A SECOND-ORDER DIFFERENTIAL APPROACH FOR UNDERDETERMINED CONVOLUTIVE SOURCE SEPARATION

Yannick DEVILLE and Stéphane SAVOLDELLI

Laboratoire d'Acoustique, de Métrologie, d'Instrumentation (LAMI),
Université Paul Sabatier, Bât. 3R1B2, 118 Route de Narbonne, 31062 Toulouse Cedex, FRANCE.
e-mail: ydeville@cict.fr

ABSTRACT

This paper concerns the underdetermined case of the convolutive source separation problem, i.e. the situation when the number of observed convolutively mixed signals is lower than the number of sources. We propose a criterion and associated algorithm which, unlike classical approaches, make it possible to perform the separation of a subset of these sources by exploiting their assumed non-stationarity properties. This approach uses the second-order statistics of the signals and adapts the filters of a direct separating system so as to cancel the "differential cross-correlation" of signals derived by this system. This new method is related to the general differential source separation concept that we proposed. Its effectiveness is shown by means of numerical tests.

1. PROBLEM STATEMENT

Blind source separation (BSS) methods aim at estimating a set of N_s source signals $x_j(n)$ from a set of N_o observed signals $y_i(n)$, which are mixtures of these source signals [1],[2]. The mixed signals $y_i(n)$ are typically provided by sensors, and in the so-called convolutive mixture model, each propagation path from source j to sensor i is represented by a filter, whose transfer function is denoted $A_{ij}(z)$ hereafter. In the $\mathcal{Z}$ domain, the overall relationship between the column vectors $X(z)$ and $Y(z)$ of sources and observations then reads:

$$Y(z) = A(z)X(z), \tag{1}$$

where the elements of the mixing matrix $A(z)$ are the transfer functions $A_{ij}(z)$. Most investigations have been performed in the case when: i) $N_o = N_s$, so that the matrix $A(z)$ is square, and ii) this matrix is invertible. The BSS problem then basically consists in determining an estimate of the inverse of $A(z)$, so as to restore the source signal vector $X(z)$ from the observation vector $Y(z)$ in (1). Various methods have been proposed to this end, based on the assumed statistical independence or uncorrelation of the source signals (see e.g. the surveys, resp. in [1] for the specific case of linear intantaneous mixtures and in [2] for the convolutive case).

As stated above, most of these investigations were performed under the assumption $N_o = N_s$. In many practical situations however, only a limited number of sensors is acceptable, due e.g. to cost constraints or physical configuration, whereas these sensors receive a larger number of sources. This underdetermined situation corresponding to $N_o < N_s$ has been considered by a few authors, mainly in the restricted case of linear instantaneous mixtures (see e.g. [3] and references therein, [4]). In a previous paper [5], we also introduced a general "differential BSS" concept for treating the underdetermined case. We defined a version of this concept by exploiting the assumed non-stationarity properties of the sources, i.e. the variations of the statistics of a subset of these sources over time. Using the "differential statistics" of the latter sources then makes it possible to separate them exactly (whereas the other sources yield some additional "noise" contributions in the outputs of the proposed system). We introduced a specific differential BSS criterion resulting from this general concept and intended for linear instantaneous mixtures. This specific approach is based on the optimization of the "differential normalized kurtosis" that we introduced to this end.

The current paper presents major extensions of our previous work on underdetermined BSS from two points of views. On the one hand, we here develop a practical approach intended for *convolutive* mixtures. On the other hand, the criterion and algorithm proposed below are based on our general differential BSS concept but are not mere convolutive extensions of the ones that we previously developed for linear instantaneous mixtures. Instead, they use other statistical signal parameters and adaptation rules. These differences stem from the well-known fact that linear instantaneous BSS can only be performed by resorting to higher-order statistics if no specific assumptions are made on the sources, whereas (strictly causal [6]) convolutive BSS may also be achieved by means of second-order statistics. By using the latter approach, we here introduce a method based on correlation (i.e. second-order) parameters and their cancellation, which is to be contrasted with our previous approach based on normalized kurtosis (i.e. fourth-order parameter) and its min/maximization.

2. PROPOSED APPROACH

2.1. Redefining the classical approach

The investigation reported in this paper is based on the classical decorrelation approach that has been used by various authors for solving the basic configuration of the convolutive BSS problem [6]-[9]. We therefore first redefine this classical method in a way which is suited to the approach that we will then use to extend it. The considered basic configuration involves two convolutive mixtures of two uncorrelated sources, defined in the $\mathcal{Z}$ domain as:

$$Y_1(z) = X_1(z) + A_{12}(z)X_2(z) \tag{2}$$
$$Y_2(z) = A_{21}(z)X_1(z) + X_2(z), \tag{3}$$

where $A_{12}(z)$ and $A_{21}(z)$ are strictly causal MA filters and their orders are (at most) equal to M. These mixed signals are provided to a separating system, which aims at restoring the source signals

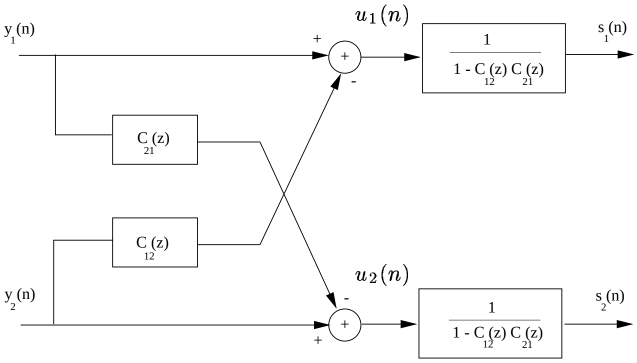

Fig. 1. Separating system based on a direct structure.

from them. The version of this system considered here is the direct structure shown in Fig. 1, where $C_{12}(z)$ and $C_{21}(z)$ are the transfer functions of strictly causal M^{th}-order MA filters. The coefficients of each of these filters evolve vs. time and the value of the k^{th} coefficient, with $k \in [1, M]$, at time n is denoted $c_{ij}(n, k)$. These coefficients are adapted so as to decorrelate the time-shifted intermediate signals $u_1(n)$ and $u_2(n)$ of the separating system, i.e. so as to fulfill the following criterion:

$$R_{u_i u_j}(n, n-k) = 0, \quad i \neq j \in \{1, 2\}, k \in [1, M], \qquad (4)$$

where $R_{vw}(m_1, m_2) = E\{v(m_1)w(m_2)\}$ denotes the cross-correlation of any couple of signals and where all the signals considered in this paper are assumed to be zero-mean for simplicity. The classical stochastic algorithm used in practice to fulfill the above criterion consists in adapting each filter coefficient at each time n according to the rule:

$$\begin{aligned} c_{ij}(n+1, k) \quad &= \quad c_{ij}(n, k) + \mu u_i(n) u_j(n-k) \\ & \qquad i \neq j \in \{1, 2\}, k \in [1, M]. \end{aligned} \qquad (5)$$

This algorithm actually implements the above criterion, as its equilibrium points correspond to:

$$\begin{aligned} E\{c_{ij}(n+1, k) - c_{ij}(n, k)\} \quad &= \quad 0 \\ & \qquad i \neq j \in \{1, 2\}, k \in [1, M], \end{aligned} \qquad (6)$$

which is clearly equivalent to (4). The motivation for selecting this criterion (and associated algorithm) may be explained as follows. The internal state of the separating system of Fig. 1 is defined by the transfer functions $C_{12}(z)$ and $C_{21}(z)$. The state of interest in the BSS problem is the so-called "separating state", which yields non-permuted separated signals on the system outputs, i.e:

$$S_i(z) = X_i(z), \quad i \in \{1, 2\} \qquad (7)$$

and which may be shown to be defined by:

$$C_{ij}(z) = A_{ij}(z), \quad i \neq j \in \{1, 2\}. \qquad (8)$$

The criterion to be used should therefore be selected so that it is fulfilled at the separating state. It may be checked easily that the criterion defined by (4) meets this requirement.

2.2. Limitations of the classical approach

Now consider the situation when the available two mixed signals $Y_1(z)$ and $Y_2(z)$ contain not only the above contributions from sources $X_1(z)$ and $X_2(z)$, but also contributions from an arbitrary number of additional sources $X_3(z)$ to $X_{N_s}(z)$, so that:

$$Y_1(z) = X_1(z) + A_{12}(z)X_2(z) + \sum_{j=3}^{N_s} A_{1j}(z)X_j(z) \qquad (9)$$

$$Y_2(z) = A_{21}(z)X_1(z) + X_2(z) + \sum_{j=3}^{N_s} A_{2j}(z)X_j(z), \qquad (10)$$

where N_s is the overall number of sources and where the additional transfer functions $A_{ij}(z)$ introduced here are also strictly causal M^{th}-order (or less) MA filters. Before focusing on the structure of Fig. 1, we consider the complete class of separating systems which process these mixed signals in a linear (convolutive) way, based on the motivations presented in [5]. The desired "optimum operation" of these systems for the considered mixed signals may be defined as follows. By linearly combining the mixed signals, only a single source contribution may be cancelled in any (non-zero) output of such systems for arbitrary mixtures. If $X_1(z)$ and $X_2(z)$ are considered as the main signals of interest, i.e. the "useful signals" to be separated, one would like these signals to appear resp. only in the outputs $S_1(z)$ and $S_2(z)$ of the separating system. $X_1(z)$ and $X_2(z)$ are then the signals that should be cancelled in system outputs. These outputs would then also contain contributions from sources $X_3(z)$ to $X_{N_s}(z)$, which are then considered as additional undesirable sources, i.e. "noise". The considered separating system would thus perform what we call the "partial BSS" of $X_1(z)$ and $X_2(z)$.

The *structure* of the separating system shown in Fig. 1 is potentially suited to this optimum operation, in the sense that it is able to achieve the partial BSS of $X_1(z)$ and $X_2(z)$: this partial BSS occurs when (8) is met, which here defines the "partial separating state". However, the overall classical approach defined in Subsection 2.1 is not able to achieve this partial BSS, due the *criterion* (and associated algorithm) that it uses, as will now be shown. By combining the mixing equations (9)-(10) and the internal equations of the separating system of Fig. 1, the intermediate signals $u_i(n)$ of the latter system may be expressed as:

$$u_i(n) = \sum_{p=1}^{N_s} \sum_{m=1}^{2M} h_{ip}(n, m) x_p(n-m), \qquad (11)$$

where $h_{ip}(n, m)$ are the time-varying coefficients of the overall filters combining the filters $A_{ij}(z)$ and $C_{ij}(z)$. The cross-correlation of these signals for any separating system state then reads:

$$\begin{aligned} R_{u_i u_j}(n, n-k) \quad = \quad & \sum_{p=1}^{N_s} \sum_{m=1}^{2M} \sum_{l=1}^{2M} \left[h_{ip}(n, m) h_{jp}(n-k, l) \right. \\ & \left. R_{x_p}(n-m, n-k-l) \right]. \end{aligned} \qquad (12)$$

At the partial separating state, the contributions of $X_1(z)$ and $X_2(z)$ in (12) may be shown to disappear, while the contributions of the noise sources remain. In other words, in the case of noisy mixed signals, the cross-correlation values $R_{u_i u_j}(n, n-k)$ are non-zero at the partial separating state. The classical approach of Subsection 2.1 cannot then converge to this state, as it adapts the separating filter coefficients so as to reach a state where the cross-correlation values $R_{u_i u_j}(n, n-k)$ are cancelled. This approach then fails to achieve partial BSS.

2.3. Proposed differential approach

We here still consider the noisy mixed signals introduced in Subsection 2.2. Based on the above results, we again use the structure of Fig. 1, but we here aim at introducing a new criterion (and associated algorithm) for adapting its filter coefficients. This criterion is designed so that the resulting method becomes able to achieve the partial BSS of $X_1(z)$ and $X_2(z)$. The proposed approach is based on the general differential BSS concept that we introduced in [5]. It requires the sources $X_1(z)$ and $X_2(z)$ to be long-term non-stationary, whereas the other sources should be long-term stationary. In other words, the useful (resp. noise) sources should have different (resp. identical) second-order statistics at times n_1 and n_2 when these times are separated by a "long" period. This long period is defined by contrast to each short period associated to a single time n_1 or n_2, over which sample statistics of the signals are measured in practice, and over which all signals should therefore be short-term stationarity. The main idea of differential BSS then consists of considering the difference between the signal statistics resp. associated to n_1 and n_2, so that the contributions of the long-term stationary (i.e. noise) sources are cancelled in the resulting differential statistical parameters. We thus get back into the classical noiseless configuration from the point of view of the resulting adaptation criterion. To apply this general idea to the specific approach introduced in this paper, we first define the "differential correlation function", which reads:

$$\begin{aligned} DR_{vw}(n_1,n_2,l_1,l_2) &= R_{vw}(n_2-l_1,n_2-l_2) \\ &\quad -R_{vw}(n_1-l_1,n_1-l_2), \quad (13) \end{aligned}$$

where n_1 and n_2 are two reference times and l_1 and l_2 are two lags. When considering the difference between the two values, resp. associated to $n=n_1$ and $n=n_2$, of the correlation involved in the classical criterion (4), we obtain:

$$\begin{aligned} R_{u_iu_j}(n_2,n_2-k) &- R_{u_iu_j}(n_1,n_1-k) \\ &= DR_{u_iu_j}(n_1,n_2,0,k). \quad (14) \end{aligned}$$

We then use the same separating filter values for $n=n_1$ and $n=n_2$, and the same principle is also applied to any time-shifted version of this couple of times. The first overall filter coefficient $h_{ip}(n,m)$ involved in (12) then has the same value for $n=n_1$ and $n=n_2$ and this common value is simply denoted $h_{ip}(n,m)$ hereafter. The coefficient $h_{jp}(n-k,l)$ in (12) leads to the same phenomenon, so that combining (12) with (14) yields:

$$\begin{aligned} DR_{u_iu_j}(n_1,n_2,0,k) &= \sum_{p=1}^{N_s}\sum_{m=1}^{2M}\sum_{l=1}^{2M}\left[h_{ip}(n,m)\right. \\ &\quad \left. h_{jp}(n-k,l)DR_{x_p}(n_1,n_2,m,k+l)\right]. \quad (15) \end{aligned}$$

This equation holds for any state and any type of sources. When the noise sources $X_3(z)$ to $X_{N_s}(z)$ have the above-mentioned long-term stationarity property, their differential autocorrelations $DR_{x_p}(.)$ contained in (15) are equal to zero, so that only the useful sources $X_1(z)$ and $X_2(z)$ remain. In other words, from the point of view of the new parameter $DR_{u_iu_j}(n_1,n_2,0,k)$ that we introduced, we actually get back in the classical 2-source to 2-sensors configuration, except that this new parameter depends on the differential auto-correlation functions of the sources, instead of the plain auto-correlations which appear in the classical approach. It may then be shown easily that $DR_{u_iu_j}(n_1,n_2,0,k)=0$ at the partial separating state whereas, if the sources $X_1(z)$ and $X_2(z)$ are long-term non-stationary, $DR_{u_iu_j}(n_1,n_2,0,k)\neq 0$ for all other states (except for some possible spurious states, which correspond to those that may exist for the classical approach). The criterion that we eventually propose for performing the partial BSS of $X_1(z)$ and $X_2(z)$ in the case of noisy mixtures therefore consists in adapting all separating filter coefficients so as to achieve:

$$DR_{u_iu_j}(n_1,n_2,0,k)=0, \quad i\neq j\in\{1,2\}, k\in[1,M]. \quad (16)$$

A practical stochastic algorithm which implements this criterion may then be derived by adapting the approach that we defined Subsection 2.1 for the classical algorithm. It reads:

$$\begin{aligned} c_{ij}(n+1,k) &= c_{ij}(n,k)+\mu_i[u_i(n_2)u_j(n_2-k) \\ &\quad -u_i(n_1)u_j(n_1-k)], \\ &\quad i\neq j\in\{1,2\}, k\in[1,M] \quad (17) \end{aligned}$$

and deserves the following comments. The classical algorithm (5) performs a sweep over the data by using an increasing time index n. The algorithm proposed here is also based on a single sweep, but each step of this sweep involves two points in the data time series, corresponding to the indices n_1 and n_2. The difference between these indices is typically kept constant (and "long", as defined above), so that the sweep is performed in parallel over two time-shifted versions of the data. The considered couple of times inside these data is then defined by a single index, denoted n in (17), which is e.g. equal to n_1 or n_2. The adaptation gains μ_1 and μ_2 in (17) typically have the same absolute value, but their signs are selected so that:

$$\begin{cases} \mu_1 DP_{x_2}(n_1,n_2)>0 \\ \mu_2 DP_{x_1}(n_1,n_2)>0 \end{cases} \quad (18)$$

where we define the differential power of any signal $v(n)$ between times n_1 and n_2 as:

$$DP_v(n_1,n_2)=R_v(n_2,n_2)-R_v(n_1,n_1). \quad (19)$$

The condition (18) is based on stability requirements, to be reported in a future paper.

3. NUMERICAL TESTS

We have validated the above approach by means of tests performed with two artificial convolutive mixtures of three synthetic random sources. The useful sources $x_1(n)$ and $x_2(n)$ consists of 100000 samples, split in two 50000-sample periods. In each of these periods, these sources are independent and each of them is stationary, binary-valued and equiprobable. They take the values ± 1 in their first period and ± 2 in the second one. The noise source $x_3(n)$ is uniformly distributed over the range $[-1,+1]$ in both periods. These three sources are mixed according to (9)-(10), where the mixing filters associated to the useful sources are set to:

$$A_{12}(z) \simeq -0.381z^{-1}+0.136z^{-2}+0.081z^{-3} \quad (20)$$

$$A_{21}(z) \simeq -0.327z^{-1}-0.184z^{-2}+0.027z^{-3}, \quad (21)$$

whereas the mixing filters associated to the noise source are:

$$A_{13}(z) = z^{-1}+\frac{2}{3}z^{-2}+\frac{1}{3}z^{-3} \quad (22)$$

$$A_{23}(z) = \frac{1}{3}z^{-1}+\frac{2}{9}z^{-2}+\frac{1}{9}z^{-3}. \quad (23)$$

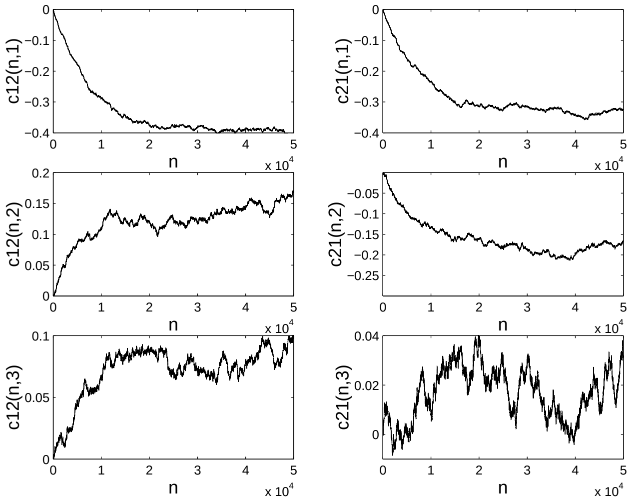

Fig. 2. Evolution of the coefficients $c_{ij}(n,k)$ of the strictly causal MA separating filters for noisy mixtures (left column: $C_{12}(z)$, right column: $C_{21}(z)$).

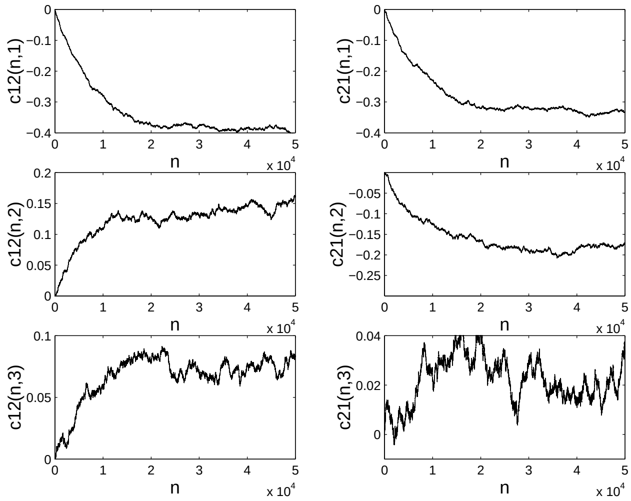

Fig. 3. Evolution of the coefficients $c_{ij}(n,k)$ of the strictly causal MA separating filters for noiseless mixtures (left column: $C_{12}(z)$, right column: $C_{21}(z)$).

Fig. 2 represents the evolution of the coefficients $c_{ij}(n,k)$ of the separating filters when the above-defined parallel sweep is performed over both 50000-sample periods and these filter coefficients are adapted by means of the proposed algorithm (17). These coefficients converge to values which are close to those of the mixing filters $A_{12}(z)$ and $A_{21}(z)$. This shows the ability of the proposed approach to achieve the partial BSS of $X_1(z)$ and $X_2(z)$ defined by (8), although the contributions from the noise source $X_3(z)$ contained in the processed mixed signals cover significantly larger ranges than those of the useful sources. This insensitivity to the presence of noise sources is also confirmed by the fact that the above separating filter coefficient trajectories are almost identical to those obtained with the proposed approach in the noiseless case, i.e. for $A_{13}(z) = 0$ and $A_{23}(z) = 0$, which are shown in Fig. 3.

4. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a criterion and associated algorithm for adapting the filters of a direct separating system so as to perform the separation of two non-stationary sources from two convolutive mixtures which also contain an arbitrary number of stationary noise sources. The proposed algorithm consists of a stochastic cancellation of the above-defined "differential cross-correlations" of intermediate signals of the separating system. This algorithm is the differential version of the classical decorrelation approach and was derived by using the general differential BSS concept that we proposed. We demonstrated its effectiveness by means of numerical tests performed with synthetic data. Our future work will esp. concern the application of this approach to real-world signals, such as noisy speech mixtures.

5. REFERENCES

[1] J.-F. Cardoso, "Blind signal separation: statistical principles", Proceedings of the IEEE, vol. 86, no. 10, pp. 2009-2025, Oct. 1998.

[2] K. J. Pope, R.E. Bogner, "Blind signal separation II. Linear, convolutive combinations", Digital Signal Processing, vol. 6 pp. 17-28, 1996.

[3] P. Comon, O. Grellier, "Non-linear inversion of underdetermined mixtures", Proceedings of the International Workshop on Independent Component Analysis and Blind Signal Separation (ICA'99), Aussois, France, Jan. 11-15, 1999, pp. 461-465.

[4] X-R Cao, R-W Liu, "General approach to blind source separation", IEEE Transactions on Signal Processing, vol. 44, no. 3, pp. 562-571, March 1996.

[5] Y. Deville, M. Benali, "Differential source separation: concept and application to a criterion based on differential normalized kurtosis", Proceedings of the 10th European Signal Processing Conference (EUSIPCO 2000), session TueAmOR1, Tampere, Finland, Sept. 4-8, 2000.

[6] N. Charkani, 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, no. 3, pp. 225-254, 1999.

[7] M.J. Al-Kindi, J. Dunlop, "Improved adaptive noise cancellation in the presence of signal leakage on the noise reference channel", Signal Processing, vol. 17, pp. 241-250, 1989.

[8] S. Van Gerven, D. Van Compernolle, "Feedforward and feedback in a symmetric adaptive noise canceler: stability analysis in a simplified case", Proceedings of the 6th European Signal Processing Conference (EUSIPCO'92), pp. 1081-1084, August 1992, Brussels, Belgium, J. Vandewalle.

[9] S. Van Gerven, D. Van Compernolle, "Signal separation by symmetric adaptive decorrelation: stability, convergence and uniqueness", IEEE Transactions on Signal Processing, vol. 43, no. 7, pp. 1602-1612, July 1995.