A Double Referenced Contrast for Blind Source Separation

Atman Jbri, Abdellah Adib, and Driss Aboutajdine

Abstract—This paper addresses the problem of blind source separation (BSS). To recover original signals, from linear instantaneous mixtures, we propose a new contrast function based on the use of a double referenced system. Our approach assumes statistical independence sources. The reference vectors will be incurred in the cumulant to evaluate the independence. The estimation of the separating matrix will be performed in two steps: whitening observations and joint diagonalization of a set of referenced cumulant matrices. Computer simulations are presented to demonstrate the effectiveness of the suggested approach.

Keywords—Blind source separation, Referenced Cumulant, Contrast, Joint Diagonalization.

I. INTRODUCTION

In this paper, we consider the blind source separation problem [1][2] which relates to separating signals without information on the signals or the signal mixtures. It finds numerous applications in diverse fields of engineering and applied sciences, e.g. seismic and astrophysics exploration [3], speech processing [5], data communications [4] and biomedical processing [6]. The data model can be formulated as follow. A set of M sensors receiving an instantaneous linear mixture of signals emitted from N ≤ M sources. The vector $\mathbf{x}(t) = [x_1(t), \cdots, x_M(t)]^T \in \mathbb{C}^{M \times 1}$ denotes the output of sensors at time $t$ which may be corrupted by an independent additive noise $\mathbf{n}(t)$. We have then

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \quad (1)$$

where the matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$ is called mixing matrix and $\mathbf{s}(t) = [s_1(t), \cdots, s_N(t)]^T \in \mathbb{C}^{N \times 1}$ the vector of sources. In this context, the objective is to find an estimate of the matrix $\mathbf{A}$ or, more precisely its inverse $\mathbf{B}$ so as to perform source separation, i.e. to recover the sources $\mathbf{s}(t)$ from the observations $\mathbf{x}(t)$. As sources are unobservable, there are some inherent indeterminacy in their estimation. That is, in general, we cannot identify the order and the power of each components of the source vector. The separation is considered achieved when the sources are estimated up to a factor and a permutation scales. That means the global matrix Q can be written as

$$\mathbf{Q} = \mathbf{BA} = \mathbf{PD}$$

where, $\mathbf{Q}$ is a weight matrix, $\mathbf{P}$ is a permutation matrix and $\mathbf{D}$ is a non-zero diagonal one. Then, the $N \times 1$ estimation vector is given by $\mathbf{y}(t) = \mathbf{Bx}(t)$.

In the past years, numerous solutions to this problem have been proposed. The contrast functions used second or higher order statistics to express the separation criteria depending on considered source assumptions [1][8][9]. Their optimization gives an estimation of the separating matrix. For example, the IADE algorithm [7] exploits different cumulants to measure the independence as follows

$$I (y) = \sum_{i,j,\ell=1}^{N} |\text{Cum}(y_i(t), y_i^*(t), y_j(t), y_j^*(t))|^2 \quad (2)$$

In this paper, we will introduce a new separation technique using a referenced system as shown in the Fig.1. The double referenced vector will be incurred in the cumulant to construct a new separation contrast, for whitened vectors, called Double Referenced Contrast for Separation DRCS. The optimization of our proposed contrast corresponds to unitary joint diagonalization criterion. The sources are assumed to satisfy the following three assumptions,

A1. The sources are zero-mean, unit power, statistically mutually independent;
A2. The sources are stationary at order $r$: the $r$-th order cumulant $\text{Cum}(s_1(t), \cdots, s_r(t))$ is an independent function of $t$;
A3. The $r$-th order cumulants of all sources have the same sign.

In the next, we denote by $\mathcal{S}$ the set of vectors satisfying these assumptions. This paper is organized as follows: In the following section, we present and demonstrate our proposed contrast and its optimization. Numerical simulations are given in the Section III, followed by conclusion.

Fig. 1. Double referenced separation system.
II. PROPOSED SEPARATION TECHNIQUE

A. Double Referenced Contrast

The contrast is a multivariate function defined on a set $Y$ of random vectors $y$ which only depends on the probability density of $y$ and whose global maxima only correspond to some separation solutions [1]. An important pre-processing step is the whitening observations, i.e. estimate a matrix $W$ such that the vector $u(t) = Wx(t)$ has the covariance $R_{uu} = E[u(t)u(t)^T]$ equals the identity [7]. The mixing problem will be unitary by a matrix $U = WA$, because

$$
R_{uu} = E[u(t)u(t)^T] = WA E[s(t)s(t)^T](WA)^T = U E[s(t)s(t)^T]U^T = UU^T = I_N
$$

(3)

In consequence, the contrast operator works on whitened vectors to estimate an unitary matrix $H$ such that $B = HW$ performs the separation. Let $\mathcal{H}$ denotes the set of unitary matrices, $\mathcal{P}$ the set of permutation matrices and $\mathcal{D}$ the set of invertible diagonal matrices. We recall and adopt the following definition of contrast,

Definition. A contrast on set $Y$ is a multivariate function $\mathcal{J}(\cdot)$ from the set $Y$ to $\mathbb{R}$ which satisfies the following three conditions:

C1. $\forall y \in Y, \forall D \in \mathcal{D}, J(Dy) = J(y)$

C2. $\forall y \in Y, \forall s \in S, J(y) \leq J(s)$

C3. $\forall H \in \mathcal{H}, \forall P \in \mathcal{P}, \exists D \in \mathcal{D}, J(Hs) = J(s) \Rightarrow H = DP$

According to this definition, the maximization of a contrast is a sufficient condition for source separation. For reference separation methods as shown in the Fig.1, we propose, under whitening constraint, the following contrast function,

$$
J_{z,v}(y) = \sum_{i,j=1}^{N} \left| \frac{\text{Cum}(y_i(t), y_j^*(t))}{z_i(t), v_j^*(t), \ldots, z_j(t), v_i^*(t)} \right|^2
$$

(4)

where the cumulant order $r \geq 4$ is pair, $z(t) = [z_1(t), \ldots, z_N(t)]^T \in \mathbb{C}^{N \times 1}$ and $v(t) = [v_1(t), \ldots, v_N(t)]^T \in \mathbb{C}^{N \times 1}$ are two reference vectors.

We note,

$$
C_{z_j,v_j}^r(y_i, y_j) \triangleq \frac{\text{Cum}(y_i(t), y_j^*(t))}{z_i(t), v_j^*(t), \ldots, z_j(t), v_i^*(t)}
$$

(5)

Our proposed contrast can be rewritten as,

$$
J_{z,v}(y) = \sum_{i,j=1}^{N} | C_{z_j,v_j}^r(y_i, y_j) |^2
$$

(6)

To demonstrate our contrast function, we must validate the three conditions cited in the definition.

Condition C1. This condition is trivial for whitened vector. In addition to diagonal structure, the matrix $D$ will be unitary because the whitened nature of any vector $a(t) = Dy(t)$ that represents an input of our proposed function. We have,

$$
J_{z,v}(Dy) = \sum_{i,j=1}^{N} | D_{ij} C_{z_j,v_j}^r(y_i, y_j) |^2 = J_{z,v}(y)
$$

(7)

Condition C2. We consider the following function,

$$
K_{z,v}(y) = \sum_{i_1, i_2=1}^{N} | C_{z_{i_1},v_{i_2}}^r(y_{i_1}, y_{i_2}) |^2
$$

(8)

By using the relation $y(t) = Qs(t)$ and exploiting the multilinearity of cumulant and independence of sources, we can write

$$
C_{z_{i_1},v_{i_2}}^r(y_{i_1}, y_{i_2}) = \sum_{\ell=1}^{N} Q_{i_1,\ell} Q_{i_2,\ell} C_{z_{\ell},v_{\ell}}(s_{\ell}, s_{\ell})
$$

(9)

Because $Q$ is an unitary matrix, the Eq.8 can be developed as,

$$
K_{z,v}(y) = \sum_{\ell_1, \ell_2=1}^{N} \left| \sum_{i=1}^{N} Q_{i,\ell_1} Q_{i,\ell_2} C_{z_{\ell_1},v_{\ell_2}}(s_{\ell_1}, s_{\ell_2}) \right|^2
$$

(10)

In other part, it is evident that $J_{z,v}(y) \leq K_{z,v}(y)$ and then the condition C2 is satisfied,

$$
J_{z,v}(y) \leq J_{z,v}(s)
$$

(11)

Condition C3. Let $Z$ and $V$ two real and positive matrices used for reference system such as,

$$
z(t) = Zu(t); v(t) = Vu(t)
$$

(12)

By using the Eq.9 with $i_1 = i_2$, the equality of Eq.11 can be expressed as,

$$
s_{\ell_1, \ell_2} = \sum_{i=1}^{N} Q_{i,\ell_1} Q_{i,\ell_2} | C_{z_{\ell_1},v_{\ell_2}}(s_{\ell_1}, s_{\ell_2})|^2 \delta_{i_1, i_2} = \sum_{\ell_1, \ell_2=1}^{N} | C_{z_{\ell_1},v_{\ell_2}}(s_{\ell_1}, s_{\ell_2})|^2 = 0
$$

(13)

where $\delta_{i_1, i_2}$ denotes the Kronecker symbol, and the matrix $T$ is defined by

$$
T_{\ell_1, \ell_2} = \begin{cases} 
Z_{\ell_1,\ell} V_{\ell,\ell}^{2p} & r = 4p + 2, \\
Z_{\ell_1,\ell} V_{\ell,\ell}^{2p-2} & r = 4p.
\end{cases}
$$

(14)

The cumulant $C_{z_{\ell_1},v_{\ell_2}}(s_{\ell_1}, s_{\ell_2})$ are reals because $r$ is pair. In addition, if the $r$-th cumulant order of sources takes the same sign, their product two by two is positive. The referenced term in Eq.14 is also positif. Then, the validity of Eq.13 is conditioned by

$$
\sum_{i=1}^{N} | Q_{i,\ell_1} Q_{i,\ell_2} |^2 - \delta_{i_1, i_2} = 0
$$

(15)

That means the global matrix $Q$ must be diagonal with possible permutation such as $Q = DP$. 

\textbf{B. Joint diagonalization criterion}

The objective of the joint diagonalization criterion (JDC) is to estimate an unitary matrix $H$ to reduce the global contribution of inter-terms (off-diagonal) relatively to auto-terms (diagonal) for a set $M$ of $N$ matrices $M_{r,\nu,i}$ [7]. It can be expressed as

$$D(H,M) = \sum_{i=1}^{N} |\text{diag}(HM(i)H^T)|^2$$  \hspace{1cm} (16)

Its application is important in the blind source separation context, where inter-terms correspond to cross-cumulants and auto-terms correspond to auto-cumulants. The maximization of independence is relied to the maximization of auto-cumulants relatively to cross-cumulants.

To express our contrast function Eq.6 as JDC, let $M_{r,\nu,i}$ a set of $N$ matrices $M_{r,\nu,i}$, $j = 1, \ldots, N$.

$$[M_{r,\nu,i}]_{1,ij} = C_{r,\nu,i}(u_{i1}, u_{i2})$$  \hspace{1cm} (17)

For any unitary matrix $H$ such that $y(t) = Hu(t)$, we have

$$C_{r,\nu,i}(y(t), y(t)) = \sum_{\ell_1,\ell_2=1}^{N} h_{\ell_1,\ell_2} \bar{h}_{\ell_1,\ell_2} C_{r,\nu,i}(u_{\ell_1}, u_{\ell_2})$$  \hspace{1cm} (18)

The contrast function can be developed as

$$J_{r,\nu}(y) = \sum_{j=1}^{N} \sum_{\ell_1,\ell_2=1}^{N} |C_{r,\nu,i}(y(t), y(t))|^2$$  \hspace{1cm} (19)

The Eq.19 demonstrates that our proposed contrast function can be formulated as joint diagonalization criterion. Then, the matrix $H$ can be performed by maximization of the JDC of $N$ referenced cumulant matrices. Finally, the separating matrix will be estimated as

$$B = HW$$  \hspace{1cm} (20)

We apply the joint diagonalization technique proposed in [7].

\textbf{III. COMPUTER SIMULATIONS}

\textbf{A. Choice of reference}

We will choose the reference vectors $z(t)$ and $v(t)$ to de-noise the cumulant matrices. Each observation $x_{i}(t)$ represents the linear mixture $x_{0,i}(t)$ perturbed by noise samples $n_{i}(t)$. We have

$$x_{i}(t) = m_{i}(t) + n_{i}(t);  i = 1, \ldots, M$$  \hspace{1cm} (21)

So, the element of cumulant matrix can be decomposed as,

$$C_{r,\nu,i}(x_{ij}, x_{ik}) = C_{r,\nu,i}(x_{0,ij}, x_{0,ik}) + C_{r,\nu,i}(n_{i}, n_{j}) + C_{r,\nu,i}(n_{i}, n_{k})$$  \hspace{1cm} (22)

Based on independence between sources and noises, we can write

$$C_{r,\nu,i}(x_{ij}, x_{ik}) = C_{r,\nu,i}(x_{0,ij}, x_{0,ik}) + C_{r,\nu,i}(n_{i1}, n_{i2})$$  \hspace{1cm} (23)

The first term corresponds to de-noised observations when the second represents the contribution of noises that can be cancelled for cross-cumulant because the independence between noise components. Then, to oppose the presence of noises we can choose

$$z_{j}(t) = x_{j}(t), v_{j}(t) = x_{j}(t); j \neq i$$  \hspace{1cm} (24)

The first term of Eq.23 can be developed as,

$$C_{r,\nu,i}(x_{ij}, x_{ik}) = C_{r,\nu,i}(x_{0,ij}, x_{0,ik}) + C_{r,\nu,i}(n_{i}, n_{j}) + C_{r,\nu,i}(n_{i}, n_{k})$$  \hspace{1cm} (25)

The second term of Eq.23 can be developed as,

$$C_{r,\nu,i}(n_{i1}, n_{i2}) = 0$$  \hspace{1cm} (26)

In consequence, the referenced cumulant takes the following expression

$$C_{r,\nu,i}(x_{ij}, x_{ik}) = C_{r,\nu,i}(x_{0,ij}, x_{0,ik})$$  \hspace{1cm} (27)

The Eq.27 demonstrates that the choice of reference, corresponding to Eq.24, performs the de-noised cumulant for any distribution of noise.

\textbf{B. Simulation Results}

Computer simulations are conducted to illustrate the performance of the proposed technique. The source signals are three linear chirps taken respectively in the frequency ranges $[10$ Hz, $300$ Hz], $[50$ Hz, $350$ Hz] and $[50$ Hz, $400$ Hz]. The mixture is obtained by a matrix randomly generated. To study the robustness of our DRCS algorithm, with respect to the noise effect, we corrupt the observed signals by an additive non-gaussian noise. The following index [10], applied on global matrix $Q = B.A$, measures the separation quality

$$P_{l} = \frac{1}{2(N-1)} \left[ \sum_{j=1}^{N} \frac{\sum_{i=1}^{N} Q_{ij}^2}{\max_{i,j}(|Q_{ij}|^2)} - 1 \right]$$  \hspace{1cm} (28)

For separation, the referenced vectors can be chosen, in general, as linear combination of components of whitened vector $u(t)$. Each referenced matrix, to be diagonalized, has $N$ auto-terms and $N(N-1)$ inter-terms. The auto-terms correspond to $C_{r,\nu,i}(y_{i}, y_{i})$ and the inter-terms correspond to $C_{r,\nu,i}(y_{i1}, y_{i2})$; $i \neq j$. In the first way, we will enrich the auto-terms and the components are taken identical

$$z(t) = u(t)$$  \hspace{1cm} (29)

In the second way, we will reduce the inter-terms by creating the diversity between components such as

$$z_{j}(t) = u_{j}(t), v_{j}(t) = u_{j}(t); j \neq i$$  \hspace{1cm} (30)

The equation Eq.30 corresponds also to the condition Eq.24 of minimization of noise effect. Because the most quantity of the inter-terms than the auto-terms, the second way will be more efficient than the proposed one. The Fig. 2 shows the
comparison, over the PI index, for the separation performance of three algorithms:

• **DRCS-1**: The reference vectors correspond to Eq.29

• **DRCS-2**: The reference vectors correspond to Eq.30: $z(t) = [u_1(t), u_2(t), u_3(t)]^T$ and $v(t) = [u_2(t), u_3(t), u_1(t)]^T$.

• **JADE** algorithm [7].

The performance index PI versus SNR is evaluated over 100 Monte-Carlo runs.

The simulation results confirm the advantage of the choice of different reference components. The **DRCS-2** performs the separation better than **DRCS-1** for all values of SNR, and JADE for $SNR \in [0dB, 20dB]$. This the important contribution and advantage of **DRCS-2** in noisy-environment. For high values of SNR ($\geq 20dB$), the JADE algorithm can improve the separation quality. The reason is the exploitation of maximum information delivered by $N^2$ cumulant matrices in contrast to our approach **DRCS** that requires only $N$ ones.

**IV. CONCLUSIONS**

In this paper, we have proposed a new separation contrast for independence sources. Our contrast works in the case of instantaneous linear mixture and exploits double referenced system. After observations whitening step, the optimization of the proposed contrast is executed by a well known joint-diagonalization procedure. The use of a different double reference vectors gives a large choice liberty. The performances illustrated by simulation results demonstrated the effectiveness of our contrast with a few number of matrices that other algorithms.

**REFERENCES**


