Fourth Order Criteria for Blind Sources Separation

A. Mansour, C. Jutten
INPG-TIRF
46, avenue Félix Viallet
38031 Grenoble Cedex

IEEE Trans. on Signal Processing, Vol 43, N8, August 1995

Contents

Abstract

In the problem of blind separation of sources, we define usually a function (cost, contrast, ...) and the solution is based on the search of the extremum of that function. The choice of cost function is then very important, both to have simple computation and to guarantee unicity and convergence to a good solution.
In this paper, in the case of instantaneous mixtures of 2 sources, we study two cost functions based on the fourth order cumulant and we prove theoretically and experimentally that the cross cumulant is a simple and good cost function.

1 Introduction

1.1 Problem description

The problem of blind separation of sources has been first introduced in 1985 by Hérault *et al.* [10] in the case of instantaneous mixtures, from the biological problem of movement coding. The algorithm was based on an independence test approximated by high-order cross-moments of every pair of outputs. The moments were introduced by means of products of odd non-linear functions in an adaptation rule. In similar problems, the role of non-linear functions has been studied by Féty [9]. However, limitations of the rule based on high-order moments have been proved by Comon *et al.* [7] and Sorouchyary [15] : if the probability density functions of sources are not even, the algorithm leads to spurious solutions.
So approximating independence test with high-order moments, although very simple, is not very efficient. In fact, let us come again to the definition of independence :
Two random variables u_i and u_j are independent if :

$$p(u_i, u_j) = p(u_i)p(u_j). \tag{1}$$

Denoting $\psi(u_i, u_j)$ the second characteristic function of $p(u_i, u_j)$, we can derive from the relation (1):

$$\psi(u_i, u_j) - \psi(u_i)\psi(u_j) = 0. \tag{2}$$

By computing Taylor expansion of (2), we get a polynomial equation, whose the coefficient of term of degree N is called cross-cumulant of order N. If the random variables are independent, cumulants of any order must be equal to zero. General expressions of cumulants can be found in Brillinger [2].

For zero mean signals, order-2 cross-cumulants reduces to covariance. At order 4, there are 3 cross-cumulants :

$$\begin{aligned} Cum_{13}(u_i, u_j) &= Mom_{13}(u_i, u_j) - 3Mom_{20}(u_i, u_j)Mom_{11}((u_i, u_j) \quad (3) \\ Cum_{22}(u_i, u_j) &= Mom_{22}(u_i, u_j) - Mom_{20}(u_i, u_j)Mom_{02}(u_i, u_j) \end{aligned}$$

$$-2Mom_{11}^2((u_i,u_j) \quad (4)$$
$$Cum_{31}(u_i,u_j) = Mom_{31}(u_i,u_j) - 3Mom_{02}(u_i,u_j)Mom_{11}((u_i,u_j). \quad (5)$$

Of course, using any order cross-cumulants is impossible, therefore we must try to define simpler but efficient criteria.

In a few studies, authors claim that second order statistics are good candidates. However, these assertion seems to be true only under specific conditions. In [1], estimation of parameters is driven by a correlation measurement, but after crossing a discriminatory, which can be for instance a hard limiter. Therefore, the hard limiter is strongly non-linear and consequently introduced high-order moments.

In the case of convolutive mixtures, Van Gersen *et al.* [16] use successfully second-order moments, but the mixtures are reduced to a delayed coefficient. On the contrary, for convolutive mixtures modelled by Finite Impulse Response (FIR) filter, Nguyen Thi *et al.* [13] experimentally observe that algorithms based on high-order statistics provide better performances than algorithms based on second order moments.

In most of works related to Blind Separation of Sources, criteria based on 4-order cumulants are used. For instance, Lacoume and Ruiz [11] estimate the parameters by maximising the quantity:

$$d = 1/(Cum_{13}^2(u_i,u_j) + Cum_{22}^2(u_i,u_j) + Cum_{31}^2(u_i,u_j)). \quad (6)$$

In [3], Cardoso proposed a method based on fourth-order moments, and then a refined version using fourth-order cumulants [3]. In [5], Comon addressed the problem by solving a polynomial system of equations expressing the cross-cumulants of outputs with respect to the cross-cumulants of observations.
In case of instantaneous mixtures as well as convolutive mixtures, Nguyen Thi *et al.* [13] proposed algorithms based on cancellation of fourth-order cross-cumulants Cum_{13} and Cum_{31}. However, experimental work [14] showed that for particular signal, spurious solutions are achieved, and it is possible to cancel these solutions by using the other cross-cumulant Cum_{22}.
Recently, Comon [6] propose another class of criteria, based on a contrast function, derived from the concept introduced by Donoho [8], and based on entropy measurement of independence.

It appears clearly that various criteria are currently used in the literature. The choice of the criteria is then a question of importance that we propose to address in this paper, in the restricted case of instantaneous mixtures of two sources. The choice is relevant to prove existence and unicity of solutions, to simplify algorithm, and also to propose efficient hardware implementations. In fact, there already exist hardware implementations [17] [4] of the sources separation algorithm proposed by Jutten and Hérault, but they suffer the same limitations as the algorithm.

1.2 Organisation of the paper

The paper is devided in 4 parts. In the second section, we introduced the model of mixtures and statistcs we will use. In the hired section, we study two cost functions based on the fourth order cross-cumulants. Theoretically result on the second cost function is proposed in section four. Finally, The section 5, shows an algorithm and experimental result.

2 Model equations

2.1 Mixture model

At any time t, we observe, with help of two sensors, two instantaneous mixtures $e_i(t)$ of the two zero-mean sources $x_i(t)$, assumed statistically independent. Denoting M the mixture matrix, we have:

$$\begin{pmatrix} e_1(t) \\ e_2(t) \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}, \tag{7}$$

2.2 Separation model

The separation is achieved by estimating a 2x2 matrix W satisfying $WM = PD$, where P is any permutation matrix and D is a diagonal matrix. The outputs of the matrix W are signals $s_i(t)$:

$$\begin{pmatrix} s_1(t) \\ s_2(t) \end{pmatrix} = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} \begin{pmatrix} e_1(t) \\ e_2(t) \end{pmatrix}, \tag{8}$$

The global matrix WM will be denoted $G = (g_{ij})$:

$$\begin{pmatrix} s_1(t) \\ s_2(t) \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}, \tag{9}$$

2.3 Equation of moments and cumulants

From the relation (9), we can express cross-moments and cross-cumulants of the outputs $s_1(t)$ and $s_2(t)$ with respect to the coefficients g_{ij} and cross-moments and cross-cumulants of the sources $x_1(t)$ and $x_2(t)$. Of course, the cross-moments and cross-cumulants of the sources are unknown. Let us denote:

$$\begin{aligned} Mom_{kl}(s_1, s_2) &= E[s_1^k(t) s_2^l(t)], & (10) \\ Cum_{kl}(s_1, s_2) &= Cum(s_1^k(t) s_2^l(t)). & \\ p_i &= E[x_i^2(t)], & (11) \\ \gamma_i &= E[x_i^4(t)], & (12) \\ \beta_i &= Cum(x_i^4). & (13) \end{aligned}$$

Then, up to the order 4, taking into account the statistical independence of the sources, we get the 10 following equations:

$$Mom_{01}(s_1,s_2) = Mom_{10}(s_1,s_2) = 0, \tag{14}$$

$$\begin{aligned} Mom_{11}(s_1,s_2) &= g_{11}g_{21}p_1 + g_{12}g_{22}p_2, &(15)\\ Mom_{20}(s_1,s_2) &= g_{11}^2p_1 + g_{12}^2p_2, &(16)\\ Mom_{02}(s_1,s_2) &= g_{21}^2p_1 + g_{22}^2p_2, &(17)\end{aligned}$$

$$\begin{aligned} Mom_{31}(s_1,s_2) &= g_{11}^3g_{21}\gamma_1 + 3g_{11}g_{12}(g_{11}g_{22}+g_{21}g_{12})p_1p_2 + g_{12}^3g_{22}\gamma_2, &(18)\\ Mom_{13}(s_1,s_2) &= g_{11}g_{21}^3\gamma_1 + 3g_{21}g_{22}(g_{11}g_{22}+g_{21}g_{12})p_1p_2 + g_{12}g_{22}^3\gamma_2, &(19)\\ Mom_{22}(s_1,s_2) &= g_{11}^2g_{21}^2\gamma_1 + (g_{11}^2g_{22}^2 + 4g_{11}g_{21}g_{12}g_{22} + g_{12}^2g_{21}^2)p_1p_2 \\ &\quad + g_{12}^2g_{22}^2\gamma_2 &(20)\end{aligned}$$

$$\begin{aligned} Cum_{31}(s_1,s_2) &= g_{11}^3g_{21}\beta_1 + g_{12}^3g_{22}\beta_2, &(21)\\ Cum_{13}(s_1,s_2) &= g_{11}g_{21}^3\beta_1 + g_{12}g_{22}^3\beta_2, &(22)\\ Cum_{22}(s_1,s_2) &= g_{11}^2g_{21}^2\beta_1 + g_{12}^2g_{22}^2\beta_2. &(23)\end{aligned}$$

3 Solutions of equations

In this section, we will study solutions of equation of the form $Criteria = 0$, where the $Criteria$ is a function of 4-order cumulants. In previous works [12], we used adaptation equations based on cancellation of $Cum_{31}(s_1,s_2)$ and $Cum_{13}(s_1,s_2)$. However, it has been experimentally shown [14] that the rule can give spurious solutions for specific sources, and that the spurious solutions can be removed by cancelling $Cum_{22}(s_1,s_2)$. For this reason, we will study here two criteria. The first one is $Cum_{31}^2(s_1,s_2) + Cum_{31}^2(s_1,s_2) = 0$ and the second one $Cum_{22}^2(s_1,s_2) = 0$.

3.1 Cost functions

3.1.1 First cost function

We consider the cost function:

$$Cum_{31}^2(s_1,s_2) + Cum_{31}^2(s_1,s_2). \tag{24}$$

If the outputs are statistically independent, each cumulant is equal to zero. Therefore, the minimum of the cost function, which is a sum of squares, corresponds exactly to zero. By using equations (21) and (22), we can write the cost function:

$$(g_{11}^3g_{21}\beta_1 + g_{12}^3g_{22}\beta_2)^2 + (g_{11}g_{21}^3\beta_1 + g_{12}g_{22}^3\beta_2)^2. \tag{25}$$

Equating it to zero, and denoting $\lambda = \sqrt[4]{\left|\frac{\beta_1}{\beta_2}\right|}$, we get six solutions:

$$g_{11} = g_{22} = 0 \tag{26}$$
$$g_{21} = g_{12} = 0 \tag{27}$$

$$g_{11} = g_{12} = 0 \tag{28}$$
$$g_{21} = g_{22} = 0 \tag{29}$$

$$\begin{aligned} g_{12} &= \lambda\, g_{11} \\ g_{22} &= -\lambda\, g_{21} \end{aligned} \tag{30}$$

$$\begin{aligned} g_{12} &= -\lambda\, g_{11} \\ g_{22} &= \lambda\, g_{21} \end{aligned} \tag{31}$$

Equations (26) and (27) are the theoretical solutions for the problem separation of sources. They lead to a diagonal matrix G, up to a permutation: they will give us the sources signals up a permutation and an amplitude coefficient. Equations (28) and (29) correspond to trivial solutions: one of the output signal equal to zero.
The two last solutions (30) and (31) are spurious solutions, depending on statistical properties of signals. We show, we can eliminate these solutions with a simple decorrelation of the output signals, in the section 3.2

3.1.2 Second cost function

Now we consider the cost function:

$$Cum_{22}^2(s_1, s_2) \tag{32}$$

If we equate the equation (23) to zero, using the definition of λ introduced in section 3.1.1. we may deduce two groups of solutions:

- If the sources have the same sign of kurtosis, we only have four solutions:

$$g_{11} = g_{22} = 0 \tag{33}$$
$$g_{21} = g_{12} = 0 \tag{34}$$

$$g_{11} = g_{12} = 0 \tag{35}$$
$$g_{21} = g_{22} = 0 \tag{36}$$

- If the sources have not the same sign of kurtosis, the solutions are:

$$\lambda^4 (g_{11}g_{21})^2 = (g_{12}g_{22})^2 \tag{37}$$

Solutions (33) to (36) are identical to solutions ((26) ... (29)) which have been converted in the last section.
It is clear that equation (37) is verified for the four solutions (33) to (36). However there exist others solutions depending on λ.

3.1.3 Conclusion

If the signals sources have kurtosis of different sign, then the two costs will give us the same solutions. When the sources have kurtosis of the same sign , the second cost (32) is better because it does not generate spurious solutions depending of sources statistics.

3.2 Decorrelation of the output signals

If we impose to the output signals to be uncorrelated then we will find from (15), the following relation between the coefficient of global matrix G and the power of the sources signals:

$$g_{11}g_{21}p_1 = -g_{12}g_{22}p_2. \tag{38}$$

Denoting $\mu = \frac{p_1}{p_2}$ and assuming the coefficients (w_{ii}) of the weight matrix are equal to one [1] then we find the following relation:

$$\begin{aligned}(m_{11}^2\mu + m_{12}^2)w_{21} + (m_{22}^2 + \mu m_{21}^2)w_{12} + (m_{21}m_{11}\mu + m_{22}m_{12})w_{12}w_{21} \\ +m_{21}m_{11}\mu + m_{12}m_{22} = 0.\end{aligned} \tag{39}$$

Obviously, in the plane (w_{12}, w_{21}), the relation (39) is the equation of an hyperbole, which has the following asymptotes:

$$w_{12} = -\frac{\mu m_{11}^2 + m_{12}^2}{\mu m_{11}m_{21} + m_{22}m_{12}} \tag{40}$$

$$w_{21} = -\frac{\mu m_{21}^2 + m_{22}^2}{\mu m_{11}m_{21} + m_{22}m_{12}}. \tag{41}$$

We can compute the spurious solutions of the first cost function ((30) and (31)). We obtain two points in the plane (w_{12}, w_{21}):

$$\left\{\begin{array}{lcl} w_{21} & = & -\frac{m_{22}+m_{21}\lambda}{m_{11}\lambda+m_{12}} \\ \\ w_{12} & = & \frac{m_{12}-m_{11}\lambda}{m_{21},\lambda-m_{22}}, \end{array}\right. \tag{42}$$

[1] This condition is not restrictive, because the separation is possible up to a diagonal matrix: in fact, we only can estimate ratio of the coefficients of W (w_{ij}/w_{ii}).

$$\begin{cases} w_{21} & = & \frac{m_{22}-m_{21}\lambda}{m_{11}\lambda-m_{12}} \\ & & \\ w_{12} & = & -\frac{m_{12}+m_{11}\lambda}{m_{22}+m_{21}\lambda}. \end{cases} \tag{43}$$

Replacing these values in (39), it is easy to see that the equation (39) does not hold except if the ratio of $Cum_4(x_i)$ by the square of the power of x_i is constant (i.e. $\beta_2 p_1^2 = \beta_1 p_2^2$). This condition between power and kurtosis of the signals is especially verified for Gaussian signals. But if only one signal is Gaussian, or for any signals, it is no more true. Moreover, the previous condition clearly gives us that the signal kurtosis have the same sign, and in that case we know that the second cost function has not any spurious solutions. Consequently, for the second cost function a simple whitening of outputs will be very efficient: if the source kurtosis have not the same sign, the spurious solutions (37) are cancelled by the decorrelation; if the source kurtosis have the same sign, we proved in section 3.1.2 that there is only good solutions.

4 Study of second cost function

In the previous section, we found the second cost is better then the first one, because of many reasons :

- It is simpler.
- It is perfect, theoretically [2], if the signals have the same sign of kurtosis.

We know the separation of sources is possible up to a diagonal matrix and a permutation then in the following we suppose $w_{ii} = 1$.
Now, we will prove the cancellation or the minimisation of the second cost leads to the same result, and we will prove that this cost have not a local minima if signals have the same sign of kurtosis.

4.1 Minimising or cancelling the cost

If the sources have the same sign of kurtosis, from the relation(23) it is clear that Cum_{22} have the same sign than the kurtosis of the source signals. Then the study of a cost function based on $Cum_{22}^2(s_1, s_2)$ or $Cum_{22}(s_1, s_2)$ give us the same result, because the sign of $Cum_{22}(s_1, s_2)$ does not change. If we look for the global extremum of the $Cum_{22}(s_1, s_2)$, we should solve:

$$\frac{\partial Cum_{22}(w_{12}, w_{21})}{\partial w_{12}} = 0. \tag{44}$$

Solving the equation gives an unique solution, see appendix 7.1:

[2] In fact, the result depends on the algorithms

$$w_{12min} = -\frac{\beta_1 g_{21}^2 m_{12} m_{11} + \beta_2 g_{22}^2 m_{21} m_{22}}{\beta_1 g_{21}^2 m_{12}^2 + \beta_2 g_{22}^2 m_{22}^2} \tag{45}$$

For finding the global extremum of the cost, we must solve the equation:

$$\frac{\partial Cum_{22}(w_{12min}, w_{21})}{\partial w_{21}} = 0 \tag{46}$$

Then we will find the following results:

$$g_{22} = 0 \tag{47}$$
$$g_{21} = 0 \tag{48}$$

$$g_{21} = -g_{22} \sqrt[3]{\frac{\beta_2 m_{11} m_{22}^2}{\beta_1 m_{12} m_{21}^2}} \tag{49}$$

The solutions (47), (48) with the relation (45) give us two global minima, corresponding to the theoretical solutions of the problem: the first one is $w_{21} = -m_{22}/m_{12}$, $w_{12} = -m_{11}/m_{21}$, and the second one is $w_{21} = -m_{21}/m_{11}$, $w_{12} = -m_{12}/m_{22}$. The last solution (49) corresponds to a maximum. Then we remark that the cancellation as well as the minimisation of that cost have the same result.

4.2 Local minima of the cost

We still suppose that the signal kurtosis are both positives [3], and we search if the second cost function has local extremum. To find them, we must solve:

$$\begin{aligned} \frac{\partial Cum_{22}(w_{12}, w_{21})}{\partial w_{12}} &= 0 \\ \frac{\partial Cum_{22}(w_{12}, w_{21})}{\partial w_{21}} &= 0 \end{aligned} \tag{50}$$

Then we will find the following solutions, see appendix 7.2:

$$\begin{aligned} g_{11} &= g_{12} = 0 \\ g_{21} &= g_{22} = 0 \end{aligned} \tag{51}$$

$$\begin{aligned} g_{11} &= g_{22} = 0 \\ g_{12} &= g_{21} = 0 \end{aligned} \tag{52}$$

[3] if the kurtosis are negatives then the minima of the cost ($cum_{22}(s_1, s_2)$) will be maximum and vice versa, but if the cost is $cum_{22}^2(s_1, s_2)$ we will find the same result

$$g_{21} = -g_{22} \sqrt[3]{\frac{\beta_2 m_{11} m_{22}^2}{\beta_1 m_{12} m_{21}^2}} \tag{53}$$

The solutions (51) and (52) give us identically the same result of the cost cancellation (see (33), . . . , (36)). Then we have the same conclusion.
The equation (53) is the global maximum (49). Then we remark that this cost function has not local extremum.
Then we resume the result of study in this section by:

- we can not separate the sources if they have a Gaussian distribution.
- we can not separate the source if the mixture matrix M is not regular.
- the cost has 3 extrema: two minima corresponding to the cost equal to zero, and a maximum.

5 Experimental results

In this section, we explain an algorithm of separation of sources based on our theoretical study. Our algorithm is an adaptive algorithm, which minimises the cost function $Cum_{22}^2(s_1, s_2)$.

We know that the $Cum_{22}^2(s_1, s_2)$ is a good cost function if the sources signals have the same sign of kurtosis. Then we will suppose that the sources signals are independent and they have the same sign of kurtosis. We can divide our algorithm in 4 stages:

1. Computation the output signals for the current value of weight matrix.
2. Estimation of the different moments, and the $Cum_{22}(s_1, s_2)$ on a block of samples.
3. Calculation of the partial derivatives of the cost with respect to the coefficients of the weight matrix (see appendix 7.3), and modification of the coefficients of the weight matrix (w_{ij}) by the vector $\vec{dw}$:

$$\vec{dw} = -cost.\frac{\vec{grad}(cost)}{\|\vec{grad}(cost)\|} \tag{54}$$

4. Up to a stop test, repeat at stage 1.

As stop condition of the algorithm, we proposed: "THE GREATEST i OUTPUT CROSSTALK [4] IS LESS THAN A THRESHOLD", OR "THE NUMBER

[4] The crosstalk on channel i is defined as (assuming source x_i is exacted on channel i at the convergence): $crosstalk_i = 10\log(\frac{E[(s_i - x_i)^2]}{E[x_i^2]})$

OF ITERATION IS GREATER THAN THE MAXIMUM ITERATION NUMBER [5]".

We tested the algorithm on independent identically distributed (i.i.d) signals (see fig 1.a, 1.b and 1.c) and with large number of samples. In this case, we achieve a -37 DB of crosstalk. If we choose a large absolute value for the wanted crosstalk then the algorithm will pass by a maximum absolute value of the crosstalk and it will turn again. Now we calculate the normalized output signals, corresponding to maximum absolute value of crosstalk, and to emphasize on the separation performance, we draw the error $(s_i - x_i)$ (fig 1.d), rather than x_i for which the comparison with $s_i$would not be very easy. Finally the two Figures 2.a and 2.b show the time evolution of the cost function and of the absolute value of crosstalk.

In other cases, where we have not a sufficient number of samples (for example when we have non stationary signals), we always can use the algorithm but its performances are not guaranteed: it will converge, but the residual crosstalk depends on many parameters (initial points, signal statistics, ...). However, the residual crosstalk of about -20 DB can be achieved with statistic estimated on 25 samples. Figure 3 shows experimental relation between the number of samples used to estimate the cross cumulant and the separation performance (showed by the maximum [6] of the crosstalk absolute value). We mark that every point in Fig 3 is the average of five experimental measure. In generally, We can observe that we have about -25 DB of the residual crosstalk, and we will work to explain the strange behavior of the graph start [7]

Another strategy, based on the search of the extrema proposed in section 4.1, can be used:

1. we fix all the parameters of the weight matrix except one, and we minimize the cost function with respect to that parameter.

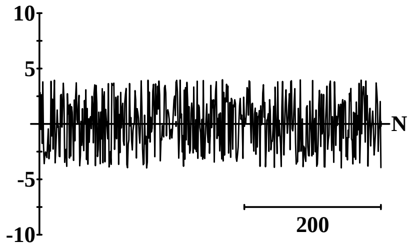
2. After that, we replace that parameter by its value (value corresponding an the minimum of the cost), and we minimize the cost with respect to the other parameter.

3. We repeat at step 1 until a stop condition.

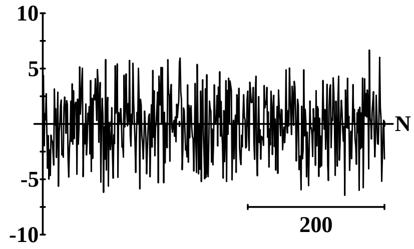
[5]Our algorithm is a cancelling adaptative algorithm, and because of the estimation error in the statistic values of the outputs signals, then we can not atteint the zero value of cost and our algorithm will not exactly converge

[6]The maximum of the crosstalk calculate for each N value with 20 iterations, where N is the sample number used to estimate the cross-cumulant.

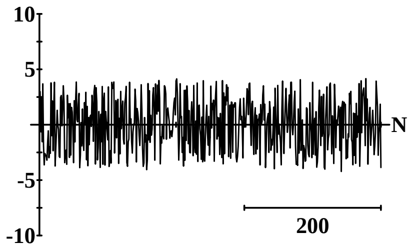
[7]Because the separation performance varies in opposite direction of the sample number.



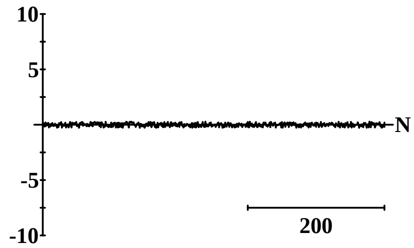
(a) Source $x_1(t)$



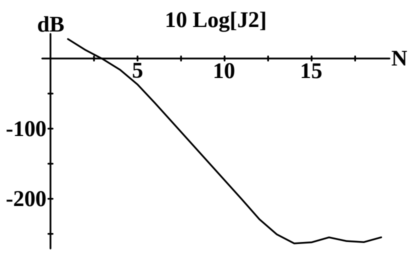
(b) mixture $e_1(t)$;



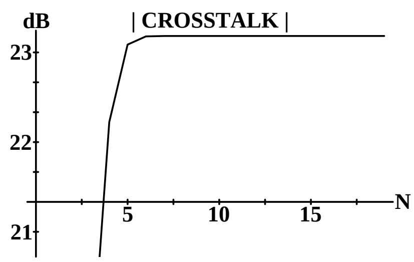
(c) output $s_1(t)$;



(d) error signal $s_1(t) - x_1(t)$;



(e) evolution of J_2;



(f) absolute value of the crosstalk;

Figure 1: Signals and performance results.

6 Conclusions

In these few papers, we study some cost functions for Blind Sources Separation, based on the fourth cross-cumulant. If we look at the P.H.D of Nguyen Thi.l [12] and the work of Xavier.O [14], Then we will remark that a cost function based on the cross-cumulant $Cum_{13}^2 + Cum_{31}^2$ is not sufficient for all signals, and if we add up the cross-cumulant Cum_{22}^2 to the old cost function we will get better the cost function. But Lacoume, J.-L. and Ruiz, P. in [11], proved that a cost function based on the square sum of the fourth cross-cumulant, is sufficient for separate our sources.
In these papers, we study the possibility to simplify the cost function. For that goal, we start our study by the comparison between two cost functions. Then we prove that a cost function based on the fourth cross-cumulant Cum_{22}^2 is a good and simpler criteria for separate the sources signals [8].

After the theoretically demonstration, we proposed an experimental algorithm based on this criteria for separate the source signals. We Tested our algorithm on i.i.d signals, and we found a good results. The performance of this algorithm is about -30 DB of crosstalk diaphonie. In the case of non stationary signals, we can not use a lot number of samples, then we need an algorithm able with a little number of samples to separate the source signals. Experimentally, we proved that even when we have a little number of samples, we can achieved a -20 DB of crosstalk diaphonie.

Finally, we proved that when the sources signals have the same sign of kurtosis, Then with a simple Whitening we able to separate our signals unless where the signals are Gaussian signals. Then in generally, the criteria based on the fourth cross-cumulant Cum_{22}^2 is a good cost function.

7 Appendix

7.1 Appendix 1

From the relations (7), (8), (9) and (23) we prove that:

$$\frac{\partial Cum_{22}(w_{12}, w_{21})}{\partial w_{12}} \geq 0.$$

From that relation we find:

$$g_{11}\beta_1 g_{21}^2 m_{21} + g_{22}^2 \beta_2 g_{12} m_{22} \geq 0.$$

Finally, from that last relation, we prove that:

$$(\beta_1 g_{21}^2 m_{21}^2 + g_{22}^2 \beta_2 m_{22}^2) w_{12} \geq -(\beta_1 g_{21}^2 m_{21} m_{11} + g_{22}^2 \beta_2 m_{22} m_{12}).$$

[8] If the sources signals have the same sign of kurtosis

If β_1 and β_2 are ≥ 0, then we will find:

$$w_{12min} \geq -\frac{\beta_1 g_{21}^2 m_{21} m_{11} + \beta_2 g_{22}^2 m_{12} m_{22}}{\beta_1 g_{21}^2 m_{12}^2 + \beta_2 g_{22}^2 m_{22}^2}. \tag{55}$$

From the value of w_{12min} in (55) we can calculate the values of global coefficients:

$$g_{11} = \frac{g_{22}^2 \beta_2 m_{22}(m_{11}m_{22} - m_{21}m_{12})}{\beta_1 g_{21}^2 m_{21}^2 + \beta_2 g_{22}^2 m_{22}^2} \tag{56}$$

$$g_{12} = \frac{g_{21}^2 \beta_1 m_{21}(m_{12}m_{21} - m_{22}m_{11})}{\beta_1 g_{21}^2 m_{21}^2 + \beta_2 g_{22}^2 m_{22}^2}. \tag{57}$$

From the relation (23) we can calculate the value of the cost :

$$Cum_{22}(w_{12min}, w_{21}) = \frac{\beta_1\beta_2(m_{11}m_{22} - m_{12}m_{21})^2 g_{22}^2 g_{21}^2}{\beta_1 g_{21}^2 m_{21}^2 + \beta_2 g_{22}^2 m_{22}^2}. \tag{58}$$

Looking for the global minimum of cost, finally we equate:

$$\frac{\partial Cum_{22}(w_{12min}, w_{21})}{\partial w_{21}} = \frac{\beta_1\beta_2 g_{21} g_{22}(m_{11}m_{22} - m_{12}m_{21})^2(\beta_1 g_{21}^3 m_{12} m_{21}^2 + \beta_2 g_{22}^3 m_{11} m_{22}^3)}{(\beta_1 g_{21}^2 m_{21}^2 + \beta_2 g_{22}^2 m_{22}^2)^2} = 0 \tag{59}$$

Then we find the solutions (47), (48) and (49).

7.2 Appendix 2

Using (23) we can prove that the system:

$$\begin{cases} \frac{\partial Cum_{22}(w_{12},w_{21})}{\partial w_{12}} & = & 0 \\ \frac{\partial Cum_{22}(w_{12},w_{21})}{\partial w_{21}} & = & 0 \end{cases} \tag{60}$$

is equivalent to:

$$\begin{cases} g_{11}\beta_1 g_{21}^2 m_{21} + g_{22}^2 \beta_2 m_{22} g_{12} & = & 0 \\ g_{21}\beta_1 g_{11}^2 m_{11} + g_{22}\beta_2 m_{12} g_{12}^2 & = & 0, \end{cases} \tag{61}$$

If we suppose that $g_{11} = 0$ (for example), then we will find:

$$g_{11} = 0 \Rightarrow \begin{cases} g_{22} & = & 0 \\ g_{12} & = & 0 \\ & & \\ m_{12} & = & m_{22} & = & 0. \end{cases} \tag{62}$$

The last solution in (62) give us a constraint in the coefficient of mixture matrix. We can eliminate that relation, because it is evident from the model (7) that

the constraint is equivalent to the case where we only have one source.
If we assume that the other coefficients of global matrix are equal to zero, we will have a similar result.
If we suppose now that (g_{ij}) and (m_{ij}) are different from zero, then we find the equation of the global maximum (49).

7.3 Appendix 3: computation of the cost derivatives.

Suppose that the two diagonal coefficients of separation matrix are equal to one ($w_{ii} = 1$), and from the relation (8), we can find these relations:

$$s_i = e_i + w_{ij} e_j \tag{63}$$

$$e_i = \frac{s_i - w_{ij} s_j}{1 - w_{ij} w_{ji}}. \tag{64}$$

From these last equations, we can calculate the partial derivatives of the output signals with respect to the coefficients of the weight matrix:

$$\begin{aligned} \frac{\partial s_i}{\partial w_{ij}} &= e_j = \frac{s_j - w_{ji} s_i}{1 - w_{ij} w_{ji}} \\ \frac{\partial s_i}{\partial w_{ji}} &= 0. \end{aligned} \tag{65}$$

Finally, if we use the relations (4) and (65), we will calculate the partial derivative of the $Cum_{22}(s_1, s_2)$ with respect to the coefficients of the weight matrix:

$$\begin{aligned} \frac{\partial Cum_{22}(s_i, sj)}{\partial w_{ij}} &= \frac{2}{1 - w_{ij} w_{ji}} [Mom_{13}(s_i, s_j) \\ &\quad - w_{ji} Mom_{22}(s_i, s_j) - 3Mom_{02}(s_i, s_j) Mom_{11}(s_i, s_j) \\ &\quad + w_{ji} Mom_{20}(s_i, s_j) Mom_{02}(s_i, s_j) + 2w_{ji} Mom_{11}^2(s_i, s_j)] \end{aligned} \tag{66}$$

References

[1] Y. Barness, J. Carlin, and M. Steinberger. Bootstrapping adaptive cross pole cancelers for satellite communications. In *International Conference on Communication*, Philadelphia, Pennsylvania, June 1982.

[2] D. R. Brillinger. *Time series data analysis and theory.* Holden Day, 1981.

[3] J. F. Cardoso. Blind identification of independent signals. In *Workshop on Higher-Order Spectral Analysis*, Vail (CO), USA, June 1989.

[4] M. H. Cohen and A. G. Andreou. Current-mode subthreshold MOS implementation of the Hérault-Jutten autoadaptative network. *IEEE Journal on Solide state Circuits*, 26(5):714–727, May 1992.

[5] P. Comon. Separation of sources using higher-order cumulants. In *SPIE Vol. 1152 Advanced Algorithms and Architectures for Signal Processing IV*, San Diego (CA), USA, August 8-10 1989.

[6] P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3):287–314, April 1994.

[7] P. Comon, C. Jutten, and J. Hérault. Blind separation of sources, Part II: Statement problem. *Signal Processing*, 24(1):11–20, November 1991.

[8] D. L. Donoho. On minimum entropy deconvolution. In *Applied Time Series Analysis II*, pages 565–608, Tulsa, 1980.

[9] L. Féty. *Méthodes de traitement d'antenne adaptées aux radiocommunications*. PhD thesis, ENST Paris, 1988.

[10] J. Hérault, C. Jutten, and B. Ans. Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. In *Actes du Xème colloque GRETSI*, pages 1017–1022, Nice, France, 20-24, May 1985.

[11] J. L. Lacoume and P. Ruiz. Sources identification: A solution based on cumulants. In *IEEE ASSP Workshop V*, Mineapolis, USA, August 1988.

[12] L. Nguyen Thi. *Séparation aveugle de sources à large bande dans un mélange convolutif.* PhD thesis, INP Grenoble, January 1993.

[13] L. Nguyen Thi, C. Jutten, and J. Caelen. Speech enhancement: Analysis and comparison of methods in various real situations. In J. Vandewalle, R. Boite, M. Moonen, and A. Oosterlinck, editors, *Signal Processing VI, Theories and Applications (EUSIPCO'92)*, pages 303–306, Brussels, Belgium, August 1992. Elsevier.

[14] X. Oliva Galvan. Blind separation of sources: Some adaptative algorithms. Technical report, Lab CEPHAG in Grenoble, April 1993.

[15] E. Sorouchyari. Blind separation of sources, Part III: Stability analysis. *Signal Processing*, 24(1):21–29, 1991.

[16] S. Van Gerven and D. Van Compernolle. Feedforward and feedback in a symmetric adaptive noise canceller: stability analysis in a simplified case. In J. Vandewalle, R. Boite, M. Moonen, and A. Oosterlinck, editors, *Signal Processing VI, Theories and Applications (EUSIPCO'92)*, pages 1081–1084, Brussels, Belgium, August 1992. Elsevier.

[17] E. Vittoz and X. Arreguit. CMOS integration of Hérault-Jutten cells for separation of sources. In *Workshop on Analog VLSI and Neural Systems*, Portland, USA, May 1989.