A general algebraic algorithm for blind extraction of one source in a MIMO convolutive mixture

Rémi Dubroca, Christophe de Luigi, Marc Castella, Eric Moreau

A General Algebraic Algorithm for Blind Extraction of one Source in a MIMO Convolutive Mixture

Rémi Dubroca, Christophe De Luigi, *Member IEEE*, Marc Castella, *Member IEEE* and Eric Moreau, *Senior Member IEEE*

***Abstract*—The paper deals with the problem of blind source extraction from a MIMO convolutive mixture. We define a new criterion for source extraction which uses higher-order contrast functions based on so called reference signals. It generalizes existing reference-based contrasts. In order to optimize the new criterion, we propose a general algebraic algorithm based on best rank-1 tensor approximation. Computer simulations illustrate the good behavior and the interest of our algorithm in comparison with other approaches.**

***Index Terms*—Contrast Functions, Blind Source Extraction, Higher Order Statistics, Tensor Decomposition, Independent Component Analysis.**

I. Introduction

THE problem of blind extraction of one source signal appears in a wide range of signal processing applications such as bio-medical and telecommunications. The aim is to restore one source signal from the observation of a set of mixed sources. Among existing mixing models, we consider a Multiple-Input / Multiple-Output (MIMO) context where non observable source signals are mixed through a multidimensional *convolutive* channel. The separation is said to be performed *blindly* when the mixing system is unknown and cannot be identified.

Such mixtures are considered in telecommunication applications like cellular communications. The mobile user signals are mixed through the air channel and observed on the different antennas of the base station. In this context, when the extraction of one mobile user signal has to be performed, the problem cannot be treated as Single-Input / Single-Output blind equalization because the other mobile user signals cannot be considered as noise: they are interfering signals. The mobile user of interest is extracted from the mixture of all users. The choice of a blind method for extraction of the source potentially avoids the necessity of a training sequence and therefore lowers bandwidth loss.

The blind source extraction problem may be considered as part of the general blind source separation (BSS) problem. Iterative methods in the context of BSS, extract the sources one by one and therefore rely on the ability to extract one source. These so-called deflation methods, presented in [16], developed in [13], [20], [9], [3], decompose the separation problem in successive stages for each source. These methods can deal both with temporally independent identically distributed (i.i.d.) sources and temporally non i.i.d. sources [19]. Hence the blind source extraction problem is a critical part in these methods, the separation relies on the ability to extract well each source.

The former method is in contrast to a global separation approach that has been studied in a convolutive context (see [4], [18], [14] for the i.i.d. case and [5] for the non i.i.d. case). The separation is not performed sequentially, but all the sources are extracted simultaneously. Such an approach faces two major obstacles in practice. First, a prewhitening is required, which is a difficult task. In addition, the optimization scheme of the criteria is likely to converge to a wrong solution because of the existence of local spurious maxima [2].

Contrast functions based on higher-order statistics have been proved to be efficient for both blind source separation or extraction [4], [17]. Here we focus on Multiple-Input / Single-Output (MISO) extraction criteria relying on higher-order statistics. One potential drawback is the high order dependence on the parameters. Recently "reference"-based contrasts have been investigated to lower this dependence or to improve the extraction performance [1], [3], [15]. These approaches impose some constraints on the reference signals.

In order to extract one source, we propose in this paper a family of contrast functions based on higher-order cumulant and using reference signals. These contrast functions impose no constraint on the reference signals. We show the usefulness of tensorial algorithms applied to independent component analysis (ICA) [6]. The optimization problem shows an equivalence to the problem of best rank-1 approximation of tensors [8], [7], [10]. This optimization scheme is also applied to the classical kurtosis contrast function [19]. The computation of these best rank-1 approximations is done through the use of the Higher-Order Power Method (HOPM) algorithm (higher-order generalization of the power method). The novel results in this paper are the following ones: we introduce a new family of reference based contrasts, we interprete tensor approximation in BSS and we propose new optimization method adapted to the proposed criteria.

The problem is formulated in Section II and Section III introduces the new contrast function. Some recalls on Tensor algebra and decomposition from [8] are done in Section IV and an optimization algorithm for our new criterion is presented in Section V. Section VI illustrates through computer simulations the good behavior of the proposed approach. Finally we conclude this paper in Section VII.

Notations. Tensors are denoted by bold-face calligraphic letters ($\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}}, \ldots$), matrices by bold-face capital letters ($\mathbf{A}, \mathbf{B}, \ldots$), vectors by bold-face lower-case letters ($\mathbf{a}, \mathbf{b}, \ldots$), and scalars by normal letters ($a, b, \ldots; \alpha, \beta, \ldots$).
In the whole paper, n stands for a generic integer ($n \in \mathbb{Z}$).

All quantities throughout the paper may be either real or complex-valued. $(.)^*$ stands for the complex conjugate, $(.)^T$ for matrix transpose, $(.)^H$ for matrix conjugate transpose. Finally $\mathrm{Cum}\{.\}$ stands for the the cumulant of any set of random variables, and $\delta_{i,j}$ stands for the Kronecker symbol, i.e. $\delta_{i,j}=1$ if $i=j$ and 0 otherwise.

II. Problem formulation

We consider a linear time-invariant (LTI) MIMO mixing model with K inputs and N outputs

$$\mathbf{x}(n)=\sum_{k\in\mathbb{Z}}\mathbf{M}(k)\mathbf{s}(n-k), \tag{1}$$

where $\mathbf{x}(n)$ is the $(N,1)$ observation vector ($N\geqslant 2$), $\mathbf{s}(n)$ is the $(K,1)$ source vector ($K\geqslant 2$), and $\mathbf{M}(n)$ is the (N,K) matrix corresponding to the impulse response of the convolutive mixing system. The transfer function of the LTI mixing system is denoted by

$$\mathbf{M}(z)=\sum_{n\in\mathbb{Z}}\mathbf{M}(n)z^{-n}, \tag{2}$$

and we assume $N\geqslant K$.

To achieve the extraction of one source blindly, that is with no access to the source signals nor to the mixing system filter, we have to make some assumptions on the LTI mixing system and the sources:

A1. The LTI mixing system is stable, left invertible. Therefore it exists a system $\mathbf{W}(n)$ such that the global LTI system with impulse response

$$\mathbf{G}(n)\triangleq\sum_{k\in\mathbb{Z}}\mathbf{W}(k)\mathbf{M}(n-k) \tag{3}$$

corresponds to the identity system.

A2. The source signals $s_i(n)$, $i\in\{1,\ldots,K\}$ are zero-mean, unit-variance, random stationary signals. Moreover they are statistically mutually independent (at least up to the order of the considered cumulants). At least one of the sources has non-zero 4-th order auto-cumulant, i.e. there exists $i\in\{1,\ldots,N\}$ such that

$$\mathrm{Cum}\{s_i(n),s_i(n),s_i(n),s_i(n)\}\neq 0.$$

In the MISO context, the aim is to estimate a $(1,N)$ row vector filter $\mathbf{w}(n)$ such that the scalar signal

$$y(n)=\sum_{k\in\mathbb{Z}}\mathbf{w}(k)\mathbf{x}(n-k) \tag{4}$$

restores one of the sources $s_i(n)$, $i\in\{1,\ldots,K\}$, up to a non-zero scalar filter. It is useful to define the global vector filter $\mathbf{g}(n)$ by

$$\mathbf{g}(n)\triangleq\sum_{k\in\mathbb{Z}}\mathbf{w}(k)\mathbf{M}(n-k), \tag{5}$$

thus

$$y(n)=\sum_{k\in\mathbb{Z}}\mathbf{g}(n-k)\mathbf{s}(k)\triangleq(\mathbf{g}\star\mathbf{s})(n). \tag{6}$$

As the source signals are unobservable, there exists some inherent undetermined factors in their estimation. They can be recovered only up to a permutation and to a scalar filtering. So, the extraction of one source is said to be achieved when there exists an index $i_0\in\{1,\ldots,K\}$ such that the filter components in $\mathbf{g}(n)$ read

$$\forall i\in\{1,\ldots,K\}\ g_i(n)\triangleq(\mathbf{g}(n))_i=g(n)\delta_{i,i_0}. \tag{7}$$

The above relation is called the "extraction condition" and expresses the fact that $y(n)$ is equal to the source signal, $s_{i_0}(n)$ up to a filtering by the scalar filter with impulse response $g(n)$. The source signal correlation sequences are denoted by

$$\gamma_i(k)\triangleq\mathrm{E}\{s_i(n)s_i^*(n-k)\},$$

where $k\in\mathbb{Z}$ and $i\in\{1,\ldots,K\}$. For any index $j\in\{1,\ldots,K\}$ and any scalar filter with impulse response $h(n)$, we define its j-norm by:

$$\|h\|_j^2\triangleq\sum_{(k_1,k_2)\in\mathbb{Z}^2}h(k_1)h^*(k_2)\gamma_j(k_2-k_1). \tag{8}$$

$\|h\|_j^2$ is the variance of the signal obtained when filtering the source $s_j(n)$ by the filter with impulse response $h(n)$. We define the (weighted) l^2-norm of the $(1,K)$ global filter $\mathbf{g}(n)$ by:

$$\|\mathbf{g}\|^2\triangleq\sum_{i=1}^{K}\|g_i\|_i^2. \tag{9}$$

Since the sources are unit power, one can restrict to the unit-norm global filter case in (7) by imposing the constraint $\mathrm{E}\{|y(n)|^2\}=1$. It is straithforward to show that this constraint is equivalent to $\|\mathbf{g}\|^2=1$. In the particular case of i.i.d. sources, this constraint is equivalent to:

$$\sum_{i=1}^{K}\sum_{k\in\mathbb{Z}}|g_i(k)|^2=1. \tag{10}$$

Moreover, in the i.i.d. case, the contrast functions (defined in the next section) allow one to equalize the sources. Consequently we define the stronger separation condition:

$$\exists\ l\in\mathbb{Z}\text{ and }\exists\ i_0\in\{1,\ldots,K\}\text{ such that }g_i(n)=\alpha\delta_{n,l}\delta_{i,i_0}. \tag{11}$$

Here the scalar filter $g(n)$ reduces to a delay l and a scaling factor $\alpha\in\mathbb{C}^*$ and (11) is actually a condition for separation and equalization.

The set of source signals satisfying assumption A2. will be denoted by $\mathcal{S}$. The set of unit norm vector filters will be denoted by $\mathcal{G}_1$. Finally, we denote by $\mathcal{Y}$ the set of output of the MISO extractor $y(n)$ when the input source signals belong to $\mathcal{S}$ and the global system belongs to $\mathcal{G}_1$.

III. Extraction criteria

A. Recalls on contrast function

A contrast is an objective function depending on the statistics of the outputs of the extracting system. Its maximization leads to the extraction of one source. In consequence, the source extraction problem becomes an optimization problem.

We can notice that a contrast is a function from $\mathcal{Y}$ to $\mathbb{R}$ but it may also be seen as a function from $\mathcal{G}_1$ to $\mathbb{R}$. In the i.i.d. case, the following definition is considered:

Definition III.1. *Contrast function for i.i.d. sources*
Let C be a function from $\mathcal{Y}$ to $\mathbb{R}$. It is called a contrast function when it satisfies the following two properties:

p1. $\exists\ (i_0, k_0) \in \{1, \ldots, K\} \times \mathbb{Z}$ *such that:*

$$\forall\ y(n) \in \mathcal{Y} \quad \mathsf{C}\{y(n)\} \leqslant \mathsf{C}\{s_{i_0}(n - k_0)\}\ . \tag{12}$$

p2. The equality holds in (12) if and only if $y(n)$ equals a multiple of $s_i(n-k)$ for some $i \in \{1, \ldots, K\}$ and some $k \in \mathbb{Z}$ (i.e. the global filter is extracting and statisfies (11)).

In the case of i.i.d. source signals, the maximization of the contrast function restores one source up to a delay and a scaling factor.

However this definition cannot be used for the non i.i.d. case because the extraction of a source can be only guaranteed up to a scalar filtering. Hence we give a weaker definition for the non i.i.d. case:

Definition III.2. *Contrast function for non i.i.d. sources*
Let C be a function from $\mathcal{Y}$ to $\mathbb{R}$. It is called a contrast function when it satisfies the following two properties:

p1. $\forall\ y(n) \in \mathcal{Y}$*:*

$$\mathsf{C}\{y(n)\} \leqslant \max_{i=1}^{K} \sup_{\substack{\mathbf{g}:||g_i||_i=1 \\ ||g_j||_j=0\ \forall j \neq i}} \mathsf{C}\{(\mathbf{g} \star \mathbf{s})(n)\}. \tag{13}$$

p2. The equality holds in (13) if and only if $y(n)$ is a filtered version of $s_i(n)$ for some $i \in \{1, \ldots, K\}$ (i.e. the global filter statisfies (7)) for some i_0).

A classical contrast function defined in [19] is the absolute value of the kurtosis:

$$\mathcal{C}_4\{y(n)\} \triangleq |\mathrm{Cum}\{y(n), y^*(n), y^*(n), y(n)\}|. \tag{14}$$

A contrast function based on referenced signal is defined in [3] by the absolute value of the fourth order cross-cumulant between the output of the extracting system and a reference signal $r(n)$:

$$\mathcal{C}_{2,4,r}\{y(n)\} \triangleq |\mathrm{Cum}\{y(n), y^*(n), r(n), r^*(n)\}|. \tag{15}$$

B. New reference based contrast function

The main goal of the paper is to propose a family of contrast functions based on reference signals generalizing the contrast in [3]. For that, we consider the following R-th order ($R \geqslant 3$) cross-cumulant where $(.)^{(*)_k}$, $k \in \{1, \ldots, S\}$ are optional complex conjugate:

$$\kappa_{S,R,\mathbf{r}}\{y(n)\} \triangleq \mathrm{Cum}\{\underbrace{y^{(*)_1}(n), \ldots, y^{(*)_S}(n)}_{S}, r_1(n), \ldots, r_{R-S}(n)\}\ . \tag{16}$$

The $r_i(n), i \in \{1, \ldots, R-S\}$, are given reference signals ($2 < S \leq R$) which are assumed to satisfied the following assumption

A3. The reference signals $r_i(n), i \in \{1, \ldots, R-S\}$ are obtained by a stable MIMO filtering of the sources or of their complex-conjugates.

Note that the case $S = 2$ is treated in [3] and the case $S = R$ is treated in [17]. In the following we consider the general case $2 < S \leq R$.

We now define the following function:

$$\mathcal{C}_{S,R,\mathbf{r}}\{y(n)\} \triangleq |\kappa_{S,R,\mathbf{r}}\{y(n)\}|. \tag{17}$$

1) Case of i.i.d. source signals:
The following supremum is assumed to satisfy the assumption

A4. $\exists (j_0, k_0)$ such that:

$$\max_{j=1}^{N} \sup_{k \in \mathbb{Z}} |\kappa_{S,R,\mathbf{r}}\{s_j(n-k)\}| = |\kappa_{S,R,\mathbf{r}}\{s_{j_0}(n-k_0)\}| < \infty. \tag{18}$$

We can now state the following proposition (for $2 < S \leqslant R$):

Proposition III.3. *In the case of i.i.d. source signals, under assumptions A1.-A4., the function $\mathcal{C}_{S,R,\mathbf{r}}$ is a contrast over the set $\mathcal{Y}$.*

The proof is given in Appendix A.

2) Case of non i.i.d. source signals:
We define the following supremum,

$$\mathcal{M}_{S,R}^{\max} \triangleq \max_{i=1}^{K} \mathcal{M}_{S,R,i} \tag{19}$$

where

$$\mathcal{M}_{S,R,i} \triangleq \sup_{\mathbf{g}:||g||_i=1} \mathsf{C}_{S,R,\mathbf{r}}\{(\mathbf{g} \star \mathbf{s})(n)\}. \tag{20}$$

We make the following assumption:

A5. For all i such that $\mathcal{M}_{S,R,i} = \mathcal{M}_{S,R}^{\max}$, the extremum $\mathcal{M}_{S,R,i}$ is reached by at least one filter of unit i-norm defined in (8).

We can state (with $2 < S \leqslant R$):

Proposition III.4. *In the case of non i.i.d. source signals, under assumptions A1.-A3. and A5., the function $\mathcal{C}_{S,R,\mathbf{r}}$ is a contrast over the set $\mathcal{Y}$.*

The proof is given in Appendix B.

We can notice that the observation signals satisfy the assumption A3. and can serve as reference signals. Note also that $r_1(n) = \ldots = r_{R-S}(n)$ is a possible choice. As we have seen in the previous propositions, we remark that the new contrast requires no condition on the reference signals except the assumption A3..

IV. Recalls on Tensor Decomposition

Tensorial algorithms applied to ICA [6] are of interest and the optimization problem has an equivalence to the problem of best rank-1 approximation of tensors [8], [7]. We recall notions on tensor decomposition which are necessary in Section V. See [8] for more details in tensor algebra.

The n-mode product of a tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times \ldots \times I_N}$ by a matrix $\mathbf{U} \in \mathbb{C}^{J_n \times I_n}$, denoted by $\boldsymbol{\mathcal{A}} \bullet_n \mathbf{U}$, is an $(I_1 \times \ldots \times J_n \times \ldots I_N)$-tensor whose entries are given by

$$(\boldsymbol{\mathcal{A}} \bullet_n \mathbf{U})_{i_1 i_2 \ldots j_n \ldots i_N} \triangleq \sum_{i_n} \mathcal{A}_{i_1 i_2 \ldots i_n \ldots i_N} \mathrm{U}_{j_n i_n}. \quad (21)$$

The outer product of N vectors $\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(N)}$ of size $I_1, \ldots, I_N$ respectively is a N-order tensor in $\mathbb{C}^{I_1 \times \ldots \times I_N}$ with elements defined by $\mathcal{A}_{i_1 \ldots i_N} = \mathrm{u}_{i_1}^{(1)} \ldots \mathrm{u}_{i_N}^{(N)}$ for all values of the indices. It is denoted by $\mathbf{u}^{(1)} \circ \ldots \circ \mathbf{u}^{(N)}$.

A N-order tensor in $\mathbb{C}^{I_1 \times \ldots \times I_N}$ is rank-1 if it can be written as the outer product of N vectors $\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(N)}$ of size $I_1, \ldots, I_N$ respectively.

The best rank-1 approximation of a N-order tensor can be described as follows. Given $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times \ldots \times I_N}$, find a scalar λ and unit-norm vectors $\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(N)}$ such that the rank-1 tensor $\tilde{\boldsymbol{\mathcal{A}}} \in \mathbb{C}^{I_1 \times \ldots \times I_N}$ defined by

$$\tilde{\boldsymbol{\mathcal{A}}} \triangleq \lambda \ \mathbf{u}^{(1)} \circ \ldots \circ \mathbf{u}^{(N)} \quad (22)$$

minimizes the least-squares cost function

$$f(\tilde{\boldsymbol{\mathcal{A}}}) = \|\boldsymbol{\mathcal{A}} - \tilde{\boldsymbol{\mathcal{A}}}\|^2. \quad (23)$$

The Frobenius norm of $\boldsymbol{\mathcal{A}}$ is defined by $\|\boldsymbol{\mathcal{A}}\| \triangleq \sqrt{\langle \boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{A}} \rangle}$ with the scalar product $\langle \boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}} \rangle$ of two tensors $\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}} \in \mathbb{C}^{I_1 \times \ldots \times I_N}$ defined by

$$\langle \boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}} \rangle \triangleq \sum_{i_1} \ldots \sum_{i_N} \mathcal{A}_{i_1 \ldots i_N} \mathcal{B}^*_{i_1 \ldots i_N}. \quad (24)$$

The minimization of the cost function $f(\tilde{\boldsymbol{\mathcal{A}}})$ is equivalent to the maximization, over the unit-norm vectors $\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(N)}$, of the function:

$$g(\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(N)}) = |\boldsymbol{\mathcal{A}} \bullet_1 \mathbf{u}^{(1)^H} \bullet_2 \ldots \bullet_N \mathbf{u}^{(N)^H}|^2. \quad (25)$$

The scalar λ is reached by:

$$\lambda = \boldsymbol{\mathcal{A}} \bullet_1 \mathbf{u}^{(1)^H} \bullet_2 \ldots \bullet_N \mathbf{u}^{(N)^H}. \quad (26)$$

A solution can be found by an ALS method which is a higher-order extension of the power method (HOPM) for matrices [8].

HOPM algorithm
Input: $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times \ldots \times I_N}$
Output: $\tilde{\boldsymbol{\mathcal{A}}} \in \mathbb{C}^{I_1 \times \ldots \times I_N}$

1) Initial values: $\mathbf{u}_0^{(n)}$, $n \in [1, \ldots, N]$.
2) Repeat until convergence:
   - $\tilde{\mathbf{u}}_{k+1}^{(1)} = \boldsymbol{\mathcal{A}} \bullet_2 \mathbf{u}_k^{(2)^H} \bullet_3 \mathbf{u}_k^{(3)^H} \bullet_4 \ldots \bullet_N \mathbf{u}_k^{(N)^H}$
     $\lambda_{k+1}^{(1)} = \|\tilde{\mathbf{u}}_{k+1}^{(1)}\|$ and $\mathbf{u}_{k+1}^{(1)} = \tilde{\mathbf{u}}_{k+1}^{(1)} / \lambda_{k+1}^{(1)}$
   - $\tilde{\mathbf{u}}_{k+1}^{(2)} = \boldsymbol{\mathcal{A}} \bullet_1 \mathbf{u}_{k+1}^{(1)^H} \bullet_3 \mathbf{u}_k^{(3)^H} \bullet_4 \ldots \bullet_N \mathbf{u}_k^{(N)^H}$
     $\lambda_{k+1}^{(2)} = \|\tilde{\mathbf{u}}_{k+1}^{(2)}\|$ and $\mathbf{u}_{k+1}^{(2)} = \tilde{\mathbf{u}}_{k+1}^{(2)} / \lambda_{k+1}^{(2)}$
   - $\ldots$
   - $\tilde{\mathbf{u}}_{k+1}^{(N)} = \boldsymbol{\mathcal{A}} \bullet_1 \mathbf{u}_{k+1}^{(1)^H} \bullet_2 \mathbf{u}_{k+1}^{(2)^H} \bullet_3 \ldots \bullet_{N-1} \mathbf{u}_{k+1}^{(N-1)^H}$
     $\lambda_{k+1}^{(N)} = \|\tilde{\mathbf{u}}_{k+1}^{(N)}\|$ and $\mathbf{u}_{k+1}^{(N)} = \tilde{\mathbf{u}}_{k+1}^{(N)} / \lambda_{k+1}^{(N)}$

   Solutions : $\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(N)}, \lambda$.
3) $\tilde{\boldsymbol{\mathcal{A}}} = \lambda \ \mathbf{u}^{(1)} \circ \ldots \circ \mathbf{u}^{(N)}$.

This iterative algorithm has to be initialized, which can be done using the Higher Order Singular Value Decomposition (HOSVD), as proposed in [8]. The stopping criterion of the algorithm is also studied in this paper.

V. Proposed algorithm

We describe a general algorithm for the optimization of the proposed family of contrast functions.

A. Optimization method

We assume that the mixing filter is a FIR filter with impulse response of length L and that $\mathbf{M}(z)$ is irreducible. Then it admits a MIMO-FIR left inverse filter of length D (see [12]), which is assumed causal with no loss of generality. The row vectors which define the impulse response can be stacked in the following $(1, ND)$ row vector:

$$\underline{\mathbf{w}} \triangleq (\mathbf{w}(0) \ldots \mathbf{w}(D-1)). \quad (27)$$

We also define the $(ND, 1)$ column vector

$$\underline{\mathbf{x}}(n) \triangleq (\mathbf{x}(n)^T \mathbf{x}(n-1)^T \ldots \mathbf{x}(n-D+1)^T)^T \ . \quad (28)$$

It is then easily seen that

$$y(n) = \underline{\mathbf{w}}\, \underline{\mathbf{x}}(n). \quad (29)$$

Let the covariance matrix be $\mathbf{R} = \mathrm{E}\{\underline{\mathbf{x}}(n) \underline{\mathbf{x}}(n)^H\}$, so we have $\mathrm{E}\{|y(n)|^2\} = \underline{\mathbf{w}} \mathbf{R} \underline{\mathbf{w}}^H$.

Now using the assumption A3., the multilinearity property of cumulants and (29), we have

$$\kappa_{S,R,\mathbf{r}}\{y(n)\} = \sum_{i_1, \ldots, i_S} \underline{w}_{i_1}^{(*)_1} \ldots \underline{w}_{i_S}^{(*)_S} \mathrm{Cum}\{\underline{x}_{i_1}^{(*)_1}(n), \ldots, \underline{x}_{i_S}^{(*)_S}(n), r_1(n), \ldots, r_{R-S}(n)\}, \quad (30)$$

where $\underline{w}_{i_k}^{(*)_k}$, $k \in \{1, \ldots, S\}$, stands for $\underline{w}_{i_k}$ conjugated in the same way as $y^{(*)_k}$ in (16). Thus this relation can be written as a S-order tensor decomposition ([8])

$$\kappa_{S,R,\mathbf{r}}\{y(n)\} = \boldsymbol{\mathcal{C}}_S \bullet_1 \underline{\mathbf{w}}^{(*)_1} \bullet_2 \ldots \bullet_S \underline{\mathbf{w}}^{(*)_S} \quad (31)$$

where the tensor $\boldsymbol{\mathcal{C}}_S$ is defined component wise as

$$(\boldsymbol{\mathcal{C}}_S)_{i_1,\ldots,i_S} = \\ \mathrm{Cum}\{\underline{x}_{i_1}^{(*)_1}(n),\ldots,\underline{x}_{i_S}^{(*)_S}(n), r_1(n),\ldots,r_{R-S}(n)\}\ . \quad (32)$$

Hence the optimization of the contrast function in (17) under the unit power constraint reads

$$\max |\boldsymbol{\mathcal{C}}_S \bullet_1 \underline{\mathbf{w}}^{(*)_1} \bullet_2 \ldots \bullet_S \underline{\mathbf{w}}^{(*)_S}| \quad \text{with} \quad \underline{\mathbf{w}}\mathbf{R}\underline{\mathbf{w}}^H = 1\ . \quad (33)$$

The covariance matrix can have some of its eigenvalues equal to zero. It is important to notice that for any row vector such that $\underline{\mathbf{w}}_0^H \in \ker \mathbf{R}$ we have

$$\underline{\mathbf{w}}_0 \mathbf{R} \underline{\mathbf{w}}_0^H = 0.$$

The solution of the above maximization problem (33) is not unique, it follows that we may impose in addition $\underline{\mathbf{w}}^H \in (\ker \mathbf{R})^{\perp}$. Thus, $\underline{\mathbf{w}}$ is projected onto the signal subspace. So, by a SVD, we decompose

$$\mathbf{R} = \mathbf{U}\mathbf{D}\mathbf{U}^H, \quad (34)$$

and we define

$$\mathbf{P} = \mathbf{U}\mathbf{D}^{1/2} \quad \text{and} \quad \mathbf{Q} = \mathbf{D}^{-1/2}\mathbf{U}^H \quad (35)$$

in order to project onto the signal subspace:

$$\underline{\tilde{\mathbf{w}}} = \underline{\mathbf{w}}\mathbf{P} \quad \text{and} \quad \tilde{\boldsymbol{\mathcal{C}}}_S = \boldsymbol{\mathcal{C}}_S \bullet_1 \mathbf{Q}^{(*)_1} \bullet_2 \ldots \bullet_S \mathbf{Q}^{(*)_S}. \quad (36)$$

We obtain $\tilde{\boldsymbol{\mathcal{C}}}_S$ and $\underline{\tilde{\mathbf{w}}}$ and finally the problem in (33) reduces to the following one:

$$\max |\tilde{\boldsymbol{\mathcal{C}}}_S \bullet_1 \underline{\tilde{\mathbf{w}}}^{(*)_1} \bullet_2 \ldots \bullet_S \underline{\tilde{\mathbf{w}}}^{(*)_S}| \quad \text{with} \quad \underline{\tilde{\mathbf{w}}}\,\underline{\tilde{\mathbf{w}}}^H = 1. \quad (37)$$

This projection onto the signal subspace is equivalent to a prewhitening step on the observations.

In Section IV, we have seen that the above maximization is equivalent to the minimization of the cost function for the best rank-1 approximation of tensor $\tilde{\boldsymbol{\mathcal{C}}}_S$ [8]. So we compute here the best rank-1 approximation of tensor $\tilde{\boldsymbol{\mathcal{C}}}_S$, using the HOPM algorithm (it can be seen as a higher order extension of the power method for best rank-1 approximation of matrix), to maximize the proposed contrast.

B. A "fixed-point" like method

It has been proposed in a previous work [3] an iterative method to improve the performances of the extraction algorithm for the case $S = 2$. This "fixed-point" like method can be also generalized to the considered family of contrasts. Here, we recall briefly the method.

We assume, in this paragraph, that all references are the same. That is for all $i \in \{1,\ldots,R-S\}$, we have $r_i(n) = r(n)$ (or possibly $r_i(n) = r^*(n)$ in the complex case). The output obtained after maximization of $\mathcal{C}_{S,R,\mathbf{r}}$ should be closer to the extracted source than the reference signal. Basically, the output which has been previously obtained by the maximization of $\mathcal{C}_{S,R,\mathbf{r}}$ serves as a new reference signal. The fixed point principle is illustrated for three iterations in the figure 1 and explained below:

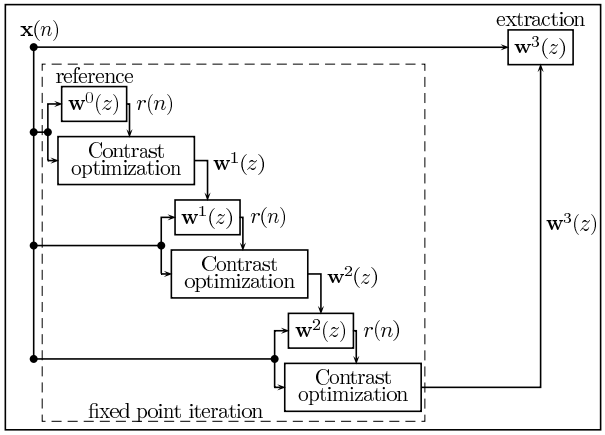

Fig. 1. Fixed point like method for 3 iterations.

"fixed-point" like algorithm (for $l_{\max}$ iterations)
Input: $\underline{\mathbf{x}}$,$\underline{\mathbf{w}}^0$.
Output: $\underline{\mathbf{w}}^{l_{\max}}$.

1) for $l \in \{1,\ldots,l_{\max}\}$:
   - fix the reference to $r(n) = \underline{\mathbf{w}}^{l-1}\underline{\mathbf{x}}(n)$,
   - $\underline{\mathbf{w}}^l = \arg\max_{\mathbf{w}} \mathcal{C}_{S,R,\mathbf{r}}\{y(n)\}$ where $y(n) = \underline{\mathbf{w}}\ \underline{\mathbf{x}}(n)$.
2) The separating filter is given by the coefficients $\underline{\mathbf{w}}^{l_{\max}}$.

VI. Simulations

We now propose computer simulations in order to illustrate the usefulness of our criterion w.r.t. other approaches.

We focus on the fourth order cumulant based contrasts in order to compare the generalized approach to the contrast functions defined in (14) and (15). In this context, we define the following contrast function for $R = 4$ and $S = 3$:

$$\mathcal{C}_{3,4,\mathbf{r}}\{y(n)\} \triangleq |\mathrm{Cum}\{y(n), y^*(n), y^*(n), r(n)\}|. \quad (38)$$

We compare three different algorithms:

1) The algorithm E_2 maximizes the quadratic criterion defined in (15). The solution is given by the eigenvector of $\tilde{\boldsymbol{\mathcal{C}}}_2$ associated to the eigenvalue with largest modulus. In practice it is obtained by the SVD of the considered matrix (second order tensor).
2) The algorithm E_3 maximizes the cubic criterion defined in (38). The solution is given by the best rank-one approximation of the tensor $\tilde{\boldsymbol{\mathcal{C}}}_3$. In practice it is obtained by a HOPM applied to the considered tensor (third order tensor).
3) The algorithm E_4 maximizes the classical quartic criterion with no reference defined in (14). The solution is given by the best rank-one approximation of the tensor $\tilde{\boldsymbol{\mathcal{C}}}_4$. In practice it is obtained by a HOPM applied to the considered tensor (fourth order tensor).

The quality of extraction is measured thanks to an error index suggested in [11] and defined by:

$$\mathrm{ind}(\mathbf{g}) \triangleq 1 - \frac{\max\limits_{i\in\{1,\ldots,N\}} \sum\limits_{k\in\mathbb{Z}} |g_i(k)|^2}{\sum\limits_{i\in\{1,\ldots,N\}} \sum\limits_{k\in\mathbb{Z}} |g_i(k)|^2}. \quad (39)$$

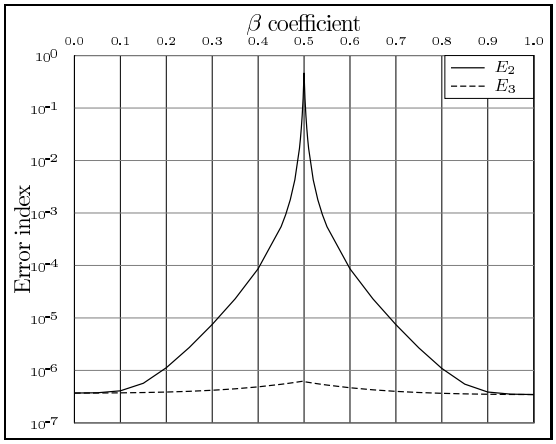

Fig. 2. Performance versus coefficient β for two sources (where the reference signal is $r(n) = \beta s_1(n) + (1-\beta)s_2(n)$).

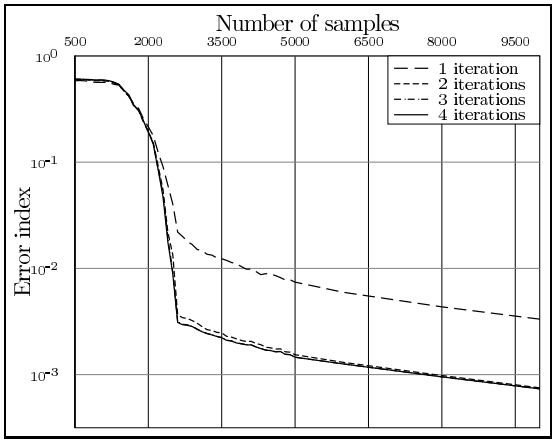

Fig. 3. Fixed point method for E_3 algorithm, complex i.i.d. case: Performance versus number of samples without noise.

This error index is non negative and takes its values between 0 and 1. The zero value represents a perfect extraction. In all simulations the mean of the error index over 200 Monte-Carlo realizations is considered. At each run, the mixing system and the sources have been drawn randomly. The reference signal is chosen as the first observation from the mixing system (except in section VI-A).
In the following simulations, we have choosen to fix in the "fixed-point" method the number of iterations to two for i.i.d. sources and five for non i.i.d. sources to compare the referenced contrasts to the classical one (except in section VI-A).

A. Experiment 1 - Influence of the reference signal

We consider the case of $K = 2$ i.i.d. real-valued binary source signals (they take their values in $\{-1, 1\}$ with equal probabilities), $N = 5$ observation signals and a real-valued mixing filter of length $L = 3$. Its coefficients are drawn randomly according to a zero-mean and unit variance normal distribution. To evaluate the influence of the reference signal, we choose it as $r(n) = \beta s_1(n) + (1-\beta)s_2(n)$ with $\beta \in [0, 1]$. In Figure 2, for $N_e = 5000$, the mean error index is plotted versus β for the E_2 and the E_3 algorithms (there is no reference in the E_4 algorithm). Better results are obtained for values of β near 0 or 1 that is to say when the reference signal is close to one source. The performance of the E_2 algorithm fails for $\beta = 0.5$ whereas the E_3 algorithm still works. This shows that our approach imposes nearly no constraint on the reference signal. E_3 is less sensitive to the chosen reference signal as E_2.

B. Experiment 2 - fixed point method

The mixing system is complex-valued and its coefficients are drawn randomly according to a zero-mean and unit variance normal distribution. We consider here a mixture of $K = 3$ i.i.d. source signals. The length of the mixing filter is $L = 3$ and the number of observations is $N = 6$.
In figure 3 the error index is drawn versus the number of sources samples N_e of i.i.d. complex-valued Quadrature Amplitude Modulation sources with 16 states (16-QAM). In this figure, we give the performance of the E_3 algorithm for different number of iterations in the "fixed-point" method. Performing more than two "fixed-point" iterations does not improve the performance sharply. Therefore, in the following experiments for i.i.d. sources, we will fix, for the E_2 and E_3 algorithms, the number of iterations of the "fixed-point" method to two.
Now we consider non i.i.d. complex-valued Continuous Phase Modulated sources. Pseudo-symbols of CPM signals are unit modulus signals defined by the relation

$$s_k(n+1) = \exp(\imath \pi h_k a_n^k)$$

where $h_k \in]0, 1[$ is a modulation index and a_n^k is the i.i.d. binary symbol sequence transmitted by th k-th user. We fix the modulation index h to 0.2. In figure 4 the error index is drawn versus the number of sources samples N_e. In this figure, we give the performance of the E_3 algorithm for different number of iterations in the "fixed-point" method. After five iterations, the performance does not improve sharply, then in the following experiments for non i.i.d. sources, we will fix, for the E_2 and E_3 algorithms, the number of iterations of the "fixed-point" method to five.

C. Experiment 3 - i.i.d. real-valued source signals

We consider real-valued binary sources. The mixing system is also real-valued and its coefficients are drawn randomly according to a zero-mean and unit variance normal distribution. We consider here a mixture of $K = 3$ i.i.d. source signals. The length of the mixing filter is $L = 3$ and the number of observations is $N = 6$. In figure 5 the error index is drawn versus the number of sources samples N_e. In this figure, we compare the performance of the E_2, E_3 and E_4 algorithms. Our approach performs better than E_4 for small amount of samples and similarly for large amount of samples. Here, even if E_4 is much more complex than E_3 (fourth order tensor versus third order tensor), the latter performs slighty better. The extraction performance of the E_2 algorithm is slightly

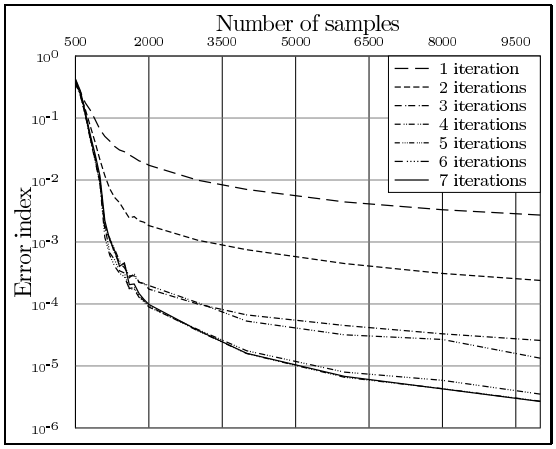

Fig. 4. Fixed point method for E_3 algorithm, complex non i.i.d. case: Performance versus number of samples without noise.

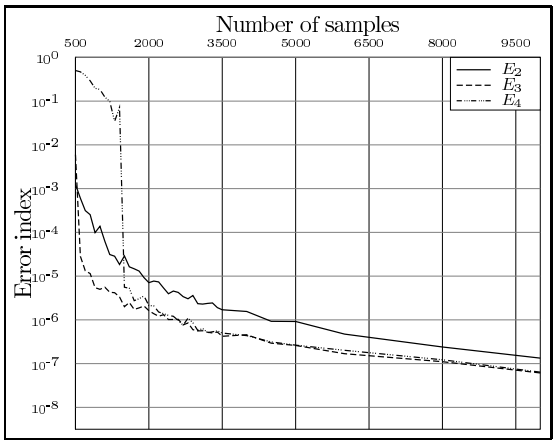

Fig. 5. Real-valued i.i.d. case: Performance versus number of samples without noise.

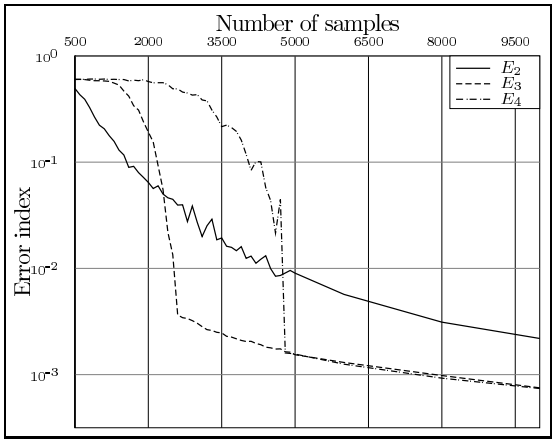

Fig. 6. Complex-valued i.i.d. case: Performance versus number of samples without noise.

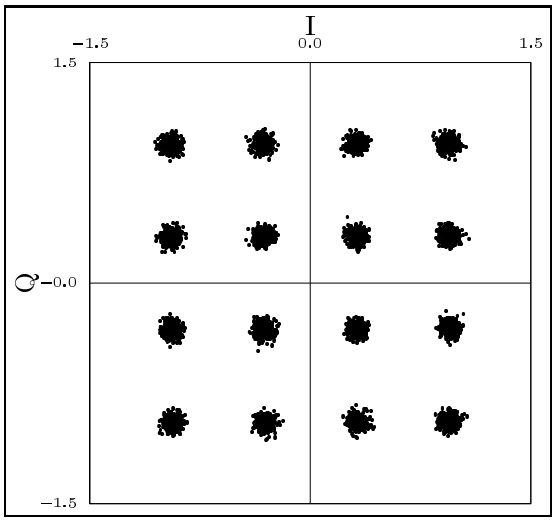

Fig. 7. Complex-valued i.i.d. case: Extracted QAM-16 constellation, error index is equal to 0.002.

lower in this configuration but it is important to note that its performances can be improved by increasing the number of iteration in the fixed point method.

D. Experiment 4 - i.i.d. complex-valued source signals

We consider complex-valued 16-QAM sources. The mixing system is complex-valued and its coefficients are drawn randomly according to a zero-mean and unit variance normal distribution. We consider here a mixture of $K = 3$ i.i.d. source signals. The length of the mixing filter is $L = 3$ and the number of observations is $N = 6$. In figure 6 the error index is drawn versus the number of sources samples N_e. In this figure, we compare the performance of the E_2, E_3 and E_4 algorithms. Our approach performs again better than E_4 for small amount of samples. E_2 is performing better for number of samples below 2000. With larger amount of samples, the E_4 and the E_3 algorithms are performing very similarly. The extraction performance of the E_2 algorithm is again slightly lower in this configuration.

In order to compare different optimization schemes, we put in the table I the performances of the cubic criterion and the contrast defined in (14) optimized using a gradient ascent method with an adaptatively adjusted step-size (initially set to one and divided by two each time the contrast does not increase), called Kurtosis. The Kurtosis and E_3 algorithms perform very similarly. The results in table II show up that 2 fixed point optimizations of the cubic criterion can be performed about 6 times faster than a single optimization of the classical kurtosis contrast (the execution times have been estimated on a dual processor Intel Xeon Dual Core 5150 running at 2.667GHz clock frequency and with 16GB RAM). Even if we have higher value of the mean error index than in the previous experiment for the three algorithms, the extraction is well performed. We illustrate this fact in the figure 7, with the constellation of a extracted 16-QAM source of 5000 samples. The algorithm used for this extraction is E_3 with $N_e = 5000$ and the error index is equal to 0.002.

E. Experiment 5 - non i.i.d. complex-valued source signals

Now we consider non i.i.d. complex-valued CPM sources. We still consider a mixture of $K = 3$ source signals, a mixing filter length of $L = 3$ and $N = 6$ observations. We fix the

Number of samples	5000	6000	7000	8000	9000	10000
Kurtosis	1.59×10^{-3}	1.29×10^{-3}	1.12×10^{-3}	9.69×10^{-4}	8.88×10^{-4}	7.59×10^{-4}
Cubic Criterion	1.69×10^{-3}	1.33×10^{-3}	1.18×10^{-3}	1.00×10^{-3}	9.18×10^{-4}	7.76×10^{-4}

TABLE I
COMPARISON OF THE AVERAGE ERROR INDEX FOR 1) ONE OPTIMIZATION OF THE KURTOSIS CONTRAST 2) 2 FIXED POINT ITERATIONS OPTIMIZATION OF CUBIC CRITERION

Number of samples	5000	6000	7000	8000	9000	10000
Kurtosis	12.38	17.38	20.55	24.41	28.53	32.71
Cubic Criterion	2.18	2.55	2.99	3.52	4.13	4.80

TABLE II
COMPARISON OF THE AVERAGE EXECUTION TIME (IN S) FOR 1) ONE OPTIMIZATION OF THE KURTOSIS CONTRAST 2) 2 FIXED POINT ITERATIONS OPTIMIZATION OF CUBIC CRITERION

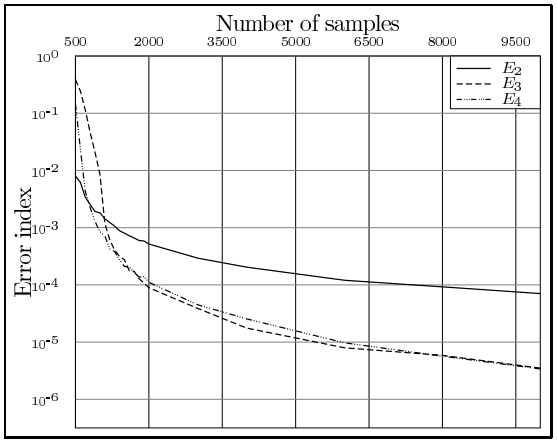

Fig. 8. Complex-valued non i.i.d. case: Performance versus number of samples without noise.

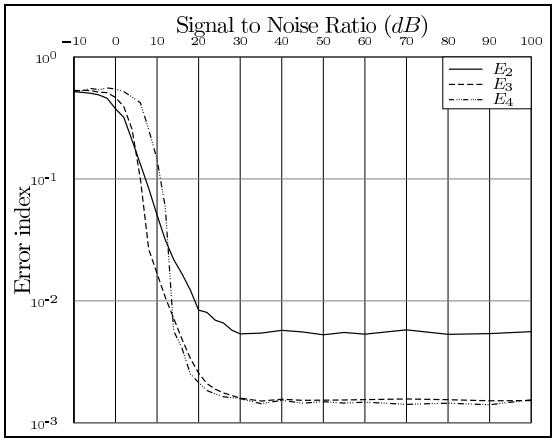

Fig. 9. Complex-valued i.i.d. case: Performance versus SNR for $N_e = 5000$.

modulation index h to 0.2. In figure 8 the error index is drawn versus the number of sources samples N_e and with no noise. In figure 8, we compare the performance of the E_2, E_3 and E_4 algorithms. For non i.i.d. complex-valued sources, E_4 and E_3 perform equally well and better than E_2.

F. Experiment 6 - additive noise signal

We consider here an additive noise signal in the model (1) that is:

$$\mathbf{x}(n) = \sum_{k \in \mathbb{Z}} \mathbf{M}(k)\mathbf{s}(n-k) + \mathbf{n}(n), \tag{40}$$

where $\mathbf{n}(n)$ is the $(N, 1)$ noise vector. The noise signals $n_i(n)$, $i \in \{1, \ldots, N\}$ are additive, complex-valued and follow the same normal law with zero-mean and N_0 variance.
We define the SNR in the worst case *i.e.* through the power ratio of the contributions of all the sources to the additive noise.

$$\text{SNR}_{dB} \quad = \quad (\|\mathbf{M}\|^2)_{dB} - (N_0)_{dB}, \tag{41}$$

where

$$\|\mathbf{M}\|^2 \quad = \quad \sum_{k \in \mathbb{Z}} \sum_{i=1}^{N} \sum_{j=1}^{K} |M_{i,j}|^2(k), \tag{42}$$

$\mathbf{M}(n)$ is the convolutive mixing system matrix.
We consider here a mixture of $K = 3$ source signals in two situations: with i.i.d. and non i.i.d. source signals. The length of the mixing filter is $L = 3$ and the number of observations is $N = 6$. In the figures 9 and 10 the error index is drawn versus the SNR with $N_e = 5000$ and we compare the performance of the E_2, E_3 and E_4 algorithms.

1) i.i.d. case: In the figure 9, we consider a mixture of 16-QAM source signals. Our approach performs better than E_4 for low SNR values, under $10\ dB$. After that point, we get back the same results as in the figure 6 for $N_e = 5000$: E_3 and E_4 are performing similarly and better than E_2.

2) non i.i.d. case: In the figure 10, we consider a mixture of CPM source signals. We fix the modulation index h to 0.2. In this case all the algorithms are performing well for SNR values above $15\ dB$. After $50\ dB$, our algorithm and E_4 are performing similarly better than the E_2. The E_3 algorithm is less computationally intensive than E_4 (fourth order tensor versus third order tensor), and the extraction thanks to E_3 is done in allmost all the situation presented here.

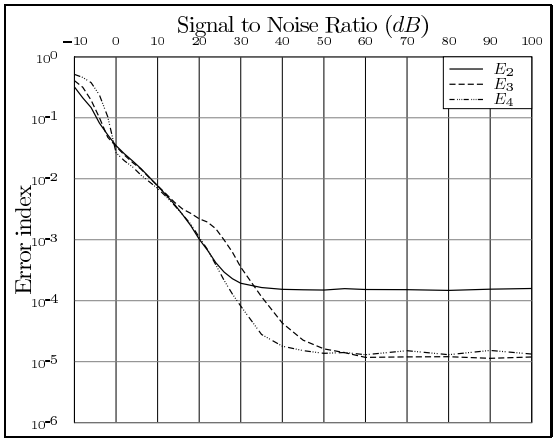

Fig. 10. Complex-valued non i.i.d. case: Performance versus SNR for $N_e = 5000$.

VII. Conclusion

We have proposed generalized contrast functions within the framework of blind extraction of one source from a MIMO convolutive mixture. This family of contrasts holds for both i.i.d. and non i.i.d. sources. In order to optimize this criterion, we have developped a new generalized algebraic algorithm based on best rank-1 tensor decomposition.

Unlike other referenced approaches, our new contrasts require no constraint on the reference signals. The new approach realizes a good compromise between generality and implementation simplicity. Computer simulations illustrate interesting features and good performance in comparison with quadratic and quartic criteria.

Appendix A
Proof of Proposition III.3

Proof: Using the assumption A3., the multilinearity property of cumulants and the i.i.d. assumption of source signals, we have

$$\kappa_{S,R,\mathbf{r}}\{y(n)\} = \sum_{j,k} \Big(\prod_{l=1}^{S} g_j^{(*)_l}(k) \Big) \kappa_{S,R,\mathbf{r}}\{s_j(n-k)\} \ . \quad (43)$$

Hence

$$\mathcal{C}_{S,R,\mathbf{r}}\{y(n)\} = \left| \sum_{j,k} \Big(\prod_{l=1}^{S} g_j^{(*)_l}(k) \Big) \kappa_{S,R,\mathbf{r}}\{s_j(n-k)\} \right| , \quad (44)$$

and thus

$$\mathcal{C}_{S,R,\mathbf{r}}\{y(n)\} \leqslant \sum_{j,k} |g_j(k)|^S \, |\kappa_{S,R,\mathbf{r}}\{s_j(n-k)\}| \ . \quad (45)$$

As $y(n)$ is unit power, we have $\sum_{j,k} |g_j(k)|^2 = 1$.

Hence $|g_j(k)|^2 \leqslant 1$ and as $S > 2$ we have for all j and k

$$|g_j(k)|^S \leqslant |g_j(k)|^2 .$$

Using this result and the assumption A4., we have

$$\begin{aligned} \mathcal{C}_{S,R,\mathbf{r}}\{y(n)\} &\leqslant \sum_{j,k} |g_j(k)|^2 \, |\kappa_{S,R,\mathbf{r}}\{s_j(n-k)\}| \\ &\leqslant \max_{j=1}^{N} \sup_{k\in\mathbb{Z}} |\kappa_{S,R,\mathbf{r}}\{s_j(n-k)\}| \sum_{j,k} |g_j(k)|^2 \\ &\leqslant \max_{j=1}^{N} \sup_{k\in\mathbb{Z}} |\kappa_{S,R,\mathbf{r}}\{s_j(n-k)\}| \ . \end{aligned} \quad (46)$$

Finally considering the equality, we have $|g_j(k)|^S = |g_j(k)|^2$ only if it exists a given couple (j_0, k_0) such that

$$\begin{cases} g_{j_0}(k_0) & = 1 \\ g_j(k) & = 0 \ \ \forall (j,k) \neq (j_0,k_0) \ . \end{cases} \quad (47)$$

(47) corresponds to the equalization condition. ■

Appendix B
Proof of Proposition III.4

Proof: We want to prove that at most one component of the norm of the global filter is non-zero.
For any $i \in \{1,\dots,K\}$, we write the i-th component of the unit norm vector filter $\mathbf{g}$: $g_i = \|g_i\|_i \tilde{g}_i$ where $\tilde{g}_i$ reads $g_i/\|g_i\|_i$ if $\|g_i\|_i \neq 0$ and $\tilde{g}_i$ is zero if $\|g_i\|_i = 0$.
Defining $\tilde{y}_i(n) \triangleq (\tilde{g}_i \star s_i)(n)$, from the assumption A3., the cumulant multilinearity and the mutual independance of the signals $\tilde{y}_i(n)$, $i \in \{1,\dots,K\}$, yield:

$$\kappa_{S,R,\mathbf{r}}\{y(n)\} = \sum_{i=1}^{K} \|g_i\|_i^S \ \kappa_{S,R,\mathbf{r}}\{\tilde{y}_i(n)\}. \quad (48)$$

Then we have the following inequality:

$$\mathcal{C}_{S,R,\mathbf{r}}\{(\mathbf{g}\star\mathbf{s})(n)\} \leqslant \sum_{i=1}^{K} \|g_i\|_i^S \ \|\kappa_{S,R,\mathbf{r}}\{\tilde{y}_i(n)\}\|. \quad (49)$$

Using (19), $\|\tilde{g}_i\|_i = 1$ and $\|\mathbf{g}\| = 1$, we get:

$$\begin{aligned} \mathcal{C}_{S,R,\mathbf{r}}\{(\mathbf{g}\star\mathbf{s})(n)\} &\leqslant \sum_{i=1}^{K} \|g_i\|_i^S \ \mathcal{M}_{S,R,i} \\ &\leqslant \mathcal{M}_{S,R}^{\max} \sum_{i=1}^{K} \|g_i\|_i^S \\ &\leqslant \mathcal{M}_{S,R}^{\max} \sum_{i=1}^{K} \|g_i\|_i^2 \\ &\leqslant \mathcal{M}_{S,R}^{\max} \end{aligned} \quad (50)$$

The last inequality holds because, as $S > 2$,

$$\sum_{i=1}^{K} \|g_i\|_i^S \leqslant \sum_{i=1}^{K} \|g_i\|_i^2 = 1 \ .$$

The previous inequality proves that the property p1. of the definition (III.2) is satisfied.
Moreover, if the equality is satisfied, we can write

$$\sum_{i=1}^{K} \|g_i\|_i^S = \sum_{i=1}^{K} \|g_i\|_i^2 = 1. \quad (51)$$

It comes that for all $i \in \{1,\dots,K\}$, at most one $\|g_i\|_i$ is non zero. Additionally this value of i is such that $\mathcal{M}_{S,R,i} =$

$\mathcal{M}_{S,R}^{\max}$. Thus, the property p2. of (III.2) is satisfied. The converse is easily deduced, and the proposition (III.2) is satisfied.

■

REFERENCES

[1] A. Adib, E. Moreau, and D. Aboutajdine, "Source Separation Contrasts using a Reference Signal," *IEEE Signal Processing Letters*, vol. 11, no. 3, pp. 312–315, March 2004.

[2] M. Castella, J. Pesquet, and A. P. Petropulu, "A family of frequency- and time-domain contrasts for blind separation of convolutive mixtures of temporally dependent signals," *IEEE Transactions on Signal Processing*, vol. 53, no. 1, pp. 107–120, January 2005.

[3] M. Castella, S. Rhioui, E. Moreau, and J.-C. Pesquet, "Quadratic Higher-Order Criteria for Iterative Blind Separation of a MIMO Convolutive Mixture of Sources," *IEEE Transactions on Signal Processing*, vol. 55, no. 1, pp. 218–232, January 2007.

[4] P. Comon, "Contrasts for Multichannel Blind Deconvolution," *IEEE Signal Processing Letters*, vol. 3, no. 7, pp. 209–211, July 1996.

[5] P. Comon and L. Rota, "Blind Separation of Independent Sources from Convolutive Mixtures," *IEICE Transactions on Fundamentals of Electronics, Communications, and Computer Sciences*, vol. E86A, no. 3, pp. 542–549, March 2003.

[6] L. De Lathauwer, P. Comon, B. De Moor, and J. Vandewalle, "Higher-order power method - application in independent component analysis," in *International Symposium on Nonlinear Theory and its Applications, NOLTA'95*, Las Vegas, USA, December 1995, pp. 91–96.

[7] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A Multilinear Singular Value Decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, April 2000.

[8] ——, "On the Best Rank-1 and Rank-(R_1,R_2, ... ,R_N) Approximation of Higher-Order Tensors," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1324–1342, May 2000.

[9] N. Delfosse and P. Loubaton, "Adaptive blind separation of independent sources: A deflation approach," *Signal Processing*, vol. 45, no. 1, pp. 59–83, July 1995.

[10] R. Dubroca, C. De Luigi, and E. Moreau, "Cubic Higher-Order Criterion and Algorithm for Blind Extraction of a Source Signal," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'09*, Taipei, Taiwan, April 2009, pp. 3149–3152.

[11] H. Ghennioui, E. M. Fadaili, N. Thirion-Moreau, A. Adib, and E. Moreau, "A Nonunitary Joint Block Diagonalization Algorithm for Blind Separation of Convolutive Mixtures of Sources," *IEEE Signal Processing Letters*, vol. 14, no. 11, pp. 860–863, November 2007.

[12] A. Gorokhov and P. Loubaton, "Subspace-Based Techniques for Blind Separation of Convolutive Mixtures with Temporally Correlated Sources," *IEEE Transactions on Circuits and Systems*, vol. 44, no. 9, pp. 813–820, September 1997.

[13] Y. Inouye and T. Sato, "Iterative Algorithms Based on Multistage Criteria for Multichannel Blind Deconvolution," *IEEE Transactions on Signal Processing*, vol. 47, no. 6, pp. 1759–1764, June 1999.

[14] B. Jelonnek, D. Boss, and K.-D. Kammeyer, "Generalized eigenvector algorithm for blind equalization," *Signal Processing*, vol. 61, no. 3, pp. 237–264, September 1997.

[15] M. Kawamoto, K. Kohno, and Y. Inouye, "Eigenvector Algorithms Incorporated with Reference Systems for Solving Blind Deconvolution of MIMO-IIR Linear Systems," *IEEE Signal Processing Letters*, vol. 14, no. 12, pp. 996–999, December 2007.

[16] P. Loubaton and P. Regalia, "Blind Deconvolution of Multivariate Signals: A Deflation Approach," in *IEEE International Conference on Communications, ICC'93*, 1993.

[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] E. Moreau and J.-C. Pesquet, "Generalized Contrasts for Multichannel Blind Deconvolution of Linear Systems," *IEEE Transactions on Signal Processing*, vol. 4, no. 6, pp. 182–183, June 1997.

[19] C. Simon, P. Loubaton, and C. Jutten, "Separation of a Class of Convolutive Mixtures: A Contrast Function Approach," *Signal Processing*, vol. 81, pp. 883–887, 2001.

[20] J. K. Tugnait, "Identification and Deconvolution of Multichannel Linear Non-Gaussian Processes Using Higher Order Statistics and Inverse Filter Criteria," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 658–672, March 1997.

Rémi Dubroca was born in Nîmes, France, in 1980. He received the M.Sc. degrees, both in Telecommunications engineering in 2004 from Institut des Sciences de l'Ingénieur de Toulon et du Var (ISITV), La Valette, France and in Signal, Image and Digital Communication (SICOM) from University of Nice, Sophia Antipolis, France. He obtained his Ph.D. degree from "Université du Sud Toulon Var" (France) in 2008. Since 2008 he has been involved with blind sources separation and independent component analysis (ICA), his research interests include both the convolutive mixtures and the contrast functions.

Christophe De Luigi (M'08) was born in Woippy, France. He graduated from the Engineering School "Institut des Sciences de l'Ingénieur de Toulon et du Var" (ISITV), La Valette du Var, France, in 1995. He received the DEA degree in 1995 in the field of applied mathematics from the University of Provence, France, and the Ph.D. degree in 2000 in the field of signal processing from the University of Toulon, France. From 2000 to 2003, he was in post-doctoral position with the University of Toulon (2 years) and with STMicroelectronics (1 year) in the field of wireless telecommunications. Since 2003, he has been an assistant professor within ISITV's Telecommunications Department. His main research interests are in statistical signal processing using higher-order statistics.

Marc Castella was born in 1976 in Courbevoie, France. In 2000, he received the "Agrégation" degree in the field of applied physics and in 2001 he received the M.Sc. degree in electrical engineering, from both the "École Normale Supérieure de Cachan" and the "Université Paris-Sud, Orsay" (France). He obtained his Ph.D. degree from "Université de Marne-la-Vallée" (France) in 2004. Since then, he has been "Maître de Conférence" in Evry (France) at "Télécom & Management SudParis" (formerly "Institut National des Télécommunications (INT)"). He is also a member of the UMR-CNRS 5157 research team SAMOVAR. His research activities have focused on the problem blind source separation in general.

Eric Moreau (M'96-SM'08) was born in Lille, France. He graduated from the "Ecole Nationale Supérieure des Arts et Métiers" (ENSAM), 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.