Blind separation of linear-quadratic mixtures of real sources using a recurrent structure

Shahram Hosseini and Yannick Deville

Université Paul Sabatier, Laboratoire d'Acoustique, Métrologie, Instrumentation,
Bat. 3R1B2, 118 route de Narbonne, 31062 Toulouse Cedex, France
hosseini@cict.fr, ydeville@cict.fr

Abstract. In this paper, we propose an approach for separating linear-quadratic mixtures of independent real sources. The method is based on parametric identification of a recurrent separating structure by means of an adaptive algorithm which uses the higher-order statistics of the outputs of this structure. We study the local stability of the recurrent structure and show experimentally that when it is stable at the separating point, it succeeds in separating the sources.

1 Introduction

Linear blind source separation (BSS) has been widely studied in the recent years and applied successfully in various domains (see for example [1] or [2] for review and more references). Recently, a new line of research has emerged that focuses on nonlinear mixtures [3]-[7]. It is well known that [3], [4], in the general case, the independence hypothesis is not sufficient for separating nonlinear mixtures because of the very large indeterminacies which make the nonlinear BSS problem ill-posed. A natural idea for reducing the indeterminacies is to constrain the structures of mixing and separating models [8].

In this paper, we study a simple linear-quadratic mixture model which may be considered as the simplest (nonlinear) version of a general polynomial model. The developed ideas can be however generalized to more complicated polynomial models which represent a Taylor series approximation of the general nonlinear models.

2 Problem statement

Suppose u_1 and u_2 are two independent random variables. Given the following nonlinear instantaneous mixture model

$$x_1 = a_{11}u_1 + a_{12}u_2 + b_1 u_1 u_2$$
$$x_2 = a_{21}u_1 + a_{22}u_2 + b_2 u_1 u_2$$

we would like to estimate u_1 and u_2 up to a permutation and a scaling factor (and possibly an additive constant). For simplicity, let's denote $s_1 = a_{11}u_1$ and $s_2 = a_{22}u_2$. The above equations can be rewritten as

$$\begin{aligned} x_1 &= s_1 - l_1 s_2 - q_1 s_1 s_2 \\ x_2 &= s_2 - l_2 s_1 - q_2 s_1 s_2 \end{aligned} \tag{1}$$

in which $l_1 = -a_{12}/a_{22}$ and $l_2 = -a_{21}/a_{11}$ represent the linear contributions of the sources in the mixture, and $q_1 = -b_1/(a_{11}a_{22})$ and $q_2 = -b_2/(a_{11}a_{22})$ represent the quadratic contributions. The negative signs are chosen for simplifying the notations of the separating structure in the following sections. Note that if $q_1 = q_2 = 0$, the model is reduced to the widely studied linear model.

A more general form of the model (1), containing the additional terms s_1^2 and s_2^2, has been studied by a few authors [9], [10]. The main specification of these works is that they consider only a special class of sources *i.e.* circular complex sources. This restriction allows them to consider $s_3 = s_1 s_2$, $s_4 = s_1^2$ and $s_5 = s_2^2$ as three additional sources. These separating methods are based on a property of circular sources, according to which all the cross-moments of the five sources s_1, s_2, s_3, s_4, s_5 are zero. Using at least 5 linear mixtures of these 5 sources, they can be separated by means, for example, of a joint diagonalization procedure, like that used in linear BSS. In the current work, however, we suppose that: 1) the sources are arbitrary real signals, and 2) only two mixtures are available.

3 Invertibility and separating structure

The first step of our investigation consists in analyzing the invertibility of the considered class of mixtures, *i.e.* in determining whether the source values s_1 and s_2 may be derived from the observed mixtures x_1 and x_2 for known coefficients l_1, l_2, q_1 and q_2. Multiplying the first equation of (1) by $-q_2$ and the second equation by q_1, then adding the results, yields

$$s_2 = [q_1 x_2 - q_2 x_1 + (q_1 l_2 + q_2) s_1]/(q_2 l_1 + q_1)$$

substituting s_2 in the second equation of (1), leads to following second-order equation for s_1

$$(q_1 l_2 + q_2) s_1^2 + (q_1 x_2 - q_2 x_1 + l_1 l_2 - 1) s_1 + (x_1 + l_1 x_2) = 0$$

Similarly, the corresponding equation with respect to s_2 is

$$(q_2 l_1 + q_1) s_2^2 + (q_2 x_1 - q_1 x_2 + l_1 l_2 - 1) s_2 + (x_2 + l_2 x_1) = 0$$

It is clear that each of these equations may have zero or two (possibly equal) real solutions. Solving the above equations, we obtain

$$\begin{aligned} \int_1 &= \frac{-(q_1 x_2 - q_2 x_1 + l_1 l_2 - 1) \pm (\Delta_1)^{1/2}}{2(q_1 l_2 + q_2)} \\ \int_2 &= \frac{-(q_2 x_1 - q_1 x_2 + l_1 l_2 - 1) \pm (\Delta_2)^{1/2}}{2(q_2 l_1 + q_1)} \end{aligned} \tag{2}$$

with $\Delta_1 = (q_1x_2 - q_2x_1 + l_1l_2 - 1)^2 - 4(q_1l_2 + q_2)(x_1 + l_1x_2)$ and $\Delta_2 = (q_2x_1 - q_1x_2 + l_1l_2 - 1)^2 - 4(q_2l_1 + q_1)(x_2 + l_2x_1)$.

The above equations define the roots $\int_1$ and $\int_2$ with respect to the available mixed variables x_1 and x_2. Let us now derive the relationship between each of these roots and the sources s_1 and s_2. Inserting (1) in Δ_1 and Δ_2 above yields

$$\begin{aligned}\Delta_1 &= [l_1l_2 - 1 + s_1(q_2 + q_1l_2) + s_2(q_1 + q_2l_1)]^2 \\ \Delta_2 &= [l_1l_2 - 1 + s_2(q_1 + q_2l_1) + s_1(q_2 + q_1l_2)]^2\end{aligned}$$

so that the roots of (2) are real. It can be easily verified that the positive signs before the square roots in (2) lead to the solution $(\int_1, \int_2) = (s_1, s_2)$ and the negative signs give the solution $(\int_1, \int_2) = (-\frac{q_1+l_1q_2}{q_2+l_2q_1}s_2 - \frac{l_1l_2-1}{q_2+l_2q_1}, -\frac{q_2+l_2q_1}{q_1+l_1q_2}s_1 - \frac{l_1l_2-1}{q_1+l_1q_2})$. Thus, the mixture structure (1) maps the two above source pairs to a unique observation point (x_1, x_2). Note that the two pairs are equal up to a permutation, a scaling factor, and an additive constant.

If the coefficients l_1, l_2, q_1 and q_2 are known, the relation (2) (with positive signs before the square roots) can be used to form a direct separating structure which restores the sources s_1 and s_2 from the mixtures x_1 and x_2. Although this direct structure could be derived for our special model, as soon as a more complicated polynomial model is considered, the solutions $\int_1$ and $\int_2$ can no longer be determined. Thus, the generalization of this method to arbitrary polynomial models seems impossible.

To avoid this limitation, we propose a recurrent structure. Such structures have been considered since the early work of Hérault and Jutten [11] and then in more complex configurations [12] - [14]. We here extend them to linear-quadratic mixtures by introducing the structure shown in Figure 1. Note that, for $q_1 = q_2 = 0$, this structure is reduced to the basic Hérault-Jutten network. It may be checked easily that, for fixed observations defined by (1), $y_1 = s_1$ and $y_2 = s_2$ corresponds to a steady state for the structure in Figure 1.
The use of this recurrent structure is more promising because it can be easily generalized to arbitrary polynomial models and its identification seems more straightforward. However, the main problem with this model is its stability. In fact, even if the model coefficients are exactly known, the computation of the structure outputs requires the realization of the following recurrent iterative model

$$\begin{aligned}y_1(n+1) &= x_1 + l_1y_2(n) + q_1y_1(n)y_2(n) \\ y_2(n+1) &= x_2 + l_2y_1(n) + q_2y_1(n)y_2(n) \qquad (3)\end{aligned}$$

where a loop on n is performed for each couple of observations (x_1, x_2) until convergence is achieved. Suppose we begin this loop with the initial conditions $y_1(0) = y_2(0) = 0$ and the observations x_1 and x_2 are obtained from the model (1). Nothing guarantees that the recurrent model (3) converges to the separating

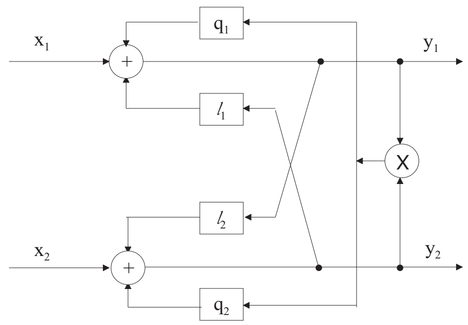

Fig. 1. Recurrent separating structure.

point $(s_1,\ s_2)$. In fact, the convergence behaviour of the model depends on the coefficient values and may be chaotic [15]. The convergence analysis is very difficult in the general case. In the following, we study only the necessary and sufficient conditions for local stability which can be considered as necessary conditions for global stability.

4 Local stability of the recurrent model

To study the local stability of the model (3) at the fixed point $(s_1,\ s_2)$, we linearize the model at this point. Let's rewrite equation (3) in vector form $\mathbf{y}(n+1) = \mathbf{f}(\mathbf{y}(n))$ where $\mathbf{y} = [y_1, y_2]^T$ and $\mathbf{f}$ is the two-dimensional function defined by (3). Linearizing this equation at the fixed point $\mathbf{s} = [s_1, s_2]^T$ yields:

$$\mathbf{y}(n+1) \simeq \mathbf{f}(\mathbf{s}) + \mathbf{J}.(\mathbf{y}(n) - \mathbf{s}) \tag{4}$$

where $\mathbf{J}$ is the Jacobian matrix of $\mathbf{f}$ defined as $\mathbf{J} = \frac{\partial \mathbf{f}(\mathbf{y}(n))}{\partial \mathbf{y}(n)}$, and evaluated at $\mathbf{s}$. From (1) and (3), it is clear that $\mathbf{f}(\mathbf{s}) = \mathbf{s}$. Defining the new centered variable $\mathbf{z} = \mathbf{y} - \mathbf{s}$, equation (4) rewrites

$$\mathbf{z}(n+1) \simeq \mathbf{J}.\mathbf{z}(n) \tag{5}$$

$\mathbf{y}(n)$ locally converges to $\mathbf{s}$, when $n \to \infty$, if and only if $\mathbf{z}(n)$ converges to zero, and therefore, if and only if the Jacobian matrix is a convergent matrix: $\lim_{n\to\infty} \mathbf{J}^n = \mathbf{0}$. A necessary and sufficient condition is that the absolute values of the two eigenvalues of $\mathbf{J}$ be smaller than one. Considering (3), the Jacobian matrix at $\mathbf{s}$ is

$$\mathbf{J} = \begin{pmatrix} q_1 s_2 & l_1 + q_1 s_1 \\ l_2 + q_2 s_2 & q_2 s_1 \end{pmatrix}$$

and its eigenvalues are

$$\lambda = \frac{1}{2}(q_1 s_2 + q_2 s_1 \pm [(q_1 s_2 + q_2 s_1)^2 + 4(l_1 l_2 + l_1 q_2 s_2 + l_2 q_1 s_1)]^{1/2})$$

Two different cases may be considered:

– The eigenvalues are complex, *i.e.* $(q_1 s_2 + q_2 s_1)^2 + 4(l_1 l_2 + l_1 q_2 s_2 + l_2 q_1 s_1) < 0$. In this case, the recurrence (3) is locally stable if and only if

$$(q_1 s_2 + q_2 s_1)^2 + |(q_1 s_2 + q_2 s_1)^2 + 4(l_1 l_2 + l_1 q_2 s_2 + l_2 q_1 s_1)| < 4$$

– The eigenvalues are real, *i.e.* $(q_1 s_2 + q_2 s_1)^2 + 4(l_1 l_2 + l_1 q_2 s_2 + l_2 q_1 s_1) \geq 0$. In this case, the local stability is achieved if and only if

$$|q_1 s_2 + q_2 s_1 \pm \sqrt{(q_1 s_2 + q_2 s_1)^2 + 4(l_1 l_2 + l_1 q_2 s_2 + l_2 q_1 s_1)}| < 2$$

Our experiments show that the local stability is very often sufficient to ensure the global stability with the initial condition $\mathbf{y}(0) = \mathbf{0}$. An example is shown in Figure 2 with $s_1 = s_2 = 1$, $q_1 = 0.5$, $q_2 = -0.5$, $l_1 = 0.6$ and $l_2 = -0.6$. The two eigenvalues are complex and their absolute values are equal to 0.9798. Figure 2.a represents the convergence trajectory with the initial conditions $y_1(0) = y_2(0) = 1.001$. In Figure 2.b, the initial conditions are $y_1(0) = y_2(0) = 0$. As can be seen, even for this extreme case where the absolute values of the eigenvalues are very close to 1, convergence is achieved.

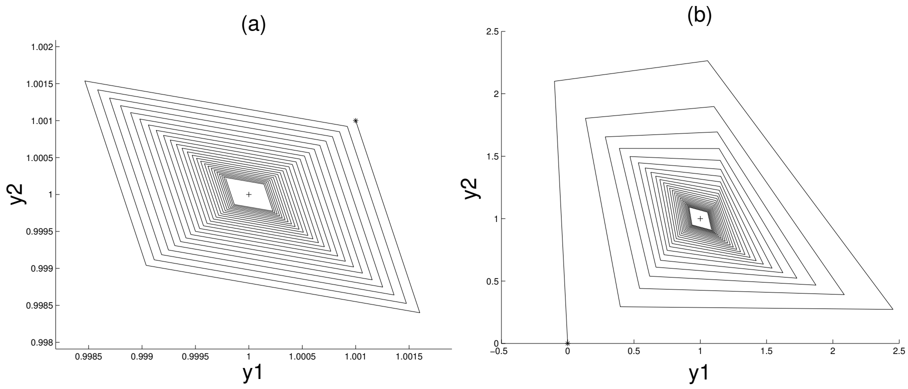

Fig. 2. Convergence trajectory (a) with initial conditions in the neighborhood of the actual sources. (b) with zero conditions.

5 Source separation algorithm

At this stage, we suppose that the sources and the mixture coefficients have numerical values ensuring the convergence of the recurrent model (3) toward the fixed point (s_1, s_2) *i.e.* to the actual sources. We suppose now the coefficients l_1, l_2, q_1 and q_2 in the recurrent separating structure of Figure 1 are unknown and we would like to identify them using only the independence of the sources.

The adaptive rule that we propose to this end is also inspired by the original work of Herault and Jutten on linear mixtures and our extensions. In their algorithm, the two parameters l_1 and l_2 are identified using two nonlinear decorrelation equations $E[f(y_1)g(y_2)] = 0$ and $E[f(y_2)g(y_1)] = 0$. In our work, having to identify 4 parameters, we need 4 nonlinear decorrelation equations. An optimal choice of the functions f and g in these equations, based on information theory, on a maximum likelihood approach, or on the method applied by Charkani and Deville to a convolutive recurrent structure [13], may be the subject of future works. In this elementary paper, we report the results obtained for a special choice of these functions *i.e.* we adapt the linear parameters l_1 and l_2 to achieve respectively $E[y_1^3\widetilde{y}_2] = 0$ and $E[y_2^3\widetilde{y}_1] = 0$, and the quadratic parameters q_1 and q_2 to obtain respectively $E[y_1^2\widetilde{y}_2] = 0$ and $E[y_2^2\widetilde{y}_1] = 0$, where $\widetilde{y}_1$ and $\widetilde{y}_2$ represent the centered versions of the outputs. For this purpose, we used the following batch algorithm, where $\mathbf{x}_1$ and $\mathbf{x}_2$ are the vectors containing the observation samples, $\mathbf{y}_1$ and $\mathbf{y}_2$ are the vectors of estimated sources, μ is a positive adaptation gain, and dot products and powers are computed component-wise.

$\mathbf{y}_1 = \mathbf{y}_2 = \mathbf{0}$
Initialize l_1, l_2, q_1 and q_2 to small random values.
Repeat
    Repeat
        $\mathbf{ynew}_1 = \mathbf{x}_1 + l_1\mathbf{y}_2 + q_1\mathbf{y}_1.\mathbf{y}_2$
        $\mathbf{ynew}_2 = \mathbf{x}_2 + l_2\mathbf{y}_1 + q_2\mathbf{y}_1.\mathbf{y}_2$
        $\mathbf{y}_1 = \mathbf{ynew}_1$
        $\mathbf{y}_2 = \mathbf{ynew}_2$
    Until convergence of the recurrent structure
    $l_1 = l_1 + \mu(\mathbf{y}_1^3)^T(\mathbf{y}_2 - mean(\mathbf{y}_2))$
    $l_2 = l_2 + \mu(\mathbf{y}_2^3)^T(\mathbf{y}_1 - mean(\mathbf{y}_1))$
    $q_1 = q_1 + \mu(\mathbf{y}_1^2)^T(\mathbf{y}_2 - mean(\mathbf{y}_2))$
    $q_2 = q_2 + \mu(\mathbf{y}_2^2)^T(\mathbf{y}_1 - mean(\mathbf{y}_1))$
Until convergence of the separating coefficients

6 Simulation results

The simulations are done using 1000 samples of two random sources uniformly distributed on $[-0.5, 0.5]$. In all the experiments, the four mixing parameters are chosen to ensure the convergence of the recurrent model (3). The distribution of the mixtures x_1 and x_2 for a sample experiment with $l_1 = -0.2$, $l_2 = 0.2$, $q_1 = -0.8$ and $q_2 = 0.8$ is shown in Figure 3.a. The highly nonlinear nature of the mixture is obvious. Figure 3.b shows the distribution of the estimated sources y_1 and y_2 applying the proposed algorithm, which indicates clearly that the independent components are retrieved.

This experiment was repeated 100 times, corresponding to 100 different seed values of random source generator. For each experiment, the Signal-to-Noise

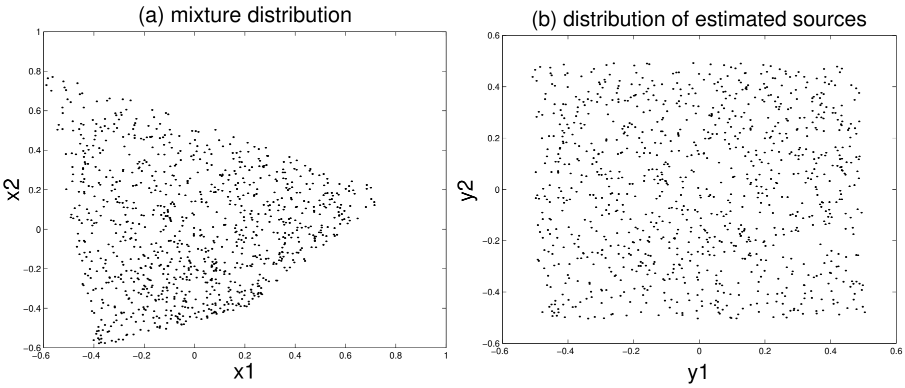

Fig. 3. (a) distribution of the two mixtures, (b) distribution of the two estimated sources.

Ratio Improvement (SNRI) of each source (in dB) is computed by

$$SNRI_i = 10 \log_{10} \frac{E[(x_i - s_i)^2]}{E[(y_i - s_i)^2]} \qquad i = 1, 2$$

and the total $SNRI$ is defined as the average of the two $SNRI_i$. The mean and the standard variation of $SNRI$ on the 100 experiments was 23.02 dB and 4.03 dB respectively. The result confirms the good performance of the separating algorithm.

Repeating the experiment with other coefficient values almost led to the same results. However, the main problem is that the adaptation procedure may sometimes bring the coefficients in the region where the recurrent network is not stable. A constrained optimization approach seems necessary to avoid these situations. More experiments with other source statistics are also necessary to confirm these elementary results.

7 Conclusion

This paper presented a first approach for separating linear-quadratic mixtures of independent real sources using a recurrent separating network. The main advantage of such a network is that it can be generalized to more complicated polynomial models. Unfortunately, the stability analysis of such networks is difficult even for the simple case studied in the paper. The reported local stability analysis allows us to determine locally necessary and sufficient conditions which impose some constraints on the mixture coefficients.

The first experimental results are promising and show that the network succeeds in separating the sources with a good performance, if its stability is guaranteed. We are currently working to complete the work reported in this paper.

Some of our objectives are: a more precise analysis of the stability of the recurrent network, determination of optimal nonlinear decorrelation functions and convergence study of parameter adaptation algorithm using an ODE analysis as in our previous investigations for similar linear approaches [13], [16], more experiments with different source statistics and different mixture coefficients, and generalization of the method to more complicated polynomial models and more sources.

References

1. A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, J. Wiley, 2001.
2. A. Cichocki and S.-I. Amari, *Adaptive Blind Signal and Image Processing - Learning Algorithms and Applications*, J. Wiley, 2002.
3. A. Hyvarinen and P. Pajunen, Nonlinear independent component analysis: Existence and uniqueness results, *Neural Networks*, vol. 12, no. 3, pp. 429-439, 1999.
4. A. Taleb and C. Jutten, Source separation in post-nonlinear mixtures, *IEEE Trans. on Signal Processing*, vol. 47, no. 10, pp. 2807-2820, 1999.
5. L. Almeida, Linear and nonlinear ICA based on mutual information, In *Proc. IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (AS-SPCC)*, pp. 117-122, Lake Louise, Canada, October 2000.
6. J. Eriksson and V. Koivunen, Blind identifiability of class of nonlinear instantaneous ICA models. In *Proc. of the XI European Signal Proc. Conf. (EUSIPCO 2002)*, volume 2, pp. 7-10, Toulouse, France, September 2002.
7. S. Hosseini and C. Jutten, On the separability of nonlinear mixtures of temporally correlated sources, *IEEE Signal Processing Letters*, vol. 10, no. 2, pp. 43-46, February 2003.
8. C. Jutten, B. Babaie-Zadeh, S. Hosseini, Three easy ways for separating nonlinear mixtures, submitted to *Signal Processing*.
9. M. Krob and M. Benidir, Blind identification of a linear-quadratic model using higher-order statistics, In *Proc. ICASSP*, vol. 4, pp. 440-443, 1993.
10. K. Abed-Meraim, A. Belouchrani, and Y. Hua, Blind identification of a linear-quadratic mixture of independent components based on joint diagonalization procedure, In *Proc. ICASSP*, pp. 2718-2721, Atlanta, USA, May 1996.
11. C. Jutten and J. Hérault, Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture, *Signal Processing*, 24:1-10, 1991.
12. T. Nguyen and C. Jutten, Blind source separation for convolutive mixtures, *Signal Processing*, vol. 45, no. 2, pp. 209-229, 1995.
13. N. Charkani and Y. Deville, Self -adaptive separation of convolutively mixed signals with a recursive structure. Part I: Stability analysis and optimization of asymptotic behaviour, *Signal Processing*, vol. 73, n. 3, pp. 225-254, 1999.
14. N. Charkani and Y. Deville, Self-adaptive separation of convolutively mixed signals with a recursive structure. Part II: Theoretical extensions and application to synthetic and real signals, *Signal Processing*, vol.75, n. 2, pp. 117-140, 1999.
15. C. Mira, L. Gardini, A. Barugola, and J.C. Cathala, *Chaotic Dynamics in two-dimensional noninvertible maps*, World Scientific, Series A on Nonlinear Sciences, vol. 20, 1996.
16. Y. Deville, Analysis of the convergence properties of self-normalized source separation neural networks, *IEEE Transactions on Signal Processing*, vol. 47, n. 5, pp. 1272-1287, May 1999.