COUPLED RANK-(L_m, L_n, ·) BLOCK TERM DECOMPOSITION BY COUPLED BLOCK SIMULTANEOUS GENERALIZED SCHUR DECOMPOSITION

Xiao-Feng Gong[*†‡], *Qiu-Hua Lin*[*], *Otto Debals*[†‡§], *Nico Vervliet*[‡§], *Lieven De Lathauwer*[†‡§]

[*]School of Information and Communication Engineering, Dalian University of Technology, China, [†]Group Science, Engineering and Technology, KU Leuven, KULAK, Belgium, [‡]Department of Electrical Engineering (ESAT), KU Leuven, Belgium, [§] iMinds Medical IT, Leuven, Belgium. E-mails: {xfgong, qhlin}@dlut.edu.cn, {otto.debals, nico.vervliet}@esat.kuleuven.be, lieven.delathauwer@kuleuven-kulak.be

ABSTRACT

Coupled decompositions of multiple tensors are fundamental tools for multi-set data fusion. In this paper, we introduce a coupled version of the rank-(L_m, L_n, ·) block term decomposition (BTD), applicable to joint independent subspace analysis. We propose two algorithms for its computation based on a coupled block simultaneous generalized Schur decomposition scheme. Numerical results are given to show the performance of the proposed algorithms.

Index Terms — Tensor, block term decomposition, coupled tensor decomposition, multi-set data fusion

1. INTRODUCTION

In the past decade, the block term decomposition (BTD) has attracted increased attention in various signal processing applications. In comparison to the well-known canonical polyadic decomposition (CPD) that writes a tensor as the sum of a minimal number of rank-1 terms, BTD decomposes a tensor into a set of terms of low multilinear rank, which is more flexible and better adapted to applications where the signals are multidimensional. Uniqueness and algorithms for various types of BTD were studied in [1-4] and applications in signal processing were reported in [5-10].

This research is funded by: (1) National natural science foundation of China (nos. 61331019 and 61379012); (2) Scientific Research Fund of Liaoning Education Department (no. L2014016); (3) China Scholarship Council; (4) a Ph.D grant of the Agency for Innovation by Science and Technology (IWT); (5) An Aspirant Grant from the Research Foundation−Flanders (FWO); (6) Research Council KU Leuven: C1 project C16/15/059-nD, CoE PFV/10/002 (OPTEC). (7) FWO: project G.0830.14N, G.0881.14N; (7) the Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO II, Dynamical systems, control and optimization, 2012-2017), (8) EU: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (no. 339804). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information.

Recently, coupled tensor decompositions have been studied in the context of multi-set data fusion, which assumes some links or shared factor matrices among target tensors. They are mainly extensions of CPD to multi-set cases, including the works in [11-20] and the references therein. In addition, the coupled rank-($L_{r,n}$, $L_{r,n}$, 1) BTD was studied in [15, 16]. Structured data fusion (SDF) was presented in [21] as a flexible framework for coupled and/or structured decompositions. The above works have shown that the coupled decomposition is able to both improve the accuracy and relax the identifiability condition compared with the decomposition of a single tensor.

In this paper, we combine the concepts of coupling and BTD, and introduce a particular coupled rank-(L_m, L_n, ·) BTD. We will show that the particular multi-set data fusion problem of joint independent subspace analysis (J-ISA) can be turned into a coupled rank-(L_m, L_n, ·) BTD. We present two algorithms for the coupled block version of so-called simultaneous generalized Schur decomposition scheme (SGSD, [22-24]). SGSD involves only unitary factors. It has found use in the analytical constant modulus algorithm [22], in the computation of CPD [23], in numerically stable representations of ill-posed CPD problems [24], etc. We here limit ourselves to the overdetermined case. The more challenging under-determined coupled rank-(L_m, L_n, ·) BTD will be addressed in the future.

Notation and definitions: vectors, matrices and tensors are denoted by lowercase boldface, uppercase boldface and uppercase calligraphic letters, respectively. The rth column vector and the (i, j)th entry of $\boldsymbol{A}$ are denoted by $\boldsymbol{a}_r$ and $a_{i,j}$, respectively. Symbols '$\otimes$', '$\odot$', '$\times_n$' and '$\circ$' denote the Kronecker product, block-wise Khatri-Rao product, mode-n product, and outer product, respectively, defined as:

$$\boldsymbol{A}\otimes\boldsymbol{B}\triangleq\begin{bmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}, \boldsymbol{A}\odot\boldsymbol{B}\triangleq[\boldsymbol{A}_1\otimes\boldsymbol{B}_1,\cdots,\boldsymbol{A}_R\otimes\boldsymbol{B}_R],$$

$$(\mathcal{T}\times_n\boldsymbol{G})_{i_1,\ldots,i_{n-1},j,i_{n+1},\ldots,i_N}\triangleq\sum\nolimits_n t_{i_1,\ldots,i_N}g_{j,i_n}, \ (\boldsymbol{a}\circ\boldsymbol{b}\circ\boldsymbol{c})_{i,j,k}\triangleq a_i b_j c_k.$$

978-1-4799-9988-0/16/$31.00 ©2016 IEEE ICASSP 2016

In the above definitions, we assume $\boldsymbol{A} \triangleq [\boldsymbol{A}_1, \cdots, \boldsymbol{A}_R]$ and $\boldsymbol{B} \triangleq [\boldsymbol{B}_1, \cdots, \boldsymbol{B}_R]$ for the block-wise Khatri-Rao product. For the mode-n product we assume that the nth dimension of $\mathcal{T}$ is equal to the number of columns of $\boldsymbol{G}$. We denote the identity matrix and the all-zero matrix as $\boldsymbol{I}_M \in \mathbb{C}^{M\times M}$ and $\boldsymbol{0}_{M,N} \in \mathbb{C}^{M\times N}$, respectively. Transpose, conjugate, conjugated transpose, Moore-Penrose pseudo inverse, and Frobenius norm are denoted as $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, $(\cdot)^\dagger$, $\|\cdot\|_F$, respectively. Mathematical expectation is denoted as $\mathrm{E}\{\}$. MATLAB notations will be used to denote submatrices of a tensor. For instance, we use $\mathcal{T}_{(:,:,k)}$ to denote the frontal slice of a tensor by fixing the third index to k.

For a given matrix $\boldsymbol{T} \in \mathbb{C}^{I\times J}$, $\mathrm{vec}(\boldsymbol{T}) \triangleq [\boldsymbol{t}_1^T, \cdots, \boldsymbol{t}_J^T]^T \in \mathbb{C}^{IJ}$ denotes column-wise vectorization of $\boldsymbol{T}$ and $\mathrm{unvec}(\cdot)$ performs the inverse. For a third-order tensor $\mathcal{T} \in \mathbb{C}^{I\times J\times K}$, notations $\boldsymbol{T}_{(1,2),3} \in \mathbb{C}^{IJ\times K}, \boldsymbol{T}_{(2,3),1} \in \mathbb{C}^{JK\times I}, \boldsymbol{T}_{(1,3),2} \in \mathbb{C}^{IK\times J}$ denote three types of matricization, defined by:

$$\left(\boldsymbol{T}_{(1,2),3}\right)_{(i-1)J+j,k} = \left(\boldsymbol{T}_{(2,3),1}\right)_{(j-1)K+k,i} = \left(\boldsymbol{T}_{(1,3),2}\right)_{(i-1)K+k,j} = t_{i,j,k}.$$

The mode-n vectors of $\mathcal{T}$ are obtained by fixing all but the nth index of $\mathcal{T}$. The mode-n rank of $\mathcal{T}$ is defined as the dimension of the subspace spanned by the mode-n vectors. It is easy to understand that the mode-1, mode-2, and mode-3 rank of $\mathcal{T}$ is equal to the rank of $\boldsymbol{T}_{(2,3),1}, \boldsymbol{T}_{(1,3),2}$ and $\boldsymbol{T}_{(1,2),3}$, respectively. Third-order tensors with mode-1, mode-2, mode-3 rank equal to L_1, L_2, L_3, respectively, are said to have multilinear rank-(L_1, L_2, L_3). Third-order tensors with mode-1 and mode-2 rank equal to L_1 and L_2, without rank constraint in the third mode, are said to have multilinear rank-(L_1, L_2, $\cdot$). Rank-(1,1,1) terms simply correspond to rank-1 terms.

2. PROBLEM FORMULATION

In this paper, we consider the following coupled rank-(L_m, L_n, $\cdot$) BTD of a set of tensors $\mathcal{T}^{(m,n)} \in \mathbb{C}^{I_m\times I_n\times K}$:

$$\mathcal{T}^{(m,n)} = \sum_{r=1}^{R} \mathcal{C}_r^{(m,n)} \times_1 \boldsymbol{A}_r^{(m)} \times_2 \boldsymbol{A}_r^{(n)*}, \quad m,n = 1,\ldots,M, \quad (1)$$

where $\mathcal{C}_r^{(m,n)} \in \mathbb{C}^{L_m\times L_n\times K}$ has mode-1 rank equal to L_m and mode-2 rank equal to L_n, and $\boldsymbol{A}_r^{(m)} \in \mathbb{C}^{I_m\times L_m}$ has full column rank. Eq. (1) suggests that each tensor $\mathcal{T}^{(m,n)}$ admits by itself a rank-(L_m, L_n, $\cdot$) BTD [1] (see Fig. 1 for an illustration). In addition, each tensor $\mathcal{T}^{(m,n)}$ is coupled with $\mathcal{T}^{(m,n')}$ in the first mode by factor matrix $\boldsymbol{A}^{(m)}$ and at the same time coupled with $\mathcal{T}^{(m',n)}$ in the second mode by $\boldsymbol{A}^{(n)*}$, $m' \neq m$, $n' \neq n$. This double coupling structure is illustrated in Fig. 2. The matrix representation of (1) is given by:

$$\boldsymbol{T}_{(1,2),3}^{(m,n)} = (\boldsymbol{A}^{(m)} \odot \boldsymbol{A}^{(n)*}) \cdot \boldsymbol{C}^{(m,n)}, \quad (2)$$

where $\boldsymbol{C}^{(m,n)} \triangleq [(\mathcal{C}_1^{(m,n)})_{(1,2),3}^T, \ldots, (\mathcal{C}_R^{(m,n)})_{(1,2),3}^T]^T \in \mathbb{C}^{RL_mL_n\times K}$. In addition, the frontal slices of $\mathcal{T}^{(m,n)}$ take the following form:

$$\mathcal{T}_{(:,:,k)}^{(m,n)} = \boldsymbol{A}^{(m)} \boldsymbol{\Sigma}_k^{(m,n)} \boldsymbol{A}^{(n)H}, \quad (3)$$

where $\boldsymbol{\Sigma}_k^{(m,n)} \in \mathbb{C}^{RL_m\times RL_n}$ is a block-diagonal matrix with blocks of size $L_m \times L_n$, containing the kth frontal slice of $\mathcal{C}_r^{(m,n)}$ as the rth block on its main diagonal.

Note that in (1) we can arbitrarily permute the terms if it is done consistently for all tensors involved. We can also post-multiply $\boldsymbol{A}^{(m)}$ and $\boldsymbol{A}^{(n)*}$ by non-singular matrices $\boldsymbol{F}_r^{(m)} \in \mathbb{C}^{L_m\times L_m}$ and $\boldsymbol{F}_r^{(n)*} \in \mathbb{C}^{L_n\times L_n}$ provided that $\mathcal{C}_r^{(m,n)}$ is replaced by $\mathcal{C}_r^{(m,n)} \times_1 \boldsymbol{F}_r^{(m)-1} \times_2 \boldsymbol{F}_r^{(n)*-1}$ for all values of m and n. The goal of coupled rank-(L_m, L_n, $\cdot$) BTD is then to solve (1) up to these trivial indeterminacies.

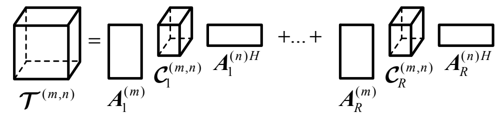


Fig. 1. The rank-(L_m, L_n, $\cdot$) BTD writes each target tensor $\mathcal{T}^{(m,n)}$ as the sum of multiple block terms of low multilinear rank.

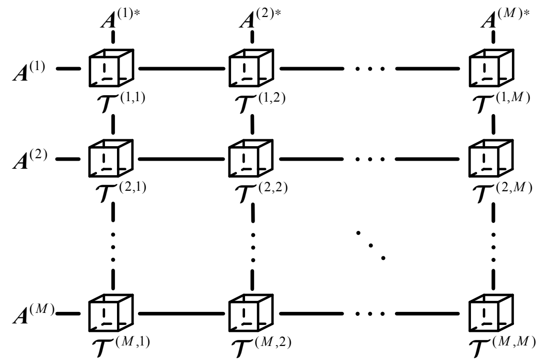


Fig. 2. The double coupling structure. The target tensors are placed at different nodes of a grid according to their indices. Each tensor is coupled with tensors along two modes by two factor matrices.

Next we explain how the coupled rank-(L_m, L_n, $\cdot$) BTD (1) is related to the J-ISA problem [25]. The multi-set data model for J-ISA is formulated as:

$$\boldsymbol{x}^{(m)}(t) = \boldsymbol{A}^{(m)}\boldsymbol{s}^{(m)}(t) = \sum_{r=1}^{R} \boldsymbol{A}_r^{(m)}\boldsymbol{s}_r^{(m)}(t), \quad m = 1,\ldots,M, \quad (4)$$

where $\boldsymbol{x}^{(m)}(t) \in \mathbb{C}^{I_m}$ denotes the observed mixture, and $\boldsymbol{s}^{(m)}(t) \in \mathbb{C}^{RL_m}$ is the latent source vector at time instant t, and $\boldsymbol{A}^{(m)} \in \mathbb{C}^{I_m\times RL_m}$ denotes the mixing matrix of the mth dataset. We partition the source vector $\boldsymbol{s}^{(m)}(t)$ into R sub-vectors $[\boldsymbol{s}_1^{(m)T}(t), \ldots, \boldsymbol{s}_R^{(m)T}(t)]^T \triangleq \boldsymbol{s}^{(m)}(t)$, with each sub-vector $\boldsymbol{s}_r^{(m)}(t)$ representing a group of L_m sources. In addition we partition the mixing matrix $[\boldsymbol{A}_1^{(m)}, \ldots, \boldsymbol{A}_R^{(m)}] \triangleq \boldsymbol{A}^{(m)}$ where $\boldsymbol{A}_r^{(m)} \in \mathbb{C}^{I_m\times L_m}$ is the sub-matrix of $\boldsymbol{A}^{(m)}$ associated with $\boldsymbol{s}_r^{(m)}(t)$.

We make the following assumptions: 1) each sub-vector $\boldsymbol{s}_r^{(m)}(t)$ contains L_m dependent sources; 2) $\boldsymbol{s}_{r_1}^{(m_1)}(t)$ and $\boldsymbol{s}_{r_2}^{(m_2)}(t)$ are independent for $r_1 \neq r_2$ and dependent for $r_1 = r_2$ regardless of the values of m_1 and m_2; 3) the sources are temporally non-stationary with zero mean and unit variance.

We calculate the cross-covariance tensor as follows:

$$\mathcal{T}_{(:,:,k)}^{(m,n)} = \mathrm{E}\{\boldsymbol{x}^{(m)}(k)[\boldsymbol{x}^{(n)}(k)]^H\} = \boldsymbol{A}^{(m)}\boldsymbol{\Sigma}_k^{(m,n)}\boldsymbol{A}^{(n)H}, \quad (5)$$

where $\mathcal{T}^{(m,n)} \in \mathbb{C}^{I_m\times I_n\times K}$, $1 \leq m,n \leq M$, K denotes the number of time-frames for which such a cross-covariance is computed. The matrix $\boldsymbol{\Sigma}_k^{(m,n)} \triangleq \mathrm{E}\{\boldsymbol{s}^{(m)}(k)[\boldsymbol{s}^{(n)}(k)]^H\}$ is block-diagonal with blocks of size $L_m \times L_n$ under the above assumptions. Comparing (5) and (3), we can see that the J-ISA data model has been converted to a coupled rank-(L_m, L_n, $\cdot$) BTD formulation by using second-order statistics.

3. ALGORITHMS

We limit ourselves to the overdetermined case where $\boldsymbol{A}^{(m)}$ and $\boldsymbol{C}^{(m,n)}$ in (2) are assumed to have full column rank for all the values of m,n, and propose two algorithms for the computation of coupled rank-(L_m, L_n, .) BTD, which extend the SGSD scheme to the coupled block case.

3.1. Coupled Block SGSD

The goal of coupled rank-(L_m, L_n, $\cdot$) BTD is to find $\boldsymbol{A}^{(m)}$, $\boldsymbol{A}^{(n)}$ and $\boldsymbol{\Sigma}_k^{(m,n)}$ that minimize the following function:

$$\eta(\Omega) = \sum_{m,n=1}^{M}\sum_{k=1}^{K}\left\|\boldsymbol{\mathcal{T}}_{(:,:,k)}^{(m,n)} - \boldsymbol{A}^{(m)}\boldsymbol{\Sigma}_k^{(m,n)}\boldsymbol{A}^{(n)H}\right\|_F^2, \quad (6)$$

where $\Omega \triangleq \{\boldsymbol{A}^{(m)}, \boldsymbol{\Sigma}_k^{(m,n)}, m,n = 1,...,M, k = 1,...,K\}$ denotes the set containing all the arguments. We let $\boldsymbol{A}^{(m)} \triangleq \boldsymbol{Q}^{(m)H}\boldsymbol{R}^{(m)}$ be a block QR decomposition of $\boldsymbol{A}^{(m)}$ and $\boldsymbol{A}^{(n)} \triangleq \boldsymbol{Z}^{(n)}\boldsymbol{L}^{(n)}$ be a block ZL decomposition of $\boldsymbol{A}^{(n)}$, where $\boldsymbol{Q}^{(m)} \in \mathbb{C}^{I_m \times I_m}$, $\boldsymbol{Z}^{(n)} \in \mathbb{C}^{I_n \times I_n}$ are unitary matrices, and $\boldsymbol{R}^{(m)} \in \mathbb{C}^{I_m \times RL_m}$, $\boldsymbol{L}^{(n)} \in \mathbb{C}^{I_n \times RL_n}$ are defined as:

$$\boldsymbol{R}^{(m)} \triangleq \begin{bmatrix} \bar{\boldsymbol{R}}^{(m)} \\ \boldsymbol{0}_{(I_m - RL_m), RL_m} \end{bmatrix}, \boldsymbol{L}^{(n)} \triangleq \begin{bmatrix} \bar{\boldsymbol{L}}^{(n)} \\ \boldsymbol{0}_{(I_n - RL_n), RL_n} \end{bmatrix}, \quad (7)$$

where $\bar{\boldsymbol{R}}^{(m)} \in \mathbb{C}^{RL_m \times RL_m}$ is an block upper-triangular matrix and $\bar{\boldsymbol{L}}^{(n)} \in \mathbb{C}^{RL_n \times RL_n}$ is a block lower-triangular matrix. We denote $\boldsymbol{\Gamma}_k^{(m,n)} \triangleq \boldsymbol{R}^{(m)}\boldsymbol{\Sigma}_k^{(m,n)}\boldsymbol{L}^{(n)H} \in \mathbb{C}^{I_m \times I_n}$, of which the top-left $RL_m \times RL_n$ submatrix is block upper-triangular. By substitution of these expressions into (6) we obtain a coupled block SGSD where the goal is to minimize the following function:

$$\eta(\Omega) = \sum_{m,n=1}^{M}\sum_{k=1}^{K}\left\|\boldsymbol{Q}^{(m)}\boldsymbol{\mathcal{T}}_{(:,:,k)}^{(m,n)}\boldsymbol{Z}^{(n)} - \boldsymbol{\Gamma}_k^{(m,n)}\right\|_F^2. \quad (8)$$

Minimizing (8) is equivalent to minimizing the following:

$$\xi\left(\boldsymbol{Q}^{(1)},...,\boldsymbol{Q}^{(M)},\boldsymbol{Z}^{(1)},...,\boldsymbol{Z}^{(M)}\right) \triangleq \sum_{m,n=1}^{M}\sum_{k=1}^{K}\left\|\boldsymbol{Q}^{(m)}\boldsymbol{\mathcal{T}}_{(:,:,k)}^{(m,n)}\boldsymbol{Z}^{(n)}\right\|_{SBLF}^2, (9)$$

where $\|\cdot\|_{SBLF}$ denotes the Frobenius norm of the strictly block lower-triangular part. Note that (9) involves only unitary factor matrices, which are optimally conditioned. This has been achieved by relaxing the block-diagonal structure of $\boldsymbol{\Sigma}_k^{(m,n)}$ in (6) to the block upper-triangular structure of $\boldsymbol{\Gamma}_k^{(m,n)}$ in (8).

When the coupled block SGSD is solved (algorithms will be presented later in subsections 3.2 and 3.3), we obtain the block upper-triangular matrices $\boldsymbol{\Gamma}_k^{(m,n)}$ .We recall that $\boldsymbol{\Gamma}_k^{(m,n)} \triangleq \boldsymbol{R}^{(m)}\boldsymbol{\Sigma}_k^{(m,n)}\boldsymbol{L}^{(n)H}$ where $\boldsymbol{R}^{(m)}, \boldsymbol{L}^{(n)}$ are given in (7), and we have:

$$\boldsymbol{\Gamma}_{k,(r,r)}^{(m,n)} = \bar{\boldsymbol{R}}_{(r,r)}^{(m)}(\boldsymbol{\mathcal{C}}_r^{(m,n)})_{:,:,k}\bar{\boldsymbol{L}}_{(r,r)}^{(n)\ H}, \quad (10)$$

where $\boldsymbol{\Gamma}_{k,(r,r)}^{(m,n)} \in \mathbb{C}^{L_m \times L_n}$, $\bar{\boldsymbol{R}}_{(r,r)}^{(m)} \in \mathbb{C}^{L_m \times L_m}$, $\bar{\boldsymbol{L}}_{(r,r)}^{(n)} \in \mathbb{C}^{L_n \times L_n}$, and $(\boldsymbol{\mathcal{C}}_r^{(m,n)})_{:,:,k}$ are the rth block on the main diagonal of $\boldsymbol{\Gamma}_k^{(m,n)}$, $\bar{\boldsymbol{R}}^{(m)}$, $\bar{\boldsymbol{L}}^{(n)}$, and $\boldsymbol{\Sigma}_k^{(m,n)}$, respectively. Note that $\bar{\boldsymbol{R}}_{(r,r)}^{(m)}$ and $\bar{\boldsymbol{L}}^{(n)}$ are invertible, and (10) implies that $\boldsymbol{\Gamma}_{k,(r,r)}^{(m,n)}$ is an estimate of $(\boldsymbol{\mathcal{C}}_r^{(m,n)})_{:,:,k}$ up to trivial indeterminacies. We arrange $\boldsymbol{\Gamma}_{k,(r,r)}^{(m,n)}$ into a matrix denoted as $\boldsymbol{C}_{est}^{(m,n)} \in \mathbb{C}^{RL_mL_n \times K}$ in the same way that $\boldsymbol{C}^{(m,n)}$ in (2) is obtained from $(\boldsymbol{\mathcal{C}}_r^{(m,n)})_{:,:,k}$. According to (10) we have the following:

$$\boldsymbol{C}_{est}^{(m,n)} = \begin{bmatrix} \bar{\boldsymbol{R}}_{(1,1)}^{(m)} \otimes \bar{\boldsymbol{L}}_{(1,1)}^{(n)\ *} & & \\ & \ddots & \\ & & \bar{\boldsymbol{R}}_{(R,R)}^{(m)} \otimes \bar{\boldsymbol{L}}_{(R,R)}^{(n)\ *} \end{bmatrix}\boldsymbol{C}^{(m,n)}. \quad (11)$$

By (2) and (11), and recall that $\boldsymbol{C}^{(m,n)}$ is assumed to have full column rank, we have the following:

$$\boldsymbol{V}^{(m,n)} \triangleq \boldsymbol{T}_{(1,2),3}^{(m,n)}\boldsymbol{C}_{est}^{(m,n)\dagger} = [\boldsymbol{A}_1'^{(m)} \otimes \boldsymbol{B}_1'^{(n)*}, \cdots, \boldsymbol{A}_R'^{(m)} \otimes \boldsymbol{B}_R'^{(n)*}], (12)$$

where $\boldsymbol{A}_r'^{(m)} \triangleq \boldsymbol{A}_r^{(m)}\bar{\boldsymbol{R}}_{(r,r)}^{(m)-1}$ and $\boldsymbol{B}_r'^{(n)} \triangleq \boldsymbol{A}_r^{(n)}\bar{\boldsymbol{L}}_{(r,r)}^{(n)-1}$ can be taken as estimates of $\boldsymbol{A}_r^{(m)}$ and $\boldsymbol{A}_r^{(n)}$ up to trivial indeterminacies. Now the problem is how to calculate $\boldsymbol{A}_r'^{(m)}$ and $\boldsymbol{B}_r'^{(n)}$ from $\boldsymbol{V}^{(m,n)}$. We partition $\boldsymbol{V}^{(m,n)}$ as $[\boldsymbol{V}_1^{(m,n)},...,\boldsymbol{V}_R^{(m,n)}]$ where $\boldsymbol{V}_r^{(m,n)} = \boldsymbol{A}_r'^{(m)} \otimes \boldsymbol{B}_r'^{(n)*} \in \mathbb{C}^{I_mI_n \times L_mL_n}$, and reshape $\boldsymbol{V}_r^{(m,n)}$ into a matrix $\boldsymbol{V}_r'^{(m,n)} \in \mathbb{C}^{I_mL_m \times I_nL_n}$ as follows:

$$[\boldsymbol{V}_r'^{(m,n)}]_{(i-1)L_m+j,(k-1)L_n+l} \triangleq [\boldsymbol{V}_r^{(m,n)}]_{(j-1)I_n+l,(i-1)L_n+k}. \quad (13)$$

By definition we have $\boldsymbol{V}_r'^{(m,n)} = \mathrm{vec}(\boldsymbol{A}_r'^{(m)})[\mathrm{vec}(\boldsymbol{B}_r'^{(n)})]^H$. Moreover, we stack the matrices $\boldsymbol{V}_r'^{(m,n)}$ for all the values of m and n into a larger rank-1 matrix $\boldsymbol{M}_r$ as follows:

$$\boldsymbol{M}_r = \begin{bmatrix} \boldsymbol{V}_r'^{(1,1)} & \cdots & \boldsymbol{V}_r'^{(1,M)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{V}_r'^{(M,1)} & \cdots & \boldsymbol{V}_r'^{(M,M)} \end{bmatrix} = \boldsymbol{v}_r\boldsymbol{v}_r'^H \in \mathbb{C}^{\sum_{m=1}^{M}I_mL_m \times \sum_{n=1}^{M}I_nL_n}, (14)$$

where $\boldsymbol{v}_r \triangleq [\mathrm{vec}(\boldsymbol{A}_r'^{(1)})^H, \cdots, \mathrm{vec}(\boldsymbol{A}_r'^{(M)})^H]^H$, and $\boldsymbol{v}_r' \triangleq [\mathrm{vec}(\boldsymbol{B}_r'^{(1)})^H, \cdots, \mathrm{vec}(\boldsymbol{B}_r'^{(M)})^H]^H$.

We can estimate $\boldsymbol{v}_r$ and $\boldsymbol{v}_r'$ as the left and right dominant singular vector of $\boldsymbol{M}_r$, respectively, from which we can calculate $\boldsymbol{A}_r'^{(m)}$ and $\boldsymbol{B}_r'^{(m)}$ for fixed r and different m.

Two algorithms for the coupled block SGSD are given in the next subsections.

3.2. Extended QZ Iteration

Here we introduce a block variant of the algorithm in [22], which alternates between updates of $\boldsymbol{Q}^{(m)}$ and $\boldsymbol{Z}^{(n)}$ to optimize the cost function in (9). In each iteration, we update $\boldsymbol{Q}^{(m)}$ as $\boldsymbol{Q}^{(m)} \leftarrow \tilde{\boldsymbol{Q}}^{(m)}\boldsymbol{Q}^{(m)}$ with $\boldsymbol{Z}^{(n)}$ fixed (vice-versa for the update of $\boldsymbol{Z}^{(n)}$). Here $\tilde{\boldsymbol{Q}}^{(m)} \in \mathbb{C}^{I_m \times I_m}$ is a unitary matrix constructed as the product of R unitary matrices:

$$\tilde{\boldsymbol{Q}}^{(m)} = \begin{bmatrix} \boldsymbol{I}_{(R-1)L_m} & \\ & \boldsymbol{H}_R^{(m)} \end{bmatrix} \cdots \begin{bmatrix} \boldsymbol{I}_{L_m} & \\ & \boldsymbol{H}_2^{(m)} \end{bmatrix}\boldsymbol{H}_1^{(m)}, \quad (15)$$

where $\tilde{\boldsymbol{H}}_r^{(m)} \in \mathbb{C}^{(I_m-(r-1)L_m)\times(I_m-(r-1)L_m)}$. We calculate $\boldsymbol{H}_1^{(m)}$ such that it minimizes the norm of the matrix formed by stacking the parts strictly below the block diagonal of the first L_n columns of $\boldsymbol{\mathcal{T}}_{(:,:,k)}^{(m,n)}$ for all values of n and k. More precisely, we concatenate $\boldsymbol{\mathcal{T}}_{(:,1:L_n,k)}^{(m,n)}$ (i.e. the submatrix of $\boldsymbol{\mathcal{T}}_{(:,:,k)}^{(m,n)}$ consisting of its first L_n columns for fixed m) for varying n and k as:

$$\boldsymbol{T}_{sub,1}^{(m)} \triangleq [\boldsymbol{\mathcal{T}}_{(:,1:L_1,1)}^{(m,1)}, \cdots, \boldsymbol{\mathcal{T}}_{(:,1:L_1,K)}^{(m,1)}, \cdots, \boldsymbol{\mathcal{T}}_{(:,1:L_M,1)}^{(m,M)}, \cdots, \boldsymbol{\mathcal{T}}_{(:,1:L_M,K)}^{(m,M)}]. \quad (16)$$

A reliable choice for minimizing $\|(\boldsymbol{T}_{sub,1}^{(m)})_{L_m+1:I_m,:}\|_F$ is to take $\boldsymbol{H}_1^{(m)H}$ equal to the matrix of the left singular vectors of of $\boldsymbol{T}_{sub,1}^{(m)}$. When $\boldsymbol{H}_1^{(m)}$ is computed and applied to $\boldsymbol{\mathcal{T}}_{(:,:,k)}^{(m,n)}$ through (8) and (15), we obtain a set of new tensors denoted as $\boldsymbol{\mathcal{T}}'^{(m,n)}$ for fixed m and varying n. Then $\boldsymbol{H}_2^{(m)}$ is designed to minimize the strictly below block-diagonal norms of the second L_n columns of $\mathcal{T}'^{(m,n)}_{(:,:,k)}$ without affecting the first L_n columns. This can be done by looking at a reduced problem where the same reasoning as in the calculation of $\mathbf{H}_1^{(m)}$ is followed for $\mathcal{T}'^{(m,n)}_{(:,:,k)}$ with the first L_n columns removed. After $\mathbf{H}_2^{(m)}$ is computed, the matrices $\mathbf{H}_3^{(m)},\ldots,\mathbf{H}_R^{(m)}$ follow sequentially. The update of $\mathbf{Z}^{(n)}$ is similar to $\mathbf{Q}^{(m)}$.

3.3. Jacobi Iteration

Here we alternate between updates of $\mathbf{Q}^{(m)}$ and $\mathbf{Z}^{(n)}$, each computed as the product of a sequence of elementary Givens matrices. For the update of $\mathbf{Q}^{(m)}$, in each step the matrices $\mathbf{Q}^{(m)}$ and $\mathbf{\Gamma}_k^{(m,n)}$ are updated by a Givens matrix $\mathbf{G}_{u,v}^{(m)} \in \mathbb{C}^{I_m \times I_m}$ as $\mathbf{Q}^{(m)} \leftarrow \mathbf{G}_{u,v}^{(m)H}\mathbf{Q}^{(m)}, \mathbf{\Gamma}_k^{(m,n)} \leftarrow \mathbf{G}_{u,v}^{(m)H}\mathbf{\Gamma}_k^{(m,n)}$, $1 \le u < v \le I_m$, where $\mathbf{G}_{u,v}^{(m)}$ is equal to the identity matrix except $(\mathbf{G}_{u,v}^{(m)})_{u,u} = c_{u,v}^{(m)}$, $(\mathbf{G}_{u,v}^{(m)})_{u,v} = -s_{u,v}^{(m)*}$, $(\mathbf{G}_{u,v}^{(m)})_{v,u} = s_{u,v}^{(m)}$, $(\mathbf{G}_{u,v}^{(m)})_{v,v} = c_{u,v}^{(m)}$, with $c_{u,v}^{(m)} \triangleq \cos\theta_{u,v}^{(m)}$ and $s_{u,v}^{(m)} \triangleq e^{i\alpha_{u,v}^{(m)}}\sin\theta_{u,v}^{(m)}$.

Denoting $\xi_m \triangleq \sum_{n=1}^{M}\sum_{k=1}^{K}\| \mathbf{G}_{u,v}^{(m)H}\mathbf{\Gamma}_k^{(m,n)} \|_{SLBF}^2$, the cost function (9) can be written as $\xi = \sum_m \xi_m$. Then for fixed m, an iteration step (u,v) consists of finding $c_{u,v}^{(m)}$ and $s_{u,v}^{(m)}$ that minimize ξ_m. Note that $\mathbf{G}_{u,v}^{(m)}$ only affects the uth and vth rows of $\mathbf{\Gamma}_k^{(m,n)}$, and that the minimization of ξ_m amounts to minimizing the strictly lower-block-triangular parts of the uth and rth columns of $\mathbf{G}_{u,v}^{(m)H}\mathbf{\Gamma}_k^{(m,n)}$ for all values of n and k.

With some technical derivations we obtain the following:

$$\xi_m = \mathbf{w}_{u,v}^{(m)T}\mathbf{M}_{u,v}^{(m)}\mathbf{w}_{u,v}^{(m)}, \tag{17}$$

where $\mathbf{M}_{u,v}^{(m)} \triangleq \sum_{n=1}^{M}\sum_{k=1}^{K}(\mathbf{M}_{k,u,v}^{(m,n)}\mathbf{M}_{k,u,v}^{(m,n)H} + \mathbf{M}'^{(m,n)}_{k,u,v}\mathbf{M}'^{(m,n)H}_{k,u,v})$, and $\mathbf{w}_{u,v}^{(m)} \triangleq [c_{u,v}^{(m)}, \mathrm{Re}(s_{u,v}^{(m)}), \mathrm{Im}(s_{u,v}^{(m)})]^T \in \mathbb{R}^3$, and $\mathbf{M}_{k,u,v}^{(m,n)} \in \mathbb{C}^{3 \times r_u L_n}$, $\mathbf{M}'^{(m,n)}_{k,u,v} \in \mathbb{C}^{3 \times r_v L_n}$ are defined as:

$$\mathbf{M}_{k,u,v}^{(m,n)} \triangleq \begin{bmatrix} (\mathbf{\Gamma}_k^{(m,n)})_{u,1:r_u L_n} \\ (\mathbf{\Gamma}_k^{(m,n)})_{v,1:r_u L_n} \\ -i(\mathbf{\Gamma}_k^{(m,n)})_{v,1:r_u L_n} \end{bmatrix}, \quad \mathbf{M}'^{(m,n)}_{k,u,v} \triangleq \begin{bmatrix} (\mathbf{\Gamma}_k^{(m,n)})_{v,1:r_v L_n} \\ -(\mathbf{\Gamma}_k^{(m,n)})_{u,1:r_v L_n} \\ -i(\mathbf{\Gamma}_k^{(m,n)})_{u,1:r_v L_n} \end{bmatrix}. \tag{18}$$

with $r_u = \lfloor (u-1)/L_m \rfloor$, $r_v = \lfloor (v-1)/L_m \rfloor$. Then $\mathbf{w}_{u,v}^{(m)}$ is taken equal to the least significant eigenvector of $\mathbf{M}_{u,v}^{(m)}$. In each iteration for the update of $\mathbf{Q}^{(m)}$, we find the optimal rotation angles $c_{u,v}^{(m)}, s_{u,v}^{(m)}$ from an EVD. The update of $\mathbf{Z}^{(m)}$ can be addressed similarly, and is not further discussed.

4. NUMERICAL RESULTS

The target tensors are constructed as:

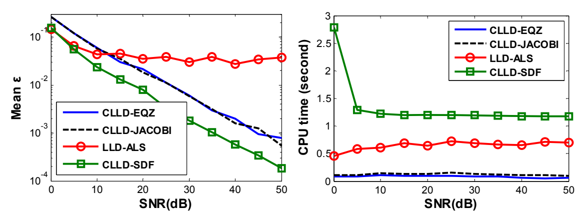
$$\mathcal{P}^{(m,n)} = \sigma_s \mathcal{T}^{(m,n)} / \| \mathcal{T}^{(m,n)} \|_F + \sigma_n \mathcal{N}^{(m,n)} / \| \mathcal{N}^{(m,n)} \|_F, \tag{19}$$

where $\mathcal{T}^{(m,n)}$ is generated by (1) with the entries of $\mathbf{A}_r^{(m)} \in \mathbb{C}^{I_m \times L_m}$, $\mathbf{C}_r^{(m,n)} \in \mathbb{C}^{L_m \times L_n}$ and $\mathcal{N}^{(m,n)} \in \mathbb{C}^{I_m \times I_n \times K}$ randomly drawn from complex normal distributions, m, n = 1,…,M. We set $M=2$, $K=200$, $R=2$, $L_1 = L_2 = L_3 = L = 2$, $I_1 = I_2 = I_3 = I = 6$. The signal-to-noise ratio (SNR) is defined with the signal level σ_s and noise level σ_n as $SNR = 20\log_{10}(\sigma_s/\sigma_n)$.

The proposed coupled rank-(L_m, L_n, ·) BTD algorithms based on the extended QZ and Jacobi iteration are denoted as CLLD-EQZ and CLLD-Jacobi, respectively. For comparison, we implement coupled rank-(L_m, L_n, ·) BTD with structured data fusion (CLLD-SDF) [21]. We also include in the comparison the computation of rank-(L_m, L_n, ·) BTD by alternating least squares (LLD-ALS, [2]) for each tensor separately. For CLLD-EQZ and CLLD-Jacobi we initialize with identity matrices. For CLLD-SDF, we initialize with the results from CLLD-EQZ. For LLD-ALS, we initialize with randomly generated factor matrices. With the obtained estimates we can reconstruct a set of tensors by (1). The average relative fitting error ε used to evaluate the performance is defined as:

$$\varepsilon = \sum_{m,n=1}^{M}\left(\left\| \mathcal{P}^{(m,n)} - \tilde{\mathcal{P}}^{(m,n)} \right\|_F^2 \Big/ \left\| \mathcal{P}^{(m,n)} \right\|_F^2 \right) \Big/ M^2, \tag{20}$$

where $\tilde{\mathcal{P}}^{(m,n)}$ is the reconstructed tensor. For CLLD-EQZ, CLLD-Jacobi and LLD-ALS, we terminate the iteration when $|\varepsilon_{cur} - \varepsilon_{prev}| / \varepsilon_{prev} \le 10^{-8}$, where ε_{cur} and ε_{prev} denote the relative fitting errors in the current and previous iteration, respectively. For CLLD-SDF, we set the tolerance parameters TolFun and TolX in the 'SDF_NLS' function of Tensorlab [26] to 0.001 and 0.03, respectively. For each SNR value, we perform 200 independent runs of all the algorithms. The results of mean ε and CPU time versus SNR are drawn in Fig. 3. We can see that the proposed algorithms provide more accurate estimates and higher computational efficiency than LLD-ALS for moderate to high SNR. CLLD-SDF gives the most accurate results, at a higher computational cost. This illustrates the interest of taking the coupling into account. It also shows that CLLD-EQZ and CLLD-Jacobi (i) provide good estimates for sufficiently high SNR and (ii) may be used as a low-cost initialization of the more expensive CLLD-SDF.



(a) Mean ε (b) CPU time

Fig. 3. Comparison of CLLD-EQZ, CLLD-JACOBI, CLLD-SDF, and LLD-ALS for SNR varying from 0dB to 50dB.

5. CONCLUSION

We have proposed two algorithms for the computation of a new coupled rank-(L_m, L_n, ·) BTD problem. The proposed algorithms are based on a coupled block version of the SGSD scheme, and can be used for the particular multi-set data fusion problem of J-ISA. Numerical results have shown that the proposed algorithms have fast computation and good accuracy, which makes them useful tools as such. When high accuracy is desired, they may be used to initialize algorithms that minimize the block-diagonal criterion (6) instead of the block-triangular criterion (8).

6. REFERENCES

[1] L. De Lathauwer, "Decompositions of a higher-order tensor in block-terms—Part II: Definitions and uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1033–1066, Sep. 2008.

[2] L. De Lathauwer, D. Nion, "Decompositions of a higher-order tensor in block-terms—Part III: Alternating least squares algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1067–1083, Sep. 2008.

[3] L. De Lathauwer, "Blind separation of exponential polynomials and the decomposition of a tensor in rank-(L_r,L_r,1) terms," *SIAM J. Matrix Anal. Appl.*, vol. 32, no. 4, pp. 1451–1474, Dec. 2011.

[4] L. Sorber, M. Van Barel, L. De Lathauwer, "Optimization-based algorithms for tensor decompositions: canonical polyadic decomposition, decomposition in rank-(L_r,L_r,1) terms and a new generalization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 695–720, Apr. 2013.

[5] L. De Lathauwer, A. de Baynast, "Blind deconvolution of DS-CDMA signals by means of decomposition in rank-(1,L,L) terms," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1562–1571, Apr. 2008.

[6] D. Nion, L. De Lathauwer, "A block component model-based blind DS-CDMA receiver," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5567–5579, Nov. 2008.

[7] L. De Lathauwer, "Block component analysis: a new concept for blind source separation," in *Proc. LVA/ICA'2012*, Tel-Aviv, Israel, Mar. 12–15, 2012, pp. 1–8.

[8] B. Hunyadi, D. Camps, L. Sorber, W. Van Paesschen, M. De Vos, S. Van Huffel, L. De Lathauwer, "Block term decomposition for modelling epileptic seizures," *EURASIP J. Adv. Signal Process*. 2014, 2014:139.

[9] O. Debals, M. Van Barel, L. De Lathauwer, "Löwner-based blind signal separation of rational functions with applications," Accepted for publication in *IEEE Trans. Signal Process.*

[10] M. Boussé, O. Debals, L. De Lathauwer, "A tensor-based method for large-scale blind source separation using segmentation," Tech. Rep. 15–59, ESAT-STADIUS, KU Leuven, Belgium, 2015.

[11] Y.-N. Yu, A. Petropulu, "PARAFAC-based blind estimation of possibly underdetermined convolutive MIMO systems," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 111–124, Jan. 2008.

[12] H. Becker, P. Comon, L. Albera, "Tensor-based preprocessing of combined EEG/MEG data," in *Proc. EUSIPCO'2012*, Bucharest, Romania, Aug. 27–31, 2012, pp. 275–279.

[13] X.-F. Gong, Y.-N. Hao, Q.-H. Lin, "Joint canonical polyadic decomposition of two tensors with one shared loading matrix," in *Proc. MLSP'2013*, Southampton, U.K., Sep. 22–25, 2013.

[14] T. Yokota, A. Cichocki, Y. Yamashita, "Linked PARAFAC / CP tensor decomposition and its fast implementation for multi-block tensor analysis," in *Proc. ICONIP'2012*, Doha, Qatar, Nov. 12–15, 2012, pp. 84–92.

[15] M. Sorensen, L. De Lathauwer, "Coupled canonical polyadic decompositions and (coupled) decompositions in multilinear rank-($L_{r,n}$, $L_{r,n}$, 1) terms—Part I: uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 36, no. 2, pp. 496–522, Apr. 2015.

[16] M. Sorensen, I. Domanov, D. Nion and L. De Lathauwer, "Coupled canonical polyadic decompositions and (coupled) decompositions in multilinear rank-($L_{r,n}$, $L_{r,n}$, 1) terms—Part II: algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 36, no. 3, pp. 1015–1045, Jul. 2015.

[17] M. Sorensen, L. De Lathauwer, "Multidimensional harmonic retrieval via coupled canonical polyadic decompositions," Tech. Rep.13-240, ESAT-STADIUS, KU Leuven, Belgium, 2013.

[18] M. Sorensen, L. De Lathauwer, "Coupled tensor decompositions in array processing," Tech. Rep. 13-241, ESAT-STADIUS, KU Leuven, Belgium, 2014.

[19] M. Sorensen, L. De Lathauwer, "Multiple invariance ESPRIT: A coupled matrix-tensor factorization approach," Tech. Rep. 13-242, ESAT-STADIUS, KU Leuven, Belgium, 2014.

[20] D. Lahat, T. Adali, C. Jutten, "Multimodal data fusion: an overview of methods, challenges, and prospects," *Proc. IEEE*, vol. 103, no.9, pp. 1449–1477, Sep. 2015.

[21] L. Sorber, M. Van Barel, L. De Lathauwer, "Structured data fusion," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 4, pp. 586–600, Jun. 2015.

[22] A.-J. van der Veen, A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Process.*, vol. 44, no. 5, pp. 1136–1155, May. 1996.

[23] L. De Lathauwer, B. De Moor, J. Vandewalle, "Computation of the canonical decomposition by means of a simultaneous generalized Schur decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 26, no. 2, pp. 295–327, Nov. 2004.

[24] A. Stegeman. "Using the simultaneous generalized Schur decomposition as a Candecomp/Parafac algorithm for ill-conditioned data," *J. Chemom., Special issue in memory of Richard Harshman*, vol. 23, no. 7-8, pp. 385-392, Jul. 2009.

[25] D. Lahat, C. Jutten, "Joint independent subspace analysis: a quasi-Newton algorithm," in *Proc. LVA/ICA'2015*, Liberec, Czech Republic, Aug. 25–28, 2015, pp. 111–118.

[26] L. Sorber, M. Van Barel, and L. De Lathauwer. *Tensorlab v2.0*, Available online, Jan. 2014. URL: http://www.tensorlab.net/.