Reference based contrast functions in a semi-blind context

Marc Castella[1] and Eric Moreau[2]

[1] Institut TELECOM; TELECOM & Management SudParis; UMR-CNRS 5157
9 rue Charles Fourier, 91011 Evry CEDEX, France
marc.castella@it-sutparis.eu
[2] University of Sud Toulon Var, ISITV, LSEET UMR-CNRS 6017
Av. G. Pompidou, BP56, 83162 La Valette du Var CEDEX, France
moreau@univ-tln.fr

Abstract. We deal with blind signal extraction in the framework of a convolutive mixture of independent sources. Considering so-called reference signals, we generalize former identifiability conditions. Based on this result, we propose to incorporate some a priori information in the references. We show the validity of reference based contrast functions in two semi-blind situations. The results are confirmed by computer simulations.

1 Introduction

This paper deals with the problem of source extraction in the general case of a convolutive mixture. This issue finds applications in many contexts such as blind equalization for digital communication signals,... Moreover, this problem is a key component in blind source separation methods based on a multi-stage (or iterative) approach (often called deflation [4, 5]).

One interesting way to tackle this problem consists in optimizing an adapted criterion called contrast function. However, finding identifiability conditions which state the precise context where source signals can be extracted remain important. This paper considers the above two points.

In this context, reference based methods have recently been introduced [1, 2]. They consider a reference signal which provides a supplementary degree of freedom and which can be used in different ways. Here we propose to incorporate some a priori information in the reference signal and we show its effectiveness. New identifiability results which generalize former conditions on the reference are first proposed. Then, we show the validity of reference based contrast functions in two semi-blind interesting situations. Finally, computer simulations illustrate our results.

2 Problem statement

2.1 Mixing model

We consider a Q-dimensional ($Q \in \mathbb{N}, Q \geq 2$) discrete-time signal which is called vector of *observations* and denoted by $\mathbf{x}(n)$ (in the whole paper, n stands for any integer: $n \in \mathbb{Z}$). It results from a linear time invariant (LTI) multichannel system $\{\mathbf{M}\}$ described by the input-output relation:

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

where $\mathbf{M}(n)$ is the sequence of (Q, N) impulse response matrices and $\mathbf{s}(n)$ is a N-dimensional ($N \in \mathbb{N}^*$) unknown and unobserved column vector, which is referred to as the vector of *sources*. The multichannel deconvolution problem consists in estimating a multivariate LTI system $\{\mathbf{W}\}$ operating on the observations, such that the vector $\mathbf{y}(n) = \sum_{k \in \mathbb{Z}} \mathbf{W}(k)\mathbf{x}(n-k) \triangleq \{\mathbf{W}\}\mathbf{x}(n)$ restores the N input sources. Different situation can occur depending on context: when no information is available on the mixing system and when the sources cannot be observed, the problem is referred to as the *blind source separation* (BSS) problem. On the contrary, when some information is assumed to be known on the sources or on the mixing system, the problem is sometimes referred to as *semi-blind source separation* (S-BSS).

The matrix transfer function $\mathbf{M}[z_1]$ of the mixing channel is given by the following z-transform (where z_1 is used instead of z for reasons that appear later)

$$\mathbf{M}[z_1] = \sum_{n \in \mathbb{Z}} \mathbf{M}(n) z_1^{-n} \ . \tag{2}$$

A similar definition holds for the matrix transfer function $\mathbf{W}[z_1]$ of the separator $\{\mathbf{W}\}$. We define the combined mixing-separating (N, N) LTI filter $\{\mathbf{G}\}$ by its impulse response $\mathbf{G}(n) = \sum_{k \in \mathbb{Z}} \mathbf{W}(k)\mathbf{M}(n-k)$.

In an iterative approach, the sources are extracted one by one. We consider one row of the separator $\{\mathbf{W}\}$, which is a $(1, Q)$ LTI row filter $\{\mathbf{w}\}$ with output $y(n)$ given by:

$$y(n) = \sum_{k \in \mathbb{Z}} \mathbf{w}(k)\mathbf{x}(n-k) \triangleq \{\mathbf{w}\}\mathbf{x}(n). \tag{3}$$

Similarly, $\{\mathbf{g}\}$ denotes the $(1, N)$ row filter given by the row of $\{\mathbf{G}\}$ which corresponds to $\{\mathbf{w}\}$. We have then:

$$y(n) = \sum_{k \in \mathbb{Z}} \mathbf{g}(k)\mathbf{s}(n-k) \triangleq \{\mathbf{g}\}\mathbf{s}(n). \tag{4}$$

We say that the separation is achieved when only one component of $\{\mathbf{g}\}$ is non zero, say the i_0th:

$$\{\mathbf{g}\} = \Big(0, \ldots, 0, \{g_{i_0}\}, 0, \ldots, 0\Big). \tag{5}$$

When the source signals are assumed to be temporally i.i.d. (independent and identically distributed) signals, a more restrictive separation condition is considered. In this case, $\{\mathbf{g}\}$ should satisfy the following condition in addition to (5):

$$\exists l \in \mathbb{N}: \quad g_{i_0}(k) = 0 \text{ if } k \neq l \text{ and } g_{i_0}(l) = \alpha \in \mathbb{C}^* \tag{6}$$

The parameters i_0, l, and α correspond respectively to the well-known indeterminacies of BSS (or S-BSS): permutation of the sources, delay and amplitude factor. In the case of non i.i.d. sources, a filtering ambiguity is included by specifying no constraint on $\{g_{i_0}\}$.

2.2 Reference based separation

In order to be able to solve the BSS or S-BSS problem, we introduce the following assumption on the sources:

A1. The source vector components $s_i(n), i \in \{1, \ldots, N\}$ are *mutually independent*, stationary and zero-mean processes with unit variance.

In addition, we assume that there exist an additional signal $r(n)$, which is referred to as a "reference" signal. This reference signal is used in order to facilitate the source separation and the following assumption is made:

A2. The signals $s_i(n), i \in \{1, \ldots, N\}$ and $r(n)$ are jointly stationary up to the fourth order.

It has been shown that $r(n)$ can generally be constructed from the observations. The corresponding constraints on $r(n)$ being quite weak, an efficient BSS procedure has been proposed based on the maximization of a reference based contrast function [1, 2]. In the following Section, we give general conditions on $r(n)$ so that the extraction of a particular source is allowed. Furthermore, it is rather natural to use the reference signal in a S-BSS context: we indeed propose in Section 4 to include in $r(n)$ some a priori information on the source which is being extracted. The validity of the corresponding separation criteria is proved.

3 Generalized identifiability conditions

Higher order statistics We consider higher order cumulants and define the following fourth order cross-cumulant:

$$C^r_{s_i s_j}(\mathbf{n}) = \mathrm{Cum}\left\{s_i(n), s_j(n-n_1)^*, r(n-n_2), r(n-n_2)^*\right\}. \tag{7}$$

Note that it exists and depends on $\mathbf{n} = (n_1, n_2)$ only according to assumption A2. We consider a multidimensional z-transform of order two w.r.t. the variables $\mathbf{n} = (n_1, n_2)$. It is defined by:

$$C^r_{s_i s_j}[\mathbf{z}] = \sum_{\mathbf{n} \in \mathbb{Z}^2} C^r_{s_i s_j}(\mathbf{n}) z_1^{-n_1} z_2^{-n_2} \tag{8}$$

where $\mathbf{z} = (z_1, z_2) \in (\mathbb{C}^*)^2$. For the signal $\mathbf{x}(n)$, the cumulants $C^r_{x_i x_j}(\mathbf{n})$ and their respective z-transforms $C^r_{x_i x_j}[\mathbf{z}]$ are defined similarly to the definitions given by (7) and (8). Let us introduce the matrices $\mathbf{C}^r_{\mathbf{x}}[\mathbf{z}]$ and $\mathbf{C}^r_{\mathbf{s}}[\mathbf{z}]$ whose (i, j) components are $C^r_{x_i x_j}[\mathbf{z}]$ and $C^r_{s_i s_j}[\mathbf{z}]$ respectively.

Cumulant decomposition Using the multilinearity property of cumulants, it can be verified that the following decomposition holds:

$$\mathbf{C}^r_{\mathbf{x}}[\mathbf{z}] = \mathbf{M}[z_1 z_2]\mathbf{C}^r_{\mathbf{s}}[\mathbf{z}]\mathbf{M}[z_1^{-1}]^{\mathrm{H}} . \tag{9}$$

Second order statistics appear formally as a particular case of the preceding results. When there is no reference signal, (7) corresponds to the correlation for which we adopt the specific notation:

$$\Gamma_{s_i s_j}(n_1) \triangleq \mathrm{Cum}\,\{s_i(n), s_j(n-n_1)^*\} \qquad \mathbf{\Gamma}_{\mathbf{s}}(n_1) = \left(\Gamma_{s_i s_j}(n_1)\right)_{(i,j)\in\{1,\ldots,N\}^2}.$$

Corresponding to (8), the power spectral matrix of the sources is defined by:

$$\mathbf{\Gamma}_{\mathbf{s}}[z_1] \triangleq \sum_{n_1\in\mathbb{Z}} \mathbf{\Gamma}_{\mathbf{s}}(n_1) z_1^{-n_1}. \tag{10}$$

Note that contrary to (9), z_2 does not appear here. Similar notations hold for the observations $\mathbf{x}(n)$ ($\mathbf{\Gamma}_{\mathbf{x}}(n_1)$ and $\mathbf{\Gamma}_{\mathbf{x}}[z_1]$ respectively). Similarly to (9), we have the well-known relation:

$$\mathbf{\Gamma}_{\mathbf{x}}[z_1] = \mathbf{M}[z_1]\mathbf{\Gamma}_{\mathbf{s}}[z_1]\mathbf{M}[z_1^{-1}]^{\mathrm{H}}. \tag{11}$$

Identifiability condition Based on (9) and (11), the following general identifiability result can be proved. The proof proceeds along similar lines as the one in [3] and it is not provided because of lack of space.

Proposition 1. *Assume that*

A3. the matrix $\mathbf{C}^r_{\mathbf{s}}[\mathbf{z}]$ is diagonal, and the first diagonal element of $\mathbf{C}^r_{\mathbf{s}}[\mathbf{z}]$ is a function distinct from all other diagonal elements.

Then, the first source of the mixture is identifiable in the following sense: for any $N \times Q$ matrix $\mathbf{W}[z_1]$ such that the $N \times N$ combined channel-equalizer z-transform matrix $\mathbf{G}[z_1] = \mathbf{W}[z_1]\mathbf{M}[z_1]$ is irreducible, if $\mathbf{W}[z_1]\mathbf{\Gamma}_{\mathbf{x}}[z_1]\mathbf{W}[z_1^{-1}]^{\mathrm{H}}$ and $\mathbf{W}[z_1 z_2]\mathbf{C}^r_{\mathbf{x}}[\mathbf{z}]\mathbf{W}[z_1^{-1}]^{\mathrm{H}}$ are both diagonal, then:

$$\mathbf{G}[z_1] = \mathbf{P}\begin{pmatrix} \alpha & 0 \ldots 0 \\ 0 & \\ \vdots & \tilde{\mathbf{Q}} \\ 0 & \end{pmatrix} \tag{12}$$

where $\alpha \in \mathbb{C}, \alpha \neq 0$ and $\mathbf{P}$ is a permutation matrix. In addition, $\tilde{\mathbf{Q}}$ is a unitary matrix.

The above result corresponds to the possibility of extracting the first source. Indeed, if $\mathbf{P}$ is the identity matrix, the first row of the above $\mathbf{G}[z_1]$ is such that the corresponding row LTI filter satisfies both (5) and (6).

4 Contrast functions

4.1 BSS contrast functions

A commonly used approach in BSS consist in finding criteria, which yield separation when they reach their maximum value: such criteria are by definition called *contrast functions*. The modulus of the fourth-order auto-cumulant is one of the most popular contrast [4, 5]. In [1], a more general expression of the following contrast has been introduced:

$$\mathcal{C}_r\{y(n)\} \triangleq |\kappa_r\{y(n)\}| \text{ where: } \kappa_r\{y(n)\} \triangleq \text{Cum}\{y(n), y(n)^*, r(n), r(n)^*\} \quad (13)$$

with $r(n)$ the reference signal. It has been proved that (13) is a valid contrast under the constraint $\mathbb{E}\left\{|y(n)|^2\right\} = 1$ and additional constraint on the reference signal $r(n)$.

Consider the specific situation where the reference $r(n)$ depends only on the source signal which is extracted (say $s_1(n)$) and is independent on the other source signals. One can then see that only the first element of the matrix $\mathbf{C}_{\mathbf{s}}^r[\mathbf{z}]$ is non zero and condition A3 of Proposition 1 is satisfied. This situation should thus be very favorable in order to extract $s_1(n)$ and we illustrate two such situations of interest. Of course, in such a situation, the original permutation ambiguity is reduced and necessarily, the first source is extracted, that is $i_0 = 1$ in (5)-(6).

4.2 S-BSS contrast function: information on the signal phase

We consider in this paragraph the situation where the reference signal reads:

$$\begin{cases} r(n) = \epsilon(n)\frac{s_1(n)}{|s_1(n)|} \text{ where:} \\ |\epsilon(n)| = 1 \text{ and } \epsilon(n) \text{ is independent of the source signals.} \end{cases} \quad (14)$$

This situation corresponds to the case where the reference contains some information on the phase of $s_1(n)$. This information is corrupted by the perturbation process $\epsilon(n)$ which randomly changes the phase.

Lemma 1. *In the case of i.i.d. source signals, if $r(n)$ is given by (14), we have (where ϵ and s_1 denote $\epsilon(n)$ and $s_1(n)$ for readability):*

$$\mathcal{C}_r\{s_j(n-k)\} = \begin{cases} 0 & \text{if } (j,k) \neq (1,0), \\ |\mathbb{E}\{\epsilon\}|^2 \left[\left|\mathbb{E}\left\{\frac{s_1^2}{|s_1|}\right\}\right|^2 + |\mathbb{E}\{|s_1|\}|^2\right] & \text{if } (j,k) = (1,0). \end{cases} \quad (15)$$

Proof. For $(j,k) \neq (1,0)$, the equality $\kappa_r\{s_j(n-k)\} = 0$ follows from independence and the vanishing property of the cumulants. Now, if $(j,k) = (1,0)$, by developing the cumulant in term of moments, we obtain:

$$\kappa_r\{s_1(n)\} = \mathbb{E}\left\{|s_1|^2|\epsilon|^2\right\} - \mathbb{E}\left\{|s_1|^2\right\}\mathbb{E}\left\{|\epsilon|^2\right\} - \left|\mathbb{E}\left\{\epsilon\frac{s_1^2}{|s_1|}\right\}\right|^2 - \mathbb{E}\{\epsilon|s_1|\}|^2$$

The result then follows from the independence of $\epsilon(n)$ and $s_1(n)$ and after simplification. □

We can now state the following result:

Proposition 2. *Considering the reference signal given by (14), the criterion $\mathcal{C}_r$ in (13) is a contrast function for i.i.d. sources under the condition $\mathbb{E}\{\epsilon(n)\} \neq 0$. In addition, the source identified by $\mathcal{C}_r$ is $s_1(n)$.*

Proof. Note that, except in the degenerate case where $s_1(n) = 0$ almost surely, $\mathcal{C}_r\{s_1(n)\} > 0$ when $\mathbb{E}\{\epsilon(n)\} \neq 0$. The proposition then follows from Lemma 1 and a straightforward application of [1, Prop. 1]. □

4.3 S-BSS: information on the signal modulus

This paragraph is concerned with the case where the reference signal has the same modulus as $s_1(n)$ but a random phase. All the results and comments in the previous section are adapted hereunder. More precisely, the reference signal reads:

$$\begin{cases} r(n) = \epsilon(n)|s_1(n)| \\ |\epsilon(n)| = 1, \mathbb{E}\{\epsilon(n)\} = 0 \text{ and } \epsilon(n) \text{ is independent of the sources.} \end{cases} \tag{16}$$

Lemma 2. *In the case of i.i.d. source signals, if $r(n)$ is given by (16), we have (where ϵ and s_1 denote $\epsilon(n)$ and $s_1(n)$ again for readability):*

$$\mathcal{C}_r\{s_j(n-k)\} = \begin{cases} 0 & \text{if } (j,k) \neq (1,0), \\ \mathbb{E}\left\{|s_1(n)|^4\right\} - \mathbb{E}\left\{|s_1(n)|^2\right\}^2 & \text{if } (j,k) = (1,0). \end{cases} \tag{17}$$

Proof. For $(j,k) \neq (1,0)$, the equality $\kappa_r\{s_j(n-k)\} = 0$ follows from independence and the vanishing property of the cumulants. Now, if $(j,k) = (1,0)$, by developing the cumulant in term of moments, we obtain:

$$\kappa_r\{s_1(n)\} = \mathbb{E}\left\{|s_1|^4|\epsilon|^2\right\} - \mathbb{E}\left\{|s_1|^2\right\}\mathbb{E}\left\{|\epsilon s_1|^2\right\} - |\mathbb{E}\{\epsilon s_1|s_1|\}|^2 - \mathbb{E}\{\epsilon s_1^*|s_1|\}|^2$$

The result then follows after simplification, using the independence of $\epsilon(n)$ and $s_1(n)$ and the assumption $\mathbb{E}\{\epsilon(n)\} = 0$. □

Proposition 3. *Considering the reference signal given by (16), the criterion $\mathcal{C}_r$ in (13) is a contrast function for i.i.d. sources if $s_1(n)$ has non constant modulus. In addition, the source identified by $\mathcal{C}_r$ is $s_1(n)$.*

Proof. Note that, if $s_1(n)$ does not have constant modulus, $\mathcal{C}_r\{s_1(n)\} > 0$ since it is the variance of $|s_1(n)|$. The proposition then follows from Lemma 1 and a straightforward application of [1, Prop. 1]. □

4.4 Generalization

As mentioned, the reference signal $r(n)$ defined by (14) or (16) satisfy the conditions of Proposition 1 because it depends on $s_1(n)$ only. Obviously, if $\widetilde{r}(n)$ is a scalar filtering of $r(n)$, it also depends on $s_1(n)$ only. More precisely, one can see that any scalar filtering of a reference given by (14) or (16) remains a valid reference signal under the same conditions as the above given ones. Finally, we would like to stress that the results in Section 4 remain valid in the case of non i.i.d. sources, although not presented here due to the lack of space.

5 Simulations

We considered three different kind of sources:

- real valued, uniformly distributed with mean zero and unit variance.
- complex valued, QAM4 and QAM16, that is the real and imaginary parts are independent taking their values with equal probability in $\{\pm 1/\sqrt{2}\}$ for QAM4 sources and in $\{\pm 1/\sqrt{5}, \pm 3/\sqrt{5}\}$ for QAM16 sources.

For each of the above choices, a set of $N = 3$ mutually independent and temporally i.i.d. sources have been generated. They have been mixed by a $Q \times N$ randomly chosen finite impulse response (FIR) filter of length 3 and with $Q = 4$ sensors. We then used the contrast $\mathcal{C}_r$ in (13) and the associated algorithm proposed in [1] to test the effectiveness of our result. The different choices for $r(n)$ are detailed next. All presented results correspond to averaged values over 1000 Monte-Carlo realizations of the mean square error (MSE) on the estimated source.

5.1 Information on the signal phase

The original reference $r(n)$ is given by (14), where:

- $\epsilon(n)$ is a binary i.i.d. Bernoulli process with $P(\epsilon(n) = 1) = p$ and $P(\epsilon(n) = -1) = 1 - p$ in the case of real-valued sources.
- $\epsilon(n) = e^{\imath\theta(n)}$ and $\theta(n)$ is uniformly distributed on an interval $[-\vartheta, \vartheta]$ in the case of complex-valued sources.

According to Section 4.4, we consider also $\widetilde{r}(n)$ as a reference, where $\widetilde{r}(n)$ has been obtained by a FIR scalar filtering of $r(n)$. The filter have three taps with randomly driven coefficients. Different values of the parameters p, ϑ and different sample sizes are considered. The results are reported in Table 1 for real-valued sources and in Table 2 for complex-valued sources.

One can observe the effectiveness of the separation when $p \neq 0.5$ and $\vartheta \neq \pi$: it should indeed be no surprise that no separation is obtained for $p = 0.5$ or $\vartheta = \pi$ since in this case, $r(n)$ is independent of $s_1(n)$.

Number of samples	Reference: $r(n)$				Reference: $\widetilde{r}(n) \leftrightarrow$ "$r(n)$+filt."			
	$p = 1$	$p = 0.9$	$p = 0.7$	$p = 0.5$	$p = 1$	$p = 0.9$	$p = 0.7$	$p = 0.5$
1000	0.0078	0.0249	0.1548	1.0780	0.0379	0.0945	0.7079	1.1340
5000	0.0015	0.0050	0.0331	1.0799	0.0130	0.0748	0.2470	1.1303
10000	7.57e-4	0.0025	0.0168	1.0762	0.0073	0.0142	0.1383	1.1335

Table 1. Average (1000 realizations) MSE on the reconstructed source (real-valued sources with uniform distribution). $r(n)$ is given by (14).

Sources	Number of samples	Reference: $r(n)$			
		$\vartheta = 0$	$\vartheta = \frac{\pi}{4}$	$\vartheta = \frac{\pi}{2}$	$\vartheta = \pi$
QAM4	1000	5.84e-5	0.0055	0.0334	1.2114
	5000	2.11e-6	0.0011	0.0067	1.2160
	10000	4.98e-7	5.40e-4	0.0034	1.2119
QAM16	1000	0.0027	0.0087	0.0395	1.2116
	5000	5.23e-4	0.0017	0.0080	1.2133
	10000	2.63e-4	8.65e-4	0.0040	1.2126

Table 2. Average (1000 realizations) MSE on the reconstructed source (complex valued sources with QAM4 or QAM16 distribution). $r(n)$ is given by (14).

5.2 Information on the signal modulus

Now the original reference $r(n)$ is given by (16), where:

- $\epsilon(n)$ is a binary i.i.d. process with $P(\epsilon(n) = 1) = P(\epsilon(n) = -1) = 1/2$ in the case of real-valued sources.
- $\epsilon(n) = e^{\imath\theta(n)}$ and $\theta(n)$ is uniformly distributed on the interval $[-\pi, \pi]$ in the case of complex-valued sources.

Similarly to the previous paragraph, we considered $r(n)$ and $\widetilde{r}(n)$ as reference signals, where $\widetilde{r}(n)$ is a scalar filtering of $r(n)$. The results are showed in Table 3. The reader can see the effectiveness of our method for uniformly distributed and QAM16 sources. On the contrary, QAM4 sources have constant modulus and thus do not satisfy the conditions of Proposition 3.

Sources		Uniform			QAM16			QAM4
Number of samples		1000	5000	10000	1000	5000	10000	10000
Reference	$r(n)$	0.0238	0.0045	0.0022	0.0548	0.0100	0.0050	1.2080
	$\widetilde{r}(n) \leftrightarrow$ "$r(n)$+filt."	0.2433	0.0651	0.0408	0.6247	0.1930	0.1190	1.2690

Table 3. Average (1000 realizations) MSE on the reconstructed source. $r(n)$ is given by (16).

References

1. Castella, M., Rhioui, S., Moreau, E., Pesquet, J.C.: Quadratic higher-order criteria for iterative blind separation of a MIMO convolutive mixture of sources. IEEE Trans. Signal Processing **55**(1) (January 2007) 218–232
2. Kawamoto, M., Kohno, K., Inouye, Y.: Eigenvector algorithms incorporated with reference systems for solving blind deconvolution of MIMO-IIR linear systems. IEEE Signal Processing Letters **14**(12) (December 2007) 996–999
3. Hua, Y., Tugnait, J.K.: Blind identifiability of FIR-MIMO systems with colored input using second order statistics. IEEE Signal Processing Letters **7**(12) (December 2000) 348–350
4. Simon, C., Loubaton, P., Jutten, C.: Separation of a class of convolutive mixtures: a contrast function approach. Signal Processing (81) (2001) 883–887
5. Tugnait, J.K.: Identification and deconvolution of multichannel linear non-gaussian processes using higher order statistics and inverse filter criteria. IEEE Trans. Signal Processing **45**(3) (March 1997) 658–672