IMPROVED CONTRAST DEDICATED TO BLIND SEPARATION IN COMMUNICATIONS

Pierre Comon[1,3] *and Eric Moreau*[2]

[1]Eurecom, B.P.193, F-06904 Sophia-Antipolis Cedex, France
[2]MS-Gessy ISITV, BP 56, F-83162 La Valette du Var Cedex , France

ABSTRACT

Contrast-based separation of sources have a number of advantages. Among others, they are optimal (in a precise sense) in presence of noise of unknown statistics. Here a new contrast is proposed that allows not only to obtain the optimal solution analytically, but also yields better performances in terms of variance of the estimated mixing matrix. This contrast needs source kurtosis to have the same sign, and is thus appropriate to multichannel blind equalization in communications.

1. INTRODUCTION

Consider the following observation model

$$\mathbf{y}(t) = \sum_{k=-\infty}^{\infty} H(t-k)\mathbf{s}(k) + \mathbf{w}(t), \tag{1}$$

where $\mathbf{y}(t)$, $\mathbf{s}(t)$ and $\mathbf{w}(t)$ are $N-$dimensional random processes with finite moments up to order 4, and with values either in the real or the complex field. $\mathbf{y}(t)$ stands for the observation, $\mathbf{s}(t)$ for the sources of interest, and $\mathbf{w}(t)$ for an additive noise. Sources $s_i(t)$ are assumed to be mutually independent up to order 4. The problem is to recover the source processes $s_i(t)$ from the sole observations $y_j(t)$. It is often referred to as the "Source Separation" problem.

A large number of approaches have been proposed to this problem, including (i) adaptive algorithms [10] [11] [12] [18] [19], and closed-form solutions based on (ii) second order statistics only [9] [15] [16], (iii) higher orders only [7] [5], or (iv) making use of both second and higher orders [2] [4] [8] [13] [17]. This is of course not an exhaustive list, but additional references may be found therein.

Some years ago, the concept of contrast has been proposed for the purpose of multichannel blind identification of linear mixtures [4]. This type of approach has been shown to have an excellent behavior compared to the other approaches recently proposed [1] [3]. The principle, originally devised for narrow-band mixtures, also extends to wide-band convolutional mixtures [5].

The interest of maximizing a contrast, over cumulant matching techniques for instance, is that all components are treated in a symmetric fashion, and 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 unknown noise. In particular, when statistics (e.g. cumulants) are estimated over a short integration length, one can consider that we have to deal with the presence of a non Gaussian noise.

Similarly, interferences may be incorporated in the non Gaussian noise effect when sources are sought to be extracted. This is always the case when an unknown transformation is identified by applying a sequence of plane rotations [2] [4].

2. NOTATION AND CONTRASTS

The present contribution concerns narrow-band multichannel mixtures, so that the following complex static model will be subsequently considered:

$$\mathbf{y} = H\,\mathbf{s} + \mathbf{w}. \tag{2}$$

If z is a complex vector-valued random variable, we define its cumulant tensors (of even order) as: $g_{\mathbf{z},ijkl} = cum\{z_i, z_j^*, z_k^*, z_l\}$, $g_{\mathbf{z},ij} = cum\{z_i, z_j^*\}$, where superscript $(^*)$ denotes complex conjugation. The kurtosis of a scalar variable s_i is defined as as $\gamma_{s,i} = g_{s,iiii}/g_{s,ii}^2$. It is assumed throughout that sources are non Gaussian, and that their kurtosis have the same sign, denoted ω.

It had been suggested in [4] to find an unknown linear transform, F, aiming at inverting H, by maximizing a contrast. One of the contrasts recommended was the sum of the squares of the marginal standardized cumulants of the output, $\mathbf{z} = F\,\mathbf{y}$. It is subsequently assumed that both H and F are unitary, which means that the standardization has been performed already in a perfect manner (by using the

[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.

second-order statistics for instance). In fact, the general problem deflates to the latter if the signal covariance is exactly known. Under this condition, the functional: $\Upsilon_2(F\mathbf{y}) = \sum_{n=1}^{N} g_{\mathbf{z},nnnn}^2$. is a contrast, since $g_{\mathbf{z},ij} = \delta_{ij}$ [4]. Yet, one can also show [14] that $\Upsilon_1(F\mathbf{y}) = \omega \sum_{n=1}^{N} g_{\mathbf{z},nnnn}$ **is a contrast**, if source kurtosis $\gamma_{s,i}$ have all the same sign ω, which is the case in radio communications. More precisely, following the definition of a contrast in [4], $\forall A$ unitary and $\forall \mathbf{z}$ with independent components we have

$$\Upsilon_1(A\mathbf{z}) = \omega \sum_{j=1}^{N} g_{\mathbf{z},jjjj} \sum_{i=1}^{N} |a_{ij}|^4 \leq \Upsilon_1(\mathbf{z}). \qquad (3)$$

Indeed, since A is unitary, $\forall j$, $\sum_{i=1}^{N} |a_{ij}|^2 = 1$ and then $\forall j$, $\sum_{i=1}^{N} |a_{ij}|^4 \leq 1$. Finally because $\forall j$, $g_{\mathbf{z},jjjj} \neq 0$, the equality holds in (3) iff $\forall j$, $\sum_{i=1}^{N} |a_{ij}|^4 = 1$. Since A is unitary, this is possible only if A is a scaled permutation.

The maximization of $\omega \Upsilon_1(F\mathbf{y})$ over the group of unitary matrices is difficult, and one can either resort to some iterative algorithm, or proceed pairwise, as originally suggested in [4]. The great advantage of the latter procedure is that the 2×2 problem is much simpler, and admits an analytical solution.

However, if Υ_2 could be maximized easily in the case of real orthogonal transforms, the complex case was much more computationally demanding, though theoretically obtainable in polynomial time [4]. From this point of view, the JADE algorithm proposed in [2] was less computationnally demanding in the complex case. Contrary to Υ_2, it is shown here that the absolute maximum of $\omega\Upsilon_1$ can be found within a very reasonable number of operations, **even in the complex case**.

3. ALGORITHM

As explained in detail in [4], the determination of a unitary mixing matrix can be carried out by determining a sequence of Givens complex rotations of the form:

$$Q = \frac{1}{\sqrt{1+\theta\theta^*}} \begin{pmatrix} 1 & \theta \\ -\theta^* & 1 \end{pmatrix} \qquad (4)$$

The contrast functions Υ_i, $i = 1, 2$, are maximized at every step and increase monotically. The procedure is similar to the one originally proposed by Jacobi for diagonalizing Hermitian matrices. The main difficulty is to find at each step the absolute maximum of the contrast with respect to variable θ.

In this section, we shall thus restrict ourselves to the case of a 2-dimensional observation in presence of an arbitrary number of sources, and look for a transform aiming at inverting the mixture (4):

$$z = Q\, y. \qquad (5)$$

If all variables were real, then the absolute maximum of $\omega\Upsilon_1$ or Υ_2 would be computable only by rooting a polynomial of degree 2 and 4, respectively. But in the complex case things are more complicated, and it is a kind of miracle that the maximum of $\omega\Upsilon_1$ can be computed just by rooting a polynomial of degree 3.

3.1. Expression of contrast Υ_1

From the multilinearity property of cumulants, Υ_1 can be rewritten as a function of θ, θ^*, and (standardized) cumulants of y. If θ and θ^* could be considered as independent variables, then Υ_1 would be the ratio of two homogeneous polynomials of global degree 8.

Introduce the auxiliary variable $\xi = \theta - 1/\theta^*$. This defines a bijection between the values of θ in the unit disk, $|\theta| \leq 1$, and $\mathbb{C}$. In fact, noting $\jmath = \sqrt{-1}$, if $\theta \stackrel{\text{def}}{=} r\, e^{\jmath\varphi}$, $0 \leq r \leq 1$, and $\xi \stackrel{\text{def}}{=} R\, e^{\jmath\phi}$, $0 \leq R$, then:

$$R = \frac{1}{r} - r, \quad r = \frac{-R + \sqrt{R^2+4}}{2}.$$

It turns out that, because of symmetry properties, contrast Υ_1 can also be expressed as a rather simple function of ξ:

$$\begin{aligned}\Upsilon_1(Q\mathbf{y}) \stackrel{\text{def}}{=} \bar{\Upsilon}_1(\xi) &= (\xi\xi^* + 4)^{-1}[a\,\xi\xi^* + 4\,\Re\{b\, e^{\jmath\beta}\,\xi\} \\ &\quad + 4\,\Re\{d\, e^{\jmath\delta}\, \frac{\xi}{\xi^*}\} + 2\,f], \qquad (6)\end{aligned}$$

with, assuming $a, f \in \mathbb{R}$ and $b, d \in \mathbb{R}^+$:

$$\begin{aligned} a &\stackrel{\text{def}}{=} g_{1111} + g_{2222}, \\ b\, e^{\jmath\beta} &\stackrel{\text{def}}{=} g_{2122} - g_{1112}, \\ d\, e^{\jmath\delta} &\stackrel{\text{def}}{=} g_{2112}, \\ f &\stackrel{\text{def}}{=} 4\, g_{1122} + g_{1111} + g_{2222}. \end{aligned}$$

For convenience, also denote $\psi \stackrel{\text{def}}{=} \phi + \beta$ and $\mu \stackrel{\text{def}}{=} \delta - 2\,\beta$. Then a simpler expression of the contrast is:

$$\bar{\Upsilon}_1(R\,e^{\jmath\phi}) = \frac{a\,R^2 + 4b\,R\cos\psi + 4d\cos(2\psi+\mu) + 2\,f}{R^2+4} \qquad (7)$$

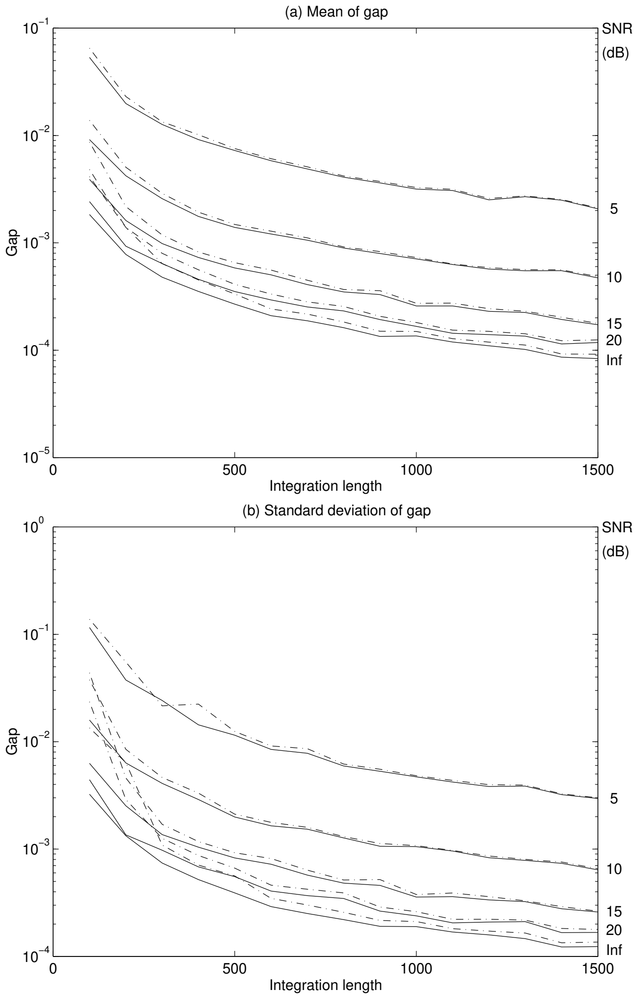
When they are not at infinity, stationary points in (R, ϕ) cancel the two partial derivatives:

$$b\cos\psi\,(R^2-4) + (2d\cos(2\psi+\mu) + f - 2a)\,R = 0 \qquad (8)$$
$$b\sin\psi\,R + 2d\sin(2\psi+\mu) = 0 \qquad (9)$$

At infinity, $\bar{\Upsilon}_1(\infty) = a$, and the solution is just $\theta = 0$, that is the identity rotation.

3.2. Computation of the phase of ξ

Generically $\beta \neq 0$, and R can be eliminated between (8) and (9). It is then very convenient to use the variable $T \stackrel{\text{def}}{=} \tan(\psi/2)$, as now demonstrated. After some



manipulations, one ends up with the fact that the optimal value of T is a root of a polynomial of degree 8, $h(T) = \sum_{i=0}^{8} m_i\, T^i$, whose coefficients are:

$$\begin{aligned} m_8 &= (-1+\cos^2\mu)d^2 \\ m_7 &= (-f+2a)d\sin\mu + 6d^2\sin\mu\,\cos\mu \\ m_6 &= (6-14\cos^2\mu)d^2 + (-8a+4f)d\cos\mu + 4b^2 \\ m_5 &= (-10a+5f)d\sin\mu - 14d^2\sin\mu\,\cos\mu \\ m_4 &= 0,\ m_3 = m_5,\ m_2 = -m_6,\ m_1 = m_7,\ m_0 = -m_8. \end{aligned}$$

Candidates for the phase of θ are consequently of the form $\phi_i = 2\arctan(T_i) - \beta$. This would be of limited interest if it was not possible to root $h(T)$ in a simple fashion. But it turns out precisely that $h(T) = T^3(T^2+1)\,p(T-\frac{1}{T})$, where $p(X)$ is the following **cubic**:

$$p(X) \stackrel{\text{def}}{=} m_8X^3 + m_7X^2 + (m_6 - 2\,m_8)\,X + m_5 + m_7 \quad (10)$$

It is indeed a **great surprise**. The polynomial $h(T)$ can thus be rooted in $\mathbb{R}$ completely analytically (*e.g.* by Cardan's method).

3.3. Computation of the modulus of ξ

Now, once we have computed a reduced number of candidates ϕ_i, the goal is to associate a modulus, R_i, with each of them. This is easy since we can use either (8) or (9), depending on whether $\sin\psi$ is small or not. If $\sin\psi$ is small, it is more suitable for conditioning reasons to utilize (8). Yet, this equation admits at most one positive real root, given by:

$$R = [-\sigma + \sqrt{\sigma^2+16}]/2; \quad \sigma = \frac{f - 2a + 2d\cos(2\psi+\mu)}{b\,\cos\psi}$$

3.4. Genericity

It has been assumed that $b \neq 0$ and $\sin\mu \neq 0$. This is true in the generic case. If one of these quantity vanishes, one can show that the solution in ϕ becomes much simpler. For reasons of space, this is not reported here, but the complete matlab code is accessible via internet[4].

4. COMPUTER RESULTS

Our performance criterion is a simplified version of the "gap" defined in [4]:

$$\begin{aligned} \epsilon(H,\hat{H}) = &\sum_i |\sum_j |D_{ij}-1|^2 + \sum_j |\sum_i |D_{ij}-1|^2 \\ &+ \sum_i |\sum_j |D_{ij}|^1 - 1| + \sum_j |\sum_i |D_{ij}|^2 - 1|. \end{aligned}$$

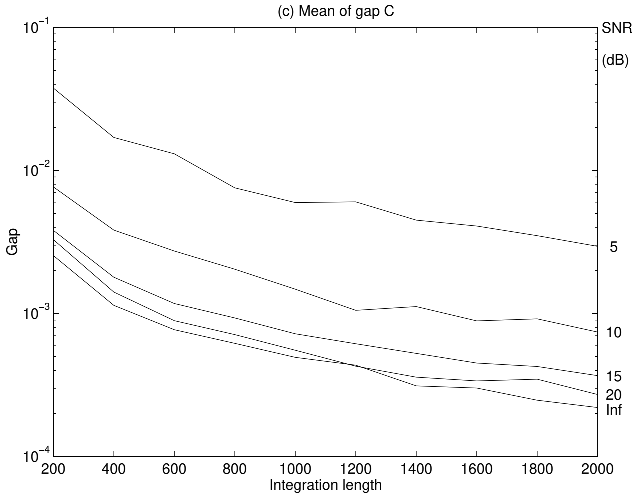
where $D \stackrel{\text{def}}{=} FH$, and $F = \hat{H}^{-1}$. Figure (a) shows the mean gap obtained with contrasts Υ_1 (solid) and Υ_2 (dashdot) in the real case, over 1000 Monte Carlo trials, for various data lengths and SNR's. Figure (b) reports the corresponding standard deviations. The sources were both PAM4 signals taking values in $\{\pm1, \pm3\}$, and the mixing matrix H was a Givens rotation of angle $\pi/4$. There is a visible advantage of Υ_1 over Υ_2.

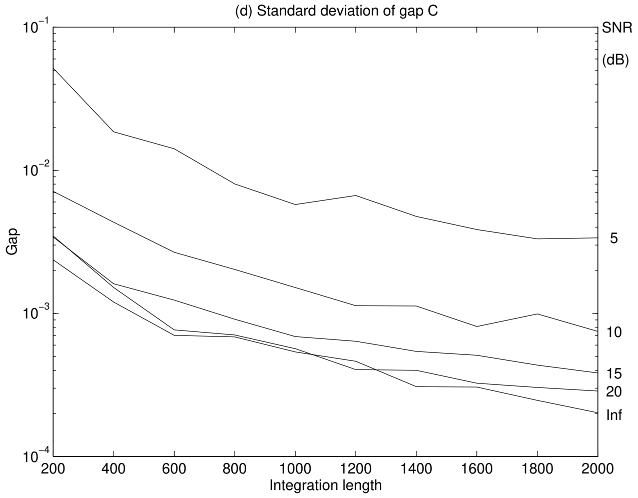
Figures (c) and (d) show the mean and standard deviation of the gap obtained with Υ_1 in the complex case, averaged over 200 Monte Carlo trials. Figure (e) is the eye diagram of a trial run with SNR 20dB and data length 500. In that case, two 16-QAM sources have been mixed by a Givens rotation of angle $\pi/5$ and phase $e^{j\pi/6}$.

5. CONCLUDING REMARKS

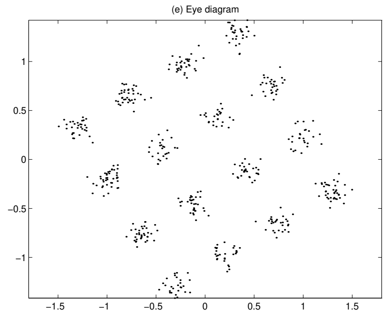
As shown by the computer simulations, the maximization of Υ_1 yields **better results** than Υ_2, for short data records and in presence of Gaussian noise. This may be very attractive in nonstationary environments. In fact, most adaptive algorithms require a fairly large

[4] Connect to http://wwwi3s.unice.fr/~comon





number of iterations before an acceptable gap between H^{-1} and F can be reached. Experiments with non Gaussian noise are being completed. Comparisons with JADE will be reported in [6].



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

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

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

[6] P. COMON, P. CHEVALIER, V. CAPDEVIELLE, "Performance of contrast-based blind source separation", Paris, to appear in *SPAWC'97*.

[7] 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.

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

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

[10] 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.

[11] 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.

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

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

[14] E. MOREAU, *Apprentissage et adaptivité, séparation auto-adaptative de sources indépendantes par un réseau de neurones*, Doctorat de l'Université de Paris-sud, Orsay, 1 fevrier 1995.

[15] 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.

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

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

[18] A. J. van der VEEN, A. PAULRAJ, "An analytical constant modulus algorithm", *IEEE Trans. SP*, vol. 44, no. 5, pp. 1136–1155, May 1996.

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