

Two Approaches for the Blind Identification of Cyclo-Stationary Signals Mixtures

Saloua Rhioui, Nadège Thirion-Moreau and Eric Moreau
 STD, ISITV, Université du Sud Toulon-Var, Avenue George Pompidou,
 BP 56 F-83162, La Valette du Var, Cedex, France
 E-mail: {rhioui, thirion, moreau}@univ.tln.fr

Abstract—This article addresses the problem of the blind identification of the mixing matrix in the case of a possibly under-determined instantaneous linear mixture of sources. The considered input signals are cyclo-stationary processes with unknown cyclic frequencies. We propose a new method consisting of the application of a particular linear operator on the correlation matrix of the observations. Then, taking advantage of the properties of the above transformed matrix, a set of rank-one matrices can be built. Combined with a classification procedure, it makes it possible to estimate the different columns of the wanted mixing matrix. This approach is also compared with the classical PARAFAC decomposition approach. Finally, computer simulations are provided in order to illustrate the behavior and the usefulness of the two proposed approaches in the context of digital communications.

Index Terms—Blind identification, under-determined mixtures, cyclo-stationary signals, second order statistics, PARAFAC decomposition, classification.

I. INTRODUCTION

We consider the blind identification of instantaneous mixtures of signals called sources. This problem finds numerous applications in various fields of engineering and applied sciences among which data communications, geophysical prospecting, astrophysics, radar, sonar, speech processing, biomedical, mechanics, etc... This problem can be simply formulated as follows: several linear mixtures of different signals called sources are observed. The purpose is then to identify the mixing system. Hence, it must be done through the observations only, which is the reason why this problem is often qualified as “blind” or “unsupervised”. Most of the approaches that are encountered in the literature assume that the sources are random independent stationary processes. Comparatively, very few works are dedicated to the case of non stationary signals and more particularly to the case of cyclo-stationary ones.

This article deals with the problem of the blind identification (eventually the blind separation) of linear mixtures of independent linearly modulated signals stemming from unknown digital communication systems. In

such a context, the baud-rates of the various transmitted signals are unknown and possibly different. Therefore, the sampled versions of the received signals belong to a particular class of non stationary signals since they appear to be cyclo-stationary sequences. The concept of cyclo-stationarity has been first introduced in array processing by Gardner [10]. It has proven to be useful for the modelling of communication signals and it has led to many breakthroughs in that field [1] [8] [9] [14]. Most of these methods, by taking into account the very specific statistical properties of the communications signals [8] [9] and the potentially knowledge of their different cyclic frequencies, generalize techniques that were developed in the context of stationary signals.

In [14], it is pointed out that the contrast function (involving higher order statistics) to be maximized cannot, in general, be estimated consistently if the cyclic frequencies of the second order statistics of the observations are unknown. However, it was shown in [8] [9] that if the second order statistics of various source signals do not share the same cyclic frequencies, then the knowledge of these latter ones is no more required. Let us notice that our developments originate from the works presented in [1]. If we are taking advantage of the same properties of a transformed correlation matrix, it is, however, performed in a rather different manner. Indeed, in [1], the number of observations is assumed to be greater than or equal to the number of sources and, moreover, the derived algorithm is based on the optimization of a given contrast function. Our approach is more direct and tackle the under-determined case. For quite obvious reasons, two different issues are considered: first, the mixing matrix identification problem is studied, then whenever it is possible *i.e.* in the over-determined case only, the source signals extraction problem is addressed. The main objective of this article is to show that the mixing matrix can be estimated in a wide context meaning that its column vectors can be nevertheless estimated whatever the number of observations is (if it is greater than or equal to two) and even if the second order cyclic frequencies of the sources are unknown but different (even though their cyclic frequencies may sometimes share the same frequency bins). That is why the under-determined case can be considered. This is carried out fully exploiting the particular structure of the correlation matrix after the application of a particular linear transformation and

Parts of this work has been presented in the IEEE International Conference on Acoustic Speech and Signal Processing (ICASSP'2006), Toulouse, France, May 2006. See reference [18].

combining it with an automatic hierarchical ascendent classification procedure [11] [21]. Such an approach is also compared with the PARAFAC decomposition one. Finally, computer simulations are provided in order to illustrate the behavior and the usefulness of the two proposed approaches in the context of digital telecommunications.

This article is organized as follows. After a brief introduction, we state the considered problem in the Section II. The important properties upon which the proposed method is based are also established in this section. The Section III is dedicated to the mixing matrix identification problem in both cases: when the cyclic frequencies are assumed known and when they are unknown. A novel algorithm is then derived: it combines a rank-one matrices selection procedure and a classification algorithm. The robustness of the proposed algorithm versus an additional noise is also studied. The Section IV is devoted to the presentation of an alternative to this method: the PARAFAC decomposition. In the following Section V, computer simulations are provided in order to illustrate the behavior and the usefulness of the two proposed approaches. Finally, in the last Section VI, a general discussion and a conclusion are proposed.

II. PROBLEM STATEMENT AND PROPOSED APPROACH

A. Problem, Recalls & Assumptions

The classical linear memoryless mixture model is considered. It reads

$$\mathbf{x}(t) = \mathbf{M}\mathbf{s}(t), \quad (1)$$

where $\mathbf{x}(t)$ is the $(M, 1)$ vector of observations, $\mathbf{s}(t)$ the $(N, 1)$ vector of sources and \mathbf{M} the (M, N) mixing matrix assumed full rank. We assume that M and N belongs to $\mathbb{N} \setminus \{0, 1\}$ and that the columns of \mathbf{M} are 2 by 2 linearly independent.

The source signals are assumed zero-mean and cyclo-stationary. Hence, their autocorrelation functions $R_{s_i}(t, \tau) = \mathbb{E}\{s_i(t)s_i^*(t - \tau)\}$, $i \in \{1, \dots, N\}$ are thus periodic in t with a period $T_i \in \mathbb{R}^{+*}$, $i \in \{1, \dots, N\}$ respectively. $\mathbb{E}\{\cdot\}$ stands for the mathematical expectation operator and T_i stands for the cyclic period of the i -th source signal $s_i(t)$. Hence, $R_{s_i}(t, \tau)$ can be decomposed into Fourier series:

$$R_{s_i}(t, \tau) = \sum_k R_{s_i}^{\text{fs}}[k, \tau] \exp(i2\pi \frac{k}{T_i} t), \quad (2)$$

where $i^2 = -1$. $R_{s_i}^{\text{fs}}[k, \tau]$ stands for the cyclic correlation function (coefficient of the Fourier series expansion) defined as

$$R_{s_i}^{\text{fs}}[k, \tau] = \frac{1}{T_i} \int_{-\frac{T_i}{2}}^{\frac{T_i}{2}} R_{s_i}(t, \tau) \exp(-2i\pi \frac{k}{T_i} t) dt. \quad (3)$$

In the following, we assume that the sources are uncorrelated and that the cyclic periods are different two by two, i.e. $T_i \neq T_j$, $\forall i, j \in \{1, \dots, N\}$ and $i \neq j$. We

define the set \mathcal{V}_i of the whole cyclic frequencies of the i -th source signal as

$$\mathcal{V}_i = \left\{ \nu_i = \frac{k}{T_i}, k \in \mathbb{Z} \right\}.$$

The correlation matrix $\mathbf{R}_x(t, \tau)$ of the observations $\mathbf{x}(t)$ is defined as

$$\mathbf{R}_x(t, \tau) = \mathbb{E}\{\mathbf{x}(t)\mathbf{x}^H(t - \tau)\}, \quad (4)$$

where $(\cdot)^H$ stands for the conjugate transpose operator. Using (1), it is easily seen that the correlation matrix in (4) admits the following decomposition

$$\mathbf{R}_x(t, \tau) = \mathbf{M}\mathbf{R}_s(t, \tau)\mathbf{M}^H, \quad (5)$$

where $\mathbf{R}_s(t, \tau)$ is the correlation matrix of the sources.

B. The proposed approach

Let us now define the following linear operator $(\cdot)^{\text{fs}}$ which operates on the matrix argument component wise

$$\mathbf{R}_x^{\text{fs}}(\nu, \tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbf{R}_x(t, \tau) \exp(-2i\pi\nu t) dt. \quad (6)$$

Since this operator is linear, using (5) in (6), we directly have

$$\mathbf{R}_x^{\text{fs}}(\nu, \tau) = \mathbf{M}\mathbf{R}_s^{\text{fs}}(\nu, \tau)\mathbf{M}^H, \quad (7)$$

where $\mathbf{R}_s^{\text{fs}}(\nu, \tau)$ is defined similarly to $\mathbf{R}_x^{\text{fs}}(\nu, \tau)$ in (6).

As the sources are uncorrelated signals, $\mathbf{R}_s(t, \tau)$ is diagonal for all t and τ . Thus, it is also the case for the matrix $\mathbf{R}_s^{\text{fs}}(\nu, \tau)$ for all ν and τ . Now, using the fact that the source signals have distinct cyclic periods, then, there exists values of ν for which $\mathbf{R}_s^{\text{fs}}(\nu, \tau)$ has a particular structure. It is given by the following proposition.

Proposition 1: For frequencies ν_o such that:

$$\text{H1. } \nu_o \in \mathcal{V}_i, \nu_o \notin \bigcup_{j, j \neq i} \mathcal{V}_j \text{ and } R_{s_i}^{\text{fs}}(\nu_o, \tau) \neq 0,$$

the diagonal matrix $\mathbf{R}_s^{\text{fs}}(\nu_o, \tau)$ possesses only one non null element at position (i, i) .

Proof: We have $\forall \nu_i \in \mathcal{V}_i, \tau \in \mathbb{R}, \forall j \in \{1, \dots, N\}$, we have:

$$\begin{aligned} R_{s_j}^{\text{fs}}(\nu_i, \tau) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_{s_j}(t, \tau) \exp(-2i\pi\nu_i t) dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \sum_k R_{s_j}^{\text{fs}}[k, \tau] \exp(2i\pi(\nu_j - \nu_i)t) dt \\ &= \sum_k R_{s_j}^{\text{fs}}[k, \tau] \underbrace{\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \exp(2i\pi(\nu_j - \nu_i)t) dt}_{\alpha(\nu_j - \nu_i)} \\ &= \sum_k R_{s_j}^{\text{fs}}[k, \tau] \alpha(\nu_j - \nu_i), \end{aligned} \quad (8)$$

where $\alpha(\nu) = 1$ if $\nu = 0$ and 0 otherwise. Thus, for such a frequency value, the diagonal matrix $\mathbf{R}_s^{\text{fs}}(\nu, \tau)$ has only one non null element that is the one at the position (i, i) .

We can now propose the following main proposition. ■

Proposition 2: For all cyclic frequencies ν and values of τ such that assumption H1 is satisfied, we have

$$\mathbf{R}_x^{\text{fs}}(\nu, \tau) = R_{s_i}^{\text{fs}}(\nu, \tau) \mathbf{m}_i \mathbf{m}_i^H \quad (9)$$

where \mathbf{m}_i stands for the i -th column vector of the matrix \mathbf{M} .

Hence, for all values of τ and for all values of ν satisfying property H1, $\mathbf{R}_x^{\text{fs}}(\nu, \tau)$ is a rank one matrix and the i -th column vector \mathbf{m}_i of the matrix \mathbf{M} can be estimated (up to the multiplication by a scalar coefficient) using an Eigenvalue Decomposition (EVD). Vector \mathbf{m}_i is nothing else than the eigenvector associated to the largest eigenvalue value of $\mathbf{R}_x^{\text{fs}}(\nu, \tau)$.

III. ALGORITHM

First of all, it has to be noticed that the knowledge of the cyclic frequencies is not necessary to achieve the identification of the mixing system (see III-B), even though such a knowledge considerably simplifies this task.

A. Known or estimated cyclic frequencies

When the cyclic frequencies $\nu_i, \forall i \in \{1, \dots, N\}$ are known, the EVD of the transformed correlation matrix $\mathbf{R}_x^{\text{fs}}(\nu_i, \tau)$ directly provides an estimate of the i -th column of the mixing matrix \mathbf{M} corresponding to the i -th source whose cyclic period is T_i . This is due to (9).

When the cyclic periods are unknown, it still remains possible to try to estimate them thanks to one of the classical estimators suggested in [3] [19] or [13] for example. The first of these estimators was proposed by Giannakis in [3]. It is based on the second order statistics and it consists, for a given signal $x(t)$, in evaluating for each frequency ν , the value of $R_x^{\text{fs}}(\nu, \tau)$, which is always null except when ν belongs to the set of the cyclic frequencies of $x(t)$. By collecting several values of $R_x^{\text{fs}}(\nu, \tau)$ for $\tau \in \{-D, \dots, D\}$ in a vector

$$\mathbf{R}_x^{\text{fs}}(\nu) = (R_x^{\text{fs}}(\nu, -D), \dots, R_x^{\text{fs}}(\nu, D)),$$

the estimator is then defined as

$$\mathcal{E}(\nu) = \|\mathbf{R}_x^{\text{fs}}(\nu)\|^2. \quad (10)$$

The estimated cyclic frequency corresponds to the maximum of the argument of $\mathcal{E}(\nu)$. The problem with such an approach is that, generally, the estimation error regarding the mixing system tends to increase when the different cyclic frequencies are just estimated and no more known. This observation explains the seek for new methods that no more explicitly rely upon the *a priori* knowledge of the different cyclic frequencies justifying the new approach that we propose.

B. Unknown cyclic frequencies: a new approach

The new method that we propose does not explicitly require to estimate the different cyclic frequencies (even though it enables their estimation). It can be decomposed into the following stages:

- Estimation of the transformed autocorrelation matrix $\mathbf{R}_x^{\text{fs}}(\nu, \tau)$ which is calculated for a sufficiently large number of frequency bins.
- Building of a rank-one matrices set thanks to the selection procedure described in Section III-B.1.
- Calculation of the corresponding estimations of the mixing matrix column vectors thanks to the properties established in Section II-B and to the resulting EVD based method suggested.
- Seek for an unique estimation of each columns of the mixing matrix thanks to the automatic ascendent hierarchical classification procedure described in Section III-B.2.
- In the over-determined case, restitution of the unknown sources using the estimated mixing matrix.

1) *Building of the rank-one matrices set:* The transformed correlation matrix $\mathbf{R}_x^{\text{fs}}(\nu, \tau)$ is calculated for a sufficiently large number of frequency bins in order to find a wide range of cyclic frequencies corresponding to the different source signals. As noticed in Section II, (9) holds for rank-one matrices only. That is why we have to start with the automatic selection of rank-one matrices.

In practice, it is performed thanks to a decision procedure implying the introduction of one or several thresholds. By EVD (because of the definite positiveness and Hermitian symmetry of the matrices $\mathbf{R}_x^{\text{fs}}(\nu, \tau)$), we have $\mathbf{R}_x^{\text{fs}}(\nu, \tau) = \mathbf{V}(\nu, \tau) \mathbf{\Lambda}(\nu, \tau) \mathbf{V}^H(\nu, \tau)$, with $\mathbf{V}(\nu, \tau)$ a unitary $M \times M$ matrix and $\mathbf{\Lambda}(\nu, \tau) = \text{Diag}\{\boldsymbol{\lambda}(\nu, \tau)\}$ a diagonal matrix with positive elements. $\boldsymbol{\lambda}(\nu, \tau) = (\lambda_1(\nu, \tau), \dots, \lambda_M(\nu, \tau))^T$, the vector of the diagonal elements of $\mathbf{\Lambda}(\nu, \tau)$. Assuming that the singular values are stacked in the decreasing order *i.e.* $\lambda_1(\nu, \tau) \geq \lambda_2(\nu, \tau) \geq \dots \geq \lambda_M(\nu, \tau) \geq 0$, then a first way to check whether a matrix is a rank-one matrix or not is given by:

Keep matrix $\mathbf{R}_x^{\text{fs}}(\nu, \tau)$ if

$$\begin{cases} \lambda_1(\nu, \tau) > \varepsilon_1 \\ \lambda_i(\nu, \tau) < \varepsilon_2 \quad \forall i \in \{2, \dots, M\} \end{cases} \quad (11)$$

where ε_1 is a positive (sufficiently) high constant and ε_2 is a positive (sufficiently) small constant.

Another possible way to proceed is provided by:

Keep matrix $\mathbf{R}_x^{\text{fs}}(\nu, \tau)$ if

$$\begin{cases} \mathcal{C}(\nu, \tau) = \frac{\lambda_1(\nu, \tau)}{\sum_{k=1}^M \lambda_k(\nu, \tau)} > 1 - \varepsilon_3 \\ \sum_{k=1}^M \lambda_k(\nu, \tau) > \varepsilon_4 \end{cases} \quad (12)$$

with ε_3 and ε_4 small positive constants.

Each of these two decision rules enables the building of a set of rank-one matrices $\mathbf{R}_x^{\text{fs}}(\nu_k, \tau)$. In the following the resulting set of matrices will denoted \mathcal{R} . It is defined as $\mathcal{R} = \{\mathbf{R}_x^{\text{fs}}(\nu_k, \tau), k = 1, \dots, K\}$ and its dimension is assumed to be equal to K . Notice that the corresponding frequencies $\{\nu_k, k = 1, \dots, K\}$ happened to be the cyclic frequencies.

2) *Classification/Estimation procedure*: Whatever the matrix belonging to \mathcal{R} , the eigenvector associated to its largest eigenvalue in a EVD decomposition corresponds to an estimate (up to a constant coefficient) of one of the columns of the mixing matrix. It is obvious that the set of the considered frequencies has to be sufficiently important in order to provide estimates of all the columns of the mixing matrix.

We assume that the dimension K of the set \mathcal{R} (see Section III-B.1) is such that $K = K_1 + K_2 + \dots + K_N$, $(K_1, \dots, K_N) \in (\mathbb{N}^*)^N$. $K_i, \forall i = \{1, \dots, N\}$ represents the number of matrices corresponding to i -th of the N sources. In theory, the corresponding $K_i, \forall i = \{1, \dots, N\}$, eigenvectors used in order to identify the i -th column of the mixing matrix should be all equal or at least collinear.

From a practical point of view, some issues still have to be discussed. The first one stems from the fact that since matrices $\mathbf{R}_x^{\text{fs}}(\nu, \tau)$ are “only” estimated, the different vectors relative to a given column of \mathbf{M} are not equal or very likely not collinear. In fact, there is a certain dispersion around their theoretical value. Moreover, the introduction of thresholds into the automatic selection procedure of rank-one matrices also induces possible errors: matrices that should not have been selected may have been added to \mathcal{R} . Finally, if we know that a selected frequency is assigned to a particular column vector, we do not know which one it is.

To solve the above problems, we propose to use an automatic hierarchical ascendent classification. A classical procedure consists of computing the centres of inertia of each clusters of points by merging to the nearest point in order to finally obtain one single point. This automatic hierarchical ascendent classification method is known under the name of unweighed pair-group method of aggregation using arithmetic averages, see e.g. [21].

Basically, weights are assigned to each points representing the number of already amalgamated points (weights being initialized to 1). The euclidian distance between two points is then measured. If this distance is lower than a given threshold, say δ , the weighted average of these two points replaces them. The new points's weight corresponds to the sum of the weights of the old ones. When a point does not have neighbours any more it is considered as a centre of inertia.

Considering the two first problems, among all the found inertia centres, only the N ones corresponding to the strongest weights are selected. The coordinates of these N centres of inertia provide estimates of the N columns of the mixing matrix \mathbf{M} , up to their order. This does not have any importance with regard to our estimation problem for which a column order indetermination remains inherent. We can finally remark that such an approach could also make it possible to estimate the number of sources (and the cyclic frequencies) when they are assumed unknown. The number of source signals corresponds to the number of centres of inertia having a sufficient weight. The cyclic frequencies can be estimated from (non zero) rank-one

matrices $\mathbf{R}_x^{\text{fs}}(\nu, \tau)$ for various values of ν .

C. Noisy case

In the noisy case, the mixing model studied in (1) is rewritten:

$$\mathbf{x}(t) = \mathbf{M}\mathbf{s}(t) + \mathbf{b}(t). \quad (13)$$

In the following, two different kind of noises $\mathbf{b}(t)$ will be considered. First, the additive noise is assumed to be a (Gaussian) stationary signal, then it is supposed to be a standard cyclo-stationary noise.

- **Stationary noise**: applying the same linear operator $(\cdot)^{\text{fs}}$ previously defined in (6), onto the correlation matrix of the observations built from the mixing equation (13), one easily obtains the following result:

$$\begin{aligned} \mathbf{R}_x^{\text{fs}}(\nu, \tau) &= \mathbf{M}\mathbf{R}_s^{\text{fs}}(\nu, \tau)\mathbf{M}^H + \mathbf{R}_b^{\text{fs}}(\nu, \tau) \\ &= \mathbf{M}\mathbf{R}_s^{\text{fs}}(\nu, \tau)\mathbf{M}^H + \mathbf{R}_b(\tau) \times \\ &\quad \underbrace{\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \exp(2i\pi\nu t) dt}_{\alpha(\nu)} \\ &= \mathbf{M}\mathbf{R}_s^{\text{fs}}(\nu, \tau)\mathbf{M}^H + \mathbf{R}_b(\tau)\alpha(\nu), \end{aligned}$$

with $\alpha(\nu) = 1$ if $\nu = 0$ and else 0, $\mathbf{R}_b^{\text{fs}}(\nu, \tau)$ is the transformed noise correlation matrix of the noise correlation matrix $\mathbf{R}_b(\tau) = \mathbb{E}\{\mathbf{b}(t)\mathbf{b}^H(t - \tau)\}$. It is easily seen that the method we propose to identify the mixing matrix is not disturbed by the addition of a stationary noise provided that the sources cyclic frequencies are not null.

- **Cyclostationary noise**: in such a case, the noise plays the role of an additional source in the initial mixture which is not a problem since our approach tackles the under-determined mixtures case. However, it is only possible if the cyclic frequency of the noise is different from those of the initial sources. If this assumption is not satisfied, a joint-diagonalization procedure still remains possible to solve the identification problem because of (7), but only under the condition that the considered problem remains an over-determined one.

IV. ANOTHER POSSIBLE APPROACH: THE PARAFAC DECOMPOSITION

As previously stated, in the over-determined case, one possible way to approach the linear decomposition given by (7) for all the frequencies, is to use a joint-diagonalization algorithm. One can find in the literature some joint-diagonalization algorithms under non unitary transformation [4] [7] [22] [23] [24] [25].

According to [5] [6] [16], the problem of the joint-diagonalization of a given set of matrices is connected to the PARAFAC decomposition problem. This last one enables, however, to consider the under-determined case.

Let τ be fixed, the K ($M \times M$) matrices $\mathbf{R}_x^{\text{fs}}(\nu_i, \tau)$, $i \in \{1, \dots, K\}$ defined in (7) are stacked in a third order tensor $\mathcal{A}(\tau) \in \mathbb{R}^{M \times M \times K\tau}$. The matrix $\mathbf{D}(\tau) \in \mathbb{R}^{K\tau \times N}$ is then defined as follows:

$$\mathbf{D}(\tau) = \begin{pmatrix} \text{diag}\{\mathbf{R}_s^{\text{fs}}(\nu_1, \tau)\} \\ \vdots \\ \text{diag}\{\mathbf{R}_s^{\text{fs}}(\nu_K, \tau)\} \end{pmatrix}$$

where $\text{diag}\{\cdot\}$ is the operator that extracts the diagonal elements of the matrix passed in argument and puts them into a $1 \times N$ vector. Then, it can be shown that $\mathcal{A}(\tau)$ can be written as

$$\mathcal{A}_{ijk} = \sum_{f=1}^N m_{if} d_{jf} m_{kf}, \quad (14)$$

where $m_{if} = (\mathbf{M})_{if}$ and $d_{jf} = (\mathbf{D}(\tau))_{jf}$. The equation (14) is recognized as the PARAFAC decomposition of the considered tensor $\mathcal{A}(\tau)$ [2] [16] [20].

It has been shown that when the decomposition is unique, the mixing matrix \mathbf{M} can be determined (up to the classical scale and permutation indetermination) even for under-determined mixtures. The uniqueness of the rank- N decomposition of the third order tensor $\mathcal{A}(\tau)$ is ensured, in our case, if $2(N + 1) \leq 2K_{\mathbf{M}} + K_{\mathbf{D}(\tau)}$ [2] [16] [20], where $K_{\mathbf{M}}$ and $K_{\mathbf{D}(\tau)}$ are the Kruskal ranks¹ of matrices \mathbf{M} and $\mathbf{D}(\tau)$ respectively. Hence, the identifiability condition reads:

- $K_{\mathbf{D}(\tau)} \geq 2(N - M + 1)$ in the under-determined case ($M < N$),
- $K_{\mathbf{D}(\tau)} \geq 2$ in the over-determined case ($M \geq N$).

Provided the decomposition uniqueness, one can find in the literature several methods to determine the decomposition. The main issue is to optimize (in the trilinear case) the following criterion [12]:

$$\mathcal{F}(\mathbf{U}, \mathbf{V}, \mathbf{W}) = \|\mathcal{T} - \sum_{f=1}^N \mathbf{u}_f \circ \mathbf{v}_f \circ \mathbf{w}_f\|_{\mathbb{F}}^2, \quad (15)$$

where $\|\cdot\|_{\mathbb{F}}$ is the Frobenius norm, defined as $\|\mathbf{A}\|_{\mathbb{F}}^2 = \sum_{i,j=1}^N |a_{ij}|^2$ with $a_{ij} = (\mathbf{A})_{ij}$. \mathbf{u}_f , \mathbf{v}_f and \mathbf{w}_f are the f -th columns of matrices \mathbf{U} , \mathbf{V} and \mathbf{W} respectively. \mathcal{T} is a third-order tensor (in our case \mathcal{T} is equal to $\mathcal{A}(\tau)$), $\mathbf{U} = \mathbf{W} = \mathbf{M}$ and $\mathbf{V} = \mathbf{D}(\tau)$ and the product \circ is defined as $\mathbf{u}_f \circ \mathbf{v}_f \circ \mathbf{w}_f = \sum_i u_{f,i} v_{f,i} w_{f,i}$ where $u_{f,i}$, $v_{f,i}$ and $w_{f,i}$ are the i -th component of vectors \mathbf{u}_f , \mathbf{v}_f and \mathbf{w}_f respectively.

The ALS algorithm (Alternating Least Squares) [15] makes it possible to carry out the minimization of the preceding criterion simply. Its guiding principle consists in minimizing \mathcal{F} by fixing successively two of the three matrices \mathbf{U} , \mathbf{V} and \mathbf{W} in order to estimate the third one. This operation is then repeated in an iterative alternating scheme.

¹The Kruskal rank of a given matrix \mathbf{A} is defined as the maximum integer k , such that any k columns drawn from $\mathbf{A}(\tau)$ are linearly independent. Generally, a matrix is full rank and Kruskal rank. Hence, in practice, $K_{\mathbf{M}} = \min(M, N)$ and $K_{\mathbf{D}(\tau)} = \min(K, N)$, where $\min(\cdot)$ denotes the minimum.

V. COMPUTER SIMULATIONS

The behaviour of the two proposed approaches is now illustrated thanks to computer simulations. We consider three mixing matrices. The first one

$$\mathbf{M}_1 = \begin{bmatrix} 1 & -0.3 & 0.6 \\ 0.6 & 1 & -0.8 \\ -0.5 & 0 & 1 \end{bmatrix}$$

corresponds to a square mixture ($M = 3, N = 3$), the second one

$$\mathbf{M}_2 = \begin{bmatrix} 1 & -0.3 & 0.6 & 0.2 \\ 0.6 & 1 & -0.8 & -0.1 \\ -0.5 & 0 & 1 & 1 \\ -0.3 & -0.6 & 0.1 & 1 \end{bmatrix}$$

corresponds to a square mixture ($M = 4, N = 4$), and the third one

$$\mathbf{M}_3 = \begin{bmatrix} 1 & -0.3 & 0.6 & 0.2 \\ 0.6 & 1 & -0.8 & -0.1 \\ -0.5 & 0 & 1 & 1 \end{bmatrix}$$

corresponds to an under-determined mixture ($M = 3, N = 4$). For simplicity reasons, we consider discrete time source signals described by the following model

$$s(n) = \sum_{k \in \mathbb{Z}} a(k)h(n - kT) \quad \forall n \in \{1, \dots, S\} \quad (16)$$

where $a(n)$ is an i.i.d. zero-mean random sequence referred to as the transmitted symbols, T is an integer related to the period symbol, $h(n)$ is a deterministic waveform signal and S is the length of the signal. In all cases, $a(n)$ is chosen to take values in the $\{-1, 1\}$ set with equal probabilities. The waveform $h(n)$ is chosen triangular. It is defined for an even cyclic period as:

$$\begin{cases} h(n) = \frac{2}{T}n & \text{if } 0 \leq n \leq \frac{T}{2} \\ h(n) = -\frac{2}{T}n + 2 & \text{if } \frac{T}{2} + 1 \leq n \leq T - 1 \\ h(n) = 0 & \text{otherwise.} \end{cases} \quad (17)$$

Finally, in the $N = 3$ case (corresponding to mixing matrix \mathbf{M}_1), the cyclic period of the three considered sources are respectively $T_1 = 4, T_2 = 10$ and $T_3 = 6$. While in the $N = 4$ case (mixing matrices \mathbf{M}_2 or \mathbf{M}_3), the cyclic period of the four considered sources are respectively $T_1 = 4, T_2 = 6, T_3 = 10$ and $T_4 = 14$.

First of all, the rank-one matrices set \mathcal{R} whose dimension is K has to be built: $\mathcal{R} = \{\mathbf{R}_x^{\text{fs}}(\nu_k, \tau), k = 1, \dots, K\}$. In Tab. I, we have given the evolution of the number of the kept matrices and the obtained performances versus the values of the two thresholds involved in the rank-one matrices detection procedure for one realization of the square 3×3 case. One can check that the results are quite similar with both detectors. That is why we have chosen, here, to use the first detector (11). With $\varepsilon_1 = 0.2$ and $\varepsilon_2 = 0.07$ and 100 Monte-Carlo runs, we have found $K = 77 \times 100$ in the square 3×3 case whereas we have found $K = 266 \times 100$ in the under-determined case with $\varepsilon_1 = 0.1$ and $\varepsilon_2 = 0.2$ and 100 Monte-Carlo runs. On the one hand, with the first method, K vectors can be estimated: actually they are the eigenvectors of matrices

$\mathbf{R}_x^{\text{fs}}(\nu_k, \tau)$ associated to their highest eigenvalue (they are displayed on the top of Fig. 1, in the square 3×3 case (resp. Fig. 6, in the under-determined case)). Then, thanks to the classification procedure, three (respectively four) columns vectors are estimated in the square mixture case (resp. in the under-determined mixture case): they are the centres of inertia of the 3 (resp. 4) clusters of the estimated columns vectors corresponding to the aggregation of the higher number of vectors. With the first method, in the square case, only 89% of the matrices of the initial set \mathcal{R} were kept to find the 3 centres of inertia (18.09% of these matrices correspond to the first column of the mixing matrix, 35.56% to its second column and 35.38% to its third column). For convenience, this subset of the initial set \mathcal{R} will be denoted in the following \mathcal{R}_c . In the under-determined case, 99.94% matrices were kept to find the 4 centres of inertia (22.86% of these matrices correspond to the first column of the mixing matrix, 22.73% to its second column, 33.93% to its third column and 20.42% to its fourth column). They are displayed on the bottom of Fig. 1 (resp. Fig. 6).

On the other hand with the PARAFAC decomposition, 3×100 (resp. 4×100) vectors can be estimated in the square 3×3 case (resp. in the under-determined case), they are stemming from the PARAFAC decomposition of the tensor $\mathcal{A}(\tau)$ (they are displayed on the top Fig. 7, in the under-determined case).

Finally, to quantify the performances of the algorithms the following performance index given in [17, eq. (42)], can be used:

$$I(\mathbf{G}) = \text{ind}_2(\mathbf{G}) \quad (18)$$

where $\mathbf{G} = \widehat{\mathbf{M}}^+ \mathbf{M}$ with $(\cdot)^+$ the Moore-Penrose generalized matrix inverse. This index measures how \mathbf{G} is close to \mathbf{DP} since we are looking for a matrix $\widehat{\mathbf{M}}$ for which there exist a permutation matrix \mathbf{P} and an invertible diagonal matrix \mathbf{D} such that

$$\widehat{\mathbf{M}} = \mathbf{DPM}. \quad (19)$$

When the performance index is given in dB, it is defined as $10 \log(I(\cdot))$.

1) *Square mixture case:* We have used the matrix \mathbf{M}_1 and sources of $S = 4620$ time samples. In Fig. 1, we plot the selected columns and the columns of the mixing system estimated by the EVD + classification based method.

First, we compare, on Fig. 2, the results obtained thanks to EVD when the cyclic frequencies are known and when they are unknown (proposed method combining rank-one detection + EVD + classification). I is displayed versus the (100 Monte-Carlo) realizations which have been preliminary sorted in the decreasing order of the obtained performance. One can observe that the proposed method provides, generally, better results than those obtained with the method based on the knowledge of the cyclic frequencies. This is certainly due to the fact that more matrices are used with our method which enables an average. It also proves that the knowledge of the cyclic frequencies is not necessary to achieve the identification

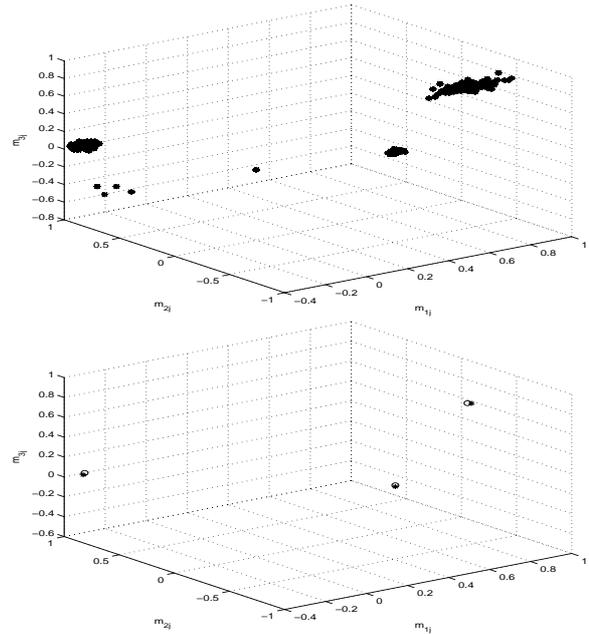


Figure 1. Top: the 77×100 estimated columns of mixing matrix \mathbf{M}_1 before classification. Bottom: the 3 estimated columns of the mixing matrix \mathbf{M}_1 after classification (*) compared with their true value (o).

of the mixing system (and the blind sources separation in this square case).

Secondly, the performance index I have been evaluated over 100 Monte-Carlo runs in order to establish a comparison between the proposed method and the PARAFAC based one. The Monte-Carlo realizations have been sorted in the decreasing order of the obtained performance. The results are displayed on Fig. 3. With regard to the PARAFAC decomposition, two cases have been considered: in the first one, all the matrices of the set \mathcal{R} (whose dimension is $K = 77 \times 100$) are used, in the second one, the results obtained thanks to the classification procedure are used in order to eliminate matrices considered as wrong in set \mathcal{R} leading to the subset \mathcal{R}_c (only 68×100 matrices are kept). We can notice that the proposed approach always provides better results than the PARAFAC based one. We can also remark that a selection of matrices thanks to the classification procedure makes it possible to considerably improve the performance of the method based on the PARAFAC decomposition.

Finally, the mean square errors (MSE) calculated on the sources have been evaluated over 100 Monte-Carlo runs for $\tau = 0$ still to establish a comparison between the proposed method and the PARAFAC based one (using matrices of the subset \mathcal{R}_c instead of those of the set \mathcal{R}). In Fig. 4, we compare the MSE calculated on the sources versus the number S of used samples. One can observe an improvement of the performance with both methods when the length S of the signals is increasing. The results remain better with the proposed approach than with the PARAFAC decomposition.

2) *The noisy case:* We consider a stationary Gaussian noise whose autocorrelation matrix $\mathbf{R}_b(\tau) =$

With first rank-one detector										
Threshold value ϵ_1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.1
Threshold value ϵ_2	0.16	0.15	0.14	0.13	0.12	0.11	0.1	0.09	0.08	0.07
Number of kept matrices	985	985	1293	1293	1293	1293	1493	1493	1493	1493
% of kept matrices	49.25	49.25	64.65	64.65	64.65	64.65	74.70	74.70	74.70	74.70
I in dB	-30.0	-30.0	-29.31	-29.31	-29.31	-29.31	-27.69	-27.69	-27.69	-27.69
With second rank-one detector										
Threshold value ϵ_3	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99	1
Threshold value ϵ_4	0.	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45
Number of kept matrices	1185	1185	985	985	785	785	785	785	785	785
% of kept matrices	59.25	59.25	49.25	49.25	39.25	39.25	39.25	39.25	39.25	39.25
I in dB	-33.18	-33.18	-30.08	-30.08	-34.33	-34.33	-34.33	-34.33	-34.33	-34.33

TABLE I. A COMPARISON OF THE TWO RANK-ONE DETECTORS: NUMBER OF KEPT MATRICES AND OBTAINED PERFORMANCES VERSUS THE VALUES OF THE TWO INVOLVED THRESHOLDS.

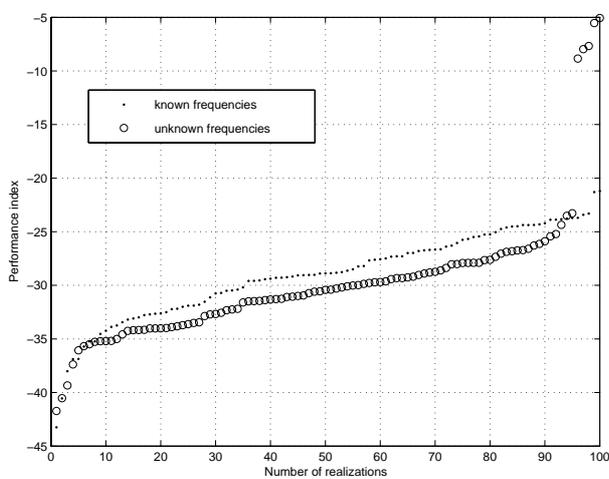


Figure 2. Square case using matrix M_1 : performance index (in dB) versus realizations (chosen among the 100 Monte-Carlo runs), comparison of the results obtained thanks to EVD when the cyclic frequencies are known (\bullet) and when they are unknown (\circ) (first proposed method combining rank-one detection + EVD + classification). The realizations are sorted in the decreasing order of the obtained performance.

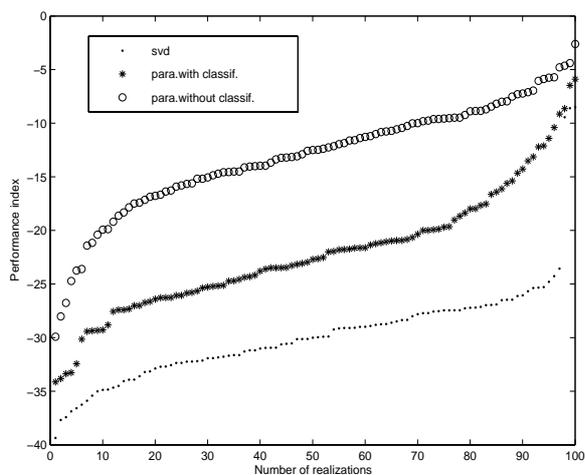


Figure 3. Performance index (in dB) versus realization (chosen among the 100 Monte-Carlo runs). The realizations are sorted in the decreasing order of the obtained performance. A comparison of the results obtained with the EVD + classification based method (\bullet), the PARAFAC decomposition using the whole set \mathcal{R} (\circ), the PARAFAC decomposition using the subset \mathcal{R}_c of the set \mathcal{R} ($*$).

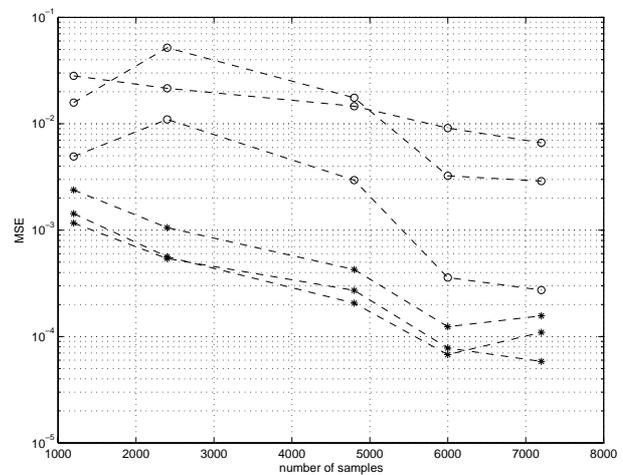


Figure 4. A comparison of the Mean Square Error (MSE) for each of the 3 sources in the over-determined case (matrix M_1 is used) versus the size S of the signals. The PARAFAC method is applied using the subset \mathcal{R}_c of the set \mathcal{R} . The mixing matrix (and then the sources) are estimated thanks to the EVD + classification based method ($*$) and the PARAFAC decomposition (\circ).

$E\{\mathbf{b}(t)\mathbf{b}^H(t - \tau)\} = \sigma_b^2 \mathbf{I}_M$, where \mathbf{I}_M is the $M \times M$ identity matrix. Using the mixing matrix M_2 , we display on the top of Fig. 5 the evolution of the performance index I versus the power of noise σ_b^2 and on its bottom the cumulative function versus the considered noise realization. One can check that the addition of a stationary noise does not affect the proposed method.

3) *Under-determined mixture case:* In Fig. 6, we plot the selected columns and the columns of the mixing system estimated by the EVD + classification based method. In Fig. 7, we plot the columns of the mixing system estimated by the PARAFAC decomposition (before and after average over 100 Monte-Carlo trials).

The MSE of the four estimated columns of the mixing matrix M_3 versus the number of time samples have also been evaluated over 100 Monte-Carlo runs for $\tau = 0$. They are displayed on Fig. 8. As previously observed, the MSE decrease when the number of used time samples increases. The obtained results remain generally better with the proposed approach than with the PARAFAC

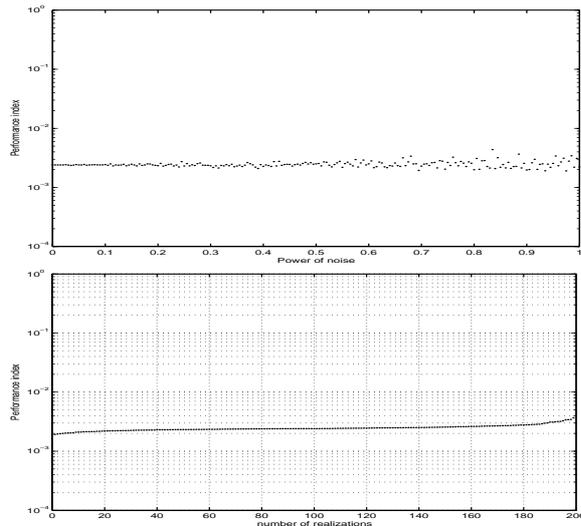


Figure 5. Top: evolution of the performance index I versus the power of noise. Bottom: cumulative function of I versus noise realizations.

decomposition. It also proves that the knowledge of the cyclic frequencies is not necessary to achieve the identification of the mixing system even in the under-determined case.

VI. CONCLUSION

In this article, we have presented a new approach in order to blindly identify the mixing matrix of a possibly under-determined mixture of sources when the inputs are cyclo-stationary signals. This approach operates into three steps: first a linear operator is applied on the observations correlation matrix, then, a rank-one matrices detection procedure is used and finally it involves a hierarchical ascendent classification. This approach has also been compared with a PARAFAC decomposition based method. Computer simulations have been provided to illustrate the good behaviour and the usefulness of the two proposed approaches in the context of digital communications.

REFERENCES

- [1] K. Abed-Meraim, Y. Xiang, J. H. Manton and Y. Hua, "Blind source separation using second-order cyclostationary statistics", *IEEE Transactions on Signal Processing*, vol. 49, no 4, pp. 694-701, April 2001.
- [2] R. Bro, "Multi way analysis in the food industry, models, algorithms and application", *PhD Thesis*, Amsterdam University, 1998.
- [3] A. V. Dandawaté and G. B. Giannakis, "Statistical tests for presence of cyclostationarity". *IEEE Transactions on Signal Processing*, vol. 42, pp. 2355-2369, September 1995.
- [4] S. Dégerine, "Sur la diagonalisation conjointe approchée par un critère des moindres carrés", in *Proc. XVIIème colloque GRETSI*, pp. 311-314, Toulouse, France, Septembre 2001.
- [5] L. De Lathauwer, "Simultaneous matrix diagonalization: the overcomplete case", in *Proc. ICA*, pp. 821-825, Nara, Japan, April 2003.

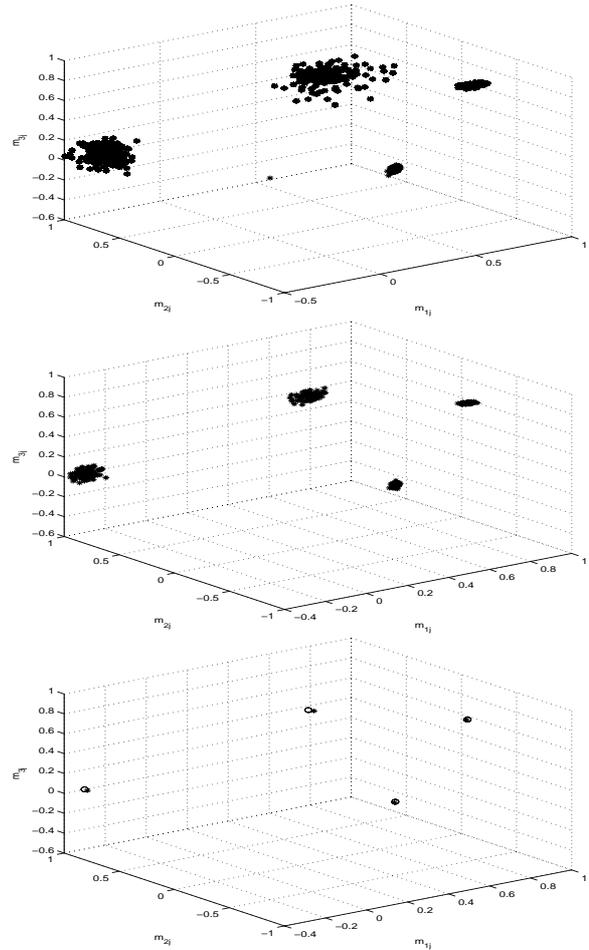


Figure 6. Top: the 266×100 vectors estimated thanks to the EVD of the matrices belonging to \mathcal{R} (column vectors of the mixing matrix M_3 estimated before classification). Middle: the 4×100 columns of the mixing matrix M_3 estimated after classification. Bottom: the 4 estimated columns of the mixing matrix M_3 after classification and an average over 100 Monte-Carlo runs (*) compared with their true value (o).

- [6] L. De Lathauwer, B. De Moor and J. Vandewalle, "Independent component analysis and (simultaneous) third-order tensor diagonalization", *IEEE Transaction on Signal Processing*, vol. 49, pp. 2262-2271, October 2001.
- [7] E. M. Fadaili, N. Thirion-Moreau and E. Moreau, "Non orthogonal joint diagonalization/zero-diagonalization for source separation based on time-frequency distributions", paratre dans *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. , April 2007.
- [8] A. Ferréol, P. Chevalier and L. Albera, "Second-order blind separation of first- and second-order cyclostationary sources - Application to AM, FSK, CPFSK and deterministic sources", *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 845-861, April 2004.
- [9] A. Ferréol, P. Chevalier and L. Albera, "On the behavior of current second and higher order blind source separation methods for cyclostationary sources - Application to AM, FSK, CPFSK and Deterministic Sources", *IEEE Transactions on Signal Processing*, vol. 48, no. 6, pp. 1712-1725, June 2000.
- [10] W. A. Gardner, "Cyclostationarity in communications and signal processing", *IEEE press*, W. A. Gardner Editor, 1993.
- [11] L. Giulieri, "Séparation aveugle de sources basée sur

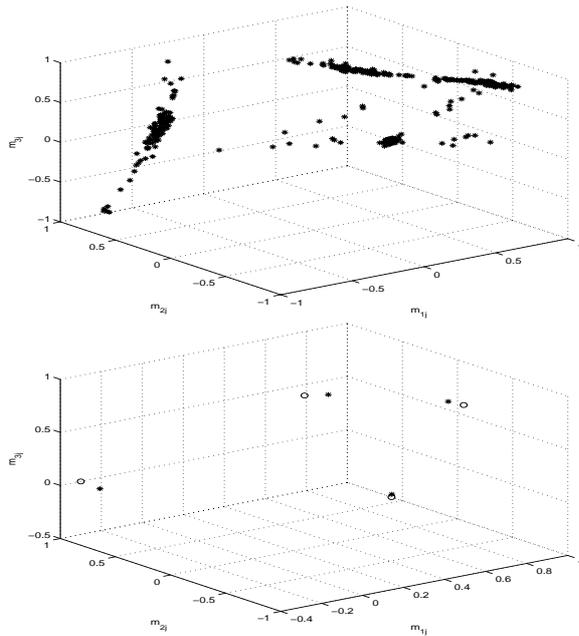


Figure 7. Top: the 4×100 estimated columns of the mixing matrix M_3 using the PARAFAC decomposition on the tensor built from the matrices belonging to the subset \mathcal{R}_c . Bottom: the 4 estimated columns (*) of the mixing matrix M_3 (after average over 100 Monte-Carlo runs) compared with their true value (\circ).

l'utilisation des transformées spatiales quadratiques", *PhD Thesis*, Université du Sud Toulon-Var, December 2003.

[12] R. A. Harshman, "Foundations of the PARAFAC procedure: model and conditions for an explanatory multi-model factor analysis", *UCLA Zorking Papers in Phonetics* col.16, no 1, 1970.

[13] S. Houcke, "Séparation autodidacte d'un mélange de sources émettant à des débits inconnus et éventuellement différents", *PhD Thesis*, Université de Marne la Vallée, 2002.

[14] P. A. Jallon, A. Chevreuil, P. Loubaton and P. Chevalier, "Separation of convolutive mixtures of cyclostationary sources: a contrast function based approach", *IEEE Transactions on Signal Processing*, vol. 45, pp. 434-444, February 1997.

[15] P. M. Koonenbergh, "Three-mode principal component analysis", *DSWO Press*, 1983.

[16] J. B. Kruskal, "Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics", *Linear Algebra Appl.*, vol. 18, pp. 95138, 1977.

[17] E. Moreau, "A generalization of joint-diagonalization criteria for source separation", *IEEE Transactions on Signal Processing*, vol. 49, no. 3, pp. 530-541, March 2001.

[18] S. Rhioui, N. Thirion-Moreau and E. Moreau, "Under-determined blind identification of cyclo-stationary signals with unknown cyclic frequencies", in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'2006)*, Toulouse, France, vol. IV, pp 553-556, May 2006.

[19] G. Scarano and G. Panci, "On cyclic approaches for blind identification of fir communications channels", in *Proc. SPAWC 1997*, pp. 81-84, April 1997.

[20] N. D. Sidiropoulos, G. B. Giannakis and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems", *IEEE Transaction on Signal Processing*, vol. 48, no. 3, pp. 810823, March 2000.

[21] P. H. A. Sneath and R. R. Sokal, "Numerical taxonomy:

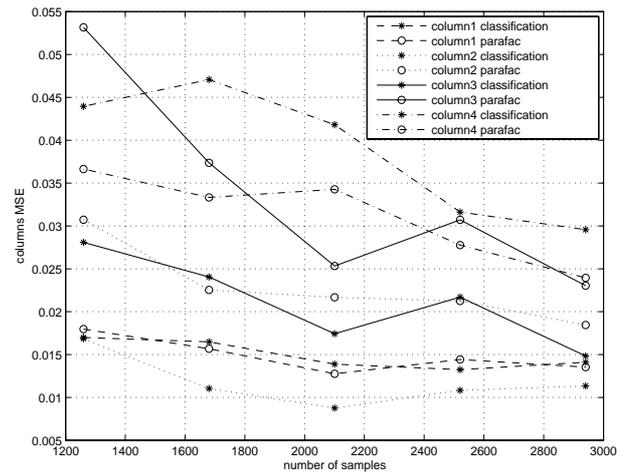


Figure 8. A comparison of the Mean Square Error (MSE) for each of the 4 columns of the estimated mixing matrix in the under-determined case (matrix M_3 is used) versus the size S of the signals. The PARAFAC method (\circ) is applied using the subset \mathcal{R}_c of the set \mathcal{R} . It is compared with the EVD + classification based method (*).

the principles and practice of numerical classification", Freeman, San Francisco, 1973.

[22] A. J. Van Der Veen, "Joint diagonalization via subspace fitting techniques", in *Proc. International Conference on Acoustic Speech and Signal Processing (ICASSP)*, Salt Lake City, USA, pp. 2773-2776, May 2001.

[23] A. Yeredor, "Non-orthogonal joint-diagonalization in the least squares sense with application in blind sources separation", *IEEE Trans. on Signal Processing*, pp. 1545-1553, vol. 50, no. 7, July 2002.

[24] A. Yeredor, A. Ziehe and K.-R. Müller, "Approximate joint diagonalization using natural gradient approach", *Lecture Notes in Computer Science (LNCS 3195): Independent Component Analysis and Blind Sources Separation*, in *Proceedings ICA*, pp. 89-96, Granada, Spain, September 2004.

[25] A. Ziehe, P. Laskov, G. Nolte and K.-R. Müller, "A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation", *Journal of Machine Learning Research*, pp. 801-818, vol. 5, July 2004.

Saloua Rhioui was born in Casablanca, Morocco, in 1978. She received the M.S. degree in telecommunications engineering from the Institut des Sciences de l'Ingénieur de Toulon et du Var, La Valette, France, and in signal, image and digital communication from the University of Nice, Sophia Antipolis, France. She received the Ph.D. degree in 2006, in the field of signal processing from the University of Sud Toulon-Var. Since 2003, she has been involved in blind source separation and independent component analysis. Her research interests include both convolutive mixtures and contrast functions.

Nadège Thirion-Moreau was born in Montbéliard, France. She graduated from the "Ecole Nationale Supérieure de Physique" (ENSPG) of the "Institut National Polytechnique de Grenoble" (INPG), Grenoble, France, in 1992. She received the DEA degree in 1992 and the Ph.D. degree in 1995, both in the field of signal processing and from the INPG. From 1996 to 1998, she has been with the "Ecole Supérieure des Procédés Electroniques et Optiques" (ESPEO), Orléans, France, as an Assistant Professor. Since 1998, she has been with the "Institut des Sciences de l'Ingénieur de Toulon et du

Var" (ISITV), La Valette, France, as an Assistant Professor, in the department of Telecommunications. Her main research interests are in deterministic and statistical signal processing including blind source separation, high-order statistics, non stationary signals, time-frequency representations, decision and classification.

Eric Moreau was born in Lille, France. He graduated from the "Ecole Nationale Supérieure des Arts et Métiers" (EN-SAM), Paris, France, in 1989 and received the "Agrégation de Physique" degree from the "Ecole Normale Supérieure de Cachan" in 1990. He received the DEA degree in 1991 and the Ph.D. degree in 1995, both in the field of signal processing and from the University of Paris-Sud, France.

From 1995 to 2001, he was assistant professor within the Telecommunications Department of the Engineering School "Institut des Sciences de l'Ingénieur de Toulon et du Var" (ISITV), La Valette, France. He is currently a Professor with the University of Toulon, France. His main research interests are in statistical signal processing using high-order statistics.