Tensor train completion of multi-way data observed along one mode

Shakir Showkat Sofi
Dept. Electrical Engineering (ESAT)
KU Leuven
Kortrijk, Belgium
shakirshowkat.sofi@kuleuven.be

Stijn Hendrikx
Dept. Electrical Engineering (ESAT)
KU Leuven
Leuven, Belgium
stijn.hendrikx@kuleuven.be

Lieven De Lathauwer
Dept. Electrical Engineering (ESAT)
KU Leuven
Kortrijk/Leuven, Belgium
lieven.delathauwer@kuleuven.be

***Abstract*—Tensor completion is an extension of matrix completion aimed at recovering a tensor from partial observations. Many problems can be framed as instances of low-rank tensor completion. Typically, the low-rank tensor completion problems are solved by optimization-based algorithms. This work shows how to use only standard linear algebra tools to compute the tensor train decomposition of an incomplete tensor with fibers observed along a single mode. From an application viewpoint, this setting is relevant when it is easier to sample or collect a dataset along a specific mode.**

***Index Terms*—tensor completion, subspace constraints, low-rank approximation**

I. Introduction

Tensor completion is a technique for recovering a partially observed tensor by leveraging the observed entries and the pattern of observation. Properties, such as the low-rankness of the tensor, facilitate the relationship between the observed and unobserved entries. Many problems can be framed as instances of tensor completion, e.g., image and video recovery, weather and traffic data imputation and traffic state estimation (see, e.g., [1]–[4]).

The low-rank assumption is key in establishing a relationship between the observed and unobserved entries of the tensor. In the literature, two main approaches are used to address the problem of low-rank tensor completion. The first approach, the rank minimization approach, aims to find a tensor with the minimum rank that matches the observations within a certain tolerance. As the rank function is non-convex, convex surrogates of rank are used to formulate a relaxed problem. The second approach explicitly imposes the tensor decomposition model on a partially observed tensor with a fixed rank and minimizes the error to fit the model parameters (see, e.g., [1]–[3], [5]–[7]).

Tensor completion problems are typically solved by optimization-based algorithms, especially when the observation pattern is random. However, if an observation pattern has some structure, better and faster algorithms with theoretical guarantees can be designed by exploiting the structure. In this work, we develop a framework for tensor completion of a tensor observed along one mode under the tensor train format. The observation pattern is fiber-wise, meaning that only some fibers are observed along a particular mode of a tensor. This structure enables us to obtain the model parameters using standard linear algebra tools. Such an observation pattern arises in many applications. For example, consider time series data collected across different spatial coordinates (latitude, longitude). Often, we cannot collect or access time series data at all locations, resulting in a spatiotemporal data tensor with only some observed fibers. There are many applications where it is easier to sample a multivariate function along a particular variable than others, which can lead to a data tensor with fiber-wise observations. Note that there is an important difference with the popular technique of (TT) cross approximation, in the sense that the latter obtains TT decomposition from a subset of fibers sampled across all modes [8]. In our method, fibers are sampled along a single mode.

In our previous work, we have shown how to compute the canonical polyadic decomposition (CPD) and the multilinear singular value decomposition (MLSVD) of a tensor that is observed along one mode [9], [10]. The present study is an extension of our research to the tensor train decomposition (TT decomposition). The TT decomposition combines the features of both the MLSVD and the CPD. The TT decomposition is stable, like the MLSVD. It overcomes the curse of dimensionality (CoD) like the CPD since it has asymptotically the same number of parameters [11], [12]. Moreover, this compact representation can be used as a "proxy" of the actual tensor that allows one to perform other computations efficiently (e.g., a constrained CPD is fitted to the TT approximation rather than the actual tensor) [13]. Therefore, extending the fiber-wise completion in the tensor train framework becomes consequential.

A. Preliminaries and Notation

We use lower-case, bold lower-case, bold capital, and calligraphic letters to denote scalars, vectors, matrices, and tensors, respectively, i.e., $x, \mathbf{x}, \mathbf{X}$, and $\mathcal{X}$, respectively. The *order* of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is the number of its *modes* or *ways*. Matlab-like indexing is used to specify a particular part of a tensor, employing commas, colons, and semicolons. For

This work was supported by the Flemish Government's AI Research Program and KU Leuven Internal Funds (iBOF/23/064, C14/22/096, IDN/19/014). Lieven De Lathauwer, Stijn Hendrikx and Shakir Showkat Sofi are affiliated with Leuven.AI - KU Leuven institute for AI, B-3000, Leuven, Belgium.

instance, the specific entry of the Nth-order tensor is denoted by $x_{i_1 i_2 \cdots i_N} = \mathcal{X}(i_1, i_2, \ldots, i_N)$ with $i_n \in \{1, 2, 3, \ldots, I_n\}$ for $n \in \{1, 2, 3, \ldots, N\}$. A mode-$n$ *fiber* of the Nth-order tensor is obtained by fixing every index except the nth. For a third-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, the mode-1 fibers $\mathbf{x}_{:i_2 i_3}$, mode-2 fibers $\mathbf{x}_{i_1 : i_3}$, and mode-3 fibers $\mathbf{x}_{i_1 i_2 :}$ are also known as column, row, and tube fibers, respectively. Similarly, the mode-1 slices $\mathbf{X}_{i_1 ::}$, mode-2 slices $\mathbf{X}_{:i_2 :}$, and mode-3 slices $\mathbf{X}_{::i_3}$, are also known as horizontal, lateral, and frontal slices, respectively. A permutation and reshaping of a tensor are reflected in the ordering of its indices and the position of the semicolon (';'), where the semicolon indicates a new mode, and the ordering of indices indicates the order in which the entries are stacked. In the nth matrix unfolding that is denoted by $\mathbf{X}_{[1,\ldots,n;n+1,\ldots,N]} \in \mathbb{R}^{I_1 \cdots I_n \times I_{n+1} \cdots I_N}$, the first n indices enumerate rows and the remaining indices enumerate columns. The rank, column space (also called the range) and kernel of a matrix $\mathbf{X}$ are denoted by $\operatorname{rank}(\mathbf{X}), \operatorname{col}(\mathbf{X}), \ker(\mathbf{X})$, respectively. The dimension of subspace X is denoted by $\dim(X)$. The product (contraction) between the last mode of tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ and the first mode of tensor $\mathcal{Y} \in \mathbb{R}^{J_1 \times J_2 \times \ldots \times J_M}$, where $I_N = J_1 = K$, yields an $(N + M - 2)$-order tensor $\mathcal{Z} = \mathcal{X} \bullet \mathcal{Y}$, the elements of which are given by:

$$z_{i_1 \ldots i_{N-1} j_2 \ldots j_M} = \sum_{k=1}^{K} x_{i_1 \ldots i_{N-1} k} y_{k j_2 \ldots j_M}.$$

B. Organization

In Section II, we briefly review the TT decomposition of a fully observed tensor. Next, in Section III, we present our approach for obtaining the TT decomposition of a tensor observed along one mode, along with the algorithm and uniqueness conditions in Sections III-B and III-C, respectively. Finally, we present convincing numerical experiments in Section IV and a brief conclusion in Section V.

II. TT DECOMPOSITION OF FULLY OBSERVED TENSOR

The TT decomposition of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is written as the contraction of a sequence of core tensors (TT cores) $\mathcal{G}^{(n)} \in \mathbb{R}^{R_{n-1} \times I_n \times R_n} (1 \leq n \leq N)$ with $R_0 = R_N = 1$ such that each entry can be expressed as the sequence of matrix products [12], [14]:

$$\mathcal{X} = \mathcal{G}^{(1)} \bullet \mathcal{G}^{(2)} \bullet \cdots \bullet \mathcal{G}^{(N-1)} \bullet \mathcal{G}^{(N)}, \tag{1}$$

or entry-wise, we can write:

$$x_{i_1 i_2 \cdots i_{N-1} i_N} = \mathbf{G}^{(1)}_{:i_1:} \mathbf{G}^{(2)}_{:i_2:} \cdots \mathbf{G}^{(N-1)}_{:i_{N-1}:} \mathbf{G}^{(N)}_{:i_N:}, \tag{2}$$

where $\mathbf{G}^{(n)}_{:i:} \in \mathbb{R}^{R_{n-1} \times R_n}$ is the ith mode-2 slice of the core tensor $\mathcal{G}^{(n)}$. The tuple of minimal integers $(R_0, R_1, \ldots, R_N)$ for which equality in (1) and (2) holds is the TT-rank of $\mathcal{X}$, denoted by $\operatorname{rank}_{TT}(\mathcal{X})$. The dense tensor $\mathcal{X}$ has a total storage complexity of $\prod_{n=1}^{N} I_n$, whereas in the TT format, the storage complexity is $\sum_{n=1}^{N} R_{n-1} I_n R_n$. A low TT rank can greatly reduce storage complexity. Additionally, we introduce a shorthand notation for TT decomposition using partial products. A tensor $\mathcal{X}$ can be represented as:

$$\mathcal{X} = \mathcal{G}^{(<n)} \bullet \mathcal{G}^{(n:m)} \bullet \mathcal{G}^{(>m)}, \quad 1 \leq n \leq m \leq N, \tag{3}$$

where the left partial product $\mathcal{G}^{(<n)} = \mathcal{G}^{(1)} \bullet \cdots \bullet \mathcal{G}^{(n-1)}$, the right partial product $\mathcal{G}^{(>m)} = \mathcal{G}^{(m+1)} \bullet \cdots \bullet \mathcal{G}^{(N)}$ and $\mathcal{G}^{(n:m)} = \mathcal{G}^{(n)} \bullet \cdots \bullet \mathcal{G}^{(m)}$.

TT decomposition of a given tensor $\mathcal{X}$ is computed by a sequence of truncated SVDs (TT-SVD) [12]. The TT-SVD algorithm sequentially computes TT cores and essentially alternates between the SVD computation step and projection step, which makes it difficult to parallelize. Inspired by the natural parallelizability of MLSVD [15], a similar method was proposed that computes the orthonormal basis for the column spaces of all matrix unfoldings independently and then utilizes them to compute the TT cores (Parallel-TTSVD) [16].

III. TENSOR TRAIN COMPLETION OF A TENSOR OBSERVED ALONG ONE MODE

In this section, we introduce a methodology analogous to the Parallel-TTSVD that can be utilised to obtain the TT decomposition of an incomplete tensor with fibers observed along a single mode. Assume a given tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ with $\operatorname{rank}_{TT}(\mathcal{X}) = (R_0, \ldots, R_N)$ of which only some of the mode-N (last mode) fibers are observed. One essential step to compute the TT cores is to find the orthonormal ranges of the matrix unfoldings. However, since the matrix unfoldings have missing entries, it is not immediately possible to utilize, e.g., SVD or QR factorization to obtain the ranges. Note that two types of observation patterns arise in unfoldings. In the $(N-1)$th matrix unfolding, i.e., $\mathbf{X}_{[1,\ldots,N-1;N]}$, the rows are either fully observed or entirely missing. If the number of observed rows is greater than or equal to R_{N-1}, we can generically obtain the last TT core ($\mathcal{G}^{(N)} \in \mathbb{R}^{R_{N-1} \times I_N \times 1}$) by computing the top R_{N-1} right-singular vectors of a matrix formed by the observed rows. The other matrix unfoldings can be seen as the horizontal stack of mode-2 slices $\tilde{\mathbf{X}}_{:i:}$ of a third-order reshaping $\tilde{\mathcal{X}} = \mathcal{X}_{[1,\ldots,n;n+1,\ldots,N-1;N]} \in \mathbb{R}^{\prod_{i=1}^{n} I_i \times \prod_{i=n+1}^{N-1} I_i \times I_N}$ of a tensor $\mathcal{X}$. That is,

$$\mathbf{X}_{[1,\ldots,n;n+1,\ldots,N]} = \left[\tilde{\mathbf{X}}_{:1:} \ \cdots \ \tilde{\mathbf{X}}_{:\prod_{i=n+1}^{N-1} I_i :} \right]. \tag{4}$$

In this case, the rows of the lateral slices (submatrices of an unfolding) of $\tilde{\mathcal{X}}$ are either fully observed or entirely missing. However, different slices may have observed rows at different row indices. A visual representation of this matrix unfolding is shown in Fig. 1.

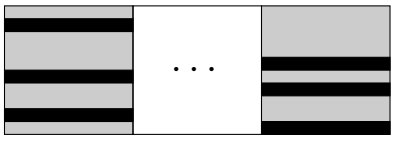
$\mathbf{X}_{[1,\ldots,n;n+1,\ldots,N]} =$ 

Fig. 1: The nth matrix unfolding of a tensor observed through fibers along a single mode, is characterized by a structured observation pattern.

The following section will discuss how to determine the orthonormal basis for the column space of such partially observed matrix unfolding.

A. The subspace constraint

For simplicity, we use $\mathbf{M} \in \mathbb{R}^{J\times K}$ to denote a low-rank matrix $\mathbf{X}_{[1,\ldots,n;n+1,\ldots,N]}$ of which we need to determine the R-dimensional column space. Let us denote a matrix that holds the orthonormal basis for $\mathrm{col}\,(\mathbf{M})$ by $\mathbf{A} \in \mathbb{R}^{J\times R}$. We denote the indices of the observed rows by a set $\alpha_l = \{i_1, \ldots, i_{J_l}\} \subseteq \{1, \ldots, J\}$, and the corresponding binary row selection matrix by $\mathbf{S}_r^{(l)} \in \mathbb{R}^{J_l\times J}$. It selects the rows indexed by α_l upon left multiplication. The cardinality of a set α_l is denoted by $|\alpha_l|$.

The central idea of the subspace constraint method is to determine the overall range of a matrix from a set of constraints derived from submatrices [17]. This method involves identifying an informationally complete set of fully observed rank-R submatrices $\{\mathbf{M}_{obs}^{(l)}\}_{l=1}^{L}$ of $\mathbf{M}$ [10]. By *informationally complete,* we mean that the set of constraints derived from the submatrices should fully characterize the desired subspace $\mathrm{col}\,(\mathbf{A})$. The lth submatrix $\mathbf{M}_{obs}^{(l)} \in \mathbb{R}^{J_l\times K_l}$ is obtained by forming the intersection of K_l columns and J_l rows with no missing entries, where $\min(J_l, K_l) \geq R$. As we assume that $\mathbf{M}_{obs}^{(l)}$ is rank R, i.e., $\mathbf{M}_{obs}^{(l)}$ has the same rank as the overall matrix $\mathbf{M}$, we can write $\mathbf{M}_{obs}^{(l)} = \mathbf{A}_{obs}^{(l)}\mathbf{B}_{obs}^{(l)\top}$, where matrices $\mathbf{A}_{obs}^{(l)} \in \mathbb{R}^{J_l\times R}$ and $\mathbf{B}_{obs}^{(l)} \in \mathbb{R}^{K_l\times R}$ are full column rank. Stacking a basis for the orthogonal complement of $\mathrm{col}\left(\mathbf{M}_{obs}^{(l)}\right)$ in a matrix $\mathbf{N}_{obs}^{(l)} \in \mathbb{R}^{J_l\times(J_l-R)}$, each column of the resulting matrix after zero padding, i.e., $\mathbf{S}_r^{(l)\top}\mathbf{N}_{obs}^{(l)} \in \mathbb{R}^{J\times(J_l-R)}$, represents a vector that is orthogonal to the desired subspace $\mathrm{col}\,(\mathbf{A})$, thus imposing constraints on what $\mathrm{col}\,(\mathbf{A})$ may be. To obtain the desired $\mathrm{col}\,(\mathbf{A})$, we find orthogonal complements for L such submatrices $\mathbf{M}_{obs}^{(l)}$ and concatenate them in a matrix $\mathbf{N} = \left[\mathbf{S}_r^{(1)\top}\mathbf{N}_{obs}^{(1)}, \cdots, \mathbf{S}_r^{(L)\top}\mathbf{N}_{obs}^{(L)}\right] \in \mathbb{R}^{J\times(\sum_{l=1}^{L}(J_l-R))}$. If the L orthogonal complements are enough to ensure the $\ker(\mathbf{N}^\top)$ is of minimal dimension, i.e., $\dim(\ker(\mathbf{N}^\top)) = R$, where the very existence of $\mathrm{col}\,(\mathbf{A})$ implies that the dimension is at least R, then the desired $\mathrm{col}\,(\mathbf{A})$ is necessarily given by $\ker(\mathbf{N}^\top)$. In what follows, we assume that our working assumptions hold, i.e., $\{\mathbf{M}_{obs}^{(l)}\}_{l=1}^{L}$ have rank R and $\dim(\ker(\mathbf{N}^\top)) = R$. Then, this set of submatrices is *informationally complete*. In the noisy case, we estimate $\mathbf{N}_{obs}^{(l)}$ from the left-singular vectors corresponding to the smallest singular values of $\mathbf{M}_{obs}^{(l)}$. Let the SVD of $\mathbf{M}_{obs}^{(l)}$ be $\mathbf{M}_{obs}^{(l)} = \mathbf{U}^{(l)}\mathbf{S}^{(l)}\mathbf{V}^{(l)\top}$. Then $\mathbf{N}_{obs}^{(l)} = \left[\mathbf{u}_{:R+1}^{(l)} \cdots \mathbf{u}_{:J_l}^{(l)}\right]$. In the noisy case, we also estimate the desired $\mathrm{col}\,(\mathbf{A})$ to be the subspace that is most orthogonal to the space spanned by the columns of $\mathbf{N}$, i.e., we estimate $\mathrm{col}\,(\mathbf{A})$ as the (approximate) kernel of $\mathbf{N}^\top$. Let the SVD of $\mathbf{N}$ be $\mathbf{U}\mathbf{S}\mathbf{V}^\top$. Then $\mathbf{A} = [\mathbf{u}_{:J-R+1} \cdots \mathbf{u}_{:J}]$.

B. Algorithm

In this section, we present our algorithm for computing the TT decomposition of a tensor observed along mode-N (last-mode). To compute the range of the nth matrix unfolding in (4), we use its subset of submatrices $\{\mathbf{S}_r^{(i_l)}\tilde{\mathbf{X}}_{:i_l:}\}_{l=1}^{L}$, which is informationally complete. After computing the orthonormal ranges of matrix unfoldings $\mathbf{X}_{[1,\ldots,n;n+1,\ldots,N]}$ for $1 \leq n \leq N-1$ independently, using the subspace constraint method, we utilize them to compute the TT cores. The pseudocode of the proposed method is outlined as Algorithm 1 (TT mode-N FiberWise).

Algorithm 1: TT mode-N FiberWise

Data: A tensor observed along mode-N $\mathcal{X} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$, TT rank $(R_0, \ldots, R_N)$ and row selection matrix $\mathbf{S}_r$

Result: TT cores $\mathcal{G}^{(1)}, \ldots, \mathcal{G}^{(N)}$

for $1 \leq n \leq N-1$ **do**
  Compute the orthonormal range $\mathbf{A}^{(n)} \in \mathbb{R}^{\prod_{i=1}^{n} I_i\times R_n}$ of nth matrix unfolding, using subspace constraint method.
end

Set $\mathcal{G}^{(1)} = \mathtt{reshape}\left(\mathbf{A}^{(1)}, [1, I_1, R_1]\right)$

for $1 \leq n \leq N-3$ **do**
  Compute $\mathbf{W}^{(n)} = \mathbf{A}^{(n)\top}\left(\mathbf{A}^{(n+1)}\right)_{[1,\ldots,n;n+1,(n+1)+1]}$
  Set $\mathcal{G}^{(n+1)} = \mathtt{reshape}\left(\mathbf{W}^{(n)}, [R_n, I_{n+1}, R_{n+1}]\right)$
end

Compute $[\sim, \sim, \mathbf{V}^{(N-1)}] = \mathtt{svd}\left(\mathbf{S}_r\mathbf{X}_{[1,\ldots,N-1;N]}, R_{N-1}\right)$

Set $\mathcal{G}^{(N)} = \mathtt{reshape}\left(\mathbf{V}^{(N-1)\top}, [R_{N-1}, I_N, 1]\right)$

Compute $\mathcal{G}^{(N-1)}$ by solving the linear system (5).

Computing the next-to-last TT core $\mathcal{G}^{(N-1)}$: In the fully observed case, the last TT core can be computed as $\mathcal{G}^{(N)} = \mathbf{S}^{(N-1)}\mathbf{V}^{(N-1)\top}$. However, for the partially observed tensors considered in this work, this is impossible since it requires the singular values of the matrix $\mathbf{X}_{[1,\ldots,N-1;N]}$, of which only some rows are observed (see, e.g., [12], [16]). To address this issue, we set the last core $\mathcal{G}^{(N)} = \mathbf{V}^{(N-1)\top}$, i.e., the orthonormal basis for the observed rows $(\mathbf{S}_r\mathbf{X}_{[1,\ldots,N-1;N]})$, where $\mathbf{S}_r$ is the row selection matrix for $(N-1)$th matrix unfolding and then we fix the scaling indeterminacies by computing the next-to-last TT core in a least-square sense. For the third-order tensor $\tilde{\mathcal{X}} = \mathcal{X}_{[1,\ldots,N-2;N-1;N]}$, the slice-wise representation, using (3), can be written as:

$$\tilde{\mathbf{X}}_{:i:} = \mathbf{G}^{(<N-1)}\mathbf{G}_{:i:}^{(N-1)}\mathbf{G}^{(>N-1)},\ i \in \{1, \ldots, I_{N-1}\},$$

where the matrices $\mathbf{G}^{(<N-1)} \in \mathbb{R}^{\prod_{i=1}^{N-2} I_i\times R_{N-2}}$ and $\mathbf{G}^{(>N-1)} = \mathbf{G}^{(N)} \in \mathbb{R}^{R_{N-1}\times I_N}$ have orthonormal columns and orthonormal rows, respectively. Let $\{\mathbf{S}_r^{(i)}\}_{i=1}^{I_{N-1}}$ be the selection matrices for the mode-2 slices of $\tilde{\mathcal{X}}$. The slices of the TT core $\mathcal{G}^{(N-1)}$ can now be computed by solving the following set of linear systems:

$$\mathbf{S}_r^{(i)}\left(\tilde{\mathbf{X}}_{:i:}\mathbf{G}^{(N)\top}\right) = \left(\mathbf{S}_r^{(i)}\mathbf{G}^{(<N-1)}\right)\mathbf{G}_{:i:}^{(N-1)},\ i \in \{1, \ldots, I_{N-1}\}. \tag{5}$$

C. Uniqueness conditions

This section will briefly discuss the generic conditions for uniqueness, up to basis transformation, of the TT cores in the

noiseless case. In the journal version of this work, we will provide an extensive algebraic discussion of the uniqueness conditions.

In Section III-A, we showed that the orthonormal range of the nth matrix unfolding, as shown in (4), can be computed from a set of its submatrices that is informationally complete. For this set of submatrices to be informationally complete, the following conditions need to be met. (i) The submatrices $\mathbf{S}_r^{(i_l)}\tilde{\mathbf{X}}_{:i_l:} \in \mathbb{R}^{|\alpha_l|\times I_N}$ have a rank equal to R_n. This means that the submatrices in Fig. 1 have at least R_n observed rows. (ii) Every row should appear in at least one of the submatrices, i.e., $\cup_{\ell=1}^{L}\alpha_\ell = \{1,\ldots,\prod_{i=1}^{n} I_i\}$, otherwise the concatenated matrix $\mathbf{N}$ will contain a row/rows with all zero entries. (iii) We should have $\dim(\ker(\mathbf{N}^\top)) = R_n$. If there exist some pairs of slices, $\tilde{\mathbf{X}}_{:i_{l_1}:}$ and $\tilde{\mathbf{X}}_{:i_{l_2}:}$, which overlap with at least R_n observed rows, then generically, condition (iii) is satisfied. Similar conditions for matrix completion were derived in previous works (see, e.g., [18]). Conditions (i), (ii), and (iii) are required to retrieve the TT cores $\left(\mathcal{G}^{(1)}\right.$ to $\left.\mathcal{G}^{(N-2)}\right)$. The next-to-last TT core $\mathcal{G}^{(N-1)}$ is obtained by solving a set of linear systems, as in (5). To uniquely solve (5), the mode-2 slices of the third-order reshaping $\tilde{\mathcal{X}} = \mathcal{X}_{[1,\ldots,N-2;N-1;N]}$ should have at least R_{N-2} observed rows. The last TT core $\mathcal{G}^{(N)}$ can be retrieved iff rank $\left(\mathbf{S}_r\mathbf{X}_{[1,\ldots,N-1;N]}\right) \geq R_{N-1}$, i.e., generically, if at least R_{N-1} mode-N fibers are observed.

IV. Experiments

In this section, we evaluate the performance of the proposed approach on both synthetic and real data. We compare our method with the tensor train weighted optimization (TT-WOPT) used in [3] and the simple low-rank tensor completion via TT (SiLRTC-TT) used in [19].

A. Synthetic data

Algorithm 1 will compute the exact TT decomposition of a noiseless tensor, of which only some mode-N fibers are observed if uniqueness conditions are met. In the presence of noise, the estimated TT decomposition will just be close to the true TT decomposition. We set up an experiment where we first generated 5th-order tensors $\mathcal{X} \in \mathbb{R}^{15\times15\times15\times15\times15}$ with $\text{rank}_{TT}(\mathcal{X}) = (1,3,3,3,4,1)$, by sampling the TT cores from the standard normal distribution. Gaussian noise $\mathcal{N}$ is added to the generated tensor to obtain a noisy tensor $\mathcal{X}_{noisy} = \mathcal{X} + \mathcal{N}$ with a fixed signal-to-noise ratio (SNR), which is defined as $\text{SNR} = 20\log_{10}\frac{\|\mathcal{X}\|}{\|\mathcal{N}\|}$. We randomly removed 30% of mode-5 fibers of the noisy tensor and ran Algorithm 1 along with the TT-WOPT and the SiLRTC-TT algorithms to compute the TT approximations $\hat{\mathcal{X}} = \mathcal{G}^{(1)} \bullet \cdots \bullet \mathcal{G}^{(5)}$. We recorded the median of the relative error, which is defined as Relative Error $= \frac{\|\mathcal{X}-\hat{\mathcal{X}}\|_F}{\|\mathcal{X}\|_F}$. Fig. 2 (left) shows the median relative error across 50 trials, plotted as a function of SNR.

Next, we generated 4th-order tensors $\mathcal{X} \in \mathbb{R}^{I\times I\times I\times I}$, where we varied I from 5 to 45 with $\text{rank}_{TT}(\mathcal{X}) = (1,3,3,4,1)$. Again, we added Gaussian noise to the generated tensors to obtain an SNR = 25 dB and randomly removed 35% of the mode-4 fibers. Fig. 2 (right) shows the median relative error and computation time across 10 trials, plotted as a function of I. It can be seen that the proposed approach is indeed fast. This is attributed to the fact that Algorithm 1 uses only standard linear algebra tools to compute the TT approximation of an incomplete tensor. Therefore, the proposed method can be used as an initialization method for optimization-based algorithms.

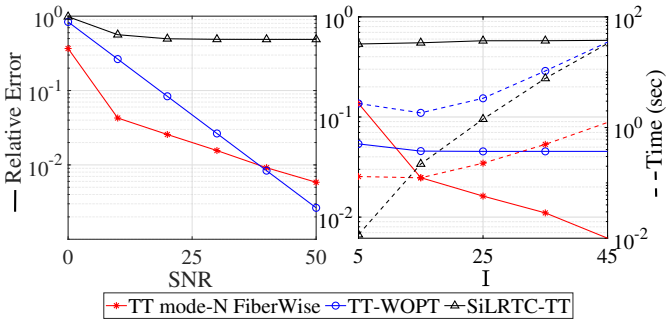


Fig. 2: In both plots we can see that the relative error of the proposed algorithm is lower than that of TT-WOPT and SiLRTC-TT at low to modest SNR values. However, TT-WOPT has a slightly lower relative error at higher SNR values. On the right, we can additionally see the computation time vs dimensions I of a tensor.

B. Multidimensional harmonic retrieval

Harmonic retrieval is a classical problem in signal processing. The goal is to estimate the parameters of a signal that is modelled as a sum of complex exponentials. Multidimensional harmonic retrieval (MHR) is the natural multidimensional extension. MHR data tensor sampled over the $(D+1)$th-order tensor grid can be modelled as:

$$x_{i_1 i_2 \cdots i_d k} = \sum_{r=1}^{R} s_r(k) \prod_{d=1}^{D} \mathrm{e}^{j(i_d-1)\mu_r^{(d)}} + n_{i_1 i_2 \cdots i_D k}, \qquad (6)$$

where $j^2 = -1$ and $s_r(k)$ is the kth complex symbol carried by the rth multidimensional harmonic. The noise $n_{i_1 i_2 \cdots i_D k}$ is modelled as zero-mean i.i.d additive Gaussian noise.

In this experiment, we generated 5th-order data tensors $\mathcal{X} \in \mathbb{C}^{10\times10\times10\times10\times20}$, by the CPD model of rank $R = 4$, using (6). The parameter D was set to 4, and we used binary phase-shift keying sources $s_r(k)$ of length $K = 20$. The parameter vectors were set as follows: $\boldsymbol{\mu}^{(1)} = [1, -0.5, 0.1, -0.8]$, $\boldsymbol{\mu}^{(2)} = [-0.5, 1, -0.9, 0.2]$, $\boldsymbol{\mu}^{(3)} = [-0.2, -0.6, 1.0, 0.4]$, and $\boldsymbol{\mu}^{(4)} = [-0.8, 0.4, 0.3, -0.1]$, which were previously used in [20]. In the first setting, we added Gaussian noise to the generated tensor with varying noise levels to achieve an SNR from 0 dB to 50 dB. We randomly removed 45% of fibers in mode-5. The estimated TT approximations were subsequently used to compute the parameter vectors $\hat{\boldsymbol{\mu}}^{(d)}$ using the ESPRIT algorithm [21]. Root mean square error (RMSE), which is defined as $\text{RMSE} = \sqrt{\frac{1}{RD}\sum_{r=1}^{R}\sum_{d=1}^{D}\left(\mu_r^{(d)} - \hat{\mu}_r^{(d)}\right)^2}$, was used to asses the accuracy. The results are summarized in Fig. 3 (left). Next, we fixed the SNR to 25 dB and studied the

effect of the missing rate on the accuracy. In Fig. 3 (right), we plotted the RMSE as a function of the missing fiber rate.

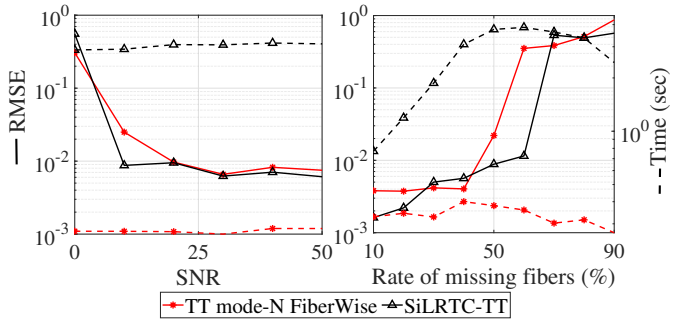


Fig. 3: In both plots we can see that the accuracy of the proposed algorithm is comparable to that of SiLRTC-TT. Note that the computation time required for the proposed algorithm is significantly less than SiLRTC-TT.

The results of this experimental study have shown that the proposed method is effective in terms of accuracy and computation time.

V. Conclusion

We introduced the framework for computing the TT decomposition of an incomplete tensor observed along one mode. The proposed method is solely based on standard linear algebra tools. The uniqueness conditions were briefly highlighted. Convincing numerical experiments showed that the proposed approach is practical. Furthermore, results suggest that the presented approach is fast in line with expectations. Hence, it is suitable for initializing optimization-based methods, which are typically more computationally expensive. Future work will show how to initialize other tensor decompositions directly using the estimated TT decomposition without constructing a full tensor.

References

[1] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, pp. 208–220, 2013.

[2] Q. Song, H. Ge, J. Caverlee, and X. Hu, "Tensor completion algorithms in big data analytics," *ACM Trans. Knowl. Discov. Data*, vol. 13, pp. 1–48, 2019.

[3] L. Yuan, Q. Zhao, L. Gui, and J. Cao, "High-order tensor completion via gradient-based optimization under tensor train format," *Signal Process. Image Commun.*, vol. 73, pp. 53–61, 2019.

[4] X. Chen, J. Yang, and L. Sun, "A nonconvex low-rank tensor completion model for spatiotemporal traffic data imputation," *Transp. Res. Part C Emerg. Technol.*, vol. 117, p. 102 673, 2020.

[5] E. Candès and B. Recht, "Exact matrix completion via convex optimization," *Commun. ACM*, vol. 55, pp. 111–119, 2012.

[6] D. Kressner, M. Steinlechner, and B. Vandereycken, "Low-rank tensor completion by Riemannian optimization," *BIT Numer. Math.*, vol. 54, pp. 447–468, 2014.

[7] S. Gandy, B. Recht, and I. Yamada, "Tensor completion and low-n-rank tensor recovery via convex optimization," *Inverse Probl.*, vol. 27, p. 025 010, 2011.

[8] I. V. Oseledets and E. Tyrtyshnikov, "TT-cross approximation for multidimensional arrays," *Linear Algebra Appl.*, vol. 432, pp. 70–88, 2010.

[9] M. Sørensen and L. De Lathauwer, "Fiber sampling approach to canonical polyadic decomposition and application to tensor completion," *SIAM J. Matrix Anal. Appl.*, vol. 40, pp. 888–917, 2019.

[10] S. Hendrikx, M. Sørensen, and L. De Lathauwer, "Multilinear singular value decomposition of a tensor with fibers observed along one mode," in *2023 IEEE Statistical Signal Processing Workshop (SSP)*, 2023, pp. 566–570.

[11] I. V. Oseledets and E. E. Tyrtyshnikov, "Breaking the curse of dimensionality, or how to use SVD in many dimensions," *SIAM J. Sci. Comput.*, vol. 31, pp. 3744–3759, 2009.

[12] I. Oseledets, "Tensor-train decomposition," *SIAM J. Sci. Comput.*, vol. 33, pp. 2295–2317, 2011.

[13] N. Vervliet, O. Debals, and L. De Lathauwer, "Exploiting efficient representations in large-scale tensor decompositions," *SIAM J. Sci. Comput.*, vol. 41, A789–A815, 2019.

[14] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, "Matrix product state representations," *Quantum Inf. Comput.*, vol. 7, pp. 401–430, 2007.

[15] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, pp. 1253–1278, 2000.

[16] T. Shi, M. Ruth, and A. Townsend, "Parallel algorithms for computing the tensor-train decomposition," *SIAM J. Sci. Comput.*, vol. 45, pp. C101–C130, 2023.

[17] D. W. Jacobs, "Linear fitting with missing data for structure-from-motion," *Comput. Vis. Image Underst.*, vol. 82, pp. 57–81, 2001.

[18] D. L. Pimentel-Alarcón, N. Boston, and R. D. Nowak, "A characterization of deterministic sampling patterns for low-rank matrix completion," *IEEE J. Sel. Top. Signal Process.*, vol. 10, pp. 623–636, 2016.

[19] J. A. Bengua, H. N. Phien, H. D. Tuan, and M. N. Do, "Efficient tensor completion for color image and video recovery: Low-rank tensor train," *IEEE Trans. Image Process.*, vol. 26, pp. 2466–2479, 2017.

[20] N. Vervliet, O. Debals, L. Sorber, and L. De Lathauwer, "Breaking the curse of dimensionality using decompositions of incomplete tensors: Tensor-based scientific computing in big data analysis," *IEEE Signal Process. Mag.*, vol. 31, pp. 71–79, 2014.

[21] R. Roy and T. Kailath, "Esprit-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, pp. 984–995, 1989.