AN ANY ORDER GENERALIZATION OF JADE FOR COMPLEX SOURCE SIGNALS

Eric MOREAU

SIS, ISITV, Univ. Toulon et du Var
av. G. Pompidou, BP 56,
F-83162 La Valette du Var, Cedex, France
moreau@isitv.univ-tln.fr

ABSTRACT

In this paper, considering the complex case, we extend some results leading to the popular JADE algorithm to cumulants of any order greater than or equal to three. We first exhibit a new contrast function which constitutes a generalization for the underlying contrast of JADE which thus appears as a particular case. Then we generalize the link between this new contrast and a joint-diagonalization criterion of a set of matrices. Moreover, in the two sources case, we show that the generalized contrast can be written as a simple quadratic form whatever the cumulant order. Finally, some computer simulations illustrate the potential advantage one can take of considering statistics of different orders for the joint-diagonalization of cumulant matrices.

1. INTRODUCTION

We consider the blind source separation problem which has found numerous solutions in the past decade. Beginning with the originate works of Hérault and Jutten, see [1] and references therein, who have proposed an adaptive (on-line) algorithm, two of the most important contributions are provided by Comon [2] and Cardoso and Souloumiac [3]. These later solutions are block (off-line) algorithms which are both closely related to contrast functions (also simply called contrasts). Such contrasts were introduced and defined in [2] and have recently found a generalization in [4]. The algorithm presented in [2] is called ICA for "Independent Component Analysis" and the one presented in [3] is called JADE for "Joint Approximate Diagonalization of Eigen-matrices".

The JADE's underlying contrast take into consideration only fourth order cumulants on the contrary to the ICA's one which remains available whatever the order of cumulants is since it is greater than or equal to three. On the other hand, the fourth order JADE's contrast has also found interesting interpretations in terms of a joint-diagonalization criteria. It is a joint diagonalization criterion (maximized w.r.t. a unitary matrix) of some cumulant matrices sets. These links are the keys for the derivation of the practical JADE algorithms.

In this paper, we are mainly interested in generalizing the underlying contrast of JADE in the complex case and its link with a joint-diagonalization criterion involving cumulants of any order greater than or equal to three. As exemplified in [5] where only real signals are considered, one can find interest in being able to choose the cumulants order or in combining statistical information of different orders.

2. PROBLEM FORMULATION

An observed signal vector $\boldsymbol{x}[n]$ is assumed to follow the linear model

$$\boldsymbol{x}[n] = \boldsymbol{G}\boldsymbol{a}[n] \tag{1}$$

where $n \in \mathbb{Z}$ is the discrete time, $\boldsymbol{a}[n]$ the $(N,1)$ vector of $N \neq 2$ unobservable *complex* input signals $a_i[n]$, $i \in \{1,\ldots,N\}$, called sources, $\boldsymbol{x}[n]$ the $(N,1)$ vector of observed signals $x_i[n]$, $i \in \{1,\ldots,N\}$ and $\boldsymbol{G}$ the (N,N) square mixing matrix assumed *invertible*.

Further, the following assumptions are considered

A1. "*Independence*" The sources $a_i[n]$, $i \in \{1,\ldots,N\}$, are zero-mean, unit power and statistically mutually independent;

A2. "*Stationarity*" $a_i[n]$, $i \in \{1,\ldots,N\}$, are random signals stationary up to order under consideration, *i.e.* $\forall i \in \{1,\ldots,N\}$, the cumulant $\mathsf{Cum}\,[\underbrace{a_i^*[n],\ldots,a_i^*[n]}_{c \text{ terms}},\underbrace{a_i[n],\ldots,a_i[n]}_{R-c \text{ terms}}]$ is an independent function of n, denoted by $\mathsf{C}_R^c[a_i]$; moreover for a considered value of c, at most one of the cumulants $\mathsf{C}_R^c[a_i]$, $i \in \{1,\ldots,N\}$, is null.

It is important now to introduce the notion of white vectors. A vector $\boldsymbol{z}[n]$ of random signals is said to be (spatially) white if its covariance matrix $\boldsymbol{R}_z = \mathsf{E}[\boldsymbol{z}\boldsymbol{z}^\dagger]$ where $(\cdot)^\dagger$ stands for the conjugate and transpose operator, equals the identity. A classical first transformation is then defined as a whitening of the observation vector $\boldsymbol{x}[n]$. This is done by applying a whitening matrix $\boldsymbol{B}$ in such a way that $\boldsymbol{B}\boldsymbol{G} = \boldsymbol{V}$ where $\boldsymbol{V}$ is a unitary matrix, *i.e.* $\boldsymbol{V}\boldsymbol{V}^\dagger = \boldsymbol{I}$. Hence, after the whitening transformation, the new "observed" vector reads

$$\boldsymbol{x}_b[n] = \boldsymbol{B}\boldsymbol{x}[n] = \boldsymbol{V}\boldsymbol{a}[n] \ . \tag{2}$$

The blind source separation problem consists now in estimating a *unitary* matrix $\boldsymbol{H}$ in such a way that the vector

$$\boldsymbol{y}[n] = \boldsymbol{H}\boldsymbol{x}_b[n] \tag{3}$$

restores one of the different sources on each of its different components.

Because the sources are inobservable and the mixture is unknown, the exact power and order of each sources can not generally be recovered. It is the reason why the separation is said to be achieved when the global unitary matrix $\boldsymbol{S}$ defined as

$$\boldsymbol{S} = \boldsymbol{H}\boldsymbol{V} \tag{4}$$

can be written

$$\boldsymbol{S} = \boldsymbol{D}\boldsymbol{P} \tag{5}$$

where $\boldsymbol{D}$ is an invertible diagonal matrix (here with unit modulus components) corresponding to arbitrary phases for the restored sources and $\boldsymbol{P}$ a permutation matrix corresponding to an arbitrary order of restitution. According to (3), (1) and (4) the output vector can be written as

$$\boldsymbol{y}[n] = \boldsymbol{S}\boldsymbol{a}[n] \ . \tag{6}$$

Because of the stationarity assumption, the explicit dependence of sources, observations and outputs vectors with the discrete time n will be now omitted whenever no confusion is possible.

Let us define some notations which will be useful in the following. Let $\mathcal{A}$ be the set of random vectors satisfying assumptions `A1` and `A2`. Let $\mathcal{U}$ be the set of unitary matrices. The subset of $\mathcal{U}$ of matrices $\boldsymbol{S}$ of the form (5) is denoted by $\mathcal{P}$ and the subset of $\mathcal{P}$ of diagonal matrices is denoted by $\mathcal{D}$. Finally the set of random vector $\boldsymbol{y}[n]$ built from (6) where $\boldsymbol{a}[n] \in \mathcal{A}$ and $\boldsymbol{S} \in \mathcal{U}$ is denoted by $\mathcal{Y}_u$.

3. CONTRAST FUNCTIONS

3.1. Recalls

Contrast functions correspond to objective functions for the source separation problem. They depend generally on the outputs of the separating system and they have to be maximized to get a separating solution. For convenience, let us recall the definition given in [4]:

Definition 1 *A contrast on $\mathcal{Y}_u$ is a multivariate mapping $\mathcal{I}(\cdot)$ from $\mathcal{Y}_u$ to the real set which satisfies the following three requirements:*
R1. $\forall \boldsymbol{y} \in \mathcal{Y}_u$, $\forall \boldsymbol{D} \in \mathcal{D}$, $\mathcal{I}(\boldsymbol{D}\boldsymbol{y}) = \mathcal{I}(\boldsymbol{y})$;
R2. $\forall \boldsymbol{a} \in \mathcal{A}$, $\forall \boldsymbol{S} \in \mathcal{U}$, $\mathcal{I}(\boldsymbol{S}\boldsymbol{a}) \leq \mathcal{I}(\boldsymbol{a})$;
R3. $\forall \boldsymbol{a} \in \mathcal{A}$, $\forall \boldsymbol{S} \in \mathcal{U}$, $\mathcal{I}(\boldsymbol{S}\boldsymbol{a}) = \mathcal{I}(\boldsymbol{a}) \Rightarrow \boldsymbol{S} \in \mathcal{P}$.

One of the first contrast can be found in [2]. Other examples of contrasts are given in [4]. On the other hand a contrast involving both cross-cumulants and auto-cumulants has been proposed in [3]. It reads

$$\mathcal{J}(\boldsymbol{y}) = \sum_{i,i_1,i_2=1}^{N} \left|\mathsf{Cum}[y_i, y_i^*, y_{i_1}^*, y_{i_2}]\right|^2 \tag{7}$$

and is called the JADE's contrast.

3.2. A generalized contrast

Here we propose a generalization of the above contrast to cumulants of any order greater than or equal to three. This is given according to the following proposition

Proposition 1 *Let R and c two integers such that $R \geq 3$ and $0 \leq c \leq R-2$, using the notation*

$$\mathsf{C}_{R,c}[\boldsymbol{y}, i, j] = \mathsf{Cum}[y_i, y_j^*, \underbrace{y_{i_1}^*, \ldots, y_{i_c}^*}_{c \text{ terms}}, \underbrace{y_{i_{c+1}}, \ldots, y_{i_{R-2}}}_{R-c-2 \text{ terms}}] \tag{8}$$

the function

$$\mathcal{J}_{R,c}(\boldsymbol{y}) = \sum_{i,i_1,\ldots,i_{R-2}=1}^{N} \left|\mathsf{C}_{R,c}[\boldsymbol{y}, i, i]\right|^2 \tag{9}$$

is a contrast on $\mathcal{Y}_u$, i.e. for white vectors $\boldsymbol{y}$.

Proof. With $\boldsymbol{S} = (S_{i,j})$, according to (6), the multi-linearity of cumulants and the independence of sources, we have

$$\begin{aligned} \mathsf{C}_{R,c}[\boldsymbol{y}, i, i] &= \sum_{\ell} |S_{i,\ell}|^2 S_{i_1,\ell}^* \cdots S_{i_c,\ell}^* \\ &\quad S_{i_{c+1},\ell} \cdots S_{i_{R-2},\ell} \ \mathsf{C}_R^{c+1}[a_\ell] \ . \end{aligned}$$

Now because $\boldsymbol{S}$ is a unitary matrix then $\forall \ell_1, \ell_2$, $\sum_m S_{m,\ell_1} S_{m,\ell_2}^* = \delta_{\ell_1,\ell_2}$ where $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise. Thus

$$\mathcal{J}_{R,c}(\boldsymbol{y}) = \sum_{\ell} \left(\sum_i |S_{i,\ell}|^4 \right) \left|\mathsf{C}_R^{c+1}[a_\ell]\right|^2 \tag{10}$$

and because $\forall \ell_1$, $\sum_{i_1} |S_{i_1,\ell_1}|^4 \leq \sum_{i_1} |S_{i_1,\ell_1}|^2 = 1$, then we have

$$\mathcal{J}_{R,c}(\boldsymbol{y}) \leq \sum_{\ell} \left|\mathsf{C}_R^{c+1}[a_\ell]\right|^2 = \mathcal{J}_{R,c}(\boldsymbol{a}) \ . \tag{11}$$

It is easily shown that the equality in (11) holds if and only if $\boldsymbol{S}$ satisfies (5). Moreover $\mathcal{J}_{R,c}(\boldsymbol{y})$ is invariant when the $y_i(n)$, $i = 1, \ldots, N$, are multiplied by a factor of modulus 1. Thus $\mathcal{J}_{R,c}(\boldsymbol{y})$ is a contrast function. $\diamondsuit$

Hence for $R = 4$ and $c = 1$ we have $\mathcal{J}_{4,1}(\boldsymbol{y}) = \mathcal{J}(\boldsymbol{y})$ which is the JADE's contrast. All other values of R and c lead to new contrasts.

3.3. Link with a Joint-diagonalization criterion

In this section the contrast $\mathcal{J}_{R,c}(\boldsymbol{y})$ is linked to a joint-diagonalization criterion of a set of matrices. Such a joint-diagonalization criterion is defined according to

Definition 2 *Considering a set of M square matrices $\boldsymbol{M}(m)$, $m = 1, \ldots, M$ denoted by $\mathcal{M}$. A joint-diagonalizer of this set is a unitary matrix that maximizes the function*

$$\mathcal{D}(\boldsymbol{H}, \mathcal{M}) = \sum_{m=1}^{M} \left(\sum_i \left|M_{i,i}^L(m)\right|^2 \right) \tag{12}$$

where

$$M_{i,i}^L(m) = \sum_{n_1,n_2} H_{i,n_1} H_{i,n_2}^* M_{n_1,n_2}(m) \ . \tag{13}$$

Now the equivalence can be stated according to the following proposition:

Proposition 2 *With $R \geq 3$, let $\mathcal{C}_{R,c}$ be the set of $M = N^{R-2}$ matrices*

$$\boldsymbol{M}(i_1, \ldots, i_{R-2}) = (M_{i,j}(i_1, \ldots, i_{R-2}))$$

defined as

$$M_{i,j}(i_1, \ldots, i_{R-2}) = \mathsf{C}_{R,c}[\boldsymbol{x}, i, j] \ . \tag{14}$$

Then, if $\boldsymbol{H}$ is a unitary matrix, we have

$$\mathcal{D}(\boldsymbol{H}, \mathcal{C}_{R,c}) = \mathcal{J}_{R,c}(\boldsymbol{H}\boldsymbol{x}) \ . \tag{15}$$

Let us notice that this proposition 2 is a generalization of one result in [3] to cumulants of any order greater than or equal to three.

According to the above proposition, we can now choose the order of cumulants (greater than or equal to three) for the joint-diagonalization of matrices. In particular third order cumulants can be used leading to the joint-diagonalization of N matrices. However even if it is sufficient to joint-diagonalize matrices of cumulants of a given order, one can find interest in combining cumulants of different orders. In particular this can lead to algorithms that are more robust w.r.t. the statistics of sources. For example one can combine third and fourth order cumulants. If third order (resp. fourth order) cumulants of the unknown sources vanish then the other fourth order (resp. third order) ones can be directly used. In the unfortunate case where both third and fourth order cumulants of the sources vanish, then one has to consider cumulants of greater order. Moreover such combination can be useful for an independent component analysis goal when one is not sure that the available data conform the initial model. Indeed in such a case cross cumulants of all orders have to be canceled.

Now we show that cumulants of different orders can be considered altogether. This is given according to the following proposition:

Proposition 3 *Let $\gamma_1, \ldots, \gamma_m$ be $m \in \mathbb{N}^*$ real non negative constants with at least one non zero. Let $S_1, \ldots, S_m$ be m integers such that $3 \leq S_1 < \cdots < S_m$ and let $c_1, \ldots, c_m$ be m integers such that $0 \leq c_i \leq S_i - 2$, $\forall i$. Finally, let*

$$\sqrt{\gamma_i}\ \mathcal{C}_{S_i,c_i} = \{\sqrt{\gamma_i}\ \boldsymbol{M}(i_1, \ldots, i_{S_i-2})\}$$

be m sets of matrices $\boldsymbol{M}(\cdot)$ of S_i order cumulants as defined in proposition 2. Then, if $\boldsymbol{H}$ is a unitary matrix, we have

$$\mathcal{D}(\boldsymbol{H}, \bigcup_{i=1}^{m} \sqrt{\gamma_i}\ \mathcal{C}_{S_i,c_i}) = \sum_{i=1}^{m} \gamma_i\ \mathcal{J}_{S_i,c_i}(\boldsymbol{H}\boldsymbol{x}) \ . \tag{16}$$

Now since it is well-known that a (non zero) non negative linear combination of contrasts is also a contrast then the joint-diagonalization of matrices of cumulants of mixed orders is again a sufficient condition for separation.

4. ABOUT THE ALGORITHM

The JADE algorithm is based on Jacobi optimization. This means that the maximization of the criterion under consideration is realized through a sequence of plane (or Givens) rotations as initiated in [2]. Each plane rotation works on a pair of the output vector $\boldsymbol{y}[n]$ and one "sweep" or iteration consists in processing the outputs through all the $N(N-1)/2$ possible pairs. Hence the N-dimensional problem is reduced to $N(N-1)/2$ problems of dimension 2. One of the main advantages is that the 2-dimensional problem is simpler and often admits an analytical solution. Thus let us now consider the only 2-dimensional problem where a plane rotation has to be determined. In the following, we parameterize it as

$$\boldsymbol{H} = \begin{pmatrix} \cos\theta & e^{\imath\phi}\sin\theta \\ -e^{-\imath\phi}\sin\theta & \cos\theta \end{pmatrix} . \tag{17}$$

For $N = 2$, it can be observed that $\mathcal{J}_{R,c}(\boldsymbol{y})$ may be written as the quadratic form

$$\mathcal{J}_{R,c}(\boldsymbol{y}) = \boldsymbol{u}_{\theta,\phi}^T \ \boldsymbol{A}_{R,c} \ \boldsymbol{u}_{\theta,\phi} \tag{18}$$

where $\boldsymbol{u}_{\theta,\phi}^T = (\cos(2\theta) \quad \sin(2\theta)\cos(\phi) \quad \sin(2\theta)\sin(\phi))$ and where $\boldsymbol{A}_{R,c} = (A_{R,c,i,j})$ is a $(3,3)$ hermitian matrix. The expression of $\boldsymbol{A}_{R,c}$ is not given here because of limitation of space.

Hence the values of θ and ϕ maximizing $\mathcal{J}_{R,c}(\boldsymbol{y})$ can be now easily found by computing the normalized (unit-norm) eigenvector of $\boldsymbol{A}_{R,c}$ associated with the largest eigenvalue exactly as in the original JADE algorithm.

It is important to notice that $\mathcal{J}_{R,c}(\boldsymbol{y})$ follows exactly the same quadratic form (18) whatever the cumulant order, in the case of two sources. This is a remarkable fact which simplify a lot the use of cumulants of different orders in the algorithm. Indeed, let us consider for example matrices of cumulants of orders $S_1, \ldots, S_m$, $m \in \mathbb{N}^*$ such that $3 \leq S_1 < \cdots < S_m$. Then according to (18)

$$\sum_{i=1}^{m} \gamma_i \mathcal{J}_{S_i,c_i}(\boldsymbol{y}) = \boldsymbol{u}_{\theta,\phi}^T \ \sum_{i=1}^{m} \gamma_i \boldsymbol{A}_{S_i,c_i} \ \boldsymbol{u}_{\theta,\phi}$$

where the γ_i's are real non-negative constants with at least one non zero.

Such a generalized algorithm for joint-diagonalization of cumulant matrices is called eJADE for "extended JADE" when considering the original implementation of JADE and adding directly the new matrices to be joint-diagonalized.

An example of computer simulations is given in Fig.1. We use both only third order cumulant matrices whose corresponding algorithm is denoted eJADE(3), and third plus fourth order cumulant matrices whose corresponding algorithm is denoted eJADE(3,4). We use a signal with parameterized third and fourth-order cumulants. It is a discrete i.i.d. signal called MS(α) which takes its values in the set $\{-1 \ , \ 0 \ , \ \alpha\}$ with the respective probability $\{\frac{1}{1+\alpha} \ , \ \frac{\alpha-1}{\alpha} \ , \ \frac{1}{\alpha(1+\alpha)}\}$. The real parameter α called "cumulant parameter" is such that $\alpha \geq 1$. Hence for a MS(α) signal $a(n)$, one easily has $\mathsf{E}[a] = 0$, $\mathsf{E}[a^2] = 1$, $\mathsf{C}_3[a] = \alpha - 1$ and $\mathsf{C}_4[a] = \alpha^2 - \alpha - 2$. The performances of the algorithms are associated to a non negative index/measure of performance [5] which is zero when the separation holds. We have plotted both the mean and the standard deviation (STD) of the estimated index (over 500 Monte Carlo runs) as a function of the cumulant parameter α. With $N = 3$, the first two considered sources are MS(α) signals while the third one is a Gaussian i.i.d. signal and a Gaussian additive noise is added to the mixture. The mixing matrix is chosen randomly with an uniform law in $[-1, 1]$. The data number is held constant to $N_d = 400$. The figure shows that the performances of eJADE(3,4) in comparison to JADE and eJADE(3) are less subject to variation w.r.t. the statistics of the sources.

REFERENCES

[1] C. Jutten and J. Herault, "Blind separation of sources, Part I: An adaptative algorithm based on neuromimetic architecture", *Signal Processing*, Vol. 24, pp 1-10, 1991.

[2] P. Comon, "Independent component analysis, a new concept?", *Signal Processing*, Vol. 36, pp 287-314, 1994.

[3] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non gaussian signals", *IEE Proceedings-F*, Vol. 40, pp 362-370, 1993.

[4] E. Moreau and N. Thirion-Moreau, "Nonsymmetrical contrasts for source separation", *IEEE Trans. Signal Processing*, Vol. 47, No. 8, pp 2241-2252, August 1999.

[5] E. Moreau, "A generalization of joint-diagonalization criteria for source separation", *IEEE Trans. Signal Processing*, accepted for publication, October 2000.

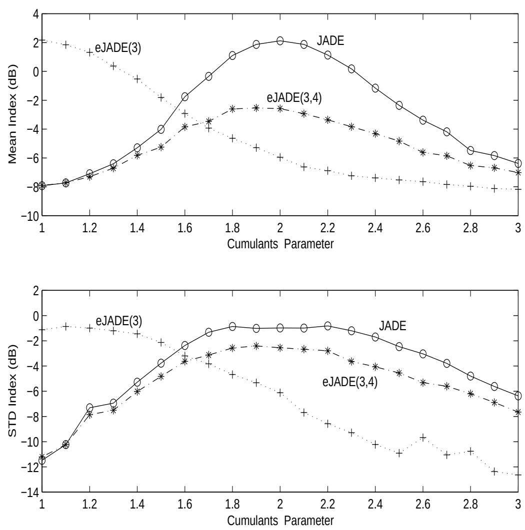

Figure 1: Mean and STD of the estimated index of JADE, eJADE(3) and eJADE(3,4) w.r.t. the cumulant parameter α. The number of data $N_d = 400$ and the noise power $P_n = 0.02$.