Constrained Cramér–Rao bounds for reconstruction problems formulated as coupled canonical polyadic decompositions

Clémence Prévost, Konstantin Usevich, Martin Haardt, Pierre Comon, David Brie

Constrained Cramér-Rao bounds for reconstruction problems formulated as coupled canonical polyadic decompositions*

Clémence Prévost[1,*], Konstantin Usevich[1], Martin Haardt[2,**], Pierre Comon[3,**], David Brie[1,**]

Abstract

We propose a theoretical performance analysis for a class of reconstruction problems, formulated as coupled canonical polyadic decompositions of two low-resolution tensor observations. We study a particular case when all the modes of the tensors are coupled. Unlike the case of a single coupling constraint, a fully-coupled model requires nonlinear constraints in some estimation scenarios. Thus we introduce two probabilistic scenarios. For each scenario, we derive the constrained Cramér-Rao bounds for the parameters and for the mean-squared error of the reconstructed tensor. We show that with a carefully chosen initialization, the maximum likelihood estimators reach the bounds, even in challenging cases (low signal-to-noise ratio or large tensor rank).

Keywords: multimodal data fusion, coupled tensor decompositions, Cramér-Rao bounds

1. Introduction

In the data fusion community, it is now commonly accepted that a same phenomenon can be partially contained in observations from several measurement devices, with different resolutions and noise contaminations [1, 2, 3]. The observations are often complementary, meaning that given dataset-specific information can be enriched from information contained in other datasets, and vice-versa. Hence data fusion was proposed to exploit the complementarity of available measurements [4].

In some engineering fields such as remote sensing or biomedical imaging, observed data often possess more than two dimensions, thus they can be represented as tensors. Several low-rank factorizations can be considered for approximating such data. A popular one is the canonical polyadic (CP) decomposition (CPD), due to its powerful uniqueness conditions. Tensor data fusion based on coupled CPD has since then proved its relevance in a wide range of applications [5, 6, 7].

*This work was partially supported by the ANR (Agence Nationale de Recherche) under grant LeaFleT (ANR-19-CE23-0021).
*Corresponding author.
**EURASIP member.
[1]Centre de Recherche en Automatique de Nancy (CRAN), Université de Lorraine, CNRS, Vandoeuvre-lès-Nancy, France. e-mail: firstname.lastname@univ-lorraine.fr.
[2]Communications Research Laboratory, Ilmenau University of Technology, Ilmenau, Germany. e-mail: firstname.lastname@tu-ilmenau.de.
[3]CNRS, GIPSA-Lab, Univ. Grenoble Alpes, F-38000 Grenoble, France. e-mail: firstname.lastname@gipsa-lab.grenoble-inp.fr.

In this paper, we consider a specific class of reconstruction problems, which aim at recovering a high-resolution tensor from tensor observations with some lower resolutions. Such problems can be found in hyperspectral super-resolution [8], biomedical imaging [9, 10], chemistry [11], or learning over graphs [12].

The work of [8], motivated by hyperspectral super-resolution, addresses the problem of reconstructing a tensor from two degraded versions. While one is degraded in two (spatial) modes, the second is degraded in the third (spectral) mode. Two scenarios were considered: fully-coupled and blind (partially coupled). Algorithms based on alternating least squares (ALS) were proposed for both scenarios and showed competitive reconstruction performance. This approach gave rise to numerous tensor-based reconstruction methods [13, 14].

Cramér-Rao bounds (CRB) are a classic tool to assess the performance of the estimators [15, 16, 17]. For coupled models, where the model parameters are subject (in part or totally) to deterministic constraints, the constrained Cramér-Rao bound (CCRB) can be used, whose versatility was shown by numerous works [18, 19, 20, 21, 22, 23]. Cramér-Rao bounds for tensor CP models have been studied in a general context. In [24, 25, 26], performance bounds for uncoupled CP models have been provided. In [27], a Bayesian framework was proposed for flexible coupling models and hybrid CRB were derived. Constrained Cramér-Rao bounds for partially coupled complex tensors admitting a CPD and a single coupling constraint were explored in [28]. The expression of the bound was based on [29], which considers a specific case where the Fisher information matrix (FIM) for the parameters is invertible. A single equality constraint between two shared CP factors was considered. In the reconstruction problems however, all the modes are coupled and the analysis of [28] is not applicable. Extending [28], a preliminary work of the authors [30] treated a special case of the degradation matrices. Another extension in the case of a single random parameter was considered in [31].

In this paper, we extend the results of [30] to the general reconstruction problem with general degradation matrices. Unlike the case of a single coupling constraint, a fully-coupled model requires nonlinear constraints in some estimation scenarios. We derive the CCRB both for the model parameters and for the mean-squared error of the reconstructed tensor, for two probabilistic scenarios; in particular, our results do not require identifiability of the individual tensors. We show that the maximum likelihood estimators reach the bounds, but their initialization should be carefully chosen since conventional initializations might lead to poor results. We propose an algorithm that gradually changes the regularization (balance) parameter between the tensors.

This paper is structured as follows. In Section 2, we recall preliminaries of tensor decompositions, their uniqueness properties, and formulate the reconstruction problem that we consider in the paper. Section 3 contains background on Cramér-Rao bounds. In particular, it provides a link between noiseless identifiability and the rank of the Fisher information matrix. Section 4 introduces the two different parameterizations and estimation scenarios that we consider in the paper. Next, in Section 5 we derive the bounds for the two scenarios. Finally, Section 6 contains our numerical results.

Contributions. In Section 3.1, we provide a clear link between generic uniqueness of the CPD, local identifiability in the noiseless case, and invertibility of the FIM. We introduce a general framework for the calculation of bounds for coupled CP models of the form (6)–(7). We introduce two estimation scenarios. The first one (see Section 4.3) evaluates the performance of the fullly coupled model for tensor reconstruction. The second one (see Section 4.4) compares the theoretical performance of the uncoupled, partially coupled and fully coupled models. In a supplementary material, we provide closed-form expressions for the CCRB in both scenarios. We assess the relative efficiency of two state-of-the-art estimators. Our numerical results show that these algorithms reach the bounds, even in challenging cases (see Section 6.4 and Section 6.6).

Notation. We follow the notations of [32, 33]. We use lower (a) or uppercase (A) plain font for scalars, boldface lowercase ($\boldsymbol{a}$) for vectors, boldface uppercase ($\boldsymbol{A}$) for matrices and calligraphic ($\boldsymbol{\mathcal{A}}$) for tensors. The elements of vectors, matrices, and tensors are denoted to as a_i, $A_{i,j}$ and $\mathcal{A}_{i_1,\ldots,i_N}$, respectively. The transpose of a matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{A}^\mathsf{T}$. The matrix $\boldsymbol{I}_N$ is the $N \times N$ identity matrix and $\boldsymbol{0}_{L\times K}$ is the $L \times K$ matrix of zeros. The symbols $\boxtimes$, $\odot$ and $\otimes$ denote the Kronecker, Khatri-Rao and outer products. We use vec$\{\cdot\}$ for the standard column-major vectorization of a matrix or a tensor. The operator diag$\{\boldsymbol{a}\}$ produces a diagonal matrix whose entries are the elements of $\boldsymbol{a}$. The operation Diag$\{\boldsymbol{A}, \boldsymbol{B}\}$ produces a block-diagonal matrix whose blocks are $\boldsymbol{A}$ and $\boldsymbol{B}$.

2. Background on tensor algebra

2.1. Preliminaries

A third-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I\times J\times K}$ is a three-dimensional array indexed by the elements $\mathcal{X}_{i,j,k}$, for $i \in \{1,\ldots,I\}$, $j \in \{1,\ldots,J\}$ and $k \in \{1,\ldots,K\}$. Each dimension of a tensor is called a mode. A mode-p fiber of $\boldsymbol{\mathcal{X}}$ is a vector obtained by fixing all but the p-th dimension. Similarly, a mode-p slab of $\boldsymbol{\mathcal{X}}$ is a matrix obtained by fixing only the p-th dimension.

Definition 2.1. ***Tensor unfoldings*** *– The mode-p unfolding of a tensor $\boldsymbol{\mathcal{X}}$, denoted by $\boldsymbol{X}^{(p)}$, is the matrix whose rows are the p-mode fibers of $\boldsymbol{\mathcal{X}}$, ordered according to the vectorization order. For a third-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I\times J\times K}$, we have $\boldsymbol{X}^{(1)} \in \mathbb{R}^{JK\times I}$, $\boldsymbol{X}^{(2)} \in \mathbb{R}^{IK\times J}$ and $\boldsymbol{X}^{(3)} \in \mathbb{R}^{IJ\times K}$.*

Definition 2.2. ***Matrix mode product*** *– The matrix p-mode product between a tensor $\boldsymbol{\mathcal{X}}$ and a matrix $\boldsymbol{M}$ is denoted by $\boldsymbol{\mathcal{X}} \bullet_p \boldsymbol{M}$ and is constructed such that each mode-p fiber of $\boldsymbol{\mathcal{X}}$ is multiplied by $\boldsymbol{M}$, e.g., the elements of the mode-1 product between $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I\times J\times K}$ and $\boldsymbol{M} \in \mathbb{R}^{L\times I}$ are accessed as $(\boldsymbol{\mathcal{X}} \bullet_1 \boldsymbol{M})_{\ell,j,k} = \sum_i \mathcal{X}_{i,j,k}\boldsymbol{M}_{i,\ell}$, $\ell \in \{1,\ldots,L\}$. Moreover, we have $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} \bullet_k \boldsymbol{M} \Leftrightarrow \boldsymbol{Y}^{(k)} = \boldsymbol{X}^{(k)}\boldsymbol{M}^\mathsf{T}$.*

Definition 2.3. ***Outer product*** *– The outer product between three vectors $\boldsymbol{a} \in \mathbb{R}^I$, $\boldsymbol{b} \in \mathbb{R}^J$, $\boldsymbol{c} \in \mathbb{R}^K$ is a rank-one tensor $\boldsymbol{\mathcal{X}} = \boldsymbol{a} \otimes \boldsymbol{b} \otimes \boldsymbol{c} \in \mathbb{R}^{I\times J\times K}$ whose elements are accessed as $\mathcal{X}_{i,j,k} = a_i b_j c_k$.*

2.2. Canonical polyadic decomposition

Definition 2.4. ***Canonical polyadic decomposition*** – *A third-order tensor admits a CPD as* $\boldsymbol{\mathcal{X}} = [\![\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}]\!]$*, where* $\boldsymbol{A} \in \mathbb{R}^{I \times R}$*,* $\boldsymbol{B} \in \mathbb{R}^{J \times R}$*,* $\boldsymbol{C} \in \mathbb{R}^{K \times R}$ *are the latent CP factors of the decomposition. When minimal, the integer* R *denotes the tensor rank of* $\boldsymbol{\mathcal{X}}$*. Each entry of* $\boldsymbol{\mathcal{X}}$ *can be expressed equivalently as*

$$\mathcal{X}_{i,j,k} = \sum_{r=1}^{R} A_{i,r} B_{j,r} C_{k,r}. \tag{1}$$

The CP factors $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ are essentially unique up to scaling and permutation ambiguities, if the rank R is not too large [32, 33]. Permutation ambiguity means that the columns of the latent CP factors can be reordered arbitrarily by any permutation matrix $\boldsymbol{\Pi} \in \mathbb{R}^{R \times R}$ as

$$\boldsymbol{\mathcal{X}} = [\![\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}]\!] = [\![\boldsymbol{A}\boldsymbol{\Pi}, \boldsymbol{B}\boldsymbol{\Pi}, \boldsymbol{C}\boldsymbol{\Pi}]\!].$$

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

$$\mathcal{X}_{i,j,k} = \sum_{r=1}^{R} (\alpha_n A_{i,r})(\beta_n B_{j,r})(\gamma_n C_{k,r}), \tag{2}$$

with $\alpha_r \beta_r \gamma_r = 1$ for $r \in \{1, \ldots, R\}$. When deriving Cramér-Rao bounds, permutation ambiguities can be neglected while a proper factor normalization is required to fix the scaling ambiguities. Throughout the paper, we correct this ambiguity by setting the first rows of the $\boldsymbol{A}$ and $\boldsymbol{B}$ factors to ones[4]. This corresponds to rescaling (2) with $\alpha_r = \frac{1}{A_{1,r}}$, $\beta_r = \frac{1}{B_{1,r}}$ and $\gamma_r = \frac{1}{\alpha_r \beta_r}$.

2.3. Uniqueness of the CPD

Definition 2.5. ***Kruskal rank*** – *The Kruskal rank of a matrix* $\boldsymbol{M}$*, denoted* $\kappa(\boldsymbol{M})$*, is defined as the maximum value* k *such that any* k *columns of* $\boldsymbol{M}$ *are linearly independent [34, 35].*

One of the most general and well-known sufficient conditions on uniqueness of the CPD is due to Kruskal [34, 36] and reads as follows:

$$\kappa(\boldsymbol{A}) + \kappa(\boldsymbol{B}) + \kappa(\boldsymbol{C}) \geq 2R + 2.$$

Stronger results are available for generic uniqueness. We say that the CPD $\boldsymbol{\mathcal{X}} = [\![\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}]\!]$ of rank R is generically unique if, for random matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ distributed according to an absolutely continuous probability distribution the CPD is unique. Equivalently, the set of $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ not leading to unique decomposition

[4] The entries in the first rows of $\boldsymbol{A}$ and $\boldsymbol{B}$ are null with probability zero, hence we can impose ones with probability one.

has measure zero. In this case, the Kruskal condition implies:

$$\min(I,R)+\min(J,R)+\min(K,R)\geq 2R+2. \tag{3}$$

In [37], another sufficient condition was provided:

$$R\leq 2^{\lfloor\log_2(J)\rfloor+\lfloor\log_2(K)\rfloor-2}. \tag{4}$$

However, it should be mentioned that (3) and (4) are only sufficient conditions ensuring generic uniqueness. The best known bounds guaranteeing generic uniqueness are given in [38, Theorem 1]. In particular, it is shown that generic uniqueness takes place for all R such $R<\lceil\frac{IJK}{I+J+K-2}\rceil$ (*i.e.* all ranks smaller than the generic rank) except few special cases and so-called unbalanced tensors, see [38] for more details.

2.4. Observation model for the reconstruction problem

We consider two tensors $\boldsymbol{\mathcal{Y}}_1\in\mathbb{R}^{I_d\times J_d\times K}$ and $\boldsymbol{\mathcal{Y}}_2\in\mathbb{R}^{I\times J\times K_d}$. As in [30], it is assumed[5] that $\boldsymbol{\mathcal{Y}}_1$ has high resolution in the third mode ($K>K_d$), while $\boldsymbol{\mathcal{Y}}_2$ possesses high resolutions in the first and second modes ($I>I_d$, $J>J_d$). Under the same acquisition conditions, $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ usually represent the same target, hence they are viewed as two degraded versions of a single tensor $\boldsymbol{\mathcal{X}}\in\mathbb{R}^{I\times J\times K}$, that is of high resolution in all three modes. We adopt the following degradation model that can be compactly written as

$$\begin{cases}\boldsymbol{\mathcal{Y}}_1 &= \boldsymbol{\mathcal{X}}\bullet_1\boldsymbol{P}\bullet_2\boldsymbol{Q}+\boldsymbol{\mathcal{E}}_1,\\ \boldsymbol{\mathcal{Y}}_2 &= \boldsymbol{\mathcal{X}}\bullet_3\boldsymbol{R}+\boldsymbol{\mathcal{E}}_2,\end{cases} \tag{5}$$

where $\boldsymbol{P}\in\mathbb{R}^{I_d\times I}$, $\boldsymbol{Q}\in\mathbb{R}^{J_d\times J}$, and $\boldsymbol{R}\in\mathbb{R}^{K_d\times K}$ have full row rank. We assume (for simplicity) that the degradation in the first and second modes is separable. The entries of the noise terms $\boldsymbol{\mathcal{E}}_1\sim\boldsymbol{\mathcal{N}}(0,\boldsymbol{\Sigma}_1)$, $\boldsymbol{\mathcal{E}}_2\sim\boldsymbol{\mathcal{N}}(0,\boldsymbol{\Sigma}_2)$ are independent and identically distributed (i.i.d.) real Gaussian tensors with zero mean and variances $\boldsymbol{\Sigma}_1=\sigma_1^2\boldsymbol{I}$ and $\boldsymbol{\Sigma}_2=\sigma_2^2\boldsymbol{I}$. Model (5) represents an ill-posed inverse problem, whose aim is to recover the tensor $\boldsymbol{\mathcal{X}}$ from the observations $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$.

Model (5) was used to tackle several reconstruction problems. For instance, in hyperspectral super-resolution [8], the matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ are blurring and downsampling matrices, while $\boldsymbol{R}$ contains the spectral response functions of the sensor used to acquiring $\boldsymbol{\mathcal{Y}}_2$. In medical imaging [10], the degradation matrices select sub-Nyquist samples (either fiber or slabs) of the target tensor in a given mode.

[5] The subscripts d in the dimensions stand for "degraded".

2.5. Reformulation as a coupled CP decomposition

Following [8], we assume that $\boldsymbol{\mathcal{X}}$ admits a CPD with rank R. The degradation model (5) becomes

$$\begin{cases} \boldsymbol{\mathcal{Y}}_1 &= [\![\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1]\!] + \boldsymbol{\mathcal{E}}_1, \\ \boldsymbol{\mathcal{Y}}_2 &= [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_2]\!] + \boldsymbol{\mathcal{E}}_2, \end{cases} \tag{6}$$

$$\text{where } \boldsymbol{A}_1 = \boldsymbol{P}\boldsymbol{A}_2, \boldsymbol{B}_1 = \boldsymbol{Q}\boldsymbol{B}_2, \boldsymbol{C}_2 = \boldsymbol{R}\boldsymbol{C}_1, \tag{7}$$

and $\boldsymbol{A}_1 \in \mathbb{R}^{I_d \times R}$, $\boldsymbol{B}_1 \in \mathbb{R}^{J_d \times R}$, $\boldsymbol{C}_1 \in \mathbb{R}^{K \times R}$, $\boldsymbol{A}_2 \in \mathbb{R}^{I \times R}$, $\boldsymbol{B}_2 \in \mathbb{R}^{J \times R}$, $\boldsymbol{C}_2 \in \mathbb{R}^{K_d \times R}$ are the factor matrices of the CPD. With this notation, the target tensor admits a CPD

$$\boldsymbol{\mathcal{X}} = [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_1]\!]. \tag{8}$$

While (6) only is an uncoupled model, the addition of the constraints in (7) make the model fully-coupled (*i.e.*, with couplings in the three modes of the tensors). In some applications, the degradation matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ are unknown, and we refer to this partially coupled case as blind. We define the blind CP model as follows:

$$\begin{cases} \boldsymbol{\mathcal{Y}}_1 &= [\![\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1]\!] + \boldsymbol{\mathcal{E}}_1, \\ \boldsymbol{\mathcal{Y}}_2 &= [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_2]\!] + \boldsymbol{\mathcal{E}}_2, \end{cases} \tag{9}$$

$$\text{where } \boldsymbol{C}_2 = \boldsymbol{R}\boldsymbol{C}_1. \tag{10}$$

2.6. Estimation

In the uncoupled case, estimation of the CP factors can be performed by applying the uncoupled ALS algorithm [39] to $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$. The identifiability of both CPDs is required. For instance, for $\boldsymbol{\mathcal{Y}}_1$, ALS minimizes the following cost function:

$$\min_{\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1} \frac{1}{\sigma_1^2} \|\boldsymbol{\mathcal{Y}}_1 - [\![\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1]\!]\|_F^2,$$

which corresponds to the Maximum Likelihood (ML) Estimator (MLE) for $\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1$. The fully-coupled problem (6)–(7) can be solved by the algorithm STEREO proposed in [8]. It is a coupled ALS algorithm minimizing the criterion

$$\min_{\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_1} \|\boldsymbol{\mathcal{Y}}_1 - [\![\boldsymbol{P}\boldsymbol{A}_2, \boldsymbol{Q}\boldsymbol{B}_2, \boldsymbol{C}_1]\!]\|_F^2 + \lambda \|\boldsymbol{\mathcal{Y}}_2 - [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{R}\boldsymbol{C}_1]\!]\|_F^2. \tag{11}$$

Assuming independent Gaussian noise and balance parameter $\lambda = \frac{\sigma_1^2}{\sigma_2^2}$, STEREO corresponds to the MLE for coupled $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$. In [8, Theorem 3], a sufficient condition for noiseless generic uniqueness of the CPD (8) recovered by STEREO was provided:

$$R \leq \min\left(2^{\lfloor \log_2(JK_d) \rfloor - 2}, I_d J_d\right). \tag{12}$$

In the proof of [8, Theorem 3], it is specified that identifiability of $\boldsymbol{\mathcal{Y}}_1$ (*i.e.*, generic uniqueness of its CPD) is not needed to establish uniqueness of the recovered target. The link between identifiability and uniqueness of the coupled CP model will be provided in Section 3.1.

In the partially coupled case (9)–(10), the matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ are unknown. In order to estimate the CP factors, we use Blind-STEREO, which is a coupled ALS algorithm that only accounts for the matrix $\boldsymbol{R}$. The criterion minimized by Blind-STEREO is

$$\min_{\substack{\boldsymbol{A}_1, \boldsymbol{B}_1 \\ \boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_1}} \|\boldsymbol{\mathcal{Y}}_1 - [\![\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1]\!]\|_F^2 + \lambda \|\boldsymbol{\mathcal{Y}}_2 - [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{R}\boldsymbol{C}_1]\!]\|_F^2, \tag{13}$$

which is the ML criterion for the partially-coupled problem if the balance parameter is $\lambda = \frac{\sigma_1^2}{\sigma_2^2}$. According to [8, Theorem 4], identifiability of $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ are required to ensure unique recovery of $\boldsymbol{\mathcal{X}}$ by Blind-STEREO in the noiseless case.

3. Cramér-Rao lower bounds for coupled models

3.1. Link between uniqueness and identifiability

First, we explain how uniqueness of the coupled CP model (6)–(7) in the noiseless case is related to the calculation of the CRB. In estimation theory, the notion of identifiability lacks a unified definition. In the literature, it is also called "observability" [40, 41]. In this paper, we propose to define it as the uniqueness of the proposed model.

Let us consider the probability density function (PDF) $\mathbf{f}_{\boldsymbol{\mathcal{Y}};\boldsymbol{\omega}}$ of the random real dataset $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^n$ parameterized by the unknown real deterministic parameter $\boldsymbol{\omega} \in \Omega \subseteq \mathbb{R}^m$. We assume that $\boldsymbol{\mathcal{Y}}$ is a random real Gaussian dataset parameterized by its mean, that is, $\boldsymbol{\mathcal{Y}} \sim \boldsymbol{\mathcal{N}}(\boldsymbol{\mu}(\boldsymbol{\omega}), \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma}$ a known, non-singular covariance matrix.

We say that the statistical model $\mathcal{F} = \{\mathbf{f}_{\boldsymbol{\mathcal{Y}};\boldsymbol{\omega}} : \boldsymbol{\omega} \in \Omega\}$ is identifiable if the mapping $\boldsymbol{\omega} \mapsto \mathbf{f}_{\boldsymbol{\mathcal{Y}};\boldsymbol{\omega}}$ is injective [42], *i.e.*, any distribution $\mathbf{f}_{\boldsymbol{\mathcal{Y}};\boldsymbol{\omega}}$ corresponds to a single parameter $\boldsymbol{\omega}$. For the case of our Gaussian dataset, the following holds true:

$$\mathbf{f}_{\boldsymbol{\mathcal{Y}};\boldsymbol{\omega}_1} = \mathbf{f}_{\boldsymbol{\mathcal{Y}};\boldsymbol{\omega}_2} \Leftrightarrow \boldsymbol{\mu}(\boldsymbol{\omega}_1) = \boldsymbol{\mu}(\boldsymbol{\omega}_2). \tag{14}$$

Thus, identifiability of the distributions is equivalent to identifiability of the means, *i.e.*, identifiability in the noiseless case.

Definition 3.1. ***Identifiability at a point*** – *The noiseless model* $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mu}(\boldsymbol{\omega})$ *is identifiable at the point* $\boldsymbol{\omega}_0$ *if*

$$(\boldsymbol{\omega} \neq \boldsymbol{\omega}_0) \Rightarrow (\boldsymbol{\mu}(\boldsymbol{\omega}) \neq \boldsymbol{\mu}(\boldsymbol{\omega}_0)) \quad \forall \boldsymbol{\omega} \in \mathbb{R}^m. \tag{15}$$

Definition 3.2. ***Local identifiability*** – *The noiseless model* $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mu}(\boldsymbol{\omega})$ *is locally identifiable at* $\boldsymbol{\omega}_0$ *if there exists an open subset* $\Omega_0 \subseteq \mathbb{R}^m$ *containing* $\boldsymbol{\omega}_0$ *such that*

$$(\boldsymbol{\omega} \neq \boldsymbol{\omega}_0) \Rightarrow (\boldsymbol{\mu}(\boldsymbol{\omega}) \neq \boldsymbol{\mu}(\boldsymbol{\omega}_0)) \quad \forall \boldsymbol{\omega} \in \Omega_0. \tag{16}$$

In our model, the Fisher information matrix (FIM) for $\boldsymbol{\omega}$ is obtained via the Slepian-Bangs formula [43]:

$$\mathbf{F}(\boldsymbol{\omega}) = \left[\tfrac{\partial \boldsymbol{\mu}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\mathsf{T}}\right]^\mathsf{T} \boldsymbol{\Sigma}^{-1} \left[\tfrac{\partial \boldsymbol{\mu}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\mathsf{T}}\right] \in \mathbb{R}^{m \times m}, \tag{17}$$

where $\frac{\partial \boldsymbol{\mu}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\mathsf{T}}$ is the Jacobian of $\boldsymbol{\mu}(\boldsymbol{\omega})$. If the FIM in (17) is non-singular, then $\boldsymbol{\mu}(\boldsymbol{\omega})$ is locally identifiable in the noiseless case [40, Theorem 5].

Conversely, a question that arises from the previous paragraph is whether local identifiability implies non-singularity of the FIM. For the case of tensor decompositions, the answer is positive. Let us consider that $\boldsymbol{\mathcal{Y}}$ is a vectorized tensor of subgeneric rank admitting a CPD as in (1), and that $\boldsymbol{\omega} = [\mathrm{vec}\{\boldsymbol{A}\}; \mathrm{vec}\{\boldsymbol{B}\}; \mathrm{vec}\{\boldsymbol{C}\}]$, $\boldsymbol{\mu}(\boldsymbol{\omega}) = \mathrm{vec}\{[\![\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}]\!]\}$. Generic uniqueness of the CPD of $\boldsymbol{\mathcal{Y}}$ implies that the rank of the Jacobian of $\boldsymbol{\mu}(\boldsymbol{\omega})$ in the generic case is equal to

$$\mathrm{rank}\left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\mathsf{T}}\right) = (I + J + K - 2)R$$

generically[6] (*i.e.*, except for a set of parameters $\boldsymbol{\omega}$ of measure zero), see [45, Sec. 3.2], [46], and [47, Def. 3.5]. Correcting the scaling ambiguities in (2) reduces the number of entries to estimate in $\boldsymbol{\omega}$, thus making the Jacobian full rank ; see Section 4.2 for more details on scaling ambiguities for coupled cases. Finally, from (17) it follows that full rank in the Jacobian implies that the FIM is invertible (and thus the CPD is locally identifiable in the noiseless case). In a nutshell, it means that correcting the scaling ambiguities allows for the FIM to be full rank. The link between uniqueness, identifiability and invertibility is summarized in Figure 1.

[6] This results is well-known for complex tensors, but it is also valid for real tensors, see [44].

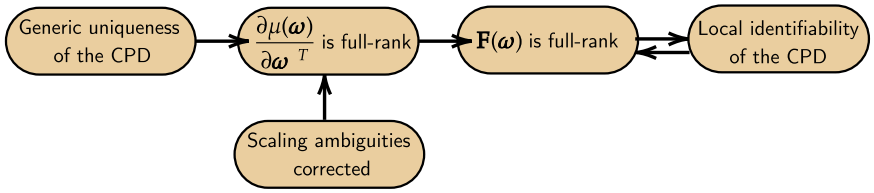

Figure 1: Link between generic uniqueness, local identifiability in the noiseless case and computing the bounds.

3.2. Coupled model with constraints

Let $\mathbf{f}_{\boldsymbol{\mathcal{Y}}_1;\boldsymbol{\omega}}$ and $\mathbf{f}_{\boldsymbol{\mathcal{Y}}_2;\boldsymbol{\omega}}$ be the PDFs of the random real datasets $\boldsymbol{\mathcal{Y}}_1 \in \mathbb{R}^{n_1}$ and $\boldsymbol{\mathcal{Y}}_2 \in \mathbb{R}^{n_2}$, parameterized by the unknown deterministic real parameter vector $\boldsymbol{\omega} \in \Omega$. A general coupled model with constraints is expressed as:

$$\begin{cases} \boldsymbol{\mathcal{Y}}_1 \sim \mathbf{f}_{\boldsymbol{\mathcal{Y}}_1;\boldsymbol{\omega}} \text{ and } \boldsymbol{\mathcal{Y}}_2 \sim \mathbf{f}_{\boldsymbol{\mathcal{Y}}_2;\boldsymbol{\omega}}, \\ \mathbf{g}(\boldsymbol{\omega}) = \mathbf{0}, \end{cases} \tag{18}$$

with $\mathbf{g}$ a non-redundant deterministic vector function differentiable $\forall \boldsymbol{\omega} \in \Omega$. Non-redundancy means that the system of equations $\mathbf{g}_i(\boldsymbol{\omega}) = \mathbf{0}$ is not reducible [28]. We assume that i) the PDFs $\mathbf{f}_{\boldsymbol{\mathcal{Y}}_1;\boldsymbol{\omega}}$ and $\mathbf{f}_{\boldsymbol{\mathcal{Y}}_2;\boldsymbol{\omega}}$ are non-redundant functions differentiable w.r.t. $\boldsymbol{\omega}$, and that their supports do not depend on $\boldsymbol{\omega}$; and that ii) the variables $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ are statistically independent.

In some cases, the model parameter $\boldsymbol{\omega} \in \Omega$ corresponds to the stacking of two parameters $\boldsymbol{\psi} \in \Psi \subseteq \mathbb{R}^{m_1}$ and $\boldsymbol{\xi} \in \Xi \subseteq \mathbb{R}^{m_2}$ $(m = m_1 + m_2)$ such that

$$\boldsymbol{\omega}^\mathsf{T} = \begin{bmatrix} \boldsymbol{\psi}^\mathsf{T} & \boldsymbol{\xi}^\mathsf{T} \end{bmatrix},$$

where $\boldsymbol{\xi}$ can be expressed as a function of $\boldsymbol{\psi}$, *i.e.*, $\boldsymbol{\xi} = \mathbf{h}(\boldsymbol{\psi})$. The function $\mathbf{h}$ is a non-redundant, differentiable function $\forall \boldsymbol{\psi} \in \Psi$. This results in the constraint

$$\mathbf{g}(\boldsymbol{\omega}) = \boldsymbol{\xi} - \mathbf{h}(\boldsymbol{\psi}) \in \mathbb{R}^{m_2}, \tag{19}$$

which can also be directly inserted in $\boldsymbol{\omega}$, leading to the following reparameterization

$$\boldsymbol{\omega}(\boldsymbol{\psi})^\mathsf{T} = \begin{bmatrix} \boldsymbol{\psi}^\mathsf{T} & \mathbf{h}(\boldsymbol{\psi})^\mathsf{T} \end{bmatrix}. \tag{20}$$

The model (18) can thus be reformulated as the following unconstrained coupled model

$$\begin{cases} \boldsymbol{\mathcal{Y}}_1 \sim \mathbf{f}_{\boldsymbol{\mathcal{Y}}_1;\boldsymbol{\psi}} \text{ and } \boldsymbol{\mathcal{Y}}_2 \sim \mathbf{f}_{\boldsymbol{\mathcal{Y}}_2;\boldsymbol{\psi}}. \end{cases} \tag{21}$$

Here, the PDFs are solely parameterized by the unknown deterministic real parameter vector $\boldsymbol{\psi} \in \Psi$, under the same assumptions (i) and (ii) on the PDFs as in model (18).

3.3. Uncoupled CRB

We consider that $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ are random real Gaussian distributed datasets parameterized by their mean, *i.e.*, $\boldsymbol{\mathcal{Y}}_1 \sim \boldsymbol{\mathcal{N}}(\boldsymbol{\mu}_1(\boldsymbol{\omega}), \boldsymbol{\Sigma}_1)$ and $\boldsymbol{\mathcal{Y}}_2 \sim \boldsymbol{\mathcal{N}}(\boldsymbol{\mu}_2(\boldsymbol{\omega}), \boldsymbol{\Sigma}_2)$ where $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are known covariance matrices. The parameter $\boldsymbol{\omega}$ is unknown real and assumed to be deterministic. The uncoupled FIM for $\boldsymbol{\omega}$ is obtained by using the Slepian-Bangs formula [43]:

$$\boldsymbol{F}(\boldsymbol{\omega}) = \begin{bmatrix} \frac{\partial \boldsymbol{\mu}_1(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\mathsf{T}} \\ \frac{\partial \boldsymbol{\mu}_2(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\mathsf{T}} \end{bmatrix}^\mathsf{T} \mathrm{Diag}\{\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2\}^{-1} \begin{bmatrix} \frac{\partial \boldsymbol{\mu}_1(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\mathsf{T}} \\ \frac{\partial \boldsymbol{\mu}_2(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\mathsf{T}} \end{bmatrix}. \tag{22}$$

If the FIM is non-singular, then the uncoupled CRB for $\boldsymbol{\omega}$ (namely $\mathbf{CRB}(\boldsymbol{\omega})$) is obtained as $\mathbf{CRB}(\boldsymbol{\omega}) = \boldsymbol{F}^{-1}(\boldsymbol{\omega})$. From Section 3.1, we see that invertibility[7] of the FIM implies local identifiability of the whole parameter $\boldsymbol{\omega}$. For uncoupled estimation, the constraint $\mathbf{g}(\boldsymbol{\omega}) = \mathbf{0}$ is ignored.

3.4. Expression for CCRB

Numerous works have addressed performance bounds on $\boldsymbol{\omega}$ under the constraint $\mathbf{g}(\boldsymbol{\omega}) = \mathbf{0}$, leading to the definition of the constrained FIM and the CCRB. In the seminal paper [29], the CCRB for $\boldsymbol{\omega}$ is expressed as

$$\mathbf{CCRB}(\boldsymbol{\omega}) = \boldsymbol{F}^{-1} - \boldsymbol{F}^{-1}\boldsymbol{G}^\mathsf{T}\left[\boldsymbol{G}\boldsymbol{F}^{-1}\boldsymbol{G}^\mathsf{T}\right]^{-1}\boldsymbol{G}\boldsymbol{F}^{-1} \succeq \mathbf{0}, \tag{23}$$

where $\boldsymbol{F} \overset{\mathsf{def}}{=} \boldsymbol{F}(\boldsymbol{\omega})$ and $\boldsymbol{G} = \left[\frac{\partial \mathbf{g}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\mathsf{T}}\right] \in \mathbb{R}^{m_2 \times m}$ is a full row-rank matrix, which is equivalent to requiring that the constraints are non-redundant. It is easy to see from (23) that the CCRB is lower than the CRB. However, this formulation explicitly requires the FIM to be non-singular, and inversion of the FIM can be costly.

In [22, 20], an alternative expression for the CCRB is

$$\mathbf{CCRB}(\boldsymbol{\omega}) = \boldsymbol{U}\left[\boldsymbol{U}^\mathsf{T}\boldsymbol{F}\boldsymbol{U}\right]^{-1}\boldsymbol{U}^\mathsf{T}, \tag{24}$$

where $\boldsymbol{U} \overset{\mathsf{def}}{=} \boldsymbol{U}(\boldsymbol{\omega}) \in \mathbb{R}^{m \times m_1}$ is a basis of $\ker(\boldsymbol{G})$. The matrix $\boldsymbol{U}^\mathsf{T}\boldsymbol{F}\boldsymbol{U}$ is called the constrained FIM.

Contrary to (23), (24) does not require invertibility of $\boldsymbol{F}$. The above expression does not depend on the choice of $\boldsymbol{U}$ either [22]. It is also noticeable that if $\boldsymbol{F}$ is invertible, then the expressions in Equation (23) and Equation (24) are equivalent [20, Corollary 1]. Here, we choose to compute (24) when we mention the CCRB.

[7] In some cases, however, the FIM can be singular (and thus, non-invertible): common practice is to resort to the Moore-Penrose pseudo-inverse of the FIM for the computation of the CRB [16, 19]. In such cases, any estimator of $\boldsymbol{\omega}$ must have infinite variance [19]: in this paper, we choose not to compute the CRB when the FIM is singular.

3.5. Reparameterized CRB

Let us now consider a reparameterization of the PDFs $\mathbf{f}_{\boldsymbol{\mathcal{Y}}_{1;\boldsymbol{\omega}}}$ and $\mathbf{f}_{\boldsymbol{\mathcal{Y}}_{2;\boldsymbol{\omega}}}$ for the unknown parameter $\boldsymbol{\psi} \in \Psi \subseteq \mathbb{R}^{m_1}$ where $\boldsymbol{\omega} = \boldsymbol{\omega}(\boldsymbol{\psi})$. We consider the particular case where $\boldsymbol{\psi}$ is a subset of parameters in $\boldsymbol{\omega}$; then, arbitrarily we can rearrange the components of $\boldsymbol{\omega}$ as in (20).

In [42, p.125], an expression for the reparameterized FIM for $\boldsymbol{\psi}$ (namely $\boldsymbol{F_c}(\boldsymbol{\psi})$) is given:

$$\boldsymbol{F_c}(\boldsymbol{\psi}) = \left[\frac{\partial \boldsymbol{\omega}(\boldsymbol{\psi})}{\partial \boldsymbol{\psi}^\mathsf{T}}\right]^\mathsf{T} \boldsymbol{F}(\boldsymbol{\omega}(\boldsymbol{\psi})) \left[\frac{\partial \boldsymbol{\omega}(\boldsymbol{\psi})}{\partial \boldsymbol{\psi}^\mathsf{T}}\right]. \tag{25}$$

Contrary to the uncoupled case, we notice that uniqueness (20) only requires identifiability of the sub-parameter $\boldsymbol{\psi}$. Additionally, we can express the reparameterized CRB for the parameter $\boldsymbol{\xi}$ (namely $\mathbf{rCRB}(\boldsymbol{\xi})$) as

$$\mathbf{rCRB}(\boldsymbol{\xi}) = \left[\frac{\partial \mathbf{h}(\boldsymbol{\psi})}{\partial \boldsymbol{\psi}^\mathsf{T}}\right]^\mathsf{T} \boldsymbol{F_c}^{-1}(\boldsymbol{\psi}) \left[\frac{\partial \mathbf{h}(\boldsymbol{\psi})}{\partial \boldsymbol{\psi}^\mathsf{T}}\right].$$

In [22], it is shown that for the parameter $\boldsymbol{\psi}$, (24) and the constrained FIM in (25) lead to the same bound[8].

4. Different parameterizations and estimation scenarios

To derive appropriate performance bounds, it is necessary to embed the problem in an appropriate probabilistic framework requiring to properly define the probabilistic model, the parameters of interest and possible associated constraints, and to fix the ambiguities resulting from the coupled CP model.

4.1. Model parameters

Following the notations of [30], we separate the CP factors into distinct parameters $\boldsymbol{\theta}_1 \in \mathbb{R}^{KR}$, $\boldsymbol{\theta}_2 \in \mathbb{R}^{K_d R}$, $\boldsymbol{\phi}_1 \in \mathbb{R}^{(I_d+J_d)R}$ and $\boldsymbol{\phi}_2 \in \mathbb{R}^{(I+J)R}$ as

$$\boldsymbol{\theta}_1^\mathsf{T} = \mathrm{vec}\{\boldsymbol{C}_1\}^\mathsf{T}, \quad \boldsymbol{\phi}_1^\mathsf{T} = \begin{bmatrix}\mathrm{vec}\{\boldsymbol{A}_1\}^\mathsf{T} & \mathrm{vec}\{\boldsymbol{B}_1\}^\mathsf{T}\end{bmatrix}, \quad \boldsymbol{\theta}_2^\mathsf{T} = \mathrm{vec}\{\boldsymbol{C}_2\}^\mathsf{T}, \quad \boldsymbol{\phi}_2^\mathsf{T} = \begin{bmatrix}\mathrm{vec}\{\boldsymbol{A}_2\}^\mathsf{T} & \mathrm{vec}\{\boldsymbol{B}_2\}^\mathsf{T}\end{bmatrix}. \tag{26}$$

These vectors are stacked into one global parameter $\boldsymbol{\omega} \in \mathbb{R}^m$ ($m = (I + J + K + I_d + J_d + K_d)R$) defined by

$$\boldsymbol{\omega}^\mathsf{T} = \begin{bmatrix}\boldsymbol{\phi}_1^\mathsf{T} & \boldsymbol{\theta}_1^\mathsf{T} & \boldsymbol{\phi}_2^\mathsf{T} & \boldsymbol{\theta}_2^\mathsf{T}\end{bmatrix}^\mathsf{T}.$$

From (7), we can see that the model parameters can be linked through non-redundant functions as $\mathbf{g_1}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \mathbf{0}$ and $\mathbf{g_2}(\boldsymbol{\phi}_1, \boldsymbol{\phi}_2) = \mathbf{0}$, where $\mathbf{g_1}$ and $\mathbf{g_2}$ are differentiable $\forall(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ (resp. $(\boldsymbol{\phi}_1, \boldsymbol{\phi}_2)$).

[8] As a result, invertibility of the constrained FIM in (24) also implies that $\boldsymbol{\psi}$ is identifiable.

4.2. General framework for the fusion problem

For the fully-coupled model (6)–(7), we wish to estimate the parameters $\boldsymbol{\phi}_2$ and $\boldsymbol{\theta}_1$, *i.e.*, the factor matrices of $\boldsymbol{\mathcal{X}}$. In order to illustrate the advantage of data fusion over uncoupled estimation, we are also interested in the performance of the uncoupled and partially-coupled models: these cases require the calculation of performance bounds for the parameters $\boldsymbol{\phi}_1$ and $\boldsymbol{\theta}_2$ as well.

As a result, we can distinguish two probabilistic scenarios, depending on whether i) we are only interested in performance bounds and an analysis for the fully-coupled problem, or ii) we want to compare the performance of the coupled CP approach to that of the uncoupled and partially-coupled approaches. Case i) allows for a bound calculation for the fully-coupled problem only and will be referred to as *scenario 1*, while ii) can encompass uncoupled and partially-coupled problems and will be referred to as *scenario 2*.

We consider that the observations $\boldsymbol{\mathcal{Y}}_1 \in \mathbb{R}^{I_d \times J_d \times K}$ and $\boldsymbol{\mathcal{Y}}_2 \in \mathbb{R}^{I \times J \times K_d}$ are random real Gaussian datasets. For all models and scenarios, $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ are distributed as in (18). Here, from the relationships between the model parameters, we can express (19) as

$$\mathbf{g}(\boldsymbol{\omega}) = \begin{bmatrix} \mathbf{g_1}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) & \mathbf{g_2}(\boldsymbol{\phi}_1, \boldsymbol{\phi}_2) \end{bmatrix}. \tag{27}$$

For each scenario and model (uncoupled, partially-coupled or fully-coupled), the expression of $\mathbf{g_1}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ and $\mathbf{g_2}(\boldsymbol{\phi}_1, \boldsymbol{\phi}_2)$ might change, resulting in different sets of constraints between the parameters. As in Section 3, the PDFs might only be parameterized by a subset of $\boldsymbol{\omega}$; in the following subsections, the expressions of these PDFs will be provided for each scenario.

Calculation of CRBs often requires inversion of a FIM, as explained in Section 3. For the FIM to be full rank, scaling ambiguities in the CPDs need to be solved [24] regarding the parameters we wish to estimate: indeed the manifold of rank-R tensors in $\mathbb{R}^{I \times J \times K}$ has dimension $(I + J + K - 2)R$. For each aforementioned scenario, we will provide different options for solving the scaling ambiguities, ensuring that the considered FIM is invertible. They will result in different sets of constraints, allowing for the calculation of the performance bounds. We will also introduce different parameterizations and distributions for the observed datasets.

4.3. Scenario 1 – Performance bounds for fully-coupled CP model

In this first scenario, we are only interested in the performance analysis for the fully-coupled CP model. This case boils down to a performance analysis for $\boldsymbol{\phi}_2$ and $\boldsymbol{\theta}_1$ only. Thus, in this scenario, we only need uniqueness of the CPD of the target tensor $\boldsymbol{\mathcal{X}}$ to calculate the bounds. As discussed in Section 2, we set $(\boldsymbol{A}_2)_{1,:} = (\boldsymbol{B}_2)_{1,:} = 1$ to fix the correct the scaling ambiguities in $\boldsymbol{\phi}_2$ (note that generically, this does not restrict the generality since entries have a probability 1 to be nonzero). This scaling option guarantees that

the FIM w.r.t. $\boldsymbol{\phi}_2$ and $\boldsymbol{\theta}_1$ is full-rank. As a result, we define the parameter $\widetilde{\boldsymbol{\phi}}_2 \in \mathbb{R}^{(I+J-2)R}$ as

$$\widetilde{\boldsymbol{\phi}}_2^{\mathsf{T}} = \begin{bmatrix} \text{vec}\{(\boldsymbol{A}_2)_{2:I,:}\}^{\mathsf{T}} & \text{vec}\{(\boldsymbol{B}_2)_{2:J,:}\}^{\mathsf{T}} \end{bmatrix}.$$

that is only composed of the unknown entries of $\boldsymbol{\phi}_2$. The full and reduced parameters can be linked through the relationship $\widetilde{\boldsymbol{\phi}}_2 = \boldsymbol{M}_2\boldsymbol{\phi}_2$. The matrix $\boldsymbol{M}_2 \in \mathbb{R}^{(I+J-2)R\times(I+J)R}$ is constructed from $\boldsymbol{I}_{(I+J)R}$ by removing the $2R$ rows corresponding to known entries of $\boldsymbol{\phi}_2$. We can recast the fully-coupled CP model in

$$\begin{cases} \boldsymbol{\mathcal{Y}}_1 &= [\![\boldsymbol{P}\boldsymbol{A}_2, \boldsymbol{Q}\boldsymbol{B}_2, \boldsymbol{C}_1]\!] + \boldsymbol{\mathcal{E}}_1, \\ \boldsymbol{\mathcal{Y}}_2 &= [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{R}\boldsymbol{C}_1]\!] + \boldsymbol{\mathcal{E}}_2, \end{cases} \tag{28}$$

that directly includes the constraints between the factor matrices. Since the entries of the noise terms $\boldsymbol{\mathcal{E}}_1$ and $\boldsymbol{\mathcal{E}}_2$ are i.i.d., $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ are distributed according to

$$\begin{cases} \mathbf{f}_{\boldsymbol{\mathcal{Y}}_1;\widetilde{\boldsymbol{\phi}}_2,\theta_1} &= \left(2\pi\sigma_1^2\right)^{\frac{-I_dJ_dK}{2}} \exp\left(-\frac{1}{2\sigma_1^2}\|\boldsymbol{\mathcal{Y}}_1 - [\![\boldsymbol{P}\boldsymbol{A}_2, \boldsymbol{Q}\boldsymbol{B}_2, \boldsymbol{C}_1]\!]\|_F^2\right), \\ \mathbf{f}_{\boldsymbol{\mathcal{Y}}_2;\widetilde{\boldsymbol{\phi}}_2,\theta_1} &= \left(2\pi\sigma_2^2\right)^{\frac{-IJK_d}{2}} \exp\left(-\frac{1}{2\sigma_2^2}\|\boldsymbol{\mathcal{Y}}_2 - [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{R}\boldsymbol{C}_1]\!]\|_F^2\right), \end{cases} \tag{29}$$

In model (28), the constraints between the factor matrices are such that $\boldsymbol{A}_1 = \boldsymbol{P}\boldsymbol{A}_2$, $\boldsymbol{B}_1 = \boldsymbol{Q}\boldsymbol{B}_2$ and $\boldsymbol{C}_2 = \boldsymbol{R}\boldsymbol{C}_1$. These equalities translate in terms of model parameters as

$$\begin{cases} \mathbf{g_1}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) &= \boldsymbol{\theta}_2 - (\boldsymbol{I}_R \boxtimes \boldsymbol{R})\,\boldsymbol{\theta}_1, \\ \mathbf{g_2}(\phi_1, \widetilde{\boldsymbol{\phi}}_2) &= \phi_1 - \begin{bmatrix} \boldsymbol{I}_R \boxtimes \boldsymbol{P} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_R \boxtimes \boldsymbol{Q} \end{bmatrix} \boldsymbol{M}_2^{\mathsf{T}}\widetilde{\boldsymbol{\phi}}_2. \end{cases} \tag{30}$$

From (30), we can see that the functions $\mathbf{g_1}$ and $\mathbf{g_2}$ are linear and thus, in this scenario, we will refer to the relationship between the model parameters as *linear constraints*.

4.4. Scenario 2 – Comparing performance bounds

Specific scaling option. In this second scenario, we want to compare performance bounds for the fully coupled problem to those in the uncoupled and partially-coupled case. This case requires the calculation of the bounds for the parameters $\boldsymbol{\phi}_2$ and $\boldsymbol{\theta}_1$, as well as for $\boldsymbol{\phi}_1$ and $\boldsymbol{\theta}_2$ for partially-coupled and uncoupled models. Contrary to scenario 1, inversion of the FIM in the partially-coupled and uncoupled case require both CPDs to be

generically unique. As a result, we also define the reduced parameter $\widetilde{\boldsymbol{\omega}} \in \mathbb{R}^{m-4R}$ as

$$\widetilde{\boldsymbol{\omega}} = \begin{bmatrix} \widetilde{\boldsymbol{\phi}}_1^{\mathsf{T}} & \boldsymbol{\theta}_1^{\mathsf{T}} & \widetilde{\boldsymbol{\phi}}_2^{\mathsf{T}} & \boldsymbol{\theta}_2^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}}.$$

We solve scaling ambiguities in $\boldsymbol{\phi}_1$ by setting the first rows of $\boldsymbol{A}_1$ and $\boldsymbol{B}_1$ to ones. Thus, we also define the reduced parameter vector $\widetilde{\boldsymbol{\phi}}_1 \in \mathbb{R}^{(I_d+J_d-2)R}$ as

$$\widetilde{\boldsymbol{\phi}}_1^{\mathsf{T}} = \begin{bmatrix} \text{vec}\{(\boldsymbol{A}_1)_{2:I_d,:}\}^{\mathsf{T}} & \text{vec}\{(\boldsymbol{B}_1)_{2:I_d,:}\}^{\mathsf{T}} \end{bmatrix}$$

that is only composed of the unknown entries of $\boldsymbol{\phi}_1$. As in the previous subsection, we can express the reduced parameter vector through the relationship $\widetilde{\boldsymbol{\phi}}_1 = \boldsymbol{M}_1 \boldsymbol{\phi}_1$, with $\boldsymbol{M}_1 \in \mathbb{R}^{(I_d+J_d-2)R \times (I_d+J_d)R}$ constructed as $\boldsymbol{M}_2$. Given (7), solving the scaling ambiguities for the coupled CP factors of $\boldsymbol{\mathcal{Y}}_1$ imposes that $(\boldsymbol{P}\boldsymbol{A}_2)_{1,:} = (\boldsymbol{Q}\boldsymbol{B}_2)_{1,:} = 1$. However, it is unlikely that the degradation matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ make the above equality valid, even if $(\boldsymbol{A}_2)_{1,:} = (\boldsymbol{B}_2)_{1,:} = 1$, as it would require that $(\boldsymbol{P})_{1,:} = \begin{bmatrix} 1 & \boldsymbol{0}_{1\times(I-1)} \end{bmatrix}$ and $(\boldsymbol{Q})_{1,:} = \begin{bmatrix} 1 & \boldsymbol{0}_{1\times(J-1)} \end{bmatrix}$. The performance analysis for this case was addressed in [30]. Here, to circumvent this limitation and address a more general case, we introduce the scaling diagonal factors

$$\boldsymbol{D}_\alpha = \text{diag}\{(\boldsymbol{P}\boldsymbol{A}_2)_{1,:}\} \text{ and } \boldsymbol{D}_\beta = \text{diag}\{(\boldsymbol{Q}\boldsymbol{B}_2)_{1,:}\} \tag{31}$$

such that $(\boldsymbol{A}_1 \cdot \boldsymbol{D}_\alpha^{-1})_{1,:} = (\boldsymbol{B}_1 \cdot \boldsymbol{D}_\beta^{-1})_{1,:} = 1$. This specific option (31) allows for the scaling ambiguities to be corrected, even after degradation by the general matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$. We also need to rescale $\boldsymbol{C}_2$ as $\boldsymbol{C}_2 = \boldsymbol{R}\boldsymbol{C}_1 \cdot (\boldsymbol{D}_\alpha \boldsymbol{D}_\beta)^{-1}$ so that $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ are degraded versions of the same tensor

$$\boldsymbol{\mathcal{X}} = [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_1(\boldsymbol{D}_\alpha \boldsymbol{D}_\beta)^{-1}]\!].$$

Model and parameterization for the fully-coupled model. The particular scaling option (31) leads to the following model with additive constraints between the CP factors:

$$\begin{cases} \boldsymbol{\mathcal{Y}}_1 &= [\![\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1]\!] + \boldsymbol{\mathcal{E}}_1, \\ \boldsymbol{\mathcal{Y}}_2 &= [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_2]\!] + \boldsymbol{\mathcal{E}}_2, \end{cases} \tag{32}$$
$$\text{subject to } \boldsymbol{A}_1 = \boldsymbol{P}\boldsymbol{A}_2 \cdot \boldsymbol{D}_\alpha^{-1}, \boldsymbol{B}_1 = \boldsymbol{Q}\boldsymbol{B}_2 \cdot \boldsymbol{D}_\beta^{-1}, \boldsymbol{C}_2 = \boldsymbol{R}\boldsymbol{C}_1 \cdot (\boldsymbol{D}_\alpha \boldsymbol{D}_\beta)^{-1}$$

for the fully coupled case. The datasets are thus distributed according to

$$\begin{cases} \mathbf{f}_{\boldsymbol{\mathcal{Y}}_1;\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_1} = \left(2\pi\sigma_1^2\right)^{\frac{-I_d J_d K}{2}} \exp\left(-\frac{1}{2\sigma_1^2}\|\boldsymbol{\mathcal{Y}}_1 - [\![\boldsymbol{P}\boldsymbol{A}_2 \cdot \boldsymbol{D}_\alpha^{-1}, \boldsymbol{Q}\boldsymbol{B}_2 \cdot \boldsymbol{D}_\beta^{-1}, \boldsymbol{C}_1]\!]\|_F^2\right), \\ \mathbf{f}_{\boldsymbol{\mathcal{Y}}_2;\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_1} = \left(2\pi\sigma_2^2\right)^{\frac{-IJK_d}{2}} \exp\left(-\frac{1}{2\sigma_2^2}\|\boldsymbol{\mathcal{Y}}_2 - [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{R}\boldsymbol{C}_1 \cdot (\boldsymbol{D}_\alpha\boldsymbol{D}_\beta)^{-1}]\!]\|_F^2\right), \end{cases} \tag{33}$$

which is a parameterization different from (29). The only case where the PDFs in (29) and (33) are equivalent is the specific case where $\boldsymbol{D}_\alpha = \boldsymbol{D}_\beta = \boldsymbol{I}_R$, addressed in [30].

In (32), we can see that the relationships linking the CP factors involve the scaling factors $\boldsymbol{D}_\alpha$ and $\boldsymbol{D}_\beta$. Rewriting these relationships in terms of the model parameters gives:

$$\begin{cases} \mathbf{g_1}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) & = \boldsymbol{\theta}_2 - \left((\boldsymbol{D}_\alpha\boldsymbol{D}_\beta)^{-1} \boxtimes \boldsymbol{R}\right)\boldsymbol{\theta}_1, \\ \mathbf{g_2}(\widetilde{\boldsymbol{\phi}}_1, \widetilde{\boldsymbol{\phi}}_2) & = \widetilde{\boldsymbol{\phi}}_1 - \boldsymbol{M}_1 \begin{bmatrix} \boldsymbol{D}_\alpha^{-1} \boxtimes \boldsymbol{P} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{D}_\beta^{-1} \boxtimes \boldsymbol{Q} \end{bmatrix} \boldsymbol{M}_2^{\mathsf{T}} \widetilde{\boldsymbol{\phi}}_2. \end{cases} \tag{34}$$

Due to the definition of $\boldsymbol{D}_\alpha$ and $\boldsymbol{D}_\beta$ in (31), (34) are *non-linear constraints* on the model parameters.

Parameterizations for uncoupled and partially-coupled models. In the uncoupled case, the datasets are distributed according to

$$\begin{cases} \mathbf{f}_{\boldsymbol{\mathcal{Y}}_1;\widetilde{\boldsymbol{\phi}}_1,\theta_1} & = \left(2\pi\sigma_1^2\right)^{\frac{-I_d J_d K}{2}} \exp\left(-\frac{1}{2\sigma_1^2}\|\boldsymbol{\mathcal{Y}}_1 - [\![\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1]\!]\|_F^2\right), \\ \mathbf{f}_{\boldsymbol{\mathcal{Y}}_2;\widetilde{\boldsymbol{\phi}}_2,\theta_2} & = \left(2\pi\sigma_2^2\right)^{\frac{-IJK_d}{2}} \exp\left(-\frac{1}{2\sigma_2^2}\|\boldsymbol{\mathcal{Y}}_2 - [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_2]\!]\|_F^2\right), \end{cases} \tag{35}$$

and follow model (6). For the partially-coupled problem, we have the following model:

$$\begin{cases} \boldsymbol{\mathcal{Y}}_1 & = [\![\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1]\!] + \boldsymbol{\mathcal{E}}_1, \\ \boldsymbol{\mathcal{Y}}_2 & = [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_2]\!] + \boldsymbol{\mathcal{E}}_2, \end{cases} \tag{36}$$
$$\text{subject to } \boldsymbol{C}_2 = \boldsymbol{R}\boldsymbol{C}_1 \cdot (\boldsymbol{D}_\alpha\boldsymbol{D}_\beta)^{-1}.$$

The datasets are distributed according to

$$\begin{cases} \mathbf{f}_{\boldsymbol{\mathcal{Y}}_1;\widetilde{\boldsymbol{\phi}}_1,\boldsymbol{\theta}_1} = \left(2\pi\sigma_1^2\right)^{\frac{-I_d J_d K}{2}} \exp\left(-\frac{1}{2\sigma_1^2}\|\boldsymbol{\mathcal{Y}}_1 - [\![\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1 \cdot (\boldsymbol{D}_\alpha\boldsymbol{D}_\beta)^{-1}]\!]\|_F^2\right), \\ \mathbf{f}_{\boldsymbol{\mathcal{Y}}_2;\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_1} = \left(2\pi\sigma_2^2\right)^{\frac{-IJK_d}{2}} \exp\left(-\frac{1}{2\sigma_2^2}\|\boldsymbol{\mathcal{Y}}_2 - [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{R}\boldsymbol{C}_1 \cdot (\boldsymbol{D}_\alpha\boldsymbol{D}_\beta)^{-1}]\!]\|_F^2\right). \end{cases} \tag{37}$$

Here, we only consider the constraint $\mathbf{g_1}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \boldsymbol{\theta}_2 - \left((\boldsymbol{D}_\alpha\boldsymbol{D}_\beta)^{-1} \boxtimes \boldsymbol{R}\right)\boldsymbol{\theta}_1$ instead of (34).

4.5. Performance for target tensor approximation

Additionally to the model parameters in (26), we also define $\boldsymbol{x} = \text{vec}\{\boldsymbol{\mathcal{X}}\} \in \mathbb{R}^{\ell}$ ($\ell = IJK$), that represents the vectorized low-rank approximation of $\boldsymbol{\mathcal{X}}$. Parameter $\boldsymbol{x}$ can be linked to the model parameters through the relationship $\mathbf{g_3}(\boldsymbol{x}, \widetilde{\boldsymbol{\psi}}) = \mathbf{0}$. In order to get the bounds for $\boldsymbol{x}$, we utilize relationships between tensor unfoldings

$$\boldsymbol{x} = \underbrace{[(\boldsymbol{C}_1 \odot \boldsymbol{B}_2) \boxtimes \boldsymbol{I}_I]}_{\boldsymbol{S}_1} \text{vec}\{\boldsymbol{A}_2\} = \underbrace{\boldsymbol{\Pi}^{(2,1)} [(\boldsymbol{C}_1 \odot \boldsymbol{A}_2) \boxtimes \boldsymbol{I}_J]}_{\boldsymbol{S_2}} \text{vec}\{\boldsymbol{B}_2\} = \underbrace{\boldsymbol{\Pi}^{(3,1)} [(\boldsymbol{B}_2 \odot \boldsymbol{A}_2) \boxtimes \boldsymbol{I}_K]}_{\boldsymbol{S}_3} \text{vec}\{\boldsymbol{C}_1\}, \quad (38)$$

where $\boldsymbol{\Pi}^{(2,1)}$ and $\boldsymbol{\Pi}^{(3,1)}$ are permutation matrices that link the entries of $\text{vec}\{\boldsymbol{X}^{(2)}\}$ (resp. $\text{vec}\{\boldsymbol{X}^{(3)}\}$) to those of $\text{vec}\{\boldsymbol{X}^{(1)}\}$. As a result, the expression of $\mathbf{g_3}(\boldsymbol{x}, \widetilde{\boldsymbol{\psi}})$ is given by

$$\mathbf{g_3}(\boldsymbol{x}, \widetilde{\boldsymbol{\psi}}) = \boldsymbol{x} - \begin{bmatrix} \boldsymbol{S}_1 & \boldsymbol{S}_2 & \boldsymbol{S}_3 \end{bmatrix} \boldsymbol{M}_3^{\mathsf{T}} \widetilde{\boldsymbol{\psi}},$$

where $\boldsymbol{M}_3 = \text{Diag}\{\boldsymbol{M}_2, \boldsymbol{I}_{KR}\}$.

5. Calculation of performance bounds

In this section, we derive performance bounds in the uncoupled, partially-coupled, and fully coupled cases. For the case of fully coupled datasets (*i.e.*, all degradation matrices are known), we address both scenarios described above. The closed-form of the matrices to be inverted and their submatrices are all available in a supplementary material. The proposed framework for computing the bounds, depending on the estimation scenario and parameter constraints, is summarized below in Figure 2.

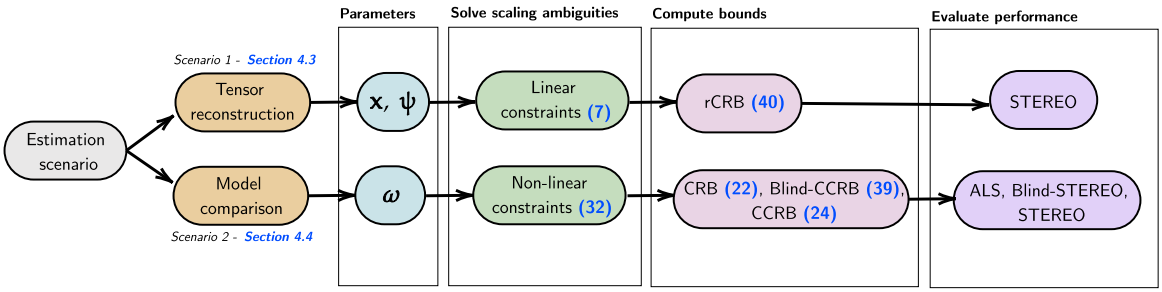

Figure 2: Pipeline of the proposed framework for CCRB derivation.

5.1. Uncoupled CRB

In the uncoupled case, the CRB for the parameter $\widetilde{\boldsymbol{\omega}}$ is obtained by inverting the uncoupled FIM. To do so, scaling ambiguities in the CPDs of $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ need to be solved so that the FIM is full rank. In practice, the FIM for $\widetilde{\boldsymbol{\omega}}$ (namely $\boldsymbol{F}(\widetilde{\boldsymbol{\omega}})$) is computed by applying (22) to the tensors $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$. The expressions of $\boldsymbol{\mu}_1(\widetilde{\boldsymbol{\omega}})$

and $\boldsymbol{\mu}_2(\widetilde{\boldsymbol{\omega}})$ are obtained from relationships between tensor unfoldings; please see [28] for a full derivation and closed-form expression of the FIM.

As in previous related works [26, 28, 30], we consider a case where the scaling ambiguities on $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ are solved, meaning that the FIM is non-singular. Thus, the CRB for $\widetilde{\boldsymbol{\omega}}$ can be obtained by inversion of the FIM: $\mathbf{CRB}(\widetilde{\boldsymbol{\omega}}) = \boldsymbol{F}^{-1}(\widetilde{\boldsymbol{\omega}})$. The CRB for each sub-parameter can be obtained by applying the block inversion lemma [48] to $\boldsymbol{F}(\widetilde{\boldsymbol{\omega}})$. Please note that the uncoupled CRB can only be computed in scenario 2.

5.2. Blind-CCRB

We now compute the CCRB associated with the blind (partially-coupled) model (36). The Blind-CCRB can only be computed in scenario 2 due to the scaling ambiguities on $\boldsymbol{A}_1, \boldsymbol{B}_1$. The Jacobian of constraints is

$$\boldsymbol{G} = \begin{bmatrix} \frac{\partial \mathbf{g_1}}{\partial \widetilde{\boldsymbol{\phi}}_1^{\mathsf{T}}} & \frac{\partial \mathbf{g_1}}{\partial \boldsymbol{\theta}_1^{\mathsf{T}}} & \frac{\partial \mathbf{g_1}}{\partial \widetilde{\boldsymbol{\phi}}_2^{\mathsf{T}}} & \frac{\partial \mathbf{g_1}}{\partial \boldsymbol{\theta}_2^{\mathsf{T}}} \end{bmatrix}, \text{ with } \frac{\partial \mathbf{g_1}}{\partial \widetilde{\boldsymbol{\phi}}_1^{\mathsf{T}}} = \mathbf{0}, \quad \frac{\partial \mathbf{g_1}}{\partial \boldsymbol{\theta}_1{}^{\mathsf{T}}} = -\boldsymbol{Z}_1, \quad \frac{\partial \mathbf{g_1}}{\partial \widetilde{\boldsymbol{\phi}}_2^{\mathsf{T}}} = -\left[\boldsymbol{Z}_2 \ \boldsymbol{Z}_3\right] \boldsymbol{M}_2^{\mathsf{T}}, \quad \frac{\partial \mathbf{g_1}}{\partial \boldsymbol{\theta}_2{}^{\mathsf{T}}} = \boldsymbol{I}_{K_d R}.$$

The matrices $\boldsymbol{Z}_1, \boldsymbol{Z}_2, \boldsymbol{Z}_3$ are given in Appendix A. As a result, we have a basis $\boldsymbol{U}$ of $\ker(\boldsymbol{G})$ such that

$$\boldsymbol{U} = \begin{bmatrix} \boldsymbol{I}_{(I+J+I_d+J_d)R-4R} \\ \begin{bmatrix} \mathbf{0} & \boldsymbol{Z}_1 & \left[\boldsymbol{Z}_2 \ \boldsymbol{Z}_3\right] \boldsymbol{M}_2^{\mathsf{T}} \end{bmatrix} \end{bmatrix}. \tag{39}$$

We thus obtain the Blind-CCRB (**Blind-CCRB**) by plugging (39) into (24).

5.3. Performance bounds for fully-coupled model

For the fully-coupled problem, we can compute the CCRB and reparameterized CRB in both scenarios.

5.3.1. Scenario 1 – linear constraints

In the first scenario, the most straightforward approach is to consider the reparameterization change for the CCRB, using model (6)–(7).

We consider the random real Gaussian distributed dataset $\boldsymbol{\mathcal{Y}}$ such that $\boldsymbol{\mathcal{Y}} \sim \boldsymbol{\mathcal{N}}(\boldsymbol{\mu}(\widetilde{\boldsymbol{\psi}}), \boldsymbol{\Sigma})$, with

$$\boldsymbol{\mathcal{Y}} = \begin{bmatrix} \mathrm{vec}\{\boldsymbol{\mathcal{Y}}_1\}^{\mathsf{T}} & \mathrm{vec}\{\boldsymbol{\mathcal{Y}}_2\}^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}}, \quad \boldsymbol{\Sigma} = \mathrm{Diag}\{\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2\},$$

$$\boldsymbol{\mu}(\widetilde{\boldsymbol{\psi}}) = \begin{bmatrix} \mathrm{vec}\{[\![\boldsymbol{P}\boldsymbol{A}_2, \boldsymbol{Q}\boldsymbol{B}_2, \boldsymbol{C}_1]\!]\} \\ \mathrm{vec}\{[\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{R}\boldsymbol{C}_1]\!]\} \end{bmatrix} = \underbrace{\begin{bmatrix} \boldsymbol{I}_K \boxtimes \boldsymbol{Q} \boxtimes \boldsymbol{P} \\ \boldsymbol{R} \boxtimes \boldsymbol{I}_{IJ} \end{bmatrix}}_{\widetilde{\boldsymbol{P}}} \mathrm{vec}\{[\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_1]\!]\}.$$

Since $\widetilde{\boldsymbol{P}}$ is constant, we only have to consider the derivatives of $\mathrm{vec}\{[\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_1]\!]\}$ w.r.t. $\widetilde{\boldsymbol{\psi}}$ from (38). As a

result, we can compute the reparameterized FIM for $\widetilde{\boldsymbol{\psi}}$ (namely $\boldsymbol{F_c}(\widetilde{\boldsymbol{\psi}})$) from the Slepian-Bangs formula as

$$\boldsymbol{F_c}(\widetilde{\boldsymbol{\psi}}) = \boldsymbol{M}_3 \begin{bmatrix} \boldsymbol{S}_1 & \boldsymbol{S}_2 & \boldsymbol{S}_3 \end{bmatrix}^{\mathsf{T}} \widetilde{\boldsymbol{P}}^{\mathsf{T}} \boldsymbol{\Sigma}^{-1} \widetilde{\boldsymbol{P}} \begin{bmatrix} \boldsymbol{S}_1 & \boldsymbol{S}_2 & \boldsymbol{S}_3 \end{bmatrix} \boldsymbol{M}_3^{\mathsf{T}}. \tag{40}$$

The reparameterized CRB for $\widetilde{\boldsymbol{\psi}}$ is obtained as $\mathbf{rCRB}(\widetilde{\boldsymbol{\psi}}) = \boldsymbol{F_c}^{-1}(\widetilde{\boldsymbol{\psi}})$. For $\boldsymbol{x}$, the reparameterized CRB is

$$\mathbf{rCRB}(\boldsymbol{x}) = \left[\frac{\partial \mathbf{g_3}}{\partial \widetilde{\boldsymbol{\psi}}^{\mathsf{T}}}\right] \mathbf{rCRB}(\widetilde{\boldsymbol{\psi}}) \left[\frac{\partial \mathbf{g_3}}{\partial \widetilde{\boldsymbol{\psi}}^{\mathsf{T}}}\right]^{\mathsf{T}}.$$

Equivalently, we can compute the CCRB using (24) with

$$\boldsymbol{G} = \begin{bmatrix} \frac{\partial \mathbf{g_2}}{\partial \boldsymbol{\phi}_1^{\mathsf{T}}} & \frac{\partial \mathbf{g_2}}{\partial \boldsymbol{\theta}_1^{\mathsf{T}}} & \frac{\partial \mathbf{g_2}}{\partial \widetilde{\boldsymbol{\phi}}_2^{\mathsf{T}}} & \frac{\partial \mathbf{g_2}}{\partial \boldsymbol{\theta}_1^{\mathsf{T}}} \\ \frac{\partial \mathbf{g_1}}{\partial \boldsymbol{\phi}_1^{\mathsf{T}}} & \frac{\partial \mathbf{g_1}}{\partial \boldsymbol{\theta}_1^{\mathsf{T}}} & \frac{\partial \mathbf{g_1}}{\partial \widetilde{\boldsymbol{\phi}}_2^{\mathsf{T}}} & \frac{\partial \mathbf{g_1}}{\partial \boldsymbol{\theta}_1^{\mathsf{T}}} \end{bmatrix}, \text{ with} \tag{41}$$

$$\frac{\partial \mathbf{g_1}}{\partial \boldsymbol{\theta}_1^{\mathsf{T}}} = -\boldsymbol{I}_R \boxtimes \boldsymbol{R}, \quad \frac{\partial \mathbf{g_1}}{\partial \boldsymbol{\theta}_2^{\mathsf{T}}} = \boldsymbol{I}_{KR}, \quad \frac{\partial \mathbf{g_2}}{\partial \widetilde{\boldsymbol{\phi}}_2^{\mathsf{T}}} = -\begin{bmatrix} \boldsymbol{I}_R \boxtimes \boldsymbol{P} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_R \boxtimes \boldsymbol{Q} \end{bmatrix} \boldsymbol{M}_2^{\mathsf{T}}, \quad \frac{\partial \mathbf{g_2}}{\partial \boldsymbol{\phi}_1^{\mathsf{T}}} = \boldsymbol{I}_{(I_d+J_d)R},$$

and the other derivatives are zero.

5.3.2. Scenario 2 – non-linear constraints

In this subsection, the non-linear constraints in (32) yield to different bounds. In (41), we now have

$$\frac{\partial \mathbf{g_1}}{\partial \boldsymbol{\theta}_1^{\mathsf{T}}} = -\boldsymbol{Z}_1, \quad \frac{\partial \mathbf{g_1}}{\partial \boldsymbol{\theta}_2^{\mathsf{T}}} = \boldsymbol{I}_{KR}, \quad \frac{\partial \mathbf{g_1}}{\partial \widetilde{\boldsymbol{\phi}}_1^{\mathsf{T}}} = \boldsymbol{0}, \quad \frac{\partial \mathbf{g_1}}{\partial \widetilde{\boldsymbol{\phi}}_2^{\mathsf{T}}} = -\boldsymbol{M}_1 \left[\boldsymbol{Z}_2 \ \boldsymbol{Z}_3\right] \boldsymbol{M}_2^{\mathsf{T}},$$
$$\frac{\partial \mathbf{g_2}}{\partial \widetilde{\boldsymbol{\phi}}_1^{\mathsf{T}}} = \boldsymbol{I}_{(I_d+J_d-2)R}, \quad \frac{\partial \mathbf{g_2}}{\partial \widetilde{\boldsymbol{\phi}}_2^{\mathsf{T}}} = -\boldsymbol{M}_1 \operatorname{Diag}\{\boldsymbol{Z}_4, \boldsymbol{Z}_5\} \boldsymbol{M}_2^{\mathsf{T}}.$$

The matrices $\boldsymbol{Z}_4$ and $\boldsymbol{Z}_5$ are given in Appendix A, and the CCRB for $\widetilde{\boldsymbol{\omega}}$ is computed using (24) as well.

We can also consider the reparameterized CRB: we assume that $\boldsymbol{\mathcal{Y}} \sim \boldsymbol{\mathcal{N}}(\boldsymbol{\mu}(\widetilde{\boldsymbol{\psi}}), \boldsymbol{\Sigma})$, with

$$\boldsymbol{\mathcal{Y}} = \begin{bmatrix} \operatorname{vec}\{\boldsymbol{\mathcal{Y}}_1\}^{\mathsf{T}} & \operatorname{vec}\{\boldsymbol{\mathcal{Y}}_2\}^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}}, \quad \boldsymbol{\Sigma} = \operatorname{Diag}\{\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2\}, \quad \boldsymbol{\mu}(\widetilde{\boldsymbol{\psi}}) = \begin{bmatrix} \operatorname{vec}\{[\![\boldsymbol{P}\boldsymbol{A}_2\boldsymbol{D}_\alpha^{-1}, \boldsymbol{Q}\boldsymbol{B}_2\boldsymbol{D}_\beta^{-1}, \boldsymbol{C}_1]\!]\} \\ \operatorname{vec}\{[\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{R}\boldsymbol{C}_1(\boldsymbol{D}_\alpha\boldsymbol{D}_\beta)^{-1}]\!]\} \end{bmatrix}.$$

The Jacobian of $\boldsymbol{\mu}(\widetilde{\boldsymbol{\psi}})$ is the matrix

$$\frac{\partial \boldsymbol{\mu}}{\partial \widetilde{\boldsymbol{\psi}}^{\mathsf{T}}} = \boldsymbol{M}_1 \begin{bmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 & \boldsymbol{X}_3 \\ \boldsymbol{X}_5 & \boldsymbol{X}_6 & \boldsymbol{X}_4 \end{bmatrix} \boldsymbol{M}_2^{\mathsf{T}}.$$

The matrices $\boldsymbol{X}_i$ $(i \in \{1, \ldots, 6\})$ are given in Appendix B. The reparameterized FIM is computed as

$$\boldsymbol{F}_{\boldsymbol{c}}(\widetilde{\boldsymbol{\psi}}) = \boldsymbol{M}_3 \left[\frac{\partial \boldsymbol{\mu}}{\partial \widetilde{\boldsymbol{\psi}}^{\mathsf{T}}}\right] \boldsymbol{\Sigma}^{-1} \left[\frac{\partial \boldsymbol{\mu}}{\partial \widetilde{\boldsymbol{\psi}}^{\mathsf{T}}}\right] \boldsymbol{M}_3^{\mathsf{T}}. \tag{42}$$

6. Computer results

All simulations were run on a MacBook Pro with 2.3 GHz Intel Core i5 and 16GB RAM. For basic tensor operations we used TensorLab 3.0 [49]. The code is implemented in MATLAB and is available online at `https://github.com/cprevost4/CCRB_Software`.

6.1. Simulations setup

The entries of the true CP factors $\boldsymbol{A}_2$, $\boldsymbol{B}_2$, $\boldsymbol{C}_1$ were generated once as i.i.d. real standard Gaussian variables, and the first rows of $\boldsymbol{A}_2$, $\boldsymbol{B}_2$ were set to ones. The true CP factors $\boldsymbol{A}_1$, $\boldsymbol{B}_1$, $\boldsymbol{C}_2$ were constructed according to the parameter constraints for each scenario. In all experiments, the degradation matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ are generated as blurring and downsampling matrices using a Gaussian filter of length q and a downsampling ratio d. For the sake of simplicity but without loss of generality, we also assume that $\boldsymbol{P} = \boldsymbol{Q}$. The degradation matrix $\boldsymbol{R}$ is a selection-and-averaging matrix that selects the common third-order slabs of $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}_2$. We refer to the Appendix C for more details on the construction of these matrices.

We simulate the performance of the coupled CP model under additive Gaussian noise. The SNR on the observed tensors in dB is defined as $SNR_i = 10 \log_{10}\left(\|\boldsymbol{\mathcal{Y}}_i\|_F^2 / \|\boldsymbol{\mathcal{E}}_i\|_F^2\right)$, $(i = 1, 2)$. We fix SNR_2 to 20dB while SNR_1 varies from 5 to 60dB, unless otherwise specified. In the following figures, we will plot our results for various values of SNR_1 while SNR_2 remains constant.

The model parameters are retrieved using MLE with at most 5000 iterations. For estimation in the uncoupled case, we use ALS [39] with random initialization for the factor matrices. For the fully-coupled case, STEREO, the algorithm proposed in [8] is used. For the blind case, we use Blind-STEREO [8]. To speed up the convergence of the coupled algorithms, the CP factors obtained by uncoupled ALS are used as initialization. For each realization, the best out of 10 initializations is picked. The scaling ambiguities are corrected during estimation so that the first rows of the $\widehat{\boldsymbol{A}}_i$, $\widehat{\boldsymbol{B}}_i$ factors are composed of ones. Then, the permutation ambiguities are corrected in the estimated factors. The correct permutation is obtained by searching for the best column permutation of the estimated $\widehat{\boldsymbol{C}}_1$ for a fixed reference $\boldsymbol{C}_1$. This is performed as follows. The columns of $\widehat{\boldsymbol{C}}_1$ are processed in turn, by starting with the one closest to the reference. Once two columns are associated, they are removed from each matrix, and we keep going with a smaller number of columns. Note that this greedy approach is known to be sub-optimal, and one could proceed optimally using the Hungarian algorithm [50] or by replacing permutation by a bistochastic matrix [51]. This permutation

is also applied to $\widehat{\boldsymbol{A}}_1$ and $\widehat{\boldsymbol{B}}_1$, and likewise for $\widehat{\boldsymbol{A}}_2$, $\widehat{\boldsymbol{B}}_2$, $\widehat{\boldsymbol{C}}_2$. For each noise realization, the squared error between the vectorized groundtruth and estimated factors is computed. Please note that the fixed entries of the factors are not taken into account. Finally, the MSE is obtained by averaging the squared errors over 500 noise realizations.

In our experiments, we consider as reference the uniform MSE (UMSE) and uniform CRB (UCRB) obtained from the MSE and CRB matrix traces, as widely considered in, e.g., [52, 53, 54]. The expressions proposed in this paper allow for calculation of the reparameterized UCRB, uniform CCRB (UCCRB) and uniform Blind-CCRB (Blind-UCCRB) by taking the trace of these matrices. Thus in the following figures, we will assess uniform efficiency of the estimators[9]. We will plot our results for the parameters $\boldsymbol{\psi}$ and $\boldsymbol{\xi}$ such that

$$\widetilde{\boldsymbol{\psi}}^{\mathsf{T}} = \begin{bmatrix} \widetilde{\boldsymbol{\phi}}_2^{\mathsf{T}} & \boldsymbol{\theta}_1^{\mathsf{T}} \end{bmatrix}, \quad \widetilde{\boldsymbol{\xi}}^{\mathsf{T}} = \begin{bmatrix} \widetilde{\boldsymbol{\phi}}_1^{\mathsf{T}} & \boldsymbol{\theta}_2^{\mathsf{T}} \end{bmatrix},$$

which correspond respectively to the CP factors of $\boldsymbol{\mathcal{X}}$, and the degraded factors.

6.2. Numerical quivalence between CCRB and reparameterized CRB

In this subsection, we illustrate the results of [22, 42] regarding the equivalence between the CCRB (3.4) and its reparameterization change (3.5). We first consider $I = J = 18$, $I_d = J_d = 6$, $K = 16$ and $K_d = 8$, and $R = 3$. In Figure 3, we show on a semi-log scale the UCCRB and its reparameterization change for $\widetilde{\boldsymbol{\omega}}$ in the fully coupled case and scenario 1 (linear constraints). In Figure 4, we consider scenario 2 and additionally plot the uncoupled UCRB, UCCRB and reparameterization changes for $\widetilde{\boldsymbol{\omega}}$ in the partially-coupled case.

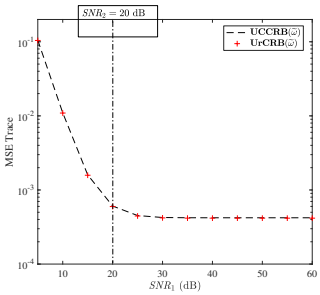

Figure 3: Scenario 1: UCCRB and its reparameterization change for $\widetilde{\boldsymbol{\omega}}$ versus SNR_1 for fixed $SNR_2 = 20$ dB.

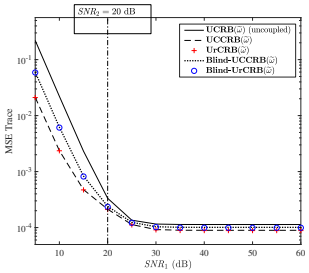

Figure 4: Scenario 2: Uniform lower bounds (uncoupled, partially-coupled, fully-coupled) for $\widetilde{\boldsymbol{\omega}}$ versus SNR_1 for fixed $SNR_2 = 20$ dB.

For both fully-coupled and blind problems, the UCCRB and its reparameterization change are numerically equivalent. Moreover, the Blind-UCCRB is above the fully-coupled UCCRB.

[9] Please note that uniform efficiency implies efficiency for each entry of the parameters.

6.3. *Asymptotic values for constrained FIM*

We notice that the uniform bounds tend to a constant for $SNR_1 > SNR_2$. To explain why such an elbow occurs, we seek for asymptotic values for the constrained FIM, *i.e.*, for $\sigma_1^2 \to \infty$. For scenario 1 (linear constraints), we resort to (40) for the constrained FIM. Developing (40) yields the following matrix:

$$\boldsymbol{F_c}(\widetilde{\boldsymbol{\psi}}) = \boldsymbol{M}_3 \begin{bmatrix} \boldsymbol{S}_1^\mathsf{T}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_1 & \boldsymbol{S}_1^\mathsf{T}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_2 & \boldsymbol{S}_1^\mathsf{T}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_3 \\ \boldsymbol{S}_2^\mathsf{T}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_1 & \boldsymbol{S}_2^\mathsf{T}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_2 & \boldsymbol{S}_2^\mathsf{T}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_3 \\ \boldsymbol{S}_3^\mathsf{T}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_1 & \boldsymbol{S}_3^\mathsf{T}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_2 & \boldsymbol{S}_3^\mathsf{T}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_3 \end{bmatrix} \boldsymbol{M}_3^\mathsf{T}, \tag{43}$$

where $\widetilde{\boldsymbol{\mathcal{P}}} = \frac{1}{\sigma_1^2}(\boldsymbol{I} \boxtimes \boldsymbol{Q}^\mathsf{T}\boldsymbol{Q} \boxtimes \boldsymbol{P}^\mathsf{T}\boldsymbol{P}) + \frac{1}{\sigma_2^2}(\boldsymbol{R}^\mathsf{T}\boldsymbol{R} \boxtimes \boldsymbol{I})$. Thus for $(i,j) \in \{1,\dots,3\}$,

$$\lim_{\sigma_1^2 \to \infty} \boldsymbol{S}_i^\mathsf{T}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_j = \frac{1}{\sigma_2^2}(\boldsymbol{S}_i^\mathsf{T}(\boldsymbol{R}^\mathsf{T}\boldsymbol{R} \boxtimes \boldsymbol{I})\boldsymbol{S}_j), \Rightarrow \lim_{\sigma_1^2 \to \infty} \boldsymbol{F_c}^{(i,j)}(\widetilde{\boldsymbol{\psi}}) = \boldsymbol{S}_i^\mathsf{T}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_j, \tag{44}$$

where $\boldsymbol{F_c}^{(i,j)}(\widetilde{\boldsymbol{\psi}})$ stands for the (i,j)-th block of $\boldsymbol{F_c}(\widetilde{\boldsymbol{\psi}})$.

For scenario 2 with non-linear constraints, developing in (24) the term corresponding to the constrained FIM $\boldsymbol{U}^\mathsf{T}\boldsymbol{F}\boldsymbol{U}$ yields

$$\begin{aligned} \lim_{\sigma_1^2 \to \infty} (\boldsymbol{U}^\mathsf{T}\boldsymbol{F}\boldsymbol{U}) = \frac{1}{\sigma_2^2}(&\boldsymbol{S}_A^\mathsf{T}\boldsymbol{S}_A + \boldsymbol{S}_B^\mathsf{T}\boldsymbol{S}_B + \boldsymbol{Z}_1^\mathsf{T}\boldsymbol{S}_C^\mathsf{T}\boldsymbol{S}_C\boldsymbol{Z}_1 + \boldsymbol{Z}_2^\mathsf{T}\boldsymbol{S}_C^\mathsf{T}\left[\boldsymbol{S}_A + \boldsymbol{S}_C\boldsymbol{Z}_2\right] \\ &+ \boldsymbol{Z}_3^\mathsf{T}\boldsymbol{S}_C^\mathsf{T}\left[\boldsymbol{S}_B + \boldsymbol{S}_C\boldsymbol{Z}_3\right] + \boldsymbol{S}_A^\mathsf{T}\boldsymbol{S}_C\boldsymbol{Z}_2 + \boldsymbol{S}_B^\mathsf{T}\boldsymbol{S}_C\boldsymbol{Z}_3), \end{aligned} \tag{45}$$

where $\boldsymbol{S}_A = (\boldsymbol{C}_2 \odot \boldsymbol{B}_2) \boxtimes \boldsymbol{I}_I$, $\boldsymbol{S}_B = \boldsymbol{\Pi}_2^{(2,1)}((\boldsymbol{C}_2 \odot \boldsymbol{A}_2) \boxtimes \boldsymbol{I}_J)$ and $\boldsymbol{S}_C = \boldsymbol{\Pi}_2^{(3,1)}((\boldsymbol{B}_2 \odot \boldsymbol{A}_2) \boxtimes \boldsymbol{I}_K)$, and $\boldsymbol{\Pi}_2^{(2,1)}$ and $\boldsymbol{\Pi}_2^{(3,1)}$ are permutation matrices that link the entries of $\mathrm{vec}\{\boldsymbol{Y}_2^{(2)}\}$ (resp. $\mathrm{vec}\{\boldsymbol{Y}_2^{(3)}\}$) to those of $\mathrm{vec}\{\boldsymbol{Y}_2^{(1)}\}$.

The asymptotic values for $\mathbf{rCRB}(\widetilde{\boldsymbol{\psi}})$ (for scenario 1) and $(\boldsymbol{U}^\mathsf{T}\boldsymbol{F}\boldsymbol{U})^{-1}$ (for scenario 2) when $\sigma_1^2 \to 0$ can be obtained by inversion of (44) and (45), respectively.

In Figure 5, we illustrate those results by plotting $\mathrm{Tr}(\mathbf{UrCRB}(\widetilde{\boldsymbol{\psi}}))$ (for scenario 1) and $\mathrm{Tr}((\boldsymbol{U}^\mathsf{T}\boldsymbol{F}\boldsymbol{U})^{-1})$ (for scenario 2) for $SNR_2 \in \{15, 30, 45\}$ dB, as well as their asymptotic values.

Figure 5: $\mathrm{Tr}(\mathbf{UrCRB}(\widetilde{\boldsymbol{\psi}}))$ (scenario 1, left); $\mathrm{Tr}((\boldsymbol{U}^\mathsf{T}\boldsymbol{F}\boldsymbol{U})^{-1})$ (scenario 2, right) and asymptotic values, versus SNR_1 for fixed SNR_2.

6.4. *Assessing the efficiency of the estimators*

In this subsection, we assess the efficiency of the estimators introduced in Section 2.6 for the reconstruction of the target tensor. We keep the same dimensions as in Section 6.2, and compare the UMSE with corresponding uniform bounds. For scenario 1, we only compare the UMSE provided by STEREO to the UCCRB obtained as in Section 5.3.1. For scenario 2, we also compare the UMSE given by Blind-STEREO to the Blind-UCCRB. In Figures 6 and 7, we show on a semi-log scale the bounds and UMSE for $\boldsymbol{x}$ in scenarios 1 and 2, respectively.

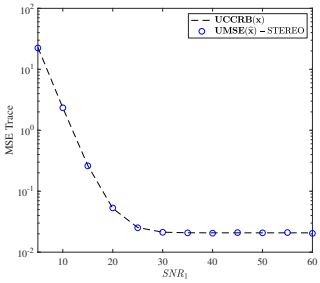

Figure 6: Scenario 1: UCCRB and UMSE from STEREO for $\boldsymbol{x}$, versus SNR_1 for fixed SNR_2.

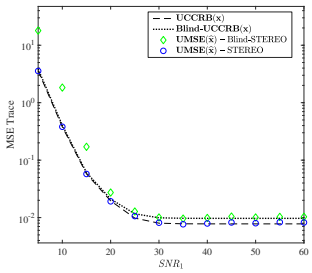

Figure 7: Scenario 2: UCCRB and Blind-UCCRB, UMSE from STEREO and Blind-STEREO for $\boldsymbol{x}$, versus SNR_1 for fixed SNR_2.

The UMSE obtained from STEREO follows the UCCRB in both scenarios. For the partially-coupled problem in scenario 2, the UMSE given by Blind-STEREO follows the Blind-UCCRB for $SNR_1 \geq 20$dB. Thus, in both scenarios, the estimators asymptotically reach their corresponding bounds. This implies that they are asymptotically efficient for each entry of the parameters.

Next, we assess performance of STEREO with respect to the two estimation scenarios. For the first scenario, we generate the model according to (28) that corresponds to the first scenario. For scenario 2, we generate model (32) with non-linear constraints between the parameters. For each scenario, we run STEREO according to the model. We also compute the CCRB as in Section 5.3.1 and Section 5.3.2 for the parameter $\boldsymbol{x}$: thus in each scenario we consider the correct CCRB as well as the CCRB obtained from the wrong model. In Figure 8, we show on a semi-log scale the UCCRB bounds and UMSE for both scenarios.

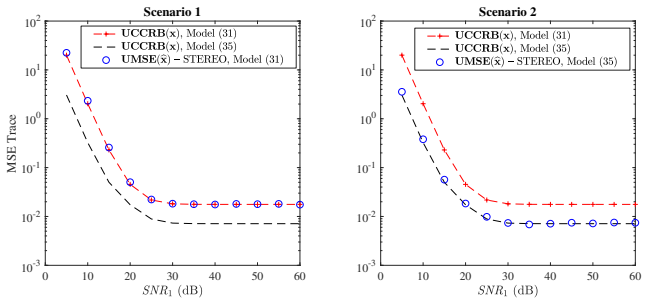

Figure 8: UMSE from STEREO and UCCRB for $\boldsymbol{x}$, scenarios 1 (left) and 2 (right), versus SNR_1 for fixed SNR_2.

In both cases, the UMSE follows the UCCRB computed from the right model. That is, for scenario 1, the UMSE reaches the UCCRB obtained with model (28), while it reaches the bound obtained with (32) in scenario 2. Thus STEREO is efficient provided that the right model is employed.

6.5. Impact of λ on the performance and modified STEREO scheme

In the following subsections, we study performance of STEREO in a case where generic uniqueness of $\boldsymbol{\mathcal{Y}}_1$ is not guaranteed, but the condition (3) for unique noiseless recovery of $\boldsymbol{\mathcal{X}}$ by STEREO is still satisfied. Such a case can be obtained by considering large tensor ranks. Contrary to Section 6.4, where the $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ are generically unique, we expect to encounter cases where the algorithm does not converge to a global minimum due to the rank being larger than (some of) the dimensions of $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$. Thus we consider a modified choice of the regularization parameter λ to circumvent these difficulties.

We first illustrate the influence of λ on the performance of STEREO with a toy example. We generate the model as in Section 6.2 with fixed $SNR_2 = 40$dB. We consider several values for the regularization parameter: $\lambda = 1 \cdot 10^7$, $\lambda = 1$ and $\lambda = 1 \cdot 10^{-4}$. They correspond to the "true" regularization parameters for $SNR_1 = 5$dB, $SNR_1 = SNR_2$ and $SNR_1 = 60$dB, respectively. In Figure 9, we plot on a semi-log scale the UCCRB for $\widetilde{\boldsymbol{\omega}}$ and UMSE obtained with different λ.

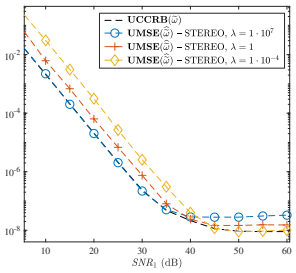

Figure 9: UCCRB for $\widetilde{\boldsymbol{\omega}}$ and UMSE for different λ.

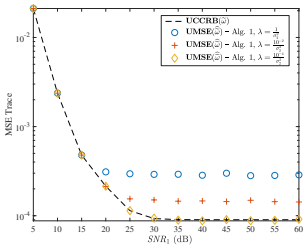

Figure 10: UCCRB for $\widetilde{\boldsymbol{\omega}}$ and UMSE at different steps of Algorithm 1.

For $\lambda = 1 \cdot 10^7$, we notice that the UMSE reaches the UCCRB for $SNR_1 \leq SNR_2$ even if λ is larger than the "true" λ. For higher SNR, the UMSE tends to a constant. For $\lambda = 1$, we can see that the UMSE is above the UCCRB for each noise level except when SNR_1 and SNR_2 have the same order. Finally, for $\lambda = 1 \cdot 10^{-4}$, while the UMSE is above the UCCRB for $SNR_1 \leq SNR_2$, it does reach the bound for higher SNR. Figure 9 shows that small values of λ lead to better performance for high SNR, which is exactly what we are aiming at.

Thus we propose a modified procedure for STEREO. For each noise level, we successively run several iterations of STEREO with decreasing values of λ. The balance parameter is initialized to $\lambda = \frac{1}{\sigma_2^2}$. Indeed, the value $\sigma_1^2 = 1$ corresponds to $SNR_1 = 0$dB; hence in our experiments, we always have $\sigma_1^2 < 1$, which

guarantees that the initial value of λ is always higher than $\frac{\sigma_1^2}{\sigma_2^2}$. We refer to this setup as "modified" and describe this new procedure in Algorithm 1.

Algorithm 1: Modified STEREO with decreasing λ
Input: $\boldsymbol{A}_{2,0}$, $\boldsymbol{B}_{2,0}$, $\boldsymbol{C}_{1,0}$
Initialize $\lambda = \frac{1}{\sigma_2^2}$;
repeat
1. $(\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_1) \leftarrow$ STEREO with 1000 iterations;
2. $\lambda \leftarrow \frac{\lambda}{10}$;
3. $(\boldsymbol{A}_{2,0}, \boldsymbol{B}_{2,0}, \boldsymbol{C}_{1,0}) \leftarrow (\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_1)$;

until $\lambda = \frac{\sigma_1^2}{\sigma_2^2}$;
Return $\boldsymbol{A}_2$, $\boldsymbol{B}_2$, $\boldsymbol{C}_1$.

Since identifiability of $\boldsymbol{\mathcal{Y}}_1$ is not guaranteed, uncoupled ALS on $\boldsymbol{\mathcal{Y}}_1$ is not guaranteed to converge. Thus in this subsection, we initialize STEREO as in [8]:

$$\begin{cases} \boldsymbol{A}_{2,0}, \boldsymbol{B}_{2,0}, \boldsymbol{C}_{2,0} = \mathbf{CPD}_R(\boldsymbol{\mathcal{Y}}_2), \\ \boldsymbol{C}_{1,0}^{\mathsf{T}} = (\boldsymbol{Q}\boldsymbol{B}_{2,0} \odot \boldsymbol{P}\boldsymbol{A}_{2,0})^{\dagger} \boldsymbol{Y}_2^{(3)}, \end{cases} \tag{46}$$

where the operation $\mathbf{CPD}_R$ returns estimated CP factors[10] with rank R. In fact, (46) boils down to considering $\lambda = \infty$. For this reason, we expect STEREO not to converge when $\frac{\sigma_1^2}{\sigma_2^2}$ is low, that is, $SNR_1 \geq SNR_2$. It should be mentioned that, for $SNR_1 < SNR_2$, Algorithm 1 might not needed, since the "true" lambda is very large. The initialization (46) with $\lambda = \infty$ is expected to provide a good estimation of the parameter $\boldsymbol{x}$ in that case.

To provide more intuition on how Algorithm 1 works, in Figure 10 we plot the UCCRB for $\widetilde{\boldsymbol{\omega}}$ and the UMSE obtained at different steps of Algorithm 1 with decreasing values of λ. We chose fixed $SNR_2 = 20$dB in our simulations. At the initialization step with $\lambda = \frac{1}{\sigma_2^2}$, the UMSE reaches the bound only for $SNR_1 \leq SNR_2$. After two steps, the UMSE reaches the UCCRB up until $SNR_1 = 25$dB and the UMSE decreases for high SNR. After two more steps of Algorithm 1, the UMSE reaches the bound for all values of SNR_1.

6.6. Performance of STEREO without identifiability of $\boldsymbol{\mathcal{Y}}_1$

To study the interest of the procedure proposed in Algorithm 1, we take $I_d = J_d = 4$, $I = J = 16$, $K_d = 10$ and $K = 20$, and $d = 4$, $q = 3$. We consider fixed $SNR_2 = 40$dB while SNR_1 varies between 5dB and 60dB. For these dimensions, the generic uniqueness of $\boldsymbol{\mathcal{Y}}_1$ is proved for $N \leq 9$ [38, Theorem 1.1], while the Kruskal condition for $\boldsymbol{\mathcal{Y}}_1$ is satisfied for $N \leq 6$. In fact, the Kruskal condition the the CPD of $\boldsymbol{\mathcal{Y}}_1$ provides a sufficient condition for unique recovery of the tensor. Condition (3) on unique recovery of $\boldsymbol{\mathcal{X}}$ by STEREO in the noiseless case gives $R \leq 16$. We address scenario 2 only, and tensor ranks $R = 10$, $R = 12$, and $R = 14$.

[10] In practice, this operation is performed using TensorLab 3.0.

We run STEREO and Algorithm 1, and average the UMSE over 500 noise realizations. In Figures 11–13, we plot on a semi-log scale the UCCRB and UMSE for $\boldsymbol{x}$ and tensor ranks $R = 10$, $R = 12$, $R = 14$, respectively.

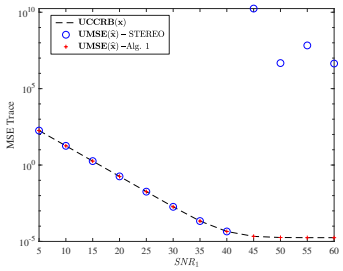

Figure 11: CCRB and MSE traces for $\boldsymbol{x}$ and tensor rank $R = 10$.

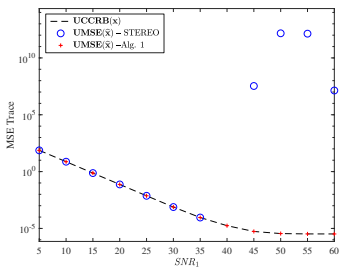

Figure 12: CCRB and MSE traces for $\boldsymbol{x}$ and tensor rank $R = 12$.

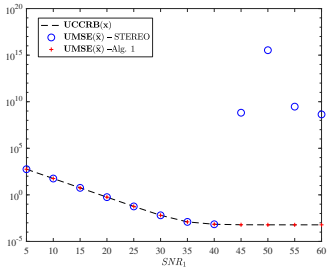

Figure 13: CCRB and MSE traces for $\boldsymbol{x}$ and tensor rank $R = 14$.

First, in Figures 11–13, we can see that for $SNR_1 \geq SNR_2$, STEREO with $\lambda = \frac{\sigma_1^2}{\sigma_2^2}$ does not converge indeed. Our guess is that the performance of the algorithm degrades when R is very large, especially when it becomes larger than (some of) the dimensions of the tensors. However, Algorithm 1 does not exhibit such behaviour: in all figures, the UMSE provided by STEREO reaches the UCCRB in this setting.

7. Conclusion

In this paper, we have provided a general framework for the calculation of the CCRB for coupled tensor models admitting CP decompositions. We have introduced two different estimation scenarios, allowing for a performance comparison of the uncoupled, partially-coupled, and fully-coupled problems. We have shown that the existing ML estimators STEREO and Blind-STEREO are efficient in optimal estimation conditions (*i.e.*, when the rank reduction allows for good estimation). In some cases, however, the tensor rank does not allow for correct estimation of the parameters by STEREO: in such cases, we have proposed an algorithm reaching the CCRB for high SNRs, contrary to the traditional STEREO.

STEREO reaches the CCRB for reconstruction of the tensor $\boldsymbol{\mathcal{X}}$. However, for scenario 2 with non-linear constraints, the CCRB is not a lower bound on constrained parameter estimation [55], and the Lehmann-unbiased CCRB (LU-CCRB) introduced in [55] is inferior to or equal to the standard CCRB. As a result, when interested in estimating the low-rank factors underlying $\boldsymbol{\mathcal{X}}$, it is sensible to seek for new constrained algorithms that reach the LU-CCRB. This matter, which is of great interest, will be explored in future works.

Appendix A. Derivatives for CCRB

We give the expression of the matrices $\boldsymbol{Z}_i$ ($i \in \{1,5\}$) for the CCRB in Sections 5.2 and 5.3:

$$
\begin{aligned}
\boldsymbol{Z}_1 &= (\boldsymbol{D}_\alpha \boldsymbol{D}_\beta)^{-1} \boxtimes \boldsymbol{R}, \\
\boldsymbol{Z}_2 &= -(\boldsymbol{D}_\alpha^2 \boldsymbol{D}_\beta)^{-1} \boxtimes \boldsymbol{R})(\boldsymbol{I}_R \odot \boldsymbol{C}_1)(\boldsymbol{I}_R \boxtimes \boldsymbol{P}_{1,:}), \\
\boldsymbol{Z}_3 &= -(\boldsymbol{D}_\alpha \boldsymbol{D}_\beta^2)^{-1} \boxtimes \boldsymbol{R})(\boldsymbol{I}_R \odot \boldsymbol{C}_1)(\boldsymbol{I}_R \boxtimes \boldsymbol{Q}_{1,:}), \\
\boldsymbol{Z}_4 &= (\boldsymbol{D}_\alpha^{-1} \boxtimes \boldsymbol{P}) - (\boldsymbol{D}_\alpha^{-2} \boxtimes \boldsymbol{P})(\boldsymbol{I}_R \odot \boldsymbol{A}_2)(\boldsymbol{I}_R \boxtimes \boldsymbol{P}_{1,:}), \\
\boldsymbol{Z}_5 &= (\boldsymbol{D}_\beta^{-1} \boxtimes \boldsymbol{Q}) - (\boldsymbol{D}_\beta^{-2} \boxtimes \boldsymbol{Q})(\boldsymbol{I}_R \odot \boldsymbol{B}_2)(\boldsymbol{I}_R \boxtimes \boldsymbol{Q}_{1,:}).
\end{aligned}
$$

Appendix B. Derivatives for reparameterized CRB

We give the expression of the matrices $\boldsymbol{X}_i$ ($i \in \{1,6\}$) used for the reparameterized CRB in Section 5.3. The permutation matrices $\boldsymbol{\Pi}_i^{(2,1)}, \boldsymbol{\Pi}_i^{(3,1)}$ link the entries of $\mathrm{vec}\{\boldsymbol{Y}_i^{(2)}\}$ (resp. $\mathrm{vec}\{\boldsymbol{Y}_i^{(3)}\}$) to those of $\mathrm{vec}\{\boldsymbol{Y}_i^{(1)}\}$.

$$
\begin{aligned}
\boldsymbol{X}_1 &= \left[\left(\boldsymbol{I}_R \odot (\boldsymbol{C}_1 \odot \boldsymbol{Q}\boldsymbol{B}_2\boldsymbol{D}_\beta^{-1})\right) \boldsymbol{D}_\alpha^{-1} \boxtimes \boldsymbol{P}\right] \left[\boldsymbol{I}_{IR} - (\boldsymbol{D}_\alpha^{-1} \boxtimes \boldsymbol{I}_I)(\boldsymbol{I}_R \odot \boldsymbol{A}_2)(\boldsymbol{I}_R \boxtimes \boldsymbol{P}_{1,:})\right], \\
\boldsymbol{X}_2 &= (\boldsymbol{I}_R \boxtimes \boldsymbol{\Pi}_1^{(2,1)}) \left[\left(\boldsymbol{I}_R \odot (\boldsymbol{C}_1 \odot \boldsymbol{P}\boldsymbol{A}_2\boldsymbol{D}_\alpha^{-1})\right) \boldsymbol{D}_\beta^{-1} \boxtimes \boldsymbol{Q}\right] \left[\boldsymbol{I}_{JR} - (\boldsymbol{D}_\beta^{-1} \boxtimes \boldsymbol{I}_J)(\boldsymbol{I}_R \odot \boldsymbol{B}_2)(\boldsymbol{I}_R \boxtimes \boldsymbol{Q}_{1,:})\right], \\
\boldsymbol{X}_3 &= (\boldsymbol{I}_R \boxtimes \boldsymbol{\Pi}_1^{(3,1)}) \left[\left(\boldsymbol{I}_R \odot (\boldsymbol{Q}\boldsymbol{B}_2\boldsymbol{D}_\beta^{-1} \odot \boldsymbol{P}\boldsymbol{A}_2\boldsymbol{D}_\alpha^{-1})\right) \boxtimes \boldsymbol{I}_K\right], \\
\boldsymbol{X}_4 &= (\boldsymbol{I}_R \boxtimes \boldsymbol{\Pi}_2^{(3,1)}) \left[(\boldsymbol{I}_R \odot (\boldsymbol{B}_2 \odot \boldsymbol{A}_2))\, (\boldsymbol{D}_\alpha \boldsymbol{D}_\beta)^{-1} \boxtimes \boldsymbol{R}\right], \\
\boldsymbol{X}_5 &= \left[\left(\boldsymbol{I}_R \odot (\boldsymbol{R}\boldsymbol{C}_1\boldsymbol{D}_\beta^{-1} \odot \boldsymbol{B}_2)\right) \boldsymbol{D}_\alpha^{-1} \boxtimes \boldsymbol{I}_I\right] \left[\boldsymbol{I}_{IR} - (\boldsymbol{D}_\alpha^{-1} \boxtimes \boldsymbol{I}_I)(\boldsymbol{I}_R \odot \boldsymbol{A}_2)(\boldsymbol{I}_R \boxtimes \boldsymbol{P}_{1,:})\right], \\
\boldsymbol{X}_6 &= (\boldsymbol{I}_R \boxtimes \boldsymbol{\Pi}_2^{(2,1)}) \left[\left(\boldsymbol{I}_R \odot (\boldsymbol{R}\boldsymbol{C}_1\boldsymbol{D}_\alpha^{-1} \odot \boldsymbol{A}_2)\right) \boldsymbol{D}_\beta^{-1} \boxtimes \boldsymbol{I}_J\right] \left[\boldsymbol{I}_{JR} - (\boldsymbol{D}_\beta^{-1} \boxtimes \boldsymbol{I}_J)(\boldsymbol{I}_R \odot \boldsymbol{B}_2)(\boldsymbol{I}_R \boxtimes \boldsymbol{Q}_{1,:})\right].
\end{aligned}
$$

Appendix C. Degradation matrices

As initially proposed in [56] and used in [8], $\boldsymbol{P}$ is constructed as $\boldsymbol{P} = \boldsymbol{S}_1\boldsymbol{T}_1$, where $\boldsymbol{T}_1$ is a blurring Toeplitz matrix and $\boldsymbol{S}_1$ is a downsampling matrix. The blurring matrix is constructed from a Gaussian blurring kernel $\boldsymbol{\phi} \in \mathbb{R}^{q\times 1}$ with a standard deviation $\sigma = \frac{q\cdot\sqrt{2\log 2}}{4}$. For $m \in \{1,\dots,q\}$ and $m' = m - \left\lceil \frac{q}{2} \right\rceil$, we have $\phi(m) = \exp\left(\frac{-m'^2}{2\sigma^2}\right)$. Thus, $\boldsymbol{T}_1 \in \mathbb{R}^{I\times I}$ can be seen as

$$
\boldsymbol{T}_1 = \begin{bmatrix}
\phi(\left\lceil \frac{q}{2} \right\rceil) & \dots & \phi(q) & 0 & \dots & 0 \\
\vdots & \ddots & & \ddots & \ddots & \vdots \\
\phi(1) & & \ddots & & \ddots & 0 \\
0 & \ddots & & \ddots & & \phi(q) \\
\vdots & \ddots & \ddots & & \ddots & \vdots \\
0 & \dots & 0 & \phi(1) & \dots & \phi(\left\lceil \frac{q}{2} \right\rceil)
\end{bmatrix}.
$$

The matrix $\boldsymbol{S}_1 \in \mathbb{R}^{I_d \times I}$, with downsampling ratio d, is made of I_d independent rows such that for $i \in \{1, \ldots, I_d\}$, $(\boldsymbol{S}_1)_{i,2+(i-1)d} = 1$ and the other coefficients are zeros.

The degradation matrix $\boldsymbol{R} \in \mathbb{R}^{K_d \times K}$ is a selection-averaging matrix. Each row represents a band in $\boldsymbol{\mathcal{Y}}_2$; coefficients are set to ones for common bands with $\boldsymbol{\mathcal{X}}$, and zeros elsewhere. The coefficients are averaged per-row. In our simulations, we average $\boldsymbol{\mathcal{X}}$ bands two by two. Below is an example of a 3×6 matrix:

$$\boldsymbol{R} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

References

[1] T. Wilderjans, E. Ceulemans, I. Van Mechelen, Simultaneous analysis of coupled data blocks differing in size: A comparison of two weighting schemes, Computational Statistics & Data Analysis 53 (4) (2009) 1086–1098.

[2] T. Stathaki, Image fusion: algorithms and applications, Elsevier, 2011.

[3] K. G. Jöreskog, Simultaneous factor analysis in several populations, Psychometrika 36 (4) (1971) 409–426.

[4] D. Lahat, T. Adali, C. Jutten, Multimodal data fusion: an overview of methods, challenges and prospects, Proc. IEEE 103 (9) (2015) 1449–1477.

[5] J. Cohen, Environmental multiway data mining, Ph.D. thesis, Université Grenoble Alpes (2016).

[6] B. Ermiş, E. Acar, A. T. Cemgil, Link prediction in heterogeneous data via generalized coupled tensor factorization, Data Mining and Knowledge Discovery 29 (1) (2015) 203–236.

[7] B. Ermiş, E. Acar, A. T. Cemgil, Link prediction via generalized coupled tensor factorisation, arXiv preprint arXiv:1208.6231 (2012).

[8] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, W.-K. Ma, Hyperspectral Super-Resolution: A Coupled Tensor Factorization Approach, IEEE Trans. Signal Process. 66 (24) (2018) 6503–6517.

[9] K. Naskovska, A. A. Korobkov, M. Haardt, J. Haueisen, Analysis of the photic driving effect via joint EEG and MEG data processing based on the coupled CP decomposition, in: 2017 25th EUSIPCO, IEEE, 2017, pp. 1285–1289.

[10] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, M. Akçakaya, Tensor Completion from Regular Sub-Nyquist Samples, arXiv e-prints (2019) arXiv:1903.00435.

[11] I. A. Godtliebsen, B. Thomsen, O. Christiansen, Tensor decomposition and vibrational coupled cluster theory, The Journal of Physical Chemistry A 117 (32) (2013) 7267–7279.

[12] C. I. Kanatsoulis, N. D. Sidiropoulos, Tex-graph: Coupled tensor-matrix knowledge-graph embedding for covid-19 drug repurposing, arXiv preprint arXiv:2010.11367 (2020).

[13] C. Prévost, K. Usevich, P. Comon, D. Brie, Hyperspectral Super-Resolution with Coupled Tucker Approximation: Identifiability and SVD-based algorithms, IEEE Transactions on Signal Processing 68 (2020) 931–946.

[14] G. Zhang, X. Fu, K. Huang, J. Wang, Hyperspectral super-resolution: A coupled nonnegative block-term tensor decomposition approach, in: IEEE CAMSAP, 2019, event-place: Guadeloupe, West Indies.

[15] H. Cramér, Mathematical Methods of Statistics, Univ. Press, Princeton, 1946.

[16] C. R. Rao, Information and accuracy attainable in the estimation of statistical parameters, Bull. Calcutta Math. Soc 37 (1945) 81–91.

[17] M. Fréchet, Sur l'extension de certaines évaluations statistiques au cas de petits échantillons, Rev. Int. Stat 11 (1943) 182–205.

[18] T. L. Marzetta, A simple derivation of the constrained multiple parameter Cramér-Rao bound, IEEE Trans. on SP 41 (6) (1993) 2247–2249.

[19] P. Stoica, T. L. Marzetta, Parameter estimation problems with singular information matrices, IEEE Trans. on SP 49 (1) (2001) 87–90.

[20] P. Stoica, B. C. Ng, On the Cramér-Rao bound under parametric constraints, IEEE SP Letters 5 (7) (1998) 177–179.

[21] T. J. Moore, B. M. Sadler, R. J. Kozick, Maximum-Likelihood Estimation, the Cramér-Rao Bound, and the Method of Scoring With Parameter Constraints, IEEE Trans. on SP 56 (3) (2008) 895–908.

[22] T. Menni, E. Chaumette, P. Larzabal, J. P. Barbot, New results on Deterministic Cramér-Rao bounds for real and complex parameters, IEEE Trans. on SP 60 (3) (2012) 1032–1049.

[23] T. Menni, J. Galy, E. Chaumette, P. Larzabal, Versatility of Constrained CRB for System Analysis and Design, IEEE Trans. on AES 50 (3) (2014) 1841–1863.

[24] S. Sahnoun, P. Comon, Joint source estimation and localization, IEEE Trans. Signal Process. 63 (10) (2015) 2485–2595.

[25] X. Liu, N. Sidiropoulos, Cramér-Rao lower bounds for low-rank decomposition of multidimensionnal arrays, IEEE Trans. Signal Process. 49 (9) (2001) 2074–2086.

[26] M. Boizard, R. Boyer, G. Favier, J. Cohen, P. Comon, Performance estimation for tensor CP decomposition with structured factors, in: Proc. ICASSP, 2015.

[27] R. Cabral Farias, J. Cohen, P. Comon, Exploring multimodal data fusion through joint decompositions with flexible couplings, IEEE Trans. Signal Process. 64 (18) (2016) 4830–4844.

[28] C. Ren, R. Cabral Farias, P.-O. Amblard, P. Comon, Performance bounds for coupled models, in: Proc. 2016 IEEE SAM, 2016, event-place: Rio de Janeiro, Brazil.

[29] J. Gorman, A. Hero, Lower bounds for parametric estimation with constraints, IEEE Trans. Inf. Theory 36 (6) (1990) 1285–1301.

[30] C. Prévost, K. Usevich, M. Haardt, P. Comon, D. Brie, Performance bounds for coupled CP model in the framework of hyperspectral super-resolution, in: Proc. IEEE CAMSAP, 2019.

[31] C. Prévost, E. Chaumette, K. Usevich, D. Brie, P. Comon, On Cramér-Rao lower bounds with random equality constraints, in: Proc. IEEE ICASSP, 2020, pp. 5355–5359.

[32] P. Comon, Tensors: A brief introduction, IEEE Signal Process. Mag. 31 (3) (2014) 44–53.

[33] T. Kolda, B. Bader, Tensor Decompositions and Applications, SIAM Review 51 (3) (2009) 455–500.

[34] J. Kruskal, Three-way arrays: Rank and uniqueness of trilinear decompositions, Lin. Alg. Appl. 18 (2) (1977) 95–138.

[35] R. Harshman, M. Lundy, Data preprocessing and the extended PARAFAC model, Research Methods for Multimode Data Analysis, H. G. Law, C. W. Snyder, J. Hattie, and R. K. McDonald, eds., Praeger Publishers, New York (1984) 216–284.

[36] J. Kruskal, Rank, decomposition, and uniqueness for 3-way and n-way arrays, Multiway Data Analysis, R. Coppi and S. Bolasco, eds., North-Holland, Amsterdam 18 (2) (1989) 7–18.

[37] L. Chiantini, G. Ottaviani, On generic identifiability of 3-tensors of small rank, SIAM Journal on Matrix Analysis and Applications 33 (3) (2012) 1018–1037.

[38] L. Chiantini, G. Ottaviani, N. Vannieuwenhoven, An algorithm for generic and low-rank specific identifiability of complex tensors, SIAM Journal on Matrix Analysis and Applications 35 (4) (2014) 1265–1287.

[39] R. Bro, Multi-way analysis in the food industry: Models, algorithms, and applications, PhD Thesis, University of Amsterdam, The Netherlands (1988).

[40] C. Jauffret, Observability and Fisher information matrix in nonlinear regression, IEEE Trans. on Aerospace and Electronic Systems 43 (2) (2007) 756–759.

[41] T. Soderstrom, Comments on "order assumption and singularity of information matrix for pulse transfer function models", IEEE Trans. on Autom. Control 20 (3) (1975) 445–447.

[42] E. Lehmann, G. Casella, Theory of Point Estimation (2nd ed.), Springer, 1998.

[43] D. Slepian, Estimation of signal parameters in the presence of noise, Trans. IRE Professional Group Inf. Theory 3 (3) (1954) 68–69.

[44] Y. Qi, P. Comon, L.-H. Lim, Semialgebraic Geometry of Nonnegative Tensor Rank, SIAM J. Matrix Anal. Appl. 37 (4) (2016) 1556–1580. doi:10.1137/16M1063708.

[45] P. Comon, Y. Qi, K. Usevich, Identifiability of an x-rank decomposition of polynomial maps, SIAM Journal on Applied Algebra and Geometry 1 (1) (2017) 388–414. doi:10.1137/16M1108388.

[46] V. Strassen, Rank and optimal computation of generic tensors, Linear Algebra and its Applications 52-53 (1983) 645–685. doi:https://doi.org/10.1016/0024-3795(83)80041-X.

[47] S. Friedland, On the generic and typical ranks of 3-tensors, Linear algebra and its applications 436 (3) (2012) 478–497.

[48] G. Seber, Matrix Handbook for Statisticians, Wiley Series in Probability and Statistics, 2008.

[49] N. Vervliet, O. Debals, L. Sorber, M. V. Barel, L. D. Lathauwer, Tensorlab 3.0, Available online, 2016.

[50] J. Munkres, Algorithms for the assignment and transportation problems, Journal of the society for industrial and applied mathematics 5 (1) (1957) 32–38.

[51] E. Sobhani, P. Comon, C. Jutten, M. Babaie-Zadeh, Corrindex: a permutation invariant performance index, Signal ProcessingHal-03230210. To appear (2021).

[52] Y. C. Eldar, Minimum variance in biased estimation: Bounds and asymptotically optimal estimators, IEEE Transactions on Signal Processing 52 (7) (2004) 1915–1930.

[53] A. O. Hero, A Cramér-Rao type lower bound for essentially unbiased parameter estimation, Tech. rep., Massachussets Inst. of tech., Lexington Lincoln Lab (1992).

[54] A. O. Hero, J. A. Fessler, M. Usman, Exploring estimator bias-variance tradeoffs using the uniform CR bound, IEEE Transactions on Signal Processing 44 (8) (1996) 2026–2041.

[55] E. Nitzan, T. Routtenberg, J. Tabrikian, Cramér-rao bound for constrained parameter estimation using Lehmann-unbiasedness, IEEE Trans. on Signal Process. 67 (3) (2018) 753–768.

[56] L. Wald, T. Ranchin, M. Mangolini, Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images, Photogrammetric Eng. and Remote Sens. 63 (6) (1997) 691–699.

Supplementary materials

Clémence Prévost, Konstantin Usevich, Martin Haardt, Pierre Comon, David Brie

This document contains supplementary materials regarding the manuscript "Constrained Cramér-Rao bounds for reconstruction problems formulated as coupled canonical polyadic decompositions". We first provide closed-form expressions for the matrices to be inverted to obtain the bounds. We also provide additional simulations regarding the influence of the tensor rank on the performance bounds on the reconstruction error of the target tensor.

1. Uncoupled CRB

We first recall the results of [1] regarding the uncoupled FIM. In practice, the FIM for $\widetilde{\boldsymbol{\omega}}$ (namely $\boldsymbol{F}(\widetilde{\boldsymbol{\omega}})$) is computed by applying the Slepian-Bangs formula to the tensors $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$:

$$\boldsymbol{F}(\widetilde{\boldsymbol{\omega}}) = \begin{bmatrix} \frac{\partial \boldsymbol{\mu}_1(\widetilde{\boldsymbol{\omega}})}{\partial \widetilde{\boldsymbol{\omega}}^\mathsf{T}} \\ \frac{\partial \boldsymbol{\mu}_2(\widetilde{\boldsymbol{\omega}})}{\partial \widetilde{\boldsymbol{\omega}}^\mathsf{T}} \end{bmatrix}^\mathsf{T} \mathrm{Diag}\{\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2\}^{-1} \begin{bmatrix} \frac{\partial \boldsymbol{\mu}_1(\widetilde{\boldsymbol{\omega}})}{\partial \widetilde{\boldsymbol{\omega}}^\mathsf{T}} \\ \frac{\partial \boldsymbol{\mu}_2(\widetilde{\boldsymbol{\omega}})}{\partial \widetilde{\boldsymbol{\omega}}^\mathsf{T}} \end{bmatrix} = \frac{1}{\sigma_1^2} \left[\frac{\partial \boldsymbol{\mu}_1^\mathsf{T}(\widetilde{\boldsymbol{\omega}})}{\partial \widetilde{\boldsymbol{\omega}}} \frac{\partial \boldsymbol{\mu}_1(\widetilde{\boldsymbol{\omega}})}{\partial \widetilde{\boldsymbol{\omega}}^\mathsf{T}} \right] + \frac{1}{\sigma_2^2} \left[\frac{\partial \boldsymbol{\mu}_2^\mathsf{T}(\widetilde{\boldsymbol{\omega}})}{\partial \widetilde{\boldsymbol{\omega}}} \frac{\partial \boldsymbol{\mu}_2(\widetilde{\boldsymbol{\omega}})}{\partial \widetilde{\boldsymbol{\omega}}^\mathsf{T}} \right]. \tag{1}$$

The expressions of the functions $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are obtained using relationships between tensor unfoldings. For $i = 1, 2$, we have

$$\boldsymbol{\mu}_i(\widetilde{\boldsymbol{\omega}}) = \underbrace{[(\boldsymbol{C}_i \odot \boldsymbol{B}_i) \boxtimes \boldsymbol{I}]}_{\boldsymbol{S}_{A_i}} \mathrm{vec}\{\boldsymbol{A}_i\} = \underbrace{\boldsymbol{\Pi}_i^{(2,1)} [(\boldsymbol{C}_i \odot \boldsymbol{A}_i) \boxtimes \boldsymbol{I}]}_{\boldsymbol{S}_{B_i}} \mathrm{vec}\{\boldsymbol{B}_i\} = \underbrace{\boldsymbol{\Pi}_i^{(3,1)} [(\boldsymbol{B}_i \odot \boldsymbol{A}_i) \boxtimes \boldsymbol{I}]}_{\boldsymbol{S}_{C_i}} \mathrm{vec}\{\boldsymbol{C}_i\}. \tag{2}$$

This yields $\frac{\partial \boldsymbol{\mu}_i(\widetilde{\boldsymbol{\omega}})}{\partial \widetilde{\boldsymbol{\omega}}^\mathsf{T}} = [\boldsymbol{S}_{A_i}\ \boldsymbol{S}_{B_i}\ \boldsymbol{S}_{C_i}]$.

Developing (1) using the above formula yields

$$\mathbf{F}(\widetilde{\boldsymbol{\omega}}) = \begin{bmatrix} \frac{1}{\sigma_1^2}\boldsymbol{S}_{A_1}^\mathsf{T}\boldsymbol{S}_{A_1} & \frac{1}{\sigma_1^2}\boldsymbol{S}_{A_1}^\mathsf{T}\boldsymbol{S}_{B_1} & \frac{1}{\sigma_1^2}\boldsymbol{S}_{A_1}^\mathsf{T}\boldsymbol{S}_{C_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \frac{1}{\sigma_1^2}\boldsymbol{S}_{B_1}^\mathsf{T}\boldsymbol{S}_{A_1} & \frac{1}{\sigma_1^2}\boldsymbol{S}_{B_1}^\mathsf{T}\boldsymbol{S}_{B_1} & \frac{1}{\sigma_1^2}\boldsymbol{S}_{B_1}^\mathsf{T}\boldsymbol{S}_{C_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \frac{1}{\sigma_1^2}\boldsymbol{S}_{C_1}^\mathsf{T}\boldsymbol{S}_{A_1} & \frac{1}{\sigma_1^2}\boldsymbol{S}_{C_1}^\mathsf{T}\boldsymbol{S}_{B_1} & \frac{1}{\sigma_1^2}\boldsymbol{S}_{C_1}^\mathsf{T}\boldsymbol{S}_{C_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{\sigma_2^2}\boldsymbol{S}_{A_2}^\mathsf{T}\boldsymbol{S}_{A_2} & \frac{1}{\sigma_2^2}\boldsymbol{S}_{A_2}^\mathsf{T}\boldsymbol{S}_{B_2} & \frac{1}{\sigma_2^2}\boldsymbol{S}_{A_2}^\mathsf{T}\boldsymbol{S}_{C_2} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{\sigma_2^2}\boldsymbol{S}_{B_2}^\mathsf{T}\boldsymbol{S}_{A_2} & \frac{1}{\sigma_2^2}\boldsymbol{S}_{B_2}^\mathsf{T}\boldsymbol{S}_{B_2} & \frac{1}{\sigma_2^2}\boldsymbol{S}_{B_2}^\mathsf{T}\boldsymbol{S}_{C_2} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{\sigma_2^2}\boldsymbol{S}_{C_2}^\mathsf{T}\boldsymbol{S}_{A_2} & \frac{1}{\sigma_2^2}\boldsymbol{S}_{C_2}^\mathsf{T}\boldsymbol{S}_{B_2} & \frac{1}{\sigma_2^2}\boldsymbol{S}_{C_2}^\mathsf{T}\boldsymbol{S}_{C_2} \end{bmatrix}. \tag{3}$$

In (3), $\mathbf{F}(\widetilde{\boldsymbol{\omega}})$ is a block-matrix of the form

$$\mathbf{F}(\widetilde{\boldsymbol{\omega}}) = \begin{bmatrix} \mathbf{F}(\widetilde{\boldsymbol{\phi}}_1) & \mathbf{F}(\widetilde{\boldsymbol{\phi}}_1, \boldsymbol{\theta}_1) & \mathbf{0} & \mathbf{0} \\ \mathbf{F}(\widetilde{\boldsymbol{\phi}}_1, \boldsymbol{\theta}_1)^\mathsf{T} & \mathbf{F}(\boldsymbol{\theta}_1) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{F}(\widetilde{\boldsymbol{\phi}}_2) & \mathbf{F}(\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_2) \\ \mathbf{0} & \mathbf{0} & \mathbf{F}(\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_2)^\mathsf{T} & \mathbf{F}(\boldsymbol{\theta}_2) \end{bmatrix}, \tag{4}$$

where for $i \in \{1,2\}$, we have

$$\mathbf{F}(\widetilde{\boldsymbol{\phi}}_i) = \begin{bmatrix} \frac{1}{\sigma_i^2}\boldsymbol{S}_{A_i}^{\mathsf{T}}\boldsymbol{S}_{A_i} & \frac{1}{\sigma_i^2}\boldsymbol{S}_{A_i}^{\mathsf{T}}\boldsymbol{S}_{B_i} \\ \frac{1}{\sigma_i^2}\boldsymbol{S}_{B_i}^{\mathsf{T}}\boldsymbol{S}_{A_i} & \frac{1}{\sigma_i^2}\boldsymbol{S}_{B_i}^{\mathsf{T}}\boldsymbol{S}_{B_i} \end{bmatrix}, \quad \mathbf{F}(\widetilde{\boldsymbol{\phi}}_i, \boldsymbol{\theta}_i) = \begin{bmatrix} \frac{1}{\sigma_i^2}\boldsymbol{S}_{A_i}^{\mathsf{T}}\boldsymbol{S}_{C_i} \\ \frac{1}{\sigma_i^2}\boldsymbol{S}_{B_i}^{\mathsf{T}}\boldsymbol{S}_{C_i} \end{bmatrix}, \quad \mathbf{F}(\boldsymbol{\theta}_i) = \frac{1}{\sigma_i^2}\boldsymbol{S}_{C_i}^{\mathsf{T}}\boldsymbol{S}_{C_i}.$$

In most performance analyses, we are only interested in the diagonal terms of the CRB, which are directly related to the optimal MSE. For $i \in \{1,2\}$, denote $\mathbf{CRB}(\widetilde{\boldsymbol{\phi}}_i)$ and $\mathbf{CRB}(\boldsymbol{\theta}_i)$ the diagonal blocks of the matrix $\mathbf{CRB}(\widetilde{\boldsymbol{\omega}}) = \mathbf{F}(\widetilde{\boldsymbol{\omega}})^{-1}$. Then from (4),

$$\mathbf{CRB}(\widetilde{\boldsymbol{\phi}}_i) = \left(\mathbf{F}(\widetilde{\boldsymbol{\phi}}_i) - \mathbf{F}(\widetilde{\boldsymbol{\phi}}_i, \boldsymbol{\theta}_i)\mathbf{F}(\boldsymbol{\theta}_i)^{-1}\mathbf{F}(\widetilde{\boldsymbol{\phi}}_i, \boldsymbol{\theta}_i)^{\mathsf{T}}\right)^{-1}, \tag{5}$$

$$\mathbf{CRB}(\boldsymbol{\theta}_i) = \left(\mathbf{F}(\boldsymbol{\theta}_i) - \mathbf{F}(\widetilde{\boldsymbol{\phi}}_i, \boldsymbol{\theta}_i)^{\mathsf{T}}\mathbf{F}(\widetilde{\boldsymbol{\phi}}_i)^{-1}\mathbf{F}(\widetilde{\boldsymbol{\phi}}_i, \boldsymbol{\theta}_i)\right)^{-1}. \tag{6}$$

2. Blind-CCRB for partially-coupled models

For the partially-coupled model, we have a matrix $\boldsymbol{U}$ such that

$$\boldsymbol{U} = \begin{bmatrix} \boldsymbol{I}_{(I+J+I_H+J_H)R-4R} \\ \begin{bmatrix} \mathbf{0} & \boldsymbol{Z}_1 & [\boldsymbol{Z}_2 \ \boldsymbol{Z}_3] \end{bmatrix} \end{bmatrix}.$$

Thus the matrix $\boldsymbol{U}^{\mathsf{T}}\boldsymbol{F}\boldsymbol{U}$ is a block matrix of the form

$$\boldsymbol{U}^{\mathsf{T}}\boldsymbol{F}\boldsymbol{U} = \begin{bmatrix} \boldsymbol{D}_{1,1} & \boldsymbol{D}_{1,2} & \mathbf{0} \\ \boldsymbol{D}_{1,2}^{\mathsf{T}} & \boldsymbol{D}_{2,2} & \boldsymbol{D}_{2,3} \\ \mathbf{0} & \boldsymbol{D}_{2,3}^{\mathsf{T}} & \boldsymbol{D}_{3,3} \end{bmatrix},$$

which subblocks are such that

$$\begin{aligned}
&\boldsymbol{D}_{1,1} = \boldsymbol{F}(\widetilde{\boldsymbol{\phi}}_1), \quad \boldsymbol{D}_{1,2} = \mathbf{F}(\widetilde{\boldsymbol{\phi}}_1, \boldsymbol{\theta}_1), \quad \boldsymbol{D}_{1,2} = \boldsymbol{F}(\boldsymbol{\theta}_1) + \boldsymbol{Z}_1^{\mathsf{T}}\boldsymbol{F}(\boldsymbol{\theta}_2)\boldsymbol{Z}_1, \\
&\boldsymbol{D}_{2,3} = \boldsymbol{Z}_1^{\mathsf{T}}\boldsymbol{F}(\widetilde{\boldsymbol{\phi}}_2) + \boldsymbol{Z}_1^{\mathsf{T}}\mathbf{F}(\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_2)\left[\boldsymbol{Z}_2 \ \boldsymbol{Z}_3\right], \\
&\boldsymbol{D}_{3,3} = \boldsymbol{F}(\widetilde{\boldsymbol{\phi}}_2) + \mathbf{F}(\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_2)\left[\boldsymbol{Z}_2 \ \boldsymbol{Z}_3\right] + \left[\boldsymbol{Z}_2 \ \boldsymbol{Z}_3\right]^{\mathsf{T}}\mathbf{F}(\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_2)^{\mathsf{T}} + \left[\boldsymbol{Z}_2 \ \boldsymbol{Z}_3\right]^{\mathsf{T}}\boldsymbol{F}(\boldsymbol{\theta}_2)\left[\boldsymbol{Z}_2 \ \boldsymbol{Z}_3\right].
\end{aligned}$$

Denote $\widetilde{\boldsymbol{D}}_{1,1}$, $\widetilde{\boldsymbol{D}}_{1,2}$, $\widetilde{\boldsymbol{D}}_{2,2}$, $\widetilde{\boldsymbol{D}}_{2,3}$, $\widetilde{\boldsymbol{D}}_{3,3}$ the blocks of $(\boldsymbol{U}^{\mathsf{T}}\boldsymbol{F}\boldsymbol{U})^{-1}$ obtained by the block-inversion lemma. Then the diagonal blocks of $\mathbf{Blind\text{-}CCRB}(\widetilde{\boldsymbol{\omega}})$ for the partially-coupled model are:

$$\begin{aligned}
&\mathbf{Blind\text{-}CCRB}(\widetilde{\boldsymbol{\phi}}_1) = \widetilde{\boldsymbol{D}}_{1,1}, \\
&\mathbf{Blind\text{-}CCRB}(\boldsymbol{\theta}_1) = \widetilde{\boldsymbol{D}}_{2,2}, \\
&\mathbf{Blind\text{-}CCRB}(\widetilde{\boldsymbol{\phi}}_2) = \widetilde{\boldsymbol{D}}_{3,3}, \\
&\mathbf{Blind\text{-}CCRB}(\boldsymbol{\theta}_2) = \boldsymbol{Z}_1\widetilde{\boldsymbol{D}}_{2,2}\boldsymbol{Z}_1^{\mathsf{T}} + \boldsymbol{Z}_1\widetilde{\boldsymbol{D}}_{2,3}\left[\boldsymbol{Z}_2 \ \boldsymbol{Z}_3\right]^{\mathsf{T}} + \left[\boldsymbol{Z}_2 \ \boldsymbol{Z}_3\right]\widetilde{\boldsymbol{D}}_{2,3}^{\mathsf{T}}\boldsymbol{Z}_1^{\mathsf{T}} + \left[\boldsymbol{Z}_2 \ \boldsymbol{Z}_3\right]\widetilde{\boldsymbol{D}}_{3,3}\left[\boldsymbol{Z}_2 \ \boldsymbol{Z}_3\right].
\end{aligned}$$

3. Fully-coupled CCRB

3.1. Scenario 1 with linear constraints

3.1.1. Standard CCRB

For the fully-coupled model, the matrix $\boldsymbol{U}^{\mathsf{T}}\boldsymbol{F}\boldsymbol{U}$ in is a block matrix of the form

$$\boldsymbol{U}^{\mathsf{T}}\boldsymbol{F}\boldsymbol{U} = \begin{bmatrix} \boldsymbol{D}_{1,1} & \boldsymbol{D}_{1,2} \\ \boldsymbol{D}_{1,2}^{\mathsf{T}} & \boldsymbol{D}_{2,2} \end{bmatrix},$$

which subblocks are such that

$$\begin{aligned}
\boldsymbol{D}_{1,1} &= \mathrm{Diag}\{\boldsymbol{I}\boxtimes\boldsymbol{P},\boldsymbol{I}\boxtimes\boldsymbol{Q}\}^{\mathsf{T}}\mathbf{F}(\widetilde{\boldsymbol{\phi}}_1)\,\mathrm{Diag}\{\boldsymbol{I}\boxtimes\boldsymbol{P},\boldsymbol{I}\boxtimes\boldsymbol{Q}\}+\mathbf{F}(\widetilde{\boldsymbol{\phi}}_2);\\
\boldsymbol{D}_{1,2} &= \mathrm{Diag}\{\boldsymbol{I}\boxtimes\boldsymbol{P},\boldsymbol{I}\boxtimes\boldsymbol{Q}\}^{\mathsf{T}}\mathbf{F}(\widetilde{\boldsymbol{\phi}}_1,\boldsymbol{\theta}_1)+\mathbf{F}(\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_2)(\boldsymbol{I}\boxtimes\boldsymbol{R}),\\
\boldsymbol{D}_{2,2} &= \mathbf{F}(\boldsymbol{\theta}_1)+(\boldsymbol{I}\boxtimes\boldsymbol{R})^{\mathsf{T}}\mathbf{F}(\boldsymbol{\theta}_2)(\boldsymbol{I}\boxtimes\boldsymbol{R}).
\end{aligned}$$

Denote $\widetilde{\boldsymbol{D}}_{1,1}$, $\widetilde{\boldsymbol{D}}_{1,2}$, $\widetilde{\boldsymbol{D}}_{2,2}$ the blocks of $(\boldsymbol{U}^{\mathsf{T}}\boldsymbol{F}\boldsymbol{U})^{-1}$ obtained by the block-inversion lemma as

$$\begin{aligned}
\widetilde{\boldsymbol{D}}_{1,1} &= \left(\boldsymbol{D}_{1,1}-\boldsymbol{D}_{1,2}\boldsymbol{D}_{2,2}^{-1}\boldsymbol{D}_{1,2}^{\mathsf{T}}\right)^{-1},\\
\widetilde{\boldsymbol{D}}_{1,2} &= -\widetilde{\boldsymbol{D}}_{1,1}\boldsymbol{D}_{1,2}\boldsymbol{D}_{2,2}^{-1},\\
\widetilde{\boldsymbol{D}}_{2,2} &= \left(\boldsymbol{D}_{2,2}-\boldsymbol{D}_{1,2}^{\mathsf{T}}\boldsymbol{D}_{1,1}^{-1}\boldsymbol{D}_{1,2}\right)^{-1}.
\end{aligned}$$

Then the diagonal blocks of $\mathbf{CCRB}(\widetilde{\boldsymbol{\omega}})$ for the fully-coupled model are:

$$\begin{aligned}
\mathbf{CCRB}(\widetilde{\boldsymbol{\phi}}_1) &= \mathrm{Diag}\{\boldsymbol{I}\boxtimes\boldsymbol{P},\boldsymbol{I}\boxtimes\boldsymbol{Q}\}\widetilde{\boldsymbol{D}}_{1,1}\,\mathrm{Diag}\{\boldsymbol{I}\boxtimes\boldsymbol{P},\boldsymbol{I}\boxtimes\boldsymbol{Q}\}^{\mathsf{T}},\\
\mathbf{CCRB}(\boldsymbol{\theta}_1) &= \widetilde{\boldsymbol{D}}_{2,2},\\
\mathbf{CCRB}(\widetilde{\boldsymbol{\phi}}_2) &= \widetilde{\boldsymbol{D}}_{1,1},\\
\mathbf{CCRB}(\boldsymbol{\theta}_2) &= (\boldsymbol{I}\boxtimes\boldsymbol{R})\widetilde{\boldsymbol{D}}_{2,2}(\boldsymbol{I}\boxtimes\boldsymbol{R})^{\mathsf{T}}.
\end{aligned}$$

3.1.2. Reparameterization change

In fact, the reparameterized FIM $\boldsymbol{F}_{\boldsymbol{c}}(\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_1)$ is a block-matrix of the form

$$\boldsymbol{F}_{\boldsymbol{c}}(\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_1)=\begin{bmatrix}\boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2} & \boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_1}\\ \boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_1}^{\mathsf{T}} & \boldsymbol{F}_{\boldsymbol{\theta}_1}\end{bmatrix},$$

with

$$\boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2}=\begin{bmatrix}\boldsymbol{S}_1^{\mathsf{T}}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_1 & \boldsymbol{S}_1^{\mathsf{T}}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_2\\ \boldsymbol{S}_2^{\mathsf{T}}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_1 & \boldsymbol{S}_2^{\mathsf{T}}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_2\end{bmatrix},\quad \boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_1}=\begin{bmatrix}\boldsymbol{S}_1^{\mathsf{T}}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_3\\ \boldsymbol{S}_2^{\mathsf{T}}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_3\end{bmatrix},\quad \boldsymbol{F}_{\boldsymbol{\theta}_1}=\boldsymbol{S}_3^{\mathsf{T}}\widetilde{\boldsymbol{\mathcal{P}}}\boldsymbol{S}_3,$$

where $\widetilde{\boldsymbol{\mathcal{P}}}=\frac{1}{\sigma_1^2}(\boldsymbol{I}\boxtimes\boldsymbol{Q}^{\mathsf{T}}\boldsymbol{Q}\boxtimes\boldsymbol{P}^{\mathsf{T}}\boldsymbol{P})+\frac{1}{\sigma_2^2}(\boldsymbol{R}^{\mathsf{T}}\boldsymbol{R}\boxtimes\boldsymbol{I})$.

Denote $\mathbf{CRB_c}(\widetilde{\boldsymbol{\phi}}_2)$ and $\mathbf{CRB_c}(\boldsymbol{\theta}_1)$ the diagonal blocks of the matrix $\mathbf{F_c}(\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_1)^{-1}$. Then we have

$$\mathbf{CRB}(\widetilde{\boldsymbol{\phi}}_2)=\left(\boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2}-\boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_1}\boldsymbol{F}_{\boldsymbol{\theta}_1}^{-1}\boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_1}^{\mathsf{T}}\right)^{-1},\tag{7}$$

$$\mathbf{CRB}(\boldsymbol{\theta}_1)=\left(\boldsymbol{F}_{\boldsymbol{\theta}_1}-\boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_1}^{\mathsf{T}}\boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2}^{-1}\boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2,\boldsymbol{\theta}_1}\right)^{-1}.\tag{8}$$

3.2. Scenario 2 with non-linear constraints

3.2.1. Standard CCRB

In this scenario, the matrix $\boldsymbol{U}^{\mathsf{T}}\boldsymbol{F}\boldsymbol{U}$ is a block matrix of the form

$$\boldsymbol{U}^{\mathsf{T}}\boldsymbol{F}\boldsymbol{U}=\begin{bmatrix}\boldsymbol{D}_{1,1} & \boldsymbol{D}_{1,2}\\ \boldsymbol{D}_{1,2}^{\mathsf{T}} & \boldsymbol{D}_{2,2}\end{bmatrix},$$

which subblocks are such that

$$\begin{aligned}
\boldsymbol{D}_{1,1} &= \mathrm{Diag}\{\boldsymbol{Z}_4, \boldsymbol{Z}_5\}^{\mathsf{T}} \mathbf{F}(\widetilde{\boldsymbol{\phi}}_1)\, \mathrm{Diag}\{\boldsymbol{Z}_4, \boldsymbol{Z}_5\} + \mathbf{F}(\widetilde{\boldsymbol{\phi}}_2) + [\boldsymbol{Z}_2\ \boldsymbol{Z}_3]^{\mathsf{T}}\, \mathbf{F}(\boldsymbol{\theta}_2)\, [\boldsymbol{Z}_2\ \boldsymbol{Z}_3] \\
&\quad + \mathbf{F}(\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_2)\, [\boldsymbol{Z}_2\ \boldsymbol{Z}_3] + [\boldsymbol{Z}_2\ \boldsymbol{Z}_3]^{\mathsf{T}}\, \mathbf{F}(\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_2)^{\mathsf{T}}, \\
\boldsymbol{D}_{1,2} &= \mathrm{Diag}\{\boldsymbol{Z}_4, \boldsymbol{Z}_5\}^{\mathsf{T}} \mathbf{F}(\widetilde{\boldsymbol{\phi}}_1, \boldsymbol{\theta}_1) + \mathbf{F}(\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_2) \boldsymbol{Z}_1 + [\boldsymbol{Z}_2\ \boldsymbol{Z}_3]^{\mathsf{T}}\, \mathbf{F}(\boldsymbol{\theta}_2) \boldsymbol{Z}_1, \\
\boldsymbol{D}_{2,2} &= \mathbf{F}(\boldsymbol{\theta}_1) + \boldsymbol{Z}_1^{\mathsf{T}} \mathbf{F}(\boldsymbol{\theta}_2) \boldsymbol{Z}_1.
\end{aligned}$$

Denote $\widetilde{\boldsymbol{D}}_{1,1}$, $\widetilde{\boldsymbol{D}}_{1,2}$, $\widetilde{\boldsymbol{D}}_{2,2}$ the blocks of $(\boldsymbol{U}^{\mathsf{T}} \boldsymbol{F} \boldsymbol{U})^{-1}$ obtained by the block-inversion lemma as

$$\begin{aligned}
\widetilde{\boldsymbol{D}}_{1,1} &= \left(\boldsymbol{D}_{1,1} - \boldsymbol{D}_{1,2} \boldsymbol{D}_{2,2}^{-1} \boldsymbol{D}_{1,2}^{\mathsf{T}} \right)^{-1}, \\
\widetilde{\boldsymbol{D}}_{1,2} &= -\widetilde{\boldsymbol{D}}_{1,1} \boldsymbol{D}_{1,2} \boldsymbol{D}_{2,2}^{-1}, \\
\widetilde{\boldsymbol{D}}_{2,2} &= \left(\boldsymbol{D}_{2,2} - \boldsymbol{D}_{1,2}^{\mathsf{T}} \boldsymbol{D}_{1,1}^{-1} \boldsymbol{D}_{1,2} \right)^{-1}.
\end{aligned}$$

Then the diagonal blocks of $\mathbf{CCRB}(\widetilde{\boldsymbol{\omega}})$ for the fully-coupled model are:

$$\begin{aligned}
\mathbf{CCRB}(\widetilde{\boldsymbol{\phi}}_1) &= \mathrm{Diag}\{\boldsymbol{Z}_4, \boldsymbol{Z}_5\} \widetilde{\boldsymbol{D}}_{1,1}\, \mathrm{Diag}\{\boldsymbol{Z}_4, \boldsymbol{Z}_5\}^{\mathsf{T}}, \\
\mathbf{CCRB}(\boldsymbol{\theta}_1) &= \widetilde{\boldsymbol{D}}_{2,2}, \\
\mathbf{CCRB}(\widetilde{\boldsymbol{\phi}}_2) &= \widetilde{\boldsymbol{D}}_{1,1}, \\
\mathbf{CCRB}(\boldsymbol{\theta}_2) &= [\boldsymbol{Z}_2\ \boldsymbol{Z}_3]\, \widetilde{\boldsymbol{D}}_{1,1}\, [\boldsymbol{Z}_2\ \boldsymbol{Z}_3]^{\mathsf{T}} + \boldsymbol{Z}_1 \widetilde{\boldsymbol{D}}_{2,2} \boldsymbol{Z}_1^{\mathsf{T}} + \boldsymbol{Z}_1 \widetilde{\boldsymbol{D}}_{1,2}^{\mathsf{T}}\, [\boldsymbol{Z}_2\ \boldsymbol{Z}_3]^{\mathsf{T}} + [\boldsymbol{Z}_2\ \boldsymbol{Z}_3]\, \widetilde{\boldsymbol{D}}_{1,2} \boldsymbol{Z}_1^{\mathsf{T}}.
\end{aligned}$$

3.2.2. Reparameterization change

The reparameterized FIM $\boldsymbol{F}_{\boldsymbol{c}}(\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_1)$ is a block-matrix of the form

$$\boldsymbol{F}_{\boldsymbol{c}}(\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_1) = \begin{bmatrix} \boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2} & \boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_1} \\ \boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_1}^{\mathsf{T}} & \boldsymbol{F}_{\boldsymbol{\theta}_1} \end{bmatrix},$$

with

$$\begin{aligned}
\boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2} &= \begin{bmatrix} \frac{1}{\sigma_1^2} \boldsymbol{X}_1^{\mathsf{T}} \boldsymbol{X}_1 + \frac{1}{\sigma_2^2} \boldsymbol{X}_5^{\mathsf{T}} \boldsymbol{X}_5 & \frac{1}{\sigma_1^2} \boldsymbol{X}_1^{\mathsf{T}} \boldsymbol{X}_2 + \frac{1}{\sigma_2^2} \boldsymbol{X}_5^{\mathsf{T}} \boldsymbol{X}_6 \\ \frac{1}{\sigma_1^2} \boldsymbol{X}_2^{\mathsf{T}} \boldsymbol{X}_1 + \frac{1}{\sigma_2^2} \boldsymbol{X}_6^{\mathsf{T}} \boldsymbol{X}_5 & \frac{1}{\sigma_1^2} \boldsymbol{X}_2^{\mathsf{T}} \boldsymbol{X}_2 + \frac{1}{\sigma_2^2} \boldsymbol{X}_6^{\mathsf{T}} \boldsymbol{X}_6 \end{bmatrix}, \\
\boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_1} &= \begin{bmatrix} \frac{1}{\sigma_1^2} \boldsymbol{X}_1^{\mathsf{T}} \boldsymbol{X}_3 + \frac{1}{\sigma_2^2} \boldsymbol{X}_5^{\mathsf{T}} \boldsymbol{X}_4 \\ \frac{1}{\sigma_1^2} \boldsymbol{X}_2^{\mathsf{T}} \boldsymbol{X}_3 + \frac{1}{\sigma_2^2} \boldsymbol{X}_6^{\mathsf{T}} \boldsymbol{X}_4 \end{bmatrix}, \\
\boldsymbol{F}_{\boldsymbol{\theta}_1} &= \frac{1}{\sigma_1^2} \boldsymbol{X}_3^{\mathsf{T}} \boldsymbol{X}_3 + \frac{1}{\sigma_2^2} \boldsymbol{X}_4^{\mathsf{T}} \boldsymbol{X}_4.
\end{aligned}$$

Denote $\mathbf{rCRB}(\widetilde{\boldsymbol{\phi}}_2)$ and $\mathbf{rCRB}(\boldsymbol{\theta}_1)$ the diagonal blocks of the matrix $\mathbf{F}_{\mathbf{c}}(\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_1)^{-1}$. Then we have

$$\mathbf{rCRB}(\widetilde{\boldsymbol{\phi}}_2) = \left(\boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2} - \boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_1} \boldsymbol{F}_{\boldsymbol{\theta}_1}^{-1} \boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_1}^{\mathsf{T}} \right)^{-1}, \tag{9}$$

$$\mathbf{rCRB}(\boldsymbol{\theta}_1) = \left(\boldsymbol{F}_{\boldsymbol{\theta}_1} - \boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_1}^{\mathsf{T}} \boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2}^{-1} \boldsymbol{F}_{\widetilde{\boldsymbol{\phi}}_2, \boldsymbol{\theta}_1} \right)^{-1}. \tag{10}$$

4. Choice of the rank

In this section, we investigate the influence of the tensor rank on the modelling error for $\boldsymbol{x}$. We assume that we want to recover a given tensor $\boldsymbol{\mathcal{X}}$ admitting a CPD with rank $R_{\text{th}} = 3$. In real applications, the observed tensors are unlikely to be low-rank tensors. Thus the proposed model only performs a low-rank approximation of the target tensor, and the appropriate tensor rank is not known *a priori*. Nevertheless, we expect the performance for the reconstruction of $\boldsymbol{\mathcal{X}}$ to vary along with the tensor rank R.

We generate the CP model with $I = J = 18$, $I_d = J_d = 6$, $K = 16$ and $K_d = 8$, and ranks $R \in \{3, \ldots, 16\}$. The first columns of $\boldsymbol{A}_2$, $\boldsymbol{B}_2$ are also set to ones. The factors $\boldsymbol{A}_1$, $\boldsymbol{B}_1$, $\boldsymbol{C}_2$ are constructed according to linear constraints, that correspond to the first scenario. The low-resolution tensors $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ are constructed from these augmented CP factors. For $R \in \{3, \ldots, 16\}$, the CCRB for $\boldsymbol{x}$ is averaged over 100 realizations of the factors $\boldsymbol{A}_i$, $\boldsymbol{B}_i$, $\boldsymbol{C}_i$ ($i \in \{1, 2\}$). This bound can be seen as an error bound on the reconstruction of the true tensor $\boldsymbol{\mathcal{X}}$. In Figure 1, we plot the averaged UCCRB as a function of SNR_1 and R for fixed SNR_2.

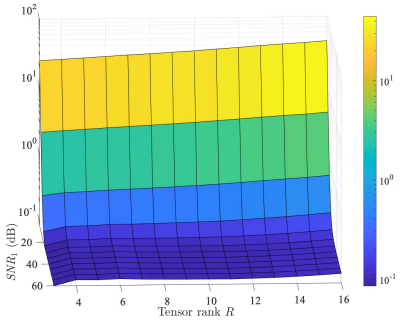

Figure 1: **UCCRB**($\boldsymbol{x}$) as a function of SNR_1 (dB) and tensor rank R.

We can see that, for all R, the UCCRB decreases when SNR_1 decreases. Moreover, for all considered SNRs, the value of the CCRB increases with R; the best theoretical performance is obtained for $R = R_{\text{th}} = 3$. This figure indicates that for low SNR, the performance for the reconstruction of the target tensor is sensitive to an overestimation of the tensor rank.

References

[1] C. Ren, R. Cabral Farias, P.-O. Amblard, P. Comon, Performance bounds for coupled models, in: Proc. 2016 IEEE SAM, 2016, event-place: Rio de Janeiro, Brazil.