Second IEEE Sensor Array and Multichannel Signal Processing Workshop, Rosslyn, VA, 4-6 August 2002

ASYMPTOTIC PERFORMANCE OF CONTRAST-BASED BLIND SOURCE SEPARATION ALGORITHMS

Laurent Albera[1,2], *Pierre Comon*[1]

(1) I3S, Algorithmes-Euclide-B, BP 121, F-06903 Sophia-Antipolis Cedex, France `[albera,comon]@i3s.unice.fr`
(2) THALES Communications, 66 rue du Fossé Blanc, BP 156, F-92231 Gennevilliers Cedex, France

ABSTRACT

For several years, contrast-based Blind Source Separation (BSS) has been successfully used in several areas, including radiocommunications. Here a functional approach relying on differential calculus theory is proposed, aiming at analyzing asymptotic performances of BBS contrast criteria: the variance of the estimated separating matrix is expressed as a function of that of estimated cumulants. As an example, this paper focuses on three widely used fourth order (FO) contrast criteria. This allows to quantify the behavior of these three separators for large samples.

1. INTRODUCTION

For more than a decade, Second Order (SO) and Higher Order (HO) blind source separation (BSS) methods [3], also referred to as Independent Component Analysis (ICA) methods [2], have been developed to separate several statistically independent sources from measurements. The purpose of this paper is to examine the asymptotic performances (*e.g.* covariance of estimate) of contrast-based algorithms. Although the subject of asymptotic analysis has already been addressed in the signal processing literature, for instance, performance of SO [6] and ML [9] estimators in antenna array processing, or behavior of SO and HO BSS algorithms in the presence of zero-mean cyclostationary sources [7], this paper proposes a functional approach allowing to compare asymptotic performances of BSS contrast criteria. As an illustration, 3 fourth order (FO) contrast criteria already compared in [5] by computer experiments, are mainly focused on, for subsequent asymptotic performance analysis.

2. ASSUMPTIONS AND NOTATION

Assume that for any fixed n, K complex outputs $x_k(n)$ of a noisy mixture of P statistically sources $s_p(n)$ are available, where $K \geq P$. The vector $\boldsymbol{x}(n)$ of the measured array outputs is given by

$$\boldsymbol{x}(n) = \boldsymbol{H}\,\boldsymbol{s}(n) + \boldsymbol{v}(n) \tag{1}$$

where $\boldsymbol{H}$, $\boldsymbol{s}(n)$, $\boldsymbol{v}(n)$ are the $K \times P$ constant mixing matrix, the source vector and the noise random vector, respectively. Note that for any fixed n, $\boldsymbol{s}(n)$ and $\boldsymbol{v}(n)$ are independent.

For the sake of convenience we need to define, for any n, the entries of the 4th order cumulant tensor $\boldsymbol{C_z}$ of a random vector, $\boldsymbol{z}(n)$, stationary and ergodic up to order 4:

$$C_{i\ell,\boldsymbol{z}}^{jk} = \mathsf{Cum}\{z_i(n),\, z_j(n)^*,\, z_k(n)^*,\, z_\ell(n)\} \tag{2}$$

Moreover, we further assume the following hypotheses:

- **A1**. For any fixed n, sources $s_p(n)$ are stationary, ergodic and uncorrelated at order 4, with values a priori in the complex field;
- **A2**. For any fixed n, noise values $v_k(n)$ are stationary, ergodic, with values a priori in the complex field too;
- **A3**. At most one fourth-order marginal source cumulant, $C_{p,\boldsymbol{s}} \stackrel{\text{def}}{=} C_{pp,\boldsymbol{s}}^{pp}$, is null;
- **A4**. The mixing matrix $\boldsymbol{H}$ is square and unitary.

3. CONTRAST-BASED BSS METHODS

The goal of BSS consists of determining a separating matrix, $\boldsymbol{U}$, called separator, such that

$$\boldsymbol{y}(n) \stackrel{\text{def}}{=} \boldsymbol{U}\,\boldsymbol{x}(n) \tag{3}$$

yields an estimate of the vector $\boldsymbol{s}(n)$. On the other hand, ICA aims at identifying the mixing matrix $\boldsymbol{H}$.

Note that assumption (**A4**) is not restrictive if a spatial prewhitening has been performed [3]. But we limit our study for the time being to the effect of fourth-order estimation errors on the separator $\boldsymbol{U}$.

Various approaches have been devised for BSS or ICA [1]. We shall focus exclusively on those maximizing a contrast measure of $\boldsymbol{y}$. Recall that contrasts are criteria $\Upsilon(\boldsymbol{U}; \boldsymbol{C_x})$ satisfying the properties below [2] [10]:

P1. *Invariance:* The contrast should not change within the set $\mathcal{T}$ of trivial matrices, which means that $\forall \boldsymbol{x} \in \mathcal{H} \cdot \mathcal{S}, \forall \boldsymbol{U} \in \mathcal{T}, \Upsilon(\boldsymbol{U}; \boldsymbol{C_x}) = \Upsilon(\boldsymbol{I}; \boldsymbol{C_x})$.

P2. *Domination:* If sources are already separated, any matrix should decrease the contrast. In other words, $\forall \boldsymbol{U} \in \mathcal{H}, \forall \boldsymbol{x} \in \mathcal{S}, \Upsilon(\boldsymbol{U}; \boldsymbol{C_x}) \leq \Upsilon(\boldsymbol{I}; \boldsymbol{C_x})$.

P3. *Discrimination:* The maximum contrast should be reached only for matrices linked to each other via trivial matrices: $\forall \boldsymbol{x} \in \mathcal{S}, \Upsilon(\boldsymbol{U}; \boldsymbol{C_x}) = \Upsilon(\boldsymbol{I}; \boldsymbol{C_x}) \Rightarrow \boldsymbol{U} \in \mathcal{T}$.

where $\mathcal{H}, \mathcal{H} \cdot \mathcal{S}, \boldsymbol{I}$ are a set of matrices, the set of processes obtained by matrix mappings $\mathcal{H}$ on processes of $\mathcal{S}$, and the identity matrix, respectively. Note that a trivial matrix is of the form $\boldsymbol{\Lambda\Pi}$ where $\boldsymbol{\Lambda}$ is an invertible diagonal matrix and $\boldsymbol{\Pi}$ a permutation.

The goal of this paper is to evaluate the asymptotic statistical properties (*e.g.* covariance) of the matrix $\boldsymbol{U}$ delivered by contrast-based algorithms.

4. ASYMPTOTIC PROPERTIES: A FUNCTIONAL APPROACH

¿From now on, it is assumed that $\Upsilon(\cdot, \boldsymbol{C})$ is of class C^2, and $\Upsilon(\boldsymbol{U}, \cdot)$ is of class C^1. This will be satisfied for criteria given in section 5. The optimal solution $\boldsymbol{U}_o$ is defined as the absolute maximum of a contrast $\Upsilon(\boldsymbol{U}; \boldsymbol{C_x})$:

$$\boldsymbol{U}_o = Arg \operatorname*{Max}_{\boldsymbol{U}} \Upsilon(\boldsymbol{U}; \boldsymbol{C_x}) \tag{4}$$

where $\boldsymbol{C_x}$ is the exact cumulant tensor of the observation. In practice, $\boldsymbol{C_x}$ is estimated by a quantity $\hat{\boldsymbol{C}}_{\boldsymbol{x}}$, which involves estimation errors on $\boldsymbol{U}$; this yields a solution $\hat{\boldsymbol{U}}$:

$$\hat{\boldsymbol{U}} = Arg \operatorname*{Max}_{\boldsymbol{U}} \Upsilon(\boldsymbol{U}; \hat{\boldsymbol{C}}_{\boldsymbol{x}}) \tag{5}$$

Both $\boldsymbol{U}_o$ and $\hat{\boldsymbol{U}}$ are maxima of Υ, and thus satisfy the stationary point equations:

$$\boldsymbol{h}(\boldsymbol{U}_o, \boldsymbol{C_x}) = \boldsymbol{0}, \quad \boldsymbol{h}(\hat{\boldsymbol{U}}, \hat{\boldsymbol{C}}_{\boldsymbol{x}}) = \boldsymbol{0} \tag{6}$$

where $\boldsymbol{h}(\cdot, \boldsymbol{C})$ denotes the gradient of $\Upsilon(\boldsymbol{U}, \boldsymbol{C})$ with respect to $\boldsymbol{U}$, in a sense subsequently defined.

Now, $\boldsymbol{h}$ is a well defined function in a P^2-dimensional linear space on the real field $\mathbb{R}$. In fact, Υ is twice continuously differentiable with respect to $\boldsymbol{U}$, and the tangent space to the variety of unitary matrices is the linear space of skew-hermitian matrices ($\boldsymbol{A}^{\mathsf{H}} = -\boldsymbol{A}$), which is indeed of dimension P^2 and admits as a basis the set of matrices $\{\boldsymbol{A}(q,r)\}$, null everywhere except in rows and columns (q,r), (r,q), such that

$$d\boldsymbol{U} = \sum_{q,r=1}^{P} d\mu_{qr} \boldsymbol{A}(q,r) \boldsymbol{U} \tag{7}$$

where $\boldsymbol{A}(q,r)_{vw} \stackrel{\text{def}}{=} \delta_{qv}\delta_{rw} - \delta_{qw}\delta_{rv}$ if $q < r$, $\boldsymbol{A}(q,r)_{vw} \stackrel{\text{def}}{=} \jmath \delta_{qv}\delta_{rw}$ if $q = r$, $\boldsymbol{A}(q,r)_{vw} \stackrel{\text{def}}{=} \jmath\,(\delta_{qv}\delta_{rw} + \delta_{qw}\delta_{rv})$ if $q > r$, and $\jmath^2 \stackrel{\text{def}}{=} -1$.

Next, $\boldsymbol{h}(\cdot,\cdot)$ is continuously differentiable, which allows to resort to the implicit functions theorem in the neighborhood of $(\boldsymbol{U}_o, \boldsymbol{C_x})$. This yields, from (6):

$$\dot{\boldsymbol{h}}_{\boldsymbol{U}}(\boldsymbol{U}_o, \boldsymbol{C_x})\, d\boldsymbol{U} + \dot{\boldsymbol{h}}_{\boldsymbol{C}}(\boldsymbol{U}_o, \boldsymbol{C_x})\, d\boldsymbol{C} = o(d\boldsymbol{U}, d\boldsymbol{C}) \tag{8}$$

Thus, in the neighborhood of $(\boldsymbol{U}_o, \boldsymbol{C_x})$, $\hat{\boldsymbol{U}} = \boldsymbol{U}_o + d\boldsymbol{U}$ can be expressed as a function of $\hat{\boldsymbol{C}}_{\boldsymbol{x}} = \boldsymbol{C_x} + d\boldsymbol{C}$. This can be rewritten in block form as [5]:

$$\boldsymbol{B}_1\, \mathbf{vec}[d\boldsymbol{U}] = \boldsymbol{B}_2\, \mathbf{vec}[d\boldsymbol{C}] \tag{9}$$

where $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ are matrices of dimension $P^2 \times P^2$ and $P^2 \times M$, respectively, built from the second derivatives of Υ, $\partial^2\Upsilon/\partial U \partial U$ and $\partial^2\Upsilon/\partial U \partial C$, stored in the proper manner. Here, M denotes the number of free parameters in $\boldsymbol{C_x}$, and, for any $P \geq 4$, is equal to $M = P(P+1)(P^2+P+1)/8$ in the complex case, which deflates to $M = P(P+1)(P+2)(P+3)/24$ in the real case.

The variance of $d\boldsymbol{U}_{vw}$ and therefore, the one of $\hat{\boldsymbol{U}}_{vw}$ can thus theoretically be accessed by the formula:

$$Var\Big\{\mathbf{vec}\Big[\hat{\boldsymbol{U}}\Big]\Big\} = \boldsymbol{B}_1^{-1}\boldsymbol{B}_2 Var\Big\{\mathbf{vec}\Big[\hat{\boldsymbol{C}}_{\boldsymbol{x}}\Big]\Big\} \boldsymbol{B}_2^{\mathsf{H}} \boldsymbol{B}_1^{-\mathsf{H}} \tag{10}$$

Nevertheless, matrix $\boldsymbol{B}_1$ is in general not full rank, because the set of $d\mu_{qr}$ does not form a free family. Its rank is $P(P-1)$, so that the inverse above should be replaced by a pseudo-inverse. Nevertheless, this covariance can be consequently still computed once we know the covariance of sample cumulants. Using McCullagh bracket notation, and noting $[\bar{2}]expr = expr + expr^*$, this covariance is given in [5] in the general case. In the symetrically distributed case in which we are interested, the covariance takes the form:

$$\begin{aligned} N\ Var\{\hat{C}_{i\ell}^{jk}, \hat{C}_{JK}^{IL}\} &= C_{i\ell IL}^{jkJK} \\ &+ [\bar{2}][4] C_{iI}^{jkJK} C_{\ell L} + [\bar{2}][4] C_{iIL}^{jkJ} C_{\ell}^{K} \\ &+ [\bar{2}][4] C_{iL}^{JK} C_{I\ell}^{jk} + [\bar{2}][4] C_{iIL}^{K} C_{\ell}^{jkJ} + C_{i\ell IL} C^{jkJK} \\ &\quad + C_{i\ell}^{JK} C_{IL}^{jk} + [8] C_{iI}^{jJ} C_{\ell L}^{kK} + [\bar{2}][4] C_{i\ell I}^{J} C_{L}^{jkK} \\ &+ [16] C_{iI}^{jJ} C^{kK} C_{\ell L} + [16] C_{iI}^{jJ} C_{L}^{k} C_{\ell}^{K} \\ &\quad + [\bar{2}][8] C_{i}^{jJK} C_{I}^{k} C_{\ell L} + [\bar{2}][8] C_{i\ell I}^{J} C^{jK} C_{L}^{k} \\ &\quad + [\bar{2}][2] C_{i\ell IL} C^{jJ} C^{kK} + [\bar{2}][2] C_{i\ell}^{JK} C_{I}^{j} C_{L}^{k} \\ &+ [16] C_{iI} C^{jJ} C_{L}^{k} C_{\ell}^{K} + [4] C_{i}^{J} C_{I}^{j} C_{L}^{k} C_{\ell}^{K} \\ &\quad + [4] C_{iI} C^{jJ} C^{kK} C_{\ell L} \end{aligned} \tag{11}$$

where N denotes the number of snapshots.

5. EXAMPLES AND ASYMPTOTIC ANALYSIS OF PARTICULAR CONTRASTS

Define the three fourth order contrast criteria below:

$$\Upsilon_1(\boldsymbol{U};\boldsymbol{C_x}) = \lambda\sum_{p=1}^{P} C_{p,\boldsymbol{y}},\ \Upsilon_2(\boldsymbol{U};\boldsymbol{C_x}) = \sum_{p=1}^{P}(C_{p,\boldsymbol{y}})^2$$

$$\Upsilon_3(\boldsymbol{U};\boldsymbol{C_x}) = \sum_{p,k,\ell=1}^{P}\left|C_{p\ell,\boldsymbol{y}}^{pk}\right|^2 \qquad (12)$$

where λ is a fixed sign. Note [5] that Υ_1 is a contrast if, for any $1 \leq p \leq P$, $C_{p,\boldsymbol{s}}$ have the same sign λ, and that Υ_3 is the contrast linked with the JADE algorithm [1].

5.1. Asymptotic results

After a first differential calculus with respect to $\boldsymbol{U}$, we obtain:

$$d\Upsilon_{1,\boldsymbol{U}} = 4\lambda\left[\sum_{q<r} d\mu_{qr}\Re\{C_{rq,\boldsymbol{y}}^{qq} - C_{qr,\boldsymbol{y}}^{rr}\} - \sum_{q>r} d\mu_{qr}\Im\{C_{qr,\boldsymbol{y}}^{rr} + C_{rq,\boldsymbol{y}}^{qq}\}\right] \qquad (13)$$

$$d\Upsilon_{2,\boldsymbol{U}} = 8\left[\sum_{q<r} d\mu_{qr}\left(C_{q,\boldsymbol{y}}\Re\{C_{rq,\boldsymbol{y}}^{qq}\} - C_{r,\boldsymbol{y}}\Re\{C_{qr,\boldsymbol{y}}^{rr}\}\right) - \sum_{q>r} d\mu_{qr}\left(C_{r,\boldsymbol{y}}\Im\{C_{qr,\boldsymbol{y}}^{rr}\} + C_{q,\boldsymbol{y}}\Im\{C_{rq,\boldsymbol{y}}^{qq}\}\right)\right] \qquad (14)$$

$$d\Upsilon_{3,\boldsymbol{U}} = 8\left[\sum_{q<r}\sum_{k,\ell} d\mu_{qr}\Re\left\{C_{q\ell,\boldsymbol{y}}^{qk}C_{rk,\boldsymbol{y}}^{q\ell} - C_{r\ell,\boldsymbol{y}}^{rk}C_{qk,\boldsymbol{y}}^{r\ell}\right\} - \sum_{q>r}\sum_{k,\ell} d\mu_{qr}\Im\left\{C_{r\ell,\boldsymbol{y}}^{rk}C_{qk,\boldsymbol{y}}^{r\ell} + C_{q\ell,\boldsymbol{y}}^{qk}C_{rk,\boldsymbol{y}}^{q\ell}\right\}\right] \qquad (15)$$

where $\Re\{z\}$ and $\Im\{z\}$ are respectively the real and imaginary parts of the complex number z.

So for each contrast, we can easily deduce from (13), (14) and (15) the function $\boldsymbol{h}_m$ defined in section (4). In particular, according to (4), (5), (6) and (13) the function $\boldsymbol{h}_1$ associated with $d\Upsilon_{1,\boldsymbol{U}}$ is described by

$$h_1(\boldsymbol{U},\boldsymbol{C})_{qr} = \begin{cases} \Re\{C_{rq,\boldsymbol{y}}^{qq} - C_{qr,\boldsymbol{y}}^{rr}\} & \text{if } q<r \\ -\Im\{C_{qr,\boldsymbol{y}}^{rr} + C_{rq,\boldsymbol{y}}^{qq}\} & \text{if } q>r \\ 0 & \text{if } q=r \end{cases} \qquad (16)$$

The implicit relation (9) rewrites:

$$d\left[d\boldsymbol{h}_m\right]_{\boldsymbol{U}}(\boldsymbol{U}_o,\boldsymbol{C_x}) = -d\left[d\boldsymbol{h}_m\right]_{\boldsymbol{C}}(\boldsymbol{U}_o,\boldsymbol{C_x}) + o(d\boldsymbol{U},d\boldsymbol{C}) \qquad (17)$$

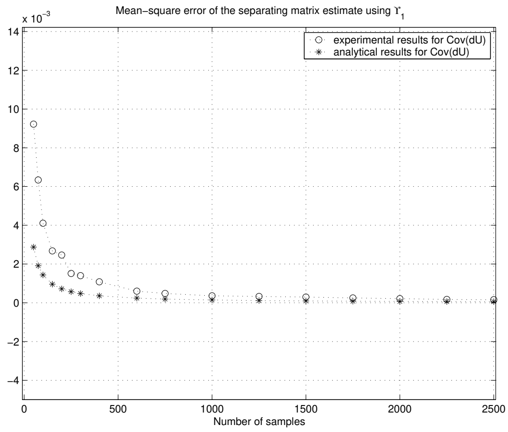

Fig. 1. Variance of estimated separating matrix $\boldsymbol{U}$ obtained by maximization of $\Upsilon_1(\boldsymbol{U})$.

where, for Υ_1 and for any $1 \leq q, r \leq P$:

$$d\left[h_1(\boldsymbol{U},\boldsymbol{C})_{qr}\right]_{\boldsymbol{U}} = \sum_{i,j=1}^{P}\Theta_{qr}^{q'r'}d\mu_{q'r'} \qquad (18)$$

$$d\left[h_1(\boldsymbol{U},\boldsymbol{C})_{qr}\right]_{\boldsymbol{C}} = \sum_{i,j,k,l=1}^{P}\Theta_{qr}^{ijkl}dC_{i\ell,\boldsymbol{x}}^{jk} \qquad (19)$$

where $\Theta_{qr}^{q'r'}$ and Θ_{qr}^{ijkl} are given in appendix. Similar (but more complicated) relations, derived for Υ_2 and Υ_3, are not reported here for reasons of space.

5.2. Simulations

Empirical variance estimates. Simulations have been run for $P = 2$ independent QPSK sources, in the presence of Gaussian complex circular noise. The mixing matrix was of the form

$$\begin{pmatrix} \cos\theta & \sin\theta\, e^{\jmath\varphi} \\ -\sin\theta\, e^{-\jmath\varphi} & \cos\theta \end{pmatrix}$$

with $\theta = \pi/7$ and $\varphi = \pi/7$. The separating matrix has been computed using algorithms reported in [2], [1], and [4]. In order to obtain reliable variance estimates, 100 independent trials have been run, and the variances of each of the four entries $\hat{U}_{ij}$ has been estimated. In figures 1 to 3, we have plotted the sum of variances $\sum_{i=1}^{2}\mathrm{Var}\{\hat{U}_{ii}\}$ as a function of the sample size.

Theoretical asymptotic variance. In order to compute the theoretical variance, it was necessary to first calculate all the cumulants of even order up to eight. For this purpose, the multilinearity property of cumulants has been used, yielding the cumulants of the two outputs of a linear transform as a function of those of its inputs. For QPSK sources, we have

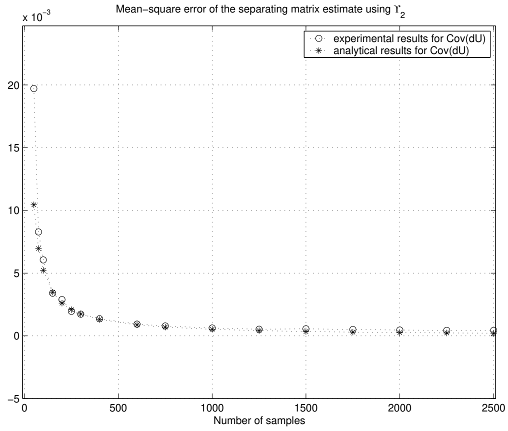

Fig. 2. Variance of estimated separating matrix $\boldsymbol{U}$ obtained by maximization of $\Upsilon_2(\boldsymbol{U})$.

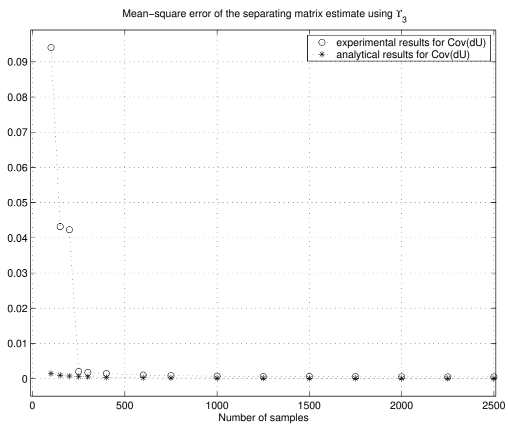

Fig. 3. Variance of estimated separating matrix $\boldsymbol{U}$ obtained by maximization of $\Upsilon_3(\boldsymbol{U})$.

the following (omitting subscript $\boldsymbol{s}$ in C_s):

$$
\begin{array}{rlll}
C_{11} = 0; & C_1^1 = 1 & & \\
C_{1111} = 1; & C_{111}^1 = 0; & C_{11}^{11} = -1 & \\
C_{111111} = 0; & C_{11111}^1 = -4; & C_{1111}^{11} = 0; & C_{111}^{111} = 4 \\
C_{11111111} = -34; & C_{1111111}^1 = 0; & C_{111111}^{11} = 34; & \\
C_{11111}^{111} = 0; & C_{1111}^{1111} = -33 & &
\end{array}
$$

General formulas are given in appendix. Since $P = 2$ is a simple case, first and second order derivatives can be computed directly in terms of $d\theta$ and $d\varphi$, and the 2×2 matrix $\boldsymbol{a}$ obtained is invertible. Thus, expressions such as (13) to (18) did not need to be used. On the other hand, expression (11) is central in this calculation. Results are reported in the figures 1 to 3, and show a good accordance with empirical results for large samples.

6. CONCLUDING REMARKS

The whole analytical calculus allows to write, for each contrast in (12), the link between the covariance of the unbiased estimated separator $\hat{\boldsymbol{U}}$ and the covariance of the unbiased estimated cumulant $\hat{\boldsymbol{C}}_{\boldsymbol{x}}$. Using also (11), the asymptotic performances of Υ_1, Υ_2 and Υ_3 can be compared to each other, and to those obtained by averaging independent trials. Two conclusions can be drawn: first, empirical performances tend to reach theoretical limits as sample sizes tend to infinity, which justifies our approach. Second, the contrast leading to the smallest variance is Υ_1.

7. APPENDIX

7.1. Multivariate high-order complex cumulants

Cumulants are given as a function of moments in statistics text books, but only in the real case [8]. Therefore, it seems useful to report here their expressions in the complex case. Again, we consider only zero-mean complex variables that are distributed symmetrically with respect to the origin. However, they do not need to be circularly distributed. Below, cumulants are denoted with κ and moments with μ. As before, superscripts correspond to variables that are complex conjugated. We have for orders 4 and 6:

$$
\begin{aligned}
\kappa_{ijk\ell} &= \mu_{ijk\ell} - [3]\mu_{ij}\mu_{k\ell} \\
\kappa_{ijk}^{\ell} &= \mu_{ijk}^{\ell} - [3]\mu_{ij}\mu_k^{\ell} \\
\kappa_{ij}^{k\ell} &= \mu_{ij}^{k\ell} - [2]\mu_i^k\mu_j^{\ell} - \mu_{ij}\mu^{k\ell}
\end{aligned}
$$

$$
\begin{aligned}
\kappa_{ijk\ell mn} &= \mu_{ijk\ell mn} - [15]\mu_{ijk\ell}\mu_{mn} + 2[15]\mu_{ij}\mu_{k\ell}\mu_{mn} \\
\kappa_{ijk\ell m}^{n} &= \mu_{ijk\ell m}^{n} - [5]\mu_{ijk\ell}\mu_m^n - [10]\mu_{ijk}^n\mu_{\ell m} \\
&\quad +2[15]\mu_{ij}\mu_{k\ell}\mu_m^n \\
\kappa_{ijk\ell}^{mn} &= \mu_{ijk\ell}^{mn} - \mu_{ijk\ell}\mu^{mn} - [8]\mu_{ijk}^m\mu_\ell^n - [6]\mu_{ij}^{mn}\mu_{k\ell} \\
&\quad +[6]\mu_{ij}\mu_{k\ell}\mu^{mn} + 2[12]\mu_{ij}\mu_k^m\mu_\ell^n \\
\kappa_{ijk}^{\ell mn} &= \mu_{ijk}^{\ell mn} - [3]\mu_{ijk}^{\ell}\mu^{mn} - [9]\mu_{ij}^{\ell m}\mu_k^n - [3]\mu_{ij}\mu_k^{\ell mn} \\
&\quad +2[9]\mu_{ij}\mu_k^{\ell}\mu^{mn} + 2[6]\mu_i^{\ell}\mu_j^m\mu_k^n
\end{aligned}
$$

and eventually for order 8:

$$
\begin{aligned}
\kappa_{ijk\ell mnpq} &= \mu_{ijk\ell mnpq} - [28]\mu_{ijk\ell mn}\mu_{pq} - [35]\mu_{ijk\ell}\mu_{mnpq} \\
&\quad +2[210]\mu_{ijk\ell}\mu_{mn}\mu_{pq} - 6[105]\mu_{ij}\mu_{k\ell}\mu_{mn}\mu_{pq} \\
\kappa_{ijk\ell mnp}^{q} &= \mu_{ijk\ell mnp}^{q} - [7]\mu_{ijk\ell mn}\mu_p^q - [21]\mu_{ijk\ell m}^q\mu_{np} \\
&\quad -[35]\mu_{ijk\ell}\mu_{mnp}^q + 2[105]\mu_{ijk}^q\mu_{\ell m}\mu_{np} \\
&\quad +2[105]\mu_{ijk\ell}\mu_{mn}\mu_p^q - 6[105]\mu_{ij}\mu_{k\ell}\mu_{mn}\mu_p^q \\
\kappa_{ijk\ell mn}^{pq} &= \mu_{ijk\ell mn}^{pq} - \mu_{ijk\ell mn}\mu^{pq} - [12]\mu_{ijk\ell m}^p\mu_n^q \\
&\quad -[15]\mu_{ijk\ell}^{pq}\mu_{mn} - [15]\mu_{ijk\ell}\mu_{mn}^{pq} - [20]\mu_{ijk}^p\mu_{\ell mn}^q \\
&\quad +2[15]\mu_{ijk\ell}\mu_{mn}\mu^{pq} + 2[30]\mu_{ijk\ell}\mu_m^p\mu_n^q \\
&\quad +2[120]\mu_{ijk}^p\mu_{\ell m}\mu_n^q + 2[45]\mu_{ij}^{pq}\mu_{k\ell}\mu_{mn} \\
&\quad -6[15]\mu_{ij}\mu_{k\ell}\mu_{mn}\mu^{pq} - 6[90]\mu_{ij}\mu_{k\ell}\mu_m^p\mu_n^q
\end{aligned}
$$

$$
\begin{aligned}
\kappa_{ijk\ell m}^{npq} &= \mu_{ijk\ell m}^{npq} - [3]\mu_{ijk\ell m}^{n}\mu^{pq} - [15]\mu_{ijk\ell}^{np}\mu_{m}^{q} \\
&-[10]\mu_{ijk}^{npq}\mu_{\ell m} - [5]\mu_{ijk\ell}\mu_{m}^{npq} - [30]\mu_{ijk}^{n}\mu_{\ell m}^{pq} \\
&+2[15]\mu_{ijk\ell}\mu_{m}^{n}\mu^{pq} + 2[30]\mu_{ijk}^{n}\mu_{\ell m}\mu^{pq} \\
&+2[60]\mu_{ijk}^{n}\mu_{\ell}^{p}\mu_{m}^{q} + 2[90]\mu_{ij}^{np}\mu_{k\ell}\mu_{m}^{q} \\
&+2[15]\mu_{i}^{npq}\mu_{jk}\mu_{\ell m} \\
&-6[45]\mu_{ij}\mu_{k\ell}\mu_{m}^{n}\mu^{pq} - 6[60]\mu_{ij}\mu_{k}^{n}\mu_{\ell}^{p}\mu_{m}^{q} \\
\kappa_{ijk\ell}^{mnpq} &= \mu_{ijk\ell}^{mnpq} - [6]\mu_{ijk\ell}^{mn}\mu^{pq} - [6]\mu_{ij}^{mnpq}\mu_{k\ell} - [16]\mu_{ijk}^{mnp}\mu_{\ell}^{q} \\
&-[16]\mu_{ijk}^{m}\mu_{\ell}^{npq} - [18]\mu_{ij}^{mn}\mu_{k\ell}^{pq} - \mu_{ijk\ell}\mu^{mnpq} \\
&+2[\bar{2}]([3]\mu_{ijk\ell}\mu^{mn}\mu^{pq} + [48]\mu_{ijk}^{m}\mu_{\ell}^{n}\mu^{pq}) \\
&+2[36]\mu_{ij}^{mn}\mu_{k\ell}\mu^{pq} + 2[72]\mu_{ij}^{mn}\mu_{k}^{p}\mu_{\ell}^{q} \\
&-6[72]\mu_{ij}\mu_{k}^{m}\mu_{\ell}^{n}\mu^{pq} - 6[24]\mu_{i}^{m}\mu_{j}^{n}\mu_{k}^{p}\mu_{\ell}^{q} \\
&-6[9]\mu_{ij}\mu_{k\ell}\mu^{mn}\mu^{pq}
\end{aligned}
$$

7.2. Expression of second order differentials

For contrast Υ_1, we give below the expressions of the coefficients of the second order differential with respect to $\boldsymbol{U}$ (omitting subscript $\boldsymbol{y}$ in $C_{i\ell,\boldsymbol{y}}^{jk}$):

for $q < r$ and $q' < r'$, $\quad \Theta_{qr}^{q'r'} =$

$$\Re\{\delta_{q'r}(C_{r'q}^{qq} - 2C_{qr}^{r'r} - C_{r'q}^{rr}) + \delta_{q'q}(C_{r'r}^{qq} + 2C_{qr}^{r'q} - C_{r'r}^{rr})$$
$$+\delta_{r'r}(C_{q'q}^{rr} + 2C_{qr}^{q'r} - C_{q'q}^{qq}) + \delta_{r'q}(C_{q'r}^{rr} - 2C_{rq}^{q'q} - C_{rq}^{qq})\}$$

for $q < r$ and $q' > r'$, $\quad \Theta_{qr}^{q'r'} =$

$$-\Im\{\delta_{r'r}(C_{q'q}^{qq} + 2C_{qr}^{q'r} - C_{q'q}^{rr}) + \delta_{q'r}(C_{r'q}^{qq} + 2C_{qr}^{r'r} - C_{r'q}^{rr})$$
$$+\delta_{r'q}(C_{q'r}^{qq} - 2C_{qr}^{q'q} - C_{q'r}^{rr}) + \delta_{q'q}(C_{r'r}^{qq} - 2C_{rq}^{r'q} - C_{r'r}^{rr})\}$$

for $q < r$ and $q' = r'$,

$$\Theta_{qr}^{q'r'} = -\Im\{(\delta_{r'r} - \delta_{r'q})(C_{rq}^{qq} + C_{qr}^{rr})\}$$

for $q > r$ and $q' < r'$, $\quad \Theta_{qr}^{q'r'} =$

$$-\Im\{\delta_{q'q}(C_{r'r}^{rr} + 2C_{rq}^{r'q} + C_{r'r}^{qq}) - \delta_{r'q}(C_{q'r}^{rr} + 2C_{rq}^{q'q} + C_{q'r}^{qq})$$
$$+\delta_{q'r}(C_{r'q}^{qq} + 2C_{qr}^{r'r} + C_{qr'}^{rr}) - \delta_{rr'}(C_{q'q}^{qq} + 2C_{qr}^{q'r} + C_{qq'}^{rr})\}$$

for $q > r$ and $q' > r'$, $\quad \Theta_{qr}^{q'r'} =$

$$-\Re\{\delta_{r'q}(C_{q'r}^{rr} - 2C_{rq}^{q'q} + C_{q'r}^{qq}) + \delta_{q'q}(C_{r'r}^{rr} - 2C_{rq}^{r'q} + C_{r'r}^{qq})$$
$$+\delta_{r'r}(C_{q'q}^{qq} - 2C_{qr}^{q'r} + C_{q'q}^{rr}) + \delta_{q'r}(C_{r'q}^{qq} - 2C_{qr}^{r'r} + C_{r'q}^{rr})\}$$

for $q > r$ and $q' = r'$,

$$\Theta_{qr}^{q'r'} = \Re\{(\delta_{r'q} - \delta_{r'r})(C_{rq}^{qq} - C_{qr}^{rr})\}$$

for $q = r$, $\quad \Theta_{qr}^{q'r'} = 0$

and those of the second order differential with respect to $\boldsymbol{C_x}$:

for $q < r$, $\quad \Theta_{qr}^{ijkl} = U_{ri}\,U_{qj}^*\,U_{qk}^*\,U_{ql} + U_{qi}\,U_{rj}^*\,U_{qk}^*\,U_{ql}$
$-U_{qi}\,U_{rj}^*\,U_{rk}^*\,U_{rl} - U_{ri}\,U_{qj}^*\,U_{rk}^*\,U_{rl}$

for $q > r$, $\quad \Theta_{qr}^{ijkl} = U_{qi}\,U_{rj}^*\,U_{rk}^*\,U_{rl} - U_{ri}\,U_{qj}^*\,U_{rk}^*\,U_{rl}$
$+U_{ri}\,U_{qj}^*\,U_{qk}^*\,U_{ql} - U_{qi}\,U_{rj}^*\,U_{qk}^*\,U_{ql}$

for $q = r$, $\quad \Theta_{qr}^{ijkl} = 0$

8. REFERENCES

[1] J. F. CARDOSO, "High-order contrasts for independent component analysis", *Neural Computation*, vol. 11, no. 1, pp. 157–192, Jan. 1999.

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

[3] P. COMON, P. CHEVALIER, "Source separation: Models, concepts, algorithms and performance", in *Unsupervised Adaptive Filtering, Vol. I, Blind Source Separation*, pp. 191–236. Wiley, 2000.

[4] P. COMON, "From Source Separation to Blind Equalization, Contrast-based Approaches", in *Int. Conf. on Image and Signal Processing (ICISP'01)*, Agadir, Morocco, May 3-5, 2001.

[5] P. COMON, P. CHEVALIER, V. CAPDEVIELLE, "Performance of contrast-based blind source separation", in *SPAWC – IEEE Sig. Proc. Advances in Wireless Com.*, Paris, April 16-18 1997, pp. 345–348.

[6] J. P. DELMAS, "Asymptotic performance of second-order algorithms", *IEEE Trans. Sig. Proc.*, 2001, submitted.

[7] A. FERREOL, P. CHEVALIER, "On the behavior of current second and higher order blind source separation methods for cyclostationary sources", *IEEE Trans. Sig. Proc.*, vol. 48, pp. 1712–1725, June 2000.

[8] P. McCULLAGH, *Tensor Methods in Statistics*, Chapman and Hall, Monographs on Statistics and Applied Probability, 1987

[9] B. OTTERSTEN M. VIBERG, T. KAILATH, "Analysis of subspace fitting and ML techniques for parameter estimation from sensor array data", *IEEE Trans. Sig. Proc.*, vol. 40, pp. 590–599, Mar. 1992.

[10] N. THIRION, E. MOREAU, "New criteria for blind signal separation", in *IEEE Workshop on Statistical Signal and Array Processing*, Pocono Manor, Pennsylvania, Usa, Aug. 14-16, 2000, pp. 344–348.