Separation of galaxy spectra measured with slitless spectroscopy

Shahram Hosseini*, Ahmed Selloum, Thierry Contini, Yannick Deville

Université de Toulouse, UPS, CNRS, CNES,
IRAP (Institut de Recherche en Astrophysique et Planétologie),
14 Av. Edouard Belin, 31400 Toulouse, France

Abstract

This paper investigates the problem of separating galaxy spectra resulting from the slitless spectroscopy system, which can e.g. be used in spatial missions like the WISP survey and the Euclid project. We first derive a physical mixing model linking observed data to source spectra, then simplify it to obtain two different approximate but realistic models. The first simplified model being entirely defined by a few parameters, we propose a semi-blind source separation method which estimates these model parameters together with the source spectra. The second simplified model is linear instantaneous, and has a special form which is used to propose a new blind source separation method exploiting the nonnegativity and spatial sparsity of data, as well as the correlation of the target source spectra in different light dispersion directions. Both methods are tested on realistic simulated data and lead to very encouraging results, confirming their effectiveness.

Keywords: Source separation; Spectrum decontamination; Astronomy; Slitless spectroscopy; Euclid mission.

1. Introduction

Source separation aims at recovering a set of unknown source signals from the observed mixtures of them. Blind Source Separation (BSS), where little information about the mixing system is available, has largely been considered in several application domains [1, 2]. In spectroscopy, the BSS methods have e.g. been used to separate source spectra in Earth observation [3, 4], chemistry [5, 6], nuclear magnetic resonance [7, 8] and astronomy [9, 10, 11].

Here, we address a new application concerning the separation of spectra measured with slitless spectroscopy. As will be explained in the following sections, the angular diversity used in the observation strategy leads to a new and specific mixing model which cannot be handled by the existing methods, and requires the development of new source separation methods presented in this paper. The proposed methods may be used in spatial missions, like in the WFC3 Infrared Spectroscopic Parallel (WISP) survey [12, 13] which is a large Hubble Space Telescope (HST) program or in the future Eulcid spatial mission [14], but they can also be used in other slitless spectroscopy applications.

A typical application is the spectra decontamination required in the Euclid project, which is a space mission of the European Space Agency (ESA), aiming at better understanding the nature of dark energy that is admitted to be responsible for accelerating the expansion of the Universe [15]. The satellite launch is currently planned for 2022. The Euclid near-infrared spectrograph will provide spectra of more than 50 million galaxies. The detection of the strong emission lines in these spectra will then permit to estimate the galaxy redshifts[1] due to the Universe expansion.

The spectroscopy is performed using *grisms*, which are combinations of prisms and diffraction gratings. A grism provides a so-called 2-dimensional (2D) spectrum of the light emitted by a celestial object (mainly galaxy or star) by differently refracting its different wavelengths, creating an effect like a rainbow where different wavelengths are mapped to different spatial positions in the grism dispersion direction. Spectroscopy in astronomy is usually done using a slit which only allows the diffraction of light coming from a small zone of the sky. Slitless spectroscopy, which will be used in the Euclid spectrograph, is however affected by the superposition of 2D spectra of different objects as shown in Fig. 1. This
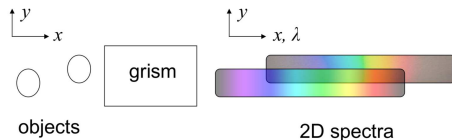

Figure 1: Superposition of the object spectra at the grism output.

contamination is the main source of error in the redshift estimation [16].

To reduce this problem, 2D spectra will be generated in several grism orientations. Thus, when the spectra of two objects are superimposed in one direction, they are usually not mixed in other directions. This observation strategy does not however totally solve the problem, because there are often other objects in the other directions which generate other mixtures. This issue may be better

---

[1]Redshift means the displacement of the object spectrum toward longer wavelengths, which may be measured by comparing the wavelength of an emission line in the observed object spectrum with the wavelength of the same emission line in a laboratory on Earth.

understood by studying the configuration shown in Fig. 2. In this example, when using a grism dispersing the light in the 0-degree direction, the spectra of objects 1 and 2 will be mixed, as shown in Fig. 2-Middle, whereas a 90-degree grism yields a mixture of the spectra of objects 1 and 3, as shown in Fig. 2-Right. Then, there are mixtures in both directions although the contaminating objects are not the same.

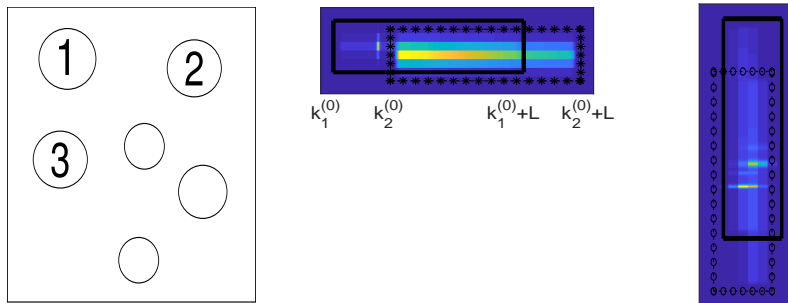Few researchers have addressed the problem of spectra decontamination in slit-

Figure 2: Left: Object spatial positions, Middle: 2D mixed spectrum dispersed in 0-degree direction by a grism, Right: mixed spectrum dispersed in 90-degree direction. Only mixed spectra containing the contribution of object 1 are shown in these figures. The boxes with solid lines in both directions contain the 2D spectrum of object 1, the boxes with asterisks and circles respectively contain the spectra of objects 2 and 3. $k_1^{(0)}$ and $k_2^{(0)}$ correspond to the first pixels of 2D spectra in 0-degree direction for objects 1 and 2, and L represents the length of spectra (in pixels).

less spectroscopy. To our knowledge, except for our conference papers [17, 18], only basic decontamination methods based on information provided by direct photometric imaging have been proposed in the literature [19, 20]. In this paper, we formulate this issue as a source separation problem. For the sake of simplicity, we only consider spectra generated in two directions (0 and 90 degrees). However, other directions may be considered without fundamental modifications in the proposed approaches, provided they have been chosen such that there are different contaminating objects in different directions.

In this application, source separation can be applied in two different ways:

1. The global approach where the observation matrix contains all the observed values in the field of view and the source matrix contains the spectra of all the objects in this field. In this case, the mixing matrix will be huge and the computation time will probably be so long that the method will be unrealizable. This approach is not retained and is not presented in the paper.

2. The local approach, used in this article, where the source separation will be applied object by object, and probably only for a certain number of objects of interest related to the objectives of the mission. In this case, the observation matrix contains the pixels related to this target object and its

3

contaminants, while the source vector contains the spectra of the object of interest and its contaminants. The main advantage of the local approach is that it is highly parallelizable so that a large number of processors may simultaneously execute a large number of local source separation algorithms.

In Section 2, we present a physical model describing the relationship between observations and sources. Two simplified mathematical models based on realistic assumptions are proposed in Section 3. In Section 4, a semi-blind source separation method exploiting available information on the instrument Point Spread Function (PSF) and on the object light intensity profiles is proposed. In Section 5, we propose a blind method which does not need this information. Simulation results are presented in Section 6 before a conclusion in Section 7.

2. Models

2.1. Object model

The telescope provides a signal corresponding to astronomical objects (mainly galaxies) and the sky background. The contribution of the sky background can be estimated and subtracted from the observed signal [21, 22]. A given object with index i may be characterized by[2]: its spectrum $s_i(\lambda)$, its center position (x_i, y_i), and its spatial light intensity profile $f_i(x - x_i, y - y_i)$.

Then, the light intensity of the object i, at a position with coordinates (x, y) and at a wavelength λ, can be written as

$$q_i(x, y, \lambda) = s_i(\lambda) f_i(x - x_i, y - y_i). \tag{1}$$

2.2. PSF model

Because of the instrument PSF, denoted as h, the object is spatially spread. The light intensity of this "new spread object" is defined by

$$
\begin{aligned}
w_i(x, y, \lambda) &= \int \int_{R^2} q_i(u, v, \lambda) h(x - u, y - v) du dv \\
&= s_i(\lambda)[f_i(x - x_i, y - y_i) * h(x, y)] = s_i(\lambda) I_i(x - x_i, y - y_i), \tag{2}
\end{aligned}
$$

where I_i represents the convolution of f_i and h.

2.3. Grism model

The grism disperses the observed image and generates 2D spectra of different orders [23]. Whereas only the first-order spectrum is considered in this paper, the other orders may also be taken into account in a similar manner.

The extension of the spectrum in the focal plane of detectors is called the *trace* and may be curved [20]. When the curvature is negligible, the trace can

---

[2]We assume hereafter that all object points have the same spectrum up to scale factors.

4

be modeled by a rectangular strip in the dispersion direction. In this paper, we assume that two grisms with the dispersion directions 0 and 90 degrees are used in the spectrograph. The 0-degree grism corresponds to a horizontal trace, in the x direction, while the 90-degree grism is associated with a vertical trace in the y direction. For an object with index i, we get at the 0-degree grism output a 2D image which reads [24, 25]

$$
\begin{aligned}
t_i^{(0)}(x, y) &= \int_{\lambda \in \Omega_\lambda} w_i\left(x - D(\lambda), y, \lambda\right) d\lambda \\
&= \int_{\lambda \in \Omega_\lambda} I_i(x - x_i - D(\lambda), y - y_i)s_i(\lambda)d\lambda,
\end{aligned}
\tag{3}
$$

where Ω_λ represents the wavelength range covered by the grism and $D(\lambda)$ is the dispersion function in the x direction, defined as the mapping of the different wavelengths to the different spatial positions on the trace. Ideally, the dispersion function should be a linear function of the wavelength. For example, for a 0-degree grism and a point-like object, if $D(\lambda) = b\lambda + c$, a given wavelength λ_1 in the spectrum is mapped to a spatial position whose abscissa is $b\lambda_1 + c + x_i$.

Similarly, we obtain at the 90-degree grism output a 2D image defined by

$$
t_i^{(90)}(x, y) = \int_{\lambda \in \Omega_\lambda} I_i(x - x_i, y - y_i - D(\lambda))s_i(\lambda)d\lambda.
\tag{4}
$$

2.4. Integration in detector pixels

The above equations are established for real (non-integer) values of spatial coordinates x and y. In practice, the signal measured in a detector pixel, for a given object i and using the 0-degree grism, is the integral of the above-mentioned $t_i^{(0)}(x, y)$ over the surface Ω_p of that pixel. The measured value for a pixel with index p is then

$$
\begin{aligned}
o_i^{(0)}(p) &= \iint_{(x,y) \in \Omega_p} t_i^{(0)}(x, y)dxdy \\
&= \iint_{(x,y) \in \Omega_p} \int_{\lambda \in \Omega_\lambda} I_i(x - x_i - D(\lambda), y - y_i)s_i(\lambda)d\lambda dxdy.
\end{aligned}
\tag{5}
$$

A similar equation may be derived for the 90-degree direction.

2.5. Contamination model

The number of contaminants for each target object can be determined from direct (i.e. non-dispersed) images of the sky. This can easily be done from the object positions and the instrument parameters. In fact, for each object, we can compute the coordinates of the box containing its 2D spectrum. If the box of the target object overlaps with those of other objects, the latter objects are

5

considered as the contaminants of the target object (see Fig. 2).

We assume that the object i is contaminated in the 0-degree (horizontal) direction by F_0 other objects with indices belonging to the set $\xi_0 = \{h_1, ..., h_{F_0}\}$, $i.e.$ a pixel p of the detector receives photons from the object i and F_0 other objects, because of the mixing effect, explained in Section 1. The measured value in the pixel p of the detector then reads

$$
\begin{aligned}
o^{(0)}(p) \;\; &= \sum_{j \in \{i\} \cup \xi_0} o_j^{(0)}(p) \\
&= \sum_{j \in \{i\} \cup \xi_0} \iint_{(x,y) \in \Omega_p} \int_{\lambda \in \Omega_\lambda} I_j(x - x_j - D(\lambda), y - y_j)s_j(\lambda)d\lambda dx dy. \quad (6)
\end{aligned}
$$

Similarly, when using the 90-degree grism, we assume that the object i is contaminated in the 90-degree (vertical) direction by F_{90} objects with indices belonging to the set $\xi_{90} = \{v_1, ..., v_{F_{90}}\}$. The measured value in the pixel p of the detector reads

$$
o^{(90)}(p) = \sum_{j \in \{i\} \cup \xi_{90}} \iint_{(x,y) \in \Omega_p} \int_{\lambda \in \Omega_\lambda} I_j(x - x_j, y - y_j - D(\lambda))s_j(\lambda)d\lambda dx dy. \quad (7)
$$

Considering the mathematical mixing model described by (6) and (7), we aim at decontaminating the spectrum of the object i, considered as the object of interest.

3. Simplified models

The physical model presented in the previous section being too complex, we here propose two simplified models which will be used respectively in our semi-blind method presented in Section 4 and in our blind method presented in Section 5.

3.1. Simplified model 1: a parametric model

We begin by discretizing the integrals in (6) and (7) using the rectangle method which yields in the 0-degree direction:

$$
o^{(0)}(p) = \sum_{j \in \{i\} \cup \xi_0} \sum_{(x_k, y_r) \in \Omega_p} \sum_{\lambda_l \in \Omega_\lambda} I_j(x_k - x_j - D(\lambda_l), y_r - y_j)s_j(\lambda_l)\Delta\lambda\Delta x\Delta y,
$$

$$
(8)
$$

where Δx, Δy and $\Delta\lambda$ are the chosen spatial and spectral discretization steps. Considering L wavelengths in Ω_λ, Eq. (8) and its 90-degree counterpart can be

rewritten as

$$o^{(0)}(p) = \sum_{j \in \{i\} \cup \xi_0} \sum_{l=1}^{L} m_j^{(0)}(p, \lambda_l) s_j(\lambda_l)$$

$$\text{with } m_j^{(0)}(p, \lambda_l) = \sum_{(x_k, y_r) \in \Omega_p} I_j(x_k - x_j - D(\lambda_l), y_r - y_j) \Delta\lambda \Delta x \Delta y,$$

$$o^{(90)}(p) = \sum_{j \in \{i\} \cup \xi_{90}} \sum_{l=1}^{L} m_j^{(90)}(p, \lambda_l) s_j(\lambda_l)$$

$$\text{with } m_j^{(90)}(p, \lambda_l) = \sum_{(x_k, y_r) \in \Omega_p} I_j(x_k - x_j, y_r - y_j - D(\lambda_l)) \Delta\lambda \Delta x \Delta y.$$

$$(9)$$

To derive a parametric model for functions I_j, defined in Eq. (2), we here assume that:

- The instrument PSF, $h(x, y)$, can be modeled by a linear combination of two circular 2-dimensional Gaussian functions. According to [21], this model, which only depends on three parameters (σ_1, σ_2 and c), is sufficiently flexible to provide an acceptable approximation of the PSF:

$$h(x, y) = c\mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_1^2 \end{bmatrix} \right) + (1 - c)\mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_2^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \right),$$

$$(10)$$

  where $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ represents a 2D Gaussian function with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$.

- The spatial light intensity profile of an object j can be modeled by a 2-dimensional Gaussian function:

$$f_j(x - x_j, y - y_j) = \mathcal{N}\left(\begin{bmatrix} x_j \\ y_j \end{bmatrix}, \begin{bmatrix} \sigma_{a_j}^2 & r_j \sigma_{a_j} \sigma_{b_j} \\ r_j \sigma_{a_j} \sigma_{b_j} & \sigma_{b_j}^2 \end{bmatrix} \right). \quad (11)$$

  The three parameters of this Gaussian function (σ_{a_j}, σ_{b_j}, r_j) depend on the object shape when modeling this shape by an ellipse, characterized by its major axis, minor axis and angle. Although there are more accurate models, like the Sérsic profile [26], the Gaussian model is very simple and can be used as an approximation of the actual profile[3].

According to (2), I_j is the convolution between h (the instrument PSF) and f_j (the spatial light profile of the object j). Since the convolution of two Gaussian

---

[3]We measured the quality of the Gaussian fit for 10 synthetic galaxy images (within a radius of 8 pixels around the galaxy center). The normalized mean square error was between 0.0073 and 0.0497 with a mean of 0.0325 and a standard deviation of 0.0182.

functions is a Gaussian function whose mean and covariance matrix are respectively the sum of means and the sum of covariance matrices of the two original Gaussian functions, we can write

$$
I_j(x - x_j, y - y_j) = c\mathcal{N}\left(\begin{bmatrix} x_j \\ y_j \end{bmatrix}, \begin{bmatrix} \sigma_{a_j}^2 + \sigma_1^2 & r_j\sigma_{a_j}\sigma_{b_j} \\ r_j\sigma_{a_j}\sigma_{b_j} & \sigma_{b_j}^2 + \sigma_1^2 \end{bmatrix}\right)
$$
$$
+ (1 - c)\mathcal{N}\left(\begin{bmatrix} x_j \\ y_j \end{bmatrix}, \begin{bmatrix} \sigma_{a_j}^2 + \sigma_2^2 & r_j\sigma_{a_j}\sigma_{b_j} \\ r_j\sigma_{a_j}\sigma_{b_j} & \sigma_{b_j}^2 + \sigma_2^2 \end{bmatrix}\right), \quad (12)
$$

which is a parametric model[4]. If the parameters of the model (12) are known, the mixing coefficients $m_j^{(0)}(p, \lambda_l)$ and $m_j^{(90)}(p, \lambda_l)$ in Eq. (9) can be computed from them and from the known dispersion function $D(\lambda)$, otherwise these mixing coefficients depend on these unknown parameters.

Now we collect $\{o^{(0)}(p)\}_{p=1...P_0}$, observed values in the 0-degree direction, and $\{o^{(90)}(p)\}_{p=1...P_{90}}$, observed values in the 90-degree direction, to define an observation vector $\mathbf{o} = \left[o^{(0)}(1), ..., o^{(0)}(P_0), o^{(90)}(1), ..., o^{(90)}(P_{90})\right]^T$. The mixing model (9) can then be written in the following matrix form:

$$
\mathbf{o} = \mathbf{Ms}, \quad (13)
$$

where

- The global vector of source spectra $\mathbf{s}$ is defined as

$$
\mathbf{s} = [\mathbf{s}_i^T, \mathbf{s}_{h_1}^T, ..., \mathbf{s}_{h_{F_0}}^T, \mathbf{s}_{v_1}^T, ..., \mathbf{s}_{v_{F_{90}}}^T]^T, \quad (14)
$$

  with $\mathbf{s}_j = [s_j(\lambda_1), s_j(\lambda_2), ..., s_j(\lambda_L)]^T$. Here, $\mathbf{s}_i$ represents the spectrum of the target source, $\mathbf{s}_{h_1}, ..., \mathbf{s}_{h_{F_0}}$ are the spectra of its contaminant objects in the 0-degree direction, and $\mathbf{s}_{v_1}, ..., \mathbf{s}_{v_{F_{90}}}$ are the spectra of its contaminants in the 90-degree direction.

- The mixing matrix $\mathbf{M}$ is defined as

$$
\mathbf{M} = \begin{bmatrix} \mathbf{M}_i^{(0)} & \mathbf{M}_{h_1}^{(0)} & ... & \mathbf{M}_{h_{F_0}}^{(0)} & \mathbf{0} & ... & \mathbf{0} \\ \mathbf{M}_i^{(90)} & \mathbf{0} & ... & \mathbf{0} & \mathbf{M}_{v_1}^{(90)} & ... & \mathbf{M}_{v_{F_{90}}}^{(90)} \end{bmatrix}, \quad (15)
$$

  where the entry (p, l) of each matrix $\mathbf{M}_j^{(0)}$ of size $P_0 \times L$ (resp. matrix $\mathbf{M}_j^{(90)}$ of size $P_{90} \times L$) is equal to $m_j^{(0)}(p, \lambda_l)$ (resp. $m_j^{(90)}(p, \lambda_l)$) defined in (9).

---

[4]It is also possible to consider $\sigma_{a_j}^2 + \sigma_1^2$, $r_j\sigma_{a_j}\sigma_{b_j}$, $\sigma_{b_j}^2 + \sigma_1^2$, $\sigma_{a_j}^2 + \sigma_2^2$ and $\sigma_{b_j}^2 + \sigma_2^2$ as new parameters.

3.2. Simplified model 2: a linear instantaneous model

This time, we make no assumption about the PSF and object light profile models. Instead, to simplify the observation model, described by (6) and (7), we use the following separability assumption:

$$I_i(x,y) = I_{i1}(x)I_{i2}(y). \tag{16}$$

This assumption is realistic for small objects because the PSF is nearly circular. For extended objects, if the object is circular, or elliptical and oriented in the x or y directions, the assumption is still valid. Otherwise, it involves some approximation error, but simulations show that this error is negligible in comparison with other errors, especially the high level of noise (see Appendix A for more details). Using this approximation, (6) becomes for a pixel with spatial index $p = (n,m)$:

$$o^{(0)}(n,m) = \sum_{j\in\{i\}\cup\xi_0} a_j^{(0)}(m)e_j^{(0)}(n), \tag{17}$$

where

$$a_j^{(0)}(m) = \int_{y|(x,y)\in\Omega_{\mathbf{P}}} I_{j2}(y - y_j)dy, \tag{18}$$

and

$$e_j^{(0)}(n) = \int_{x|(x,y)\in\Omega_{\mathbf{P}}} \int_{\lambda\in\Omega_\lambda} I_{j1}(x - x_j - D(\lambda))s_j(\lambda)d\lambda dx. \tag{19}$$

Using this approximation, the coefficient $a_j^{(0)}(m)$ does not depend on the x-coordinate of the pixel (*i.e. n*): the integral defining $a_j^{(0)}(m)$ has the same value for all detector pixels located on a horizontal line. In a similar way, $e_j^{(0)}(n)$ does not depend on m.

The same approach applied to the mixing equation (7) leads to the following approximation:

$$o^{(90)}(n,m) = \sum_{j\in\{i\}\cup\xi_{90}} a_j^{(90)}(n)e_j^{(90)}(m), \tag{20}$$

where

$$a_j^{(90)}(n) = \int_{x|(x,y)\in\Omega_{\mathbf{P}}} I_{j1}(x - x_j)dx, \tag{21}$$

and

$$e_j^{(90)}(m) = \int_{y|(x,y)\in\Omega_{\mathbf{P}}} \int_{\lambda\in\Omega_\lambda} I_{j2}(y - y_j - D(\lambda))s_j(\lambda)d\lambda dy. \tag{22}$$

We can collect all the measured values in P pixels situated in a rectangular $N \times M$ zone, receiving radiations from the object i and F_0 contaminating objects in the 0-degree direction, to form an $M \times N$ matrix $\mathbf{X}^{(0)}$ whose entry (m,n) is $o^{(0)}(n,m)$. Then, we can write the mixing equation (17) in the following matrix form:

$$\mathbf{X}^{(0)} = \mathbf{A}^{(0)}\mathbf{E}^{(0)}, \tag{23}$$

9

where

$$\mathbf{A}^{(0)} = \begin{bmatrix} a_i^{(0)}(1) & a_{h_1}^{(0)}(1) & ... & a_{h_{F_0}}^{(0)}(1) \\ \vdots & \vdots & \vdots & \vdots \\ a_i^{(0)}(M) & a_{h_1}^{(0)}(M) & ... & a_{h_{F_0}}^{(0)}(M) \end{bmatrix},$$

$$\mathbf{E}^{(0)} = \begin{bmatrix} e_i^{(0)}(1) & ... & e_i^{(0)}(N) \\ e_{h_1}^{(0)}(1) & ... & e_{h_1}^{(0)}(N) \\ \vdots & \vdots & \vdots \\ e_{h_{F_0}}^{(0)}(1) & ... & e_{h_{F_0}}^{(0)}(N) \end{bmatrix}. \tag{24}$$

Similarly, we get the following matrix model in the 90-degree direction:

$$\mathbf{X}^{(90)} = \mathbf{A}^{(90)}\mathbf{E}^{(90)}, \tag{25}$$

where the element (n, m) of $\mathbf{X}^{(90)}$ is $o^{(90)}(n, m)$, the element (n, j) of $\mathbf{A}^{(90)}$ is $a_j^{(90)}(n)$, and the element (j, m) of $\mathbf{E}^{(90)}$ is $e_j^{(90)}(m)$.

In the following two sections, we propose two source separation methods: a semi-blind method based on the parametric model presented in section 3.1, and a blind method based on the linear-instantaneous mixing model of section 3.2. We emphasize that the word *blind* is used here according to its meaning in the source separation community and refers to the fact that the mixing matrix is unknown. This does not mean that the methods are able to estimate the spectra of objects that have not been detected in photometric images of the sky: our methods need to know the number of sources in the mixing model and their approximate positions: information which must be extracted from photometric images.

4. Semi-blind method

This method is based on the simplified model 1, described in Section 3.1. Consider the vector of unknown parameters

$$\boldsymbol{\theta} = [\boldsymbol{\theta}_{PSF}, \boldsymbol{\theta}_j | j \in \{i\} \cup \xi_0 \cup \xi_{90}], \tag{26}$$

where

- $\boldsymbol{\theta}_{PSF} = [\sigma_1, \sigma_2, c]$ is the vector of the PSF parameters,

- $\boldsymbol{\theta}_j = [x_j, y_j, \sigma_{a_j}, \sigma_{b_j}, r_j]$ is the vector of the position and the spatial light profile parameters of object j.

If $\boldsymbol{\theta}$ is exactly known, and subject to the validity of the model described in (13), we can calculate the mixing matrix $\mathbf{M}$ from these parameters and the known dispersion function $D(\lambda)$, thanks to Eq. (12), (9) and (15). Then, the

problem of estimating the global vector of sources $\mathbf{s}$, defined in (14), can be formulated as the minimization of the standard squared Euclidean norm:

$$J_1 = ||\mathbf{o} - \mathbf{Ms}||_2^2, \tag{27}$$

which leads to the simple least squares estimate

$$\hat{\mathbf{s}} = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{o}. \tag{28}$$

185 The light flux, measured based on the number of photons received on the detector or the number of electrons generated by it, cannot be negative. To take into account this non-negativity of sources, we can use the non-negative least squares method presented in [27]. A simpler approach is to replace negative values in $\hat{\mathbf{s}}$ by zero.

190 In some cases, the mixing matrix $\mathbf{M}$ may be ill-conditioned, i.e. its condition number (defined as the ratio of its largest singular value to the smallest) is too large. In our application, this situation can occur if, for example, two contaminants are spatially very close and have almost the same profile. In this case, better results may be obtained by adding a regularization term to the cost
195 function J_1. Using a smoothing constraint, we can $e.g.$ minimize the criterion $||\mathbf{o} - \mathbf{Ms}||_2^2 + ||\mathbf{\Gamma s}||_2^2$, where the Tikhonov matrix $\mathbf{\Gamma}$ is the following difference matrix:

$$\mathbf{\Gamma} = \beta \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \ddots \\ 0 & -1 & 2 & -1 & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & -1 & 2 \end{bmatrix}, \tag{29}$$

and the parameter β determines the contribution of the smoothness constraint. The least squares solution then becomes: $\hat{\mathbf{s}} = (\mathbf{M}^T \mathbf{M} + \mathbf{\Gamma}^T \mathbf{\Gamma})^{-1} \mathbf{M}^T \mathbf{o}$.

Finally, the contributions of the target object i in the 2D spectra observed in both directions may be estimated by

$$\hat{\mathbf{o}}_i^{(0)} = \mathbf{M}_i^{(0)} \hat{\mathbf{s}}_i, \;\; \hat{\mathbf{o}}_i^{(90)} = \mathbf{M}_i^{(90)} \hat{\mathbf{s}}_i, \tag{30}$$

200 where $\hat{\mathbf{s}}_i$ is the estimated spectrum of the object i corresponding to the first part of the vector $\hat{\mathbf{s}}$ (see Eq. (14)), and $\mathbf{M}_i^{(0)}$ and $\mathbf{M}_i^{(90)}$ correspond to the first part of the mixing matrix $\mathbf{M}$ (see Eq. (15)).

In practice, the parameters of the model proposed in Section 3.1 are not perfectly known so that the above method may lead to unsatisfactory results.
205 Here, we present a semi-blind method which estimates the parameter vector $\boldsymbol{\theta}$ together with the source signals.

Note that even though $\boldsymbol{\theta}$ is unknown, its variation domain, defined by the extreme values $\boldsymbol{\theta}_{min}$ and $\boldsymbol{\theta}_{max}$, is supposed to be known thanks to our knowledge on the physical constraints, the instrument characteristics, and the availability
210 of direct images.

11

The idea is to minimize $J_2 = ||\mathbf{o} - \mathbf{M}(\boldsymbol{\theta})\mathbf{s}||_2^2$ under the constraint $\boldsymbol{\theta} \in [\boldsymbol{\theta}_{min}, \boldsymbol{\theta}_{max}]$, where $\mathbf{M}(\boldsymbol{\theta})$ represents the value of the mixing matrix $\mathbf{M}$ for a given value of the parameter vector $\boldsymbol{\theta}$. A simple and fast solution is to consider $\boldsymbol{\theta}$ as a master variable, whereas $\mathbf{s}$ becomes a salve variable. In each step of the loop for updating $\boldsymbol{\theta}$, the slave variable $\mathbf{s}$ is set to its optimal value, i.e. to its value which minimizes the criterion J_2 with respect to $\mathbf{s}$ for the considered value of $\boldsymbol{\theta}$. This optimum is nothing but the least squares solution, i.e.

$$\hat{\mathbf{s}}_{/\theta} = [\mathbf{M}(\boldsymbol{\theta})^T \mathbf{M}(\boldsymbol{\theta})]^{-1} \mathbf{M}(\boldsymbol{\theta})^T \mathbf{o}. \tag{31}$$

Setting $\mathbf{s} = \hat{\mathbf{s}}_{/\theta}$ in J_2, the cost function to be minimized (only with respect to $\boldsymbol{\theta}$) becomes

$$J_3(\boldsymbol{\theta}) = ||\mathbf{o} - \mathbf{M}(\boldsymbol{\theta})[\mathbf{M}(\boldsymbol{\theta})^T \mathbf{M}(\boldsymbol{\theta})]^{-1} \mathbf{M}(\boldsymbol{\theta})^T \mathbf{o}||_2^2. \tag{32}$$

We use an iterative trust-region-reflective algorithm[5] to solve the following constrained optimization problem:

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in [\boldsymbol{\theta}_{min}, \boldsymbol{\theta}_{max}]}{argmin} J_3(\boldsymbol{\theta}). \tag{33}$$

After the convergence, the final value of $\widehat{\boldsymbol{\theta}}$ estimated by this iterative algorithm is used to calculate the mixing matrix $\mathbf{M}(\widehat{\boldsymbol{\theta}})$. Finally, we deduce from this matrix the estimator of the global vector of sources: $\hat{\mathbf{s}} = [\mathbf{M}(\widehat{\boldsymbol{\theta}})^T \mathbf{M}(\widehat{\boldsymbol{\theta}})]^{-1} \mathbf{M}(\widehat{\boldsymbol{\theta}})^T \mathbf{o}$. The non-negativity of sources may also be taken into account as described in the non-blind method.

The contributions of the object i in the 2D spectra observed in both directions may then be estimated using (30).

Once more, it is also possible to add a regularization term to the cost function J_2. In this case, $\hat{\mathbf{s}}_{/\theta} = [\mathbf{M}(\boldsymbol{\theta})^T \mathbf{M}(\boldsymbol{\theta}) + \boldsymbol{\Gamma}^T \boldsymbol{\Gamma}]^{-1} \mathbf{M}(\boldsymbol{\theta})^T \mathbf{o}$, and the cost function J_3 becomes $||\mathbf{o} - \mathbf{M}(\boldsymbol{\theta})\hat{\mathbf{s}}_{/\theta}||_2^2 + ||\boldsymbol{\Gamma}\hat{\mathbf{s}}_{/\theta}||_2^2$. The above algorithm may then be used with this new cost function.

5. Blind method

In this method, we make no assumption about the model of the PSF and about the spatial light profile of objects. The method is based on the second simplified model described in Section 3.2 and Eq. (23) and (25). We here provide some general comments about the mixture.

1. A pixel, according to its position, can receive radiations from a different number of objects. For example, in Fig. 2-Middle, the leftmost pixels are illuminated only by object 1, while those in the middle receive photons from objects 1 and 2.

---

[5]The `fmincon` MATLAB® function was used to this purpose.

2. Thanks to direct high-resolution photometric images of the sky, we approximately know the position (x_j, y_j) of each object. The parameters of the dispersion function $D(\lambda)$ (see Eq. (3)) and the length L of the 2D spectrum (in pixels) corresponding to the wavelength range of the grism are also known. Thus, we can compute the index $k_j^{(0)}$ of the first pixel of the 2D spectrum of a given object j, with respect to a reference point, in the 0-degree direction. An example of $k_1^{(0)}$, $k_2^{(0)}$ and L is provided in Fig. 2-Middle. Thus, we can write

$$e_j^{(0)}(n) = \begin{cases} z_j^{(0)}(n - k_j^{(0)}) & n \in]k_j^{(0)}, L + k_j^{(0)}] \\ 0 & \text{otherwise} \end{cases} , \qquad (34)$$

where $e_j^{(0)}(n)$ is defined in (19) and $z_j^{(0)}(n)$ stands for the non-zero part of $e_j^{(0)}$ after correcting its shift $k_j^{(0)}$ with respect to the reference point. As a result, some elements of $\mathbf{E}^{(0)}$ in (23)-(24) are known to be zero.
Similarly, some elements of $\mathbf{E}^{(90)}$ in (25) are known to be zero, and we can define

$$e_j^{(90)}(m) = \begin{cases} z_j^{(90)}(m - k_j^{(90)}) & m \in]k_j^{(90)}, L + k_j^{(90)}] \\ 0 & \text{otherwise} \end{cases} , \qquad (35)$$

where $e_j^{(90)}(m)$ is defined in (22) and $z_j^{(90)}(m)$ is the non-zero part of $e_j^{(90)}$ after correcting its shift with respect to the reference point.

The mixing model described by (23) and (25) corresponds to a linear instantaneous mixture, suggesting the use of Blind Source Separation (BSS) methods, which have been applied to astronomical data in several applications. For example, methods based on Independent Component Analysis (ICA) [1, 2, 28] have been used to separate the Cosmic Microwave Background (CMB) from the foreground sources [29, 30, 31, 32, 33], separate the components of galaxy images provided by the Hubble space telescope [34], or separate artefacts from astrophysical image data [35]. Other methods, based on the Non-negative Matrix Factorization (NMF) principle [36, 37, 38], have been e.g. used to separate particle spectra in the interstellar dust [10] or to separate stellar spectra in the dense fields [11].

In our application, however, the source spectra are not always independent so that the BSS methods based on ICA cannot be used[6]. Data involved in the mixtures being non-negative, we could be tempted to use the NMF principle. Nevertheless, the solution of the classical NMF is not unique and the NMF algorithms may converge towards spurious minima. In fact, in our tests on simulated data, classical unconstrained NMF provided poor results. Here, we propose a new method, based on the specific model defined by (23) and

---

[6]We measured the correlation coefficient between several pairs of spectra, and for some of them the result was significantly different from zero.

(25), which constrains the optimization algorithm, and so increases its chance to converge towards the right solution. This method exploits the fact that the target source spectrum to be decontaminated is shared by two different mixtures corresponding to its contamination in the 0- and 90-degree directions. In the following, the notation $\mathbf{H}(a : b, c : d)$ describes a sub-matrix made up of the rows from a to b and the columns from c to d of a matrix $\mathbf{H}$.

From (19), (22), (34) and (35), it is clear that both signals $z_i^{(0)}$ and $z_i^{(90)}$ correspond to "convolved"[7] versions of the target spectrum $s_i(\lambda)$ so that they are highly correlated. Thus, we propose to estimate the unknown matrices $\mathbf{A}^{(0)}$, $\mathbf{E}^{(0)}$, $\mathbf{A}^{(90)}$, $\mathbf{E}^{(90)}$ by minimizing the following cost function:

$$J_4 = 0.5||\mathbf{X}^{(0)} - \mathbf{A}^{(0)}\mathbf{E}^{(0)}||_2^2 + 0.5||\mathbf{X}^{(90)} - \mathbf{A}^{(90)}\mathbf{E}^{(90)}||_2^2 - \alpha\rho, \qquad (36)$$

subject to

$$\mathbf{A}^{(0)}, \ \mathbf{E}^{(0)}, \ \mathbf{A}^{(90)}, \ \mathbf{E}^{(90)} \geq 0$$

$$\mathbf{E}^{(0)}(j, n) = 0 \quad \text{if } n \notin]k_j^{(0)}, L + k_j^{(0)}] \quad \forall j \in \{i, \xi_0\}$$

$$\mathbf{E}^{(90)}(j, m) = 0 \quad \text{if } m \notin]k_j^{(90)}, L + k_j^{(90)}] \quad \forall j \in \{i, \xi_{90}\},$$

where α is a positive regularization parameter and ρ is the empirical correlation coefficient between the row vectors

$$\mathbf{z}_i^{(0)} = \mathbf{E}^{(0)}(1, 1 + k_i^{(0)} : L + k_i^{(0)}) \qquad (37)$$

and

$$\mathbf{z}_i^{(90)} = \mathbf{E}^{(90)}(1, 1 + k_i^{(90)} : L + k_i^{(90)}), \qquad (38)$$

i.e. the non-zero portions of the first rows of $\mathbf{E}^{(0)}$ and $\mathbf{E}^{(90)}$, which are related to the spectrum of the target object i. Thus, the regularization in (36) aims to provide a solution which maximizes the correlation between $\mathbf{z}_i^{(0)}$ and $\mathbf{z}_i^{(90)}$. Here $\rho = \sigma_{12}/(\sigma_1\sigma_2)$ where σ_{12} is the empirical covariance, defined as

$$\sigma_{12} = \frac{1}{L}\mathbf{z}_i^{(0)}.\mathbf{z}_i^{(90)T} - \left(\frac{1}{L}\mathbf{z}_i^{(0)}.\mathbf{1}\right)\left(\frac{1}{L}\mathbf{z}_i^{(90)}.\mathbf{1}\right), \qquad (39)$$

and σ_1^2 and σ_2^2 are the empirical variances of $\mathbf{z}_i^{(0)}$ and $\mathbf{z}_i^{(90)}$, defined as

$$\sigma_1^2 = \frac{1}{L}\mathbf{z}_i^{(0)}.\mathbf{z}_i^{(0)T} - \left(\frac{1}{L}\mathbf{z}_i^{(0)}.\mathbf{1}\right)^2,$$

$$\sigma_2^2 = \frac{1}{L}\mathbf{z}_i^{(90)}.\mathbf{z}_i^{(90)T} - \left(\frac{1}{L}\mathbf{z}_i^{(90)}.\mathbf{1}\right)^2, \qquad (40)$$

with $\mathbf{1}$ a column vector whose L elements are equal to one.

---

[7]It is not a standard convolution because in (19) and (22) there is $-D(\lambda)$ instead of $-\lambda$.

To update a matrix $\mathbf{R}$, we use the following rule in each iteration of a projected gradient algorithm [37]:

$$\mathbf{R} \leftarrow \max \left\{ \mathbf{R} - \mu \frac{\partial J_4}{\partial \mathbf{R}}, 0 \right\}, \tag{41}$$

with μ a positive small gradient step size.

Thus, matrices $\mathbf{A}^{(0)}$ and $\mathbf{A}^{(90)}$ are updated using the above rule and these gradients [39]:

$$\frac{\partial J_4}{\partial \mathbf{A}^{(0)}} = \mathbf{D}^{(0)} \mathbf{E}^{(0)T} \quad , \quad \frac{\partial J_4}{\partial \mathbf{A}^{(90)}} = \mathbf{D}^{(90)} \mathbf{E}^{(90)T}. \tag{42}$$

where $\mathbf{D}^{(0)} = \mathbf{A}^{(0)} \mathbf{E}^{(0)} - \mathbf{X}^{(0)}$ and $\mathbf{D}^{(90)} = \mathbf{A}^{(90)} \mathbf{E}^{(90)} - \mathbf{X}^{(90)}$.

To update matrices $\mathbf{E}^0$ and $\mathbf{E}^{90}$ except their first row (which contains the common target source), we use these gradients [39]:

$$\frac{\partial J_4}{\partial \mathbf{E}^{(0)}(2:end, 1:end)} = [\mathbf{A}^{(0)}(1:end, 2:end)]^T \mathbf{D}^{(0)},$$

$$\frac{\partial J_4}{\partial \mathbf{E}^{(90)}(2:end, 1:end)} = [\mathbf{A}^{(90)}(1:end, 2:end)]^T \mathbf{D}^{(90)}. \tag{43}$$

265 From the above discussion, we know which elements of matrices $\mathbf{E}^{(0)}(2:end, 1:end)$ and $\mathbf{E}^{(90)}(2:end, 1:end)$ must be zero. Thus, after applying the update rules, these elements are replaced by zeros.

For the first row of these matrices, we have to compute the gradients with respect to $\mathbf{z}_i^{(0)}$ and $\mathbf{z}_i^{(90)}$ using the definitions of $\mathbf{z}_i^{(0)}$, $\mathbf{z}_i^{(90)}$ and ρ. At first, we compute the derivatives of ρ with respect to $\mathbf{z}_i^{(0)}$ and $\mathbf{z}_i^{(90)}$. Since $\rho = \sigma_{12}/(\sigma_1 \sigma_2)$, we can write

$$\frac{\partial \rho}{\partial \mathbf{z}_i^{(0)}} = \frac{\sigma_1 \sigma_2 \frac{\partial \sigma_{12}}{\partial \mathbf{z}_i^{(0)}} - \sigma_2 \sigma_{12} \frac{\partial (\sigma_1^2)^{\frac{1}{2}}}{\partial \mathbf{z}_i^{(0)}}}{\sigma_1^2 \sigma_2^2}$$

$$= \frac{\sigma_1 \sigma_2 \frac{\partial \sigma_{12}}{\partial \mathbf{z}_i^{(0)}} - \frac{1}{2} \sigma_2 \sigma_{12} \sigma_1^{-1} \frac{\partial \sigma_1^2}{\partial \mathbf{z}_i^{(0)}}}{\sigma_1^2 \sigma_2^2} = \frac{\sigma_1^2 \frac{\partial \sigma_{12}}{\partial \mathbf{z}_i^{(0)}} - \frac{1}{2} \sigma_{12} \frac{\partial \sigma_1^2}{\partial \mathbf{z}_i^{(0)}}}{\sigma_1^3 \sigma_2}. \tag{44}$$

From (39), we get

$$\frac{\partial \sigma_{12}}{\partial \mathbf{z}_i^{(0)}} = \frac{1}{L} \mathbf{z}_i^{(90)} - \left(\frac{1}{L} \mathbf{1}^T \right) \left(\frac{1}{L} \mathbf{z}_i^{(90)} . \mathbf{1} \right) = \frac{1}{L} (\mathbf{z}_i^{(90)} - \bar{z}_i^{(90)} \mathbf{1}^T), \tag{45}$$

where $\bar{z}_i^{(90)}$ represents the mean of $\mathbf{z}_i^{(90)}$. From (40), we obtain

$$\frac{\partial \sigma_1^2}{\partial \mathbf{z}_i^{(0)}} = \frac{2}{L} \mathbf{z}_i^{(0)} - 2 \left(\frac{1}{L} \mathbf{1}^T \right) \left(\frac{1}{L} \mathbf{z}_i^{(0)} . \mathbf{1} \right) = \frac{2}{L} (\mathbf{z}_i^{(0)} - \bar{z}_i^{(0)} \mathbf{1}^T), \tag{46}$$

15

where $\bar{z}_i^{(0)}$ is the mean of $\mathbf{z}_i^{(0)}$. The derivative of ρ with respect to $\mathbf{z}_i^{(90)}$ can be computed in a similar way.

As a result, using (37) and (38), the gradients of J_4 with respect to $\mathbf{z}_i^{(0)}$ and $\mathbf{z}_i^{(90)}$ read

$$\frac{\partial J_4}{\partial \mathbf{z}_i^{(0)}} = [\mathbf{A}^0(1:end,1)]^T \mathbf{D}^0(1:end,1+k_i^{(0)}:L+k_i^{(0)})$$
$$- \alpha \frac{\sigma_1^2(\mathbf{z}_i^{(90)} - \bar{z}_i^{(90)}\mathbf{1}^T) - \sigma_{12}(\mathbf{z}_i^{(0)} - \bar{z}_i^{(0)}\mathbf{1}^T)}{L\sigma_1^3\sigma_2}, \tag{47}$$

$$\frac{\partial J_2}{\partial \mathbf{z}_i^{(90)}} = [\mathbf{A}^{90}(1:end,1)]^T \mathbf{D}^{90}(1:end,1+k_i^{(90)}:L+k_i^{(90)})$$
$$- \alpha \frac{\sigma_2^2(\mathbf{z}_i^{(0)} - \bar{z}_i^{(0)}\mathbf{1}^T) - \sigma_{12}(\mathbf{z}_i^{(90)} - \bar{z}_i^{(90)}\mathbf{1}^T)}{L\sigma_2^3\sigma_1}. \tag{48}$$

$\mathbf{z}_i^{(0)}$ and $\mathbf{z}_i^{(90)}$ are updated using these gradients. Then, the non-zero portions of the first rows of $\mathbf{E}^{(0)}$ and $\mathbf{E}^{(90)}$, corresponding to $\mathbf{z}_i^{(0)}$ and $\mathbf{z}_i^{(90)}$, defined in (37) and (38), are set to the updated values and the other elements of these first rows are set to zero.

After estimating $\mathbf{z}_i^{(0)}$ and $\mathbf{z}_i^{(90)}$ using this method, the contributions of the target object in the 2D spectra observed in both directions may be estimated by (see (23), (25), (37) and (38)):

$$\hat{\mathbf{X}}_i^{(0)} = \mathbf{A}^{(0)}(1:end,1)\mathbf{E}^{(0)}(1,1+k_i^{(0)}:L+k_i^{(0)}),$$
$$\hat{\mathbf{X}}_i^{(90)} = \mathbf{A}^{(90)}(1:end,1)\mathbf{E}^{(90)}(1,1+k_i^{(90)}:L+k_i^{(90)}). \tag{49}$$

This method is summarized in Algorithm 1.

A simpler version of the blind method may be derived by assuming that the signals $\mathbf{z}_i^{(0)}$ and $\mathbf{z}_i^{(90)}$ are equal (instead of highly correlated): $\mathbf{z}_i^{(0)} = \mathbf{z}_i^{(90)} = \mathbf{z}_i$. In this case, the third term of the cost function (36), *i.e.* $\alpha\rho$, will be removed. Then, the gradient of (36) with respect to $\mathbf{z}_i$ (supposing $\alpha = 0$) reads

$$\frac{\partial J_4}{\partial \mathbf{z}_i} = [\mathbf{A}^{(0)}(1:end,1)]^T \mathbf{D}^{(0)}(1:end,1+k_i^{(0)}:L+k_i^{(0)})$$
$$+ [\mathbf{A}^{(90)}(1:end,1)]^T \mathbf{D}^{(90)}(1:end,1+k_i^{(90)}:L+k_i^{(90)}). \tag{50}$$

The third step of Algorithm 1 then becomes:
3) Compute the gradient (50). Update the portions of the first rows of $\mathbf{E}^{(0)}$ and $\mathbf{E}^{(90)}$ using the rule (41) and the definitions of $\mathbf{z}_i^{(0)} = \mathbf{z}_i^{(90)} = \mathbf{z}_i$ provided in (37) and (38).
The initialization of $\mathbf{A}^{(0)}, \mathbf{A}^{(90)}, \mathbf{E}^{(0)}$ and $\mathbf{E}^{(90)}$ may be done using random values. A better idea may be to initialize them using the results of the non-blind

16

Algorithm 1 : Blind method

- Initialize $\mathbf{A}^{(0)}, \mathbf{A}^{(90)}, \mathbf{E}^{(0)}, \mathbf{E}^{(90)}$.

repeat

    1) Compute $\mathbf{D}^{(0)} = \mathbf{A}^{(0)}\mathbf{E}^{(0)} - \mathbf{X}^{(0)}$ and $\mathbf{D}^{(90)} = \mathbf{A}^{(90)}\mathbf{E}^{(90)} - \mathbf{X}^{(90)}$, then the gradients (42). Update $\mathbf{A}^{(0)}, \mathbf{A}^{(90)}$ using the rule (41).

    2) Compute $\mathbf{D}^{(0)}, \mathbf{D}^{(90)}$, then the gradients (43). Update $\mathbf{E}^{(0)}, \mathbf{E}^{(90)}$ except their first rows using the rule (41). Set to zero the elements specified in (34), (35).

    3) Compute the gradients (47) and (48). Update the portions of the first rows of $\mathbf{E}^{(0)}$ and $\mathbf{E}^{(90)}$ using the rule (41) and the definitions of $\mathbf{z}_i^{(0)}$ and $\mathbf{z}_i^{(90)}$ provided in (37) and (38).

until convergence

- Estimate the contributions of the object i in the 2D spectra using (49).
---

method, described at the beginning of Section 4, by assuming point-like objects (*i.e.* $\sigma_{a_j} = \sigma_{b_j} = r_j = 0$ $\forall j$ in (12)), and by using estimated values of the object positions and the PSF parameters. The rough estimated values of $\mathbf{M}$ and $\mathbf{s}$ provided by the non-blind method (see Eq. (27) and (28) and the paragraph before them) may then be used to get a first estimate of the entries of matrices $\mathbf{A}^{(0)}, \mathbf{A}^{(90)}, \mathbf{E}^{(0)}$ and $\mathbf{E}^{(90)}$ using (18), (19), (21) and (22). This initial estimate is then improved by the blind method.

It is also possible to add a smoothing criterion to the cost function, like that used in the semi-blind method.

6. Test results

In this section, we present some simulation results using data that we generated by the first version of the TIPS simulator[8], first considering a simple scenario with only 4 sources, then in a realistic dense field.

6.1. A simple scenario with 4 objects

In the first example, we consider a simple simulated scenario, shown in Fig. 3, where four realistic galaxies are placed in the sky such that in each of the 0- and 90-degree directions there are mixtures of two spectra. In fact, from this figure it is clear that when the light is dispersed in the 0-degree direction there will be two mixtures: the first one contains 2D spectra of objects 1 and 4 while the second one contains spectra of objects 2 and 3. Similarly, in the 90-degree direction, the spectra of objects 1 and 3 and those of objects 2 and 4 are mixed.

---

[8]This public version of TIPS was available here [40]. The new versions of this simulator, which are used to generate the official internal Euclid data, have been modified with respect to that first version.
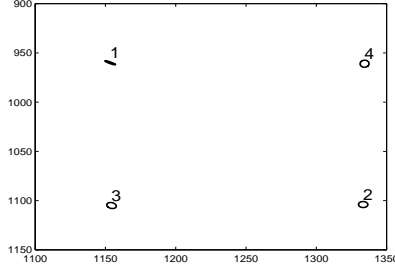
17

Figure 3: Spatial positions of 4 galaxies used in the simulation (coordinates: pixel indices).

6.1.1. Noiseless case

At first, from a source catalog defining magnitudes, shapes, locations and
[315] real spectra of 4 galaxies, and using the TIPS simulator, we generate observed
noiseless 2D spectra in both the 0 and 90-degree directions. The boxes contain-
ing the 4 above-mentioned mixed 2D spectra are shown in Fig. 4.

To decontaminate each galaxy spectrum, we use the two mixtures containing
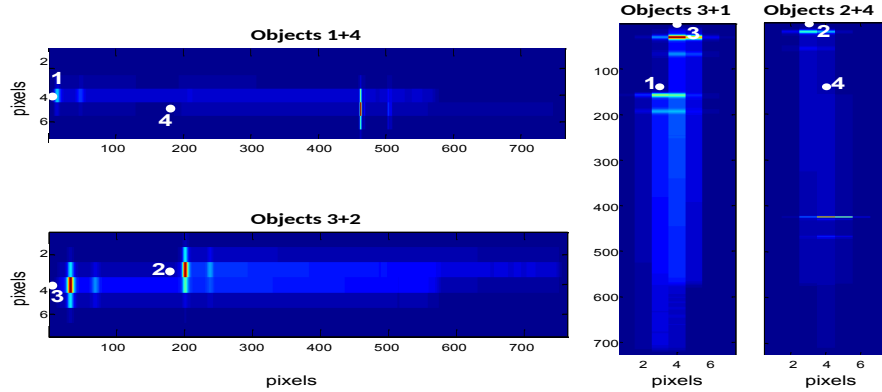

Figure 4: Noiseless mixed 2D spectra in the 0-degree (left) and 90-degree (right) directions.
The numbers label the beginning of the spectrum of each object in each mixture.

its contributions. Semi-blind and blind methods are used to this purpose.
[320] To illustrate the results of the semi-blind method, we used the mixture of
object 3 with object 2 in the 0-degree direction and with object 1 in the 90-
degree direction, shown in Fig. 4, to construct the vector **o** in Eq. (13). After
applying the semi-blind method, without a smoothing constraint, we obtained
the estimated 2D spectrum of object 3 in both directions.
[325] To check the quality of this estimation, we determine the contribution of ob-
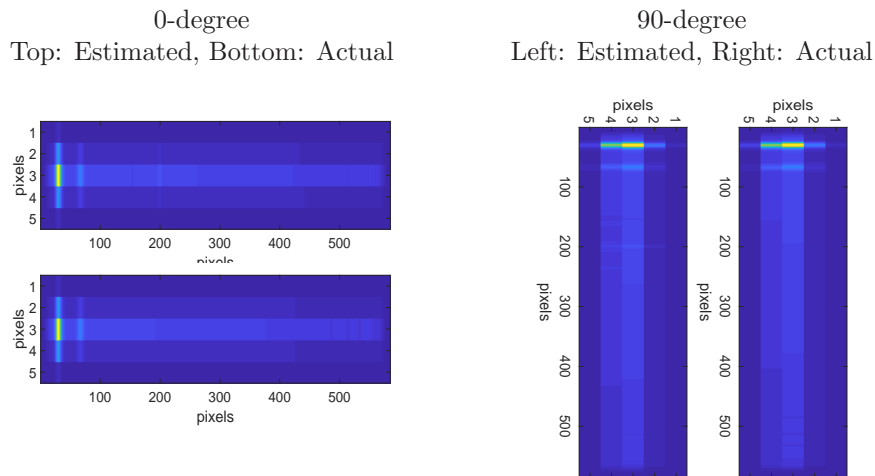ject 3 in the 2D mixed spectra, by putting *only* this object at the input of the

18

Figure 5: Estimated 2D spectrum of object 3 in both directions by the semi-blind method, compared with the actual spectrum.

TIPS simulator and measuring the corresponding actual 2D spectrum of this object (in both directions) at its output. Fig. 5 shows the actual and estimated spectra, which are very similar.

To better interpret the results, we can present and study the so-called 1D spectrum, defined here as the central row of the 2D spectrum (*i.e.* the row, in the dispersion direction, containing the center of an object and corresponding to the maximum intensity). The first three rows of Fig. 6 shows a comparison between:

- The actual 1D spectrum of object 3, which corresponds to the 3rd row of its actual 2D spectrum represented in Fig. 5, for both directions: the amplitude of the 1D spectrum is related to the color (or the gray scale) of the 2D spectrum.

- The observed (i.e. mixed) 1D spectrum of object 3, which corresponds to the 4th row of its observed 2D spectrum, in both directions, represented in Fig. 4. We can easily see the contamination of object 3 by object 2 in the 0-degree direction and by object 1 in the 90-degree direction, mainly characterized by the presence of spurious emission lines.

- The estimated 1D spectrum of object 3 in both directions, obtained by the semi-blind method, which confirms its effectiveness.

Our tests with different values of $\boldsymbol{\theta}_{min}$ and $\boldsymbol{\theta}_{max}$ in the semi-blind method show that although the performance slightly varies with these values, it is not so sensitive to it unless the random initial values of the parameters, which are chosen in the interval $[\boldsymbol{\theta}_{min}, \boldsymbol{\theta}_{max}]$, are too far from their actual values (especially for the object positions). If wider intervals are chosen, convergence is generally slows down and the computation time increases, because in this case the initial
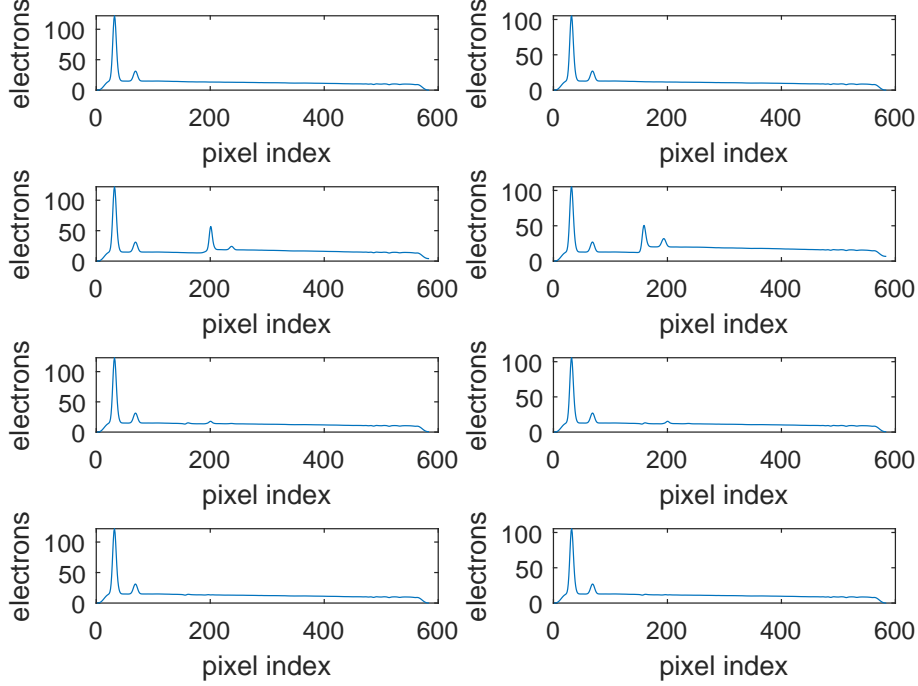
19

Figure 6: 1D spectra in the noiseless case. Left: 0-degree direction, Right: 90-degree direction. First row: actual spectra of object 3, Second row: observed mixed spectra, Third row: estimated spectra of object 3 using the semi-blind method, Forth row: estimated spectra of object 3 using the blind method.

random values are usually further from their actual values.

We then applied our blind method, randomly initialized, to the same mixture as used above. The correlation constraint parameter and the gradient step size were $\alpha = 10$ and $\mu = 0.001$. The last row of Fig. 6 shows the estimated 1D spectrum of object 3 obtained by this method in both directions. Once more, the decontamination is satisfactory. To compare the performances of the two methods, the Signal-to-Interference Ratio (SIR), in dB, for each galaxy before and after decontamination is computed by

$$SIR_{\text{before}}(i) = \frac{1}{2}\sum_{r=\{0,90\}} 10 \log_{10} \frac{\left\|\mathbf{Actual}_i^{(r)}\right\|_2^2}{\left\|\mathbf{Observed}_i^{(r)} - \mathbf{Actual}_i^{(r)}\right\|_2^2} \tag{51}$$

$$SIR_{\text{after}}(i) = \frac{1}{2}\sum_{r=\{0,90\}} 10 \log_{10} \frac{\left\|\mathbf{Actual}_i^{(r)}\right\|_2^2}{\left\|\mathbf{Estimated}_i^{(r)} - \mathbf{Actual}_i^{(r)}\right\|_2^2} \tag{52}$$

where **Actual**$_i^{(r)}$, **Observed**$_i^{(r)}$ and **Estimated**$_i^{(r)}$ are respectively the actual (and noiseless), observed and estimated 2D spectra related to object i in the r-degree direction.

To decontaminate each of the galaxy spectra in the scenario shown in Fig. 4, we applied our semi-blind and blind methods to the couple of mixed spectra where this galaxy spectrum is present. Results are summarized in Table 1 and confirm the effectiveness of both methods. The blind method provides better

Table 1: SIR (in dB) before and after decontamination of noiseless mixture, and the computation time.

galaxy	SIR_{before}	Semi-Blind		Blind	
		SIR_{after}	Time	SIR_{after}	Time
1	1.93 dB	25.55 dB	9.76 min	33.03 dB	0.32 min
2	2.72 dB	28.08 dB	20.76 min	33.35 dB	0.21 min
3	1.80 dB	27.11 dB	9.84 min	33.82 dB	0.18 min
4	3.41 dB	27.59 dB	15.25 min	20.96 dB	0.26 min

results for three objects and requires much less computation time than the semi-blind method.

6.1.2. Noisy data

We then repeated the same simulation using noisy data. Four realistic mixed 2D spectra simulated by the TIPS simulator are shown in Fig. 7. The high level of noise may be remarked by comparing this figure with Fig. 4. The estimated
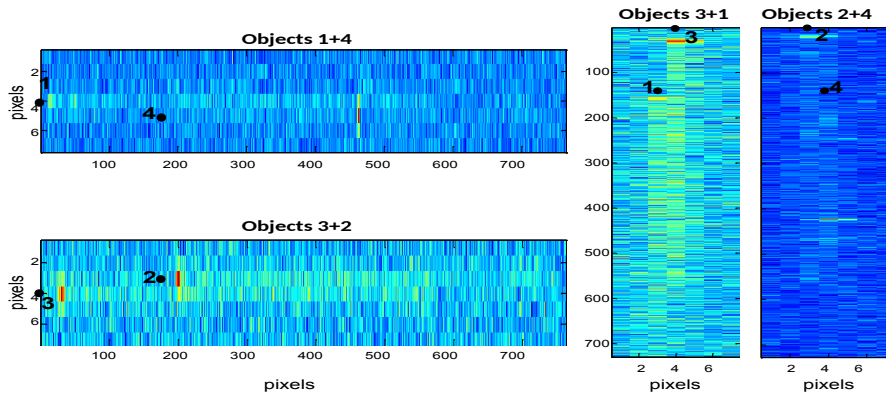

Figure 7: Observed noisy 2D spectra in the 0-degree (left) and 90-degree (rigth) directions.

1D spectra of object 3 in both directions are compared with the actual spectrum in Fig. 8. The decontamination using both methods is quite successful although

21

the estimated sources are very noisy. Note that our methods are designed to separate, and not to denoise, source spectra. Results are summarized in Table
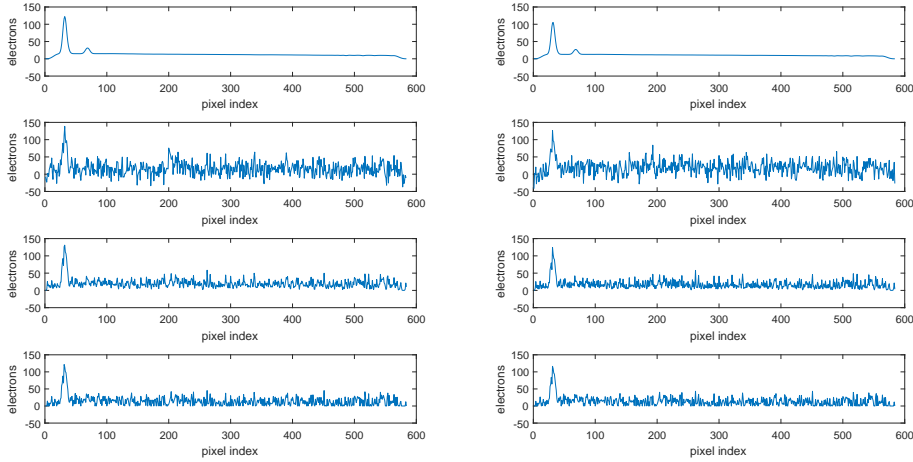

Figure 8: 1D spectra in the noisy case. Left: 0-degree direction, Right: 90-degree direction. First row: actual spectra of object 3, Second row: observed mixed spectra, Third row: estimated spectra of object 3 using the semi-blind method, Forth row: estimated spectra of object 3 using the blind method.

2. As can be seen, the SIR is improved by both methods. Note that even when the decontamination is perfectly achieved, according to (52), the SIR after decontamination is equal to the output signal to noise ratio, which is small because of the high noise level. The two methods provide globally comparable results but the blind method requires much less computation time.

6.2. Simulation results for a realistic dense field

6.2.1. Data simulation

Here, we test our methods in a scenario where the spatial density of simulated extragalactic objects is much more realistic. Thus, 1000 objects are randomly distributed over a 1000×1000 pixel scene. The object positions are shown in Fig. 9. Using a source catalog defining magnitudes, shapes, locations and real

Table 2: SIR (in dB) before and after decontamination of noisy mixture, and the computation time.

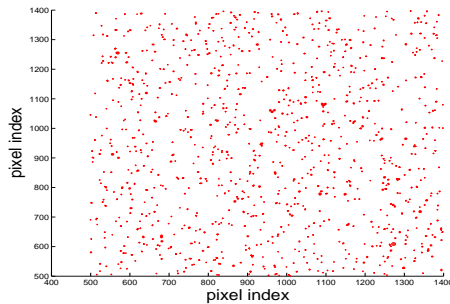|		Semi-Blind		Blind	
galaxy	SIR_{before}	SIR_{after}	Time	SIR_{after}	Time
1	-8.20 dB	2.68 dB	5.13 min	2.09 dB	0.40 min
2	-7.21 dB	3.71 dB	6.09 min	3.71 dB	0.22 min
3	-6.99 dB	3.64 dB	6.29 min	4.11 dB	0.28 min
4	-7.67 dB	2.61 dB	5.03 min	3.70 dB	0.25 min

Figure 9: Spatial configuration of galaxies in a realistic dense field.

spectra of 1000 stellar objects, the TIPS simulator generated mixtures of 2D spectra in 0 and 90-degree directions. From a scientific point of view, the objects of interest in this work are assumed to have an H_α emission line[9] in the interval [1200 - 2000] nm whose flux must be greater than 3×10^{-16} $erg/s/cm^2$.[10] Among the 1000 objects in the above field, only 35 objects satisfy these conditions. In the following, the decontamination results using our methods will be presented for these 35 objects.

6.2.2. Contaminant detection

Given an object of interest to decontaminate, we must at first determine the objects which contaminate it. This may be easily done from the object positions and the instrument parameters. In fact, for each object, we can compute the box containing its 2D spectrum. If the box of the target object overlaps with that of other objects, the latter objects are considered as the contaminants of the target object.

6.2.3. Decontamination results

We tested different versions of our semi-blind and blind methods to decontaminate 35 objects of interest in two different configurations: first in a noiseless case, then in a noisy one where strong and realistic noise was added to mixtures. With the semi-blind method, the best results were obtained by adding a smoothing constraint with $\beta = 0.2$ in (29). The best results of the blind method were obtained by assuming $\mathbf{z}_i^{(0)} = \mathbf{z}_i^{(90)}$ and by initializing matrices using the non-blind estimation as explained at the end of Section 5.

For each of the 35 target objects, the contamination performances were computed using the SIR formulas defined in (51) and (52). The mean and standard deviation of 35 SIR values before and after decontamination in the noiseless case

---

[9]H_α is the main emission line in the star-forming galaxy spectra, created by a hydrogen atom when an electron falls from the third lowest to the second lowest energy level.

[10]These values, initially considered for the objects of interest in the Euclid mission, have slightly changed since the work reported in this paper.

are provided in Table 3 for the semi-blind and blind methods. Both methods provide satisfactory results, leading to a nearly 19-dB average SIR improvement.

As an example, Fig. 10 shows 1D spectra related to the first target object in

Table 3: Mean and standard deviation of SIR (in dB) before and after decontamination of noiseless mixture in the dense field.

		SIR_{after}			
SIR_{before}		Semi-blind		Blind	
mean	std	mean	std	mean	std
-5.52 dB	7.49 dB	13.55 dB	7.55 dB	13.59 dB	7.40 dB

the noiseless case. As can be seen, the emission line related to a contaminating object in the 0-degree direction, and the continuum due to contaminating objects in both directions are largely removed by both methods.
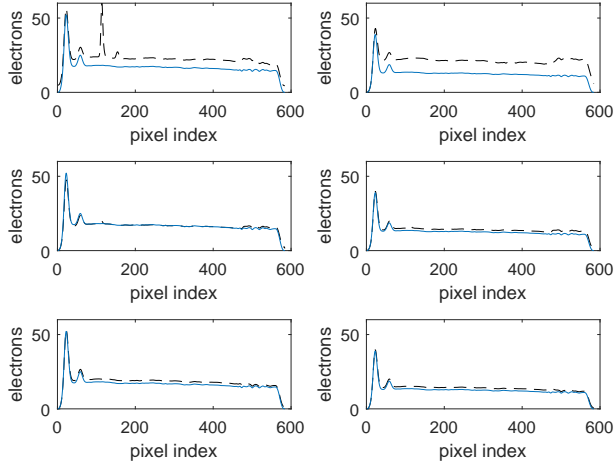

Figure 10: 1D spectra in the noiseless case for one of the objects of interest in the dense field scenario. Left: 0-degree direction, Right: 90-degree direction. Solid line: actual spectra of the target object, Dashed line in the first row: observed mixed spectra, Dashed line in the second row: estimated spectra using the semi-blind method, Dashed line in the third row: estimated spectra using the blind method.

The SIRs were also computed in the noisy case for each of the 35 target objects. The histograms of the SIR improvement, defined as the difference between the SIRs after and before decontamination, are shown in Fig. 11. As can be seen, the SIR is improved for all the objects using the blind method, and for all the objects except one using the semi-blind method. This improvement, and therefore the quality of decontamination, is different for different objects, and mainly depends on the object brightness and the level of contamination

in two directions. The mean and standard deviation of 35 SIR values before and after decontamination in this noisy case are listed in Table 4. The blind method provides slightly better results than the semi-blind one: the average SIR is improved about 16dB by the blind method and about 14dB by the semi-blind method.

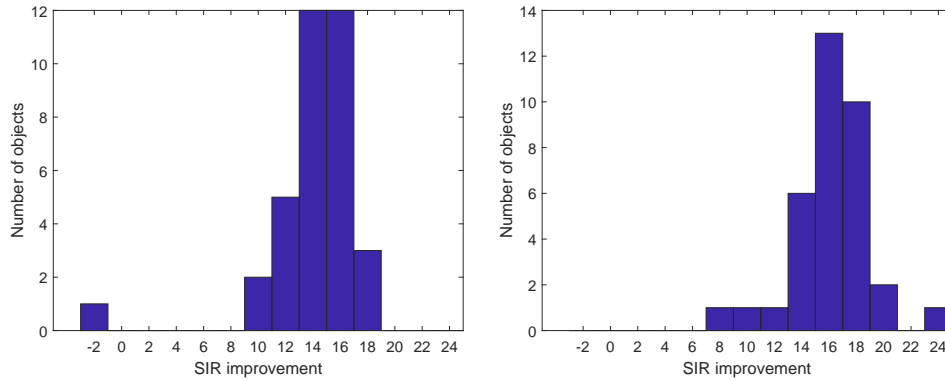Fig. 12 illustrates the decontamination results for the first target object in the

Figure 11: Histograms of the SIR improvement (with counts as label) for the 35 target objects using the semi-blind (left figure) and blind (right figure) methods in the noisy case.

noisy case, and confirms the effectiveness of both methods.

Table 4: Mean and standard deviation of SIR (in dB) before and after decontamination of noisy mixture in the dense field.

SIR_{before}		SIR_{after}			
		Semi-blind		Blind	
mean	std	mean	std	mean	std
-14.15 dB	4.30 dB	-0.07 dB	5.02 dB	2.10 dB	2.81 dB

For both methods, the computation time is different for different objects, and increases with the number of contaminants. The blind method is 20 to 100 times faster than the semi-blind method and in our tests it leads often to better results. This may be explained as follows. First, note that the semi-blind method is basically based on a "convolutive" model (see Eq. (6) and (7)) and is supposed to directly estimate the object spectra (denoted by s_j). In other words, it must achieve both source separation and "deconvolution". On the other hand, the blind method is based on an instantaneous model and aims at estimating the signals e_j, defined by Eq. (19) and (22), which are the results of "convolution" between the spectrum of an object and its spatial profile in the
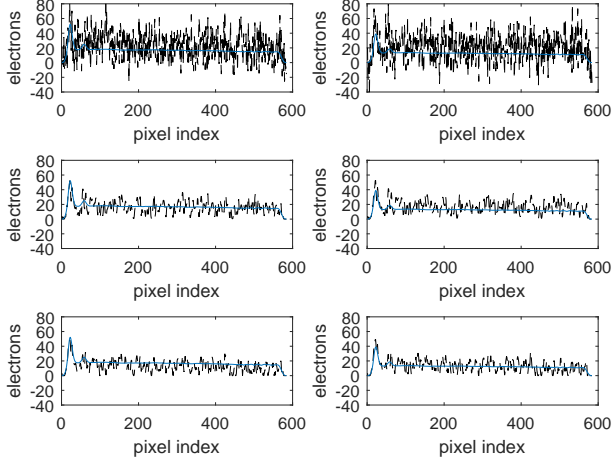
Figure 12: 1D spectra in the noisy case for one of the objects of interest in the dense field scenario. Left: 0-degree direction, Right: 90-degree direction. Solid line: actual spectra of the target object, Dashed line in the first row: observed mixed spectra, Dashed line in the second row: estimated spectra using the semi-blind method, Dashed line in the third row: estimated spectra using the blind method.

dispersion direction. This method does not aim to "deconvolve" spectra: its task is logically simpler. This difference could explain the better performance of the blind method. Moreover, the mixing matrix in the semi-blind method is much larger than that of the blind method. For example, consider a scenario where the observed area of the sky contains 3×500 pixels in each of the 0 and 90-degree directions, with 3 objects each containing 500 wavelength samples (a target object and a contaminant in each direction). In the mixing equation of the semi-blind method (Eq. (13)), the observation vector $\mathbf{o}$ will be of size 3000×1, the source vector $\mathbf{s}$ of size 1500×1, and the mixing matrix $\mathbf{M}$ of size 3000×1500. On the other hand, for each of the mixing models used in the blind method (Eq. (23) and (25)), the observation matrix $\mathbf{X}$ in each direction will be of size 3×500, the source matrix $\mathbf{E}$ in each direction of size 2×500 and the mixing matrix $\mathbf{A}$ in each direction of size 3×2 only. This may explain why the blind method is much faster.

7. Conclusion

In this paper, we investigated the problem of spectra decontamination in slitless spectroscopy. The proposed methods may be used in space missions like the WISP survey and the Euclid project, and more generally in other slitless spectroscopy applications.

We first established a physical model linking observed data to source spectra in two dispersion directions, then simplified it to get a parametric model and a linear instantaneous model. Based on the parametric model, we proposed

26

a semi-blind method which estimates source spectra together with the model parameters. A blind method, based on the linear instantaneous model, was also proposed. This method exploits the non-negativity and spatial sparsity of data, and the correlation between spectra of the target source in two dispersion directions.

Both methods were tested on noiseless and noisy realistic simulated data, first using a simple scenario with 4 objects, then with a realistic dense field. The obtained results are very encouraging and confirm the effectiveness of the proposed methods. Quantitatively, in all our tests, the signal to interference ratio is improved by both methods for almost all the considered objects. From an astronomer point of view, one of the most important things is the ability of the decontamination method to eliminate emission lines and continuum of contaminating objects, so as to facilitate the detection of the emission lines belonging to the object of interest. As can be seen in our tests, both methods succeed in accomplishing this task satisfactorily, even though the blind method provides slightly better results.

Several issues may be considered in future investigations. If some information is available about the noise covariance matrix or about the spectral or spatial variations of the PSF, it may be used to enhance our algorithms. In this paper, only first-order spectra generated by grisms were considered. For very bright objects, zero-order and second-order spectra are not negligible and should be taken into account. In practice, the dispersed spectra may be slightly curved. in this case, this curvature should be considered and corrected before applying our methods. Finally, in our work, we considered only two grisms. If several grisms with different orientations are available, merging their information may improve the estimation quality.

Appendix A: Testing the plausibility of the simplified model 2

In Section 3.2, we used the separability assumption (16) to derive the simplified model 2, described by Eq. (17) and (20), which is a linear instantaneous model. While the precision of the separability assumption is difficult to check directly, we can perform a test to check the plausibility of the resulting linear instantaneous model. According to Eq. (17), the convolved spectrum $e_j^{(0)}$, defined by Eq. (19), does not depend on the y-coordinate of the pixel (i.e. m): the integral defining it has the same value for all detector pixels located on a vertical line. This effect may be measured in the following manner. If the mixing model

27

(17) is valid, for a unique non-contaminated object i, Eq. (17) reads:

$$o^{(0)}(n, m) = a_i^{(0)}(m)e_i^{(0)}(n). \tag{53}$$

As a result, in different rows of the box containing the observed 2D spectrum related to this object, we should have the same signal up to a scale factor. Consequently, the observed spectra on different rows will be proportional. This effect may be e.g. observed in Fig. 5 (0-degree, bottom), which shows the actual 2D spectrum of object 3. To check this effect, we performed the following test. We put a single object at the input of the TIPS simulator and obtained at the output its (non-contaminated) 2D spectrum, then measured the correlation coefficient between observations in different rows of the box. Repeating this test with objects with different shapes and characteristics, we computed 84 correlation coefficients. 76.19% of the results were greater than 0.99, 14.29% were between 0.95 and 0.99, and the other 9.52% were between 0.90 and 0.95. Thus, we can consider that the linear instantaneous model, based on the separability assumption, is quite realistic.

References

[1] A. Hyvärinen, J. Karhunen, E. Oja, *Independent Component Analysis*, Wiley, 2001.

[2] P. Comon and C. Jutten Eds., *Handbook of blind source separation. Independent component analysis and applications*, Academic Press, Oxford, 2010.

[3] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, J. Chanussot, Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354-379, April 2012.

[4] S. Karoui, Y. Deville, S. Hosseini, A. Ouamri, Blind spatial unmixing of multispectral images : New methods combining sparse component analysis, clustering and non-negativity constraints, *Pattern Recognition*, vol. 45, Issue 12, pp. 4263-4278, Dec. 2012.

[5] L. T. Duarte, S. Moussaoui and C. Jutten, Source Separation in Chemical Analysis : Recent achievements and perspectives, *IEEE Signal Processing Magazine*, vol. 31, no. 3, pp. 135-146, May 2014.

[6] S. Hosseini, Y. Deville, L. T. Duarte, A. Selloum, Extending NMF to blindly separate linear-quadratic mixtures of uncorrelated sources, in proceedings of 26th IEEE International Workshop on Machine Learning for Signal Processing (MLSP), Vietri sul Mare, Salerno, Italy, Sep. 2016.

[7] D. Nuzillard and J. M. Nuzillard, Application of blind source separation to 1-D and 2-D nuclear magnetic resonance spectroscopy, *IEEE Signal Processing Letters*, vol. 5, no. 8, pp. 209-211, Aug. 1998.

[8] I. Toumi, S. Caldarelli, B. Torrésani, A review of blind source separation in NMR spectroscopy, *Progress in Nuclear Magnetic Resonance Spectroscopy*, Vol. 81, Pp. 37-64, Aug. 2014.

[9] Jorge Igual, Raul Llinares, An informed source separation of astrophysical ice analogs, *Digital Signal Processing*, Vol. 17, Issue 5, pp. 947-964, Sep. 2007.

[10] O. Berné, C. Joblin, Y. Deville, J. D. Smith, M. Rapacioli, J. P. Bernard, J. Thomas, W. Reach, and A. Abergel, Analysis of the emission of very small dust particles from Spitzer spectro-imagery data using blind signal separation methods, *Astronomy & Astrophysics*, vol. 469, 2007.

[11] I. Meganem, S. Hosseini, and Y. Deville, Chapter 6. Separation of stellar spectra based on nonnegativity and parametric modelling of mixing operator, in *Non-negative matrix factorization techniques : Advances in theory and applications*, G. R. Naik Ed, Springer-Verlag, Berlin Heidelberg, 2016.

[12] WFC3 Infrared Spectroscopic Parallel Survey Homepage http://wisps.ipac.caltech.edu/Home.html

[13] H. Atek et al., The WFC3 Infrared Spectroscopic Parallel (WISP) Survey, *The Astrophysical Journal*, Vol. 723, Issue 1, pp. 104-115, 2010.

[14] Euclid Consortium page, http://www.euclid-ec.org/

[15] NASA Science: astrophysics, http://science.nasa.gov/astrophysics/focus-areas/what-is-dark-energy/

[16] Euclid Definition Study Report (Redbook), http://arxiv.org/ftp/arxiv/papers/1110/1110.3193.pdf, ESA, 2011.

[17] A. Selloum, S. Hosseini, T. Contini, Y. Deville, Semi-blind separation of galaxy spectra from a mixture obtained by slitless spectroscopy, in proceedings of *23rd European Signal Processing Conference (EUSIPCO)*, pp. 1671-1675, Nice, France, September 2015.

[18] A. Selloum, S. Hosseini, Y. Deville, T. Contini, Mixing model in slitless spectroscopy and resulting blind methods for separating galaxy spectra, in proceedings of 26th IEEE International Workshop on Machine Learning for Signal Processing (MLSP), Vietri sul Mare, Salerno, Italy, September 2016.

[19] J. R. Walsh, M. Kümmel, S. S. Larsen, Recent developments of the ACS spectral extraction software aXe, The 2005 HST calibration workshop, Space Telescope Science Institute, 2005.

[20] M. Kümmel, J. R. Walsh, N. Pirzkal, H. Kuntschner, A. Pasquali, The Slitless Spectroscopy Data Extraction Software, *Publications of the Astronomical Society of the Pacific (PASP)*, 121:59-72, Jan. 2009.

[21] J. Zoubian, *Observations cosmologiques avec un télescope grand champ spatial : Simulations pixels du spectromètre sans fente d'EUCLID*, PhD thesis (in French), May 2012.

[22] G. Aldering, SNAP sky background at the north ecliptic pole, LBNL report (LBNL-51157), 2001.

[23] C. Palmer, E. Loewen, *Diffraction grating handbook*, Richardson Gratings, 2014.

[24] W. Freudling et al., The Hubble Legacy Archive NICMOS grism data, *Astronomy & Astrophysics* 490, 1165-1179, 2008.

[25] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, Pergamon Press, 1970.

[26] A.W. Graham, S.P. Driver, A Concise Reference to (Projected) Sérsic $R^{1/n}$ Quantities, Including Concentration, Profile Slopes, Petrosian Indices, and Kron Magnitudes, *Publications of the Astronomical Society of Australia*, Volume 22, Issue 2, pp. 118-127, 2005.

[27] C.L Lawson, R.J. Hanson, Solving Least Squares Problems, SIAM, 1974.

[28] Y. Deville, Blind Source Separation and Blind Mixture Identification Methods, *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp. 1-33, 2016.

[29] C. Baccigalupi, L. Bedini, C. Burigana, G. De Zotti, A. Farusi, D. Maino, M. Maris, F. Perrotta, E. Salerno, L. Toffolatti, and A. Tonazzini, Neural networks and the separation of cosmic microwave background and astrophysical signals in sky maps, *Monthly Notices of the Royal Astronomical Society*, vol. 318, pp. 769-780, 2000.

[30] Y. Moudden, J.-F. Cardoso, J.-L. Starck, J. Delabrouille, Blind component separation in wavelet space: Application to CMB analysis. *EURASIP Journal on Advances in Signal Processing*, vol. 15, Aug. 2005.

[31] M. T. Özgen, E. E. Kuruoglu, D. Herranz, Astrophysical image separation by blind time-frequency source separation methods, *Digital Signal Processing*, vol. 19, Issue 2, pp. 360-369, 2009.

[32] E. E. Kuruoglu, Bayesian source separation for cosmology, *IEEE Signal Processing Magazine*, vol. 27, no. 1, pp. 43-54, Jan. 2010.

[33] Planck Collaboration: P. A. R. Ade et al., Planck 2013 results. XII. Diffuse component separation, *Astronomy & Astrophysics*, vol. 571, 2014.

[34] D. Nuzillard, and A. Bijaoui, Blind source separation and analysis of multispectral astronomical images, *Astronomy and Astrophysics Supplement Series*, vol. 147, pp. 129-138, 2000.

[35] M. Funaro, E. Oja, H. Valpola, Independent component analysis for artefact separation in astrophysical images, *Neural Networks*, vol. 16, Issues 34, pp. 469-478, 2003.

[36] D. D. Lee, H. S. Seung, Learning the parts of objects by non-negative matrix factorization, *Nature*, vol. 401, pp. 788-791, Oct. 1999.

[37] C.-J. Lin, Projected gradient methods for non-negative matrix factorization, *Neural Computation*, 19:2756-2779, 2007.

[38] A. Cichocki, R. Zdunek, A. H. Phan, and S.-I. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*, Hoboken, NJ, USA: Wiley, 2009.

[39] K. B. Petersen, M. S. Pedersen, *The Matrix Cookbook*, http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3274/pdf/imm3274.pdf, Technical University of Denmark, 2012.

[40] TIPS Simulator page, http://projects.lam.fr/projects/tips/