

Blind source separation in convolutive mixtures: a hybrid approach for colored sources

Frédéric ABRARD and Yannick DEVILLE

Laboratoire d'Acoustique Métrologie Instrumentation, Université Paul Sabatier,
118 route de Narbonne, 31062 Toulouse Cedex, France.
`abrard@cict.fr`, `ydeville@cict.fr`

Abstract. This paper deals with the blind source separation problem in convolutive mixtures when the sources are colored processes. In this case, classical methods first extract the innovation processes of the sources and then color them, which yields two successive filter approximations. On the contrary, we propose here a new concept allowing to directly extract estimates of the colored sources in one step.

1 Introduction

In this paper, we focus on the 2-source to 2-sensor blind source separation problem, as shown in Fig. 1. The two colored sources X_1 and X_2 result from two white non-Gaussian signals P_1 and P_2 , colored by two ARMA filters D_1 and D_2 . They are transferred through a mixing matrix which consists of filters $A_{ij}(z)$. To simplify the notations, all the filters $D_i(z)$ and $A_{ij}(z)$ are supposed to be MA. According to this configuration, we can write the two measured signals Y_1

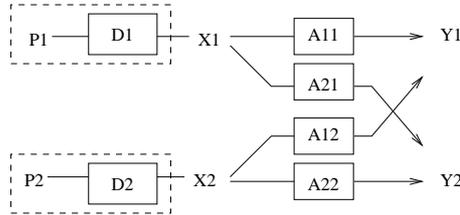


Fig. 1. Source generation and mixing matrix.

and Y_2 as:

$$\begin{aligned} Y_1(t) &= A_{11}(t) * X_1(t) + A_{12}(t) * X_2(t) \\ Y_2(t) &= A_{21}(t) * X_1(t) + A_{22}(t) * X_2(t) \end{aligned} \quad (1)$$

where $*$ denotes the convolution operator.

Blind source separation (BSS) consists in estimating the two sources X_1 and X_2 from the two observed signals Y_1 and Y_2 without any knowledge of

the mixing filters or sources properties, except that the sources are assumed to be independent. It is well known that classical methods such as kurtosis maximization allow to extract the white process P_1 or P_2 even if the sources are colored [1], [2], [3], [4].

This may easily be seen by deriving from the \mathcal{Z} -transform of (1) that:

$$[Y(z)] = [A(z)][D(z)][P(z)] \quad (2)$$

which becomes:

$$[Y(z)] = [V(z)][P(z)] \quad (3)$$

We see in (3) that there is no distinction between the coloration and the propagation stages. Moreover, each linear combination of several filtered i.i.d signals makes the result closer to the Gaussian density than the signal whose normalized kurtosis has the highest absolute value among all the original signals thus filtered and combined [5], [2]. Therefore, maximizing the absolute value of the output normalized kurtosis by means of the separation system used in [3] will directly give the innovation process P_i whose normalized kurtosis has the highest absolute value.

Some authors [3], [4] worked on this problem and proposed a solution which consists in first estimating the white process and then building a post-processing filter which will try to artificially color the signal in order to restore the contribution of one source on one sensor. This method has the main drawback to consist of two steps and to combine two successive approximations: The first inverse filters and the coloration filter.

We here propose a new approach which directly estimates the different contributions of the two sources on one sensor in one step.

2 Preliminary version

To extract the sources we need a separation system which will recompose the two observations in order to cancel one source. This could be performed by the direct separation system showed in figure 2. We use higher-order statistics in our

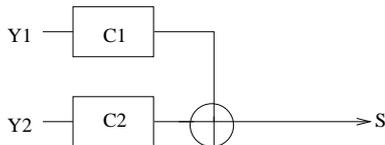


Fig. 2. Preliminary separation system.

separation criterion. We choose to maximize the absolute value of the normalized

kurtosis of the output S of the separation system. This parameter is expressed as:

$$|k(S)| = \left| \frac{\text{cum}_4(S)}{[\text{cum}_2(S)]^2} \right| \quad (4)$$

It is well known that by maximizing the expression in (4) vs the coefficients of C_1 and C_2 we extract a white process P_i up to a scale factor and delay [3]. The output is then:

$$S(z) = \alpha_i z^{-p_i} P_i(z) \quad (5)$$

with α_i constant, z^{-p_i} delay operator and i source index.

Our idea came from the observation that maximizing the kurtosis leads to:

- the separation of two non-Gaussian signals, provided they are statistically independent at order four,
- the fourth-order whiteness of the extracted signals.

We here aim at avoiding the above whitening effect in order to extract the colored sources (or the resulting observed signals, which each consist of the contribution of only one source on a sensor). According to this idea, we add another filter, whose transfer function is denoted B , which will artificially whiten the output S . The maximization of the kurtosis of the resulting whitened signal U is then intended to restore $U = P_1$ or $U = P_2$, which allows S to become equal to X_1 or X_2 , or to the observed versions of these signals. The new separation system thus obtained is showed in Fig. 3.

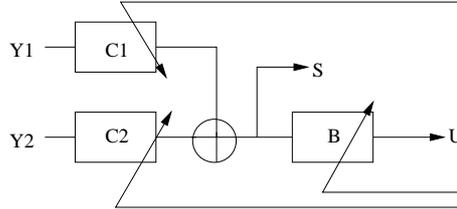


Fig. 3. Separation system including a whitening filter.

We adapt the three filters C_1 , C_2 and B by maximizing $|k(U)|$ vs the respective coefficients of the different filters.

But this system is not constrained enough to guarantee that it converges to filter values such that its output S extracts a source signal itself, or its observed version contained in a sensor signal. We can show this by writing in \mathcal{Z} -domain:

$$S(z) = (C_1(z).A_{11}(z) + C_2(z).A_{21}(z)).X_1(z) + (C_1(z).A_{12}(z) + C_2(z).A_{22}(z)).X_2(z) \quad (6)$$

Supposing that $S(z)$ contains only $X_1(z)$, i.e. the maximization of the kurtosis cancelled the contribution from $X_2(z)$, then:

$$C_1(z).A_{12}(z) + C_2(z).A_{22}(z) = 0 \quad (7)$$

which means

$$C_2(z) = -\frac{C_1(z).A_{12}(z)}{A_{22}(z)} \quad (8)$$

Combining this filter value with (6) yields:

$$S = \left(C_1(z).(A_{11}(z) - \frac{A_{12}(z).A_{21}(z)}{A_{22}(z)}) \right) .X_1(z) \quad (9)$$

There is no other constraint here to fix the value of $C_1(z)$ since, in practice the whiteness effect is achieved by the filter $B(z)$. S is then a version of X_1 transferred through an arbitrary time varying filter, which yields an uncontrolled frequency distortion for this source. On the contrary, we want this filter to take a specific value, such that S is equal to X_1 itself or to the contribution of X_1 in one of the sensor signals. To this end, we now introduce a modified version of this approach.

3 Proposed approach

3.1 First idea

In the particular case when the signal S only contains a contribution from one source transferred through an unknown filter H , we can write $S(t) = H(t)*X_1(t)$. Combining this expression with (1) yields:

$$Y_1(t) - S(t) = (A_{11}(t) - H(t)) * X_1(t) + A_{12}(t) * X_2(t) \quad (10)$$

and therefore:

$$E [(Y_1(t) - S(t))^2] = E [(A_{11}(t) - H(t)) * X_1(t)]^2 + E [(A_{12}(t) * X_2(t))^2] \quad (11)$$

The expression in (11) is minimized when $H(t) = A_{11}(t)$. In this case, we extract $S(t) = A_{11} * X_1(t)$, which is the contribution of the source X_1 measured on sensor 1. In other words, we thus get an adequately colored version of the innovation process of this source.

We then choose to adapt the coefficients of C_1 and C_2 by maximizing the new criterion defined as:

$$F_{C_1, C_2} = |k(U)| - \lambda E [(Y_1(t) - S(t))^2] \quad (12)$$

with $\lambda > 0$, while the whitening filter B is adapted so as to maximize $|k(U)|$.

This aims at combining two effects i.e:

- separating the sources by kurtosis maximization,
- extracting the signal $S(t) = A_{11} * X_1(t)$ by minimizing $E [(Y_1(t) - S(t))^2]$.
Note that the other source may then easily be derived, as $Y_1(t) - S(t) = A_{12}(t) * X_2(t)$.

But in this intuitive approach we considered that the output S only contains one source and that whatever happens no other source could appear, which is not guaranteed here.

When both sources are present in S , the function $E [(Y_1(t) - S(t))^2]$ is clearly minimized vs C_1 and C_2 when $S = Y_1$, which may be far from what we expected before. We see here the danger to reach the solution $S = Y_1$. However, in this case the value of the kurtosis $|k(B * Y_1)|$ is not maximum because Y_1 still includes the two sources.

Therefore, we now provide a more general theoretical analysis, which shows in which conditions the proposed approach works.

3.2 Theoretical considerations

Considering a point defined by fixed values of (C_1, C_2, B) , for any signal S the condition for our cost function to be lower at that point than at the desired convergence point may be expressed as:

$$|k(B * S)| - \lambda E [(Y_1 - S)^2] < |k(P_1)| - \lambda E [(A_{12} * X_2)^2] \quad (13)$$

Then we have two different cases:

1. If $E [(A_{12} * X_2)^2] - E [(Y_1 - S)^2] < 0$ then (13) yields the following constraint on λ :

$$\lambda > \frac{|k(P_1)| - |k(B * S)|}{E [(A_{12} * X_2)^2] - E [(Y_1 - S)^2]} \quad (14)$$

The numerator $|k(P_1)| - |k(B * S)|$ is always positive, so the expression (14) is true for all $\lambda > 0$.

2. If $E [(A_{12} * X_2)^2] - E [(Y_1 - S)^2] > 0$ then (13) yields the constraint:

$$\lambda < \frac{|k(P_1)| - |k(B * S)|}{E [(A_{12} * X_2)^2] - E [(Y_1 - S)^2]} \quad (15)$$

The condition (15) means that, for some positive values of λ , there may exist points where the considered cost function takes higher values than at the desired convergence point. The latter point then does not correspond to the maximum of this function. This results from the hybrid nature of this cost function (12): whereas its term $|k(U)|$ alone is maximized exactly at the desired convergence point, the additional term $-\lambda E [(Y_1(t) - S(t))^2]$ that we introduced in this function shifts the position of the global maximum of the overall resulting function. It may be shown that the value of λ controls the magnitude of this shift, and therefore the accuracy of the extraction of the source signals. The behavior of the proposed separation system then depends as follows on λ :

1. For $\lambda = 0$ the maximum of the cost function coincides exactly with the desired convergence point. However, $\lambda = 0$ is not acceptable because the system is then underdetermined, as explained above.
2. For a very small λ , the above underdetermination disappears but the system is still ill-conditioned.
3. For intermediate values of λ , the system becomes well-conditioned and the maximum of its cost function is still close to the desired convergence point. This is the range of λ to be used.
4. If λ is further increased, the position of the maximum may shift significantly from the desired position. Especially, if λ is very high, (12) shows that the cost function has almost the same behavior as $-\lambda E [(Y_1(t) - S(t))^2]$, so that its maximum almost corresponds to $S(t) = Y_1(t)$.

The proposed approach thus provides an original alternative to classical solutions: whereas the latter methods perform an approximation by using an artificial coloring step, the solution proposed in this paper avoids this approximation by extracting the colored sources directly, but the specific cost function defined above entails another type of approximation, i.e. in the position of the convergence point.

A quantitative assessment of the above-defined ranges of values of λ and of the accuracy of the proposed approach is provided in the next section.

4 Experimental results

We here present several results obtained with the following settings:

Two binary white processes with $|k(P1)| = 1.9987$ and $|k(P2)| = 2$. The impulse responses of the coloration filters correspond to the following coefficient arrays: $D_1 = [1 \ 0.8 \ 0.5 \ 0.3 \ 0.1]$ and $D_2 = [1 \ 0.9 \ 0.7 \ 0.45 \ 0.4]$. Similarly, the mixing matrix is built as:

$$\begin{aligned} A_{11} &= [0.5 \ -0.9 \ 0.4] & A_{12} &= [-0.6 \ 0.6 \ 0.3] \\ A_{21} &= [0.4 \ -0.7 \ 0.3] & A_{22} &= [0.7 \ 0.9 \ 0.5] \end{aligned} \quad (16)$$

The length of the separating and whitening filters is set to 9.

We adapt the coefficients with an extended version of the modified gradient ascent algorithm described in [6] in the case of linear instantaneous mixtures. This algorithm here reads:

$$\begin{aligned} B^j(n+1) &= B^j(n) + \mu \tanh \left(\text{sign}[k(U)] \frac{\partial k(U)}{\partial B^j} \right) \\ C_1^j(n+1) &= C_1^j(n) + \mu \tanh \left(\text{sign}[k(U)] \frac{\partial k(U)}{\partial C_1^j} - \frac{\partial \lambda E[(Y_1 - S)^2]}{\partial C_1^j} \right) \\ C_2^j(n+1) &= C_2^j(n) + \mu \tanh \left(\text{sign}[k(U)] \frac{\partial k(U)}{\partial C_2^j} - \frac{\partial \lambda E[(Y_1 - S)^2]}{\partial C_2^j} \right) \end{aligned} \quad (17)$$

Denoting \hat{e}_l the l^{th} estimated source and e_l the contribution of this source on

the first sensor, we evaluate the accuracy of the separation by the criterion:

$$E(\hat{e}_l) = \frac{\sum_n |\hat{e}_l(n) - e_l(n)|^2}{\sum_n |e_l(n)|^2} \quad (18)$$

The following results are obtained after 500 iterations, where each iteration includes an update of all three filters. The cumulants values are computed on a 300-sample data window. With several values of λ we get:

value of λ	source restored	accuracy $E(\hat{e}_1)$	accuracy $E(\hat{e}_2)$
3	no	0.779	0.4118
2.5	yes (after erratic convergence)	0.2407	0.1272
2	yes	0.1587	0.0839
1	yes	0.0817	0.0432
0.85	yes	0.0824	0.0435
0.5	yes	0.093	0.0486
0.2	yes	0.1276	0.0674
.05	no, without any convergence	2.2639	1.1966

We see that, as predicted in the previous sections, with λ over 2.5 we do not extract a source. Moreover, our experimental tests show that for all $\lambda > 2.5$ the output we get is equal to Y_1 . Besides, as we mentioned, when λ is below a certain value we do not extract the right solution nor Y_1 : the filters then keep on evolving but do not converge. Fig. 4 to 7 show the results obtained with $\lambda = 0.85$.

5 Conclusions and future work

In this paper, we introduced a new approach for extracting colored sources from their convolutive mixtures. This approach consists of a separating structure and adaptation algorithms for its filters. This structure makes it possible to directly extract a colored source, whereas classical methods use a two-step approach including an artificial "recoloring" step, which is likely to yield only an approximate extraction of the sources. The proposed adaptation algorithms yield specific convergence properties, which entail another type of (slight) approximation. In our future investigations, we will aim at improving these convergence properties by developing modified versions of these algorithms. Anyway, the numerical results reported above show that this first version of our approach already succeeds in separating the considered signals.

References

1. J. A. CADZOW. Blind deconvolution via cumulant extrema. *IEEE signal processing magazine*, pages 24–42, May 1996.
2. J. S. SEO. *Blind fault detection and source identification using higher order statistics for impacting systems*. PhD thesis, University of Southampton, Faculty of Engineering and Applied Science, Institute of Sound and Vibration Research, November 2000.

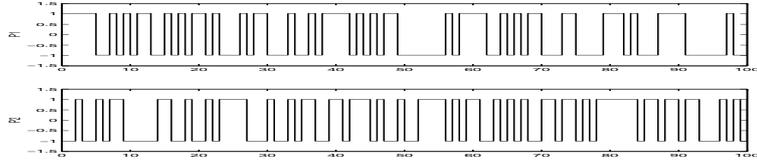


Fig. 4. White processes P_1 (upper plot) and P_2 (lower plot).

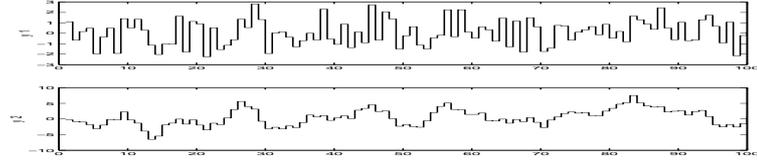


Fig. 5. Sensor signals Y_1 (upper plot) and Y_2 (lower plot).

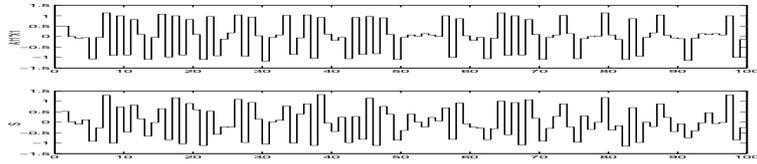


Fig. 6. Signals to be compared: $A_{11} * X_1$ (upper plot) and S (lower plot).

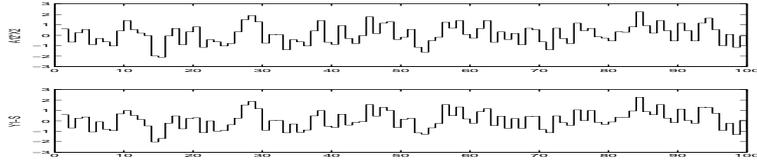


Fig. 7. Comparison between $A_{12} * X_2$ (upper plot) and $Y_1 - S$ (lower plot).

3. J. K. TUGNAIT. Identification and deconvolution of multichannel linear non-gaussian processes using higher order statistics and inverse filter criteria. *IEEE Trans. on Signal Processing*, 45(3):658–672, March 1997.
4. C. SIMON PH. LOUBATON C. VIGNAIT C. JUTTEN G. D'URSO. Blind source separation of convolutive mixtures by maximization of fourth-order cumulants: the non-iid case. *Proc. Asilomar*, November 98.
5. D. DONOHO. On minimum entropy deconvolution. In *Applied Time series Analysis II* ed. D. Findley, Academic Press, New York, pages 556–608, 1981.
6. F. ABRARD Y. DEVILLE M. BENALI. Numerical and analytical solution to the differential source separation problem. In *Proceedings of EUSIPCO*, Tampere, Finland, 2000.