# An Algebraic Non Orthogonal Joint Block Diagonalization Algorithm for Blind Separation of Convolutive Mixtures of Sources

Hicham Ghennioui[2,3], El Mostafa Fadaili[1],
Nadège Thirion-Moreau[2], Abdellah Adib[3,4], and Eric Moreau[2]

[1] IBISC, CNRS FRE 2873 40 rue du Pelvoux, F-91020 Evry-Courcouronnes, France
[2] STD, ISITV, av. G. Pompidou, BP56, F-83162 La Valette du Var Cedex, France
ghennioui@gmail.com, {fadaili,thirion,moreau}@univ-tln.fr
[3] GSCM-LRIT, FSR, av. Ibn Battouta, BP1014, Rabat, Maroc
[4] DPG, IS, av. Ibn Battouta, BP703, Rabat, Maroc
adib@israbat.ac.ma

**Abstract.** This paper deals with the problem of the blind separation of convolutive mixtures of sources. We present a novel method based on a new non orthogonal joint block diagonalization algorithm ($\mathsf{NO}-\mathsf{JBD}$) of a given set of matrices. The main advantages of the proposed method are that it is more general and a preliminary whitening stage is no more compulsorily required. The proposed joint block diagonalization algorithm is based on the algebraic optimization of a least mean squares criterion. Computer simulations are provided in order to illustrate the effectiveness of the proposed approach in three cases: when exact block-diagonal matrices are considered, then when they are progressively perturbed by an additive Gaussian noise and finally when estimated correlation matrices are used. A comparison with a classical orthogonal joint block-diagonalization algorithm is also performed, emphasizing the good performances of the method.

## 1 Introduction

In the signal processing community, many works have been recently dedicated to the study of the problem of joint decomposition of matrices or tensors because of their numerous applications especially in blind source separation and array processing [1]-[14].

Here, we are interested in the problem of the blind separation of convolutive mixtures of sources. That is why this communication is dedicated to the so-called joint block-diagonalization of matrices problem. In such a decomposition, the wanted matrices are block diagonal ones[1]. Such a problem has been already considered in [1][4][7] but under the constraint that the joint-block diagonalizer is an orthogonal (unitary in the complex case) matrix. Our purpose, here, is to

[1] A block diagonal matrix is a block matrix in which the off-diagonal block terms are zero matrices and the diagonal matrices are square.

M.E. Davies et al. (Eds.): ICA 2007, LNCS 4666, pp. 193–200, 2007.
© Springer-Verlag Berlin Heidelberg 2007

discard this unitary constraint. To that aim, we show how the (non necessarily orthogonal) joint-block diagonalizer can be algebraically estimated by minimizing a least mean squares criterion, leading to a new non-orthogonal joint block-digonalization algorithm. Some computer simulations are provided in order to illustrate the good behaviour of the proposed algorithm. Then, it is shown how this algorithm finds application in blind source separation where it is applied, here, to a set of observations correlation matrices at different time delays.

The rest of this communication is organized as follows. The problem statement and the proposed joint block-diagonalization algorithm are both introduced in the Section 2. In the Section 3, we show how this algorithm can be applied to solve the problem of blind separation of convolutive mixtures of sources. Computer simulations are provided in both sections to illustrate the effectiveness of the proposed algorithm and to compare it with another one based on an orthogonal joint block-diagonalization.

## 2 Non-orthogonal Joint Block-Diagonalization Problem

### 2.1 Problem Statement

The non-orthogonal joint block-diagonalization problem is stated in the following way: let us consider a set $\mathcal{M}$ of $N_m$, $N_m \in \mathbb{N}^*$ square invertible matrices $\mathbf{M}_i \in \mathbb{R}^{M\times M}$, $i \in \{1,\ldots,N_m\}$ which all admit the following decomposition:

$$\mathbf{M}_i = \mathbf{A}\mathbf{D}_i\mathbf{A}^T \;, \quad \text{or} \quad \mathbf{D}_i = \mathbf{B}\mathbf{M}_i\mathbf{B}^T \;, \quad \forall i \in \{1,\ldots,N_m\} \tag{1}$$

where $\mathbf{D}_i = \begin{pmatrix} \mathbf{D}_{i1} & \ldots & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & \ldots & \mathbf{D}_{ir} \end{pmatrix}$, $\forall i \in \{1,\ldots,N_m\}$, are $N \times N$ block diagonal matrices with $\mathbf{D}_{ij}, i \in \{1,\ldots,N_m\}, j \in \{1,\ldots,r\}$ are $n_j \times n_j$ square matrices so that $n_1+\ldots+n_r = N$ (in our case, we will assume that all the matrices have the same size *i.e* $N = r \times n_j, \forall j \in \{1,\ldots,r\}$) and where $\mathbf{0}$ denotes the $n_j \times n_j$ null matrix. $\mathbf{A}$ is the $M \times N$ $(M \geq N)$ full rank matrix and $\mathbf{B}$ is its pseudo-inverse (or generalized Moore-Penrose inverse).

The non-orthogonal joint block-diagonalization problem consists in estimating the matrix $\mathbf{A}$ and the matrices $\mathbf{D}_{ij}$, $i \in \{1,\ldots,N_m\}, j \in \{1,\ldots,r\}$ (or more simply the matrix $\mathbf{B}$ only) from the matrices set $\mathcal{M}$. The case of an orthogonal matrix $\mathbf{A}$ has been already considered in [7] where a first solution is proposed.

### 2.2 Joint Block-Diagonalization Algorithm

In this communication, we propose to consider the following cost function

$$\mathcal{C}_{BD}(\mathbf{C}) = \sum_{k=1}^{N_m} \|\mathsf{OffBdiag}\{\mathbf{C}^T\mathbf{M}_k\mathbf{C}\}\|^2, \tag{2}$$

where the operator $\mathsf{OffBdiag}\{\cdot\}$ denotes the zero-block-diagonal matrix and $\mathbf{C} = \mathbf{B}^T$. Thus:

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} & \dots & \mathbf{M}_{1r} \\ \mathbf{M}_{21} & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \vdots \\ \mathbf{M}_{r1} & \mathbf{M}_{r2} & \dots & \mathbf{M}_{rr} \end{pmatrix} \Rightarrow \mathsf{OffBdiag}\{\mathbf{M}\} = \begin{pmatrix} \mathbf{0} & \mathbf{M}_{12} & \dots & \mathbf{M}_{1r} \\ \mathbf{M}_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathbf{M}_{r1} & \mathbf{M}_{r2} & \dots & \mathbf{0}, \end{pmatrix}. \quad (3)$$

Let $\mathbf{C} = [\mathbf{C}_1, \cdots, \mathbf{C}_r]$, where $\mathbf{C}_j, j \in \{1, \cdots, r\}$, are $r$ block matrices of dimension $M \times n_j$ . The cost function (2) can be rewritten as:

$$\mathcal{C}_{BD}(\mathbf{C}) = \sum_{k=1}^{N_m} \sum_{i,j=1(i \neq j)}^{r} \|\mathbf{C}_i^T \mathbf{M}_k \mathbf{C}_j\|^2 = \sum_{k=1}^{N_m} \sum_{m=1}^{n_i} \sum_{n=1}^{n_j} \sum_{i,j=1(i \neq j)}^{r} |(\mathbf{c}_i^m)^T \mathbf{M}_k \mathbf{c}_j^n|^2 \quad (4)$$

where $\mathbf{c}_j^n$, $\forall n \in \{1, \dots, n_j\}$ stand for the $n_j$ column vectors of matrices $\mathbf{C}_j$, $\forall j \in \{1, \dots, r\}$. Then:

$$\begin{aligned} \mathcal{C}_{BD}(\mathbf{C}) &= \sum_{k=1}^{N_m} \sum_{m,n=1}^{n_i,n_j} \sum_{i,j=1(i \neq j)}^{r} ((\mathbf{c}_i^m)^T \mathbf{M}_k \mathbf{c}_j^n)((\mathbf{c}_i^m)^T \mathbf{M}_k \mathbf{c}_j^n)^T \\ &= \sum_{k=1}^{N_m} \sum_{m,n=1}^{n_i,n_j} \sum_{i,j=1(i \neq j)}^{r} (\mathbf{c}_i^m)^T (\mathbf{M}_k \mathbf{c}_j^n (\mathbf{c}_j^n)^T \mathbf{M}_k^T) \mathbf{c}_i^m \\ &= \sum_{m=1}^{n_i} \sum_{i=1}^{r} (\mathbf{c}_i^m)^T \left[ \sum_{j=1(j \neq i)}^{r} \sum_{n=1}^{n_j} \sum_{k=1}^{N_m} \mathbf{M}_k \mathbf{c}_j^n (\mathbf{c}_j^n)^T \mathbf{M}_k^T \right] \mathbf{c}_i^m \\ &= \sum_{m=1}^{n_i} \sum_{i=1}^{r} (\mathbf{c}_i^m)^T \mathbf{Q}_i(\mathbf{C}_{\bar{i}}) \mathbf{c}_i^m \end{aligned} \quad (5)$$

where $\mathbf{Q}_i(\mathbf{C}_{\bar{i}}) = \sum_{j=1(j \neq i)}^{r} \sum_{n=1}^{n_j} \sum_{k=1}^{N_m} \mathbf{M}_k \mathbf{c}_j^n (\mathbf{c}_j^n)^T \mathbf{M}_k^T$ is a symmetric matrix.

As $\mathbf{c}_j^n (\mathbf{c}_j^n)^T$ is rank one, $\forall j = 1, \dots, r$, and $\forall n = 1, \dots, n_j$, the matrix $\mathbf{Q}_i(\mathbf{C}_{\bar{i}})$ possesses $N - (r-1)n_j = n_j$ eigenvectors associated with null eigenvalues. Then, the minimization of this quadratic form under the unit norm constraint can be achieved by taking the $n_j$ unit eigenvectors associated with the $n_j$ smallest eigenvalues of $\mathbf{Q}_i(\mathbf{C}_{\bar{i}})$. However since matrix $\mathbf{Q}_i$ for a given $i$ also depends on column vectors of matrix $\mathbf{C}$, we propose to use an iterative procedure. The proposed non-orthogonal joint block-diagonalization (denoted by $\mathsf{NO-JBD}$) writes:

$\forall i \in \{1, \dots, r\}$ with $l \in \mathbb{N}^*$ and given $\mathbf{C}_{\bar{i}}^{(0)}$ an initial matrix, do (a) and (b)

(a) Calculate $\mathbf{Q}_i(\mathbf{C}_{\bar{i}}^{(l)})$

(b) Find the $n_i$ lowest eigenvalues $\lambda_i^{m(l)}, m \in \{1, \dots, n_i\}$ and the associated eigenvectors $\mathbf{c}_i^{m(l)}, m \in \{1, \dots, n_i\}$ of matrix $\mathbf{Q}_i(\mathbf{C}_{\bar{i}}^{(l)})$

Stop after a given number of iterations or when $|\lambda_i^{m(l)} - \lambda_i^{m(l-1)}| \leq \varepsilon$ where $\varepsilon$ is a given small positive threshold.

### 2.3 Computer Simulations

We present simulations to illustrate the effectiveness of the proposed algorithm. We consider a set $\mathbf{D}$ of $N_m = 11$ (resp. 31, 56, 96) matrices, randomly chosen (according to a Gaussian law) of mean 0 and variance 1. Initially these matrices are exactly block-diagonal, then random noise matrices of mean 0 and variance $\sigma_b^2$ are added. A signal to noise ratio can be defined as $\mathsf{SNR} = 10\log(\frac{1}{\sigma_b^2})$. To measure the quality of the separation, the following performance index (which is an extension of the one introduced in [10]) is used:

$$I(\mathbf{G}) = \frac{1}{r(r-1)}\left[\sum_{i=1}^{r}\left(\sum_{j=1}^{r}\frac{\|(\mathbf{G})_{i,j}\|^2}{\max\limits_{\ell}\|(\mathbf{G})_{i,\ell}\|^2}-1\right)+\sum_{j=1}^{r}\left(\sum_{i=1}^{r}\frac{\|(\mathbf{G})_{i,j}\|^2}{\max\limits_{\ell}\|(\mathbf{G})_{\ell,j}\|^2}-1\right)\right]$$

where $(\mathbf{G})_{i,j}\forall i,j \in \{1,\ldots,r\}$ is the $(i,j)$-th (square) block matrix of $\mathbf{G} = \hat{\mathbf{C}}^T\mathbf{A}$. All the displayed results have been averaged over 30 Monte-Carlo trials. On the Fig. 1, the performance index of algorithm $\mathsf{NO-JBD}$ is displayed versus the number of used matrices (left) and versus the $\mathsf{SNR}$ (right). These curves illustrate the good behaviour of the algorithm since $I \approx -110$ dB at high $\mathsf{SNR}$.

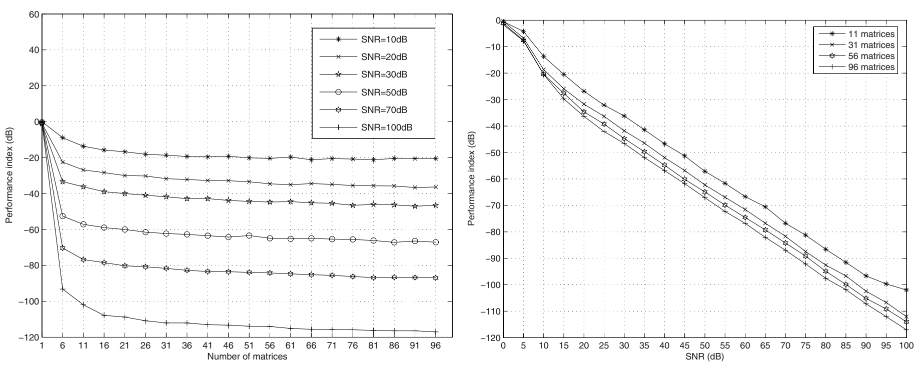


**Fig. 1.** Left: performance index versus number of matrices, right: performance index versus $\mathsf{SNR}$

## 3 Separation of Convolutive Mixtures of Sources

### 3.1 Model and Assumptions

We consider a convolutive finite-duration impulse response ($\mathsf{FIR}$) model given by

$$x_i(t) = \sum_{j=1}^{n}\sum_{\ell=0}^{L} h_{ij}(\ell)s_j(t-\ell) + n_j(t),\ \forall i = 1,\ldots,m \tag{6}$$

where $s_j(t)$, $\forall j = 1, \ldots, n$ are the $n$ sources, $x_i(t)$, $i = 1, \ldots, m$, are the $m > n$ observed signals, $h_{ij}(t)$ is the real transfer function between the $j$-th source and $i$-th sensor with an overall extent of $(L+1)$ taps. $n_i(t)$, $\forall i = 1, \ldots, m$ are additive noises. Our developments are based on the two following assumptions:

**Assumption A:** Each source signal is a real temporally coherent signal. Moreover they are uncorrelated two by two, *i.e.*, for all pairs of sources $(s_i(t), s_j(t))$ with $i \neq j$, for all time delay $\tau_{ij}$, we have $R_{ij}(t, \tau_{ij}) = 0$, where $R_{ij}(t, \tau)$ denotes the cross-correlation function between the sources $s_i(t)$ and $s_j(t)$. It is defined as follows: $R_{ij}(t, \tau) = \mathsf{E}\{s_i(t) s_j(t+\tau)\}$, where $\mathsf{E}\{.\}$ stands for the mathematical expectation.

**Assumption B:** The noises $n_i(t), i = 1, \ldots, m$, are assumed real stationary white random signals, mutually uncorrelated, independent from the sources, with the same variance $\sigma_n^2$. The noises correlation matrix can be written as:

$$\mathbf{R}_n(\tau) = \mathsf{E}\{\mathbf{n}(t)\mathbf{n}^T(t+\tau)\} = \sigma_n^2 \delta(\tau) \mathbf{I}_m \tag{7}$$

where $\delta(\tau)$ stands for the Delta impulse, $\mathbf{I}_m$ for the $m \times m$ identity matrix and $(.)^T$ for the transpose operator.

Let us now recall how the convolutive mixing model can be reformulated into an instantaneous one [4][7].

Considering the vectors $\mathbf{S}(t)$, $\mathbf{X}(t)$ and $\mathbf{N}(t)$ respectively defined as:

$$\begin{aligned}
\mathbf{S}(t) &= [s_1(t), \ldots, s_1(t-(L+L')+1), \ldots, s_n(t-(L+L')+1)]^T \\
\mathbf{X}(t) &= [x_1(t), \ldots, x_1(t-L'+1), \ldots, x_m(t-L'+1)]^T \\
\mathbf{N}(t) &= [n_1(t), \ldots, n_1(t-L'+1), \ldots, n_m(t-L'+1)]^T
\end{aligned}$$

and the $(M \times N)$ matrix $\mathbf{A}$, where $M = mL'$ and $N = n(L+L')$:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \ldots & \mathbf{A}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{m1} & \ldots & \mathbf{A}_{mn} \end{pmatrix}$$

where

$$\mathbf{A}_{ij} = \begin{pmatrix} h_{ij}(0) & \ldots\ldots & h_{ij}(L) & 0 & \ldots & 0 \\ 0 & \ddots\ \ddots & \ddots & \ddots\ \ddots & & \vdots \\ \vdots & \ddots\ \ddots & \ddots & \ddots\ \ddots & & 0 \\ 0 & \ldots\ 0 & h_{ij}(0) & \ldots\ldots & & h_{ij}(L) \end{pmatrix} \tag{8}$$

are $(L' \times (L+L'))$ matrices, the model described by Eq. (6) can be written in matrix form as:

$$\mathbf{X}(t) = \mathbf{A}\mathbf{S}(t) + \mathbf{N}(t) \tag{9}$$

In order to have an over-determined model, $L'$ must be chosen such that $mL' \geq n(L+L')$. We assume, here, that the matrix $\mathbf{A}$ is full rank. Because of the Assumption A, all the components of $\mathbf{S}(t)$ are temporally coherent signals. Moreover, two different components of this vector are correlated at least in one non null time delay. With regard to the noise vector $\mathbf{N}(t)$, the Assumption B holds for each of its components involving that its correlation matrix $\mathbf{R}_{\mathbf{N}}(\tau)$ reads:

$$
\begin{aligned}
\mathbf{R}_N(\tau) &= \mathsf{E}\{\mathbf{N}(t)\mathbf{N}^T(t+\tau)\} \\
&= \begin{pmatrix}
\sigma_n^2\tilde{\mathbf{I}}_{L'}(\tau) & \mathbf{0}_{L'} & \dots & \mathbf{0}_{L'} \\
\mathbf{0}_{L'} & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \mathbf{0}_{L'} \\
\mathbf{0}_{L'} & \dots & \mathbf{0}_{L'} & \sigma_n^2\tilde{\mathbf{I}}_{L'}(\tau)
\end{pmatrix}
\end{aligned}
\tag{10}
$$

where $\tilde{\mathbf{I}}_{L'}(\tau)$ is the $L' \times L'$ matrix which contains ones on the $\tau^{th}$ superdiagonal if $0 \leq \tau < L'$ or on the $|\tau|^{th}$ subdiagonal if $-L' \leq \tau \leq 0$ and zeros elsewhere. Then, we have:

$$
\mathbf{R}_X(t,\tau) - \mathbf{R}_N(\tau) = \mathbf{A}\mathbf{R}_S(t,\tau)\mathbf{A}^T = \mathbf{R}_Y(t,\tau) \tag{11}
$$

Because sources signals are spatially uncorrelated and temporally coherent, the matrices $\mathbf{R}_S(t,\tau)$, $\forall\tau$ are block diagonal matrices. To recover the mixing matrix $\mathbf{A}$, the matrices $\mathbf{R}_Y(t,\tau)$, $\forall\tau$ and $\forall t$ can be joint block diagonalized without any unitarity constraint about the wanted matrix $\mathbf{A}$.

Notice that in this case, the recovered sources after inversion of the system are obtained up to a permutation and up to a filter but we will not discuss about these indeterminations in this communication.

### 3.2 Computer Simulations

We present simulations to illustrate the effectiveness of the proposed algorithm in the blind source separation context and to establish a comparison with another algorithm ($\mathsf{O-JBD}$) for the orthogonal joint block diagonalization of matrices. While our algorithm is directly applied on the correlation matrices of the observations, the second algorithm is applied after a pre-whitening stage on the correlation matrices of the pre-whitened observations. We consider $m = 4$ mixtures of $n = 2$ speech source signals sampled at 8 kHz, $L = 2$ and $L' = 4$. These signal sources are mixed according to the following transfer function matrix whose components are randomly generated:

$$
\mathbf{A}[z] = \begin{pmatrix}
0.9772 + 0.2079z^{-1} - 0.0439z^{-2} & -0.6179 + 0.7715z^{-1} + 0.1517z^{-2} \\
-0.2517 - 0.3204z^{-1} + 0.9132z^{-2} & -0.1861 + 0.4359z^{-1} - 0.8805z^{-2} \\
0.0803 - 0.7989z^{-1} - 0.5961z^{-2} & 0.5677 + 0.6769z^{-1} + 0.4685z^{-2} \\
-0.7952 + 0.3522z^{-1} + 0.4936z^{-2} & -0.2459 + 0.8138z^{-1} - 0.5266z^{-2}
\end{pmatrix}
$$

where $\mathbf{A}[z]$ stands for the $z$ transform of $\mathbf{A}(t)$. On the Fig. 2, we have displayed the performance index versus the number of matrices (left) and versus the $\mathsf{SNR}$. One can check that the obtained performance are better with the $\mathsf{NO-JBD}$ algorithm than with the $\mathsf{O-JBD}$ algorithm. One can also evaluate the block-diagonalization error defined as:

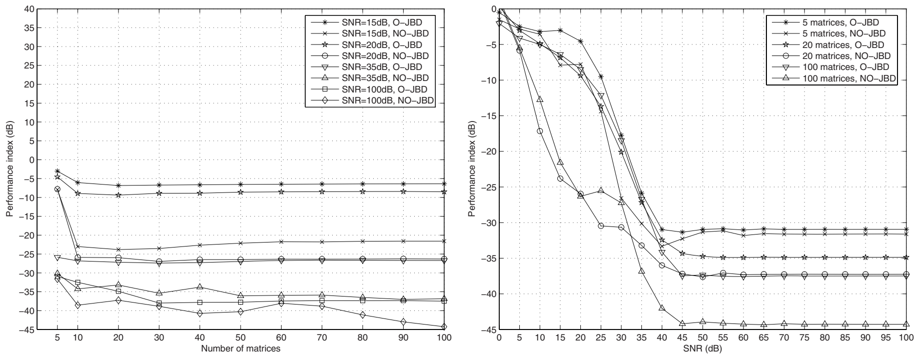


**Fig. 2.** Left: performance index versus number of matrices, right: performance index versus SNR

$\mathcal{E} = 10\log_{10}\{\frac{1}{N_m}\sum_{k=1}^{N_m}\|\mathsf{OffBdiag}\{\mathbf{BR}_Y(t,\tau_k)\mathbf{B}^T\|_F^2\}$ where $\mathbf{B}$ is the pseudo-inverse of the mixing matrix $\mathbf{A}$ and $\|.\|_F$ denotes the Frobenius norm. Finally, a comparaison of the block-diagonalization error with the $\mathsf{NO}-\mathsf{JBD}$ and $\mathsf{O}-\mathsf{JBD}$ algorithms versus the number of matrices (resp. SNR) is given in the left of Fig. 3 (resp. its right).

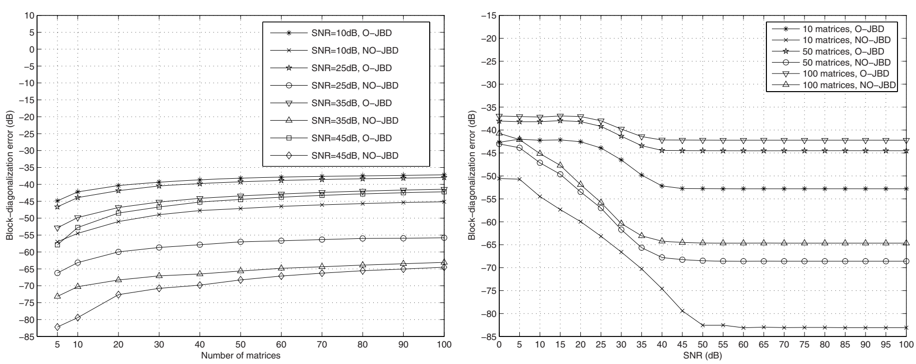


**Fig. 3.** Left: block-diagonalization error versus number of matrices, right: block-diagonalization error versus SNR

## 4 Discussion and Conclusion

In this paper, we have proposed a new joint block diagonalization algorithm for the separation of convolutive mixtures of sources that does not rely upon a unitary constraint. We have illustrated the usefulness of the proposed approach thanks to computer simulations: the considered algorithm has been applied to source separation using the correlation matrices of speech sources evaluated over different time delays.

## References

1. Abed-Meraïm, K., Belouchrani, A., Leyman, R.: Time-frequency signal analysis and processing: a comprehensive reference. In: Boashash, B. (ed.) chapter in Blind source separation using time-frequency distributions, Prentice-Hall, Oxford (2003)
2. Belouchrani, A., Abed-Meraïm, K., Amin, M., Zoubir, A.: Joint anti-diagonalization for blind source separation. In: Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'2001), Salt Lake City, Utah (May 2001)
3. Belouchrani, A., Abed-Meraïm, K., Cardoso, J.-F., Moulines, E.: A blind source separation technique using second order statistics. IEEE Trans. on Signal Processing 45, 434–444 (1997)
4. Bousbiah-Salah, H., Belouchrani, A., Abed-Meraim, K.: Blind separation of non stationary sources using joint block diagonalization. In: Proc. IEEE Workshop on Statistical Signal Processing, pp. 448–451 (2001)
5. Cardoso, J.-F., Souloumiac, A.: Blind Beamforming for non Gaussian signals. IEEE Proceedings-F 40, 362–370 (1993)
6. Comon, P.: Independant component analysis, a new concept? Signal Processing 36, 287–314 (1994)
7. DeLathauwer, L., Févotte, C., De Moor, B., Vandewalle, J.: Jacobi algorithm for joint block diagonalization in blind identification. In: 23rd Symposium on Information Theory in the Benelux, Louvain-la-Neuve, Belgium (May 2002)
8. Fadaili, E.-M., Thirion-Moreau, N., Moreau, E.: Algorithme de zéro-diagonalisation conjointe pour la séparation de sources déterministes, dans les Proc. du 20ème colloque GRETSI, Louvain-La-Neuve, Belgique, Septembre 2005, pp. 981–984 (2005)
9. Fadaili, E.-M., Thirion-Moreau, N., Moreau, E.: Non orthogonal joint diagonalization/zero-diagonalization for source separation based on time-frequency distributions. To appear in IEEE Trans. on Signal Processing 55(5) (2007)
10. Moreau, E.: A generalization of joint-diagonalization criteria for source separation. IEEE Trans. Signal Processing 49(3), 530–541 (2001)
11. Pham, D.-T.: Joint approximate diagonalization of positive definite matrices. SIAM Journal on Matrix Analysis and Appli 22(4), 1136–1152 (2001)
12. Yeredor, A.: Non-orthogonal joint diagonalization in the least square sense with application in blind source separation. IEEE Transactions on signal processing 50(7), 1545–1553 (2002)
13. Yeredor, A., Ziehe, A., Müller, K.R.: Approximate joint diagonalization using a natural gradient approach. In: Puntonet, C.G., Prieto, A.G. (eds.) ICA 2004. LNCS, vol. 3195, pp. 89–96. Springer, Heidelberg (2004)
14. Ziehe, A., Laskov, P., Nolte, G.G., Müller, K.-R.: A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation Journal of Machine Learning Research, No. 5, pp. 801–818 (July 2004)