Tucker Decomposition Based on a Tensor Train of Coupled and Constrained CP Cores

Maxence Giraud, Vincent Itier, Remy Boyer, Yassine Zniyed, André de Almeida

Tucker Decomposition Based on a Tensor Train of Coupled and Constrained CP Cores

Maxence Giraud* *Student Member, IEEE*, Vincent Itier†‡, Rémy Boyer* *Senior Member, IEEE*, Yassine Zniyed§ and André L.F. de Almeida¶ *Senior Member, IEEE*

***Abstract*—Many real-life signal-based applications use the Tucker decomposition of a high dimensional/order tensor. A well-known problem with the Tucker model is that its number of entries increases exponentially with its order, a phenomenon known as the "curse of the dimensionality". The Higher-Order Orthogonal Iteration (HOOI) and Higher-Order Singular Value Decomposition (HOSVD) are known as the gold standard for computing the range span of the factor matrices of a Tucker Decomposition but also suffer from the curse. In this paper, we propose a new methodology with a similar estimation accuracy as the HOSVD with non-exploding computational and storage costs. If the noise-free data follows a Tucker decomposition, the corresponding Tensor Train (TT) decomposition takes a remarkable specific structure. More precisely, we prove that for a Q-order Tucker tensor, the corresponding TT decomposition is constituted by $Q-3$ 3-order TT-core tensors that follow a Constrained Canonical Polyadic Decomposition. Using this new formulation and the coupling property between neighboring TT-cores, we propose a JIRAFE-type scheme for the Tucker decomposition, called TRIDENT. Our numerical simulations show that the proposed method offers a drastically reduced complexity compared to the HOSVD and HOOI while outperforming the Fast Multilinear Projection (FMP) method in terms of estimation accuracy.**

***Index Terms*—Tensor, Tucker decomposition, Constrained CPD, Tensor Train, Multilinear Algebra**

I. Introduction

Low-rank tensor decompositions are increasingly used to solve a variety of difficult problems with multilinear data, and this in various areas such as signal processing and telecommunications [1], [2], [3], [4], [5], [6], [7], [8], [9], hyperspectral image reconstruction [10], [11] and machine learning [1], [12], [13]. A popular generalization of the Singular Value Decomposition (SVD) [14] to tensors is the Tucker Decomposition [15], [16], [17]. Large-scale tensors suffer from the "curse of dimensionality" [18] as complexity and storage costs scale exponentially with the order of the tensor. Recently, tensor networks [4], [19], [20] such as the Tensor Train (TT) decomposition [21] have been introduced to mitigate the curse of dimensionality. The TT decomposition represents a tensor as a cascade of lower-order tensors and has been exploited in several recent works such as signal processing applications [22], [23], [24]. This format is very popular due to its low storage compactness and has a good trade-off between optimality (with the TT-SVD algorithm) and numerical stability from an algorithmic point of view. In this paper, we propose to use the TT decomposition formulation of the Tucker decomposition [25] to efficiently estimate the factor matrices. We go further than [25] by proving that the TT-cores follow a Constrained Canonical Polyadic Decomposition. This theoretical result suggests us a new algorithm called TRIDENT (**T**ucke**r** Decompos**i**tion base**d** on a T**e**nsor Trai**n** of Coupled and constrain**t** CP cores).

Notations: In this paper, we represent vectors, matrices, and higher-order tensors by lower case bold letters ($\mathbf{x}$), upper case bold letters ($\mathbf{X}$) and upper case calligraphic letters ($\mathcal{X}$) respectively. To indicate an entry, we write $x_{i_1,\ldots,i_n}$. The transpose, the transpose of the inverse, and the Moore-Penrose inverse of the matrix $\mathbf{X}$ are respectively denoted by $\mathbf{X}^{\mathrm{T}}$, $\mathbf{X}^{-\mathrm{T}}$ and $\mathbf{X}^{\dagger}$. The n^{th} unfolding of a tensor $\mathcal{X}$ is written as $\mathcal{X}^{(n)}$. The identity tensor is written as $\mathcal{I}_{f,d}$ where d is the order and f is the size of the dimensions. The symbols $\circ$ and $\otimes$ denote respectively the outer product between vectors and the Kronecker multiplication between matrices respectively. The Khatri-Rao product written as $\odot$ is defined as the column-wise Kronecker product [17]. We write the n-mode product and the $\binom{m}{n}$-mode product respectively as $\times_n$ and $\times_n^m$ [17].

II. Background

In this section, we provide a few definitions and properties that will be useful in the formulation of the proposed method.

Definition 1: The Tucker decomposition of a Q-order tensor $\mathcal{X}$ of dimensions $I_1 \times \ldots \times I_Q$ with multilinear ranks $\{T_1, \ldots, T_Q\}$ is defined as [16]

$$\mathcal{X} = \mathcal{C} \times_1 \mathbf{F}_1 \times_2 \mathbf{F}_2 \times_3 \cdots \times_Q \mathbf{F}_Q, \tag{1}$$

where $\mathcal{C}$ is a Q-order tensor of size $T_1 \times \cdots \times T_Q$, and $\mathbf{F}_q$ is a factor matrix of size $I_q \times T_q$, for $1 \leq q \leq Q$. The CP decomposition [15], [26] expression can be interpreted as a Tucker decomposition with the core tensor being the identity tensor and the multilinear ranks all equal to a single value called the canonical rank [17].

Since the CP format is a special case of the Tucker format, we can rewrite the Tucker decomposition as a CPD. A useful property of such refactoring is given below.

This work has been partially supported by the ANR-20-THIA-0013 program "AI_PhD@Lille" and the CornelIA (CO-construction RespoNsable Et durable d'une Intelligence Artificielle) CPER project.

* Univ. Lille, UMR 9189 CRIStAL, F-59000 Lille, France
† IMT Nord Europe, Institut Mines-Télécom, Centre for Digital Systems, F-59000 Lille, France
‡ Univ. Lille, CNRS, Centrale Lille, Institut Mines-Télécom, UMR 9189 CRIStAL, F-59000 Lille, France
§ Univ. Toulon, Aix Marseille Université, CNRS, LIS, UMR 7020, France
¶ Department of Teleinformatics Engineering, Federal University of Fortaleza, Brazil
Corresponding author: maxence.giraud@univ-lille.fr

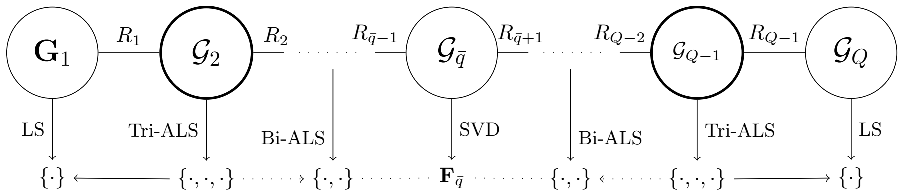


Fig. 1: Big picture of the TRIDENT algorithm. Starting from $\mathcal{G}_2$ and $\mathcal{G}_{Q-1}$ we can estimate iteratively the Tucker factors and their associated latent variables thanks to the coupling property of the TT-cores.

Property 1: Defining $\mathcal{Y} = \text{reshape}(\mathbf{I}_c; (a, b, c))$ with $ab = c$, as a reshaping of the identity matrix into a 3-order tensor, we have [27]

$$\mathcal{Y} = \mathcal{I}_{3,ab} \times_1 \mathbf{\Phi}_{a,b} \times_2 \mathbf{\Psi}_{a,b} \tag{2}$$

with $\mathbf{\Psi}_{a,b} = \mathbf{I}_a \otimes \mathbf{1}_b^{\mathrm{T}} \in \mathbb{R}^{a \times ab}$ and $\mathbf{\Phi}_{a,b} = \mathbf{1}_a^{\mathrm{T}} \otimes \mathbf{I}_b \in \mathbb{R}^{b \times ab}$.

Definition 2: The Tensor-Train decomposition of a tensor $\mathcal{X}$ with TT-ranks $\{R_1, ..., R_{Q-1}\}$ is defined as [21]

$$\mathcal{X} = \mathbf{G}_1 \times_2^1 \mathcal{G}_2 \times_3^1 \mathcal{G}_3 \times_4^1 \cdots \times_Q^1 \mathbf{G}_Q \tag{3}$$

where $\mathcal{G}_j$ of dimensions $R_{j-1} \times I_{j-1} \times R_j$, for $1 \leq q \leq Q$, are called the TT-cores, with $R_0 = R_Q = 1$.

III. New link between the Tucker and the Tensor-Train decomposition

A. Established link

The proposed method relies on the relationship between the Tucker decomposition and the TT established recently in [25], where a high-order tensor estimation problem is solved via trains of coupled third-order CP and Tucker tensors.

Theorem 1: Consider a tensor $\mathcal{X}$ of dimensions $I_1 \times \cdots \times I_Q$ expressed in a Tucker format with multilinear ranks $\{T_1, ..., T_Q\}$. The set of TT-cores $\{\mathcal{G}_q\}$ associated with the TT decomposition of $\mathcal{X}$ are given by [25]

$$\begin{aligned}
&\mathbf{G}_1 = \mathbf{F}_1 \mathbf{M}_1^{-1}, \\
&\mathcal{G}_q = \mathcal{T}_q \times_1 \mathbf{M}_{q-1} \times_2 \mathbf{F}_q \times_3 \mathbf{M}_q^{-\mathrm{T}} \qquad (1 < q < \bar{q}), \\
&\mathcal{G}_{\bar{q}} = \mathcal{C}_{\bar{q}} \times_1 \mathbf{M}_{\bar{q}-1} \times_2 \mathbf{F}_{\bar{q}} \times_3 \mathbf{M}_{\bar{q}}^{-\mathrm{T}}, \\
&\mathcal{G}_q = \bar{\mathcal{T}}_q \times_1 \mathbf{M}_{q-1} \times_2 \mathbf{F}_q \times_3 \mathbf{M}_q^{-\mathrm{T}} \qquad (\bar{q} < q < Q), \\
&\mathbf{G}_Q = \mathbf{M}_{Q-1} \mathbf{F}_Q^{-\mathrm{T}}.
\end{aligned} \tag{4}$$

The TT-ranks associated with the TT-cores are defined as $R_q = \min\left(\prod_{i=1}^q T_i, \prod_{i=q+1}^Q T_i\right)$, with $\bar{q}$ being the smallest q that verifies $\prod_{i=1}^q T_i < \prod_{i=q+1}^Q T_i$. The factor matrices are denoted by $\mathbf{F}_q$ and have dimensions $I_q \times T_q$. The matrices $\mathbf{M}_q$ are change-of-basis matrices of dimensions $R_q \times R_q$ which are latent variables shared between neighboring TT-cores. Finally, $\mathcal{T}_q$ and $\bar{\mathcal{T}}_q$ are proper reshapings of the identity matrix to match the dimensions of the related factors in Eq. (5). This relationship leads to the FMP algorithm [25], which allows estimating the Tucker factors using the TT decomposition.

B. The coupled constrained CP formulation

In order to estimate all the Tucker factors using the equivalence between Tucker and TT decomposition, we propose to reformulate this equivalence using Property 1. The tensors $\mathcal{T}_q$ and $\bar{\mathcal{T}}_q$ from Eq. (4) are reshapings of the identity matrix defined as

$$\begin{aligned}
\mathcal{T}_q &= \text{reshape}(\mathbf{I}_{R_q}; (R_{q-1}, T_q, R_q)) \qquad (1 < q < \bar{q}), \\
\bar{\mathcal{T}}_q &= \text{reshape}(\mathbf{I}_{R_q}; (R_{q-1}, T_q, R_q)) \qquad (\bar{q} < q < Q).
\end{aligned} \tag{5}$$

Using Theorem 1 and Eq. (5), we can decompose the TT-cores in Eq. (4) as constrained CPDs, the structure of which are defined by the following lemma.

Lemma 1: Using the constrained CP formulation of Property 1, we can rewrite the second and fourth equations (for $1 < q < \bar{q}$ and $\bar{q} < q < Q$, respectively) of the TT-cores defined in Theorem 1 as

$$\begin{aligned}
\mathcal{G}_q &= \mathcal{I}_{3,R_q} \times_1 \mathbf{M}_{q-1} \mathbf{\Psi}_{R_{q-1},T_q} \times_2 \mathbf{F}_q \mathbf{\Phi}_{R_{q-1},T_q} \times_3 \mathbf{M}_q^{-\mathrm{T}} \\
\mathcal{G}_q &= \mathcal{I}_{3,R_q} \times_1 \mathbf{M}_{q-1} \times_2 \mathbf{F}_q \mathbf{\Phi}_{R_q,T_q} \times_3 \mathbf{M}_q^{-\mathrm{T}} \mathbf{\Psi}_{R_q,T_q}.
\end{aligned}$$

Proof. Using Property 1, we can rewrite $\mathcal{T}_q$ and $\bar{\mathcal{T}}_q$ as

$$\begin{aligned}
\mathcal{T}_q &= \mathcal{I}_{3,R_q} \times_1 \mathbf{\Phi}_{R_{q-1},T_q} \times_2 \mathbf{\Psi}_{R_{q-1},T_q}, \\
\bar{\mathcal{T}}_q &= \mathcal{I}_{3,R_q} \times_2 \mathbf{\Phi}_{R_q,T_q} \times_3 \mathbf{\Psi}_{R_q,T_q}.
\end{aligned} \tag{6}$$

We can obtain the above two expressions for $\mathcal{G}_q$ (for $1 < q < \bar{q}$ and $\bar{q} < q < Q$, respectively) by substituting Eq. (6) into Eq. (4) using the fact that the order of the mode-n multiplication is irrelevant. □

IV. Estimation scheme

A. TRIDENT algorithm

In this section, we introduce a new algorithm which efficently solve the problem of interest

$$\begin{aligned}
&\min_{\mathbf{F}_1,\ldots,\mathbf{F}_Q,\mathbf{M}_1,\ldots,\mathbf{M}_{Q-1}} \|\mathbf{G}_1 - \mathbf{F}_1 \mathbf{M}_1^{-1}\|_F^2 \\
&+ \sum_{q=2}^{\bar{q}-1} \|\mathcal{G}_q - \mathcal{I}_{3,R_q} \times_1 \mathbf{M}_{q-1} \mathbf{\Psi}_{R_{q-1},T_q} \times_2 \mathbf{F}_q \mathbf{\Phi}_{R_{q-1},T_q} \times_3 \mathbf{M}_q^{-\mathrm{T}}\|_F^2 \\
&+ \|\mathcal{G}_{\bar{q}} - \mathcal{C}_{\bar{q}} \times_1 \mathbf{M}_{\bar{q}-1} \times_2 \mathbf{F}_{\bar{q}} \times_3 \mathbf{M}_{\bar{q}}^{-\mathrm{T}}\|_F^2 \\
&+ \sum_{q=\bar{q}+1}^{Q-1} \|\mathcal{G}_q - \mathcal{I}_{3,R_q} \times_1 \mathbf{M}_{q-1} \times_2 \mathbf{F}_q \mathbf{\Phi}_{R_q,T_q} \times_3 \mathbf{M}_q^{-\mathrm{T}} \mathbf{\Psi}_{R_q,T_q}\|_F^2 \\
&+ \|\mathbf{G}_Q - \mathbf{M}_{Q-1} \mathbf{F}_Q^{-\mathrm{T}}\|_F^2.
\end{aligned}$$

An overview of the algorithm is depicted in Figure 1. Using the formulation of the equivalence between the TT and Tucker decomposition and incorporating Lemma 1, the

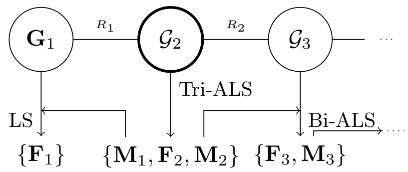


Fig. 2: Zoomed picture of the TRIDENT algorithm for the first three TT-cores.

Tucker factor matrices are estimated after the TT-SVD step. We first estimate both $\mathbf{F}_2$ and $\mathbf{F}_{Q-1}$ as well as the latent variables associated with those factors using an alternating least squares method (ALS) operating on the corresponding CCPD-decomposed TT-cores. Then, by propagating iteratively the latent matrices, we estimate the remaining factors using either a Least Squares (LS) method or an ALS. For $1 < q < \bar{q}$, the matrix factors of the TT-cores are found via a constrained Tri-ALS estimation scheme, each iteration being composed of the following estimation steps

$$\begin{aligned}
\mathbf{M}_{q-1} &= \frac{1}{T_q}\mathbf{G}_q^{(1)}\left(\left(\mathbf{M}_q^{-\mathrm{T}} \odot \mathbf{F}_q\mathbf{\Phi}_{R_{q-1},T_q}\right)^{\mathrm{T}}\right)^{\dagger}\mathbf{\Psi}^{\mathrm{T}}_{R_{q-1},T_q},\\
\mathbf{F}_q &= \frac{1}{R_{q-1}}\mathbf{G}_q^{(2)}\left(\left(\mathbf{M}_q^{-\mathrm{T}} \odot \mathbf{M}_{q-1}\mathbf{\Psi}_{R_{q-1},T_q}\right)^{\mathrm{T}}\right)^{\dagger}\mathbf{\Phi}^{\mathrm{T}}_{R_{q-1},T_q},\\
\mathbf{M}_q^{-\mathrm{T}} &= \mathbf{G}_q^{(3)}\left(\left(\mathbf{F}_q\mathbf{\Phi}_{R_{q-1},T_q} \odot \mathbf{M}_{q-1}\mathbf{\Psi}_{R_{q-1},T_q}\right)^{\mathrm{T}}\right)^{\dagger},
\end{aligned} \tag{7}$$

where we have used the fact that $\mathbf{\Psi}_{a,b}\mathbf{\Psi}_{a,b}^T = b\mathbf{I}_a$ and $\mathbf{\Phi}_{a,b}\mathbf{\Phi}_{a,b}^T = a\mathbf{I}_b$. Similarly, for $\bar{q} < q < Q$, we get

$$\begin{aligned}
\mathbf{M}_{q-1} &= \mathbf{G}_q^{(1)}\left(\left(\mathbf{M}_q^{-\mathrm{T}}\mathbf{\Psi}_{R_q,T_q} \odot \mathbf{F}_q\mathbf{\Phi}_{R_q,T_q}\right)^{\mathrm{T}}\right)^{\dagger},\\
\mathbf{F}_q &= \frac{1}{R_{q+1}}\mathbf{G}_q^{(2)}\left(\left(\mathbf{M}_q^{-\mathrm{T}}\mathbf{\Psi}_{R_q,T_q} \odot \mathbf{M}_{q-1}\right)^{\mathrm{T}}\right)^{\dagger}\mathbf{\Phi}^{\mathrm{T}}_{R_q,T_q},\\
\mathbf{M}_q^{-\mathrm{T}} &= \frac{1}{T_q}\mathbf{G}_q^{(3)}\left(\left(\mathbf{F}_q\mathbf{\Phi}_{R_q,T_q} \odot \mathbf{M}_{q-1}\right)^{\mathrm{T}}\right)^{\dagger}\mathbf{\Psi}^{\mathrm{T}}_{R_q,T_q}.
\end{aligned} \tag{8}$$

Applying this Tri-ALS estimation scheme to the TT-cores $\mathcal{G}_2$ and $\mathcal{G}_{Q-1}$ allows us to estimate the associated Tucker factors and latent matrices. We can then propagate these latent matrices using the coupled LS criterion to perform a Bi-ALS estimation on other cores (following the Tri-ALS scheme described above but skipping the propagated change-of-basis matrix). The propagation of the latent matrices is illustrated in Figure 2. As we do not have a theoretical convergence guarantee for ALS, we use the criterion $\lambda = \frac{\|\langle \widehat{\mathbf{F}}^i \rangle - \langle \widehat{\mathbf{F}}^{i-1} \rangle\|_F}{\|\langle \widehat{\mathbf{F}}^{i-1} \rangle\|_F}$ where i is the current iteration and $\langle \cdot \rangle$ the spanned subspace, to declare convergence. The value of λ is set experimentally to optimize the trade-off between the number of iterations and the reconstruction error. Finally, a SVD is applied to the second unfolding of $\mathcal{G}_{\bar{q}} \times_1 \widehat{\mathbf{M}}_{\bar{q}-1}^{-1} \times_3 \widehat{\mathbf{M}}_{\bar{q}}$, and compute the left singular vectors as the Tucker factor $\widehat{\mathbf{F}}_{\bar{q}}$. A summary of the complete estimation scheme is provided in Algorithm 1.

B. Computational Complexity

The computational complexity of our algorithm is directly related to that of the TT-SVD, which is of order $\mathcal{O}(TI^Q)$ [25] with T the largest multilinear rank, I and Q represent respectively the largest dimension and the order of the original tensor. As the order of the tensor increases the complexity of the ALS becomes insignificant compared to that of the TT-SVD and thus both FMP and TRIDENT have similar computational complexities. On the other hand, the computational complexities of the HOOI and the HOSVD scale linearly by an additional factor Q, resulting in $\mathcal{O}(QTI^Q + N_{\text{iter}}QT^{Q+1})$ and $\mathcal{O}(QTI^Q)$, respectively.

Algorithm 1: TRIDENT

input : Q-order tensor $\mathcal{X}$ and multilinear ranks $\{T_1, ..., T_Q\}$

output: Factor matrices $\mathbf{F}_i$ and core tensor $\mathcal{C}$ of the Tucker decomposition

1 Compute $\bar{q}$ and the TT-ranks using the multilinear ranks and Theorem 1.

2 Compute the TT decomposition:
$\{\mathbf{G}_1, \mathcal{G}_2, ..., \mathcal{G}_{Q-1}, \mathbf{G}_Q\} = \text{TT-SVD}(\mathcal{X})$

3 Compute the factors using the ALS described in Eq. (7) and (8):
$\{\widehat{\mathbf{M}}_1, \widehat{\mathbf{F}}_2, \widehat{\mathbf{M}}_2\} = \text{Tri-ALS}(\mathcal{G}_2)$
$\{\widehat{\mathbf{M}}_{Q-2}, \widehat{\mathbf{F}}_{Q-1}, \widehat{\mathbf{M}}_{Q-1}\} = \text{Tri-ALS}(\mathcal{G}_{Q-1})$
for $3 \leq q \leq \bar{q} - 1$ **do**
  $\{\widehat{\mathbf{F}}_q, \widehat{\mathbf{M}}_q\} = \text{Bi-ALS}(\mathcal{G}_q, \widehat{\mathbf{M}}_{q-1})$
end
for $Q - 1 \geq q \geq \bar{q} + 1$ **do**
  $\{\widehat{\mathbf{F}}_q, \widehat{\mathbf{M}}_q\} = \text{Bi-ALS}(\mathcal{G}_q, \widehat{\mathbf{M}}_{q-1})$
end
$\widehat{\mathbf{F}}_1 = \mathbf{G}_1\widehat{\mathbf{M}}_1$
$\widehat{\mathbf{F}}_Q = \mathbf{G}_Q^{\dagger}\widehat{\mathbf{M}}_{Q-1}$

4 Compute $\widehat{\mathbf{F}}_{\bar{q}}$ as the left singular bases of the second unfolding of $\mathcal{G}_{\bar{q}} \times_1 \widehat{\mathbf{M}}_{\bar{q}-1}^{-1} \times_3 \widehat{\mathbf{M}}_{\bar{q}}$.

5 *(Optional)* Compute the core tensor
$\mathcal{C} = \mathcal{X} \times_1 \mathbf{F}_1^{\dagger} \times_2 ... \times_{\bar{q}} \mathbf{F}_{\bar{q}}^{\mathrm{T}} \times_{\bar{q}+1} ... \times_Q \mathbf{F}_Q^{\dagger}$;

C. Remarks on uniqueness

For $1 < q < \bar{q}$, the model considered in Lemma 1 does not enjoy the essential uniqueness property of the CPD. By adopting the analysis made in [27], we can estimate the factors of the constrained CPD up to the ambiguity relationship $\mathbf{T}_{M_q}^{\mathrm{T}} = \left(\mathbf{T}_{M_{q-1}} \otimes \mathbf{T}_{F_q}\right)^{-1}$, so that $\widehat{\mathbf{F}}_q = \mathbf{F}_q\mathbf{T}_{F_q}, \widehat{\mathbf{M}}_q = \mathbf{M}_q\mathbf{T}_{M_q}, \widehat{\mathbf{M}}_{q-1} = \mathbf{M}_{q-1}\mathbf{T}_{M_{q-1}}$. Indeed, the Tucker model has no essential uniqueness but only uniqueness up to a change-of-basis matrix, which is the case in the considered model. Let $\mathbf{F}$ be a generic Tucker factor and $\mathbf{\Phi}$ its associated contraint matrix, we define $\widetilde{\mathbf{F}} = \widehat{\mathbf{F}}\mathbf{\Phi}^T$. We have $\widetilde{\mathbf{F}}\mathbf{\Phi} = R\mathbf{F}\mathbf{T}_F$. The projector on $\widetilde{\mathbf{F}}$ is given by

$$\widetilde{\mathbf{F}}\mathbf{\Phi}(\widetilde{\mathbf{F}}\mathbf{\Phi})^{\dagger} = R\left(\mathbf{F}\mathbf{T}_F\right)\frac{1}{R}\left(\mathbf{F}\mathbf{T}_F\right)^{\dagger} = \mathbf{F}\mathbf{T}_F\mathbf{T}_F^{-1}\mathbf{F}^{\dagger} = \mathbf{F}\mathbf{F}^{\dagger}.$$

V. Numerical Simulations

In order to assess the performances of the TRIDENT algorithm, we compare it with the HOOI [28], HOSVD [29]

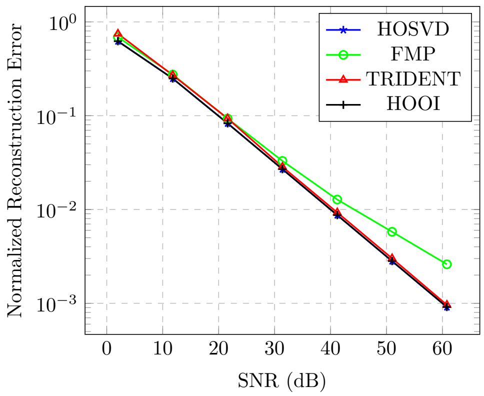
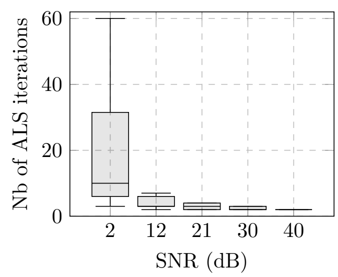


Fig. 3: Normalized reconstruction error of the different methods as a function of the SNR.

Tensor Order	HOOI	HOSVD	FMP	TRIDENT
6	0.19	0.071	0.031	0.037
7	1.48	0.91	0.216	0.224
8	13.6	10.2	1.95	1.96
9	148	106	18.2	18.3

TABLE I: Comparison of time complexity (in s) with increasing order of the tensor with an SNR set at 30dB, an 8-order hypercubic tensor with dimensions equal to 8, and multilinear ranks all equal to 2.



(a) Bi-ALS

(b) Tri-ALS

Fig. 4: Boxplots of the number of iterations for ALS convergence for different values of SNR.

and the FMP [25] for three different experiments. These experiments are carried out for Q-order tensors, where Q ranges from 4 to 9. We look at how closely the algorithm is able to reconstruct the original tensor. In our experiments, we generate the original tensor as $\mathcal{X} = \mathcal{T}/\|\mathcal{T}\|_F + \sigma\mathcal{B}/\|\mathcal{B}\|_F$, where $\mathcal{T}$ is the reconstructed tensor from a single realization of a zero-mean Gaussian distribution of the Tucker factors and $\mathcal{B}$ is an Additive White Gaussian Noise (AWGN). We consider hypercubic tensors with dimensions equal to 8 and multilinear ranks all equal to 2. The results are averaged over 300 independent Monte Carlo runs. We point out that similar experiments have been carried out for several dimensions (ranging from 6 to 10) and ranks (ranging from 2 to 4) and we observed the same conclusions (the results are not displayed in this paper due to space limitations). The code [1] was written in a non-optimized way to ensure a fair comparison.

Reconstruction experiment: We first compare the reconstruction error for different SNRs, defined as SNR=$-10\log_{10}(\sigma^2)$. To measure the reconstruction error, we consider the normalized reconstruction error defined as $\|\mathcal{X} - \hat{\mathcal{X}}\|_F/\|\mathcal{X}\|_F$. The results are shown in Figure 3 considering an 8-order tensor. We can observe that the proposed method achieves similar results to the HOOI and HOSVD while performing better than the competing FMP method at higher SNRs (over 30dB).

Time complexity: In a second experiment, we verify our computational complexity analysis by evaluating the time complexity with an increasing tensor order. The results are averaged over 300 runs with ranks and dimensions as stated at the beginning of Section V. As shown in Table I, the proposed method asymptotically converges to the results of FMP, while the time complexities of HOOI and HOSVD grow much faster. These results consider an SNR=30dB, but the impact of the noise on the time complexity is only significant for lower-order tensors (6-order or less). For higher-order tensors, the time complexity of the TT-SVD is the main limiting factor. For a 9-order tensor, we observe a 6-fold difference between the HOSVD and TRIDENT. When comparing HOOI and TRIDENT, an 8-fold is observed, which is consistent with our analysis considering the non-dominating terms.

[1]https://gitlab.univ-lille.fr/maxence.giraud/trident

Convergence of ALS: We finally investigate the convergence of the ALS stage of our algorithm. More specifically the number of steps until convergence depends on the SNR for both the Tri-ALS and all the Bi-ALS. For these experiments, we used the same conditions as stated at the beginning of section V, as well as an 8-order tensor for the original tensor. The stopping criterion for the ALS is set to $\lambda = 10^{-6}$ which was experimentally verified as the best choice considering the tradeoff between reconstruction error and execution time. The results are reported in Figure 4. We can see that the convergence is achieved within tens of iterations (or even faster) while exhibiting a small variance. We observe a SNR higher than 20dB the convergence is achieved within less than 20 iterations, while for a SNR higher than 5dB, 40 iterations are enough for convergence. Moreover, the variance of the convergence speed (measured in terms of the number of iterations) yields a confident result.

VI. Conclusion

This paper proposed a new JIRAFE-like algorithm to compute the Tucker decomposition with low storage and computational costs. The curse of dimensionality can be mitigated by the proposed TRIDENT estimator, which is based on a new algebraic equivalence between the Tucker model and a structured Tensor Train (TT) decomposition. In particular, we prove that for a Q-order tensor, $Q-3$ TT-cores follow a Constrained CPD model. Our numerical simulations show that the TRIDENT estimator has similar estimation accuracy for most of the SNR range as the HOOI and HOSVD methods with considerably lower time complexity. Perspectives include an adaptation of the proposed algorithm by replacing the TT-SVD with the TT-HSVD to obtain a faster implementation [30] as well as resorting to enhancements to the iterative part of the algorithm using the approaches proposed in [31], [32].

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 Transactions on Signal Processing*, vol. 65, no. 13, pp. 3551–3582, 2017.

[2] Y. Zniyed, R. Boyer, A. L. F. de Almeida, and G. Favier, "Tensor-train modeling for mimo-ofdm tensor coding-and-forwarding relay systems," in *2019 27th European Signal Processing Conference (EUSIPCO)*, 2019, pp. 1–5.

[3] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. PHAN, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 145–163, 2015.

[4] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 145–163, 2015.

[5] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "A constrained factor decomposition with application to mimo antenna systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2429–2442, 2008.

[6] G. T. de Araújo, A. L. F. de Almeida, and R. Boyer, "Channel estimation for intelligent reflecting surface assisted mimo systems: A tensor modeling approach," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 3, pp. 789–802, 2021.

[7] N. Sidiropoulos, R. Bro, and G. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Transactions on Signal Processing*, vol. 48, no. 8, pp. 2377–2388, 2000.

[8] H. Zheng, C. Zhou, Z. Shi, Y. Gu, and Y. D. Zhang, "Coarray tensor direction-of-arrival estimation," *IEEE Transactions on Signal Processing*, vol. 71, pp. 1128–1142, 2023.

[9] H. Zheng, C. Zhou, Z. Shi, and A. L. F. de Almeida, "Subttd: Doa estimation via sub-nyquist tensor train decomposition," *IEEE Signal Processing Letters*, vol. 29, pp. 1978–1982, 2022.

[10] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W.-K. Ma, "Hyperspectral super-resolution: Combining low rank tensor and matrix structure," in *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2018, pp. 3318–3322.

[11] R. A. Borsoi, C. Prévost, K. Usevich, D. Brie, J. C. M. Bermudez, and C. Richard, "Coupled tensor decomposition for hyperspectral and multispectral image fusion with inter-image variability," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 3, pp. 702–717, 2021.

[12] M. Signoretto, L. De Lathauwer, and J. A. Suykens, "A kernel-based framework to tensorial data analysis," *Neural networks*, vol. 24, no. 8, pp. 861–874, 2011.

[13] Y. Panagakis, J. Kossaifi, G. G. Chrysos, J. Oldfield, M. A. Nicolaou, A. Anandkumar, and S. Zafeiriou, "Tensor methods in computer vision and deep learning," *Proceedings of the IEEE*, vol. 109, no. 5, pp. 863–890, 2021.

[14] G. H. Golub and C. F. Van Loan, *Matrix computations*. JHU press, 2013.

[15] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *Journal of Mathematics and Physics*, vol. 6, no. 1-4, pp. 164–189, 1927.

[16] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.

[17] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.

[18] 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 Processing Magazine*, vol. 31, no. 5, pp. 71–79, 2014.

[19] A.-H. Phan, A. Cichocki, A. Uschmajew, P. Tichavský, G. Luta, and D. P. Mandic, "Tensor networks for latent variable analysis: Novel algorithms for tensor train approximation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 11, pp. 4622–4636, 2020.

[20] A. Cichocki, N. Lee, I. V. Oseledets, A.-H. Phan, Q. Zhao, and D. Mandic, "Low-Rank Tensor Networks for Dimensionality Reduction and Large-Scale Optimization Problems: Perspectives and Challenges PART 1," vol. 9, no. 4-5, pp. 249–429.

[21] I. V. Oseledets, "Tensor-train decomposition," *SIAM Journal on Scientific Computing*, vol. 33, no. 5, pp. 2295–2317, 2011.

[22] Y. Liu, *Tensors for Data Processing. Theory, Methods, and Applications*. Elsevier, Academic Press, 2022.

[23] Y. Liu, J. Liu, Z. Long, and C. Zhu, *Tensor computation for data analysis*. Springer, 2022.

[24] A. Cichocki, "Era of big data processing: A new approach via tensor networks and tensor decompositions," *arXiv preprint arXiv:1403.2048*, 2014.

[25] Y. Zniyed, R. Boyer, A. L. de Almeida, and G. Favier, "High-order tensor estimation via trains of coupled third-order CP and Tucker decompositions," *Linear Algebra and its Applications*, vol. 588, pp. 304–337, Mar. 2020.

[26] R. A. Harshman, "Foundations of the PARAFAC procedure: models and conditions for an "explanatory" multi-modal factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, no. 1, p. pages, 1970.

[27] A. L. de Almeida, G. Favier, and J. C. M. Mota, "PARAFAC-based unified tensor modeling for wireless communication systems with application to blind multiuser equalization," *Signal Processing*, vol. 87, no. 2, pp. 337–351, Feb. 2007.

[28] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "On the Best Rank-1 and Rank-(R1,R2,. . .,RN) Approximation of Higher-Order Tensors," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1324–1342, 2020.

[29] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000.

[30] Y. Zniyed, R. Boyer, A. L. F. de Almeida, and G. Favier, "A tt-based hierarchical framework for decomposing high-order tensors," *SIAM Journal on Scientific Computing*, vol. 42, no. 2, pp. A822–A848, 2020.

[31] A.-H. Phan, P. Tichavský, and A. Cichocki, "Fast alternating LS algorithms for high order candecomp/parafac tensor factorizations," *IEEE Transactions on Signal Processing*, vol. 61, no. 19, pp. 4834–4846, 2013.

[32] C. Battaglino, G. Ballard, and T. G. Kolda, "A practical randomized CP tensor decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 39, no. 2, pp. 876–901, 2018.