2023 IEEE 9th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP) | 979-8-3503-4452-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/CAMSAP58249.2023.10403496

Coupled Matrix Tensor Factorization via a Semi-Algebraic Solution based on Simultaneous Matrix Diagonalization (SECSI-CMTF)

Alla Manina[1], *Mikus Janis Grasis*[1], *André L. F. de Almeida*[2], and *Martin Haardt*[1]
[1]Communications Research Laboratory, Technische Universität Ilmenau, Ilmenau, Germany
[2]Wireless Telecommunications Research Group, Federal University of Ceará, Fortaleza, Brazil
Email: {alla.manina, mikus.grasis, martin.haardt}@tu-ilmenau.de, andre@gtel.ufc.br

***Abstract*—Recent research has shown that the joint analysis of heterogeneous data can be beneficial to understand the underlying structure of the data compared to a separate analysis. This research direction has gained high interest due to the technological progress, where massive amounts of data from multiple sources are collected, e.g., multi-modal data from a patient such as EEG (electroencephalogram), MAG (magnetoencephalogram), and other data gathered from laboratory tests. This task of data fusion is challenging due to the heterogeneous structure of the data. In this study, we perform a joint CP decomposition of a heterogeneous data set, i.e., a matrix coupled with a three-dimensional tensor along the first mode, via a new formulation of the coupled matrix and tensor factorization (CMTF) based on the SEmi-Algebraic framework for approximate CP decomposition via SImultaneous matrix diagonalization (SECSI). In comparison with the traditional alternating and gradient-based optimization algorithms, the proposed SECSI-CMTF algorithm shows an accurate and robust performance with a significantly increased computational speed. The results are evaluated on synthetic data set and compared to other state-of-the-art approaches, also in ill-conditioned scenarios and in scenarios with different SNRs.**

***Index Terms*—tensor decomposition, CP, coupled matrix-tensor factorization, semi-algebraic framework, SECSI, simultaneous diagonalization**

I. Introduction

Due to its high significance, rigorous requirements, and huge potential, data fusion from multiple sources has always been attracting a lot of research interest. From its structure, the data can be gathered in different formats, i.e., matrices or higher-order tensors, forming the heterogeneous data sets and this raises a challenge. The joint analysis of massive heterogeneous data requires reliable mathematical tools that can provide sufficient processing for further interpretation to reveal the underlying structure of the data, while providing an acceptable complexity-accuracy trade-off.

The task of the joint analysis of heterogeneous data sets first started with the collective matrix factorization [1] and has been successfully applied in bioinformatics [2], [3], social network analysis [4], and signal processing [5]. Furthermore, this problem was reformulated for the joint processing of heterogeneous data sets, i.e., a matrix coupled with a higher-order tensor [6]–[10]. The joint factorization of such data has already been successfully applied, e.g., in [7], where the authors provided the quantitative comparison of histology and in vitro samples with the goal of differentiating between cancerous and healthy states of brain, breast, and bone tissues. Additionally, this approach found its application in social networks [11], [12]. The authors in [6] formulated the CMTF problem where the heterogeneous data sets are modeled by fitting outer-product models to higher-order tensors and matrices in a coupled manner with the squared Euclidean distance as the loss function. They used a gradient-based optimization approach. Furthermore, this approach has been extended in [8] to reveal the common and distinct components of the aforementioned heterogeneous data set structure by incorporating sparsity penalties on component weights. The approach in [8] is also a gradient-based optimization approach and has already a successful application in metabolomics.

Our contribution is summarized as follows. In this paper, we introduce a robust semi-algebraic framework to compute the CMTF via an extension of SECSI [13]. This leads to a significantly reduced computational complexity as compared to the state-of-the-art CMTF approaches [6]. SECSI-CMTF demonstrates an accurate performance and a significantly increased computational speed on simulated data.

The data model from [6] is highlighted in Section II. In Section III, we introduce our proposed approach. In Section IV, we evaluate the proposed approach via simulations, and Section V concludes our findings.

The following notation is used in the paper. Scalars are denoted as lower-case and capital italic letters, i.e., a, A. The bold-faced lower-case and capital letters $\boldsymbol{a}$ and $\boldsymbol{A}$ stand for vectors and matrices, respectively, and bold-faced calligraphic letters $\boldsymbol{\mathcal{A}}$ stand for tensors. We denote the transposition, Hermitian transpose, matrix inversion, and Moore-Penrose pseudo matrix inversion as ${}^{\mathsf{T}}$, ${}^{\mathsf{H}}$, ${}^{-1}$, and ${}^{+}$, respectively. The symbols $\circ$, $\diamond$, and $\oslash$ stand for the outer product, Khatri-Rao product, and element-wise division. The higher-order norm of a tensor and the Frobenius norm of a matrix are denoted as $\|\cdot\|_{\mathrm{H}}$ and $\|\cdot\|_{\mathrm{F}}$, respectively. We use the Kruskal operator [14] to denote the CP model of an N-way tensor

$\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$, i.e., $\boldsymbol{\mathcal{A}} = [\![\boldsymbol{F}_1, \boldsymbol{F}_2, \ldots, \boldsymbol{F}_N]\!]$, where $\boldsymbol{F}_n \in \mathbb{C}^{I_n \times R}$, $n = 1, \ldots, N$, are the factor matrices of $\boldsymbol{\mathcal{A}}$. The n-mode product of a tensor $\boldsymbol{\mathcal{A}}$ with a matrix $\boldsymbol{B} \in \mathbb{C}^{I_n \times R}$ is denoted as $\boldsymbol{\mathcal{A}} \times_n \boldsymbol{B}$. Finally, the N-way super-diagonal tensor which holds ones on the super-diagonal and zeros elsewhere is denoted by $\boldsymbol{\mathcal{I}}_{N,R} \in \mathbb{C}^{R \times R \times \cdots \times R}$.

II. Data Model

The objective function of the joint factorization of heterogeneous data sets, e.g., consisting of a tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ coupled with a matrix $\boldsymbol{Y} \in \mathbb{C}^{I_1 \times M}$ can be formulated as

$$f(\boldsymbol{F}_1, \boldsymbol{F}_2, \boldsymbol{F}_3, \boldsymbol{D}) = \|\boldsymbol{\mathcal{X}} - [\![\boldsymbol{F}_1, \boldsymbol{F}_2, \boldsymbol{F}_3]\!]\|_{\mathrm{F}}^2 + \|\boldsymbol{Y} - \boldsymbol{F}_1 \boldsymbol{D}^{\mathsf{H}}\|_{\mathrm{F}}^2, \tag{1}$$

where the first mode is assumed to be coupled, i.e., the matrix $\boldsymbol{F}_1 \in \mathbb{C}^{I_1 \times R}$ is shared by both factorizations [6]. This formulation fits the CANDECOMP/PARAFAC model of $\boldsymbol{\mathcal{X}}$ and factorizes $\boldsymbol{Y}$ according to the corresponding coupled matrix $\boldsymbol{F}_1$ that is assumed to be the same along the first mode for both, $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{Y}$ [15], [16]. The remaining matrices $\boldsymbol{F}_n \in \mathbb{C}^{I_n \times R}$, $n = 2, \ldots, N$, are the n-mode factor matrices of the tensor $\boldsymbol{\mathcal{X}}$, and $\boldsymbol{D} \in \mathbb{C}^{M \times R}$ is a factor matrix along the second mode of $\boldsymbol{Y}$.

This formulation as in (1) was proposed in [6]. However, according to [6], it can be considered as a special case of the approach introduced by [9] for multi-way multi-block data analysis. This research direction was also of interest in [7], [11], [12], [17]. Moreover, due to its high significance in many applications, e.g., in metabolomics, the CMTF model has been extended to the unsupervised identification of the common (shared) and individual (unshared) structures across multiple data sets in [8].

III. SECSI-CMTF

A. Joint subspace estimation

Let us consider a three-dimensional tensor $\boldsymbol{\mathcal{X}}_0 \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ and a matrix $\boldsymbol{Y}_0 \in \mathbb{C}^{I_1 \times M}$, which are coupled along the first mode.[1] In this case, the tensor and the matrix are defined as

$$\boldsymbol{\mathcal{X}}_0 = \boldsymbol{\mathcal{I}}_{3,R} \times_1 \boldsymbol{F}_1 \times_2 \boldsymbol{F}_2 \times_3 \boldsymbol{F}_3 \in \mathbb{C}^{I_1 \times I_2 \times I_3}, \tag{2}$$

$$\boldsymbol{Y}_0 = \boldsymbol{F}_1 \boldsymbol{D}^{\mathsf{H}} \in \mathbb{C}^{I_1 \times M}, \tag{3}$$

where $\boldsymbol{F}_n \in \mathbb{C}^{I_n \times R}$, $n = 1, \ldots, N$, are the factor matrices of the tensor. Similarly, the matrices $\boldsymbol{F}_1 \in \mathbb{C}^{I_1 \times R}$ and $\boldsymbol{D} \in \mathbb{C}^{M \times R}$ are the factor matrices of $\boldsymbol{\mathcal{Y}}_0$. Here, $\boldsymbol{F}_1$ is a coupled factor matrix along the first mode, i.e., it is the same for $\boldsymbol{\mathcal{X}}_0$ and a $\boldsymbol{Y}_0$. In practice, we observe a noise-corrupted model as

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{X}}_0 + \boldsymbol{\mathcal{N}} \text{ and } \boldsymbol{Y} = \boldsymbol{Y}_0 + \boldsymbol{N}, \tag{4}$$

where $\boldsymbol{\mathcal{N}}$ and $\boldsymbol{N}$ are the additive noise tensor and a matrix, respectively. In our proposed framework, we also consider the case when the tensor and the matrix can have different SNRs.

[1] However, the coupling can also occur in the second or third mode and the resulting framework can be extended to the case with multiple coupled matrices.

For notational simplicity, we show the derivations for the noiseless case [13]. The truncated HOSVD (T-HOSVD) is defined as

$$\boldsymbol{\mathcal{X}}_0 = \boldsymbol{\mathcal{S}}^{[\mathrm{s}]} \times_1 \boldsymbol{U}_1^{[\mathrm{s}]} \times_2 \boldsymbol{U}_2^{[\mathrm{s}]} \times_3 \boldsymbol{U}_3^{[\mathrm{s}]}, \tag{5}$$

where $\boldsymbol{\mathcal{S}}^{[\mathrm{s}]} \in \mathbb{C}^{R \times R \times R}$ is the truncated core tensor and $\boldsymbol{U}_n^{[\mathrm{s}]} \in \mathbb{C}^{I_n \times R}$, $n = 1, \ldots, N$, are the unitary subspace matrices of a tensor that span the n-mode column space of the corresponding unfolding of $\boldsymbol{\mathcal{X}}$. For the first coupled mode, the common unitary subspace matrix $\boldsymbol{U}_1^{[\mathrm{s}]}$ is obtained via

$$\left[[\boldsymbol{\mathcal{X}}_0]_{(1)}, \boldsymbol{Y}_0 \right] = \boldsymbol{U}_1^{[\mathrm{s}]} \cdot \boldsymbol{\Sigma}_1^{[\mathrm{s}]} \cdot \boldsymbol{V}_1^{[\mathrm{s}]^{\mathsf{H}}}, \tag{6}$$

where $\boldsymbol{U}_1^{[\mathrm{s}]} \in \mathbb{C}^{I_1 \times R}$, $\boldsymbol{\Sigma}_1^{[\mathrm{s}]} \in \mathbb{C}^{R \times R}$ and $\boldsymbol{V}_1^{[\mathrm{s}]^{\mathsf{H}}} \in \mathbb{C}^{R \times (I_2 I_3 + M)}$. For the latter, note that

$$\boldsymbol{V}_1^{[\mathrm{s}]^{\mathsf{H}}} = \left[\boldsymbol{V}_{I_2 I_3}^{[\mathrm{s}]^{\mathsf{H}}}, \quad \boldsymbol{V}_M^{[\mathrm{s}]^{\mathsf{H}}} \right], \tag{7}$$

where $\boldsymbol{V}_{I_2 I_3}^{[\mathrm{s}]^{\mathsf{H}}} \in \mathbb{C}^{R \times I_2 I_3}$ and $\boldsymbol{V}_M^{[\mathrm{s}]^{\mathsf{H}}} \in \mathbb{C}^{R \times M}$ span the row space for the 1-mode unfolding of $\boldsymbol{\mathcal{X}}_0$ and $\boldsymbol{Y}_0$, respectively. We will make use of this fact in further derivations.

B. CPD via Simultaneous Matrix Diagonalization (SMD)

In this section, we introduce the SECSI-CMTF approach step-by-step for one tensor coupled with a matrix along the first mode. The suggested approach is based on the computation of the CP decomposition via Simultaneous Matrix Diagonalization (SMD) [13]. The initial step is the coupled truncated HOSVD described in the previous section. From the fundamental link between the HOSVD and the CP decomposition, we have [13]

$$\begin{aligned} \boldsymbol{\mathcal{X}}_0 &= \left(\boldsymbol{\mathcal{S}}^{[\mathrm{s}]} \times_3 \boldsymbol{U}_3^{[\mathrm{s}]} \right) \times_1 \boldsymbol{U}_1^{[\mathrm{s}]} \times_2 \boldsymbol{U}_2^{[\mathrm{s}]} \\ &= \left(\boldsymbol{\mathcal{I}}_{3,R} \times_3 \left(\boldsymbol{U}_3^{[\mathrm{s}]} \cdot \boldsymbol{T}_3 \right) \right) \times_1 \left(\boldsymbol{U}_1^{[\mathrm{s}]} \cdot \boldsymbol{T}_1 \right) \times_2 \left(\boldsymbol{U}_2^{[\mathrm{s}]} \cdot \boldsymbol{T}_2 \right), \end{aligned} \tag{8}$$

where $\boldsymbol{F}_n = \boldsymbol{U}_n^{[\mathrm{s}]} \cdot \boldsymbol{T}_n$, $n = 1, \ldots, N$. The matrices $\boldsymbol{T}_n \in \mathbb{C}^{R \times R}$, $n = 1, \ldots, N$, are the transform matrices, which diagonalize the core tensor $\boldsymbol{\mathcal{S}}^{[\mathrm{s}]}$. In order to obtain the tensor in terms of the transform matrices, we multiply (8) by the matrices $\boldsymbol{U}_1^{[\mathrm{s}]^{\mathsf{H}}}$ and $\boldsymbol{U}_2^{[\mathrm{s}]^{\mathsf{H}}}$ along the 1- and 2-mode, respectively, resulting in a non-symmetric SMD [13]

$$\boldsymbol{\mathcal{S}}_3 = \boldsymbol{\mathcal{F}}_3 \times_1 \boldsymbol{T}_1 \times_2 \boldsymbol{T}_2, \tag{9}$$

where $\boldsymbol{\mathcal{S}}_3 = \boldsymbol{\mathcal{S}}^{[\mathrm{s}]} \times_3 \boldsymbol{U}_3^{[\mathrm{s}]} \in \mathbb{C}^{R \times R \times I_3}$ and $\boldsymbol{\mathcal{F}}_3 = \boldsymbol{\mathcal{I}}_{3,R} \times_3 \boldsymbol{F}_3 \in \mathbb{C}^{R \times R \times I_3}$. The symmetric one is obtained by multiplying (9) by the pivoting slice $\boldsymbol{p} \in \{1, 2, \ldots, I_3\}$ from the right-hand-side (rhs) and left-hand-side (lhs), respectively

$$\begin{aligned} \mathbf{I}: \boldsymbol{S}_{3,k}^{\mathrm{rhs}} &= \boldsymbol{S}_{3,k} \cdot \boldsymbol{S}_{3,p}^{-1} \\ &= \boldsymbol{T}_1 \cdot \operatorname{diag}\{\boldsymbol{F}_3(k,:) \oslash \boldsymbol{F}_3(p,:)\} \cdot \boldsymbol{T}_1^{-1} \end{aligned} \tag{10}$$

and

$$\begin{aligned} \mathbf{II}: \boldsymbol{S}_{3,k}^{\mathrm{lhs}} &= \left(\boldsymbol{S}_{3,p}^{-1} \cdot \boldsymbol{S}_{3,k} \right)^{\mathsf{T}} \\ &= \boldsymbol{T}_2 \cdot \operatorname{diag}\{\boldsymbol{F}_3(k,:) \oslash \boldsymbol{F}_3(p,:)\} \cdot \boldsymbol{T}_2^{-1}, \end{aligned} \tag{11}$$

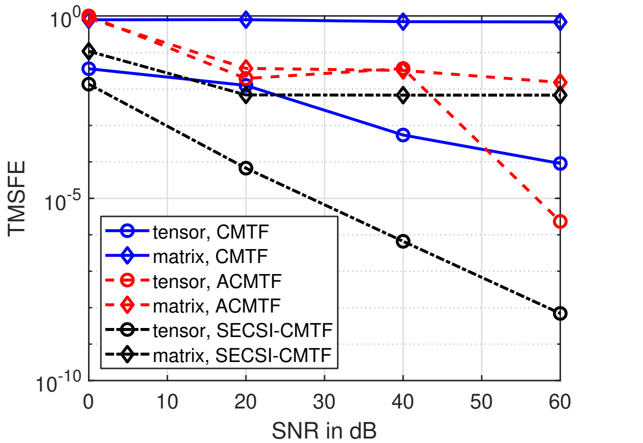

Figure 1: TMSFE as a function of SNR for $\hat{\boldsymbol{\mathcal{X}}} \in \mathbb{R}^{55\times150\times409}$ and $\hat{\boldsymbol{Y}} \in \mathbb{R}^{55\times409}$, $R = 8$, $\mathrm{SNR}_{\mathrm{matrix}} = 35\,\mathrm{dB}$, $\mathrm{SNR}_{\mathrm{tensor}} = 0:60\,\mathrm{dB}$, $\boldsymbol{F}_1$ is coupled with $\rho_1 = 0.97$.

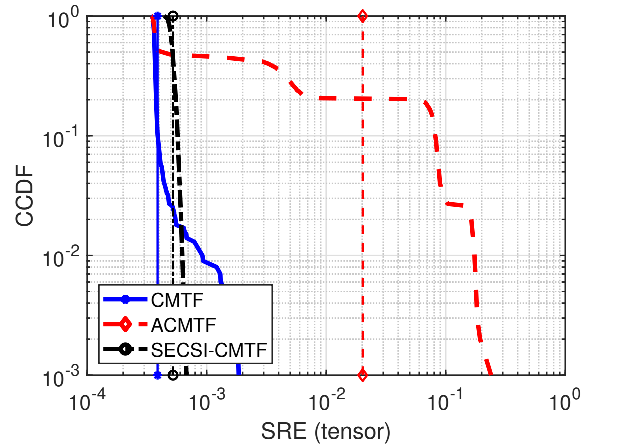

Figure 2: CCDF (Complementary Cumulative Distribution Function) of the SRE for $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{55\times150\times409}$, $R = 8$, $\mathrm{SNR} = 6\,\mathrm{dB}$, $\boldsymbol{F}_1$ is coupled with $\rho_1 = 0.97$. The vertical lines represent the mean value of the error for each curve.

where $\boldsymbol{S}_{3,k}$ stands for the k-th slice of the tensor $\boldsymbol{\mathcal{S}}_3$ along the third mode, $\boldsymbol{F}_3(k,:)$ is the k-th row of the matrix $\boldsymbol{F}_3$, and $k = 1,\ldots,I_3$. By applying the same procedure to the rest of the modes, six different SMDs are obtained as shown in Table I [13].

For more details regarding the solutions of the SMDs, we refer the reader to [13].

C. Estimation of the Factor Matrices

After solving the right-hand side SMD, we get the estimate for the first (coupled) matrix

$$\hat{\boldsymbol{F}}_{1,\mathrm{I}} = \boldsymbol{U}_1^{[\mathrm{s}]} \cdot \boldsymbol{T}_1, \tag{12}$$

where $\boldsymbol{T}_1$ is the same for both coupled tensor and coupled matrix. From the knowledge of this matrix and the shared subspace matrices $\boldsymbol{U}_1^{[\mathrm{s}]}$ and $\boldsymbol{V}_M^{[\mathrm{s}]^{\mathrm{H}}}$ from (6) and (7), we get an estimate of $\hat{\boldsymbol{Y}}$

$$\hat{\boldsymbol{Y}} = \boldsymbol{U}_1^{[\mathrm{s}]} \cdot \boldsymbol{T}_1 \cdot \hat{\boldsymbol{D}}^{\mathrm{H}} \leftarrow \hat{\boldsymbol{D}}^{\mathrm{H}} = \boldsymbol{T}_1^{-1} \cdot \boldsymbol{\Sigma}_1^{[\mathrm{s}]} \cdot \boldsymbol{V}_M^{[\mathrm{s}]^{\mathrm{H}}}. \tag{13}$$

Overall, we get two estimates of $\hat{\boldsymbol{Y}}$ from the corresponding SMDs I and III from Table I. The estimate of $\hat{\boldsymbol{F}}_{3,\mathrm{I}}$ is obtained form the diagonal elements of the jointly diagonalized matrices [13]. Since $\hat{\boldsymbol{F}}_{1,\mathrm{I}}$ and $\hat{\boldsymbol{F}}_{3,\mathrm{I}}$ are already estimated, the

Table I: SMDs

No.	Symmetric SMD
I	$\boldsymbol{S}_{3,k}^{\mathrm{rhs}} = \boldsymbol{T}_1 \cdot \mathrm{diag}\{\boldsymbol{F}_3(k,:) \oslash \boldsymbol{F}_3(p,:)\} \cdot \boldsymbol{T}_1^{-1}$
II	$\boldsymbol{S}_{3,k}^{\mathrm{lhs}} = \boldsymbol{T}_2 \cdot \mathrm{diag}\{\boldsymbol{F}_3(k,:) \oslash \boldsymbol{F}_3(p,:)\} \cdot \boldsymbol{T}_2^{-1}$
III	$\boldsymbol{S}_{2,k}^{\mathrm{rhs}} = \boldsymbol{T}_1 \cdot \mathrm{diag}\{\boldsymbol{F}_2(k,:) \oslash \boldsymbol{F}_2(p,:)\} \cdot \boldsymbol{T}_1^{-1}$
IV	$\boldsymbol{S}_{2,k}^{\mathrm{lhs}} = \boldsymbol{T}_3 \cdot \mathrm{diag}\{\boldsymbol{F}_2(k,:) \oslash \boldsymbol{F}_2(p,:)\} \cdot \boldsymbol{T}_3^{-1}$
V	$\boldsymbol{S}_{1,k}^{\mathrm{rhs}} = \boldsymbol{T}_2 \cdot \mathrm{diag}\{\boldsymbol{F}_1(k,:) \oslash \boldsymbol{F}_1(p,:)\} \cdot \boldsymbol{T}_2^{-1}$
VI	$\boldsymbol{S}_{1,k}^{\mathrm{lhs}} = \boldsymbol{T}_3 \cdot \mathrm{diag}\{\boldsymbol{F}_1(k,:) \oslash \boldsymbol{F}_1(p,:)\} \cdot \boldsymbol{T}_3^{-1}$

remaining estimate for $\hat{\boldsymbol{F}}_{2,\mathrm{I}}$ is computed via a Least Squares (LS) fit as follows

$$\hat{\boldsymbol{F}}_{2,\mathrm{I}} = \left[\boldsymbol{\mathcal{X}}_0\right]_{(2)} \cdot \left[\left(\hat{\boldsymbol{F}}_{3,\mathrm{I}} \diamond \hat{\boldsymbol{F}}_{1,\mathrm{I}}\right)^{\mathrm{T}}\right]^{+}. \tag{14}$$

From solving the remaining SMDs, the remaining four sets of factor matrix estimates are obtained. Overall, we obtain six sets of estimates. The final solution is chosen for both, the tensor and the matrix, separately and is based on the chosen selection criterion. In this paper, we consider the REConstructed Paired Solutions (REC PS) for the selection of the final factor matrix estimates. REC PS solves all SMDs and calculates the reconstruction error (RSE) for solutions coming from the same (paired) SMD. The pair of solutions, which leads to the smallest value of the RSE is then chosen as the final solution. The reconstruction error is calculated as

$$\mathrm{RSE} = \frac{\left\|\hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{X}}\right\|_{\mathrm{H}}^2}{\left\|\boldsymbol{\mathcal{X}}\right\|_{\mathrm{H}}^2}, \tag{15}$$

where $\hat{\boldsymbol{\mathcal{X}}}$ is the estimated tensor and $\boldsymbol{\mathcal{X}}$ denotes the noise-corrupted tensor. The same procedure applies to the final estimate selection of $\hat{\boldsymbol{Y}}$, where they are chosen from all available estimates independently.

IV. Simulations Results

In this section, we evaluate the proposed approach by conducting simulations with synthetic data. As benchmark algorithms, we use the traditional Coupled Matrix Tensor Factorization algorithm (CMTF) [8] and the Advanced CMTF (ACMTF) [18], which are available in the MATLAB CMTF toolbox (available from https://www.models.life.ku.dk).

First, we consider a real-valued scenario of a tensor $\boldsymbol{\mathcal{X}}$ coupled with a matrix $\boldsymbol{Y}$ along the first mode (1-mode) according to (2) and (3), respectively. All of the factor matrices, i.e.,

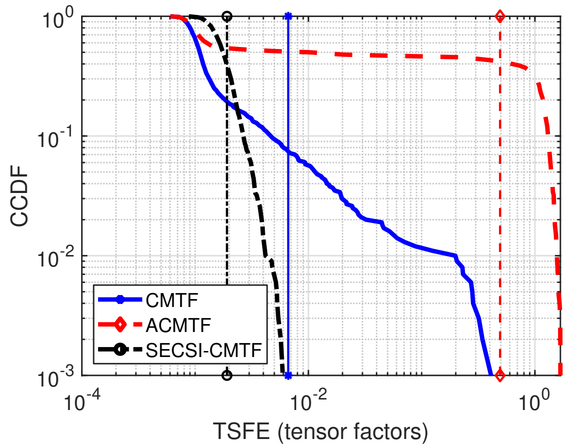

Figure 3: CCDF of the TSFE from the tensor factors, $\hat{\boldsymbol{\mathcal{X}}} \in \mathbb{R}^{55\times150\times409}$, $R = 8$, $\mathrm{SNR} = 6\,\mathrm{dB}$, $\boldsymbol{F}_1$ is coupled with $\rho_1 = 0.97$.

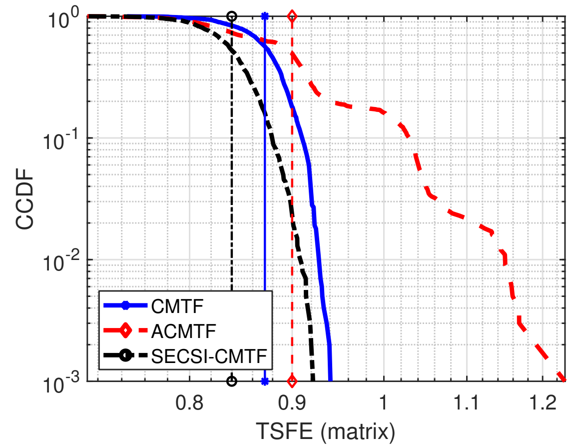

Figure 4: CCDF of the TSFE for $\hat{\boldsymbol{Y}} \in \mathbb{R}^{55\times409}$, $R = 8$, $\mathrm{SNR} = 6\,\mathrm{dB}$, $\boldsymbol{F}_1$ is coupled with $\rho_1 = 0.97$.

$\boldsymbol{F}_n$, $n = 1, \ldots, N$, and $\boldsymbol{D}$, are formed from i. i. d. zero mean Gaussian distributed random entries with variance one. Then, the synthetic data is generated via adding i. i. d. zero mean Gaussian noise with variance σ_{n}^2. As the accuracy measures, we use the squared reconstruction error (SRE) defined in (15) and the Total relative Mean Square Factor Error (TMSFE)

$$\mathrm{TMSFE} = \mathbb{E}\left\{ \sum_{\hat{\boldsymbol{F}}_n = \hat{\boldsymbol{F}}_1, \hat{\boldsymbol{F}}_2, \hat{\boldsymbol{F}}_3} \min_{\boldsymbol{P}\in\mathcal{M}_{PD}(R)} \frac{\|\hat{\boldsymbol{F}}_n \cdot \boldsymbol{P} - \boldsymbol{F}_n\|_{\mathrm{F}}^2}{\|\boldsymbol{F}_n\|_{\mathrm{F}}^2} \right\},$$

where $\mathcal{M}_{PD}(R)$ is a set of permuted diagonal matrices $\boldsymbol{P}$ of size $R \times R$, which correct for the inherent scaling and permutation ambiguity in the CP-model, R is the tensor rank, and $\boldsymbol{F}_n$ are the n-mode factor matrices, $n = 1, \ldots, N$.

In Fig. 1, we compare the TMSFE of CMTF, ACMTF, and SECSI-CMTF for a real-valued tensor $55 \times 150 \times 409$ with the rank $R = 8$ and matrix $\boldsymbol{Y}$, where $\boldsymbol{F}_1 \in \mathbb{R}^{55\times R}$ and $\boldsymbol{D} \in \mathbb{R}^{409\times R}$. In this scenario, we fix the SNR of $\boldsymbol{Y}$ to $\mathrm{SNR} = 35\,\mathrm{dB}$ and vary the SNR of $\boldsymbol{\mathcal{X}}$ from $0\,\mathrm{dB}$ to $60\,\mathrm{dB}$. To challenge the algorithms even more, we add correlation in the coupled mode with the correlation parameter $\rho_1 = 0.97$ via

$$\boldsymbol{F} \leftarrow \boldsymbol{F} \cdot \boldsymbol{R}(\rho), \;\; \boldsymbol{R}(\rho) = (1-\rho) \cdot \boldsymbol{I}_R + {}^{\rho}\!/_{R} \cdot \boldsymbol{1}_{R\times R},$$

where $\boldsymbol{1}_{R\times R}$ denotes a matrix of ones and ρ is a parameter to control the collinearity of the resulting matrix $\boldsymbol{F}$. Since the common mode is highly-correlated, it is ill-conditioned and, therefore, the approximate coupled matrix-tensor decomposition is difficult to calculate. The results have been averaged over 100 Monte Carlo trials. As can be observed from Fig. 1, the joint subspace estimation improves the estimates in low SNRs, and, thus, SECSI-CMTF shows a more accurate performance than CMTF and ACMTF. The average CPU for CMTF, ACMTF, and SECSI-CMTF was $45\,\mathrm{min.}$, $2\,\mathrm{h.}$, and $5\,\mathrm{sec.}$, respectively.

Next, in Fig. 2, we compare the SRE of CMTF, ACMTF, and SECSI-CMTF for the same simulation set up with 1,000 Monte-Carlo trials, except for SNR. Now we consider a low SNR case, i.e., $\mathrm{SNR} = 6\,\mathrm{dB}$, and it is fixed for both coupled objects. The TMSFE of the tensor factors $\boldsymbol{F}_1$, $\boldsymbol{F}_2$, and $\boldsymbol{F}_3$ and the matrix $\hat{\boldsymbol{Y}}$ with coupling via $\boldsymbol{F}_1$ are depicted in Fig. 3 and Fig. 4, respectively. SECSI-CMTF shows a more robust performance than CMTF with a significantly enhanced computational speed (CPU), i.e., the average CPU is $5\,\mathrm{sec.}$ and $13\,\mathrm{min.}$, respectively, while ACMTF decomposes the coupled objects with the average CPU being $49\,\mathrm{min.}$

V. Conclusions

In this paper, we have presented a new approach for coupled matrix tensor factorizations (CMTFs) based on solving the SMDs for a tensor coupled with a matrix along one mode. The developed algorithm has been compared with the traditional CMTF and ACMTF algorithms. The results of the simulations have shown that the proposed SECSI-CMTF outperforms the benchmark algorithms in critical scenarios, i.e., low SNRs and ill-conditioning in the coupled mode.

Moreover, while SECSI-CMTF provides a semi-algebraic solution, the CMTF and ACMTF are both dependent on the maximum number of function evaluations and inner iterations of the external optimization algorithms, which are both set manually. Thus, SECSI-CMTF has a significantly smaller computational complexity in all of the simulation scenarios. In contrast to the approaches in [8], [18], the proposed SECSI-CMTF algorithm supports matrix-tensor decompositions for complex-valued data.

References

[1] A. P. Singh and G. J. Gordon, "Relational learning via collective matrix factorization," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '08. New York, NY, USA: Association for Computing Machinery, 2008, p. 650–658.

[2] O. Alter, P. O. Brown, and D. Botstein, "Generalized singular value decomposition for comparative analysis of genome-scale expression data sets of two different organisms," *Proceedings of the National Academy of Sciences*, vol. 100, no. 6, pp. 3351–3356, 2003.

[3] L. Badea, "Extracting gene expression profiles common to colon and pancreatic adenocarcinoma using simultanous nonnegative matrix factorization," *Pacific Symposium on Biocomputing*, pp. 267–278, 2008.

[4] M. Jiang, P. Cui, R. Liu, Q. Yang, F. Wang, W. Zhu, and S. Yang, "Social contextual recommendation," in *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, ser. CIKM '12. New York, NY, USA: Association for Computing Machinery, 2012, p. 45–54.

[5] A. Yeredor, "Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation," *IEEE Transactions on Signal Processing*, vol. 50, no. 7, pp. 1545–1553, 2002.

[6] E. Acar, T. G. Kolda, and D. M. Dunlavy, "All-at-once optimization for coupled matrix and tensor factorizations," *In KDD Workshop on Mining and Learning with Graphs*, 2011.

[7] E. Acar, G. E. Plopper, and B. Yener, "Coupled analysis of in vitro and histology tissue samples to quantify structure-function relationship," *PLOS ONE*, vol. 7, no. 3, pp. 1–14, 03 2012.

[8] E. Acar, A. Lawaetz, M. Rasmussen, and R. Bro, "Structure-revealing data fusion model with applications in metabolomics," vol. 2013, 07 2013, pp. 6023–6026.

[9] A. Smilde, J. Westerhuis, and R. Boqué, "Multiway multiblock component and covariates regression models," *Journal of Chemometrics*, vol. 14, pp. 301–331, 05 2000.

[10] A. Banerjee, S. Basu, and S. Merugu, "Multi-way clustering on relation graphs," in *SDM*, 2007.

[11] B. Ermis, E. Acar, and T. Cemgil, "Link prediction via generalized coupled tensor factorisation," *In ECML/PKDD Workshop on Collective Learning and Inference on Structured Data*, 2012.

[12] Y.-R. Lin, J. Sun, P. Castro, R. Konuru, H. Sundaram, and A. Kelliher, "Metafac: Community discovery via relational hypergraph factorization," in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '09. New York, NY, USA: Association for Computing Machinery, 2009, p. 527–536.

[13] 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.

[14] T. G. Kolda, "Multilinear operators for higher-order decompositions," Sandia National Laboratories, Tech. Rep. SAND2006-2081, April 2006.

[15] J. D. Caroll 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.

[16] R. A. Harshman, "Foundations of the parafac procedure: Models and conditions for an explanatory multi-model factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, pp. 1–84, 1970.

[17] T. Wilderjans, E. Ceulemans, H. Kiers, and K. Meers, "The LMPCA program: A graphical user interface for fitting the linked-mode parafac-pca model to coupled real-valued data," *Behavior research methods*, vol. 41, pp. 1073–82, 11 2009.

[18] E. Acar, E. E. Papalexakis, G. Gurdeniz, M. A. Rasmussen, A. J. Lawaetz, M. Nilsson, and R. Bro, "Structure-revealing data fusion," *BMC Bioinformatics*, vol. 15, 2014.