Tensor Train Representation of MIMO Channels using the JIRAFE Method

Yassine Zniyed, Remy Boyer, André de Almeida, Gérard Favier

Tensor Train Representation of MIMO Channels using the JIRAFE Method

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), 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

MIMO technology has been subject of increasing interest in both academia and industry for future wireless standards. However, its performance benefits strongly depend on the accuracy of the channel at the base station. In a recent work, a fourth-order channel tensor model was proposed for MIMO systems. In this paper, we extend this model by exploiting additional spatial diversity at the receiver, which induces a fifth order tensor model for the channel. For such high orders, there is a crucial need to break the initial high-dimensional optimization problem into a collection of smaller coupled optimization sub-problems. This paper exploits new results on the equivalence between the canonical polyadic decomposition (CPD) and the tensor train (TT) decomposition for the multi-path scenario. Specifically, we propose a Joint dImensionality Reduction And Factor rEtrieval (JIRAFE) method to find the transmit and receive spatial signatures as well as the complex path gains (which also capture the polarization effects). Monte Carlo simulations show that our proposed TT-based representation of the channel is more robust to noise and computationally more attractive than available competing tensor-based methods, for physical parameters estimation.

Keywords: MIMO systems, CPD, tensor train decomposition, dimensionality reduction, factor retrieval.

1. Introduction

MIMO systems have been subject of intense research due to their great potential to provide substantial energy efficiency and data rate gains [1]. Hence, for MIMO channel modeling and estimation, it is important to accurately estimate path directions in azimuth and elevation, along with the polarization

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)

and amplitude parameters at both sides of the link. In multiuser MIMO systems, the knowledge of the channel parameters at the base station (angles of arrival, angles of departure, path gains, and polarization parameters) can be efficiently exploited to realize beamforming designs and deal with multiuser interference. In this context, adopting a parametric approach to model/estimate the MIMO channel enables the use of limited feedback in frequency division duplexing (FDD) systems to provide the base station with the downlink channel parameters for subsequent transmit signal design. In a recent paper [2], a tensor-based approach for dual-polarized MIMO channel estimation has been proposed by recasting the MIMO channel as a fourth-order tensor. The authors assumed a MIMO system with a uniform rectangular array (URA) at the transmitter (*e.g.* base station) and a uniform linear array (ULA) at the receiver (*e.g.* user equipment). The identifiability of the channel parameters is thoroughly discussed and a channel estimation algorithm is proposed, the core of which relies on the Alternating Least Squares (ALS) algorithm [3]. Despite being an attractive solution, its computational complexity may still be high, especially when downlink channel estimation is carried out at the user equipment with limited processing capabilities. In this work, URA is considered also at the reception. Note that this particular array geometry is highly relevant, not only for wireless communications, but also for modern radio-interferometry-based telescopes [4]. Exploiting azimuth and elevation diversities at both ends of the link increases the tensor order to five [2]. Even if the ALS-based method proposed in [2] remains a possible solution, our approach is inspired by the tensor network theory [5], and more precisely some new results [6, 7, 9] on the equivalence between the CPD and the TTD. Note that Theorem 2 in this work provides a sensibly different result from [6, 7]. Indeed, the ALS algorithm turns out to be often inefficient for high-order tensors. The main drawbacks as illustrated in the simulation part of this work are ill-converging problems [10] and a high computational complexity cost. To mitigate this dimensionality problem, the decomposition of the channel tensor is carried out using a Joint dImensionality Reduction And Factor rEtrieval (JIRAFE) principle [6, 7, 8]. The acronym JIRAFE encompasses a flexible and generic family of algorithms which has already been successfully applied in the context of multidimensional harmonic retrieval [11]. More precisely, the fifth-order channel tensor is first decomposed as a graph-based connected lower-order tensors, called cores [12]. The coupled structure of these cores is described in two scenarios of interest, *i.e.*, when only few (< 4) propagation paths are dominant and the case where the multi-path propagation condition becomes more severe (≥ 4). While the first case is based on some preliminary results given in [6], the core structure for the more challenging situations is described in this work. The second step of the JIRAFE method is dedicated to the factor retrieval for which we exploit the Vandermonde rectification strategy proposed in [13]. Our detailed contributions can be summarized as follows:

1. From a fundamental perspective, the equivalence between CPD and TTD has been presented in [6, 7] for full column rank factors. This equivalence

is deeply reformulated in the sense that the structure of the TT-cores changes if the full column rank factor assumption is violated. This is precisely the case of the MIMO channel tensor considered in this work.

2. Comparatively with the channel model considered in [2], the one proposed in this work exploits an URA at the reception, inducing an increase of spatial diversity. This case has been first mentioned in [2], but in this work we detail the model when URAs are considered at both the transmitter and the receiver.
3. The MIMO channel is represented under a TT format, instead of the usual CPD representation, and the structure of the TT-cores is highlighted separately under the assumptions of full column rank and full row rank for the matrix factors.
4. The TT structure characterized by properties of coupling between two adjacent cores containing the same latent matrices, is exploited for dimensionality reduction and channel parameters estimation using the JIRAFE (Joint dImensionality Reduction And Factor rEtrieval) scheme.

2. Tensor-based channel modeling

2.1. Canonical Polyadic Decomposition (CPD) of the channel tensor

2.1.1. Expression of the channel tensor

The steering vectors for the k-th path for an URA in transmission of size $M_{T_x} \times M_{T_y}$ and in reception of size $M_{R_x} \times M_{R_y}$ are respectively, $\mathbf{a}_T(k) = \mathbf{a}_{T_x}(k) \otimes \mathbf{a}_{T_y}(k)$, $\mathbf{a}_R(k) = \mathbf{a}_{R_x}(k) \otimes \mathbf{a}_{R_y}(k)$ where

$$\mathbf{a}_X(k) = [1, \exp(j\omega_X(k)), \cdots, \exp(j\omega_X(k)(M_X - 1))]^T \tag{1}$$

with $X \in \{T_x, T_y, R_x, R_y\}$. Note that $\otimes$ and $\odot$ denote respectively Kronecker and Khatri-Rao products. The steering matrices for K paths in transmission and in reception are respectively, $\mathbf{A}_T = \mathbf{A}_{T_x} \odot \mathbf{A}_{T_y}$ and $\mathbf{A}_R = \mathbf{A}_{R_x} \odot \mathbf{A}_{R_y}$ in which

$$\mathbf{A}_X = \begin{bmatrix}\mathbf{a}_X(1) & \cdots & \mathbf{a}_X(K)\end{bmatrix}.$$

Now, define $\beta_k^{(p,q)}$ as the k-th entry of the vector $\boldsymbol{\beta}^{(p,q)}$, with $1 \leq k \leq K$, where $\beta_k^{(p,q)}$ is the generalized (complex) path-loss parameter for the k-th path and for the (p,q)-th subchannel. Note that $p \in \{V_r, H_r\}$ refers to the vertical (V) polarized and horizontal (H) polarized receive antennas, and $q \in \{V_t, H_t\}$ refers to the V-polarized and H-polarized transmit antennas. In the noise-free scenario, the channel matrix is given by

$$\mathbf{H} = (\mathbf{A}^*_{T_x} \odot \mathbf{A}^*_{T_y} \odot \mathbf{A}_{R_x} \odot \mathbf{A}_{R_y})\mathbf{B}^T$$

with $\mathbf{B} = [\boldsymbol{\beta}^{(V_r,V_t)}\boldsymbol{\beta}^{(V_r,H_t)}\boldsymbol{\beta}^{(H_r,V_t)}\boldsymbol{\beta}^{(H_r,H_t)}]^T \in \mathbb{C}^{4\times K}$. From this matrix unfolding of the channel tensor $\boldsymbol{\mathcal{H}}$, we can conclude that it follows a fifth-order CPD of canonical rank K, and size $M_{T_x} \times M_{T_y} \times M_{R_x} \times M_{R_y} \times 4$, given by

$$\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{I}}_{5,K} \times_1 \mathbf{A}^*_{T_x} \times_2 \mathbf{A}^*_{T_y} \times_3 \mathbf{A}_{R_x} \times_4 \mathbf{A}_{R_y} \times_5 \mathbf{B} + \boldsymbol{\mathcal{N}}. \tag{2}$$

The additive term $\boldsymbol{\mathcal{N}}$ encompasses the background noise and the estimation error due to the pre-estimation of the unstructured channel obtained by sending known pilot sequences from the transmit antennas. In the massive MIMO case, using orthogonal pilot sequences would lead to very long pilot sequences of the order of the number of the transmit antennas. However, as shown recently in [14], Kronecker-structured pilots can be used, which allows to significantly reduce the length of the pilot sequences. The noise tensor $\boldsymbol{\mathcal{N}}$ can then be modeled as zero-mean circularly complex Gaussian random variables.

Note that the considered channel tensor model is an extension of the one presented in [2] due to the URA assumption at both the transmitter and the receiver. We draw attention to the fact that the focus of our work is on the extraction of the MIMO channel parameters based on the unstructured noisy channel tensor.

2.1.2. Model assumptions

For identifiability concerns, we assume the following constraints.

1. The steering matrices are all of full column rank, which implies $K \leq \min\{M_{T_x}, M_{T_y}, M_{R_x}, M_{R_y}\}$[1].
2. Two scenarios in terms of the number of propagation paths are of interest:
   (a) $K < 4$ (few dominant paths), then $\mathbf{B}$ is a full column rank matrix.
   (b) $K \geq 4$, then $\mathbf{B}$ is a full row rank factor matrix.

2.2. Tensor train decomposition (TTD) of the channel tensor

The idea of the TTD [12] is to break the dimensionality/order of $\boldsymbol{\mathcal{H}}$ into 3-order TT-cores and two matrices according to

$$\boldsymbol{\mathcal{H}} \stackrel{\text{TTD}}{=} \mathbf{G}_1 \times_2^1 \boldsymbol{\mathcal{G}}_2 \times_3^1 \boldsymbol{\mathcal{G}}_3 \times_4^1 \boldsymbol{\mathcal{G}}_4 \times_5^1 \mathbf{G}_5, \tag{3}$$

where $\mathbf{G}_1 \in \mathbb{C}^{M_{T_x} \times K}$, $\boldsymbol{\mathcal{G}}_2 \in \mathbb{C}^{K \times M_{T_y} \times K}$, $\boldsymbol{\mathcal{G}}_3 \in \mathbb{C}^{K \times M_{R_x} \times K}$, $\boldsymbol{\mathcal{G}}_4 \in \mathbb{C}^{K \times M_{R_y} \times K}$, and $\mathbf{G}_5 \in \mathbb{C}^{K \times 4}$. The product $\times_p^q$ is defined as in [6]. As mentioned in the previous section, $\boldsymbol{\mathcal{H}}$ follows a 5-order CPD of canonical rank-K. In this context, the TT-cores can be analytically relied to the desired factors of the CPD. This is the subject of the next section.

3. Joint Dimensionality Reduction and Factor Retrieval (JIRAFE)

3.1. JIRAFE: dimensionality reduction

In the following, we present two theorems on the structure of the TT-cores resulting from the TT-SVD [12] algorithm applied to eq. (2).

[1] The assumption of a small number K of paths compared to the number of transmit/receive antennas is usually made in massive MIMO scenarios, especially in millimeter-wave systems due to the poor scattering propagation and the high number of antennas [15, 16].

1. Theorem 1 is a generalization of the result given in [6], *i.e.*, when there exists few ($K < 4$) dominant propagation paths. In this case, all the factors are full-column rank.
2. Theorem 2 modifies Theorem 1 for the more challenging scenario where there is four or more propagation paths. In this case, the last factor is full-row rank.

Theorem 1. *When all the factors are full-column rank ($K < 4$), the TT-cores are given by*

$$\mathbf{G}_1 = \mathbf{A}_{T_x}^* \mathbf{M}_1^{-1}, \quad \boldsymbol{\mathcal{G}}_2 = \boldsymbol{\mathcal{I}}_{3,K} \times_1 \mathbf{M}_1 \times_2 \mathbf{A}_{T_y}^* \times_3 \mathbf{M}_2^{-T}, \tag{4}$$

$$\boldsymbol{\mathcal{G}}_3 = \boldsymbol{\mathcal{I}}_{3,K} \times_1 \mathbf{M}_2 \times_2 \mathbf{A}_{R_x} \times_3 \mathbf{M}_3^{-T}, \tag{5}$$

$$\boldsymbol{\mathcal{G}}_4 = \boldsymbol{\mathcal{I}}_{3,K} \times_1 \mathbf{M}_3 \times_2 \mathbf{A}_{R_y} \times_3 \mathbf{M}_4^{-T}, \quad \textit{and} \;\; \mathbf{G}_5 = \mathbf{M}_4 \mathbf{B}^T \tag{6}$$

where, for $1 \leq k \leq 4$, $\mathbf{M}_k \in \mathbb{C}^{K \times K}$ are nonsingular transformation matrices. This means that the TT-ranks are all equal to the canonical rank K, where K is the number of paths.

Theorem 2. *If the last factor is full row rank,* i.e., *$K \geq 4$, then the TT-cores $\{\mathbf{G}_1, \boldsymbol{\mathcal{G}}_2, \boldsymbol{\mathcal{G}}_3\}$ verify the same factorizations as in Theorem 1 but the two last TT-cores are given by:*

$$\boldsymbol{\mathcal{G}}_4 = \boldsymbol{\mathcal{I}}_{3,K} \times_1 \mathbf{M}_3 \times_2 \mathbf{A}_{R_y} \times_3 \mathbf{M}^{-T}\mathbf{B}, \quad \textit{and} \quad \mathbf{G}_5 = \mathbf{M} \tag{7}$$

where $\mathbf{M} \in \mathbb{C}^{4 \times 4}$ is a nonsingular transformation matrix. This means that the TT-ranks are equal to $(K, K, K, 4)$.

PROOF. Both theorems rely on constructive proofs based on the algebraic structure of the TT-SVD algorithm applied to a 5-order CPD tensor. Depending on the rank of $\mathbf{B}$, the reasoning in both cases will be the same for all but the two last TT-cores. We recall that $\mathbf{M}_k$ are *change-of-basis* matrices that appear due to the use of the SVD to extract dominant subspaces [6, 7].

Remark 1. The two above theorems show that the TT-core structure mixes physical quantities, *i.e.*, $\{\mathbf{A}_{T_x}^*, \mathbf{A}_{T_y}^*, \mathbf{A}_{R_x}, \mathbf{A}_{R_y}, \mathbf{B}\}$ and latent matrices, *i.e.*, $\{\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \mathbf{M}_4, \mathbf{M}\}$, and that each TT-core is coupled with its two neighbor TT-cores *via* the latent matrices.

3.2. JIRAFE: CPD factors retrieval

When the TT-cores have been estimated thanks to the TT-SVD algorithm for instance, the aim of the second step of JIRAFE is to propose an estimation strategy exploiting the TT-core structures given in Theorems 1 and 2. The use of the TTD representation and the JIRAFE approach is mainly motivated by three reasons: (i) it allows to break the dimensionality of the original 5-order tensor to smaller 3-order tensors using closed-form solutions, (ii) the parameters estimation can be done using the smaller 3-order tensors when their structure is derived with a lower computational cost, and (*iii*) we will show in the next section that the JIRAFE approach will have a better robustness compared to other state-of-art algorithms. We refer to [7] for more details about the TTD and the JIRAFE approach.

3.2.1. Few dominant paths scenario ($K < 4$)

When $K < 4$, we have to minimize the following criterion with respect to the physical quantities $\{\mathbf{A}^*_{T_x}, \mathbf{A}^*_{T_y}, \mathbf{A}_{R_x}, \mathbf{A}_{R_y}, \mathbf{B}\}$ and the latent quantities $\{\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \mathbf{M}_4\}$:

$$\mathcal{C}_1 = ||\hat{\mathbf{G}}_1 - \mathbf{A}^*_{T_x}\mathbf{M}_1^{-1}||_F^2 + ||\hat{\mathbf{G}}_5 - \mathbf{M}_4\mathbf{B}^T||_F^2 \tag{8}$$
$$+ ||\hat{\boldsymbol{\mathcal{G}}}_2 - \boldsymbol{\mathcal{I}}_{3,K} \times_1 \mathbf{M}_1 \times_2 \mathbf{A}^*_{T_y} \times_3 \mathbf{M}_2^{-T}||_F^2 \tag{9}$$
$$+ ||\hat{\boldsymbol{\mathcal{G}}}_3 - \boldsymbol{\mathcal{I}}_{3,K} \times_1 \mathbf{M}_2 \times_2 \mathbf{A}_{R_x} \times_3 \mathbf{M}_3^{-T}||_F^2 \tag{10}$$
$$+ ||\hat{\boldsymbol{\mathcal{G}}}_4 - \boldsymbol{\mathcal{I}}_{3,K} \times_1 \mathbf{M}_3 \times_2 \mathbf{A}_{R_y} \times_3 \mathbf{M}_4^{-T}||_F^2. \tag{11}$$

This criterion is the sum of coupled LS criteria. At this point, we choose deliberately to promote a local/sequential (but fast) optimization method (see Algorithm 1) instead of a global/optimal optimization strategy based for instance on the Lagrangian minimization [17]. In Algorithm 1, we denote by Tri-ALS, the ALS algorithm applied to a 3-order tensor, the acronym TR_1A stands Toeplitz Rank-1 Approximation and is dedicated to a Vandermonde rectification strategy presented in [13], while KRF denotes a non-iterative method called Khatri-Rao Factorization proposed in [18]. KRF algorithm recovers 3-order CPD factors assuming that one factor is known and full column rank. It computes K SVDs of rank-one matrices to recover the remaining two other factors. Since the focus of this work is on the new TT representation for the channel tensor $\boldsymbol{\mathcal{H}}$, we could adopt any state-of-art method for estimating the channel, such as least squares or matched filtering techniques.

Algorithm 1 JIRAFE for few dominant propagation paths

Input: 5-order rank-K tensor $\boldsymbol{\mathcal{H}}$, TT-ranks: (K, K, K, K).
Output: Estimated CPD factors: $\hat{\mathbf{A}}^*_{T_x}, \hat{\mathbf{A}}^*_{T_y}, \hat{\mathbf{A}}_{R_x}, \hat{\mathbf{A}}_{R_y}, \hat{\mathbf{B}}$.

1: <u>Dimensionality reduction:</u> $[\hat{\mathbf{G}}_1, \hat{\boldsymbol{\mathcal{G}}}_2, \hat{\boldsymbol{\mathcal{G}}}_3, \hat{\boldsymbol{\mathcal{G}}}_4, \hat{\mathbf{G}}_5] \hookleftarrow \text{TT-SVD}(\boldsymbol{\mathcal{H}}, R)$,
2: <u>CPD factors retrieval</u>: $[\hat{\mathbf{M}}_1, \hat{\mathbf{A}}^*_{T_y}, \hat{\mathbf{M}}_2^{-T}] \hookleftarrow \text{Tri-ALS}(\hat{\boldsymbol{\mathcal{G}}}_2, K)$,
3: $[\hat{\mathbf{A}}_{R_x}, \hat{\mathbf{M}}_3^{-T}] \hookleftarrow \text{KRF}(\hat{\boldsymbol{\mathcal{G}}}_3, \hat{\mathbf{M}}_2, K)$,
4: $[\hat{\mathbf{A}}_{R_Y}, \hat{\mathbf{M}}_4^{-T}] \hookleftarrow \text{KRF}(\hat{\boldsymbol{\mathcal{G}}}_4, \hat{\mathbf{M}}_3, K)$,
5: $\hat{\mathbf{A}}^*_{T_x} = \hat{\mathbf{G}}_1\hat{\mathbf{M}}_1$, and $\hat{\mathbf{B}} = \hat{\mathbf{G}}_5^T\hat{\mathbf{M}}_4^{-T}$;
6: <u>Rectification:</u> $[\hat{\mathbf{A}}^*_{T_x}, \hat{\mathbf{A}}^*_{T_y}, \hat{\mathbf{A}}_{R_x}, \hat{\mathbf{A}}_{R_y}] \hookleftarrow \text{TR}_1\text{A}(\hat{\mathbf{A}}^*_{T_x}, \hat{\mathbf{A}}^*_{T_y}, \hat{\mathbf{A}}_{R_x}, \hat{\mathbf{A}}_{R_y})$.

Remark 2. Note that the steering factors have a Vandermonde structure. It then makes sense to use a Tri-ALS algorithm that takes into account the structure of these factors. In the simulations, we will use a class of ALS-based methods that is called RecALS, for Rectified ALS [13]. Note that other methods for angle estimation could also be applied. The chosen RecALS method is a 3-order ALS that integrates a rectification strategy of the Vandermonde structure of the factor matrices. This strategy is also applied on the Vandermonde factors resulting from the KRF algorithm. In the simulations, and considering the use of TR_1A to rectify the Vandermonde structure, we will call the proposed method JIRAFE.

3.2.2. More general multi-path scenario ($K \geq 4$)

Based on Theorem 2, we propose a second algorithm which minimizes the following criterion over the physical quantities $\{\mathbf{A}^*_{T_x}, \mathbf{A}^*_{T_y}, \mathbf{A}_{R_x}, \mathbf{A}_{R_y}, \mathbf{B}\}$ and over the latent quantities $\{\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \mathbf{Q}\}$:

$$\mathcal{C}_2 = ||\hat{\mathbf{G}}_1 - \mathbf{A}^*_{T_x}\mathbf{M}_1^{-1}||_F^2 + ||\hat{\mathbf{G}}_5 - \mathbf{Q}^{\dagger T}\mathbf{B}^T||_F^2 \tag{12}$$

$$+ ||\hat{\boldsymbol{\mathcal{G}}}_2 - \boldsymbol{\mathcal{I}}_{3,K} \times_1 \mathbf{M}_1 \times_2 \mathbf{A}^*_{T_y} \times_3 \mathbf{M}_2^{-T}||_F^2 \tag{13}$$

$$+ ||\hat{\boldsymbol{\mathcal{G}}}_3 - \boldsymbol{\mathcal{I}}_{3,K} \times_1 \mathbf{M}_2 \times_2 \mathbf{A}_{R_x} \times_3 \mathbf{M}_3^{-T}||_F^2 \tag{14}$$

$$+ ||\hat{\boldsymbol{\mathcal{G}}}_4 - \boldsymbol{\mathcal{I}}_{3,K} \times_1 \mathbf{M}_3 \times_2 \mathbf{A}_{R_y} \times_3 \mathbf{Q}||_F^2 \tag{15}$$

where $\mathbf{Q} = \mathbf{M}^{-T}\mathbf{B}$ and $\hat{\mathbf{G}}_5 = \mathbf{M}$.

Note that the Algorithm 2 is deduced from Algorithm 1 by replacing the TT-ranks by $(K, K, K, 4)$ in the dimensionality reduction. In addition, lines 4 and 5 become $[\hat{\mathbf{A}}_{R_Y}, \hat{\mathbf{Q}}] \hookleftarrow \text{KRF}(\hat{\boldsymbol{\mathcal{G}}}_4, \hat{\mathbf{M}}_3, 4)$ and $\hat{\mathbf{A}}^*_{T_x} = \hat{\mathbf{G}}_1\hat{\mathbf{M}}_1$, and $\hat{\mathbf{B}} = \hat{\mathbf{G}}_5^T\hat{\mathbf{Q}}$, respectively.

It is worth noting that it has been proven in [6, 7] that the JIRAFE method estimates the CPD factors up to the same trivial ambiguities as for the ALS algorithm. In contrast to the scheme presented in [6], the proposed JIRAFE algorithm replaces the Bi-ALS algorithm by the non-iterative KRF estimator, which allows to mitigate potential ill-convergence problems.

4. Simulation Results

In this section, we show the interest of using the TTD of Section 2.2 over the CPD through the JIRAFE-based proposed algorithms. The Vandermonde factors $\mathbf{A}_{T_x}, \mathbf{A}_{T_y}, \mathbf{A}_{R_x}$, and $\mathbf{A}_{R_y}$ are generated, respectively, based on single random realizations of the angular frequencies $\omega_{T_x}(k), \omega_{T_y}(k), \omega_{R_x}(k)$ and $\omega_{R_y}(k)$ following a uniform distribution in $]0, \pi]$. The factor $\mathbf{B}$ is drawn from a complex Gaussian distribution with zero mean and unit variance. The scenario for the simulations consists of a receiver and a transmitter both with 10×8 URAs, implying that $\boldsymbol{\mathcal{H}}$ satisfies a 5-order rank-K CPD, of dimensions $10 \times 8 \times 10 \times 8 \times 4$. For the simulations, we consider the noisy channel tensor $\boldsymbol{\mathcal{H}}$ in eq. (2). This tensor can be given using a pre-estimation step based on a supervised approach such as the one proposed in [14]. The considered MSE concerns the estimation

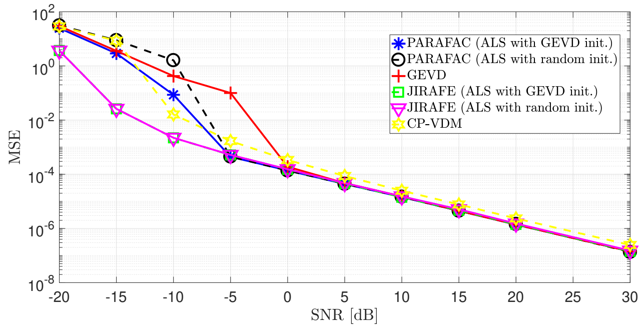

Fig. 1: MSE vs SNR in dB with Alg. 1 for $K = 3$.

error over the angular frequencies, *i.e.*,

$$\texttt{MSE} = \sum_{k=1}^{K} \Big(\big(\omega_{T_x}(k) - \hat{\omega}_{T_x}(k)\big)^2 + \big(\omega_{T_y}(k) - \hat{\omega}_{T_y}(k)\big)^2 \tag{16}$$

$$+ \big(\omega_{R_x}(k) - \hat{\omega}_{R_x}(k)\big)^2 + \big(\omega_{R_y}(k) - \hat{\omega}_{R_y}(k)\big)^2 \Big), \tag{17}$$

the signal to noise ratio (SNR) is defined as

$$\text{SNR [dB]} = 10 \log \frac{||\boldsymbol{\mathcal{H}}||_F^2}{||\boldsymbol{\mathcal{N}}||_F^2}.$$

The depicted MSE is calculated by averaging the results over 1000 independent Monte Carlo runs, truncated from 5% worst and 5% best MSEs to eliminate the influence of ill-convergence experiments and outliers. At each Monte Carlo run, the noise tensor $\boldsymbol{\mathcal{N}}$ changes. For the CPD computation, we consider the TensorLab toolbox functions [19]. In this work, we assume both, random and eigendecomposition-based initializations in the ALS. The proposed method is compared to three state-of-art algorithms, the ALS-based solution proposed in [2], called PARAFAC, which uses an ALS algorithm, with both random and eigendecomposition-based initializations, followed by closed-form solutions to estimate the parameters from the factors, the generalized eigenvalue decomposition (GEVD) [20] followed by a rectification step to retrieve the Vandermonde structure as in PARAFAC, and the so-called CP-VDM, for CPD with Vandermonde factor matrix, proposed in [21]. All algorithms are applied to the more general model in eq. (2). In Fig. 1, we fix $K = 3$, *i.e.*, the last factor has a full column rank. One may remark that JIRAFE and PARAFAC have same robustness to noise for a wide range of positive SNR. For negative SNRs, JIRAFE is the most robust estimator, for both initialization cases. This can be justified by the noise reduction property of the truncated SVD when the TT-SVD is applied. The same remark can be made for Fig. 2 where $K = 4$. In this figure, we can see that both, JIRAFE and PARAFAC, are competing for high

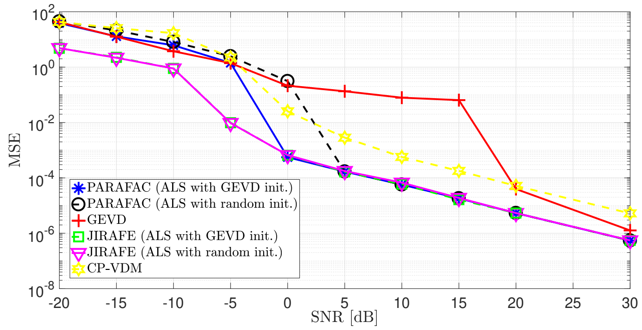

Fig. 2: MSE vs SNR in dB with Alg. 2 for $K = 4$.

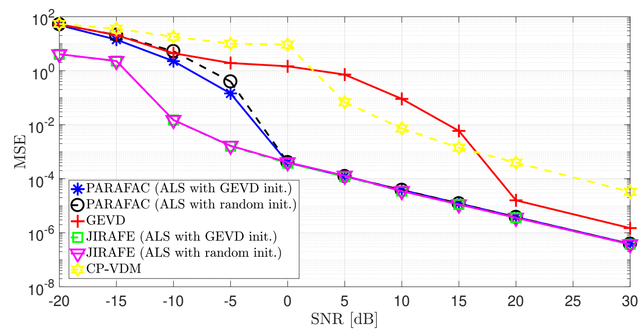

Fig. 3: MSE vs SNR in dB with Alg. 2 for $K = 5$.

SNRs. Meanwhile, JIRAFE is more robust than the other estimators for low SNRs. It is worth noting that for JIRAFE, both initializations give the same robustness to noise, which is not the case for PARAFAC where we can remark that the initialization has an influence on the robustness for low SNRs. Another important remark is that the optimization-based solutions, *i.e.*, ALS-based solutions in our case, namely PARAFAC and JIRAFE, have a better robustness compared to algebraic solutions such as GEVD or CP-VDM which seem not to be optimal but have a very low computational cost as we will see in the next experiments. In Fig. 3, the number of paths is fixed at $K = 5$, which means that the last factor is full row rank. We have a similar behavior for JIRAFE and PARAFAC as in the last experiment. On the other hand, CP-VDM and GEVD have difficulties when the rank increases. It is worth noting that breaking the dimensionality of tensor $\boldsymbol{\mathcal{H}}$ helps to improve the convergence of the ALS algorithm, since with JIRAFE, the ALS is applied to 3-order tensors instead of the original fifth-order tensor. In Fig. 4, we plot the mean number of iterations for JIRAFE and PARAFAC using a GEVD initialization, which was the most robust solution for PARAFAC. The bars represent the standard deviation for each SNR. We notice that 3-order ALS for JIRAFE needs only 1 iteration to

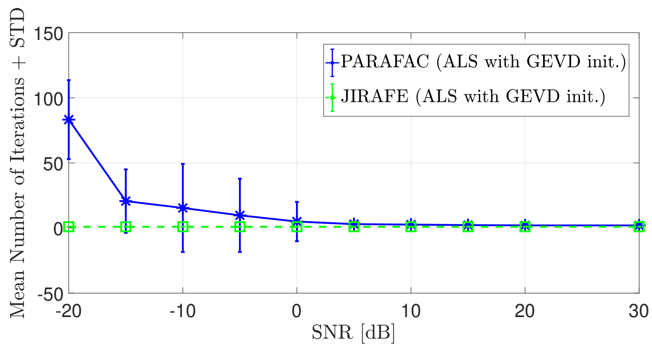

Fig. 4: Mean number of iteration for $K = 2$.

Table 1: Comparison of the computation time with SNR= 15dB.

Canonical rank K	2	3	4	5
PARAFAC (GEVD init.)	$0,09$ (s)	$0,11$ (s)	$0,13$ (s)	$0,16$ (s)
PARAFAC (random init.)	$0,13$ (s)	$0,14$ (s)	$0,16$ (s)	$0,20$ (s)
GEVD	$0,02$ (s)	$0,03$ (s)	$0,03$ (s)	$0,03$ (s)
JIRAFE (GEVD init.)	$0,04$ (s)	$0,05$ (s)	$0,06$ (s)	$0,07$ (s)
JIRAFE (random init.)	$0,04$ (s)	$0,06$ (s)	$0,07$ (s)	$0,09$ (s)
CP-VDM	$0,01$ (s)	$0,02$ (s)	$0,02$ (s)	$0,02$ (s)

converge after applying a GEVD, compared to PARAFAC which needs several iterations after the same GEVD initialization especially for low SNRs, while keeping in mind that the computational cost of a 3-order ALS iteration, that needs $O(3K^2M^2)$ flops, is very low compared to that of a 5-order ALS with $O(5K^2M^4)$ flops, where $M = \max(M_{T_x}, M_{T_y}, M_{R_x}, M_{R_y}) = 10$. This shows that breaking the dimensionality improves indeed the convergence of the ALS. Tab. 1 gives the average computation time for each method. The proposed JIRAFE with GEVD initialization provides the best tradeoff between noise robustness and computational complexity.

5. Conclusion

In this paper, an extension of a MIMO channel is considered using URAs both at the transmitter and the receiver, which leads to a fifth-order channel tensor. A TT-based representation has been derived for this tensor, highlighting the coupling between two adjacent core tensors via the latent matrices. For a multi-path scenario, a new JIRAFE-based method has been proposed for channel parameters estimation. This method allows to break the dimensionality of the original fifth-order CPD into a train of third-order tensors. Simulation results show the effectiveness of the proposed TT-based channel representation in terms of noise robustness and computation time for retrieving the physical channel parameters.

References

[1] M. Jo, D. Araujo, T. Maksymyuk, A. L. F. de Almeida, T. F. Maciel, and J. C. M. Mota, "Massive MIMO: Survey and future research topics," *IET Communications*, May 2016.

[2] C. Qian, X. Fu, N. D. Sidiropoulos, Y. Yang, "Tensor-based channel estimation for dual-polarized massive MIMO systems," *IEEE Trans. Signal Processing*, vol. 66, no. 24, pp. 6390 –6402, 2018.

[3] 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, pp. 1-84, 1970.

[4] V. Ollier, M.N. El Korso, R. Boyer, and P. Larzabal, *et al.*, "French SKA White Book-The French community towards the Square Kilometer Array," Published by the SKA-France, Chapter 4 "Technological developments", 2017.

[5] A. Cichocki, "Era of Big Data Processing: A New Approach via Tensor Networks and Tensor Decompositions," *CoRR*, 2010

[6] Y. Zniyed, R. Boyer, A. L. F. de Almeida and G. Favier, "High-Order CPD estimation with dimensionality reduction using a Tensor Train model," *Proc. EUSIPCO*, 2018.

[7] Y. Zniyed, R. Boyer, A. L. F. 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, March 2020, pp. 304-337, https://hal.archives-ouvertes.fr/hal-01815214.

[8] Y. Zniyed, S. Miron, R. Boyer and D. Brie, "Uniqueness of tensor train decomposition with linear dependencies," *Proc. IEEE CAMSAP*, 2019.

[9] A. Phan, A. Cichocki, I. Oseledets, G. G. Calvi, S. Ahmadi-Asl, and D. P. Mandic, "Tensor Networks for Latent Variable Analysis: Higher Order Canonical Polyadic Decomposition," *IEEE Trans. on Neural Networks and Learning Systems*, pp. 1-15, 2019.

[10] 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.

[11] Y. Zniyed, R. Boyer, A. L. F. de Almeida and G. Favier, "Multidimensional Harmonic Retrieval Based on Vandermonde Tensor Train," *Elsevier, Signal Processing*, vol. 163, pp. 75-86, 2019.

[12] I. Oseledets, "Tensor-Train decomposition," *SIAM Journal on Scientific Computing*, vol. 33, pp. 2295-2317, 2011.

[13] R. Boyer and P. Comon, "Rectified ALS Algorithm for Multidimensional Harmonic Retrieval," *Proc. IEEE SAM Workshop*, 2016.

[14] C. Qian , X. Fu, and N. D. Sidiropoulos, "Algebraic Channel Estimation Algorithms for FDD Massive MIMO Systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 5, pp. 961-973, 2019.

[15] D. C. Araujo, A. L. F. de Almeida, and J. C. M. Mota, "Compressive sensing based channel estimation for massive MIMO systems with planar arrays", *Proc. IEEE CAMSAP*, Cancun, Mexico, 2015.

[16] A. Alkhateeb, O. El Ayach, G. Leus, and R. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, Oct 2014.

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

[18] A. Y. Kibangou, G. Favier, "Non-iterative solution for PARAFAC with a Toeplitz matrix factor", *Proc. EUSIPCO*, 2009.

[19] N. Vervliet, O. Debals, L. Sorber, M. Van Barel and L. De Lathauwer, "Tensorlab 3.0", Mar. 2016, Available online, URL: http://www.tensorlab.net.

[20] E. Sanchez and B. R. Kowalski, "Tensorial resolution: a direct trilinear decomposition", *Journal of Chemometrics*, vol. 4, pp. 29-45, 1990.

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