ExPACO: detection of an extended pattern under nonstationary correlated noise by patch covariance modeling

Olivier Flasseur, Loïc Denis, Éric Thiébaut, Thomas Olivier, Corinne Fournier

ExPACO: detection of an extended pattern under nonstationary correlated noise by patch covariance modeling

Olivier Flasseur†, Loïc Denis†, Éric Thiébaut‡, Thomas Olivier†, and Corinne Fournier†

† Univ Lyon, UJM-Saint-Etienne, CNRS, Institut d Optique Graduate School,
Laboratoire Hubert Curien UMR 5516, F-42023, Saint-Etienne, France
‡ Univ Lyon, Univ Lyon1, ENS de Lyon, CNRS,
Centre de Recherche Astrophysique de Lyon UMR 5574, F-69230, Saint-Genis-Laval, France

***Abstract*—In several areas of imaging, it is necessary to detect the weak signal of a known pattern superimposed over a background. Because of its temporal fluctuations, the background may be difficult to suppress. Detection of the pattern then requires a statistical modeling of the background. Due to difficulties related to (i) the estimation of the spatial correlations of the background, and (ii) the application of an optimal detector that accounts for these correlations, it is common practice to neglect them.**

In this work, spatial correlations at the scale of an image patch are locally estimated based on several background images. A fast algorithm for the computation of detection maps is derived. The proposed approach is evaluated on images obtained from a holographic microscope.

***Index Terms*—matched filter, patch, shrinkage covariance estimator, correlation**

I. Introduction

The detection of a faint pattern modeled by a few parameters is a common image processing task encountered in many fields (microscopy, astronomy, radar imaging, etc.). In particular, in holographic microscopy, small spherical objects such as cocci bacteria, bubbles or droplets can be imaged by recording the diffraction patterns they produce under a coherent illumination. The resulting image, called *a hologram*, can be modeled by Fourier optics or Mie scattering theory. The diffraction pattern produced by each object depends only on a few parameters (3D object location and size) [1]. Hologram reconstruction requires the detection and localization of each diffraction pattern in the data. In astronomy, the detection and accurate localization of very faint sources is crucial to detect exoplanets or to study the gravitational interaction of stars in the vicinity of the central black hole of our galaxy [2].

These detection and localization tasks become very difficult when the amplitude of the pattern of interest is small compared to the background fluctuations and under the presence of nonstationary and correlated background structures.

Our contributions: We derive an algorithm for the detection of a shift-invariant pattern in an image. This algorithm accounts for the local correlations of the background. It uses a parameter-free shrinkage covariance estimator for local patch covariance modeling. Computation of a detection map by direct application of the generalized likelihood ratio test has a prohibitive computational complexity when the pattern is spatially extended. An efficient algorithm based on fast Fourier transforms is derived to approximate the detection map.

This work has been supported by grants from CNRS (RESSOURCES project, Défi Imag'In), région Auvergne-Rhône-Alpes (project DIAGHOLO) and ANR-11-LABX-0063 / ANR-11-IDEX-0007.

II. Statistical modeling of the background

We focus on the case where several images of the background are recorded, in the absence of the objects of interest, before performing the detection tasks. Because of the temporal fluctuations of the background, subtracting an average background or even a linear combination of the background images is not sufficient to efficiently remove the background from the image: spatially-structured residuals remain and degrade the detection performance.

To overcome this limitation, a statistical modeling of the background fluctuations can be built in order to better decide at detection time which part of the signal may be ascribed to the pattern to detect and which is more likely due to a typical temporal fluctuation of the background.

Background fluctuations are generally nonstationary and spatially correlated. Capturing those fluctuations at the scale of an N-pixels image from only a few tens up to a few hundreds of background images is very difficult: the covariance matrix has $\mathcal{O}(N^2)$ terms ($\mathcal{O}(\cdot)$ corresponds to Bachmann–Landau asymptotic notation). Either some structure has to be assumed for the covariance matrix or a more local modeling is necessary.

In the domain of statistical modeling of natural images, patches (i.e. small windows of a few pixels wide) have emerged as an adequate scale at which spatial correlations can be captured. A multivariate Gaussian modeling of stacks of similar patches is used in NL-Bayes algorithm [3] in order to estimate denoised images. In the EPLL (Expected Patch Log Likelihood) model [4], the statistics of natural images is captured as a product of Gaussian mixture models defined over each patch that can be extracted from the image. There are numerous other works that base the modeling of images on a decomposition into image patches, from non-local approaches

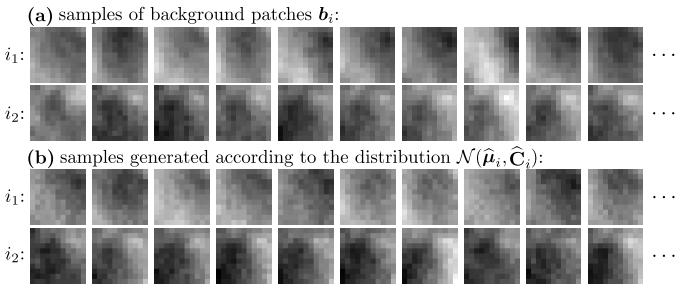

Fig. 1. Illustration of the local model for background patches: **(a)** sample 11×11 patches from experimental backgrounds, extracted at 2 different locations; **(b)** patches generated from the multivariate Gaussian model locally learned from 60 background samples.

[5], fields of experts [6], higher-order Markov random fields [7], patch dictionaries [8], up to the recent generalized use of convolutional neural networks that learn a model based on an input patch (whose size depends on the so-called receptive field, i.e., the depth and structure of the network).

In the context of exoplanet detection by direct imaging [2], a weak point-like source has to be detected over a spatially structured background (the host star is only imperfectly masked out by a coronagraph, which causes this undesirable background). We have recently proposed an exoplanet detection algorithm, called PACO [9], [10], based on a local modeling of the PAtch COvariances. In this paper, we follow the same statistical modeling and consider the adaptation of the method to the detection of spatially Extended patterns (ExPACO).

The background component $\boldsymbol{b} \in \mathbb{R}^N$ is decomposed into patches. The patch $\boldsymbol{b}_i \in \mathbb{R}^K$ extracted around pixel i is modeled by a multivariate Gaussian: $\boldsymbol{b}_i \sim \mathcal{N}(\boldsymbol{\mu}_i, \mathbf{C}_i)$, where both the mean $\boldsymbol{\mu}_i$ and the covariance matrix $\mathbf{C}_i$ are estimated locally from the set of background images. $\boldsymbol{\mu}_i$ is estimated by the sample mean: $\widehat{\boldsymbol{\mu}}_i = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{b}_{i,t}$, where the notation $\boldsymbol{b}_{i,t}$ indicates the K-pixels patch centered at pixel i, in the t-th background image. To estimate the covariance matrix of size $K \times K$ from T background images, the sample covariance matrix $\widehat{\mathbf{S}}_i = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{b}_{i,t} - \widehat{\boldsymbol{\mu}}_i)(\boldsymbol{b}_{i,t} - \widehat{\boldsymbol{\mu}}_i)^{\mathrm{t}}$ can be used when $T \gg K$. When $T \approx K$, $\widehat{\mathbf{S}}_i$ has a large variance and when $T < K$, $\widehat{\mathbf{S}}_i$ becomes rank-deficient. Ledoit and Wolf [11] and Chen *et al.* [12] improve the sample covariance matrix by shrinkage towards a matrix proportional to the identity. In [9], we extend the formula defining the shrinkage estimator as the convex combination of $\widehat{\mathbf{S}}_i$ and a diagonal covariance matrix $\widehat{\mathbf{D}}_i$ defined by $[\widehat{\mathbf{D}}_i]_{m,n} = 0$ if $m \neq n$ and $[\widehat{\mathbf{D}}_i]_{m,m} = [\widehat{\mathbf{S}}_i]_{m,m}$:

$$\widehat{\mathbf{C}}_i = \widehat{\rho}_i \widehat{\mathbf{D}}_i + (1 - \widehat{\rho}_i)\widehat{\mathbf{S}}_i \,, \tag{1}$$

where the shrinkage coefficient $\widehat{\rho}_i$ is obtained by clipping to the $[0, 1]$ range the value:

$$\widehat{\rho}_i = \frac{\mathrm{Tr}(\widehat{\mathbf{S}}_i^2) + \mathrm{Tr}^2(\widehat{\mathbf{S}}_i) - 2\sum_{k=1}^{K}[\widehat{\mathbf{S}}_i]_{kk}^2}{(T+1)(\mathrm{Tr}(\widehat{\mathbf{S}}_i^2) - \sum_{k=1}^{K}[\widehat{\mathbf{S}}_i]_{kk}^2)} \,. \tag{2}$$

Figure 1**(a)** shows some patches extracted at different locations of background images acquired with a holographic microscope (see section V and Fig.2 for details). In Figure 1**(b)**, some random realizations drawn according to the local Gaussian model $\mathcal{N}(\widehat{\boldsymbol{\mu}}_i, \widehat{\mathbf{C}}_i)$ learned from 60 background patches are displayed. Patches from two different locations i_1 and i_2 are shown on two different rows. Fluctuations around the mean background differ according to the location. Generated patches are quite similar to the patches extracted from the background images at the corresponding location.

III. Detection of an extended pattern

The joint detection / localization problem of a pattern $\boldsymbol{m}(x_0, y_0)$ centered at the 2D location (x_0, y_0) in an observed image $\boldsymbol{f}$ corrupted by a background $\boldsymbol{b}$ can be formulated as a binary hypothesis test:

$$\begin{cases} \mathcal{H}_0: & \boldsymbol{f} = \boldsymbol{b} \,, \\ \mathcal{H}_1: & \boldsymbol{f} = \boldsymbol{b} + \alpha\, \boldsymbol{m}(x_0, y_0) \;\; \text{with} \;\; \alpha > 0 \,. \end{cases} \tag{3}$$

The estimation of the 2D location (x_0, y_0) of the pattern and of its amplitude α is necessary in order to decide between the two hypotheses. The neg-log-likelihood of parameters α, x_0 and y_0 under $\mathcal{H}_1$, with our patch-based modeling and an independence assumption between patch, is:

$$-\log \mathrm{p}(\boldsymbol{f}|\mathcal{H}_1, \alpha, x_0, y_0) = \tfrac{1}{2}\sum_i \boldsymbol{r}_i^{\mathrm{t}} \mathbf{C}_i^{-1} \boldsymbol{r}_i + c, \tag{4}$$

where the residual patch $\boldsymbol{r}_i$ is obtained by removing the average background $\boldsymbol{\mu}_i$ and the contribution of the pattern $\boldsymbol{m}_i$ to the observed patch $\boldsymbol{f}_i$: $\boldsymbol{r}_i = \boldsymbol{f}_i - \boldsymbol{\mu}_i - \alpha\, \boldsymbol{m}_i(x_0, y_0)$, and where c is a constant that depends only on the sum of the log of the covariance determinants.

At a given location (x_0, y_0), the maximum likelihood estimate of the amplitude of the pattern is given by:

$$\widehat{\alpha}(x_0, y_0) = \frac{\max(b(x_0, y_0), 0)}{a(x_0, y_0)} = \frac{[b(x_0, y_0)]_+}{a(x_0, y_0)} \,, \tag{5}$$
$$\text{with} \begin{cases} a(x_0, y_0) = \sum_i \boldsymbol{m}_i(x_0, y_0)^{\mathrm{t}} \widehat{\mathbf{C}}_i^{-1} \boldsymbol{m}_i(x_0, y_0) \\ b(x_0, y_0) = \sum_i (\boldsymbol{f}_i - \widehat{\boldsymbol{\mu}}_i)^{\mathrm{t}} \widehat{\mathbf{C}}_i^{-1} \boldsymbol{m}_i(x_0, y_0) \,. \end{cases}$$

To decide in favor of hypothesis $\mathcal{H}_0$ or $\mathcal{H}_1$, for a fixed location (x_0, y_0), the generalized likelihood test (GLRT) is:

$$\log \frac{\mathrm{p}(\boldsymbol{f}|\mathcal{H}_1, \widehat{\alpha}, x_0, y_0)}{\mathrm{p}(\boldsymbol{f}|\mathcal{H}_0)} = \frac{[b(x_0, y_0)]_+^2}{a(x_0, y_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta \,,$$

which can be recast for $\eta \geq 0$ as a test on the signal-to-noise ratio of pattern $\boldsymbol{m}$ (SNRT), which follows $\mathcal{N}(0, 1)$ under $\mathcal{H}_0$, see [9]:

$$\frac{\widehat{\alpha}(x_0, y_0)}{\sigma_{\alpha(x_0, y_0)}} = \frac{b(x_0, y_0)}{\sqrt{a(x_0, y_0)}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \sqrt{\eta} = \tau \,. \tag{6}$$

The maximum likelihood location of the pattern is obtained by maximization of the SNRT over all possible locations.

So far, in our modeling only a single pattern has been considered. If several patterns are superimposed, each centered on a different 2D location, a greedy approach similar to the matching pursuit can be applied: patterns are detected one at a time by forming the SNRT, and after each detection the detected pattern is subtracted from the data so that the next pattern can be detected by applying the SNRT on the residuals.

IV. ExPACO: FAST COMPUTATION OF DETECTION MAPS

Localization of a pattern requires to maximize the SNRT. Since the SNRT is a non-convex function of the 2D location (x_0, y_0), with many local maxima observed in practice, it is necessary to systematically evaluate the SNRT over a grid to identify the global maximum. If the pattern[1] $\boldsymbol{m}$ is extended, patches $\boldsymbol{m}_i$ extracted from the reference pattern are all non-zero and the sums in equation (6) involve N terms if patches overlap, or N/K if patches do not overlap. In the following discussion, we consider that patches overlap.

Evaluating the SNRT defined in equation (6) for a given location requires $\mathcal{O}(NK^2)$ scalar operations if the inverse matrices $\widehat{\mathbf{C}}_i^{-1}$ are precomputed (which requires $\mathcal{O}(NK^3)$ operations). Therefore, to produce a SNRT map, i.e. to compute the SNRT for N locations (x_0, y_0) spanning the whole pixel grid, $\mathcal{O}(N^2K^2)$ operations are required[2]. This high computational complexity prevents a direct application of the SNRT based on the background modeling described in section II with patch covariances.

In this paper, we propose a fast algorithm to compute SNRT detection maps in the case of shift-invariant models $\boldsymbol{m}$. The algorithm is based on a reformulation of (6) that involves discrete correlations. These correlations are computed using fast Fourier transforms. We describe in turn how the numerator and the denominator of (6) can be efficiently computed.

A. Fast computation of $b(x_0, y_0)$ for all pixel shifts:

When the model is shifted, restrictions of the model to each patch $\boldsymbol{m}_i$ are modified while patches $\boldsymbol{u}_i \equiv \widehat{\mathbf{C}}_i^{-1}(\boldsymbol{f}_i - \widehat{\boldsymbol{\mu}}_i)$ remain unchanged. Let $\mathbf{U} = \begin{pmatrix}\boldsymbol{u}_1 & \cdots & \boldsymbol{u}_N\end{pmatrix}$ be the $K \times N$ matrix collecting all transformed patches $\boldsymbol{u}_i$. By application of a singular value decomposition (SVD), matrix $\mathbf{U}$ can be decomposed into a sum of K rank-one matrices:

$$\mathbf{U} = \sum\nolimits_{k=1}^{K} \boldsymbol{v}_k \boldsymbol{\beta}_k^{\mathrm{t}} \,, \tag{7}$$

where $\{\boldsymbol{v}_k\}_{k=1..K}$ are the left singular vectors (i.e. modes) and $[\boldsymbol{\beta}_k]_i$ is the coefficient of each patch $\boldsymbol{u}_i$ related to mode $\boldsymbol{v}_k$ (obtained as the product of the k-th singular value and of the i-th entry of the k-th right singular vector).

The computation of scalar products $\boldsymbol{m}_i^{\mathrm{t}}\boldsymbol{v}_k$ for all the patches $\boldsymbol{m}_i$ that can be extracted from model $\boldsymbol{m}$ is readily obtained by a 2D correlation[3]: $\boldsymbol{m}_i^{\mathrm{t}}\boldsymbol{v}_k = [\boldsymbol{m} \star \boldsymbol{v}_k]_i$, where the notation $\star$ denotes a 2D discrete correlation.

Since the transformed patch $\boldsymbol{u}_i$ can be expanded as $\boldsymbol{u}_i = \sum_{k=1}^{K}[\boldsymbol{\beta}_k]_i \, \boldsymbol{v}_k$, we obtain that $\boldsymbol{u}_i^{\mathrm{t}}\boldsymbol{m}_i = \boldsymbol{m}_i^{\mathrm{t}} \sum_{k=1}^{K}[\boldsymbol{\beta}_k]_i \, \boldsymbol{v}_k = \sum_{k=1}^{K}[\boldsymbol{\beta}_k]_i \boldsymbol{m}_i^{\mathrm{t}}\boldsymbol{v}_k = \sum_{k=1}^{K}[\boldsymbol{\beta}_k]_i [\boldsymbol{m} \star \boldsymbol{v}_k]_i$. Translating the model $\boldsymbol{m}$ leaves the weights $\boldsymbol{\beta}_k$ unchanged (they depend only on the $\boldsymbol{u}_i$) but shifts the term $\boldsymbol{m} \star \boldsymbol{v}_k$. When computing the sum over all positions i of the product $[\boldsymbol{\beta}_k]_i[\boldsymbol{m} \star \boldsymbol{v}_k]_i$ for all shifts, a correlation is performed:

$$b(x_0, y_0) = \sum\nolimits_{k=1}^{K} \left[\boldsymbol{\beta}_k \star (\boldsymbol{m}(0,0) \star \boldsymbol{v}_k)\right]_{x_0, y_0} \,. \tag{8}$$

With this formulation, the computational complexity is reduced to that of the computation of the vectors $\boldsymbol{u}_i$ ($\mathcal{O}(NK^2)$ operations if matrix inverses are precomputed), of the SVD of a $N \times K$ matrix ($\mathcal{O}(NK^2)$ operations) and of $2K$ discrete correlations (that can be computed with FFTs in $\mathcal{O}(N \log N)$ each). The total complexity (including the precomputation of matrix inverses) is then $\mathcal{O}(NK(K^2 + \log N))$, which is much better than the original complexity: $\mathcal{O}(N^2K^2)$.

B. Fast computation of $a(x_0, y_0)$ for all pixel shifts:

Computation of this term is more challenging because the $\boldsymbol{m}_i$ terms are shifted with respect to the location of the covariance matrices $\widehat{\mathbf{C}}_i$. It is not possible to derive a linear expansion and identify discrete correlations. We resort to an approximation in order to compute efficiently this term for all shifts.

Let $\mathbf{M}$ be the $K \times N$ matrix that collects all patches extracted from the centered model $\boldsymbol{m}(0,0)$: $\mathbf{M} = \begin{pmatrix}\boldsymbol{m}_1(0,0) & \cdots & \boldsymbol{m}_N(0,0)\end{pmatrix}$. To reduce the complexity, we approximate each patch $\boldsymbol{m}_i$ by a scaled version of an element taken from a $K \times P$ codebook $\mathbf{W}$:

$$\mathbf{M} \approx \begin{pmatrix} | & & | \\ \boldsymbol{w}_1 & \cdots & \boldsymbol{w}_P \\ | & & | \end{pmatrix} \mathbf{Q}^{\mathrm{t}} \tag{9}$$

where the code $\mathbf{Q}$ is a $N \times P$ matrix such that for all row, only a single entry is non-zero. If $q(i)$ is the value of the non-zero entry of the i-th row of $\mathbf{Q}$ and $p(i)$ is the column number corresponding to that entry, then the patch $\boldsymbol{m}_i$ is approximated by $q(i)\boldsymbol{w}_{p(i)}$.

Construction of the codebook can be performed using a modified version of the k-means clustering algorithm where distances to cluster centers are evaluated by normalized correlation and the update of a cluster center corresponds to a truncated SVD where only the left and right singular vectors corresponding to the largest singular value are kept. This corresponds to a particular case of K-SVD sparse coding algorithm [8] where the sparsity is equal to 1. Note that the approximation can be made exact by setting $P = N$, but in practice P will be chosen to be several orders of magnitude smaller than N.

Next, we build P maps $\boldsymbol{e}_1$ to $\boldsymbol{e}_P$ such that $[\boldsymbol{e}_k]_i = \boldsymbol{w}_k^{\mathrm{t}}\widehat{\mathbf{C}}_i^{-1}\boldsymbol{w}_k$. The k-th column $\boldsymbol{q}_k$ of the code $\mathbf{Q}$ indicates which vectors $\boldsymbol{m}_i$ are best represented by $\boldsymbol{w}_k$. If we replace each vector $\boldsymbol{m}_i$ in the sum $\sum_i \boldsymbol{m}_i^{\mathrm{t}}\widehat{\mathbf{C}}_i^{-1}\boldsymbol{m}_i$ by its closest vector in the codebook, we obtain the approximation:

$$\sum\nolimits_{i=1}^{N} \boldsymbol{m}_i^{\mathrm{t}}\widehat{\mathbf{C}}_i^{-1}\boldsymbol{m}_i \approx \sum\nolimits_{k=1}^{P} (\boldsymbol{q}_k^2)^{\mathrm{t}}\boldsymbol{e}_k \,, \tag{10}$$

where the square is applied element-wise on the vector $\boldsymbol{q}_k$. When the model $\boldsymbol{m}$ is shifted, the maps $\boldsymbol{q}_k$ indicating the

[1] for compactness, we drop in this paragraph the spatial location (x_0, y_0) in the notation $\boldsymbol{m}(x_0, y_0)$

[2] since the number K of pixels in a patch is much smaller than the total number N of pixels in an image, precomputing the inverse of covariance matrices in $\mathcal{O}(NK^3)$ is negligible compared to producing the map in $\mathcal{O}(N^2K^2)$ and is K-fold improvement compared to inverting linear systems at each location, which would lead to a total complexity of $\mathcal{O}(N^2K^3)$.

[3] in the 2D correlation, $\boldsymbol{m}$ has the size of an image while $\boldsymbol{v}_k$ has the size of a patch, the resulting correlation at a pixel i thus involves only the patches $\boldsymbol{m}_i$ and $\boldsymbol{v}_k$

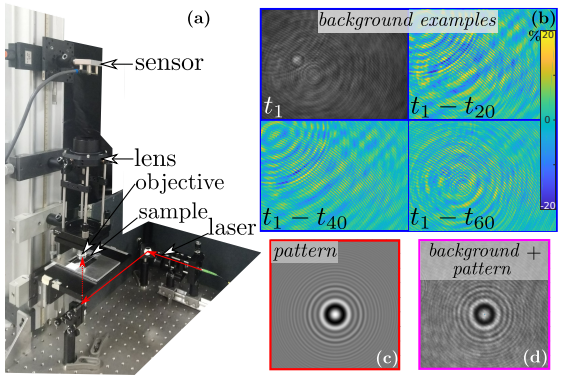

Fig. 2. **(a)** our holographic microscope; **(b)** some difference of background images; **(c)** example of a diffraction pattern to be detected in the image; **(d)** a hologram of a polystyrene microscopic bead of 2.1 microns radius.

location of patches $\boldsymbol{m}_i$ best represented by $\boldsymbol{w}_k$ are shifted accordingly. Values of the denominator of the SNRT for any pixel shift of the model can then be obtained by discrete correlations:

$$a(x_0, y_0) \approx \left[\sum_{k=1}^{P} \boldsymbol{q}_k \star \boldsymbol{e}_k \right]_{x_0, y_0}. \tag{11}$$

The complexity of the procedure is the following: application of the k-means algorithm for a fixed number of iterations is performed in $\mathcal{O}(PN)$, computation of the maps $\{\boldsymbol{e}_k\}_{k=1..P}$ requires $\mathcal{O}(PNK^2)$ operations, computation of the P correlations is performed in $\mathcal{O}(PN \log N)$ with FFTs. The total complexity is thus $\mathcal{O}(PN(\log N + K^2))$. This corresponds to a strong improvement compared to the original complexity (in $\mathcal{O}(N^2K^2)$) if we choose a codebook size $P \ll N$. Once the (approximate) location of the pattern has been identified, a local optimization based on the exact evaluation of (6) can be performed (i.e. no bias is incurred, only the location of the pattern may be missed if the approximation is too coarse and the local optimization then leads to the wrong local optimum).

V. Results

In this section, we assess both the accuracy and the detection performance of ExPACO. For this purpose, we use experimental background images recorded with the holographic microscope presented in Fig.2**(a)**. Difference of 4 background images out of the 61 that we collected are shown in Fig.2**(b)**. Fluctuations in the background are due to small variations of the optical path lengths between optical surfaces (because of mechanical vibrations or thermal inhomogeneities of the air) and are difficult to fully model. The diffraction pattern created by a spherical bead is displayed in Fig.2**(c-d)**. The bead size and the refractive index contrast between the bead and the surrounding medium (approximately 0.08 here) is sufficient to obtain quite contrasted diffraction fringes. Identification of the diffraction pattern becomes critical when this index contrast or the object size drops. To evaluate the performance of ExPACO, we added a model of the diffraction pattern created by a microscopic bead to some background images, the other background images were kept to estimate the background statistics $\widehat{\boldsymbol{\mu}}_i$ and $\widehat{\mathbf{C}}_i$.

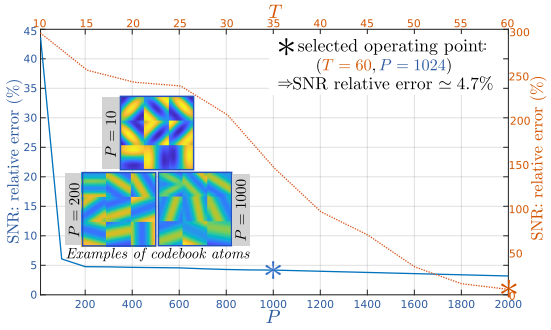

Fig. 3. Evolution of the relative error induced by our fast approximation as a function of the number P of atoms in the codebook $\mathbf{W}$ and of the number T of background images available to estimate $\{\widehat{\boldsymbol{\mu}}_i\}_{i=1..N}$ and $\{\widehat{\mathbf{C}}_i\}_{i=1..N}$.

Fig.3 evaluates the accuracy of the approximation presented in section IV to compute efficiently SNR maps. Square patches of size 7×7 are used. Since the evaluation of the SNR by direct application of equation (6) has a prohibitive complexity, we compare values of the SNR on only 2000 different locations (x_0, y_0). Application of (6) takes 12 hours to compute these 2000 locations on 20 CPU cores, while our fast approximation produces the SNR map for one million different locations in about 2 minutes (computation of the codebook which must be done only once for a given pattern $\boldsymbol{m}$ takes another 2 minutes). As expected, Fig.3 shows a reduction of the approximation error of the SNR when the size of the codebook P increases (blue curve). With $P \approx 200$ elements, the codebook $\mathbf{W}$ is large enough to capture most of the geometrical structures of the pattern and to obtain an approximation error below 5%.

The impact of the number of background images in the estimation of the background statistics ($\widehat{\boldsymbol{\mu}}_i$ and $\widehat{\mathbf{C}}_i$) is assessed by comparing values of the SNR obtained when $\widehat{\boldsymbol{\mu}}_i$ and $\widehat{\mathbf{C}}_i$ are estimated from an increasing number of backgrounds (the reference SNR is obtained when 60 backgrounds are used). With 7×7 patches, covariance matrices are 49×49. With less than 50 background samples, the obtained SNR map differs significantly from the SNR map obtained with all backgrounds.

In the following experiments, we set $T = 60$, $P = 1024$, and add pattern with a very low amplitude on the remaining 61st background image. As a baseline, we consider a background model with a diagonal covariance matrix: this corresponds to setting $\widehat{\rho}_i = 0$ for all i in equation (2). Only the structures that are in the mean background $\boldsymbol{\mu}$ are taken into account in this baseline method, denoted *diagonal covariance* in the following.

In Fig.4, we compare SNR maps obtained with a diagonal covariance and with ExPACO for the 4 different patch shapes shown at the top of the figure. Each of the patches has 25 pixels: the first patch is a 5×5 full patch, the remaining 3 patches are patches with holes: a 19×19 patch, a 29×29 patch and a 39×39 patch. The rationale in increasing the size of the patch while keeping the same number of pixels is to capture longer-range correlations without increasing the size of the covariance matrices $\mathbf{C}_i$. In each SNR map, a single peak is expected at the location circled in pink, corresponding to the position (x_0, y_0) of the pattern that has been added to the background image (contrast between the background

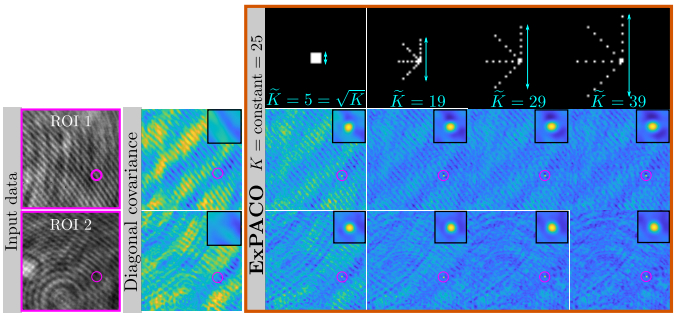

Fig. 4. Detection of a pattern in a structured background: diagonal covariance *vs* ExPACO with various patch shapes.

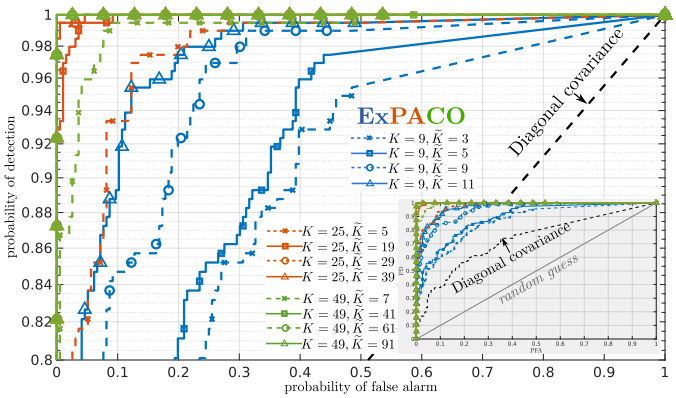

Fig. 5. ROC curves of the diagonal covariance detector and of ExPACO for various patch sizes and patch shapes. The inset shows an unzoomed to better locate the diagonal covariance detector with respect to ExPACO.

and the pattern is about 140). This peak is visible only with ExPACO. Moreover, with more extended patches, the SNR displays much fewer undesirable fluctuations: typical background structures are better modeled and are thus less likely to be mistaken for the pattern.

The influence of the patch size and of the number of pixels within a patch is more systematically studied through receiver operating curves (ROC) curves. Patterns are injected one at a time, at different locations, on various backgrounds. Fig.5 reports the evolution of the ROC curve with the number of pixels K in the patch and with the spatial extent $\widetilde{K}$ of the patch. A clear improvement is observed with respect to a detector based on a diagonal covariance assumption. The detection performance of ExPACO improves when larger patches and more pixels are taken into account, since this allows to capture longer-range correlations. In these simulations where the pattern is 140 times fainter than the background, using $K = 25$ pixels spread over areas of 19×19 to 39×39 pixels leads to the detection of 100% of the patterns without false alarms. In contrast, the diagonal covariance detector reaches only 15% of detections without false alarms. The number of pixels K and the spatial extension $\widetilde{K}$ should not be too large: a large K leads to a large covariance matrix which can not be accurately estimated from a limited number of background images (the shrinkage step strongly bias the covariance towards a diagonal covariance matrix if $K \gg T$); if $\widetilde{K}$ is too large, shorter-range correlations are lost and the

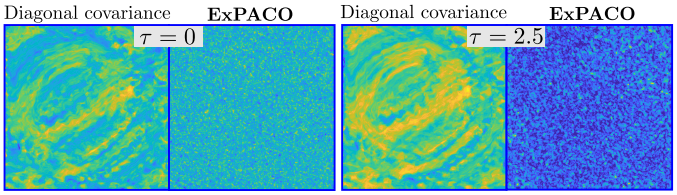

Fig. 6. False alarm rates, for two different thresholds.

model is less local (hence less adapted to highly nonstationary backgrounds). We found in our experiments that the choices $K \geq 110$ (pixels) $\simeq 2\,T$ and $\widetilde{K} \geq 50$ (patch width in pixels) both lead to a degradation of ExPACO performance.

A desirable property for detectors is to lead to a constant false alarm rate. In the case of nonstationary backgrounds, this requires a robustness to the differences in the structures found in the background. Fig.6 shows maps of the false alarm rate reached by the diagonal covariance estimator and by ExPACO. These maps were built by reporting the fraction of positive SNR tests in the absence of patterns (i.e. under $\mathcal{H}_0$), for different background images (1 background is selected and the 60 remaining are used to estimate the $\widehat{\boldsymbol{\mu}}_i$ and $\widehat{\mathbf{C}}_i$, in a leave-one-out rotation). Contrary to the diagonal covariance detector, the false alarm rate of ExPACO is stationary.

VI. Conclusion

We introduced ExPACO, a fast algorithm to detect extended patterns in nonstationary and correlated backgrounds. A local model of the spatial correlations is learned from a few tens of background images. By accounting for these correlations, the detection performance is improved compared to a standard detector. Patches with holes are shown to lead to a good trade-off between the size of the covariance matrices to estimate and the range of the correlations that are captured.

References

[1] C. Wang, X. Zhong, D. B. Ruffner, A. Stutt, L. A. Philips, M. D. Ward, and D. G. Grier, "Holographic characterization of protein aggregates," *Journal of pharmaceutical sciences*, 2016.

[2] G. Chauvin, S. Desidera, A.-M. Lagrange, A. Vigan *et al.*, "Discovery of a warm, dusty giant planet around HIP 65426," *A&A*, 2017.

[3] M. Lebrun, A. Buades, and J.-M. Morel, "A nonlocal bayesian image denoising algorithm," *SIAM Journal on Imaging Sciences*, 2013.

[4] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," 2011.

[5] A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling & Simulation*, 2005.

[6] S. Roth and M. J. Black, "Fields of experts: A framework for learning image priors," 2005.

[7] Y. Chen, R. Ranftl, and T. Pock, "Insights into analysis operator learning: From patch-based sparse models to higher order MRF," *IEEE TIP*, 2014.

[8] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE TIP*, 2006.

[9] 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.

[10] O. Flasseur *et al.*, "An unsupervised patch-based approach for exoplanet detection by direct imaging," in *IEEE ICIP*, 2018.

[11] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *J. Multivariate Anal.*, 2004.

[12] Y. Chen, A. Wiesel, Y. C. Eldar, and A. O. Hero, "Shrinkage algorithms for MMSE covariance estimation," *IEEE Trans. on Sig. Proc.*, 2010.