CANONICAL POLYADIC DECOMPOSITION OF A TENSOR THAT HAS MISSING FIBERS: A MONOMIAL FACTORIZATION APPROACH

Mikael Sørensen and Nicholas D. Sidiropoulos *

University of Virginia
{ms8tz, nikos}@virginia.edu

Lieven De Lathauwer †

KU Leuven
Lieven.DeLathauwer@kuleuven.be

ABSTRACT

The Canonical Polyadic Decomposition (CPD) is one of the most basic tensor models used in signal processing and machine learning. Despite its wide applicability, identifiability conditions and algorithms for CPD in cases where the tensor is incomplete are lagging behind its practical use. We first present a tensor-based framework for bilinear factorizations subject to monomial constrains, called monomial factorizations. Next, we explain that the CPD of a tensor that has missing fibers can be interpreted as a monomial factorization problem. Finally, using the monomial factorization interpretation, we show that CPD recovery conditions can be obtained that only rely on the observed fibers of the tensor.

Index Terms— Tensor, canonical polyadic decomposition, monomial, missing data. subsampling.

1. INTRODUCTION

The CPD of a tensor has found many applications in signal processing and machine learning; see [1] and references therein. In many applications the tensor is incomplete due to missing observations, corrupt data or subsampling. It has also been recognized that incomplete tensors, obtained by random sampling, play an important role in the context of large scale CPD computations [2, 3, 4]. Several optimization-based methods to compute the CPD of an incomplete tensor have been proposed (e.g., [5, 6, 2, 7, 1]), but lack a theoretical foundation. As an alternative to random sampling, we have recently proposed a structured subsampling approach in which only a subset of the fibers of the tensor are considered [8]. In particular, we showed that if the fibers in one of the modes of the tensor are sampled in a structured way, then CPD recovery can be ensured, despite missing data. We also mention that a tensor subsampling method that relies on regular sampling has been proposed in [9]. In this paper we first provide in Section 2 a new monomial factorization approach to bilinear factorizations exhibiting monomial structure in one mode. Based on this framework, in Section 3 we will extend the tensor subsampling framework in [8] to cases where fibers in several modes of the tensor are considered. This result will explain that by considering fibers in several modes, more relaxed CPD recovery conditions can be obtained compared to results that only consider fibers in a single mode. Before diving into the details, a brief review of CPD is provided next.

1.1. Canonical Polyadic Decomposition (CPD)

Consider the CPD of the tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$:

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{s}_r = \sum_{r=1}^{R} \mathbf{G}^{(r)} \circ \mathbf{s}_r, \tag{1}$$

where R denotes the rank of $\mathcal{X}$, $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_R] \in \mathbb{C}^{I\times R}$, $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_R] \in \mathbb{C}^{J\times R}$, $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_R] \in \mathbb{C}^{K\times R}$ are the CPD factor matrices of $\mathcal{X}$ and '$\circ$' denotes the outer product, e.g., $(\mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{s}_r)_{ijk} = a_{ir} b_{jr} s_{kr}$. Note that $\mathbf{G}^{(r)} = \mathbf{a}_r \mathbf{b}_r^T$ is a rank-1 matrix. This fact will be exploited in Section 3. A key feature of the CPD that will also be used in Section 3 is that it is unique under mild conditions, i.e., $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{S}$ are unique (up to intrinsic column scaling and permutation ambiguities); see [10, 11, 12, 13, 14, 15, 16, 17, 18, 19] and references therein. In this paper we will consider the following three matrix representations of (1):

$$\mathbf{X}^{(1)} = (\mathbf{B} \odot \mathbf{S})\, \mathbf{A}^T \in \mathbb{C}^{JK\times I}, \tag{2}$$

$$\mathbf{X}^{(2)} = (\mathbf{A} \odot \mathbf{S})\, \mathbf{B}^T \in \mathbb{C}^{IK\times J}, \tag{3}$$

$$\mathbf{X}^{(3)} = (\mathbf{A} \odot \mathbf{B})\, \mathbf{S}^T \in \mathbb{C}^{IJ\times K}, \tag{4}$$

where '$\odot$' denotes the Khatri-Rao (columnwise Kronecker) product and '$(\cdot)^T$' denotes the transpose. The rows of $\mathbf{X}^{(1)}$ correspond to the mode-1 fibers $\{\mathbf{x}_{\bullet jk}\}$ of $\mathcal{X}$, defined as $(\mathbf{x}_{\bullet jk})_{ijk} = x_{ijk}$. Likewise, $\mathbf{X}^{(2)}$ and $\mathbf{X}^{(3)}$ are obtained by stacking the mode-2 fibers $\{\mathbf{x}_{i\bullet k}\}$ and mode-3 fibers $\{\mathbf{x}_{ij\bullet}\}$ of $\mathcal{X}$, defined as $(\mathbf{x}_{i\bullet k})_{ijk} = x_{ijk}$ and $(\mathbf{x}_{ij\bullet})_{ijk} = x_{ijk}$, respectively.

*This work was supported by NSF ECCS-1608961 and IIS-1704074.

†Research supported by: (1) Research Council KU Leuven: C1 project C16/15/059-nD, (2) F.W.O.: the Fund for Scientific Research–Flanders (Belgium), EOS Project no 30468160 (SeLMA), (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.

2. MONOMIAL FACTORIZATION

Consider bilinear factorizations of the form

$$\mathbf{X} = \mathbf{A}\mathbf{S}^T \in \mathbb{C}^{I\times K}, \quad (5)$$

in which $\mathbf{S} \in \mathbb{C}^{K\times R}$ has full column rank and the columns of $\mathbf{A} \in \mathbb{C}^{I\times R}$ satisfy N monomial relations of the form

$$a_{p_{1,n},r}\cdots a_{p_{L,n},r} - a_{s_{1,n},r}\cdots a_{s_{L,n},r} = 0, \quad (6)$$

where $a_{m,r}$ denotes the mth entry of the rth column of $\mathbf{A}$ and L denotes the degree of the monomials in (6).

2.1. Block term decomposition approach

Define the vectors $\mathbf{b}_r^{(n)} = [a_{p_{1,n},r} \dots a_{p_{L,n},r}]^T \in \mathbb{C}^L$ and $\mathbf{c}_r^{(n)} = [a_{s_{1,n},r} \dots a_{s_{L,n},r}]^T \in \mathbb{C}^L$. Then relation (6) can be related to the matrix $\mathbf{A}_L(\mathbf{b}_r^{(n)}, \mathbf{c}_r^{(n)}) \in \mathbb{C}^{L\times L}$:

$$\mathbf{A}_L(\mathbf{b}_r^{(n)}, \mathbf{c}_r^{(n)}) = \begin{bmatrix} b_{1,r}^{(n)} & 0 & \cdots & 0 & (-1)^L \cdot c_{1,r}^{(n)} \\ c_{2,r}^{(n)} & b_{2,r}^{(n)} & \ddots & & 0 \\ 0 & c_{3,r}^{(n)} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_{L,r}^{(n)} & b_{L,r}^{(n)} \end{bmatrix}. \quad (7)$$

It can be verified that if $\prod_{l=1}^{L} a_{p_{l,n},r} \neq 0$ or $\prod_{l=1}^{L} a_{s_{l,n},r} \neq 0$, then $\mathbf{A}_L(\mathbf{b}_r^{(n)}, \mathbf{c}_r^{(n)})$ is a rank-$(L-1)$ matrix when the monomial relation (6) is satisfied. Using the mapping (7) and the fact that (5) is bilinear, we obtain

$$\begin{aligned}\mathbf{Y}^{(n)} &:= [\mathrm{vec}(\mathbf{A}_L(\mathbf{u}_1^{(n)}, \mathbf{v}_1^{(n)})) \dots \mathrm{vec}(\mathbf{A}_L(\mathbf{u}_K^{(n)}, \mathbf{v}_K^{(n)}))] \\ &= \mathbf{M}^{(n)}\mathbf{S}^T \in \mathbb{C}^{L^2\times K}, \quad n \in \{1,\dots,N\}, \quad (8)\end{aligned}$$

where $\mathrm{vec}(\cdot)$ denotes the column vector obtained by stacking the columns of its input matrix, $\mathbf{u}_k^{(n)} = [x_{p_{1,n},k} \dots x_{p_{L,n},k}]^T \in \mathbb{C}^L$, $\mathbf{v}_k^{(n)} = [x_{s_{1,n},k} \dots x_{s_{L,n},k}]^T \in \mathbb{C}^L$ and $\mathbf{M}^{(n)} = [\mathrm{vec}(\mathbf{A}_L(\mathbf{b}_1^{(n)}, \mathbf{c}_1^{(n)})) \dots \mathrm{vec}(\mathbf{A}_L(\mathbf{b}_R^{(n)}, \mathbf{c}_R^{(n)}))] \in \mathbb{C}^{L^2\times R}$. The key observation is that each equation in (8) corresponds to a Block Term Decomposition (BTD) [20], in which the columns of $\mathbf{M}^{(n)}$ correspond to vectorized rank-$(L-1)$ matrices $\mathbf{m}_r^{(n)} = \mathbf{A}_L(\mathbf{b}_r^{(n)}, \mathbf{c}_r^{(n)})$. Overall, the collection of all N equations in (8) yields a coupled BTD [21, 22]. Consequently, uniqueness conditions and algorithms developed for coupled BTD can also be used to solve the monomial factorization problem (5).

2.2. Null space approach

As an alternative to the previously discussed coupled BTD approach to monomial factorizations, a null space approach will briefly be discussed. In short, the monomial factorization problem (6) can, under certain conditions (not discussed here), be reduced to a CPD problem that in the exact (noiseless) case can be solved by means of an eigenvalue decomposition (EVD). Let $\mathbf{e}_i^{(I)} \in \mathbb{C}^I$ denote a unit vector with unit entry at the ith position and zeros elsewhere. Furthermore, let $\mathbf{W} = [\mathbf{w}_1,\dots,\mathbf{w}_R] = \mathbf{S}^{-T}$. Since $b_{l,r}^{(n)} = \mathbf{e}_{p_{l,n}}^{(I)T}\mathbf{X}\mathbf{w}_r$ and $c_{l,r}^{(n)} = \mathbf{e}_{s_{l,n}}^{(I)T}\mathbf{X}\mathbf{w}_r$, we conclude from (6) that

$$\begin{aligned}&a_{p_{1,n}}\cdots a_{p_{L,n}} - a_{s_{1,n}}\cdots a_{s_{L,n}} = \\ &(\mathbf{e}_{p_{1,n}}^{(I)T}\mathbf{X}\mathbf{w}_r)\cdots(\mathbf{e}_{p_{L,n}}^{(I)T}\mathbf{X}\mathbf{w}_r) - (\mathbf{e}_{s_{1,n}}^{(I)T}\mathbf{X}\mathbf{w}_r)\cdots(\mathbf{e}_{s_{L,n}}^{(I)T}\mathbf{X}\mathbf{w}_r) = \\ &\mathbf{p}_L^{(n)T}\cdot(\mathbf{w}_r\otimes\cdots\otimes\mathbf{w}_r) = 0, \quad r\in\{1,\dots,R\}, \quad (9)\end{aligned}$$

where '$\otimes$' denotes the Kronecker product and $\mathbf{p}_L^{(n)} := \mathbf{X}^T\mathbf{e}_{p_{1,n}}^{(I)} \otimes\cdots\otimes \mathbf{X}^T\mathbf{e}_{p_{L,n}}^{(I)} - \mathbf{X}^T\mathbf{e}_{s_{1,n}}^{(I)} \otimes\cdots\otimes \mathbf{X}^T\mathbf{e}_{s_{L,n}}^{(I)} \in \mathbb{C}^{R^L}$. Stacking yields

$$\mathbf{P}^{(N,L)}\cdot(\mathbf{w}_r\otimes\cdots\otimes\mathbf{w}_r) = \mathbf{0}, \quad r\in\{1,\dots,R\}, \quad (10)$$

where $\mathbf{P}^{(N,L)} = [\mathbf{p}_L^{(1)},\dots,\mathbf{p}_L^{(N)}]^T \in \mathbb{C}^{N\times R^L}$. From (10) we know that there exist at least R linearly independent vectors $\{\mathbf{w}_r\otimes\cdots\otimes\mathbf{w}_r\}$, each with property $\mathbf{w}_r\otimes\cdots\otimes\mathbf{w}_r \in \ker(\mathbf{P}^{(N,L)})\cap\pi_S^{(L)}$, where $\pi_S^{(L)}$ denotes the subspace of vectorized R^L symmetric tensors. Thus, if the dimension of $\ker(\mathbf{P}^{(N,L)})\cap\pi_S^{(L)}$ is minimal (i.e., R) and the columns of $\mathbf{R}\in\mathbb{C}^{N\times R}$ form a basis for $\ker(\mathbf{P}^{(N,L)})\cap\pi_S^{(L)}$, then there exists a nonsingular change-of-basis matrix $\mathbf{F}\in\mathbb{C}^{R\times R}$ such that

$$\mathbf{R} = (\underbrace{\mathbf{W}\odot\cdots\odot\mathbf{W}}_{L\text{ times}})\mathbf{F}^T. \quad (11)$$

Clearly, (11) corresponds to an $(L+1)$-th order tensor $\mathcal{R} = \sum_{r=1}^{R}\mathbf{w}_r\circ\cdots\mathbf{w}_r\circ\mathbf{f}_r$, whose CPD is unique and can be computed via an EVD [19].

3. CPD OF TENSOR THAT HAS MISSING FIBERS

Consider the incomplete version of the tensor (1):

$$\mathcal{Y} = \mathcal{D} * \mathcal{X} = \mathcal{D} * \left(\sum_{r=1}^{R}\mathbf{a}_r\circ\mathbf{b}_r\circ\mathbf{s}_r\right) \in \mathbb{C}^{I\times J\times K}, \quad (12)$$

where '$*$' denotes the Hadamard (element-wise) product, $y_{ijk} = (\mathcal{D}*\mathcal{X})_{ijk} = d_{ijk}x_{ijk}$ and entry d_{ijk} of $\mathcal{D}\in\{0,1\}^{I\times J\times K}$ is equal to one if x_{ijk} is observed and zero otherwise. In this section we will demonstrate how the monomial factorization framework can be used to obtain uniqueness conditions in cases where $\mathcal{D}$ is structured. More precisely, we consider the case where the tensor $\mathcal{X}$ has missing fibers. Let $\mathbf{D}^{(1)}\in\{0,1\}^{J\times K}$ denote the mode-1 fiber observation matrix in which $d_{jk}^{(1)} = 1$ if fiber $\mathbf{x}_{\bullet jk}$ is observed. Likewise, let $\mathbf{D}^{(2)}\in\{0,1\}^{I\times K}$ and $\mathbf{D}^{(3)}\in\{0,1\}^{I\times J}$ denote the mode-2 and mode-3 fiber observation matrices in which $d_{ik}^{(2)} = 1$ if fiber $\mathbf{x}_{i\bullet k}$ is observed and $d_{ij}^{(3)} = 1$ if fiber $\mathbf{x}_{ij\bullet}$ is observed. The missing fiber versions of (2)–(4) are given by

$$\mathbf{Y}^{(1)} = \mathrm{Diag}(\mathrm{vec}(\mathbf{D}^{(1)T}))\mathbf{X}^{(1)} \in \mathbb{C}^{JK\times I}, \quad (13)$$

$$\mathbf{Y}^{(2)} = \mathrm{Diag}(\mathrm{vec}(\mathbf{D}^{(2)T}))\mathbf{X}^{(2)} \in \mathbb{C}^{IK\times J}, \quad (14)$$

$$\mathbf{Y}^{(3)} = \mathrm{Diag}(\mathrm{vec}(\mathbf{D}^{(3)T}))\mathbf{X}^{(3)} \in \mathbb{C}^{IJ\times K}, \quad (15)$$

where $\text{Diag}(\text{vec}(\mathbf{D}^{(n)T}))$ denotes a diagonal matrix that holds the vector $\text{vec}(\mathbf{D}^{(n)T})$ on its diagonal. In [8] we considered the problem of finding the CPD from a subset of observable mode-3 fibers in $\mathbf{Y}^{(3)}$ as briefly reviewed in Section 3.1. Using the monomial factorization approach, we will in Section 3.2 explain that joint exploitation of observable fibers in several modes can lead to improved uniqueness conditions.

3.1. Exploiting fibers in a single mode

Assume that we observe F_3 mode-3 fibers, i.e., $\mathbf{Y}^{(3)} \neq \mathbf{0}$ and $\mathbf{D}^{(3)}$ contains F_3 nonzero entries. As in ordinary CPD, we can exploit the rank-1 property of $\mathbf{G}^{(r)}$ in (1). More precisely, j_1-th column of $\mathbf{G}^{(r)}$ is proportional to its j_2-th column, i.e., $\mathbf{a}_r b_{j_1,r} \propto \mathbf{a}_r b_{j_2,r}$. This property can be expressed in terms of the monomial relation

$$\begin{vmatrix} g^{(r)}_{i_1 j_1} & g^{(r)}_{i_1 j_2} \\ g^{(r)}_{i_2 j_1} & g^{(r)}_{i_2 j_2} \end{vmatrix} = g^{(r)}_{i_1 j_1} g^{(r)}_{i_2 j_2} - g^{(r)}_{i_2 j_1} g^{(r)}_{i_1 j_2} = 0, \qquad (16)$$

where '$|\cdot|$' denotes the determinant. We will now make use of the monomial factorization framework. The combination of (9) with $[a_{p_{1,n}}, a_{p_{2,n}}, a_{s_{1,n}}, a_{s_{2,n}}] = [g^{(r)}_{i_1 j_1}, g^{(r)}_{i_2 j_2}, g^{(r)}_{i_2 j_1}, g^{(r)}_{i_1 j_2}]$ and (16) yields

$$\begin{vmatrix} (\mathbf{e}^{(I)}_{i_1} \otimes \mathbf{e}^{(J)}_{j_1})^T \mathbf{Y}^{(3)} \mathbf{w}_r & (\mathbf{e}^{(I)}_{i_1} \otimes \mathbf{e}^{(J)}_{j_2})^T \mathbf{Y}^{(3)} \mathbf{w}_r \\ (\mathbf{e}^{(I)}_{i_2} \otimes \mathbf{e}^{(J)}_{j_1})^T \mathbf{Y}^{(3)} \mathbf{w}_r & (\mathbf{e}^{(I)}_{i_2} \otimes \mathbf{e}^{(J)}_{j_2})^T \mathbf{Y}^{(3)} \mathbf{w}_r \end{vmatrix}$$
$$= \mathbf{p}^{(n)} (\mathbf{w}_r \otimes \mathbf{w}_r) = 0, \qquad (17)$$

where $\mathbf{p}^{(n)} = (\mathbf{e}^{(I)}_{i_1} \otimes \mathbf{e}^{(J)}_{j_1})^T \mathbf{Y}^{(3)} \otimes (\mathbf{e}^{(I)}_{i_2} \otimes \mathbf{e}^{(J)}_{j_2})^T \mathbf{Y}^{(3)} - (\mathbf{e}^{(I)}_{i_2} \otimes \mathbf{e}^{(J)}_{j_1})^T \mathbf{Y}^{(3)} \otimes (\mathbf{e}^{(I)}_{i_1} \otimes \mathbf{e}^{(J)}_{j_2})^T \mathbf{Y}^{(3)}$, and the superscript '$n$' in the row-vector $\mathbf{p}^{(n)} \in \mathbb{C}^{1 \times R^2}$ takes all four subscripts i_1, i_2, j_1 and j_2 into account. Define

$$\mathbf{\Phi} = \Big\{ (i_1, i_2, j_1, j_2) \mid d^{(3)}_{i_1 j_1} = d^{(3)}_{i_2 j_1} = d^{(3)}_{i_1 j_2} = d^{(3)}_{i_2 j_2} = 1,$$
$$1 \leq i_1 < i_2 \leq I,\ 1 \leq j_1 < j_2 \leq J \Big\}. \qquad (18)$$

In words, $\mathbf{\Phi}$ contains all quadruples (i_1, i_2, j_1, j_2) from which a monomial relation of the form (16) can be constructed, given only $\mathbf{Y}^{(3)}$. Stacking yields (cf. Eq. (10)):

$$\mathbf{P}^{(N_3,2)} (\mathbf{w}_r \otimes \mathbf{w}_r) = \mathbf{0}, \qquad (19)$$

where $\mathbf{P}^{(N_3,2)} = [\mathbf{p}^{(1)T}, \ldots, \mathbf{p}^{(N_3)T}]^T \in \mathbb{C}^{N_3 \times R^2}$ and N_3 denotes the number of elements in (18). Hence, if the dimension of $\ker(\mathbf{P}^{(N_3,2)}) \cap \pi^{(2)}_S$ is minimal (i.e., R), then $\mathbf{W} = \mathbf{S}^{-T}$ is unique (up to intrinsic ambiguities). This also means that $\mathbf{S}$ and $\mathbf{Z} = \mathbf{Y}^{(3)} \mathbf{W} = \text{Diag}(\text{vec}(\mathbf{D}^{(3)T}))(\mathbf{A} \odot \mathbf{B})$ are unique. The remaining matrices $\mathbf{A}$ and $\mathbf{B}$ can now be obtained from $\text{Diag}(\text{vec}(\mathbf{D}^{(3)T}))(\mathbf{A} \odot \mathbf{B})$ via rank-1 matrix completion. In more detail, observe that the r-th column of $\mathbf{Z}$ can be reshaped into an incomplete $(I \times J)$ rank-1 matrix

$$\mathbf{Z}^{(r)} = \mathbf{D}^{(3)} * (\mathbf{a}_r \mathbf{b}_r^T), \quad r \in \{1, \ldots, R\}. \qquad (20)$$

The incomplete matrix $\mathbf{Z}^{(r)}$ can be interpreted as a bipartite graph, denoted by $G^{(r)}$. The two groups of vertices associated with $G^{(r)}$ are the row indices $1, \ldots, I$ and the column indices $1, \ldots, J$. Let $\mathcal{E}^{(r)} = \{(i,j) \mid z^{(r)}_{i,j} \neq 0\}$ denote the edge set associated with the bipartite graph $G^{(r)}$. If $G^{(r)}$ is connected and has the property

$$\begin{cases} \forall i \in \{1, \ldots, I\},\ \exists j^{'} \in \{1, \ldots, J\} :\ (i, j^{'}) \in \mathcal{E}^{(r)}, \\ \forall j \in \{1, \ldots, J\},\ \exists i^{'} \in \{1, \ldots, I\} :\ (i^{'}, j) \in \mathcal{E}^{(r)}, \end{cases} \qquad (21)$$

then the vectors $\mathbf{a}_r$ and $\mathbf{b}_r$ can be obtained from (20) via a rank-1 factorization of the incomplete matrix $\mathbf{Z}^{(r)}$; see [8] for further details. To summarize, if

$$\begin{cases} \mathbf{S} \text{ has full column rank}, \\ \ker(\mathbf{P}^{(N_3,2)}) \cap \pi^{(2)}_S \text{ is an } R\text{-dimensional subspace}, \\ G^{(r)} \text{ is connected and with property (21)}, \forall r \in \{1, \ldots, R\}, \end{cases} \qquad (22)$$

then the rank of $\mathcal{X}$ is R and the CPD of $\mathcal{X}$ is unique.

Note that the sufficient CPD uniqueness condition (22) is far from necessary and it can easily be improved upon by the use of tensorization methods [19]. In [8] it was explained that as few as $F_3 = I + J + R$ fibers can be sufficient for CPD uniqueness.

3.2. Exploiting fibers in several modes

Using the monomial factorization framework in Section 2, we will now demonstrate that by jointly taking the observable fibers in $\mathbf{Y}^{(2)}$ and $\mathbf{Y}^{(3)}$ into account, improved uniqueness conditions can be obtained. Let $\mathcal{X}^{(2)} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{s}_r \circ \mathbf{b}_r = \sum_{r=1}^{R} \mathbf{H}^{(r)} \circ \mathbf{b}_r \in \mathbb{C}^{I \times K \times J}$ denote the tensorized version of (3), in which $\mathbf{H}^{(r)} = \mathbf{a}_r \mathbf{s}_r^T \in \mathbb{C}^{I \times K}$. Observe that the j-th column of $\mathbf{G}^{(r)}$ in (1) is proportional to the k-th column of $\mathbf{H}^{(r)}$, i.e., $\mathbf{a}_r b_{j,r} \propto \mathbf{a}_r s_{k,r}$. This property can be expressed in terms of the monomial relation:[1]

$$\begin{vmatrix} g^{(r)}_{i_1 k} & h^{(r)}_{i_1 j} \\ g^{(r)}_{i_2 k} & h^{(r)}_{i_2 j} \end{vmatrix} = g^{(r)}_{i_1 k} h^{(r)}_{i_2 j} - g^{(r)}_{i_2 k} h^{(r)}_{i_1 j} = 0, \qquad (23)$$

where $1 \leq i_1 < i_2 \leq I$, $1 \leq j \leq J$ and $1 \leq k \leq K$. Let us assume that the CPD factor matrices $\mathbf{B}$ and $\mathbf{S}$ have full column rank $(J, K \geq R)$. W.l.o.g. we can now assume that $\mathbf{B}$ and $\mathbf{S}$ are nonsingular, i.e., $J = K = R$. (Note that if the matrices $\mathbf{B}$ and $\mathbf{S}$ do not have full column rank, then higher-order minors can be considered [16].) Similar to (9) and (16), we are now looking for nonsingular matrices $\mathbf{W} = \mathbf{S}^{-T}$ and $\mathbf{V} = \mathbf{B}^{-T}$ so that

$$\begin{cases} \mathbf{Y}^{(3)} \mathbf{w}_r = \text{Diag}(\text{vec}(\mathbf{D}^{(3)T}))(\mathbf{a}_r \otimes \mathbf{b}_r), \\ \mathbf{Y}^{(2)} \mathbf{v}_r = \text{Diag}(\text{vec}(\mathbf{D}^{(2)T}))(\mathbf{a}_r \otimes \mathbf{s}_r). \end{cases} \qquad (24)$$

[1] We note in passing that this approach is similar to double coupled CPD [23]. However, there are also notable differences. First, in the proposed monomial factorization approach, we can exploit both the rank-1 structures within $\mathbf{Y}^{(n)}$ (e.g. via (16)) and the rank-1 structures between $\mathbf{Y}^{(m)}$ and $\mathbf{Y}^{(n)}$ (e.g. via (23)) whereas in double coupled CPD only the latter is exploited. Second, the monomial factorization formulation is different, e.g., it reduces the problem to a CPD problem. Third, we also consider the incomplete case where data is missing.

The combination of (23) and (24) yields

$$\begin{vmatrix} (\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_j^{(J)})^T \mathbf{Y}^{(3)} \mathbf{w}_r & (\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_k^{(K)})^T \mathbf{Y}^{(2)} \mathbf{v}_r \\ (\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_j^{(J)})^T \mathbf{Y}^{(3)} \mathbf{w}_r & (\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_k^{(K)})^T \mathbf{Y}^{(2)} \mathbf{v}_r \end{vmatrix} = \mathbf{q}^{(n)} (\mathbf{w}_r \otimes \mathbf{v}_r) = 0, \quad (25)$$

where $\mathbf{q}^{(n)} = (\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_j^{(J)})^T \mathbf{Y}^{(3)} \otimes (\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_k^{(K)})^T \mathbf{Y}^{(2)} - (\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_j^{(J)})^T \mathbf{Y}^{(3)} \otimes (\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_k^{(K)})^T \mathbf{Y}^{(2)}$, and the superscript '$n$' in the row-vector $\mathbf{q}^{(n)} \in \mathbb{C}^{1 \times R^2}$ takes all the subscripts i_1, i_2, j and k into account. Define

$$\Xi_{(1)}^{(2,3)} = \{(i_1, i_2, j, k) \mid d_{i_1 k}^{(2)} = d_{i_2 k}^{(2)} = d_{i_1 j}^{(3)} = d_{i_2 j}^{(3)} = 1, \; 1 \le i_1 < i_2 \le I, \; 1 \le j \le J, \; 1 \le k \le K\}. \quad (26)$$

In words, $\Xi_{(1)}^{(2,3)}$ contains all quadruples (i_1, i_2, j, k) from which a monomial relation of the form (23) can be constructed, given only $\mathbf{Y}^{(2)}$ and $\mathbf{Y}^{(3)}$. Stacking yields (cf. (10)):

$$\mathbf{Q}^{(N_{2,3},2,3)} (\mathbf{w}_r \otimes \mathbf{v}_r) = \mathbf{0}, \quad (27)$$

where $\mathbf{Q}^{(N_{2,3},2,3)} = [\mathbf{q}^{(1)T}, \ldots, \mathbf{q}^{(N_p)T}]^T \in \mathbb{C}^{(N_{2,3},2,3) \times R^2}$ in which $N_{2,3}$ denotes the number of elements in (26). Based on (27) a uniqueness condition[2] that takes into account that fibers in two modes are observed can immediately be derived. However, a more relaxed uniqueness condition can be obtained by combining (19) and (27). This approach jointly exploits the rank-1 structure within $\mathbf{Y}^{(3)}$ and the rank-1 structure between $\mathbf{Y}^{(2)}$ and $\mathbf{Y}^{(3)}$. (The inclusion of the rank-1 structure within $\mathbf{Y}^{(2)}$ can be done in a similar manner.) Observe that

$$\begin{cases} (\mathbf{I}_R \otimes \mathbf{P}^{(N,2)}) \mathbf{\Pi}^T (\mathbf{w}_r \otimes \mathbf{w}_r \otimes \mathbf{v}_r) = \mathbf{0}, \\ (\mathbf{I}_R \otimes \mathbf{Q}^{(N_{2,3},2,3)}) (\mathbf{w}_r \otimes \mathbf{w}_r \otimes \mathbf{v}_r) = \mathbf{0}, \end{cases} \quad (28)$$

where $\mathbf{\Pi} \in \mathbb{C}^{R^3 \times R^3}$ is the permutation matrix with property $\mathbf{\Pi} (\mathbf{w}_r \otimes \mathbf{v}_r \otimes \mathbf{w}_r) = \mathbf{w}_r \otimes \mathbf{w}_r \otimes \mathbf{v}_r$. From (28) it is clear that (19) and (27) can be combined as follows

$$\mathbf{\Gamma} (\mathbf{w}_r \otimes \mathbf{w}_r \otimes \mathbf{v}_r) = \mathbf{0}, \quad (29)$$

where $\mathbf{\Gamma} = \begin{bmatrix} (\mathbf{I}_R \otimes \mathbf{P}^{(N_3,2)}) \mathbf{\Pi}^T \\ \mathbf{I}_R \otimes \mathbf{Q}^{(N_{2,3},2,3)} \end{bmatrix} \in \mathbb{C}^{(N_3 + N_{2,3})R \times R^3}$. From (29) it is in turn clear that if the subspace $\ker(\mathbf{\Gamma}) \cap \pi_S^{(2)} \times \mathbb{C}^R$ is R-dimensional (which is minimal since $\mathbf{\Gamma} (\mathbf{W} \odot \mathbf{W} \odot \mathbf{V}) = \mathbf{0}$), then $\mathbf{S} = \mathbf{W}^{-T}$ and $\mathbf{B} = \mathbf{V}^{-T}$ can be obtained via a CPD. More precisely, let the columns of the matrix $\mathbf{R} \in \mathbb{C}^{R^3 \times R}$ constitute a basis for $\ker(\mathbf{\Gamma}) \cap \pi_S^{(2)} \times \mathbb{C}^R$. Then there exists a nonsingular change-of-basis matrix $\mathbf{F} \in \mathbb{C}^{R \times R}$ such that

$$\mathbf{R} = (\mathbf{W} \odot \mathbf{W} \odot \mathbf{V}) \mathbf{F}^T. \quad (30)$$

Clearly, (30) corresponds to a (partially symmetric) tensor $\mathcal{R} = \sum_{r=1}^R \mathbf{w}_r \circ \mathbf{w}_r \circ \mathbf{v}_r \circ \mathbf{f}_r$, whose CPD is unique. Finally, $\mathbf{A}$ follows from (12). Briefly, let $\mathbf{d}_i \in \{0,1\}^{JK}$ with $(\mathbf{d}_i)_p = 1$ if entry $(\mathbf{x}^{(1)})_{pi}$ in (2) is observed and zero otherwise. Then the i-th column $\mathbf{y}_i^{(1)}$ of $\mathbf{Y}^{(1)}$ admits the factorization

$$\mathbf{y}_i^{(1)} = \mathbf{d}_i * ((\mathbf{B} \odot \mathbf{S}) \mathbf{A}^T \mathbf{e}_i^{(I)}) = \text{Diag}(\mathbf{d}_i)(\mathbf{B} \odot \mathbf{S}) \mathbf{A}^T \mathbf{e}_i^{(I)}, \quad (31)$$

where $\text{Diag}(\mathbf{d}_i)$ denotes the diagonal matrix that holds $\mathbf{d}_i$ on its diagonal. Hence, if the matrix $\text{Diag}(\mathbf{d}_i)(\mathbf{B} \odot \mathbf{S})$ has full column rank, which is necessary for CPD uniqueness, then the i-th row of $\mathbf{A}$ is unique, i.e., $\mathbf{e}_i^{(I)T} \mathbf{A} = (\text{Diag}(\mathbf{d}_i)(\mathbf{B} \odot \mathbf{S}))^\dagger \mathbf{y}_i^{(1)}$, where $(\cdot)^\dagger$ denotes the left-inverse. To summarize, if

$$\begin{cases} \mathbf{B} \text{ and } \mathbf{S} \text{ have full column rank,} \\ \ker(\mathbf{\Gamma}) \cap \pi_S^{(2)} \times \mathbb{C}^R \text{ is an } R\text{-dimensional subspace,} \\ \text{Diag}(\mathbf{d}_i)(\mathbf{B} \odot \mathbf{S}) \text{ has full column rank } \forall i \in \{1, \ldots I\}, \end{cases} \quad (32)$$

then the rank of $\mathcal{X}$ is R and the CPD of $\mathcal{X}$ is unique.

Comparing the sufficient CPD uniqueness condition (22) (and other related conditions that only rely on fibers in a single mode) with the sufficient CPD uniqueness condition (32), it is clear that the connectivity constraint on $G^{(r)}$ in the former condition has been dropped. In other words, by jointly considering fibers in at least two modes, the structure of the fiber observation matrix $\mathbf{D}^{(3)}$ in (15) can be relaxed.

Another notable difference between a condition that only relies on fibers in a single mode and a condition that exploits fibers in several modes is that the latter can lead to a relaxed bound on R. As an example, consider the case where $I = 4$, $J = K = R$, $F_2 = 16$ mode-2 fibers $\{\mathbf{x}_{i \bullet k}\}_{i,k \in \{1,2,3,4\}}$ are observed and $F_3 = 16$ mode-3 fibers $\{\mathbf{x}_{ij \bullet}\}_{i,j \in \{1,2,3,4\}}$ are observed. An immediate way to establish CPD uniqueness is to use a single-mode fiber condition. For example, we know from [8] that if $R \le F_3 - I - J = 8$, then $\mathbf{S}$ is expected to be unique (up to column scaling and permutation ambiguities). If $G^{(r)}$ is connected and with property (21) for all $r \in \{1, \ldots, R\}$, then CPD uniqueness can be ensured from (20). The very basic two-mode fibers observation condition (32) relaxes the bound to $R \le 11$.

To summarize, the joint exploitation of fibers in two modes of a tensor for CPD recovery is a novel and nontrivial extension of existing results based on fibers in a single mode (e.g., [8]). The extension to joint exploitation of observable fibers in all three modes is analogous. However, due to space considerations, this extension will first be discussed in a full length journal paper version.

4. CONCLUSION

We first presented a new tensor decomposition approach to bilinear factorization enjoying monomial constraints. Based on this approach we developed uniqueness conditions for the CPD of a tensor in which only a subset of its fibers are observed. In particular, we showed that by jointly exploiting fibers in several modes, more relaxed conditions on the rank and the missing data pattern of the tensor compared to [8] can be obtained without sacrificing the uniqueness of its CPD.

[2] If only relation (27) is exploited, then in the subsequent uniqueness condition (32), the assumption that $\ker(\mathbf{\Gamma}) \cap \pi_S^{(2)} \times \mathbb{C}^R$ is R-dimensional has to be replaced by the more restrictive assumption that $\ker(\mathbf{Q}^{(N_{2,3},2,3)})$ is R-dimensional.

5. REFERENCES

[1] 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 Processing*, vol. 65, no. 13, pp. 3551–3582, July 2017.

[2] N. Vervliet, O. Debals, L. Sorber, and L. De Lathauwer, "Breaking the curse of dimensionality using decompositions of incomplete tensors," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 71–79, Jan. 2014.

[3] N. D. Sidiropoulos, E. E. Papalexakis, and C. Faloutsos, "PArallel RAndomly COMPressed Cubes: A Scalable Distributed Architecture for Big Tensor Decomposition," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 57–70, Sept. 2014.

[4] N. Vervliet and L. De Lathauwer, "A randomized block sampling approach to canonical polyadic decomposition of large-scale tensors," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 2, pp. 284–295, 2016.

[5] G. Tomasi and R. Bro, "PARAFAC and missing values," *Chemometr. Intell. Lab.*, vol. 75, no. 2, pp. 163–180, 2005.

[6] E. Acar, D. M. Dunlavy, T. G. Kolda, and M. Mørup, "Scalable tensor factorizations for incomplete data," *Chemometr. Intell. Lab.*, vol. 106, no. 1, pp. 41–56, 2011.

[7] T. Maehara, K. Hayashi, and K.-.I Kawarabayash, "Expected tensor decomposition with stochastic gradient descent," in *Proc. AAAI'16, Phoenix, AZ, USA, 2016*.

[8] M. Sørensen and L. De Lathauwer, "Fiber sampling approach to canonical polyadic decomposition and tensor completion," *Submitted to SIAM J. Matrix Anal. Appl.*

[9] C. I. Kanatsoulis, N. D. Sidiropoulos, M. Akcakaya, and X. Fu, "Regular sampling of tensor signals: Theory and application to fMRI," *Submitted to ICASSP 2019*.

[10] 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.

[11] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with applications to arithmetic complexity and statistics," *Linear Algebra and its Applications*, vol. 18, pp. 95–138, 1977.

[12] T. Jiang and N. D. Sidiropoulos, "Kruskal's permutation lemma and the identification of CANDECOMP/PARAFAC and bilinear model with constant modulus constraints," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2625–2636, Sept. 2004.

[13] Lieven De Lathauwer, "A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 28, no. 3, pp. 642–666, 2006.

[14] L. Chiantini and G. Ottaviani, "On generic identifiability of of 3-tensors of small rank," *SIAM J. Matrix Anal. Appl.*, vol. 33, no. 3, pp. 101–8–1037, 2012.

[15] C. Bocci, L. Chiantini, and G. Ottaviani, "Refined methods for the identifiability of tensors," *Annali de Matematica*, vol. 193, no. 6, pp. 1691–1702, Dec. 2014.

[16] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part I: Basic results and uniqueness of one factor matrix," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 855–875, 2013.

[17] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part II: Overall uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 876–903, 2013.

[18] I. Domanov and L. De Lathauwer, "Generic uniqueness conditions for the canonical polyadic decomposition and INDSCAL," *SIAM J. Matrix Anal. Appl.*, vol. 36, no. 4, pp. 1567–1589, 2015.

[19] I. Domanov and L. De Lathauwer, "Canonical polyadic decomposition of third-order tensors: Relaxed uniqueness conditions and algebraic algorithm," *Linear Algebra Appl.*, vol. 513, pp. 342–375, Jan. 2017.

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

[21] 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 J. Matrix Anal. Appl.*, vol. 36, no. 2, pp. 496–522, 2015.

[22] 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 J. Matrix Anal. Appl.*, vol. 36, no. 3, pp. 1015–1045, 2015.

[23] X.-F. Gong, Q.-H. Lin, F.-Y. Cong, and L. De Lathauwer, "Double coupled canonical polyadic decomposition for joint blind source separation," *IEEE Trans. Signal Processing*, vol. 66, pp. 3475–3490, 2018.