017, *AJ*, 155, 19

Keppler, M., Benisty, M., Müller, A., et al. 2018, *A&A*, 617, A44

Krzanowski, W. J., & Hand, D. J. 2009, *ROC Curves for Continuous Data* (1st ed.; London/Boca Raton, FL: Chapman and Hall/CRC Press)

Kuzuhara, M., Tamura, M., Kudo, T., et al. 2013, *ApJ*, 774, 11

Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., & Artigau, É. 2007, *ApJ*, 660, 770

Lagrange, A. M., Bonnefoy, M., Chauvin, G., et al. 2010, *Sci.*, 329, 57

Lozi, J., Guyon, O., Jovanovic, N., et al. 2018, *Proc. SPIE*, 10703, 1070359

Macintosh, B., Graham, J. R., Barman, T., et al. 2015, *Sci.*, 350, 64

Macintosh, B., Poyneer, L., Sivaramakrishnan, A., & Marois, C. 2005, *Proc. SPIE*, 5903, 59030J

Marois, C., Correia, C., Galicher, R., et al. 2014, *Proc. SPIE*, 9148, 91480U

Marois, C., Doyon, R., Racine, R., & Nadeau, D. 2000, *PASP*, 112, 91

Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., & Nadeau, D. 2006, *ApJ*, 641, 556

Marois, C., Macintosh, B., Barman, T., et al. 2008, *Sci.*, 322, 1348

Martinache, F., Guyon, O., Jovanovic, N., et al. 2014, *PASP*, 126, 565

Mawet, D., Milli, J., Wahhaj, Z., et al. 2014, *ApJ*, 792, 97

Mawet, D., Pueyo, L., Lawson, P., et al. 2012, *Proc. SPIE*, 8442, 844204

Meeker, S. R., Mazin, B. A., Walter, A. B., et al. 2018, *PASP*, 130, 065001

Racine, R., Walker, G. A. H., Nadeau, D., Doyon, R., & Marois, C. 1999, *PASP*, 111, 587

Roddier, F., Gilli, J. M., & Lund, G. 1982, *JOpt*, 13, 263

Soummer, R., & Ferrari, A. 2007, *ApJL*, 663, L49

Soummer, R., Ferrari, A., Aime, C., & Jolissaint, L. 2007, *ApJ*, 669, 642

Soummer, R., Pueyo, L., & Larkin, J. 2012, *ApJL*, 755, L28

Sparks, W. B., & Ford, H. C. 2002, *ApJ*, **578**, 543

Stangalini, M., Causi, G. L., Pedichini, F., et al. 2018, *ApJ*, **868**, 6

Stangalini, M., Pedichini, F., Ambrosino, F., Centrone, M., & Moro, D. D. 2016, *Proc. SPIE*, **9909**, 99097P

Tanner, Jr., W. P., & Swets, J. A. 1954, *Psychological Review*, **6**, 401

van Eyken, J. C., Strader, M. J., Walter, A. B., et al. 2015, *ApJS*, **219**, 14

Walter, A. B., Mazin, B. A., Bockstiegel, C., et al. 2018, *Proc. SPIE*, **10702**, 107020V