

Enhanced Computation of the Coupled Block-Term Decomposition in Multilinear Rank Terms

Ildar Safiullin¹, Liana Khamidullina¹, Alexey A. Korobkov², and Martin Haardt¹

¹Communication Research Laboratory, Ilmenau University of Technology, Ilmenau, Germany

²Kazan National Research Technical University n.a. A.N. Tupolev-KAI, Russia

e-mail: ildar.safiullin@tu-ilmenau.de, liana.khamidullina@tu-ilmenau.de, korobkov@inbox.ru, martin.haardt@tu-ilmenau.de

Abstract—Coupled tensor decompositions proved to be a valuable tool for many signal processing applications, including biomedical data analysis, source separation, data fusion, and many others. In this paper, we present an algorithm to calculate the coupled (BTD) of multiple three-dimensional tensors with a coupled mode and possibly an unknown rank structure. The proposed approach is composed of three main parts, the first is an extension of the linear regression of global eigenvalues (LaRGE) technique to estimate the number of blocks and sum of multilinear ranks in the coupled tensors. The second part accounts for the calculation of the preliminary multilinear factors of the coupled BTD by means of the semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalizations (SECSI). The last part contains the final refinement procedures that return the estimated BTD factors. Moreover, we provide some synthetic data simulations showing that the proposed algorithm demonstrates more stable and accurate results than the schemes from the literature.

Index Terms—Coupled factorization, block-term decomposition, data fusion, tensor decomposition.

I. INTRODUCTION

The apparent popularity of tensors in all possible signal processing areas is an eloquent indication that they are evidently an effective and promising tool for many applications, from statistics and communications to linguistics and big data analysis. Apart from their mild uniqueness conditions and data structure preserving qualities, tensor decompositions provide effective denoising and feature extraction capabilities [1].

The focus of this paper falls on the block-term decompositions (BTD), to be more specific, on the coupled decomposition in rank- $(L_r, L_r, 1)$ terms. Although, it might be not yet as widely used as, for example, the Higher-Order Singular Value Decomposition (HOSVD) [2], [3] or the Canonical Polyadic Decomposition (CPD) (sometimes also referred to as CANDCOMP or PARAFAC) [4], [5], the BTDs prove to be particularly interesting for applications where the data have a heterogeneous rank structure and possess collinearity in some of the modes. There are some variations in the definitions of the block-term decomposition based on the ranks of the block-terms. These definitions, uniqueness conditions, and the algebraic and optimization-based algorithms are well described

The authors gratefully acknowledge the support of the German Research Foundation (DFG) under grant no. HA 2239/14-1 (AdAMMM) and grant no. ZH 640/2-1 (AdAMMM).

in [6]–[11]. From the applications point of view, the rank- $(L_r, L_r, 1)$ block-term decomposition appears to be the most attractive one. It has found its applications in communications, biomedical data processing, image and graph analysis, and many others [12]–[22]. Commonly, when it comes to applications, the question that arises before applying any low-rank approximation concerns the rank of the tensor. Given that it is rarely known beforehand, one should find a way to define it. Some techniques to estimate the factors and the ranks of a single block-term tensor are introduced in [23]–[26].

Another direction, in which the BTDs are evolving, is the computation of coupled decompositions and data fusion. Many applications that have multiple related data sets to be analyzed can genuinely benefit from coupled factorizations. They often provide more flexible uniqueness conditions and improved feature extraction capabilities, and thus deeper insights into the data. The studies on the uniqueness and the algorithms for coupled decompositions, including the coupled BTD, can be found in [27]–[30].

In this paper, we propose a new approach to estimate the factors of the coupled block-term decomposition of multiple tensors with a common mode based on the coupled extension of the semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalizations (SECSI) [31], [32]. Moreover, we show how the multilinear rank structure of the coupled rank- $(L_r, L_r, 1)$ decomposition can be estimated based on the extension of the linear regression of global eigenvalues (LaRGE) and clustering techniques. The simulation results show that compared to the algorithm in [30], [33], which requires a known rank structure to perform the decomposition, the proposed algorithm is more accurate, numerically stable, and can perform the decomposition even if the ranks are not available beforehand.

Notation. Matrices and vectors are denoted by upper-case (\mathbf{A}) and lower-case (\mathbf{a}) bold-faced letters, respectively. Bold-faced calligraphic letters denote tensors (\mathcal{A}). The superscripts $\{\cdot\}^T$ and $\{\cdot\}^H$ denote the transpose and Hermitian transpose, respectively. The Khatri-Rao and outer products are denoted as \diamond and \circ , respectively. Additionally, we denote the higher-order norm of a tensor \mathcal{A} by $\|\mathcal{A}\|_H$, and the Frobenius norm of a matrix \mathbf{A} by $\|\mathbf{A}\|_F$. The n -mode unfolding of the tensor \mathcal{A} is denoted as $[\mathcal{A}]_{(n)}$ (we use the reverse cyclical ordering of the columns [2]). The n -mode product between a tensor \mathcal{A} and a

matrix B is written as $\mathcal{A} \times_n B$. $\mathcal{I}_{d,R}$ denotes a super-diagonal d -dimensional tensor of size $R \times \dots \times R$ with elements equal to one if all d indices are equal and zero otherwise, and $\mathbf{1}_R$ is a column vector of ones of size R .

II. DATA MODEL

Let us consider a set of M noise-corrupted three-dimensional tensors $\mathcal{X}^{(m)} \in \mathbb{C}^{I_1^{(m)} \times I_2^{(m)} \times I_3}$ with the third mode in common. Then the coupled rank- $(L_r, L_r, 1)$ BTD of M tensors can be written as follows

$$\mathcal{X}^{(m)} = \mathcal{X}_0^{(m)} + \mathcal{N}^{(m)} = \sum_{r=1}^R \left(\mathbf{A}_r^{(m)} \cdot \mathbf{B}_r^{(m)\text{T}} \right) \circ \mathbf{c}_r + \mathcal{N}^{(m)} \quad (1)$$

$$= \mathcal{I}_{3, \Sigma L_r} \times_1 \mathbf{A}^{(m)} \times_2 \mathbf{B}^{(m)} \times_3 \mathbf{C}^o + \mathcal{N}^{(m)}, \quad (2)$$

where $m \in \{1, \dots, M\}$, and R is a number of block-terms with ranks L_r , $r \in \{1, \dots, R\}$ (we assume that the multilinear ranks $L_r^{(m)}$ are equal for all M tensors ($L_r^{(1)} = L_r^{(2)} = \dots = L_r^{(M)} = L_r$), which is usually the case for the coupled data sets¹). For simplicity of notation, ΣL_r refers to $\sum_{r=1}^R L_r$. Moreover, $\mathcal{X}_0^{(m)} \in \mathbb{C}^{I_1^{(m)} \times I_2^{(m)} \times I_3}$ is a noise-free BTD tensor, and $\mathcal{N}^{(m)} \in \mathbb{C}^{I_1^{(m)} \times I_2^{(m)} \times I_3}$ is an additive noise tensor. The factors $\mathbf{A}_r^{(m)} \in \mathbb{C}^{I_1^{(m)} \times L_r}$ can be stacked in a matrix $\mathbf{A}^{(m)}$ as $\mathbf{A}^{(m)} = [\mathbf{A}_1^{(m)}, \mathbf{A}_2^{(m)}, \dots, \mathbf{A}_R^{(m)}] \in \mathbb{C}^{I_1^{(m)} \times \Sigma L_r}$, and the individual factors $\mathbf{B}_r^{(m)} \in \mathbb{C}^{I_2^{(m)} \times L_r}$ are represented as a submatrices of $\mathbf{B}^{(m)} = [\mathbf{B}_1^{(m)}, \mathbf{B}_2^{(m)}, \dots, \mathbf{B}_R^{(m)}] \in \mathbb{C}^{I_2^{(m)} \times \Sigma L_r}$. The common vectors \mathbf{c}_r are stacked in the matrix $\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_R] \in \mathbb{C}^{I_3 \times R}$. Then, it can be seen from (2), that the BTD in (1) may be viewed as a CP decomposition with CP-rank ΣL_r and with repeated or linearly depended columns in \mathbf{C}^o

$$\mathbf{C}^o = [\mathbf{c}_1 \cdot \mathbf{1}_{L_1}^T, \mathbf{c}_2 \cdot \mathbf{1}_{L_2}^T, \dots, \mathbf{c}_R \cdot \mathbf{1}_{L_R}^T] \in \mathbb{C}^{I_3 \times \Sigma L_r}. \quad (3)$$

In this paper, we use this link between the two decompositions to first calculate the initial estimates in the coupled CPD form and subsequently convert them to coupled BTD factors.

III. RANK STRUCTURE ESTIMATION

Since in most of the tensor decomposition applications the rank structure is usually unknown, its correct estimation becomes essential in relation to low-rank tensor decompositions. While a number of algorithms for CPD rank estimation are available in the literature, there are not yet many options for block-term decompositions. Moreover, to the best of our knowledge, there are no techniques devoted to the estimation of ranks in coupled BTD models. Therefore, in this contribution, we present an extension of the LaRGE technique [34], [35] to estimate the rank structure of the coupled rank- $(L_r, L_r, 1)$ decomposition. The LaRGE algorithm estimates the CP-rank of a tensor by performing a linear regression of its noise eigenvalues, which commonly follow an exponential decline, i.e., a linear decline on a logarithmic scale. Since the signal

eigenvalues do not possess the same property, the linear approximation will fail when reaching signal eigenvalues, which will indicate the detection of the rank. In the following, we present the extension of LaRGE for coupled BTD rank estimation.

Given a set of M noise-corrupted three-dimensional tensors $\mathcal{X}^{(m)} \in \mathbb{C}^{I_1^{(m)} \times I_2^{(m)} \times I_3}$ as in (1), the coupled HOSVD of $\mathcal{X}^{(m)}$ can be written as

$$\mathcal{X}^{(m)} = \mathcal{S}^{(m)} \times_1 \mathbf{U}_1^{(m)} \times_2 \mathbf{U}_2^{(m)} \times_3 \mathbf{U}_3, \quad (4)$$

where $\mathcal{S}^{(m)} \in \mathbb{C}^{I_1^{(m)} \times I_2^{(m)} \times I_3}$ are the core tensors, $\mathbf{U}_1^{(m)} \in \mathbb{C}^{I_1^{(m)} \times I_1^{(m)}}$ and $\mathbf{U}_2^{(m)} \in \mathbb{C}^{I_2^{(m)} \times I_2^{(m)}}$ are composed of the left singular vectors of the 1- and 2-mode unfoldings of $\mathcal{X}^{(m)}$ computed from $[\mathcal{X}^{(m)}]_{(n)} = \mathbf{U}_n^{(m)} \mathbf{\Sigma}_n^{(m)} \mathbf{V}_n^{(m)\text{H}}$, where $\mathbf{\Sigma}_n^{(m)}$ contains the n -mode singular values $\sigma_{(n),i}^{(m)}$, $n \in \{1, 2\}$ on its main diagonal. $\mathbf{U}_3 \in \mathbb{C}^{I_3 \times I_3}$ is the common unitary matrix calculated from the SVD of the concatenated 3-mode unfoldings of $\mathcal{X}^{(m)}$ [32], and therefore, $\sigma_{(3),i}^{(m)}$ is common for all tensors. For model order estimation, we are interested in the first $J = \min\{I_3, I_n^{(m)}\}$ singular values.

Due to the asymmetric rank structure, we divide the rank estimation processes into two separate procedures. In the first run, only the semi-global eigenvalues $\tilde{\lambda}_{(1,2),i}^{(m)}$, $i = \{1, \dots, J\}$

$$\tilde{\lambda}_{(1,2),i}^{(m)} = \left(\sigma_{(1),i}^{(m)} \right)^2 \cdot \left(\sigma_{(2),i}^{(m)} \right)^2 \quad (5)$$

from the 1- and 2-mode are used to estimate ΣL_r . To take into account the information from all coupled tensors, we compute the geometric mean of M semi-global eigenvalues in (5) as follows

$$\tilde{\lambda}_{(1,2),i} = \sqrt[M]{\tilde{\lambda}_{(1,2),i}^{(1)} \cdot \tilde{\lambda}_{(1,2),i}^{(2)} \cdot \dots \cdot \tilde{\lambda}_{(1,2),i}^{(M)}}. \quad (6)$$

Since there are rank deficiencies in the 3-mode, we consider it separately to estimate the number of block-terms R using the 3-mode eigenvalues $\tilde{\lambda}_{(3),i}^{(m)} = \left(\sigma_{(3),i}^{(m)} \right)^2$ averaged along all M tensors as

$$\tilde{\lambda}_{(3),i} = \sqrt[M]{\tilde{\lambda}_{(3),i}^{(1)} \cdot \tilde{\lambda}_{(3),i}^{(2)} \cdot \dots \cdot \tilde{\lambda}_{(3),i}^{(M)}}. \quad (7)$$

Then, the linear regression is applied separately to $\lambda_{(1,2),i} = \ln \left(\tilde{\lambda}_{(1,2),i} \right)$ and $\lambda_{(3),i} = \ln \left(\tilde{\lambda}_{(3),i} \right)$ to estimate ΣL_r and R , respectively. For notational convenience, in the following, we simply use λ_i which will refer either to $\lambda_{(1,2),i}$ or $\lambda_{(3),i}$, depending on which rank, ΣL_r or R , is to be estimated. The estimation starts from the smallest noise eigenvalue ($i = J$), and subsequently finds a linear prediction $\hat{\lambda}_i$ by moving up the scale. For each step k in the prediction, the relative prediction error is computed as

$$\delta_{J-k} = \frac{\lambda_{J-k} - \hat{\lambda}_{J-k}}{\left| \hat{\lambda}_{J-k} \right|} = \frac{\Delta_{J-k}}{\left| \hat{\lambda}_{J-k} \right|}, \quad (8)$$

followed by the standard deviation of the approximation error

$$\sigma_{J-k} = \sqrt{\frac{1}{k} \sum_{i=J}^{J-k} \left(\Delta_i - \frac{1}{k} \sum_{i=J}^{J-k} \Delta_i \right)^2}, \quad (9)$$

¹Ranks can be different across the block-terms r .

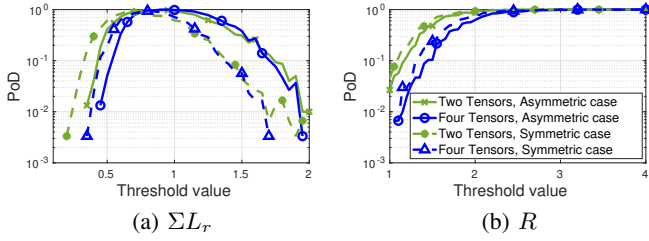


Fig. 1: PoD vs. Threshold. Simulation parameters: $\Sigma L_r = 9$, $R = 3$, $L_r = [2, 3, 4]$ (asymmetric), SNR = -11 dB; $\Sigma L_r = 20$, $R = 4$, $L_r = [5, 5, 5, 5]$ (symmetric), SNR = -5 dB. 300 Monte-Carlo trials.

where Δ_i is the absolute prediction error. Then, by computing the ratio between the relative prediction error and the standard deviation of the approximation errors in the previous steps as $\text{PESDR}_k = \frac{\Delta_{J-k}}{\sigma_{J-k-1}}$, the algorithm can detect the first signal eigenvalue (corresponding either to ΣL_r or R) if the Prediction Error to Standard Deviation Ratio (PESDR) exceeds the predefined threshold ρ for the first time. For a more detailed description of the LaRGE algorithm, we refer the reader to [34], [35]. Fig. 1 shows the probabilities of detection (PoD) with respect to a threshold ρ for the estimation of ΣL_r (Fig. 1(a)) and R (Fig. 1(b)), respectively. For this simulation, we generated two and four coupled BTD tensors $\mathcal{X}_0^{(1)} \in \mathbb{C}^{50 \times 60 \times 70}$, $\mathcal{X}_0^{(2)} \in \mathbb{C}^{52 \times 60 \times 70}$, $\mathcal{X}_0^{(3)} \in \mathbb{C}^{54 \times 60 \times 70}$, and $\mathcal{X}_0^{(4)} \in \mathbb{C}^{55 \times 60 \times 70}$ with the 3-mode in common, and added zero mean circularly symmetric complex Gaussian (ZMCSCG) noise. The SNR is -11 dB in the asymmetric case and -5 dB in the symmetric case. For higher SNRs, the PoD tends to one for all ρ s in the considered range. As it can be observed, the threshold value does not depend on the number of tensors and is approximately equal to 0.75 for ΣL_r and to 3.0 for R .

IV. CPD-BASED COMPUTATION OF THE COUPLED RANK- $(L_r, L_r, 1)$ BTD

We divide the proposed BTD algorithm into three main parts, an initial estimation step based on the simultaneous matrix diagonalizations (SECSI), clustering, and final refinements, which will be explained in the following.

Initial coupled BTD estimates via coupled CPD: In order to compute the initial coupled BTD estimates, we employ the SECSI framework for the coupled CPD. The coupled decomposition with rank ΣL_r is performed on the BTD tensors as in (2). SECSI utilizes the simultaneous matrix diagonalization (SMD) [36] approach that defines several heuristics to enhance the accuracy of the CP decomposition [31], [32]. The main steps of SECSI include the computation of the truncated HOSVD and a subsequent estimation of the transformation matrices that represent the connection between the CPD and truncated HOSVD. Assuming that the coupled HOSVD has been previously computed in the rank estimation step in (4), it can be directly truncated without calculating the full HOSVD again. Assuming the noiseless case, the coupled truncated HOSVD of M tensors $\mathcal{X}_0^{(m)}$ with the third mode in common

is written as

$$\mathcal{X}_0^{(m)} = \mathcal{S}^{[s(m)]} \times_1 U_1^{[s(m)]} \times_2 U_2^{[s(m)]} \times_3 U_3^{[s]}, \quad (10)$$

where $\mathcal{S}^{[s(m)]} \in \mathbb{C}^{\Sigma L_r \times \Sigma L_r \times \Sigma L_r}$ are the core tensors, and a set of truncated unitary matrices $U_1^{[s(m)]} \in \mathbb{C}^{I_1^{(m)} \times \Sigma L_r}$, $U_2^{[s(m)]} \in \mathbb{C}^{I_2^{(m)} \times \Sigma L_r}$, and $U_3^{[s]} \in \mathbb{C}^{I_3 \times \Sigma L_r}$ span the column space of the corresponding n -mode unfolding of $\mathcal{X}_0^{(m)}$. Given that these unitary matrices share the same column space as the CPD factors in (2), the CPD factor matrices can be expressed through three non-singular transform matrices $T_1^{(m)}$, $T_2^{(m)}$, and T_3 as follows

$$\begin{aligned} \mathbf{A}^{(m)} &= U_1^{[s(m)]} T_1^{(m)}, & \mathbf{B}^{(m)} &= U_2^{[s(m)]} T_2^{(m)}, \\ \mathbf{C}^o &= U_3^{[s]} T_3. \end{aligned} \quad (11)$$

The inverses of the transform matrices diagonalize the core tensors $\mathcal{S}^{[s(m)]}$. Thus, if we consider, for instance, a tensor $\mathcal{S}_3^{(m)}$ such that $\mathcal{S}_3^{(m)} = \mathcal{S}^{[s(m)]} \times_3 U_3^{[s]}$, and a tensor $\mathcal{D}_3 = \mathcal{I}_{3, \Sigma L_r} \times_3 \mathbf{C}^o$, then $\mathcal{S}_3^{(m)} = \mathcal{D}_3 \times_1 T_1^{(m)} \times_2 T_2^{(m)}$, where every i th 3-mode slice of \mathcal{D}_3 is a diagonal matrix $\mathbf{D}_{3,i}$ with the i th row of \mathbf{C}^o on its main diagonal. The previous equation can be rewritten as

$$\mathcal{S}_{3,i}^{(m)} = T_1^{(m)} \mathbf{D}_{3,i} T_2^{(m)T}, \quad (12)$$

where $\mathcal{S}_{3,i}^{(m)}$ is the i th 3-mode slice of $\mathcal{S}_3^{(m)}$, $i \in \{1, \dots, I_3\}$. By multiplying (12) from either the right-hand side (rhs) or the left-hand side (lhs) by an arbitrary pivot slice $(\mathcal{S}_{3,p}^{(m)})^{-1}$, $p \in \{1, \dots, I_3\}$, the asymmetric SMD problem in (12) can be transformed to a symmetric one

$$\begin{aligned} \mathcal{S}_{3,i}^{(m)\text{rhs}} &= \mathcal{S}_{3,i}^{(m)} \left(\mathcal{S}_{3,p}^{(m)} \right)^{-1} \\ &= T_1^{(m)} \left(\mathbf{D}_{3,i} \mathbf{D}_{3,p}^{-1} \right) \left(T_1^{(m)} \right)^{-1}, \end{aligned} \quad (13)$$

$$\begin{aligned} \mathcal{S}_{3,i}^{(m)\text{lhs}} &= \left(\left(\mathcal{S}_{3,p}^{(m)} \right)^{-1} \mathcal{S}_{3,i}^{(m)} \right)^T \\ &= T_2^{(m)} \left(\mathbf{D}_{3,i} \mathbf{D}_{3,p}^{-1} \right) \left(T_2^{(m)} \right)^{-1}. \end{aligned} \quad (14)$$

Consequently, the transform matrices $T_1^{(m)}$ and $T_2^{(m)}$ can be estimated by means of SMD [37]. The matrix \mathbf{C}^o can be then retrieved from the jointly diagonalized matrices $\mathbf{D}_{3,i}$ up to one scaling ambiguity per column, and for each SMD problem (rhs or lhs), the factors $\mathbf{A}^{(m)}$ and $\mathbf{B}^{(m)}$ are obtained either from (11) or from a Least Squares (LS) fit as $\mathbf{A}^{(m)} = [\mathcal{X}_0^{(m)}]_{(1)} \cdot (\mathbf{B}^{(m)} \diamond \mathbf{C}^o)^{T+}$ or $\mathbf{B}^{(m)} = [\mathcal{X}_0^{(m)}]_{(2)} \cdot (\mathbf{C}^o \diamond \mathbf{A}^{(m)})^{T+}$. Similar SMD problems can be constructed by multiplying the core tensors $\mathcal{S}^{[s(m)]}$ by $U_1^{[s(m)]}$ or $U_2^{[s(m)]}$ in the 1-mode or the 2-mode, respectively. This, in general, will lead in total to $2 + 4M$ SMD problems, and thus $4 + 2M$ estimates of the third factor matrix and six sets of estimates of non-coupled factor matrices [32]. Then the final best set of factor estimates can be chosen according to a predefined heuristic [31]. However, we have observed in practice, that in contrast to the CPD model data, the BTD constructed data always lead to the final estimates chosen from the 3-mode SMDs. Therefore, to estimate the BTD factors by means of SECSI, we only solve the SMDs in (12). Due to the page limitations, we refer the reader to [31], [32] for more details on the SECSI and coupled SECSI frameworks.

Clustering: The solution of SMDs in (13) or (14) leads to initial estimates $\hat{\mathbf{A}}^{(m)}$, $\hat{\mathbf{B}}^{(m)}$, and $\hat{\mathbf{C}}^{o(m)}$ in the CPD form. Therefore, in the next step, the matrix $\hat{\mathbf{C}}^{o(m)}$ with collinear columns has to be brought to the proper BTD form, which can be done by k -means clustering. By applying the clustering to $\hat{\mathbf{C}}^{o(m)}$, we retrieve permutation matrices $\mathbf{P}^{(m)}$ which assign each column in $\hat{\mathbf{C}}^{o(m)}$ to one of the R clusters. Additionally, the k -means finds the optimal number of elements in each cluster (L_r). Therefore, it can be also employed in cases when the multilinear ranks are not known. For the cases with a known rank structure, clustering with size constraints can be used [38]. By multiplying $\hat{\mathbf{A}}^{(m)}$, $\hat{\mathbf{B}}^{(m)}$, and $\hat{\mathbf{C}}^{o(m)}$ by the permutation matrices $\mathbf{P}^{(m)}$, we rearrange their columns in a way that the columns corresponding to r th block are grouped together and follow the correct order as in the description of (1). These permuted factors are denoted as $\tilde{\mathbf{A}}^{(m)}$, $\tilde{\mathbf{B}}^{(m)}$, and $\tilde{\mathbf{C}}^{o(m)}$.

Final refinements: To return the final BTB estimates of $\mathbf{A}_r^{(m)}$, $\mathbf{B}_r^{(m)}$, and \mathbf{c}_r , the collinear columns in $\tilde{\mathbf{C}}^{o(m)}$ have to be reduced. This can be performed by one ALS iteration as follows. First, recalculate $\tilde{\mathbf{A}}^{(m)}$ and $\tilde{\mathbf{B}}^{(m)}$ with a current $\tilde{\mathbf{C}}^{o(m)}$ as $\tilde{\mathbf{A}}^{(m)} = [\mathcal{X}^{(m)}]_{(1)} \cdot (\tilde{\mathbf{B}}^{(m)} \diamond \tilde{\mathbf{C}}^{o(m)})^{\text{T}+}$, and $\tilde{\mathbf{B}}^{(m)} = [\mathcal{X}^{(m)}]_{(2)} \cdot (\tilde{\mathbf{C}}^{o(m)} \diamond \tilde{\mathbf{A}}^{(m)})^{\text{T}+}$. Then the submatrices in $\tilde{\mathbf{C}}^{o(m)}$ are reduced to vectors as follows [8]

$$\tilde{\mathbf{C}}^{(m)} = [\mathcal{X}^{(m)}]_{(3)} \cdot \left[\left(\tilde{\mathbf{A}}_1^{(m)} \diamond \tilde{\mathbf{B}}_1^{(m)} \right) \mathbf{1}_{L_1}, \dots, \left(\tilde{\mathbf{A}}_R^{(m)} \diamond \tilde{\mathbf{B}}_R^{(m)} \right) \mathbf{1}_{L_R} \right]^{\text{T}+}. \quad (15)$$

Since there are M similar estimates of \mathbf{C} , we choose the one with the smallest reconstruction error, and use the obtained $\tilde{\mathbf{A}}^{(m)}$, $\tilde{\mathbf{B}}^{(m)}$, and $\tilde{\mathbf{C}}$ as an initialization of the Nonlinear Least Squares (NLS) algorithm in [30].

V. NUMERICAL RESULTS

In this section, we compare the performances of the proposed coupled BTB algorithm with the algorithm from the Tensorlab toolbox which is based on the Structured Data Fusion via Nonlinear Least Squares (SDF-NLS) in [30]. For the simulations, we have generated two coupled BTB tensors $\mathcal{X}_0^{(1)} \in \mathbb{C}^{50 \times 60 \times 70}$ and $\mathcal{X}_0^{(2)} \in \mathbb{C}^{54 \times 60 \times 70}$ according to the model in (1), and added the noise tensors $\mathcal{N}^{(m)}$ containing ZMCSCG entries with variance σ_n^2 . Accordingly, the SNR is defined as $1/\sigma_n^2$. The 3-mode is in common.

We use two metrics to compare the performances, squared reconstruction error (SRE) and squared factor reconstruction error (SFRE)

$$\text{SRE} = \frac{\|\hat{\mathcal{X}}_0 - \mathcal{X}_0\|_{\text{H}}^2}{\|\mathcal{X}_0\|_{\text{H}}^2}, \quad \text{SFRE} = \frac{\|\hat{\mathbf{F}} \cdot \mathbf{P}_c - \mathbf{F}\|_{\text{F}}^2}{\|\mathbf{F}\|_{\text{F}}^2}, \quad (16)$$

where $\hat{\mathcal{X}}_0$ is a tensor reconstructed from the estimated BTB factors, \mathbf{F} and $\hat{\mathbf{F}}$ denote the original and estimated factors \mathbf{C} or $(\mathbf{A}_r^{(m)} \mathbf{B}_r^{(m)\text{T}})$, respectively. The matrix \mathbf{P}_c resolves the permutation and scaling. The errors for $\mathbf{A}^{(m)}$ and $\mathbf{B}^{(m)}$ are shown together due to inherent matrix product ambiguities [7]. The complementary cumulative distribution functions (CCDFs)

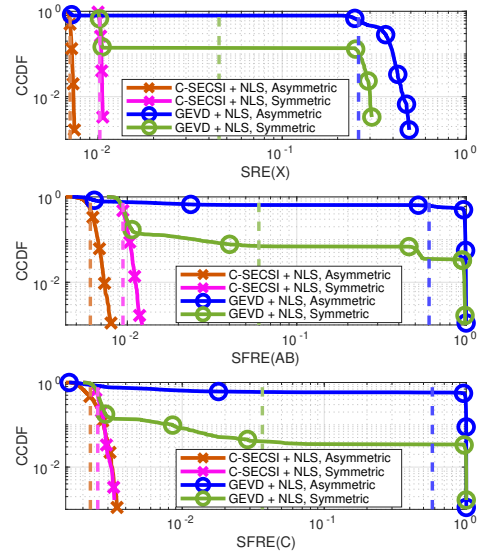


Fig. 2: CCDF vs. Reconstruction errors. SNR = 20 dB. Red and magenta lines with crosses denote the proposed algorithm, and blue and green lines with circles denote the Tensorlab [30], [33] algorithm. Simulation parameters: $R = 3$, $L_r = [2, 3, 4]$ (asymmetric); $R = 4$, $L_r = [5, 5, 5, 5]$ (symmetric). 300 Monte-Carlo trials.

of SRE and SFRE are presented in Fig. 2. The errors for $(\mathbf{A}_r^{(m)} \mathbf{B}_r^{(m)\text{T}})$ are averaged across all values of r . C-SECSI-NLS corresponds to the NLS initialized with the proposed coupled SECSI-based approach, and GEVD-NLS denotes the generalized eigenvalue decomposition based initialization. As it can be seen, the proposed algorithm significantly outperforms the GEVD-NLS. It can be explained by the fact that in contrast to the GEVD in [33], the SECSI framework uses an SMD-based approach which takes all the slices of a tensor into account and chooses the best estimate instead of considering only one pencil in GEVD, which makes its performance more stable.

VI. CONCLUSIONS

In this paper, we have introduced a new algorithm to estimate the factors of the coupled rank- $(L_r, L_r, 1)$ BTB of multiple tensors with a common mode based on the coupled extension of the semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalizations (SECSI). Additionally, we have shown how the multilinear rank structure of the coupled rank- $(L_r, L_r, 1)$ decomposition can be estimated based on the extension of the LaRGE scheme and clustering techniques. The numerical results demonstrate that compared to the GEVD-based approach in [30], [33], the proposed SECSI-based initialization guarantees the convergence, and therefore is more accurate and numerically stable. Moreover, in contrast to [30], [33], our approach allows performing the decomposition even if the ranks are not available beforehand. Therefore it can be further used in real data applications with unknown rank structure, for instance, in EEG-MEG, since these signals are usually recorded simultaneously and thus exhibit coupling.

REFERENCES

- [1] T. G. Kolda and B. W. Bader, "Tensor Decompositions and Applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.
- [2] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A Multilinear Singular Value Decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000.
- [3] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, pp. 279–311, 1966.
- [4] J. D. Carroll and J. J. Chang, "Analysis of individual differences in multidimensional scaling via an N -way generalization of "Eckart-Young" decomposition," *Psychometrika*, vol. 35, pp. 283–319, 1970.
- [5] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, pp. 1–84, 1970.
- [6] L. De Lathauwer, "Decompositions of a Higher-Order Tensor in Block Terms—Part I: Lemmas for Partitioned Matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1022–1032, 2008.
- [7] L. De Lathauwer, "Decompositions of a Higher-Order Tensor in Block Terms—Part II: Definitions and Uniqueness," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1033–1066, 2008.
- [8] L. De Lathauwer and D. Nion, "Decompositions of a Higher-Order Tensor in Block Terms—Part III: Alternating Least Squares Algorithms," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1067–1083, 2008.
- [9] D. Nion and L. De Lathauwer, "A link between the decomposition of a third-order tensor in rank- $(L, L, 1)$ terms and Joint Block Diagonalization," in *Proc. 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec. 2009, pp. 89–92.
- [10] L. Sorber, M. Van Barel, and 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 Journal on Optimization*, vol. 23, no. 2, pp. 695–720, 2013.
- [11] L. De Lathauwer, "Blind Separation of Exponential Polynomials and the Decomposition of a Tensor in Rank- $(L_r, L_r, 1)$ Terms," *SIAM Journal on Matrix Analysis and Applications*, vol. 32, no. 4, pp. 1451–1474, 2011.
- [12] P. M. R. de Oliveira and V. Zarzoso, "Block Term Decomposition of ECG Recordings for Atrial Fibrillation Analysis: Temporal and Inter-Patient Variability," *Journal of Communication and Information Systems*, vol. 34, no. 1, pp. 111–119, 2019.
- [13] L. N. Ribeiro, A. L. F. de Almeida, and V. Zarzoso, "Enhanced block term decomposition for atrial activity extraction in atrial fibrillation ECG," in *Proc. 9th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Jul. 2016, pp. 1–5.
- [14] V. Zarzoso, "Parameter estimation in block term decomposition for noninvasive atrial fibrillation analysis," in *Proc. IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec. 2017.
- [15] L. De Lathauwer and A. de Baynast, "Blind Deconvolution of DS-CDMA Signals by Means of Decomposition in Rank- $(1, L, L)$ Terms," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1562–1571, 2008.
- [16] J.-X. Yang, X.-F. Gong, H. Li, Y.-G. Xu, and Z.-W. Liu, "Using coupled multilinear rank- $(L, L, 1)$ block term decomposition in multi-static-multi-pulse MIMO radar to localize targets," in *Proc. of International Symposium on Neural Networks*, Jul. 2019, pp. 565–574.
- [17] Y. Cheng, M. Riesmeyer, J. Haueisen, and M. Haardt, "Using the Multi-Linear Rank- $(L_r, L_r, 1)$ Decomposition for the Detection of the 200 Hz Band Activity in Somatosensory Evoked Magnetic Fields and Somatosensory Evoked Electrical Potentials," *IEEE Access*, vol. 9, pp. 106 232–106 244, 2021.
- [18] X. Han, L. Albera, A. Kachenoura, H. Shu, and L. Senhadji, "Block term decomposition with rank estimation using group sparsity," in *Proc. IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec. 2017.
- [19] E. Gujral and E. E. Papalexakis, "OnlineBTD: Streaming Algorithms to Track the Block Term Decomposition of Large Tensors," in *Proc. IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*, 2020, pp. 168–177.
- [20] E. Gujral, R. Pasricha, and E. Papalexakis, "Beyond Rank-1: Discovering Rich Community Structure in Multi-Aspect Graphs," in *Proc. of The Web Conference 2020*, New York, USA, 2020, p. 452–462.
- [21] G. Zhang, X. Fu, K. Huang, and J. Wang, "Hyperspectral Super-Resolution: A Coupled Nonnegative Block-Term Tensor Decomposition Approach," in *Proc. IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec. 2019, pp. 470–474.
- [22] G. Zhang, X. Fu, J. Wang, X.-L. Zhao, and M. Hong, "Spectrum Cartography via Coupled Block-Term Tensor Decomposition," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3660–3675, 2020.
- [23] A. A. Rontogiannis, E. Kofidis, and P. V. Giampouras, "Block-Term Tensor Decomposition: Model Selection and Computation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 3, pp. 464–475, 2021.
- [24] —, "Block-Term Tensor Decomposition: Model Selection and Computation," in *Proc. 28th European Signal Processing Conference (EUSIPCO 2020)*, Jan. 2021, pp. 1976–1980.
- [25] A. A. Rontogiannis, P. V. Giampouras, and E. Kofidis, "Rank-Revealing Block-Term Decomposition for Tensor Completion," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2021)*, Jun. 2021, pp. 2915–2919.
- [26] J. H. d. M. Goulart, P. M. R. de Oliveira, R. C. Farias, V. Zarzoso, and P. Comon, "Alternating Group Lasso for Block-Term Tensor Decomposition and Application to ECG Source Separation," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2682–2696, 2020.
- [27] X.-F. Gong, Q.-H. Lin, O. Debals, N. Vervliet, and L. De Lathauwer, "Coupled rank- (L_m, L_n, \bullet) block term decomposition by coupled block simultaneous generalized Schur decomposition," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016)*, Mar. 2016, pp. 2554–2558.
- [28] M. Sørensen and 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 Journal on Matrix Analysis and Applications*, vol. 36, pp. 496–522, Apr. 2015.
- [29] M. Sørensen, I. Domanov, 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 Journal on Matrix Analysis and Applications*, vol. 36, pp. 1015–1045, Jan. 2015.
- [30] L. Sorber, M. Van Barel, and L. De Lathauwer, "Structured Data Fusion," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 4, pp. 586–600, 2015.
- [31] F. Roemer and M. Haardt, "A semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalizations (SECSI)," *Signal Processing*, vol. 93, no. 9, pp. 2722–2738, 2013.
- [32] A. Manina, M. Grasis, L. Khamidullina, A. Korobkov, J. Haueisen, and M. Haardt, "Coupled CP Decomposition of EEG and MEG Magnetometer and Gradiometer Measurements via the Coupled SECSI Framework," in *Proc. 55th Asilomar Conference on Signals, Systems, and Computers*, 2021.
- [33] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, "Tensorlab 3.0," Mar. 2016. [Online]. Available: <https://www.tensorlab.net/>
- [34] A. A. Korobkov, M. K. Diugoriva, J. Haueisen, and M. Haardt, "Multi-dimensional model order estimation using LineAr Regression of Global Eigenvalues (LaRGE) with applications to EEG and MEG recordings," in *Proc. 28th European Signal Processing Conference (EUSIPCO 2020)*, Jan. 2021, pp. 1005–1009.
- [35] B. Ahmad, L. Khamidullina, A. A. Korobkov, A. Manina, J. Haueisen, and M. Haardt, "Joint Model Order Estimation for Multiple Tensors with a Coupled Mode and Applications to the Joint Decomposition of EEG, MEG Magnetometer, and Gradiometer tensors," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2022)*, May 2022.
- [36] L. De Lathauwer, "A Link between the Canonical Decomposition in Multilinear Algebra and Simultaneous Matrix Diagonalization," *SIAM Journal on Matrix Analysis and Applications*, vol. 28, no. 3, pp. 642–666, Jan. 2006.
- [37] T. Fu and X. Gao, "Simultaneous Diagonalization With Similarity Transformation for Non-Defective Matrices," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006)*, May 2006, pp. 1137–1140.
- [38] S. Zhu, D. Wang, and T. Li, "Data clustering with size constraints," *Knowledge-Based Systems*, vol. 23, no. 8, pp. 883–889, 2010.