Multidimensional Harmonic Retrieval Based on Vandermonde Tensor Train

Yassine Zniyed, Remy Boyer, André L. F. de Almeida, Gérard Favier

Multidimensional Harmonic Retrieval Based on Vandermonde Tensor Train[1]

Yassine Zniyed[a], Rémy Boyer[b], André L. F. de Almeida[c], Gérard Favier[d]

[a] *Laboratoire des Signaux et Systèmes (L2S), Université Paris-Sud (UPS), CNRS, CentraleSupélec, Gif-sur-Yvette, France.*
[b] *CRIStAL, Université de Lille, Villeneuve d'Ascq, France.*
[c] *Department of Teleinformatics Engineering, Federal University of Fortaleza, Brazil.*
[d] *Laboratoire I3S, Université Côte d'Azur, CNRS, Sophia Antiplois, France.*

Abstract

Multidimensional Harmonic Retrieval (MHR) is at the heart of important signal-based applications. The exploitation of the large number of available measurement diversities for data fusion increases inexorably the tensor order/dimensionality. The need to mitigate the "curse of dimensionality" in this case is crucial. To efficiently cope with this massive data processing problem, a new scheme called JIRAFE (Joint dImensionality Reduction And Factors rEtrieval) is proposed to estimate the parameters of interest in the MHR problem, namely, the MP angular-frequencies, of the associated P-order rank-M Canonical Polyadic Decomposition (CPD). Our methodology consists of two main steps. The first one breaks the high-order measurement tensor into a collection of graph-connected 3-order tensors, each following a 3-order CPD of rank-M, also called Tensor Train (TT)-cores. This result is based on a model property equivalence between the CPD and the Tensor Train decomposition (TTD) with coupled TT-cores. The second step makes use of a Vandermonde based rectified Alternating Least Squares (RecALS) algorithm to estimate the factors of interest, by enforcing the desired matrix structure. We show that our methodology has several advantages in terms of flexibility, robustness to noise, computational cost and automatic pairing of the parameters of interest with respect to the tensor order.

Keywords: Multidimensional harmonic retrieval, structured Canonical Polyadic Decomposition, Vandermonde factors, High-order tensors, Tensor-Train.

[1] Conference version of part of this work appeared in [1].

Email addresses: yassine.zniyed@l2s.centralesupelec.fr (Yassine Zniyed), remy.boyer@univ-lille.fr (Rémy Boyer), andre@gtel.ufc.br (André L. F. de Almeida), favier@i3s.unice.fr (Gérard Favier)

1. Introduction

Multidimensional Harmonic Retrieval (MHR) [2, 3] is a classical signal processing problem that has found several applications in spectroscopy [4], wireless communications [5], sensor array processing [6, 7], to mention a few. The Multidimensional Harmonic (MH) model can be viewed as the tensor-based generalization of the one-dimensional harmonic one, resulting from the sampling process over a multidimensional regular grid. As a consequence, the MH model can be expressed in tensor form as a constrained Canonical Polyadic Decomposition (CPD) [8, 9, 10] with structured Vandermonde factor matrices. Unlike the sum of M one-dimensional harmonics, which is parametrized by M angular-frequencies only, the P-dimensional harmonic model needs the estimation of a large number (PM) of angular-frequencies of interest. We can easily note that the number of unknown parameters and the order of the associated data tensor grow with P. For instance, in the problem of dual-polarized MIMO (Multiple-Input Multiple-Output) channel estimation, the data tensor order is five [11]. Moreover, it is likely that the joint exploitation of multi-diversity/modality sensing technologies for data fusion [12, 13, 14] further increases the data tensor order. This trend is usually called the "curse of dimensionality" [15, 16, 17] and the challenge here is to reformulate a high-order tensor as a set of low-order tensors. In this context, we observe an increasing interest for the tensor network theory (see [15] and references therein). Tensor network provides a useful and elegant graphical representation of a high-order tensor into a factor graph where the nodes are low-order tensors, called cores, and the edges encode their dependencies, *i.e.*, their common dimensions, often called "rank". In addition, tensor network allows to perform scalable/distributed computations over the cores [15]. In the tensor network framework, Hierarchical/tree Tucker [18, 19] and Tensor Train (TT) [20] are two popular representations of a high-order tensor into a graph-connected low-order (at most 3) tensors. In this work, we focus our effort on the TT formalism for its simplicity and compactness in terms of storage cost. Unlike the hierarchical Tucker model, TT is exploited in many practical and important contexts as, for instance, tensor completion [21], blind source separation [22], and machine learning [23], to mention a few. In the context of the MHR problem, this strategy has at least two advantages. First, it is well-known that the convergence of the Alternating Least Squares (ALS) algorithm becomes more and more difficult when the order increases [24, 25, 26]. To deal with this problem, applying ALS on lower-order tensors is preferable. The second argument is to exploit some latent coupling properties between the cores [27, 23] to propose new efficient estimators.

The Maximum Likelihood estimator [28, 29] is the optimal choice from an estimation point of view, since it is statistically efficient, *i.e.*, its Mean Squared Error (MSE) reaches the Cramér-Rao Bound (CRB) in the presence of noise. The main drawback of the maximum likelihood estimator is its prohibitive complexity cost. This limitation is particularly severe in the context of a high-order data tensor. To overcome this problem, several low- complexity methods can

be found in the literature. These methods may not reach, sometimes, the CRB, but they provide a significant gain in terms of the computational cost compared to the maximum likelihood estimator. There are essentially two main families of methods. The first one is based on the factorization of the data to estimate the well-known signal/noise subspace such as the Estimation of Signal Parameters *via* Rotational Invariance Techniques (ESPRIT) [30], the ND-ESPRIT [31], the Improved Multidimensional Folding technique [32], and the CP-VDM [33]. The second one is based on the uniqueness property of the CPD. Indeed, factorizing the data tensor thanks to the ALS algorithm [34] allows a direct identification of the unknown parameters by inspection of the factor matrices. However, the ALS algorithm [25], as well as most of its variants, do not take into account the Vandermonde structure of the factor matrices. Potentially, discarding this physical *a priori* knowledge on the MH model may degrade the estimation performance of the unknown parameters [35]. So, recently, a new family of estimation methods has been introduced in [36]. This approach, called RecALS for Rectified ALS, modifies the ALS method by integrating a rectification/reinforcement strategy on the Vandermonde structure of the factor matrices.

In this work, we propose a new scheme called JIRAFE for Joint dImensionality Reduction And Factors rEtrieval. JIRAFE is composed of two main steps. The first one is the dimensionality reduction of a high-order harmonic model thanks to a Tensor Train decomposition (TTD) [20]. We show how a model equivalence between the CPD and TTD [1] can be exploited to design an efficient optimization strategy exploiting a sum of coupled Least Squares (LS) problems[2]. The second step is dedicated to the estimation of the unknown parameters using a polynomial rooting procedure. This scheme belongs to the RecALS family and can be viewed as an alternative to the one proposed in [36]. Our main contributions can be summarized as follows:

- **Vandermonde based Tensor Train decomposition (VTTD).** The classic MHR problem is reformulated under a Vandermonde based TT format, instead of the usual CPD formulation.
- **Optimization strategy for VTTD.** The structure of the TT-cores is provided and in particular the latent coupling property existing between the TT-cores is explained and exploited in the context of a sequential optimization strategy of a sum of coupled LS problems. This is the first step of JIRAFE, *i.e.*, the dimensionality reduction step.
- **New rectification strategy for Vandermonde factors.** A new rectification strategy for Vandemonde factors is presented and used in the second step of JIRAFE, *i.e.*, the retrieval step.

In the sequel, the paper is organized as follows. In Section 2, we present the

[2]Note that the idea of rewriting a CPD into the TT format was briefly mentioned in [20, 27], without discussing the structure of the TT-cores resulting from a decomposition algorithm.

MH model as a structured CPD. In Section 3, the TT decomposition and the dimensionality reduction strategy using the TT-SVD algorithm will be detailed. Section 4 gives the algorithmic description of the new proposed retrieval scheme. Simulations showing the effectiveness of the proposed method are given in Section 5, by evaluating the robustness and the computational time of the proposed method compared to some state-of-art algorithms. Section 6 is dedicated to the conclusion.

2. Background and Problem Statement

2.1. Notations

The notations used throughout this paper are the following ones: the superscripts $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^{-1}$, $(\cdot)^{-T}$, and rank$(\cdot)$ denote, respectively, the complex conjugate, the transpose, the inverse, the transpose-inverse and the rank. The Khatri-Rao and outer products are referred to by $\odot$ and $\circ$, respectively. The operator diag$(\cdot)$ forms a diagonal matrix from its vector argument. unfold$_p(\cdot)$ is the p-mode unfolded matrix [25]. The Frobenius norm is denoted by $||\cdot||_F$. Scalars, vectors, matrices and tensors are represented by x, $\mathbf{x}$, $\mathbf{X}$ and $\boldsymbol{\mathcal{X}}$, respectively. $[\mathbf{X}]_{:,n}$ refers to the n-th column of matrix $\mathbf{X}$. $\underline{\mathbf{X}}$ and $\overline{\mathbf{X}}$ are matrices where the last and first rows have been deleted from matrix $\mathbf{X}$ (same definition is applied for vectors). $\boldsymbol{\mathcal{I}}_{p,M}$ denotes the p-order identity tensor of size $M \times \cdots \times M$, and $\boldsymbol{\mathcal{I}}_{2,M} = \mathbf{I}_M$. `CritStop` is a stop criterion for the iterative algorithms that will be chosen in the simulations. We also define the reshape function that will be useful in the sequel. For a P-order tensor $\boldsymbol{\mathcal{X}}$ of size $N_1 \times \cdots \times N_P$, we have the unfolding matrix $\mathbf{X}_{(p)}$ of size $(N_1 \cdots N_p) \times (N_{p+1} \cdots N_P)$ defined as:

$$\mathbf{X}_{(p)} = \text{reshape}\left(\boldsymbol{\mathcal{X}}; \prod_{s=1}^{p} N_s, \prod_{s=p+1}^{P} N_s\right), \tag{1}$$

for a general matrix unfolding formula, we refer to (15) in [37].

Definition 1. The p-mode product $\times_p$ between a P-order tensor $\boldsymbol{\mathcal{A}}$ of size $N_1 \times \cdots \times N_P$ and a matrix $\mathbf{B}$ of size $M_p \times N_p$, is a P-order tensor of size $N_1 \times \cdots N_{p-1} \times M_p \times N_{p+1} \times \cdots \times N_P$ denoted by

$$[\boldsymbol{\mathcal{A}} \times_p \mathbf{B}]_{n_1,\ldots,n_{p-1},m_p,n_{p+1},\ldots,n_P} = \sum_{k=1}^{N_p} [\boldsymbol{\mathcal{A}}]_{n_1,\ldots,n_{p-1},k,n_{p+1},\ldots,n_P} [\mathbf{B}]_{m_p,k}. \tag{2}$$

Definition 2. The contraction product $\times_q^p$ between two tensors $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ of size $N_1 \times \cdots \times N_Q$ and $M_1 \times \cdots \times M_P$, where $N_q = M_p$, is a tensor of order $Q + P - 2$ defined by [15]

$$[\boldsymbol{\mathcal{A}} \times_q^p \boldsymbol{\mathcal{B}}]_{n_1,\ldots,n_{q-1},n_{q+1},\ldots,n_Q,m_1,\ldots,m_{p-1},m_{p+1},\ldots,m_P} \tag{3}$$

$$= \sum_{k=1}^{N_q} [\boldsymbol{\mathcal{A}}]_{n_1,\ldots,n_{q-1},k,n_{q+1},\ldots,n_Q} [\boldsymbol{\mathcal{B}}]_{m_1,\ldots,m_{p-1},k,m_{p+1},\ldots,m_P}. \tag{4}$$

Definition 3. A P-order tensor of size $N_1 \times \ldots \times N_P$ belonging to the family of rank-M CPD [8] admits the following decomposition:

$$\boldsymbol{\mathcal{X}} = \sum_{m=1}^{M} [\mathbf{F}_1]_{:,m} \circ [\mathbf{F}_2]_{:,m} \circ \ldots \circ [\mathbf{F}_P]_{:,m}, \tag{5}$$

or equivalently

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{P,M} \times_1 \mathbf{F}_1 \times_2 \mathbf{F}_2 \times_3 \ldots \times_P \mathbf{F}_P, \tag{6}$$

where the p-th mode factor $\mathbf{F}_p$ is of size $N_p \times M$, $1 \leq p \leq P$.

2.2. Generalized Vandermonde Canonical Polyadic Decomposition

The MH model assumes that the measurements can be modeled as the superposition of M undamped exponentials sampled on a P-dimensional grid according to [2]

$$[\boldsymbol{\mathcal{X}}]_{n_1 \ldots n_P} = \sum_{m=1}^{M} \alpha_m \prod_{p=1}^{P} z_{m,p}^{n_p - 1}, \quad 1 \leq n_p \leq N_p \tag{7}$$

in which the m-th complex amplitude is denoted by α_m and the pole is defined by $z_{m,p} = e^{i\omega_{m,p}}$ where $\omega_{m,p}$ is the m-th angular-frequency along the p-th dimension, and we have $\mathbf{z}_p = \begin{bmatrix} z_{1,p} & z_{2,p} & \ldots & z_{M,p} \end{bmatrix}^T$. Note that the tensor $\boldsymbol{\mathcal{X}}$ is expressed as the linear combination of M rank-1 tensors, each of size $N_1 \times \ldots \times N_P$ (the size of the grid), and follows a generalized Vandermonde CPD [38]:

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{A}} \times_1 \mathbf{V}_1 \times_2 \ldots \times_P \mathbf{V}_P \tag{8}$$

where $\boldsymbol{\mathcal{A}}$ is a $M \times \ldots \times M$ diagonal tensor with $[\boldsymbol{\mathcal{A}}]_{m,\ldots,m} = \alpha_m$ and

$$\mathbf{V}_p = \begin{bmatrix} \mathbf{v}(z_{1,p}) & \ldots & \mathbf{v}(z_{M,p}) \end{bmatrix}$$

is a $N_p \times M$ rank-M Vandermonde matrix, where

$$\mathbf{v}(z_{m,p}) = \begin{bmatrix} 1 & z_{m,p} & z_{m,p}^2 & \ldots & z_{m,p}^{N_p - 1} \end{bmatrix}^T.$$

We define a noisy MH tensor model of order P as

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} + \sigma \boldsymbol{\mathcal{E}} \tag{9}$$

where $\sigma\boldsymbol{\mathcal{E}}$ is the noise tensor, σ is a positive real scalar, and each entry $[\boldsymbol{\mathcal{E}}]_{n_1 \ldots n_P}$ follows an *i.i.d.* circular Gaussian distribution $\mathcal{CN}(0, 1)$, and $\boldsymbol{\mathcal{X}}$ has a canonical rank equal to M.

Remark 1. The addressed MHR problem aims to estimate the PM angular-frequencies correctly paired in terms of "source/dimension" given by the complex argument of the entries of the vectors $\{\mathbf{z}_1, \ldots, \mathbf{z}_P\}$, from the noisy observation tensor $\boldsymbol{\mathcal{Y}}$, when the order of $\boldsymbol{\mathcal{X}}$ is strictly greater than three.

It is important to stress that the paring and estimation operations have to be jointly performed. Indeed, searching the tensor generated from a given scheduling of PM parameters of interest that are the closest to the observed tensor suffers from the well-known combinatorial explosion, especially for large values of P.

3. Dimensionality reduction based on a train of low-order tensors

3.1. Tensor Train Decomposition

Tensor Train was first introduced in [20] for the linear algebra community. The principle of the TTD is to factorize a P-order tensor $\boldsymbol{\mathcal{X}}$ of size $N_1 \times \cdots \times N_P$ into a train of $P-2$ tensors/cores of order 3 and two matrices. As illustrated on Fig 1, the TT-cores are the nodes in the graph and the edges are the common dimensions between two consecutive cores, called TT-ranks. The formal definition of the TTD is given hereafter.

Definition 4. A P-order tensor $\boldsymbol{\mathcal{X}}$ admits a decomposition into a train of tensors if

$$[\boldsymbol{\mathcal{X}}]_{n_1,\cdots,n_P} = \sum_{r_1,\cdots,r_{P-1}=1}^{R_1,\cdots,R_{P-1}} [\mathbf{G}_1]_{n_1,r_1} [\boldsymbol{\mathcal{G}}_2]_{r_1,n_2,r_2} \cdots [\boldsymbol{\mathcal{G}}_{P-1}]_{r_{P-2},n_{P-1},r_{P-1}} [\mathbf{G}_P]_{r_{P-1},n_P}, \tag{10}$$

or equivalently

$$\boldsymbol{\mathcal{X}} = \mathbf{G}_1 \times_2^1 \boldsymbol{\mathcal{G}}_2 \times_3^1 \boldsymbol{\mathcal{G}}_3 \times_4^1 \ldots \times_{P-1}^1 \boldsymbol{\mathcal{G}}_{P-1} \times_P^1 \mathbf{G}_P, \tag{11}$$

where the TT-cores $\mathbf{G}_1$, $\boldsymbol{\mathcal{G}}_p$ $(2 \leq p \leq P-1)$ and $\mathbf{G}_P$ are, respectively, of size $N_1 \times R_1$, $R_{p-1} \times N_p \times R_p$ and $R_{P-1} \times N_P$. $\{R_1, \ldots, R_{Q-1}\}$ are referred to by the TT-ranks with $1 \leq n_p \leq N_p$ and $1 \leq p \leq P$.

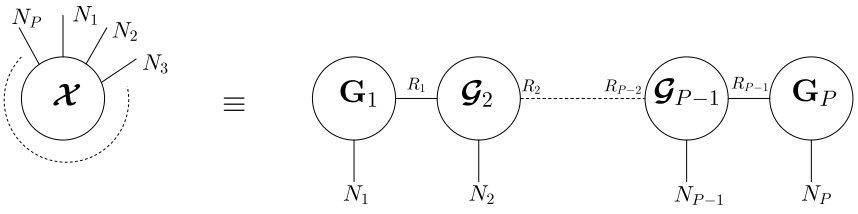

Fig. 1: Graph formalism of the TTD for a P-order tensor

Remark 2. It is important to note that there exists a large number of other graph topologies. In this case, we refer to the tensor network theory (see [15] and references therein), as for instance, the Hierarchical/tree Tucker decomposition

[18, 19]. It appears that, in practice, more complicated graph topologies are not easy to exploit and, to the best of our knowledge, there is no efficient decomposition algorithms in contrast to the TTD.

3.2. Vandermonde based TT decomposition

Due to the simple graph-based formalism of the TT, it is straightforward to rewrite the diagonal tensor $\boldsymbol{\mathcal{A}}$ introduced in eq. (8) as the following TTD

$$\boldsymbol{\mathcal{A}} = \mathbf{A} \times_2^1 \boldsymbol{\mathcal{I}}_{3,M} \times_3^1 \ldots \times_{P-1}^1 \boldsymbol{\mathcal{I}}_{3,M} \times_P^1 \mathbf{I}_M \tag{12}$$

where $\mathbf{A}$ is a $M \times M$ diagonal matrix with $[\mathbf{A}]_{m,m} = \alpha_m$. A graph-based visualization is given on Fig. 2.

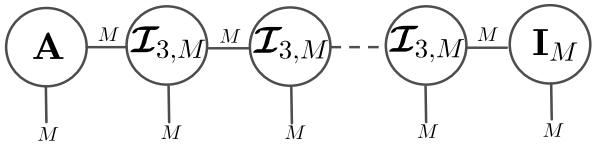

Fig. 2: A possible TTD of tensor $\boldsymbol{\mathcal{A}}$.

Plugging the TTD of tensor $\boldsymbol{\mathcal{A}}$ into eq. (8), leads to

$$\begin{aligned}
\boldsymbol{\mathcal{X}} &= \boldsymbol{\mathcal{A}} \times_1 \mathbf{V}_1 \times_2 \ldots \times_P \mathbf{V}_P && (13)\\
&= \left(\mathbf{A} \times_2^1 \boldsymbol{\mathcal{I}}_{3,M} \times_3^1 \ldots \times_{P-1}^1 \boldsymbol{\mathcal{I}}_{3,M} \times_P^1 \mathbf{I}_M\right) \times_1 \mathbf{V}_1 \times_2 \ldots \times_P \mathbf{V}_P && (14)\\
&= (\mathbf{V}_1\mathbf{A}) \times_2^1 \boldsymbol{\mathcal{V}}_2 \times_3^1 \ldots \times_{P-1}^1 \boldsymbol{\mathcal{V}}_{P-1} \times_P^1 \mathbf{V}_P^T && (15)
\end{aligned}$$

where $\boldsymbol{\mathcal{V}}_p = \boldsymbol{\mathcal{I}}_{3,M} \times_2 \mathbf{V}_p$ is of size $M \times N_p \times M$. A representation of $\boldsymbol{\mathcal{V}}_p$ and $\boldsymbol{\mathcal{X}}$ are given on Figs 3 and 4, respectively.

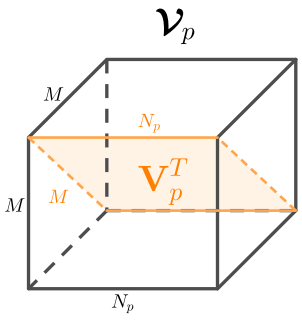

Fig. 3: Representation of the 3-order $M \times N_p \times M$ tensor $\boldsymbol{\mathcal{V}}_p$.

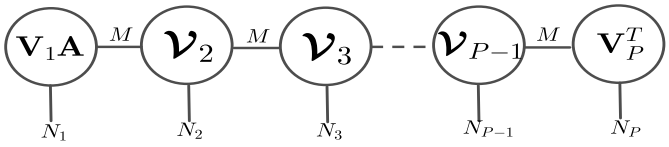

Fig. 4: VTTD of tensor $\boldsymbol{\mathcal{X}}$ corresponding to eq. (15).

According to eq. (15), the generalized Vandermonde CPD of tensor $\boldsymbol{\mathcal{X}}$ is equivalent to a train of $(P-2)$ 3-order tensors that follow a constrained rank-M CPD with a Vandermonde factor on its 2nd mode.

3.3. TTD based on the TT-SVD algorithm

In this section, we present the retrieved TT-cores structure when the TT-SVD algorithm [20] is applied to the original tensor in eq. (8). We first recall that the TT-SVD algorithm is a sequential algorithm that retrieves the TT-cores by extracting the dominant subspaces of the unfolding matrices of the original tensor using the truncated SVD. This algorithm retrieves the original TT-cores up to nonsingular transformation matrices [1]. In Fig. 5, we present the TT-SVD of a 4-order tensor. Note that, at each step, the matrix of right singular vectors is reshaped into a matrix $\mathbf{V}_{(2)}^{(p)}$ having one mode in one dimension and a combination of remaining modes in the other dimension. Applying the SVD to $\mathbf{V}_{(2)}^{(p-1)}$ generates matrices $\mathbf{U}^{(p)}$ and $\mathbf{V}^{(p)}$ containing the left and right singular vectors, respectively. Note that the singular values are absorbed in $\mathbf{V}^{(p)}$. The TT-cores $\boldsymbol{\mathcal{G}}_p$ are recovered from the reshaping of $\mathbf{U}^{(p)}$. In the following theorem, we give the retrieved TT-cores structure when applying the TT-SVD algorithm to the multidimensional harmonic model in eq. (8).

Theorem 1. *Applying the TT-SVD algorithm to eq.* (15) *with full column rank factors, we have the following result:*

$$\mathbf{G}_1 = \mathbf{V}_1\mathbf{A}\mathbf{M}_1^{-1} \tag{16}$$

$$\boldsymbol{\mathcal{G}}_p = \boldsymbol{\mathcal{I}}_{3,M} \times_1 \mathbf{M}_{p-1} \times_2 \mathbf{V}_p \times_3 \mathbf{M}_p^{-T}, \quad \textit{where } 2 \leq p \leq P-1, \tag{17}$$

$$\mathbf{G}_P = \mathbf{M}_{P-1}\mathbf{V}_P^T \tag{18}$$

where $\mathbf{M}_p \in \mathbb{C}^{M\times M}$ *is a nonsingular transformation matrix.*

Proof. Here, we prove the result given in the preliminary article [1]. Our methodology is based on a constructive proof. The aim is to apply the TT-SVD algorithm to the model of interest presented in eq. (7) and to provide the algebraic structure of the TT-cores resulting from the decomposition. Let $\boldsymbol{\mathcal{X}}$ be a P-order rank-M constrained CPD tensor following eq. (7) of size $N_1 \times \cdots \times N_P$ with full column rank factors. Applying the TT-SVD algorithm to $\boldsymbol{\mathcal{X}}$ consists of applying sequentially the SVD to extract the dominant subspaces at each step. In the following, we give the expression of the matrix unfoldings and the SVD factors at each step.

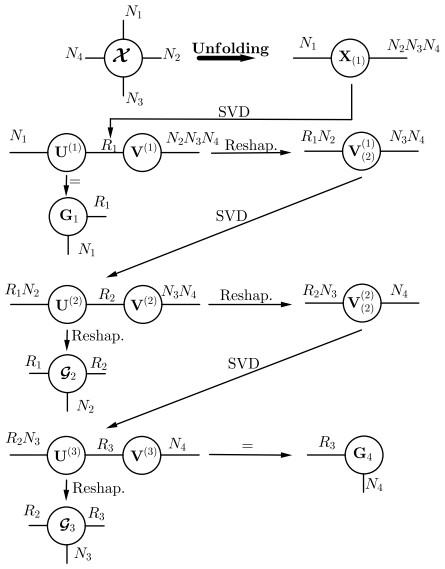

Fig. 5: TT-SVD applied to a 4-order tensor $\boldsymbol{\mathcal{X}}$.

- The first unfolding $\mathbf{X}_{(1)}$ of size $N_1 \times (N_2 \cdots N_P)$ is given by:

$$\mathbf{X}_{(1)} = \text{reshape}\big(\boldsymbol{\mathcal{X}}; N_1, \prod_{s=2}^{P} N_s\big) \tag{19}$$

$$\stackrel{\Delta}{=} \mathbf{V}_1 \mathbf{A} (\mathbf{V}_P \odot \mathbf{V}_{P-1} \odot \cdots \odot \mathbf{V}_2)^T \tag{20}$$

$$\stackrel{\text{SVD}}{=} \mathbf{U}^{(1)} \mathbf{V}^{(1)} \tag{21}$$

where $\mathbf{U}^{(1)}$ and $\mathbf{V}^{(1)}$ are the left and right singular vectors matrices, respectively. We recall that the diagonal singular values matrix is absorbed in $\mathbf{V}^{(1)}$. These matrices can be expressed as

$$\mathbf{U}^{(1)} \mathbf{M}_1 = \mathbf{V}_1 \mathbf{A}, \tag{22}$$

$$\mathbf{V}^{(1)} = \mathbf{M}_1 (\mathbf{V}_P \odot \mathbf{V}_{P-1} \odot \cdots \odot \mathbf{V}_2)^T \tag{23}$$

where $\mathbf{M}_1$ is a $M \times M$ transformation matrix and $\mathbf{A}$ is defined in Section 3.2.
Note that in terms of rank we have the equality

$$\text{rank}\mathbf{X}_{(1)} = \text{rank}(\mathbf{V}_1 \mathbf{A}) = \text{rank}\mathbf{G}_1 = M. \tag{24}$$

It is worth noting that the Khatri-Rao product of matrices $\mathbf{V}_p$ does not decrease the rank [39], *i.e.,* the rank of the Khatri-Rao product matrix also equals M. From eq. (22), the expression of the first TT-core is

$$\mathbf{G}_1 = \mathbf{U}^{(1)} = \mathbf{V}_1 \mathbf{A} \mathbf{M}_1^{-1}. \tag{25}$$

- From eq. (23), reshaping matrix $\mathbf{V}^{(1)}$ provides

$$\mathbf{V}^{(1)}_{(2)} = \text{reshape}\big(\mathbf{V}^{(1)}; MN_2, \prod_{s=3}^{P} N_s\big) \tag{26}$$

$$\triangleq (\mathbf{V}_2 \odot \mathbf{M}_1)(\mathbf{V}_P \odot \mathbf{V}_{P-1} \odot \cdots \odot \mathbf{V}_3)^T \tag{27}$$

$$\overset{\text{SVD}}{=} \mathbf{U}^{(2)}\mathbf{V}^{(2)} \tag{28}$$

where $\mathbf{U}^{(2)}$ and $\mathbf{V}^{(2)}$ are of rank M, and can be expressed as

$$\mathbf{U}^{(2)}\mathbf{M}_2 = \mathbf{V}_2 \odot \mathbf{M}_1, \tag{29}$$

$$\mathbf{V}^{(2)} = \mathbf{M}_2(\mathbf{V}_P \odot \mathbf{V}_{P-1} \odot \cdots \odot \mathbf{V}_3)^T. \tag{30}$$

The second TT-core can then be expressed as:

$$\boldsymbol{\mathcal{G}}_2 = \text{reshape}\big(\mathbf{U}^{(2)}; M, N_2, M\big) \tag{31}$$

$$= \boldsymbol{\mathcal{I}}_{3,M} \times_1 \mathbf{M}_1 \times_2 \mathbf{V}_2 \times_3 \mathbf{M}_2^{-T}. \tag{32}$$

- Following the same reasoning, the expression of $\mathbf{V}^{(p-1)}_{(2)}$ at the p-th step is

$$\mathbf{V}^{(p-1)}_{(2)} = \text{reshape}\big(\mathbf{V}^{(p-1)}; MN_p, \prod_{s=3}^{P} N_s\big) \tag{33}$$

$$\triangleq (\mathbf{V}_p \odot \mathbf{M}_{p-1})(\mathbf{V}_P \odot \mathbf{V}_{P-1} \odot \cdots \odot \mathbf{V}_{p+1})^T \tag{34}$$

$$\overset{\text{SVD}}{=} \mathbf{U}^{(p)}\mathbf{V}^{(p)} \tag{35}$$

where

$$\mathbf{U}^{(p)}\mathbf{M}_p = \mathbf{V}_p \odot \mathbf{M}_{p-1}, \tag{36}$$

$$\mathbf{V}^{(p)} = \mathbf{M}_p(\mathbf{V}_P \odot \mathbf{V}_{P-1} \odot \cdots \odot \mathbf{V}_{p+1})^T, \tag{37}$$

where $\mathbf{M}_p$ is a $M \times M$ transformation matrix and the TT-core expression is given by

$$\boldsymbol{\mathcal{G}}_p = \text{reshape}\big(\mathbf{U}^{(p)}; M, N_p, M\big) \tag{38}$$

$$= \boldsymbol{\mathcal{I}}_{3,M} \times_1 \mathbf{M}_{p-1} \times_2 \mathbf{V}_p \times_3 \mathbf{M}_p^{-T}. \tag{39}$$

- At the last step, and from eq. (37), the SVD factor $\mathbf{V}^{(P-1)}$ and the last TT-core expressions are given by

$$\mathbf{G}_P = \mathbf{V}^{(P-1)} = \mathbf{M}_{P-1}\mathbf{V}_P^T. \tag{40}$$

Theorem 1 can be proved using (25), (32), (39) and (40).

The meaning of the above result is that applying the TT-SVD algorithm to eq. (8) generates 3-order TT-cores that follow a CPD, with TT-ranks all equal to the canonical rank M (see Fig. 6). Furthermore, each CPD-train core $\boldsymbol{\mathcal{G}}_p$ has a common coupled factor with the two consecutive TT-cores $\boldsymbol{\mathcal{G}}_{p-1}$ and $\boldsymbol{\mathcal{G}}_{p+1}$. Theorem 1 shows that applying the TT-SVD algorithm to eq. (8) allows to retrieve the exact TT-cores defined in eq. (15), up to transformation matrices $\mathbf{M}_p$. These matrices $\mathbf{M}_p$ are *change-of-basis* matrices that are related to the estimation of the dominant subspace using the SVD and TT-SVD, as shown in the previous proof.

Remark 3. One should be cautious about reducing the above theorem to the well-known TT model ambiguities. Indeed, each entry of a tensor following a TT model is the product of P matrices, each of which obtained as a reshaping of a TT-core [20]. In fact, the transformation matrices involved in the TT-SVD algorithm coincide with the TT model ambiguities only if the factors are full column rank[3]. Finally, the above theorem shows that the TT-SVD involves lattent but crucial information in the transformation matrices that must be estimated.

Remark 4. Under mild conditions, it is well-known that the factors of a P-order CPD are unique up to trivial ambiguities (common column permutation and scaling [40]). In [1], it is proved that the factors based on eq. (17) can also be estimated with the same trivial ambiguities.

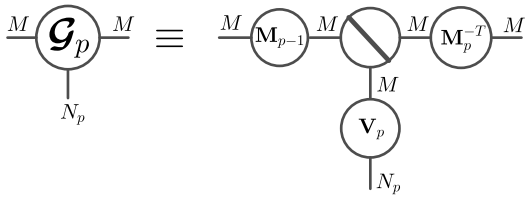

Fig. 6: 3-order CPD of the p-th TT-core. Matrices $\mathbf{M}_{p-1}$ and $\mathbf{M}_p$ are latent quantities.

4. Factor retrieval scheme

4.1. JIRAFE with Vandermonde-based rectification

The proposed estimator is based on the JIRAFE principle. JIRAFE, meaning Joint dImensionality Reduction And Factors rEtrieval, is composed of two main steps.

[3]For a full row rank factor as for instance in wireless communications [11], the TT-cores computed by the TT-SVD algorithm have less intuitive and more complicated expressions. This is the subject of current research.

1. The first one is the computation of the TTD of the initial tensor. By doing this, the initial P-order tensor is broken down into P graph-connected third-order tensors, called TT-cores as represented by Fig. 4. This dimensionality reduction is an efficient way to mitigate the "curse of dimensionality". To reach this goal, the TT-SVD [20] presented above is used as a first step.
2. The second step is dedicated to the factorization of the TT-cores. Recall the main result given by Theorem 1, *i.e.*, if the initial tensor follows a P-order CPD of rank M, then the TT-cores for $2 \leq p \leq P-1$ follow coupled 3-order CPD of rank M. Consequently, the JIRAFE minimizes the following criterion over the physical quantities $\{\mathbf{V}_1, \ldots, \mathbf{V}_P\}$ and over the latent quantities $\{\mathbf{M}_1, \ldots, \mathbf{M}_{P-1}\}$:

$$\mathcal{C} = ||\mathbf{G}_1 - \mathbf{V}_1 \mathbf{A} \mathbf{M}_1^{-1}||_F^2 + ||\mathbf{G}_P - \mathbf{M}_{P-1} \mathbf{V}_P^T||_F^2 \tag{41}$$

$$+ \sum_{p=2}^{P-1} ||\boldsymbol{\mathcal{G}}_p - \boldsymbol{\mathcal{I}}_{3,M} \times_1 \mathbf{M}_{p-1} \times_2 \mathbf{V}_p \times_3 \mathbf{M}_p^{-T}||_F^2. \tag{42}$$

   The above cost function is the sum of coupled LS criterions. The aim of JIRAFE is to recover the original tensor factors using only the 3-order tensors $\boldsymbol{\mathcal{G}}_p$ based on the result given in eq. (17). Equation (42) is expressed as a sum of dependent positive terms due to the coupling properties existing through the matrices $\{\mathbf{M}_1, \ldots, \mathbf{M}_{P-1}\}$ as demonstrated in Theorem 1. Note that the coupling properties between tensors are usually due to physical constraints [41, 12]. In our case, these properties are a result of the sequential structure of the TT-SVD algorithm. Matrices $\{\mathbf{M}_1, \ldots, \mathbf{M}_{P-1}\}$ can be considered as latent matrices, *i.e.*, they do not have a physical meaning, but are essential from an estimation point of view. These remarkable coupling properties take place between a given TT-core and the two other TT-cores connected to it in the graph-based representation (see Fig. 6). Minimizing independently all the positive terms in eq. (42) is a simple procedure but this also means that the structure of the problem of interest is completely eluded. On the other hand, finding jointly $\{\mathbf{V}_1, \ldots, \mathbf{V}_P\}$ and $\{\mathbf{M}_1, \ldots, \mathbf{M}_{P-1}\}$ is not trivial and highly time consuming [42, 12]. Consequently, the JIRAFE approach adopts a straightforward sequential methodology, described in Fig. 7, to minimize the cost function $\mathcal{C}$. Any 3-order algorithm existing in the literature dedicated to the computation of a 3-order CPD can be exploited as for instance the popular ALS algorithm [34]. However, in the context of the MHR problem, the RecALS method introduced in [36] is used and extended. The idea of the RecALS is to associate the ALS algorithm with a Vandermonde rectification strategy. In [36], a link between a Vandermonde vector and the rank-1 factorization of a Toeplitz matrix is proposed. As a cheaper alternative, the Shift Invariant Property (SIP) is proposed and described in the next section.

Based on the new Vandermonde constrained Tensor Train modelisation in

eq. (15), and the CPD structure of the TT-cores in Theorem 1, the idea of the proposed scheme is to replace the estimation of the high P-order tensor by a sequential estimation procedure that operates on 3-order tensors only. We recall that the RecALS algorithm is an ALS-based solution that is efficient for 3-order tensors but becomes a delicate estimator for high-order tensors. In addition, the ALS-based techniques may require several iterations to converge, and convergence is increasingly difficult when the order of the tensor increases, and it is not even guaranteed. In the proposed solution, the use of iterative algorithms, such as ALS, becomes easier after applying a dimensionality reduction to the original tensor using a CPD-train model, since they are applied to (smaller) 3-order tensors. The new proposed solution is called VTT-RecALS algorithm, and its pseudo-code is presented in Algorithm 1. The VTT-RecALS algorithm

Algorithm 1 VTT-RecALS

Input: $\boldsymbol{\mathcal{Y}}$, M, `CritStop`
Output: Estimated parameters: $\{\mathbf{z}_1, \ldots, \mathbf{z}_P\}$.

1: Dimensionality reduction:

$$[\mathbf{G}_1, \boldsymbol{\mathcal{G}}_2, \cdots, \boldsymbol{\mathcal{G}}_{P-1}, \mathbf{G}_P] = \text{TT-SVD}(\boldsymbol{\mathcal{X}}, M).$$

2: Factor retrieval:
3: For $p = 2$,

$$[\hat{\mathbf{M}}_1, \hat{\mathbf{V}}_2, \hat{\mathbf{M}}_2, \mathbf{z}_2] = \text{RecALS}_3(\boldsymbol{\mathcal{G}}_2, M, \texttt{CritStop}).$$

4: **for** $p = 3 \cdots P-1$ **do**
5: $\quad[\hat{\mathbf{V}}_p, \hat{\mathbf{M}}_p, \mathbf{z}_p] = \text{RecALS}_2(\boldsymbol{\mathcal{G}}_p, \hat{\mathbf{M}}_{p-1}, M, \texttt{CritStop})$
6: **end for**
7: $\hat{\mathbf{V}}_1\hat{\mathbf{A}} = \mathbf{G}_1\hat{\mathbf{M}}_1$, and $\hat{\mathbf{V}}_P = \mathbf{G}_P^T\hat{\mathbf{M}}_{P-1}^{-T}$

is actually divided into two parts. The first part is dedicated to dimensionality reduction, *i.e.*, breaking the dimensionality of the high P-order tensor into a set of 3-order tensors using Theorem 1. The second part is dedicated to the factors retrieval from the TT-cores using the RecALS algorithm presented in the next section. It is worth noting that the factors $\hat{\mathbf{V}}_p$ are estimated up to a trivial (common) permutation ambiguity [40]. As noted in [1], since all the factors are estimated up to a unique column permutation matrix, the estimated angular-frequencies are automatically paired.

4.2. Shift Invariance Principle (SIP)

In this section, we propose a new rectification strategy for Vandermonde factors, which is an alternative to Toeplitz Rank-1 Approximation (TR_1A) proposed in [36]. The rectification strategy is called shift invariance principle (SIP),

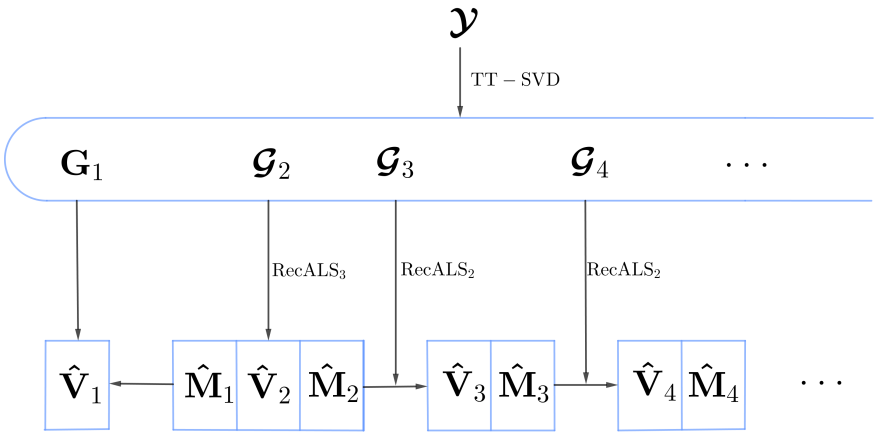

Fig. 7: VTT-RecALS representation

which is inspired from the notion of pencil of matrices (see [43] for instance). It is a rectification strategy that copes with the multidimensionality of the MH model, and is integrated into the VTT-RecALS algorithm to rectify the Vandermonde factors.

4.2.1. The SIP criterion

Note that in a noiseless scenario, each Vandermonde factor matrix $\mathbf{V_p}$ in (8) satisfies the following equality [44, 45]

$$\overline{\mathbf{V}}_p = \underline{\mathbf{V}}_p \text{diag}(\mathbf{z}_p),$$

where $\mathbf{V}_p$ is the p-th factor in eq. (8). Let us now consider the following cost function:

$$\min_{\mathbf{z}_p} \mathcal{C}(\mathbf{z}_p) \quad \text{where} \quad \mathcal{C}(\mathbf{z}_p) = ||\overline{\mathbf{V}}_p - \underline{\mathbf{V}}_p \text{diag}(\mathbf{z}_p)||_F^2 = \sum_{m=1}^{M} P(z_{m,p}) \tag{43}$$

in which $P(z_{m,p}) = ||\overline{\mathbf{v}}(z_{m,p}) - \underline{\mathbf{v}}(z_{m,p}) z_{m,p}||^2$. Minimizing eq. (43) with respect to $\mathbf{z}_p$ is equivalent to minimizing each positive term $P(z_{m,p})$ in the sum. In addition, $P(z_{m,p})$ is not a function of $z_{m',p}$ for $m' \neq m$. Thus, minimizing eq. (43) with respect to $\mathbf{z}_p$ is equivalent to solve M independent problems of the following form:

$$\min_{z_{m,p}} P(z_{m,p}) \quad \text{where} \quad P(z_{m,p}) = -z_{m,p}^* \cdot Q(z_{m,p}) \tag{44}$$

and

$$Q(z_{m,p}) = a_{m,p} z_{m,p}^2 - b_{m,p} z_{m,p} + a_{m,p}^* \tag{45}$$

in which $a_{m,p} = \overline{\mathbf{v}}(z_{m,p})^H \underline{\mathbf{v}}(z_{m,p})$, $b_{m,p} = ||\overline{\mathbf{v}}(z_{m,p})||^2 + ||\underline{\mathbf{v}}(z_{m,p})||^2$, and $Q(z_{m,p})$ is a second degree polynomial where the argument of the two roots is

$$\hat{\omega}_{m,p} = \angle \hat{z}_{m,p} = \angle \left(\frac{b_{m,p} \pm \sqrt{b_{m,p}^2 - 4|a_{m,p}|^2}}{2a_{m,p}} \right) = \angle \left(\frac{1}{a_{m,p}} \right), \qquad (46)$$

where $\hat{\omega}_{m,p}$ is the estimate of the m-th angular-frequency along the p-th dimension.

The result in eq. (46) is integrated in the RecALS$_3$ algorithm. We denote by RecALS$_3$, the RecALS applied to a 3-order tensor, while RecALS$_2$ denotes the RecALS applied to a 3-order tensor using the knowledge of one factor. The RecALS$_3$ algorithm used in Algorithm 1 is summarized in Algorithm 2. RecALS$_2$ has a similar algorithmic description as Algorithm 2, removing step 3, since $\mathbf{M}_{p-1}$ becomes an a priori known input. These algorithms are applied to the resultant 3-order TT-cores to recover the 3 factors with a Vandermonde 2nd mode factor.

Algorithm 2 Rectified Tri-ALS (RecALS$_3$)

Input: $\boldsymbol{\mathcal{G}}_p$, M, `CritStop`

Output: Estimated parameters: $\{\omega_{1,p}, \ldots, \omega_{M,p}\}$.

1: **Initialize:** $\hat{\mathbf{V}}_p$, $\hat{\mathbf{M}}_p^{-T}$
2: **while** `CritStop` is false **do**
3: $\quad \hat{\mathbf{M}}_{p-1} = \text{unfold}_1 \boldsymbol{\mathcal{G}}_p \cdot \left((\hat{\mathbf{M}}_p^{-T} \odot \hat{\mathbf{V}}_p)^T \right)^\dagger$
4: $\quad \hat{\mathbf{V}}_p = \text{unfold}_2 \boldsymbol{\mathcal{G}}_p \cdot \left((\hat{\mathbf{M}}_p^{-T} \odot \hat{\mathbf{M}}_{p-1})^T \right)^\dagger$
5: $\quad$ **for** $m = 1 \cdots M$ **do**
6: $\quad\quad a_{m,p} = \overline{\mathbf{v}}(z_{m,p})^H \underline{\mathbf{v}}(z_{m,p})$
7: $\quad\quad \omega_{m,p} = \angle \left(\frac{1}{a_{m,p}} \right)$
8: $\quad\quad z_{m,p} = e^{i\omega_{m,p}}$
9: $\quad$ **end for**
10: $\quad \hat{\mathbf{V}}_p := [\mathbf{v}(z_{1,p}) \ldots \mathbf{v}(z_{M,p})]$
11: $\quad \hat{\mathbf{M}}_p^{-T} = \text{unfold}_3 \boldsymbol{\mathcal{G}}_p \cdot \left((\hat{\mathbf{V}}_p \odot \hat{\mathbf{M}}_{p-1})^T \right)^\dagger$
12: **end while**

4.2.2. Comparison with other RecALS scheme

It is worth noting that, for each parameter $\omega_{m,p}$, the TR$_1$A method, proposed in [36], for a P-order rank-M tensor of size $N \times \cdots \times N$, is based on (i) the rank-1 diagonalization to obtain the dominant eigen-vector $\mathbf{u}$ of a rank-1 $N \times N$ Toeplitz matrix and (ii) the computation of the angle of the product $[\mathbf{u}]_1 [\mathbf{u}]_2^*$. This cost is evaluated to $O(N+1)$. So, for the entire set of parameters of interest, the final computation cost for the TR$_1$A method is evaluated to

$O(N \cdot P \cdot M + P \cdot M)$. In the SIP methodology, only MP inner products have to be computed. Each inner product implies $N-1$ sums and multiplications, thus the complexity is evaluated to $O(N-1)$. The overall cost of the SIP method is thus $O(N \cdot P \cdot M - P \cdot M)$. We can see that the additional term $O(PM)$ is involved in the TR_1A method with respect to the SIP one. This quantity may be large for high values of P.

5. Simulation results

This section is organized as follows.

- In section 5.1, the interest of the VTT approach is illustrated and studied.
  (a) Paragraph 5.1.1 compares the VTT-RecALS-SIP algorithm with the CPD-RecALS-SIP algorithm to evaluate the interest of using the VTT instead of the CPD.
  (b) Paragraph 5.1.2 studies the impact of the parameters P and N on the estimation accuracy of the VTT-RecALS-SIP algorithm.
- Section 5.2 is dedicated to the comparison of the VTT-RecALS-SIP with the state-of-art estimators.
  (a) In terms of robustness to noise in paragraph 5.2.1.
  (b) In terms of computational time in paragraph 5.2.2.

It is worth noting that the study of the impact of the parameters dimension, rank and order, is important, since they can be related to physical quantities in realistic applications. For instance, in the problem of dual-polarized MIMO channel estimation [11], the dimension, rank and order represent respectively, the number of sensors/antennas at the transmission and reception, the number of dominant paths and the number of spatial diversities at the reception and transmission.
The simulations were performed on a PC equipped with Matlab2016b, an i7, 2.10GHz processor and 8Gb RAM.
Note that the $N \times M$ Vandermonde factors are generated based on a single realization of $\omega_{m,p}$ following a uniform distribution in $]0, \pi[$. Let $f\big(\hat{\boldsymbol{\mathcal{X}}}^{(t)}\big) = ||\boldsymbol{\mathcal{X}} - \hat{\boldsymbol{\mathcal{X}}}^{(t)}||_F$, where $\hat{\boldsymbol{\mathcal{X}}}^{(t)}$ denotes the estimated tensor at the t-th iteration. The convergence test, noted as `CritStop`, for RecALS algorithms is chosen such that $\frac{\left|f\big(\hat{\boldsymbol{\mathcal{X}}}^{(t)}\big) - f\big(\hat{\boldsymbol{\mathcal{X}}}^{(t+1)}\big)\right|}{f\big(\hat{\boldsymbol{\mathcal{X}}}^{(t)}\big)} < \epsilon$, or when the number of iterations exceeds 1000. The signal to noise ratio (SNR) is defined as

$$\text{SNR [dB]} = 10 \log\Big(\frac{||\boldsymbol{\mathcal{X}}||_F^2}{||\sigma \boldsymbol{\mathcal{E}}||_F^2}\Big),$$

with $\boldsymbol{\mathcal{E}}$ drawn from a complex circular *i.i.d.* Gaussian distribution with zero mean and unit variance. The plotted CRB is the one calculated in [46]. The rank M is supposed to be perfectly known in the simulations.
The `MSE` is defined as $\texttt{MSE} = \frac{1}{MC} \sum_{k=1}^{MC} \sum_{p=1}^{P} \sum_{m=1}^{M} (\omega_{m,p} - \hat{\omega}_{m,p}^{(k)})^2$, where $\hat{\omega}_{m,p}^{(k)}$

is the estimation of $\omega_{m,p}$ at the k-th run and MC is the number of Monte-Carlo runs. The depicted MSE is the error over the angular frequencies, and is obtained by averaging the results over 1000 independent Monte Carlo runs, truncated from 5% worst and 5% best MSEs to eliminate ill-convergence experiments and outliers.

5.1. Advantages of the TT approach

5.1.1. VTT over CPD

In this section, we show the interest of using the VTT model over the CPD in terms of noise robustness. The aim of this section is to compare VTT and CPD, through the comparison of VTT-RecALS-SIP and CPD-RecALS-SIP, using the proposed solution SIP presented in Section 4.2.
In Fig. 8, the algorithms VTT-RecALS-SIP and CPD-RecALS-SIP are applied to a 6-order tensor. The rank is fixed at $M = 3$ and the dimension $N = 6$. We can remark that both methods are efficient for a wide range of SNR, with a better robustness for the VTT-RecALS-SIP for low SNR. Similar behavior was found for a rank $M = 2$. We also noticed that for $N = 8$, $M = 2$ and $P = 6$, *i.e.*,

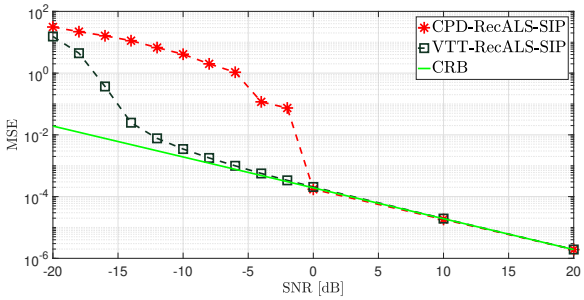

Fig. 8: MSE vs SNR in dB for $P = 6$ with $M = 3$, $N = 6$.

for a lower number (PM) of parameters to be estimated, and a higher number of samples compared to the last experiment, the MSE of the VTT-RecALS-SIP gets closer to the CRB for low SNR, meanwhile the CPD-RecALS-SIP keeps the same behavior. It is worth noting that the effectiveness of VTT-RecALS-SIP over the CPD-RecALS-SIP in the low SNR range can be justified by the noise reduction property of the truncated SVD when the TT-SVD is applied.
Note that the well-known threshold effect in the MSE curves indicates the limit SNR between the two regimes, *i.e.*, when the estimator fails and succeeds to estimate the parameters of interest [47, 48, 49]. So, this is a key quantity to assess the quality of an estimator in a practical context (see for instance [50, 51]). Thanks to the dimensionality reduction step performed here with VTT, this SNR threshold is improved at least by 10 dB compared to the CPD solution in Fig. 8.

5.1.2. Impact of parameters P and N on the VTT-RecALS-SIP

The purpose of next simulations to evaluate the impact of parameters P and N on the behavior of the VTT-RecALS-SIP algorithm. First, in Fig. 9, we fix $P = 6$, $M = 2$, SNR = 5dB and vary the dimension N. Note that the MSE continuously decreases when the dimension is increased, which is predictable, since the number (MP) of parameters is fixed and the number N^P of samples grows with dimension N. In Fig. 10, we fix $N = 6$, $M = 2$, SNR = 5dB and

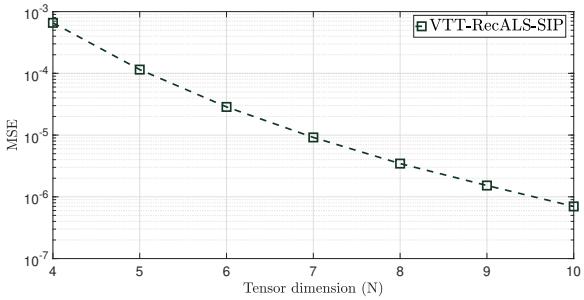

Fig. 9: MSE vs tensor dimension for $P = 6$ with $M = 2$, SNR = 5dB (impact of N).

vary the tensor order P. Here, both, the number (MP) of parameters and N^P of samples grow, but since this latter grows faster with P, the MSE of VTT-RecALS-SIP linearly decreases, which shows that the proposed method becomes more efficient when the order of the MH tensor increases.

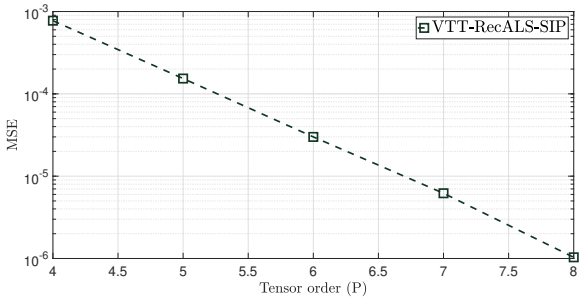

Fig. 10: MSE vs tensor order for $N = 6$ with $M = 2$, SNR = 5dB (impact of P).

5.2. VTT-RecALS-SIP versus the state-of-art estimators

After showing the interest of the VTT-RecALS-SIP over the CPD-RecALS-SIP, in this section, we compare the proposed VTT-RecALS-SIP algorithm with different state-of-art schemes such as the CPD-RecALS-TR_1A [36], the ND-ESPRIT [31] and the CP-VDM [28]. Like for the first experiments, different values of parameters are considered. In Tab. 1, we give the chosen values in

each figure. Note that only one parameter is changed from an experiment to another.

Table 1: Summary of chosen parameters in Section 5.2

	Dimension N	Rank M	Order P
Fig. 11	6	2	4
Fig. 12	6	3	4
Fig. 13	6	2	6
Fig. 14	8	2	6

5.2.1. Robustness to noise

In Fig. 11, we fix $P = 4$, $M = 2$ and $N = 6$. We can remark that the VTT-RecALS-SIP has the MSE closest to the CRB for positive SNR, keeping in mind that $P = 4$ is not a very high order, which is thus not the most interesting case. Increasing the rank $M = 3$, in Fig. 12 compared to Fig.

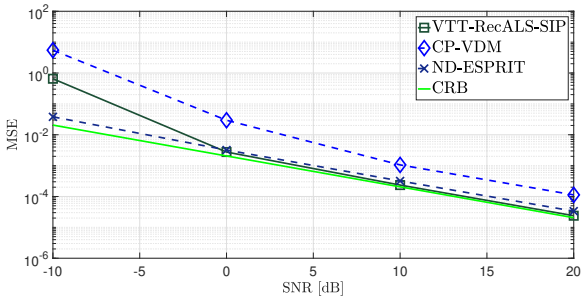

Fig. 11: MSE vs SNR in dB for $P = 4$ with $M = 2$, $N = 6$.

11, and for a relatively "small" order $P = 4$, gives a comparable behavior between the VTT-RecALS-SIP and the ND-ESPRIT algorithms, the difference of the computational time of both methods is evaluated in the next section. Meanwhile the gap with CP-VDM becomes more pronounced. More interesting cases with a higher order P are considered in the next experiments. In the following figures, we choose $P = 6$. In Fig. 13, we fix $M = 2$ and $N = 6$. We remark that, as in the previous section, the VTT-RecALS-SIP becomes more robust when the order increases, and is efficient for a wide range of SNR. This justifies that VTT-RecALS-SIP is well designed for high-order tensors. A last scenario for robustness simulations is with $P = 6$, $N = 8$ and $M = 2$, considered in Fig. 14. The ND-ESPRIT is not depicted here due to its too high computational complexity for high-order tensors. In this figure, we compare the VTT-RecALS-SIP algorithm with CPD-RecALS and CPD-RecALS-TR$_1$A. We recall that the CPD-RecALS is the RecALS algorithm applied to a P-order

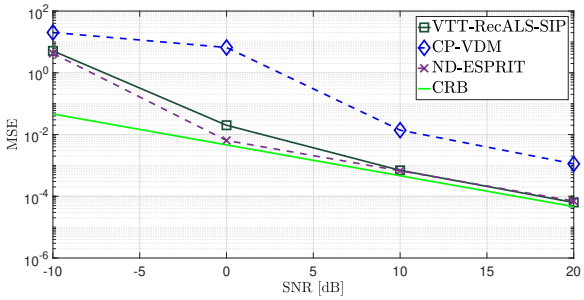

Fig. 12: MSE vs SNR in dB for $P = 4$ with $M = 3$, $N = 6$.

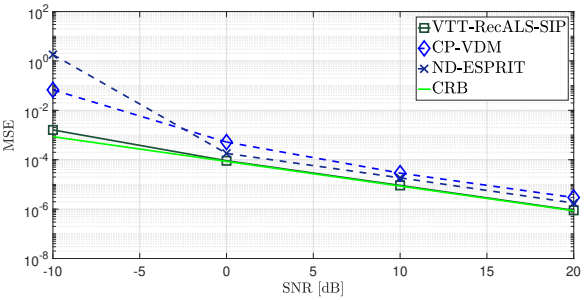

Fig. 13: MSE vs SNR in dB for $P = 6$ with $M = 2$, $N = 6$.

tensor using a naive rectification, by dividing each column by its first entry, to refine the Vandermonde structure, whereas CPD-RecALS-TR$_1$A uses the TR$_1$A rectification. Note that the CPD-RecALS-TR$_1$A is efficient for positive SNR, but is computationally intense, which means that VTT-RecALS-SIP is the best tradeoff between noise robustness and computational complexity in this case.

5.2.2. Computational times

In this section, the computational time is evaluated using the native "Tic-Toc" functions of MatLab. In the following figures, we generate a 6-order rank-2 tensor of size $N_1 \times \cdots \times N_6$, we fix SNR = 5dB, and we vary the number of the measurements as the product of the dimensions $N_1 N_2 \cdots N_6$. In Fig. 15, we start with a tensor of size $4 \times \cdots \times 4$, having 4^6 measurements, and we increase each N_i to $N_i = 6$, having at the end 6^6 measurements. The same methodology is used in Fig. 16, changing each dimension from $N_i = 6$ to $N_i = 8$.
Note that in Fig. 15 the computational time of ND-ESPRIT grows faster than the other algorithms. For larger number of measurements in Fig. 16, the ND-ESPRIT is removed due to its too high computational time. Note that, like all the ALS-based estimators, CPD-RecALS-TR$_1$A has also an intense computational time since it is an iterative algorithm applied to high-order tensors. For example, for a 6-order tensor of size $8 \times \cdots \times 8$ which corresponds to 8^6

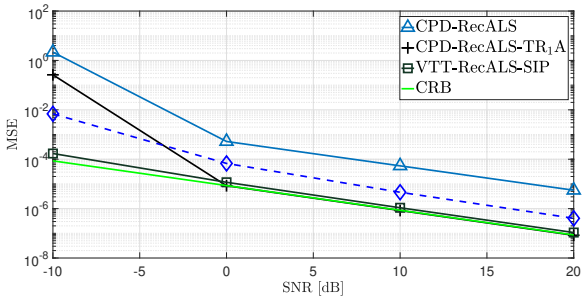

Fig. 14: MSE vs SNR in dB for $P = 6$ with $M = 2$, $N = 8$.

measurements, we have an interesting computational gain of more than 50 for VTT-RecALS-SIP compared to CPD-RecALS-TR$_1$A for the same noise robustness (see Fig. 14). On the other side, we have a comparable computational time for VTT-RecALS-SIP and CP-VDM which is a non-iterative algorithm, unlike VTT-RecALS-SIP. This can be justified by the dimensionality reduction step, and by the fact that the computational time of VTT-RecALS-SIP is approximately the one of TT-SVD algorithm when $P \gg 1$. As a conclusion, we

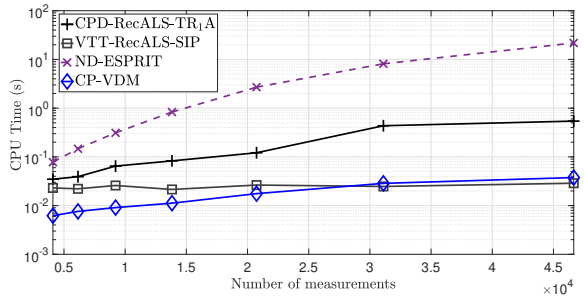

Fig. 15: CPU time versus the number of measurements for $P = 6$, $M = 2$, $N_i = (4, 6)$, and SNR= 5dB.

can say that VTT-RecALS-SIP offers the best tradeoff between noise robustness and computational complexity.

6. Conclusion

Multidimensional Harmonic Retrieval (MHR) is at the heart of many important signal-based applications. The MHR problem admits a natural formulation into the tensor (*a.k.a.* multi-way array) framework, usually called generalized Vandermonde CPD. Joint exploitation of multi-diversity/modality sensing technologies for data fusion increases inexorably the tensor order/dimensionality. Thus, efficient estimation schemes have to face to the well-known "curse of

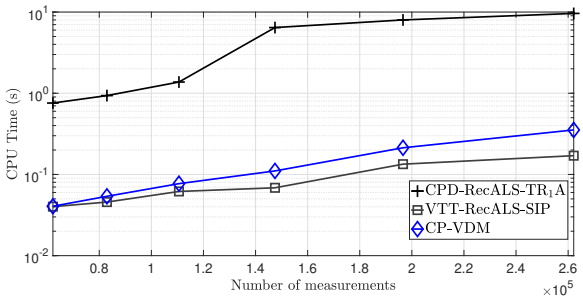

Fig. 16: CPU time versus the number of measurements for $P = 6$, $M = 2$, $N_i = (6, 8)$, and SNR= 5dB.

dimensionality". The challenge here is to reformulate a high-order tensor as a set of low-order tensors, called cores or nodes into the graph-based formalism. Splitting the initial multidimensional optimization problem into a sum of low dimensionality optimization problems for each node of the graph has at least two advantages. Firstly, ill-converging problems for high dimensional optimization are considerably mitigated. Secondly, thanks to the graph-based formalism, some lattent coupling properties between the nodes of the graph can be revealed. As a consequence, new optimization strategies taking the coupling relations into account can be designed. In this work, a new scheme, called VTT-RecALS-SIP, belonging to the JIRAFE (Joint dImensionality Reduction And Factors rEtrieval) family, is proposed for the MHR problem. The aim of the first step of the VTT-RecALS-SIP scheme is to reduce the dimension of the initial Least-Square optimization problem or equivalently to recover the nodes in the popular and simple graph called Tensor Train. We show that due to the MHR problem structure, each node in the TT is associated to a partially structured VTT-core coupled with its two neighbors (in a graph-based sense) VTT-cores. In other words, the initial difficult multidimensional LS optimization problem is now reformulated as a more tractable and flexible equivalent optimization problem, *i.e.*, as the sum of coupled low dimensional LS optimization problems. The second step is dedicated to the Vandermonde-based factor retrieval, *i.e.*, the estimation of the parameters of interest automatically paired. To reach this goal, a new rectified ALS algorithm is proposed and adapted to the exploitation of the coupling properties between the VTT-cores. Specifically, the Vandermonde rectification exploits the Shift Invariance Property (SIP). Numerical simulations show the effectiveness of the proposed VTT-RecALS-SIP method in terms of noise robustness and computational cost compared to other state-of-art methods.

References

[1] Y. Zniyed, R. Boyer, A. L. F. de Almeida and G. Favier, *High-Order CPD estimation with dimensionality reduction using a Tensor Train model*, EU-

SIPCO , Sep 2018, Rome, Italy.

[2] T. Jiang, N. D. Sidiropoulos and J. M. F. T. Berge, *Almost-sure identifiability of multidimensional harmonic retrieval*, IEEE Transactions on Signal Processing, vol. 49, no. 9, pp. 1849-1859, Sep. 2001.

[3] N. D. Sidiropoulos, *Generalizing Carathéodory's uniqueness of harmonic parameterization to N dimensions*, IEEE Transactions on Information Theory, vol. 47, no. 4, pp. 1687-1690, May 2001.

[4] Y. Li, J. Razavilar and K.J.R. Liu, *A high-resolution technique for multidimensional NMR spectroscopy*, IEEE Transactions on Biomedical Engineering, vol.45, pp.78-86, Jan. 1998

[5] D. Nion and N. D. Sidiropoulos, *Tensor algebra and multidimensional harmonic retrieval in signal processing for mimo radar*, IEEE Transactions on Signal Processing, vol. 58, no. 11, pp. 5693-5705, 2010.

[6] N. D. Sidiropoulos, R. Bro and G. B. Giannakis, *Parallel Factor Analysis in Sensor Array Processing*, IEEE Transactions on Signal Processing, vol. 48, pp. 2377-2388, 2000.

[7] M. D. Zoltowski, M. Haardt and C. P. Mathews, *Closed-form 2-D angle estimation with rectangular arrays in element space or beamspace via unitary ESPRIT*, IEEE Transactions on Signal Processing, vol. 44, no. 2, pp. 316-328, Feb. 1996.

[8] 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.

[9] J. Carroll and J.-J. Chang, *Analysis of individual divergences in multidimensional scaling via an n-way generalization of 'Eckart-Young' decomposition*, Psychometrika, vol. 35, pp. 283-319, 1970.

[10] F. L. Hitchcock, *Multiple invariants and generalized rank of a p-way matrix or tensor*, Journal of Mathematics and Physics, vol. 7, pp. 39-79, 1927.

[11] C. Qian, X. Fu, N. D. Sidiropoulos and Ye Yang, *Tensor-Based Channel Estimation for Dual-Polarized Massive MIMO Systems*, arXiv:1805.02223v2, 2018.

[12] E. Acar, R. Bro, and A.K. Smilde. *Data fusion in metabolomics using coupled matrix and tensor factorizations.* Proceedings of the IEEE, vol. 103, no. 9, pp:1602-1620, 2015

[13] R. C. Farias , J.E. Cohen, and P. Comon, *Exploring multimodal data fusion through joint decompositions with flexible couplings.* IEEE Transactions on Signal Processing, vol. 64, no. 18, pp:4830-4844, 2016

[14] D. Lahat, T. Adali and C. Jutten, *Multimodal Data Fusion: An Overview of Methods, Challenges and Prospects*. Proceedings of the IEEE, vol. 103, no. 9, pp.1449-1477, 2015

[15] A. Cichocki, *Era of Big Data Processing: A New Approach via Tensor Networks and Tensor Decompositions*, CoRR http://arxiv.org/pdf/1403.2048.pdf.

[16] A. Phan, A. Cichocki, A. Uschmajew, P. Tichavsky, G. Luta and D. Mandic, *Tensor Networks for Latent Variable Analysis. Part I: Algorithms for Tensor Train Decomposition*, arXiv:1609.09230v1, 2016.

[17] 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.

[18] W. Hackbusch and S. Kühn, *A new scheme for the tensor representation*, Journal of Fourier Analysis and Applications, vol. 15, pp. 706-722, 2009.

[19] I. V. Oseledets and E. E. Tyrtyshnikov, *Breaking the curse of dimensionality, or how to use SVD in many dimensions*, SIAM Journal on Scientific Computing, vol. 31, pp. 3744-3759, 2009.

[20] I. V. Oseledets, *Tensor-train decomposition*, SIAM Journal on Scientific Computing, vol. 33, pp. 2295-2317, 2011.

[21] D. Kressner, M. Steinlechner and B. Vandereycken, *Low-rank tensor completion by Riemannian optimization*, BIT Numerical Mathematics, vol. 54, pp. 447-468, 2014.

[22] M. Boussé and O. Debals and L. De Lathauwer, *A Tensor-Based Method for Large-Scale Blind Source Separation Using Segmentation*, IEEE Transactions on Signal Processing, vol. 65, pp. 346-358, 2016.

[23] N. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. Papalexakis and C. Faloutsos, *Tensor Decomposition for Signal Processing and Machine Learning*, IEEE Transactions on Signal Processing, vol. 65, pp. 3551 - 3582, 2017.

[24] F. Roemer and M. Haardt. *A closed-form solution for Parallel Factor (PARAFAC) Analysis*, In IEEE International Conference on Acoustics, Speech and Signal Processing, 2008.

[25] R. Bro. *PARAFAC. Tutorial and applications*, Chemometrics and Intelligent Laboratory Systems, vol. 38, pp. 149-171, 1997.

[26] N. Li, S. Kindermann and C. Navasca. *Some convergence results on the regularized Alternating Least-Squares method for tensor decomposition*, Linear Algebra and its Applications, vol. 438, pp. 796-812, 2013.

[27] 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*, CoRR, arXiv:1609.00893, 2016.

[28] P. M. Kroonenberg and J. de Leeuw, *Principal component analysis of three-mode data by means of alternating least squares algorithms*, Psychometrika, vol. 45, no. 1, pp. 69-97, Mar. 1980.

[29] M.P. Clark and L.L. Scharf, *Two-dimensional modal analysis based on maximum likelihood*, IEEE Transactions on Signal Processing, vol. 42, no. 6, pp. 1443-1452, Jun. 1994

[30] R. Roy and T. Kailath, *ESPRIT-Estimation of signal parameters via rotational invariance techniques*, IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-37, pp. 984-995, Jul. 1989.

[31] S. Sahnoun, K. Usevich and P. Comon, *Multidimensional ESPRIT for Damped and Undamped Signals: Algorithm, Computations, and Perturbation Analysis*, IEEE Transactions on Signal Processing, vol. 65, pp. 5897-5910, 2017.

[32] J. Liu and X. Liu, *An eigenvector-based approach for multidimensional frequency estimation with improved identifiability*, IEEE Transactions on Signal Processing, vol.54, no.12, pp.4543-4556, Dec. 2006

[33] M. Sørensen and L. De Lathauwer, *Blind Signal Separation via Tensor Decomposition With Vandermonde Factor: Canonical Polyadic Decomposition*, IEEE Transactions on Signal Processing, vol. 61, pp. 5507-5519, 2013.

[34] 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.

[35] J. H. Goulart, M. Boizard, R. Boyer, G. Favier and P. Comon, *Tensor CP Decomposition with structured factor matrices: Algorithms and Performance*, IEEE Journal of Selected Topics in Signal Processing, vol. 10, pp. 757-769, 2016.

[36] R. Boyer and P. Comon, *Rectified ALS Algorithm for Multidimensional Harmonic Retrieval*, Sensor Array and Multichannel Signal Processing Workshop (SAM), 2016.

[37] G. Favier and A.L.F. de Almeida, *Tensor Space-Time-Frequency Coding With Semi-Blind Receivers for MIMO Wireless Communication Systems*, IEEE Transactions on Signal Processing, vol. 62, pp. 5987-6002, 2014.

[38] J. M. Papy, L. De Lathauwer and S. Van Huffel, *Exponential data fitting using multilinear algebra: the single-channel and multi-channel case*, Wiley Online Library, vol. 12, pp. 809-826, 2005.

[39] J. T. Berge, *The k-rank of a Khatri-Rao product*, Unpublished Note, Heijmans Institute of Psychological Research, University of Groningen, The Netherlands.

[40] M. Sørensen and L. De Lathauwer. *New Uniqueness Conditions for the Canonical Polyadic Decomposition of Third-Order Tensors*, SIAM Journal on Matrix Analysis and Applications, vol. 36, pp. 1381-1403, 2015.

[41] E. Acar, T. G. Kolda and D. M. Dunlavy, *All-at-once Optimization for Coupled Matrix and Tensor Factorizations*, arXiv:1105.3422v1, 2011.

[42] J. E. Cohen, Environmental Multiway Data Mining. Signal and Image processing. Université Grenoble Alpes, 2016.

[43] Y. Hua and T. K. Sarkar, *Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise*, IEEE Trans. Acoust., Speech, Signal Processing, vol. 38, no. 5, pp. 814-824, May 1990.

[44] R. Badeau, B. David and G. Richard, *A new perturbation analysis for signal enumeration in rotational invariance techniques*, IEEE Transactions on Signal Processing, vol. 54, pp. 450-458, 2006.

[45] J. M. Papy, L. De Lathauwer and S. Van Huffel, *A Shift Invariance-Based Order-Selection Technique for Exponential Data Modelling*, IEEE Signal Processing Letters, vol. 14, pp. 473-476, 2007.

[46] R. Boyer, *Deterministic asymptotic Cramér-Rao bound for the multidimensional harmonic model*, Signal Processing, Elsevier, vol. 88, pp. 2869-2877, 2008.

[47] A. Renaux, L. Najjar-Atallah, P. Larzabal and P. Forster, *A useful form of the Abel bound and its application to estimator threshold prediction*, IEEE Transactions on Signal Processing, vol. 55, pp. 2365-2369, May 2007.

[48] L. Knockaert, *The Barankin bound and threshold behavior in frequency estimation*, IEEE Transactions on Signal Processing, vol. 45, no. 9, pp. 2398-2401, Sep. 1997.

[49] E. Chaumette, J. Galy, A. Quinlan, P. Larzabal, *A New Barankin Bound Approximation for the Prediction of the Threshold Region Performance of Maximum-Likelihood Estimators*, IEEE Transactions on Signal Processing, vol. 56, no. 11, pp. 5319-5333, Nov. 2008.

[50] N. D. Tran, A. Renaux, R. Boyer, S. Marcos and P. Larzabal, *Weiss-Weinstein bound for MIMO radar with colocated linear arrays for SNR threshold prediction*, Elsevier Signal Processing, May 2012.

[51] C. D. Richmond, *Mean-squared error and threshold SNR prediction of maximum-likelihood signal parameter estimation with estimated colored noise covariances*, IEEE Transactions on Information Theory, vol. 52, no. 5, pp. 2146-2164, May 2006.