Coupled Matrix-Tensor Factorizations — The Case of Partially Shared Factors

Lieven De Lathauwer*† and Eleftherios Kofidis§¶
*KU Leuven – E.E. Dept. (ESAT) – STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium
†Group Science, Engineering and Technology, KU Leuven – Kulak, E. Sabbelaan 53, 8500 Kortrijk, Belgium
E-mail: Lieven.DeLathauwer@kuleuven-kulak.be
§Department of Statistics and Insurance Science, University of Piraeus, 18534 Piraeus, Greece
¶Computer Technology Institute & Press "Diophantus" (CTI), Patras, Greece
E-mail: kofidis@unipi.gr

***Abstract*—Coupled matrix-tensor factorizations have proved to be a powerful tool for data fusion problems in a variety of applications. Uniqueness conditions for such coupled decompositions have only recently been reported, demonstrating that coupling through a common factor can ensure uniqueness beyond what is possible when considering separate decompositions. In view of the increasing interest in application scenarios involving more general notions of coupling, we revisit in this paper the uniqueness question for the important case where the factors common to the tensor and the matrix only share some of their columns. Related computational aspects and numerical examples are also discussed.**

I. Introduction

This is the era of Big Data, with an unprecedentedly large number of sources generating huge amounts of data of all kinds and dimensions, which often need to be *jointly* processed and analyzed to mine as much of the hidden information as possible. Data fusion, defined in [1] as the analysis of several data sets such that different data sets can interact and inform each other, is well established as the most promising data analysis approach in an increasing number of applications involving diverse sets of inter-related data. These include brain imaging [2]–[7], metabolomics [8], array processing [9], [10], multidimensional harmonic retrieval [11], [12], link prediction [13], collaborative filtering and data mining [14], [15], among many others. Diversity is the key to data fusion [1], [16]. For example, combining measurements from electroencephalography (EEG) and functional magnetic resonance imaging (fMRI), which are known to be complementary in their spatio-temporal resolutions, has shown considerable gains in neuroimaging studies when compared with the results of employing only one of the modalities [3], [17].

Research supported by: (1) Research Council KU Leuven: C1 project C16/15/059-nD, (2) F.W.O.: project G.0830.14N, G.0881.14N, (3) 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.

This work has been partly supported by the University of Piraeus Research Center.

Heterogeneity in the data sets, also manifested in their number of dimensions, with some of them being two-dimensional and others intrinsically multi-dimensional, very often suggests to model them with matrices and tensors, i.e., multi-way arrays [18], [19], respectively. In the EEG-fMRI fusion example, classical approaches adopt a channels$\times$time$\times$frequency tensor model for the EEG observations whereas the fMRI signal is commonly represented as a matrix with its dimensions corresponding to voxels and time [17]. Their fusion relies on the *coupling* of the EEG tensor and the fMRI matrix along their common mode (way or dimension), time. That could be also the subject mode in a multi-subject experiment [20].

Coupling of tensors and/or matrices along one or more modes is reflected in sharing corresponding latent factors in their decomposition.[1] Uniqueness conditions for such (coupled) decompositions have only recently been reported [24] and include both the canonical polyadic decomposition (CPD) and block-term decompositions (BTD) of multilinear rank-$(L_{r,n}, L_{r,n}, 1)$ [25], [26]. It was demonstrated that coupling through one or more common factors that are shared among the tensors can ensure uniqueness beyond what is possible when considering separate decompositions [24, Section 4.6]. Thus, uniqueness can be ensured via coupling even in cases where individual decompositions are not unique [9]. Both (semi-)algebraic [26]–[28] and optimization-based [21], [29]–[31] algorithms for performing such factorizations have recently been developed. Notably, a library of tensor decompositions and factor structures to choose from is offered in Tensorlab [32].

Coupled matrix-tensor factorizations (CMTF) are a special case, where both matrices and higher-order tensors are decomposed in a coupled manner [24]. Such models have been shown to provide a powerful tool for data fusion problems in a variety of applications, including metabolomics [8], estimation of missing data [33], brain imaging [34], [35], and data mining [36], among many others. Methods for joint compression [37], [38] and scalable decomposition of large

[1]Decompositions with shared factors can be viewed as constrained ones [21], [22] and, in a way, generalize joint matrix diagonalization ideas [23] to multi-block data sets.

coupled matrices and tensors (see [31], [39] and references therein) have been also developed.

Recently, there has been an increasing interest in the more general case of coupling through factors that are not equal (fully shared) but are only related to each other (correlated or similar) in some way (see, e.g., [26], [40]–[44]). Examples of this appear in numerous applications, where there are sources of variation that are only captured by some of the measuring modalities. These include, for example, joint analysis of EEG, fMRI and structural MRI (sMRI) signals [35] and multimodal measurements of chemical compounds in chemometrics [45]. Another example is found in the joint analysis of brain imaging data from both patients and controls (healthy subjects), where there are sources that only appear in one of the groups studied [2]. Recent performance analysis results (in the form of constrained Cramér-Rao bounds) show the parameter estimation accuracy to improve compared to an equivalent uncoupled model [46].[2] Algorithms for such coupled factorizations have been recently reported in [18], [21], [32], [45], [47]–[53] and may or may not assume the number of shared components *a-priori* known. For example, the *Advanced CMTF (ACMTF)* scheme [45] lets the model itself "reveal" which components are shared in each coupled mode by including the weights of the components in the optimization problem through a sparsity constraint. This scheme and its improved variants [42], [48] have been successfully applied in a number of neuroimaging problems [34], [35].

However, to the best of the authors' knowledge, no theoretical analysis of such a CMTF, namely with partially shared factors, has been reported yet.[3] The aim of this paper is to explore the uniqueness question when the factors common to the tensor and the matrix only share some (not all) of their columns. Related computational aspects are also briefly discussed and numerical examples are provided to illustrate the theory.

II. Problem Formulation

Let a third-order tensor $\boldsymbol{\mathcal{X}}^{(1)} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ and a matrix $\boldsymbol{X}^{(2)} \in \mathbb{R}^{J \times I_3}$ which are coupled in their last modes (of dimension I_3) through a (partially) shared latent factor. Namely, let the tensor have a CPD of rank R_1, $\boldsymbol{\mathcal{X}}^{(1)} = [\![\boldsymbol{A}^{(1)}, \boldsymbol{B}^{(1)}, \boldsymbol{C}^{(1)}]\!]$ and the matrix a rank-R_2 factorization (with $R_2 \geq 2$), $\boldsymbol{X}^{(2)} = \boldsymbol{A}^{(2)}\boldsymbol{C}^{(2)\mathrm{T}}$, where $\boldsymbol{A}^{(1)} \in \mathbb{R}^{I_1 \times R_1}$, $\boldsymbol{B}^{(1)} \in \mathbb{R}^{I_2 \times R_1}$, $\boldsymbol{C}^{(1)} \in \mathbb{R}^{I_3 \times R_1}$, $\boldsymbol{A}^{(2)} \in \mathbb{R}^{J \times R_2}$, and $\boldsymbol{C}^{(2)} \in \mathbb{R}^{I_3 \times R_2}$, with $\boldsymbol{C}^{(1)}, \boldsymbol{C}^{(2)}$ sharing at least one of their columns. The case of full sharing, that is, $\boldsymbol{C}^{(1)} = \boldsymbol{C}^{(2)} = \boldsymbol{C}$ (with $R_1 = R_2$), was studied in [24] as a special case of the coupled CPD (CCPD) of a set of tensors. Stacking then the mode-3 matricized tensor and the matrix one above the other results in

$$\begin{bmatrix} \boldsymbol{X}^{(1)}_{(3)} \\ \boldsymbol{X}^{(2)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}^{(1)} \odot \boldsymbol{A}^{(1)} \\ \boldsymbol{A}^{(2)} \end{bmatrix} \boldsymbol{C}^{\mathrm{T}}, \tag{1}$$

where $\odot$ denotes the Khatri-Rao product. Note that the full-rank decomposition of $\boldsymbol{X}^{(2)}$ is inherently non-unique [55]. Nevertheless, (1) shows that this is overcome through coupling with a higher-order tensor whose CPD can be unique under mild conditions. As explained in [24], for $\boldsymbol{A}^{(2)}$ to be uniquely determined from (1), the common factor $\boldsymbol{C}$ must have full column rank.

Consider now a more relaxed coupling condition, where the factors $\boldsymbol{C}^{(1)}, \boldsymbol{C}^{(2)}$ above only share F of their columns, with $1 \leq F < R_2$, and write (without loss of generality)

$$\boldsymbol{C}^{(1)} = \begin{bmatrix} \boldsymbol{C} & \boldsymbol{c}^{(1)}_{F+1} & \boldsymbol{c}^{(1)}_{F+2} & \cdots & \boldsymbol{c}^{(1)}_{R_1} \end{bmatrix} \tag{2}$$

$$\boldsymbol{C}^{(2)} = \begin{bmatrix} \boldsymbol{C} & \boldsymbol{c}^{(2)}_{F+1} & \boldsymbol{c}^{(2)}_{F+2} & \cdots & \boldsymbol{c}^{(2)}_{R_2} \end{bmatrix}, \tag{3}$$

where $\boldsymbol{C} \in \mathbb{R}^{I_3 \times F}$ is their common part. The question now is what are the implications of such a *partial* coupling for the (non)uniqueness of the matrix factorization above given that the tensor CPD is unique. It should be noted that the following analysis and discussion are valid for tensors $\boldsymbol{\mathcal{X}}^{(1)}$ of any order, although a 3rd-order tensor is assumed here, for the sake of simplicity.

III. On the (Non)Uniqueness of a CMTF with Partially-Shared Factors

Take the simplest (yet interesting and frequently encountered) case where $\boldsymbol{C}^{(2)}$ has only one unshared column with $\boldsymbol{C}^{(1)}$, that is, $R_2 = F + 1$ in (3). Note that the tensor CPD uniqueness implies that the first F columns of $\boldsymbol{C}^{(2)}$ are uniquely determined. The uncertainty is thus restricted to its R_2th column, which, in view of the full column rank assumption, can be expressed as a linear combination of the columns of $\boldsymbol{C}$ (its share of the common part) and a vector $\boldsymbol{q}$ (the part orthogonal to $\boldsymbol{C}$), which satisfies $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{q} = \boldsymbol{0}$. Thus,

$$\boldsymbol{c}^{(2)}_{R_2} = \boldsymbol{C}\boldsymbol{\alpha} + \beta\boldsymbol{q}, \tag{4}$$

with *unknown* coefficients $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{R_2-1} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{(R_2-1)\times 1}$ and $\beta \neq 0$. The matrix $\boldsymbol{C}^{(2)}$ can then be decomposed as

$$\begin{aligned} \boldsymbol{C}^{(2)\mathrm{T}} &= \underbrace{\begin{bmatrix} \boldsymbol{I}_{R_2-1} & \boldsymbol{0}_{(R_2-1)\times 1} \\ \boldsymbol{\alpha}^{\mathrm{T}} & \beta \end{bmatrix}}_{\boldsymbol{\Sigma} \in \mathbb{R}^{R_2 \times R_2}} \underbrace{\begin{bmatrix} \boldsymbol{C} & \boldsymbol{q} \end{bmatrix}^{\mathrm{T}}}_{\boldsymbol{C}_q \in \mathbb{R}^{I_3 \times R_2}} \\ &= \boldsymbol{\Sigma}\boldsymbol{C}_q^{\mathrm{T}}, \end{aligned} \tag{5}$$

and this is valid whether this is the true C-factor of $\boldsymbol{X}^{(2)}$ or an estimate thereof computed by a CCPD algorithm. Clearly, $\boldsymbol{\Sigma}$ is nonsingular (with determinant β) and $\boldsymbol{C}_q$ has full column rank. Hence

$$\begin{aligned} \boldsymbol{A}^{(2)} &= \boldsymbol{X}^{(2)} \left(\boldsymbol{C}^{(2)\mathrm{T}}\right)^{\dagger} = \boldsymbol{X}^{(2)} \left(\boldsymbol{C}_q^{\mathrm{T}}\right)^{\dagger} \boldsymbol{\Sigma}^{-1} \\ &= \boldsymbol{G}\boldsymbol{\Sigma}^{-1}, \end{aligned} \tag{6}$$

where $(\cdot)^{\dagger}$ denotes the Moore-Penrose inverse and

$$\boldsymbol{G} = \boldsymbol{X}^{(2)} \left(\boldsymbol{C}_q^{\mathrm{T}}\right)^{\dagger} = \begin{bmatrix} \boldsymbol{g}_1 & \boldsymbol{g}_2 & \cdots & \boldsymbol{g}_{R_2} \end{bmatrix} \in \mathbb{R}^{J \times R_2}. \tag{7}$$

Note that, since $\boldsymbol{\Sigma}$ is nonsingular, the vector $\boldsymbol{q}$, being the only unknown column of $\boldsymbol{C}_q$, can be computed from the range of

[2] Notice, however, that, as demonstrated in [33] for missing data estimation, there may exist cases (with weakly coupled components) for which data fusion is not preferable to the uncoupled analysis.

[3] *Matrix*-based methods for distinguishing common and distinct components in multi-block data sets were recently reviewed in [54], in a unifying subspace-based framework.

$\boldsymbol{C}^{(2)}$ (see eq. (5)), which in turn can be determined (based on the above assumptions) from the row space of $\boldsymbol{X}^{(2)}$ [55].[4] Indeed, the columns of $\boldsymbol{C}_q$ form a basis for the range of $(\boldsymbol{X}^{(2)})^{\mathrm{T}}$ and this space is also spanned by the first R_2 right singular vectors of $\boldsymbol{X}^{(2)}$, say $\boldsymbol{V}_1 := \boldsymbol{V}(:, 1:R_2)$, where $\boldsymbol{V}$ is its right singular matrix and Matlab indexing notation has been used. Thus, there exists an $R_2 \times R_2$ nonsingular matrix, $\boldsymbol{T} = \left[\begin{array}{cc} \boldsymbol{T}(:, 1:R_2-1) & \boldsymbol{t}_{R_2} \end{array} \right]$, such that $\boldsymbol{C}_q = \boldsymbol{V}_1 \boldsymbol{T}$. Hence $\boldsymbol{T}(:, 1:R_2-1) = \boldsymbol{V}_1^{\mathrm{T}} \boldsymbol{C}$ while $\boldsymbol{t}_{R_2}$ can de determined as the right singular vector of $\boldsymbol{C}^{\mathrm{T}} \boldsymbol{V}_1$ corresponding to its zero singular value. Then the direction $\boldsymbol{q}$ is found as $\boldsymbol{q} = \boldsymbol{V}_1 \boldsymbol{t}_{R_2}$. Thus, the uniqueness of the tensor CPD dictates the direction for $\boldsymbol{q}$. In other words, whatever is (the estimate of) $\boldsymbol{c}_{R_2}^{(2)}$, its $\boldsymbol{q}$ direction will be the same since it is exclusively determined by $\boldsymbol{C}$ and $\boldsymbol{X}^{(2)}$. Therefore $\boldsymbol{C}_q$ and hence $\boldsymbol{G}$ in (7) can be considered as known (and essentially unique in view of the coupling with the unique tensor CPD). Write (the resulting estimate of) $\boldsymbol{A}^{(2)}$ as $\boldsymbol{A}^{(2)} = \left[\begin{array}{cccc} \boldsymbol{a}_1^{(2)} & \boldsymbol{a}_2^{(2)} & \cdots & \boldsymbol{a}_{R_2}^{(2)} \end{array} \right]$. Since

$$\boldsymbol{\Sigma}^{-1} = \left[\begin{array}{cc} \boldsymbol{I}_{R_2-2} & \boldsymbol{0}_{(R_2-2)\times 1} \\ -\frac{1}{\beta}\boldsymbol{\alpha}^{\mathrm{T}} & \frac{1}{\beta} \end{array} \right],$$

(6) implies that $\boldsymbol{A}^{(2)}$ can be expressed as follows

$$\begin{aligned} \boldsymbol{a}_1^{(2)} &= \boldsymbol{g}_1 - \frac{\alpha_1}{\beta}\boldsymbol{g}_{R_2} = \boldsymbol{g}_1 - \alpha_1 \boldsymbol{a}_{R_2}^{(2)} \\ \boldsymbol{a}_2^{(2)} &= \boldsymbol{g}_2 - \frac{\alpha_2}{\beta}\boldsymbol{g}_{R_2} = \boldsymbol{g}_2 - \alpha_2 \boldsymbol{a}_{R_2}^{(2)} \\ &\vdots \\ \boldsymbol{a}_{R_2-1}^{(2)} &= \boldsymbol{g}_{R_2-1} - \frac{\alpha_{R_2-1}}{\beta}\boldsymbol{g}_{R_2} = \boldsymbol{g}_{R_2-1} - \alpha_{R_2-1} \boldsymbol{a}_{R_2}^{(2)} \\ \boldsymbol{a}_{R_2}^{(2)} &= \frac{1}{\beta}\boldsymbol{g}_{R_2}, \end{aligned} \tag{8}$$

namely in terms of the *unknown* coefficients $\boldsymbol{\alpha}$, β and the *known* matrix $\boldsymbol{G}$. For the latter, note that $\boldsymbol{G}(:, 1{:}R_2-1) = \boldsymbol{X}^{(2)}(\boldsymbol{C}^{\mathrm{T}})^{\dagger}$ and $\boldsymbol{g}_{R_2} = \boldsymbol{X}^{(2)} \frac{\boldsymbol{q}}{\|\boldsymbol{q}\|_2^2}$. These results can be summarized as follows:

Theorem 1. *Consider a third-order tensor $\boldsymbol{\mathcal{X}}^{(1)} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ having a unique CPD of rank $R_1 \geq 2$ with factor matrices $\boldsymbol{A}^{(1)} \in \mathbb{R}^{I_1 \times R_1}$, $\boldsymbol{B}^{(1)} \in \mathbb{R}^{I_2 \times R_1}$, and $\boldsymbol{C}^{(1)} \in \mathbb{R}^{I_3 \times R_1}$, and a matrix $\boldsymbol{X}^{(2)} \in \mathbb{R}^{J \times I_3}$ with rank-R_2 decomposition $\boldsymbol{X}^{(2)} = \boldsymbol{A}^{(2)} \boldsymbol{C}^{(2)\mathrm{T}}$. If $\boldsymbol{C}^{(2)}$ has only one unshared component (column) with $\boldsymbol{C}^{(1)}$, the indeterminacies in the matrix $\boldsymbol{A}^{(2)}$ are as in (8) where $\boldsymbol{G}$ is known and given by (7) and $\alpha_1, \alpha_2, \ldots, \alpha_{R_2-1}$ and $\beta \neq 0$ are unknown.*

Remarks.

1) Coupling of a matrix with a unique CPD through a common factor as in (1) removes the indeterminacies in the matrix factorization (except of course for permutation and scaling of columns) [24, Section 4.6]. When there is an unshared column in the C-factor, one gets a one-dimensional indeterminacy (in the direction of $\boldsymbol{g}_{R_2}$) in all the columns of $\boldsymbol{A}^{(2)}$ but the one corresponding to the unshared component. The latter is uniquely determined, up to unknown scaling. Hence if not all components in the coupled mode are shared, the coupling does not make the matrix $\boldsymbol{A}^{(2)}$ unique but it significantly reduces the indeterminacies in it. Further reductions could be possible if some (limited) *a-priori* information were available, for instance, allowing β and hence $\boldsymbol{a}_{R_2}^{(2)}$ to be also determined.
2) Regarding the compensation of the indeterminacies in $\boldsymbol{c}_{R_2}^{(2)}$ by those in $\boldsymbol{A}^{(2)}$, it is also of interest to see the unknown coefficients in (8). Namely, the larger is $|\beta|$ with respect to the $|\alpha|$'s, that is, the more significant is the new 'information' provided by the unshared component, the less significant is the unknown part of the columns of $\boldsymbol{A}^{(2)}$ and hence the closer to being known is $\boldsymbol{A}^{(2)}$. A possibly useful (and unique) approximation to $\boldsymbol{A}^{(2)}$ can be computed by projecting its first $R_2 - 1$ columns on the orthogonal complement of $\boldsymbol{g}_{R_2}$.[5]
3) Following a similar argument, one can show that an analogous result holds for the case that $\boldsymbol{C}^{(2)}$ has more than one (i.e., $R_2 - F > 1$) unshared components with $\boldsymbol{C}^{(1)}$. In such a case, however, there is an $(R_2 - F)$-dimensional indeterminacy in the columns of $\boldsymbol{A}^{(2)}$. For example, with two unshared columns, $\boldsymbol{C}_q = \left[\begin{array}{ccc} \boldsymbol{C} & \boldsymbol{q}_{R_2-1} & \boldsymbol{q}_{R_2} \end{array} \right]$ with $\boldsymbol{C}^{\mathrm{T}} \boldsymbol{q}_r = \boldsymbol{0}$, $r = R_2 - 1, R_2$ and the vectors $\boldsymbol{q}_{R_2-1} \perp \boldsymbol{q}_{R_2}$ can each lie in one of $R_2 - F = 2$ orthogonal directions, which are however known (given by the $R_2 - F$ right singular vectors of $\boldsymbol{C}^{\mathrm{T}} \boldsymbol{V}_1$ corresponding to its zero singular values). The last $R_2 - F$ columns of $\boldsymbol{G} = \boldsymbol{X}^{(2)} (\boldsymbol{C}_q^{\mathrm{T}})^{\dagger}$ are given by $\boldsymbol{g}_r = \boldsymbol{X}^{(2)} \frac{\boldsymbol{q}_r}{\|\boldsymbol{q}_r\|_2^2}$, $r = R_2 + 1, R_2$ and are therefore determined only up to a column *permutation*. Eq. (8) then takes the form

$$\begin{aligned} &\boldsymbol{A}^{(2)}(:, 1{:}R_2-2) = \\ &\qquad \boldsymbol{G}(:, 1{:}R_2-2) - \boldsymbol{g}_{R_2-1} \frac{\boldsymbol{\alpha}_{R_2-1}^{\mathrm{T}}}{\beta_{R_2-1}} - \boldsymbol{a}_{R_2}^{(2)} \boldsymbol{\alpha}_{R_2}^{\mathrm{T}} \\ \boldsymbol{a}_{R_2-1}^{(2)} &= \frac{1}{\beta_{R_2-1}} \boldsymbol{g}_{R_2-1} - \gamma_{R_2-1} \boldsymbol{a}_{R_2}^{(2)} \\ \boldsymbol{a}_{R_2}^{(2)} &= \frac{1}{\beta_{R_2}} \boldsymbol{g}_{R_2}, \end{aligned} \tag{9}$$

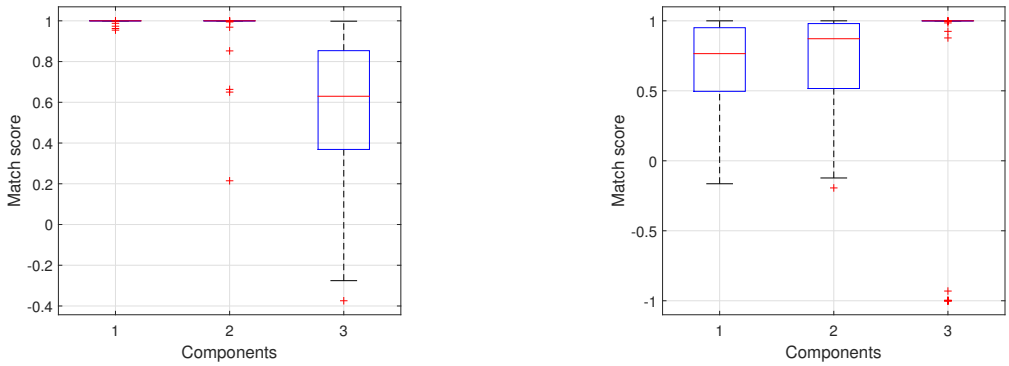
   with *unknown* coefficients $\boldsymbol{\alpha}_{R_2-1}, \boldsymbol{\alpha}_{R_2} \in \mathbb{R}^{F \times 1}$, $\gamma_{R_2-1} \in \mathbb{R}$, and $\beta_{R_2-1}, \beta_{R_2} \neq 0$.
4) Coupling through equality of the shared columns as in (2), (3) was assumed to keep the presentation simple. It should be noted that the above analysis and discussion are still valid in cases where the shared parts of the coupled factors are related through some other, non-identity function, such as a linear transformation [56].

IV. Numerical Examples

Consider the CMTF of a $6 \times 6 \times 6$ tensor and a 36×6 matrix (so that they have the same number of entries), both

[4] This vector is unique within the R_2-dimensional column space of $\boldsymbol{C}^{(2)}$. Moreover, $\beta \boldsymbol{q}$ is the projection of $\boldsymbol{c}_{R_2}^{(2)}$ onto the null space of $\boldsymbol{C}^{\mathrm{T}}$.

[5] Moreover, $\boldsymbol{g}_{R_2} \perp \boldsymbol{g}_i$, $i = 1, 2, \ldots, R_2 - 1$ if all nonzero singular values of $\boldsymbol{X}^{(2)}$ are equal.



(a) $\boldsymbol{C}^{(2)}$ (b) $\boldsymbol{A}^{(2)}$

Fig. 1. Results of the CMTF for the factors of the matrix. Component 3 is unshared in the C-factor.

of rank $R_1 = R_2 = 3$, which are coupled in their last modes. A large number of such tensor-matrix pairs were generated, with the entries of their factors drawn independently from the zero mean standard Gaussian distribution. The nonlinear least squares method from the *Structured Data Fusion (SDF)* framework [21] (constrained so that (2), (3) are satisfied, with $F = 2$) was used to compute their coupled decompositions. The Tensorlab `sdf_nls` function [31], [32]) was employed, with multi-start initialization. Fig. 1 depicts (in the form of boxplots) the match score (Tucker's phi coefficient [57]) for the $R_2 = 3$ columns of the factors of the matrix. The results show that, as expected, there is a strong indeterminacy in the estimation of the unshared column of $\boldsymbol{C}^{(2)}$ (see Fig. 1a), which is, however, compensated by a scaling-only indeterminacy in the corresponding column of $\boldsymbol{A}^{(2)}$ (see Fig. 1b). However, as shown in (8) and demonstrated in Fig. 1b, this induces a (one-dimensional) indeterminacy in the other columns of $\boldsymbol{A}^{(2)}$ resulting in their poor estimation.

V. Conclusions

In view of the increasing interest in application scenarios involving coupling among tensors and matrices more general than with equal latent factors, the important case where the factors common to the tensor and the matrix in a CMTF only share some of their columns was addressed in this paper. It was shown, via simple linear algebra arguments, how the indeterminacies in the unshared columns of the coupling factor are compensated by indeterminacies of a special character in the columns of the other factor of the matrix. Numerical examples were presented to illustrate the theory.

References

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

[2] V. D. Calhoun and T. Adalı, "ICA for fusion of brain imaging data," in *Signal Processing Techniques for Knowledge Extraction and Information Fusion*, D. Mandic *et al.*, Eds. Springer, 2008, ch. 12, pp. 221–240.

[3] T. Adalı, Y. Levin-Schwartz, and V. D. Calhoun, "Multimodal data fusion using source separation: Application to medical imaging," *Proc. IEEE*, vol. 103, no. 9, pp. 1494–1506, Sep. 2015.

[4] E. Karahan, P. A. Rojas-López, M. L. Bringas-Vega, P. A. Valdés-Hernández, and P. A. Valdes-Sosa, "Tensor analysis and fusion of multimodal brain images," *Proc. IEEE*, vol. 103, no. 9, pp. 1531–1559, Sep. 2015.

[5] X. Chen, Z. J. Wang, and M. J. McKeown, "Joint blind source separation for neurophysiological data analysis," *IEEE Signal Process. Mag.*, pp. 86–107, May 2016.

[6] X. Fu, K. Huang, O. Stretcu, H. A. Song, E. Papalexakis, P. Talukdar, T. Mitchell, N. Sidiropoulos, C. Faloutsos, and B. Poczos, "BRAIN-ZOOM: High resolution reconstruction from multi-modal brain signals," in *SIAM Int'l Conf. Data Mining (SDM-2017)*, Houston, TX, Apr. 2017.

[7] K. Naskovska, A. A. Korobkovy, M. Haardt, and J. Haueisen, "Analysis of the photic driving effect via joint EEG and MEG data processing based on the coupled CP decomposition," in *25th European Signal Processing Conf. (EUSIPCO-2017)*, Kos, Greece, 28 Aug.–2 Sep. 2017.

[8] E. Acar, R. Bro, and A. K. Smilde, "Data fusion in metabolomics using coupled matrix and tensor factorizations," *Proc. IEEE*, vol. 103, no. 9, pp. 1602–1620, Sep. 2015.

[9] M. Sørensen and L. De Lathauwer, "Coupled tensor decompositions for applications in array signal processing," in *5th IEEE Int'l Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP-2013)*, St. Martin, France, Dec. 2013.

[10] ——, "Multiple invariance ESPRIT for nonuniform linear arrays: A coupled canonical polyadic decomposition approach," *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3693–3704, Jul. 2016.

[11] ——, "Multidimensional harmonic retrieval via coupled canonical polyadic decomposition — Part I: Model and identifiability," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 517–527, Jan. 2017.

[12] ——, "Multidimensional harmonic retrieval via coupled canonical polyadic decomposition — Part II: Algorithm and multirate sampling," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 528–539, Jan. 2017.

[13] A. S. Zamzam, V. N. Ioannidis, and N. D. Sidiropoulos, "Coupled graph tensor factorization," in *50th Asilomar Conf. Signals, Systems and Computers (ACSSC-2016)*, Pacific Grove, CA, Nov. 2016.

[14] E. E. Papalexakis, C. Faloutsos, and N. D. Sidiropoulos, "Tensors for data mining and data fusion: Models, applications, and scalable algorithms," *ACM Trans. Intelligent Systems and Technology*, vol. 8, no. 2, pp. 16:1–16:44, 2016.

[15] F. M. Almutairi, N. D. Sidiropoulos, and G. Karypis, "Context-aware recommendation-based learning analytics using tensor and coupled matrix factorization," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 5, pp. 729–741, Aug. 2017.

[16] T. Adalı, Y. Levin-Schwartz, and V. D. Calhoun, "Multimodal data fusion using source separation: Two effective models based on ICA and IVA and their properties," *Proc. IEEE*, vol. 103, no. 9, pp. 1478–1493, Sep. 2015.

[17] B. Hunyadi, P. Dupont, W. Van Paesschen, and S. Van Huffel, "Tensor decompositions and data fusion in epileptic electroencephalography and functional magnetic resonance imaging data," *WIREs Data Mining Knowl. Discov.*, vol. 7, pp. 1–15, Jan./Feb. 2017.

[18] A. Cichocki, D. Mandic, A.-H. Phan, C. Caiafa, G. Zhou, Q. Zhao,

and L. De Lathauwer, "Tensor decompositions for signal processing applications — From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, pp. 145–163, Mar. 2015.

[19] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, Jul. 2017.

[20] W. Swinnen, B. Hunyadi, E. Acar, S. Van Huffel, and M. De Vos, "Incorporating higher dimensionality in joint decomposition of EEG and fMRI," in *22nd European Signal Processing Conf. (EUSIPCO-2014)*, Lisbon, Portugal, Sep. 2014.

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

[22] J. E. Cohen, K. Usevitch, and P. Comon, "A tour of constrained tensor canonical polyadic decomposition," 2016. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01311795/document

[23] G. Chabriel, M. Kleinsteuber, E. Moreau, H. Shen, P. Tichavsky, and A. Yeredor, "Joint matrices decompositions and blind source separation," *IEEE Signal Process. Mag.*, pp. 34–43, May 2014.

[24] 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, no. 2, pp. 496–522, Apr. 2015.

[25] X.-F. Gong, Q.-H. Lin, O. Debals, N. Vervliet, and L. De Lathauwer, "Coupled rank-$(L_m, L_n, \cdot)$ block term decomposition by coupled block simultaneous generalized Schur decomposition," in *41st Int'l Conf. Acoustics, Speech and Signal Processing (ICASSP-2016)*, Shanghai, China, Mar. 2016.

[26] X.-F. Gong, Q.-H. Lin, F.-Y. Cong, and L. De Lathauwer, "Double coupled canonical polyadic decomposition for joint blind source separation," ESAT-STADIUS, KU Leuven, Belgium, Tech. Rep., 2017.

[27] 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, no. 2, pp. 1015–1045, Jul. 2015.

[28] K. Naskovska and M. Haardt, "Extension of the semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalization to the efficient calculation of coupled CP decompositions," in *50th Asilomar Conf. Signals, Systems and Computers (ACSSC-2016)*, Pacific Grove, CA, Nov. 2016.

[29] 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, no. 4, pp. 1073–1082, 2009.

[30] E. Acar, T. G. Kolda, and D. M. Dunlavy, "All-at-once optimization for coupled matrix and tensor factorizations," in *9th Workshop on Mining and Learning with Graphs (MLG-2011)*, San Diego, CA, Aug. 2011.

[31] N. Vervliet, O. Debals, and L. De Lathauwer, "Tensorlab 3.0 — numerical optimization strategies for large-scale constrained and coupled matrix/tensor factorization," in *50th Asilomar Conf. Signals, Systems and Computers (ACSSC-2016)*, Pacific Grove, CA, Nov. 2016.

[32] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, *Tensorlab user guide — Release 3.0*, Mar. 2016. [Online]. Available: http://www.tensorlab.net/userguide3.pdf

[33] E. Acar, M. A. Rasmussen, F. Savorani, T. Næs, and R. Bro, "Understanding data fusion within the framework of coupled matrix and tensor factorizations," *Chemometrics and Intelligent Laboratory Systems*, vol. 129, pp. 53–63, 2013.

[34] E. Acar, Y. Levin-Schwartz, V. Calhoun, and T. Adalı, "Tensor-based fusion of EEG and fMRI to understand neurological changes in schizophrenia," in *50th IEEE Int'l Symp. Circuits and Systems (ISCAS-2017)*, Baltimore, MD, May 2017.

[35] ——, "ACMTF for fusion of multi-modal neuroimaging data and identification of biomarkers," in *25th European Signal Processing Conf. (EUSIPCO-2017)*, Kos, Greece, 28 Aug.–2 Sep. 2017.

[36] C. Faloutsos, "Tensor analysis — applications and algorithms," in *SIAM Annual Meeting*, Pittsburgh, PA, Jul. 2017, presentation. [Online]. Available: http://users.wfu.edu/ballard/SIAM-AN17/faloutsos.pdf

[37] L. De Lathauwer and J. Vandewalle, "Dimensionality reduction in higher-order signal processing and rank-$(R_1, R_2, \ldots, R_N)$ reduction in multilinear algebra," *Linear Algebra and its Applications*, vol. 391, pp. 31–55, Nov. 2004, Special issue on linear algebra in signal and image processing.

[38] J. E. Cohen, R. C. Farias, and P. Comon, "Joint tensor compression for coupled canonical polyadic decompositions," in *24th European Signal Processing Conf. (EUSIPCO-2016)*, Budapest, Hungary, 29 Aug.–2 Sep. 2016.

[39] D. Choi, J.-G. Jang, and U. Kang, "Fast, accurate, and scalable method for sparse coupled matrix-tensor factorization," Nov. 2017, arXiv:1708.08640v5 [cs.NA]. [Online]. Available: https://arxiv.org/pdf/1708.08640.pdf

[40] Y. K. Yılmaz, A. T. Cemgil, and U. Şimşekli, "Generalised coupled tensor factorisation," in *25th Neural Information Processing Systems (NIPS-2011)*, Granada, Spain, Dec. 2011.

[41] U. Şimşekli, A. T. Cemgil, and B. Ermiş, "Learning mixed divergences in coupled matrix and tensor factorization models," in *40th Int'l Conf. Acoustics, Speech and Signal Processing (ICASSP-2015)*, Brisbane, Australia, Apr. 2015.

[42] B. Rivet, M. Duda, A. Guérin-Dugué, C. Jutten, and P. Comon, "Multimodal approach to estimate the ocular movements during EEG recordings: A coupled tensor factorization method," in *Int'l Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC-2015)*, Milan, Italy, Aug. 2015.

[43] S. A. Khan, E. Leppäaho, and S. Kaski, "Bayesian multi-tensor factorization," *Machine Learning*, vol. 105, pp. 233–253, 2016.

[44] R. C. Farias, J. E. Cohen, and P. Comon, "Exploring multimodal data fusion through joint decompositions with flexible couplings," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4830–4844, Sep. 2016.

[45] E. Acar, E. E. Papalexakis, G. Gürdeniz, M. A. Rasmussen, A. J. Lawaetz, M. Nilsson, and R. Bro, "Structure-revealing data fusion," *Bioinformatics*, 2014. [Online]. Available: http://www.biomedcentral.com/1471-2105/15/239

[46] C. Ren, R. C. Farias, P.-O. Amblard, and P. Comon, "Performance bounds for coupled models," in *IEEE Workshop on Sensor Array and Multichannel Signal Processing (SAM-2016)*, Rio de Janeiro, Brazil, Jul. 2016.

[47] W. Liu, J. Chan, J. Bailey, C. Leckie, and K. Ramamohanarao, "Mining labelled tensors by discovering both their common and discriminative subspaces," in *SIAM Int'l Conf. Data Mining (SDM-2013)*, Austin, TX, May 2013.

[48] E. Acar, M. Nilsson, and M. Saunders, "A flexible modeling framework for coupled matrix and tensor factorizations," in *22nd European Signal Processing Conf. (EUSIPCO-2014)*, Lisbon, Portugal, Sep. 2014.

[49] M. Genicot, P.-A. Absil, R. Lambiotte, and S. Sami, "Coupled tensor decomposition: a step towards robust components," in *24th European Signal Processing Conf. (EUSIPCO-2016)*, Budapest, Hungary, 29 Aug.–2 Sep. 2016.

[50] T. Yokota and A. Cichocki, "Linked Tucker2 decomposition for flexible multi-block data analysis," in *21st Int'l Conf. Neural Information Process. (ICONIP-2014)*, Kuching, Malaysia, Nov. 2014.

[51] G. Zhou, A. Cichocki, Y. Zhang, and D. P. Mandic, "Group component analysis for multiblock data: Common and individual feature extraction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 11, pp. 2426–2439, Nov. 2016.

[52] G. Zhou, Q. Zhao, Y. Zhang, T. Adalı, S. Xie, and A. Cichocki, "Linked component analysis from matrices to high-order tensors: Applications to biomedical data," *Proc. IEEE*, vol. 104, no. 2, pp. 310–331, Feb. 2016.

[53] F. Sedighin, M. Zadeh, B. Rivet, and C. Jutten, "A new algorithm for multimodal soft coupling," in *13th Int'l Conf. Latent Variable Analysis and Signal Separation (LVA/ICA-2017)*, Grenoble, France, Feb. 2017.

[54] A. K. Smilde, I. Måge, T. Næs, T. Hankemeier, M. A. Lips, H. A. L. Kiers, E. Acar, and R. Bro, "Common and distinct components in data fusion," *J. Chemometrics*, 2017. [Online]. Available: https://doi.org/10.1002/cem.2900

[55] R. Piziak and P. L. Odell, "Full rank factorization of matrices," *Mathematics Magazine*, vol. 72, no. 3, pp. 193–201, Jun. 1999.

[56] S. Van Eyndhoven, B. Hunyadi, L. De Lathauwer, and S. Van Huffel, "Flexible fusion of electroencephalography and functional magnetic resonance imaging: Revealing neural-hemodynamic coupling through structured matrix-tensor factorization," in *25th European Signal Processing Conf. (EUSIPCO-2017)*, Kos, Greece, 28 Aug.–2 Sep. 2017.

[57] U. Lorenzo-Seva and J. M. F. ten Berge, "Tucker's congruence coefficient as a meaningful index of factor similarity," *Methodology*, vol. 2, no. 2, pp. 57–64, 2006.