Detection of a Moving Source in Speckle Noise. Application to Exoplanet Detection

Isabelle Smith, *Student Member, IEEE*, André Ferrari, *Member, IEEE*, and Marcel Carbillet

***Abstract*—Astronomical instruments able to detect the direct light of extra solar planets are currently under development. This paper focuses on instruments that will acquire a set of successive images where the planet (the *source* in more general purposes) moves in a known manner on a speckled background. Performant signal processing tools are required to account for the very low signal-to-noise ratio of the data. In the astrophysical context, the background arises mainly from the light scattered by the parent star. An accurate—but general—data model has been proposed in previous works to statistically describe optical images taking into account the spatial correlation of the wavefront complex amplitude. First, an estimator of the position and the intensity of the potential source is proposed. Because of several kinds of numerical constraints, it is derived from a simplified Gaussian data model. Under reasonable constraints on the motion of the source, the estimators are proved to be consistent, even under the accurate data model. For the detection test, we propose to threshold a linear statistics that arises from the intensity estimation. The threshold needs to be precisely related to the probability of false alarm (PFA) and the probability of detection (PD). Under the detailed model, the distribution of the data is only reachable through its moment generating function. We propose therefore to estimate analytically PFA and PD using the saddlepoint approximation. To evaluate the quality of these estimations, a Monte Carlo analysis is applied to monodimensional simulated data. The saddlepoint approximation proves to be very accurate, unlike the Gaussian approximation or even a low-order Gram-Charlier series approximation.**

***Index Terms*—Astronomy, estimation, object detection, signal detection, speckle.**

I. Introduction and Summary

THIS paper is devoted to the detection/estimation problem of a moving source from a set of successive optical images. It focuses on a low signal-to-noise ratio case where a detailed modeling of the image formation is required and all the motion parameters are known, except the initial position. The images are degraded by spatially correlated speckle noise [1], static aberrations, nonuniform background, photocounting noise, camera read-out noise, etc.

Estimators of the intensity and the initial position of the source are to be derived, as well as a detection test where the threshold is to be accurately related to the probability of false alarm (PFA). However, the true data model distribution being intractable, the derivation of classical estimators and a classical composite test such as the generalized likelihood ratio test (GLRT) is unreachable. In this context, the solution generally relies on suboptimal simple estimators and detector. This limitation is also imposed in general by the large amount of data to process. Here the estimators and a linear detection test arise from a Gaussian assumption on the data. The performances of the test must also be derived using a detailed data model. Unlike the direct distribution, the moment generating function (MGF) of the speckle model is reachable. Then the receiver operating characteristic (ROC) curve and the PFA-threshold relation of all linear tests for the precise model can be achieved using the saddlepoint approximation.

These results are applied to ground-based direct imaging of exoplanets, which is a current driving motivation in astronomy. This motivation is materialized by future planet finders such as the very large telescope (VLT) instrument Spectro-Polarimetric High-Contrast Exoplanet REsearch (SPHERE) [2] and the Gemini instrument Gemini planet imager (GPI) [3]. Both instruments will include extreme adaptive optics (AO) and high-contrast coronagraphy. AO's purpose is to compensate in real-time thanks to deformable mirrors the deformation of the wavefronts due to the terrestrial atmosphere and optical aberrations of the telescope. Then coronagraphy aims at suppressing the star light diffracted by the primary mirror at the planet location. These setups are highly necessary for the direct detection of exoplanets: a very high intensity contrast between the star and the exoplanet can be expected in the infrared band (i.e., contrast ratios of 10^6 to 10^{10} depending on the type of the exoplanet). Without extreme AO and a performant coronagraph (that would reduce the diffracted light by a factor of typically 10^7, as integrated in the field of interest), the exoplanet would be too deeply embedded in the light coming from the star and scattered by the terrestrial atmosphere, the diffraction of the primary mirror and the optical aberrations. As a consequence, no algorithm could be reasonably applied for exoplanet detection [4]. We will focus here on the instrument SPHERE [2]. The main difficulty arises from the combination of residual speckles from the turbulent atmosphere (uncorrected by the AO system) and static speckles from the optical system aberrations. These diffraction patterns being similar to the unresolved exoplanet profile, detection algorithms proposed in the literature try to discriminate the planet from the background by the detection of its motion. The motion arises with alt-azimuthal telescopes mounts as the telescope moves to counterbalance the Earth one-day period rotation [5]. It produces an apparent field-of-view global rotation. In particular, the source rotates around the center of the camera in a determistic and known manner. Note that on a period of one night,

Manuscript received December 18, 2007; revised October 17, 2008. First published November 11, 2008; current version published February 13, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Maria Hansson-Sandsten.

The authors are with the Laboratoire Fizeau, UNS/OCA/CNRS, Parc Valrose, 06108 Nice Cedex 2, France (e-mail: isabelle.smith@unice.fr; ferrari@unice.fr; marcel.carbillet@unice.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2008.2009273

the orbital motion of the exoplanet around the star is negligible for most star-exoplanet systems. This effect is therefore not accounted for in this paper. Previous works try to suppress and stabilize the background using differential processing of successive images [6], [7]. The detection will involve typically 1000 images of 300×300 pixels for a 3-h run. It is important to note that the simulation of realistic datasets using for example the software package CAOS [8] or the software package SPHERE [9] is extremely time demanding and does not allow one to obtain probability of detection (PD) and PFA using Monte Carlo simulations.

The conventions used in the paper are the following: boldscript is used for column vectors (e.g., $\boldsymbol{u}$), $\boldsymbol{u}^t$ denotes the transpose of $\boldsymbol{u}$, $\boldsymbol{u}^H$ denotes the complex conjugate of $\boldsymbol{u}^t$, $\boldsymbol{u}_\ell$ denotes the ℓth component of $\boldsymbol{u}$ and $\boldsymbol{v} = e^{\boldsymbol{u}}$ means $\boldsymbol{v}_\ell = e^{\boldsymbol{u}_\ell}\ \forall\ \ell$. Then, identical notations are used for random variables and their realizations and uppercases are only used for matrices (*e.g.,* A). $A_{\ell,m}$ is the (ℓ, m)th component of A, $Tr(A)$ is its trace, $|A|$ its determinant, I is the identity matrix of the relevant dimension, and $D_{\boldsymbol{s}}$ refers to the diagonal matrix whose diagonal terms are the components of $\boldsymbol{s}$.

This paper is organized as follows. Section II presents the detailed statistical model used to describe the observed data. Section III deals with the estimation algorithm: first the estimators are derived using a simplified data model, then their statistical properties under the simplified model and the detailed model are discussed, and they are finally applied to step-by-step astrophysical simulations. Section IV is devoted to the choice of the detection test and the evaluation of the relations PFA-threshold and PD-threshold using the true data model. Analytical expressions derived in Section IV are validated in Section V using monodimensional simulations.

II. Data Model

The point of this section is to develop a precise statistical model for the data. This detailed model is derived in the specific context of exoplanet direct imaging but is fully relevant in the general framework of speckle imaging [1]. The dataset consists of N successive images $\boldsymbol{x}(k)$, $k = 1 \ldots N$. The sum of such images are illustrated by the left image of Fig. 1, as it may appear within the astrophysical context described in the introduction. The individual images have been simulated for the IRDIS (Infra-Red Dual-beam Imaging and Spectroscopy) facility of SPHERE [2], using the software package SPHERE [9] developed within the CAOS problem-solving environment [10] and assuming the standard simulation parameters [11]. This package, developed in the framework of the SPHERE consortium, simulates the propagation of the light emitted by the physical scene (star + exoplanet) and reaching the CCD camera. It relies on a detailed physical step-by-step modeling and accounts for the analyses and observations made on the prototype instruments currently developed for SPHERE. For the astrophysical problematic handled here, no real data are available yet. The simulated images are realistic enough for the purpose of this paper.

In the right image of Fig. 1, the integrated motion of the source is shown for different initial positions r_0. Field rotation

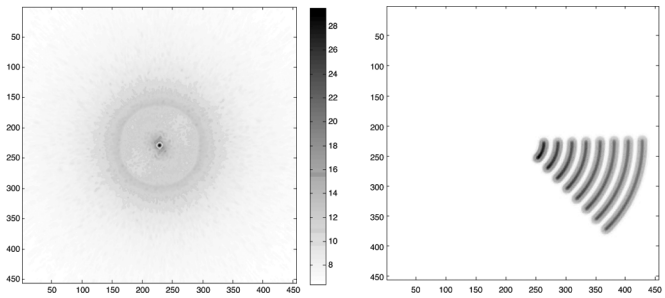

Fig. 1. The left image is the time integration of $N = 450$ simulated data images: $\sum_{k=1}^{N} \boldsymbol{x}(k)$. The right image represents the time integration of $N = 450$ images as if there were just exoplanets and no star. It shows $\sum_{k=1}^{N} \boldsymbol{p}(k, r_0)$ for eight different initial positions r_0, all initial positions being located on the first diagonal of the image. Field-of-view motion is a rotation that applies identically on the whole image and with a time-varying angular speed. Both figures are represented at the power of 0.2.

is a global circular rotation, not uniform in time (rotation rate is not constant). The dataset presented in Fig. 1 is simulated from a quite complex physical step-by-step modeling, and needs now to be described in a probabilistic language.

A. Distribution of the Incoming Light Intensity

The light intensity in the focal plane of the telescope that leads to image k is represented by an $M \times 1$ vector $\boldsymbol{i}(k)$, where M is the number of pixels of the detector. $\boldsymbol{i}(k)$ is the sum of the star light $\boldsymbol{u}(k)$ and a potential exoplanet light. The star illuminates all pixels because of numerous diffusion and diffraction effects, unlike the exoplanet whose total intensity is negligible with respect to the star total intensity. The star contribution is called the *background* and the exoplanet contribution, which is the signal of interest, concentrated in a region of a few pixels, is called the *source*.

The background intensity vector $\boldsymbol{u}(k)$ is associated to the complex amplitude vector $\boldsymbol{\psi}(k)$ through $\boldsymbol{u}_\ell(k) = |\boldsymbol{\psi}_\ell(k)|^2$, $\ell = 1 \ldots M$. $\boldsymbol{\psi}(k)$ arises from the propagation of the star light wavefront through the atmosphere, the AO system and the coronagraph. A straightforward extension of the high flux model derived in [12] to the multivariate case leads to a decomposition of the complex amplitude $\boldsymbol{\psi}(k)$ in two terms: $\boldsymbol{\psi}(k) = \boldsymbol{\mu} + \boldsymbol{\phi}(k)$, where

- the first term, $\boldsymbol{\mu} \in \mathbb{C}^M$, models the static impulse response of the coronagraph and the *unknown* static aberrations (e.g., lens defaults) of the optical system. This deterministic term plays a central role in the detection performance degradation [13];
- the second term models the residuals of atmospheric turbulence that are not corrected by the AO system and that propagate through the coronagraph. This term is assumed to be a complex zero-mean circular Gaussian vector, so that $\mathsf{E}[\boldsymbol{\phi}(k)\boldsymbol{\phi}(k)^H] = \Sigma$ and $\mathsf{E}[\boldsymbol{\phi}(k)\boldsymbol{\phi}(k)^t] = 0$, see [1] and [14].

The complex amplitude of the companion planet can be modeled like the star response, but reduces to a simple expression: for extreme AO and precise optical devices, the peak of this response can be considered as deterministic, and the fluctuating outer part can be neglected with respect to the local background.

This instantaneous response can be obtained by simulation, calibration or analytical expression as in [15]. Note that when a coronagraph is used the response is not shift-invariant.

The exoplanet moves with respect to the static optical aberrations because alt-azimuthal mounts make the field-of-view rotating *en bloc* (see Fig. 1) with respect to the coronagraph and the camera. Its position on the image at any time k is a known function of its unknown position r_0 on the first image: field rotation only depends on the position of the star in the sky in the local frame of observation [5]. Recall that during a night, the orbital motion of the explanet around its parent star is neglected.

Consequently the presence of a source results in the deterministic response $\alpha_0 \boldsymbol{p}(k, r_0)$ where α_0 is the *unknown intensity* of the possible source, and $\boldsymbol{p}(k, r_0)$, with $\boldsymbol{p}_\ell(k, r_0) \geq 0 \; \forall \; \ell$, is the *known instrumental response* of the source at time k, where r_0 is the *unknown position* of the source on the first image.

Finally, we will consider that

$$\begin{aligned} \boldsymbol{i}(k) &= \boldsymbol{u}(k) + \alpha_0 \boldsymbol{p}(k, r_0) \\ \text{with } \boldsymbol{u}_\ell(k) &= |\boldsymbol{\psi}_\ell(k)|^2 \text{ and } \boldsymbol{\psi}(k) \sim \mathcal{N}_c(\boldsymbol{\mu}, \Sigma) \end{aligned} \tag{1}$$

and with $\alpha_0 > 0$ if there is a source (H_1 hypothesis) or $\alpha_0 = 0$ if there is no source (H_0 null hypothesis). We make the general assumption that the successive vectors $\boldsymbol{\psi}(k)$, $k = 1 \ldots N$ are independent.

Gaussian circularity for a scalar random variable implies that its real and imaginary parts are independent with identical variance. Consequently (1) shows that $\boldsymbol{u}_\ell(k)$ is proportional to a random variable distributed according to a noncentral χ^2 distribution with two degrees of freedom [16]. The multivariate distribution of $\boldsymbol{u}(k)$ is much more complicated to derive. It can be obtained by noting that $\boldsymbol{u}(k)$ is the diagonal of the $M \times M$ matrix $\boldsymbol{\psi}(k)\boldsymbol{\psi}(k)^H$, so that the distribution of $\boldsymbol{u}(k)$ is the *ad-hoc* marginal of the distribution of $\boldsymbol{\psi}(k)\boldsymbol{\psi}(k)^H$. But the rank of $\boldsymbol{\psi}(k)\boldsymbol{\psi}(k)^H$ is 1 so its probability density function (pdf) does not exist. More specifically, $\boldsymbol{\psi}(k)\boldsymbol{\psi}(k)^H$ has a noncentral complex Wishart distribution [17]–[19]. However, its MGF is fortunately tractable, and the MGF of its marginal distribution is simply computed by evaluating the MGF of $\boldsymbol{\psi}(k)\boldsymbol{\psi}(k)^H$ in a suitable way.

More specifically, since $U(k) = \boldsymbol{\psi}(k)\boldsymbol{\psi}(k)^H$ has a noncentral complex Wishart distribution, [20] gives

$$\mathsf{E}[e^{Tr(SU(k))}] = \frac{e^{Tr(S(I-\Sigma S)^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^H)}}{|I - \Sigma S|}. \tag{2}$$

The distribution of $U(k)$ is related to the distribution of $\boldsymbol{u}(k)$ through

$$\mathsf{E}[e^{Tr(D_{\boldsymbol{s}} U(k))}] = \mathsf{E}\left[e^{\sum_{\ell=1}^{M} \boldsymbol{s}_\ell \boldsymbol{u}_\ell(k)}\right] = h_{\boldsymbol{u}(k)}(\boldsymbol{s}) \tag{3}$$

so applying this marginalization to the MGF (2) gives:

$$h_{\boldsymbol{u}(k)}(\boldsymbol{s}) = \frac{e^{\boldsymbol{\mu}^H D_{\boldsymbol{s}}(I-\Sigma D_{\boldsymbol{s}})^{-1}\boldsymbol{\mu}}}{|I - \Sigma D_{\boldsymbol{s}}|}. \tag{4}$$

According to (1) the MGF of $\boldsymbol{i}(k)$ is finally

$$h_{\boldsymbol{i}(k)}(\boldsymbol{s}) = e^{\alpha_0 \boldsymbol{p}(k, r_0)^t \boldsymbol{s}} h_{\boldsymbol{u}(k)}(\boldsymbol{s}). \tag{5}$$

B. Distribution of the Photo-Electrons

Using a usual semiclassical approach, the photocounting effect has then to be considered when light is converted to an electronic signal by the detector [21].

Let us denote as $\boldsymbol{n}(k)$ the vector of photocounts associated to the intensity $\boldsymbol{i}(k)$. Conditioned upon the vector of intensities, the random variables $\boldsymbol{n}_\ell(k)$ are independent and distributed according to Poisson distributions $\mathcal{P}$ with means $\boldsymbol{i}_\ell(k)$, $\ell = 1 \ldots M$:

$$\boldsymbol{n}_\ell(k) | \boldsymbol{i}_\ell(k) \sim \mathcal{P}(\boldsymbol{i}_\ell(k)). \tag{6}$$

Tractable expressions of $\Pr(\boldsymbol{n}(k) = \boldsymbol{q})$ are obviously difficult to obtain. However, many interesting properties regarding the distribution of $\boldsymbol{n}(k)$ can be derived in the multivariate case. We will focus hereafter on the MGF of $\boldsymbol{n}(k)$ which is related to the MGF of $\boldsymbol{i}(k)$ by [22]:

$$g_{\boldsymbol{n}(k)}(\boldsymbol{z}) = \mathsf{E}\left[\prod_{\ell=1}^{M} \boldsymbol{z}_\ell^{\boldsymbol{n}_\ell(k)}\right] = h_{\boldsymbol{i}(k)}(\boldsymbol{z} - \boldsymbol{1}) \tag{7}$$

C. Distribution of the Preprocessed Digital Images

In addition to the electrons due to the photon process, electrons appear from the read-out noise and the dark current processes. A classical CCD processing consists of calibrating these additive noises and subtracting them to the final digital images. The preprocessed images $\boldsymbol{x}(k)$ are finally modeled by:

$$\boldsymbol{x}(k) = \boldsymbol{n}(k) + \boldsymbol{e}(k) \text{ where } \boldsymbol{e}(k) \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 I) \tag{8}$$

and the set $\{\boldsymbol{n}(k), \boldsymbol{e}(k)\}_{k=1 \ldots N}$ is independent. Then

$$h_{\boldsymbol{x}(k)}(\boldsymbol{s}) = \mathsf{E}[e^{\boldsymbol{x}(k)^t \boldsymbol{s}}] = e^{\sigma^2 \|\boldsymbol{s}\|^2/2} g_{\boldsymbol{n}(k)}(e^{\boldsymbol{s}}). \tag{9}$$

It is worthy to note that adding a continuous random noise to the number of photo-electrons results in a real valued dataset.

III. Estimation Algorithm

A. Simplified Data Model

The distribution of $\boldsymbol{x}(k)$ being intractable, the derivation of a test statistics will rely on the simplified data model

$$\boldsymbol{x}(k) = \boldsymbol{d}_0 + \alpha_0 \boldsymbol{p}(k, r_0) + \boldsymbol{\epsilon}(k) \tag{10}$$

where $\boldsymbol{\epsilon}(k) \sim \mathcal{N}\left(\boldsymbol{0}, \rho_0^2 I\right)$, $\boldsymbol{d}_0$ denotes a stationary deterministic unknown instrumental response and $\alpha_0 \boldsymbol{p}(k, r_0)$ is the same deterministic signature of the source as in the more realistic model.

The Gaussian distribution can be justified for the star light contribution under a high flux assumption. It is proven in Appendix I that when the deterministic part of the star light ($\boldsymbol{\mu}_\ell^2, \ell = 1, \ldots, M$) is "large," the distribution (7) of $\boldsymbol{n}$ properly standardized will converge in distribution to a Gaussian independent distributed vector. The Gaussian approximation for the planet contribution, which is Poisson distributed according to (1) and (6), is classically related to the assumption that α_0 is "large," and again with a proper standardization.

Note that the model (10) can be written as:

$$\boldsymbol{x} = H(r_0)\boldsymbol{\theta}_0 + \boldsymbol{e}, \quad \boldsymbol{\theta}_0^t = \left(\boldsymbol{d}_0^t, \alpha_0\right) \tag{11}$$

where $\boldsymbol{x}$ and $\boldsymbol{e}$ denote respectively the vertical concatenations of the vectors $\{\boldsymbol{x}(k)\}_{k=1,\ldots,N}$ and of the vectors $\{\boldsymbol{\epsilon}(k)\}_{k=1,\ldots,N}$, and $H(r_0)$ is the vertical concatenation of the N matrices $(I, \boldsymbol{p}(k, r_0))$. Equation (11) shows that as soon as r_0 is known the simplified data model reduces to a Gaussian linear one.

B. Maximum Likelihood Estimator

For the simplified model (10), for a given initial position r, the likelihood maximization reduces to a classical chi-square minimization and gives $\hat{\boldsymbol{\theta}}(r) = (H(r)^t H(r))^{-1} H^t(r)\boldsymbol{x}$. Using the block matrix inversion lemma, and after some computations, we get

$$\hat{\alpha}(r) = \sum_k \boldsymbol{w}(k, r)^t \boldsymbol{x}(k) \tag{12}$$

$$\hat{\boldsymbol{d}}(r) = \frac{\sum_k \boldsymbol{x}(k) - \hat{\alpha}(r)\boldsymbol{p}(k, r)}{N} \tag{13}$$

with

$$\boldsymbol{w}(k, r) = \frac{N\boldsymbol{p}(k, r) - \sum_k \boldsymbol{p}(k, r)}{c_N(r)} \tag{14}$$

$$c_N(r) = N \sum_k ||\boldsymbol{p}(k, r)||^2 - ||\sum_k \boldsymbol{p}(k, r)||^2 \tag{15}$$

$$= \sum_k \sum_{l<k} ||\boldsymbol{p}(k, r) - \boldsymbol{p}(l, r)||^2. \tag{16}$$

The MLE estimation of r_0 is then achieved by maximizing the concentrated log-likelihood

$$\ln L(\hat{\boldsymbol{\theta}}(r), r, \rho^2) = -\frac{||\boldsymbol{x}||^2 - \boldsymbol{x}^t H(r)\hat{\boldsymbol{\theta}}(r)}{2\rho^2} - \frac{MN}{2}\ln(2\pi\rho^2). \tag{17}$$

Substituting (12) and (13) in (17), we obtain after a development $\hat{r}_{ML} = \arg\max_{r=1,\ldots,M} \left\{c_N(r)\hat{\alpha}(r)^2\right\}$. The Maximum Likelihood of α is $\hat{\alpha}_{ML} = \hat{\alpha}(\hat{r}_{ML})$. Since $c_N(r)$ is positive according to (16), if we assume that $\hat{\alpha}_{ML} > 0$ for a given dataset $\{\boldsymbol{x}(k)\}_k$, it is shown in Appendix I that:

$$\hat{r}_{ML} = \arg \max_{r=1,\ldots,M} \left\{\sqrt{c_N(r)}\hat{\alpha}(r)\right\} \tag{18}$$

still with

$$\hat{\alpha}_{ML} = \hat{\alpha}(\hat{r}_{ML}). \tag{19}$$

It is very important to note that

$$\begin{aligned} \text{if } c_N(\arg\max_r(\sqrt{c_N(r)}\hat{\alpha}(r))) &\approx c_N(\arg\max_r(\hat{\alpha}(r))) \\ \text{then } \hat{r}_{ML} &\approx \arg\max_r(\hat{\alpha}(r)) \text{ and} \\ \hat{\alpha}_{ML} &\approx \max_r(\hat{\alpha}(r)) \end{aligned} \tag{20}$$

with $c_N(r)$ given by (15). The proof looks like the one given in Appendix I. The equality $c_N(r) = c_N(r')$ occurs for example if the set $\{\boldsymbol{p}(k, r')\}_k$ is the set $\{\boldsymbol{p}(k, r)\}_k$ but where all pixels of the field $\boldsymbol{p}(k, r)$ are translated by the common displacement vector $\vec{r'} - \vec{r}$ for any k. More generally it can be any rotation, reflection or any combination of these isometries.

In the field-rotation case for example, if an initial position r is relabeled by polar coordinates $r \to (\rho, \phi)$, for r and r' such that $\rho(r) = \rho(r')$, $\{\boldsymbol{p}(k, r)\}_k = Rot_{\phi(r')-\phi(r)}[\{\boldsymbol{p}(k, r')\}_k]$ so that $c_N(\rho, \phi)$ does not depend on ϕ. Concerning the dependence on ρ, for example in the simulation illustrated in Fig. 1 $c_N(r)$ varies of the order of 30% between the inner and the outer part of the image. But a neighborhood in which $\hat{r}_{ML}$ belongs with a quite concentrated probability has a typical size of a few pixels,[1] and in such a narrow window $c_N(r) \approx c_N$. This numerical assessment is supported by the estimation results shown in Section III-D where the estimation (20) is used.

Additional remarks are appropriate for the ML estimates

- The identifiability condition for the parameters is $c_N(r) \neq 0$, which is the case as soon as at least two profiles do not totally overlap, as seen in (16).
- After some algebra, like the one that leads to (16) from (15), it can be shown that $\hat{\alpha}(r)$ (12) equals

$$\hat{\alpha}(r) = \frac{\sum_{k=1}^{N} \sum_{l<k} (\boldsymbol{p}(k, r) - \boldsymbol{p}(l, r))^t (\boldsymbol{x}(k) - \boldsymbol{x}(l))}{c_N(r)}. \tag{21}$$

  Consequently $\hat{\alpha}(r)$ can be viewed as the output of a properly normalized matched filter applied on *all* the available differential images $\left(E = \left\{(k, l)_i \in [1, N]^2\right\}_{i \in [1, N^2]}\right)$. But $\hat{\alpha}(r)$ is not equal in general to such a combination of differential images on any subset $\mathcal{S} \subset E$. For example, if we define a set of successive indices by $\mathcal{S} = \{(k, l)_i \text{ such as } k_i < l_i \text{ and } l_i < k_{i+1}\}_i$, then $c_N(r)^{-1} \sum_{(k,l) \in \mathcal{S}} (\boldsymbol{p}(k, r) - \boldsymbol{p}(l, r))^t (\boldsymbol{x}(k) - \boldsymbol{x}(l))$ does not converges to $\hat{\alpha}(r)$ with probability 1.

C. Consistency of the Estimator

Sufficient conditions can be derived for the consistency of the estimators (19) of α_0 and r_0 under the simplified model (10), and of α_0 when r_0 is known under the accurate model of Section II.

1) r_0 Unknown, Simple Model: The model (10) is not linear with respect to the unknown parameters when r_0 is unknown. Furthermore, due to the motion of the source $\boldsymbol{p}(k, r)$, it does not correspond to a standard independent and identically distributed model. Asymptotic properties of MLE's for the independent but not identically distributed case have been intensively studied in the literature, see for example [23] and [24]. However, due to the simplicity of the model (10), sufficient conditions for the consistency and asymptotic normality of $\left(\hat{\alpha}_{ML}, \hat{\boldsymbol{d}}_{ML}, \hat{r}_{ML}, \hat{\rho}^2_{ML}\right)$ can be obtained from the general result derived in [25].

Following [25], the main assumption is the existence of a nonrandom matrix $A_N^{-1} \to 0$ such as, if $\mathcal{J}_N$ denotes the opposite of the Hessian matrix of the log-likelihood function, $A_N^{-1}\mathcal{J}_N\left(A_N^{-1}\right)^t$ converges to a definite positive matrix W.

Computation of the $(M + 3) \times (M + 3)$ matrix $\mathcal{J}_N$ for the model (10) is straightforward. Expression of this matrix is given in (22), shown at the bottom of the next page, assuming

[1]This occurs because of the strong decay of the axi-symmetric profile of the source and because speckles have the same typical profile.

the ordering $(\alpha, \boldsymbol{d}, r, \sigma^2)$ and the notation shown in (22), where $(i,j) \in \{(0,0),(1,0),(1,1),(2,0)\}$ and where $\boldsymbol{p}^{(i)}(k,r)$ denotes the ith derivative of $\boldsymbol{p}(k,r)$ with respect to r.

Choosing $A_N = \sqrt{N}I$ and using standard results on the convergence of empirical moments of i.i.d. Gaussian samples, it can be easily proved that the matrix $A_N^{-1}\mathcal{J}_N\left(A_N^{-1}\right)^t$ converges to a matrix W if and only if the sums $\boldsymbol{S}_N^i$ and $S_N^{i,j}$ converge to some limits $\boldsymbol{S}_\infty^i$ and $S_\infty^{i,j}$. This constitutes the first set of conditions for consistency. The second set of conditions is less straightforward and arises from the constraint that W must be positive definite. For example, application of Sylvester criterion to the limit of (22) shows that consistency requires: $S_\infty^{0,0} > \left\|\boldsymbol{S}_\infty^0\right\|^2$ (note that $S_\infty^{0,0} \geq \left\|\boldsymbol{S}_\infty^0\right\|^2$ is always verified). This second set of conditions prevents simple cases to lead to consistence. For example, if $\boldsymbol{p}(N,r) \rightarrow \boldsymbol{p}$, application of Cesaro rule implies $S_\infty^{0,0} = \left\|\boldsymbol{S}_\infty^0\right\|^2 = \|\boldsymbol{p}\|^2$.

Note that in the specific case where r_0 is known, these conditions can be checked from an exact calculation since the model (11) would be Gaussian and linear with respect to $\boldsymbol{\theta}_0$. The MLE of $\boldsymbol{\theta}_0$ would be an efficient estimator: the unbiasedness is straightforward and the covariance of $\hat{\boldsymbol{\theta}}$ is shown to be equal to the Cramer-Rao lower bound with $\mathsf{cov}(\hat{\boldsymbol{\theta}}) = \rho^2\left(H^t(r_0)H(r_0)\right)^{-1}$. In particular, denoting in (12) $\hat{\alpha}_{ML} = \hat{\alpha}(r_0)$,

$$\mathsf{var}(\hat{\alpha}_{ML}) = \frac{N}{c_N(r_0)}\rho_0^2 \tag{23}$$

so that, dropping the dependence in r, $\lim_{N\rightarrow\infty}\mathsf{var}(\hat{\alpha}_{ML}) = 0$ if $N^{-1}c_N \rightarrow \infty$. The more general conditions derived above lead to this necessary condition because $N^{-1}c_N = N\left(S_N^{0,0} - \left\|\boldsymbol{S}_N^0\right\|^2\right)$.

2) r_0 Known, Accurate Model: Actually, the estimator $\hat{\alpha}_{ML}$ given in (19) is applied to data that cannot be modeled by a Gaussian distribution: its expectation and variance should be derived from the more precise data model presented in Section II.

However, the Taylor expansion of the $\hat{\alpha}_{ML}$ log-MGF from which these moments can be derived is only tractable if r_0 is known. In that case, the estimate is simply linear: $\hat{\alpha}_{ML} = \hat{\alpha}(r_0) = \sum_k \boldsymbol{w}(k,r_0)^t\boldsymbol{x}(k)$ (12).

If $h_i(z,r)$ denotes the MGF of the statistics $\hat{\alpha}(r)$ under H_i (see (28)),

$$h_1(z,r) = \mathsf{E}[e^{z\hat{\alpha}(r)}] = \prod_{k=1}^{N} h_{\boldsymbol{x}(k)}(z\boldsymbol{w}(k,r)). \tag{24}$$

Substituting (4), (5) and (7) in (24) gives:

$$\ln(h_1(z,r)) = \alpha\sum_k (e^{z\boldsymbol{w}(k,r)} - \boldsymbol{1})^t\boldsymbol{p}(k,r_0) + \ln(h_0(z,r)) \tag{25}$$

with:

$$\ln(h_0(z,r)) = \frac{N\sigma^2z^2}{2c_N(r)} - \sum_k \ln|I - \Sigma(D_{e^z\boldsymbol{w}(k,r)} - I)|$$
$$+\boldsymbol{\mu}^H\sum_k (D_{e^z\boldsymbol{w}(k,r)} - I)(I - \Sigma(D_{e^z\boldsymbol{w}(k,r)} - I))^{-1}\boldsymbol{\mu}. \tag{26}$$

Thanks to the stationarity of the background $\boldsymbol{\mu}$ and thanks to the fact that $\sum_k \boldsymbol{w}(k,r) = \boldsymbol{0}$ for any r, the computation of the two first terms of the Taylor expansion of $\ln(h_1(z,r_0))$ proves that $\hat{\alpha}_{ML}$ is still unbiased, and that

$$\begin{aligned}\mathsf{var}[\hat{\alpha}_{ML}] = {} & N\sigma^2c_N(r_0)^{-1} + \alpha\sum_{k=1}^{N}\sum_{\ell=1}^{M}\boldsymbol{p}_\ell(k,r_0)\boldsymbol{w}_\ell(k,r_0)^2 \\ & + \sum_{\ell=1}^{M}(|\boldsymbol{\mu}_\ell|^2 + \Sigma_{\ell,\ell})\sum_{k=1}^{N}\boldsymbol{w}_\ell(k,r_0)^2 \\ & + \sum_{\substack{1\leq\ell\leq M\\ 1\leq m\leq M}}\left(|\Sigma_{\ell,m}|^2 + 2Re(\boldsymbol{\mu}_\ell^*\boldsymbol{\mu}_m\Sigma_{\ell,m})\right) \\ & \times\sum_k \boldsymbol{w}_\ell(k,r_0)\boldsymbol{w}_m(k,r_0).\end{aligned} \tag{27}$$

The first term in (27) is the same as in the simplified model, with $\sigma^2 = \rho_0^2$. The other terms quantify the performance degradations related to the speckle noise and photo-counting noise. If $\forall\, k\|\boldsymbol{p}(k,r_0)\| < \infty$, the second, third and fourth term in (27) contribute to $\mathsf{var}[\hat{\alpha}_{ML}]$ as $O\left(N^3c_N^{-2}\right)$ as $N \rightarrow \infty$, which is slightly more restrictive than $N^2c_N^{-2} \rightarrow 0$ required in the simple model.

Finally, under conditions barely more restrictive than the ones given in Section III-C-I, $\hat{\alpha}_{ML}$ derived in Section III-B is still consistent under the true model with r_0 known.

D. Results

Fig. 2 shows the estimation $\hat{\alpha}(r)$ computed for a large fraction of pixels r from (12), and using the images that lead to Fig. 1. As discussed in Section III-B, $c_N(r) \approx c_N$ locally is quite relevant, so then $\hat{\alpha}_{ML}$ and $\hat{r}_{ML}$ are estimated by (20).

As stated at the beginning of Section II, the simulated data shown on Fig. 1 were obtained from a detailed physical simu-

$$\boldsymbol{S}_N^i = \frac{1}{N}\sum_{k=1}^{N}\boldsymbol{p}^{(i)}(k,r), \quad S_N^{i,j} = \frac{1}{N}\sum_{k=1}^{N}\boldsymbol{p}^{(i)}(k,r)^t\boldsymbol{p}^{(j)}(k,r)$$

$$\frac{\sigma^2}{N}\mathcal{J}_N = \begin{pmatrix} S_N^{0,0} & \boldsymbol{S}_N^{0\,t} & -\frac{1}{N}\sum_k \boldsymbol{p}^{(1)}(k,r)^t\boldsymbol{\epsilon}(k) + \alpha S_N^{0,1} & \frac{1}{N\sigma^2}\sum_k \boldsymbol{p}(k,r)^t\boldsymbol{\epsilon}(k) \\ \boldsymbol{S}_N^0 & I & \alpha\boldsymbol{S}_N^{1\,t} & \frac{1}{N\sigma^2}\sum_k \boldsymbol{\epsilon}(k) \\ -\frac{1}{N}\sum_k \boldsymbol{p}^{(1)}(k,r)^t\boldsymbol{\epsilon}(k) + \alpha S_N^{0,1} & \alpha\boldsymbol{S}_N^1 & -\frac{\alpha}{N}\sum_k \boldsymbol{\epsilon}(k)^t\boldsymbol{p}^{(2)}(k,r) + \alpha^2 S_N^{1,1} & \frac{\alpha}{N\sigma^4}\sum_k \boldsymbol{p}^{(1)}(k,r)^t\boldsymbol{\epsilon}(k) \\ \frac{1}{N\sigma^2}\sum_k \boldsymbol{p}(k,r)^t\boldsymbol{\epsilon}(k) & \frac{1}{N\sigma^2}\sum_k \boldsymbol{\epsilon}(k) & \frac{\alpha}{N\sigma^4}\sum_k \boldsymbol{p}^{(1)}(k,r)^t\boldsymbol{\epsilon}(k) & -\frac{M}{2\sigma^2} + \frac{1}{\sigma^4 N}\sum_k \boldsymbol{\epsilon}(k)^t\boldsymbol{\epsilon}(k) \end{pmatrix} \tag{22}$$

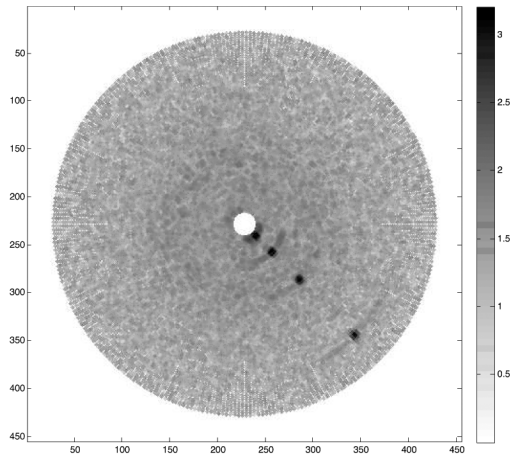

Fig. 2. Map of $\hat{\alpha}(r)$ given by (12) obtained from the data of Fig. 1 at the power of 0.2. The four simulated exoplanets located at $0''.2, 0''.5, 1''$ and $2''$ from the center of the image, and 1.6 10^{-6} times less bright than the central star, give rise to the four local maxima of $\hat{\alpha}(r)$.

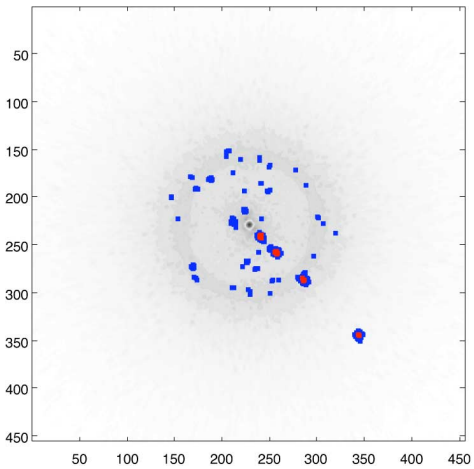

Fig. 3. Thresholded map (two levels $\zeta_1 = 170$ and $\zeta_2 = 280$, to compare to $\max_r \hat{\alpha}(r) = 287$) of $\hat{\alpha}(r)$ where $\hat{\alpha}(r)$ was presented in Fig. 2. The map is superimposed on the log of the data $\sum \boldsymbol{x}(k)$ presented in Fig. 1. The four simulated exoplanets (located at $0''.2, 0''.5, 1''$ and $2''$ from the star (center of the image), and 1.6 10^{-6} times less bright than the central star) are correctly detected. The second threshold gives all these four good detections and no false alarm.

lation devoted to the SPHERE instrument, not from any of the two mathematical models presented in this paper. But although the good properties derived above shouldn't stricty apply, the estimators still give good results on these simulated data.

The estimation of r_0 falls within one pixel of the true position. The quality of the estimation of α_0 is more difficult to estimate because the input in the simulation is not simply α_0. As a first sight, the number of photo-electrons N_γ due to the source on a single exposure is estimated by $\hat{N}_\gamma = \hat{\alpha}_{ML} \times \|\boldsymbol{p}(k,r)\|_1$ (it is the same value for any k and any r in general), and this estimate is compared to the physical flux of the exoplanet used in the simulation, decreased by the different gain factors of the instrument. Then, $\hat{N}_\gamma$ is within a few percents of N_γ.

IV. Detection Algorithm and Detection Level

A. Detection Algorithm

According to the model developed in Section II, the source detection problem consists in deciding among one of the two hypotheses

$$H_0 : \alpha_0 = 0, \quad H_1 : \alpha_0 > 0. \tag{28}$$

A decision test needs to be derived, and some PFA and PD have to be computable to give some sense to the test. In the adopted frequentist approach, a natural test rejects hypothesis H_0 if $T(\boldsymbol{x}) > \zeta$ where $T(\boldsymbol{x})$ is some well chosen statistics. PFA and PD are defined by $P_{FA}(\zeta) = P(T(\boldsymbol{x}) > \zeta | H_0)$ and $P_D(\zeta) = P(T(\boldsymbol{x}) > \zeta | H_1)$.

In general, a very low PFA is required and the PFA is very sensitive to the shape of the right far tail of the distribution of $T(\boldsymbol{x})$. Its distribution has therefore to be precisely known using the model derived in Section II.

A natural choice for $T(\boldsymbol{x})$ is the GLRT statistics. Obviously its computation will be intractable for the accurate model of Section II since the likelihood itself is not tractable. In addition, note that even if an approximate GLRT statistics could be derived, the computation of its PFA would be difficult. A possibility to compute the PFA could be to use the Wilks's theorem [26, Sec. 6.3.1] that states that "under quite general conditions" the asymptotic (as $N \to \infty$) distribution of the GLRT statistics is chi-squared with degree of freedom equal to the difference between the number of free parameters under both hypotheses. However, this result is not applicable because

1) the measurements are independent but *not identically distributed* under H_1. This assumption is central in the proof of the asymptotic GLRT distribution, [26];
2) under H_0 the parameter r_0 is not identifiable. The difficulty related to the unindentifiablity of parameters under H_0 when the GLRT is used has been raised in [27] for a data model similar to (10).

An other possibility would be to use the GLRT of the simple model given in (10). However, its performance property does not apply under the more accurate model and furthermore it is a too complex function of the data to compute its PFA under the accurate model of Section II.

In order to be able to compute accurately the PFA, we, hence, decide to use the natural suboptimal linear statistics $T(\boldsymbol{x}, r) = \hat{\alpha}(r) = \sum_k \boldsymbol{w}^t(k,r)\boldsymbol{x}(k)$ as given in (12). The test has then to be applied to each pixel r:

$$H_0 \text{ rejected at } r \text{ if } \hat{\alpha}(r) > \zeta \tag{29}$$

and PFA and PD become functions of the pixel r on which the test is applied: $P_{FA}(r,\zeta)$, $P_D(r,\zeta)$.

Fig. 3 shows a thresholded map of $\hat{\alpha}(r)$, where $\hat{\alpha}(r)$ is presented in Fig. 2.

B. Saddlepoint Approximation

The purpose of this section is the evaluation of PFA and PD to set the practical value of the threshold ζ as a function of a required PFA and to quantify the performances of the detection scheme (29). Despite the simplicity of the detector $\hat{\alpha}(r)$, the computation of these cumulative probabilities is not straightforward because, as mentioned in Section II, the multivariate pdf

of the data is not available. As explained in Section IV-A, the data have to be precisely modeled: in particular, the distribution cannot be approximated by a simple Gaussian distribution *a priori* . The data MGF is tractable and lead under H_1 to the test statistics MGF given by (25), and by (26) under H_0.

Taking benefit of the existence of an analytic expression of the MGF of the statistics $\hat{\alpha}$, we propose to use the saddlepoint approximation to compute PFA and PD. In particular, the application of this standard approximation to the computation of cumulative probabilities is synthetically presented in [28] where the Laplace transform of the pdf $\tilde{h}_i(z) = h_i(-z)$ is used instead of the MGF. Using only the first term of the developments, the following expressions derived from the saddlepoint approximation are used

$$P_{FA}(\zeta, r) \approx -\left(2\pi\phi_0^{(2)}(z_0, r)\right)^{-1/2} e^{\phi_0(z_0,r)} \tag{30}$$

$$P_D(\zeta, r) \approx 1 - \left(2\pi\phi_1^{(2)}(z_1, r)\right)^{-1/2} e^{\phi_1(z_1,r)} \tag{31}$$

$$\phi_i(z, r) = \frac{\tilde{h}_i(z, r)e^{\zeta z}}{n(i)z}, \quad n(0) = -1, \;\; n(1) = 1 \tag{32}$$

$$\phi_i^{(j)}(z_0, r) = \left.\frac{\partial^j \phi_i(z, r)}{\partial z^j}\right|_{z=z_0} \tag{33}$$

where the saddlepoints $z_0 < 0$ and $z_1 > 0$ satisfy

$$\phi_i'(z_i, r) = 0, \quad z_0 < 0, \;\; z_1 > 0. \tag{34}$$

Remark the following.

- As mentioned in [28], $z_0 < 0$ and $z_1 > 0$ is a natural but unnecessary consequence of our interest in cases where $\mathsf{E}[\hat{\alpha}(r)] < \zeta$ under H_0 and $\mathsf{E}[\hat{\alpha}(r)] > \zeta$ under H_1.
- The computation of the first and second order derivatives of $\phi_0(z, r)$ and $\phi_1(z, r)$ are obtained from (26) and (25) using $(\ln|A(z)|)' = -\mathrm{Tr}(A(z)^{-1}A(z)')$ and $(A(z)^{-1})' = -A(z)^{-1}A(z)'A(z)^{-1}$. These lengthy equations are not given herein.
- The computation of PFA and PD from (30) and (31) does not necessarily involve the resolution of the equation $\phi_0'(z_0, r) = 0$. For example, to plot PFA or PD as a function of ζ, a large set of thresholds has to be covered. But a reexpression of the (32)–(34) parametrizes ζ by z_i

$$\zeta(z_i, r) = \frac{1}{z_i} - \frac{d\ln(\tilde{h}_i(z, r))}{dz}|_{z=z_i}. \tag{35}$$

  Therefore, ranging ζ reduces to ranging z_i by hand if the functions $\zeta(z_i, \cdot)$, $PFA(\zeta)$, etc. are smooth enough.
- If $\sigma^2 = 0$, the test statistics $\hat{\alpha}(r)$ is a linear combination of discrete random variables. Although the forms (30) and (31) are relevant when approximating the cumulative probabilities of continuous random variables, the validity in the discrete case has been recently discussed in [29].

As presented in [28], note that another classical technique to approximate PFA and PD relies on the Gram-Charlier or on the Edgeworth expansion of the pdf of $\hat{\alpha}$ about a Gaussian distribution. The *k*th coefficient of this series requires the *k*th-order cumulant of $\hat{\alpha}$: it can be computed from the Taylor expansion of the log-MGF (25) and (26). A comparison between the saddlepoint approximation and the Gram-Charlier approximation will be discussed in Section V.

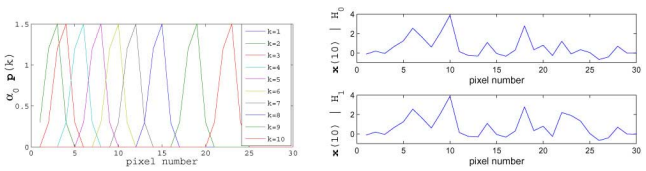

Fig. 4. Left: source profiles $\alpha_0 \boldsymbol{p}(k, r_0)$ for $k = 1 \ldots 10$ (note $\boldsymbol{p}(1, r_0) \approx \boldsymbol{p}(2, r_0)$). Right: realization of $\boldsymbol{x}(10)$ under H_0 (top) and H_1 (bottom).

V. Simulation Results

A. Simulation Cases

The saddlepoint approximation used to derive PFA and PD for (25) and (26) is validated using data simulated under the model described in Section II. To estimate the true cumulative probability from classical Monte Carlo integrations, the number of statistics $\hat{\alpha}$ needs to be *at least* 1 over this probabilty. Since each statistics results from the process of a dataset made of $M \times N$ pixels, and since the PFA of interest are typically as low as 10^{-4}, such Monte Carlo simulations are very time consuming. Consequently, the simulations are 1D, unlike the astronomical images presented on Fig. 1, and with low M and N.

Four simulation cases are used. The parameters used for the reference case (called "Data 1") are as follows.

- $M = 30$ and $N = 10$.
- The $\boldsymbol{\psi}(k)$ are simulated according to a first order circular Gaussian autoregressive process with covariance matrix $\Sigma_{i,j} = \rho^{|i-j|}e^{j\phi(j-i)}$ and with parameters $\boldsymbol{\mu} = 0.08 \times \mathbf{1}$, $\phi = \pi/4$ and $\rho = 0.85$. A typical correlation pattern size induced by Σ can be defined as the -3 dB width of $|\Sigma_{i,j}|$ seen as a function of $|i - j|$. It is equal to 4 pixels for $\rho = 0.85$.
- The source is uniformly accelerated from an initial position $r_0 = 1$ and its profile is chosen static with a characteristic size similar to the correlation pattern size induced by Σ
- The variance of the additive Gaussian noise is $\sigma^2 = 0.25$.
- The datasets are simulated under H_0 ($\alpha_0 = 0$) and H_1 with $\alpha_0 = 0.3$.
- 10^6 independent datasets are simulated for both H_0 and H_1.

The left plot of Fig. 4 shows the source profile $\alpha_0 \boldsymbol{p}(k, r_0)$ for all $k = 1, \ldots, N$ and the right plot of Fig. 4 shows a realization of $\boldsymbol{x}_{10}$ under both hypothesis. Under H_1, the source profile can be seen between pixels 20 and 25.

The detection statistics $\hat{\alpha}(r)$ is then computed at r_0, assumed to be known, from (12). Fig. 5 shows the histograms of $\hat{\alpha}(r_0)$ obtained under H_1 and H_0 for all simulated datasets.

Three other simulation cases are used. Each differs from the reference case Data 1 by two parameters maximum.

- Data 2: new number of exposures $N' = 3N$; to keep the same small number of pixels M, the source is given a slow uniform speed.

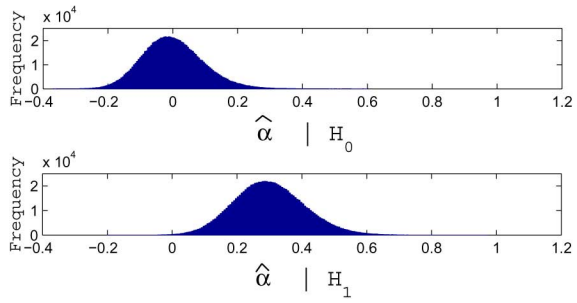

Fig. 5. Histograms of $\hat{\alpha}(r_0)$ under H_0 (top) and under H_1 (bottom).

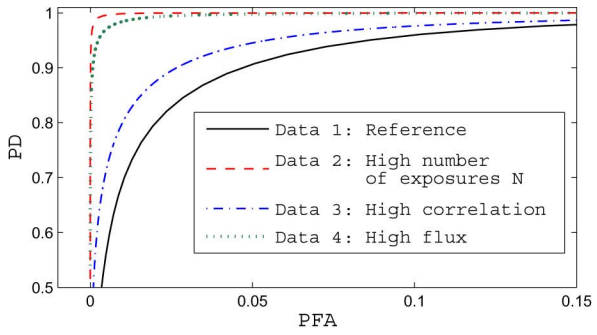

Fig. 6. ROC curves of the detection test given in (29) for the four simulation cases. They are computed from the empirical cumulative probabilities, considered as the "true" PFA and PD in the comparison analysis that follows.

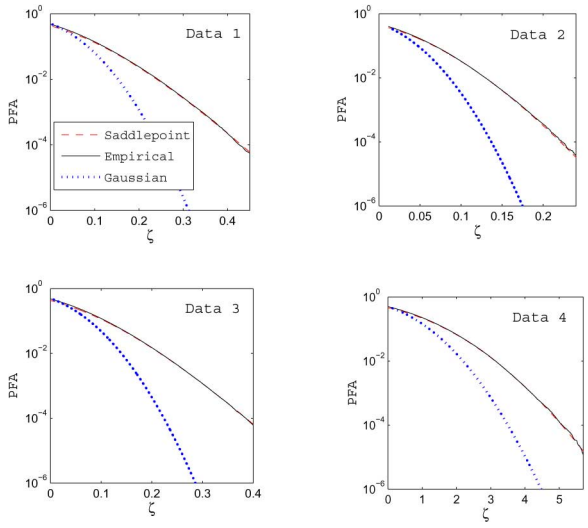

Fig. 7. PFA computed under the Saddlepoint approximation and the Gaussian approximation with empirical estimations of its mean and variance, and compared to the empirical distribution.

- Data 3: new correlation parameter $\rho' = 0.95$ that gives a correlation pattern size of 14 pixels.
- Data 4: new background mean $\boldsymbol{\mu}' = 100\,\boldsymbol{\mu}$, so that typically 100 photoelectrons can be expected per pixel instead of a few; new source flux $\alpha_0' = 20\alpha_0$.

Fig. 6 shows the receiver operating characteristic (ROC) curve associated to the detection scheme (29) for these four different simulation cases. Each curve is simply estimated from the 10^6 samples of $\hat{\alpha}$ under both hypotheses.

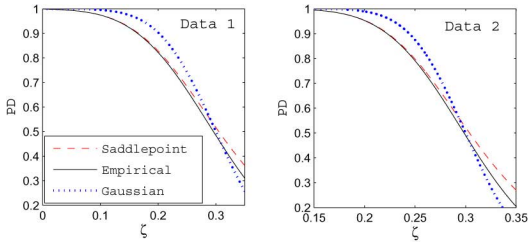

Fig. 8. $PD(r_0)$ computed under the Saddlepoint approximation and the Gaussian approximation, and compared to the empirical distribution.

B. Evaluation of PFA and PD Estimations

Fig. 7 shows $PFA(r_0)$ as a function of the threshold, for the detection process described in Section IV-A. It is computed from different estimations:

- the saddlepoint approximation as described in Section IV-B;
- the empirical distribution of $\hat{\alpha}$ under H_0;
- a normal distribution, with moments that are estimated by the sample moments, which are equal, up to a relative precision of 10^{-4}, to the true moments $\mathsf{E}[\hat{\alpha}] = \alpha_0$ and $\mathsf{var}(\hat{\alpha})$ given by (27).

In all cases, from about 10^{-4} to 10^{-1}, the saddlepoint plot perfectly matches the empirical plot, unlike the normal approximation. Actually, the correct behavior of the saddlepoint approximation and the misbehavior of the normal approximation extends for PFA even lower than 10^{-10}. The relative error of the normal approximation for a given PFA is quite stable for all studied cases. This shows that the saddlepoint plot may be considered as being almost a perfect estimate of the PFA if the PFA of interest is lower than 10^{-1} and that the normal approximation should be avoided by all mean.

Fig. 8 shows $PD(r_0)$ as a function of the threshold, computed as in Fig. 7, but using the second expression (31) to compute the saddlepoint approximation. The saddlepoint approximation diverges from the true PD as the threshold leaves the tail of the distribution of $\hat{\alpha}$, but remains a better approximation than the normal approximation for $PD < 0.6$. This qualitative behavior and the quantitative values are quite stable for the two cases shown, despite the fact that they represent quite different detectability cases, as seen on Fig. 6. The saddlepoint approximation can be considered as a almost perfect estimate for $PD > 0.8$ and should still be preferred to the normal approximation for $PD > 0.6$.

Fig. 9 still shows PD as a function of the threshold, but using also a development around the normal approximation, as done with the Gram-Charlier or the Edgeworth series. However the use of two additional terms in the Gram-Charlier series (the first being the previous Gaussian approximation) does not make an effective difference. Since the calculation of the Gram-Charlier terms involves the computation of high order cumulants, it is too computationally demanding to look further in the development. Also note that when truncated after the third term, the Gram-Charlier series and the Edgeworth series are identical.

Finally, Fig. 10 shows that for cases where one may expect a high detection probability for a reasonable PFA (see 6 Data 4), the saddlepoint approximation remains a very satisfactory

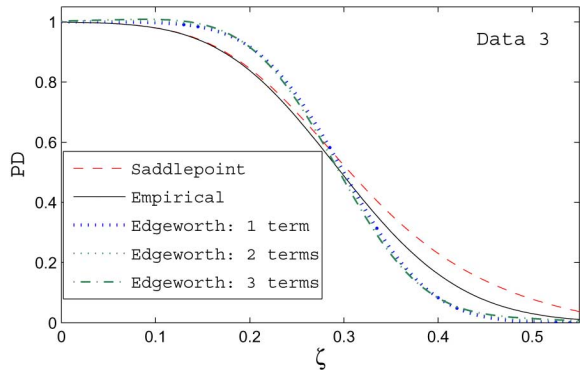

Fig. 9. $PD(r_0)$ computed under the Saddlepoint approximation and the Gram-Charlier series truncated after the first term (Gaussian approximation), second term, and third term, and compared to the empirical distribution. Focus on a linear range for PD.

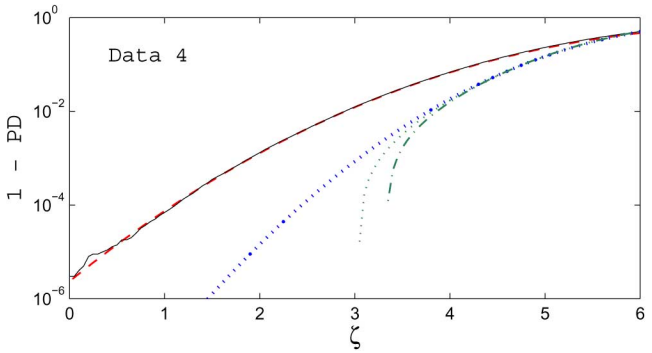

Fig. 10. $PD(r_0)$ computed under the Saddlepoint approximation and the Gram-Charlier series truncated after the first term (Gaussian approximation), second term and third term, and compared to the empirical distribution. Focus on the extreme range $PD \approx 1$.

way to estimate PD. The normal approximation is also a bad approximation. This was expected since the range analysed here is equivalent to the one analysed for PFA. Also note that a further development of the Gram-Charlier series gives even worse estimations than the simple gaussian approximation.

The poor performances of the normal approximation for $\hat{\alpha}$ that is observed in the previous Figs. 7–10 can be directly seen from the random samples of $\hat{\alpha}$ shown on Fig. 5. Application of the Kolmogorov-Smirnov and Chi-Squared tests to $\hat{\alpha}|H_0$ and $\hat{\alpha}|H_1$ rejects the normal hypothesis with a probability higher than $0.9999\ldots$

Tools commonly used to compare an empirical distribution to the normal distribution are the skewness γ_1 and the (excess) kurtosis γ_2: both parameters are 0 for a normal distribution. Here, the empirical distributions of $\hat{\alpha}$ are slightly right skewed, with $\hat{\gamma}_1(H_0) \simeq 0.35$ under H_0, $\hat{\gamma}_1(H_1) \simeq 0.24$ under H_1, $\hat{\gamma}_2(H_0) \simeq 0.33$ and $\hat{\gamma}_2(H_1) \simeq 0.18$. Some sense can be given to these values by relating them to corresponding values for the Poisson distribution. It is quite relevant since $\mathcal{P}(\lambda)$ has a skewness $\gamma_1(\lambda) = \lambda^{-1/2}$ and a kurtosis $\gamma_2(\lambda) = \lambda^{-1}$ that have the same sign as $\hat{\gamma}_1$ and $\hat{\gamma}_2$, and it tends to a Gaussian distribution as $\lambda \to \infty$. We observe that

- $\hat{\gamma}_1(H_0) > \gamma_1(10) \simeq 0.32$ and $\hat{\gamma}_2(H_0) > \gamma_2(10) = 0.1$. So under H_0 the empirical distribution is in the skewness and kurtosis sense "less Gaussian" than $\mathcal{P}(10)$.
- $\hat{\gamma}_1(H_1) > \gamma_1(20) \simeq 0.22$, and the sample kurtosis was even greater than $\gamma_2(10) > \gamma_2(20)$. So under H_1, the empirical distribution is in that same sense "much less Gaussian" than $\mathcal{P}(20)$.

VI. Conclusion and Outlook

The goal of this paper was the detection-estimation of a moving source from a set of images. The quite realistic model mainly relied on the modeling of the incoming complex amplitude by a circular (correlated) Gaussian distribution, and also took into account the Poisson process and common electronic noises.

The derivation of an intensity and a position estimator for the potential source, and the derivation of a statistics for the detection task were intractable from the realistic model. However, under a high flux assumption and after renormalization, such a model was shown to reduce to a simple Gaussian model. Applying a simple likelihood approach on this model, practical estimators were derived, and shown to have some consistent properties even under the quite realistic model.

As a detection test, the intensity estimator was chosen as the test statistics to threshold, for a given initial position. Thanks to the linearity of this detector, and thanks to the tractability of the MGF of the accurate model, it was possible to estimate the Probability of False Alarms and the Probability of Detection by using the saddlepoint approximation.

An application of this estimation on a 1-D data model proved that the saddlepoint approximation was very accurate, and confirmed the fact that when the high flux approximation is not reached, a Gaussian approximation or even a Gram-Charlier series approximation does not lead to a satisfactory estimation of PFA and PD.

However, implicit assumptions made in this work should be relaxed to test the robustness of the statistical properties of the estimators, the performance of the detector and the accuracy of the estimations of PFA and PD. In particular, the effect of time variation and/or randomness should be studied for the mean of the background $\boldsymbol{\mu}$. Furthermore, the computation of PFA and PD depends on $\boldsymbol{\mu}$ and the covariance matrix Σ, and they were assumed to be known here. Actually, they have to be estimated either from the data themselves (with some additional priors) or from some calibration measures. These issues will be adressed in some future work.

Appendix

Gaussian Approximation of the Detailed Data Model: This appendix aims at showing that under a *high flux approximation* , the distribution of the vector $\boldsymbol{n}(k)$ that verifies (7) with (4) and (5), converges after a proper normalization to a standard normal distribution. Like the parameters of the distribution of $\boldsymbol{n}(k)$, the convergence does not depend on k, so the dependence on k will be dropped from now.

The proof is related to the following notations and lemma.

Definition 1: Consider a $M \times 1$ random vector $\boldsymbol{v}$ and a set $\mathcal{B}$ of integers in $[1, M]$ that can be repeated. Denote the cardinal of $\mathcal{B}$ by $c_{\mathcal{B}}$ and the set of coordinates indexed by all $m \in \mathcal{B}$ by

$\boldsymbol{v}_m : m \in B$. Finally, define as $\mathcal{I}$ the set of the different indices k contained in $\mathcal{B}$, c its cardinal and r_k the number of occurrences of the index k in $\mathcal{B}$, so that $c_{\mathcal{B}} = \sum_{k=1}^{c} r_k$. Denoting by $\boldsymbol{z}$ a $c \times 1$ vector, let's define the cumulant $\kappa(\boldsymbol{v}_m : m \in \mathcal{B})$ as

$$\kappa(\boldsymbol{v}_m : m \in \mathcal{B}) = \kappa_{r_k : k \in \mathcal{I}}(\boldsymbol{v}_k : k \in \mathcal{I}) \quad (36)$$

$$= (-1)^{c_{\mathcal{B}}} \frac{\partial^{c_{\mathcal{B}}} \ln \tilde{h}_{\mathcal{I}}(\boldsymbol{z})}{\partial \boldsymbol{z}_1^{r_1} \partial \boldsymbol{z}_2^{r_2} \dots \partial \boldsymbol{z}_c^{r_c}} |_{\boldsymbol{z}=\boldsymbol{0}} \quad (37)$$

with $r_k : k \in \mathcal{I}$ the list of orders of the cumulant, and $\tilde{h}_{\mathcal{I}}$ the Laplace transform of the joint pdf of the c random scalars $\boldsymbol{v}_k$, $k \in \mathcal{I}$.

Lemma 1: Consider a set $\mathcal{B}$ of integers in $[1, M]$ that can be repeated, V a $M \times M$ matrix and $\boldsymbol{n}$ a $M \times 1$ random vector that satisfies $\boldsymbol{n}_\ell | \boldsymbol{u}_\ell \sim \mathcal{P}(\boldsymbol{u}_\ell)$ and $\boldsymbol{n}_\ell | \boldsymbol{u}$ and $\boldsymbol{n}_{\ell'} | \boldsymbol{u}$ are independent for $\ell \neq \ell'$. The joint cumulant of $(V\boldsymbol{n})_m : m \in \mathcal{B}$ conditioned on the vector $\boldsymbol{u}$ equals

$$\kappa((V\boldsymbol{n})_m : m \in \mathcal{B} | \boldsymbol{u}) = \sum_{m \in B} \nu_m \boldsymbol{u}_m \quad (38)$$

where the ν_m only depend on $\mathcal{B}$ and V.

Proof: The multilinearity of cumulants implies that $\kappa((V\boldsymbol{n})_m : m \in B | \boldsymbol{u})$ is a linear combination of the cumulants $\kappa(\boldsymbol{n}_m : m \in \mathcal{B}_i | \boldsymbol{u})$ where $\mathcal{B}_i$ is any set of integers in $[1, M]$ such that its cardinal is the same as $\mathcal{B}$. See, for example, [30] for the expression using tensors.

There are two types of set $\mathcal{B}_i$ as follows.

- If $\mathcal{B}_i$ contains at least two different indexes, $\kappa(\boldsymbol{n}_m : m \in \mathcal{B}_i | \boldsymbol{u}) = 0$.
- If all the subscripts in $\mathcal{B}_i$ equal ℓ, $\boldsymbol{n}_\ell | \boldsymbol{u}_\ell$ being Poisson distributed with parameter $\boldsymbol{u}_\ell$ we have

$$\kappa(\boldsymbol{n}_m : m \in \mathcal{B}_i | \boldsymbol{u}) = \kappa_{c_{\mathcal{B}}}(\boldsymbol{n}_\ell | \boldsymbol{u}_\ell) = \boldsymbol{u}_\ell. \quad (39)$$

■

Theorem 1: Let's assume a vector $\boldsymbol{n}$ satisfies (7) with (4) and (5), with $\boldsymbol{\mu} = v\boldsymbol{\mu}^\circ, v \in \mathbb{R}$. The properly standardized vector $\boldsymbol{n}$ converges in distribution to $\mathcal{N}(\boldsymbol{0}, I)$ when $v^2 \to \infty$.

Proof: First $\boldsymbol{n}$ needs to be standardized. From (1) and (6), the mean and the covariance matrix of $\boldsymbol{n}$ are related to those of $\boldsymbol{u}$ through $\mathsf{E}[\boldsymbol{n}] = \mathsf{E}[\boldsymbol{u}]$ and $\mathsf{cov}(\boldsymbol{n}) = \mathsf{cov}(\boldsymbol{u}) + D_{\mathsf{E}[\boldsymbol{u}]}$. But those ones are computed in [22]. Replacing $\boldsymbol{\mu}$ by $v\boldsymbol{\mu}^\circ$ shows finally that all the cumulants of $\boldsymbol{n}$ are linear functions of v^2

$$\mathsf{E}[\boldsymbol{n}] = \boldsymbol{\alpha} + v^2\boldsymbol{\beta}, \quad \mathsf{cov}(\boldsymbol{n}) = A + v^2 B \quad (40)$$

where the vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and the matrices A and B are not functions of v

$$\boldsymbol{\alpha} = (\Sigma_{11}, \dots, \Sigma_{MM})^t, \quad \boldsymbol{\beta} = ((\boldsymbol{\mu}_1^\circ)^2, \dots, (\boldsymbol{\mu}_M^\circ)^2)^t \quad (41)$$

$$A = \left(|\Sigma_{\ell,\ell'}|^2\right)_{\ell,\ell'=1,\dots,M} + D_{(\Sigma_{11},\dots,\Sigma_{MM})^t} \quad (42)$$

$$B = 2\Re[D^*_{\boldsymbol{\mu}^\circ} \Sigma D_{\boldsymbol{\mu}^\circ}] + D_{(\boldsymbol{\mu}_1^\circ)^2,\dots,(\boldsymbol{\mu}_M^\circ)^2}. \quad (43)$$

The matrices A and B are symmetric positive defined (e.g., A is the covariance of $\boldsymbol{n}$ when $v = 0$). Consequently, there exist a non singular matrix V and a diagonal matrix $D_{\boldsymbol{a}}$ such that, [31]

$$VAV^t = D_{\boldsymbol{a}}, \quad VBV^t = I.$$

Then, $\boldsymbol{n}$ can be standardized in a vector $\tilde{\boldsymbol{n}}$ ($\mathsf{E}[\tilde{\boldsymbol{n}}] = 0$ and $\mathsf{cov}(\tilde{\boldsymbol{n}}) = I$) in such a way that v^2 appears unmixed in the normalization matrices

$$\tilde{\boldsymbol{n}} = \left(D_{\boldsymbol{a}} + v^2 I\right)^{-1/2} V(\boldsymbol{n} - \boldsymbol{\alpha} - v^2\boldsymbol{\beta}). \quad (44)$$

To prove the asymptotic Gaussianity of $\tilde{\boldsymbol{n}}$ as $v^2 \to \infty$, we need to show that all cumulants of total order higher than 2 tend to 0 and that the lower order cumulant converge.

The second property directly arises from the standardization: for any v, any k and any $k' \neq k$, $\kappa(\tilde{\boldsymbol{n}}_k) = 0$, $\kappa_2(\tilde{\boldsymbol{n}}_k) = 1$ and $\kappa(\tilde{\boldsymbol{n}}_k, \tilde{\boldsymbol{n}}_{k'}) = 0$.

Then, let's define a set $\mathcal{S}$ of integers in $[1, M]$ such that $\mathrm{Card}(\mathcal{S}) > 2$, and show that $\kappa(\tilde{\boldsymbol{n}}_k : k \in \mathcal{S}) \to 0$. This can be done using the "law of total cumulance," [32]

$$\kappa(\tilde{\boldsymbol{n}}_k : k \in \mathcal{S}) = \sum_{\pi} \kappa(\kappa(\tilde{\boldsymbol{n}}_m : m \in B | \boldsymbol{u}) : B \in \pi) \quad (45)$$

where the sum is over all partitions π of $\mathcal{S}$ and B runs through the whole list of "blocks" of the partition π. Substituting (44) in (45), since $(D_{\boldsymbol{s}} V\boldsymbol{n})_m = \boldsymbol{s}_m (V\boldsymbol{n})_m$, since $\kappa(a_1, \dots, a_N) = 0$ for $a_1, \dots, a_N$ N deterministic variables with $N > 1$, and by multilinearity of the cumulant, we have

$$\kappa(\tilde{\boldsymbol{n}}_k : k \in \mathcal{S}) = \sum_{\pi} \kappa\left((f_B(\boldsymbol{a}) g_B(\boldsymbol{u})) : B \in \pi\right) \quad (46)$$

$$\text{where } f_B(\boldsymbol{a}) = \prod_{m \in B} (\boldsymbol{a}_m + v^2)^{-1/2} \quad (47)$$

$$g_B(\boldsymbol{u}) = \kappa((V\boldsymbol{n})_m : m \in B | \boldsymbol{u}). \quad (48)$$

From multilinearity and lemma 1

$$\begin{aligned} &\kappa(\tilde{\boldsymbol{n}}_k : k \in \mathcal{S}) \\ &= \sum_{\pi} \left(\prod_{B \in \pi} f_B(\boldsymbol{a}) \right) \kappa\left(g_B(\boldsymbol{u}) : B \in \pi\right) \\ &= f_{\mathcal{S}}(\boldsymbol{a}) \sum_{\pi} \kappa\left(\sum_{m \in \mathcal{B}} \nu_m \boldsymbol{u}_m : B \in \pi \right) \\ &= f_{\mathcal{S}}(\boldsymbol{a}) \sum_{\pi} \sum_{m \in \mathcal{B}} \nu_{m_1} \dots \nu_{m_{c_\pi}} \kappa(\boldsymbol{u}_{m_1}, \dots, \boldsymbol{u}_{m_{c_\pi}}) \end{aligned} \quad (49)$$

where the last equality arises from multilinearity and denoting by c_π the number of blocks B in π.

But the cumulants of $\boldsymbol{u}$ are related to the coefficients of the Taylor expansion of the log-MGF, [30]. Replacing in the logarithm of (4) $\boldsymbol{\mu}$ by $v\boldsymbol{\mu}^\circ$ shows that all the cumulants of $\boldsymbol{u}$ are linear functions of v^2. So $\kappa\left(\sum_{m \in \mathcal{B}} \nu_m \boldsymbol{u}_m : B \in \pi\right)$ is a linear function of v^2 and finally we have

$$\kappa(\tilde{\boldsymbol{n}}_k : k \in \mathcal{S}) = \frac{cv^2 + d}{\prod_{k \in S} \sqrt{\boldsymbol{a}_k + v^2}}.$$

This equation proves that if $\mathrm{Card}(\mathcal{S}) > 2$ $\kappa(\tilde{\boldsymbol{n}}_k : k \in \mathcal{S}) \to 0$ when $v^2 \to \infty$. ■

According to this result, when v is large, the distribution of $\boldsymbol{n}$ can be approximated by a Gaussian distribution with mean $\boldsymbol{\alpha} + v^2\boldsymbol{\beta} \approx v^2\boldsymbol{\beta}$ and covariance $A + v^2 B \approx v^2 B$ where $\boldsymbol{\beta}$ is given by (41) and B is given by (43).

It is worthy to note that

$$\Re[D^*_{\boldsymbol{\mu}^\circ}\Sigma D_{\boldsymbol{\mu}^\circ}] \leq ||\Sigma||_F ||D_{\left((\boldsymbol{\mu}_1^\circ)^2,\ldots,(\boldsymbol{\mu}_M^\circ)^2\right)}||_F \leq ||\Sigma||_F ||\boldsymbol{\beta}||_1. \tag{50}$$

Hence if $||\Sigma||_F||\boldsymbol{\beta}||_1 \ll 1$, the first term in B (43) can be neglected with respect to the second one, which is diagonal. Therefore, when the residual speckles are very low $\boldsymbol{n}$ will be approximately independent, in addition to being Gaussian.

Positivity Assumption on the Intensity Estimator $\hat{\alpha}_{ML}$: This appendix is part of the ML derivation of the intensity and position estimator derived in Section III-B.

Lemma 2: Denote $\hat{\alpha}_{ML}$ and $\hat{r}_{ML}$ the ML intensity and position estimators for the Gaussian data model (10). If we assume that $\hat{\alpha}_{ML} > 0$, then $\hat{r}_{ML}$ reduces to

$$\hat{r}_{ML} = \arg \max_{r=1,\ldots,M} \left\{ \sqrt{c_N(r)}\hat{\alpha}(r) \right\}. \tag{51}$$

Proof: Recall $\hat{\alpha}(r)$ is given by (12) and $c_N(r) \geq 0$ is given by (16). Define

$$f_1(r) = c_N(r)\hat{\alpha}(r)^2; \quad f_2(r) = \sqrt{c_N(r)}\hat{\alpha}(r) \tag{52}$$

and r_1 and r_2 their respective maxima

$$f_i(r_i) \geq f_i(r) \quad \forall\, r. \tag{53}$$

With these notations, $\hat{r}_{ML} = r_1$ and $\hat{\alpha}_{ML} = \hat{\alpha}(r_1)$. From the assumption $\hat{\alpha}_{ML} > 0$ and (53) applied to f_2 for $r = r_1$,

$$f_2(r_2) \geq f_2(r_1) > 0; \quad f_1(r_2) \geq f_1(r_1) > 0. \tag{54}$$

Finally, (53) applied to f_1 for $r = r_2$ and (54) lead to $r_1 = r_2$ if f_1 is bijective. f_1 is bijective with probability 1 since r takes a finite number of values and $\hat{\alpha}$ is a real random variable not fully correlated from a pixel to another. ■

REFERENCES

[1] J. Goodman, *Speckle Phenomena in Optics*. Greenwood Village, CO: Roberts and Company, 2007.

[2] F. Wildi *et al.*, R. Tyson and M. Lloyd-Hart, Eds., "The SPHERE exoplanet imager: Status report at PDR," in *SPIE Proc.*, Sep. 2007, vol. 6691.

[3] B. Macintosh *et al.*, "The Gemini planet imager," in *Advances in Adaptative Optics II*, B. Ellerbroek and D. B. Calia, Eds. Philadelphia, PA: SPIE, 2006.

[4] C. Aime and R. Soumer, "The effect of a coronagraph on the statistics of adaptative optics pinned speckles," in *Astronomy With High Contrast Imaging III: Instrumental Techniques, Modeling and Data Processing*, ser. EAS. Les Ulis, France: EDP Sciences, 2004, vol. 12.

[5] G. Avila and K. Wirenstrand, Field and Pupil Rotations for the VLT 8 m Unit Telescopes European Southern Observatory, 2005, Tech. Rep.

[6] C. Marois, D. Lafrenière, R. Doyon, B. Macintosh, and D. Nadeau, "Angular differential imaging: A powerful high-contrast imaging technique," *The Astrophys. J.*, vol. 641, pp. 556–564, Apr. 2006.

[7] L. Mugnier, J.-F. Sauvage, T. Fusco, and G. Rousset, "Multi-channel planet detection algorithm for angular differential imaging," in *Adaptive Optics: Analysis and Methods*, B. Ellerbroek, Ed., Vancouver, Canada, Jun. 2007, Optical Society of America.

[8] M. Carbillet, C. Vérinaud, B. Femenía, A. Riccardi, and L. Fini, "Modeling astronomical adaptive optics: I. The software package CAOS," *Month. Notices of Royal Astronom. Soc.*, 2004.

[9] M. Carbillet *et al.*, "The software package sphere: A numerical tool for end-to-end simulations of the VLT instrument SPHERE," in *Adaptive Optics Systems*, N. Hubin, C. E. Max, and P. L. Wizinowich, Eds. Philadelphia, PA: SPIE, 2008.

[10] M. Carbillet *et al.*, "Caos: A numerical simulation tool for astronomical adaptive optics (and beyond)," in *SPIE Proc.*, 2004, vol. 5490, pp. 637–637.

[11] A. Boccaletti *et al.*, "End-to-end simulations of adaptive-optics-assisted coronagraphic differential imaging: Estimation of performance for SPHERE," in *Adapt. Opt. Syst.*, N. Hubin, C. E. Max, and P. L. Wizinowich, Eds. Philadelphia, PA: SPIE, 2008.

[12] C. Aime and R. Soummer, "Influence of speckle and Poisson noise on exoplanet detection with a coronagraph," presented at the Eur. Signal Process. Conf., 2004.

[13] R. Soummer, A. Ferrari, C. Aime, and L. Jolissaint, "Speckle noise and dynamic range in coronagraphic images," *The Astrophys. J.*, vol. 669, pp. 642–656, Nov. 2007.

[14] B. Picinbono, *Random Signals and Systems*. Englewood Cliffs, NJ: Pearson US Imports & PHIPEs, 1993.

[15] A. Ferrari, "Analytical analysis of lyot coronographs response," *The Astrophys. J.*, vol. 657, pp. 1201–1209, Mar. 2007.

[16] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. New York: Wiley, 1995, vol. 2.

[17] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, ser. Wiley Series in Probability and Mathematical Statistics. New York: Wiley, 1982.

[18] G. Letac and H. Massam, A "Tutorial on noncentral Wishart distributions," Lab. Stat. Prob., Toulouse, France, Tech. Rep., 2004.

[19] J.-Y. Tourneret, A. Ferrari, and G. Letac, "The noncentral Wishart distribution: Properties and application to speckle imaging," in *Proc. IEEE Statist. Signal Process. Workshop*, 2005.

[20] A. Ferrari, "Modeling for direct imaging of exoplanets data processing," in *Astronomy With High Contrast Imaging III: Instrumental Techniques, Modeling and Data Processing*, ser. EAS. Les Ulis, France: EDP Sciences, 2006, vol. 22.

[21] J. Goodman, *Statistical Optics*. New York: Wiley, 1985.

[22] A. Ferrari, G. Letac, and J.-Y. Tourneret, "Multivariate mixed Poisson distributions," in *Proc. EUSIPCO*, Vienna, Austria, 2004.

[23] A. Philippou and G. Roussas, "Asymptotic normality of the maximum likelihood estimate in the independent not identically distributed case," *Ann. Instit. Statist. Math.*, vol. 27, no. 1, pp. 45–55, 1975.

[24] B. Hoadley, "Asymptotic properties of maximum likelihood estimators for the independent not identically distributed case," *The Ann. Math. Statist.*, vol. 42, no. 6, pp. 1977–1991, 1971.

[25] T. J. Sweeting, "Uniform asymptotic normality of the maximum likelihood estimator," *The Ann. Statist.*, vol. 8, no. 6, pp. 1375–1381, 1980.

[26] P. J. Bickel and K. A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2000, vol. I.

[27] A. Swindlehurst and P. Stoica, "Maximum likelihood methods in radar array signal processing," *Proc. IEEE*, vol. 86, no. 2, pp. 421–441, Feb. 1998.

[28] C. W. Helstrom, *Elements of Signal Detection and Estimation*. Englewood Cliffs, NJ: PTR Prentice-Hall, 1995.

[29] S. A. Aldosari and J. Moura, "Detection in sensor networks: The saddlepoint approximation," *IEEE Trans. Signal Process.*, vol. 55, pp. 327–340, Jan. 2007.

[30] P. McCullagh, *Tensor Methods in Statistics*, ser. Monographs on statistics and applied probability. London, U.K.: Chapman and Hall, 1987.

[31] G. H. Golub and C. F. V. Loan, *Matrix Computations*, ser. Johns Hopkins Series in the Mathematical Sciences, 2nd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1989.

[32] D. Brillinger, "The calculation of cumulants via conditioning," *Ann. Instit. Statist. Math.*, vol. 21, pp. 215–218, 1969.

Isabelle Smith (S'08) received the M.Sc. degree in physics from the University of Paris 7 (Jussieu), France, in 2006.

She spent one year of graduate study at Cornell University, Ithaca, NY, and one with the Meudon observatory. Since then, she has been enrolled in a 3-year Ph.D. program with the University of Nice Sophia-Antipolis, France, in the Fizeau Astrophysics Laboratory. Her main interest has been estimation and detection in astrophysical contexts using analytical approaches.

André Ferrari (S'91–M'93) received the Ingénieur degree from École Centrale de Lyon, France, in 1988, the M.Sc. degree in 1989, and the Ph.D. degree from the University of Nice Sophia-Antipolis (UNS), France, in 1992, all in electrical and computer engineering.

He is currently a Professor with the UNS. He is a member of the Fizeau Laboratory (UNS, the Centre National de la Recherche Scientifique—CNRS, and the Observatoire de la Côte d'Azur—OCA), where his research activity is centered around statistical signal processing and modeling, with a particular interest in processing of optical data with applications to astrophysics.

Marcel Carbillet received the D.E.A. (M.Sc.) degree in 1993 and the Ph.D. degree in 1996, both from the Université de Nice Sophia Antipolis (UNS, France) and in the field of optical astronomy.

He then spent one year with the Université de Genève, Switzerland, and seven years with the Osservatorio di Arcetri, Italy. He currently has a permanent position of Maître de Conférences at the Fizeau Laboratory (UNS, the Observatoire de la Côte d'Azur—OCA, and the Centre National de la Recherche Scientifique— CNRS). His research activities are focused on astronomical adaptive optics (AO), high-resolution imaging (especially post-AO data), and high-contrast imaging (especially concerning exoplanets).