APPROXIMATE TENSOR DIAGONALIZATION BY INVERTIBLE TRANSFORMS

Mikael Sørensen, Pierre Comon, Sylvie Icart and Luc Deneire

Laboratoire I3S, CNRS/UNSA
Les Algorithmes - Euclide-B, 06903 Sophia Antipolis, France
phone: +33-4 9294 2751
{sorensen, pcomon, icart, deneire}@i3s.unice.fr

ABSTRACT

Multilinear techniques are increasingly used in Signal Processing and Factor Analysis. In particular, it is often of interest to transform a tensor into another that is as diagonal as possible or to simultaneously transform a set matrices into a set of matrices that are close to diagonal. In this paper we propose a parameterization of the general linear group. Based on this parameterization Jacobi-type procedures for congruent diagonalization and PARAFAC decomposition problems are developed. Comparisons with an existing congruent diagonalization algorithm is reported.

1. INTRODUCTION

In signal processing the problem of congruent diagonalization of a set of matrices is a frequently occuring problem. For instance, it occurs in Blind Source Separation (BSS) of instantaneous mixtures [11], [13], [16], in analytical constant modulus algorithms for blind separation of communication signals [15] and in frequency estimation [12]. It also appears in under-determined BSS [9], [10], e.g. via a PARallel FACtor (PARAFAC) decomposition method, such as described in [8].

The Jacobi-like sweeping procedure for tensors was introduced in [5]. We propose in this paper Jacobi-like procedures that proceed pairwise, and which yield one or several nonsingular matrices at each step of the iterative congruent diagonalization. In the *Tensor Approximate Diagonalization* (TAD) problem, each nonsingular matrix acts on every mode of the tensor, such as to minimize the sum of squares of the off-diagonal entries. Similary, In the *Tensor Approximate Fitting* (TAF) problem, each nonsingular matrix acts on every mode of the tensor, so as to minimize the fit. On the other hand, in the *Joint Approximate Congruent Diagonalization* problem (JACD) a single nonsingular matrix is sought so as to minimize the sum of squares of the off-diagonal entries of all the matrices of interest.

The general framework we introduce is usable in at least three different problems: first the JACD of Hermitian matrices, should they be positive definite as in [14] or not; second, the JACD of invertible matrices of general form; third, the TAD and TAF problems.

This work has been partially supported by contract ANR-06-BLAN-0074 "Decotes". The work of Mikael Sørensen is supported by the EU by a Marie-Curie Fellowship (EST-SIGNAL program : http://est-signal.i3s.unice.fr) under contract No MEST-CT-2005-021175.

1.1 Notations and Definitions

Let vectors, matrices and tensors be denoted by lower case boldface, upper case boldface and upper case calligraphic letters respectively. Let $\circ$ and $\odot$ denote the outer and Khatri-Rao product respectively and let $\det(\cdot)$, $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, $(\cdot)^\dagger$, $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ denote the determinant, transpose, conjugate, conjugate-transpose, Moore-Penrose pseudoinverse, real part and imaginary part of a matrix, respectively. Furthermore, $\mathrm{GL}_m(\mathbb{C})$, $\mathrm{SL}_m(\mathbb{C})$, $\mathrm{SO}_m(\mathbb{C})$ and $\mathrm{S}_m(\mathbb{C})$ will denote the set of $m \times m$ nonsingular matrices, nonsingular matrices with determinant equal to one, unitary matrices with determinant equal to one, and Hermitian matrices, respectively. Moreover, let $\mathbf{A} \in \mathbb{C}^{I\times J}$, then $\mathrm{Vec}(\mathbf{A}) \in \mathbb{C}^{IJ}$ will denote a column vector with the property $(\mathrm{Vec}(\mathbf{A}))_{i+(j-1)I} = (\mathbf{A})_{ij}$. Similary, let $\mathbf{A} \in \mathbb{C}^{I\times I}$, then $\mathrm{Vecd}(\mathbf{A}) \in \mathbb{C}^{I}$ denotes a column vector with the property $(\mathrm{Vecd}(\mathbf{A}))_i = (\mathbf{A})_{ii}$. Finally, let $\mathbf{A} \in \mathbb{C}^{I\times J}$, then $D_k(\mathbf{A}) \in \mathbb{C}^{J\times J}$ denotes the diagonal matrix holding row k of $\mathbf{A}$ on its diagonal.

The Frobenius norm of a tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ is defined to be

$$\|\mathcal{X}\|_F = \sqrt{\sum_{i,j,k} \left|\mathcal{X}_{ijk}\right|^2}.$$

A rank-1 tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ is equal to the outer product of some non-zero vectors $\mathbf{a} \in \mathbb{C}^I$, $\mathbf{b} \in \mathbb{C}^J$, $\mathbf{c} \in \mathbb{C}^K$ such that $(\mathcal{X})_{ijk} = (\mathbf{a})_i (\mathbf{b})_j (\mathbf{c})_k$. The rank of a tensor $\mathcal{X}$ is equal to the minimal number of rank-1 tensors needed to compose $\mathcal{X}$. Assume that the rank of $\mathcal{X}$ is R, then it can be written as

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r,$$

where $\{\mathbf{a}_r\} \subset \mathbb{C}^I$, $\{\mathbf{b}_r\} \subset \mathbb{C}^J$ and $\{\mathbf{c}_r\} \subset \mathbb{C}^K$. This decomposition is sometimes referred to as the PARAFAC decomposition. If we stack the vectors $\{\mathbf{a}_r\}$, $\{\mathbf{b}_r\}$ and $\{\mathbf{c}_r\}$ into the matrices

$$\begin{aligned}
\mathbf{A} &= [\ \mathbf{a}_1, \ \cdots \ , \mathbf{a}_R\ ] \in \mathbb{C}^{I\times R}\\
\mathbf{B} &= [\ \mathbf{b}_1, \ \cdots \ , \mathbf{b}_R\ ] \in \mathbb{C}^{J\times R}\\
\mathbf{C} &= [\ \mathbf{c}_1, \ \cdots \ , \mathbf{c}_R\ ] \in \mathbb{C}^{K\times R}
\end{aligned}$$

and let $\mathbf{X}^{(k)} \in \mathbb{C}^{I\times J}$ denote the matrix such that $\left(\mathbf{X}^{(k)}\right)_{ij} = (\mathcal{X})_{ijk}$. Then

$$\mathbf{X}^{(k)} = \mathbf{A} D_k(\mathbf{C}) \mathbf{B}^T.$$

The rest of the paper is organized as follows. The proposed parameterization of the general linear group is introduced in section 2. Then pairwise procedures for the TAF problem is discussed in section 3. Next a pairwise procedure for the JACD problems of a set of Hermitian matrices are described in section 4. Finally the proposed JACD algorithm is compared to an existing one in section 5.

2. PARAMETERIZATION OF $\mathrm{GL}_m(\mathbb{C})$

Let $\mathbf{A} \in \mathrm{GL}_m(\mathbb{C})$, then

$$\mathbf{A} = \Lambda\overline{\mathbf{A}},$$

where $\overline{\mathbf{A}} \in \mathrm{SL}_m(\mathbb{C})$ and $\Lambda = \sqrt[m]{\det(\mathbf{A})}\,\mathbf{I}_m$. Next we propose a normalized QL factorization for any nonsingular matrix.

Define the QL factorization $\overline{\mathbf{A}} = \mathbf{QL}$, where $\mathbf{Q} \in \mathrm{SO}_m(\mathbb{C})$ is unitary and $\mathbf{L}$ lower triangular. We obtain the product

$$\mathbf{A} = \Lambda\mathbf{QL},$$

where $\det(\mathbf{L}) = 1$. Assume that $\mathbf{A} \in \mathrm{GL}_2(\mathbb{C})$, then the following parameterization is possible:

$$\mathbf{A} = \begin{bmatrix} \sqrt[2]{\det(\mathbf{A})} & 0 \\ 0 & \sqrt[2]{\det(\mathbf{A})} \end{bmatrix} \begin{bmatrix} c & se^{i\phi} \\ -se^{-i\phi} & c \end{bmatrix} \begin{bmatrix} a & 0 \\ b & \frac{1}{a} \end{bmatrix},$$

where $a, b \in \mathbb{C}$, $c = \cos(\theta)$, $s = \sin(\theta)$ and $\theta, \phi \in \mathbb{R}$. In fact, $\mathbf{A}$ being nonsingular, we necessarily have that $a \neq 0$. Based on this factorization, we propose to factorize a matrix $\mathbf{A} \in \mathrm{GL}_k(\mathbb{C})$ into a product of *elementary matrices* that differ from the Identity only in four places. This idea has been extensively used for parameterizing unitary matrices with the so-called Givens rotations [6]. Here we extend the idea to general invertible matrices of $\mathrm{GL}_k(\mathbb{C})$. Consider the elementary matrix

$$\mathbf{A}[p,q] = \Lambda[p,q]\mathbf{Q}[p,q]\mathbf{L}[p,q],$$

where

$$\mathbf{A}[p,q]_{mn} = \begin{cases} 1 & \text{if } m = n \text{ and } m \notin \{p,q\} \\ \mathbf{A}_{pp} & \text{if } m = n = p \\ \mathbf{A}_{qq} & \text{if } m = n = q \\ \mathbf{A}_{pq} & \text{if } m = p \text{ and } n = q \\ \mathbf{A}_{qp} & \text{if } m = q \text{ and } n = p \\ 0 & \text{otherwise} \end{cases}$$

and

$$\Lambda[p,q]_{mn} = \begin{cases} 1 & \text{if } m = n \text{ and } m \notin \{p,q\} \\ \sqrt[k]{\det(\mathbf{A}[p,q])} & \text{if } m = n = p \\ \sqrt[k]{\det(\mathbf{A}[p,q])} & \text{if } m = n = q \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbf{Q}[p,q]_{mn} = \begin{cases} 1 & \text{if } m = n \text{ and } m \notin \{p,q\} \\ \cos(\theta) & \text{if } m = n = p \\ \cos(\theta) & \text{if } m = n = q \\ \sin(\theta)e^{i\phi} & \text{if } m = p \text{ and } n = q \\ -\sin(\theta)e^{-i\phi} & \text{if } m = q \text{ and } n = p \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbf{L}[p,q]_{mn} = \begin{cases} 1 & \text{if } m = n \text{ and } m \notin \{p,q\} \\ a & \text{if } m = n = p \\ \frac{1}{a} & \text{if } m = n = q \\ b & \text{if } m = q \text{ and } n = p \text{ and } m > n \\ 0 & \text{otherwise} \end{cases}$$

It can be shown that any $\mathbf{A} \in \mathrm{GL}_m(\mathbb{C})$ can be factorized as the product of the above three kinds of elementary matrices, viz, diagonal, lower triangular and unitary or orthogonal, e.g. we can set $\phi = 0$ in $\mathbf{Q}[p,q]$. Hence we propose to factorize a matrix $\mathbf{A} \in \mathrm{GL}_m(\mathbb{C})$ into a product of the above parameterization Givens-type matrices

$$\mathbf{A} = \prod_{s=1}^{S} \prod_{p=1}^{m-1} \prod_{q=p+1}^{m} \mathbf{A}[p,q,s],$$

where S is equal to number of sweeps executed by the iterative procedure and the third index indicate the sweep number.

3. TENSOR APPROXIMATE FITTING

Given a tensor $\mathcal{T} \in \mathbb{C}^{I \times J \times K}$ with rank R satisfying the inequalities $R \leq \min(IJ, K)$ and $R(R-1) \leq I(I-1)J(J-1)/2$. Then under mild conditions it was shown in [8] that PARAFAC decomposition problem can be reformulated as a simultaneous matrix diagonalization problem of a set of R complex symmetric matrices of dimension $R \times R$.

Moreover, given a tensor $\mathcal{T} \in \mathbb{C}^{I \times J \times K}$ with rank $R \leq \min(I, J, K)$ then by a dimension reduction step as explained in [7] the PARAFAC decomposition problem can be reformulated to the estimation of the parameters of a tensor $\mathcal{X} \in \mathbb{C}^{R \times R \times K}$. This is typically done by minimizing the function

$$\begin{aligned} f(\mathbf{U},\mathbf{V},\mathbf{W}) &= \sum_{k=1}^{K} \|\mathbf{X}^{(k)} - \mathbf{U}D_k(\mathbf{W})\mathbf{V}^T\|_F^2 \\ &= \sum_{k=1}^{K} \|\mathrm{Vec}\left(\mathbf{X}^{(k)}\right) - (\mathbf{V} \odot \mathbf{U})\,\mathrm{Vecd}\,(D_k(\mathbf{W}))\|_F^2. \end{aligned} \quad (1)$$

The optimal $D_k(\mathbf{W})$ can be expressed as follows:

$$D_k(\mathbf{W}) = \mathrm{diag}\left((\mathbf{V} \odot \mathbf{U})^{\dagger}\,\mathrm{Vec}\left(\mathbf{X}^{(k)}\right)\right), \forall k. \quad (2)$$

Inserting (2) into (1) we obtain

$$f_{\mathbf{W}}(\mathbf{U},\mathbf{V}) = \sum_{k=1}^{K} \|\mathbf{X}^{(k)} - \mathbf{U}\mathrm{diag}\left((\mathbf{V} \odot \mathbf{U})^{\dagger}\,\mathrm{Vec}\left(\mathbf{X}^{(k)}\right)\right)\mathbf{V}^T\|_F^2.$$

By assuming that $\mathbf{X}^{(k)} = \mathbf{U}D_k(\mathbf{W})\mathbf{V}^T$ for all k and $\mathbf{U},\mathbf{V} \in \mathrm{GL}_R(\mathbb{C})$, then obviously

$$D_k(\mathbf{W}) = \mathrm{diag}\left(\mathbf{U}^{-1}\mathbf{X}^{(k)}\mathbf{V}^{-T}\right). \quad (3)$$

Inserting (3) into (1) we obtain the cost function

$$\begin{aligned} g(\mathbf{U},\mathbf{V}) &= \sum_{k=1}^{K} \|\mathbf{X}^{(k)} - \mathbf{U}\mathrm{diag}\left(\mathbf{U}^{-1}\mathbf{X}^{(k)}\mathbf{V}^{-T}\right)\mathbf{V}^T\|_F^2 \quad (4) \\ &= g(\mathbf{U}\Lambda_{\mathbf{U}},\mathbf{V}\Lambda_{\mathbf{V}}), \end{aligned}$$

where $\Lambda_{\mathbf{U}}$ and $\Lambda_{\mathbf{V}}$ are arbitrary nonsingular diagonal matrices. A first approach to reduce the number of unknowns would be to use (2). A simpler approach would be to make use of (4). This can be seen as a non-symmetric version of the cost function introduced by Afsari in [1] for JACD of Hermitian matrices but applied on the TAF problem.

Under the assumption that $D_k(\mathbf{W})$ is nonsingular for some k, say $k = 1$, then as proposed in [7] another simplification can be considered which will result in a different optimization problem. This will yield a joint diagonalization problem by a congruence transform

$$\mathbf{Y}^{(k)} = \mathbf{X}^{(k)}\mathbf{X}^{(1)^{-1}} = \mathbf{U}D_k(\mathbf{W})D_1(\mathbf{W})^{-1}\mathbf{U}^{-1}. \quad (5)$$

By the same argument as before we can from (5) obtain the following cost function

$$\begin{aligned} h(\mathbf{U}) &= \sum_{k=1}^{K} \|\mathbf{Y}^{(k)} - \mathbf{U}\mathrm{diag}\left(\mathbf{U}^{-1}\mathbf{Y}^{(k)}\mathbf{U}\right)\mathbf{U}^{-1}\|_F^2 \quad (6) \\ &= h(\mathbf{U}\Lambda_{\mathbf{U}}), \end{aligned}$$

where $\Lambda_{\mathbf{U}}$ is an arbitrary nonsingular diagonal matrix. Due to the invariance property of the cost functions (4) and (6) the term $\Lambda[p,q]$ vanish and in the lower triangular matrices we can set $a = 1$. Furthermore, since $h(\mathbf{U}\Pi) = h(\mathbf{U})$, where Π is an arbitrary permutation matrix the parameterization $\mathbf{A}[p,q]$ could be reduced to

$$\mathbf{A}[p,q]_{mn} = \begin{cases} 1 & \text{if } m = n \\ a & \text{if } m = p \text{ and } n = q \\ b & \text{if } m = q \text{ and } n = q \\ 0 & \text{otherwise} \end{cases}$$

where $1 - ab \neq 0$ and $a,b \in \mathbb{C}$.

Remark that minimizing $f_{\mathbf{W}}(\mathbf{U},\mathbf{V})$, $g(\mathbf{U},\mathbf{V})$ and $h(\mathbf{U})$ will in general yield different solutions, unless the tensor considered is exactly diagonalizable by invertible transforms. An outline of a general sweeping procedure addressing this kind of problems can be seen on algorithm 1.

4. JOINT APPROXIMATE CONGRUENT DIAGONALIZATION OF HERMITIAN MATRICES

Various BSS problems are often addressed by minimizing the off-diagonal terms of a set of Hermitian matrices. Let $\{\mathbf{T}^{(k)}\}_{k=1}^{K} \subset \mathrm{S}_m(\mathbb{C})$ be the matrices we attempt to diagonalize. Then the minimization of the

Algorithm 1 Outline of sweeping procedure for TAF.

```
Initialize: U = I_m, V = I_m
Step 1: Repeat until convergence
for p = 1 to m − 1 do
  for q = p + 1 to m do
    calculate optimal A_U[p,q] and A_V[p,q]
    U ← U A_U[p,q] , V ← V A_V[p,q]
  end for
end for
Step 2: Check if algorithm has converged. If not, then
go to Step 1.
```

Algorithm 2 Outline of sweeping procedure for minimization of the off-term elements of a set of Hermitian matrices.

```
Initialize: U = I_m
Step 1: Repeat until convergence
for p = 1 to m − 1 do
  for q = p + 1 to m do
    calculate Q[p,q] according to subsection 4.1.
    T^(k) ← Q[p,q]^H T^(k) Q[p,q]
    U ← Q[p,q]^H U
    calculate L[p,q] according to subsection 4.2.
    T^(k) ← L[p,q] T^(k) L[p,q]^H
    U ← L[p,q] U
  end for
end for
Step 2: Check if algorithm has converged. If not, then
go to Step 1.
```

cost function

$$J(\mathbf{U}) = \sum_{k=1}^{K} \|\mathbf{U}\mathbf{T}^{(k)}\mathbf{U}^H - \mathrm{diag}\left(\mathbf{U}\mathbf{T}^{(k)}\mathbf{U}^H\right)\|_F^2 \quad (7)$$

have been applied in [11], [16]. This can be interpreted as an extension of the JADE algorithm [3] to the case of general invertible transforms.

The updating rules are $\mathbf{T}^{(k)} \leftarrow \mathbf{A}[p,q]\mathbf{T}^{(k)}\mathbf{A}[p,q]^H$ and $\mathbf{U} \leftarrow \mathbf{A}[p,q]\mathbf{U}$. Let us parameterize $\mathbf{A}[p,q] \in \mathbb{C}^{m\times m}$ as $\mathbf{A}[p,q] = \sqrt[m]{\det(\mathbf{A}[p,q])}\mathbf{Q}[p,q]\mathbf{L}[p,q]$. Then since

$$J(\alpha\mathbf{U}) = \alpha^2\alpha^{*2}J(\mathbf{U}), \quad \forall\alpha \in \mathbb{C}, \quad (8)$$

we can infer from equation (8) that the scalar $\sqrt[m]{\det(\mathbf{A}[p,q])}$ does not contribute to any diagonalization of the matrices $\{\mathbf{T}^{(k)}\}$. . Hence we will only consider the term $\mathbf{Q}[p,q]\mathbf{L}[p,q] \in \mathrm{SL}_m(\mathbb{C})$.

Due to the computational complexity of the problem, we calculate and update only one matrix at a time: either the unitary matrix or the triangular one. An outline of the proposed sweeping procedure can be seen on algorithm 2.

4.1 Algebraic Solution for the Unitary Matrix

The algebraic solution for the unitary Jacobi-subproblem was solved in [3], [4] known as the JADE problem and therefore the JADE algorithm can be used to solve the problem. Let $\mathbf{U} = \mathbf{Q}[p,q]$, then due to the invariance property of the Frobenius norm wrt. any unitary operator, the unitary Jacobi-subproblem is equivalent to maximizing the quadratic form

$$f(\mathbf{Q}[p,q]) = \sum_{k=1}^{K} \|\text{diag}\left(\mathbf{Q}[p,q]^H \mathbf{T}^{(k)} \mathbf{Q}[p,q]\right)\|_F^2 =$$

$$\mathbf{s}^T \frac{1}{2} \begin{bmatrix} 2\gamma_1 & \text{Re}\{\gamma_2\} & \text{Im}\{\gamma_2\} \\ \text{Re}\{\gamma_2\} & \gamma_5 + 2\text{Re}\{\gamma_4\} & 2\text{Im}\{\gamma_4\} \\ \text{Im}\{\gamma_2\} & 2\text{Im}\{\gamma_4\} & \gamma_5 - 2\text{Re}\{\gamma_4\} \end{bmatrix} \mathbf{s},$$

where

$$\begin{aligned} \mathbf{s} &= \left[\cos(2\theta), \sin(2\theta)\cos(\phi), \sin(2\theta)\sin(\phi)\right]^T \\ \gamma_1 &= \sum_{k=1}^{K} \left|\mathbf{T}_{pp}^{(k)}\right|^2 + \left|\mathbf{T}_{qq}^{(k)}\right|^2 \\ \gamma_2 &= \sum_{k=1}^{K} \mathbf{T}_{qq}^{(k)*}\mathbf{T}_{pq}^{(k)} + \mathbf{T}_{qp}^{(k)*}\mathbf{T}_{qq}^{(k)} - \mathbf{T}_{pp}^{(k)*}\mathbf{T}_{pq}^{(k)} - \mathbf{T}_{qp}^{(k)*}\mathbf{T}_{pp}^{(k)} \\ \gamma_3 &= \sum_{k=1}^{K} 2\text{Re}\{\mathbf{T}_{pp}^{(k)*}\mathbf{T}_{qq}^{(k)}\} + \left|\mathbf{T}_{pq}^{(k)}\right|^2 + \left|\mathbf{T}_{qp}^{(k)}\right|^2 \\ \gamma_4 &= \sum_{k=1}^{K} \mathbf{T}_{qp}^{(k)*}\mathbf{T}_{pq}^{(k)} \\ \gamma_5 &= \gamma_1 + \gamma_3 \end{aligned}$$

Hence the problem amounts to find an eigenvector associated to the largest eigenvalue of the above matrix.

4.2 Algebraic Solution for the Lower Triangular Matrix

Let $\mathbf{U} = \mathbf{L}[p,q]$ and $\alpha = ab$, then

$$J(\mathbf{L}[p,q]) = \beta_2\alpha\alpha^* + \beta_1\alpha + \beta_1^*\alpha^* + \beta_0 \triangleq f(\alpha,\alpha^*),$$

where

$$\begin{aligned} \beta_2 &= \sum_{k=1}^{K} 2\left|\mathbf{T}_{pp}^{(k)}\right|^2 \\ \beta_1 &= \sum_{k=1}^{K} \left(\mathbf{T}_{pp}^{(k)}\mathbf{T}_{qp}^{(k)*} + \mathbf{T}_{pp}^{(k)}\mathbf{T}_{pq}^{(k)}\right) \\ \beta_0 &= \sum_{k=1}^{K} 2\left|\mathbf{T}_{qp}^{(k)}\right|^2 \end{aligned}$$

Let us apply the formal partial derivatives [2] , then the stationary points of $f(\alpha,\alpha^*)$ can be found to be

$$\frac{\partial f(\alpha,\alpha^*)}{\partial \alpha^*} = \beta_2\alpha + \beta_1^* = 0 \Leftrightarrow \alpha = -\frac{\beta_1^*}{\beta_2},$$

which exists when $\beta_2 \neq 0$. Furthermore, when it exists it is a global minimum since

$$\frac{\partial^2 f(\alpha,\alpha^*)}{\partial \alpha^* \partial \alpha} = \beta_2 > 0.$$

When a solution exists the pair $a = 1$ and $b = -\frac{\beta_1^*}{\beta_2}$ will be used. On the other hand, when no solution exists then we make use of $a = 1$ and $b = 0$.

5. COMPUTER RESULTS

The comparison of the proposed algorithm, here called SL, and the method by Li [11], called FAJD, will be on the test data $\{\mathbf{T}_k\}_{k=1}^{20} \subset \mathrm{S}_5(\mathbb{C})$, where $\mathbf{T}_k = \mathbf{U}\mathbf{D}_k\mathbf{U}^H + \beta\mathbf{E}_k^H\mathbf{E}_k$, $\mathbf{U} \in \mathbb{C}^{5\times5}$, $\mathbf{E}_k \in \mathrm{S}_5(\mathbb{C})$, $\beta \in \mathbb{R}$ and $\mathbf{D}_k$ are diagonal matrices with $(\mathbf{D}_k)_{ii} \in \mathbb{R}$. The real and imaginary parts of the involved matrices are randomly drawn elements from a uniform distribution with support over the interval $[-1,1]$.

In order to avoid degenerate solutions Li proposed to add the term $\log(\det(\mathbf{U}))$ to the functional $J(\mathbf{U})$. This additional term will penalize ill-conditioned solutions consisting of small singular values. A measure on how well-conditioned a given solution is, is given by the ratio

$$\kappa(\mathbf{U}) = \frac{\sigma_{max}(\mathbf{U})}{\sigma_{min}(\mathbf{U})},$$

where $\sigma_{max}(\mathbf{U})$ and $\sigma_{min}(\mathbf{U})$ denotes the maximal and minimal singular values of $\mathbf{U}$ respectively.

In the SL method an iteration is equal to one sweep while in the FAJD method an iteration corresponds to an update of all the parameters of the transform matrix.

Let

$$snr = 10\log\left(\frac{\sum_{k=1}^{20} \|\mathbf{U}\mathbf{D}_k\mathbf{U}^H\|_F^2}{\sum_{k=1}^{20} \|\beta\mathbf{E}_k\|_F^2}\right)$$

and

$$J_n(\mathbf{U}) = \frac{\sum_{k=1}^{20} \|\mathbf{U}\mathbf{T}_k\mathbf{U}^H - \text{diag}\left(\mathbf{U}\mathbf{T}_k\mathbf{U}^H\right)\|_F^2}{\sum_{k=1}^{20} \|\mathbf{U}\mathbf{T}_k\mathbf{U}^H\|_F^2}.$$

Then the mean and median convergence of the algorithms over 100 trials after 100 iterations as measured by J_n when snr varies from 0 to 20 with a hop factor 5 can be seen on figure 1 and 2 respectively. Furthermore the mean and median condition numbers of the solutions can be seen on figures 3 and 4 respectively.

6. SUMMARY

A new parameterization of the general linear group has been proposed. Based on this parameterization sweeping algorithms for PARAFAC estimation and congruent diagonalization of a set of matrices have been proposed. Preliminary comparisons with an existing congurent diagonalization algorithm have been reported.

REFERENCES

[1] B. Afsari , "Sensitivity Analysis for the Problem of Matrix Joint Diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 30, pp. 1148–1171 . 2008.

[2] D.H. Brandwood , "A Complex Gradient Operator and its Applications in Adaptive Antenna Theory," *IEE Proceedings-F and H*, vol. 130, pp. 11–16 . 1983.

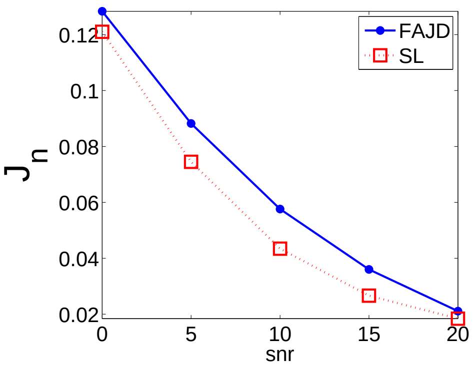

Figure 1: The mean diagonalization of a set of matrices on the form $\mathbf{T}_k = \mathbf{U}\mathbf{D}_k\mathbf{U}^H + \beta\mathbf{E}_k^H\mathbf{E}_k$ when β is varying.

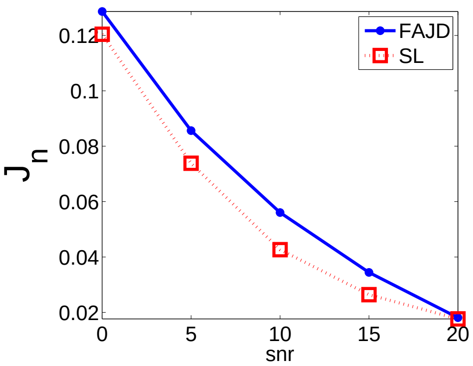

Figure 2: The median diagonalization of a set of matrices on the form $\mathbf{T}_k = \mathbf{U}\mathbf{D}_k\mathbf{U}^H + \beta\mathbf{E}_k^H\mathbf{E}_k$ when β is varying.

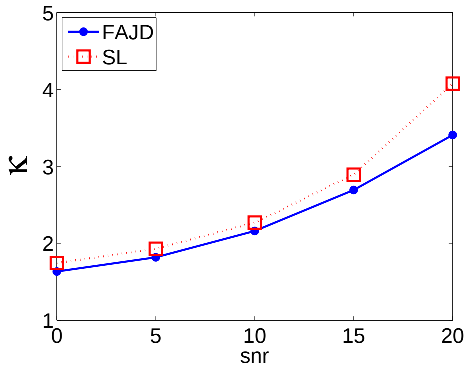

Figure 3: The mean condition number of the diagonalization matrix $\mathbf{U}$ when *snr* is varying.

[3] J.-F. Cardoso and A. Souloumiac , "Blind Beamforming for non-Gaussian signals," *IEE Proceedings-F*, vol. 140, pp. 362–370 . 1993.

[4] J.-F. Cardoso and A. Souloumiac , "Jacobi angles for simultaneous diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 17, pp.161–164 . 1996.

[5] P. Comon, *Independent Component Analysis*, in Higher Order Statistics, pp. 29–38. Elsevier, 1992.

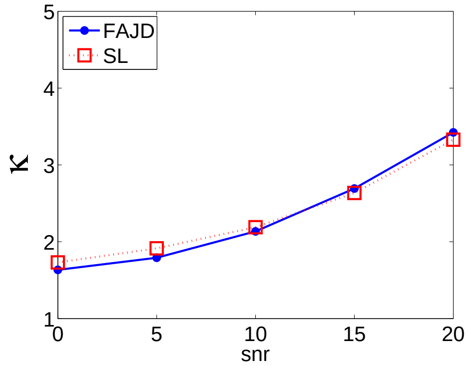

Figure 4: The median condition number of the diagonalization matrix $\mathbf{U}$ when *snr* is varying.

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

[7] L. De Lathauwer and B. De Moor and J. Vandewalle , "Computation of the Canonical Decomposition by means of a Simultaneous Generalized Schur Decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 26, pp. 295–327. 2004.

[8] L. De Lathauwer , "A Link between the Canonical Decomposition in multilinear algebra and Simultaneous Matrix Diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 28, pp. 642–666. 2006.

[9] L. De Lathauwer and J. Castaing and J.-F. Cardoso, "Fourth-Order Cumulant-Based Blind Identification of Underdetermined Mixtures," *IEEE Trans. Signal Processing*, vol. 55, pp. 2965–2973. 2007.

[10] L. De Lathauwer and J. Castaing, "Blind Identification of Underdetermined Mixtures by Simultaneous Matrix Diagonalization," *IEEE Trans. Signal Processing*, vol. 56, pp. 1096–1105. 2008.

[11] X.-L. Li and X.-D. Zhang, "Nonorthogonal joint diagonalization Free of Degenerate Solution ," *IEEE Trans. Signal Processing*, vol. 55, pp. 1803–1814. 2007.

[12] O. J. Micka and A. J. Weiss, "Estimating Frequencies of Exponentials in Noise Using Joint Diagonalization," *IEEE Trans. Signal Processing*, vol. 47, pp. 341–348. 1999.

[13] D.-T. Pham and J.-F. Cardoso, "Blind Separation of Instantaneous Mixtures of non Stationary Sources," *IEEE Trans. Signal Processing*, vol. 49, pp. 1837–1848. 2001.

[14] D.-T. Pham, "Joint Approximate Diagonalization of Positive Definite Hermitian Matrices," *SIAM J. Matrix Anal. Appl.*, vol. 22, pp. 1136–1152. 2001.

[15] A.-J. Van Der Veen and A. Paulraj, "An Analytical Constant Modulus Algorithm," *IEEE Trans. Signal Processing*, vol. 44, pp. 1136–1155. 1996.

[16] A. Ziehe and P. Laskov and G. Nolte and K.-L. Müller, "A Fast Algorithm for Joint Diagonalization with Non-orthogonal Transformations and its Application to Blind Source Separation," *Journal of Machine Learning Research*, vol. 5, pp. 777–800. 2004.