Non Unitary Joint Block Diagonalization of Complex Matrices Using a Gradient Approach

Hicham Ghennioui[1,2], Nadège Thirion-Moreau[1], Eric Moreau[1], Abdellah Adib[2,3], and Driss Aboutajdine[2]

[1] STD, ISITV, av. G. Pompidou, BP56, F-83162 La Valette du Var Cedex, France
ghennioui@gmail.com, {thirion,moreau}@univ-tln.fr
[2] GSCM-LRIT, FSR, av. Ibn Battouta, BP1014, Rabat, Maroc
[3] DPG, IS, av. Ibn Battouta, BP703, Rabat, Maroc
adib@israbat.ac.ma, aboutaj@ieee.org

Abstract. This paper addresses the problem of the non-unitary approximate joint block diagonalization ($\mathsf{NU}-\mathsf{JBD}$) of matrices. Such a problem occurs in various fields of applications among which blind separation of convolutive mixtures of sources and wide-band signals array processing. We present a new algorithm for the non-unitary joint block-diagonalization of complex matrices based on a gradient-descent algorithm whereby the optimal step size is computed algebraically at each iteration as the rooting of a 3rd-degree polynomial. Computer simulations are provided in order to illustrate the effectiveness of the proposed algorithm.

1 Introduction

In the recent years, the problem of the joint decomposition of matrices or tensors sets have found interesting solutions through signal processing applications in blind source separation and array processing.

One of the first considered problem was the joint-diagonalization of matrices under the unitary constraint, leading to the nowadays well-known JADE [4] and SOBI [2] algorithms. The following works have addressed either the problem of the joint-diagonalization of tensors [5][7][12] or the problem of the joint-diagonalization of matrices but discarding the unitarity constraint [6][10][14] [15][16][17].

A second type of matrices decomposition has proven to be useful in blind source separation, telecommunications and cryptography. It consists in joint zero-diago-nalizing several matrices either under the unitary constraint [1] or not [9][10]. Most of the proposed (unitary) joint-diagonalization and/or zero-diagonalization algorithms have been applied to the problem of the blind separation of instantaneous mixtures of sources.

Finally, a third particular type of matrices decomposition arises in both the wide-band sources localization in correlated noise fields and the blind separation

M.E. Davies et al. (Eds.): ICA 2007, LNCS 4666, pp. 201–208, 2007.

of convolutive mixtures of sources problems. It is called joint block-diagonalization since the wanted matrices are block diagonal matrices[1] in such a decomposition. Such a problem has been considered in [3][8] where the block-diagonal matrices under consideration have to be positive definite and hermitian matrices and the required joint-block diagonalizer is a unitary matrix.

In this paper, our purpose is to discard this unitary constraint. To that aim, we generalize the non unitary joint-diagonalization approach proposed in [16] to the non-unitary joint block-diagonalization of several complex hermitian matrices. The resulting algorithm is based on a gradient-descent approach whereby the optimal step size is computed algebraically at each iteration as the rooting of a 3rd-degree polynomial. The main advantage of the proposed algorithm is that it is relatively general since the only needed assumption about the complex matrices under consideration is their hermitian symmetry. Finally, the use of the optimal step size speeds up the convergence.

The paper is organized as follows. We state the considered problem in the Section 2. In the Section 3, we present the algebraical derivations leading to the proposed non-unitary joint block-diagonalization algorithm. Computer simulations are provided in the Section 4 in order to illustrate the behaviour of the proposed approach.

2 Problem Statement

The non-unitary 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 matrices $\mathbf{M}_i \in \mathbb{C}^{M\times M}$, $i \in \{1,\ldots,N_m\}$ which all admit the following decomposition:

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

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 assume that all the matrices have the same size *i.e.* $N = r\times n_j, j \in \{1,\ldots,r\}$) 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-unitary joint bloc-diagonalization problem consists in estimating the matrix $\mathbf{A}$ and the matrices $\mathbf{D}_{ij}$, $i \in \{1,\ldots,N_m\}, j \in \{1,\ldots,r\}$ from only the matrices set $\mathcal{M}$. The case of a unitary matrix $\mathbf{A}$ has been considered in [8] where a first solution is proposed.

[1] A block diagonal matrix is a square diagonal matrix in which the diagonal elements are square matrices of any size (possibly even), and the off-diagonal elements are 0. A block diagonal matrix is therefore a block matrix in which the blocks off the diagonal are the zero matrices and the diagonal matrices are square.

3 Non-Unitary Joint Block-Diagonalization Using a Gradient Approach

In this section, we present a new algorithm to solve the problem of the non-unitary joint block-diagonalization. We propose to consider the following cost function

$$\mathcal{C}_{BD}(\mathbf{B}) = \sum_{i=1}^{N_m} \|\mathsf{OffBdiag}\{\mathbf{B}\mathbf{M}_i\mathbf{B}^H\}\|_F^2, \tag{2}$$

where $\|.\|_F$ stands for the Frobenius norm and the operator $\mathsf{OffBdiag}\{\cdot\}$ denotes the zero block-diagonal matrix. Thus:

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} & \dots & \mathbf{M}_{1r} \\ \mathbf{M}_{21} & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \vdots \\ \mathbf{M}_{r1} & \mathbf{M}_{r2} & \dots & \mathbf{M}_{rr} \end{pmatrix} \Rightarrow \mathsf{OffBdiag}\{\mathbf{M}\} = \begin{pmatrix} \mathbf{0} & \mathbf{M}_{12} & \dots & \mathbf{M}_{1r} \\ \mathbf{M}_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathbf{M}_{r1} & \mathbf{M}_{r2} & \dots & \mathbf{0} \end{pmatrix} \triangleq \mathbf{E} \tag{3}$$

Our aim is to minimize the cost function (2).

To make sure that the found matrix $\mathbf{B}$ keeps on being invertible, it is updated according to the following scheme (see [17]):

$$\mathbf{B}^{(m)} = (\mathbf{I} + \mathbf{W}^{(m-1)})\mathbf{B}^{(m-1)} \quad \forall m = 1, 2, \dots, \tag{4}$$

where $\mathbf{B}^{(0)}$ is some initial guess, $\mathbf{B}^{(m)}$ denotes the estimated matrix $\mathbf{B}$ at the m-th iteration, $\mathbf{W}^{(m-1)}$ is a sufficiently small (in terms of Frobenius norm) zero-block diagonal matrix and $\mathbf{I}$ is the identity matrix.

Denoting $\mathbf{M}_i^{(m)} = \mathbf{B}^{(m-1)}\mathbf{M}_i\mathbf{B}^{(m-1)H}$ $\forall i = 1, \dots, N_m$ and $\forall m = 1, 2, \dots$, where $(\cdot)^H$ stands for the transpose conjugate operator, then at the m-th iteration, the cost function can be expressed versus $\mathbf{W}^{(m-1)}$ rather than $\mathbf{B}^{(m)}$. We now have:

$$\mathcal{C}_{BD}(\mathbf{W}^{(m-1)}) = \sum_{i=1}^{N_m} \|\mathsf{OffBdiag}\{(\mathbf{I} + \mathbf{W}^{(m-1)})\mathbf{M}_i^{(m)}(\mathbf{I} + \mathbf{W}^{(m-1)})^H\}\|_F^2 \tag{5}$$

or more simply $\mathcal{C}_{BD}^{(m)}(\mathbf{W}) \triangleq \sum_{i=1}^{N_m} \|\mathsf{OffBdiag}\{(\mathbf{I} + \mathbf{W})\mathbf{M}_i^{(m)}(\mathbf{I} + \mathbf{W})^H\}\|_F^2$.

At each iteration, the wanted matrix $\mathbf{W}$ is then updated according to the following adaptation rule:

$$\mathbf{W}^{(m)} = -\mu\nabla\mathcal{C}_{BD}(\mathbf{W}^{(m-1)}) \quad \forall m = 1, 2, \dots \tag{6}$$

where μ is the step size or adaptation coefficient and where $\nabla\mathcal{C}_{BD}(\mathbf{W}^{(m-1)})$ stands for the complex gradient matrix defined, like in [13], by:

$$\nabla\mathcal{C}_{BD}(\mathbf{W}^{(m-1)}) = 2\frac{\partial\mathcal{C}_{BD}(\mathbf{W}^{(m-1)})}{\partial\mathbf{W}^{(m-1)*}} \quad \forall m = 1, 2 \dots \tag{7}$$

where $(\cdot)^*$ is the complex conjugate operator. We now have to calculate the complex gradient matrix $\nabla\mathcal{C}_{BD}^{(m)}(\mathbf{W}) = 2\frac{\partial\mathcal{C}_{BD}^{(m)}(\mathbf{W})}{\partial\mathbf{W}^*}$.

3.1 Gradient of the Cost Function $\mathcal{C}_{BD}^{(m)}(\mathbf{W})$

Let $\mathbf{D}_i^{(m)}$ and $\mathbf{E}_i^{(m)}$ respectively denote the block-diagonal and zero block-diagonal matrices extracted from the matrix $\mathbf{M}_i^{(m)}$ ($\mathbf{M}_i^{(m)} = \mathbf{E}_i^{(m)} + \mathbf{D}_i^{(m)}$). As $\mathbf{W}$ is a zero-block diagonal matrix too, the cost function $\mathcal{C}_{BD}^{(m)}(\mathbf{W})$ can be expressed as:

$$\begin{aligned}
\mathcal{C}_{BD}^{(m)}(\mathbf{W}) &= \sum_{i=1}^{N_m} \|\mathsf{OffBdiag}\{\mathbf{M}_i^{(m)}\} + \mathsf{OffBdiag}\{\mathbf{M}_i^{(m)}\mathbf{W}^H\} + \mathsf{OffBdiag}\{\mathbf{W}\mathbf{M}_i^{(m)}\} \\
&+ \mathsf{OffBdiag}\{\mathbf{W}\mathbf{M}_i^{(m)}\mathbf{W}^H\}\|_F^2 \\
&= \sum_{i=1}^{N_m} \|\mathbf{E}_i^{(m)} + \mathbf{D}_i^{(m)}\mathbf{W}^H + \mathbf{W}\mathbf{D}_i^{(m)} + \mathbf{W}\mathbf{E}_i^{(m)}\mathbf{W}^H\|_F^2 \\
&= \sum_{i=1}^{N_m} \mathsf{tr}\{(\mathbf{E}_i^{(m)} + \mathbf{D}_i^{(m)}\mathbf{W}^H + \mathbf{W}\mathbf{D}_i^{(m)} + \mathbf{W}\mathbf{E}_i^{(m)}\mathbf{W}^H)^H(\mathbf{E}_i^{(m)} + \mathbf{D}_i^{(m)}\mathbf{W}^H \\
&+ \mathbf{W}\mathbf{D}_i^{(m)} + \mathbf{W}\mathbf{E}_i^{(m)}\mathbf{W}^H)\}
\end{aligned} \tag{8}$$

where $\mathsf{tr}\{.\}$ stands for the trace operator. Then, using the linearity property of the trace and assuming to simplify the derivations that the considered matrices are hermitian, we finally find that:

$$\begin{aligned}
\mathcal{C}_{BD}^{(m)}(\mathbf{W}) = \sum_{i=1}^{N_m} & \mathsf{tr}\{\mathbf{E}_i^{(m)H}\mathbf{E}_i^{(m)}\} + 2\mathsf{tr}\{\mathbf{E}_i^{(m)H}(\mathbf{D}_i^{(m)}\mathbf{W}^H + \mathbf{W}\mathbf{D}_i^{(m)})\} \\
&+ \mathsf{tr}\{\mathbf{W}\mathbf{D}_i^{(m)H}\mathbf{D}_i^{(m)}\mathbf{W}^H + \mathbf{D}_i^{(m)H}\mathbf{W}^H\mathbf{W}\mathbf{D}_i^{(m)}\} \\
&+ 2\mathsf{tr}\{\mathbf{E}_i^{(m)H}\mathbf{W}\mathbf{E}_i^{(m)}\mathbf{W}^H\} \\
&+ \mathsf{tr}\{\mathbf{W}\mathbf{D}_i^{(m)H}\mathbf{W}\mathbf{D}_i^{(m)} + \mathbf{D}_i^{(m)H}\mathbf{W}^H\mathbf{D}_i^{(m)}\mathbf{W}^H\} \\
&+ 2\mathsf{tr}\{\mathbf{W}\mathbf{E}_i^{(m)H}\mathbf{W}^H(\mathbf{D}_i^{(m)}\mathbf{W}^H + \mathbf{W}\mathbf{D}_i^{(m)})\} \\
&+ \mathsf{tr}\{\mathbf{W}\mathbf{E}_i^{(m)H}\mathbf{W}^H\mathbf{W}\mathbf{E}_i^{(m)}\mathbf{W}^H\}
\end{aligned} \tag{9}$$

Using now the following properties [11]

$$\mathsf{tr}\{\mathbf{PQR}\} = \mathsf{tr}\{\mathbf{RPQ}\} = \mathsf{tr}\{\mathbf{QRP}\} \tag{10}$$

$$\frac{\partial \mathsf{tr}\{\mathbf{P}\mathbf{X}^H\}}{\partial \mathbf{X}^*} = \mathbf{P} \tag{11}$$

$$\frac{\partial \mathsf{tr}\{\mathbf{P}\mathbf{X}\}}{\partial \mathbf{X}^*} = \mathbf{0} \tag{12}$$

$$d\mathsf{tr}\{\mathbf{P}\} = \mathsf{tr}\{d\mathbf{P}\} \tag{13}$$

$$d\mathsf{tr}\{\mathbf{P}\mathbf{X}^H\mathbf{Q}\mathbf{X}\} = \mathsf{tr}\{\mathbf{P}d\mathbf{X}^H\mathbf{Q}\mathbf{X} + \mathbf{P}\mathbf{X}^H\mathbf{Q}d\mathbf{X}\} \tag{14}$$

$$\frac{\partial \mathsf{tr}\{\mathbf{P}\mathbf{X}^H\mathbf{Q}\mathbf{X}\}}{\partial \mathbf{X}^*} = \mathbf{QXP} \tag{15}$$

It finally leads to the following result:

$$\begin{aligned}\nabla\mathcal{C}_{BD}^{(m)}(\mathbf{W}) = 4\sum_{i=1}^{N_m}\Big(&\mathbf{E}_i^{(m)H}\mathbf{D}_i^{(m)} + \mathbf{W}\mathbf{D}_i^{(m)H}\mathbf{D}_i^{(m)} + \mathbf{E}_i^{(m)H}\mathbf{W}\mathbf{E}_i^{(m)} \\ &+ \mathbf{W}\mathbf{E}_i^{(m)H}\mathbf{W}^H\mathbf{D}_i^{(m)} + \mathbf{W}\mathbf{E}_i^{(m)}\mathbf{W}^H\mathbf{W}\mathbf{E}_i^{(m)H} + \mathbf{D}_i^{(m)}\mathbf{W}^H\mathbf{D}_i^{(m)H} \\ &+ \mathbf{D}_i^{(m)}\mathbf{W}^H\mathbf{W}\mathbf{E}_i^{(m)H} + \mathbf{W}\mathbf{D}_i^{(m)}\mathbf{W}\mathbf{E}_i^{(m)H}\Big). \qquad (16)\end{aligned}$$

3.2 Seek of the Optimal Step Size

The expression (16) is then used in the gradient descent algorithm (6). To accelerate its convergence, the optimal step size μ is computed algebraically at each iteration. To that aim, one has to calculate $\mathcal{C}_{BD}^{(m)}(\mathbf{W} \leftarrow -\mu\nabla\mathcal{C}_{BD}^{(m)}(\mathbf{W}))$, but here we use $\mathcal{C}_{BD}^{(m)}(\mathbf{W} \leftarrow \mu\mathbf{F}^{(m)} = -\mu\mathsf{OffBdiag}\{\nabla\mathcal{C}_{BD}^{(m)}(\mathbf{W})\})$. $\mathbf{F}^{(m)}$ is the anti-gradient matrix. We use $\mathsf{OffBdiag}\{\nabla\mathcal{C}_{BD}^{(m)}(\mathbf{W})\}$ instead of $\nabla\mathcal{C}_{BD}^{(m)}(\mathbf{W})$ because $\mathbf{W}$ is a sufficiently small (in terms of norm) zero block-diagonal matrix and thus only the off block-diagonal terms are involved in the descent of the criterion. We now have to seek for the optimal step μ ensuring the minimization of the cost function $\mathcal{C}_{BD}^{(m)}(\mu\mathbf{F}^{(m)})$. This step is determined by the rooting of the 3rd-degree polynomial (18) which is obtained as the derivative of the 4rd-degree polynomial $\mathcal{C}_{BD}^{(m)}(\mu\mathbf{F}^{(m)})$ with respect to μ:

$$\mathcal{C}_{BD}^{(m)}(\mu\mathbf{F}^{(m)}) = a_0^{(m)} + a_1^{(m)}\mu + a_2^{(m)}\mu^2 + a_3^{(m)}\mu^3 + a_4^{(m)}\mu^4, \qquad (17)$$

$$\frac{\partial\mathcal{C}_{BD}^{(m)}(\mu\mathbf{F}^{(m)})}{\partial\mu} = 4a_4^{(m)}\mu^3 + 3a_3^{(m)}\mu^2 + 2a_2^{(m)}\mu + a_1^{(m)}, \qquad (18)$$

where the coefficients have been found to be equal to:

$$a_0^{(m)} = \sum_{i=1}^{N_m}\mathsf{tr}\{\mathbf{E}_i^{(m)H}\mathbf{E}_i^{(m)}\} \qquad (19)$$

$$a_1^{(m)} = \sum_{i=1}^{N_m}\mathsf{tr}\{\mathbf{E}_i^{(m)H}(\mathbf{D}_i^{(m)}\mathbf{F}^H + \mathbf{F}\mathbf{D}_i^{(m)}) + (\mathbf{D}_i^{(m)}\mathbf{F}^H + \mathbf{F}\mathbf{D}_i^{(m)})^H\mathbf{E}_i^{(m)}\} \qquad (20)$$

$$\begin{aligned}a_2^{(m)} = \sum_{i=1}^{N_m}\mathsf{tr}\Big\{&\mathbf{E}_i^{(m)H}\mathbf{F}\mathbf{E}_i^{(m)}\mathbf{F}^H + \mathbf{F}\mathbf{E}_i^{(m)H}\mathbf{F}^H\mathbf{E}_i^{(m)} \\ &+ (\mathbf{D}_i^{(m)}\mathbf{F}^H + \mathbf{F}\mathbf{D}_i^{(m)})^H(\mathbf{D}_i^{(m)}\mathbf{F}^H + \mathbf{F}\mathbf{D}_i^{(m)})\Big\} \qquad (21)\end{aligned}$$

$$\begin{aligned}a_3^{(m)} = \sum_{i=1}^{N_m}\mathsf{tr}\Big\{&(\mathbf{D}_i^{(m)}\mathbf{F}^H + \mathbf{F}\mathbf{D}_i^{(m)})^H\mathbf{F}\mathbf{E}_i^{(m)}\mathbf{F}^H \\ &+ \mathbf{F}\mathbf{E}_i^{(m)H}\mathbf{F}^H(\mathbf{D}_i^{(m)}\mathbf{F}^H + \mathbf{F}\mathbf{D}_i^{(m)})\Big\} \qquad (22)\end{aligned}$$

$$a_4^{(m)} = \sum_{i=1}^{N_m}\mathsf{tr}\{\mathbf{F}^{(m)}\mathbf{E}_i^{(m)H}\mathbf{F}^{(m)H}\mathbf{F}^{(m)}\mathbf{E}_i^{(m)}\mathbf{F}^{(m)H}\}. \qquad (23)$$

The optimal step μ corresponds to the root of the polynomial (18) attaining the absolute minimum in the polynomial (17).

3.3 Summary of the Proposed Algorithm

The proposed non-unitary joint block-diagonalization based on a gradient algorithm denoted by $\mathsf{JBD}_{\mathsf{NU,G}}$ is now presented below:

Denote the N_m square matrices as $\mathbf{M}_1^{(0)}, \mathbf{M}_2^{(0)}, \ldots, \mathbf{M}_{N_m}^{(0)}$
Given initial estimates $\mathbf{W}^{(0)} = \mathbf{0}$ and $\mathbf{B}^{(0)} = \mathbf{I}$
For $m = 1, 2, \ldots$
  For $i = 1, \ldots, N_m$
    Compute $\mathbf{M}_i^{(m)}$ as

$$\mathbf{M}_i^{(m)} = \mathbf{B}^{(m-1)}\mathbf{M}_i^{(m-1)}\mathbf{B}^{(m-1)H}$$

    Compute $\nabla\mathcal{C}_{BD}^{(m)}(\mathbf{W})$ whose expression is given by equation (16)
  EndFor
Set $\mathbf{F}^{(m)} = -\mathsf{OffBdiag}\{\nabla\mathcal{C}_{BD}^{(m)}(\mathbf{W})\}$
Compute the coefficients $a_0^{(m)}, \ldots, a_4^{(m)}$ thanks to (19), (20), (21), (22) and (23)
Set the optimal step μ by the research of the root of the polynomial (18) attaining the absolute minimum in the polynomial (17)
Set $\mathbf{W}^{(m)} = \mu\mathbf{F}^{(m)}$ and $\mathbf{B}^{(m)} = (\mathbf{I} + \mathbf{W}^{(m-1)})\mathbf{B}^{(m-1)}$
EndFor

4 Computer Simulations

In this section, we perform simulations to illustrate the behaviour of the proposed algorithm. We consider a set $\mathbf{D}$ of $N_m = 11$ (resp. 31, 101) matrices, randomly chosen (according to a Gaussian law of mean 0 and variance 1). Initially these matrices are exactly block-diagonal, then matrices with random entries chosen from a Gaussian law of mean 0 and variance σ_b^2 are added. The signal to noise ratio (SNR) is then defined by $\mathsf{SNR} = 10\log(\frac{1}{\sigma_b^2})$. We use the following performance index which is an extension of that introduced in [12]:

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

where $(\mathbf{G})_{i,j} \forall i, j \in \{1, \ldots, r\}$ is the (i,j)-th matrix block (square) of $\mathbf{G} = \hat{\mathbf{B}}\mathbf{A}$. The displayed results are averaged over 30 Monte-Carlo trials. In this example, they were obtained considering $M = N = 12$, $r = 3$ and real and symmetric matrices. On the left of Fig. 1 we display the performance index obtained with the proposed algorithm versus the number of used matrices for different values of the SNR. On its right we have plotted the evolution of the performance index versus the SNR.

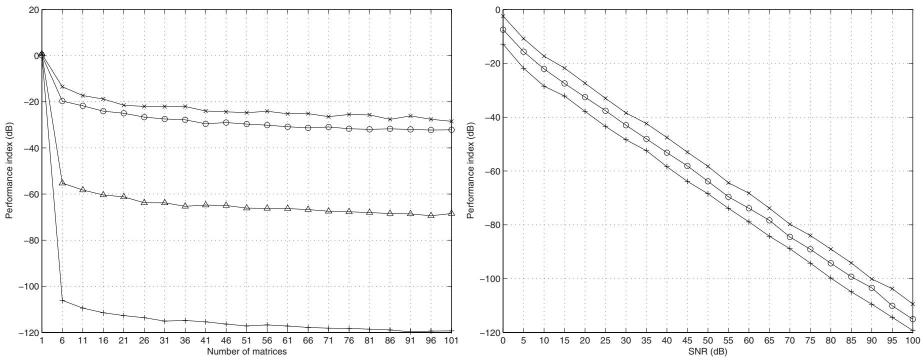

Fig. 1. Left: performance index versus number N_m of used matrices for different values of the SNR (SNR=10 dB ($\times$), 20 dB ($\circ$), 50 dB ($\triangle$) and 100 dB ($+$)). Right: performance index versus SNR for different size of the matrices set to be joint block-diagonalized (N_m=11 ($\times$), 31 ($\circ$), 101 ($+$)).

5 Discussion and Conclusion

In this paper, we have proposed a new algorithm (named $\mathsf{JBD}_{\mathsf{NU,G}}$) based on a gradient approach to perform the non-unitary joint block-diagonalization of a given set of complex matrices. One of the main advantages of this algorithm is that it applies to complex hermitian matrices. This algorithm finds application in blind separation of convolutive mixtures of sources and in array processing. In the context of blind sources separation, it should enable to achieve better performances by discarding the unitary constraint. In fact, starting with a pre-whitening stage is a possible way to amount to a unitary square mixture of sources to be able to use unitary joint-decomposition algorithms. But such a pre-whitening stage imposes a limit on the attainable performances that can be overcome thanks to non-unitary algorithms.

References

1. 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)
2. Belouchrani, A., Abed-Meraïm, K., Cardoso, J.-F., Moulines, E.: A blind source separation technique using second order statistics. IEEE Transactions on Signal Processing 45, 434–444 (1997)
3. Bousbiah-Salah, H., Belouchrani, A., Abed-Meraïm, K.: Blind separation of non stationary sources using joint block diagonalization. In: Proc. IEEE Workshop on Statistical Signal Processing, pp. 448–451 (2001)
4. Cardoso, J.-F., Souloumiac, A.: Blind beamforming for non Gaussian signals. IEEE Proceedings-F 40, 362–370 (1993)

5. Comon, P.: Independant component analysis, a new concept? Signal Processing 36, 287–314 (1994)
6. Dégerine, S.: Sur la diagonalisation conjointe approchée par un critère des moindres carrés. In: Proc. 18ème Colloque GRETSI, Toulouse, Septembre 2001, pp. 311–314 (2001)
7. DeLathauwer, L.: Signal processing based on multilinear algebra. PhD Thesis, Université Catholique de Leuven, Belgique (September 1997)
8. 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)
9. Fadaili, E.-M., Thirion-Moreau, N., Moreau, E.: Algorithme de zéro-diagonalisation conjointe pour la séparation de sources déterministes. In: dans les Proc. du 20ème colloque GRETSI, Louvain-La-Neuve, Belgique, Septembre 2005, pp. 981–984 (2005)
10. 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 Transactions on Signal Processing, 55(4) (April 2007)
11. Joho, M.: A systematic approach to adaptive algorithms for multichannel system identification, inverse modeling and blind identification. PHD Thesis, Swiss Federal Institute of Technology, Zürich (December 2000)
12. Moreau, E.: A generalization of joint-diagonalization criteria for source separation. IEEE Trans. Signal Processing 49(3), 530–541 (2001)
13. Petersen, K.B., Pedersen, M.S.: The matrix cookbook (January 5, 2005)
14. Pham, D.-T.: Joint approximate diagonalization of positive definite matrices. SIAM Journal on Matrix Analysis and Applications 22(4), 1136–1152 (2001)
15. 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)
16. Yeredor, A., Ziehe, A., Muller, 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)
17. 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, 801–818 (July 2004)