

# DYNAMIC SIGNAL MIXTURES AND BLIND SOURCE SEPARATION

*D. Obradovic*

Siemens AG, Central Technology Department, Information and Communications

Otto-Hahn-Ring 6, 81739 Munich, Germany,

Dragan.Obradovic@mchp.siemens.de

## ABSTRACT

Methods for blind source separation (BSS) from linear instantaneous signal mixtures have drawn a significant attention due to their ability to recover original independent non-Gaussian sources without analyzing their temporal statistics. Hence, original voices or images (modulo permutation and linear scaling) are extracted from their mixtures without modeling the dynamics of the signals. The typical methods for performing blind source separation are Linear Independent Component Analysis (ICA) and the InfoMax method. Linear ICA directly penalizes a suitably chosen measure of the statistical dependence between the extracted signals. These measures are either obtained from the Information theoretic postulates such as the mutual information or from the cumulant expansion of the associated probability density functions. The InfoMax method is based on the entropy maximization of the non-linear transformation of the separated signals.

This paper analyzes extensions of the instantaneous blind source separation techniques to the case of linear dynamic signal mixtures. Furthermore, the paper introduces a novel method based on combining Time Delayed Decorrelation (TDD) with the minimization of the cumulant cost function. TDD is used to obtain an acceptable initial condition for the cumulant based cost function optimization in order to reduce the numerical complexity of the latter method. This combined approach is illustrated on two examples including a real life cocktail party example.

**Keywords:** higher order statistics, signal reconstruction

## 1. INTRODUCTION

The problems of independent feature extraction and blind signal separation are closely related. The goal of feature extraction is the identification of statistically independent variables which "span" the original data space. A typical example is PCA in the case of linearly mixed gaussian input variables. Hence, in the latter case, the uniqueness of extracted features is not of primary interest but their independence. On the contrary, blind signal separation tries to extract original sources from the available signal mixtures. The only available information is that the original sources were statistically independent while, differently from the standard blind channel equalization problem, their

probability density functions are not known. Consequently, it is obvious that the Blind Source Separation can be seen as the independent feature extraction problem with the constraint that the obtained features are unique modulo allowed transformation such as linear scaling, delay, and permutation.

In order to pose BSS as a learning problem, an appropriate cost function that provides a direct or an indirect measure of statistical dependence has to be defined. The typical approaches are:

- i) minimization of the mutual information between the outputs of the linear transformation [1],[2]
- ii) maximization of entropy at the output of non-linearities applied individually to the output of the demixing transformation [3],[4],[5]
- iii) minimization of the magnitude of the non-diagonal cumulant elements (cross-cumulants) [1],[2] of the joint distribution at the output of the demixing transformation, and
- iv) simultaneous minimization of cross-correlation between the output signals at different delays (TDD)[6].

The cumulant cost function penalizes the cross-cumulant elements at the cumulant orders 2-4 [1],[2]:

$$D_{\text{cum}} \approx J(\mathbf{M}) = \sum_{i=1}^4 \sum_{\text{nondiag}} [C_{\text{nondiag}}^{(i)}]^2 \quad (1)$$

In the case of instantaneous mixing, suitable preprocessing and scaling significantly simplify the cost function in (1):

$$\min J_r(\mathbf{R}) = \min \sum_{i=1}^n -[K^r_i(\vec{y}_1)]^2; \quad \mathbf{R}^T = \mathbf{R}^{-1} \quad (2)$$

$$\vec{y}_1 = \mathbf{R} \cdot \mathbf{D}_1^{-0.5} \cdot \mathbf{N} \cdot \vec{x}$$

since the minimization of all cross-cumulants (non-diagonal elements of cumulant tensors) is substituted with maximization of significantly fewer diagonal elements [7]. The scaling matrices  $\mathbf{D}_1$  and  $\mathbf{N}$  are obtained from the singular value decomposition of the covariance matrix of  $\vec{x}$ .

In the following part of the paper, two different application of the BSS from of linear dynamic mixtures are presented. The first application is the so called Blind Deconvolution of a white noise signal filtered through an unknown linear filter. The second

application is the separation of two signals mixed through an unknown transfer function matrix.

## 2. BLIND DECONVOLUTION OF THE SINGLE-INPUT SINGLE-OUTPUT SYSTEM

Let a white non-Gaussian noise sequence  $\{z_i\}$  be filtered through an unknown stable invertible single-input single-output system  $S$  producing an output  $\{x_i\}$ . Due to convolution, the output of  $S$  is:

$$x_i = \sum_{j=0}^i s_j \cdot z_{i-j} \quad (3)$$

where  $\{s_i\}$  is the corresponding impulse response. The goal is to identify a stable linear system described by its finite-length impulse response  $h = [h_0, h_1, \dots, h_N]$  which when subjected to the input sequence  $\{x_i\}$  produces a white noise sequence  $\{y_i\}$ . The blind deconvolution problem can then be easily converted into a standard instantaneous blind source separation problem where the demixing matrix  $M$  is of dimension  $N \times N$  and it has the following structure of a causal convolution matrix:

$$M = \begin{bmatrix} h_0 & 0 & 0 & \dots & 0 \\ h_1 & h_0 & 0 & \dots & 0 \\ h_2 & h_1 & h_0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ h_N & h_{N-1} & h_{N-2} & \dots & h_0 \end{bmatrix} \quad (4)$$

Hence, the blind deconvolution problem with the non-gaussian white noise at the input can be seen as a standard instantaneous BSS problem with a restriction on the demixing matrix. The demixing is achieved when the outputs are decorrelated, i.e. white.

## 3. MULTI-CHANNEL DYNAMIC SIGNAL SEPARATION

The multi-channel blind source separation problem of dimension two is depicted in Figure 1.

The mixing system  $S(q)$ , where  $q$  stands for a unit delay, is assumed to be stable and to have a stable inverse, i.e. that it is a minimum-phase system. Moreover, the input signals  $z_1(t)$  and  $z_2(t)$  are assumed to be statistically independent and non-Gaussian. The signals  $x_1(t)$  and  $x_2(t)$  are inputs into a demixing system  $M(q)$  whose parameters are trained to maximize the statistical independence between the output signals  $y_1(t)$  and  $y_2(t)$ .

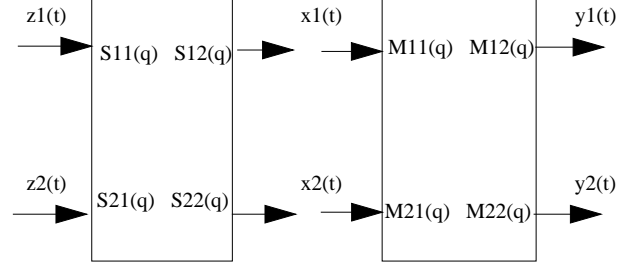


Figure 1. Mixing - Demixing Scheme

There are several issues that are specific to the dynamic BSS:

(i) *stability of the demixing system*. This is trivially achieved when  $M(q)$  is restricted to have a finite impulse response. Otherwise, stability of IIR system  $M(q)$  can be guaranteed during the learning phase by projection into the unit circle. Possible non-causality of the inverse of  $S(q)$  might be handled by appropriate time shift (delay) of the input sequence  $x(t)$ .

(ii) *Uniqueness of the separated signals  $y(t)$* . In the case of the instantaneous (static) mixtures, the original sources are recovered modulo perturbation and scaling. In the case of dynamical demixing, the non-uniqueness is even more present. It is obvious that if  $y_1(t)$  and  $y_2(t)$  are independent, than any linearly filtered versions of these signals will still be independent. Hence, additional information is needed in order to reduce the inherent non-uniqueness of the problem.

(iii) *gaussianization of the data*. The cumulant based algorithms for static blind source separation effectively minimize the higher-order cross-cumulants corresponding to the output signals  $y(t)$ . On the other hand, linear filtering leads to the "gaussianization" of the data where the higher order cumulants tend to zero. Consequently, this can lead to spurious solutions where the cost function achieves the minimum while the separation did not take place. In order to prevent this degenerative case, we will restrict the structure of the demixing transfer function matrix  $M(q)$ .

From now on, we will assume that  $M_{11}(q) = 1$  and  $M_{22}(q) = 1$ . This assumption will be exact if the mixing elements  $S_{11}(q)$  and  $S_{22}(q)$  were also equal to identity. Otherwise, we will assume that  $S_{11}(q)$  and  $S_{22}(q)$  have stable inverses which will enable us to scale diagonal elements of  $M(q)$  to one. This assumption will substantially reduce the non-uniqueness of the BSS solution and would effectively avoid a possibility of extensive gaussianization of the output signals. Although the restriction on the diagonal elements of  $M(q)$  may initially look very restrictive, there are plenty of real life examples where it is satisfied. A typical example is the denoising of speech signals based on two-microphone recording where one microphone is pointed towards the speaker while the other microphone is pointed in the opposite direction and its primary role is too obtain the noise signal.

Theoretically speaking, any of the BSS algorithms for the static (instantaneous) case can be applied to the dynamic signal separation but on the expense of the increased numerical

complexity. The information maximization method suffers from the lack of clear relationship between the input and output joint probability density functions which makes the learning rule more complex than in the instantaneous mixing problem. Similarly, the simplified cumulant based method in (2) cannot be applied since there is no dynamic preprocessing that preserves the Frobenius norm of the cumulant tensors. Hence, the cumulant approach must be based on direct evaluation of the cross-cumulants as defined in (1) which can be numerically expensive. Therefore, in order to use cumulant BSS method, an appropriate initialization is needed.

The TDD method in the case of simultaneous decorrelation at two different time delays is related to an appropriate matrix eigenvalue problem [6]. We will use this approach to initiate the cross-cumulant minimization problem. Hence, our combined approach is:

- 1) repeat the TDD method based on the eigenvalue problem in the frequency domain for several pairs of delays and pick the solution where the two cross-correlation terms are minimal
- 2) start the cross-cumulant minimization with the initial FIR parameters from the previous step.

This approach is illustrated in the following section on two examples. The first example is an artificial example where the quality of the extracted signals can be explicitly checked while the second example is a real life cocktail party example based on the recordings obtained from [8].

## 4. EXAMPLES

In the first example, the mixing transfer function matrix is defined as:

$$S(q) = \begin{bmatrix} 1 & 1 - 0.5q + 0.24q^2 \\ 3.5 - q - 1q^2 & 1 \end{bmatrix} \quad (5)$$

As stated before,  $q$  stands for a unit delay. Hence, a transformation to the standard  $z$  domain is carried out by substituting  $q$  by  $z^{-1}$ . The original independent sources are two voice signals from 2 different speakers sampled at 8000Hz. The transfer function matrix is stable and has a stable inverse.

The TDD eigenvalue method was run for several delays and the best results were obtained for the delay of 96 samples while the window size for frequency evaluation was equal to 256 points. The first 6 elements of the obtained impulse response for  $M12(q)$  and  $M21(q)$  were used as the initial values for the cross-cumulant minimization approach.

The quality of the obtained demixing transfer function matrix  $M(q)$  can be checked by computing the impulse response of the product  $S(q)M(q)$  which is depicted in Figure 2. The scaling  $I - M12(q)M21(q)$  was not carried out because it does not affect the independence if it is achieved and, in this case, it did not influence the hearing quality of the recovered voices. From

Figure 2 it is visible that all the elements of the transfer functions  $M21(q)$  and  $M12(q)$  are close to zero what indicates that the original signals were successfully recovered.

The second example is taken from [8]. In this example two speakers were recorded speaking simultaneously. The first speaker was counting from one to ten in English while the other was doing the same but in Spanish. The recording was performed in a normal office room with distance between speaker and the microphones of 60cm in square ordering.

The recovery of original voices was carried out successfully. The mixed and recovered signals are depicted in Figure 3.

## 5. CONCLUSION

Blind Signal Separation from dynamic mixtures is considerably more complex than the problem of instantaneous signal mixtures. In this paper we have mentioned the existing methods for instantaneous BSS and discussed additional difficulties arising in the case of dynamic mixtures. A novel method based on combining time delayed decorrelation and the cross-cumulant minimization is introduced. The TDD method is repeated for different delays until suitable values for the FIR elements of the demixing transfer function matrix are found. The obtained transfer function matrix is then used as the initial condition for the further cross-cumulant minimization. The effectiveness of this combined method is illustrated on two examples including a real life cocktail-party example.

## 6. REFERENCES

- [1] P. Comon, Independent Component Analysis, A new concept? *Signal Processing*, **36**, 287-314, 1994.
- [2] G. Deco and D. Obradovic, *An Information-Theoretic Approach to Neural Computing*. Springer Verlag, New York, 1996
- [3] A.J. Bell and T.J. Sejnowski, An Information-Maximization Approach to Blind Separation and Blind Deconvolution, *Neural Computation*, **7**, Number 6, 1129-1160, 1995
- [4] B.A. Pearlmutter and L.C., Parra, A context-sensitive generalization of ICA, *Proc. ICONIP'96*, Japan, 1996.
- [5] T. Lee, M. Girolami, A.J. Bell, and T.J. Sejnowski, An unifying Information-Theoretic Framework for Independent Component Analysis, *International Journal on Mathematical and Computer Modeling*. In press.
- [6] F. Ehlers and H. Schuster, Blind separation of convolutive mixtures and an application in automatic speech recognition in noisy

environment, *IEEE Trans. Signal Processing*, 45 (10), 2608-2609, 1997.

*International Conference on Acoustics, Speech and Signal Processing* ", Vol 2, 1249-1252, Seattle, 1998.

- [7] D. Obradovic and G. Deco, Information Maximization and Independent component analysis: Is there a difference?, *Neural Computation*. In press.
- [8] T-W. Lee, A. Ziehe, R. Orglmeister and T.J. Sejnowski, Combining time-delayed decorrelation and ICA: towards solving the cocktail party problem, *Proceedings of "IEEE*

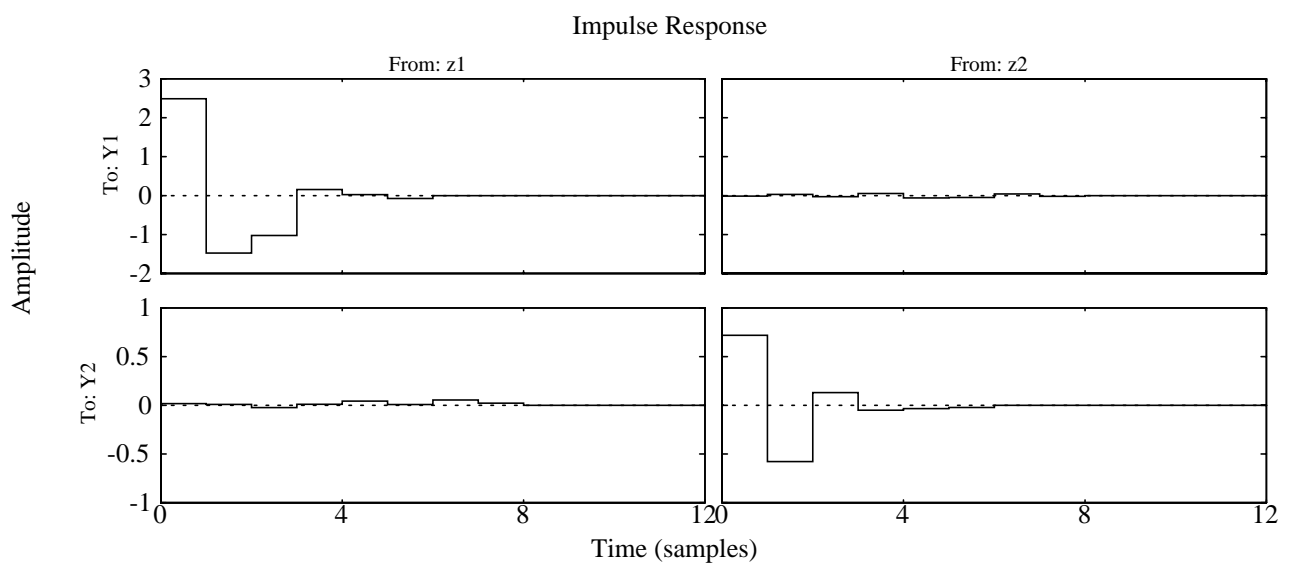


Figure 2. Impulse response of the combined mixing-demixing systems

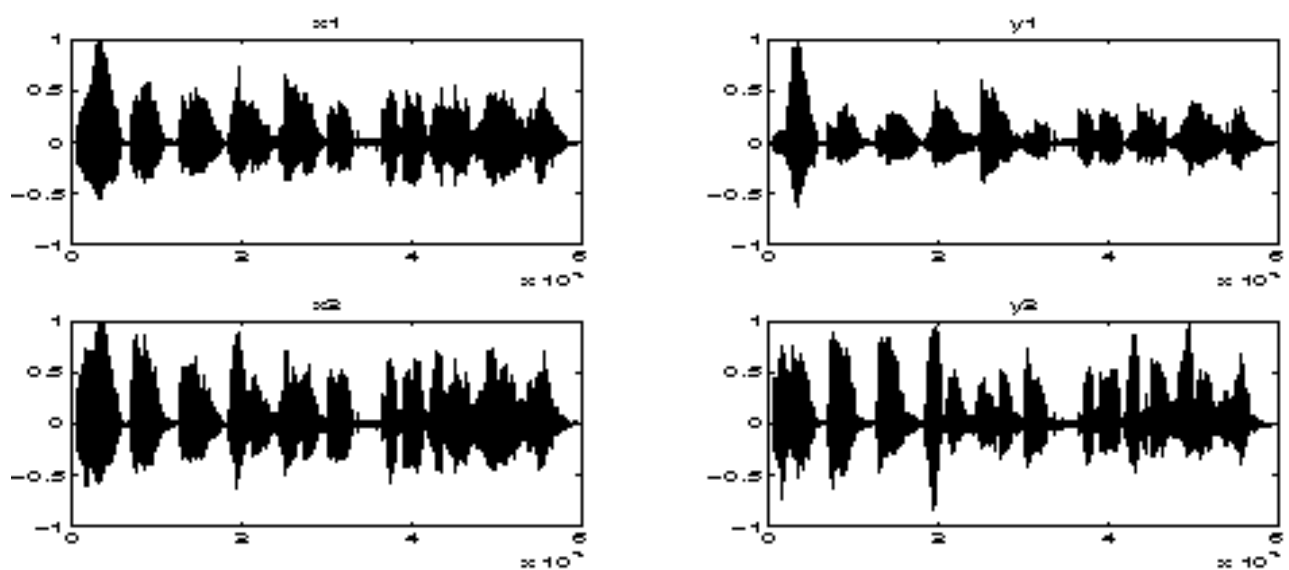


Figure 3. Mixed and recovered signals in the cocktail party example