

Separation of Delayed Parameterized Sources

Hassan Mortada¹, Vincent Mazet¹, Charles Soussen² and Christophe Collet¹

¹ICUBE, Université de Strasbourg - CNRS, 67412 Illkirch, France

²Université de Lorraine and CNRS, CRAN, 54518 Vandœuvre-lès-Nancy, France

E-mail: {mortada, vincent.mazet, c.collet}@unistra.fr, charles.soussen@univ-lorraine.fr

Abstract—This paper addresses the delayed (or anechoic) source separation problem in the case of parameterized deterministic sources. An alternating least square scheme is proposed to estimate the source parameters, the mixing coefficients and the delays. For the challenging delay parameter we adapt a sparse approximation strategy. A first algorithm considers discrete delays; then an extension, inspired by the recent sparse deconvolution literature, allows for continuous delay estimation. Numerical simulations demonstrate the effectiveness of the proposed algorithms compared to state-of-the-art methods for highly correlated Gaussian sources.

Keywords—anechoic source separation, alternating least squares, continuous delay estimation, decomposition of spectroscopic signals, correlated sources.

I. INTRODUCTION

Separation of delayed sources, also known as anechoic source separation, is a special case of convolutive source separation [1] where each delayed source occurs at most once in each mixture and is invariant to translation. Some applications include, *e.g.*, audio signal processing [2], underwater acoustics [3]. This work is inspired by the study of galaxy kinematics [4] where data are multispectral images, where each pixel is related to a spectrum with emission lines. These lines must be estimated but they undergo different delays in each spectrum due to the redshift and internal gas motions (Doppler effect). This justifies the delayed source separation model, where sources and mixtures are respectively associated to emission lines and spectra.

Since the delayed source separation problem is ill-posed, strong assumptions are made to discriminate the sources. Most of the methods rely on the assumption that the sources are independent, like in the shifted independent component analysis method [5]. The source independence can be promoted by different ways, for instance with second order statistic methods either in temporal [6], [7] or frequency domains [3], [8]. Another strategy is to analyze the data in a domain where the sources are disjoint or uncorrelated, usually the time-frequency domain [2], [9], [10]. However, the independent and uncorrelated source assumptions do not hold in many real-world problems [11], so classical approaches cannot be applied. An early attempt to separate correlated sources was proposed in [12] using sparsity constraints on the sources: each mixture is sparsified, then the sparse representations are matched together

to retrieve the sources and the related delays. However, this two-step method may suffer from error propagations. In [13], [14], the uncorrelated source assumption has been relaxed, by introducing the concept of partially correlated sources. This technique, however, is limited to the instantaneous source separation problem.

In the literature, the difficult situation where sources are highly correlated and delayed is not covered, thus new algorithms have to be proposed. In this work, we propose an original method that takes advantage of prior knowledge on the shape of the source signals. We assume that the sources can be modeled by parameterized functions, whose parameters correspond to shape information such as the width of an emission line. The separation of parameterized sources might appear to be simpler than blind source separation since the shape of the sources is known beforehand up to a few parameters. However, parameterized sources are often correlated especially if the sources are modeled with the same function. For instance, emission lines in the galaxy kinematic problem are modeled by Gaussian functions [4], yielding highly correlated sources.

Estimating the source parameters in a unique mixture shares similarities with sparse deconvolution, where time-shifted copies of a *perfectly known* waveform are gathered into a dictionary [15], [16]. We extend this framework to the case of numerous sources and mixtures, by considering different waveforms (*i.e.* sources) whose shape parameters are unknown and have to be estimated. Furthermore, the sparse solution is designed to comply with the delayed source separation model, in particular the fact that sources occur at most once in a mixture. In addition, our approach exploits recent results on sparse deconvolution that allow for continuous delay estimation [15], [16]. Finally, the dictionary does not need to be stored.

This paper is organized as follows. The delayed and parameterized source separation problem and the optimization problem are formulated in section II. Section III presents the Alternative Least Squares (ALS) framework, then two ALS algorithms are proposed for discrete and continuous delay estimation in section IV. In section V, numerical simulations illustrate the performance of the proposed algorithms.

II. PROBLEM STATEMENT

In delayed source separation, each mixture $\mathbf{x}_i(t)$ is modeled as a linear combination of J sources $\mathbf{s}_j(t)$:

$$\mathbf{x}_i(t) = \sum_{j=1}^J a_{ij} \mathbf{s}_j(t - c_{ij}) \quad i = 1, \dots, I \quad (1)$$

where a_{ij} and c_{ij} are respectively the amplitude and the delay from source j to mixture i . Hereafter, we consider the over-determined case where $I > J$ (this is a typical setting in multispectral imaging). Recall that the sources we are considering are assumed to follow a known parameterized model $\mathbf{s}_j(t; w_j)$ where w_j represents the shape parameters of source j . The noisy observations $\mathbf{x}_i(t)$ are thus modeled by:

$$\mathbf{x}_i(t) = \sum_{j=1}^J a_{ij} \mathbf{s}_j(t - c_{ij}; w_j) + \mathbf{n}_i(t) \quad i = 1, \dots, I \quad (2)$$

where the noise $\mathbf{n}_i(t)$ gathers both the observation noise and the modeling errors.

Source separation is recast as a parameter estimation problem, where the parameters to estimate are the amplitudes a_{ij} , the delays c_{ij} and the shape parameters w_j . Assuming that the noise is i.i.d. Gaussian, source separation can be formulated as the least squares optimization problem:

$$\min_{\mathbf{A}, \mathbf{C}, \mathbf{w}} \left\{ E(\mathbf{A}, \mathbf{C}, \mathbf{w}) = \sum_{i=1}^I \varepsilon(\mathbf{a}_i, \mathbf{c}_i, \mathbf{w}) \right\} \quad (3)$$

where the error related to mixture i is given by:

$$\varepsilon(\mathbf{a}_i, \mathbf{c}_i, \mathbf{w}) = \left\| \mathbf{x}_i(t) - \sum_{j=1}^J a_{ij} \mathbf{s}_j(t - c_{ij}; w_j) \right\|_2^2. \quad (4)$$

In (3)-(4), t is sampled over a discrete grid of length N . Without loss of generality, we assume that the sampling step is equal to 1. Matrices \mathbf{A} and \mathbf{C} and vector \mathbf{w} are given by:

- $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_I] \in \mathbb{R}^{J \times I}$ with $\mathbf{a}_i = [a_{i1}, \dots, a_{iJ}]^T$;
- $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_I] \in \mathbb{R}^{J \times I}$ with $\mathbf{c}_i = [c_{i1}, \dots, c_{iJ}]^T$;
- $\mathbf{w} = [w_1, \dots, w_J]^T \in \mathbb{R}^{J \times 1}$.

Minimization of criterion (3) is discussed in the next section.

III. ALTERNATING LEAST SQUARES

ALS is a classical descent strategy [17], consisting of minimizing criterion E with respect to groups of variables alternatively. In delayed source separation, several approaches [2], [5], [8] make use of an ALS strategy, where the sources are updated while fixing the delays and amplitudes to the values found in the previous iteration and *vice versa*. However, the update stages significantly differ in the proposed algorithm from these works. Specifically, the source estimation stage is replaced by their shape estimation, whereas the delays are estimated using an effective sparse deconvolution solver. It should be noticed that the ALS scheme does not guarantee to converge towards the global minimizer of (3), since each step of the scheme is a local minimization with respect to a group of the parameters and E is non-convex. However, all

Algorithm 1: Alternative Least Squares.

Inputs: mixtures \mathbf{x}_i , stopping parameter ρ
Outputs: $\hat{\mathbf{C}}, \hat{\mathbf{A}}, \hat{\mathbf{w}}$
Initialization: $\hat{\mathbf{A}}, \hat{\mathbf{C}}, \hat{\mathbf{w}}$
1 $\hat{E} \leftarrow E(\hat{\mathbf{A}}, \hat{\mathbf{C}}, \hat{\mathbf{w}})$
2 **do**
3 $E^{\text{old}} \leftarrow \hat{E}$
4 $\hat{\mathbf{w}} \leftarrow \underset{\mathbf{w}}{\text{argmin}} E(\hat{\mathbf{A}}, \hat{\mathbf{C}}, \mathbf{w})$
5 **for** $i=1:I$ **do**
6 $(\hat{\mathbf{a}}_i, \hat{\mathbf{c}}_i) \leftarrow \underset{\mathbf{a}_i, \mathbf{c}_i}{\text{argmin}} \varepsilon(\mathbf{a}_i, \mathbf{c}_i, \hat{\mathbf{w}})$
7 **end**
8 $\hat{E} \leftarrow E(\hat{\mathbf{A}}, \hat{\mathbf{C}}, \hat{\mathbf{w}})$
9 **while** $|E^{\text{old}} - \hat{E}|/E^{\text{old}} \geq \rho$

the simulations we have performed led to estimations close to the ground-truth parameters.

The ALS framework is presented in Algorithm 1, where the stopping criterion is related to the relative decrease of the approximation error. In the first stage, E is minimized with respect to \mathbf{w} (line 4 of Algorithm 1) using a non-linear least-squares solver, namely the Levenberg-Marquardt algorithm. In the second stage, E is minimized with respect to \mathbf{A} and \mathbf{C} jointly. From (3), this is a separable problem leading to I independent sub-problems (lines 5-7 of Algorithm 1):

$$\min_{\mathbf{A}, \mathbf{C}} E(\mathbf{A}, \mathbf{C}, \mathbf{w}) \Leftrightarrow \forall i, \min_{\mathbf{a}_i, \mathbf{c}_i} \varepsilon(\mathbf{a}_i, \mathbf{c}_i, \mathbf{w}) \quad (5)$$

The implementation of each sub-problem is detailed next.

IV. SPARSE APPROACH FOR AMPLITUDE AND DELAY ESTIMATION

Hereafter, index i is dropped for brevity. The i -th sub-problem in the right-hand side of (5) can be rewritten as:

$$\begin{aligned} & \min_{\mathbf{a}, \mathbf{c}} \left\| \mathbf{x}(t) - \sum_{j=1}^J a_j \mathbf{s}_j(t - c_j; w_j) \right\|_2^2 \\ & = \min_{\mathbf{a}, \mathbf{c}} \left\| \mathbf{x}(t) - \sum_{j=1}^J \mathbf{s}_j(t; w_j) * (a_j \delta_{c_j}(t)) \right\|_2^2 \end{aligned} \quad (6)$$

where $*$ is the convolution operator and $\delta_{c_j}(t) = \delta(t - c_j)$ refers to the Kronecker function. Eq. (6) involves only one mixture $\mathbf{x}(t) = \mathbf{x}_i(t)$. In addition, the reader should keep in mind that c_j (resp., a_j) is the delay of the j -th source in mixture i , and has to be understood as c_{ij} (resp., a_{ij}) in (4).

Eq. (6) reads as a sparse approximation problem since $a_j \delta_{c_j}(t)$ are 1-sparse vectors. Therefore, our strategy to estimate the delays c_j and amplitudes a_j is inspired by sparse approximation algorithms.

A. Discrete Delay Estimation

To address each sparse approximation problem (6), we impose the delays c_j to be located on a discrete grid of sampling step Δ . In other words, the delays take the form $c_j = \ell \Delta$ with

Algorithm 2: OMP-like implementation for minimization of $\varepsilon(\mathbf{a}, \mathbf{c}, \mathbf{w})$ w.r.t. (\mathbf{a}, \mathbf{c}) : case of discrete delays

Inputs: mixture \mathbf{x} , sources $\mathbf{s}_j(t; w_j)$, sampling step Δ

Outputs: \mathbf{a} and \mathbf{c}

Initialization: $\mathcal{S} = \emptyset$, $\mathcal{J} = \emptyset$, $\mathbf{c} = \mathbf{0}$

```

1 for  $k = 1 : J$  do
2    $(\hat{\ell}, \hat{j}) \leftarrow \underset{\substack{\ell \in \mathbb{N} \\ j \notin \mathcal{J}}}{\text{argmax}} \frac{\langle \mathbf{s}_j * \delta_{\ell\Delta}, \mathbf{x} - \mathbf{S}\mathbf{a} \rangle}{\|\mathbf{s}_j\|_2}$ 
3    $\mathcal{J}(k) \leftarrow \hat{j}$ 
4    $c_{\hat{j}} \leftarrow \hat{\ell}\Delta$ 
5   Include  $\mathbf{s}_{\hat{j}}(t - c_{\hat{j}}; w_{\hat{j}})$  as new column in  $\mathbf{S}$ 
6    $\mathbf{a} \leftarrow \mathbf{S}^\dagger \mathbf{x}$ 
7 end
8 Permute  $\mathbf{a}$  with respect to  $\mathcal{J}$ 

```

$\ell \in \mathbb{N}$. The sparsity level J being small, we resort to greedy algorithms because of their simplicity. Specifically, we propose to adapt the classical Orthogonal Matching Pursuit (OMP) algorithm [18] (Algorithm 2), which is an iterative forward selection algorithm. In the proposed adaptation, one source (indexed by \hat{j}) is selected per iteration, and the related delay $c_{\hat{j}}$ is computed using the OMP selection rule, *i.e.*, selecting the most correlated delayed source version to the current residual. \mathcal{J} gathers the indices of the selected sources, and \dagger denotes the pseudo-inverse operator, $\mathbf{S}^\dagger = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T$. The amplitudes $a_{\hat{j}}$ are updated by solving a linear least squares problem, and an index permutation (with respect to \mathcal{J}) is applied in step 8. This permutation is required since the order in which the delayed version are arranged in \mathbf{S} does not match the predefined ordering of sources.

B. Continuous Delay Estimation

The delays can be estimated more accurately by increasing the grid resolution, *i.e.*, decreasing the value of Δ . However, this increases the computational burden and might affect negatively the performance of the OMP algorithm for correlated sources. We choose an alternating strategy, in which the OMP-like implementation of Subsection IV-A is adapted to deal with continuous delays (off-the-grid delays). Recently, Fyhn *et al.* [16, Section IV] proposed a continuous extension of OMP, named IBOMP, for sparse deconvolution. Their approach makes use of an interpolation scheme first introduced in [15]. Specifically, this approach requires to preset a grid of step Δ , and then to interpolate the delayed versions of the waveform for consecutive delays that lay on the grid.

IBOMP can be straightforwardly adapted to our context because the structure of the delayed unmixing problem is very close to that of sparse deconvolution problems. Specifically, each iteration of Algorithm 2 almost identifies to an iteration of OMP for sparse deconvolution (with waveform $\mathbf{s}_{\hat{j}}$), yielding the selection of source \hat{j} and related discrete (coarse) estimation delay $\hat{\ell}\Delta$. Then, the refined estimation of $c_{\hat{j}}$ becomes $c_{\hat{j}} + \eta$

where $\eta \in [-\Delta/2, \Delta/2]$ is a real value. The proposed strategy consists in replacing step 5 of Algorithm 2 by:

Include $\mathbf{s}_{\hat{j}}(t - c_{\hat{j}} - \eta; w_{\hat{j}})$ as new column in \mathbf{S}

The authors in [15], [16] advocate the use of polar interpolation which empirically outperforms other interpolation techniques in terms of delay estimation. This strategy exploits the fact that delayed versions of a source $\mathbf{s}_{\hat{j}}(t)$ form a manifold which lies on the surface of a hypersphere (because the ℓ_2 -norm of the sources is preserved under delay variations). Therefore, a segment on this manifold can be approximated by an arc, which leads to polar interpolation. The resulting equation has the following closed-form [15, Eq. (16)]:

$$\mathbf{s}_{\hat{j}}(t - c_{\hat{j}} - \eta; w_{\hat{j}}) = (\alpha \quad \beta \quad \gamma) \begin{pmatrix} \mathbf{s}_{\hat{j}}(t - c_{\hat{j}} - \frac{\Delta}{2}; w_{\hat{j}}) \\ \mathbf{s}_{\hat{j}}(t - c_{\hat{j}}; w_{\hat{j}}) \\ \mathbf{s}_{\hat{j}}(t - c_{\hat{j}} + \frac{\Delta}{2}; w_{\hat{j}}) \end{pmatrix} \quad (7)$$

where α , β and γ are scalar values calculated as in [16].

V. RESULTS

The performance of the proposed algorithm is evaluated on synthetic data. We are inspired by the galaxy kinematic problem [4] implying that: (i) all the sources are modeled by a Gaussian peak: $\mathbf{s}_j(t; w_j) = \exp(-t^2/2w_j^2)$; (ii) the shape parameters and amplitudes take positive values, while the delays slowly evolve from one mixture to another; (iii) the sources do not necessarily appear in each mixture. To consider the last point, a stopping criterion is added in Algorithm 2 (line 1): the loop can break in less than J iterations when the residual $\varepsilon(\mathbf{a}, \mathbf{c}, \mathbf{w})$ becomes lower than a threshold (set to the noise variance in our experiments).

In the first experiment, $I = 30$ mixtures are generated with $N = 200$ samples, $J = 3$ sources, widths $\mathbf{w}^* = [1.99 \ 3.78 \ 8.95]$ and a Signal to Noise Ratio (SNR)¹ equal to 10 dB. Fig. 1 (a),(e) show that the proposed algorithm (with continuous delay estimation and step $\Delta = 2$) is able to separate properly the sources and their parameters (in particular, $\hat{\mathbf{w}} = [1.89 \ 3.81 \ 9.05]$). We also compare with [19], which specifically decomposes a sequence of spectroscopic signals into a sum of Gaussians, using an MCMC (Monte Carlo Markov Chain) algorithm stated in a Bayesian framework. The estimation is also satisfactory [Fig. 1 (b),(f)], with $\hat{\mathbf{w}}^{\text{MCMC}} = [2.05 \ 3.97 \ 8.99]$ and the reconstruction of the mixtures is clearly correct. On the contrary, classical delayed source separation methods [2], [5] fail to retrieve the sources since they do not match the Gaussian shape [Fig. 1 (g)-(h)]. Furthermore, the reconstruction of the mixtures fails as well [Fig. 1 (c)-(d)]. This can be explained by the fact that independent and uncorrelated source assumptions are not valid in our application.

Next, we present a comparison between both versions of the proposed algorithm (discrete and continuous delay estimation) and the method of [19] that gave comparable results,

¹Defined as 10 times the log-ratio of the mean energy of the noiseless spectra and the noise variance.

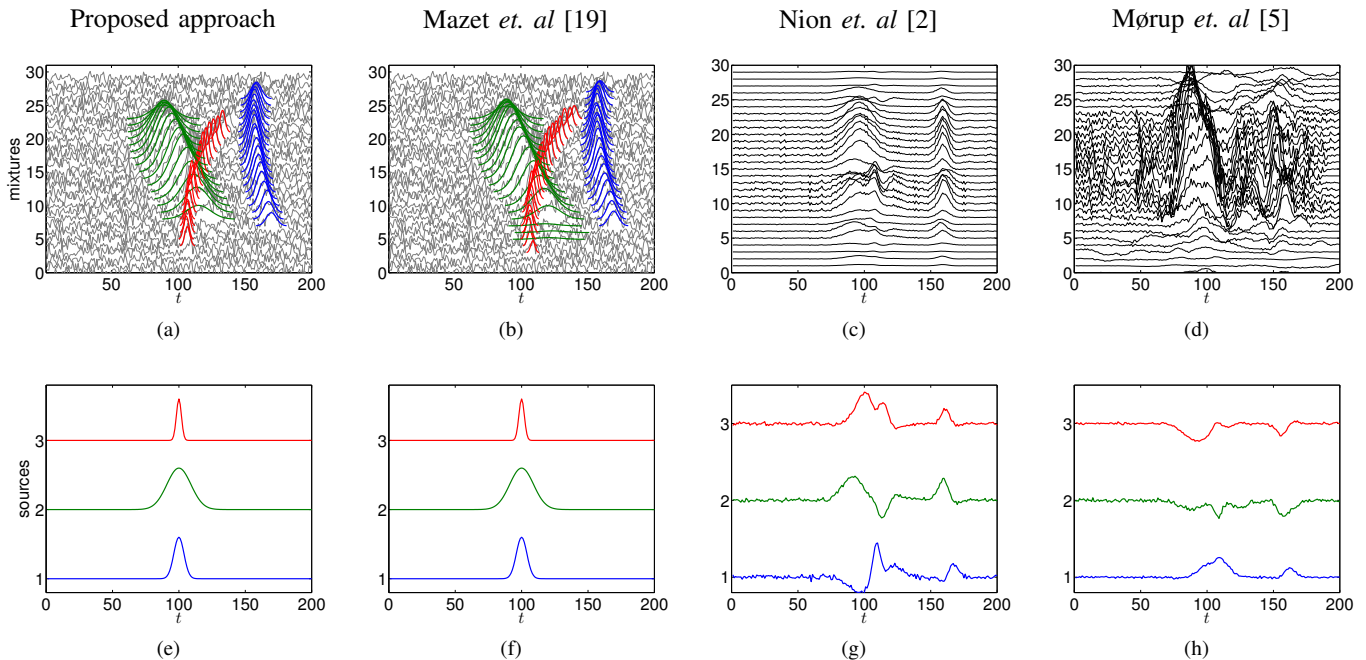


Fig. 1. Example of delayed source separation. Fig.(a)-(b) show the synthetic data (gray) and the sources with estimated parameters (colored) by the proposed approach (continuous delay estimation) and Mazet *et. al* [19] method. (c)(d): mixture reconstruction by the methods of Nion *et. al* [2] and Mørup *et. al* [5]. The bottom sub-figures display the estimated sources for each of these methods.

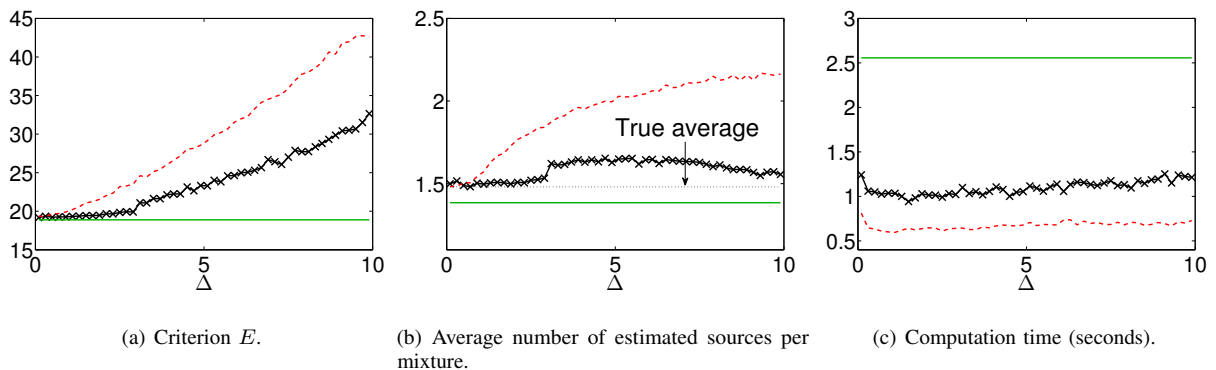


Fig. 2. Comparison of both ALS versions with the algorithm of [19]. Scores are averaged over 50 trials. Red line (- -): discrete version; black line (×): continuous version; green line (solid line): algorithm of [19].

contrary to [2], [5]. 50 datasets with random parameters were generated, each with $I = 40$ mixtures, an average SNR equal to 15 dB, $J = 3$ sources. The stopping criterion threshold is set to $\rho = 10^{-6}$. The data were processed for different grid steps Δ , except for the algorithm of [19] which does not require to set Δ . Fig. 2 shows averaged results over 50 trials. Similar comparison results were found for different noise levels (SNR $\in [0, 30]$ dB). The algorithms are compared in terms of the value of criterion E , that is, the quality of approximation of the mixture signals [Fig. 2 (a)]. Results show that the continuous version improves the estimation quality compared to the discrete version especially, for the largest Δ 's and gives very close results to [19]. For instance, good performance of the continuous version can be obtained even for a coarse $\Delta = 2.9$. However, the residual norm depends on the number of

estimated peaks (a high number of peaks decreases the residual at the price of an overestimation). The proposed algorithm tends to slightly use more than the actual number of peaks needed whereas the algorithm of [19] tends to underestimate this number, as shown in Fig. 2 (b). The continuous version is more expensive in time than the discrete one [Fig. 2 (c)] because of the extra interpolation step. Obviously, the computation time increases with $\frac{1}{\Delta}$. Nevertheless, both proposed algorithms are much faster than the algorithm of [19] (≈ 2.5 times faster for all the values of Δ).

VI. CONCLUSION

This paper presents an ALS strategy to deal with the delayed source separation of parameterized sources. The source estimation step is recast to their shape parameter estimation.

For the challenging delay estimation, we extend the sparse approximation algorithm to the case of numerous mixtures and multiple sources. This algorithm allows continuous delay estimation, giving more accurate estimates than dictionary based approaches (discrete estimation). The proposed ALS method with continuous delay estimation outperforms state-of-the-art delayed source separation methods when sources are highly correlated. Furthermore, it is as effective as the best competitors with much better computation time. In practice, it seems that setting only one sub-iteration of the Levenberg-Marquardt algorithm (line 4 in Algorithm 1) is sufficient to provide accurate outputs; The rationale behind this choice is that in the first iterations, the estimates \hat{A} and \hat{C} are not accurate, therefore there is no need to compute a precise solution of step 4 of Algorithm 1. As a perspective, we would like to include a smooth regularization on the delays to model their slow evolution throughout the mixtures. Another perspective is to extend the proposed algorithm to deal with sources whose shape parameters evolve through the mixtures. These two situations arise in multispectral images of galaxies.

REFERENCES

- [1] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. Academic Press, 2010.
- [2] D. Nion, B. Vandewoestyne, S. Vanaverbeke, K. Abeele, H. Gersem, and L. D. Lathauwer, "A time-frequency technique for blind separation and localization of pure delayed sources," in *LVA/ICA 2010*.
- [3] W. Ouedraogo, B. Nicolas, B. Oudompheng, J. I. Mars, and C. Jutten, "A frequency method for blind separation of an anechoic mixture," in *Proc. EUSIPCO*, 2014.
- [4] E. Villeneuve and H. Carfantan, "Nonlinear deconvolution of hyperspectral data with MCMC for studying the kinematics of galaxies," *IEEE Trans. Image Process.*, vol. 23, pp. 4322–4335, 2014.
- [5] M. Mørup, K. H. Madsen, and L. K. Hansen, "Shifted independent component analysis," in *ICA*, 2007.
- [6] G. Chabriel and J. Barrère, "An instantaneous formulation of mixtures for blind separation of propagating waves," *IEEE Trans. Signal Process.*, vol. 54, pp. 49–58, 2006.
- [7] N. Jiang and D. Farina, "Covariance and time-scale methods for blind separation of delayed sources," *IEEE Trans. on Biomedical Engineering*, vol. 58, pp. 550–556, 2011.
- [8] A. Yeredor, "Blind source separation with pure delay mixtures," in *ICA*, 2001.
- [9] L. Omlor and M. A. Giese, "Anechoic blind source separation using Wigner marginals," *J. Mach. Learn. Res.*, vol. 12, pp. 1111–1148, 2011.
- [10] Ö. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Trans. Signal Process.*, vol. 52, pp. 1830–1847, 2004.
- [11] S. Moussaoui, D. Brie, A. Mohammad-Djafari, and C. Carteret, "Separation of non-negative mixture of non-negative sources using a Bayesian approach and MCMC sampling," *IEEE Trans. Signal Process.*, vol. 54, pp. 4133–4145, 2006.
- [12] J. Duan, "Restoration and separation of piecewise polynomial signals. Application to Atomic Force Microscopy," PhD thesis, Université Henri Poincaré, Nancy, 2010.
- [13] W. Naanaa and J. M. Nuzillard, "Blind source separation of positive and partially correlated data," *Signal Process.*, vol. 85, pp. 1711–1722, 2005.
- [14] J. Bobin, J. Rapin, A. Larue, and J.-L. Starck, "Sparsity and adaptivity for the blind separation of partially correlated sources," *IEEE Trans. Signal Process.*, vol. 63, pp. 1199–1213, 2015.
- [15] C. Ekanadham, D. Tranchina, and E. P. Simoncelli, "Recovery of sparse translation-invariant signals with continuous basis pursuit," *IEEE Trans. Signal Process.*, vol. 59, pp. 4735–4744, 2011.
- [16] K. Fyhn, M. F. Duarte, and S. H. Jensen, "Compressive parameter estimation for sparse translation-invariant signals using polar interpolation," *IEEE Trans. Signal Process.*, vol. 63, pp. 870–881, 2015.
- [17] J. De Leeuw, "Convergence of the majorization method for multidimensional scaling," *J. Classification*, vol. 5, 1988.
- [18] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Asilomar Conf. on Signals, Systems and Computers*, vol. 1, 1993, pp. 40–44.
- [19] V. Mazet, S. Faisan, S. Awali, M. Gaveau, and L. Poisson, "Unsupervised joint decomposition of a spectroscopic signal sequence," *Signal Process.*, vol. 109, pp. 193–205, 2015.