ALGORITHMS FOR JOINT BLOCK DIAGONALIZATION

K. Abed-Meraim and A. Belouchrani***

* TSI Dept., Telecom Paris, 37-39 rue Dareau, 75014 Paris, France
** EEE Dept., Polytechnic National School of Algiers, EL Harrach 16200, Algeria
E-mail: abed@tsi.enst.fr, belouchrani@hotmail.com

ABSTRACT

We present, in this paper, several algorithms for the joint block diagonalization (JBD) of a set of matrices. In particular, we will show and explain how the JBD can be achieved (up to a permutation matrix) using a Jacobi-like joint diagonalization algorithm. Two simple techniques are proposed to reduce the permutation indeterminacy to a 'block permutation' indeterminacy (the latter being inherent to the JBD problem). Finally, a comparative study of the considered JBD methods is provided.

1. INTRODUCTION

Estimating the joint eigenstructure of several matrices is a problem that arises in many multivariate signal processing applications, e.g., joint diagonalization for source separation [3], joint eigendecomposition for parameter estimation and pairing [9, 6], joint block-diagonalization (JBD) for source localization and convolutive source separation [7, 5], joint Schur decomposition for multidimensional harmonic retrieval [1, 10] and blind system identification [2].

In this paper, we focus on the JBD problem. This problem has been first treated in [4] for a set of positive definite symmetric matrices. In [7], a new JBD approach of singular (not necessarily positive) matrices has been introduced to solve the problem of DOA (Direction Of Arrival) estimation. Herein, we present 'exact' and 'approximate' iterative JBD algorithms. These algorithms are Jacobi-like techniques that minimize a squared error cost function iteratively by means of Givens rotations. An advantage of the Jacobi methods is their inherent parallelism, which allows efficient implementations on certain parallel architectures [11]. Another virtue of the Jacobi methods is their favorable rounding-error properties, in the sense that small relative perturbation in the matrices entries cause small relative perturbations in the entries of their engenstructures [12]. Finally, we show how the JBD of a set of matrices can be achieved via the joint diagonalization algorithm in [3].

2. PROBLEM FORMULATION

Consider a set of K matrices, $\mathbf{M}_1, \cdots, \mathbf{M}_K$, $\mathbf{M}_i \in \mathbb{C}^{n\times n}, i = 1, \cdots, K$, that have the following decomposition:

$$\mathbf{M}_i = [\mathbf{E}_1, \cdots, \mathbf{E}_r] \begin{bmatrix} \mathbf{D}_{i1} & \cdots & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & \cdots & \mathbf{D}_{ir} \end{bmatrix} \begin{bmatrix} \mathbf{E}_1^H \\ \vdots \\ \mathbf{E}_r^H \end{bmatrix} \quad (1)$$

where $\mathbf{E} = [\mathbf{E}_1, \cdots, \mathbf{E}_r]$ is unitary and $\mathbf{D}_{ij}$, $j = 1, \cdots, r$ are r $m_j \times m_j$ square matrices with $m_1 + \cdots + m_r = n$.

The matrices $\mathbf{M}_i$, $i = 1, \cdots, K$ are said to be jointly block diagonalizable under unitary transform $\mathbf{E}$, i.e., $\mathbf{E}^H\mathbf{M}_i\mathbf{E}$ are block diagonal matrices for $i = 1, \cdots, K$. $\mathbf{E}^H$ denotes the transpose conjugate of $\mathbf{E}$.

The problem of JBD consists in estimating the matrices $\mathbf{E}$ and $\mathbf{D}_{ij}$, $i = 1, \cdots, K$, $j = 1, \cdots, r$ given the matrices $\mathbf{M}_i$, $i = 1, \cdots, K$. Note that the JBD decomposition is not unique since if $\{\mathbf{E}_1, \cdots, \mathbf{E}_r, \mathbf{D}_{11}, \cdots, \mathbf{D}_{Kr}\}$ is a solution then $\mathbf{E}'_i = \mathbf{E}_i\mathbf{U}_i$ and $\mathbf{D}'_{ij} = \mathbf{U}_i^H\mathbf{D}_{ij}\mathbf{U}_i$ is also another admissible solution, where $\mathbf{U}_i, i = 1, \cdots, r$ are unitary matrices. In other words, matrix $\mathbf{E}$ can be determined only up to a block diagonal unitary matrix. However, is most practical applications this indeterminacy is inherent and does not affect the final result of the considered problem.

In practice, the matrices $\mathbf{M}_1, \cdots, \mathbf{M}_K$ are given by some sample estimated statistics that are corrupted by estimation errors due to noise and finite sample size effects. Thus, they are only "approximately" simultaneously block diagonalizable. This suggests that a viable JBD algorithm must provide a kind of an "average eigenstructure" when it is applied to a set of nearly joint block diagonalizable matrices. An optimal solution based on a least-squares approach has been proposed in [8] on which the Jacobi-like JBD algorithms of next section are based.

3. JACOBI-LIKE JBD ALGORITHMS

Solving the JBD problem in the least-squares sense consists in choosing the $n \times n$ unitary matrix $\mathbf{E}$ and the $m_i \times m_i$ matrices $\mathbf{D}_{ki}$ such that one minimizes the Frobenius norm of the difference between the data matrices $\mathbf{M}_k$ and the true matrices given by (1), i.e.

$$\min_{\mathbf{E}_i, \mathbf{D}_{ki}} \sum_{k=1}^{K} \|\mathbf{M}_k - \sum_{i=1}^{r} \mathbf{E}_i\mathbf{D}_{ki}\mathbf{E}_i^H\|^2 \quad (2)$$

This least-squares fitting problem leads to (see [8] for more details) the following criterion

$$\max_{\mathbf{E}} \sum_{k=1}^{K} \|\mathrm{bdiag}(\mathbf{E}^H\mathbf{M}_k\mathbf{E})\|^2 \quad (3)$$

where bdiag($\mathbf{M}$) is a block diagonal matrix constructed from $\mathbf{M} \stackrel{\text{def}}{=} [\mathbf{M}_{ij}]_{1\le i,j\le r}$ ($\mathbf{M}_{ij}$ being $m_i \times m_j$ matrices) in the following way:

$$\mathrm{bdiag}(\mathbf{M}) = \begin{bmatrix} \mathbf{M}_{11} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{M}_{rr} \end{bmatrix}. \quad (4)$$

In other words, the JBD of $\{\mathbf{M}_1, \cdots, \mathbf{M}_K\}$ consists in maximizing under unitary transform $\mathbf{E}$ the sum of the norms of their block diagonal entries. This is equivalent to minimizing the sum of the norms of their block off-diagonal entries. It is clear that the latter is minimum (zero) when all matrices are exactly block diagonal. Criterion (3) is referred to as the JBD criterion in the sequel.

To minimize the JBD criterion (3), we choose here to compute the unitary matrix $\mathbf{E}$ as products of Givens rotations, i.e.,

$$\mathbf{E} = \prod_{\text{nb. of sweeps}} \prod_{1 \leq p < q \leq n} \mathbf{\Theta}_{(qp)}$$

where the elementary Givens rotations $\mathbf{\Theta}_{(qp)}$ are defined as unitary matrices where all diagonal elements are 1 except for the two elements equal to c in rows (and columns) p and q. Likewise, all off-diagonal elements of $\mathbf{\Theta}_{(qp)}$ are 0 except for the two elements s and $-\bar{s}$ at (p,q) and (q,p) respectively, where $\bar{s}$ denotes the conjugate of s. The scalar numbers c and s are given by

$$\left\{ \begin{array}{rcl} c & = & \cos(\theta) \\ s & = & \sin(\theta)\exp(i\alpha) \end{array} \right.$$

where $i = \sqrt{-1}$. In the sequel, we describe a procedure to choose the rotation angles θ and α at a particular iteration such that the cost function (3) is increased to its maximum. To this end, we need to specify the orthogonal transformation

$$\mathbf{M}' = \mathbf{\Theta}_{(qp)}^H \mathbf{M} \mathbf{\Theta}_{(qp)} \quad (5)$$

for any given matrix $\mathbf{M} \in \mathbb{C}^{n \times n}$. First notice that these orthogonal transformation changes only the pth and qth rows and pth and qth columns of $\mathbf{M}$. Also, when p and q sweep the same diagonal block, i.e., $1 \leq p,q \leq m_1$, or $m_1 < p,q \leq m_1 + m_2$, $\cdots$, or $\sum_{i=1}^{r-1} m_i < p,q \leq n$, then the value of the JBD criterion (3) is unchanged. Therefore, we consider only indices (p,q) that sweep two different diagonal blocks, e.g., for $r = 2$ we select p and q in the range $1 \leq p \leq m_1 < q \leq n$.

The changed entries of $\mathbf{M}'$ in its block-diagonal matrices are given by:

$$\begin{array}{rcl} \mathbf{M}'(p,j) & = & c\mathbf{M}(p,j) - s\mathbf{M}(q,j), \quad j \neq p \\ \mathbf{M}'(j,p) & = & c\mathbf{M}(j,p) - \bar{s}\mathbf{M}(j,q), \quad j \neq p \\ \mathbf{M}'(p,p) & = & c^2\mathbf{M}(p,p) + |s|^2\mathbf{M}(q,q) \\ & & -\bar{s}c\mathbf{M}(p,q) - sc\mathbf{M}(q,p) \\ \mathbf{M}'(q,j) & = & \bar{s}\mathbf{M}(p,j) + c\mathbf{M}(q,j), \quad j \neq q \\ \mathbf{M}'(j,q) & = & s\mathbf{M}(j,p) + c\mathbf{M}(j,q), \quad j \neq q \\ \mathbf{M}'(q,q) & = & c^2\mathbf{M}(q,q) + |s|^2\mathbf{M}(p,p) \\ & & +\bar{s}c\mathbf{M}(q,p) + sc\mathbf{M}(p,q) \end{array} \quad (6)$$

3.1 Exact JBD algorithm

The proposed method consists of maximizing iteratively the JBD criterion (3) by successive Givens rotations, starting from $\mathbf{E} = \mathbf{I}$. The rotations $\mathbf{\Theta}_{(qp)}$ are computed such that (3) is maximum, i.e., at each iteration, the angle parameters (θ, α) are given by:

$$\begin{array}{rcl} (\theta,\alpha) & = & \arg\max \mathscr{C}(\theta,\alpha) \\ \mathscr{C}(\theta,\alpha) & \stackrel{\text{def}}{=} & \sum_{k=1}^{K} \|\text{bdiag}(\mathbf{M}'_k)\|^2 \end{array} \quad (7)$$

where $\mathbf{M}'_k$ is defined as in (5). The exact JBD algorithm can be summarized as follows (using informal notation)[1]:

$\mathbf{E} = \mathbf{I}$
for $k = 1, \ldots$, nb. of iterations
for $(p,q) \in \mathscr{I}$
$\mathbf{\Theta}_{(qp)} = \arg\max_{\theta,\alpha} \mathscr{C}(\theta,\alpha)$
$\mathbf{E} := \mathbf{E}\mathbf{\Theta}_{(qp)}$ and $\mathbf{M}_k := \mathbf{\Theta}_{(qp)}^H \mathbf{M}_k \mathbf{\Theta}_{(qp)}$, $k = 1, \cdots, K$

where $\mathscr{I}$ denotes the set of 'admissible' indices, i.e.

$$\mathscr{I} = \left\{ (p,q) \mid \exists\ 0 \leq i_p < i_q < r, \ \sum_{j=1}^{i_p} m_j < p \leq \sum_{j=1}^{i_p+1} m_j \right. \\ \left. \text{and } \sum_{j=1}^{i_q} m_j < q \leq \sum_{j=1}^{i_q+1} m_j \right\} \quad (8)$$

After some straightforward derivations, the maximization of $\mathscr{C}(\theta,\alpha)$ is shown to be equivalent to the maximization of the linear-quadratic form

$$\max_{\|\mathbf{v}\|=1} (\mathbf{v}^T \mathbf{G} \mathbf{v} + \mathbf{g}^T \mathbf{v}) \quad (9)$$

where

$$\mathbf{v} = [\cos(2\theta), \sin(2\theta)\cos(\alpha), \sin(2\theta)\sin(\alpha)]^T \quad (10)$$

and $\mathbf{G}$ (resp. $\mathbf{g}$) is a 3×3 real-valued matrix (resp. a 3×1 real-valued vector) the expressions of which (omitted here due to space limitation) are known explicit functions of the entries of $\mathbf{M}_1, \cdots, \mathbf{M}_K$. Using a Lagrange multiplier, the maximization of (9) leads to:

$$2(\mathbf{G} + \lambda\mathbf{I})\mathbf{v} + \mathbf{g} = \mathbf{0} \quad (11)$$

and hence

$$\mathbf{v} = -\frac{1}{2}(\mathbf{G} + \lambda\mathbf{I})^{-1}\mathbf{g} \quad (12)$$

where λ is a real scalar chosen in such a way that

$$\|\mathbf{v}\|^2 = 1 \quad \Longleftrightarrow \quad \frac{1}{4}\sum_{i=1}^{3} \frac{|\mathbf{u}_i^T\mathbf{g}|^2}{(\lambda_i + \lambda)^2} = 1 \quad (13)$$

$\{\mathbf{u}_i\}_{1 \leq i \leq 3}$ and $\{\lambda_i\}_{1 \leq i \leq 3}$ being the eigenvectors and eigenvalues of $\mathbf{G}$. Or

$$\mathbf{v} = -\frac{1}{2}(\mathbf{G} - \lambda_i\mathbf{I})^{\#}\mathbf{g} + c_i\mathbf{u}_i, \quad i = 1,2,3 \quad (14)$$

which corresponds to the case where λ coincides with the opposite of an eigenvalue of $\mathbf{G}$, i.e., $-\lambda_i$, $i = 1,2,3$. $^{\#}$ denotes the pseudo-inverse operator and c_i is a real constant chosen such that $\|\mathbf{v}\| = 1$ and (9) is maximum[2]. It is given by $c_i = \text{sign}(\mathbf{u}_i^T\mathbf{g})(1 - \|(\mathbf{G} - \lambda_i\mathbf{I})^{\#}\mathbf{g}\|^2/4)^{\frac{1}{2}}$.

[1] Note that the criterion is calculated with the current values of the matrices $\mathbf{M}_1, \cdots, \mathbf{M}_K$ which are updated at each iteration using the above orthogonal transformation. For notational simplicity, we keep using the generic notation $\mathbf{M}_1, \cdots, \mathbf{M}_K$ to denote the updated matrices.

[2] This solution exists only if $\|(\mathbf{G} - \lambda_i\mathbf{I})^{\#}\mathbf{g}\|/2 \leq 1$. Note that, when it exists, both c_i and $-c_i$ satisfy $\|\mathbf{v}\| = 1$. We then choose the one that maximizes (9) which is given by $\text{sign}(c_i) = \text{sign}(\mathbf{u}_i^T\mathbf{g})$.

As we can see from (13), finding λ involves a 6th order polynomial rooting. In general, equation (13) has multiple real-valued roots λ leading to multiple solutions for vector $\mathbf{v}$. Among the set of all possible solutions, we select the one associated with the vector $\mathbf{v}$ that maximizes the value of $\mathscr{C}(\theta,\alpha)$, i.e., the vector $\mathbf{v}$ that maximizes (9).

Note that in the real data case, the matrix $\mathbf{G}$ is 2×2 and vector $\mathbf{g}$ is 2×1. In this case, (13) becomes a 4-th order polynomial equation.

Once $\mathbf{v} = [v(1), v(2), v(3)]^T$ is obtained, the Givens rotation parameters c and s are computed according to (10) as

$$\begin{aligned} c &= \sqrt{\frac{v(1)+1}{2}} \\ s &= \frac{v(2)+iv(3)}{2c} \end{aligned} \tag{15}$$

Next, by means of a slight approximation of the JBD criterion, we present an alternative solution to (9) where no polynomial rooting is required.

3.2 Approximate JBD algorithm

To simplify the JBD algorithm, we choose here to approximate $|\mathbf{M}'(p,p)|^2$ (up to a scalar constant independent from the rotation parameters (θ,α)) by

$$|\mathbf{M}'(p,p)|^2 \approx |[\mathbf{M}\mathbf{\Theta}_{(qp)}](p,p)|^2 + |[\mathbf{\Theta}_{(qp)}^H\mathbf{M}](p,p)|^2.$$

This can be shown to be a first order approximation for $|\mathbf{M}'(p,p)|^2$ in the vicinity of the optimal point. Effectively, by writing $\mathbf{\Theta}_{(qp)} = \mathbf{I} + \varepsilon_{qp}$ (with $\|\varepsilon_{qp}\| \ll 1$ in the neighborhood of the convergence point), we obtain as a first order approximation

$$|\mathbf{M}'(p,p)|^2 \approx |[\mathbf{M} + \varepsilon_{qp}^H\mathbf{M} + \mathbf{M}\varepsilon_{qp}](p,p)|^2 \tag{16}$$

$$\approx |\mathbf{M}(p,p)|^2 + 2\Re e\{\mathbf{M}(p,p)[\varepsilon_{qp}^H\mathbf{M}](p,p) + \mathbf{M}(p,p)[\mathbf{M}\varepsilon_{qp}](p,p)\} \tag{17}$$

$$\approx |[\mathbf{M}\mathbf{\Theta}_{(qp)}](p,p)|^2 + |[\mathbf{\Theta}_{(qp)}^H\mathbf{M}](p,p)|^2 - |\mathbf{M}(p,p)|^2 \tag{18}$$

since at the first order, $|[\mathbf{M}\mathbf{\Theta}_{(qp)}](p,p)|^2$ and $|[\mathbf{\Theta}_{(qp)}^H\mathbf{M}](p,p)|^2$ are approximated respectively by

$$|\mathbf{M}(p,p)|^2 + 2\Re e\{\mathbf{M}(p,p)[\mathbf{M}\varepsilon_{qp}](p,p)\}$$
$$\text{and}\quad |\mathbf{M}(p,p)|^2 + 2\Re e\{\mathbf{M}(p,p)[\varepsilon_{qp}^H\mathbf{M}](p,p)\}$$

With this approximation, the JBD criterion becomes of the form

$$\mathscr{C}(\theta,\alpha) \approx \tilde{\mathbf{g}}^T\mathbf{v} \tag{19}$$

and its maximization leads to the explicit expressions:

$$\begin{aligned} \alpha &= \arctan\left(\frac{\Im m(a)}{\Re e(a)}\right), \\ \theta &= \frac{1}{2}\arctan\left(-2\frac{\Re e(e^{-i\alpha}a)}{b}\right) + \frac{1-\mathrm{sgn}(b)}{2}\frac{\pi}{2} \\ a &= \sum_{k=1}^{K}\left(\sum_{j\in\mathscr{I}_p}\mathbf{M}_k(p,j)\overline{\mathbf{M}}_k(q,j) + \mathbf{M}_k(j,p)\overline{\mathbf{M}}_k(j,q)\right. \\ &\quad \left. - \sum_{j\in\mathscr{I}_q}\mathbf{M}_k(p,j)\overline{\mathbf{M}}_k(q,j) + \mathbf{M}_k(j,p)\overline{\mathbf{M}}_k(j,q)\right) \\ b &= \sum_{k=1}^{K}\left(\sum_{j\in\mathscr{I}_p}(|\mathbf{M}_k(p,j)|^2 + |\mathbf{M}_k(j,p)|^2 - |\mathbf{M}_k(q,j)|^2\right. \\ &\quad - |\mathbf{M}_k(j,q)|^2) - \sum_{j\in\mathscr{I}_q}(|\mathbf{M}_k(p,j)|^2 + \\ &\quad \left.|\mathbf{M}_k(j,p)|^2 - |\mathbf{M}_k(q,j)|^2 - |\mathbf{M}_k(j,q)|^2)\right) \end{aligned}$$

where $\Re e(a)$ and $\Im m(a)$ denote the real part and imaginary part of a, respectively and

$$\begin{aligned} \mathscr{I}_p &= \{j \mid \sum_{l=1}^{i_p} m_l < j \leq \sum_{l=1}^{i_p+1} m_l\} \\ \mathscr{I}_q &= \{j \mid \sum_{l=1}^{i_q} m_l < j \leq \sum_{l=1}^{i_q+1} m_l\} \end{aligned}$$

i_p and i_q are defined as in (8).

One can notice that an iteration of the approximate JBD algorithm is much cheaper than that of the exact JBD. Unfortunately, because of the approximation, the former algorithm may require much more iterations to converge. Therefore, its overall complexity can be higher than that of the exact JBD algorithm. Next, we present a third approach based on the joint diagonalization algorithm in [3] that presents a better compromise between (the per iteration) complexity and the convergence rate.

4. JBD VIA JOINT DIAGONALIZATION (JD)

In [3], a Jacobi-like algorithm has been introduced to perform the joint diagonalization[3] of a set of matrices $\mathbf{M}_1,\cdots,\mathbf{M}_K$ under a common unitary transform $\mathbf{E}$. Matrix $\mathbf{E}$ is estimated as a product of Givens rotations but contrary to the JBD algorithm, the optimization of the rotation angles leads to a quadratic (instead of linear-quadratic) criterion of the form $\mathbf{v}^T\tilde{\mathbf{G}}\mathbf{v}$ where $\tilde{\mathbf{G}}$ is a 3×3 (or 2×2 in the real case) real-valued matrix and $\mathbf{v}$ is the vector defined in (10). Therefore, $\mathbf{v}$ corresponds to the unit-norm least eigenvector of $\tilde{\mathbf{G}}$ that can be computed explicitly [3].

Moreover, it is shown that optimizing the JD criterion by successive Givens rotations leads to solving the same problem for 2×2 matrices: i.e., the optimization of $\mathbf{\Theta}_{(qp)}$ is equivalent to the JD of the set of 2×2 matrices

$$\mathbf{M}_k^{(qp)} = \begin{bmatrix} M_k(p,p) & M_k(p,q) \\ M_k(q,p) & M_k(q,q) \end{bmatrix}, \quad k = 1,\cdots,K. \tag{20}$$

The iterative algorithm is stopped when at a given sweep (iteration) we have[4]:

$$\mathbf{\Theta}_{(qp)} = \mathbf{I} \qquad \forall\ 1 \leq p < q \leq n \tag{21}$$

which means that the JD criterion (i.e. $\sum_k \|\mathbf{E}^H\mathbf{M}_k\mathbf{E} - \mathrm{diag}(\mathbf{E}^H\mathbf{M}_k\mathbf{E})\|^2$) cannot be further decreased.

[3] JD can be seen as a particular case of JBD when $m_i = 1$, $\forall\, i$.

[4] $\mathbf{\Theta}_{(qp)} = \mathbf{I}$ is equivalent to $s = 0$. In practice, we use a small threshold ε and test if for all $1 \leq p < q \leq n$, $|s| < \varepsilon$.

Based on this, we claim that the JBD can be achieved (up to a permutation matrix) by applying the JD algorithm to the considered set of matrices $\mathbf{M}_1, \cdots, \mathbf{M}_K$. More precisely, we have the following conjuncture:

Conjuncture 1 *Let $\mathbf{M}_1, \cdots, \mathbf{M}_K$ be a set of K matrices satisfying* (1). *Then, optimizing the JD criterion for $\mathbf{M}_1, \cdots, \mathbf{M}_K$ using the Jacobi-like algorithm in [3] provides a set of block diagonal matrices up to an unknown permutation matrix* $\mathbf{P}$.

This conjuncture comes from the following observation: if the 2×2 matrices in (20) are non-diagonal for values of p and q corresponding to non-diagonal block entries (i.e. $(p,q) \in \mathscr{I}$), then there exists a unitary matrix that transforms them onto diagonal matrices (because $\mathbf{M}_1, \cdots, \mathbf{M}_K$ satisfy (1)). Consequently, the JD criterion can be further decreased and the stopping condition (21) is not verified in this case.

Now, performing the JBD up to an unknown permutation matrix is not satisfactory as the 'admissible' permutations are only those which preserve the block-diagonal structure (e.g., permuting the diagonal blocks or permuting entries within the same diagonal block is admissible). To get rid of this undesirable permutation matrix, two solutions have been considered:

- In the first one, the permutation $\mathbf{P}$ is decomposed as a product of elementary permutations[5] $\mathbf{P}_{(qp)}$. The latter is considered, at a given sweep, only if it increases the JBD criterion, i.e.

$$\sum_{k=1}^{K} \|\text{bdiag}(\mathbf{P}_{(qp)}^T \mathbf{M}_k \mathbf{P}_{(qp)})\|^2 > \sum_{k=1}^{K} \|\text{bdiag}(\mathbf{M}_k)\|^2$$

- The second approach consists in determining the permutation from the positions of the non-zero entries of $\mathbf{M}_k - \text{bdiag}(\mathbf{M}_k)$. This leads to a simple solution (details are omitted due to space limitation) when the considered set of matrices satisfies (1) exactly. However, when matrices $\mathbf{M}_1, \cdots, \mathbf{M}_K$ are only approximately JB diagonalizable, this approach requires a non-trivial thresholding to decide whether a given matrix entry is assimilated to zero or not.

5. SIMULATIONS

We provide here a simulation example to compare the convergence speed (in terms of number of sweeps) of the three proposed JBD algorithms. We consider $K = 4$ matrices of size $n = 6$ that satisfy exactly (1) (noiseless case) with block-diagonal sizes $m_1 = m_2 = 3$. We evaluate the off block diagonal squared norm, i.e $\sum_k \|\mathbf{M}_k - \text{bdiag}(\mathbf{M}_k)\|^2$, for each sweep over 100 Monte-Carlo runs. At each run the matrices are generated randomly according to (1) (i.e. at each run, we generate randomly the block diagonal entries $\mathbf{D}_{ki}, k = 1, \cdots, 4$ and $i = 1, 2$ as well as the unitary matrix $\mathbf{E}$). In this example, the convergence rate of the JD-based algorithm coincides with that of the exact JBD (E-JBD) algorithm and is much larger than that of the approximate JBD (A-JBD) algorithm. Also, as shown by figure 1, the JBD criterion reaches a lower value with E-JBD and JD algorithms compared to A-JBD.

[5] $\mathbf{P}_{(qp)}$ is defined in such a way that for a given vector $\mathbf{x}$, $\tilde{\mathbf{x}} = \mathbf{P}_{(qp)}\mathbf{x}$ iff $\tilde{x}(k) = x(k)$, for $k \notin \{p,q\}$, $\tilde{x}(p) = x(q)$ and $\tilde{x}(q) = x(p)$.

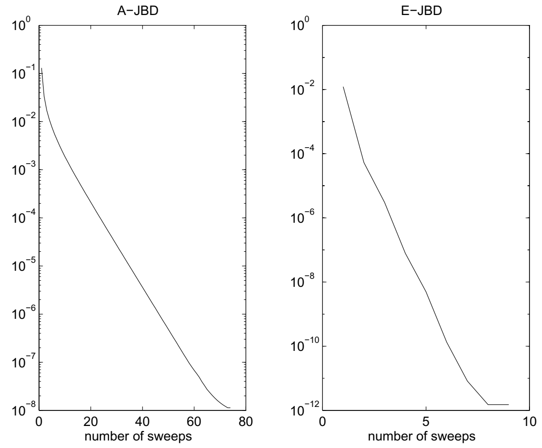

Fig. 1: Convergence speed proposed JBD algorithms.

REFERENCES

[1] M. Haardt, K. Hueper, J.B. Moore, and J.A. Nossek, "Simultaneous Schur decomposition of several matrices to achieve automatic pairing in multidimensional harmonic retrieval problem", *Proc. EUSIPCO*, pp. 531–534, 1996.

[2] Alle-Jan van der Veen, A. Paulraj, "An analytical constant modulus algorithm", *IEEE-SP*, May 1996.

[3] J.F. Cardoso and A. Souloumiac, "Jacobi angles for simultaneous diagonalization", *In SIAM Journal of Matrix Analysis and Applications*, Vol. 17, No 1, Jan. 1996.

[4] B. D. Flury and B. E. Neuenschwander, "Simultaneous diagonalization algorithm with applications in multivariate statistics", *Int. Series of Num. Math.*, vol. 119, 1994.

[5] H. Bousbia-Saleh, A. Belouchrani and K. Abed-Meraim, "Jacobi-like algorithm for blind signal separation of convolutive mixtures", *IEE Elec. Letters*, vol. 37, n 16, pp. 1049-1050, Aug. 2001.

[6] L. De Lathauwer, "Signal Processing Based on Multilinear Algebra", *Ph.D. Thesis, K.U.Leuven*, Sept. 1997.

[7] A. Belouchrani, A. Moeness, and K. Abed-Meraim, "Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices", *IEEE-SP Letters*, pp. 266–268, Sep. 1997.

[8] A. Belouchrani, K. Abed-Meraim, and Y. Hua, "Jacobi-like Algorithms for Joint Block Diagonalization: Application to Source Localization", *in Proc. IEEE Int. Work. on Intelligent Sig. Proc. and Comm. Sys. (Melbourne)*, 1998.

[9] Y. Hua and K. Abed-Meraim, "Techniques of eigenvalues estimation and association", *Digital Signal Processing*, vol. 7, no. 4, pp. 253-259, October 1997.

[10] K. Abed-Meraim and Y. Hua, "A Least-Squares Approach to Joint Schur Decomposition", *Proc. ICASSP (Seattle, USA)*, IV: 2541–2544, 1998.

[11] P.J. Eberlein, "On one-sided Jacobi methods for parallel computation", *SIAM J. Algebraic Discrete Meth.* vol. 8, pp. 790–796, 1987.

[12] G. H. Golub and C. F. Van Loan, "Matrix computations", *The Johns Hopkins University Press*, 1989.