SPARSITY-BASED SPATIAL-SPECTRAL RESTORATION OF MUSE ASTROPHYSICAL HYPERSPECTRAL DATA CUBES

Sébastien Bourguignon[⋆], *Hervé Carfantan*[†], *Éric Slezak*[⋆], *David Mary*[◇]

[⋆,◇]Université de Nice Sophia Antipolis, CNRS, Observatoire de la Côte d'Azur, [⋆]Cassiopée, [◇]Fizeau, Boulevard de l'Observatoire, BP 4229, F-06304 Nice, France
[†]Université de Toulouse, UPS-OMP / CNRS; IRAP; 14, av. Edouard Belin, F-31400 Toulouse, France

ABSTRACT

We consider the restoration of extragalactic deep field hyperspectral imaging data, in the context of the forthcoming MUSE instrument. Joint spatial-spectral restoration is addressed by taking into account the three-dimensional point spread function (PSF) of the instrument and the noise statistical distribution, with strong spectral variations for both of them. Since objects of interest have limited spatial extensions, restoration is formulated for sub-cubes with restricted spatial coverage. Prior information is incorporated by means of sparsity constraints in the spectral domain, using a specific dictionary with physically meaningful elementary features. We decompose the too high-dimensional underlying optimization problem into two steps by exploiting the separability property of the PSF. First, spectra are processed independently, where sparsity performs spectral dimension reduction. Then, a computationally tractable three-dimensional restoration problem is solved. Simulations reveal the interest of this approach, where restoration efficiently performs the separation of two close objects and the unmixing of their spectra.

***Index Terms*—** Spatial-spectral processing, data restoration, sparse approximation, astrophysical hyperspectral imaging.

1. INTRODUCTION

The forthcoming MUSE integral-field spectrograph will provide massive hyperspectral observations of extragalactic deep fields, with 300×300 pixels and up to $4\,000$ wavelengths [1]. One of the major challenges of MUSE is to detect and characterize very distant, and therefore faint, emitting sources. Objects of interest may cover only a few pixels at the instrument spatial resolution (0.2×0.2 arcsec^2 per pixel), or even be smaller than the pixel size (*unresolved* sources). Because data will be collected in a very noisy environment, only one or a few spectral features may be detectable, such as very significant emission lines or breaks at unknown wavelengths. Such very localized (both spatially and spectrally) informational content, when observed through MUSE, will be spread on spatial and spectral neighborhoods in the cube due to the three-dimensional point spread function (PSF) of the instrument.

Data restoration in this context has a twofold objective. First, it aims at improving data quality by performing deconvolution and, up to a certain point, noise reduction. Second, by re-concentrating information at its original spatial and spectral location, it aims at getting better detection statistics than would be obtained by analyzing the data as a collection of distinct images or spectra. In this paper, restoration is addressed as an inverse problem [2], which takes into account instrumental and noise characteristics, which is crucial in our case since both the PSF and the noise distribution vary with wavelength. Since most information in MUSE deep field observations is localized at relatively small spatial scales, sub-cubes with restricted spatial dimensions are considered. Doing so, sophisticated methods with higher complexity can be developed.

Given the very high level of noise affecting the data, reconstruction has to be constrained with prior assumptions. Relevant prior information in deep field hyperspectral data can be expressed spectrally, where strong constraints can be incorporated in accordance to astrophysical knowledge [3]. In a recent work [4], a sophisticated prior model for galaxy *spectra* was developed, based on sparsity constraints in a specifically designed, highly redundant, dictionary. Restoration was then combined with the detection of salient spectral features. This same model is considered here, jointly with a spatial-spectral instrumental model, which better fits the observational context than an approach operating individually on each spectrum. It results, however, in a considerable increase in the complexity of the estimation problem. Therefore, we exploit the implicit dimension reduction operated by spectral sparsity constraints to build a suboptimal but tractable three-dimensional restoration method.

Section 2 details observational specificities and formulates the restoration inverse problem. In Section 3, sparsity-based prior spectral information is motivated in our context and details are given about the dictionary design. Section 4 describes the proposed approach: sparse approximations are first obtained for each spectrum in order to perform dimension reduction. Then, spatial-spectral restoration is performed in spectrally reduced dimension. Simulation results and conclusions are finally given in Section 5.

2. OBSERVATIONAL SPECIFICITIES AND PROBLEM FORMULATION

2.1. Three-dimensional PSF

We consider MUSE as a linear system, characterized by its three-dimensional PSF, that is, its impulse response in both spatial and spectral domains. In the following, $\boldsymbol{r} \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$ denote a couple of spatial coordinates and wavelength, respectively. We assume that the PSF H is separable into a spatial, bi-dimensional, *field* spread function (FSF) F and a spectral *line* spread function (LSF) L. At every element $(\boldsymbol{r}_0, \lambda_0)$ of the cube, the PSF then reads:

$$H_{\boldsymbol{r}_0,\lambda_0}(\boldsymbol{r},\lambda) = F_{\boldsymbol{r}_0,\lambda_0}(\boldsymbol{r}) L_{\boldsymbol{r}_0,\lambda_0}(\lambda) \in \mathbb{R}. \quad (1)$$

Both FSF and LSF are not translation invariant, so that the PSF cannot be written under the usual form $H(\boldsymbol{r}-\boldsymbol{r}_0, \lambda-\lambda_0)$. In particular, the blurring effect due to the FSF and, to a lesser extent, to the LSF is expected to decrease as wavelength increases.

This work was partially supported by ANR project 08-BLAN-0253-01 DAHLIA - Dedicated Algorithms for HyperspectraL Imaging in Astronomy.

2.2. Spectrally variable noise

The noise level is also expected to vary with wavelength. Indeed, noise is mainly due to powerful parasite emissions in the atmosphere at specific wavelengths. The quantum efficiency of the instrument is also wavelength-dependent. Least, a 20 nm bandwidth around 590 nm is reserved for a laser reference star adaptive optics system, where no signal can be detected. Figure 1 plots the equivalent noise variance combining these three effects as a function of wavelength. Spatial variations of the noise level are also expected, due to the Poisson nature of light emission: noise variance is higher at pixels where light flux is higher. In the following, we consider that noise samples are zero-mean, Gaussian[1] and independent, but with different variances for each data point, that are supposed to be known.

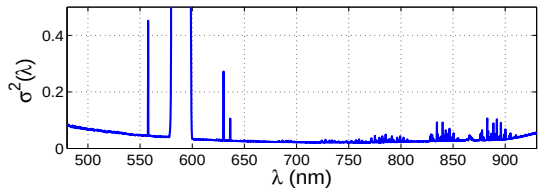


Fig. 1. Noise variance as a function of wavelength. $\sigma(\lambda)$ corresponds to light flux, in erg.s^{-1}.cm^{-2}.pixel^{-1} ($\times 10^{-20}$).

2.3. Notations and inverse problem formulation

Let $\boldsymbol{y}_k$ denote the spectrum in the hyperspectral data cube corresponding to the k^{th} pixel, and $\boldsymbol{Y}$ the column vector collecting all $\boldsymbol{y}_k$ (data cube). Notation Y_j designs the scalar value corresponding to the j^{th} component of $\boldsymbol{Y}$, that is, j indexes a single spatio-spectral coordinate. We consider the linear formulation:

$$\boldsymbol{Y} = \mathbf{H}_C \boldsymbol{X} + \boldsymbol{\epsilon}, \tag{2}$$

where $\boldsymbol{Y}$ collects the observations, $\boldsymbol{X}$ collects the data to be restored, $\mathbf{H}_C$ represents the instrument's PSF for the whole cube and $\boldsymbol{\epsilon}$ is an error term, accounting for noise and model errors. In the following, subscript C (for cube) characterizes matrices operating on spatial-spectral data such as $\mathbf{Y}$, whereas no subscript will be used for matrices operating only on spectra such as $\boldsymbol{y}_k$. Column $\boldsymbol{h}_j$ of matrix $\mathbf{H}_C$ is the instrument response to an impulsion located at the spatio-spectral coordinate j. That is, coefficients in $\boldsymbol{h}_j$ describe how the energy concentrated at coordinate j is spread over neighbor spatial and spectral elements in the cube. Elements of $\mathbf{H}_C$ are obtained by discretizing the PSF (1) at MUSE' spatial and spectral resolutions and mapping the results into corresponding coordinates in the vectorized formulation. Thanks to the PSF separability (1), matrix $\mathbf{H}_C$ can be written $\mathbf{H}_C = \mathbf{L}_C \mathbf{F}_C$, where $\mathbf{L}_C$ and $\mathbf{F}_C$ are related to the LSF and FSF, respectively. Moreover, due to the ordering of data in vector $\boldsymbol{Y}$, the LSF matrix $\mathbf{L}_C$ is block-diagonal, each bloc $\mathbf{L}_k$ corresponding to the LSF for the k^{th} pixel. Note that both FSF and LSF are very localized in space and wavelength compared to the respective dimensions of the data – current models use a 7×7 pixel FSF and 11 point LSF – so matrices $\mathbf{L}_C$, $\mathbf{F}_C$ and $\mathbf{H}_C$ are very sparse. These matrices are approximately square, depending on the boundary assumptions that are made on the "convolutive" model.

The least squares solution of (2) is known to have a poor behavior for noisy data, due to the bad conditioning of $\mathbf{H}_C$, which generates uncontrolled noise amplification. Efficient restoration then relies on the incorporation of appropriate prior knowledge in order to constrain the solution [2].

[1]For the considered integration times (from few minutes to one hour), collected light flux is high enough for the Poisson statistics of the data to be approximated by additive Gaussian noise.

3. REGULARIZATION WITH SPARSITY-BASED PRIOR SPECTRAL INFORMATION

3.1. Motivations for spectral sparsity constraints

Some comments are now necessary in order to build appropriate prior information. Let us remark that MUSE data are expected to show more informational coherence spectrally than spatially. Indeed, data observed at a given pixel correspond to the spectrum of the part of the source intercepted by the corresponding pixel (or to a mix of several spectra in case of object overlapping). Consequently, relevant prior information can be designed in the spectral domain, according to the spectral energy distribution related to the astrophysical processes [3]. On the contrary, the spatial domain shows much more diversity, the field of view being composed of thousands of objects with different sizes, shapes and brightness profiles.

We consider spectral prior information through sparsity constraints: any spectrum is approximately synthesized by a few number of *atoms* chosen in an overcomplete *dictionary*. Reasons for using sparsity-based models here are threefold.

- Prior information is efficiently expressed by means of an adapted dictionary, driven by physical knowledge.
- Restoration is coupled to the *detection* of relevant spectral features in the data, identified as the active atoms.
- Sparsity implicitly operates dimension reduction, which will be used in our three-dimensional restoration scheme for computational efficiency.

3.2. Spectral dictionary with physically motivated atoms

In a recent work [4], we considered the design of a dictionary $\mathbf{D}$ of elementary spectral features, so that any galaxy spectrum $\boldsymbol{x}$ is modeled as the linear combination of a few atoms of $\mathbf{D}$: $\boldsymbol{x} \simeq \mathbf{D}\boldsymbol{u}$ where $\boldsymbol{u}$ is sparse, that is, most coefficients in $\boldsymbol{u}$ are zero. More precisely, $\boldsymbol{x}$ is decomposed into three components, each of which is supposed to have a sparse decomposition in a specific dictionary: a line spectrum, a continuum and a series of breaks in the continuum. The line dictionary is the biggest one. It is composed of discretized splines, centered along the wavelength axis, and with 11 different widths. Widths range from 0.13 nm, which is the instrument spectral resolution (in this case, lines are delta functions) to 138 nm, which is considered as the maximum possible linewidth. Positive or negative decomposition coefficients here characterize emission or absorption lines, respectively. The continuum dictionary is composed of low-frequency sine waves with different phase shifts. Finally, the break dictionary is composed of step functions, where steps are centered along the wavelength axis. The resulting dictionary is highly redundant, with approximately 8 times more atoms than data points. Such an increased dimension is necessary to efficiently model spectral features with enough precision, in particular for spectral lines. Recall, however, that a sparse solution is searched in this dictionary, so that most decomposition coefficients are zero.

3.3. Regularized restoration

We now consider the restoration problem (2), where prior spectral sparsity constraints of § 3.2 are incorporated. Let K the number of

pixels in $\boldsymbol{X}$. We suppose that each spectrum in $\boldsymbol{X}$ satisfies $\boldsymbol{x}_k = \mathbf{D}\boldsymbol{u}_k$, where $\boldsymbol{u}_k$ is sparse. If $\boldsymbol{U}$ denotes the sparse representation of the cube collecting all $\boldsymbol{u}_k$, then $\boldsymbol{X} = \mathbf{D}_C\boldsymbol{U}$, where $\mathbf{D}_C$ is a block-diagonal matrix composed of K blocks $\mathbf{D}$. Then, our problem reads:

$$\text{find sparse } \boldsymbol{U} \text{ fitting } \boldsymbol{Y} = \mathbf{H}_C\mathbf{D}_C\boldsymbol{U} + \boldsymbol{\epsilon}.$$

In this paper, we consider the *Basis Pursuit De-Noising* approach [5] for such problem, which can be written:

$$\widehat{\boldsymbol{U}} = \arg\min_{\boldsymbol{U}} \|\boldsymbol{Y} - \mathbf{H}_C\mathbf{D}_C\boldsymbol{U}\|^2_{\boldsymbol{\Sigma}} + \gamma \|\boldsymbol{U}\|_1 , \qquad (3)$$

where the ℓ^1-norm $\|\boldsymbol{U}\|_1 = \sum_j |U_j|$ enforces sparsity and the data misfit term is weighted by the appropriate noise covariance matrix:

$$\|\boldsymbol{Y} - \mathbf{H}_C\mathbf{D}_C\boldsymbol{U}\|^2_{\boldsymbol{\Sigma}} = (\boldsymbol{Y} - \mathbf{H}_C\mathbf{D}_C\boldsymbol{U})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y} - \mathbf{H}_C\mathbf{D}_C\boldsymbol{U}).$$

Such weighted norm corresponds to the neg-log-likelihood of model (2) for centered Gaussian noise $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ [2].

4. A TWO-STEP RESTORATION PROCEDURE

4.1. Necessity for preliminary reduction of dimensions

Restoration of the full data cube by (3) is obviously intractable in practice. Since the objects of interest and the FSF are very small compared to the whole field of view, one can consider sub-cubes with much smaller spatial dimensions (typically, a few tens or hundreds of pixels). However, even for a small data cube with 25×25 pixels, $\boldsymbol{Y}$ has approx. $2.5\,10^6$ data and $\boldsymbol{U}$ has approx. $2\,10^7$ unknowns, which are still too high values for efficient implementation of (3). Hence, we consider the following two-step procedure :

i) First, for each spectrum $\boldsymbol{y}_k$, a sparse approximation in $\mathbf{D}$ is estimated, which yields a set of selected atoms, say Ω_k.

ii) Then, problem (3) is addressed by restricting the support of each unknown $\boldsymbol{u}_k$ composing $\boldsymbol{U}$ to the set of atoms that were selected *at least* in one spectrum at step i), that is $\Omega = \cup_k\Omega_k$.

This is a suboptimal scheme in the sense that it is not guaranteed to give the same results as (3). However, it substantially reduces the problem size so that the solution can be computed in practice. Typically, in the former example, if 100 atoms are selected in Ω, the number of unknowns decreases to approx. $6\,10^4$. These two steps are now described in § 4.2 and § 4.3.

4.2. Sparse approximation of individual spectra

Thanks to the PSF separability, initial model (2) reads, with notations of § 2.3: $\boldsymbol{Y} = \mathbf{L}_C\mathbf{F}_C\boldsymbol{X} + \boldsymbol{\epsilon} = \mathbf{L}_C\boldsymbol{V} + \boldsymbol{\epsilon}$, where $\boldsymbol{V} = \mathbf{F}_C\boldsymbol{X}$ is the spatially spread version of $\boldsymbol{X}$. As matrix $\mathbf{L}_C$ is block-diagonal, one can equivalently consider independently each pixel spectrum $\boldsymbol{y}_k = \mathbf{L}_k\boldsymbol{v}_k + \boldsymbol{\epsilon}_k$. Therefore, we propose in a first step to compute a sparse approximation of each spatially spread spectrum $\boldsymbol{v}_k$ by:

$$\widehat{\boldsymbol{u}}_k = \arg\min_{\boldsymbol{u}_k} \|\boldsymbol{y}_k - \mathbf{L}_k\mathbf{D}\boldsymbol{u}_k\|^2_{\boldsymbol{\Sigma}_k} + \gamma \|\boldsymbol{u}_k\|_1 , \qquad (4)$$

where such "one-dimensional" problem is computationally reasonable [4]. Note that a spectral feature present in "true" data $\boldsymbol{X}$ also appears in $\boldsymbol{V}$ but with lower amplitude, due to spatial spreading. Consequently, such individual spectrum restoration must be performed with a lower regularization parameter γ than for (3), so that most of the atoms actually present in $\boldsymbol{X}$ are detected, even if it results in a high number of false detections. This is of little consequence, as the aim of this step is dimension reduction and not the restoration of $\boldsymbol{X}$.

4.3. Spatial-spectral restoration in reduced spectral dimension

The former estimation procedure is applied to all spectra $\boldsymbol{y}_k$ in data cube $\boldsymbol{Y}$. This yields K sparse estimates $\widehat{\boldsymbol{u}}_k$. We consider the set of atoms that are detected at least in one $\widehat{\boldsymbol{u}}_k$. Let $\mathbf{D}^\Omega$ be composed of the columns of $\mathbf{D}$ indexed by $\Omega = \cup_k \text{supp}\,(\widehat{\boldsymbol{u}}_k)$, where $\text{supp}\,(\widehat{\boldsymbol{u}}_k)$ denotes the support (the indices of non-zero components) of $\widehat{\boldsymbol{u}}_k$. Then, let $\mathbf{D}^\Omega_C$ the block-diagonal matrix composed of K blocks $\mathbf{D}^\Omega$. We now search a sparse approximation of $\boldsymbol{X}$ in dictionary $\mathbf{D}^\Omega_C$, under observation model (2):

$$\widehat{\boldsymbol{U}}^\Omega = \arg\min_{\boldsymbol{U}^\Omega} \left\|\boldsymbol{Y} - \mathbf{H}_C\mathbf{D}^\Omega_C\boldsymbol{U}^\Omega\right\|^2_{\boldsymbol{\Sigma}} + \gamma \left\|\boldsymbol{U}^\Omega\right\|_1 . \qquad (5)$$

Both (3) and (5) formulate spatial-spectral restoration, but the number of unknowns in (5) is highly reduced. In particular, $\mathbf{D}^\Omega_C$ has often more rows than columns in practice. Hence, it is not a *redundant* dictionary. A sparse solution is still searched, however, which aims at removing the false alarms obtained at the first step (see § 4.2).

The ℓ^1-norm penalization term in (5) brings the searched sparsity, but it also introduces bias on non-zero amplitudes. Amplitudes are subsequently re-estimated by minimizing the least-squares term in (5), restricted to the non-zero components of $\widehat{\boldsymbol{U}}^\Omega$. Finally, the restored data are computed by $\widehat{\boldsymbol{X}} = \mathbf{D}^\Omega_C\widehat{\boldsymbol{U}}^\Omega$.

4.4. Interpretation of the solution

The solution obtained with the two former steps is a suboptimal approximation of (3) with reduced spectral dimension, which however preserves a rather flexible approach. In particular:

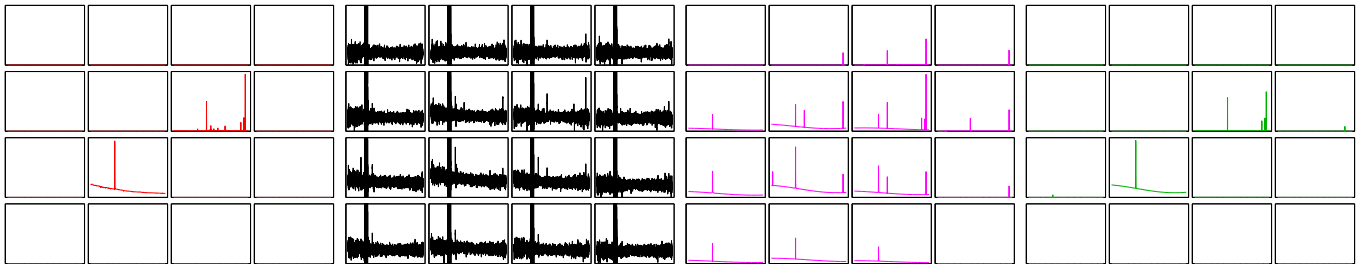
- *for unresolved sources*: suppose that a spectral feature is truly present at only one pixel and is detected at several ones because of spatial spreading. Then, the sparsity constraint at step ii) should favor the cancellation of such false detections;
- *for resolved sources*: on the contrary, if a spectral feature is truly present at several pixels and if it is detected at step i) only at the brightest one because of noise, then the "three-dimensional" observation model included in step ii) should favor its detection on these pixels.

From this point of view, such approach seems more appropriate to our problem than joint spatial-spectral regularization based on *structured sparsity* (see e.g. [6]). It is also computationally less expensive.

The main limitation of our approach holds in that spectral features should be detected at least in one spectrum processed at step i), whereas the full spatial-spectral approach (3) would probably allow the detection of fainter features. This is why step i) is performed with a deliberately high number of false alarms, in order to favor the detection of faint spectral features. Then, the sparsity constraint in the restricted support at step ii) aims at removing such false detections.

4.5. Implementation details

Boundary assumptions and matrix dimensions Let N_r and N_c be the respective numbers of rows and columns in the considered spatial window, $N_{\text{pix}} = N_rN_c$ and N_λ the number of wavelengths. Suppose that the FSF and the LSF have a support of $(F_r + 1) \times (F_c + 1)$ and $L + 1$ points, respectively. Since we use small spatial windows, in order to limit boundary degradations, we consider the most complete "convolutive" model where $\boldsymbol{X}$ collects all elements in the cube contributing to data $\boldsymbol{Y}$. It is then composed of $M_{\text{pix}} = (N_r + F_r) \times (N_c + F_c)$ pixels and of $M_\lambda = N_\lambda + L$ wavelengths. Typical support sizes of MUSE' FSF and LSF are 7×7 and



a) original spectra ($\mathbf{X}$) b) "convolved" and noisy data ($\mathbf{Y}$) c) individual spectral restoration ($\mathbf{D}\widehat{\boldsymbol{u}}_k$) d) spatial-spectral approach ($\widehat{\boldsymbol{X}} = \mathbf{D}_c^{\Omega}\widehat{\boldsymbol{U}}^{\Omega}$)

Fig. 2. Simulated data with 4×4 pixels $\times$ 3 463 wavelengths and restoration results.

11, respectively. Current MUSE simulations consider $N_\lambda = 3\,463$, for which dictionary $\mathbf{D}$ has $M_\lambda = 3\,473$ lines and $K_\lambda = 26\,140$ columns.

Matrices normalization The data misfit term in the above optimization problems can be written in a generic formulation as:

$$\|\boldsymbol{y} - \mathbf{H}\mathbf{D}\boldsymbol{u}\|_{\boldsymbol{\Sigma}}^2 = \left\|\boldsymbol{\Sigma}^{-1/2}\boldsymbol{y} - \boldsymbol{\Sigma}^{-1/2}\mathbf{H}\mathbf{D}\boldsymbol{u}\right\|^2$$

where diagonal matrix $\boldsymbol{\Sigma}$ has very different elements, corresponding to the variances plotted in Figure 1. It was shown that the columns of equivalent dictionary $\boldsymbol{\Sigma}^{-1/2}\mathbf{H}\mathbf{D}$ have very different norms, so that normalization is necessary for obtaining identical false alarm rates on all coefficients [4]. Consequently, the proper data misfit terms that are considered in Equations (4) and (5) are of the form $\left\|\boldsymbol{\Sigma}^{-1/2}\boldsymbol{y} - \overline{\mathbf{D}}\overline{\boldsymbol{u}}\right\|^2$, where $\overline{\mathbf{D}}$ is the normalized-column version of $\boldsymbol{\Sigma}^{-1/2}\mathbf{H}\mathbf{D}$ and $\overline{\boldsymbol{u}}$ is obtained by appropriately weighting $\boldsymbol{u}$ by the corresponding column norms.

Optimization We use an Iterative Coordinate Descent (ICD) algorithm for solving (4) and (5), which consists in iterating componentwise minimizations until convergence. Such a strategy has been shown to efficiently optimize sparsity-enhancing functionals [4] when no fast transform can be used to compute matrix-vector products (our case). Note that whereas the dictionary size in (4) is $N_\lambda \times K_\lambda$, it yields $(N_{\text{pix}}N_\lambda) \times (M_{\text{pix}}\text{Card}(\Omega))$ in (5). Hence, the dictionary in (5) cannot be stored in practice. An ICD procedure is then particularly suited to such a problem, since every scalar optimization only requires access to the corresponding column of the involved matrix. Hence, each column is computed as often as necessary, at a low cost since only a few of its elements are non-zero.

5. EXPERIMENTAL RESULTS

For visualization purposes, a scene with reduced (4×4) spatial dimension is simulated, composed of two close unresolved sources, see Figure 2a). The associated spectra result from astrophysical simulations, in particular they are not built from the dictionary detailed in § 3.2 that is used for reconstruction. The scene was convolved by a 3×3 pixel, circular Gaussian FSF with 2 pixel FWHM (full width at half maximum), and by an 11 point Gaussian LSF with FWHM equal to 2 spectral resolution elements. Gaussian noise is added with the variance spectral distribution of Figure 1, giving the simulated observed data in Figure 2b). In the following, pixel coordinates are given under the form (row number, column number).

Our estimation scheme was applied to the data of Figure 2b). For the first step (see § 4.2), parameter γ was tuned empirically so as to preserve the detection of the three spectral lines at the very right extremity of the spectrum at (2,3), while minimizing false detections. Results are given in Figure 2c), showing false detections of spectral lines for most pixels. In particular, both restored spectra at (2,3) and (3,2) are contaminated by the spatial spreading of lines from the other pixel. The number of detected components here is $\text{Card}\,\Omega = 33$, so that problem (5) has 1 188 unknowns, instead of $9.4\,10^5$ for problem (3). Spatial-spectral reconstruction operating in such reduced dimension (see § 4.3) yields the results plotted in Figure 2d) – we only show the 16 central pixels of the 6×6 restored scene. Interferences in Figure 2c) are rather well canceled by taking into account spatial spreading. Similarly, information is well re-concentrated at the two "active" pixels, since the restoration at other pixels is almost void of any detection.

Such an approach yielded encouraging results in terms of close objects separation and indirect spectral unmixing, at a reasonable computational cost. Several applications of the proposed methodology obviously deserve more attention. In particular, quantitative performance analysis in terms of correct detections vs. false alarms should be performed. In the context of MUSE data, further work also includes the use of spectral sparsity based dimension reduction for addressing more specific problems, such as the restoration of both, possibly mixed, unresolved and extended sources.

6. REFERENCES

[1] R. Bacon *et al.*, "Probing unexplored territories with MUSE: a second generation instrument for the VLT," in *Proc. SPIE*, July 2006, vol. 6269 of *Ground-based and Airborne Instrumentation for Astronomy*.

[2] J. Idier, Ed., *Bayesian Approach to Inverse Problems*, ISTE Ltd and John Wiley & Sons Inc, Apr. 2008.

[3] J. Tennyson, *Astronomical Spectroscopy*, Imperial College Press, Aug. 2005.

[4] S. Bourguignon, D. Mary, and É. Slezak, "Sparsity-based restoration of astrophysical spectra: models and algorithms," *accepted for publication in IEEE J. Selected Topics Sig. Proc*, Jan. 2011.

[5] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.

[6] M. Kowalski and B. Torrésani, "Sparsity and persistence: mixed norms provide simple signal models with dependent coefficients," *Signal, Image and Video Processing*, vol. 3, no. 3, pp. 251–264, 2009.