Sources separation in hyperspectral images for protoplanets detection and characterization

Rémi Julo, Florent Chatelain, Mickaël Bonnefoy, Olivier J J Michel, Olivier Flasseur, Philippe Delorme, Sebastián Jorquera

Sources separation in hyperspectral images for protoplanets detection and characterization

R. Julo[1,2], F. Chatelain[2], M. Bonnefoy[1], O.J.J. Michel[2], O. Flasseur[3], P. Delorme[1], S. Jorquera[4,5,6]

[1]Univ. Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France, [2]Univ. Grenoble Alpes, G. INP, CNRS, Gipsa-lab, 38400 Saint-Martin-d'Hères, France, [3]Univ. Lyon 1, ENS Lyon, CNRS, CRAL, 69230 Lyon, France, [4]Univ. de Chile, 7550000 Las Condes, Chile, [5]Univ. Côte d'Azur, 06304 Nice, France, [6]NPF, 2340000 Valparaíso, Chile

***Abstract*—We propose a new method to attenuate the nuisance of the flux halo of stars in hyperspectral images and recover the spatio-spectral sparse signals of forming extrasolar planets. The proposed approach takes into account the punctuality of the sources and the spatio-spectral point spread function of the instrument. We show that our linear and parametric method based on polynomial modulation avoids distortions in the planets spectra compared to present algorithm involving frequency filtering. Simulated data demonstrate the enhancements. We also study the influence of model complexity on the quality of the results. We verify the improvement on real data of observations of the PDS 70 system with the MUSE instrument.**

***Index Terms*—Hyperspectral images, Detection, Spectral unmixing, Protoplanets characterization, Polynomial modulation.**

I. Introduction

Young analogues of Jupiter form within gas-rich disks surrounding stars during their first 10^6 years [1]. The accretion of the gas (mostly Hydrogen and Helium) onto the planet embryo (protoplanet) is producing narrow emission lines at specific wavelengths within the optical and near-infrared spectral range (0.4-2.5μm). These sparse signals have been recently targeted by ground-based telescopes equipped with Integral Field Spectrographs (IFS), allowing to record hyperspectral images of the sky [2], [3]. Actually, the contrast between the emission lines of the protoplanets and the thermal continuum emission from the star is higher at these specific wavelengths and ease the detection of the planets. The emission lines can also be used to study the poorly understood accretion mechanisms occurring onto the protoplanets [4], [5].

The Multi Unit Spectroscopic Explorer (MUSE) of the Very Large Telescope (VLT) [6] is a powerful instrument for this task. It is an IFS operating at optical wavelengths (0.47-0.93μm) and medium-spectral resolution ($R_\lambda = \lambda/\Delta\lambda$ from 2000 to 4000). It allows to target the bright Hα emission line (6583Å) of protoplanets. The instrument is equipped with adaptive optics, in order to restore the angular resolution of the telescope [7], [8] required for detecting protoplanets close to their host stars. MUSE produces stacks of monochromatic images whose pixels correspond to 0.025" size on sky and offer a hyperspectral data diversity paving the way to leverage innovative stellar halo subtraction methods. While main state-of-the-art detection approaches to tackle the presence of the halo rely on data based differential methods that require difficult registration or rescaling steps, MUSE led to the direct detection of two protoplanets (the only two known) around the young star PDS 70 thanks to techniques using this spectral diversity [9]. However, present halo subtraction methods are based on non-linear and non-parametric models, which distort planetary spectra and limit the efficiency and robustness of the detections obtained on these data.

In this article, we start describing the data and the detection problem as a sources separation problem. We then explain the details of our new robust method (allowing simultaneously better modeling of the star contribution and better preservation of the sought signals of protoplanets). We conclude with simulations results highlighting our solution improvements, and with on-sky data results to connect with astronomy reality.

We acknowledge support from the French National Research Agency (ANR) through project grant ANR-20-CE31-0012 and the Programmes Nationaux de Planétologie et de Physique Stellaire (PNP and PNPS).

II. Sources mixture problem

A. Notations and Definitions

$\boldsymbol{a}$, $\boldsymbol{A}$, $\boldsymbol{\mathcal{A}}$ respectively denote a vector, a matrix and a cube. The set of the first ℓ integers is $[\ell] = \{1, \dots, \ell\}$.

A hyperspectral image is a data cube. It is both a stack of monochromatic images (spatial matrices) at wavelengths λ_i for $i \in [\ell]$, ℓ being the number of wavelengths, and a matrix of *spaxels* (spectral pixels), i.e. vectors of spectra acquired at positions $(x, y) \in [n] \times [m]$, n and m being respectively the number of rows and columns of the monochromatic images. For a hyperspectral cube $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{\ell \times n \times m}$, the spaxel at position (x, y) is denoted $\boldsymbol{a}_{xy} \in \mathbb{R}^{\ell}$, and the ith monochromatic image is denoted $\boldsymbol{A}_i \in \mathbb{R}^{n \times m}$.

Finally, $\otimes$ denote the outer product between vectors or matrices, $*$ the Fredholm operator (convolution generalization to a potentially variant input), and $\cdot$ the Hadamard (entrywise) product between vectors of the same size.

B. Observation model

Stars and planets are not resolved by the optical imaging system, and are consequently considered as point sources. We define $\boldsymbol{\mathcal{A}}$ as the perfect observation cube of a number u of point sources (e.g., unresolved planets or stars) at position

(x_j, y_j) with spectrum $\check{\boldsymbol{a}}^{(j)}$ respectively, for $j \in [u]$. This ideal cube thus expresses as:

$$\boldsymbol{\mathcal{A}} = \sum_{j=1}^{u} \check{\boldsymbol{a}}^{(j)} \otimes \boldsymbol{\Delta}^{(j)}\,, \tag{1}$$

where $\boldsymbol{\Delta}^{(j)} \in \mathbb{R}^{n \times m}$ is the indicator matrix which equals 1 at position (x_j, y_j), 0 otherwise.

However, due to diffraction, the observation cube $\boldsymbol{\mathcal{O}}$ is a degraded version of $\boldsymbol{\mathcal{A}}$. This is blurred by the spatio-spectral Point Spread Function (PSF) of the instrument, assuming that the optical imaging system is linear. As a complete modeling of the PSF is complex, the following spatio-spectral separation, which gives realistic results, is commonly used: $\mathrm{PSF}_i(i', x, y) = \mathrm{LSF}_i(i') \times \mathrm{FSF}_i(x, y)$, with LSF the spectral *Line Spread Function* and FSF the spatial *Field Spread Function*, both wavelength dependent in i. This allows to understand the observation cube wavelength by wavelength. We note $\boldsymbol{a}^{(j)} = \mathrm{LSF} * \check{\boldsymbol{a}}^{(j)} \in \mathbb{R}^{\ell}$ the spectrum of the jth source blurred by the LSF operator. $\boldsymbol{\phi}_i^{(j)} = \mathrm{FSF}_i * \boldsymbol{\Delta}^{(j)} \in \mathbb{R}^{n \times m}$ is the FSF matrix[1] at the ith wavelength and centered in (x_j, y_j). Then, a noiseless observation of the ith monochromatic image of the cube $\boldsymbol{\mathcal{O}}$, degraded by the instrument response, reads:

$$\boldsymbol{O}_i = \sum_{j=1}^{u} a_i^{(j)} \boldsymbol{\phi}_i^{(j)}, \quad \forall i \in [\ell]\,, \tag{2}$$

where $a_i^{(j)}$ is the ith entry of the jth source spectrum $\boldsymbol{a}^{(j)}$.

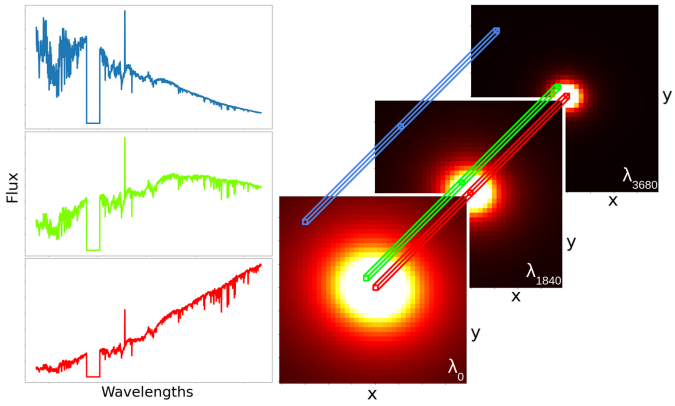

Fig. 1. Stellar cube after spatio-spectral spread, from PDS 70 star observation by MUSE simulation. On the right, three monochromatic images are shown at three different spectral channels λ_0, λ_{1840} and λ_{3680}. On the left, three spaxels are shown at different distances from center, highlighting spectra distortion caused by spatial spread evolution with wavelength.

In the perfect case where the instrument is diffraction limited with a circular aperture, the FSF is theoretically an Airy disk. However, due to the complex deformation caused by the adaptive optic system, modeling the FSF is difficult. We rather define the *Chromatic Spread Function* (CSF), which illustrates the (position-dependent) spectra deformation as highlighted in Figure 1. Let $\boldsymbol{\Phi}^{(j)} = (\boldsymbol{\phi}_i)_{i \in [l]} \in \mathbb{R}^{\ell \times n \times m}$ be the cube of the FSF matrices centered around the jth source. Then the CSF vector of the jth source at position (x, y), denoted by $\boldsymbol{\alpha}_{xy}^{(j)}$, is defined as the (x, y) spaxel of the FSF cube $\boldsymbol{\Phi}^{(j)}$. This leads us to represent the cube $\boldsymbol{\mathcal{O}}$ as a matrix of spaxels, defined as:

$$\boldsymbol{o}_{xy} = \sum_{j=1}^{u} \boldsymbol{a}^{(j)} \cdot \boldsymbol{\alpha}_{xy}^{(j)}, \quad \forall (x, y) \in [n] \times [m]\,. \tag{3}$$

For the considered problem, the astrophysical sources reduce to a single star hosting $u_p \geq 0$ protoplanets (see Fig. 2). The data cube can be decomposed as $\boldsymbol{\mathcal{O}} = \boldsymbol{\mathcal{S}} + \boldsymbol{\mathcal{P}}$, where $\boldsymbol{\mathcal{S}}$ and $\boldsymbol{\mathcal{P}}$ are the stellar and the (proto)planets cube respectively.

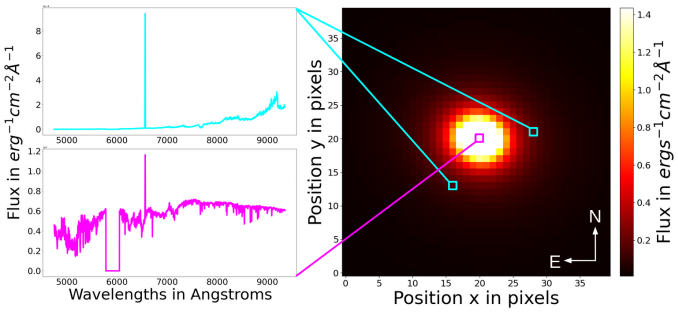

Fig. 2. Real PDS 70 system image by MUSE (spectral integration of all the wavelengths) with PDS 70 b theoretical observed planet spectrum in cyan and PDS 70 theoretical observed star spectrum in magenta. Magenta square indicates PDS 70 position while cyan squares indicate PDS 70 bc positions.

To complete our model, we finally consider an electronic noise which is i.i.d., centered, and symmetrically distributed. We assume here that the photon noise (following a Poisson distribution depending on flux) is not significant where we expect to find protoplanets. All these assumptions result in equation (4), where the observed hyperspectral data cube $\boldsymbol{\mathcal{D}}$ is composed of the cube of the planets $\boldsymbol{\mathcal{P}}$ (the sought sources), the cube of the star $\boldsymbol{\mathcal{S}}$ (the strong nuisance component) and the cube of observation noise $\boldsymbol{\mathcal{E}}$:

$$\boldsymbol{\mathcal{D}} = \boldsymbol{\mathcal{S}} + \boldsymbol{\mathcal{P}} + \boldsymbol{\mathcal{E}}. \tag{4}$$

III. Estimated Nuisance Subtraction

We present in this part all the details of our method, the aim of which is to solve the problem of spectrum distortion after subtraction, using a linear parametric model of the CSF.

In hyperspectral data, the estimation $\hat{\boldsymbol{\mathcal{P}}}$ of the cube $\boldsymbol{\mathcal{P}}$ amounts to detect the positions of a few possible planets and to characterize them from their spectrum. We assume the presence of the Hα line at 6563Å, in both the stellar spectrum and protoplanets spectra. This line, exhibiting an elevated flux compared to the continuum (low-frequency component of the spectrum signal), serves as the most reliable indicator of accretion and consequently the formation of planets. We primarily focus on detecting this signal which is both spatially and spectrally sparse, as illustrated in Fig. 2. The main challenges lie in the fact that the Hα line is also present in the spectrum of the star (due to similar phenomena), associated to a flux which is much higher than the flux of the protoplanets.

[1]The FSF is L_1-normalized such that the flux of the source is preserved.

Estimation methods (as state-of-the-art ones) are based on the stellar halo subtraction:

$$\hat{\boldsymbol{\mathcal{P}}} = \boldsymbol{\mathcal{D}} - \hat{\boldsymbol{\mathcal{S}}}\,. \tag{5}$$

i.e. the direct subtraction of the estimated nuisance $\hat{\boldsymbol{\mathcal{S}}}$. According to (4), perfect stellar halo estimation $\hat{\boldsymbol{\mathcal{S}}} = \boldsymbol{\mathcal{S}}$ leads to a planet estimate $\hat{\boldsymbol{\mathcal{P}}} = \boldsymbol{\mathcal{P}} + \boldsymbol{\mathcal{E}}$ which is still corrupted by the observation noise $\boldsymbol{\mathcal{E}}$. This latter is in general much fainter than the stellar halo and can also be reduced by usual filtering methods. Based on the previous data description (3), estimating $\hat{\boldsymbol{\mathcal{S}}}$ comes to estimate $\hat{\boldsymbol{s}}$ the star's degraded spectrum $\boldsymbol{s}$ and $\hat{\boldsymbol{\alpha}}_{xy}$ the instrument CSF $\boldsymbol{\alpha}_{xy}$, allowing spaxel by spaxel halo estimation and avoiding any spatial model of the FSF:

$$\hat{\boldsymbol{s}}_{xy} = \hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{\alpha}}_{xy}\,. \tag{6}$$

On the one hand, flux integration of the entire field on each wavelength is a faithful estimate $\hat{\boldsymbol{s}}$ of $\boldsymbol{s}$ according to (2). It is assumed here that the total flux of the planets remains low compared to that of the star. However, an arbitrary selection of the noisiest spaxels is discarded beforehand and a robust estimate is computed based on their median for each wavelength. On the other hand, $\hat{\boldsymbol{\alpha}}_{xy}$ is the key coefficient of the method to be learned for each spaxel. This is where we want to focus our efforts to improve the current method.

A. State-of-the-art

PDS 70 b and PDS 70 c were discovered subtracting a spaxel-by-spaxel estimation of the star [9]. Following Eq. (6), the method use a "reference spectrum" $\boldsymbol{r}$ (obtained by spatial averaging of each monochromatic image) multiplied by an "alpha coefficient":

$$\hat{\boldsymbol{\alpha}}_{xy} = \mathrm{SG}_{d,w}\left(\boldsymbol{d}_{xy}./\boldsymbol{r}\right), \tag{7}$$

with $\mathrm{SG}_{d,w}$ the Savitzky-Golay filter of degree d and spectral window size w, $./$ the Hadamard (entrywise) division and $\boldsymbol{d}_{xy}$ the data spaxel.

However, this method precludes any physical interpretations from estimated spaxels because of the strong distortion of estimated spaxels of planets. It intrinsically removes continua of spaxels and it notably oversubtracts stellar flux around the Hα line depending on the relative line-continuum flux ratios of the star and the planets (because of the restricted spectral localization of the Savitzky-Golay window). Moreover, from a statistical perspective, a significant limitation of this approach stems from the non-linear division operation which is prone to numerical instabilities. A small variation in the denominator (i.e., the estimated stellar spectrum) can result in significant changes, leading to unstable estimates that overfit the observed data.

B. Linear parametric estimation of the CSF

Assuming smoothness of the CSF, we propose to model it at each position (x, y) by a polynomial function of the wavelengths λ_i:

$$\boldsymbol{\alpha}_{xy} = \left(\sum_{k=0}^{d} \lambda_i^k \beta_{xy}^{(k)}\right)_{i\in[\ell]}. \tag{8}$$

Following (6), this amounts to modulating the reference stellar spectrum $\hat{\boldsymbol{s}}$ by a polynomial of degree d :

$$\boldsymbol{s}_{xy} = \left(\sum_{k=0}^{d} \lambda_i^k \hat{s}_i \beta_{xy}^{(k)}\right)_{i\in[\ell]}. \tag{9}$$

The degree, or *order*, d of the polynomial is a hyperparameter controlling the complexity of the model (see section III-C for more details).

With this parametric model, the estimation of the stellar contribution is reduced to the estimation of the polynomial coefficients $\beta_{xy}^{(k)}$. This expresses as a linear regression model:

$$\boldsymbol{d}_{xy} = \boldsymbol{X}\boldsymbol{\beta}_{xy} + \boldsymbol{\varepsilon}_{xy}, \tag{10}$$

where the features corresponding to the column vectors of the design matrix $\boldsymbol{X} \in \mathbb{R}^{\ell\times d+1}$ is defined as $\boldsymbol{x}_k = \boldsymbol{\lambda}^{\cdot k} \cdot \hat{\boldsymbol{s}} \in \mathbb{R}^{\ell}$ and $\boldsymbol{\lambda}^{\cdot k}$ denotes the vector of wavelengths raised to the power of k, for $k = 0, \ldots, d$. Here $\boldsymbol{\beta}_{xy} \in \mathbb{R}^{d+1}$ is the vector of the unknown regression coefficients. Its estimation can be directly obtained by ordinary least squares (OLS) method:

$$\hat{\boldsymbol{\beta}}_{xy} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} ||\boldsymbol{d}_{xy} - \boldsymbol{X}\boldsymbol{\beta}||_2^2 = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{d}_{xy}\,. \tag{11}$$

However, this polynomial regression problem with equally spaced wavelengths is ill-conditioned as the degree d of the polynomial increases. In fact, the Gram matrix $\boldsymbol{X}^T\boldsymbol{X}$ becomes ill-conditioned, even for moderate or low degree[2]. As a result, the OLS estimator (11) is calculated with considerable error and is unreliable. Then, the use of Legendre orthogonal polynomials allows us to mitigate the ill-conditioning problem by cancelling multi-collinearity between different polynomial terms. The orthogonality also allows the coefficients of each term to be estimated independently, simplifying the interpretation of the model. This results in a change of polynomial basis and the new features read now, for $k = 0, \ldots, d$:

$$\boldsymbol{x}_k = \boldsymbol{\pi}_k(\boldsymbol{\lambda}) \cdot \hat{\boldsymbol{s}},$$

with $\boldsymbol{\pi}_k(\boldsymbol{\lambda}) \in \mathbb{R}^{\ell}$ a vector where each entry corresponds to the value of the kth (shifted) Legendre polynomial π_k [10] evaluated at the respective wavelength in $\boldsymbol{\lambda}$. Finally, these features are centered, except for $\boldsymbol{x}_0$, and L_1-normalized:

$$\boldsymbol{x}'_0 = \frac{\boldsymbol{x}_0}{||\boldsymbol{x}_0||_1} \quad \text{and} \quad \boldsymbol{x}'_k = \frac{\boldsymbol{x}_k - \overline{\boldsymbol{x}_k}}{||\boldsymbol{x}_k||_1}, \quad \forall k \in [d]\,. \tag{12}$$

This makes it possible to construct a *quasi*-orthogonal design matrix $\boldsymbol{X}$ for estimating polynomial coefficients in a stable way[3] from (11) in the unconstrained case. An additional advantage of this parameterization is that the coefficients have easier interpretations: the first coefficient $\beta_{xy}^{(0)}$ represents the total flux of the source, $\beta_{xy}^{(1)}$ controls the linear modulation of its spectrum and the other coefficients control higher-frequency modulations. This allows to possibly enforce box

[2]For equally spaced regression variables in the $[0, 1]$ range, the Gram matrix for the polynomial regression problem becomes equivalent to an Hilbert matrix which is known to be very ill-conditioned.

[3]This is (asymptotically with ℓ) an orthogonal basis when the reference spectrum $\hat{\boldsymbol{s}}$ is a non-zero constant.

constraints on these parameters in the least-squares fit to regularize the solution, using standard constrained optimization algorithm as in [11]. For instance one can ensure that the flux of the estimated stellar halo spatially decreases with distance from the star ($0 \leq \beta_{xy}^{(0)} \leq t_{xy}$ for a spatially dependent decreasing threshold t_{xy}), or that the linear trend in the modulation decreases with wavelengths ($\beta_{xy}^{(1)} \leq 0$) as physically expected.

Hereafter, we assume for simplicity that the coefficients have been estimated through (unconstrained) OLS as given in (11), but using the Legendre based design matrix. Then the stellar halo estimate reads:

$$\hat{\boldsymbol{s}}_{xy} = \boldsymbol{\Pi}_{\boldsymbol{X}} \boldsymbol{d}_{xy} \,, \tag{13}$$

where $\boldsymbol{\Pi}_{\boldsymbol{X}} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$ is the orthogonal projection matrix onto the image of $\boldsymbol{X}$.

C. Orthogonal decomposition of the estimation

Leveraging the observation model (4) and the stellar halo estimate (13), we can derive the orthogonal projection for each component of the data:

$$\hat{\boldsymbol{s}}_{xy} = \boldsymbol{\Pi}_{\boldsymbol{X}} \boldsymbol{s}_{xy} + \boldsymbol{\Pi}_{\boldsymbol{X}} \boldsymbol{p}_{xy} + \boldsymbol{\Pi}_{\boldsymbol{X}} \boldsymbol{e}_{xy} \,. \tag{14}$$

This shows the contribution of each component to halo estimation. First, part of the halo which is orthogonal to the modulated reference spectrum cannot be estimated. Last, noise or planet components may also contribute to the estimate of the star's spaxel if they can be overfitted by the modulation model. One benefit of the proposed approach over state-of-the-art methods is our ability to interpret and quantify the various sources of estimation error. By subtraction (5), we obtain now the decomposition of the estimate of the planets' spaxels:

$$\hat{\boldsymbol{p}}_{xy} = \boldsymbol{\Pi}_{\boldsymbol{X}}^{\perp} \boldsymbol{s}_{xy} + \boldsymbol{\Pi}_{\boldsymbol{X}}^{\perp} \boldsymbol{p}_{xy} + \boldsymbol{\Pi}_{\boldsymbol{X}}^{\perp} \boldsymbol{e}_{xy} \,, \tag{15}$$

where the matrix $\boldsymbol{\Pi}_{\boldsymbol{X}}^{\perp} = \mathbf{I}_{\ell} - \boldsymbol{\Pi}_{\boldsymbol{X}}$ is the projection onto the orthogonal complement to the image of $\boldsymbol{X}$.

These results are used in section IV for tuning the polynomial degree d and evaluating the estimation performances on simulated datasets.

IV. Results

We consider the PDS 70 system observed with the MUSE instrument. The set of wavelengths is $\Lambda = \{\lambda_0 + i\lambda_{\text{step}}\}_{i\in[\ell]}$ with $\lambda_0 \approx 4749.71$Å and $\lambda_{\text{step}} = 1.25$Å according to data headers. While data cubes have more than 300 rows and 300 columns, we zoom on a smaller cube with $n = 40$ and $m = 40$ (1" in the sky) centered around the PDS 70 star where the sought planets PDS 70 b and c are supposed to be.

For the generation of synthetic cubes, we consider first simple cases where Hα contrasts (i.e., ratio of planet's spectrum flux to that of the star at Hα line wavelengths) are 100 times bigger than the estimated real ones (1.7×10^{-3} and 8.3×10^{-4} respectively for PDS 70 b and PDS 70 c, from state-of-the-art results). The simulated FSFs are Moffat functions, and their Full Width at Half Maximum (FWHM) decreases inversely linearly with wavelength. The noise $\boldsymbol{\mathcal{E}}$ is finally generated with i.i.d. $\mathcal{N}(0,\sigma^2)$ entries (with values according to previous noise estimations [9]).

A. Hyperparameter tuning for the polynomial order d

The polynomial order d controls the bias/variance trade-off of the star and planets estimation. For small d, the model is simpler and may underfit the stellar spectrum. Conversely, for large d, the model is more flexible and may overfit the potential planets. For the usual squared loss, the optimal value for d is obtained by minimizing the mean squared error (MSE):

$$\text{MSE} = \mathbb{E}\left(||\hat{\boldsymbol{p}}_{xy} - \boldsymbol{p}_{xy}||^2\right). \tag{16}$$

Straightforward computations yield:

$$\text{MSE} = ||\boldsymbol{\Pi}_{\boldsymbol{X}}\boldsymbol{p}_{xy}||^2 + ||\boldsymbol{\Pi}_{\boldsymbol{X}}^{\perp}\boldsymbol{s}_{xy}||^2 + \mathbb{E}(||\boldsymbol{\Pi}_{\boldsymbol{X}}^{\perp}\boldsymbol{e}_{xy}||^2)\,, \tag{17}$$

where $\mathbb{E}(||\boldsymbol{\Pi}_{\boldsymbol{X}}^{\perp}\boldsymbol{e}_{xy}||^2) = (\ell - (d+1))\sigma^2$.

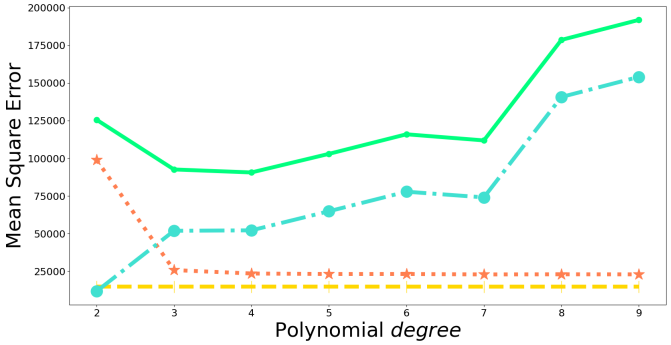

Fig. 3. Mean Squared Error decomposition as given in (17) at Hα line spectral channel. The MSE (solid light green line) is the sum of the overfitting error of the planets $||\boldsymbol{\Pi}_{\boldsymbol{X}}\boldsymbol{p}_{xy}||^2$ (dash-dotted turquoise line), of the underfitting error of the star $||\boldsymbol{\Pi}_{\boldsymbol{X}}^{\perp}\boldsymbol{s}_{xy}||^2$ (dotted light orange line) and of the underfitting error of the noise $\mathbb{E}\left(||\boldsymbol{\Pi}_{\boldsymbol{X}}^{\perp}\boldsymbol{e}_{xy}||^2\right)$ (dashed yellow line).

Fig. 3 shows the evolution of the MSE as a function of d. On this simulation, a polynomial degree $d = 4$ is the best choice for Hα line recovery. Note however that this optimal value depend on the FSF/CSF of the instrument. Their true behavior is likely more complicated than what we achieved in our simulation, using a model of FSF with FWHM decreasing inversely with the wavelength.

B. Spectral simulation results

The brightness of the planets in our simulation highlights the self-subtraction effect at the wavelengths of the Hα line (see Fig. 4) for the state-of-the-art approach described in section III-A. The filter parameters (linear fits on a sliding window of 101 MUSE spectral bins) are fixed as in [9]. Figure 5 illustrates that our method can efficiently correct the self-subtraction with $d = 4$ as suggested in figure 3. This side effect is corrected by the limited complexity of the polynomial model compared to state-of-the-art filter. Moreover, while continuum (low-frequency spectrum component, in infrared for planets) is completely removed by this low-pass filter, our method keeps part of it. It is distorted and therefore unusable for

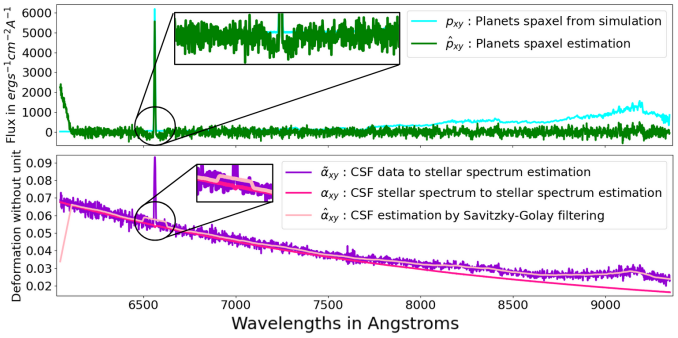

Fig. 4. Simulated PDS 70 b spaxel subtraction results with the **state-of-the-art method**. Top panel compares the estimated planets spaxel with the ground truth. Bottom panel compares the estimated CSF with the deformation relative to the stellar spaxel and that relative to the data spaxel. The self-subtraction effect and its origin in the CSF estimate are circled and zoomed.

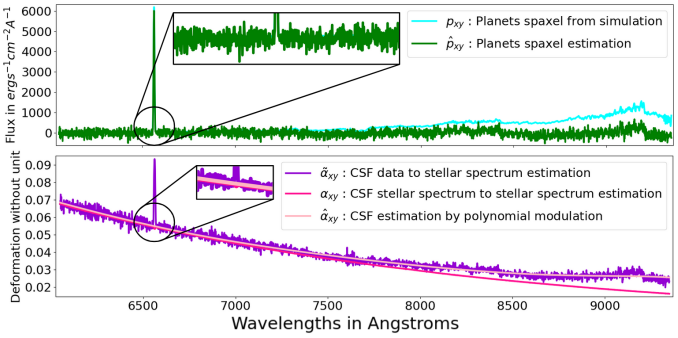

Fig. 5. Simulated PDS 70 b planet spaxel subtraction results with the **proposed method**. The legend is similar to Fig. 4. Correction of the self-subtraction effect and correction of its origin are clearly visible in the same zooms (respectively in green and pink).

physical interpretations but it is still additional information for detection. The continuum is not completely preserved because the stellar halo estimate (13) overfits the planets contribution.

C. Real data detections

We used four on-sky observations of PDS 70 presented in [9] to verify the accuracy of our method. We merge them by median to reduce electronic noise.

Fig. 3 shows MSE results for a simplistic simulation (with Moffat FWHM decreasing exactly inversely with wavelength) to highlight the main effects of the subtraction. With a slightly more realistic simulation (with Moffat FWHM not decreasing exactly inversely with wavelength as built by multiple Moffat fits on real PDS70 monochromatic images), the optimum degree moves to $d = 7$ allowing a little more flexibility to fit instrumental imperfections.

As with the state-of-the-art method, the two planets PDS 70 b and PDS 70 c are well detected, but stellar flux residues are less dispersed with our new method. After applying spectral matched-filtering with an Hα line model and spatial matched-filtering with an FSF model, the signal due to the two planets is magnified. The difference in robustness between the two methods becomes even more pronounced. As illustrated in Figure 6, we can see that the number of false alarms to detect the two planets is considerably reduced (respectively 108/13 false alarms for old/new method with a threshold of 0.033, left-hand PDS 70 c detection joining the false alarms below).

V. Conclusion

Our proposed method demonstrates better subtraction results in both simulated scenarios and our available data. It exhibits increased robustness by mitigating self-subtraction in the Hα line. Part of the continuum is also preserved. Furthermore, it paves the way to many more improvements for the future of planets detection using spectral diversity. Spectral and spatial regularisation would help to obtain a better estimation, better stellar spectrum estimation would improve results on particular at Hα line wavelengths, and a recursive process should push the exploitation of this data even further.

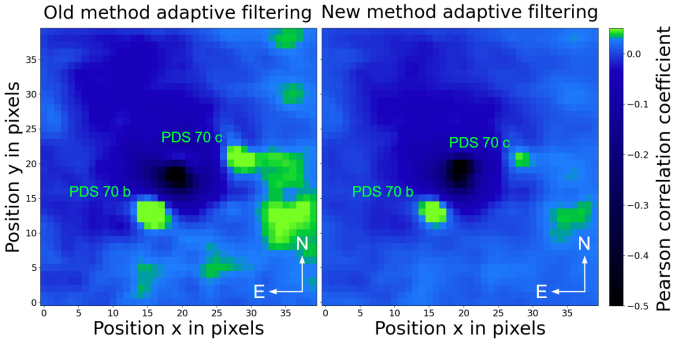

Fig. 6. Spatial adaptive filtering of the post stellar halo subtraction residuals for real MUSE datacube of PDS 70 system. The matching template cube is obtained from a Hα line model combined with a spatial model of the FSF.

References

[1] Y. Aoyama et al., "Theoretical Model of Hydrogen Line Emission from Accreting Gas Giants," *The astrophysic journal*, vol. 866, 2018.

[2] T. Uyama et al., *The Astronomic Journal*, vol. 154, 2017.

[3] C. Xie et al., "Searching for proto-planets with muse," *Astronomy & Astrophysics*, vol. 644, 2020.

[4] S. C. Ringqvist et al., "Resolved near-uv hydrogen emission lines at 40-myr super-jovian protoplanet delorme 1 (ab)b. indications of magnetospheric accretion," *Astronomy & Astrophysics*, vol. 669, 2023.

[5] D. Demars et al., "Emission line variability of young 10-30 M_{Jup} companions. I. The case of GQ Lup b and GSC 06214-00210 b," *Astronomy & Astrophysics*, vol. 676, 2023.

[6] R. Bacon et al., "The MUSE second-generation VLT instrument," in *Proceedings of the SPIE*, vol. 7735, 2010.

[7] R. Arsenault et al., "Eso adaptive optics facility," in *Adaptive Optics Systems*, vol. 7015, 2008.

[8] S. Ströbele et al., "GALACSI system design and analysis," in *Adaptive Optics Systems III*, vol. 8447, 2012.

[9] S. Y. Haffert et al., "Two accreting proto-planets around the young star pds 70," *Nature Astronomy*, 2019.

[10] G. Arfken and H. Weber, *Mathematical Methods For Physicists*, six ed. Elsevier Science, 2005, ch. 12.

[11] P. B. Stark and R. L. Parker, "Bounded-variable least-squares: an algorithm and applications," *Computational Statistics*, vol. 10, 1995.