Comments On Blind Beamforming for Multiple Non-Gaussian Signals and the Constant-Modulus Algorithm

Eric Moreau and Pierre Comon

Abstract— **In a recent paper [1], Sheinvald compared a cumulant matching criterion to three simplified criteria, that he claimed to be equivalent. We show that two simplified criteria can be viewed as direct consequences of a result already published, and that their equivalence must be revised. We also give additional references of closely related works that have been omitted, some of them presenting Jacobi-like source separation algorithms, presented as original in [1].**

I. Introduction

In [1], a method for blindly separating multiple cochannel non Gaussian signals received by a sensor array has been presented. The optimization criteria are very similar to those independently introduced in [4][6][7]. When comparing the various cost criteria, Sheinvald assumes that the second order whitening is perfect, and that noise is absent. We shall subsequently make the same assumptions, so that the mixing matrix $\boldsymbol{A}$ can be assumed to be unitary, as well as the separating matrix $\boldsymbol{W}$ and the matrix of the global system $\boldsymbol{U} = \boldsymbol{W}\,\boldsymbol{A}$.

More precisely, denote $\boldsymbol{s} = (s_1, \ldots, s_q)^T$ the vector of the q zero-mean source signals, $\boldsymbol{x} = (x_1, \ldots, x_q)^T$ the vector of the q observations, and $\boldsymbol{y} = (y_1, \ldots, y_q)^T$ the vector of the q estimated sources (although the estimated sources are denoted by $\hat{s}_i$ in [1], they are denoted by y_i in this paper, $1 \le i \le q$, in order to avoid confusion with sample statistics). The relation between the output $\boldsymbol{y}$ and the source $\boldsymbol{s}$ is given by

$$\boldsymbol{y} = \boldsymbol{U}\boldsymbol{s} \tag{1}$$

while the relation between the observation $\boldsymbol{x}$ and $\boldsymbol{y}$ is given by

$$\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x}\ . \tag{2}$$

Next, given a random vector $\boldsymbol{z}$ of size q, denote the fourth-order cumulant tensor $\kappa_{\mathbf{z},\,ijk\ell} = \mathsf{Cum}[z_i, z_j, z_k^*, z_\ell^*]$, and $\hat{\kappa}_{\mathbf{z},\,ijk\ell}$ a consistent estimate of it. For the sake of convenience, the exact and estimated marginal cumulants are denoted by κ_{z_i} and $\hat{\kappa}_{z_i}$, respectively, instead of $\kappa_{\mathbf{z},\,iiii}$ and $\hat{\kappa}_{\mathbf{z},\,iiii}$.

After a first whitening transformation, Sheinvald uses a cumulant matching quadratic cost criterion that can be written as

$$\widehat{\mathcal{C}}(\boldsymbol{y}) = \sum_{i,j,k,\ell} \left|\hat{\kappa}_{\mathbf{y},ijk\ell} - \kappa_{\mathbf{s},ijk\ell}\right|^2 . \tag{3}$$

We say that two optimization criteria $\mathcal{G}_1(\cdot)$ and $\mathcal{G}_2(\cdot)$ are equivalent, which will be denoted as $\mathcal{G}_1 \equiv \mathcal{G}_2$, if there exist two real numbers a and b such that $\mathcal{G}_2(\cdot) = a\,\mathcal{G}_1(\cdot) + b$, $a > 0$.

Sheinvald claims (in his Lemma 1) that, under certain hypotheses, $\widehat{\mathcal{C}}(\boldsymbol{y})$ is equivalent to three simpler functions, that we shall call $\widehat{\mathcal{I}}_p(\boldsymbol{y})$, $p \in \{1, 2, 3\}$, to fix the ideas. Their expressions are given by:

$$\widehat{\mathcal{I}}_1(\boldsymbol{y}) = -\sum_{i=1}^{q} \kappa_{s_i}\hat{\kappa}_{y_i} \tag{4}$$

$$\widehat{\mathcal{I}}_2(\boldsymbol{y}) = -\mathrm{sign}(\kappa_s)\sum_{i=1}^{q} \hat{\kappa}_{y_i} \tag{5}$$

$$\widehat{\mathcal{I}}_3(\boldsymbol{y}) = -\sum_{i=1}^{q} \hat{\kappa}_{y_i}^2 \tag{6}$$

Expression (5) requires all the sources to have an identical kurtosis, *i.e.* $\forall i$, $\kappa_{s_i} = \kappa_s$, whereas expression (4) requires κ_{s_i} to be all known.

In this correspondence we revisit the above results and their proof. We also point out additional references of very closely related works in which functions $\mathcal{I}_2(\boldsymbol{y}) = -\mathrm{sign}(\kappa_s)\sum_i \kappa_{y_i}$ and $\mathcal{I}_3(\boldsymbol{y}) = -\sum_i \kappa_{y_i}^2$ had already been proposed, and in which Jacobi-like algorithms had already been devised.

II. Exact cumulants

In this section, because exact cumulants are considered, the functions in (3), (4), (5) and (6) are respectively written $\mathcal{C}(\boldsymbol{y})$, $\mathcal{I}_1(\boldsymbol{y})$, $\mathcal{I}_2(\boldsymbol{y})$ and $\mathcal{I}_3(\boldsymbol{y})$. We assume that sources are independent (at order four), that is:

$$\kappa_{\mathbf{s},\,ijk\ell} = \kappa_{s_i}\delta[i,j,k,\ell] \tag{7}$$

where $\delta[i,j,k,\ell] = 1$ if $i = j = k = \ell$, and 0 otherwise. First define the square of the Euclidean norm of the output cumulant tensor as:

$$\mathcal{T}(\boldsymbol{y}) = \sum_{i,j,k,\ell} \left|\kappa_{\mathbf{y},ijk\ell}\right|^2 . \tag{8}$$

It has been shown in [3][4] that if $\boldsymbol{V}$ is an unitary matrix then for any $\boldsymbol{y}$ we have $\mathcal{T}(\boldsymbol{V}\boldsymbol{y}) = \mathcal{T}(\boldsymbol{y})$, which means that function $\mathcal{T}(\cdot)$ is invariant under unitary transforms of its vector argument. This property is only due to the multilinearity of cumulants and unitarity of $\boldsymbol{V}$. Now developing the square in $\mathcal{C}(\boldsymbol{y}) = ||\kappa_{\mathbf{y}} - \kappa_{\mathbf{s}}||^2$ and using the independence of sources leads immediately to:

$$\mathcal{C}(\boldsymbol{y}) = \mathcal{T}(\boldsymbol{y}) - \mathcal{I}_3(\boldsymbol{s}) + 2\,\mathcal{I}_1(\boldsymbol{y})\ . \tag{9}$$

Recall that in this section, only exact cumulants are considered. As a consequence, if source cumulants are known, $\mathcal{T}(\boldsymbol{y}) = \mathcal{T}(\boldsymbol{s}) = -\mathcal{I}_3(\boldsymbol{s})$ is constant and obviously $\mathcal{C} \equiv \mathcal{I}_1$. On the other hand if they are unknown, from (9), one can write the cost criterion as:

$$\mathcal{C}(\boldsymbol{y}) = \mathcal{I}_3(\boldsymbol{y}) - \mathcal{I}_3(\boldsymbol{s}) + \sum_{i=1}^{q}\left(\kappa_{y_i} - \kappa_{s_i}\right)^2 . \tag{10}$$

Submitted as a correspondence on February 1999; accepted September, 1999. E.Moreau is with MS-GESSY, ISITV, av. G. Pompidou, BP56, F-83162 La Valette du Var Cedex, France, moreau@isitv.univ-tln.fr; P.Comon is with Lab. I3S, Algorithmes/Euclide-B, 2000 route des Lucioles, Sophia-Antipolis, F-06410 Biot, France, comon@i3s.unice.fr.

If $\mathcal{I}_3(\boldsymbol{y}) - \mathcal{I}_3(\boldsymbol{s})$ were constant, this would show that $\mathcal{C}(\boldsymbol{y})$ is a convex function of κ_{s_i}, and that it reaches its minimum when $\kappa_{s_i} = \kappa_{y_i}$. But it is not the case since $\mathcal{I}_3(\boldsymbol{y}) - \mathcal{I}_3(\boldsymbol{s})$ directly depends on source kurtoses κ_{s_i}.

To illustrate this fact it is sufficient to consider two sources with different kurtoses and the separating solution $y_1 = s_2$ and $y_2 = s_1$. Indeed in this case $\mathcal{I}_3(\boldsymbol{y}) - \mathcal{I}_3(\boldsymbol{s})$ is minimized (and is zero) while $\sum_i (\kappa_{y_i} - \kappa_{s_i})^2$ is not. Thus $\mathcal{C} \not\equiv \mathcal{I}_3$.

The properties of $\mathcal{I}_3(\boldsymbol{y})$ have been already given in works related to contrast functions [3][4], that have not been referred to in [1]. Let us recall that a contrast is an objective function for the blind separation of source signals in the sense that its global maximization solves the problem.

More precisely, it has been proved that the only unitary matrices $\boldsymbol{U}$ that minimize $\mathcal{I}_3(\boldsymbol{y}) = -\sum_i \kappa_{y_i}^2$, which one can also write explicitly as $\sum_i [\sum_p |U_{ip}|^4 \kappa_{s_p}]^2$, are of the form $\boldsymbol{U} = \boldsymbol{\Lambda P}$, where $\boldsymbol{\Lambda}$ is diagonal with unit modulus entries and $\boldsymbol{P}$ is a permutation (see [4] for the real case, and [3] or the appendix of [4] for the complex case).

III. Sample cumulants

In the previous section, for the sake of simplicity, the relevant properties have been derived assuming the cumulants were exact. Now for consistency with [1], we show how these properties hold true if output cumulants are replaced with sample ones, and point out awkwardnesses in [1].

In practice, the independence of sources is expressed by:

$$\hat{\kappa}_{\mathbf{s},\,ijk\ell} = \hat{\kappa}_{s_i} \delta[i,j,k,\ell], \tag{11}$$

that is, the sample source cumulant tensor is diagonal. Similarly to the previous section, one defines the Froebenius norm of the estimated cumulant tensor as $\widehat{\mathcal{T}}(\boldsymbol{y}) = \sum_{i,j,k,\ell} |\hat{\kappa}_{\mathbf{y},ijk\ell}|^2$. The invariance of $\widehat{\mathcal{T}}(\boldsymbol{V y})$ over the group of unitary matrices holds true, because it is only due to multilinearity, which is still valid for estimated cumulants.

Let us first examine differently but following the same spirit, the derivations in [1]. Hence developing the square of $\widehat{\mathcal{C}}(\boldsymbol{y})$ in (3) yields:

$$\widehat{\mathcal{C}}(\boldsymbol{y}) = \widehat{\mathcal{T}}(\boldsymbol{y}) - \mathcal{I}_3(\boldsymbol{s}) + 2\,\widehat{\mathcal{I}}_1(\boldsymbol{y}) \ , \tag{12}$$

or equivalently using (2) and the invariance property of $\widehat{\mathcal{T}}(\boldsymbol{y})$, $\widehat{\mathcal{C}}(\boldsymbol{y}) = \widehat{\mathcal{T}}(\boldsymbol{x}) - \mathcal{I}_3(\boldsymbol{s}) + 2\,\widehat{\mathcal{I}}_1(\boldsymbol{y})$. Notice that $\widehat{\mathcal{T}}(\boldsymbol{x})$ does not depend neither on $\boldsymbol{U}$ nor on κ_{s_i} reminding that $\boldsymbol{x}$ is now given. Thus it can be considered as constant. If source cumulants κ_{s_i} are known, $\mathcal{I}_3(\boldsymbol{s})$ is constant, and $\widehat{\mathcal{C}} \equiv \widehat{\mathcal{I}}_1$; this is equation (26) of [1] which is correct.

Now assume that source cumulants κ_{s_i} are unknown. The cost criterion can, alternatively, be expressed as:

$$\widehat{\mathcal{C}}(\boldsymbol{y}) = \widehat{\mathcal{T}}(\boldsymbol{x}) + \widehat{\mathcal{I}}_3(\boldsymbol{y}) + \sum_{i=1}^{q} (\hat{\kappa}_{y_i} - \kappa_{s_i})^2 \tag{13}$$

where $\widehat{\mathcal{I}}_3(\boldsymbol{y})$ depends on $\boldsymbol{U}$ but not on κ_{s_i}. This shows that this time $\kappa_{s_i} = \hat{\kappa}_{y_i}$ is the minimum of $\widehat{\mathcal{C}}(\boldsymbol{y})$ with respect to κ_{s_i}, when $\boldsymbol{U}$ is fixed. Plugging back into (12) yields

$$\min_{\kappa_{s_i}} \widehat{\mathcal{C}} \equiv \widehat{\mathcal{I}}_3 \tag{14}$$

which is written as

$$\widehat{\mathcal{C}} \equiv \widehat{\mathcal{I}}_3 \ ; \tag{15}$$

in [1] (cf. result (28) of [1]). Lastly, if source cumulants are unknown but equal to each other, we have from (12): $\widehat{\mathcal{C}}(\boldsymbol{y}) \equiv q\,\kappa_s^2 - 2\,\kappa_s \sum_i \hat{\kappa}_{y_i}$. The minimization of this quantity with respect to κ_s leads to $\kappa_s = \frac{1}{q}\sum_i \hat{\kappa}_{y_i}$. Plugging back into the cost criterion gives $\widehat{\mathcal{C}}(\boldsymbol{y}) \equiv -\kappa_s \sum_i \hat{\kappa}_{y_i}$. So this yields $\widehat{\mathcal{C}} \equiv \widehat{\mathcal{I}}_2$. In [1], the proof of this result was restricted to the case where κ_s was known, which was unnecessary.

But it turns out that the above result (15) is not correct, because the fact that $\widehat{\mathcal{C}}(\boldsymbol{y}) \equiv \mathcal{I}_3(\boldsymbol{y})$ for fixed $\boldsymbol{U}$ does not mean that $\mathrm{Min}_U\, \widehat{\mathcal{C}}(\boldsymbol{y}) = \mathrm{Min}_U\, \mathcal{I}_3(\boldsymbol{y})$, contrary to what was claimed in [1]. In fact, this would require the following property:

$$\underset{\boldsymbol{u}}{\mathrm{Min}}\, f(\boldsymbol{z},\boldsymbol{u}) = \underset{\boldsymbol{u}}{\mathrm{Min}}\, \underset{\boldsymbol{z}}{\mathrm{Min}}\, f(\boldsymbol{z},\boldsymbol{u}) \tag{16}$$

that is unfortunately false when $\boldsymbol{z}$ is itself a function of $\boldsymbol{u}$. Yet in the present problem, $\hat{\kappa}_{s_p} = \sum_i |U_{ip}|^4 \hat{\kappa}_{y_i}$ play the role of z_p and $\boldsymbol{u}$ the role of $\boldsymbol{U}$.

The correct way to address the problem is to minimize $\widehat{\mathcal{I}}_3(\boldsymbol{y})$ with respect to $\boldsymbol{U}$. Along the same lines as in [4][8][9] , and thanks to assumption (11), one finds that $\boldsymbol{U}$ can be determined only up to a multiplicative factor of the form $\boldsymbol{\Lambda P}$, which makes sense. On the other hand, the properties presented in [1] would yield that this inherent indeterminacy disappears, which is necessarily false, as in the case of exact cumulants.

IV. Numerical algorithm

Lastly, a Jacobi-like algorithm has been proposed in [1]. Yet, the algorithm described in [4] also performs a Jacobi-like procedure for maximizing the contrast through the product of Givens (plane) rotations; this technique had been first proposed in 1989 for source separation [2], but this was ignored in [1].

Later, in [7] (cf. [8] for the complex case) it was shown that function $\sum_{i=1}^{q} |\kappa_{y_i}|$ is also a contrast. It is also shown therein that in the specific case where the source kurtosis all have the same sign, *i.e.* $\forall i$, $\mathrm{sign}(\kappa_i) = \varepsilon$, then the above contrast can be written as $\mathcal{J}(\boldsymbol{y}) = \varepsilon \sum_{i=1}^{q} \kappa_{y_i}$. Thus the function $-\mathcal{I}_2(\boldsymbol{y})$ in (5) is equal to $\mathcal{J}(\boldsymbol{y})$. Now it is also pointed out in [7] that if the sources are moreover real[1] then immediately $\mathcal{J}(\boldsymbol{y}) = \varepsilon \sum_{i=1}^{q} \mathsf{E}\{|y_i|^4\} + C$ where $C = 3q\varepsilon$ is a constant term. Hence its relation with a "2-2" constant modulus cost function is a direct consequence of the "orthonormalization" first transformation.

Finally the contrasts [3]-[9] have been derived for cumulants of any order larger than or equal to 3, under the assumption that the corresponding source cumulants of the

[1] The case of complex circular (or proper) sources, *i.e.* $\forall i$, $\mathsf{E}\{s_i^2\} = 0$, considered in [1] follows directly the same derivation.

same order are non zero. These results can be extended to contrasts of the form $\sum_i f(|\kappa_{y_i}|)$, where f is convex [9]; in particular, they apply to $\sum_i |\kappa_{y_i}|^\alpha$, for any real exponent $\alpha \geq 1$, which covers the cases $\alpha = 1, 2$ studied in [1]. It was also shown that a class of complex signals, more general than circular (or proper) signals, can be used in these contrasts. We must eventually insist that a contrast maximization algorithm has been given in [5] for the optimization of $\mathcal{J}(\boldsymbol{y})$, with a polynomial complexity (that is, within a finite number of arithmetic operations).

The above remarks show that the algorithm proposed in [1] is very similar to the contrast maximization algorithms already independently proposed in [4][5].

References

[1] J. Sheinvald, "On blind beamforming for multiple non-Gaussian signals and the constant-modulus algorithm", *IEEE Trans. Signal Processing*, Vol. 46, No. 7, pp 1878-1885, July 1998.

[2] P. Comon, "Separation of Stochastic Processes", *Proc. Workshop on Higher-Order Spectral Analysis*, Vail, Colorado, June 28-30, 1989, pp 174-179.

[3] P. Comon, "Analyse en composantes indépendantes et identification aveugle", *Traitement du Signal*, Vol. 7, No. 5, pp 435-450, Dec. 1990.

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

[5] P. Comon and E. Moreau, "Improved Contrast Dedicated to Blind Separation in Communications", in Proc. *ICASSP'97, IEEE International Conference on Acoustics, Speech and Signal Processing*, Munich, Germany, Vol. 5, pp 3453-3456, April 1997.

[6] J.F. Cardoso, "Blind beamforming for non Gaussian signals", *IEE Proceedings series F*, vol.140, no.6, pp 362-370, December 1993.

[7] E. Moreau and O. Macchi, "High order contrasts for self-adaptive source separation", *International Journal of Adaptive Control and Signal Processing*, Vol. 10, No. 1, pp 19-46, January 1996.

[8] E. Moreau, "Criteria for complex sources separation", in Proc. *EUSIPCO'96, European Signal Processing Conference*, Trieste, Italy, Vol. II, pp 931-934, September 1996.

[9] E. Moreau and J.-C. Pesquet, "Generelized contrasts for multichannel blind deconvolution of linear systems", *IEEE Signal Processing Letters*, vol. 4, no. 6, pp 182-183, June 1997.