# PERFORMANCE OF CONTRAST-BASED BLIND SOURCE SEPARATION

*Pierre Comon*[1,3], *Pascal Chevalier*[2] *and Véronique Capdevielle*[2]

[1]Eurecom, B.P.193, F-06904 Sophia-Antipolis Cedex, France
[2]Thomson CSF Communications, B.P.156, F-92231 Gennevilliers Cedex, France

## ABSTRACT

A source separator is proposed based on a contrast functional recently introduced. This contrast is not squaring the sample cumulants, which allows a reduced variance on the estimated separating matrix. In fact, it is shown by extensive computer experiments how this contrast compares with JADE, and with the contrast squaring sample fourth-order cumulants. Then a procedure is outlined that allows to compute the asymptotic theoretical performances.

## 1. INTRODUCTION AND NOTATION

**Notation.** The following narrow band model, classical in the context of array processing, is considered:

$$\mathbf{x} = A\,\mathbf{m} + \mathbf{v} = \mathbf{s} + \mathbf{v}, \qquad (1)$$

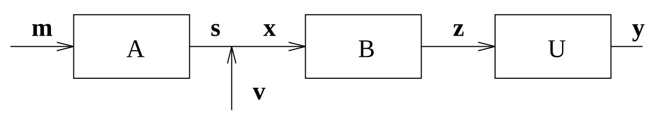
where $\mathbf{m}$ and $\mathbf{v}$ are independent random vectors of dimension $P$ and $N$, respectively, with values a priori in the complex field, and with finite moments up to order 4; $\mathbf{x}$ stands for the observation, $\mathbf{s}$ for the signal part in the observation (part of interest), $\mathbf{m}$ for source complex envelopes, $\mathbf{v}$ for an additive noise with i.i.d. components, not necessarily Gaussian ($c_v$ denotes the noise kurtosis), and $A$ for a $N \times P$ unknown matrix, $P \leq N$. As usual in the source separation problem, sources $m_i(t)$ are assumed to be non Gaussian, for at least $P-1$ of them, and to be mutually independent up to order 4. The goal is to recover sources $\mathbf{m}$, and possibly the columns $\mathbf{a}_i$ of $A$.

Denote $C_{i\ell,\mathbf{z}}^{jk} = cum\{z_i, z_j^*, z_k^*, z_\ell\}$, $C_{i,\mathbf{z}}^{j} = cum\{z_i, z_j^*\}$, the cumulants of variable $\mathbf{z}$ of orders 4 and 2, respectively; superscripts $(^*)$ and $(^\dagger)$ represent respectively complex conjugation and complex transposition. For convenience, one denotes $c_p = C_{pp,\mathbf{m}}^{pp}$, the source kurtosis. Also denote $W = U\,B$, aiming at inverting $A$, so that $\mathbf{y} = W\,\mathbf{x}$ is an LS estimate of the source vector, up to the standard indeterminacies [5]. However, the spatial matched filter (SMF), $W_{smf}$, is generally a better estimate of the sources, once $W$ has been estimated; in subsequent computer simulations, the SMF will be used.

In practice, the whitening matrix $B$ is first computed from an estimate of the source covariance, $\hat{R}_s$. Next, once $B$ is fixed, the unitary part $U$ is estimated by maximizing one of the two functionals, aiming at maximizing the statistical independence of components $y_i$:

$$\Upsilon_2(U) = \sum_{n=1}^{P} \left(C_{pp,\mathbf{y}}^{pp}\right)^2, \quad \Upsilon_1(U) = \mu \sum_{n=1}^{P} C_{pp,\mathbf{y}}^{pp}. \qquad (2)$$

It is known from [6] that $\Upsilon_2$ is a contrast, and from [8] that $\Upsilon_1$ is a contrast if $c_p$ have the same sign, $1 \leq p < P$. In [8], it has been shown that the absolute maximum of $\Upsilon_1$ can be computed analytically with a reduced complexity.




**Survey.** Approaches that have been proposed to this problem include: (i) adaptive algorithms [14] [13] [12] [19], and closed-form solutions based on (ii) second order statistics only [17] [16] [11], (iii) higher orders only [9] [7], or (iv) making use of both second and higher orders [10] [15] [18] [2] [6] [8]. The present submission is related to the three latter references only.

Approaches based on contrast $\Upsilon_2$ [6], referred to as "**C2**" in the remaining, have already shown an excellent behavior compared to the other approaches [4]. The goal of this paper is to report on performances of "**C1**" approaches, based on $\Upsilon_1$ [8], and compare them with JADE [2] denoted "**J2**", and with **C2**.

**Contrasts.** The main interest of maximizing a contrast, over cumulant matching techniques for instance, is that the solution can pretend to some optimality in presence of noise with unknown statistics (even non Gaussian). Thus, contrasts are attractive in presence of noise, or when statistics (e.g. cumulants) are estimated over short data records. In addition, interferences may be incorporated in the non Gaussian noise effect when sources are sought to be extracted.

[3]Also with I3S-CNRS, Sophia-Antipolis, F-06560 Valbonne `comon@alto.unice.fr`. This work has been supported in part by DRET, under contract 953441400.

## 2. ACTUAL PERFORMANCE

In this section, we analyse, by computer simulations, the performance of **C1** and compare it to that of **J2** and **C2**, recalling that it has been shown in [4] [5] that the methods **J2** and **C2** have approximately the same performance in most situations, provided that the number of sources $P$ is not over-estimated, while the performance of **J2** degrades with respect to that of **C2** when P is over- estimated. Note that the figures 3 and 4 are the results of the average over 10 independent realizations.

### 2.1. Performance measure

The performance is measured in terms of the following matrix, where $\mathbf{w}_i^\dagger$ are rows of $W_{smf}$:

$$\mathrm{SINR}_j[\mathbf{w}_i] = \frac{var\{(\mathbf{w}_i^\dagger \mathbf{a}_j)m_j\}}{var\{\sum_{p\neq j}(\mathbf{w}_i^\dagger \mathbf{a}_p)m_p\} + var\{\mathbf{w}_i^\dagger \mathbf{v}\}}. \quad (3)$$

The output $y_{i_o}$ that estimates source $m_j$ the best is the one that maximizes $\mathrm{SINR}_j[\mathbf{w}_i]$ over all indices $i$. Assuming this estimate for every source leads to a $P$-dimensional performance vector, SINRM, corresponding to a maximal SINR for every source. This is indeed the most natural criterion to quantitatively evaluate the performance of source estimates, with the goal of comparing source separators [5] [4].

Note that with the measure (3), in contrast with the gap proposed in [6] [8], it might happen that the same output estimates two different source envelopes.

### 2.2. Results

**Summary.** After a great number of computer simulations, we are able to conclude that, provided (i) the number of sources in not over-estimated, (ii) the background noise is Gaussian, (iii) no source is Gaussian and their kurtosis have the same sign, then the performance of **C1** is the same as that of **J2** and **C2**, whatever the kind of sources and the observation duration. In these cases, the illustration of the **J2** and **C2** methods performance with respect to the optimal ones, presented in [5] [4] for several sources scenari, still holds for **C1**. However, in the other cases, the **C1** method behaviour may significantly differ from **J2** and **C2**, as it is discussed and illustrated in the next sections.

In the figures below, the line types associated to the methods are: solid for **J2**, dashdotted for **C2**, and dashed for **C1**, and the vertical scale is always in dB.

**Sources with different fourth-order cumulant signs.** In the presence of sources with different kurtosis signs, for a given value of the constant $\mu$ in criterion (2) and for a given source scenario, there is, for each source taken separately, a value of its kurtosis $c_i$ over which or under which the separation of this source from the other by the **C1** method fails. This value is, in particular, a function of $\mu$, of the SNR (Signal to Noise Ratio), and of the modulus and sign of the source kurtosis, $c_i$, $1 < i < P$. This result is illustrated in figure 1 which shows, for $P = 2$, an Uniformly Linear Array (ULA) of $N = 4$ sensors, and a Gaussian noise, the variation of the steady-state $\mathrm{SINRM}_1$ [5], (it is the same for the source 2), at the output of **C1** as a function of $c_2$, for $c_1 = -2$ and for sources whose input SNR is 10 dB. Note the optimal separation of sources as the constant $\mu$ corresponds to the sign of the highest source kurtosis modulus. In the other cases or when the sources have kurtosis with the same modulus but with a different sign, the method **C1** fails in separating the sources.

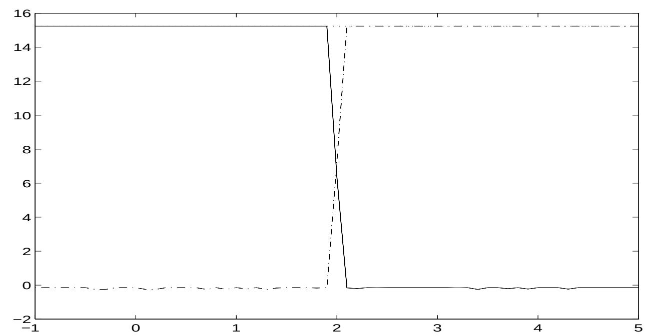



Figure 1: $\mathrm{SINRM}_1$ as a function of $c_2$, with $P = 2$, $N = 4$, ULA, $\theta_1 = 0$, $\theta_2 = 20$ degrees, SNR = 10 dB, $c_1 = -2$, $\mu = -1$: solid, $\mu = +1$: dashdotted

**Non Gaussian background noise.** In the presence of a non Gaussian background noise, the steady state performance of **C1** becomes very poor as soon as the absolute value of the noise kurtosis $c_v$ becomes greater than a threshold which depends on several parameters, and which increases with the SNR and the source kurtosis. However, the threshold associated with **C1** is much higher than that associated with **J2** or **C2**, which proves that **C1** is much more robust to a non Gaussian noise than **J2** and **C2**. This result is illustrated in figure 2 which shows the variation of the $\mathrm{SINRM}_1$ at the output of **C1**, **C2** and **J2**, as a function of $c_v$, for $P = 2$ with SNR = 5 dB and for several values of $c_1 = c_2 = c$. Note the very high robustness of **C1** to the non Gaussian noise with respect to **J2** and **C2**.

**Presence of a Gaussian source.** When one of the sources is Gaussian or quasi- Gaussian, the performance of **C1** degrades with respect to that of **J2** and **C2** and the performance degradation increases with the input SNR of the sources and as the number of independent snapshots $K$, used to estimate the data statistics, decreases. These results are illustrated in figure 3 which shows the variations of the $\mathrm{SINRM}_1$ at

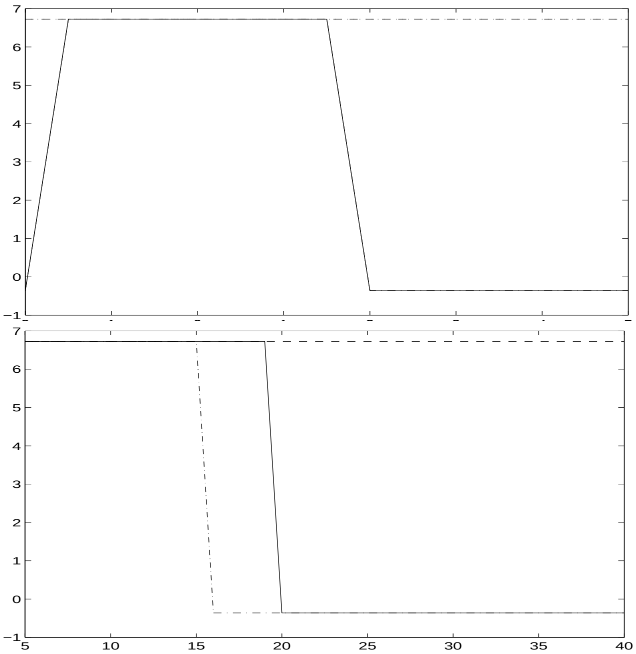



Figure 2: $\text{SINRM}_1$ at the output of **C1**, **J2** and **C2** as a function of $c_v$, ULA, $N = 4$, $P = 2$, $\theta_1 = 0$, $\theta_2 = 10$ degrees, SNR = 5 dB, $c = -0.1$: top, $c = -1$: bottom

the output of **C1**, **J2** and **C2** as a function of $K$, for an ULA of 4 sensors receiving 2 sources, the second one being Gaussian, and for several values of the input SNR of the sources.

**Over-estimation of the source number.** When the number of sources is over-estimated, the convergence speed of the methods **C1**, **J2** and **C2** decreases. In this case, the method **C2** remains the most powerful whereas **J2** becomes the slowest. Thus, the performance of **C1**, although lower than that of **C2**, is still better than **J2**'s. Figure 4 illustrates this result for $P = 2$ sources, by showing the variations of the $\text{SINRM}_1$ (it is the same for $\text{SINRM}_2$) at the output of **C1**, **J2** and **C2** as a function of $K$, when the number of estimated sources is 2 or 3.

## 3. ASYMPTOTIC ANALYSIS

It has been shown by Cardoso that **J2** and **C2** ultimately reach the same performance, but he did not compute explicitly the variance of the estimated matrix $U$, as a function of the data length $K$. From this variance, on can also compute an asymptotical expression for the SINR matrix, that could serve as a reference in the calculation of actual performance. This cannot be reported for reasons of space.

But the principle can be outlined as follows: station-

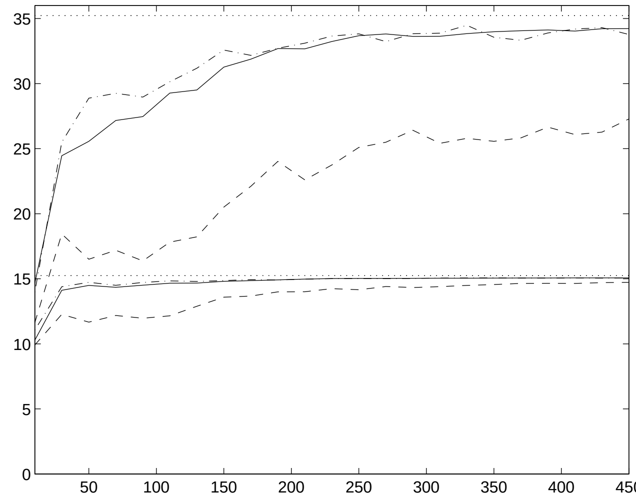



Figure 3: $\text{SINRM}_1$ at the output of **C1**, **J2** and **C2** as a function of $K$, with ULA $N = 4$, $P = 2$, $\theta_1 = 0$, $\theta_2 = 20$ degrees, SNR = 10, 30 dB, $c_1 = -1$, $c_2 = 0$

ary matrices of a criterion $\Upsilon(U)$ satisfy a relation:

$$h(U, \gamma) = 0, \tag{4}$$

where $\gamma$ denotes the set of sample cumulants of $\mathbf{z}$ that are utilized in the computation of $\Upsilon$. A first order expansion can be explicitly computed because everything is known in $h(\cdot, \cdot)$, and leads to:

$$G\mathbf{vec}[U] = H\mathbf{vec}[\gamma]. \tag{5}$$

The variance of $U_{ij}$ can thus be accessed by the formula:

$$Var\{\mathbf{vec}[U]\} = G^{-}H\,Var\{\mathbf{vec}[\gamma]\}\,H^{\dagger}G^{-\dagger}. \tag{6}$$

This covariance can be computed once we know the covariance of sample cumulants. Using McCullagh bracket notation, and noting $[\bar{2}]expr = expr + expr^*$, this covariance takes the general form:

$$\begin{aligned}
K\ Var\{\hat{C}_{i\ell}^{jk}, \hat{C}_{IL}^{JK}\} &= C_{i\ell IL}^{jkJK} \\
&+[\bar{2}][4]C_{iI}^{jkJK}C_{\ell L} + [\bar{2}][4]C_{iIL}^{jkJ}C_{\ell}^{K} \\
&+[\bar{2}][2]\left([8]C_{iI}^{jkJ}C_{\ell L}^{K} + [2]C_{i}^{jkJK}C_{\ell IL} + [2]C_{i\ell}^{kJK}C_{IL}^{j}\right) \\
&+[\bar{2}][4]C_{iL}^{JK}C_{I\ell}^{jk} + [\bar{2}][4]C_{iIL}^{K}C_{\ell}^{jkJ} + C_{i\ell IL}C^{jkJK} \\
&\quad + C_{i\ell}^{JK}C_{IL}^{jk} + [8]C_{iI}^{jJ}C_{\ell L}^{kK} + [\bar{2}][4]C_{i\ell I}^{J}C_{L}^{jkK} \\
&+[16]C_{iI}^{jJ}C^{kK}C_{\ell L} + [16]C_{iI}^{jJ}C_{L}^{k}C_{\ell}^{K} \\
&\quad +[\bar{2}][8]C_{i}^{jJK}C_{I}^{k}C_{\ell L} + [\bar{2}][8]C_{i\ell I}^{J}C^{jK}C_{L}^{k} \\
&\quad +[\bar{2}][2]C_{i\ell IL}C^{jJ}C^{kK} + [\bar{2}][2]C_{i\ell}^{JK}C_{I}^{j}C_{L}^{k} \\
&+[32]C_{iI}^{j}C_{L}^{kJ}C_{\ell}^{K} + [32]C_{i}^{jJ}C_{I}^{kK}C_{\ell L}
\end{aligned}$$

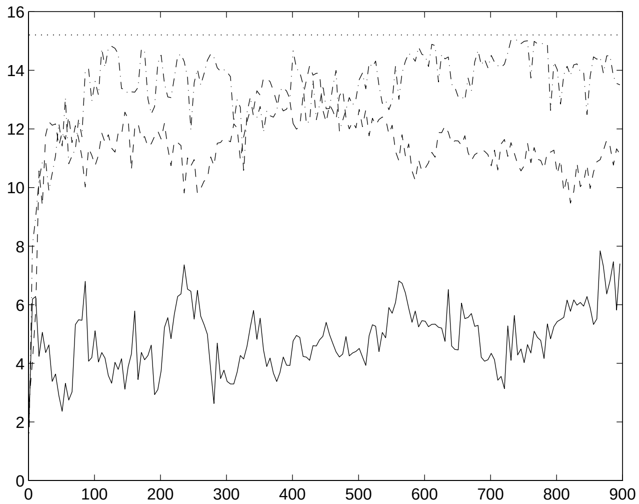



Figure 4: $\mathrm{SINRM}_1$ at the output of **C1**, **J2** and **C2** as a function of $K$, with ULA, $N = 4$, $P = 2$, $\hat{P} = 2$ or 3, $\theta_1 = 0$, $\theta_2 = 20$ degrees, SNR = 10 dB, $c_1 = c_2 = -1$.

$$
\begin{aligned}
&+[16] C_{iI}^{j} C^{kJK} C_{\ell L} + [16] C_{i}^{jJ} C_{IL}^{k} C_{\ell}^{K} \\
&+[\bar{2}][8] C_{i\ell I} C_{L}^{jK} C^{kJ} + [\bar{2}][8] C_{i\ell}^{J} C_{L}^{jK} C_{I}^{k} \\
&+[\bar{2}][4] C_{i\ell I} C^{jJK} C_{L}^{k} + [\bar{2}][4] C_{i\ell}^{J} C_{IL}^{j} C^{kK} \\
+&[16] C_{iI} C^{jJ} C_{L}^{l} C_{\ell}^{K} + [4] C_{i}^{J} C_{I}^{j} C_{L}^{k} C_{\ell}^{K} \\
&+[4] C_{iI} C^{jJ} C^{kK} C_{\ell L}
\end{aligned}
$$

Recall that in the circular case, the covariance of sample cumulants has been given in [3]. One can check out that the above indeed deflates to the same kind of formula in the latter case:

$$
\begin{aligned}
K\ Var\{\hat{C}_{i\ell}^{jk}, \hat{C}_{IL}^{JK}\} = & C_{i\ell IL}^{jkJK} \\
&+[\bar{2}][4] C_{iIL}^{jkJ} C_{\ell}^{K} \\
&+[\bar{2}][4] C_{iL}^{JK} C_{I\ell}^{jk} + C_{i\ell}^{JK} C_{IL}^{jk} + [8] C_{iI}^{jJ} C_{\ell L}^{kK} \\
&+[16] C_{iI}^{jJ} C_{L}^{k} C_{\ell}^{K} + [\bar{2}][2] C_{IL}^{jk} C_{i}^{J} C_{\ell}^{K} \\
&+[4] C_{i}^{J} C_{I}^{j} C_{L}^{k} C_{\ell}^{k}
\end{aligned}
$$

## 4. CONCLUDING REMARKS

The previous analysis shows that although the performances of **C1** are different from those of **J2** or **C2** in some situations; they remain very promising in most practical cases. Besides, the implementation of **C1** seems less costly than that of **C2** [8], which may compensate its somewhat less attractive performances with respect to **C2**. Comparisons with the theoretical asymptotic limits will be reported in a future paper.

## 5. REFERENCES


[1] J. F. CARDOSO, "On the performance of source separation algorithms", in *Proc. EUSIPCO*, Edinburgh, Sept. 1994, pp. 776–779.

[2] J. F. CARDOSO, A. SOULOUMIAC, "Blind beamforming for non-Gaussian signals", *IEE Proceedings - Part F*, vol. 140, no. 6, pp. 362–370, Dec. 1993.

[3] J. F. CARDOSO, E. MOULINES, "Asymptotic Performance Analysis of Direction-Finding Algorithms based on Fourth-Order Cumulants", *IEEE Trans. SP*, pp. 214–224, Jan. 1995.

[4] P. CHEVALIER, "Méthodes aveugles de filtrage d'antennes", *Revue d'Electronique et d'Electricité, SEE*, , no. 3, pp. 48–58, Sept 1995.

[5] P. CHEVALIER, "On the performance of higher order blind sources separation methods", in *IEEE-ATHOS Workshop on Higher-Order Statistics*, Begur, Spain, 12–14 June 1995, pp. 30–34.

[6] P. COMON, "Independent Component Analysis, a new concept ?", *Signal Processing, Elsevier*, vol. 36, no. 3, pp. 287–314, Apr. 1994, Special issue on HOS.

[7] P. COMON, "Contrasts for multichannel blind deconvolution", *Signal Processing Letters*, vol. 3, no. 7, pp. 209–211, July 1996.

[8] P. COMON, E. MOREAU, "Improved contrast dedicated to blind separation in communications", *ICASSP'97*, Munich, Apr. 1997.

[9] L. DeLATHAUWER, B. DeMOOR, J. VANDEWALLE, "Independent component analysis based on higher-order statistics only", in *SSAP Workshop*, Corfu, Greece, June 1996, pp. 356–359.

[10] N. DELFOSSE, P. LOUBATON, "Adaptive blind separation of independent sources: a deflation approach", *Signal Processing*, vol. 45, pp. 59–83, 1995.

[11] Z. DING, "An outer-product decomposition algorithm for multichannel blind identification", in *IEEE SSAP Workshop*, Corfu, Greece, June 1996, pp. 132–135.

[12] D. GESBERT, P. DUHAMEL, S. MAYRARGUE, "Subspace-based adaptive algorithms for the blind equalization of multichannel FIR filters", in *Proc. EUSIPCO*, Edinburgh, Sept. 1994, pp. 712–715.

[13] Y. INOUYE, T. HABE, "Blind equalization of multichannel linear time-invariant systems", *Inst. Elec. Inf. and Com. Eng.*, vol. 24, no. 5, pp. 9–16, May 1995.

[14] O. MACCHI, E. MOREAU, "Self-Adaptive Source Separation. Part I", to appear in *IEEE Trans. SP*.

[15] A. MANSOUR, C. JUTTEN, "Fourth-order criteria for blind sources separation", *IEEE Trans. SP*, vol. 43, no. 8, pp. 2022–2025, Aug. 1995.

[16] E. MOULINES, P. DUHAMEL, J. F. CARDOSO, S. MAYRAGUE, "Subspace methods for the blind identification of multichannel FIR filters", *IEEE Trans. on SP*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[17] D.T.M. SLOCK, C.B. PAPADIAS, "Further results on blind identification and equalization of multiple FIR channels", *ICASSP95*, Detroit, Michigan, May 1995.

[18] A. SWAMI, G. GIANNAKIS, S. SHAMSUNDER, "Multichannel ARMA processes", *IEEE Trans. on SP*, vol. 42, pp. 898–913, Apr. 1994.

[19] D. YELLIN, E. WEINSTEIN, "Multichannel signal separation: Methods and analysis", *IEEE Trans. SP*, vol. 44, no. 1, pp. 106–118, Jan. 1996.