Proceedings of the IASTED International Conference
SIGNAL PROCESSING AND COMMUNICATIONS
September19-22,2000, Marbella, Spain

Discussion of Simple Algorithms and Methods to Separate Nonstationary Signals.

Ali MANSOUR and Noboru OHNISHI
Bio-Mimetic Control Research Center (RIKEN),
2271-130, Anagahora, Shimoshidami, Moriyama-ku, Nagoya 463 (JAPAN)
email: mansour@ieee.org, ohnishi@nagoya.riken.go.jp

Abstract

In the last decade, many researchers have investigated the blind separation of sources and many algorithms have been proposed to solve this problem for the case of an instantaneous mixture (memoryless mixture) [1].

In general, high-order statistics (i.e., fourth order) are used. However, it has been shown that algorithms and criteria can be simplified by adding special assumptions [2].

In this paper, we outline the investigation of the separation of nonstationary signals using only second-order statistics. For the case of independent nonstationary (at least using second-order statistics) sources such speech signals where the power of the signals is considered time variant, we prove, using geometrical information, that the decorrelation of the output signals at any time leads to the separation of the independent sources. In other words, for these kinds of sources, any algorithm can separate the sources if at the convergence of this algorithm the covariance matrix of the output signals becomes a diagonal matrix at any time. Finally, some algorithms are proposed and the experimental results are discussed and shown.

keywords: Decorrelation, Second-order Statistics, Whiteness, Blind separation of sources, Natural gradient, Kull-back divergence, Hadamard inequality, Jacobi Diagonalization, Cyclic Jacobi Diagonalization, Joint Diagonalization.

320-046

1 Introduction

The blind separation of sources is a recent and important problem in the signal processing field. It involves retrieving unknown sources of unknown mixtures from observation using multisensors. The authors maintain two fundamental assumptions [2].

- **H1**: The sources are unknown and statistically independent from each other.
- **H2**: The channel model is known: as instantaneous (or memoryless) [3, 4, 5, 6], convolutive [7], or non-linear mixture [8, 9].

For the instantaneous mixture, one must assume that the mixture matrix $\mathbf{M}$ is a full-rank non-singular matrix [10, 11]. For the other kinds of mixtures, the authors maintain similar assumptions. For the instantaneous mixture, many algorithms have been proposed by different researchers [12, 13, 14, 15]. All of these algorithms are based on high-order statistics and in most cases fourth-order cumulants or moments are used.

After further assumptions [16, 17], researchers proposed algorithms and criteria based solely on second-order statistics, for example, those concerning the subspace properties of the channel [18, 19], the correlation properties of the sources (i.e., the samples of each source are correlated) [20, 21], or the nonstationary properties of the sources [22, 23].

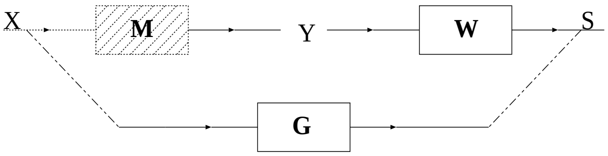

Figure 1: Mixture Model.

In this paper, we assume the following **H3**: the sources are independent nonstationary at least for second-order statistics such speech signals where the power of the signals can be considered time variant. Our first goal is to prove, using geometrical information, that for such signals, the decorrelation of the output signals at any time implies the separation of the sources. Therefore, the separation of nonstationary signals is possible using only second-order statistics. Finally, simple algorithms for speech or music signals and the performances are also discussed.

2 Channel Model

Let $X(n)$ be a $p \times 1$ zero-mean random vector denoting the source vector at time n. Let $Y(n)$ denote the observed (or mixture) signals (see Fig. 1) at time n. According to the instantaneous model,

$$Y(n) = \mathbf{M} \quad X(n), \tag{1}$$

where $\mathbf{M} = (m_{ij})$ is a $p \times p$ full-rank (non-singular) matrix which represents the unknown mixture.

Let $\mathbf{W} = (w_{ij})$ denote the $p \times p$ weight matrix. The estimated sources are given by

$$S(n) = \mathbf{W} \quad Y(n) = \mathbf{WM} \quad X(n) = \mathbf{G} \quad X(n), \tag{2}$$

where $\mathbf{G} = \mathbf{WM}$ is the global matrix. It is obvious that by only using the source independence assumption and model (1), we cannot exactly retrieve the sources ($S(n) \neq X(n)$). Generally, we can separate the sources up to a permutation and scales [24]. The separation is considered to be achieved when the global matrix becomes

$$\mathbf{G} = \mathbf{WM} = \mathbf{P\Delta}, \tag{3}$$

where $\mathbf{P}$ is any $p \times p$ permutation matrix and Δ is any $p \times p$ diagonal full-rank matrix.

3 Decorrelation and Separation

In this section, it is proved that one can separate nonstationary signals using only the second-order statistics of the estimated signals (i.e., the decorrelation of the covariance matrix of the output signals). To simplify this idea and to explain the geometrical solutions of this problem, let us first consider the case of two sensors and two sources.

3.1 First Case: Two Sources

In this subsection, we consider that there are two sensors and two sources (i.e., $p = 2$). In the previous section, it was mentioned that the separation is achieved when the global matrix becomes the product of any permutation matrix and any non-singular diagonal matrix, as in (3), thus one can use the value of $w_{ii} = 1$ without any loss of generality. Using (3), the global matrix can be rewritten as

$$\mathbf{G} = \begin{pmatrix} m_{11} + m_{21}w_{12} & m_{12} + m_{22}w_{12} \\ m_{21} + m_{11}w_{21} & m_{12}w_{21} + m_{22} \end{pmatrix}. \tag{4}$$

Supposing that one can achieve decorrelation of the output signals $S(n)$ and using assumption **H1**, it is possible to prove that the coefficients of the weight matrix satisfy the following condition:

$$\begin{aligned} &\mathrm{E}\{s_1(n) \quad s_2(n)\} = 0 \Longrightarrow \\ &\quad (m_{11} + m_{21}w_{12})(m_{21} + m_{11}w_{21})P_1 + \\ &\qquad (m_{21} + m_{11}w_{21})(m_{12}w_{21} + m_{22})P_2 = 0, \end{aligned} \tag{5}$$

where $\mathrm{E}\{x(n)\}$ is the expectation of $x(n)$ and $P_i = \mathrm{E}\{x_i^2(n)\}$ is the power of the i-th source $x_i(n)$. When the sources are stationary then the powers P_i stet constant. In this case, condition (5) is the equation of a hyperbola. At the convergence, the point (w_{12}, w_{21}) can be any point on the hyperbola. Therefore, separation cannot be achieved by using only second-order statistics.

In the general case, using assumptions **H1** and **H2**, one can also assume hereafter the following **H4**: the ratio of two signal powers P_i is also time variant (the two powers P_i cannot have a linear relationship). Since condition (5) must be satisfied for any value of $P_i > 0$, the weight matrix coefficients must satisfy the following conditions:

$$(m_{11} + m_{21}w_{12})(m_{21} + m_{11}w_{21}) = 0, \tag{6}$$

$$(m_{21} + m_{11}w_{21})(m_{12}w_{21} + m_{22}) = 0. \tag{7}$$

The solutions of equations (6) and (7) must be considered for the following three cases

- The coefficients of the mixture matrix are nonzero ($m_{ij} \neq 0$). Using equations (6) and (7), the coefficient W_{ij} can be evaluated as

$$w_{12} = -\frac{m_{11}}{m_{21}} \quad \text{and} \quad w_{21} = -\frac{m_{22}}{m_{12}}, \tag{8}$$

Or

$$w_{12} = -\frac{m_{12}}{m_{22}} \quad \text{and} \quad w_{21} = -\frac{m_{21}}{m_{11}}. \tag{9}$$

  In both (8) and (9), the separation of sources can be achieved (i.e., the global matrix $\mathbf{G}$ satisfies equation (3)).

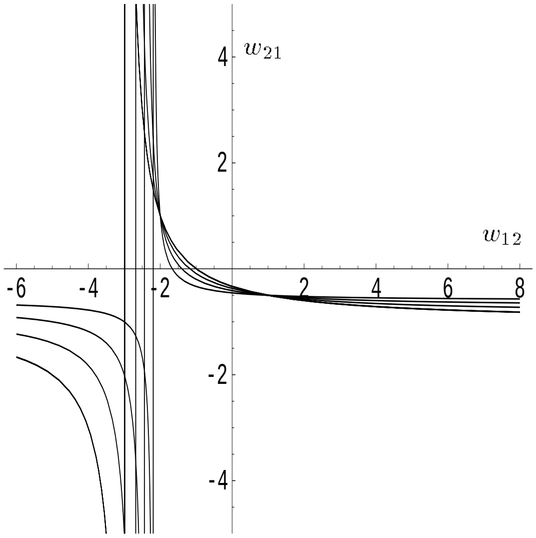

Figure 2: A set of hyperbolas, with the same mixing matrix and different stationary sources.

- One coefficient of the mixture matrix is equal to zero (for example $m_{11} = 0$). Using (6) and (7), we can write

$$w_{12} = 0 \text{ and } w_{21} = -\frac{m_{22}}{m_{12}}. \tag{10}$$

  In this case separation is also achieved.

- If more than one coefficient of the mixture matrix are equal to zero then $\mathbf{M}$ will become a permutation matrix, under the assumption that $\mathbf{M}$ is a full-rank nonsingular matrix. In this case, there is no mixture problem.

Figure 2 shows hyperbolas corresponding to the solutions of equation (5) for mixing matrix $\mathbf{M} = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}$ and different stationary sources. All of the hyperbolas have two intersection points corresponding to (8) and (9).

3.2 General Case

Let $\mathbf{\Lambda}$ denote the covariance matrix of the sources. Using assumption **H1**, we can deduce that $\mathbf{\Lambda}$ is a diagonal matrix, $\mathbf{\Lambda} = \text{diag}(P_1, \dots, P_p)$. After the decorrelation of the output signals $S(n)$, their covariance matrix becomes a diagonal one:

$$\text{E}\{S(n) \quad S(n)^T\} = \mathbf{G\Lambda G}^T = \mathbf{D}, \tag{11}$$

where $\mathbf{D} = (d_{ij})$ is any diagonal matrix. From last equation (11), we can deduce that $\mathbf{G}$ is an orthogonal matrix and we can prove that

$$g_{il}^2 P_l = d_{ii}, \tag{12}$$

$$\sum_l g_{il} g_{jl} P_l = 0 \quad \forall l, \text{ and } i \neq j. \tag{13}$$

Generally the orthogonality of $\mathbf{G}$ is not great enough to separate the sources. In the case of nonstationary signals, the covariance matrix $\mathbf{\Lambda}$ changes with time. This means that equation (13) must hold for any value of P_i (her P_i are assumed to be independently changing with time). Thus we can deduce that

$$g_{il} g_{jl} = 0 \quad \forall l, \text{ and } i \neq j. \tag{14}$$

Equation (14) implies the following:

- **P1: All columns of G have at most one nonzero coefficient.**

- **P2: All the rows of G have at least one nonzero coefficient.**: In fact, let G_i (respectively W_i) denotes the i-th row of $\mathbf{G}$ (respectively of $\mathbf{W}$) and let us put $w_{ii} = 1$, as in the previous sub-section. Using equation (2), one can write

$$G_i = W_i.\mathbf{M}. \tag{15}$$

  Using equation (15), and the conditions that $w_{ii} = 1$ (i.e., $W_i \neq 0$) and $\mathbf{M}$ is a full-rank matrix, we can deduce that G_i cannot be a zero vector and proposition **P2** is valid.

- Propositions **P1** and **P2** imply the following:

  P3: Each column of G has only one nonzero coefficient or G satisfies the condition (3).

  P3 simply means that separation can be achieved using second-order statistics.

4 Algorithms & Experimental Results

In this section, we discuss three possible approaches to the blind separation of nonstationary sources by using only second-order statistics

4.1 Jacobi Diagonalization

The first approach is based on the Jacobi Diagonalization [25] and the Joint Diagonalization [26]. Let us denote by $\mathbf{R} = (r_{ij})$ a $p \times p$ full rank matrix and let $\mathbf{J}(m, n, \theta)$ be a Givens[1] rotations matrix.

By definition the Off function of a matrix $\mathbf{R}$ is:

$$\text{Off}(\mathbf{R}) = \sqrt{\sum_{i=1}^{p} \sum_{j=1, j\neq i}^{p} r_{ij}^2} \tag{16}$$

[1] The Givens rotations $\mathbf{J}(m, n, \theta) = (J_{ij})$ are similar to identity matrix except for the four elements $J_{mm} = J_{nn} = \cos\theta$ and $J_{mn} = -J_{nm} = \sin\theta$. The Givens rotations are also denoted by Jacobi rotations.

It is clear that the Off($\mathbf{R}$) is equal to zero when $\mathbf{R}$ is a diagonal matrix. The Jacobi method seeks for a set of Givens rotations matrix $\mathbf{J}(m,n,\theta)$ that minimize the Off function of $\mathbf{J}^T(m,n,\theta)\mathbf{R}\mathbf{J}(m,n,\theta)$. Using the same idea, the Cyclic Jacobi method [25] applied to a symmetric matrix $\mathbf{R}$ gives an orthogonal matrix $\mathbf{V}$ such that Off($\mathbf{V}^T\mathbf{R}\mathbf{V}$) $\leq tol||\mathbf{R}||_F$, here $tol > 0$ is the tolerance and $||\mathbf{R}||_F$ is the Frobenius norm[2].

According to the previous section, one can separate non-stationary sources (speech or music) from an instantaneous mixture by looking for a weight matrix $\mathbf{W}$ that can diagonalize the covariance matrix of the output signals. Unfortunately, the Cyclic Jacobi method can not directly be used to achieve our goal because the sources are assumed to be a second order non-stationary signals, therefore the covariance matrix of such signals are time variant.
On the other hand, using the joint diagonalization algorithm proposed by cardoso and soulamic [26], one can jointly diagonalize a set of q covariance matrix $\mathbf{R}_i = \mathrm{E}\{S(n)S(n)^T\}$, here $1 \leq i \leq q$. The joint diagonalization algorithm is a modified version of the cyclic Jacobi method that minimize the following function with respect to a matrix $\mathbf{V}$:

$$\mathrm{JOff}(\mathbf{R}_1, \cdots, \mathbf{R}_q) = \sum_i Off(\mathbf{V}^T\mathbf{R}_i\mathbf{V}) \qquad (17)$$

It is obvious that $JOff(\mathbf{R}_1,\cdots,\mathbf{R}_q) = 0$ when $\mathbf{V}^T\mathbf{R}_i\mathbf{V}$ is a diagonal matrix for every i. Because the estimation error and the noise, one can not minimize $JOff(\mathbf{R}_1,\cdots,\mathbf{R}_q)$ to the lower limit (i.e 0).

In our experimental study, the number q of the covariance matrices $\mathbf{R}_i$ has been chosen between 10 and 25. The covariance matrices $\mathbf{R}_i$ have been estimated according to the adaptive estimator of [27] over some sliding windows of 500 to 800 samples and shifted 100 to 200 samples for each $\mathbf{R}_i$. All the previous limits have been determined by an experimental study using our data base signals.

In addition, we should mention that we used a threshold to reduce the silence effect: When ever the observation signals at time n_0 is less than the predefined threshold ϵ, it will not be considered as input signals: If the observation signals at time n_0 is less than the predefined threshold ϵ that means two things:

1. That the sources are in common silence period, i.e we are receiving just noise signals.

2. The samples of the sources at time n_0 have some relationship: For example, in the case of two sources, the first observation signal should be $y_1(n_0) = m_{11}s_1(n_0) + m_{12}s_2(n_0) < \epsilon$. Using the independence assumption **H1**, one can consider, without loss of generality, that the probability to have such instant n_0 is so small and it has no effective effect on the signal statistics or on the behavior of the algorithm.

[2] The Frobenius norm of a $p \times p$ matrix $\mathbf{R} = (r_{ij})$ is $||\mathbf{R}||_F = \sqrt{\sum_{i=1}^{p}\sum_{j=1}^{p} r_{ij}^2}$.

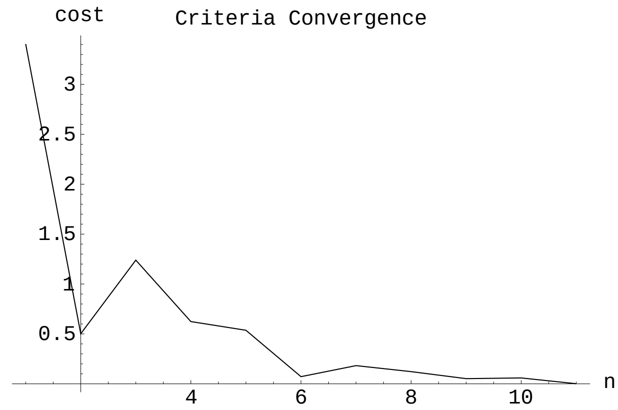

Figure 3: Evaluation of the cost function with respect to the iteration number.

We conducted many experiments and found that the crosstalk was between -17 dB and -25 dB. Fig. 3 shows the evaluation of the cost function with respect to the iteration number. The experimental study shows that the convergence of this algorithm are obtained in few iterations. Fig. 4 shows the experimental results of the separation of two speech sources.

Finally, we should mention that the first one who suggest the separation by multi-diagonalization of the covariance matrix was Fety [20]. The approach of Fety have been the subject of research and discussion of many other researchers: It has been discussed and improved by Comon *et al.* [28, 29, 30]. Recently, Belouchrani *et al.* presented an algorithm based on the approach of Fety and the Joint Diagonalization [31, 32, 33] to separate stationary correlated (in time) and independent (in space) sources signals from an instantaneous mixture. In [33] Belouchrani *et al.* discuss the performances of their algorithm and prove the convergence of such approach.

4.2 Kull-back divergence

The second approach is based on the Kull-back distance. The Kull-back distance (or divergence) of two probability density functions (pdf) f_x and f_y is given by [34]

$$\delta(f_x, f_y) = \int f_x(u) \log\left(\frac{f_x(u)}{f_y(u)}\right) \mathrm{d}u. \qquad (18)$$

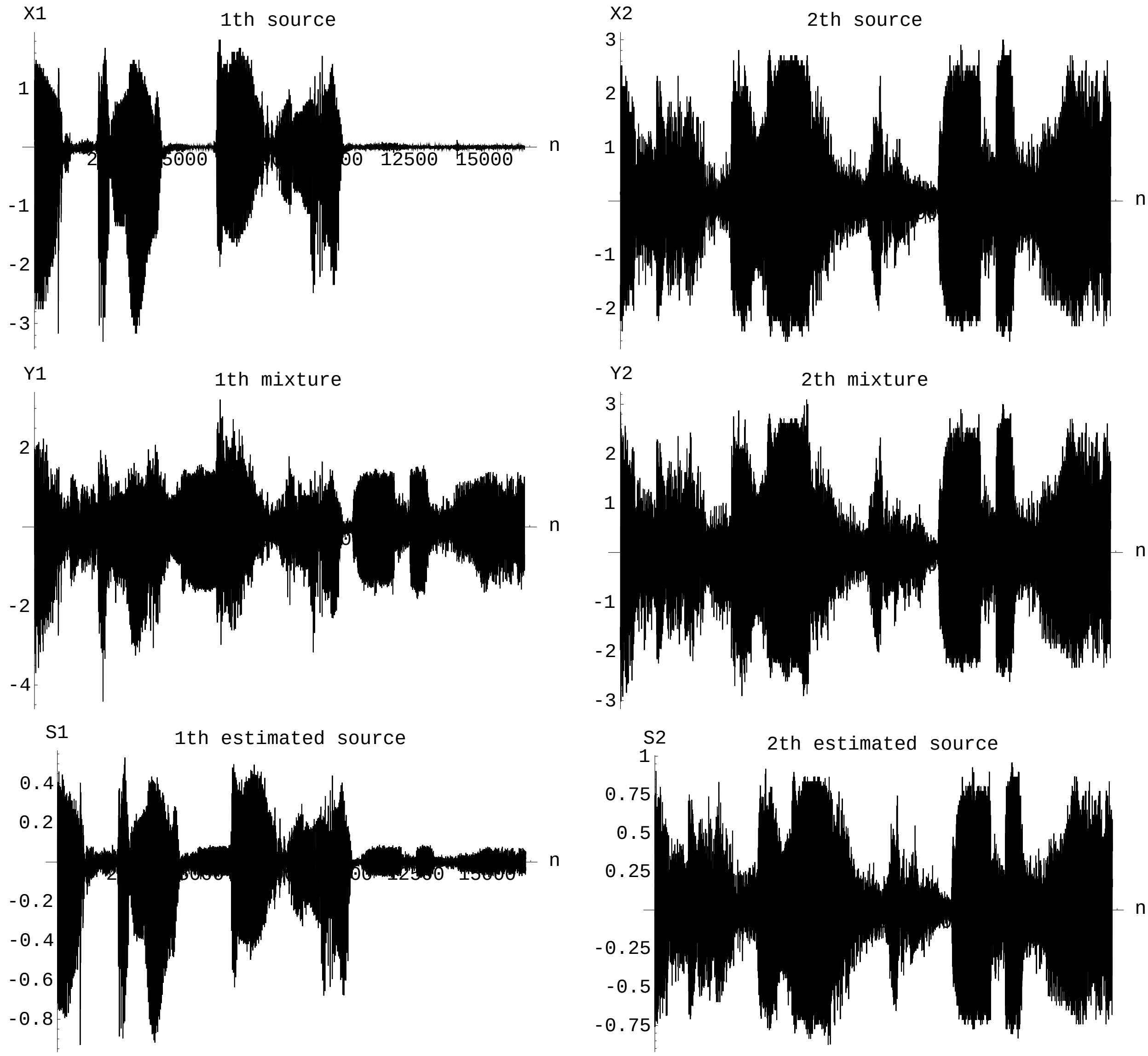

Figure 4: First column contains the signals of the first channel (i.e., first source, first mixture signal and the first estimated source), the second column contains the signals of the second channel.

It is known [35], that the kull-back divergence between two random zero mean Gaussian vectors V_1 and V_2 is given by

$$\delta(\mathbf{R},\ \mathbf{I}) = \frac{1}{2}(\mathrm{Trace}\{\mathbf{R}\} - \log\det(\mathbf{R})) \geq 0, \qquad (19)$$

where $\mathbf{I}$ is the $p \times p$ identity matrix, and $\mathbf{R} = \mathrm{E}\{S(n)\ \ S(n)^T\}$ is the $p \times p$ covariance matrix of the estimated sources $S(n)$. One of the kull-back divergence properties is that

$$\delta(\mathbf{R},\ \mathbf{I}) = 0 \quad \text{iff} \quad \mathbf{R} = \mathbf{I}. \qquad (20)$$

Thus the minimization of divergence (19) makes the matrix $\mathbf{R}$ close to an identity matrix (i.e., a diagonal matrix) and induces the separation of the sources, as we explained in the previous section.

The minimization of divergence (19) is achieved according to the natural gradient [36, 37]. In this case the weight matrix $\mathbf{W}$ can be updated at iteration $(k+1)$ by

$$\mathbf{W}_{k+1} = \mathbf{W}_k - \lambda\{\mathbf{R} - \mathbf{I}\}\mathbf{W}_k, \qquad (21)$$

where $0 < \lambda < 1$ is a scale parameter. $\mathbf{R}$ is estimated of $\mathbf{R}$ in slide windows of a small number of samples, according to the method described in [27].

The advantage of this approach is that the algorithm and the updating rules are simple. However the convergence point of this criterion (19) is a $\mathbf{W}^*$ that makes the matrix $\mathbf{R}$ close to an identity matrix (i.e., a special diagonal matrix). It is obvious that this condition is more restrictive than the initial condition described in the previous section where $\mathbf{R}$ must simply be a di-

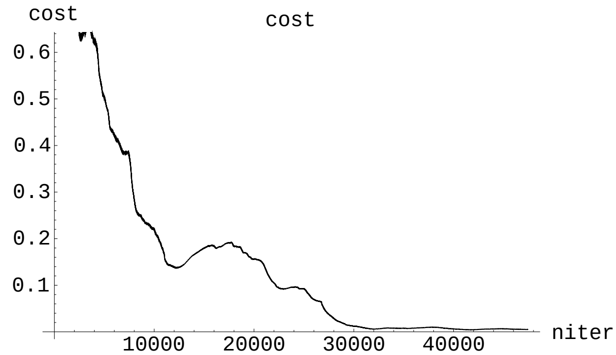

Figure 5: Evaluation of the cost function with respect to the iteration number.

agonal matrix. We conducted many experiments and found that the crosstalk was between -15 dB and -23 dB. The evaluation of the cost function with respect to the iteration number is shown in Fig. 5. The mixing matrix used was $\mathbf{M} = \begin{pmatrix} 1 & -0.6 \\ 0.4 & 1 \end{pmatrix}$.

Fig. 6 shows the experimental results of the separation of two speech sources.

4.3 Hadamard's inequality

The last approach is based on Hadamard's inequality, Hadamard's inequality [38] of an arbitrary positive semidefinite matrix $\mathbf{R} = (r_{ij})$ is given by

$$\prod_{i=1}^{p} r_{ii} \geq \det\{\mathbf{R}\}, \tag{22}$$

where the equality holds if and only if the matrix $\mathbf{R}$ is a diagonal matrix. Using equation (22), it can be proven that:

$$\sum_{i=1}^{p} \log r_{ii} - \log\det\{\mathbf{R}\} \geq 0. \tag{23}$$

Using this property, some authors [22, 23, 39] suggest the separation of nonstationary signals by minimizing a modified version of Hadamard's inequality (23) of the estimated source's covariance matrix $\mathbf{R} = \mathrm{E}\{S(n) \quad S(n)^T\}$ with respect to the weight matrix $\mathbf{W}$

$$\min_{\mathbf{W}} \sum_{i=1}^{p} \log \mathrm{E}\{s_i^2(n)\} - \log\det\{\mathrm{E}\{S(n) \quad S^T(n)\}\}, \tag{24}$$

The experimental results of this kind of algorithm are discussed in [22, 40].

5 Conclusion

In this paper, we proved that second-order statistics are sufficient to separate the instantaneous mixture of independent nonstationary signals and that the decorrelation is equivalent to the separation when the sources satisfy assumptions **H1** to **H4**. The study was divided into two parts,

- In the case of two sources, using the geometrical information of the mixing signals, we prove that one can decorrelate the stationary signals or separate the nonstationary signals by using only second-order statistics.
- For the general case, we proved that the diagonalization of the autocorrelation matrix can separate nonstationary signals.

Finally, the application of these theoretical results in a real world situation was discussed by examining three possible approaches. In addition, we should mention that the first algorithm converge in few iterations but it needs more computation effort than the second one. In the other hand, the experimental study shows that the convergence of the second one needs much more iteration to converge than the first one. The comparison among these three algorithms and theirs performances will be the subject of a submitted paper [41].

References

[1] A. Mansour, A. Kardec Barros, and N. Ohnishi, "Blind separation of sources: Methods, assumptions and applications.," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E83-A, no. 8, pp. 1498–1512, 2000, Special Section on Digital Signal Processing in IEICE EA.

[2] C. Jutten and J. F. Cardoso, "Separation of sources: Really blind ?," in *International symposium on nonlinear theory and its applications*, Las Vegas, Nevada, U. S. A., December 1995.

[3] J. F. Cardoso and P. Comon, "Tensor-based independent component analysis," *Signal Processing*, vol. 5, pp. 673–676, 1990, Theories and Applications.

[4] A. Mansour and C. Jutten, "Fourth order criteria for blind separation of sources," *IEEE Trans. on Signal Processing*, vol. 43, no. 8, pp. 2022–2025, August 1995.

[5] G. Puntonet, C., A. Mansour, and C. Jutten, "Geometrical algorithm for blind separation of sources," in *Actes du XVème colloque GRETSI*, Juan-Les-Pins, France, 18-21 September 1995, pp. 273–276.

[6] C. G. Puntonet, A. Prieto, C. jutten, M. Rodriguez-Alvarez, and J. Ortega, "Separation of sources: A geometry-based procedure for reconstruction of n-valued signals," *Signal Processing*, vol. 46, no. 3, pp. 267–284, 1995.

[7] L. Nguyen Thi and C. Jutten, "Blind sources separation for convolutive mixtures," *Signal Processing*, vol. 45, no. 2, pp. 209–229, 1995.

[8] M. Krob and M. Benidir, "Fonction de contraste pour l'identification aveugle d'un modéle linéaire quadratique," in *Actes du XIVème colloque GRETSI*, Juan-Les-Pins, France, September 1993, pp. 101–104.

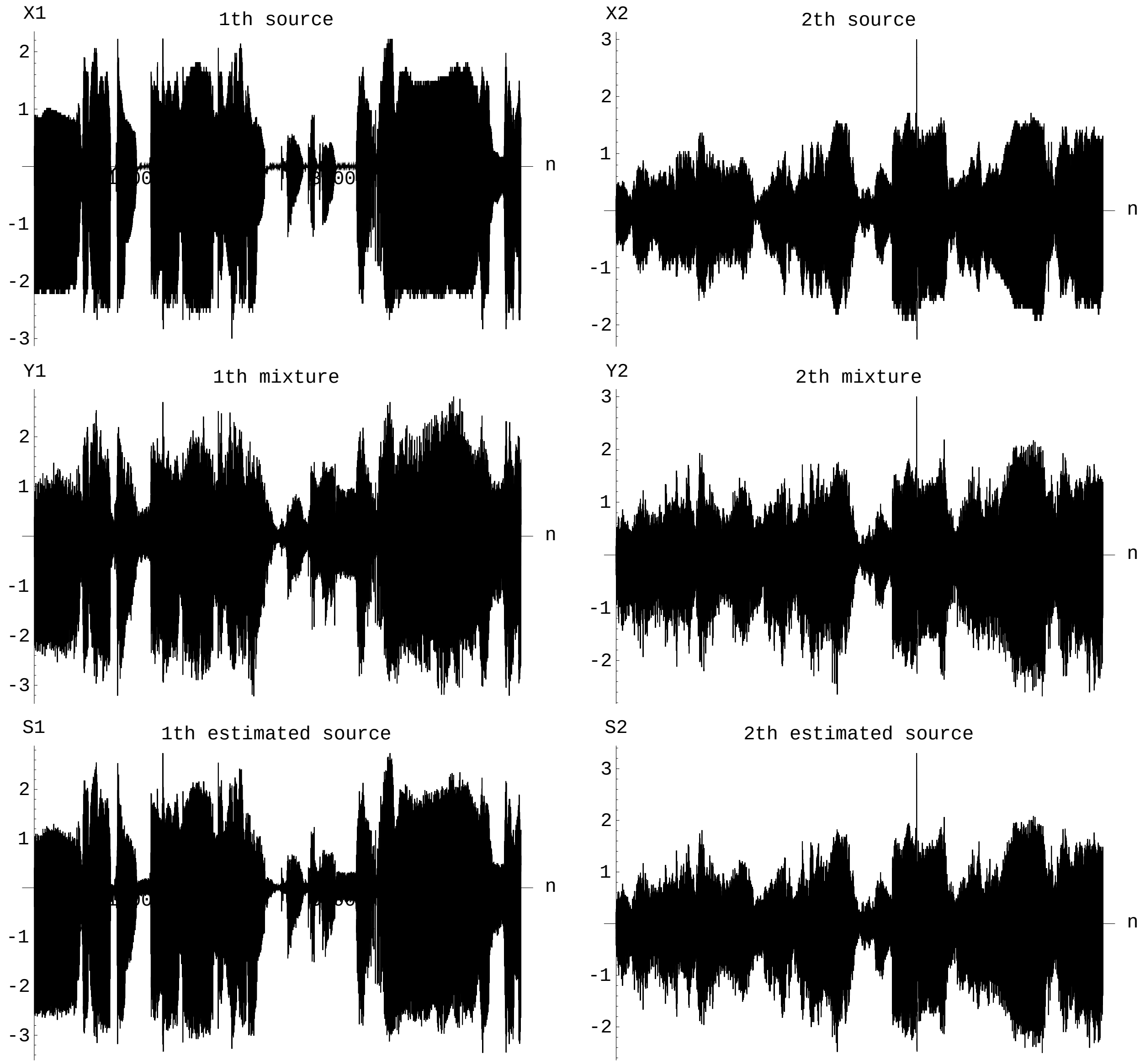

Figure 6: First column contains the signals of the first channel (i.e., first source, first mixture signal and the first estimated source), the second column contains the signals of the second channel.

[9] A. Taleb and Ch. Jutten, "Batch algorithm for source separation in postnonlinear mixtures," in *First International Workshop on Independent Component Analysis and signal Separation (ICA99)*, Aussois, France, 11-15 January 1999, pp. 155–160.

[10] 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.

[11] S. I. Amari, A. Cichoki, and H. H. Yang, "A new learning algorithm for blind signal separation," in *Neural Information Processing System 8*, Eds. D.S. Toureyzky et. al., 1995, pp. 757–763.

[12] N. Delfosse and P. Loubaton, "Adaptive blind separation of independent sources: A deflation approach," *Signal Processing*, vol. 45, no. 1, pp. 59–83, July 1995.

[13] O. Macchi and E. Moreau, "Self-adaptive source separation by direct and recursive networks," in *in Proc. International Conference on Digital Signal Processing (DSP'95)*, Limasol, Cyprus, May 1993, pp. 1154–1159.

[14] A. Mansour and N. Ohnishi, "Multichannel blind separation of sources algorithm based on cross-cumulant and the levenberg-marquardt method.," *IEEE Trans. on Signal Processing*, vol. 47, no. 11, pp. 3172–3175, November 1999.

[15] A. Mansour, "A batch subspace ica algorithm.," in *10th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, Pocono Manor Inn, Pennsylvania, USA, 14 - 16 August 2000, pp. 63–67.

[16] A. Belouchrani and J. F. Cardoso, "Maximum likelihood source separation for discrete sources," in *Signal Processing VII, Theories and Applications*, M.J.J.

Holt, C.F.N. Cowan, P.M. Grant, and W.A. Sandham, Eds., Edinburgh, Scotland, September 1994, pp. 768–771, Elsevier.

[17] F. Gamboa and E. Gassiat, "Source separation when the input sources are discrete or have constant modulus," *IEEE Trans. on Signal Processing*, vol. 45, no. 12, pp. 3062–3072, December 1997.

[18] A. Gorokhov and P. Loubaton, "Subspace based techniques for second order blind separation of convolutive mixtures with temporally correlated sources," *IEEE Trans. on Circuits and Systems*, vol. 44, pp. 813–820, September 1997.

[19] A. Mansour, C. Jutten, and P. Loubaton, "An adaptive subspace algorithm for blind separation of independent sources in convolutive mixture," *IEEE Trans. on Signal Processing*, vol. 48, no. 2, pp. 583–586, February 2000.

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

[21] S. I. Amari, "Ica of temporally correlated signals - learning algorithm," in *First International Workshop on Independent Component Analysis and signal Separation (ICA99)*, Aussois, France, 11-15 January 1999, pp. 13–18.

[22] K. Matsuoka, M. Oya, and M. Kawamoto, "A neural net for blind separation of nonstationary signals," *Neural Networks*, vol. 8, no. 3, pp. 411–419, 1995.

[23] M. Kawamoto, K. Matsuoka, and M. Oya, "Blind separation of sources using temporal correlation of the observed signals," *IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E80-A, no. 4, pp. 111–116, April 1997.

[24] P. Comon, "Independent component analysis, a new concept?," *Signal Processing*, vol. 36, no. 3, pp. 287–314, April 1994.

[25] G. H. Golub and C. F. Van Loan, *Matrix computations*, The johns hopkins press- London, 1984.

[26] J. F. Cardoso and A. Soulamic, "Jacobi angles for simultaneous diagonalization," *SIAM*, vol. 17, no. 1, pp. 161–164, 1996.

[27] A. Mansour, A. Kardec Barros, and N. Ohnishi, "Comparison among three estimators for high order statistics.," in *Fifth International Conference on Neural Information Processing (ICONIP'98)*, Kitakyushu, Japan, 21-23 October 1998, pp. 899–902.

[28] P. Comon and Lacoume J. L., "Statistiques d'ordres supérieurs pour le traitement du signal," in *Ecole Predoctorale de physique, Traitement du signal-Développements Récents*, Les Houches, France, August-September 1993.

[29] P. Comon, "Remarque sur la diagonilisation tensorielle par la méthode de jacobi," in *Actes du XIVème colloque GRETSI*, Juan-Les-Pins, France, September 1993, pp. 125–128.

[30] J. L. Lacoume, P. O. Amblard, and P. Comon, *Statistiques d'ordre supérieur pour le traitement du signal*, Mason, Paris, 1997.

[31] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso, and E. Moulines, "Second-order blind separation of correlated sources," in *Int. Conf. on Digital Sig.*, Nicosia, Cyprus, july 1993, pp. 346–351.

[32] A. Belouchrani and K. Abed-Meraim, "Séparation aveugle au second ordre de sources corrélées," in *Actes du XIVème colloque GRETSI*, Juan-Les-Pins, France, September 1993, pp. 309–312.

[33] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso, and E. Moulines, "A blind separation technique using second-order statistics," *IEEE Trans. on Signal Processing*, vol. 45, pp. 434–444, 1997.

[34] M. Basseville, "Distance measures for signal processing and pattern recognition," *Signal Processing*, vol. 18, no. 4, pp. 349–369, December 1989.

[35] B. Laheld, *Séparation auto-adaptative de sources. Implantations et performances*, Ph.D. thesis, ENST Paris, 1994.

[36] S. I. Amari, "Natural gradient works efficiently in learning," *Neural Computation*, vol. 10, no. 4, pp. 251–276, 1998.

[37] J. F. Cardoso and B. Laheld, "Equivariant adaptive source separation," *IEEE Trans. on Signal Processing*, vol. 44, no. 12, December 1996.

[38] B. Noble and J. W. Daniel, *Applied linear algebra*, Prentice-Hall, 1988.

[39] H. C. Wu and J. C. Principe, "Simultaneous diagonalization in the frequency domain (sdif) for source separation," in *First International Workshop on Independent Component Analysis and signal Separation (ICA99)*, Aussois, France, 11-15 January 1999, pp. 245–250.

[40] M. Kawamoto, A. Kardec Barros, A. Mansour, K. Matsuoka, and N. Ohnishi, "Real world blind separation of convolved non-stationary signals.," in *First International Workshop on Independent Component Analysis and signal Separation (ICA99)*, Aussois, France, 11-15 January 1999, pp. 347–352.

[41] A. Mansour, M. Kawamoto, and N. Ohnishi, "Blind separation for instantaneous mixture of speech signals: Algorithms and performances.," in *Intelligent Systems and Technologies for the Next Millennium (TENCON 2000)*, Kuala Lumpur, Malaysia, 24-27 September 2000, To appear.