Fast and robust exo-planet detection in multi-spectral, multi-temporal data

Éric Thiébaut[a], Loïc Denis[b], Laurent M. Mugnier[c], André Ferrari[d], David Mary[d], Maud Langlois[a], Faustine Cantalloube[c], and Nicholas Devaney[e]

[a]Univ. Lyon, Univ. Lyon 1, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230, Saint-Genis-Laval, France
[b]Univ Lyon, UJM-Saint-Étienne, CNRS, Institut d'Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, Saint-Étienne, France
[c]ONERA – The French Aerospace Lab, F-92322 Châtillon, France
[d]Observatoire de la Côte d'Azur, Laboratoire Lagrange, Université de Nice, France
[e]School of Physics, National University of Ireland, Galway, Ireland

ABSTRACT

Exo-planet detection is a signal processing problem that can be addressed by several detection approaches. This paper provides a review of methods from detection theory that can be applied to detect exo-planets in coronographic images such as those provided by SPHERE and GPI. In a first part, we recall the basics of signal detection and describe how to derive a fast and robust detection criterion based on a heavy tail model that can account for outliers in the residuals. In a second part, we derive detectors that handle jointly several wavelengths and exposures and focus on an approach that prevents from interpolating the data, thereby preserving the statistics of the original data.

Keywords: Exo-planet detection; multi-variate data processing.

The detection of exo-planets in coronographic images is one of the most challenging data processing task for new instruments like SPHERE and GPI. The difficulty is that the planets must be sought in images where the stellar leakage forms speckles which largely dominate the planetary signal.[1,2] In this context, robust detection methods may be more suitable as they can take into account the imperfect speckle suppression achieved by image processing methods based on image differences (ADI, SDI) or principal component analysis (PCA) for instance. We recently proposed[3] to use Cauchy distributions (rather than Gaussian ones) to achieve robust detection and showed that the corresponding Locally Most Powerful test (LMP) can then be computed on large images by means of FFTs. We have generalized this to other distribution than Cauchy.[4] In this contribution, we first show how to generalize the LMP test to multi-spectral and multi-temporal data, in the case of Gaussian noise this closely follows what has been proposed for DARWIN.[5]

1. ROBUST DETECTION OF A KNOWN PATTERN IN A SINGLE IMAGE

1.1 Detection Principles

Detecting an object given the measurements $y \in \mathbb{R}^M$ amounts to deciding between the following two hypotheses:

$$\mathscr{H}_0\colon \quad y = n\,, \tag{1}$$

$$\mathscr{H}_1\colon \quad y = \alpha\, m(\theta) + n\,, \tag{2}$$

where $n \in \mathbb{R}^M$ is a nuisance term (noise and irrelevant background), $m(\theta) \in \mathbb{R}^M$ is the footprint of the object and α is the amplitude of the signal. The footprint of the object non-linearly depends on the parameters θ. For example, in the case of the detection of exo-planets, θ is the position of the planet and $m_i(\theta) \propto h(\xi_i, \theta)$ with ξ_i the position of the i-th pixel and $h(\xi, \theta)$ the point spread function (PSF) at detector position ξ for a source at position θ.

E-mail: `eric.thiebaut@univ-lyon1.fr`

Given the object parameters, here α and θ, the following likelihood ratio test (LRT) is optimal to decide between the two hypotheses:

$$\text{LRT:} \quad \frac{p(y \,|\, \mathscr{H}_1)}{p(y \,|\, \mathscr{H}_0)} = \frac{p_n(y - \alpha\, m(\theta))}{p_n(y)} \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \gamma\,, \tag{3}$$

where $p(y \,|\, \mathscr{H}_k)$ is the likelihood of the data under hypothesis $\mathscr{H}_k$, p_n is the probability density function of the nuisance terms and $\gamma > 0$ is a chosen threshold. In words, the detection is decided if it is γ times more likely that there is an object (with given parameters α and θ) rather than no object. The LRT, also known as Neyman-Pearson *clairvoyant* test, minimizes the probability of error for any fixed probability of detection but requires to know the object parameters, here α and θ. In practice none of these parameters are known and the test must be relaxed to account for this. Substituting the unknown parameters by their maximum likelihood values leads to the generalized likelihood ratio test (GLRT):

$$\text{GLRT:} \quad \frac{\max\limits_{\alpha,\theta} p(y \,|\, \alpha, \theta; \mathscr{H}_1)}{p(y \,|\, \mathscr{H}_0)} = \frac{p(y \,|\, \widehat{\alpha}, \widehat{\theta}; \mathscr{H}_1)}{p(y \,|\, \mathscr{H}_0)} \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \gamma\,, \tag{4}$$

where:

$$\big\{\widehat{\alpha}, \widehat{\theta}\,\big\} = \max_{\alpha,\theta} p(y \,|\, \alpha, \theta; \mathscr{H}_1)\,, \tag{5}$$

are the maximum likelihood estimators for the parameters α and θ.

1.2 Fast Detection for Independent Gaussian Noise

Assuming the nuisance terms follow a centered Gaussian distribution, say $n \sim \mathcal{N}(0, \Gamma)$, and taking the logarithm (which is a monotonous function) of the likelihood ratio, the LRT writes:

$$\text{LRT:} \quad \log\!\left(\frac{p(y \,|\, \mathscr{H}_1)}{p(y \,|\, \mathscr{H}_0)}\right) = \alpha\, m(\theta)^{\mathrm{t}}\, W\, y - \frac{\alpha^2}{2}\, m(\theta)^{\mathrm{t}}\, W\, m(\theta) \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \tau\,, \tag{6}$$

with $\tau = \log \gamma$ and $W = \Gamma^{-1}$ the inverse of the covariance matrix of the nuisance terms. The expression in the left hand side of the test can be maximized in α to yield the maximum likelihood amplitude conditioned to the knowledge of θ:

$$\widehat{\alpha}(\theta) = \max_{\alpha} p(y \,|\, \alpha, \theta; \mathscr{H}_1) = \frac{m(\theta)^{\mathrm{t}}\, W\, y}{m(\theta)^{\mathrm{t}}\, W\, m(\theta)}\,. \tag{7}$$

Then replacing α by $\widehat{\alpha}(\theta)$ yields the following GLRT:

$$\text{GLRT:} \quad \max_{\theta} \left\{ T_{\mathsf{GLR}}(\theta) = \frac{\big|m(\theta)^{\mathrm{t}}\, W\, y\big|}{\sqrt{m(\theta)^{\mathrm{t}}\, W\, m(\theta)}} \right\} \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \eta\,, \tag{8}$$

with $\eta = \sqrt{2\tau}$. The criterion $T_{\mathsf{GLR}}(\theta)$ is in general not convex in θ, thus solving the above problem requires to implement a global optimization strategy. An important feature would then to be able to quickly compute the following terms:

$$a(\theta) = m(\theta)^{\mathrm{t}}\, W\, m(\theta)\,, \tag{9a}$$

$$b(\theta) = m(\theta)^{\mathrm{t}}\, W\, y\,. \tag{9b}$$

In the case of stationary PSF*, *i.e.* $h(\xi, \theta) = h(\xi - \theta)$, and independent noise, *i.e.* $W = \mathrm{diag}(w)$, the above terms amount to computing correlation products at position θ. Sampling these correlation products at $\theta = \xi_i$ for all pixel indices i can be done very quickly by means of the fast Fourier transform (FFT):

$$a(\theta) = \sum_i h^2(\xi_i - \theta)\, w_i \qquad \Longrightarrow \qquad a(\xi_i) = [(h \otimes h) \circledast w]_i\,, \tag{10a}$$

$$b(\theta) = \sum_i h(\xi_i - \theta)\, w_i\, y_i\,. \qquad \Longrightarrow \qquad b(\xi_i) = [h \circledast (w \otimes y)]_i\,, \tag{10b}$$

*or more generally for the detection of a shift invariant footprint where w_i is the weight of the i-th pixel, $\otimes$ denotes the componentwise multiplication and $\circledast$ denotes the discrete 2D correlation over the grid of pixels. To avoid aliasing in these computations, it is necessary to properly zero-pad[3,6] the *images* w, y and h. The most likely position of a source, at the pixel precision, is then $\xi_{\widehat{\imath}}$ where $\widehat{\imath}$ is given by:

$$\widehat{\imath} = \arg\max_i \frac{|a(\xi_i)|}{\sqrt{b(\xi_i)}} .$$

1.3 Accounting for the Positivity of the Brightness

For exo-planet detection, the parameter α represents the brightness of the source and necessarily $\alpha \geq 0$. The maximum likelihood nonnegative amplitude conditioned to the knowledge of the position θ is now given by:

$$\widehat{\alpha}(\theta) = \max_{\alpha \geq 0} p(y \,|\, \alpha, \theta; \mathscr{H}_1) = \frac{\max\bigl(0, m(\theta)^{\mathrm{t}}\, W\, y\bigr)}{m(\theta)^{\mathrm{t}}\, W\, m(\theta)} = \frac{\bigl(m(\theta)^{\mathrm{t}}\, W\, y\bigr)_+}{m(\theta)^{\mathrm{t}}\, W\, m(\theta)} , \tag{11}$$

where $(\ldots)_+ \stackrel{\text{def}}{=} \max(0, \ldots)$. Replacing α by $\widehat{\alpha}(\theta)$ yields the following GLRT:

$$\text{GLRT:} \quad \max_\theta \left\{ T_{\mathsf{GLR}}(\theta) = \frac{\bigl(m(\theta)^{\mathrm{t}}\, W\, y\bigr)_+}{\sqrt{m(\theta)^{\mathrm{t}}\, W\, m(\theta)}} \right\} \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \eta . \tag{12}$$

For a shift-invariant PSF and independent noise, the same tricks as before can be used to quickly locate the most likely position at the pixel precision:

$$\widehat{\imath} = \arg\max_i \frac{\max(a(\xi_i), 0)}{\sqrt{b(\xi_i)}} .$$

1.4 Locally Most Powerful Test

Assuming Gaussian distribution for the nuisance terms is however not very robust against false detections due to outliers for instance. A heavy tailed distribution as the Cauchy one is preferable. The GLRT becomes however very costly to apply as it requires to compute $\widehat{\alpha}(\theta)$ for every trial location θ while there is no closed form solution for the optimal amplitude $\widehat{\alpha}(\theta)$.

Rao[7] has proposed the following test:

$$\frac{1}{I_\alpha(\theta)} \left(\frac{\partial \log p(y \,|\, \alpha, \theta)}{\partial \alpha} \right)^2 \bigg|_{\alpha=0} \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \eta , \tag{13}$$

for some threshold η and where $I_\alpha(\theta)$ is Fisher's information for the parameter α knowing α and θ:

$$I_\alpha(\theta) \;=\; -\mathbb{E}\left\{ \frac{\partial^2 \log p(y \,|\, \alpha, \theta)}{\partial \alpha^2} \right\} \;=\; \mathbb{E}\left\{ \left(\frac{\partial \log p(y \,|\, \alpha, \theta)}{\partial \alpha} \right)^2 \right\} \geq 0 , \tag{14}$$

where $\mathbb{E}\{\ldots\}$ denotes expectation. The advantage of Rao's test is that the expression to compute does not depend on the value of α as it is evaluated for $\alpha = 0$.

When α is known to be nonnegative, Rao's test becomes:

$$I_\alpha(\theta)^{-\frac{1}{2}} \, \frac{\partial \log p(y \,|\, \alpha, \theta)}{\partial \alpha} \bigg|_{\alpha=0} \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \sqrt{\eta} , \tag{15}$$

and it can be proven that this test is the *locally most powerful test* (LMPT).[7] That is to say, in the regime of weak signals ($\alpha \to 0^+$) it maximizes the probability of detection for any given probability of false alarm.

1.5 Robust Test

To implement a robust yet general test based on the LMPT, we have recently proposed[3,4] to assume that, after a normalization, the nuisance terms are independent and identically distributed (i.i.d.). Let:

$$t_i \overset{\text{def}}{=} \frac{y_i - \alpha\, m_i(\theta)}{s_i}\,, \tag{16}$$

be the normalized residuals for some $s_i > 0$, then our assumptions lead to write the likelihood of the data as:

$$p(y\,|\,\alpha,\theta;\mathscr{H}_1) = \prod_i \exp(-\varphi(t_i))\,, \tag{17}$$

where $\varphi(t)$ is the co-log-likelihood of the normalized residuals. The first derivative of the likelihood of the data with respect to α writes:

$$\frac{\partial \log p(y\,|\,\alpha,\theta;\mathscr{H}_1)}{\partial \alpha} = \sum_i \frac{m_i(\theta)}{s_i}\,\varphi'\!\left(\frac{y_i - \alpha\, m_i(\theta)}{s_i}\right),$$

which gives one of the terms of the LMPT:

$$\left.\frac{\partial \log p(y\,|\,\alpha,\theta;\mathscr{H}_1)}{\partial \alpha}\right|_{\alpha=0} = \sum_i \frac{m_i(\theta)}{s_i}\,\varphi'(y_i/s_i)\,. \tag{18}$$

The second derivative with respect to α of the log-likelihood is:

$$\frac{\partial^2 \log p(y\,|\,\alpha,\theta;\mathscr{H}_1)}{\partial \alpha^2} = -\sum_i \frac{m_i^2(\theta)}{s_i^2}\,\varphi''(t_i)\,,$$

which is needed to compute Fisher's information on parameter α:

$$I_\alpha(\theta) = -\mathbb{E}\left\{\frac{\partial^2 \log p(y\,|\,\alpha,\theta;\mathscr{H}_1)}{\partial \alpha^2}\right\} = \sum_i \mathbb{E}\{\varphi''(t_i)\}\,\frac{m_i^2(\theta)}{s_i^2}\,.$$

From our assumption that the t_i are i.i.d., it follows that $\mathbb{E}\{\varphi''(t_i)\} = \beta$ whatever i, where β is given by:

$$\beta = \int_{-\infty}^{+\infty} \varphi''(t)\,\exp(-\varphi(t))\,\mathrm{d}t\,. \tag{19}$$

Finally, Fisher's information on the parameter α writes:

$$I_\alpha(\theta) = \beta \sum_i \frac{m_i^2(\theta)}{s_i^2}\,, \tag{20}$$

which, incidentally, does not depend on α.

Putting all together:

$$T_{\mathsf{LMP}}(\theta) = \frac{\sum_i \frac{m_i(\theta)}{s_i}\,\varphi'(y_i/s_i)}{\sqrt{\beta \sum_i \frac{m_i^2(\theta)}{s_i^2}}} \tag{21}$$

Let us introduce the following pixelwise weights (accounting for zero-padding to avoid aliasing issues[3,6]):

$$w_i = \begin{cases} 1/s_i & \text{if } i\text{-th data is valid;} \\ 0 & \text{otherwise;} \end{cases} \tag{22}$$

then computing $T_{\mathsf{LMP}}(\theta)$ for any $\theta \in \{\xi_i\}_{i=1,\ldots,M}$ (the regular grid of observed pixels) can be computed as:

$$T_{\mathsf{LMP}}(\theta = \xi_i) = \frac{\sum_{i'} \frac{m_{i'}(\xi_i)}{s_{i'}}\,\varphi'(y_{i'}/s_{i'})}{\sqrt{\beta \sum_{i'} \frac{m_{i'}^2(\xi_i)}{s_{i'}^2}}} = \frac{[m \circledast (w \otimes z)]_i}{\sqrt{\beta\,[(m \otimes m) \circledast (w \otimes w)]_i}} \tag{23}$$

where $z \in \mathbb{R}^M$ is a separable non-linear transform of the data:

$$(\forall i) \quad z_i \overset{\text{def}}{=} \varphi'(y_i/s_i)\,. \tag{24}$$

2. DETECTION IN MULTI-VARIATE DATA

For multi-exposure data, the apparent position of a planet in the field of view of the instrument may vary with the time. In the case of a pupil-stabilized instrument like SPHERE or GPI, this motion is however deterministic and we denote by $\zeta(t,\theta)$ the apparent position of the planet at time t assuming the planet was at position θ at some given reference date t_{ref}; hence $\zeta(t_{\mathsf{ref}},\theta)=\theta$. For multi-spectral, multi-temporal images (*e.g.*, as provided by SPHERE IFS), the model of available data under hypothesis $\mathscr{H}_1$ becomes:

$$y_{j,k,\ell} = f(\lambda_\ell)\, h_{k,\ell}(\xi_j, \zeta_k(\theta)) + n_{j,k,\ell}\,, \tag{25}$$

where $f(\lambda_\ell)$ is the spectral energy distribution (SED) of the planet at the effective wavelength λ_ℓ of the ℓ-th spectral channel, $h_{k,\ell}(\xi,\zeta)$ is the PSF in the ℓ-th spectral channel during the k-th exposure, ξ_j is the position of the j-th pixel of the considered data frame, $\zeta_k(\theta)=\zeta(t_k,\theta)$ is the apparent position of the planet at the date t_k of the observation and, as before, $n_{j,k,\ell}$ is a nuisance term. Compared to the previous Section, the data index $i\sim(j,k,\ell)$ has now three components: j, k and ℓ are respectively the indexes of the pixel, of the exposure and of the spectral channel.

Using matrix notation, the direct model can be expressed as:

$$y = H(\theta)\, f + n\,,$$

where f is a vector whose components are the SED of the planet in the different spectral channels, $f_\ell = f(\lambda_\ell)$ $(\forall \ell)$, and $H(\theta)$ is a tensor whose coefficients are $H_{j,k,\ell}(\theta) = h_{k,\ell}(\xi_j,\zeta_k(\theta))$. The tensor product above reads:

$$[H(\theta)\, f]_{j,k,\ell} = H_{j,k,\ell}(\theta)\, f_\ell\,.$$

2.1 Gaussian Noise Distribution

Assuming the nuisance terms have a centered Gaussian distribution, the co-log-likelihood of the data writes:

$$-\log p(y\,|\,f,\theta;\mathscr{H}_1) = (1/2)\,\|y - H(\theta)\, f\|_W^2 + \text{const.}\,, \tag{26}$$

and the maximum likelihood estimator of the planetary SED knowing θ solves the following *normal equations*:

$$\widehat{f}(\theta) \;=\; \arg\min_f \|y - H(\theta)\, f\|_W^2 \;=\; A(\theta)^{-1}\, b(\theta)\,, \tag{27}$$

with:

$$A(\theta) = H(\theta)^{\mathrm{t}}\, W\, H(\theta) \quad\iff\quad (\forall \ell, \forall \ell')\quad A_{\ell,\ell'}(\theta) = \sum_{j,k,j',k'} H_{j,k,\ell}(\theta)^{\mathrm{t}}\, W_{j,k,\ell,j',k',\ell'}\, H_{j',k',\ell'}(\theta)\,, \tag{28a}$$

$$b(\theta) = H(\theta)^{\mathrm{t}}\, W\, y \quad\iff\quad (\forall \ell)\quad b_\ell(\theta) = \sum_{j,k,j',k',\ell'} H_{j,k,\ell}(\theta)^{\mathrm{t}}\, W_{j,k,\ell,j',k',\ell'}\, y_{j',k',\ell'}\,. \tag{28b}$$

After simplifications, the (squared) criterion in the GLRT writes:

$$T_{\mathsf{GLR}}^2(\theta) = \|y\|_W^2 - \min_f \|y - H(\theta)\, f\|_W^2 = b(\theta)^{\mathrm{t}}\, A(\theta)^{-1}\, b(\theta) = b(\theta)^{\mathrm{t}}\, \widehat{f}(\theta)\,. \tag{29}$$

Quite interestingly, $\widehat{f}(\theta)$ being a linear estimator with respect to the data, *cf.* Eq. (27), it is straightforward to establish that:

$$\mathrm{Cov}\{\widehat{f}(\theta)\} = A(\theta)^{-1}\,, \tag{30}$$

is the covariance of the estimator $\widehat{f}(\theta)$. Furthermore, it can be shown[8] that $\widehat{f}(\theta)$ is the *best linear unbiased estimator* (BLUE) of the SED knowing the planet position θ at the reference time. Using the covariance and the expression in Eq. (27), the GLRT criterion can be rewritten as:

$$T_{\mathsf{GLR}}^2(\theta) = \widehat{f}(\theta)^{\mathrm{t}}\; \mathrm{Cov}\{\widehat{f}(\theta)\}^{-1}\, \widehat{f}(\theta)\,. \tag{31}$$

Thus $T_{\mathsf{GLR}}^2(\theta)$ is the quadratic sum of the SNR of the estimator of the SED in the different spectral channels. The most likely detection occurs for the position θ where the total SNR of the SED is maximized. This result generalizes to correlated Gaussian noise an observation made by Mugnier et al.[9] for multi-spectral data with independent Gaussian noise. Note that, to evaluate the GLRT criterion, one has to compute $A(\theta)$ and $b(\theta)$ thus the maximum likelihood planet SED given its position θ is a by-product of the detection.

2.2 Independent Gaussian Noise Approximation

For correlated Gaussian noise, the nonnegativity constraint is more difficult to apply even though constrained optimization can be used to derive the maximum likelihood of the nonnegative SED.[5] Assuming independent Gaussian noise, the weighting tensor W becomes *diagonal*, *i.e.* $W_{j,k,\ell,j',k',\ell'} = \delta_{j,j'}\,\delta_{k,k'}\,\delta_{\ell,\ell'}\,w_{j,k,\ell}$, and the problem to solve becomes independent in λ (that is index ℓ). Accounting for the positivity of the SED, the most likely planet SED knowing the position θ of the planet at t_{ref} is directly derived from what precedes:

$$\widehat{f}_\ell(\theta) \;=\; \min_{f_\ell \geq 0} \|y - H(\theta)\,f\|_W^2 \;=\; \frac{\max\bigl(b_\ell(\theta), 0\bigr)}{a_\ell(\theta)}\,, \tag{32}$$

with:

$$a_\ell(\theta) = \sum_{j,k} H_{j,k,\ell}^2(\theta)\,w_{j,k,\ell}\,, \tag{33a}$$

$$b_\ell(\theta) = \sum_{j,k} H_{j,k,\ell}(\theta)\,w_{j,k,\ell}\,y_{j,k,\ell}\,. \tag{33b}$$

The criterion for the GLRT becomes:

$$T_{\mathsf{GLR}}^2(\theta) = \sum_\ell \frac{\max\bigl(b_\ell(\theta), 0\bigr)^2}{a_\ell(\theta)}\,. \tag{34}$$

The covariance of the estimator $\widehat{f}_\ell(\theta)$ in Eq. (32) is diagonal. However, due to the positivity, $1/a_\ell(\theta)$ is not equal to the variance of $\widehat{f}_\ell(\theta)$ and $T_{\mathsf{GLR}}^2(\theta)$ is therefore not strictly the quadratic sum of the SNR in every spectral channel. Now it remains to find the maximum likelihood planet position (under hypothesis $\mathscr{H}_1$):

$$\widehat{\theta} = \arg\max_\theta T_{\mathsf{GLR}}(\theta)\,.$$

2.3 Fast Computation under the Independent Gaussian Noise Approximation

Provided we consider planets not too close to the coronographic mask, we can assume that the off-axis PSF is stationary. Then:

$$h_{k,\ell}(\xi, \zeta) = h_{k,\ell}(\xi - \zeta)\,, \tag{35}$$

assuming that the pixel and planet positions, ξ and ζ, are expressed in the same coordinate system. In the previous equations, the coefficients of the tensor H become:

$$H_{j,k,\ell}(\theta) = h_{k,\ell}(\xi_j - \zeta_k(\theta))\,, \tag{36}$$

and it appears that computing $a_\ell(\theta)$ (resp. $b_\ell(\theta)$) involves computing the correlation of the squared PSF with the weights (resp. of the PSF with the weighted data). In order to exploit fast computations and following the developments in Section 2.3, we propose to compute the following maps (for every exposure and spectral channel):

$$a_{j,k,\ell} = \sum_{j'} h_{k,\ell}^2(\xi_{j'} - \xi_j)\,w_{j',k,\ell} = [(\boldsymbol{h}_{k,\ell} \otimes \boldsymbol{h}_{k,\ell}) \circledast \boldsymbol{w}_{k,\ell}]_j\,, \tag{37a}$$

$$b_{j,k,\ell} = \sum_{j'} h_{k,\ell}(\xi_{j'} - \xi_j)\,w_{j',k,\ell}\,y_{j',k,\ell} = \bigl[\boldsymbol{h}_{k,\ell} \circledast \bigl(\boldsymbol{w}_{k,\ell} \otimes \boldsymbol{y}_{k,\ell}\bigr)\bigr]_j\,. \tag{37b}$$

which amounts to assuming that the planet position at time t_k coincides with the j-th pixel: $\zeta_k(\theta) = \xi_j$. As before, these expressions are discrete correlations which can be quickly computed by means of the FFT after proper zero-padding to avoid aliasing.[6] The boldface notation above indicates 2D maps sampled on the grid of the pixels: $[\boldsymbol{h}_{k,\ell}]_j = h_{k,\ell}(\xi_j)$, $\bigl[\boldsymbol{y}_{k,\ell}\bigr]_j = y_{j,k,\ell}$ and $\bigl[\boldsymbol{y}_{k,\ell}\bigr]_j = y_{j,k,\ell}$. As such maps are rather smooth[†], the value of

[†]because they involve correlation by the PSF of the squared PSF; otherwise subdivide pixels

$a_\ell(\theta)$ and $b_\ell(\theta)$, defined in Eq. (33a) and (33b), can be approximated by simple interpolation of the precomputed maps and then summed for the time index k:

$$a_\ell(\theta) \approx \sum_k \underbrace{\sum\nolimits_j S_j(\zeta_k(\theta))\, a_{j,k,\ell}}_{\text{interpolation of } a_{j,k,\ell} \text{ at } \zeta_k(\theta)}, \tag{38a}$$

$$b_\ell(\theta) \approx \sum_k \underbrace{\sum\nolimits_j S_j(\zeta_k(\theta))\, b_{j,k,\ell}}_{\text{interpolation of } b_{j,k,\ell} \text{ at } \zeta_k(\theta)}, \tag{38b}$$

where $a_\ell(\theta)$ and $b_\ell(\theta)$ were defined in Eq. (33a) and Eq. (33b) while $S_j(\xi)$ denotes the coefficients of a spatial interpolation operator $\mathbf{S}(\xi)$ which interpolates at position ξ a map sampled on the grid ξ_j of pixels. For the sake of simplicity, we assumed that the same interpolation operator is used for $a_\ell(\theta)$ and $b_\ell(\theta)$, although different interpolation operators could be used for improved accuracy. Interpolation for a given position is a fast operation[10] and the GLRT criterion given in Eq. (34) can be quickly estimated for any assumed position θ of the planet at t_{ref}. This provides a fast means to perform planet detection in multi-spectral and multi-temporal data without data interpolation.

2.4 Robust Detection in Multi-variate Data

To apply the LMPT defined in Section 1.4 to multivariate data, we can rewrite the direct model of the data as:

$$y_{j,k,\ell} = \alpha\, f(\lambda_\ell)\, h_{k,\ell}(\xi_j, \zeta_k(\theta)) + n_{j,k,\ell} \tag{39}$$

where $\alpha \geq 0$ is the brightness of the planet and, now, $f(\lambda)$ is its normalized SED. If $f(\lambda)$ is known, the LMP test directly applies:

$$T_{\mathsf{LMP}}(\theta) = \frac{\sum_{i,k,\ell} f(\lambda_\ell)\, H_{j,k,\ell}(\theta)\, w_{j,k,\ell}\, \varphi'(y_{j,k,\ell}\, w_{j,k,\ell})}{\sqrt{\beta \sum_{i,k,\ell} f(\lambda_\ell)^2\, H^2_{j,k,\ell}(\theta)\, w^2_{j,k,\ell}}} \tag{40}$$

and the same approach as in section 2.3 can be followed to perform a fast computation to locate an exo-planet within pixel accuracy:

$$\hat{\theta} \approx \arg\max_\theta \frac{\sum_k \sum_j S_j(\zeta_k(\theta)) b_{j,k}}{\sqrt{\sum_k \sum_j S_j(\zeta_k(\theta)) a_{j,k}}}, \tag{41}$$

with:

$$a_{j,k} = \sum_\ell \sum_{j'} \beta\, f(\lambda_\ell)^2\, h^2_{k,\ell}(\xi_{j'} - \xi_j)\, w^2_{j',k,\ell} = \sum_\ell \beta\, f(\lambda_\ell)^2 \left[(\boldsymbol{h}_{k,\ell} \otimes \boldsymbol{h}_{k,\ell}) \circledast (\boldsymbol{w}_{k,\ell} \otimes \boldsymbol{w}_{k,\ell})\right]_j, \tag{42}$$

$$b_{j,k} = \sum_\ell \sum_{j'} f(\lambda_\ell)\, h_{k,\ell}(\xi_{j'} - \xi_j)\, w_{j',k,\ell}\, \varphi'(y_{j',k,\ell}\, w_{j',k,\ell}) = \sum_\ell f(\lambda_\ell) \left[\boldsymbol{h}_{k,\ell} \circledast \left(\boldsymbol{w}_{k,\ell} \otimes \varphi'\left(\boldsymbol{y}_{k,\ell} \otimes \boldsymbol{w}_{k,\ell}\right)\right)\right]_j. \tag{43}$$

2.5 Detection with Speckle Removal

In practice, coronographic images are plagued by stellar leakage in the form of a pattern of speckles.[1,2] As most of these speckles are quasi-static, a simple way to get rid of them is to work on image differences[9,11] possibly after proper image registration to make the speckles coincident.[12] In order to account for the speckles, the model of the images must be modified as follows:

$$\mathscr{H}_0\colon \quad y = G\,x + n\,, \tag{44}$$

$$\mathscr{H}_1\colon \quad y = G\,x + H(\theta)\,f + n\,, \tag{45}$$

were $G\,x$ denotes the background signal due to the speckles which are linearly parameterized by x while $H(\theta)\,f$ is the contribution of a planet at position θ and n the centered noise. For instance, assuming quasi-static speckles, $(G\,x)_{j,k,\ell} = x_{j,\ell}$ where $x_{j,\ell}$ is the pattern of the speckles at the j-th pixel in the ℓ-th spectral channel whatever the exposure index k, hence $G_{j,k,\ell,j',\ell'} = \delta_{j,j'}\,\delta_{\ell,\ell'}$. Note that we can impose some physical prior knowledge of the speckle field through G, in particular the fact that $G\,x$ must be spatially band-limited. Using matrix notation, the co-log-likelihood of the images now writes:

$$-\log p(y \,|\, x, f, \theta; \mathscr{H}_1) = (1/2)\,\|y - G\,x - H(\theta)\,f\|_W^2 + \text{const.}\,, \tag{46}$$

and the criterion for the GLRT has to be rewritten as:

$$T^2_{\mathsf{GLR}}(\theta) = \min_x \|y - G\,x\|_W^2 - \min_{x,f} \|y - G\,x - H(\theta)\,f\|_W^2\,. \tag{47}$$

Let us first consider the best background parameters x given f and θ under hypothesis $\mathscr{H}_1$ (it will be sufficient to take $f \equiv 0$ to obtain the background parameters under $\mathscr{H}_0$):

$$\widehat{x}(f,\theta) = \arg\min_x \|y - G\,x - H(\theta)\,f\|_W^2 = \big(G^{\mathrm{t}}\,W\,G\big)^{-1}\,G^{\mathrm{t}}\,W\,\big(y - H(\theta)\,f\big)\,, \tag{48}$$

which, after some simple algebra, yields:

$$\begin{aligned}\min_x \|y - G\,x - H(\theta)\,f\|_W^2 &= \|y - G\,\widehat{x}(f,\theta) - H(\theta)\,f\|_W^2 \\ &= \|y - H(\theta)\,f\|_Q^2\,,\end{aligned} \tag{49}$$

with:

$$Q = W - W\,P\,, \tag{50}$$

and:

$$P = G\,\big(G^{\mathrm{t}}\,W\,G\big)^{-1}\,G^{\mathrm{t}}\,W\,, \tag{51}$$

which is a projection[‡]. Considering Eq. (49), it turns out that removing the speckles in a maximum likelihood sense amounts to processing the images y with a planet footprint $H(\theta)$ and modified statistical weights $Q = W - W\,P$. At this point it is interesting to realize that:

$$\begin{aligned}y - G\,\widehat{x}(f,\theta) - H(\theta)\,f &= y - P\,\big(y - H(\theta)\,f\big) - H(\theta)\,f \\ &= r - K(\theta)\,f\,,\end{aligned}$$

where $K(\theta) = (I - P)\,H(\theta)$ and:

$$r = y - P\,y = y - G\,\widehat{x}_{\mathscr{H}_0}\,, \tag{52}$$

are the residual images after subtracting the contribution of the speckles corresponding to the maximum likelihood under hypothesis $\mathscr{H}_0$ (no planets): $\widehat{x}_{\mathscr{H}_0} = \widehat{x}(f = 0, \theta)$. Therefore:

$$\min_x \|y - G\,x - H(\theta)\,f\|_W^2 = \|r - K(\theta)\,f\|_W^2\,, \tag{53}$$

is another form of Eq. (49) which shows that it is also equivalent to process the residuals r assuming a planet footprint given by $K(\theta) = (I - P)\,H(\theta)$ and unchanged statistical weights W. Note that $K(\theta)$ is the projection of the footprint $H(\theta)$ onto the orthogonal of the space spanned by projector P.

From Eq. (49) and Eq. (53), we deduce that:

$$\begin{aligned}\min_x \|y - G\,x\|_W^2 - \min_x \|y - G\,x - H(\theta)\,f\|_W^2 &= \|y\|_Q^2 - \|y - H(\theta)\,f\|_Q^2 = 2\,f^{\mathrm{t}}\,H(\theta)^{\mathrm{t}}\,Q\,y - f^{\mathrm{t}}\,H(\theta)^{\mathrm{t}}\,Q\,H(\theta)\,f \\ &= \|r\|_W^2 - \|r - K(\theta)\,f\|_W^2 = 2\,f^{\mathrm{t}}\,K(\theta)^{\mathrm{t}}\,W\,r - f^{\mathrm{t}}\,K(\theta)^{\mathrm{t}}\,W\,K(\theta)\,f\,.\end{aligned}$$

The two alternative expressions above take the same usual form:

$$\min_x \|y - G\,x\|_W^2 - \min_x \|y - G\,x - H(\theta)\,f\|_W^2 = 2\,b(\theta)^{\mathrm{t}}\,f - f^{\mathrm{t}}\,A(\theta)\,f\,, \tag{54}$$

[‡]which is easily found by checking that it is idempotent, *i.e.* that $P^2 = P$

with:

$$A(\theta) = H(\theta)^{\mathrm{t}}\, Q\, H(\theta) = K(\theta)^{\mathrm{t}}\, W\, K(\theta)\,, \tag{55a}$$

$$b(\theta) = H(\theta)^{\mathrm{t}}\, Q\, y = K(\theta)^{\mathrm{t}}\, W\, r = H(\theta)^{\mathrm{t}}\, W\, r\,, \tag{55b}$$

where the above equalities follow from the identities:

$$Q\, y = W\,(y - P\, y) = W\, r\,, \tag{56}$$

$$Q\, H(\theta) = W\, K(\theta)\,, \tag{57}$$

which result from the definitions of P in Eq. (51), of $Q = W\,(I - P)$ in Eq. (50) of $K(\theta) = (I - P)\, H(\theta)$ and from which it is easy to show that: $P^{\mathrm{t}}\, W = W\, P = P^{\mathrm{t}}\, W\, P$.

Maximizing the expression in Eq. (54) with respect to f yields the maximum likelihood estimator of the planet SED knowing its position θ:

$$\widehat{f}(\theta) = A(\theta)^{-1}\, b(\theta)\,,$$

and the criterion for the GLRT becomes:

$$T^2_{\mathsf{GLR}}(\theta) = b(\theta)^{\mathrm{t}}\, A(\theta)^{-1}\, b(\theta)\,.$$

These expressions are exactly the same as those obtained when there was no spurious background, *cf.* Eq. (27) and (29), only the expressions of $A(\theta)$ and $b(\theta)$ have changed.

2.6 Fast Computations with Speckle Removal

We assume independent Gaussian noise, a shift-invariant PSF and, now, quasi-static speckles. The projection operator P can be derived from observing that (whatever u of suitable size):

$$P\, u = G\, \widehat{x}(u)\,,$$

with:

$$\widehat{x}(u) = \arg\min_{x} \|u - G\, x\|^2_W = \arg\min_{x} \sum_{j,k,\ell} w_{j,k,\ell}\, (u_{j,k,\ell} - x_{j,\ell})^2\,.$$

Noting that the problem is separable in j and ℓ, the solution is a trivial weighted average:

$$\widehat{x}_{j,\ell}(u) = \frac{\sum_k w_{j,k,\ell}\, u_{j,k,\ell}}{\sum_k w_{j,k,\ell}}\,,$$

and $[P\, u]_{j,k,\ell} = [G\, \widehat{x}(u)]_{j,k,\ell} = \widehat{x}_{j,\ell}(u)$ whatever j, k and ℓ. Hence computing the residuals $r = y - P\, y$ poses no problems:

$$r_{j,k,\ell} = y_{j,k,\ell} - \frac{\sum_{k'} w_{j,k',\ell}\, y_{j,k',\ell}}{\sum_{k'} w_{j,k',\ell}} = \sum_{k'} c_{j,k,k',\ell}\, y_{j,k',\ell}\,, \tag{58}$$

with

$$c_{j,k,k',\ell} = \delta_{k,k'} - \frac{w_{j,k',\ell}}{\sum_{k''} w_{j,k'',\ell}}\,. \tag{59}$$

Then taking the last expression in Eq. (55b), the right hand side term of the normal equations for the SED writes:

$$b_\ell(\theta) = \sum_{j,k} h_{k,\ell}(\xi_j - \zeta_k(\theta))\, w_{j,k,\ell}\, r_{j,k,\ell} \approx \sum_k \sum_j S_j(\zeta_k(\theta))\, b_{j,k,\ell}\,, \tag{60}$$

which involves the interpolation at $\zeta_k(\theta)$ of:

$$b_{j,k,\ell} = \sum_{j'} h_{k,\ell}(\xi_{j'} - \xi_j)\, w_{j',k,\ell}\, r_{j',k,\ell} = [\boldsymbol{h}_{k,\ell} \circledast (\boldsymbol{w}_{k,\ell} \otimes \boldsymbol{r}_{k,\ell})]_j\,, \tag{61}$$

which is inexpensive to compute.

For the left hand side term of the normal equations for the SED, we use the first expression in Eq. (55a) to write:

$$a_\ell(\theta) = \sum_{j,k,k'} h_{k,\ell}(\xi_j - \zeta_k(\theta))\, w_{j,k,\ell}\, c_{j,k,k',\ell}\, h_{k',\ell}(\xi_j - \zeta_{k'}(\theta))\,, \tag{62}$$

which does not simplify without additional approximations. The most simple one consists in assuming that $c_{j,k,k',\ell} \approx \delta_{k,k'}$ which is valid in the limit of a large number of exposures. Then:

$$a_\ell(\theta) \approx \sum_{j,k} h_{k,\ell}^2(\xi_j - \zeta_k(\theta))\, w_{j,k,\ell} \approx \sum_k \sum_j S_j(\zeta_k(\theta))\, a_{j,k,\ell}\,, \tag{63}$$

which involves the interpolation at $\zeta_k(\theta)$ of:

$$a_{j,k,\ell} = \sum_{j'} h_{k,\ell}^2(\xi_{j'} - \xi_j)\, w_{j',k,\ell} = [(\boldsymbol{h}_{k,\ell} \otimes \boldsymbol{h}_{k,\ell}) \circledast \boldsymbol{w}_{k,\ell}]_j\,, \tag{64}$$

which is inexpensive to compute.

3. SUMMARY AND PERSPECTIVES

We recalled the basics of signal detection and described how to derive a fast and robust detection criterion based on a heavy tail model that can account for outliers in the residuals. We derived detectors that handle jointly several wavelengths and exposures and focused on an approach that avoids interpolating the data, thereby preserving the statistics of the original data. The next stage is to apply these methods to simulated and real data.

ACKNOWLEDGMENTS

The research leading to these results has received support from the DETECTION project funded by the French CNRS (Mission pour l'Interdisciplinarité, DEFI IMAGIN).

REFERENCES

[1] Janson, M., Brandner, W., Henning, T., and Zinnecker, H., "Early comeon+ adaptive optics observation of gq lupi and its substellar companion," *Astron. Astrophys.* **453**, 609–614 (2006).
[2] Hinkley, S., Oppenheimer, B. R., Soummer, R., Sivaramakrishnan, A., Roberts, Jr., L. C., Kuhn, J., Makidon, R. B., Perrin, M. D., Lloyd, J. P., Kratter, K., and Brenner, D., "Temporal evolution of coronagraphic dynamic range and constraints on companions to vega," *Astrophys. J.* **654**, 633–640 (2007).
[3] Denis, L. and Thiébaut, É., "Détection robuste et rapide d'un motif connu dans une image," in [*XXVème Colloque GRETSI*], 374, GRETSI (2015).
[4] Denis, L., Ferrari, A., Mary, D., Mugnier, L., and Thiébaut, É., "Fast and robust detection of a known pattern in an image," in [*Eusipco 2016 Conference*], (2016).
[5] Thiébaut, É. and Mugnier, L., "Maximum a posteriori planet detection and characterization with a nulling interferometer," in [*IAU Colloq. 200: Direct Imaging of Exoplanets: Science & Techniques*], Aime, C. and Vakili, F., eds., 547–552, Cambridge University Press (2006).
[6] Soulez, F., Denis, L., Thiébaut, É., Fournier, C., and Goepfert, C., "Inverse problem approach in particle digital holography: out-of-field particle detection made possible," *J. Opt. Soc. America A* **24**(12), 3708–3716 (2007).
[7] Kay, S. M., [*Fundamentals of Statistical Signal Processing: Detection Theory*], vol. 2 of *Signal Processing Series*, Prentice-Hall (1998).
[8] Kay, S. M., [*Fundamentals of Statistical Signal Processing: Estimation Theory*], vol. 1 of *Signal Processing Series*, Prentice-Hall (1998).
[9] Mugnier, L. M., Cornia, A., Sauvage, J.-F., Rousset, G., Fusco, T., and Védrenne, N., "Optimal method for exoplanet detection by angular differential imaging," *J. Opt. Soc. America A* **26**(6), 1326–1334 (2009).
[10] Thévenaz, P., Blu, T., and Unser, M., "Interpolation revisited," *IEEE Trans. Medical Imag.* **19**(7), 739–758 (2000).
[11] Smith, I., Ferrari, A., and Carbillet, M., "Detection of a moving source in speckle noise. application to exoplanet detection," *IEEE Trans. Signal Process.* **57**(3), 904–915 (2009).
[12] Thiébaut, É., Devaney, N., Langlois, M., and Hanley, K., "Exploiting physical constraints for multi-spectral exoplanet detection," in [*SPIE Conf. on Astronomical Telescopes & Instrumentation*], **9909**, 65, SPIE Proc. (2016).