Spatially Constrained Online Dictionary Learning for Source Separation

Argheesh Bhanot, Céline Meillier, Fabrice Heitz, Laura Harsan

Spatially Constrained Online Dictionary Learning for Source Separation

Argheesh Bhanot, Céline Meillier, Fabrice Heitz and Laura Harsan

Abstract

Whether in medical imaging, astronomy or remote sensing, the data are increasingly complex. In addition to the spatial dimension, the data may contain temporal or spectral information that characterises the different sources present in the image. The compromise between spatial resolution and temporal/spectral resolution is often at the expense of spatial resolution, resulting in a potentially large mixing of sources in the same pixel/voxel. Source separation methods must incorporate spatial information to estimate the contribution and signature of each source in the image. We consider the particular case where the position of the sources is approximately known thanks to external information that may come from another imaging modality or from *a priori* knowledge. We propose a spatially constrained dictionary learning source separation algorithm that uses *e.g.* high resolution segmentation map or regions of interest defined by an expert to regularise the source contribution estimation. The originality of the proposed model is the replacement of the sparsity constraint classically expressed in the form of an ℓ_1 penalty on the localisation of sources by an indicator function exploiting the external source localisation information. The model is easily adaptable to different applications by adding or modifying the constraints on the sources properties in the optimisation problem. The performance of this algorithm has been validated on synthetic and quasi-real data, before being applied on real data previously analysed by other methods of the literature in order to compare the results. To illustrate the potential of the approach, different applications have been considered, from scintigraphic data to astronomy or fMRI data.

Index Terms

Source separation, dictionary learning, fMRI, scintigraphy, hyperspectral data.

A. Bhanot, C. Meiller and F. Heitz are with IMAGeS team at ICube, UMR 7357, Université de Strasbourg, CNRS. e-mail: (bhanot@unistra.fr; meillier@unistra.fr; fabrice.heitz@unistra.fr)

L. Harsan is with IMIS team at ICube, UMR 7357, Université de Strasbourg, CNRS. e-mail: (harsan@unistra.fr).

I. INTRODUCTION

The issue of source separation, or unmixing, is well known to the signal and image processing community. It concerns a very large number of applications and can occur under different conditions of source mixing. A large part of the literature is devoted to blind source separation (BSS) [1]–[3]. BSS methods allow to solve *cocktail party* problems for which P signals (or images) composed of a mixture of R sources are observed, without any *a priori* on the properties of the sources. If the number of sources involved in the mixtures is not known *a priori*, then they must also be estimated [4], [5]. Early BSS methods mainly comprised Independent Component Analysis (ICA) and sparse decomposition analysis. Many variants of the ICA approach have been proposed in the literature to solve BSS problems. All of them are based on the general principle of spatial independence of the sources, which makes it possible to estimate their temporal (or spectral) signatures.

For instance, in brain functional networks detection in functional Magnetic Resonance Imaging (fMRI) data, ICA is widely used to separate spatial sources by assuming the independence of the temporal signals associated with each spatial source, *i.e.* functional network. Spatial ICA has proven effective in [6], [7] for fMRI data, but the main drawback of ICA approaches is the unknown number of sources which is set arbitrarily and may lead to a large number of nuisance sources, that must be screened manually or by a semi-automatic method [8].

In contrast to the BSS problem, many unmixing problems involve a dictionary of pre-defined sources. For example, in hyperspectral imaging for remote sensing, libraries of light spectra corresponding to the different materials that may be observed in the scene are available, so only the proportion of the different materials in each pixel is estimated. Between these two extreme cases, there are a large number of unmixing problems where some information on the form or location of the sources or the type of mixture is known. *Sum-to-one* and *positivity* constraints on the coefficients of the mixing matrix are classic in signal and image processing. In remote sensing applications, hyperspectral data linear unmixing is for instance carried out by methods based on nonnegative matrix factorization [9], [10]. In recent years, sparse decomposition methods have been widely used to solve source separation problems. The sparsity constraint is another way to reduce the set of solutions. It can be combined with the two latter constraints. The sparsity may concern the mixing itself, *i.e.* for a given observed signal, only a few number of sources is involved, or the decomposition of the sources on a dictionary (wavelet, discrete cosine transform, or custom atoms) [11]. Recently for fMRI application, sparse analysis based on dictionary learning methods have proven to be promising [12]–[15]. Dictionary learning methods take into account the spatial sparsity of the functional networks in the form of ℓ_1 constraints on the mixing matrix in the minimisation problem.

In this paper, we are interested in the problem of source separation in spatially structured data: 2D or 3D images that contain temporal information (fMRI, scintigraphy) or light spectrum information (hyperspectral imaging). We consider two categories of this kind of unmixing problems. In the first one, for a given pixel/voxel, different sources contribute to the mix in the sense that the spatial resolution is not fine enough to allow spatial separation of the sources. This is the case, for example, with unmixing problems in remote sensing [10] or fMRI applications [15], [16]. In this case, the mixing matrix is a matrix of proportions where for a given pixel, the sum of the contributions of each source is equal to 1. In the second category of unmixing problems, the mixing is additive, the signals of the different sources are superimposed and their sum forms the observed mixing signal. Decomposition of scintigraphic image sequences into tissue images and their time-activity curves, or unmixing of light sources in hyperspectral data in astronomy are examples of this second category of problems. In this case, the *sum-to-one* constraint is not relevant, the coefficients of the mixing matrix are the intensity of the contribution of each source in the mixture. Since the observed signals are observed in the form of images, the constraints that can be defined in the optimisation problem should be related to the location of the sources and not to their shape. We consider the case where no information on the temporal or spectral signature of the different sources, or on their dependence is available.

We propose a dictionary learning method that introduces sparsity constraints on the spatial localisation of sources from external knowledge. Additional constraints on the mixing matrix (positivity and *sum-to-one* constraints) can be added or removed depending on the application. To illustrate the potential of the approach, different applications have been considered, from scintigraphic data to astronomy or fMRI data. All these data are of very different natures, as well as the *a priori* information available on the location of the sources. We thus show that our algorithm is adaptable to different types of data and different types of *a priori* knowledge on the location of sources. In the case of multimodal observations, information regarding the possible spatial location of sources is usually derived from a high spatial resolution image that does not provide the second dimension, namely temporal or spectral information. Unlike multimodal image fusion problems, such as pansharpening [17], our goal is not to produce a spatially and spectrally or temporally well-resolved image. We rather aim at exploiting a segmentation information from a high spatial resolution image in order to improve the unmixing of spectral or temporal sources at a lower resolution image level. In some applications, such a segmentation map is not available, but approximate spatial location information can be provided by an expert who can define regions of interest (ROIs) (see for example the unmixing method for highly realistic simulated renography dataset in [18], [19]).

This paper is organised as follow: section II presents the observation model and the dictionary learning approach for unmixing spatially structured data based on the spatial sparsity constraint. The performance

and adaptability of the algorithm to different applications are illustrated on synthetic data in section III and on real data in section IV. Comparison of our approach with application-specific state-of-the-art methods are proposed on three different applications: unmixing in scintigraphic image sequences, fMRI and hyperspectral astronomic datacubes.

II. Model and method

The classical linear model used in source separation may be written as:

$$\mathbf{Y} \simeq \mathbf{U}\mathbf{A}, \tag{1}$$

where $\mathbf{Y} \in \mathbb{R}^{N\times P}$ is the observed data. In spatio-spectral separation, N can be interpreted as the spectral length and in spatio-temporal separation N is the length of the temporal signals. P is the number of voxels or pixels, depending upon the dataset. Matrix $\mathbf{U} \in \mathbb{R}^{N\times R}$ contains the temporal/spectral signatures where R is the number of sources. Matrix $\mathbf{A} \in \mathbb{R}^{R\times P}$, usually called mixing or abundance matrix, codes the fraction of the R components contributions at each voxel or pixel.

A. Constrained optimisation formulation

Given the observation model (1), the following minimisation problem:

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 \tag{2}$$

does not have a unique solution because of the joint estimation of $\mathbf{A}$ and $\mathbf{U}$, and the ill-posedness of the problem. In order to restrain the number of solutions, we introduce some standard constraints on matrix $\mathbf{A}$ such as the *positivity constraint* $\mathbf{A} \in \mathbb{R}^+$ and the *sum-to-one constraint* $\sum \mathbf{A}[.,i] = 1$, with $i = 1:P$, as these are the proportions voxel by voxel. The form of matrix $\mathbf{A}$ may also be constrained by some extra information from high resolution (HR) segmentations of sources or source locations knowledge: we know which ROIs may contribute to a given voxel, i.e. present a non-zero proportion at this voxel. When the total number of sources R is high, this *a priori* knowledge allows to constrain the possible solutions of the minimisation problem. In the most general setting, the unmixing problem is recast as:

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2 + \mathcal{I}_{\mathbb{R}^+}(\mathbf{A}) + \mathcal{I}_S(\mathbf{A}) + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}), \tag{3}$$

where the first term is the data fidelity term, the second term is a Tikhonov regularisation controlled by parameter μ_σ set to 10^{-4} to prevent bad conditioning (see section II-B). The third term is a positivity constraint where $\mathcal{I}_{\mathbb{R}^+}(\mathbf{A}) = \infty$ if at least one of the elements of $\mathbf{A}$ is negative, and 0 otherwise. The fourth term in eq. (3) codes an optional *sum-to-one* constraint on each column of matrix $\mathbf{A}$, $\mathcal{I}_S(\mathbf{A}) = \infty$ if at least one column of $\mathbf{A}$ does not sum to one, and 0 otherwise. The last term $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A})$ is the indicator

function on the set $M(\tilde{\mathbf{A}})$ of matrices having a structure similar to a given binary "structure matrix" $\tilde{\mathbf{A}}$, *i.e.* $\mathbf{A} \in M(\tilde{\mathbf{A}})$ if and only if $\mathbf{A} \in \mathbb{R}^{R\times P}$ and coefficient $\mathbf{A}_{i,j} = 0$ if $\tilde{\mathbf{A}}_{i,j} = 0$. $\tilde{\mathbf{A}}$ is a binary matrix, where element $(\tilde{\mathbf{A}})_{r,i} = 1$ if, according to a priori knowledge about spatial localisation of the sources, the r^{th} region of interest could exist in the i^{th} voxel, and 0 otherwise. This results in $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}) = \infty$ if at least one element of $\mathbf{A}$ is non-zero while it is zero in $\tilde{\mathbf{A}}$, and 0 otherwise.

Combining sets $\mathbb{R}^+ \cap S = S^+$, the optimisation problem can be rewritten as follows:

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2 + \mathcal{I}_{S^+}(\mathbf{A}) + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}). \tag{4}$$

Estimating jointly $\mathbf{U}$ and $\mathbf{A}$ in eq. (4) is a typical problem of dictionary learning (DL). But, unlike conventional DL algorithms, there is no sparsity regularisation term in the form of an ℓ_1 penalty: it is the sources localisation information coded in the structural term $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A})$ which enforces the sparse decomposition of each voxel. A classical way to solve the joint estimation problem is to optimise alternatively the cost function eq. (3) along $\mathbf{U}$ and $\mathbf{A}$.

B. Estimation of the temporal / spectral signatures matrix $\mathbf{U}$

Considering that $\mathbf{A}$ is fixed, problem (3) becomes:

$$\min_{\mathbf{U}} \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\ \|_F^2 + \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2. \tag{5}$$

The Tikhonov regularisation term $\frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2$ is introduced to improve the conditioning of problem (2), μ_σ is set to 10^{-4} to prevent collinearity between columns of $\mathbf{U}$.

The solution of (5) is the ridge estimator defined by:

$$\hat{\mathbf{U}} = \mathbf{Y}\mathbf{A}^T(\mathbf{A}\mathbf{A}^T + \mu_\sigma \mathbf{I}_R)^{-1}, \tag{6}$$

where $\mathbf{I}_R$ is the $R \times R$ identity matrix.

C. Estimation of the abundance / mixing matrix $\mathbf{A}$

Consider that $\mathbf{U}$ is fixed, then problem eq. (3) becomes $\min_{\mathbf{A}} f(\mathbf{A})$, where:

$$f(\mathbf{A}) = \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \mathcal{I}_{S^+}(\mathbf{A}) + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}). \tag{7}$$

Note that this function is separable according to the pixels/voxels $i \in \{1,\ldots, \mathrm{P}\}$, which leads to: $\min_{\mathbf{a}_i} f(\mathbf{a}_i)$, with:

$$f(\mathbf{a}_i) = \frac{1}{2}\|\mathbf{y}_i - \mathbf{U}\mathbf{a}_i\|_F^2 + \mathcal{I}_{S^+}(\mathbf{a}_i) + \mathcal{I}_{M(\tilde{\mathbf{a}}_i)}(\mathbf{a}_i), \tag{8}$$

where $\mathbf{a}_i$ is a column vector from the matrix $\mathbf{A}$ (and with an abuse of notation for $f(.)$). The set of all the vectors with a structure similar to $\mathbf{a}_i$ is given by $\tilde{\mathbf{a}}_i$, where $\tilde{\mathbf{a}}_i$ is a column of $\tilde{\mathbf{A}}$. The regularisation terms in eq. (8) can be summarised as:

$$g(\mathbf{a}_i) = \mathcal{I}_{M(\tilde{\mathbf{a}}) \cap S^+}(\mathbf{a}_i). \tag{9}$$

Note that the objective function eq. (8) is convex since the first term is convex and differentiable and $g(a_i)$ is convex but non differentiable. The proof of convexity for $g(\mathbf{a}_i)$ is given in Appendix A. Minimisation of the objective function given by eq. (8) belongs to the class of problems on which the proximal gradient methods can be applied. Different algorithms are available, for example, alternating direction method of multipliers (ADMM) [20], projected gradient, also known as iterative shrinkage-thresholding algorithm (ISTA) or FISTA (Fast ISTA) [21]. Algorithm FISTA was preferred for its rapid convergence: its implementation is given in Algorithm 1.

Initialisation of $\mathbf{A}^{(0)}, \omega^{(1)}$
while *STOPPING CRITERIA* $\neq$ *TRUE* **do**
  Minimisation problem w.r.t $\mathbf{U}$
  $\mathbf{U}^{(l+1)} = \mathbf{Y}\mathbf{A}^{(l)^T}(\mathbf{A}^{(l)}\mathbf{A}^{(l)^T} + \mu_\sigma \mathbf{I}_R)^{-1}$
  Parallel minimisation w.r.t. the columns a_i of $\mathbf{A}$ by employing FISTA [21]
  for $k \leftarrow 1$ **to** $proxsteps$ **do**
    $\mathbf{a}_i^{(k)} = prox_g(\boldsymbol{\omega}^{(k)} - \lambda \nabla f(\boldsymbol{\omega}^{(k)}))$
    $t^{(k+1)} = \frac{1+\sqrt{1+4(t^{(k)})^2}}{2}$
    $\boldsymbol{\omega}^{(k+1)} = \mathbf{a}_i^{(k)} + \left(\frac{t^{(k)}-1}{t^{(k+1)}}\right)(\mathbf{a}_i^{(k)} - \mathbf{a}_i^{(k-1)})$
  end
  $\mathbf{a}_i^{(l+1)} = \mathbf{a}_i^{(proxsteps)}$
end
return $\mathbf{A}$, $\mathbf{U}$

Algorithm 1: Alternate optimisation algorithm to estimate $\mathbf{A}$ and $\mathbf{U}$

In Algorithm 1, $\nabla f(\mathbf{a}_i)$ is the gradient of $f(\mathbf{a}_i)$, given by $\mathbf{U}^T(\mathbf{U}\mathbf{a}_i - \mathbf{y}_i)$. The step size λ is set equal to the inverse of the Lipschitz constant of $\nabla f(\mathbf{a}_i)$ i.e. $1/L$, where $L = \|\mathbf{U}^T\mathbf{U}\|_F$. $t^{(k+1)}$ is an auxiliary variable which helps in the fast convergence of FISTA, $\boldsymbol{\omega}$ calculates intermediate values based on a special linear combination of the last two points and $prox$ refers to the proximal operator [21]. In our case, the proximal operator is just the projection of $\mathbf{a}_i$ in the positive orthant, with the vector normalised

to-sum-to-one. This projection also forces the elements of abundance matrix $(\mathbf{A})_{r,i}$ to be non-zero only at positions where the region of interest r projects on pixel/voxel i ($\mathcal{I}_{M(\tilde{\mathbf{a}})}$ constraint). The proximal operator of the function g is:

$$prox_g(\mathbf{y}) = \underset{\mathbf{x}\in M(\tilde{\mathbf{a}})\cap S^+}{\operatorname{argmin}} \|\mathbf{x}-\mathbf{y}\|_2 = \mathcal{P}_{M(\tilde{\mathbf{a}})\cap S^+}(\mathbf{y}), \tag{10}$$

where $\mathcal{P}$ is the projection operator on set $M(\tilde{\mathbf{a}}) \cap S^+$. The orthogonal projection of a vector $\mathbf{y} \in \mathbb{R}^R$ on $M(\tilde{\mathbf{a}}) \cap S^+$ is obtained using the projection onto convex sets (POCS) method [22]. POCS algorithm alternates projection onto the simplex $S^+ = \mathbb{R}^+ \cap S$ and projection onto the set $M(\tilde{\mathbf{a}})$ of vectors having the same structure as $\tilde{\mathbf{a}}_i$. Only a few iterations are required for convergence of the POCS algorithm.

Convergence towards a global minimum of DL algorithms cannot be proven. In practice, a good initialisation of $\mathbf{A}$ and the presence of pure pixels (as in remote sensing applications) in each region guarantee a good joint estimation of $\mathbf{U}$ and $\mathbf{A}$. Previous work [16] has demonstrated the importance of well-defining the spatial constraint on abundance $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A})$ to ensure an acceptable estimate of abundances and spectral or temporal signatures.

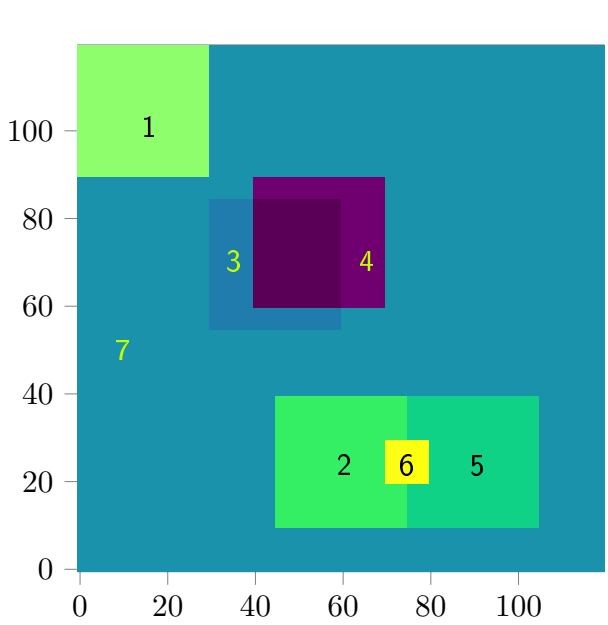
III. Evaluation on synthetic datasets

In this section we evaluate the unmixing performance of our algorithm on two different synthetic datasets. Dataset I was created to show unmixing of signals/spectra taking into account different situations that could occur in real applications such as fMRI or astronomical data unmixing. Dataset II is an example of realistic synthetic data in scintigraphy used in [18], for which the authors have proposed an unmixing method based on prior knowledge of the location of the regions of interest. This method, called Robust Unmixing of Dynamic Sequences Using Regions of Interest (RUDUR), has been compared in [18] to different region of interest-based algorithms commonly used in scintigraphy and has shown to be better at estimating spatial maps and temporal signals. The performance of our algorithm is compared to RUDUR in paragraph III-B.

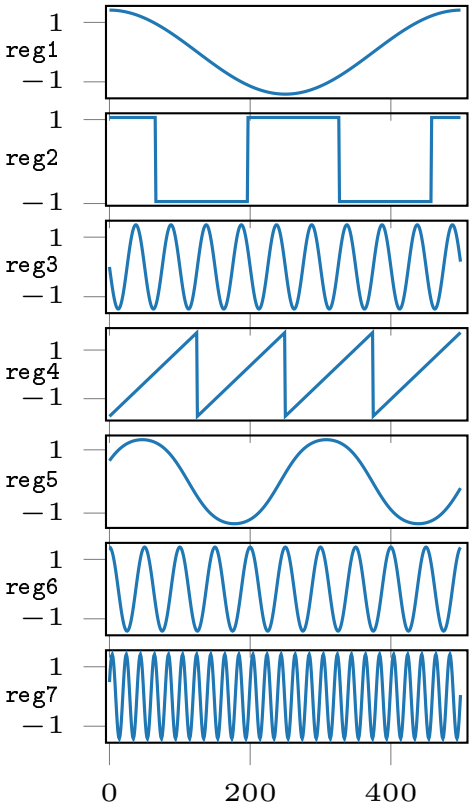
A. Dataset I

1) Data description: Unmixing algorithms are often sensitive to the assumption of pure pixels (i.e. each source or region has an abundance of 1 in at least one pixel of the image). To challenge this hypothesis, synthetic temporal data were simulated. Seven temporal signatures are mixed in a 120×120 pixels image. Ground truth signals and locations for the different regions are presented in Fig. 1. In Fig. 1a, we see a region 6 superimposed on two regions (2 and 5). Two other regions (3 and 4), partially covering each other, are included. Region 7 and region 1 are comprised of pixels not belonging to any

other region. Data was generated for different SNRs ranging from -20dB to 20dB with a zero mean Gaussian white noise.



(a) Localisation map for the 7 different sources. Region 6 is a difficult region, which overlaps region 2 and region 5. Region 7 and 1 do not intersect any other region. Regions 3 and 4 are partially overlapping.



(b) Nature of signals

Fig. 1: Localisation map and temporal signatures used to build synthetic data.

2) Algorithm details: The ground truth is given by the localisation map in Fig. 1a. To initialise $\mathbf{A}^{(0)}$, each region was dilated with a 7 pixels square structuring element and then the proportion for each region over each pixel was calculated, respecting the *sum-to-one* condition. The dilatation was done to introduce the uncertainty in the localisation of regions; as the localisation is seldom precise when dealing with real data. The algorithm used for unmixing is given in Algorithm 1. 400 steps were adopted for FISTA, in combination with 50 steps of alternate optimisation. The weighting parameter in the Tikhonov regularisation was set to 10^{-4} as no more smoothing was required. A standard normalisation was applied to the data before processing: $Y^{(i)} = \frac{Y^{(i)} - \mu^{(i)}}{\sigma_Y^{(i)}}$, where μ is the mean of the temporal signal of the i^{th} pixel and σ is the standard deviation of the timecourse of the i^{th} pixel.

3) Results and discussion: We observe that the timecourses and the abundances for the seven regions are well estimated even if the abundances are not perfectly initialised. The mean squared errors (MSE) for the estimation of the timecourses are displayed in table I, The MSE in estimating the abundances are given in table III. From table I and table III we see that for most of the regions the errors decrease as SNR increases. The spectral angle distances (SAD)(formula in Appendix B) given in table II follow a similar trend, proving the effectiveness of the unmixing method. The estimated timecourses for each region were normalised by standard deviation before calculating the SAD and MSE. We also generated synthetic data where region 6 was completely included in region 5. In this case, due to noise it was impossible to correctly estimate the timecourse of the region included in the other (and therefore its abundance).

	-20dB	-10dB	0dB	10dB	20dB
reg1	9.5e-02	1.0e-02	1.1e-03	1.1e-04	1.1e-05
reg2	1.3e-01	1.3e-02	1.2e-03	1.2e-04	1.3e-05
reg3	2.2e-01	2.1e-02	2.4e-03	3.2e-03	4.0e-03
reg4	2.3e-01	2.2e-02	2.5e-03	3.2e-03	4.0e-03
reg5	1.3e-01	1.3e-02	1.3e-03	1.3e-04	1.3e-05
reg6	1.4e+00	3.4e-01	4.2e-02	4.3e-03	6.3e-04
reg7	9.6e-03	9.6e-04	9.6e-05	9.6e-06	1.0e-06

TABLE I: Region wise mean squared errors for $\mathbf{U}$ for different SNRs

	-20dB	-10dB	0dB	10dB	20dB
reg1	1.8e+01	5.8e+00	1.9e+00	5.9e-01	1.9e-01
reg2	2.0e+01	6.5e+00	2.0e+00	6.4e-01	2.0e-01
reg3	2.7e+01	8.3e+00	2.8e+00	3.2e+00	3.6e+00
reg4	2.7e+01	8.6e+00	2.9e+00	3.2e+00	3.6e+00
reg5	2.0e+01	6.5e+00	2.1e+00	6.5e-01	2.1e-01
reg6	7.1e+01	3.4e+01	1.2e+01	3.7e+00	1.4e+00
reg7	5.6e+00	1.8e+00	5.6e-01	1.8e-01	5.7e-02

TABLE II: Region wise spectral angles for U for different SNRs

Fig. 2 illustrates the convergence of the algorithm for different SNRs. Fig. 2 plots $\frac{\|\mathbf{Y}-\mathbf{UA}\|_F}{\|\eta\|_F}$, as a function of the number of steps in the alternate minimisation and η denotes the white Gaussian noise

	-20dB	-10dB	0dB	10dB	20dB
reg1	5.8e-04	6.1e-05	4.7e-06	3.2e-07	2.8e-08
reg2	2.2e-03	2.1e-4	3.8e-05	1.0e-04	1.3e-04
reg3	3.2e-03	3.4e-04	4.9e-05	5.6e-05	7.0e-05
reg4	3.4e-03	3.8e-04	5.4e-05	5.9e-05	7.1e-05
reg5	2.1e-03	2.0e-04	4.6e-05	1.1e-04	1.3e-04
reg6	9.2e-04	8.9e-05	7.0e-05	2.1e-04	2.6e-04
reg7	5.4e-03	5.4e-04	6.3e-05	9.4e-05	1.2e-04

TABLE III: Region wise mean squared errors for A for different SNRs

present in the dataset at different SNRs. In the very first steps of the alternate optimisation we see that the curves decrease sharply and ultimately settle around a particular value when convergence is achieved. In Fig. 2 the -20dB and -10dB curves converge to values less than 1 since some noise remains in the estimated timecourses. The 0^{th} iteration shows the ratio $\frac{\|\mathbf{Y}-\mathbf{UA}\|_F}{\|\eta\|_F}$ calculated with the initial enlarged $\mathbf{A}$ and $\mathbf{U}$, estimated with least squares.

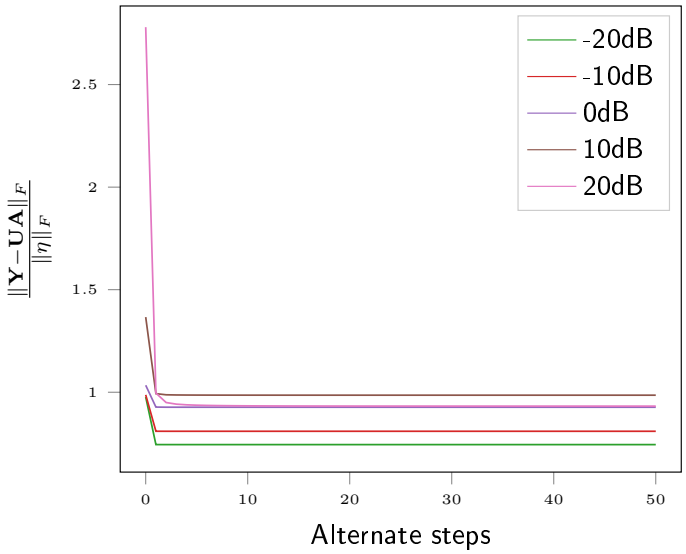


Fig. 2: Convergence curves plotted for different values of SNR for Dataset I. The curves represent $\frac{\|\mathbf{Y}-\mathbf{UA}\|_F}{\|\eta\|_F}$ plotted against the number of iterations.

B. Dataset II

In an effort to objectively evaluate the performances of our approach, we propose to test and compare our method on a more realistic synthetic data set of scintigraphy images created for the evaluation of the performances of the state-of-the-art RUDUR method [18]. We have reused the RUDUR code, as it is distributed by the authors [18].

1) Data description: This second dataset provides a first test case in which the method is confronted with a real application, on physical model-based simulations of scintigraphic images, with ground truth. It enables a comparison with a recent reference method in the field [18]. Region of interest-based unmixing methods are common in scintigraphy [5], [19], [23]. All standard methods are source separation methods. They estimate the time activity curves (TACs) and emissions of a tracer (a radioactive element) in the different body organs. A dataset of scintigraphic data has been made available at [24]. This dataset is based on a Monte Carlo simulation of scintillation camera imaging [25]. The datacube comprises images of size 21×26, with N = 60. The dataset, containing R = 3 regions, is shown in Fig. 3. In Fig. 3 the first row (a) shows the ground truths for the spatial maps. In the second row (b) the ROI initialisation is presented. This initialisation is not the same as in [18] as our method needs strictly greater initial ROIs than the regions which need to be unmixed.

2) Algorithm details: RUDUR algorithm has a soft constraint on the source locations which allows the regions to unmix data even if the ROIs selected lie in the interior of the actual locations of the regions. In our model, we use a hard constraint on the locations of the regions, through the regularization term $\mathcal{I}_{M(\tilde{\mathbf{A}})}$, so initial ROIs should be stricly enclosing the regions for which we want to estimate the timecourses. To achieve this, the binary mask of ROIs used in [18] have been dilated with a 5 pixels square structuring element.

We should note that this application corresponds to a an additive case of unmixing, so the sum-to-one constraint was dropped off in our algorithm. Further, as scintigraphy timecourses should be strictly positive (representing the emission of the tracer), eq. (3) was changed to:

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2 + \mathcal{I}_{\mathbb{R}^+}(\mathbf{A}) + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}) + \mathcal{I}_{\mathbb{R}^+}(\mathbf{U}).$$

Due to the addition of $\mathcal{I}_{\mathbb{R}^+}(\mathbf{U})$, the constraint of positivity on the TACs, the least squares given in step 4 of Algorithm 1 to solve for $\mathbf{U}$ had to be replaced by FISTA steps to estimate $\mathbf{U}$ in each alternate step. The initialisation of the algorithm was done with the help of least squares using the initial dilated ROIs.

As the sum-to-one condition is dropped, the proximal operator of the function g in Algorithm 1 changes to:

$$prox_g(\mathbf{y}) = \underset{\mathbf{x}\in M(\tilde{\mathbf{a}})\cap R^+}{\operatorname{argmin}} \|\mathbf{x}-\mathbf{y}\|_2 = \mathcal{P}_{\mathbb{R}^+\cap M(\tilde{\mathbf{a}})}(\mathbf{y}), \tag{11}$$

where $\mathcal{P}$ is now the projection operator on the set $\mathbb{R}^+ \cap M(\tilde{\mathbf{a}})$.

3) Results and discussion: We ran our algorithm with 500 steps of alternate optimisation. At each iteration l, convergence is monitored by the optimisation gain κ_l defined as:

$$\kappa_l \;=\; \frac{\|\mathbf{Y}-\mathbf{U}^{(l)}\mathbf{A}^{(l)}\|_F - \|\mathbf{Y}-\mathbf{U}^{(l-1)}\mathbf{A}^{(l-1)}\|_F}{\|\mathbf{Y}-\mathbf{U}^{(l-1)}\mathbf{A}^{(l-1)}\|_F}. \tag{12}$$

which decreases to 10^{-16} at the 500th alternate step. Estimated temporal signals and spatial maps were normalised by the criteria given in [18] for comparison with the provided ground truth. The normalised mean squared error (NMSE) and normalised mean absolute error (NMAE) for the estimated spatial maps and time activity curves were calculated. In addition, we also provide spectral angle distances (SAD) for the estimated time activity curves. The formulae for NMAE, NMSE and SAD are given in Appendix B. The quantitative results on the synthetic scintigraphy data are given in table IV. We observe that the signals estimated by our method are close to those obtained by RUDUR, with a better estimation by our algorithm for U2 and U3. The NMSE and NMAE for spatial maps calculated using the two methods are given in table V. The errors on the spatial maps were calculated by restricting the pixels of the estimated maps to the initial ROIs. Errors are lower in the case of RUDUR, for the spatial maps.

	U1	U2	U3
NMSE Alt opt	0.049	0.046	0.011
NMSE Rudur	0.031	0.047	0.015
NMAE Alt opt	0.173	0.142	0.091
NMAE Rudur	0.130	0.143	0.105
SAD Alt opt	12.181	12.049	6.128
SAD Rudur	9.561	12.362	6.633

TABLE IV: Errors (NMAE and NMSE) and spectral angle differences between the estimated timecourses and the ground truth for the scintigraphy dataset.

The spatial maps unmixed by our method are presented in Fig. 3 (c) and those estimated by RUDUR are displayed in Fig. 3 (d). Visually the maps are very similar although the errors are lower in the case of RUDUR. The last row(e) shows the TACs estimated by the proposed method (in blue). We observe

	A1	A2	A3
NMSE Alt opt	0.189	0.084	0.139
NMSE Rudur	0.147	0.064	0.070
NMAE Alt opt	0.433	0.304	0.419
NMAE Rudur	0.340	0.261	0.344

TABLE V: Errors between the estimated spatial maps for the different regions and the ground truth. The errors given here were calculated by restricting the pixels of the estimated spatial maps to the initial ROIs.

in Fig. 3 that our solution is near the target solution (in dashed black) for all the sources and is close to the solution provided by RUDUR.

IV. Application on semi-real and real data

In this section we discuss the application of our algorithm to sources unmixing of (semi-real) resting state (rs) 3D fMRI data and on a real dataset of hyperspectral astronomic images. In these two applications, in addition to the 3D + time or hyperspectral data, a high resolution segmentation map is available as a standard, with a registration on the low resolution images containing the sources to be separated.

A. Semi real rs-fMRI dataset

Resting state functional Magnetic Resonance Imaging (rs-fMRI) has been widely used for studying brain functional connectivity [26]. Rs-fMRI allows the observation of changes in cerebral activity by analysing the blood-oxygen-level-dependent (BOLD) signal [27]. At rest, only spontaneous activity is measured and a set of anatomical regions with the same fluctuations are considered part of a common network. We apply the algorithm on a single subject rs-fMRI. The motivation to work on single subject has been detailed in [16]. Benefit of integrating a high resolution (HR) anatomical atlas in the single-subject case has also been demonstrated in this previous work.

In order to test the proposed unmixing method, we use data acquired in a preclinical study with an Alzheimer mouse model at ICube laboratory under the IRIS platform, the studies on this data are given in [28], [29]. The data consist of a 3D+t rs-fMRI and a 3D anatomical image registered to the rs-fMRI image. The anatomical image has a dimension of $256 \times 256 \times 34$ and $0.08299 \times 0.07812 \times 0.4$ mm resolution. Functional images have a spatio-temporal dimension of $147 \times 87 \times 27 \times 500$ with $0.1445 \times 0.2299 \times 0.5$ mm spatial resolution and 2s for the temporal resolution. The high resolution (HR) segmentation map

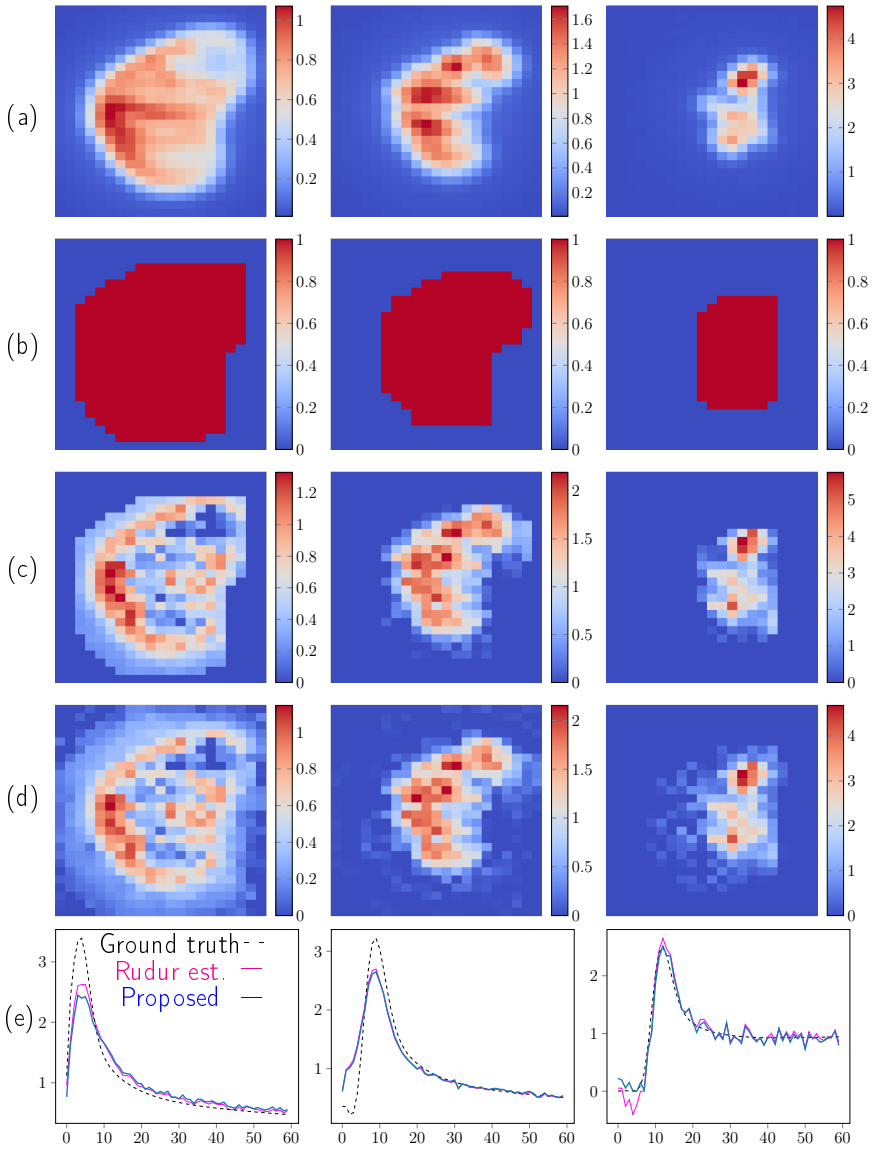


Fig. 3: (a) Scintigraphic data. Ground truth for spatial maps, (b) Initial ROIs , (c) Spatial maps estimated by the proposed algorithm, (d) spatial maps estimated by RUDUR, (e) TACs estimated by RUDUR and our method.

comes from the mouse Allen Brain Atlas (ABA) [30], shown in Fig. 4. This very HR atlas provides a 3D MRI volume (template) and a structural annotation volume, both at $25 \times 25 \times 25$ μm resolution with $228 \times 160 \times 264$ voxels. The annotations identify more than 600 different anatomical structures in the mouse brain.

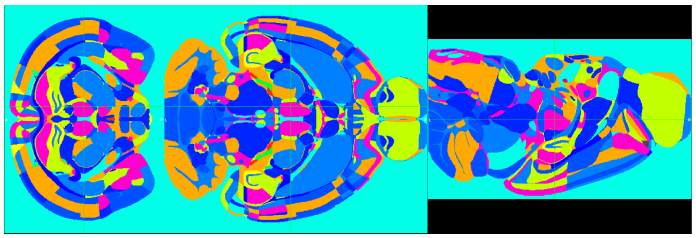


Fig. 4: Rs-fMRI data unmixing. 3D representation of the segmentation map associated with the Allen Mouse Brain Atlas [30]. Each colour represents a label associated with an anatomical region.

1) Validation dataset: A validation dataset is created by introducing a set of synthetic temporal signatures in seven small regions of the real Azheimer mice dataset. The seven regions, labelled ACAd1, ACAd5, ACAd6a, ACAv1, ACAv5, ORBl1 and PL1, have been arbitrarily chosen in the prefrontal cortex. A first synthetic signal is obtained by averaging the real signals of the regions ACAd1 and ACAd5 which were already highly correlated in the real data. This signal is then modified to create signals with arbitrary high correlation or anti-correlation for the regions ACAd1, ACAd5, ACAd6a, ACAv1, ACAv5, ORBl1 and PL1 (see blue lines in plots of figure Fig. 6). These correlations do not have a physical significance, they are used as a ground truth for evaluation of the proposed algorithm performances [16]. The seven synthetic signals are shown in Fig. 6.

The classical fMRI preprocessing pipeline of slice timing and co-registration is applied on the rs-fMRI dataset. Also, the confounding signals are regressed before analysing the data. The next step consists in registering the spatially well-resolved ABA template to the artificially augmented anatomical image (which is already perfectly aligned with the rs-fMRI data). The registration of the ABA mouse template to the anatomical images provides the deformation field that is applied to the HR segmentation map to transport the different labelled regions on the augmented rs-fMRI data. The spatial resolution of rs-fMRI data is augmented by subdividing each original voxel into $3 \times 6 \times 2$ high resolution voxels. The registration of the anatomical image to the augmented rs-fMRI leads to an increase in its own resolution.

Synthetic signals are introduced in the standardised artificially augmented fMRI data, which are then reduced to the initial low resolution. These synthetic signals are thus mixed with the real signals in the voxels containing a portion of the seven selected regions. Since the atlas has a much higher spatial resolution than the fMRI or structural MRI data (up to a factor of 20 in one of the dimensions), the temporal signatures of the different anatomical regions are highly mixed within each low resolution fMRI voxel. Let us note that the pure pixel assumption is not verified in the regions where the signals were added. The minimum and maximum number of overlapping regions on the voxels of each region is given

in table VI.

	ACAd1	ACAd5	ACAd6a	ACAv1	ACAv5	PL1	ORBl1
Min	4	2	4	4	6	4	2
Max	9	12	18	12	17	10	10

TABLE VI: Minimum and maximum number of regions overlapping on the voxels for fMRI regions where signals were added.

2) Algorithm details: Finally, DL is performed at the (low) resolution of the initial fMRI data $\mathbf{Y} \in \mathbb{R}^{N\times P}$, where N = 490 temporal samples and P = 21024 voxels after extracting the brain. The initial abundance matrix $A^{(0)} \in \mathbb{R}^{R\times P}$ is constructed as follows. Let's say that each voxel $i \in \{1,\ldots, \mathrm{P}\}$ was subdivided into J high resolution voxels during the artificial augmentation step. For each voxel i of $\mathbf{Y}$ and all regions $r \in \{1,\ldots,R\}$, the element $(A^{(0)})_{r,i}$ will contain the proportion of high resolution voxels in voxel i, occupied by region r. If region r is not transported to the low-resolution voxel i then $(A^{(0)})_{r,i} = 0$. Matrix $\tilde{A}$ which supports the spatial constraint $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A})$ in eq. (4) is defined as: $(\tilde{A})_{r,i} = 1$ when $(A^{(0)})_{r,i} > 0$ and 0 elsewhere.

3) Results and discussion: The proposed DL method is applied to the validation data set. Empirically, the algorithm convergences to an acceptable solution for $\mathbf{A}$ and $\mathbf{U}$ after 500 iterations, see Fig. 5, corresponding to a gain on the optimization $\kappa_l < 10^{-3}$ (eq. (12)). For the estimation of $\mathbf{A}$, the FISTA algorithm requires a stopping criterion or a maximum number of iterations. In our implementation, FISTA is stopped when $\|\mathbf{a}_i^{(l-1)} - \mathbf{a}_i^{(l)}\|_2 < 10^{-8}$ or $l > 100$.

Fig. 6 shows the estimated timecourses in dashed red against the ground truth signals in blue. The mean squared error (MSE) of the estimated timecourses are given on the plots. Despite the strong mixing in the voxels of the seven regions, our algorithm provides a very good estimate of the synthetic timecourses introduced in the data. In neurosciences applications, these timecourses are then used to build functional brain networks.

B. Hyperspectral astronomic data

In order to confront our algorithm with an unmixing problem on real data, we finally consider the problem of unmixing sources in hyperspectral astronomic data. Some datasets in this domain comprise of hyperspectral datacubes and an external information on the spatial location of the sources. We studied datacubes from the MUSE instrument, the Multi Unit Spectroscopic Explorer [31], installed at the Very

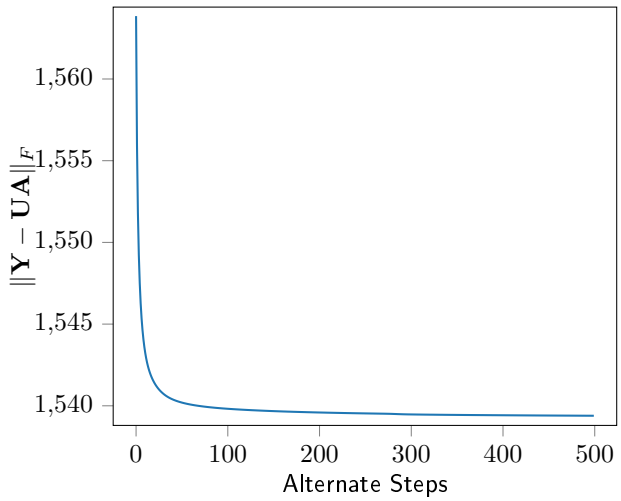


Fig. 5: Rs-fMRI data. The figure shows that $\|\mathbf{Y} - \mathbf{UA}\|_F$ converges; as is seen by the flattening of the curve in the last iterations.

Large Telescope, which produces hyperspectral observations of the deep sky. In these hyperspectral images, we can observe hundreds or even thousands of galaxies. Depending on their age, chemical composition, type, distance, etc, these galaxies have different spectra. These spectra may contain emission lines, continuous components, etc. One of the main objectives of MUSE data analysis is the detection of very distant galaxies, which therefore emit very low light flux. Spectrum of distant galaxies consists of a single emission line, the Lyman-α emission line, which is a marker of the strong presence of hydrogen in the galaxy. They are difficult to detect due to their distance and their very faint intensity compared to closer galaxies, and a strong noise affects the data. Moreover if two galaxies are aligned in the direction of observation, their spectra will blend inside pixels of MUSE images.

Recently two articles [32], [33] were published around a MUSE dataset called Ultra Deep Field (UDF) which corresponds to an area of the sky previously observed by the Hubble Space Telescope (HST). The HST observation is a spatially well-resolved image, of spatial resolution 0.1×0.1 arcsec, for which there is a segmentation map presented in [34]. Due to the difference in resolution of the MUSE data, which is only 0.7×0.7 arcsec, two distinct sources in the HST image, may overlap in the MUSE data creating a mix in the spectra. In [32], [33], the information provided in the Rafelski catalogue is exploited to perform the deblending and prove that MUSE, despite its lower spatial resolution, allows, thanks to the spectral information, to unmix two spatially close or even superimposed sources. We have selected the same portion of the image that is presented in figure 21 of [32] where the objects identified by

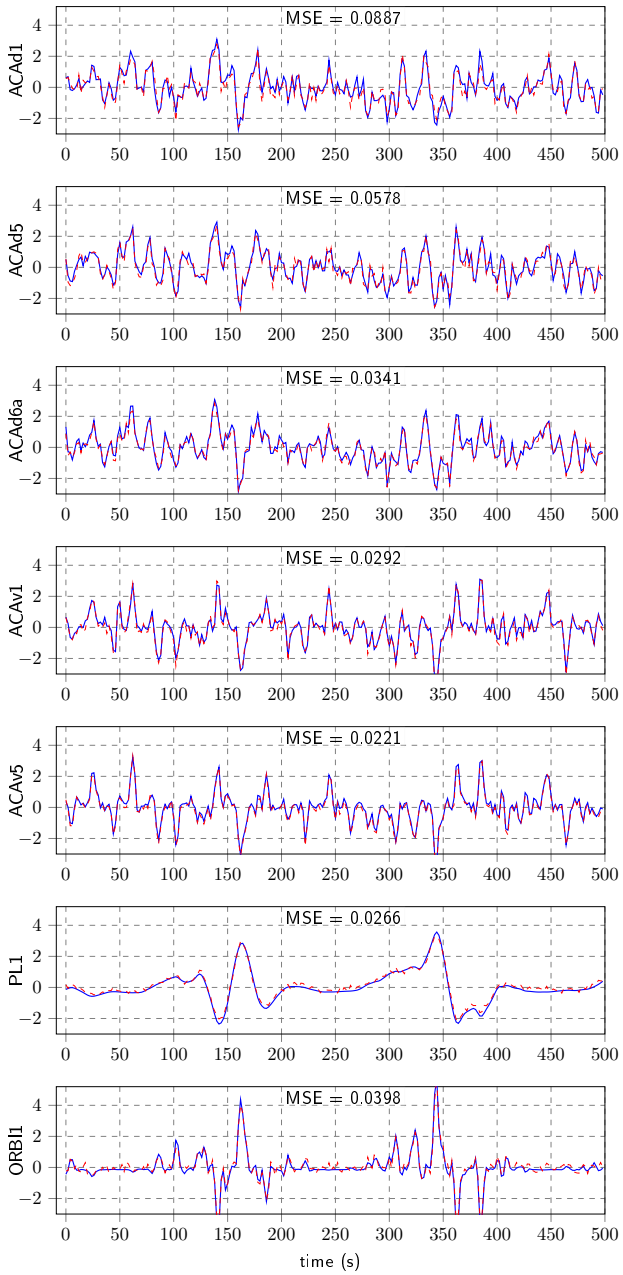


Fig. 6: Rs-fMRI data. The plot shows samples corresponding to the first 500 seconds of the synthetic signals (in blue) and their corresponding estimated timecourses (in dashed red)

ID#4451, ID#4460 and ID#4465 in [34] are spatially superimposed in the MUSE observation. This gives a 25×25 pixels image with spectra composed of 3681 samples from 4750 to 9350 Angström (1 Ang = 0.1 nm). A total of 9 galaxies are present in this field of view, with three of them that are spatially close in the HST segmentation map represented at the middle of the first line in Fig. 7. The

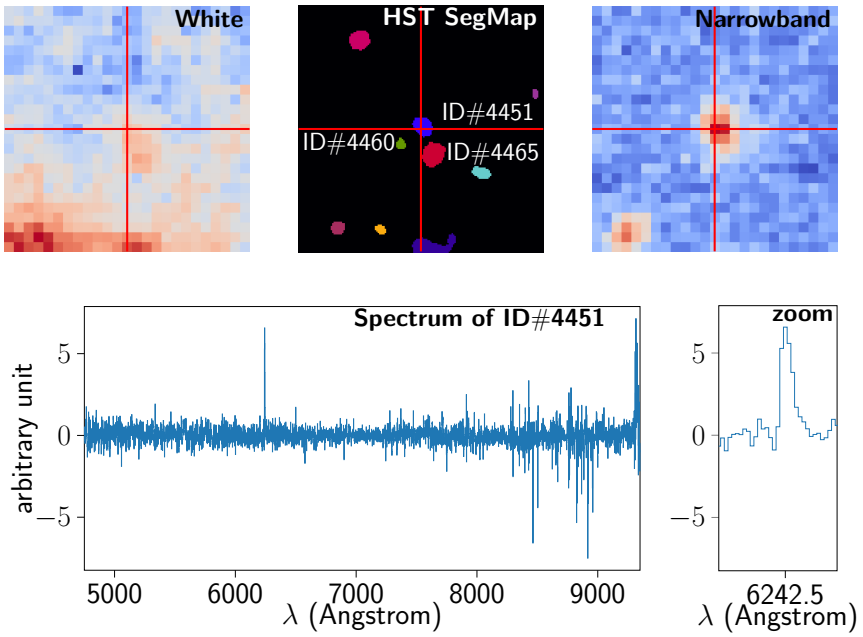


Fig. 7: UDF Hyperspectral astronomic data. On top, from left to right, MUSE reconstructed white light image, HST Rafelski segmentation map and narrowband image centered on $\lambda = 6242.5$ Ang (position of the emission line in estimated spectrum of source ID#4451). The central Rafelski source denoted by red crosshair is ID#4451. Bottom, from left to right: our estimated spectrum of source ID#4451 over the whole wavelength range and zoom on the Lyα emission line estimated at $\lambda = 6242.5$.

source ID#4465 is brighter than galaxies ID#4451 and ID#4460. Its contribution is visible on the white light image, obtained by averaging the datacube with respect to the wavelength axis. A visible source on the white light image indicates that its spectrum contains a continuous component plus, possibly, some emission lines. Contribution of source ID#4451 is embedded in source ID#4465's one. The objective of this section is to show that knowing the spatial location of such a blended source provides enough information to unmix spectra of the different superimposed sources with our algorithm.

For defining the spatial constraint $\mathcal{I}_{M(\tilde{\mathbf{A}})}$ required by our model, the HST segmentation map provided by [34] is used. This map is perfectly registered on the MUSE data. Then by degrading the spatial resolution from 0.1 arcsec to 0.7 arcsec, we obtain binary masks for all the objects present in the field. For the three central sources, the obtained binary masks are shown on the first line of Fig. 8. The algorithm used for unmixing galaxy spectra corresponds exactly to the version described in section II. One hundred alternated optimisation steps allows to reach a gain eq. (12) equal to 10^{-10}. It should be noted that the background is considered here as a source, its mask is available in the segmentation map and it is

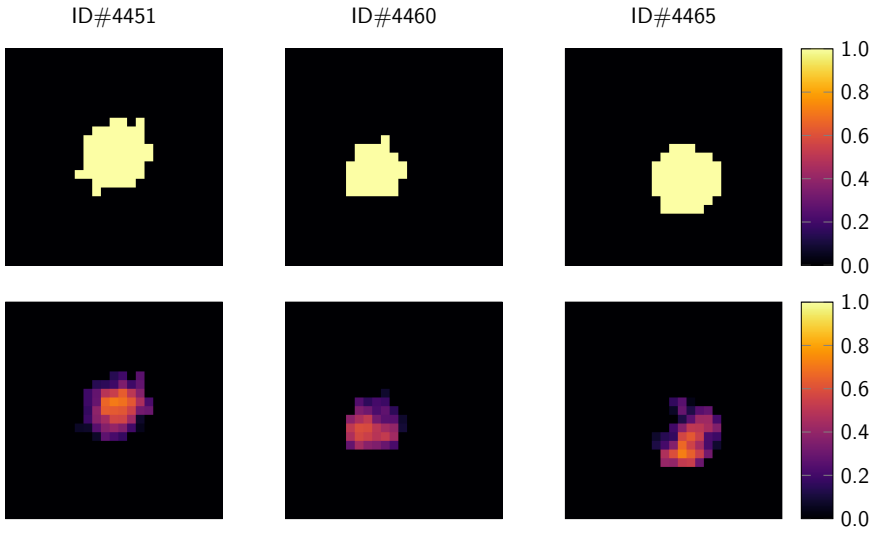


Fig. 8: Hyperspectral astronomic data. On top, from left to right, binary mask of sources ID#4451, ID#4460 and ID#4465. Bottom, from left to right, estimated abundance map of sources ID#4451, ID#4460 and ID#4465

processed in the same way as for the galaxies, to degrade its resolution to the resolution of the MUSE data.

It is impossible to compare quantitatively the results obtained with a ground truth, since no such information exists for the MUSE data. However, same conclusions as in [32] can be drawn about the spectrum estimated by our algorithm for source #4451: at wavelength $\lambda = 6242.5$ Ang, there is an emission line corresponding to object #4451 of Rafelski's catalogue. This emission line has the characteristics of the Lyman-α line (Lyα), namely an asymmetric profile as illustrated in Fig. 7. These result are very similar to the ones presented in Figure 21 in [32] that is reproduced in Fig. 9 by courtesy of the authors. Similarity between results presented in Fig. 7 and Fig. 9 confirms the interest of our generic approach to solve this particular type of unmixing problem.

C. Discussion

The last application shows the ability of the proposed algorithm to unmix hyperspectral data with a spatially constrained dictionary learning algorithm. It also underlines the necessity of having some external knowledge about the localisation of sources in higher spatial resolution to improve classical dictionary learning unmixing algorithm. In the case of MUSE data, such information coming from the segmentation of another telescope observation (here the Hubble Space Telescope) is rarely available. Moreover, a good segmentation of this additional image is also required. Besides, this *a priori* knowledge

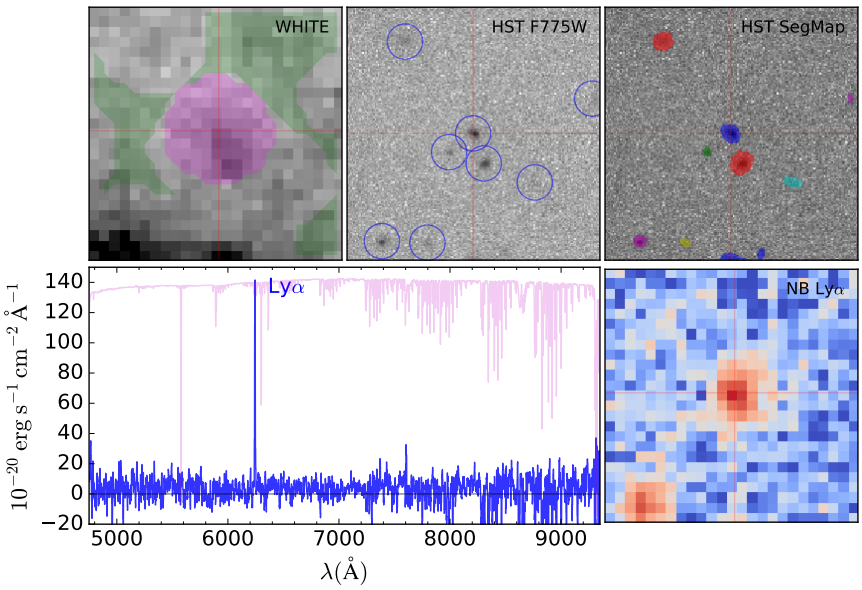


Fig. 9: Hyperspectral astronomic data. Reproduction of figure 21 from paper [32] with the authors' permission and pending approval from A&A.

of sources spatial localisation must be precisely registered to the spectral/temporal data to unmix. As for the fMRI application, in [16] the authors showed that the estimated values were sensitive to the quality of the localisation maps provided. As these maps are used to define the regularisation term $\mathcal{I}_{M(\tilde{a}_i)}$, their precision influences the estimation of the temporal/spectral signatures and their abundances in strongly mixed voxels: the better the precision of the localisation map, the better the estimation of the spectral/temporal information associated to each source.

From an algorithmic point of view, the strong points of the proposed approach are the small number of parameters to be set and its genericity. The algorithm has a single intrinsic parameter μ_σ used in the Tikhonov regularisation, it has only to be set to a very low arbitrary value, as explained in section II-B.

With the scintigraphy application, it has been shown that it is easy to adapt the unmixing problem by adding or removing constraints. In some applications, the proportions of each region have a natural smooth evolution from one pixel to another. A constraint for spatial smoothing within the regions could be introduced in a future variant of the algorithm. However, depending upon the nature of the added constraint, it would increase the complexity of the algorithm. For the current version, the constraints that are in the form of indicator functions, would lead to the introduction of nested loops and consequently to a significant increase in calculation time.

Concerning the execution time of the algorithm, the main factors are the stopping criteria of the different

nested iterative algorithms and the size of the images to be unmixed. Furthermore, the calculation time depends much more on the spatial dimension of the image than on the temporal/spectral dimension as the complexity lies in the estimation of $\mathbf{A}$, and to a lesser extent on the number of regions. The algorithm approached convergence for the astronomical data in a few minutes on a laptop, whereas 3 days CPU time was necessary for the fMRI data on a server with 32 cores, with a processing speed of each core around 1603 Mhz. The applications we presented do not require real-time processing but a GPU implementation would lead to a significant gain in speed as the code is highly parallelisable.

V. Conclusion

A method to unmix data consisting of an image and a temporal or spectral dimension has been presented. The proposed method proves effective in unmixing problems where some prior information related to the ROIs is available or where such information can be obtained from a registered high resolution image. It should be noted that the method does not take into consideration the morphology, the local structure or the texture of the sources, but only their approximate locations. The spatial constraints are classically expressed as an ℓ_1-penalty to promote sparsity of the mixture in each voxel. Problems with such constraints are generally solved by dictionary learning algorithms. The originality of the proposed approach lies in the replacement of this penalty by a constraint on localisation of the different regions of interests. In some of the examples presented in the article, it has been shown that the algorithm can be easily adapted for specific applications by introducing or dropping constraints of *sum-to-one* on spatial maps or positivity on timecourses. State-of-the-art results have been obtained on three very different applications. The comprehensiveness of the method makes it possible to easily adapt it to other fields of application such as remote sensing applications for which incorporation of spatial information has been proved to enhance hyperspectral image unmixing performance [35]. The code will be available soon online for the community.

Appendix

The convexity of $\mathcal{I}_{M(\tilde{\mathbf{a}}_i)\cap S^+}(\mathbf{a}_i)$ can be proven if we can prove that the set $M(\tilde{\mathbf{a}}_i)\cap S\cap\mathbb{R}^+$ is convex. This can be easily verified by following the assumption that each voxel contains the contribution of at

least one region of interest. By definition:

$$\begin{aligned}\mathcal{I}_{\mathbb{R}^+\cap M(\tilde{\mathbf{a}}_i)\cap S}(\mathbf{a}_i)&\\ &= +\infty \text{ if } \mathbf{a}_i \text{ has a non-zero value where } \tilde{\mathbf{a}}_i \text{ is } 0\\ &= +\infty \text{ if the sum of } \mathbf{a}_i \text{ is not equal to } 1\\ &= +\infty \text{ if at least one element in } \mathbf{a}_i \text{ is negative}\\ &= 0 \text{ otherwise}\end{aligned}$$

where $(\tilde{\mathbf{a}}_i)$ is a binary vector where element $(\tilde{\mathbf{a}}_i)_r = 1$ if the r^{th} region of the segmentation map intersects the i^{th} voxel, and 0 otherwise. A set formed by the intersection $M(\tilde{\mathbf{a}}_i) \cap S \cap \mathbb{R}^+$ is convex if for all $t \in [0, 1]$ and for all $\mathbf{v}, \mathbf{w} \in M(\tilde{\mathbf{a}}_i) \cap S \cap \mathbb{R}^+$, a line segment defined :

$$\mathbf{z} = t\mathbf{v} + (1 - t)\mathbf{w} \tag{13}$$

lies in $M(\tilde{\mathbf{a}}_i) \cap S \cap \mathbb{R}^+$.

Proof that $\mathbf{z} \in \mathbb{R}^+$ and $\mathbf{z} \in S$ is straightforward.

Proof that the line segment $\mathbf{z}$ belongs to subspace $M(\tilde{\mathbf{a}}_i)$ is detailed in the following.

Let $\mathbf{u} \in \mathbb{R}^+$. By introducing the complementary vector $(\mathbf{1} - (\tilde{\mathbf{a}}_i))$ of $\tilde{\mathbf{a}}_i$, where the notation $\mathbf{1}$ is a vector of ones of the same size as $\tilde{\mathbf{a}}_i$, we have the equivalence:

$$\mathbf{u} \in M(\tilde{\mathbf{a}}_i) \iff \mathbf{u}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) = 0 \tag{14}$$

Implication $\mathbf{u} \in M(\tilde{\mathbf{a}}_i) \implies \mathbf{u}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) = 0$ is straightforward (vector $\mathbf{u}$ lying in $M(\tilde{\mathbf{a}}_i)$ has its coefficients $(\mathbf{u})_r = 0$ when $1 - (\tilde{\mathbf{a}}_i)_r = 1$). It is easy to prove that the reverse is true provided that $\mathbf{u} \in \mathbb{R}^+$. As a consequence, vectors $\mathbf{v}$ and $\mathbf{w}$ lying in $M(\tilde{\mathbf{a}}_i)$ verify:

$$\mathbf{v}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) = 0 \quad \text{and} \quad \mathbf{w}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) = 0 \tag{15}$$

which yields:

$$\begin{aligned}&t\mathbf{v}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) + (1 - t)\mathbf{w}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) = 0,\\ \iff \quad &\mathbf{z}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) = 0,\end{aligned}$$

which, since $\mathbf{z} \in \mathbb{R}^+$ and according to (14), implies that $\mathbf{z} \in M(\tilde{\mathbf{a}}_i)$.

The formulae for the different criteria used to analyse the experimental results are given below:

$$NMSE(U_r^{GT}, U_r) = \frac{\sum_{n=1}^{N}(U_{r,n}^{GT} - U_{r,n})^2}{\sum_{n=1}^{N}(U_{r,n}^{GT})^2} \tag{16}$$

$$NMAE(U_r^{GT}, U_r) = \frac{\sum_{n=1}^{N} |U_{r,n}^{GT} - U_{r,n}|}{|\sum_{n=1}^{N} (U_{r,n}^{GT})|} \tag{17}$$

$$SAD(U_r^{GT}, U_r) = \cos^{-1} \frac{\sum_{n=1}^{N} U_{r,n}^{GT} U_{r,n}}{\sqrt{\sum_{n=1}^{N} U_{r,n}^{GT\,2}} \sqrt{\sum_{n=1}^{N} U_{r,n}^{2}}} \tag{18}$$

where U_r^{GT} is the ground truth temporal signal for r^{th} region, U_r is the estimated signal for the r^{th} region, N is the length of the temporal signal. NMAE and NMSE can be written in a similar manner to evaluate spatial maps.

ACKNOWLEDGMENT

The authors would like to thank IRIS platform [36] at ICube for the Alzheimer's data and equally R. Bacon for providing MUSE dataset.

REFERENCES

[1] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37–52, 1987.

[2] C. Jutten and J. Herault, "Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture," *Signal processing*, vol. 24, no. 1, pp. 1–10, 1991.

[3] P. Comon, "Independent component analysis, a new concept?," *Signal processing*, vol. 36, no. 3, pp. 287–314, 1994.

[4] J. M. Bioucas-Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *2009 First workshop on hyperspectral image and signal processing: Evolution in remote sensing*, pp. 1–4, IEEE, 2009.

[5] V. Šmídl and O. Tichý, "Automatic regions of interest in factor analysis for dynamic medical imaging," in *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 158–161, IEEE, 2012.

[6] M. J. McKeown, T.-P. Jung, S. Makeig, G. Brown, S. S. Kindermann, T.-W. Lee, and T. J. Sejnowski, "Spatially independent activity patterns in functional MRI data during the Stroop color-naming task," *Proceedings of the National Academy of Sciences*, vol. 95, no. 3, pp. 803–810, 1998.

[7] J. Xu, M. Potenza, and V. Calhoun, "Spatial ICA reveals functional activity hidden from traditional fMRI GLM-based analyses," *Frontiers in Neuroscience*, vol. 7, p. 154, 2013.

[8] M. Sourty, L. Thoraval, D. Roquet, J.-P. Armspach, and J. Foucher, "Towards an automated selection of spontaneous co-activity maps in functional magnetic resonance imaging," in *Medical Imaging 2015: Biomedical Applications in Molecular, Structural, and Functional Imaging* (B. Gimi and R. C. Molthen, eds.), vol. 9417, pp. 144 – 151, International Society for Optics and Photonics, SPIE, 2015.

[9] Z. Yang, G. Zhou, S. Xie, S. Ding, J. Yang, and J. Zhang, "Blind spectral unmixing based on sparse nonnegative matrix factorization," *IEEE Transactions on Image Processing*, vol. 20, pp. 1112–1125, April 2011.

[10] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and 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, 2012.

[11] V. Abolghasemi, S. Ferdowsi, and S. Sanei, "Blind separation of image sources via adaptive dictionary learning," *IEEE Transactions on Image Processing*, vol. 21, pp. 2921–2930, June 2012.

[12] A. Abraham, E. Dohmatob, B. Thirion, D. Samaras, and G. Varoquaux, "Extracting brain regions from rest fMRI with total-variation constrained dictionary learning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 607–615, Springer, 2013.

[13] H. Eavani, R. Filipovych, C. Davatzikos, T. D. Satterthwaite, R. E. Gur, and R. C. Gur, "Sparse dictionary learning of resting state fMRI networks," in *Pattern Recognition in NeuroImaging (PRNI), 2012 International Workshop on*, pp. 73–76, IEEE, 2012.

[14] G. Varoquaux, A. Gramfort, F. Pedregosa, V. Michel, and B. Thirion, "Multi-subject dictionary learning to segment an atlas of brain spontaneous activity," in *Information Processing in Medical Imaging*, (Berlin, Heidelberg), pp. 562–573, Springer Berlin Heidelberg, 2011.

[15] A. Mensch, G. Varoquaux, and B. Thirion, "Compressed online dictionary learning for fast resting-state fMRI decomposition," in *International Symposium on Biomedical Imaging*, 13th International Symposium on Biomedical Imaging, (Prague, Czech Republic), pp. 1282–1285, IEEE, IEEE, Apr. 2016.

[16] A. Bhanot, C. Meillier, F. Heitz, and L. Harsan, "Online dictionary learning for single-subject fMRI data unmixing," in *2019 27th European Signal Processing Conference (EUSIPCO)*, pp. 1–5, IEEE, 2019.

[17] L. Loncan, L. B. De Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simoes, *et al.*, "Hyperspectral pansharpening: A review," *IEEE Geoscience and remote sensing magazine*, vol. 3, no. 3, pp. 27–46, 2015.

[18] M. Filippi, M. Desvignes, and E. Moisan, "Robust unmixing of dynamic sequences using regions of interest," *IEEE transactions on medical imaging*, vol. 37, no. 1, pp. 306–315, 2017.

[19] H. Benali, I. Buvat, F. Frouin, J. Bazin, and R. Di Paola, "Foundations of factor analysis of medical image sequences: a unified approach and some practical implications," *Image and Vision Computing*, vol. 12, no. 6, pp. 375–385, 1994.

[20] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[21] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[22] S. Boyd, J. Dattorro, *et al.*, "Alternating projections," *EE392o, Stanford University*, 2003.

[23] K. Nijran and D. Barber, "The importance of constraints in factor analysis of dynamic studies," in *Information processing in medical imaging*, pp. 521–529, Springer, 1988.

[24] M. Šámal, A. Piepsz, G. Brolin, J. Heikkinen, and J. Valoušek, "Database of dynamic renal scintigraphy." http://www.dynamicrenalstudy.org/, 2012. [Online; accessed 3-March-2020].

[25] G. Brolin, K. Sjögreen Gleisner, and M. Ljungberg, "Dynamic (99m)Tc-MAG3 renography: images for quality control obtained by combining pharmacokinetic modelling, an anthropomorphic computer phantom and Monte Carlo simulated scintillation camera imaging.," *Physics in Medicine and Biology*, vol. 58, no. 10, pp. 3145–3161, 2013.

[26] M. H. Lee, C. D. Smyser, and J. S. Shimony, "Resting-state fMRI: a review of methods and clinical applications," *American Journal of Neuroradiology*, vol. 34, no. 10, pp. 1866–1872, 2013.

[27] N. K. Logothetis and B. A. Wandell, "Interpreting the BOLD signal," *Annu. Rev. Physiol.*, vol. 66, pp. 735–769, 2004.

[28] L. Degiorgis, M. Karatas, M. Sourty, T. Bienert, M. Reisert, C. Mathis, A.-L. Boutillier, F. Blanc, J.-P. Armspach, and L.-A. Harsan, "Longitudinal alterations of resting-state functional connectivity in alzheimer's disease in a tauopathy mouse model," in *Proceedings of the ISMRM Paris, France*, 2018.

[29] L. Degiorgis, *MRI analysis of brain connectivity in a mouse model of Alzheimer's disease*. PhD thesis, Strasbourg, 2019.

[30] S. W. O. et al., "A mesoscale connectome of the mouse brain," *Nature*, vol. 508, pp. 207–214, 2014.

[31] R. Bacon, M. Accardo, L. Adjali, H. Anwand, S. Bauer, I. Biswas, J. Blaizot, D. Boudon, S. Brau-Nogue, J. Brinchmann, *et al.*, "The MUSE second-generation VLT instrument," in *Ground-based and Airborne Instrumentation for Astronomy III*, vol. 7735, p. 773508, International Society for Optics and Photonics, 2010.

[32] R. Bacon, S. Conseil, D. Mary, J. Brinchmann, M. Shepherd, M. Akhlaghi, P. M. Weilbacher, L. Piqueras, L. Wisotzki, D. Lagattuta, *et al.*, "The MUSE Hubble Ultra Deep Field Survey-I. Survey description, data reduction, and source detection," *Astronomy & Astrophysics*, vol. 608, p. A1, 2017.

[33] H. Inami, R. Bacon, J. Brinchmann, J. Richard, T. Contini, S. Conseil, S. Hamer, M. Akhlaghi, N. Bouché, B. Clément, *et al.*, "The MUSE Hubble Ultra Deep Field Survey-II. Spectroscopic redshifts and comparisons to color selections of high-redshift galaxies," *Astronomy & Astrophysics*, vol. 608, p. A2, 2017.

[34] M. Rafelski, H. I. Teplitz, J. P. Gardner, D. Coe, N. A. Bond, A. M. Koekemoer, N. Grogin, P. Kurczynski, E. J. McGrath, M. Bourque, *et al.*, "UVUDF: ultraviolet through near-infrared catalog and photometric redshifts of galaxies in the Hubble Ultra Deep Field," *The Astronomical Journal*, vol. 150, no. 1, p. 31, 2015.

[35] L. Wang, C. Shi, C. Diao, W. Ji, and D. Yin, "A survey of methods incorporating spatial information in image classification and spectral unmixing," *International Journal of Remote Sensing*, vol. 37, no. 16, pp. 3870–3910, 2016.

[36] Accueil, "Plateforme IRIS: Imagerie, robotique et innovation en santé,." https://plateforme.icube.unistra.fr/iris/index.php?title=Accueil&oldid=649, 2020. [Online; accessed 14-May-2020].