Finding Meaningful Detections: False Discovery Rate Control in Correlated Detection Maps

Olivier Flasseur†, Loïc Denis‡, Éric Thiébaut†, Maud Langlois†

† Univ. Lyon, Univ. Lyon1, ENS de Lyon, CNRS,
Centre de Recherche Astrophysique de Lyon UMR 5574, F-69230, Saint-Genis-Laval, France

‡ Univ. Lyon, UJM-Saint-Etienne, CNRS, Institut d Optique Graduate School,
Laboratoire Hubert Curien UMR 5516, F-42023, Saint-Etienne, France

***Abstract*—The detection of faint sources is a key step in several areas of signal and image processing. The reliability of the detection depends on two key components: (i) the detection criterion used to derive detection maps in which the signature of a source takes the form of a detection peak, and (ii) the extraction procedure identifying the meaningful detections.**

In this work, the expected false discovery rate guides the selection of meaningful detections. A procedure is designed to account for correlations in the detection maps. This prevents the issue of the multiple detections of a single source and corrects the number of effective independent tests performed. The proposed approach is evaluated on an astrophysical application: the detection of exoplanets by high-contrast imaging.

***Index Terms*—detection, FDR, correlated data, matched filter**

I. Introduction

Many applications require detecting and localizing a pattern of known shape in a signal or an image, for example, the echo of a modulated wave to estimate the time-of-flight in radar [1], the point spread function of the microscope to perform super-resolution in fluorescence microscopy, or an unresolved exoplanet [2]. Under the assumption of additive Gaussian background noise, the detection and localization are typically performed with a matched filter. This matched filter can be interpreted as a binary test assessing the absence or presence of a pattern centered at any of the possible sample locations.

Controlling the false alarm rate, i.e., the proportion of wrongful detections of the pattern solely due to noise is often crucial. Since not only one location is tested at a time but all possible locations are considered, false alarms must be studied in the context of multiple hypothesis testing.

To control the false alarms when multiple hypotheses are tested, several approaches are possible. First, the probability of reporting at least a false alarm in one of the many tests performed can be controlled (the so-called *family-wise error rate*: FWER). However, as the number of tests increases, this requires setting increasingly restrictive detection thresholds. A more relevant target is the proportion of erroneous detections (the so-called *false discoveries*) relative to the total number of reported detections: the *false discovery rate* (FDR). The seminal paper [3] shows that the application of a linear step-up (LSU) procedure, based on the analysis of the ranked p-values associated with each detection, guarantees the FDR to match a target value, in expectation. The application of this procedure to detection maps obtained by matched filtering is however not straightforward because of the spatial correlations in the maps [4].

Our contributions: this paper describes a procedure to identify meaningful detections from a signal or an image, based on two key ingredients:

- a greedy detection technique based on a variant of matching pursuit,
- an estimation of the effective number of independent tests to adapt an LSU procedure and decide how many detections are meaningful according to a target FDR.

We illustrate that our method effectively controls the FDR in an astrophysical application: the detection of exoplanets and background sources in high-contrast imaging.

II. Background on multiple hypothesis testing

The problem of detecting the presence of a pattern can be expressed, in its most simplified form, as the binary hypothesis:

$$\text{(P1)}\qquad \begin{cases} \mathcal{H}_0: & \boldsymbol{s} = \boldsymbol{b} & \text{(pure background)} \\ \mathcal{H}_1: & \boldsymbol{s} = \boldsymbol{b} + \boldsymbol{m}\,, & \text{(background + pattern)} \end{cases}$$

where $\boldsymbol{s}$ corresponds to an observed signal, $\boldsymbol{b}$ is the background, i.e. a random realization of a stochastic model of the signal in the absence of the pattern, and $\boldsymbol{m}$ is the (deterministic) pattern whose presence is sought. In this formulation of the detection problem, the background noise is assumed not to depend on the presence of the pattern, which is often the case when considering the detection of weak signals relatively to a much stronger background. The pattern $\boldsymbol{m}$ is also considered to be known beforehand so that the only difficulty is to decide whether the noisy observed signal $\boldsymbol{s}$ corresponds to a random realization of pure background or to the superimposition of background and the sought pattern.

Under a pdf of the background $\mathrm{p}_{\boldsymbol{B}}(\boldsymbol{b})$, the likelihood ratio test is the most powerful test for the binary hypothesis (P1), as shown by Neyman and Pearson [5]:

$$\mathscr{L}(\boldsymbol{s}) = \frac{\mathrm{p}_{\boldsymbol{S}}(\boldsymbol{s}|\mathcal{H}_1)}{\mathrm{p}_{\boldsymbol{S}}(\boldsymbol{s}|\mathcal{H}_0)} = \frac{\mathrm{p}_{\boldsymbol{B}}(\boldsymbol{s}-\boldsymbol{m})}{\mathrm{p}_{\boldsymbol{B}}(\boldsymbol{s})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta\,. \tag{1}$$

This work has been supported by the Région Auvergne-Rhône-Alpes under project DIAGHOLO.

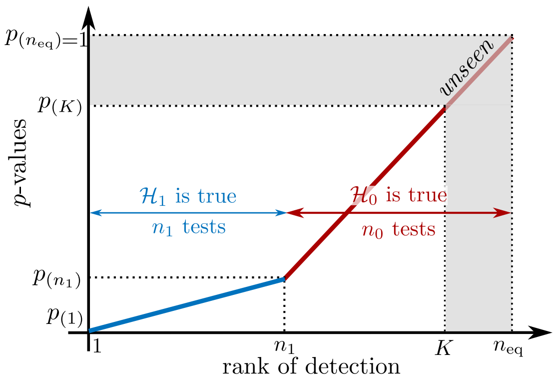

Fig. 1. Ordered p-values versus their rank of detection.

The probability of false alarm PFA (*aka* type I error), corresponds to the probability P of wrongfully deciding for $\mathcal{H}_1$ while $\mathcal{H}_0$ was true: $\mathrm{PFA}(\eta) = \mathrm{P}_{\boldsymbol{B}}(\mathscr{L}(\boldsymbol{b}) \geq \eta)$. The p-value associated to an observed signal $\boldsymbol{s}$ corresponds to $\mathrm{PFA}(\mathscr{L}(\boldsymbol{s}))$, i.e. the probability to observe by chance deviations in the background leading, in the absence of pattern, to likelihood ratio values greater or equal to $\mathscr{L}(\boldsymbol{s})$, the actual value of the likelihood ratio for the signal $\boldsymbol{s}$.

Let us now consider a multiple hypothesis test built by running in parallel the binary hypothesis test (P1) on signals $\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_n$. Out of these n signals, an unknown number n_1 of signals contain the pattern. Random realizations of the background in each of these signals are assumed to be independent.

To select the detection threshold η, one may consider the probability that at least one false alarm occurs when the n tests are conducted, i.e. the FWER. Under our assumption of independent realizations of backgrounds, $\mathrm{FWER} = 1 - (1 - \mathrm{PFA}(\eta))^n$. To control the FWER at a prescribed level q, the detection threshold must be set according to $\eta = \mathrm{PFA}^{-1}[1 - (1-q)^{1/n}]$. When the number n of tests increases, this leads to tests with very limited power: $\forall q < 1$, $\lim_{n\to\infty} \eta = \infty$ (very high detection thresholds must be set to guarantee that the probability of experiencing one or more false alarms remains small, at the cost of a drop of the probability of detection).

To address the issue of a loss of statistical power when n is large, several data-adaptive procedures have been proposed in the literature to set the detection threshold η. Rather than controlling the probability of committing at least one false detection, Benjamini and Hochberg [3] describe a method to control the FDR defined by:

$$\mathrm{FDR} = \mathbb{E}\left[\frac{V}{\max(R,1)}\right], \tag{2}$$

where V is the random variable corresponding to the number of false alarms and R is the random variable for the total number of detections, i.e. the number of positive tests $\mathscr{L}(\boldsymbol{s}_i) \geq \eta$.

Benjamini-Hochberg's method is an LSU procedure that first sorts the p-values of the test statistics in increasing order $p_{(1)}$ to $p_{(n)}$. The largest rank k such that $p_{(i)} \leq i\frac{q}{n}$ for all $i \leq k$ is iteratively identified. If no $k > 0$ exists, $\mathcal{H}_0$ is accepted for each of the n tests, otherwise, the hypothesis $\mathcal{H}_1$ is accepted

---

Algorithm 1: Estimation of $\widehat{n}_0$ – see [7]

Calculate $\forall k \in [\![1;n]\!]$, $S_k = (1 - p_{(k)})/(n+1-k)$, the k^{th} slope estimate.

Start with $k = 2$, proceed towards larger k, stop when $S_k < S_{k-1}$, and use: $\widehat{n}_0 = \min(1/S_k + 1, n)$.

---

for the tests corresponding to the first k p-values $p_{(1)}$ to $p_{(k)}$. It is shown in [3] that this procedure controls the FDR at a level $\mathrm{FDR} = \frac{n_0}{n}q$ in the case of independent backgrounds, where n_0 is the number of true $\mathcal{H}_0$ hypotheses and n the total number of independent tests. To reach a prescribed FDR level, it is therefore necessary to estimate n_0.

The estimation of n_0 can be performed by an analysis of the ordered p-values versus their rank of detection. A change of slope is expected between the n_1 (respectively n_0) tests under the $\mathcal{H}_1$ (respectively $\mathcal{H}_0$) hypothesis, see Figure 1. An estimate $\widehat{n}_0$ of n_0 is obtained through the inverse of the slope of the linear part associated with the largest p-values. Several adaptive methods have been proposed (see e.g. [6] for a review) to detect the change of slope and to estimate n_0. These strategies are generally based on the seminal procedure of [7], see Algorithm 1. Once the estimate $\widehat{n}_0$ is obtained, the LSU procedure is then applied at level $q' = q\frac{n}{\widehat{n}_0} \geq q$ to control the FDR at the prescribed level q, where the n tests are independent.

When the tests are not independent, these procedures no longer control the FDR in the sense that the actual FDR can exceed the prescribed level q. Benjamini and Hochberg have shown that performing an LSU procedure at level $q'' = \frac{q'}{\sum_{k=1}^{n} k^{-1}}$ provides an FDR that never exceeds the prescribed level q [8]. In practice, this rule is very conservative and the prescribed level q is rarely reached, causing an unnecessary loss of power.

In the context of a joint detection and localization problem in signal processing, the family of hypotheses covers a set of possible locations of the pattern:

$$\text{(P2)}\ \begin{cases} \mathcal{H}_0: & \boldsymbol{s} = \boldsymbol{b} & \text{(pure background)} \\ \mathcal{H}_1: & \boldsymbol{s} = \boldsymbol{b} + \alpha_1 \boldsymbol{m}_1\,, & \text{(background + pattern at loc. 1)} \\ \quad\ldots & & \\ \mathcal{H}_n: & \boldsymbol{s} = \boldsymbol{b} + \alpha_n \boldsymbol{m}_n\,, & \text{(background + pattern at loc. } n\text{)} \end{cases}$$

where $\boldsymbol{m}_i$ represents the pattern centered at the i^{th} location and where the amplitude α of the pattern is not known in advance and must thus be estimated to perform the detection. The likelihood of a given hypothesis $\mathcal{H}_i$ can be approximated by replacing α by the maximum likelihood estimate under $\mathcal{H}_i$: $\mathrm{p}(\boldsymbol{s}|\mathcal{H}_i) \approx \max_\alpha \mathrm{p}_{\boldsymbol{B}}(\boldsymbol{s} - \alpha \boldsymbol{m}_i)$.

Under the assumption of a Gaussian background, $\mathrm{p}_{\boldsymbol{B}}(\boldsymbol{b}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are respectively the mean and the covariance matrix of the background. The generalized likelihood ratio (i.e. the likelihood ratio when the unknown α is replaced by its maximum likelihood estimate) is then:

$$2\log \mathrm{GLRT}_i: \quad \frac{(\boldsymbol{m}_i^{\mathrm{t}} \boldsymbol{\Sigma}^{-1} \boldsymbol{s})^2}{\boldsymbol{m}_i^{\mathrm{t}} \boldsymbol{\Sigma}^{-1} \boldsymbol{m}_i} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau^2\,, \tag{3}$$

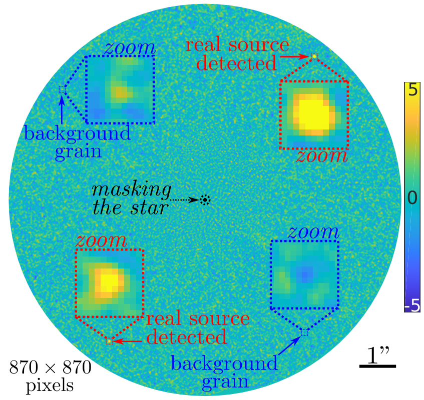

Fig. 2. Example of a detection map obtained with the PACO algorithm [2].

where τ is a detection threshold. When the amplitude α of the pattern is known *a priori* to be non-negative, the test takes the form:

$$\text{S/N}_i : \quad \frac{\boldsymbol{m}_i^{\mathrm{t}}\boldsymbol{\Sigma}^{-1}\boldsymbol{s}}{\sqrt{\boldsymbol{m}_i^{\mathrm{t}}\boldsymbol{\Sigma}^{-1}\boldsymbol{m}_i}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau \,, \tag{4}$$

which can be interpreted as the signal-to-noise ratio of the pattern amplitude, i.e. the ratio $\widehat{\alpha}/\sigma_\alpha$ where σ_α is the standard deviation of the estimator of α. The test in equation (4) also corresponds to the matched filter for detecting $\boldsymbol{m}$.

If the pattern $\boldsymbol{m}_i$ has a limited extension (in the sense that most elements of $\boldsymbol{m}$ are 0 except close to the i^{th} location) and if the correlation length in the covariance matrix $\boldsymbol{\Sigma}$ is also small, tests $\mathcal{H}_i$ and $\mathcal{H}_j$ can be considered almost independent at locations i and j that are far apart. In contrast, tests at two close locations are not independent and thus require to adapt the standard procedure to control the FDR. Figure 2 shows a detection map obtained by computing a criterion of the form of equation (4) from astronomical high-contrast observations. We describe in more detail this application in Section IV. Spatial correlations can be observed: the two detection peaks are not limited to single pixels but are spread over a larger area, outside those two peaks, tests are correlated which is visible by the grainy aspect of the detection map.

To extract meaningful detections from a detection map, it is therefore necessary to automatically set a relevant detection threshold τ, using a method that is adapted to the presence of *correlations* between tests, and to extract a *single detection* at each detection peak (although several neighboring pixels might simultaneously reach values above the threshold).

III. Proposed algorithm

There exist strong connections between multiple hypothesis testing in signal processing and sparse representations, see for example [9] for an application in synthetic aperture radar imaging. Let $\mathbf{M}$ be the matrix whose n columns are the patterns $\boldsymbol{m}_1$ to $\boldsymbol{m}_n$. The multiple detection problem, when expressed under the form of a sparse decomposition problem, takes the form:

$$\arg\min_{\boldsymbol{\alpha}} \; \|\mathbf{M}\boldsymbol{\alpha} - \boldsymbol{s}\|^2_{\boldsymbol{\Sigma}^{-1}} + \tau^2 \|\boldsymbol{\alpha}\|_0 \,, \tag{5}$$

where $\boldsymbol{\alpha}$ is the vector of the n amplitudes related to each one of the n patterns $\boldsymbol{m}_1$ to $\boldsymbol{m}_n$ and $\|\boldsymbol{\alpha}\|_0$ is the ℓ_0 pseudo-norm, i.e. the number of non-zero entries in vector $\boldsymbol{\alpha}$. The solution to this minimization problem is equivalent to running the tests GLRT_i defined in equation (3) in parallel with the same threshold τ^2 if the non-zero values in $\boldsymbol{\alpha}$ correspond to indices of patterns that do not overlap. In the case of pattern overlapping, the superimposition of patterns is considered with this latter formulation.

A slightly different form is often used, where meaningful detections are controlled via the total number of detections K:

$$\arg\min_{\boldsymbol{\alpha}} \; \|\mathbf{M}\boldsymbol{\alpha} - \boldsymbol{s}\|^2_{\boldsymbol{\Sigma}^{-1}} \quad \text{s.t.} \quad \|\boldsymbol{\alpha}\|_0 \leq K \,. \tag{6}$$

Formulation (6) is typically (approximately) solved using greedy algorithms that iteratively build the solution, starting from a zero amplitude vector $\boldsymbol{\alpha}$, until K patterns are selected.

While the ℓ_0 formulation lends itself to the design of algorithms adapted to the extraction of patterns in both the non-overlapping and partially overlapping cases, it remains to solve the issue of the automatic selection of the total number of detections K. We propose in this paragraph an extension of the Benjamini-Hochberg procedure to address this issue.

We suggest that an equivalent number n_{eq} of “independent” tests (roughly corresponding to the number of spatially correlated peaks of a detection map) can be estimated from an incomplete representation of the diagram p-values versus rank of detection, in the case of dependent tests. For this purpose, problem (6) is solved with a greedy approach. The number of patterns K to detect is set to a value much larger than the expected number of sources in order to exhibit the expected change of slope between the tests of the two hypotheses in the diagram, see Figure 1. If no prior on the number of sources is known, the greedy iterative procedure can be run until no more source of positive flux is detected. Given the K detected sources, the K available p-values are sorted in increasing order and the estimate $\widehat{n}_1 = K - \widehat{n}_0$ of sources is obtained with Algorithm 1. Then, an estimate $\widehat{n}_{\text{eq}}$ of the equivalent number of independent tests is computed (see Figure 1):

$$\widehat{n}_{\text{eq}} = \frac{1 - p_{(\widehat{n}_1)}}{p_{(K)} - p_{(\widehat{n}_1)}} \, (K - \widehat{n}_1) + \widehat{n}_1 \,. \tag{7}$$

Given the estimates $\widehat{n}_1$ and $\widehat{n}_{\text{eq}}$, the LSU procedure is applied at a level $q' = q\frac{\widehat{n}_0}{\widehat{n}_{\text{eq}}} = q\frac{\widehat{n}_{\text{eq}} - \widehat{n}_1}{\widehat{n}_{\text{eq}}}$ to control the FDR at level q.

IV. Results

In this section, we evaluate the performance of the proposed procedure both on synthetic and real data. For all tests, we consider $K = 300$ detections for the estimation of n_{eq} and n_1, which seems reasonable given the number of sources we expect to detect.

A. Application on synthetic data

We first test the proposed procedure on simulated data to evaluate its ability to control the FDR of dependent tests. We have numerically injected 100 synthetic sources (mean S/N of detection about 4) in each of 100 centered random fields of 1000×1000 pixels with unit variance. Detection maps are

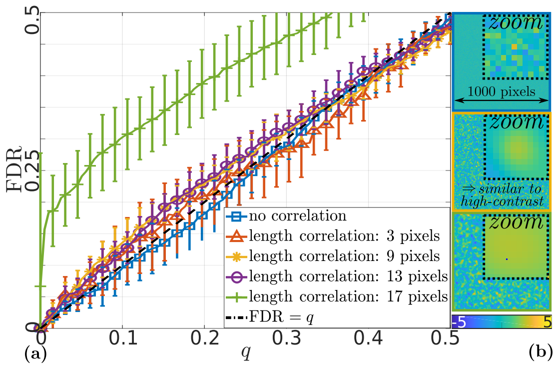

Fig. 3. Influence of the typical length of correlation. (a) FDR as a function of the prescribed level q; (b) example of spatially correlated images.

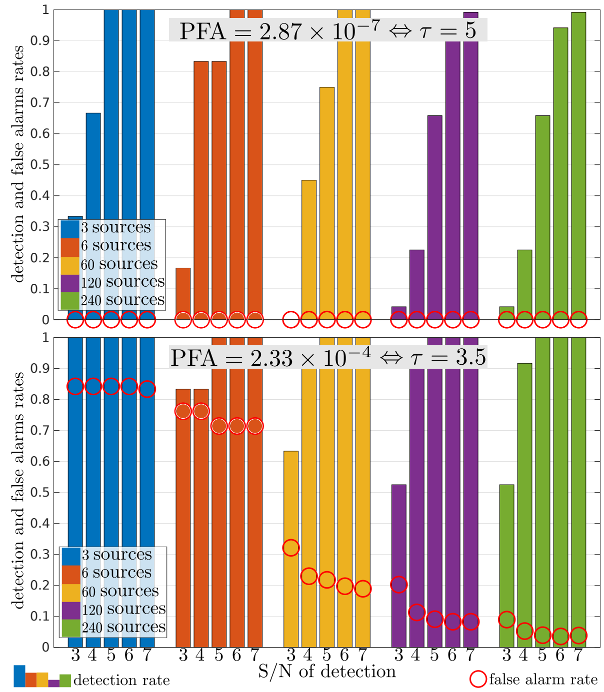

Fig. 4. Detection and false alarm rates as a function of the number of sources and their S/N of detection with a detection procedure controlling the PFA.

obtained by evaluating the criterion (4) in each point of the field of view, and the meaningful detections are extracted with the procedure proposed in Section III to control the FDR.

Figure 3 represents the resulting mean FDR as a function of the prescribed level q for various typical lengths of correlation in the detection maps. The FDR is well controlled for low and moderate lengths of correlation while it exceeds the prescribed level q when the length of correlation is higher than about 15 pixels. An interesting feature is that the FDR is well controlled for the typical length of correlation (about 9 pixels) encountered on the detection maps from high-contrast imaging of exoplanets, see zoom insets in Figure 2.

B. Application to real data

We now test the proposed procedure on real data in the context of the detection of exoplanets by high-contrast imaging [10]. High-contrast imaging consists in observing a star and its close environment hosting potential exoplanets and background stars on short temporal exposures. While the observations are conducted with a coronagraph to mask most of the light from the star, the detection of the exoplanets remains extremely difficult due to very high-contrast between the host star and the exoplanets (typically greater than 10^3 in the recorded images). In this context, we have recently proposed a processing algorithm, named `PACO` (for `PA`tch `CO`variances), dedicated to the detection of exoplanets in high-contrast imaging [2], [11]. `PACO` follows the general detection framework presented in Section II since it includes whitening the spatial correlations of the data and a matched filter approach. `PACO` also implements several features to deal with the specificities of the data. The distribution of the underlying detection criterion in the absence of exoplanets is well controlled: it follows closely a centered Gaussian law with unit variance. We have shown on hundreds of datasets from the European SPHERE instrument operating at the Very Large Telescope that `PACO` offers a significantly better detection sensitivity than existing processing methods.

In this section, we illustrate how the detection of the exoplanets can be conducted using the procedure proposed in section III to control the FDR. To quantify the performance, we first perform a numerical experiment: we inject realistic exoplanet signatures in a real dataset obtained from the SPHERE instrument where no real source was ever detected. The number of synthetic sources added to the data varies between 3 and 240 and their detection S/N in the detection maps produced by `PACO` varies between 3 and 7.

We first consider a detection strategy based on a control of the PFA. Sources are detected successively in decreasing S/N order by a matching pursuit algorithm until a fixed detection threshold τ (corresponding to a prescribed PFA) is reached. Figure 4 reports the detection and false alarm rates for each tested case and emphasizes the limitation of a detection procedure based on the control of the PFA. While there is no gain to decrease the detection threshold when the sources are rare and bright, it could be beneficial to adapt the detection threshold when sources of weak intensity are numerous. For example, the detection rate goes from about 5% to 55% when the detection threshold is set to 3.5 instead of 5 in the case with 240 very faint sources in the field of view (leftmost green bars). At the same time, the increase of the false alarm rate is limited: from 0% to 10%.

We then perform the detection by controlling the FDR with the proposed procedure. Figure 5 represents the achieved FDR as a function of the prescribed level q for different numbers of sources and for two levels of source intensities. An example of the graph p-values versus rank of detection is also given for a fixed value of q. This shows that the FDR is relatively well controlled in the sense that it almost never exceeds the prescribed level q. The conservative discrepancy (between 5% and 10%) can be attributed to the non-stationarity of the correlation structures of the detection maps: the typical length of correlation at the center of the field of view is quite different from farther from the star. Using a single equivalent number

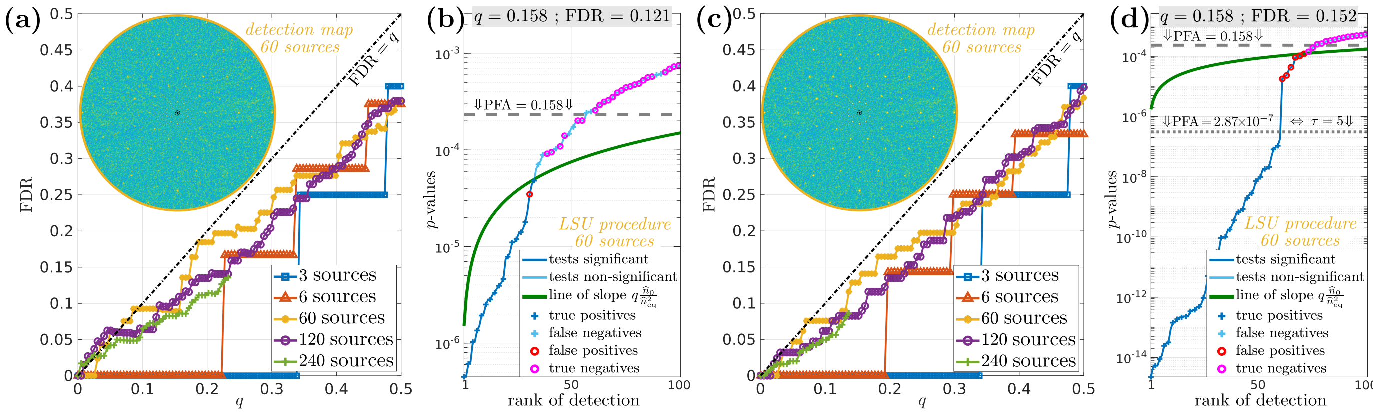

Fig. 5. (a,c): FDR as a function of the prescribed level q for different numbers of sources injected in real data. (b,d): Examples of the graph p-values versus rank of detection for $q = 0.158$, the detections are classified as true/false positives/negatives. The tests fulfilling $p_{(i)} \le iq\frac{\widehat{n}_0}{\widehat{n}_{\mathrm{eq}}^2}$ (i.e. p-values below the green curve) are declared significant. The results are given for two levels of intensity: the mean S/N of detection of the sources in the PACO detection maps is about 4 (a,b) and 6 (c,d), see the insets in (a,c).

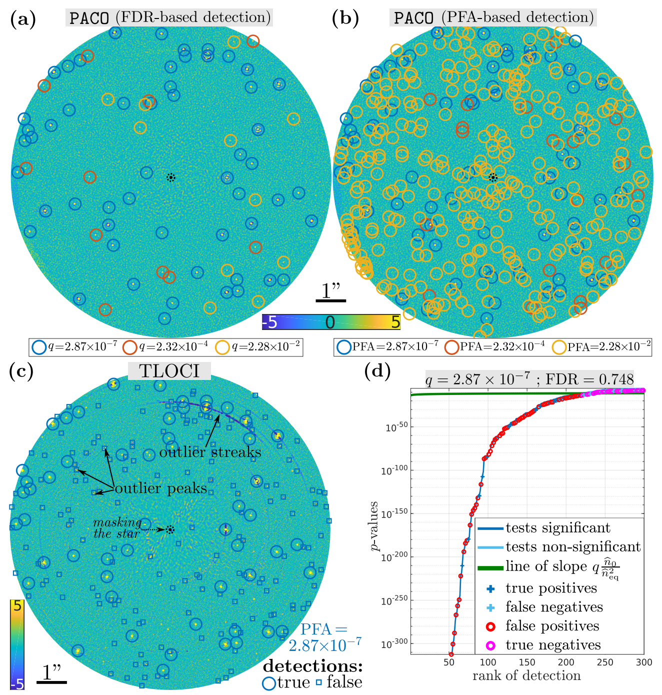

Fig. 6. (a,b) Detection map obtained with PACO, controlling the FDR (a) or the PFA (b). The detection map obtained with the widely used method TLOCI (c) displays an unexpected number of false alarms due to a mismodeling of outliers and spatial correlations in the data (false alarms were labeled by an expert), the FDR procedure (d) understandably fails in this case.

$\widehat{n}_{\mathrm{eq}}$ of independent tests is only a simplified consideration for these correlations. The estimate $\widehat{n}_1$ of the number of sources is also satisfactory since it is between $0.78n_1$ and $1.05n_1$ depending on the level of intensity of the sources.

Given the successful validation of our procedure, we now consider another dataset from the SPHERE instrument that contains many real sources. Figure 6(a,b) gives the detection results obtained with the PACO algorithm under control of either the PFA or the FDR. This shows the benefit of the proposed procedure to control the number of errors among the number of discoveries instead of the number of errors among the total number of tests that are wrongfully assumed independent. Beyond the design of the procedure extracting the meaningful detections, the modelization of the background noise of the data also plays a critical role in the global performance of a detection method. As illustrated by Figure 6(c,d), a widely used algorithm for the processing of high-contrast data (TLOCI) produces many more false alarms than expected. This is due to a mismodeling of the correlations and outliers in SPHERE observations, thus preventing the automatic and reliable identification of the sources, whatever the extraction procedure.

V. Conclusion

In this paper, we have investigated the extraction of meaningful detections from spatially correlated detection maps by controlling the FDR. We have shown that an estimate of the equivalent number of independent tests and of the number of sources led to the accurate control of the FDR both in simulated cases and on experimental data.

References

[1] O. Besson, E. Chaumette, and F. Vincent, “Adaptive detection of a Gaussian signal in Gaussian noise,” in *IEEE CAMSAP*, 2015.

[2] O. Flasseur, L. Denis, É. Thiébaut, and M. Langlois, “Exoplanet detection in angular differential imaging by statistical learning of the nonstationary patch covariances - The PACO algorithm,” *A&A*, 2018.

[3] Y. Benjamini and Y. Hochberg, “Controlling the false discovery rate: a practical and powerful approach to multiple testing,” *Journal of the Royal statistical society: series B (Methodological)*, 1995.

[4] C. Meillier, F. Chatelain, O. Michel, and H. Ayasso, “Error control for the detection of rare and weak signatures in massive data,” in *European Signal Processing Conference*, 2015.

[5] J. Neyman and E. S. Pearson, “On the problem of the most efficient tests of statistical hypotheses,” *Philosophical Transactions of the Royal Society of London*, 1933.

[6] Y. Benjamini, A. M. Krieger, and D. Yekutieli, “Adaptive linear step-up procedures that control the false discovery rate,” *Biometrika*, 2006.

[7] Y. Benjamini and Y. Hochberg, “On the adaptive control of the false discovery rate in multiple testing with independent statistics,” *Journal of educational and Behavioral Statistics*, 2000.

[8] Y. Benjamini and D. Yekutieli, “The control of the false discovery rate in multiple testing under dependency,” *Annals of statistics*, 2001.

[9] S. Lobry, L. Denis, and F. Tupin, “Multitemporal SAR image decomposition into strong scatterers, background, and speckle,” *IEEE Journal of Selected Topics in Applied Earth Obs. and Remote Sensing*, 2016.

[10] L. Pueyo, “Direct imaging as a detection technique for exoplanets,” *Handbook of Exoplanets*, 2018.

[11] O. Flasseur, L. Denis, É. Thiébaut, and M. Langlois, “Robustness to bad frames in angular differential imaging: a local weighting approach,” *A&A*, 2020.