UNIQUENESS OF TENSOR TRAIN DECOMPOSITION WITH LINEAR DEPENDENCIES

Yassine Zniyed[1], *Sebastian Miron*[2], *Rémy Boyer*[3], *David Brie*[2]

[1] Laboratoire des Signaux et Systèmes, CentraleSupelec, Gif-Sur-Yvette, France
[2] CRAN, Université de Lorraine, CNRS, Vandœuvre-lès-Nancy, France
[3] Laboratoire CRIStAL, Univeristé de Lille, Villeneuve d'Ascq, France

ABSTRACT

With the increase in measurement/sensing technologies, the collected data are intrinsically multidimensional in a large number of applications. This can be interpreted as a growth of the dimensionality/order of the associated tensor. There exists therefore a crucial need to derive equivalent and alternative models of a high-order tensor as a graph of low-order tensors. In this work we consider a " train " graph, *i.e.*, a Q-order tensor will be represented as a Tensor Train (TT) composed of $Q-2$ 3-order core tensors and two core matrices. In this context, it has been shown that a canonical rank-R CPD model can always be represented exactly by a TT model whose cores are canonical rank-R CPD. This model is called TT-CPD. We generalize this equivalence to the PARALIND model in order to take into account potential linear dependencies in factors. We derive and discuss here uniqueness conditions for the case of the TT-PARALIND model.

Index Terms— Tensor Train, PARALIND, identifiability

1. INTRODUCTION

Canonical Polyadic Decomposition (CPD) [1] is one of the most used tensor decompositions in signal processing. The CPD and its variants are attractive tools due to their ability to decompose tensors into physically interpretable quantities, called factors. Its uniqueness has been studied in several state-of-art articles such as [2, 3, 4]. Uniqueness and compactness are two of the advantages that make the CPD widespread. Indeed, the CPD is usually unique under mild conditions and its storage cost grows linearly with respect to the order. Recently, tensor networks (TNs) [5] have been subject of increasing interest, especially for high-order tensors, allowing more flexible tensor modelling. TNs split high-order ($Q > 3$) tensors into a set of lower-order tensors. Tensor train decomposition (TTD) [6] is one the most compact and simple TNs. Indeed, TTD breaks a high Q-order tensor into a set of Q lower-order tensors, called TT-cores. These TT-cores have orders at most equal to 3. In this sense, TNs are able to break the "curse of dimensionality".

In a recent work [7], an equivalence between the CPD and the TTD was proposed. In fact, it has been shown that a Q-order CPD of rank-R is equivalent to a train of 3-order CPD(s) of rank-R. This result makes it easier to jointly reduce the dimension and estimate the CPD factors using the TT-cores when the original tensor has a high order. Otherwise, when Q is high, the CPD factors estimation becomes a difficult task using ALS-based techniques [8]. At the same time, the existing results on the equivalence between CPD and TTD are based on the assumption that the CPD factor matrices are all full column rank, in which case, estimating the factor matrices from the TT-cores is straightforward. Posteriorly to [7], in the unpublished work [9], another TT-based method has been proposed for CPD factors estimation. This latter only assumes that the TT-cores have a CP decomposition, without specifying the coupling properties between the TT-cores, which leads to a straightforward factor estimation up to same column permutation and scaling indeterminacies whose product equals 1, as in [7, 10]. Dealing with the CPD ambiguities is not required in [7, 10], in contrast to [9]. Moreover, none of these works have discussed the partial identifiability, which is one of the contributions of this work.

In this work, we focus on the case where linear dependencies are present between the columns on the factor matrices leading to high-order PARALIND (PARAllel profiles with LINear Dependences) model [11]. PARALIND is a variant of the CPD with constrained factor/loading matrices, that models a linearly dependent factor $\boldsymbol{P}$ as a product of a full column rank matrix $\tilde{\boldsymbol{P}}$ and an interaction matrix $\boldsymbol{\Phi}$. Matrix $\boldsymbol{\Phi}$ introduces the linear dependency and rank deficiency in $\boldsymbol{P}$. Linear dependencies in factor matrices are of great interest in real scenarios and can be encountered in chemometrics applications [11] or in array signal processing [12], to mention a few. In this work, some new equivalence results between the TTD and PARALIND are presented. The TT-cores structure is exposed when the Q-order PARALIND has only two full column rank factor matrices. Partial and full uniqueness conditions for the new TT-PARALIND model are also studied.

The notations used in this paper are as follows. Scalars, vectors, matrices and tensors are represented by x, $\boldsymbol{x}$, $\boldsymbol{X}$ and $\boldsymbol{\mathcal{X}}$, respectively. The symbols $(\cdot)^{\mathsf{T}}$ and $(\cdot)^{-1}$ denote, respectively, the transpose and the inverse. $\boldsymbol{\mathcal{I}}_{k,R}$ denotes the k-order identity tensor of size $R \times \cdots \times R$, and $\boldsymbol{\mathcal{I}}_{2,R} = \boldsymbol{I}_R$. The matrix $\text{unfold}_k \boldsymbol{\mathcal{X}}$ of size $N_k \times N_1 \cdots N_{k-1} N_{k+1} \cdots N_Q$ refers to the k-mode unfolding of $\boldsymbol{\mathcal{X}}$ of size $N_1 \times \cdots \times N_Q$. The n-mode product is denoted by $\underset{n}{\bullet}$. The contraction $\overset{p}{\underset{q}{\bullet}}$ between two tensors $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ of size $N_1 \times \cdots \times N_Q$ and $M_1 \times \cdots \times M_P$, with $N_q = M_p$ is a tensor of order $(Q + P - 2)$ such that

$$[\boldsymbol{\mathcal{A}} \overset{p}{\underset{q}{\bullet}} \boldsymbol{\mathcal{B}}]_{n_1,\ldots,n_{q-1},n_{q+1},\ldots,n_Q,m_1,\ldots,m_{p-1},m_{p+1},\ldots,m_P}$$
$$= \sum_{k=1}^{N_q} [\boldsymbol{\mathcal{A}}]_{n_1,\ldots,n_{q-1},k,n_{q+1},\ldots,n_Q} [\boldsymbol{\mathcal{B}}]_{m_1,\ldots,m_{p-1},k,m_{p+1},\ldots,m_P}.$$

2. EQUIVALENCE BETWEEN PARALIND AND TTD

2.1. Tensor-Train Decomposition (TTD)

Definition 1. *A Q-order tensor of size $N_1 \times \ldots \times N_Q$ that follows a Tensor Train decomposition (TTD) [6] of TT-ranks $\{R_1, \ldots, R_{Q-1}\}$ admits the following definition:*

$$\boldsymbol{\mathcal{X}} = \boldsymbol{G}_1 \overset{1}{\underset{2}{\bullet}} \boldsymbol{\mathcal{G}}_2 \overset{1}{\underset{3}{\bullet}} \boldsymbol{\mathcal{G}}_3 \overset{1}{\underset{4}{\bullet}} \ldots \overset{1}{\underset{Q-1}{\bullet}} \boldsymbol{\mathcal{G}}_{Q-1} \overset{1}{\underset{Q}{\bullet}} \boldsymbol{G}_Q, \tag{1}$$

where the TT-cores $\boldsymbol{G}_1, \boldsymbol{\mathcal{G}}_q$, and $\boldsymbol{G}_Q$ are, respectively, of dimensions $N_1 \times R_1, R_{q-1} \times N_q \times R_q$, and $R_{Q-1} \times N_Q$, for $2 \le q \le Q-1$, and we have $\text{rank}\{\boldsymbol{G}_1\} = R_1$, $\text{rank}\{\boldsymbol{G}_Q\} = R_{Q-1}$, $\text{rank}\{\text{unfold}_1\boldsymbol{\mathcal{G}}_q\} = R_{q-1}$, *and* $\text{rank}\{\text{unfold}_3\boldsymbol{\mathcal{G}}_q\} = R_q$.

It is straightforward to see that the TTD of $\boldsymbol{\mathcal{X}}$ in eq. (1) is not unique since

$$\boldsymbol{\mathcal{X}} = \boldsymbol{A}_1 \overset{1}{\underset{2}{\bullet}} \boldsymbol{\mathcal{A}}_2 \overset{1}{\underset{3}{\bullet}} \boldsymbol{\mathcal{A}}_3 \overset{1}{\underset{4}{\bullet}} \ldots \overset{1}{\underset{Q-1}{\bullet}} \boldsymbol{\mathcal{A}}_{Q-1} \overset{1}{\underset{Q}{\bullet}} \boldsymbol{A}_Q,$$

where

$$\boldsymbol{A}_1 = \boldsymbol{G}_1 \boldsymbol{U}_1^{-1},$$
$$\boldsymbol{A}_Q = \boldsymbol{U}_{Q-1} \boldsymbol{G}_Q,$$
$$\boldsymbol{\mathcal{A}}_q = \boldsymbol{U}_{q-1} \overset{1}{\underset{2}{\bullet}} \boldsymbol{\mathcal{G}}_q \overset{1}{\underset{3}{\bullet}} \boldsymbol{U}_q^{-1}.$$

For $1 \le q \le Q-1$, $\boldsymbol{U}_q$ are square nonsingular matrices of dimension $R_q \times R_q$. In practice, the TTD is performed thanks to the state-of-art TT-SVD algorithm [6]. It is a sequential algorithm that recovers the TT-cores $\boldsymbol{\mathcal{G}}_q$ based on $(Q-1)$ SVDs applied to several matrix-based reshapings using the original tensor $\boldsymbol{\mathcal{X}}$. This algorithm allows to recover the true TT-cores up to a post and pre-multiplication by transformation (*change-of-basis*) matrices due to the extraction of dominant subspaces when using the SVD. In the next section, we will derive the structure of the estimated TT-cores when the original tensor $\boldsymbol{\mathcal{X}}$ follows a CPD with linear dependencies between the columns of the loading matrices.

2.2. PARALIND-TTD equivalence

Consider Q-order tensor $\boldsymbol{\mathcal{X}}$ of size $N_1 \times \cdots \times N_Q$ that follows a rank-R CPD:

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{Q,R} \underset{1}{\bullet} \boldsymbol{P}_1 \underset{2}{\bullet} \boldsymbol{P}_2 \ldots \underset{Q}{\bullet} \boldsymbol{P}_Q, \tag{2}$$

where the loading matrices $\boldsymbol{P}_q$ are of size $N_q \times R$. It was shown in [7, 10] that if the loading matrices $\boldsymbol{P}_q$ are full-column rank for $1 \le q \le Q$, then they can be recovered from the TT-cores by order-3 CPD decompositions.

In this section we study the case where linear dependencies are present between the columns of the loading matrices of (2). Thus, a loading matrix $\boldsymbol{P}_q$ can be expressed as:

$$\boldsymbol{P}_q = \tilde{\boldsymbol{P}}_q \boldsymbol{\Phi}_q, \tag{3}$$

where $\tilde{\boldsymbol{P}}_q$ is full column rank of size $N_q \times R_q$ ($R_q \le R$) and $\boldsymbol{\Phi}_q$ is a rank deficient matrix of size $R_q \times R$ containing the dependency pattern between the columns of $\tilde{\boldsymbol{P}}_q$. This CPD model with linear dependencies is also known as *PARALIND (PARAllel profiles with LINear Dependences)* [11].

Theorem 1 (PARALIND - TTD equivalence)**.** *Decomposing tensor $\boldsymbol{\mathcal{X}}$ in (2) into a TT format, where $\boldsymbol{P}_1$ and $\boldsymbol{P}_Q$ are full column rank matrices, and $\boldsymbol{P}_q$ $(2 \le q \le Q-1)$ follow (3), recovers the estimated TT-cores such that*

$$\boldsymbol{G}_1 = \boldsymbol{P}_1 \boldsymbol{U}_1^{-1},$$
$$\boldsymbol{\mathcal{G}}_q = \boldsymbol{\mathcal{I}}_{3,R} \underset{1}{\bullet} \boldsymbol{U}_{q-1} \underset{2}{\bullet} (\tilde{\boldsymbol{P}}_q \boldsymbol{\Phi}_q) \underset{3}{\bullet} \boldsymbol{U}_q^{-\mathsf{T}}, \; 2 \le q \le Q-1$$
$$\boldsymbol{G}_Q = \boldsymbol{U}_{Q-1} \boldsymbol{P}_Q^{\mathsf{T}},$$

where, for $1 \le q \le Q-1$, $\boldsymbol{U}_q$ is a square $R \times R$ nonsingular matrix. The TT-cores $\boldsymbol{G}_1, \boldsymbol{\mathcal{G}}_q$, and $\boldsymbol{G}_Q$ are, respectively, of dimensions $N_1 \times R, R \times N_q \times R$, and $R \times N_Q$, given TT-ranks all equal to R.

Proof. Note that tensor $\boldsymbol{\mathcal{I}}_{Q,R}$ in eq. (2) can be expressed as

$$\boldsymbol{\mathcal{I}}_{Q,R} = \mathbf{I}_R \overset{1}{\underset{2}{\bullet}} \boldsymbol{\mathcal{I}}_{3,R} \overset{1}{\underset{3}{\bullet}} \cdots \overset{1}{\underset{Q-1}{\bullet}} \boldsymbol{\mathcal{I}}_{3,R} \overset{1}{\underset{Q}{\bullet}} \mathbf{I}_R, \tag{4}$$

replacing eq. (4) into eq. (2), we get

$$\boldsymbol{\mathcal{X}} = (\mathbf{I}_R \overset{1}{\underset{2}{\bullet}} \boldsymbol{\mathcal{I}}_{3,R} \overset{1}{\underset{3}{\bullet}} \cdots \overset{1}{\underset{Q}{\bullet}} \mathbf{I}_R) \underset{1}{\bullet} \boldsymbol{P}_1 \underset{2}{\bullet} \boldsymbol{P}_2 \underset{3}{\bullet} \ldots \underset{Q}{\bullet} \boldsymbol{P}_Q$$
$$= (\mathbf{I}_R \overset{1}{\underset{2}{\bullet}} \boldsymbol{\mathcal{I}}_{3,R} \overset{1}{\underset{3}{\bullet}} \cdots \overset{1}{\underset{Q}{\bullet}} \mathbf{I}_R) \underset{1}{\bullet} \boldsymbol{P}_1 \underset{2}{\bullet} \tilde{\boldsymbol{P}}_2 \boldsymbol{\Phi}_2 \underset{3}{\bullet} \ldots \underset{Q}{\bullet} \boldsymbol{P}_Q$$

Before introducing the ambiguity matrices $\boldsymbol{U}_q$, tensor $\boldsymbol{\mathcal{X}}$ can then be expressed into a TT format as

$$\boldsymbol{\mathcal{X}} = \underbrace{\boldsymbol{P}_1}_{\boldsymbol{A}_1} \overset{1}{\underset{2}{\bullet}} \underbrace{(\boldsymbol{\mathcal{I}}_{3,R} \underset{2}{\bullet} \tilde{\boldsymbol{P}}_2 \boldsymbol{\Phi}_2)}_{\boldsymbol{\mathcal{A}}_2} \overset{1}{\underset{3}{\bullet}} \cdots \overset{1}{\underset{Q-2}{\bullet}} \underbrace{(\boldsymbol{\mathcal{I}}_{3,R} \underset{2}{\bullet} \tilde{\boldsymbol{P}}_{Q-2} \boldsymbol{\Phi}_{Q-2})}_{\boldsymbol{\mathcal{A}}_{Q-2}}$$
$$\overset{1}{\underset{Q-1}{\bullet}} \underbrace{(\boldsymbol{\mathcal{I}}_{3,R} \underset{2}{\bullet} \tilde{\boldsymbol{P}}_{Q-1} \boldsymbol{\Phi}_{Q-1})}_{\boldsymbol{\mathcal{A}}_{Q-1}} \overset{1}{\underset{Q}{\bullet}} \underbrace{\boldsymbol{P}_Q^T}_{\boldsymbol{A}_Q}. \tag{5}$$

One may note that for $2 \le q \le Q-1$, the considered TT-cores $\boldsymbol{A}_1$, $\boldsymbol{\mathcal{A}}_q$ and $\boldsymbol{A}_Q$ verify the definition of the TTD given in Definition 1, *i.e.,* $\text{rank}\{\boldsymbol{A}_1\} = \text{rank}\{\boldsymbol{A}_Q\} = \text{rank}\{\text{unfold}_1\boldsymbol{\mathcal{A}}_q\} = \text{rank}\{\text{unfold}_3\boldsymbol{\mathcal{A}}_q\} = R$, which justify that matrices $\boldsymbol{P}_1$ and $\boldsymbol{P}_Q$ must be of full column rank. By identifying the TT-cores $\boldsymbol{\mathcal{A}}_q$ in eq. (5), introducing the pre- and post-multiplication ambiguity matrices $\boldsymbol{U}_q$ presented in 2.1, and using the following equivalence

$$\boldsymbol{\mathcal{G}}_q = \boldsymbol{U}_{q-1} \overset{1}{\underset{2}{\bullet}} \boldsymbol{\mathcal{A}}_q \overset{1}{\underset{3}{\bullet}} \boldsymbol{U}_q^{-1} = \boldsymbol{\mathcal{A}}_q \underset{1}{\bullet} \boldsymbol{U}_{q-1} \underset{3}{\bullet} \boldsymbol{U}_q^{-T},$$

theorem 1 is proven. □

3. UNIQUENESS OF THE PARALIND-TTD

One of the most popular condition for the uniqueness of the CPD decomposition is the Kruskal's condition [2] relying on the concept of "Kruskal-rank", or simply krank. The krank of an $N \times R$ matrix $\boldsymbol{P}$, denoted by $\text{krank}\{\boldsymbol{P}\}$, is the maximum value of $\ell \in \mathbb{N}$ such that every ℓ columns of $\boldsymbol{P}$ are linearly independent. By definition, the krank of a matrix is less than or equal to its rank. Kruskal proved [2] that the condition

$$\text{krank}\{\boldsymbol{P}_1\} + \text{krank}\{\boldsymbol{P}_2\} + \text{krank}\{\boldsymbol{P}_3\} \geq 2R + 2 \quad (6)$$

is sufficient for uniqueness of the CPD decomposition in (2), with $Q = 3$. Furthermore, it becomes a necessary and sufficient condition in the cases $R = 2$ or 3 (see [3]). Herein, by uniqueness, we understand "essential uniqueness", meaning that if another set of matrices $\bar{\boldsymbol{P}}_1$, $\bar{\boldsymbol{P}}_2$ and $\bar{\boldsymbol{P}}_3$ verify (6), then there exists a permutation matrix $\boldsymbol{\Pi}$ and three invertible diagonal scaling matrices ($\boldsymbol{\Delta}_1$, $\boldsymbol{\Delta}_2$, $\boldsymbol{\Delta}_3$) satisfying $\boldsymbol{\Delta}_1\boldsymbol{\Delta}_2\boldsymbol{\Delta}_3 = \boldsymbol{I}_R$, where $\boldsymbol{I}_R$ is the R-th-order identity matrix, such that

$$\bar{\boldsymbol{P}}_1 = \boldsymbol{P}_1\boldsymbol{\Pi}\boldsymbol{\Delta}_1, \quad \bar{\boldsymbol{P}}_2 = \boldsymbol{P}_2\boldsymbol{\Pi}\boldsymbol{\Delta}_2, \quad \bar{\boldsymbol{P}}_3 = \boldsymbol{P}_3\boldsymbol{\Pi}\boldsymbol{\Delta}_3.$$

The uniqueness condition (6) has been generalised to Q-order CPDs in [13]. It states that the loading matrices $\boldsymbol{P}_q$ $(q = 1, \ldots, Q)$ in (2) can be uniquely estimated from $\boldsymbol{\mathcal{X}}$ if

$$\sum_{q=1}^{Q} \text{krank}\{\boldsymbol{P}_q\} \geq 2R + (Q-1). \quad (7)$$

This condition is sufficient but not necessary for the uniqueness of the CPD decomposition.

Based on Kruskal's uniqueness condition as well as on the results derived in [4], we formulate in the following a partial and a full uniqueness condition for the PARALIND-TTD of a Q-order tensor.

Theorem 2 (Partial uniqueness of TT-PARALIND)**.** *The loading matrix $\boldsymbol{P}_q$ can be uniquely recovered from the estimated TT decomposition of $\boldsymbol{\mathcal{X}}$ if there exist q_1 and q_2 ($q_1 \neq q_2 \neq q$), such that:*

$$\begin{cases} \text{rank}\{\boldsymbol{P}_{q_1}\} = \text{rank}\{\boldsymbol{P}_{q_2}\} = R, \\ \text{rank}\{\boldsymbol{P}_q\} \geq 2. \end{cases}$$

Proof. In the CPD (2) the order of the factor matrices is arbitrary and can be changed by a simple index permutation. Thus, in the following we will suppose, without loss of generality, that $q_1 = 1$ and $q_2 = Q$. The fact that $\text{rank}\{\boldsymbol{P}_1\} = \text{rank}\{\boldsymbol{P}_Q\} = R$ implies that the square matrices $\boldsymbol{U}_q$ in theorem 1 are all full rank R. Therefore, the $\boldsymbol{\mathcal{G}}_q$ tensor can be uniquely recovered from $\boldsymbol{\mathcal{X}}$ by the TT-SVD algorithm.

According to theorem 1, the tensor $\boldsymbol{\mathcal{G}}_q$ can be expressed as:

$$\boldsymbol{\mathcal{G}}_q = \boldsymbol{\mathcal{I}}_{3,R} \underset{1}{\bullet} \boldsymbol{U}_{q-1} \underset{2}{\bullet} \boldsymbol{P}_q \underset{3}{\bullet} \boldsymbol{U}_q^{-\mathsf{T}}. \quad (8)$$

Following Kruskal's uniqueness condition (6), the factor matrices in (8) can be recovered from $\boldsymbol{\mathcal{G}}_q$ if

$$\text{krank}\{\boldsymbol{U}_{q-1}\} + \text{krank}\{\boldsymbol{P}_q\} + \text{krank}\{\boldsymbol{U}_q^{-\mathsf{T}}\} \geq 2R+2. \quad (9)$$

However, in our case we are only interested in recovering $\boldsymbol{P}_q$, which allows to relax Kruskal's condition. It was proven in [4] that the matrix $\boldsymbol{P}_q$ can be be uniquely estimated from $\boldsymbol{\mathcal{G}}_q$ if

$$\text{krank}\{\boldsymbol{U}_{q-1}\} + \text{rank}\{\boldsymbol{P}_q\} + \text{krank}\{\boldsymbol{U}_q^{-\mathsf{T}}\} \geq 2R+2. \quad (10)$$

As $\boldsymbol{U}_{q-1}$ and $\boldsymbol{U}_q$ are full rank square matrices, and $\text{rank}\{\boldsymbol{P}_q\} \geq 2$, (9) is verified, which completes the proof.

□

Theorem 3 (Full TT-PARALIND uniqueness)**.** *The loading matrices $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_Q$ can be uniquely recovered from the estimated TT-cores $\boldsymbol{G}_1, \boldsymbol{\mathcal{G}}_2, \ldots, \boldsymbol{\mathcal{G}}_{Q-1}, \boldsymbol{G}_Q$ if:*

$$\begin{cases} \text{rank}\{\boldsymbol{P}_1\} = \text{rank}\{\boldsymbol{P}_Q\} = R \\ \text{rank}\{\boldsymbol{P}_q\} \geq 2, \quad 2 < q < Q-1 \\ \text{krank}\{\boldsymbol{P}_2\}, \text{krank}\{\boldsymbol{P}_{Q-1}\} \geq 2. \end{cases}$$

Proof. This result is a consequence of theorem 2. The uniqueness of factor matrices $\boldsymbol{P}_2, \ldots, \boldsymbol{P}_{Q-1}$ can be proven by repeatedly applying theorem 2 to the different TT-cores $\boldsymbol{\mathcal{G}}_q$, $2 \leq q \leq Q-1$. Meanwhile, condition (10) does not guarantee uniqueness of the change-of-basis matrices $\boldsymbol{U}_{q-1}$ and $\boldsymbol{U}_q$. In order to guarantee this, Kruskal's condition (9) would need to be verified, which we do not require.

Thus, the condition $\text{krank}\{\boldsymbol{P}_2\}, \text{krank}\{\boldsymbol{P}_{Q-1}\} \geq 2$ implies uniqueness of the CPD decomposition of TT-cores $\boldsymbol{\mathcal{G}}_2$ and $\boldsymbol{\mathcal{G}}_{Q-1}$ and consequently, the uniqueness of the $R \times R$ non-singular matrices $\boldsymbol{U}_1$ and $\boldsymbol{U}_{Q-1}$. From theorem 1 we get:

$$\boldsymbol{P}_1 = \boldsymbol{G}_1\boldsymbol{U}_1 \text{ and } \boldsymbol{P}_Q = \boldsymbol{G}_Q^{\mathsf{T}}\boldsymbol{U}_{Q-1}^{-\mathsf{T}}.$$

Thus, the unique recovery of $\boldsymbol{G}_1$ and $\boldsymbol{G}_Q$ from $\boldsymbol{\mathcal{X}}$ along with uniqueness of $\boldsymbol{U}_1$ and $\boldsymbol{U}_{Q-1}$ implies uniqueness of factor matrices $\boldsymbol{P}_1$ and $\boldsymbol{P}_Q$, which completes the proof.

□

4. DISCUSSION

4.1. More restrictive conditions

Compared to Kruskal's condition (7) for order-Q CPD, the uniqueness condition of theorem 3 is more restrictive. For example, in the case of a fourth-order tensor ($Q = 4$), the condition of theorem 3 implies $\sum_{q=1}^{4} \text{krank}\{\boldsymbol{P}_q\} \geq 2R+4$, while Kruskal's condition requires $\sum_{q=1}^{4} \text{krank}\{\boldsymbol{P}_q\} \geq 2R + 3$. This is a direct consequence of imposing simultaneous (partial) uniqueness on all the order-3 TT-cores. More restrictive uniqueness conditions is the price to pay for having a numerically efficient algorithm, that guarantees recovery of the loading matrices for a wide variety of scenarios.

4.2. Estimation scheme architecture

It is worth noting that, from an algorithmic point of view, the estimation of the loading matrices $\boldsymbol{P}_q$ can be done either in parallel or sequentially. For a parallel estimation scheme, the conditions of theorem 3 are sufficient. In [7], a sequential scheme was proposed, based on a sequential retrieval of both matrices $\boldsymbol{P}_q$ and $\boldsymbol{U}_q$. It requires at each step the knowledge of $\boldsymbol{U}_{q-1}$ for decomposing $\boldsymbol{\mathcal{G}}_q$. To use a similar sequential scheme for the TT-PARALIND model, it is necessary to also ensure the uniqueness of matrices $\boldsymbol{U}_q$. This can be done by replacing condition $\text{rank}\{\boldsymbol{P}_q\} \geq 2$ $(2 < q < Q-1)$ in theorem 3 by a stronger one, $\text{krank}\{\boldsymbol{P}_q\} \geq 2$ $(2 < q < Q-1)$.

4.3. Perspectives

1. The condition $\text{rank}\{\boldsymbol{P}_1\} = \text{rank}\{\boldsymbol{P}_Q\}$ in theorem 1 requires the knowledge of the indices of full-rank modes of tensor $\boldsymbol{\mathcal{X}}$, which are then arbitrarily fixed to 1 and Q; once these two modes are fixed, the order in which the remaining modes are processed is arbitrary. It is certainly possible to obtain a condition involving only one full rank matrix, but in this case the order in which the other modes are processed must be carefully chosen to guarantee the required rank conditions for the TT-SVD algorithm. This aspect is currently under investigation.

2. A very promising application domain of these results is the low-rank approximation of high-dimensional probability mass functions. In this case, these uniqueness results are of upmost importance as the linear dependencies in the model could account for the random variables correlations. A potential application is represented by the flow cytometry data analysis, as shown in [14].

5. CONCLUSION

The factorisation of a high-order tensor into a collection of low-order tensors, called cores, is an important research topic. Indeed, this family of methods called tensor Networks is an efficient way to mitigate the well-known "curse of dimentionality" problem. In this work, we prove that a Q-order PARALIND of rank R can be reformulated as a $Q-2$ train of tensors possibly column-deficiency and two full column rank matrices. The condition of partial and full uniqueness are exposed and discussed.

6. REFERENCES

[1] 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.

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

[3] J. M. F. ten Berge and N. D. Sidiropoulos, "On uniqueness in CANDECOMP/PARAFAC," *Psychometrika*, vol. 67, no. 3, pp. 399–409, Sept. 2002.

[4] Xijing Guo, Sebastian Miron, David Brie, and Alwin Stegeman, "Uni-mode and partial uniqueness conditions for CANDECOMP/PARAFAC of three-way arrays with linearly dependent loadings," *SIAM Journal on Matrix Analysis and Applications*, vol. 33, no. 1, pp. 111–129, 2012.

[5] A. Cichocki, "Era of big data processing: A new approach via tensor networks and tensor decompositions," *CoRR*, 2014.

[6] I. V. Oseledets, "Tensor-Train decomposition," *SIAM J. Scientific Computing*, vol. 33, no. 5, pp. 2295–2317, 2011.

[7] Y. Zniyed, R. Boyer, A. L.F. de Almeida, and G. Favier, "High-order cpd estimation with dimensionality reduction using a tensor train model," in *26th European Signal Processing Conference (EUSIPCO)*, 2018.

[8] R. Bro, N. D. Sidiropoulos, and G. B. Giannakis, "A fast least squares algorithm for separating trilinear mixtures," in *ICA99 - Int. Workshop on Independent Component Analysis and Blind Separation.*, 1999.

[9] Anh-Huy Phan, Andrzej Cichocki, Ivan Oseledets, Salman Ahmadi Asl, Giuseppe Calvi, and Danilo Mandic, "Tensor networks for latent variable analysis: Higher order canonical polyadic decomposition," *arXiv:1809.00535*, 2018.

[10] Y. Zniyed, R. Boyer, A. L. F. de Almeida, and G. Favier, "Multidimensional harmonic retrieval based on vandermonde tensor train," *Elsevier Signal Processing*, vol. 163, pp. 75–86, 2019.

[11] Rasmus Bro, Richard A. Harshman, Nicholas D. Sidiropoulos, and Margaret E. Lundy, "Modeling multi-way data with linearly dependent loadings," *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 23, no. 7-8, pp. 324–340, 2009.

[12] L. Xu, T. Jing, Y. Longxiang, and Z. Hongbo, "PARALIND-based identifiability results for parameter estimation via uniform linear array," *EURASIP Journal on Advances in Signal Processing*, 2012.

[13] Nicholas D. Sidiropoulos and Rasmus Bro, "On the uniqueness of multilinear decomposition of N-way arrays," *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 14, no. 3, pp. 229–239, 2000.

[14] David Brie, Rémi Klotz, Sebastian Miron, Saïd Moussaoui, Christian Mustin, Philippe Bécuwe, and Stéphanie Grandemange, "Joint analysis of flow cytometry data and fluorescence spectra as a non-negative array factorization problem," *Chemometrics and Intelligent Laboratory Systems*, vol. 137, pp. 21–32, 2014.