



HAL
open science

HIGH-DIMENSIONAL, LOW-RANK TENSOR APPROXIMATION: CRAMÉR-RAO LOWER BOUNDS AND APPLICATION TO MIMO CHANNELS

Clémence Prévost, Pierre Chainais

► **To cite this version:**

Clémence Prévost, Pierre Chainais. HIGH-DIMENSIONAL, LOW-RANK TENSOR APPROXIMATION: CRAMÉR-RAO LOWER BOUNDS AND APPLICATION TO MIMO CHANNELS. 2023. hal-04302405

HAL Id: hal-04302405

<https://hal.science/hal-04302405v1>

Preprint submitted on 23 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **HIGH-DIMENSIONAL, LOW-RANK TENSOR APPROXIMATION:**
2 **CRAMÉR-RAO LOWER BOUNDS AND APPLICATION TO**
3 **MIMO CHANNELS***

4 C. PRÉVOST, P. CHAINAIS[†]

5 **Abstract.** Tensor factorization has been steadily used in the past decade to represent high-
6 dimensional data. In particular, the canonical polyadic (CP) decomposition (CPD) is very appreci-
7 ated for its modeling power and remarkable uniqueness properties. However, computing the CPD is
8 challenging when the order of the tensor becomes high: numerical issues and high needs for storage
9 and processing can lead the standard algorithms to diverge. To circumvent this limitation, the equiv-
10 alence between the CPD and the Tensor Train Decomposition (TTD) is exploited. This approach is
11 implemented in a new algorithm called Dimensionality Reduction, joint Estimation of the Ambiguity
12 Matrices and the CP FACTors (DREAMFAC). A global coupled optimization scheme is proposed
13 to break the curse of dimensionality and estimate the CP factors. DREAMFAC performs better
14 than state-of-the-art methods. It avoids the usual propagation of the estimation error in the factors
15 of the TTD. In particular, DREAMFAC reaches the Cramér-Rao lower bounds associated with the
16 considered coupled CP-TT model, which is not the case for the state-of-the-art sequential proce-
17 dure. Performances are illustrated on the problem of estimating the channels in a dual-polarized
18 MIMO system. Numerical experiments show the competitive performance of the proposed method
19 for recovery of the CP factors and estimation of the channel parameters, even with very low SNR.

20 **Key words.** Canonical polyadic decomposition, tensor train decomposition, coupled optimiza-
21 tion, factor retrieval, MIMO systems

22 **MSC codes.** 15A23, 15A69

23 **1. Introduction.** Tensors provide a faithful representation of higher-order ob-
24 servations by preserving their multidimensional structure [7, 18]. For instance, color
25 images can be seen as data cubes with two spatial dimensions (the pixels) and a
26 spectral dimension (the colours) encoding the red, green and blue channels. Tensor
27 factorization has been steadily used in the past decade to model such data, due to
28 its capabilities to capture the interactions between a set of latent factors. It has been
29 successfully applied to various problems in, e.g., signal processing and machine learn-
30 ing [34], brain signal processing [1], video completion [41] or telecommunications [29].
31 The notion of low rank in tensors is not unified: various tensor decompositions are
32 available, all carrying different properties and rank definitions. Perhaps the most nat-
33 ural generalization of the concept of matrix rank to tensors is the canonical polyadic
34 (CP) decomposition (CPD) [10]. The CP rank of a tensor is defined as the minimal
35 number K of rank-one tensors that, when linearly combined, lead to a perfect recov-
36 ery of that tensor. A N -th order rank-one tensor is given by the outer product of
37 N vectors. The CP decomposition has gained a lot of interest for data processing
38 and analysis, due to its remarkable uniqueness properties under mild conditions [38].
39 However, the set of low-rank tensors is not closed, therefore computing the CPD is
40 an ill-posed problem [20]. In practice, this task is usually carried out by suboptimal
41 iterative algorithms. Furthermore, the storage and processing cost of tensors increase
42 exponentially with their order N . This limitation is known as the curse of the di-
43 mensionality [28]. The Tensor Train (TT) decomposition (TTD) has been recently
44 introduced [27]. This decomposition has two advantages. First, it exploits a stable
45 numerical estimation procedure which avoids the iterative algorithms used to com-

*Submitted to the editors November 22, 2023.

Funding: This work was partly supported by the ANR project “Chaire IA Sherlock” ANR-20-CHIA-0031-01 hold by P. Chainais, as well as by the national support within the *programme d’investissements d’avenir* ANR-16-IDEX-0004 ULNE and Région HDF.

[†]Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRISAL, F-59000 Lille, France (clemence.prevast[at]univ-lille.fr, pierre.chainais[at]centralelille.fr).

46 pute the CPD. Second, it breaks the curse of dimensionality by operating on a set
 47 of matrices and third-order tensors, called TT-cores. In particular, its storage cost is
 48 linear in N . Numerous works were devoted to the task of dimensionality reduction
 49 in the CP model. In [43], the authors provided an equivalence between the CPD
 50 and the TTD through a set of non-singular change-of-basis matrices shared among
 51 the TT-cores. An algorithm, called Joint dImensionality Reduction And Factor Es-
 52 timation (JIRAFE), was proposed. It recovered the factors of the CPD underlying
 53 a high-order tensor from its estimated TT-cores. A JIRAFE-like procedure was also
 54 recently extended to constrained CPD in [13]. These algorithms were based on a
 55 series of local and sequential optimization problems to obtain the CP factors at a
 56 low computational cost. The change-of-basis matrices appearing in the TTD were
 57 estimated once and propagated through the other optimization problems, without
 58 enforcing coherence between the TT-cores. For this reason, JIRAFE may lead to a
 59 sub-optimal estimation of the CP factors estimated in last position of the tensor train.

60 This work introduces a new algorithm for joint dimensionality reduction and esti-
 61 mation of the CP factors. A global optimization strategy is proposed. In contrast with
 62 JIRAFE, it ensures the coherence between TT-cores by considering coupled updates
 63 for the change-of-basis matrices. This approach is implemented in a new algorithm
 64 called Dimensionality Reduction, joint Estimation of the Ambiguity Matrices and the
 65 CP FACTors (DREAMFAC). Cramér-Rao Bound (CRB) [8, 12, 30] are used to evalu-
 66 ate and compare the relative performance of JIRAFE and DREAMFAC. Cramér-Rao
 67 bounds for tensor CP models have been extensively studied in the literature, includ-
 68 ing performance bounds for uncoupled CP models [4, 21, 32], Bayesian frameworks [6],
 69 and constrained models [31]. To take into account the coupling between the TT-
 70 cores, a coupled model that enforces constraints on the change-of-basis matrices is
 71 introduced. For such models the constrained Cramér-Rao bound (CCRB) can be
 72 used, whose versatility was shown by numerous works [22–24, 39, 40]. An application
 73 to the problem of harmonic retrieval in MIMO systems [2] permits to illustrate the
 74 behaviour of DREAMFAC on a realistic low-rank decomposition problem. Indeed,
 75 for MIMO channels modeling and estimation, it is important to accurately estimate
 76 channels parameters at the base station (angles of arrival, angles of departure, path
 77 gains, and polarization parameters) to perform beamforming and deal with multi-
 78 user interferences. In [29], a tensor-based approach for dual-polarized MIMO channel
 79 estimation was proposed. The MIMO channel was recast as a fourth-order tensor
 80 admitting a CPD. The authors proposed an Alternating Least Squares (ALS) to es-
 81 timate CP factors. In [42], the authors adapted the JIRAFE procedure to treat the
 82 MIMO channel estimation problem. Therefore this problem can be used as a relevant
 83 benchmark.

84 The paper is organized as follows. Section 2 introduces important background
 85 on tensors and their low-rank decompositions. Section 3 describes the problem at
 86 hand, the JIRAFE procedure and its limitations. The proposed approach DREAM-
 87 FAC is also detailed. Section 4 contains the derivation of the Cramér-Rao bounds
 88 for the proposed coupled model. It includes the detailed probabilistic framework,
 89 and closed-form expressions for the matrices to invert. Section 5 gathers an exten-
 90 sive set of numerical experiments to illustrate the performance of DREAMFAC with
 91 respect to the state of the art. They include simulations highlighting the robustness
 92 of DREAMFAC compared to JIRAFE, comparison of the algorithms' performance to
 93 the CRB, and results on the selected application to the estimation of MIMO channels.

94
 95 *Notation.* The following notations [7, 18] are used: lower (a) or uppercase (A)
 96 plain font for scalars, boldface lowercase (\mathbf{a}) for vectors, boldface uppercase (\mathbf{A}) for
 97 matrices and calligraphic (\mathcal{A}) for tensors. The elements of vectors, matrices, and
 98 tensors are denoted by a_d , A_{d_1, d_2} and $\mathcal{A}_{d_1, \dots, d_N}$, respectively. The transpose of a

99 matrix \mathbf{A} is denoted by \mathbf{A}^\top . The matrix \mathbf{I}_N is the $N \times N$ identity matrix and $\mathbf{0}_{L \times K}$
 100 is the $L \times K$ matrix of zeros. The symbols \boxtimes , \odot and \otimes denote the Kronecker, Khatri-
 101 Rao and outer products. The operator $\text{vec}\{\cdot\}$ stands for the standard column-major
 102 vectorization of a matrix or a tensor. The operation $\text{Diag}\{\mathbf{A}, \mathbf{B}\}$ produces a block-
 103 diagonal matrix whose blocks are \mathbf{A} and \mathbf{B} .

104 2. Background on low-rank tensor models.

105 **2.1. Preliminaries on tensors.** A tensor $\mathcal{X} \in \mathbb{R}^{D_1 \times \dots \times D_N}$ is an N -dimensional
 106 array indexed by the elements $\mathcal{X}_{d_1, \dots, d_N}$, for $d_n \in \{1, \dots, D_n\}$ ($n \in \{1, \dots, N\}$). Each
 107 dimension of a tensor is called a mode. A mode- p fiber of \mathcal{X} is a vector obtained by
 108 fixing all but the p -th dimension.

109 **DEFINITION 2.1** (Tensor unfoldings). *The mode- p unfolding of a tensor \mathcal{X} , de-*
 110 *noted by $\mathbf{X}^{(p)}$, is the matrix whose rows are the p -mode fibers of \mathcal{X} , ordered according*
 111 *to the vectorization order. For a tensor $\mathcal{X} \in \mathbb{R}^{D_1 \times \dots \times D_N}$, e.g., $\mathbf{X}^{(1)} \in \mathbb{R}^{D_N \dots D_2 \times D_1}$.*

112 **DEFINITION 2.2** (Matrix mode product). *The matrix p -mode product between a*
 113 *tensor \mathcal{X} and a matrix \mathbf{M} is denoted by $\mathcal{X} \bullet_p \mathbf{M}$ and is constructed such that each*
 114 *mode- p fiber of \mathcal{X} is multiplied by \mathbf{M} , e.g., the elements of the mode-1 product between*
 115 *$\mathcal{X} \in \mathbb{R}^{D_1 \times \dots \times D_N}$ and $\mathbf{M} \in \mathbb{R}^{L \times D_1}$ are accessed as*

$$116 \quad (2.1) \quad (\mathcal{X} \bullet_1 \mathbf{M})_{\ell, d_2, \dots, d_N} = \sum_{i=1}^{D_1} \mathcal{X}_{i, d_2, \dots, d_N} \mathbf{M}_{\ell, i}, \quad \ell \in \{1, \dots, L\}.$$

117 *Moreover, it holds that $\mathcal{Y} = \mathcal{X} \bullet_p \mathbf{M} \Leftrightarrow \mathbf{Y}^{(p)} = \mathbf{X}^{(p)} \mathbf{M}^\top$.*

118 **DEFINITION 2.3** (Tensor contraction product). *The contraction product on modes*
 119 *p, q between two tensors $\mathcal{X} \in \mathbb{R}^{D_1 \times \dots \times D_N}$ and $\mathcal{Y} \in \mathbb{R}^{J_1 \times \dots \times J_M}$ with $D_p = J_q$ is denoted*
 120 *by $\mathcal{X} \bullet_p^q \mathcal{Y}$. It produces a tensor of order $N + M - 2$ such that*

$$121 \quad (2.2) \quad (\mathcal{X} \bullet_p^q \mathcal{Y})_{d_1, \dots, d_{p-1}, d_{p+1}, \dots, d_N, j_1, \dots, j_{q-1}, j_{q+1}, \dots, j_M} =$$

$$122 \quad (2.3) \quad \sum_{\ell=1}^{D_p=J_q} (\mathcal{X})_{d_1, \dots, d_{p-1}, \ell, d_{p+1}, \dots, d_N} (\mathcal{Y})_{j_1, \dots, j_{q-1}, \ell, j_{q+1}, \dots, j_M}.$$

123

124 **DEFINITION 2.4** (Outer product). *The outer product between N vectors $\mathbf{a}_n \in \mathbb{R}^{D_n}$*
 125 *($n \in \{1, \dots, N\}$) is a rank-one tensor $\mathcal{X} = \mathbf{a}_1 \otimes \dots \otimes \mathbf{a}_N \in \mathbb{R}^{D_1 \times \dots \times D_N}$ whose elements*
 126 *are accessed as $\mathcal{X}_{d_1, \dots, d_N} = (\mathbf{a}_1)_{d_1} \dots (\mathbf{a}_N)_{d_N}$.*

127 **2.2. The canonical polyadic decomposition.** For all N -dimensional tensor
 128 $\mathcal{X} \in \mathbb{R}^{D_1 \times \dots \times D_N}$, there exists an integer K such that it admits a canonical polyadic
 129 (CP) decomposition (CPD) as

$$130 \quad (2.4) \quad \mathcal{X} = \llbracket \mathbf{A}_1, \dots, \mathbf{A}_N \rrbracket,$$

131 where $\mathbf{A}_n \in \mathbb{R}^{D_n \times K}$ ($n \in \{1, \dots, N\}$) are called the CP factors. When minimal, the
 132 integer K is the rank of the tensor \mathcal{X} . Then each entry of \mathcal{X} can be expressed as

$$133 \quad (2.5) \quad \mathcal{X}_{d_1, \dots, d_N} = \sum_{k=1}^K (A_1)_{d_1, k} \dots (A_N)_{d_N, k}.$$

134

135 The CP factors are essentially unique up to scaling and permutation ambiguities,
 136 if the rank K is not too large [7, 18]. The permutation ambiguity means that the

137 columns of the latent CP factors can be reordered arbitrarily by any permutation
138 matrix $\mathbf{\Pi} \in \mathbb{R}^{K \times K}$ as

$$139 \quad (2.6) \quad \mathcal{X} = \llbracket \mathbf{A}_1, \dots, \mathbf{A}_N \rrbracket = \llbracket \mathbf{A}_1 \mathbf{\Pi}, \dots, \mathbf{A}_N \mathbf{\Pi} \rrbracket.$$

140 The scaling ambiguity means that the individual factors can be scaled as

$$141 \quad (2.7) \quad \mathcal{X}_{d_1, \dots, d_N} = \sum_{k=1}^K \sum_{r=1}^K (\lambda_{1,k}(A_1)_{d_1,k}) \dots (\lambda_{N,k}(A_N)_{d_N,k}).$$

142 where $\lambda_{1,k} \dots \lambda_{N,k} = 1$, for all $k \in \{1, \dots, K\}$.

144 **2.3. The tensor train decomposition (TTD).** The TTD [27] factorizes a
145 tensor $\mathcal{X} \in \mathbb{R}^{D_1 \times \dots \times D_N}$ as a series of matrix-tensor products and contractions between
146 third-order tensors as

$$147 \quad (2.8) \quad \mathcal{X} = \mathbf{G}_1 \overset{1}{\underset{2}{\bullet}} \mathcal{G}_2 \dots \overset{1}{\underset{3}{\bullet}} \mathcal{G}_{N-1} \overset{1}{\underset{3}{\bullet}} \mathbf{G}_N,$$

149 where $\mathbf{G}_1 \in \mathbb{R}^{D_1 \times K_1}$, $\mathbf{G}_N \in \mathbb{R}^{K_{N-1} \times D_N}$, and $\mathcal{G}_n \in \mathbb{R}^{K_{n-1} \times D_n \times K_n}$, for $n \in \{2, \dots, N-1\}$,
150 are referred to as TT-cores. The integers K_1, \dots, K_{N-1} are called the TT-ranks. A fast
151 and efficient way to estimate the TT-cores is to resort to the TT-SVD algorithm [27], a
152 procedure that sequentially extracts dominant singular vectors from tensor unfoldings.
153 As a result, the TTD in (2.8) is not unique. In fact, due to the use of the SVD, we
154 can replace two successive TT-cores \mathcal{G}_n and \mathcal{G}_{n+1} by \mathcal{G}'_n and \mathcal{G}'_{n+1} such that

$$155 \quad (2.9) \quad \mathcal{G}'_n = \mathcal{G}_n \overset{1}{\underset{3}{\bullet}} \mathbf{M}_n^{-1}, \quad \mathcal{G}'_{n+1} = \mathbf{M}_n \overset{1}{\underset{2}{\bullet}} \mathcal{G}_{n+1},$$

157 where $\mathbf{M}_n \in \mathbb{R}^{K_n \times K_n}$ is a non-singular change-of-basis matrix. This means that the
158 multiplicative ambiguities in the TTD correspond to post- and pre-multiplications by
159 nonsingular matrices.

160 **2.4. Equivalence between the CPD and the TTD.** There exists an equiv-
161 alence between the CP and TT decompositions. If the TT-ranks are such that
162 $K_1 = \dots = K_{N-1} = K \leq \min(D_1, \dots, D_N)$, the TTD can be used to efficiently
163 estimate the rank- K CP factors of a higher-order tensor [43]. In particular, using the
164 TT-SVD, one can obtain TT factors such that:

$$165 \quad (2.10) \quad \mathbf{G}_1 = \mathbf{A}_1 \mathbf{M}_1^{-1}, \quad \mathcal{G}_n = \llbracket \mathbf{M}_{n-1}, \mathbf{A}_n, \mathbf{M}_n^{-\top} \rrbracket, \quad \mathbf{G}_N = \mathbf{M}_{N-1} \mathbf{A}_N^{\top},$$

167 where $\mathbf{M}_n \in \mathbb{R}^{K_n \times K_n}$ are non-singular change-of-basis matrices. These matrices can
168 be related to the permutation and scaling ambiguity matrices of the CPD, see the
169 proof of [43, Theorem 6] for more details.

170 3. Joint estimation and dimensionality reduction.

171 **3.1. The model and the optimization problem.** Let \mathcal{X} a tensor of rank K
172 that admits a CPD of the form

$$173 \quad (3.1) \quad \mathcal{X} = \llbracket \mathbf{A}_1, \dots, \mathbf{A}_N \rrbracket, \quad \text{where } \mathbf{A}_n \in \mathbb{R}^{D_n \times K} \quad \forall n \in \{1, \dots, N\}.$$

175 Let $\mathcal{Y} \in \mathbb{R}^{D_1 \times \dots \times D_N}$ the noisy observation according to the following model

$$176 \quad (3.2) \quad \mathcal{Y} = \mathcal{X} + \mathcal{E},$$

178 where the tensor \mathcal{E} represents isotropic white Gaussian noise. The model (3.2) can
 179 be rewritten

180 (3.3)
$$\mathcal{Y} = \llbracket \mathbf{A}_1, \dots, \mathbf{A}_N \rrbracket + \mathcal{E}.$$

182 The problem of estimating the CP factors \mathbf{A}_n based on model (3.3) has found numer-
 183 ous interests over the past decade. It has been applied, among others, to problems
 184 in signal processing and machine learning [34], brain signal processing [1], video com-
 185 pletion [41] or telecommunications [29]. The popularity of the CPD lies in its high
 186 regularization power, mild uniqueness properties and versatility to model a variety of
 187 problems, from denoising to component analysis.

188 The most popular way to estimate the CP factors is an iterative Alternating
 189 Least-Squared (ALS) algorithm [9]. However, this procedure becomes exponentially
 190 expensive as the order N increases. One iteration of CP-ALS requires $\mathcal{O}(K^2 \prod_n D_N)$
 191 flops. Furthermore, CP-ALS becomes less robust and is prone to convergence issues for
 192 high-order tensors. These limitations fall under the so-called curse of dimensionality
 193 [28]. Consequently, one may need to perform tensor dimensionality reduction to
 194 exploit the benefits of model (3.3).

195 Directly solving (3.3) using high-dimensional CP-ALS is costly. To circumvent
 196 this limitation, it is possible to perform low-rank denoising by considering the follow-
 197 ing two optimization problems. To ease the notation, let us denote $\mathbb{A} = \{\mathbf{A}_1, \dots, \mathbf{A}_N\}$,
 198 $\mathbb{M} = \{\mathbf{M}_1, \dots, \mathbf{M}_{N-1}\}$ and $\mathbb{G} = \{\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_{N-1}, \mathbf{G}_N\}$.

199 First, a low rank- K factorized approximation is obtained by solving

200 (3.4)
$$\min_{\mathbb{G}} \|\mathcal{Y} - \mathbf{G}_1 \underset{2}{\bullet} \mathbf{G}_2 \dots \underset{N-1}{\bullet} \mathbf{G}_{N-1} \underset{N}{\bullet} \mathbf{G}_N\|_F^2.$$

201 This step can be interpreted as a denoising of \mathcal{Y} using the TT-SVD with rank K . This
 202 operation provides the best rank- K approximation of \mathcal{Y} in the least-squared sense,
 203 while being less costly than high-dimensional CP-ALS.

204 Then, the links between the TT-cores \mathbb{G} and the CP factors \mathbb{A} (2.10) are formulated
 205 through the optimization problem

206 (3.5)
$$\min_{\mathbb{A}, \mathbb{M}} \sum_{n=2}^{N-1} (\|\mathbf{G}_n - \llbracket \mathbf{M}_{n-1}, \mathbf{A}_n, \mathbf{M}_n^{-\top} \rrbracket\|_F^2)$$

 207
$$+ \|\mathbf{G}_1 - \mathbf{A}_1 \mathbf{M}_1^{-1}\|_F^2 + \|\mathbf{G}_N - \mathbf{M}_{N-1} \mathbf{A}_N^{\top}\|_F^2,$$

209 that accounts for the ambiguities of the TT-SVD through the \mathbb{M} matrices. The n -th
 210 ambiguity matrix \mathbf{M}_n appears twice in (3.5), once in the term related to \mathbf{G}_n and once
 211 in that related to \mathbf{G}_{n+1} . As a result, (3.5) is a coupled optimization problem.

212 **3.2. State-of-the-art and its limitations.** In [43], the authors proposed to
 213 promote a fast local/sequential optimization method as a sub-optimal solution instead
 214 of minimizing (3.4)–(3.5). This procedure was called Joint dDimensionality Reduction
 215 And Factor Estimation (JIRAFE) [43]. The first step of JIRAFE estimates the TT-
 216 cores using the TT-SVD algorithm with ranks $K_1 = \dots = K_{N-1} = K$ by solving
 217 (3.4). Then, CP-ALS is performed on a single \mathbf{G}_n ($n \in \{2, \dots, N-1\}$), e.g. on \mathbf{G}_2 ,
 218 by minimizing the criterion

219 (3.6)
$$\min_{\mathbf{M}_1, \mathbf{A}_2, \mathbf{M}_2} \|\mathbf{G}_2 - \llbracket \mathbf{M}_1, \mathbf{A}_2, \mathbf{M}_2^{-\top} \rrbracket\|_F^2,$$

221 Since the three CP factors \mathbf{M}_1 , \mathbf{A}_2 , \mathbf{M}_2 underlying \mathbf{G}_2 are estimated thanks to an
 222 alternate least-squares algorithm, this step was referred to as Tri-ALS. In JIRAFE, the

223 resulting estimate of \mathbf{M}_2 is then fixed and propagated to other \mathcal{G}_n , $n \in \{3, \dots, N-1\}$.
 224 Then, for each \mathcal{G}_n , the cost function to minimize is

$$225 \quad (3.7) \quad \min_{\mathbf{A}_n, \mathbf{M}_n} \|\mathcal{G}_n - \llbracket \mathbf{M}_{n-1}, \mathbf{A}_n, \mathbf{M}_n^{-\top} \rrbracket\|_F^2,$$

227 where \mathbf{M}_{n-1} is supposed to be known. This operation was termed Bi-ALS. Finally,
 228 \mathbf{A}_1 and \mathbf{A}_N were obtained using the links provided in (2.10). Algorithm 3.1 provides a
 229 complexity gain of approximately N_{iter} flops with respect to CP-ALS. The JIRAFE
 230 procedure is summarized in Algorithm 3.1 below.

Algorithm 3.1 JIRAFE

input: Observation tensor \mathcal{Y} , CP-rank K
output: CP factors $\mathbf{A}_1, \dots, \mathbf{A}_N$
 Estimate \mathbf{G}_1 , \mathbf{G}_N and $\mathcal{G}_n \forall n \in \{2, \dots, N-1\}$ using TT-SVD on \mathcal{Y}
repeat
 Estimate \mathbf{M}_1 , \mathbf{A}_2 , \mathbf{M}_2^{-1} using (3.6) (Tri-ALS)
 for $n = 3, \dots, N-1$ **do**
 Estimate \mathbf{A}_n , \mathbf{M}_n^{-1} using (3.7) with \mathbf{M}_{n-1} known (Bi-ALS)
 end for
until convergence
 Estimate \mathbf{A}_1 and \mathbf{A}_N using (2.10)

231 However, JIRAFE suffers from a major limitation due to the sub-optimal esti-
 232 mation of the ambiguity matrices \mathbf{M}_n . For instance, \mathbf{M}_2 was estimated only from
 233 \mathcal{G}_2 even though it was shared between the CPD of \mathcal{G}_2 and \mathcal{G}_3 . More generally, each
 234 matrix \mathbf{M}_n is shared between the CPD of \mathcal{G}_n and \mathcal{G}_{n+1} . As such, some important in-
 235 formation contained in (3.5) is ignored by JIRAFE so that the coherence between the
 236 TT-cores is lost. As a result, the performance of JIRAFE is highly dependent on the
 237 choice of the TT-core considered for the initial Tri-ALS step. By first estimating the
 238 CP factors underlying \mathcal{G}_2 , the estimation error on the $\mathbf{M}_n, \mathbf{A}_n$ for $n > 2$ is expected
 239 to increase with n . As a result, \mathbf{M}_{N-1} and \mathbf{A}_N shall inherit from the estimation
 240 errors of previous CP factors. Conversely, if Tri-ALS was performed first on \mathcal{G}_{N-1}
 241 before propagating estimations downto $n = 1$, \mathbf{A}_1 and \mathbf{M}_1 would not be estimated
 242 accurately. Therefore, the following subsection introduces a new estimation method
 243 to circumvent these limitations.

244 **3.3. Proposed approach.** This section proposes a new approach to solve prob-
 245 lem (3.5). It fully takes into account the coupling induced by the \mathbf{M}_n by carefully
 246 considering the global minimization of the cost function of (3.5). In place of a se-
 247 quential estimation of factors that solves local optimization sub-problems, the global
 248 optimization is considered. Each iteration of the algorithm will update every factor,
 249 thus ensuring to reach the global minimum of (3.5). This is made possible thanks
 250 to a block-coordinate descent that alternates between all the parameters of interest.
 251 The resulting algorithm is called Dimensionality Reduction, joint Estimation of the
 252 Ambiguity Matrices and the CP Factors (DREAMFAC).

253 First, DREAMFAC estimates the \mathbf{G}_1 , \mathbf{G}_N and $\mathcal{G}_n \forall n \in \{2, \dots, N-1\}$ thanks
 254 to a TT-SVD applied to \mathcal{Y} . This step consists in solving for (3.4), hence this step is
 255 similar in JIRAFE and DREAMFAC. Then, the coupling constraints on the ambiguity

256 matrices lead to the following coupled model:

$$\begin{aligned}
 & \left\{ \begin{array}{l} \mathbf{G}_1 \mathbf{M}_1 = \mathbf{A}_1 \\ \mathbf{G}_2^{(1)\top} = \mathbf{M}_1 (\mathbf{M}_2^{-\top} \odot \mathbf{A}_2)^\top \end{array} \right. \\
 & \left\{ \begin{array}{l} \mathbf{G}_n^{(3)} \mathbf{M}_n = (\mathbf{A}_n \odot \mathbf{M}_{n-1}), \quad \forall 2 \leq n \leq N-2, \\ \mathbf{G}_{n+1}^{(1)\top} = \mathbf{M}_n (\mathbf{M}_{n+1}^{-\top} \odot \mathbf{A}_{n+1})^\top, \quad \forall 2 \leq n \leq N-2, \end{array} \right. \\
 & \left\{ \begin{array}{l} \mathbf{G}_{N-1}^{(3)} \mathbf{M}_{N-1} = (\mathbf{A}_{N-1} \odot \mathbf{M}_{N-2}) \\ \mathbf{M}_{N-1} \mathbf{A}_N^\top = \mathbf{G}_N \end{array} \right.
 \end{aligned}
 \tag{3.8}$$

261 Considering that the ambiguity matrices \mathbf{M}_n are fixed, the \mathbf{A}_n can be updated from
 262 (3.8) by solving a simple least-squares problem, as in JIRAFE. Considering that the
 263 CP factors \mathbf{A}_n are fixed, the coherence between TT-cores leads to new coupled updates
 264 for the \mathbf{M}_n matrices:

$$\begin{aligned}
 & \mathbf{M}_1 = \arg \min_{\mathbf{M}_1} \|\mathbf{G}_1 \mathbf{M}_1 - \mathbf{A}_1\|_F^2 + \|\mathbf{G}_2^{(1)\top} - \mathbf{M}_1 (\mathbf{M}_2^{-\top} \odot \mathbf{A}_2)^\top\|_F^2, \\
 & + \|\mathbf{G}_n^{(1)\top} - \mathbf{M}_{n-1} (\mathbf{M}_n^{-\top} \odot \mathbf{A}_n)^\top\|_F^2, \\
 & \tag{3.10}
 \end{aligned}$$

$$\begin{aligned}
 & \mathbf{M}_n = \arg \min_{\mathbf{M}_n} \|\mathbf{G}_n^{(3)} \mathbf{M}_n - (\mathbf{A}_n \odot \mathbf{M}_{n-1})\|_F^2 \\
 & + \|\mathbf{G}_{n+1}^{(1)\top} - \mathbf{M}_n (\mathbf{M}_{n+1}^{-\top} \odot \mathbf{A}_{n+1})^\top\|_F^2, \quad \forall 2 \leq n \leq N-2, \\
 & \tag{3.11}
 \end{aligned}$$

$$\mathbf{M}_{N-1} = \arg \min_{\mathbf{M}_{N-1}} \|\mathbf{G}_{N-1}^{(3)} \mathbf{M}_{N-1} - (\mathbf{A}_{N-1} \odot \mathbf{M}_{N-2})\|_F^2 + \|\mathbf{G}_N - \mathbf{M}_{N-1} \mathbf{A}_N^\top\|_F^2.$$

271 The solution to this mixture of least-squares problem is explicit. Appendix A gathers
 272 the resulting closed-form expressions for the ambiguity matrices. In the proposed
 273 approach, the $(\mathbf{M}_n)_{1 \leq n \leq N-1}$ and $(\mathbf{A}_n)_{2 \leq n \leq N-1}$ matrices are all updated in the same
 274 loop, termed Multi-ALS. This new procedure, called Dimensionality Reduction, joint
 275 Estimation of the Ambiguity Matrices and the CP FACTors (DREAMFAC), is sum-
 276 marized in Algorithm 3.2 below.

Algorithm 3.2 DREAMFAC

input: Observation tensor \mathcal{Y} , CP-rank K

output: CP factors $\mathbf{A}_1, \dots, \mathbf{A}_N$

Estimate \mathbf{G}_1 , \mathbf{G}_N and $\mathbf{G}_n \forall n \in \{2, \dots, N-1\}$ using TT-SVD on \mathcal{Y}

repeat

 Estimate the \mathbf{M}_n using (3.9)–(3.11), and the \mathbf{A}_n using (3.7) (Multi-ALS)

 Estimate \mathbf{A}_1 and \mathbf{A}_N using (2.10)

until convergence

277 Note that the coherence between the TT-cores is ensured thanks to the cou-
 278 pled estimation of the ambiguity matrices \mathbf{M}_n . Hence the information comprised in
 279 model (3.8) is fully exploited by this new approach. As a by-product, in contrast
 280 with JIRAFE, the performance of DREAMFAC is independent on the starting TT-
 281 core. Figure 1 shows a graphical representation of model (3.8) and summarizes the
 282 differences between JIRAFE and DREAMFAC.

283 Finally, turning to the algorithmic complexity, Algorithm 3.2 requires at most the
 284 same number of updates as JIRAFE. While updating the \mathbf{A}_n matrices still requires

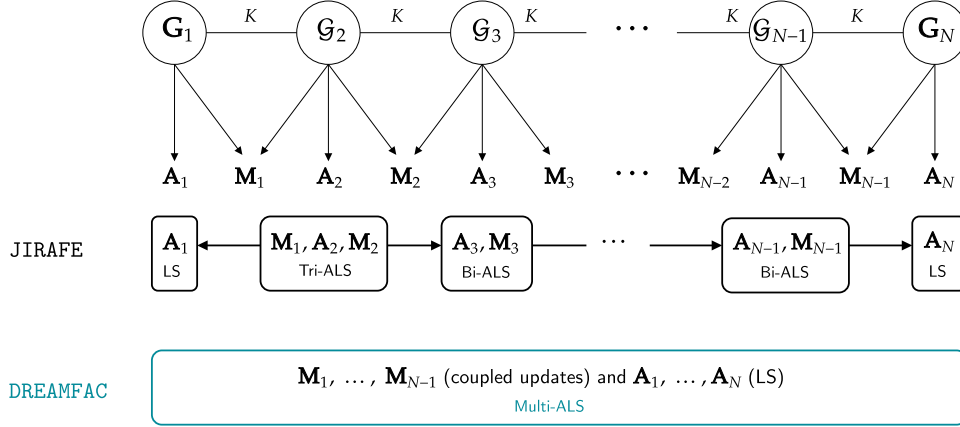


FIG. 1. Graphical representation of JIRAFE (black) and DREAMFAC (blue).

285 $\mathcal{O}(K^2 D_n)$ flops as in JIRAFE, the coupled updates of the \mathbf{M}_n 's can be made faster
 286 by using a fast solver for the Sylvester equations [3, 14, 35]. Such solvers require only
 287 $\mathcal{O}(2K^3)$ flops, which is smaller¹ than $\mathcal{O}(K^2 D_n)$ as soon as $K \ll D_n$, which is most
 288 likely.

289 4. Cramér-Rao bounds for the proposed approach.

290 **4.1. General probabilistic framework.** To derive appropriate performance
 291 bounds, it is necessary to embed the problem in an appropriate probabilistic frame-
 292 work. This requires to properly define the probabilistic model and the parameters of
 293 interest. A general coupled model has the form

$$294 \quad (4.1) \quad \begin{cases} \mathbf{g}_n \sim \mathbf{f}_\omega(\mathbf{g}_n) & \forall n \in \{1, \dots, N\}, \\ \mathbf{h}_n(\omega) = 0, \end{cases}$$

295

296 where \mathbf{g}_n denote the observations. The functions $\mathbf{f}_\omega(\mathbf{g}_n)$ are the probability density
 297 functions (PDF) of the random real datasets \mathbf{g}_n , parameterized by the unknown
 298 deterministic real parameter vector ω . Several assumptions will be necessary: i) the
 299 PDFs $\mathbf{f}_\omega(\mathbf{g}_n)$ are non-redundant functions differentiable w.r.t. ω , and their support
 300 as functions of \mathbf{g}_n do not depend on ω ; ii) the \mathbf{g}_n are statistically independent. The
 301 constraints in the model are described by the functions \mathbf{h}_n that are non-redundant
 302 deterministic vector functions, everywhere differentiable with respect to ω .

303 The model (3.8) can be rewritten equivalently under the form (4.1). First, the
 304 link between the TT-cores and the $\mathbf{A}_n, \mathbf{M}_n$ matrices is such that

$$305 \quad (4.2) \quad \mathbf{G}_1 = \mathbf{A}_1 \mathbf{M}_{1,1}^{-1}, \quad \mathcal{G}_n = \llbracket \mathbf{M}_{n-1,n}, \mathbf{A}_n, \mathbf{M}_{n,n}^{-\top} \rrbracket, \quad \mathbf{G}_N = \mathbf{M}_{N-1,N} \mathbf{A}_N^\top.$$

307 This paper proposes an approach to solve the approximation problem (??) in the
 308 least-squares sense based on model (3.3). Therefore, it is assumed that the obser-
 309 vations are Gaussian. To be more precise, let us note $\mathbf{g}_n = \text{vec}\{\mathcal{G}_n\}$ for $n \in \{1, N\}$
 310 and $\mathbf{g}_n = \text{vec}\{\mathcal{G}_n\}$ for $n \in \{2, \dots, N-1\}$. The vectors \mathbf{g}_n are random real Gaussian
 311 distributed datasets parameterized by their mean, *i.e.*,

$$312 \quad (4.3) \quad \mathbf{g}_n \sim \mathcal{N}(\boldsymbol{\mu}_n(\omega), \sigma_n^2 \mathbf{I}) \quad \text{where} \quad \boldsymbol{\mu}_n = \text{vec}\{\llbracket \mathbf{M}_{n-1,n}, \mathbf{A}_n, \mathbf{M}_{n,n}^{-\top} \rrbracket\},$$

¹In many applications, e.g., harmonic retrieval in MIMO channels, the number of elements K is often very small.

314 and σ_n^2 denote the variances of the Gaussian noise on \mathbf{g}_n .

315 The unknown real deterministic parameter $\boldsymbol{\omega}$ is such that

316 (4.4)
$$\boldsymbol{\omega} = [\text{vec}\{\mathbf{A}_1\}^\top \text{vec}\{\mathbf{M}_{1,1}\}^\top \text{vec}\{\mathbf{M}_{1,2}\}^\top \text{vec}\{\mathbf{A}_2\}^\top \dots \text{vec}\{\mathbf{A}_N\}^\top]^\top.$$

318 Second, the coherence between the TT-cores is ensured by the set of constraints

319 (4.5)
$$\mathbf{M}_{n,n} = \mathbf{M}_{n,n+1} \quad \forall n \in \{1, \dots, N-1\}.$$

321 Reshaping (4.5) under the form of vector functions depending on $\boldsymbol{\omega}$ yields

322 (4.6)
$$\mathbf{h}_n(\boldsymbol{\omega}) = \text{vec}\{\mathbf{M}_{n,n}\} - \text{vec}\{\mathbf{M}_{n,n+1}\} \quad \forall n \in \{1, \dots, N-1\}.$$

324 As a result, the model (4.3) under constraints (4.6) characterizes (3.8) under the form
325 of a general coupled model.

326 **4.2. Calculation of the Constrained Cramér-Rao bound.** Evaluating the
327 performance of the coupled model (4.3) under constraints (4.6) is necessary to compare
328 the relative performance of JIRAFE and DREAMFAC for the estimation of $\boldsymbol{\omega}$. The
329 standard tool for this task is the constrained Cramér-Rao bound (CCRB) [22, 40],
330 defined² as

331 (4.7)
$$\text{CCRB}(\boldsymbol{\omega}) = \mathbf{U} [\mathbf{U}^\top \mathbf{F} \mathbf{U}]^{-1} \mathbf{U}^\top,$$

332 where $\mathbf{F} \triangleq \mathbf{F}(\boldsymbol{\omega})$ is the Fisher Information Matrix (FIM) of the model, \mathbf{J}_h is the
333 Jacobian matrix related to the constraints $\mathbf{h}(\boldsymbol{\omega}) = \mathbf{0}$ and \mathbf{U} is a basis of $\ker(\mathbf{J}_h)$.

334 Additionally to the model parameter $\boldsymbol{\omega}$, let us define $\mathbf{x} = \text{vec}\{\boldsymbol{\mathcal{X}}\} \in \mathbb{R}^\ell$ ($\ell =$
335 $\Pi_n D_n$), that represents the vectorized low-rank approximation of $\boldsymbol{\mathcal{Y}}$. The parameter
336 \mathbf{x} can be linked to the model parameters through the relationship $\mathbf{x} - \mathbf{h}_\mathcal{X}(\boldsymbol{\omega}) = \mathbf{0}$,
337 where $\mathbf{h}_\mathcal{X}$ is a non-redundant deterministic vector function, everywhere differentiable
338 with respect to $\boldsymbol{\omega}$ such that

339 (4.8)
$$\mathbf{h}_\mathcal{X}(\boldsymbol{\omega}) = \text{vec}\{\llbracket \mathbf{A}_1, \dots, \mathbf{A}_N \rrbracket\}.$$

341 The above parameterization allows us to compute the CCRB on the reconstruction
342 of \mathbf{x} [19, p.125] as

343 (4.9)
$$\text{CCRB}(\mathbf{x}) = \left[\frac{\mathbf{h}_\mathcal{X}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \right]^\top \text{CCRB}(\boldsymbol{\omega}) \left[\frac{\mathbf{h}_\mathcal{X}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \right].$$

345 **4.2.1. Expression of the FIM.** The FIM for $\boldsymbol{\omega}$ first needs to be computed in
346 order to obtain the CCRB.

347 The FIM for $\boldsymbol{\omega}$ is a block-diagonal matrix of the form

348 (4.10)
$$\mathbf{F}(\boldsymbol{\omega}) = \begin{bmatrix} \mathbf{F}_1(\boldsymbol{\omega}) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \ddots & & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & & \mathbf{0} & \mathbf{F}_N(\boldsymbol{\omega}) \end{bmatrix},$$

350 where the blocks $\mathbf{F}_n(\boldsymbol{\omega})$ encode the contributions of each \mathbf{g}_n to the estimation of $\boldsymbol{\omega}$.
351 They are obtained by using the Slepian-Bangs formula [36]:

352 (4.11)
$$\mathbf{F}_n(\boldsymbol{\omega}) = \frac{1}{\sigma_n^2} \left[\frac{\partial \boldsymbol{\mu}_n(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \right]^\top \left[\frac{\partial \boldsymbol{\mu}_n(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \right].$$

²If \mathbf{F} is invertible, then (4.7) and the alternative expression for the CCRB provided in [15] are equivalent [40, Corollary 1].

354 The expression of $\left[\frac{\partial \boldsymbol{\mu}_n(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top}\right]$ is obtained simply for $n = 1$ and $n = N$:

$$355 \quad (4.12) \quad \frac{\partial \boldsymbol{\mu}_1(\boldsymbol{\omega})}{\partial \text{vec}\{\mathbf{A}_1\}^\top} = \mathbf{M}_{1,1}^{-\top} \boxtimes \mathbf{I}_{D_1},$$

$$356 \quad (4.13) \quad \frac{\partial \boldsymbol{\mu}_1(\boldsymbol{\omega})}{\partial \text{vec}\{\mathbf{M}_{1,1}\}^\top} = \mathbf{I}_K \boxtimes \mathbf{A}_1,$$

$$357 \quad (4.14) \quad \frac{\partial \boldsymbol{\mu}_1(\boldsymbol{\omega})}{\partial \text{vec}\{\mathbf{M}_{N-1,N}\}^\top} = \mathbf{A}_N \boxtimes \mathbf{I}_K,$$

$$358 \quad (4.15) \quad \frac{\partial \boldsymbol{\mu}_N(\boldsymbol{\omega})}{\partial \text{vec}\{\mathbf{A}_N\}^\top} = \boldsymbol{\Pi}_N \left(\mathbf{M}_{N-1,N}^{-1} \boxtimes \mathbf{I}_{D_N} \right),$$

360 where $\boldsymbol{\Pi}_N$ is a permutation matrix linking the entries of $\text{vec}\{\mathbf{G}_N^\top\}$ to those of $\text{vec}\{\mathbf{G}_N\}$.

361 For $n \in \{2, \dots, N-1\}$, $\left[\frac{\partial \boldsymbol{\mu}_n(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top}\right]$ can be computed using relationships between tensor
362 unfoldings:

$$363 \quad (4.16) \quad \mathbf{g}_n = [(\mathbf{M}_{n,n}^{-\top} \odot \mathbf{A}_n) \boxtimes \mathbf{I}_K] \text{vec}\{\mathbf{M}_{n-1,n}\}$$

$$364 \quad (4.17) \quad = \boldsymbol{\Pi}_n^{(2,1)} [(\mathbf{M}_{n,n}^{-\top} \odot \mathbf{M}_{n-1,n}) \boxtimes \mathbf{I}_{D_n}] \text{vec}\{\mathbf{A}_n\}$$

$$365 \quad (4.18) \quad = \boldsymbol{\Pi}_n^{(3,1)\top} [\mathbf{I}_K \boxtimes (\mathbf{A}_n \odot \mathbf{M}_{n-1,n})] \text{vec}\{\mathbf{M}_{n,n}\},$$

367 where $\boldsymbol{\Pi}_n^{(2,1)}$ and $\boldsymbol{\Pi}_n^{(3,1)}$ are permutation matrices that link the entries of $\text{vec}\{\mathbf{G}_n^{(2)}\}$

368 (resp. $\text{vec}\{\mathbf{G}_n^{(3)}\}$) to those of $\mathbf{g}_n = \text{vec}\{\mathbf{G}_n^{(1)}\}$. To ease the notation, we define the

369 matrices $\mathbf{S}_{1,n}, \mathbf{S}_{2,n}, \mathbf{S}_{3,n}$ ($n \in \{2, \dots, N-1\}$) as

$$370 \quad (4.19) \quad \mathbf{S}_{1,n} = [(\mathbf{M}_{n,n}^{-\top} \odot \mathbf{A}_n) \boxtimes \mathbf{I}_K],$$

$$371 \quad (4.20) \quad \mathbf{S}_{2,n} = \boldsymbol{\Pi}_n^{(2,1)} [(\mathbf{M}_{n,n}^{-\top} \odot \mathbf{M}_{n-1,n}) \boxtimes \mathbf{I}_{D_n}],$$

$$372 \quad (4.21) \quad \mathbf{S}_{3,n} = \boldsymbol{\Pi}_n^{(3,1)\top} [\mathbf{I}_K \boxtimes (\mathbf{A}_n \odot \mathbf{M}_{n-1,n})].$$

374 As a result, we have

$$375 \quad (4.22) \quad \left[\frac{\partial \boldsymbol{\mu}_n(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top}\right] = \begin{cases} \begin{bmatrix} \mathbf{M}_{1,1}^{-\top} \boxtimes \mathbf{I}_{D_1} & \mathbf{I}_K \boxtimes \mathbf{A}_1 \\ \mathbf{S}_{1,n} & \mathbf{S}_{2,n} & \mathbf{S}_{3,n} \end{bmatrix} & \text{for } n = 1, \\ \begin{bmatrix} \mathbf{A}_N \boxtimes \mathbf{I}_K & \boldsymbol{\Pi}_N \left(\mathbf{M}_{N-1,N}^{-1} \boxtimes \mathbf{I}_{D_N} \right) \end{bmatrix} & \text{for } n = N. \end{cases} \quad \forall n \in \{2, \dots, N-1\},$$

377 The closed-form expressions for the $\mathbf{F}_n(\boldsymbol{\omega})$ are provided in Appendix B.

378 **4.2.2. Expression of the CCRB.** The Jacobian matrix \mathbf{J}_h is obtained by
379 deriving the functions $\mathbf{h}_n(\boldsymbol{\omega})$ with respect to all the elements of $\boldsymbol{\omega}$ into a block-matrix.
380 Given the constraints on $\boldsymbol{\omega}$ in (4.6), it holds that

$$381 \quad (4.23) \quad \frac{\partial \mathbf{h}_n(\boldsymbol{\omega})}{\partial \text{vec}\{\mathbf{A}_n\}^\top} = \mathbf{0}, \quad \forall n \in \{1, \dots, N-1\},$$

$$382 \quad (4.24) \quad \frac{\partial \mathbf{h}_n(\boldsymbol{\omega})}{\partial \text{vec}\{\mathbf{M}_{m-1,m}\}^\top} = \begin{cases} \mathbf{I}_{K^2} & \text{if } m = n \\ \mathbf{0} & \text{if } m \neq n, \end{cases}$$

$$383 \quad (4.25) \quad \frac{\partial \mathbf{h}_n(\boldsymbol{\omega})}{\partial \text{vec}\{\mathbf{M}_{m,m}\}^\top} = \begin{cases} -\mathbf{I}_{K^2} & \text{if } m = n \\ \mathbf{0} & \text{if } m \neq n. \end{cases}$$

384

385 As a result, \mathbf{J}_h is such that

(4.26)

$$386 \quad \mathbf{J}_h = \begin{bmatrix} \frac{\partial \mathbf{h}_1(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \\ \vdots \\ \frac{\partial \mathbf{h}_{N-1}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_{K^2} & -\mathbf{I}_{K^2} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \mathbf{0} & \mathbf{0} & \vdots & \mathbf{I}_{K^2} & -\mathbf{I}_{K^2} & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \mathbf{0} & \mathbf{0} & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{I}_{K^2} & -\mathbf{I}_{K^2} & \mathbf{0} \end{bmatrix}.$$

387

388 The matrix \mathbf{U} can then be obtained by simply solving $\mathbf{J}_h \mathbf{U} = \mathbf{0}$. Therefore \mathbf{U} is the
 389 identity matrix of size $K(2N - 2 + \sum_n D_N)$. The CCRB submatrices for the \mathbf{A}_n and
 390 the $\mathbf{M}_{n-1,n}$ are obtained by developing expression (4.7):

391 (4.27) $\quad \text{CCRB}(\mathbf{A}_n) = \mathbf{F}(\mathbf{A}_n)^{-1}, \quad \text{CCRB}(\mathbf{M}_{n-1,n}) = \mathbf{F}(\mathbf{M}_{n-1,n})^{-1}.$

393 **4.3. Reparameterized CRB for reconstruction of the low-rank tensor.**
 394 The expression for $\mathbf{h}_{\mathcal{X}}(\boldsymbol{\omega})$ is necessary to compute the reparameterized CRB in (4.9).
 395 The relationships between tensor unfoldings give

396 (4.28) $\quad \mathbf{x} = [(\mathbf{A}_N \odot \dots \odot \mathbf{A}_2) \boxtimes \mathbf{I}_{D_1}] \text{vec}\{\mathbf{A}_1\}$
 397 $\quad = \dots$
 398 (4.29) $\quad = \mathbf{\Pi}_{\mathcal{X}}^{(N,1)} [(\mathbf{A}_{N-1} \odot \dots \odot \mathbf{A}_1) \boxtimes \mathbf{I}_{D_N}] \text{vec}\{\mathbf{A}_N\},$

400 where the $\mathbf{\Pi}_{\mathcal{X}}^{(n,1)}$ are permutation matrices that link the entries of the $\text{vec}\{\mathbf{X}^{(n)}\}$
 401 vectors to those of $\text{vec}\{\mathbf{X}^{(1)}\} = \mathbf{x}$. The matrices $\mathbf{S}_{n,\mathcal{X}}$ are defined such that

402 (4.30) $\quad \mathbf{S}_{n,\mathcal{X}} = \mathbf{\Pi}_{\mathcal{X}}^{(n,1)} [(\mathbf{A}_N \odot \dots \odot \mathbf{A}_{n+1} \odot \mathbf{A}_{n-1} \odot \dots \odot \mathbf{A}_1) \boxtimes \mathbf{I}_{D_n}],$

404 finally yielding

405 (4.31) $\quad \left[\frac{\mathbf{h}_{\mathcal{X}}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \right] = [\mathbf{S}_{1,\mathcal{X}} \quad \mathbf{0} \quad \mathbf{0} \quad \mathbf{S}_{2,\mathcal{X}} \quad \mathbf{0} \quad \mathbf{0} \quad \dots \quad \mathbf{S}_{N,\mathcal{X}}].$

406 Therefore,

407 (4.32) $\quad \text{CCRB}(\mathbf{x}) = \text{Diag}\{\mathbf{S}_{1,\mathcal{X}}^\top \mathbf{F}(\mathbf{A}_1)^{-1} \mathbf{S}_{1,\mathcal{X}}, \dots, \mathbf{S}_{N,\mathcal{X}}^\top \mathbf{F}(\mathbf{A}_N)^{-1} \mathbf{S}_{N,\mathcal{X}}\}.$

409 Given the expressions of the constrained Cramér-Rao bounds in Equation (4.27) and
 410 Equation (4.32), the performance of the proposed approach can now be evaluated
 411 numerically.

412 5. Simulations.

413 **5.1. Recovery of the CP factors.** A 7-order ($N = 7$) tensor \mathcal{X} with $D_1 =$
 414 $\dots = D_N = 6$ and $K = 3$ is taken as a reference. This tensor admits a reference
 415 CPD as in (3.1) with i.i.d. entries generated from the normal distribution. The noisy
 416 tensor \mathcal{Y} was generated using isotropic white Gaussian noise to yield a SNR of 20dB.

417 The proposed approach is compared to two JIRAFE-like algorithms. In the first
 418 one, Tri-ALS is initially performed on \mathcal{G}_2 and the factors is propagated towards the
 419 higher n . In the second one, Tri-ALS is performed on \mathcal{G}_{N-1} with propagation towards
 420 the lower n . The TT-cores are estimated using TT-SVD with $K = 3$. The CP factors
 421 were initialized as $\mathbf{A}_n = \mathbf{U}_n (\boldsymbol{\Sigma}_n)^{\frac{1}{2}}$ using the SVD of the n -th mode unfolding of \mathcal{Y}

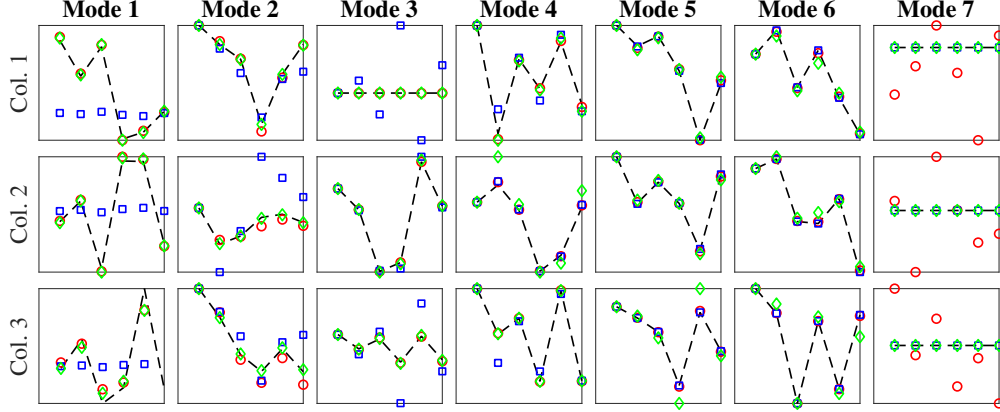


FIG. 2. Columns of the reference (black dashed lines) and estimated factors as returned by JIRAFE with forward propagation (red circles), JIRAFE with backward propagation (blue squares), and DREAMFAC (green diamonds).

422 with rank K , namely $\mathbf{Y}^{(n)} = \mathbf{U}_n \Sigma_n \mathbf{V}_n$. The algorithms have a maximum of 1000
 423 iterations, and the results are averaged over 100 noise realizations.

424 Figure 2 shows columns of the reference and estimated factors. The columns of the
 425 \mathbf{A}_n factors for $n \in \{2, \dots, N-1\}$ are correctly estimated by all the algorithms, except
 426 for a few outliers in \mathbf{A}_6 . Algorithm 3.2 usually provides a slightly better estimate
 427 than the JIRAFE procedures. While DREAMFAC also correctly estimates \mathbf{A}_1 and
 428 \mathbf{A}_N , the JIRAFE-like procedures provided an incorrect estimation of \mathbf{A}_N (resp. \mathbf{A}_1).

429 To further investigate the propagation of the error through the TT-cores, the
 430 normalized mean square error between the estimated $\hat{\mathbf{A}}_n$ and the reference \mathbf{A}_n is
 431 considered:

$$432 \quad (5.1) \quad \text{NMSE} = \frac{\|\hat{\mathbf{A}}_n - \mathbf{A}_n\|_F^2}{\|\mathbf{A}_n\|_F^2}.$$

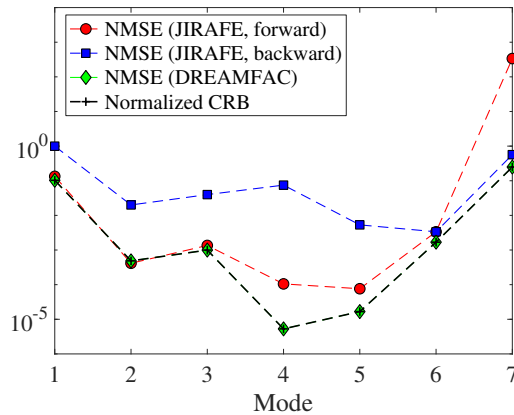


FIG. 3. NMSE provided by JIRAFE with forward propagation (red dots), backward propagation (blue squares), and DREAMFAC (green diamonds), and normalized CRB.

433 Figure 3 shows in semi-log scale the NMSE for each factor, as provided by the

three algorithms. As a comparison, the value of the uniform CCRB, normalized by $\|\mathbf{A}_n\|_F^2$, is also displayed. For the two JIRAFE-like procedures, the highest NMSE corresponded to the last estimated factor, resp. \mathbf{A}_N for the forward procedure and \mathbf{A}_1 for the backward procedure. The NMSE provided by DREAMFAC remained smaller for both \mathbf{A}_1 and \mathbf{A}_N . On average, it was also smaller than that provided by the JIRAFE algorithms, and reaches the optimal NMSE expected from the normalized CRB.

5.2. Relative efficiency of the algorithms. This subsection assesses the efficiency of JIRAFE and DREAMFAC by comparing their performance to the constrained Cramér-Rao bound obtained in Section 4.

The reference tensor \mathcal{X} is a 7th-order tensor with $D_1 = \dots = D_N = 6$ and $K = 3$. The entries of the true CP factors \mathbf{A}_n were generated once as i.i.d. real standard Gaussian variables. The model is simulated under additive Gaussian noise. We assume that the noise level was the same for all TT-cores: this assumption is reasonable since the \mathcal{G}_n are all estimated from \mathcal{Y} using TT-SVD. on the TT-cores. The SNR on the observed tensors in dB is defined as $SNR_i = 10 \log_{10} (\|\mathcal{G}_i\|_F^2 / \|\mathcal{E}_i\|_F^2)$, ($i = 1, \dots, N$).

The model parameters are retrieved using JIRAFE and DREAMFAC. The CP factors and ambiguity matrices are initialized randomly. The permutation ambiguities in the estimated factors are corrected using the Hungarian algorithm [25] to make comparisons aligned with the reference. The experiments show the uniform MSE and uniform CCRB obtained from the MSE and CCRB matrix traces, as widely considered in, e.g., [11, 16, 17]. The expressions for the bounds proposed in this paper permit the computation of the uniform CCRB by taking the trace of these matrices.

Figure 4 shows in semi-log scale the uniform bounds and MSEs for all the entries of the \mathbf{A}_n and \mathbf{M}_n , $\boldsymbol{\omega}$ and \mathcal{X} as a function of the SNR. It is noticeable that the uniform MSE produced by DREAMFAC reaches the uniform CCRB. Therefore the proposed approach is optimal for estimation of the parameters and reconstruction of \mathcal{X} . JIRAFE yields a higher MSE, which was expected since it is a suboptimal optimization method. The two algorithms depict the same kind of behavior with respect to the SNR. This is particularly visible for estimation of the ambiguity matrices. DREAMFAC permits to gain 5dB for the estimation of the \mathbf{A}_n , and 3dB for the reconstruction of \mathcal{X} compared to JIRAFE.

Figure 5 shows in semi-log scale the uniform CCRB for \mathbf{A}_1 , \mathbf{A}_3 , \mathbf{A}_5 and \mathbf{A}_7 , as well as the uniform MSEs, as a function of the SNR. The results for \mathbf{A}_2 , \mathbf{A}_4 and \mathbf{A}_6 are similar. The uniform MSE obtained with DREAMFAC reaches the CCRB for each factor, and its scale barely varies through the \mathbf{A}_n . The MSE obtained with JIRAFE was higher, and progressively increased for high values of n . Therefore DREAMFAC is also efficient for the estimation of each CP factor.

5.3. Application to channel estimation in dual-polarized MIMO systems. The problem of channel estimation in dual-polarized massive MIMO aims at recovering the channel parameters at the base station (angles of arrival, angles of departure, path gains, and polarization parameters). In [29], a tensor-based approach for this task was proposed. The MIMO channel was recast as a fourth-order tensor admitting a CPD. In [42], the authors adapted the JIRAFE procedure to treat the channel estimation problem with receiver and transmitter rectangular arrays, hence the channel was viewed as a fifth-order tensor.

5.3.1. Model description. The steering vectors for the k -th path between a Uniform Rectangular Array (URA) transmitter of size $M_T^x \times M_T^y$ and a URA receiver of size $M_R^x \times M_R^y$ are such that

$$(5.2) \quad \mathbf{a}_T(k) = \mathbf{a}_T^x(k) \boxtimes \mathbf{a}_T^y(k), \quad \mathbf{a}_R(k) = \mathbf{a}_R^x(k) \boxtimes \mathbf{a}_R^y(k),$$

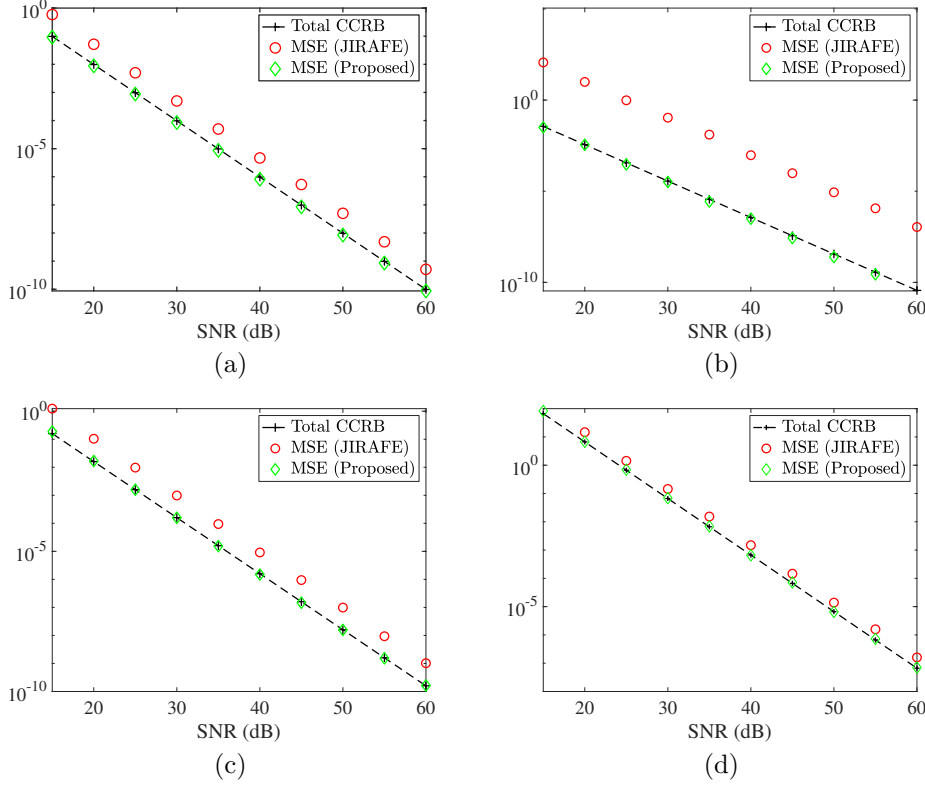


FIG. 4. Uniform CCRB and MSE provided by JIRAFE and DREAMFAC, for (a) estimation of all the \mathbf{A}_n , (b) estimation of all the \mathbf{M}_n , (c) estimation of $\boldsymbol{\omega}$ and (d) reconstruction of \mathcal{X} .

486 with e.g., $\mathbf{a}_T^x(k) = [1, \exp(j\omega_T^x(k)), \dots, \exp(j\omega_T^x(k)(M_T^x - 1))]^\top$, and likewise for
 487 $\mathbf{a}_T^y(k)$, $\mathbf{a}_R^x(k)$, $\mathbf{a}_R^y(k)$. As a result, for K paths, the steering matrices in transmission
 488 and in reception are

$$489 \quad (5.3) \quad \mathbf{A}_T = \mathbf{A}_T^x \odot \mathbf{A}_T^y, \quad \mathbf{A}_R = \mathbf{A}_R^x \odot \mathbf{A}_R^y,$$

491 with e.g., $\mathbf{A}_T^x = [\mathbf{a}_T^x(1), \dots, \mathbf{a}_T^x(K)]$.

492 The path-loss matrix $\mathbf{B} \in \mathbb{C}^{4 \times K}$ contains the path-loss parameters. Let the (p, q) -
 493 th subchannel correspond to $p \in \{V_R, H_R\}$ for the vertical (V) polarized and horizontal
 494 (H) polarized receive antennas, and $q \in \{V_T, H_T\}$ for the V-polarized and H-polarized
 495 transmit antennas. For the k -th path and for the (p, q) -th subchannel, $\beta_k^{(p,q)}$ is the
 496 path-loss parameter. Therefore, noting $\beta^{(p,q)} = [\beta_1^{(p,q)}, \dots, \beta_K^{(p,q)}]$,

$$497 \quad (5.4) \quad \mathbf{B} = \begin{bmatrix} \beta^{(V_R, V_T)} & \beta^{(V_R, H_T)} & \beta^{(H_R, V_T)} & \beta^{(H_R, H_T)} \end{bmatrix},$$

498 As a result, the channel tensor $\mathcal{H} \in \mathbb{C}^{M_T^x \times M_T^y \times M_R^x \times M_R^y \times 4}$ can be written as

$$499 \quad (5.5) \quad \mathcal{H} = \llbracket \mathbf{A}_R^x, \mathbf{A}_R^y, (\mathbf{A}_T^x)^\star, (\mathbf{A}_T^y)^\star, \mathbf{B} \rrbracket + \mathcal{E},$$

501 which corresponds to a fifth-order CPD of rank K . The noise term \mathcal{E} encompasses
 502 the background noise and the estimation error due to the pre-estimation of the un-
 503 structured channel, and can be modeled as zero-mean circularly complex Gaussian
 504 random variables.

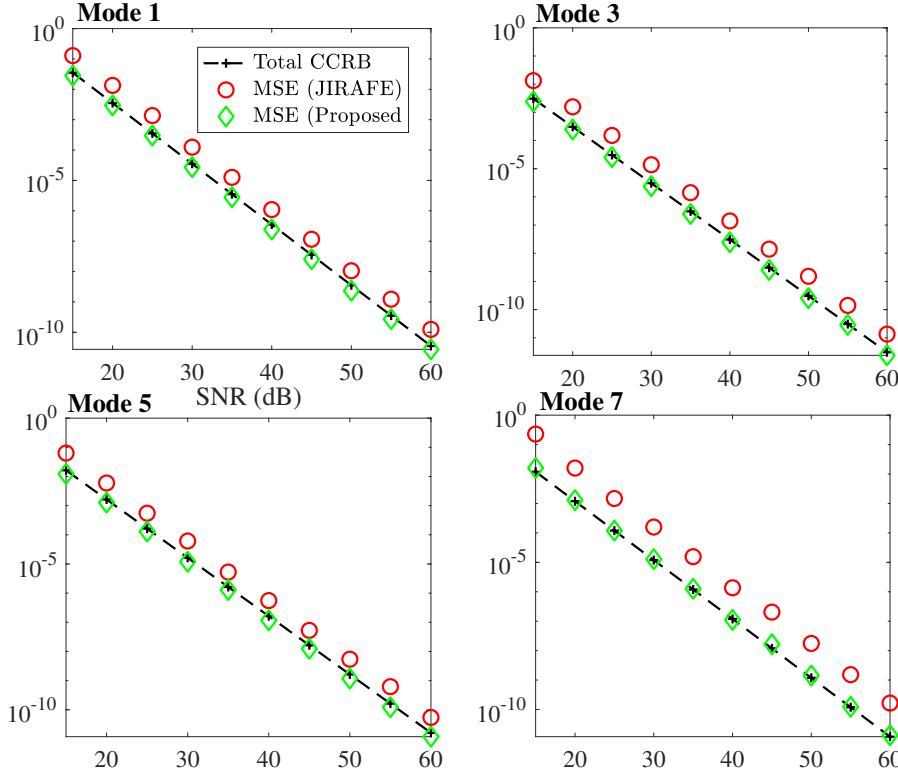


FIG. 5. Uniform CCRB and uniform MSE provided by JIRAFE and DREAMFAC for the estimation of the \mathbf{A}_n individually.

505 The main assumption of the model (5.5) is that $K \leq \min(M_T^x, M_T^y, M_R^x, M_R^y)$, *i.e.*,
 506 the steering matrices are full column-rank. Furthermore, it is assumed that there are
 507 few dominant paths, *i.e.*, $K < 4$ and \mathbf{B} is a full column-rank matrix. The assumption
 508 of a small number of paths is usually made in massive MIMO scenarios.

509 **5.3.2. Results.** The matrices $\mathbf{A}_T^x, \mathbf{A}_T^y, \mathbf{A}_R^x, \mathbf{A}_R^y$ were generated once based on
 510 single random realizations of the angular frequencies $\omega_T^x(k), \omega_T^y(k), \omega_R^x(k), \omega_R^y(k)$ fol-
 511 lowing a uniform distribution on $]0, \pi]$. The factor \mathbf{B} was drawn from a complex
 512 Gaussian distribution with zero mean and unit variance. The following dimensions
 513 were considered: $M_T^x = M_R^x = 10$ and $M_T^y = M_R^y = 8$, and $K = 3$.

514 Figure 6 shows in semi-log scale true factors and the estimated factors provided
 515 by DREAMFAC and JIRAFE, with 30dB SNR. All factors are recovered correctly
 516 by DREAMFAC. The factors estimated by JIRAFE seem coherent with the reference
 517 factors, but with the wrong scale and angle.

518 To further assess the performance of the proposed approach, the MSE between

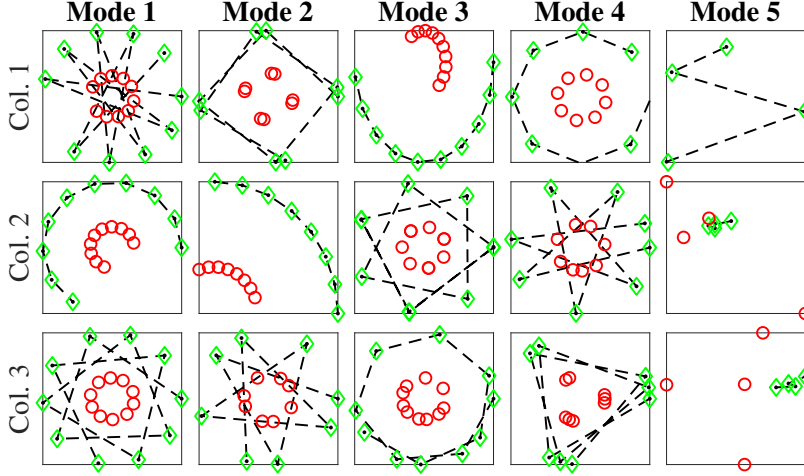


FIG. 6. Columns of the reference factors (black dashed lines) and estimated factors as returned by JIRAFE (red circles) and DREAMFAC (green diamonds).

519 the true and estimated factors, and between the angular frequencies, were considered:

(5.6)

$$520 \quad \text{MSE}_R^x = \sum_{k=1}^K \left(\mathbf{A}_R^x(:, k) - \hat{\mathbf{A}}_R^x(:, k) \right)^2, \quad \text{MSE}_R^y = \sum_{k=1}^K \left(\mathbf{A}_R^y(:, k) - \hat{\mathbf{A}}_R^y(:, k) \right)^2,$$

(5.7)

$$521 \quad \text{MSE}_T^x = \sum_{k=1}^K \left(\mathbf{A}_T^x(:, k) - \hat{\mathbf{A}}_T^x(:, k) \right)^2, \quad \text{MSE}_T^y = \sum_{k=1}^K \left(\mathbf{A}_T^y(:, k) - \hat{\mathbf{A}}_T^y(:, k) \right)^2,$$

(5.8)

$$522 \quad \text{MSE}_\omega = \sum_{k=1}^K \left(\omega_T^x(k) - \hat{\omega}_T^x(k) \right)^2 + \left(\omega_T^y(k) - \hat{\omega}_T^y(k) \right)^2 \\ 523 \quad + \left(\omega_R^x(k) - \hat{\omega}_R^x(k) \right)^2 + \left(\omega_R^y(k) - \hat{\omega}_R^y(k) \right)^2.$$

525 The MSE were evaluated over 10 values of the SNR in $[-20, 30]$ dB, and calculated by
 526 averaging the results over 500 independent noise realizations. Computing the complex
 527 CCRB is a delicate task that requires the calculation of Wirtinger derivatives [26].
 528 For this reason, calculation of the CRB associated with model (5.5) is relegated to
 529 future works.

530 DREAMFAC was compared to four tensor-based methods. The first one was
 531 CP-ALS [9] followed by closed-form solutions to estimate the parameters from the
 532 factors. The second one was the so-called CP-VDM, for CPD with Vandermonde factor
 533 matrix, proposed in [37]. The third one was based on the generalized eigenvalue
 534 decomposition (GEVD) [33]. The fourth one was JIRAFE followed by a Vandermonde
 535 rectification strategy [5] (termed JIRAFE-VDM) to enforce the structure of the steering
 536 matrices. This step is crucial to estimate the factors with the correct scale and
 537 angle, as exemplified in Figure 6. All algorithms used at most 1000 iterations.

538 Figure 7 shows in semi-log scale the averaged MSE for recovery of the steering
 539 matrices. For a SNR superior or equal to 5dB, all approaches but GEVD yield the
 540 same MSE. Thanks to the Vandermonde rectification, JIRAFE could achieve good

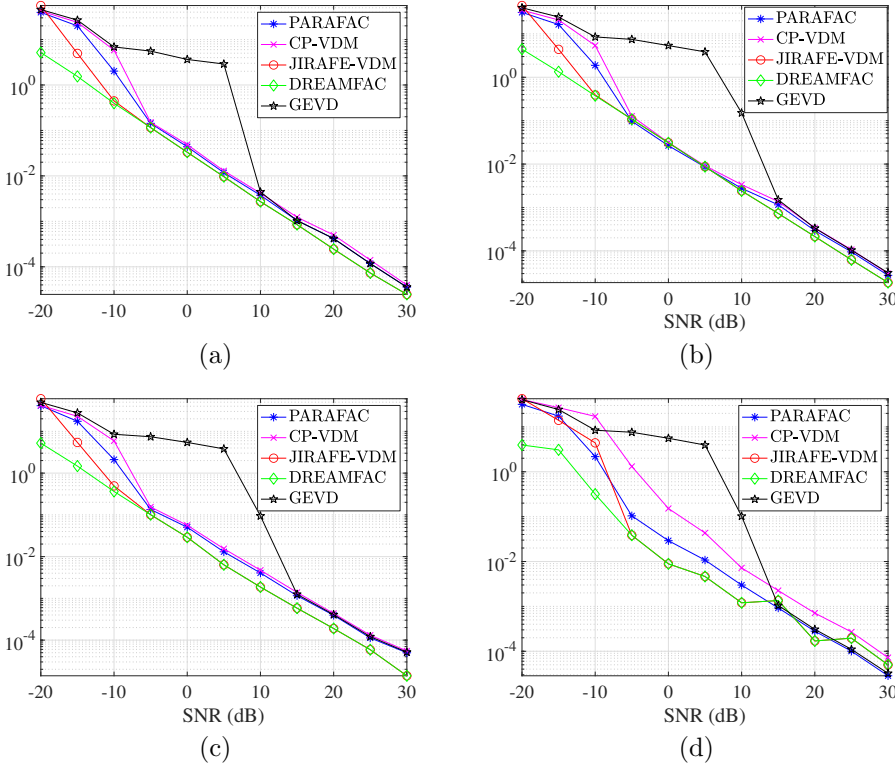


FIG. 7. (a) MSE_R^x , (b) MSE_R^y , (c) MSE_T^x , (d) MSE_T^y as a function of the SNR.

541 performance for higher values of the SNR. For lower values of the SNR, DREAMFAC
 542 yields the lowest MSE. Its performance is better than that of JIRAFE, with a gain of
 543 approximately 3dB on the quality of the estimation.

544 Figure 8 shows in semi-log scale the averaged MSE_ω provided by the algorithms,
 545 as a function of the SNR. For estimation of the angular frequencies, DREAMFAC
 546 yields the best MSE for a low SNR. Its gain with respect to JIRAFE for low SNR
 547 is of approximately 5dB. The proposed approach has similar performance to other
 548 approaches for a SNR superior or equal to 5dB.

549 **6. Conclusion.** This paper proposes a new approach called DREAMFAC for
 550 Dimensionality Reduction, joint Estimation of the Ambiguity Matrices and the CP
 551 Factors. It relies on a global coupled optimization scheme, instead of a local and
 552 sequential strategy in previous approaches. DREAMFAC performs favorably with
 553 respect to the state-of-the-art of estimation of the CP factors. We have derived con-
 554 strained Cramér-Rao bounds to evaluate the potential optimum performance under
 555 mild assumptions, as well as for comparison with state-of-the-art methods. Numerical
 556 experiments show that the proposed algorithm reaches the Cramér-Rao bound,
 557 as expected since it fully exploits the information of the full coupled model.

558 The performance of DREAMFAC is exemplified on the realistic problem of complex
 559 harmonics retrieval in dual-polarized MIMO channels. The proposed approach
 560 performs well on this task, providing better estimation of the base station parameter
 561 at low SNR than state-of-the-art methods. As a conclusion, this paper clarifies what
 562 are the best performance of a low-rank tensor CP factorization using tensor-train.

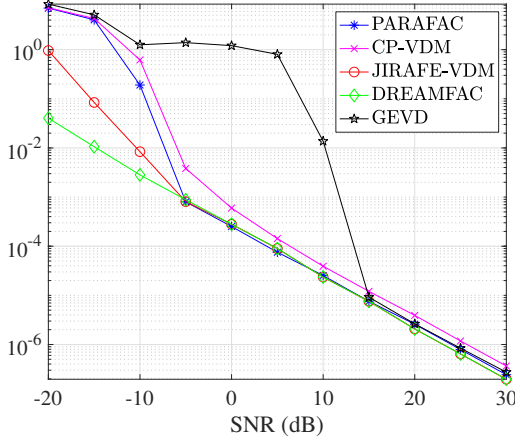


FIG. 8. *MSE on the estimation of the angular frequencies.*

563 The proposed approach, DREAMFAC, is an optimal solution to this problem.

564 **Appendix A. Closed-form expression for the coupled M_n in the pro-**
 565 **posed approach.**

566 Solutions the least-squared problems (3.9)–(3.11) can be obtained by solving the
 567 following Sylvester equations:

$$568 \quad (\text{A.1}) \quad \mathbf{M}_1 (\mathbf{M}_2^{-1} \mathbf{M}_2^{-\top} \square \mathbf{A}_2^{\top} \mathbf{A}_2) + (\mathbf{G}_1^{\top} \mathbf{G}_1) \mathbf{M}_1 = \mathbf{G}_1^{\top} \mathbf{A}_1 + \mathbf{G}_2^{(1)\top} (\mathbf{A}_2 \odot \mathbf{M}_2),$$

(A.2)

$$569 \quad \mathbf{M}_{n-1} (\mathbf{M}_n^{-1} \mathbf{M}_n^{-\top} \square \mathbf{A}_n^{\top} \mathbf{A}_n) + (\mathbf{G}_{n-1}^{(3)\top} \mathbf{G}_{n-1}^{(3)}) \mathbf{M}_{n-1} \\ 570 \quad = \mathbf{G}_n^{(1)\top} (\mathbf{M}_n^{-\top} \odot \mathbf{A}_n) + \mathbf{G}_{n-1}^{(3)\top} (\mathbf{A}_{n-1} \odot \mathbf{M}_{n-2}),$$

(A.3)

$$571 \quad \mathbf{M}_n (\mathbf{M}_{n+1}^{-1} \mathbf{M}_{n+1}^{-\top} \square \mathbf{A}_{n+1}^{\top} \mathbf{A}_{n+1}) + (\mathbf{G}_n^{(3)\top} \mathbf{G}_n^{(3)}) \mathbf{M}_n \\ 572 \quad = \mathbf{G}_{n+1}^{(1)\top} (\mathbf{M}_{n+1}^{-\top} \odot \mathbf{A}_{n+1}) + \mathbf{G}_n^{(3)\top} (\mathbf{A}_n \odot \mathbf{M}_{n-1}),$$

$$573 \quad (\text{A.4}) \quad \mathbf{M}_{N-1} (\mathbf{A}_N^{\top} \mathbf{A}_N) + (\mathbf{G}_{N-1}^{(3)\top} \mathbf{G}_{N-1}^{(3)}) \mathbf{M}_{N-1} \\ 574 \quad = \mathbf{G}_N \mathbf{A}_N + \mathbf{G}_{N-1}^{(3)\top} (\mathbf{A}_{N-1} \odot \mathbf{M}_{N-2}).$$

576 Fast solvers (see [35]) can be used to solve (A.1)–(A.4).

577 **Appendix B. Closed-form expressions for the uncoupled FIM.**

578 The FIM matrices $\mathbf{F}_n(\boldsymbol{\omega})$ introduced in (4.11) are block-matrices such as

(B.1)

579
$$\mathbf{F}_1(\boldsymbol{\omega}) = \frac{1}{\sigma_1^2} \begin{bmatrix} \mathbf{M}_{1,1}^{-1} \mathbf{M}_{1,1}^{-\top} \boxtimes \mathbf{I}_{d_1} & \mathbf{M}_{1,1}^{-1} \boxtimes \mathbf{A}_1 \\ \mathbf{M}_{1,1}^{-\top} \boxtimes \mathbf{A}_1^\top & \mathbf{I}_R \boxtimes \mathbf{A}_1^\top \mathbf{A}_1 \end{bmatrix},$$

(B.2)

580
$$\mathbf{F}_n(\boldsymbol{\omega}) = \frac{1}{\sigma_n^2} \begin{bmatrix} \mathbf{S}_{1,n}^\top \mathbf{S}_{1,n} & \mathbf{S}_{1,n}^\top \mathbf{S}_{2,n} & \mathbf{S}_{1,n}^\top \mathbf{S}_{3,n} \\ \mathbf{S}_{2,n}^\top \mathbf{S}_{1,n} & \mathbf{S}_{2,n}^\top \mathbf{S}_{2,n} & \mathbf{S}_{2,n}^\top \mathbf{S}_{3,n} \\ \mathbf{S}_{3,n}^\top \mathbf{S}_{1,n} & \mathbf{S}_{3,n}^\top \mathbf{S}_{2,n} & \mathbf{S}_{3,n}^\top \mathbf{S}_{3,n} \end{bmatrix} \text{ for } n \in \{2, \dots, N-1\},$$

(B.3)

581
$$\mathbf{F}_N(\boldsymbol{\omega}) = \frac{1}{\sigma_N^2} \begin{bmatrix} \mathbf{A}_N^\top \mathbf{A}_N \boxtimes \mathbf{I}_R & (\mathbf{A}_N^\top \boxtimes \mathbf{I}_R) \boldsymbol{\Pi}_N \left(\mathbf{M}_{N-1,N}^{-1} \boxtimes \mathbf{I}_{D_N} \right) \\ \left(\mathbf{M}_{N-1,N}^{-\top} \boxtimes \mathbf{I}_{D_N} \right) \boldsymbol{\Pi}_N^\top (\mathbf{A}_N \boxtimes \mathbf{I}_R) & \mathbf{M}_{N-1,N}^{-\top} \mathbf{M}_{N-1,N}^{-1} \boxtimes \mathbf{I}_{D_N} \end{bmatrix}.$$

583 It is possible to identify the blocks containing the contribution of the \mathbf{A}_n and the
584 $\mathbf{M}_{n-1,n}$ to the estimation of $\boldsymbol{\omega}$). These blocks, denoted to as $\mathbf{F}(\mathbf{A}_n)$ and $\mathbf{F}(\mathbf{M}_{n-1,n})$,
585 are such that

586 (B.4)
$$\mathbf{F}(\mathbf{A}_n) = \begin{cases} \mathbf{I}_R \boxtimes \mathbf{A}_1^\top \mathbf{A}_1 & \text{for } n = 1, \\ \mathbf{S}_{2,n}^\top \mathbf{S}_{2,n} & \text{for } n \in \{2, \dots, N-1\}, \\ \mathbf{A}_N^\top \mathbf{A}_N \boxtimes \mathbf{I}_R & \text{for } n = N, \end{cases}$$

587 (B.5)
$$\mathbf{F}(\mathbf{M}_{n-1,n}) = \begin{cases} \mathbf{S}_{1,n}^\top \mathbf{S}_{1,n} & \text{for } n \in \{2, \dots, N-1\}, \\ \mathbf{M}_{N-1,N}^{-\top} \mathbf{M}_{N-1,N}^{-1} \boxtimes \mathbf{I}_{D_N} & \text{for } n = N. \end{cases}$$

589 REFERENCES

590 [1] E. ACAR, C. AYKUT-BINGOL, H. BINGOL, R. BRO, AND B. YENER, *Multiway analysis of epilepsy*
591 *tensors*, *Bioinformatics*, 23 (2007), pp. i10–i18.
592 [2] D. C. ARAÚJO, T. MAKSYMUK, A. L. F. DE ALMEIDA, T. MACIEL, J. C. M. MOTA, AND
593 M. JO, *Massive mimo: survey and future research topics*, *IET Communications*, 10 (2016),
594 pp. 1938–1946.
595 [3] R. BARTELS AND G. STEWART, *Solution of the matrix equation $AX+XB=C$* , *Commun. ACM*,
596 15 (1972), pp. 820–826.
597 [4] M. BOIZARD, R. BOYER, G. FAVIER, J. COHEN, AND P. COMON, *Performance estimation for*
598 *tensor CP decomposition with structured factors*, in *Proc. ICASSP*, 2015.
599 [5] R. BOYER AND P. COMON, *Rectified als algorithm for multidimensional harmonic retrieval*, in
600 2016 IEEE SAM Signal Processing Workshop, IEEE, 2016, pp. 1–5.
601 [6] R. CABRAL FARIAS, J. COHEN, AND P. COMON, *Exploring multimodal data fusion through joint*
602 *decompositions with flexible couplings*, *IEEE Trans. Signal Process.*, 64 (2016), pp. 4830–
603 4844.
604 [7] P. COMON, *Tensors: A brief introduction*, *IEEE Signal Process. Mag.*, 31 (2014), pp. 44–53.
605 [8] H. CRAMÉR, *Mathematical Methods of Statistics*, Univ. Press, Princeton, 1946.
606 [9] L. DE LATHAUWER AND D. NION, *Decompositions of a higher-order tensor in block terms—part*
607 *III: Alternating least squares algorithms*, *SIAM journal on Matrix Analysis and Applica-*
608 *tions*, 30 (2008), pp. 1067–1083.
609 [10] L. DOMANOV, I. AND DE LATHAUWER, *Canonical polyadic decomposition of third-order tensors:*
610 *Reduction to generalized eigenvalue decomposition*, *SIAM Journal on Matrix Analysis and*
611 *Applications*, 35 (2014), pp. 636–660.
612 [11] Y. C. ELДАР, *Minimum variance in biased estimation: Bounds and asymptotically optimal*
613 *estimators*, *IEEE Transactions on Signal Processing*, 52 (2004), pp. 1915–1930.
614 [12] M. FRÉCHET, *Sur l’extension de certaines évaluations statistiques au cas de petits échantillons*,
615 *Rev. Int. Stat.*, 11 (1943), pp. 182–205.
616 [13] M. GIRAUD, V. ITIER, R. BOYER, Y. ZNIYED, AND A. L. F. DE ALMEIDA, *Tucker decomposition*
617 *based on a tensor train of coupled and constrained cp cores*, *IEEE Signal Process. Letters*,
618 (2023).
619 [14] G. GOLUB, S. NASH, AND C. V. LOAN, *A Hessenberg-Schur method for the problem $AX+XB=C$* ,
620 *IEEE Trans. Autom. Control*, 24 (1979), pp. 909–913.

- 621 [15] J. GORMAN AND A. HERO, *Lower bounds for parametric estimation with constraints*, IEEE
622 Trans. Inf. Theory, 36 (1990), pp. 1285–1301.
- 623 [16] A. O. HERO, *A Cramér-Rao type lower bound for essentially unbiased parameter estimation*,
624 tech. report, Massachussets Inst. of tech., Lexington Lincoln Lab, 1992.
- 625 [17] A. O. HERO, J. A. FESSLER, AND M. USMAN, *Exploring estimator bias-variance tradeoffs using*
626 *the uniform CR bound*, IEEE Transactions on Signal Processing, 44 (1996), pp. 2026–2041.
- 627 [18] T. G. KOLDA AND B. W. BADER, *Tensor Decompositions and Applications*, SIAM Review, 51
628 (2009), pp. 455–500.
- 629 [19] E. LEHMANN AND G. CASELLA, *Theory of Point Estimation (2nd ed.)*, Springer, 1998.
- 630 [20] N. LI, S. KINDERMANN, AND C. NAVASCA, *Some convergence results on the regularized alter-*
631 *nating least-squares method for tensor decomposition*, Linear Algebra and its Applications,
632 438 (2013), pp. 796–812.
- 633 [21] X. LIU AND N. SIDIROPOULOS, *Cramér-Rao lower bounds for low-rank decomposition of multi-*
634 *dimensional arrays*, IEEE Trans. Signal Process., 49 (2001), pp. 2074–2086.
- 635 [22] T. MENNI, E. CHAUMETTE, P. LARZABAL, AND J. P. BARBOT, *New results on Deterministic*
636 *Cramér-Rao bounds for real and complex parameters*, IEEE Trans. on SP, 60 (2012),
637 pp. 1032–1049.
- 638 [23] T. MENNI, J. GALY, E. CHAUMETTE, AND P. LARZABAL, *Versatility of Constrained CRB for*
639 *System Analysis and Design*, IEEE Trans. on AES, 50 (2014), pp. 1841–1863.
- 640 [24] T. J. MOORE, B. M. SADLER, AND R. J. KOZICK, *Maximum-Likelihood Estimation, the Cramér-*
641 *Rao Bound, and the Method of Scoring With Parameter Constraints*, IEEE Trans. on SP,
642 56 (2008), pp. 895–908.
- 643 [25] J. MUNKRES, *Algorithms for the assignment and transportation problems*, Journal of the society
644 for industrial and applied mathematics, 5 (1957), pp. 32–38.
- 645 [26] E. OLLILA, V. KOIVUNEN, AND J. ERIKSSON, *On the cramér-rao bound for the constrained*
646 *and unconstrained complex parameters*, in 2008 5th IEEE Sensor Array and Multichannel
647 Signal Processing Workshop, IEEE, 2008, pp. 414–418.
- 648 [27] I. V. OSELEDETS, *Tensor-train decomposition*, SIAM Journal on Scientific Computing, 33
649 (2011), pp. 2295–2317.
- 650 [28] I. V. OSELEDETS AND E. E. TYRTYSHNIKOV, *Breaking the curse of dimensionality, or how to use*
651 *sud in many dimensions*, SIAM Journal on Scientific Computing, 31 (2009), pp. 3744–3759.
- 652 [29] C. QIAN, X. FU, N. D. SIDIROPOULOS, AND Y. YANG, *Tensor-based channel estimation for*
653 *dual-polarized massive mimo systems*, IEEE Trans. Signal Process., 66 (2018), pp. 6390–
654 6403.
- 655 [30] C. R. RAO, *Information and accuracy attainable in the estimation of statistical parameters*,
656 Bull. Calcutta Math. Soc, 37 (1945), pp. 81–91.
- 657 [31] C. REN, R. CABRAL FARIAS, P.-O. AMBLARD, AND P. COMON, *Performance bounds for coupled*
658 *models*, in Proc. 2016 IEEE SAM, 2016. event-place: Rio de Janeiro, Brazil.
- 659 [32] S. SAHNOUN AND P. COMON, *Joint source estimation and localization*, IEEE Trans. Signal
660 Process., 63 (2015), pp. 2485–2595.
- 661 [33] E. SANCHEZ AND B. R. KOWALSKI, *Tensorial resolution: a direct trilinear decomposition*, Jour-
662 nal of Chemometrics, 4 (1990), pp. 29–45.
- 663 [34] N. D. SIDIROPOULOS, L. DE LATHAUWER, X. FU, K. HUANG, E. E. PAPALEXAKIS, AND
664 C. FALOUTSOS, *Tensor decomposition for signal processing and machine learning*, IEEE
665 Transactions on Signal Processing, 65 (2017), pp. 3551–3582.
- 666 [35] V. SIMONCINI, *Computational methods for linear matrix equations*, SIAM Review, 58 (2016),
667 pp. 377–441, <https://doi.org/10.1137/130912839>.
- 668 [36] D. SLEPIAN, *Estimation of signal parameters in the presence of noise*, Trans. IRE Professional
669 Group Inf. Theory, 3 (1954), pp. 68–69.
- 670 [37] M. SØRENSEN AND L. DE LATHAUWER, *Blind signal separation via tensor decomposition with*
671 *vandermonde factor: Canonical polyadic decomposition*, IEEE Transactions on Signal Process-
672 ing, 61 (2013), pp. 5507–5519.
- 673 [38] A. STEGEMAN AND N. D. SIDIROPOULOS, *On kruskal’s uniqueness condition for the cande-*
674 *comp/parafac decomposition*, Linear Algebra and its applications, 420 (2007), pp. 540–552.
- 675 [39] P. STOICA AND T. L. MARZETTA, *Parameter estimation problems with singular information*
676 *matrices*, IEEE Trans. on SP, 49 (2001), pp. 87–90.
- 677 [40] P. STOICA AND B. C. NG, *On the Cramér-Rao bound under parametric constraints*, IEEE SP
678 Letters, 5 (1998), pp. 177–179.
- 679 [41] Q. ZHAO, G. ZHOU, L. ZHANG, A. CICHOCKI, AND S.-I. AMARI, *Bayesian robust tensor factor-*
680 *ization for incomplete multiway data*, IEEE Trans. Neural Networks and Learning Systems,
681 27 (2015), pp. 736–748.
- 682 [42] Y. ZNIYED, R. BOYER, A. DE ALMEIDA, AND G. FAVIER, *Tensor train representation of mimo*
683 *channels using the jirafe method*, Signal Processing, 171 (2020), p. 107479.
- 684 [43] Y. ZNIYED, R. BOYER, A. F. DE ALMEIDA, AND G. FAVIER, *High-order tensor estimation*
685 *via trains of coupled third-order cp and tucker decompositions*, Linear Algebra and its
686 Applications, 588 (2020), pp. 304–337.