



Constrained Cramér-Rao lower bounds for CP-based hyperspectral super-resolution

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

► **To cite this version:**

Clémence Prévost, Konstantin Usevich, Martin Haardt, Pierre Comon, David Brie. Constrained Cramér-Rao lower bounds for CP-based hyperspectral super-resolution. 2020. hal-03083709

HAL Id: hal-03083709

<https://hal.archives-ouvertes.fr/hal-03083709>

Preprint submitted on 19 Dec 2020

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

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

Constrained Cramér-Rao lower bounds for CP-based hyperspectral super-resolution

Clémence Prévost, *Student Member, IEEE*, Konstantin Usevich, *Member, IEEE*, Martin Haardt, *Fellow, IEEE*, Pierre Comon, *Fellow, IEEE*, David Brie, *Member, IEEE*.

Abstract—We propose a theoretical performance analysis for the hyperspectral super-resolution task, formulated as a coupled canonical polyadic decomposition. We introduce two probabilistic scenarios along with different parameterizations, then derive constrained Cramér-Rao lower bounds (CCRB) for the proposed scenarios. We then illustrate the versatility of the CCRB throughout a set of experiments, including its usefulness to design the hyperspectral measurement system. We also assess the relative performance of existing estimators and use the CCRB as a tool to design more efficient algorithms.

Index Terms—hyperspectral super-resolution, multimodal data fusion, coupled tensor decompositions, Cramér-Rao bounds.

I. INTRODUCTION

Hyperspectral super-resolution (HSR) [1] is a problem of growing interest in the signal processing community. It consists in fusing a multispectral image (MSI), which has a good spatial resolution but few spectral bands [2], and a hyperspectral image (HSI), whose spatial resolution is lower than that of the MSI. The aim is to recover a super-resolution image (SRI), which possesses both high spatial and spectral resolutions. This problem lies in the framework of multimodal data fusion [3] between heterogeneous datasets. The datasets have different parametric models with shared variables, linked through possibly non-linear deterministic relations.

Many methods were developed for solving the HSR problem. Early matrix-based approaches include coupled nonnegative matrix factorization [4] (CNMF), methods based on solving Sylvester equations [5], Bayesian approaches (HySure [6]), FUMI [7], to name a few. Most of these methods are based on a coupled low-rank factorization of the matricized hyperspectral and multispectral images.

In [8], a promising tensor-based method was proposed, making use of the inherent 3D nature of spectral images. Assuming that the SRI itself admits a low-rank canonical polyadic (CP) decomposition (CPD), the HSR problem is reformulated as a coupled CP approximation. An alternating least squares (ALS) algorithm called Super-resolution TENSOR REconstruction (STEREO) is proposed, achieving high reconstruction performance. In some cases, the spatial degradation operator is unknown, therefore blind algorithms are needed. A blind version of STEREO was proposed in [8] that also uses an ALS algorithm for a coupled CP model.

This work was partially supported by the ANR (Agence Nationale de Recherche) under grant LeaFleT (ANR-19-CE23-0021).

C. Prévost, K. Usevich and D. Brie are with Centre de Recherche en Automatique de Nancy (CRAN), Université de Lorraine, CNRS, Vandœuvre-lès-Nancy, France. e-mail: firstname.lastname@univ-lorraine.fr.

M. Haardt is with the Communications Research Laboratory, Ilmenau Institute of Technology, Ilmenau, Germany. e-mail: firstname.lastname@tu-ilmenau.de.

P. Comon is with CNRS, GIPSA-Lab, Univ. Grenoble Alpes, F-38000 Grenoble, France. e-mail: firstname.lastname@gipsa-lab.grenoble-inp.fr.

In parallel, Cramér-Rao bounds (CRB) for tensor CP models have been studied in a general context. In [9], [10], [11], performance bounds for uncoupled CP models have been provided. In [12], a Bayesian framework was proposed for flexible coupling models and hybrid Cramér-Rao bounds were derived. Constrained Cramér-Rao bounds (CCRB) for partially coupled complex tensors admitting a CPD and possibly non-linear couplings were explored in [13]. The expression of the bound was based on the work in [14], which considers a specific case where the Fisher information matrix (FIM) for the parameters is invertible.

Only recently the performance bounds have been studied for the HSR problem [15], [16]. Based on the above works, in [15], constrained performance bounds for the estimation of the latent CP factors were provided in a simple case. This work was extended to the case of constraints involving a random parameter in [16].

In this paper, we study performance bounds for coupled CP tensors in the HSR framework: in particular, we adopt a degradation model that is more general than that of [15], and thus, that is more fair to the specific acquisition scenario of HSIs and MSIs. We introduce two probabilistic scenarios, leading to different choices of parameterizations. For each scenario, we derive the CCRB for the model parameters, based on the results in [17], [18], extending that of [14]; in particular, they do not require identifiability of the model at hand. We use the CCRB as a tool to study the impact of the design of the acquisition of the HSI on the theoretical performance of the model for the estimation of the latent CP factors.

Notation. In this paper, we follow [19], [20] for tensor notations. We use the following fonts: lower (a) or uppercase (A) plain font for scalars, boldface lowercase (\mathbf{a}) for vectors, boldface uppercase (\mathbf{A}) for matrices and calligraphic (\mathcal{A}) for tensors. The elements of vectors, matrices, and tensors are denoted to as a_i , $A_{i,j}$ and $\mathcal{A}_{i_1, \dots, i_N}$, respectively. For a matrix \mathbf{A} , we denote its transpose by \mathbf{A}^T . We use the notation \mathbf{I}_N for the $N \times N$ identity matrix and $\mathbf{0}_{L \times K}$ for the $L \times K$ matrix of zeros. The symbols \boxtimes and \odot denote the Kronecker and Khatri-Rao products, respectively. The operator \otimes stands for the vector outer product. We use $\text{vec}\{\cdot\}$ for the standard column-major vectorization of a matrix or a tensor. For a vector \mathbf{a} , the operator $\text{diag}\{\mathbf{a}\}$ produces a diagonal matrix whose diagonal entries are the elements of \mathbf{a} . For two matrices \mathbf{A} and \mathbf{B} , the operation $\text{Diag}\{\mathbf{A}, \mathbf{B}\}$ produces a block-diagonal matrix whose diagonal blocks are \mathbf{A} and \mathbf{B} .

II. BACKGROUND ON TENSORS AND HSR

A. Tensor algebra preliminaries

A third-order tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ can be viewed as a three-dimensional array indexed by the elements $\mathcal{X}_{i,j,k}$, for

$i \in \{1, \dots, I\}$, $j \in \{1, \dots, J\}$ and $k \in \{1, \dots, K\}$. Each dimension of a tensor is called a mode. A mode- p fiber of \mathcal{X} is a vector obtained by fixing all but the p -th dimension.

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

Definition II.2. Matrix mode product – The matrix p -mode product between a tensor \mathcal{X} and a matrix M is denoted by $\mathcal{X} \bullet_p M$ and is evaluated such that each mode- p fiber of \mathcal{X} is multiplied by M . For instance, the elements of the mode-1 product between $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ and $M \in \mathbb{R}^{L \times I}$ are accessed as $(\mathcal{X} \bullet_1 M)_{\ell, j, k} = \sum_i \mathcal{X}_{i, j, k} M_{i, \ell}$, $\ell \in \{1, \dots, L\}$.

Moreover, we have $\mathcal{Y} = \mathcal{X} \bullet_k M \Leftrightarrow \mathbf{Y}^{(k)} = \mathbf{X}^{(k)} M^T$.

Definition II.3. Outer product – The outer product between three vectors $\mathbf{a} \in \mathbb{R}^I$, $\mathbf{b} \in \mathbb{R}^J$, $\mathbf{c} \in \mathbb{R}^K$ is a rank-one tensor $\mathcal{X} = \mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} \in \mathbb{R}^{I \times J \times K}$ whose elements are accessed as $\mathcal{X}_{i, j, k} = a_i b_j c_k$.

B. Canonical polyadic decomposition

Definition II.4. Canonical polyadic decomposition – A third-order tensor admits a CPD as $\mathcal{X} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$, where $\mathbf{A} \in \mathbb{R}^{I \times N}$, $\mathbf{B} \in \mathbb{R}^{J \times N}$, $\mathbf{C} \in \mathbb{R}^{K \times N}$ are the latent CP factors of the decomposition. When minimal, the integer N denotes the tensor rank of \mathcal{X} . Each entry of \mathcal{X} can be expressed equivalently as

$$\mathcal{X}_{i, j, k} = \sum_{n=1}^N A_{i, n} B_{j, n} C_{k, n}. \quad (1)$$

The CPD enjoys powerful uniqueness conditions: indeed, the CP factors \mathbf{A} , \mathbf{B} , \mathbf{C} are essentially unique up to scaling and permutation ambiguities, if the rank N is not too large [19], [20]. Permutation ambiguity means that the columns of the latent CP factors can be reordered arbitrarily by any permutation matrix $\mathbf{\Pi} \in \mathbb{R}^{N \times N}$ as

$$\mathcal{X} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket = \llbracket \mathbf{A}\mathbf{\Pi}, \mathbf{B}\mathbf{\Pi}, \mathbf{C}\mathbf{\Pi} \rrbracket.$$

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

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

with $\alpha_n \beta_n \gamma_n = 1$ for $n \in \{1, \dots, N\}$. 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 \mathbf{A} and \mathbf{B} factors to ones. This corresponds to rescaling eq. (2) with $\alpha_n = \frac{1}{A_{1, n}}$, $\beta_n = \frac{1}{B_{1, n}}$ and $\gamma_n = \frac{1}{\alpha_n \beta_n}$.

Property II.5. Tensor unfoldings under CP model – The unfoldings of \mathcal{X} admitting a CPD (1) can be expressed as

$$\begin{aligned} \mathbf{X}^{(1)} &= (\mathbf{C} \odot \mathbf{B}) \mathbf{A}^T \in \mathbb{R}^{JK \times I}, \\ \mathbf{X}^{(2)} &= (\mathbf{C} \odot \mathbf{A}) \mathbf{B}^T \in \mathbb{R}^{IK \times J}, \\ \mathbf{X}^{(3)} &= (\mathbf{B} \odot \mathbf{A}) \mathbf{C}^T \in \mathbb{R}^{IJ \times K}. \end{aligned}$$

C. Uniqueness of the CPD

Definition II.6. Kruskal rank – The Kruskal rank of a matrix M , denoted $\kappa(M)$, is defined as the maximum value k such that any k columns of M are linearly independent [21], [22].

One of the most general and well-known sufficient conditions on uniqueness of the CPD is due to Kruskal [21], [23] and reads as follows:

$$\kappa(\mathbf{A}) + \kappa(\mathbf{B}) + \kappa(\mathbf{C}) \geq 2N + 2.$$

Stronger results are available for generic uniqueness. We say that the CPD $\mathcal{X} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$ of rank N is generically unique if, for random matrices \mathbf{A} , \mathbf{B} , \mathbf{C} distributed according to an absolutely continuous probability distribution, we have

$$\min(I, N) + \min(J, N) + \min(K, N) \geq 2N + 2. \quad (3)$$

Equivalently, the set of \mathbf{A} , \mathbf{B} , \mathbf{C} not leading to unique decomposition has measure zero. In this case, the Kruskal condition implies (3).

In [24], another sufficient condition was provided:

$$N \leq 2^{\lceil \log_2(J) \rceil + \lceil \log_2(K) \rceil - 2}. \quad (4)$$

However, it should be mentioned that (3) and (4) are only sufficient conditions ensuring generic uniqueness.

The best known bounds on generic uniqueness are given in [25, Theorem 1]. In particular, it is shown that generic uniqueness takes place for all N such $N < \lceil \frac{IJK}{I+J+K-2} \rceil$ (i.e. all ranks smaller than the generic rank) except few special cases and so-called unbalanced tensors, see [25] for more details.

D. Coupled CP Model for hyperspectral super-resolution

We consider two tensors $\mathcal{Y}_1 \in \mathbb{R}^{I_H \times J_H \times K}$ and $\mathcal{Y}_2 \in \mathbb{R}^{I \times J \times K_M}$. In HSR, \mathcal{Y}_1 and \mathcal{Y}_2 represent the HSI and MSI, respectively. The spectral resolution of the MSI is lower than that of the HSI ($K_M \ll K$), while its spatial resolution is higher ($I > I_H$, $J > J_H$). Under the same acquisition conditions, the MSI and HSI usually represent the same target, hence \mathcal{Y}_1 and \mathcal{Y}_2 are viewed as two degraded versions of a single SRI $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$. The HSR problem thus consists in recovering the SRI \mathcal{X} from \mathcal{Y}_1 and \mathcal{Y}_2 .

We adopt the following degradation model that can be compactly written as contraction of $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ with matrices \mathbf{P} , \mathbf{Q} , \mathbf{R} :

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

where $\mathbf{P} \in \mathbb{R}^{I_H \times I}$, $\mathbf{Q} \in \mathbb{R}^{J_H \times J}$, and $\mathbf{R} \in \mathbb{R}^{K_M \times K}$ have full row rank. The matrix \mathbf{R} is the spectral degradation matrix containing the spectral response functions [26] of each band for the multispectral sensor, and \mathbf{P} , \mathbf{Q} are the spatial degradation matrices, i.e., we assume (for simplicity) that the spatial degradation is separable. For instance, the commonly accepted Wald's protocol [27] uses separable Gaussian blurring and downsampling in both spatial dimensions. The entries of the noise terms $\mathcal{E}_1 \sim \mathcal{N}(0, \mathbf{\Sigma}_1)$, $\mathcal{E}_2 \sim \mathcal{N}(0, \mathbf{\Sigma}_2)$ are independent and identically distributed (i.i.d.) real Gaussian tensors with zero mean and variances $\mathbf{\Sigma}_1 = \sigma_1^2 \mathbf{I}$ and $\mathbf{\Sigma}_2 = \sigma_2^2 \mathbf{I}$.

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

$$\begin{cases} \mathcal{Y}_1 = [\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1] + \boldsymbol{\varepsilon}_1, \\ \mathcal{Y}_2 = [\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_2] + \boldsymbol{\varepsilon}_2, \end{cases} \quad (6)$$

$$\text{where } \mathbf{A}_1 = \mathbf{P}\mathbf{A}_2, \mathbf{B}_1 = \mathbf{Q}\mathbf{B}_2, \mathbf{C}_2 = \mathbf{R}\mathbf{C}_1, \quad (7)$$

and $\mathbf{A}_1 \in \mathbb{R}^{J_H \times N}$, $\mathbf{B}_1 \in \mathbb{R}^{J_H \times N}$, $\mathbf{C}_1 \in \mathbb{R}^{K \times N}$, $\mathbf{A}_2 \in \mathbb{R}^{I \times N}$, $\mathbf{B}_2 \in \mathbb{R}^{J \times N}$, $\mathbf{C}_2 \in \mathbb{R}^{K_M \times N}$ are the factor matrices of the CPD. With this notation, the SRI admits a CPD

$$\mathcal{X} = [\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_1].$$

While (6) only is an uncoupled model, the addition of the constraints in (7) lead to the fully-coupled HSR.

In some applications, the spatial degradation matrices \mathbf{P} and \mathbf{Q} are unknown, and we refer to this case as blind-HSR. Similarly to model (6)–(7), we define the blind CP model as follows:

$$\begin{cases} \mathcal{Y}_1 = [\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1] + \boldsymbol{\varepsilon}_1, \\ \mathcal{Y}_2 = [\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_2] + \boldsymbol{\varepsilon}_2, \end{cases} \quad (8)$$

$$\text{where } \mathbf{C}_2 = \mathbf{R}\mathbf{C}_1, \quad (9)$$

thus blind-HSR only accounts for the knowledge of the spectral degradation matrix \mathbf{R} .

E. Estimation

In the uncoupled case, estimation of the CP factors can be performed by applying the uncoupled ALS algorithm [28] to \mathcal{Y}_1 and \mathcal{Y}_2 . The identifiability of both CPDs is required. For instance, for \mathcal{Y}_1 , ALS minimizes the following cost function:

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

which corresponds to the Maximum Likelihood (ML) Estimator (MLE) for $\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1$.

The fully-coupled HSR problem (6)–(7) can be solved by the algorithm STEREO proposed in [8]. It is a coupled ALS algorithm minimizing the criterion

$$\min_{\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_1} \|\mathcal{Y}_1 - [\mathbf{P}\mathbf{A}_2, \mathbf{Q}\mathbf{B}_2, \mathbf{C}_1]\|_F^2 + \lambda \|\mathcal{Y}_2 - [\mathbf{A}_2, \mathbf{B}_2, \mathbf{R}\mathbf{C}_1]\|_F^2. \quad (10)$$

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

$$N \leq \min \left(2^{\lfloor \log_2(JK_M) \rfloor - 2}, I_H J_H \right). \quad (11)$$

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

In the blind-HSR, the spatial degradation matrices \mathbf{P} and \mathbf{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 spectral degradation matrix \mathbf{R} . The criterion minimized by Blind-STEREO is

$$\min_{\substack{\mathbf{A}_1, \mathbf{B}_1, \\ \mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_1}} \|\mathcal{Y}_1 - [\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1]\|_F^2 + \lambda \|\mathcal{Y}_2 - [\mathbf{A}_2, \mathbf{B}_2, \mathbf{R}\mathbf{C}_1]\|_F^2, \quad (12)$$

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

III. CRAMÉR-RAO LOWER BOUNDS FOR COUPLED MODELS

A. 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” [29], [30]. In this paper, we propose to define it as the uniqueness of the proposed model.

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

We say that the statistical model $\mathcal{F} = \{\mathbf{f}_{\mathcal{Y};\omega} : \omega \in \Omega\}$ is identifiable if the mapping $\omega \mapsto \mathbf{f}_{\mathcal{Y};\omega}$ is injective [31], *i.e.*, any distribution $\mathbf{f}_{\mathcal{Y};\omega}$ corresponds to a single parameter ω . For the case of our Gaussian dataset, the following holds true:

$$\mathbf{f}_{\mathcal{Y};\omega_1} = \mathbf{f}_{\mathcal{Y};\omega_2} \Leftrightarrow \boldsymbol{\mu}(\omega_1) = \boldsymbol{\mu}(\omega_2). \quad (13)$$

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

Definition III.1. Identifiability at a point – *The noiseless model $\mathcal{Y} = \boldsymbol{\mu}(\omega)$ is identifiable at the point ω_0 if*

$$(\omega \neq \omega_0) \Rightarrow (\boldsymbol{\mu}(\omega) \neq \boldsymbol{\mu}(\omega_0)) \quad \forall \omega \in \mathbb{R}^m. \quad (14)$$

Definition III.2. Local identifiability – *The noiseless model $\mathcal{Y} = \boldsymbol{\mu}(\omega)$ is locally identifiable at ω_0 if there exists an open subset $\Omega_0 \subseteq \mathbb{R}^m$ containing ω_0 such that*

$$(\omega \neq \omega_0) \Rightarrow (\boldsymbol{\mu}(\omega) \neq \boldsymbol{\mu}(\omega_0)) \quad \forall \omega \in \Omega_0. \quad (15)$$

In the above model, the Fisher information matrix (FIM) for ω is obtained via the Slepian-Bangs formula [32]:

$$\mathbf{F}(\omega) = \left[\frac{\partial \boldsymbol{\mu}(\omega)}{\partial \boldsymbol{\omega}^\top} \right]^\top \boldsymbol{\Sigma}^{-1} \left[\frac{\partial \boldsymbol{\mu}(\omega)}{\partial \boldsymbol{\omega}^\top} \right] \in \mathbb{R}^{m \times m}, \quad (16)$$

where $\frac{\partial \boldsymbol{\mu}(\omega)}{\partial \boldsymbol{\omega}^\top}$ is the Jacobian of $\boldsymbol{\mu}(\omega)$. If the FIM in (16) is non-singular, then $\boldsymbol{\mu}(\omega)$ is locally identifiable in the noiseless case [29, Theorem 5].

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 \mathcal{Y} is a vectorized tensor of subgeneric rank admitting a CPD as in (1), and that $\omega = [\text{vec}\{\mathbf{A}\}; \text{vec}\{\mathbf{B}\}; \text{vec}\{\mathbf{C}\}]$, $\boldsymbol{\mu}(\omega) = \text{vec}\{[\mathbf{A}, \mathbf{B}, \mathbf{C}]\}$.

Generic uniqueness of the CPD of \mathcal{Y} implies that the rank of the Jacobian of $\boldsymbol{\mu}(\boldsymbol{\omega})$ in the generic case is equal to

$$\text{rank} \left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \right) = (I + J + K - 2)N$$

generically¹ (*i.e.*, except for a set of parameters $\boldsymbol{\omega}$ of measure zero), see [34, Sec. 3.2], [35], and [36, Def. 3.5]. Thus, the Jacobian is full rank once the scaling ambiguities in (2) are corrected; see section IV-B for more details on scaling ambiguities for coupled cases. Finally, from (16) it follows that full rank in the Jacobian implies that the FIM is invertible.

B. Coupled model with constraints

Let $\mathbf{f}_{\mathcal{Y}_1;\boldsymbol{\omega}}$ and $\mathbf{f}_{\mathcal{Y}_2;\boldsymbol{\omega}}$ be the PDFs of the random real datasets $\mathcal{Y}_1 \in \mathbb{R}^{n_1}$ and $\mathcal{Y}_2 \in \mathbb{R}^{n_2}$, parameterized by the unknown deterministic real parameter vector $\boldsymbol{\omega} \in \Omega$.

A general coupled model with constraints is expressed as:

$$\begin{cases} \mathcal{Y}_1 \sim \mathbf{f}_{\mathcal{Y}_1;\boldsymbol{\omega}} \text{ and } \mathcal{Y}_2 \sim \mathbf{f}_{\mathcal{Y}_2;\boldsymbol{\omega}}, \\ \mathbf{g}(\boldsymbol{\omega}) = \mathbf{0}, \end{cases} \quad (17)$$

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 [13].

We assume that:

- (i) the PDFs $\mathbf{f}_{\mathcal{Y}_1;\boldsymbol{\omega}}$ and $\mathbf{f}_{\mathcal{Y}_2;\boldsymbol{\omega}}$ are non-redundant functions differentiable w.r.t. $\boldsymbol{\omega}$, and that their supports do not depend on $\boldsymbol{\omega}$;
- (ii) the variables \mathcal{Y}_1 and \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

$$\boldsymbol{\omega} = \begin{bmatrix} \boldsymbol{\psi} \\ \boldsymbol{\xi} \end{bmatrix},$$

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

$$\mathbf{g}(\boldsymbol{\omega}) = \boldsymbol{\xi} - \mathbf{h}(\boldsymbol{\psi}) \in \mathbb{R}^{m_2}, \quad (18)$$

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

$$\boldsymbol{\omega}(\boldsymbol{\psi}) = \begin{bmatrix} \boldsymbol{\psi} \\ \mathbf{h}(\boldsymbol{\psi}) \end{bmatrix}. \quad (19)$$

The model (17) can thus be reformulated as the following unconstrained coupled model

$$\begin{cases} \mathcal{Y}_1 \sim \mathbf{f}_{\mathcal{Y}_1;\boldsymbol{\psi}} \text{ and } \mathcal{Y}_2 \sim \mathbf{f}_{\mathcal{Y}_2;\boldsymbol{\psi}}. \end{cases} \quad (20)$$

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 (17).

¹This result is well-known for complex tensors, but it is also valid for real tensors, see [33].

C. Uncoupled CRB

We consider that \mathcal{Y}_1 and \mathcal{Y}_2 are random real Gaussian distributed datasets parameterized by their mean, *i.e.*, $\mathcal{Y}_1 \sim \mathcal{N}(\boldsymbol{\mu}_1(\boldsymbol{\omega}), \boldsymbol{\Sigma}_1)$ and $\mathcal{Y}_2 \sim \mathcal{N}(\boldsymbol{\mu}_2(\boldsymbol{\omega}), \boldsymbol{\Sigma}_2)$ 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 [32]:

$$\mathbf{F}(\boldsymbol{\omega}) = \begin{bmatrix} \frac{\partial \boldsymbol{\mu}_1(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \\ \frac{\partial \boldsymbol{\mu}_2(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \end{bmatrix}^\top \text{Diag}\{\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2\}^{-1} \begin{bmatrix} \frac{\partial \boldsymbol{\mu}_1(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \\ \frac{\partial \boldsymbol{\mu}_2(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \end{bmatrix}. \quad (21)$$

If the FIM is non-singular, then the uncoupled CRB for $\boldsymbol{\omega}$ (namely $\text{CRB}(\boldsymbol{\omega})$) is obtained as $\text{CRB}(\boldsymbol{\omega}) = \mathbf{F}^{-1}(\boldsymbol{\omega})$. From section III-A, we see that invertibility of the FIM implies local identifiability of the whole parameter $\boldsymbol{\omega}$.

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 [37], [38]. In such cases, any estimator of $\boldsymbol{\omega}$ must have infinite variance [38]: in this paper, we choose not to compute the CRB when the FIM is singular. For uncoupled estimation, the constraint $\mathbf{g}(\boldsymbol{\omega}) = \mathbf{0}$ is ignored.

D. 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 constrained CRB (CCRB). The versatility of the CCRB has been illustrated in a number of studies [38], [39], [40], [41], extending the results of [14] and proving its utility for performance analysis and design of a measurement system [18].

In the seminal paper [14], the CCRB for $\boldsymbol{\omega}$ is expressed as

$$\text{CCRB}(\boldsymbol{\omega}) = \mathbf{F}^{-1} - \mathbf{F}^{-1} \mathbf{G}^\top \left[\mathbf{G} \mathbf{F}^{-1} \mathbf{G}^\top \right]^{-1} \mathbf{G} \mathbf{F}^{-1} \succeq \mathbf{0}, \quad (22)$$

where $\mathbf{F} \stackrel{\text{def}}{=} \mathbf{F}(\boldsymbol{\omega})$ and $\mathbf{G} = \left[\frac{\partial \mathbf{g}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \right] \in \mathbb{R}^{m_2 \times m}$ is a full row-rank matrix, which is equivalent to requiring that the constraints are non-redundant. We can express \mathbf{G} as

$$\mathbf{G} = \begin{bmatrix} \frac{\partial \mathbf{g}(\boldsymbol{\omega})}{\partial \boldsymbol{\psi}^\top} & \frac{\partial \mathbf{g}(\boldsymbol{\omega})}{\partial \boldsymbol{\xi}^\top} \end{bmatrix} = \begin{bmatrix} -\frac{\partial \mathbf{h}(\boldsymbol{\psi})}{\partial \boldsymbol{\psi}^\top} & \mathbf{I}_{m_2} \end{bmatrix}. \quad (23)$$

It is easy to see from eq. (22) 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 [17], [40], an alternative expression for the CCRB is

$$\text{CCRB}(\boldsymbol{\omega}) = \mathbf{U} \left[\mathbf{U}^\top \mathbf{F} \mathbf{U} \right]^{-1} \mathbf{U}^\top, \quad (24)$$

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

Contrary to eq. (22), eq. (24) does not require invertibility of \mathbf{F} . The above expression does not depend on the choice of \mathbf{U} either [17]. It is also noticeable that if \mathbf{F} is invertible, then the expressions in eq. (22) and eq. (24) are equivalent [40, Corollary 1]. Here, we choose to compute eq. (24) when we mention the CCRB.

E. Reparameterized CRB

Let us now consider a reparameterization of the PDFs $\mathbf{f}_{\mathbf{y}_1;\omega}$ and $\mathbf{f}_{\mathbf{y}_2;\omega}$ for the unknown parameter $\psi \in \Psi \subseteq \mathbb{R}^{m_1}$ where $\omega = \omega(\psi)$. We consider the particular case where ψ is a subset of parameters in ω ; then, arbitrarily we can rearrange the components of ω as in eq. (19).

In [31, p.125], an expression for the reparameterized FIM for ψ (namely $\mathbf{F}_c(\psi)$) is given:

$$\mathbf{F}_c(\psi) = \left[\frac{\partial \omega(\psi)}{\partial \psi^\top} \right]^\top \mathbf{F}(\omega(\psi)) \left[\frac{\partial \omega(\psi)}{\partial \psi^\top} \right]. \quad (25)$$

Contrary to the uncoupled case, we notice that uniqueness of model (19) only requires identifiability of the sub-parameter ψ . Additionally, we can express the reparameterized CRB for the parameter ξ (namely $\text{CRB}_c(\xi)$) as

$$\text{CRB}_c(\xi) = \left[\frac{\partial \mathbf{h}(\psi)}{\partial \psi^\top} \right]^\top \mathbf{F}_c^{-1}(\psi) \left[\frac{\partial \mathbf{h}(\psi)}{\partial \psi^\top} \right].$$

In [17], it is shown that for the parameter ψ , eq. (24) and eq. (25) lead to the same bound. Indeed, expressing the parameter ω as in eq. (19) is equivalent to introducing the set of constraints in eq. (18). We can check that

$$\begin{aligned} \frac{\partial \omega(\psi)}{\partial \psi^\top} &= \left[\frac{\mathbf{I}_{m_1}}{\frac{\partial \mathbf{h}(\psi)}{\partial \psi^\top}} \right] \in \mathbb{R}^{m \times m_1}, \\ \frac{\partial \mathbf{g}(\omega)}{\partial \omega^\top} &= \left[\frac{\partial \mathbf{g}(\omega)}{\partial \psi^\top} \quad \frac{\partial \mathbf{g}(\omega)}{\partial \xi^\top} \right] = \left[-\frac{\partial \mathbf{h}(\psi)}{\partial \psi^\top} \quad \mathbf{I}_{m_2} \right] \in \mathbb{R}^{m_2 \times m}. \end{aligned}$$

Therefore, $\mathbf{U} \stackrel{\text{def}}{=} \frac{\partial \omega(\psi)}{\partial \psi^\top}$ is a basis of $\ker \left(\frac{\partial \mathbf{g}(\omega)}{\partial \omega^\top} \right)$, which shows that the constrained FIM in eq. (24) and eq. (25) are equivalent².

IV. DIFFERENT PARAMETERIZATIONS AND ESTIMATION SCENARIOS

To derive appropriate performance bounds, it is necessary to embed the HSR 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.

A. Model parameters

We first define the model parameters³ that describe the coupled CP model in eq. (6). Here, we follow the notations of [15]. Since the spatial and spectral degradations are considered to be separable, and never occur simultaneously in model (5), we separate the CP factors into distinct parameters $\theta_1 \in \mathbb{R}^{K_N}$, $\theta_2 \in \mathbb{R}^{K_M N}$, $\phi_1 \in \mathbb{R}^{(I_H+J_H)N}$ and $\phi_2 \in \mathbb{R}^{(I+J)N}$ as

$$\begin{aligned} \theta_1^\top &= \text{vec}\{\mathbf{C}_1\}^\top, & \phi_1^\top &= [\text{vec}\{\mathbf{A}_1\}^\top \quad \text{vec}\{\mathbf{B}_1\}^\top], \\ \theta_2^\top &= \text{vec}\{\mathbf{C}_2\}^\top, & \phi_2^\top &= [\text{vec}\{\mathbf{A}_2\}^\top \quad \text{vec}\{\mathbf{B}_2\}^\top]. \end{aligned} \quad (26)$$

²As a result, invertibility of the constrained FIM in eq. (24) also implies that ψ is identifiable

³To provide a link with the notation of section III, we can define the parameters $\psi \in \mathbb{R}^{(I+J+K)N}$ and $\xi \in \mathbb{R}^{(I_H+J_H+K_M)N}$ such that

$$\psi = [\phi_2^\top \quad \theta_1^\top]^\top, \quad \xi = [\phi_1^\top \quad \theta_2^\top]^\top.$$

The above vectors can be stacked into one global parameter $\omega \in \mathbb{R}^m$ ($m = (I + J + K + I_H + J_H + K_M)N$) defined by

$$\omega^\top = [\phi_1^\top \quad \theta_1^\top \quad \phi_2^\top \quad \theta_2^\top]^\top.$$

From eq. (7), we can see that the model parameters can be linked together through non-redundant functions as

$$\mathbf{g}_1(\theta_1, \theta_2) = \mathbf{0}, \quad \mathbf{g}_2(\phi_1, \phi_2) = \mathbf{0},$$

where \mathbf{g}_1 and \mathbf{g}_2 are differentiable $\forall(\theta_1, \theta_2)$ (resp. (ϕ_1, ϕ_2)).

B. General framework for the fusion problem

For the HSR problem formulated as a fully-coupled CPD, we wish to estimate the parameters ϕ_2 and θ_1 , *i.e.*, the factor matrices of the SRI. In order to illustrate the advantage of data fusion over uncoupled estimation, we are also interested in the performance of the uncoupled and blind-HSR models: these cases require the calculation of performance bounds for the parameters ϕ_1 and θ_2 as well. As a result, we can distinguish two probabilistic scenarios, regarding if i) we are only interested in performance bounds and an analysis for the fully-coupled HSR problem, or ii) we want to compare the performance of the coupled CP approach to that of the uncoupled and blind approaches. Case i) allows for a bound calculation for the fully-coupled HSR problem only and will be referred to as *scenario 1*, while ii) can encompass uncoupled and Blind-HSR and will be referred to as *scenario 2*.

We consider that the low-resolution tensors $\mathcal{Y}_1 \in \mathbb{R}^{I_H \times J_H \times K}$ and $\mathcal{Y}_2 \in \mathbb{R}^{I \times J \times K_M}$ are random real Gaussian datasets. For all models and scenarios, \mathcal{Y}_1 and \mathcal{Y}_2 are distributed as in (17). Here, from the relationships between the model parameters, we can express eq. (18) as

$$\mathbf{g}(\omega) = [\mathbf{g}_1(\theta_1, \theta_2) \quad \mathbf{g}_2(\phi_1, \phi_2)]. \quad (27)$$

For each scenario and estimation framework (uncoupled, blind-HSR or fully-coupled HSR), the expression of $\mathbf{g}_1(\theta_1, \theta_2)$ and $\mathbf{g}_2(\phi_1, \phi_2)$ might change, resulting in different sets of constraints between the parameters. As in section III, the PDFs might only be parameterized by a subset of ω ; 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 III. For the FIM to be full rank, scaling ambiguities in the CPDs need to be solved [9] regarding the parameters we wish to estimate: indeed the manifold of rank- N tensors in $\mathbb{R}^{I \times J \times K}$ has dimension $(I + J + K - 2)N$. For each aforementioned scenario, we will provide different scaling options, allowing for the calculation of the performance bounds. We will also introduce different parameterizations and distributions for the observed datasets.

C. Scenario 1 – Performance bounds for fully-coupled HSR

In this first scenario, we are only interested in the performance analysis for the fully-coupled HSR problem. This case boils down to a performance analysis for ϕ_2 and θ_1 only. Thus, in this scenario, we only need uniqueness of the CPD of the SRI tensor to calculate the bounds. As discussed in section II, we set $(\mathbf{A}_2)_{1,:} = (\mathbf{B}_2)_{1,:} = 1$ to fix the correct the scaling ambiguities in ϕ_2 .

As a result, we define the reduced parameter

$$\tilde{\omega}_2^\top = [\text{vec}\{(\mathbf{A}_2)_{2:I,:}\}^\top \quad \text{vec}\{(\mathbf{B}_2)_{2:J,:}\}^\top] \in \mathbb{R}^{(I+J-2)N}.$$

that is only composed of the unknown entries of ϕ_2 . The full and reduced parameters can be linked through the relationship $\tilde{\phi}_2 = \mathbf{M}_2 \phi_2$. Here, the matrix $\mathbf{M}_2 \in \mathbb{R}^{(I+J-2)N \times (I+J)N}$ is a selection matrix constructed from $\mathbf{I}_{(I+J)N}$ by removing the $2N$ rows corresponding to known entries of ϕ_2 .

In that case, we can express the coupled CP model as

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

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., \mathbf{Y}_1 and \mathbf{Y}_2 are distributed according to

$$\begin{cases} \mathbf{f}_{\mathbf{Y}_1; \tilde{\phi}_2, \theta_1} &= (2\pi\sigma_1^2)^{\frac{-I_H J_H K}{2}} \\ &e\left(-\frac{1}{2\sigma_1^2} \|\mathbf{Y}_1 - [\mathbf{P}\mathbf{A}_2, \mathbf{Q}\mathbf{B}_2, \mathbf{C}_1]\|_F^2\right), \\ \mathbf{f}_{\mathbf{Y}_2; \tilde{\phi}_2, \theta_1} &= (2\pi\sigma_2^2)^{\frac{-I_J K M}{2}} \\ &e\left(-\frac{1}{2\sigma_2^2} \|\mathbf{Y}_2 - [\mathbf{A}_2, \mathbf{B}_2, \mathbf{R}\mathbf{C}_1]\|_F^2\right), \end{cases} \quad (29)$$

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

$$\begin{cases} \mathbf{g}_1(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) &= \boldsymbol{\theta}_2 - (\mathbf{I}_N \boxtimes \mathbf{R}) \boldsymbol{\theta}_1, \\ \mathbf{g}_2(\phi_1, \tilde{\phi}_2) &= \phi_1 - \begin{bmatrix} \mathbf{I}_N \boxtimes \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \boxtimes \mathbf{Q} \end{bmatrix} \mathbf{M}_2^\top \tilde{\phi}_2. \end{cases} \quad (30)$$

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*.

D. Scenario 2 – Comparing performance bounds for the HSR problem

1) *Specific scaling option:* In this second scenario, we want to compare performance bounds for the fully coupled problem to those in the uncoupled and blind case. This case requires the calculation of the bounds for the parameters ϕ_2 and θ_1 , as well as for ϕ_1 and θ_2 for blind and uncoupled HSR. Contrary to scenario 1, inversion of the FIM in the blind and uncoupled case require both CPDs to be generically unique. As a result, we also define the reduced parameter $\tilde{\omega} \in \mathbb{R}^{m-4N}$ as

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

We solve scaling ambiguities in ϕ_1 by setting the first rows of \mathbf{A}_1 and \mathbf{B}_1 to ones. As a result, we define the reduced parameter vector

$$\tilde{\phi}_1^\top = [\text{vec}\{(\mathbf{A}_1)_{2:I_H, :}\}^\top \quad \text{vec}\{(\mathbf{B}_1)_{2:I_H, :}\}^\top] \in \mathbb{R}^{(I_H+J_H-2)N},$$

that is only composed of the unknown entries of ϕ_1 . As in the previous subsection, we can express the reduced parameter vector through the relationship $\tilde{\phi}_1 = \mathbf{M}_1 \phi_1$, with $\mathbf{M}_1 \in \mathbb{R}^{(I_H+J_H-2)N \times (I_H+J_H)N}$ constructed as \mathbf{M}_2 .

Given eq. (7), solving the scaling ambiguities for the coupled CP factors of \mathbf{Y}_1 imposes that $(\mathbf{P}\mathbf{A}_2)_{1,:} = (\mathbf{Q}\mathbf{B}_2)_{1,:} = 1$. However, in a realistic coupled framework, it is unlikely that the spatial degradation matrices \mathbf{P} and \mathbf{Q} make the above equality valid, even if $(\mathbf{A}_2)_{1,:} = (\mathbf{B}_2)_{1,:} = 1$, as it would require that $(\mathbf{P})_{1,:} = [1 \quad \mathbf{0}_{1 \times (I-1)}]$ and $(\mathbf{Q})_{1,:} =$

$[1 \quad \mathbf{0}_{1 \times (J-1)}]$. The performance analysis for this case was addressed in [15].

Here, to circumvent this limitation and address a more general case, we introduce the scaling diagonal factors

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

such that $(\mathbf{A}_1 \cdot \mathbf{D}_\alpha^{-1})_{1,:} = (\mathbf{B}_1 \cdot \mathbf{D}_\beta^{-1})_{1,:} = 1$. We also need to rescale \mathbf{C}_2 as $\mathbf{C}_2 = \mathbf{R}\mathbf{C}_1 \cdot (\mathbf{D}_\alpha \mathbf{D}_\beta)^{-1}$ so that \mathbf{Y}_1 and \mathbf{Y}_2 are degraded versions of the SRI

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

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

$$\begin{cases} \mathbf{Y}_1 &= [\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1] + \boldsymbol{\mathcal{E}}_1, \\ \mathbf{Y}_2 &= [\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_2] + \boldsymbol{\mathcal{E}}_2, \end{cases} \quad (32)$$

$$\text{subject to } \mathbf{A}_1 = \mathbf{P}\mathbf{A}_2 \cdot \mathbf{D}_\alpha^{-1}, \mathbf{B}_1 = \mathbf{Q}\mathbf{B}_2 \cdot \mathbf{D}_\beta^{-1}, \\ \mathbf{C}_2 = \mathbf{R}\mathbf{C}_1 \cdot (\mathbf{D}_\alpha \mathbf{D}_\beta)^{-1}$$

for the fully coupled case.

The datasets are thus distributed according to

$$\begin{cases} \mathbf{f}_{\mathbf{Y}_1; \tilde{\phi}_2, \theta_1} &= (2\pi\sigma_1^2)^{\frac{-I_H J_H K}{2}} \\ &e\left(-\frac{1}{2\sigma_1^2} \|\mathbf{Y}_1 - [\mathbf{P}\mathbf{A}_2 \cdot \mathbf{D}_\alpha^{-1}, \mathbf{Q}\mathbf{B}_2 \cdot \mathbf{D}_\beta^{-1}, \mathbf{C}_1]\|_F^2\right), \\ \mathbf{f}_{\mathbf{Y}_2; \tilde{\phi}_2, \theta_1} &= (2\pi\sigma_2^2)^{\frac{-I_J K M}{2}} \\ &e\left(-\frac{1}{2\sigma_2^2} \|\mathbf{Y}_2 - [\mathbf{A}_2, \mathbf{B}_2, \mathbf{R}\mathbf{C}_1 \cdot (\mathbf{D}_\alpha \mathbf{D}_\beta)^{-1}]\|_F^2\right), \end{cases} \quad (33)$$

which is a parameterization different from eq. (29). The only case where the PDF in eq. (29) and eq. (33) are equivalent is the specific case where $\mathbf{D}_\alpha = \mathbf{D}_\beta = \mathbf{I}_N$, addressed in [15].

In eq. (32), we can see that the relationships linking the CP factors involve the scaling factors \mathbf{D}_α and \mathbf{D}_β . Rewriting these relationships in terms of the model parameters gives:

$$\begin{cases} \mathbf{g}_1(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) &= \boldsymbol{\theta}_2 - ((\mathbf{D}_\alpha \mathbf{D}_\beta)^{-1} \boxtimes \mathbf{R}) \boldsymbol{\theta}_1, \\ \mathbf{g}_2(\tilde{\phi}_1, \tilde{\phi}_2) &= \tilde{\phi}_1 - \mathbf{M}_1 \begin{bmatrix} \mathbf{D}_\alpha^{-1} \boxtimes \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_\beta^{-1} \boxtimes \mathbf{Q} \end{bmatrix} \mathbf{M}_2^\top \tilde{\phi}_2. \end{cases} \quad (34)$$

Due to the definition of \mathbf{D}_α and \mathbf{D}_β in (31), we refer to (34) as *non-linear constraints* on the model parameters.

3) *Parametrizations for uncoupled and blind-HSR:* In the uncoupled case, the datasets are distributed according to

$$\begin{cases} \mathbf{f}_{\mathbf{Y}_1; \tilde{\phi}_1, \theta_1} &= (2\pi\sigma_1^2)^{\frac{-I_H J_H K}{2}} \\ &e\left(-\frac{1}{2\sigma_1^2} \|\mathbf{Y}_1 - [\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1]\|_F^2\right), \\ \mathbf{f}_{\mathbf{Y}_2; \tilde{\phi}_2, \theta_2} &= (2\pi\sigma_2^2)^{\frac{-I_J K M}{2}} \\ &e\left(-\frac{1}{2\sigma_2^2} \|\mathbf{Y}_2 - [\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_2]\|_F^2\right), \end{cases} \quad (35)$$

and follow model (6).

For the Blind-HSR problem, we have the following model:

$$\begin{cases} \mathbf{Y}_1 &= [\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1] + \boldsymbol{\mathcal{E}}_1, \\ \mathbf{Y}_2 &= [\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_2] + \boldsymbol{\mathcal{E}}_2, \end{cases} \quad (36)$$

subject to $C_2 = RC_1 \cdot (D_\alpha D_\beta)^{-1}$.

The datasets are distributed according to

$$\begin{cases} \mathbf{f}_{\mathcal{Y}_1; \tilde{\phi}_1, \theta_1} &= (2\pi\sigma_1^2)^{-\frac{I_H J_H K}{2}} \\ & e^{-\left(-\frac{1}{2\sigma_1^2} \|\mathcal{Y}_1 - [\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1 \cdot (D_\alpha D_\beta)^{-1}]\|_F^2\right)}, \\ \mathbf{f}_{\mathcal{Y}_2; \tilde{\phi}_2, \theta_1} &= (2\pi\sigma_2^2)^{-\frac{I_{JK} M}{2}} \\ & e^{-\left(-\frac{1}{2\sigma_2^2} \|\mathcal{Y}_2 - [\mathbf{A}_2, \mathbf{B}_2, \mathbf{RC}_1 \cdot (D_\alpha D_\beta)^{-1}]\|_F^2\right)}. \end{cases} \quad (37)$$

For the blind case, we only consider the constraint $\mathbf{g}_1(\theta_1, \theta_2) = \theta_2 - ((D_\alpha D_\beta)^{-1} \boxtimes \mathbf{R}) \theta_1$ instead of (34).

E. Performance of the reconstruction of the SRI

Additionally to the model parameters in (26), we also define $\mathbf{x} = \text{vec}\{\mathcal{X}\} \in \mathbb{R}^\ell$ ($\ell = IJK$), that represents the vectorized SRI. Parameter \mathbf{x} can be linked to the model parameters through the relationship

$$\mathbf{g}_3(\mathbf{x}, \tilde{\phi}_2, \theta_1) = \mathbf{0}.$$

In order to get the bounds for \mathbf{x} , we utilize relationships between tensor unfoldings

$$\mathbf{x} = \underbrace{[(\mathbf{C}_1 \odot \mathbf{B}_2) \boxtimes \mathbf{I}_I]}_{S_1} \text{vec}\{\mathbf{A}_2\} \quad (38)$$

$$= \underbrace{\Pi^{(2,1)} [(\mathbf{C}_1 \odot \mathbf{A}_2) \boxtimes \mathbf{I}_J]}_{S_2} \text{vec}\{\mathbf{B}_2\} \quad (39)$$

$$= \underbrace{\Pi^{(3,1)} [(\mathbf{B}_2 \odot \mathbf{A}_2) \boxtimes \mathbf{I}_K]}_{S_3} \text{vec}\{\mathbf{C}_1\}, \quad (40)$$

where $\Pi^{(2,1)}$ and $\Pi^{(3,1)}$ are permutation matrices that link the second (resp. third) unfolding of \mathcal{X} to the first unfolding. As a result, the expression of $\mathbf{g}_3(\mathbf{x}, \tilde{\phi}_2, \theta_1)$ is given by

$$\mathbf{g}_3(\mathbf{x}, \tilde{\phi}_2, \theta_1) = \mathbf{x} - [\mathbf{S}_1 \quad \mathbf{S}_2 \quad \mathbf{S}_3] \mathbf{M}_3^\top \begin{bmatrix} \tilde{\phi}_2 \\ \theta_1 \end{bmatrix},$$

where $\mathbf{M}_3 = \text{Diag}\{\mathbf{M}_2, \mathbf{I}_{KN}\}$.

V. PERFORMANCE BOUNDS FOR THE HSR PROBLEM

In this section, we derive performance bounds for the HSR problem in the uncoupled, blind, and fully coupled cases. For the case of fully coupled datasets (*i.e.*, all degradation matrices are known), we address both scenarios described above.

A. Uncoupled CRB

In the uncoupled case, the CRB for the parameter $\tilde{\omega}$ is obtained by inverting the uncoupled FIM. To do so, scaling ambiguities in the CPDs of \mathcal{Y}_1 and \mathcal{Y}_2 need to be solved so that the FIM is full rank.

In practice, the FIM for $\tilde{\omega}$ (namely $F(\tilde{\omega})$) is computed by applying (21) to the tensors \mathcal{Y}_1 and \mathcal{Y}_2 :

$$F(\tilde{\omega}) = \begin{bmatrix} \frac{\partial \mu_1(\tilde{\omega})}{\partial \tilde{\omega}^\top} \\ \frac{\partial \mu_2(\tilde{\omega})}{\partial \tilde{\omega}^\top} \end{bmatrix}^\top \text{Diag}\{\Sigma_1, \Sigma_2\}^{-1} \begin{bmatrix} \frac{\partial \mu_1(\tilde{\omega})}{\partial \tilde{\omega}^\top} \\ \frac{\partial \mu_2(\tilde{\omega})}{\partial \tilde{\omega}^\top} \end{bmatrix}. \quad (41)$$

Here, the expressions of $\mu_1(\tilde{\omega})$ and $\mu_2(\tilde{\omega})$ are obtained from relationships between tensor unfoldings; please see [13] for a full derivation.

As in previous related works [11], [13], [15], we consider a case where the scaling ambiguities on \mathcal{Y}_1 and \mathcal{Y}_2 are solved, meaning that the FIM is non-singular. Thus, the CRB for $\tilde{\omega}$ can be obtained by inversion of the FIM: $\text{CRB}(\tilde{\omega}) = F^{-1}(\tilde{\omega})$. The CRB for each sub-parameter can be obtained by applying the block inversion lemma [42] to $F(\tilde{\omega})$. Please note that the uncoupled CRB can only be computed in scenario 2.

B. Blind-CCRB

We now address the case of blind-HSR and compute the CCRB associated with model (36). The Blind-CCRB can only be computed in scenario 2 due to the correction of scaling ambiguities on $\mathbf{A}_1, \mathbf{B}_1$.

We apply formula (24) to model (36), with $F \stackrel{\text{def}}{=} F(\tilde{\omega})$,

$$\mathbf{G} = \begin{bmatrix} \frac{\partial \mathbf{g}_1}{\partial \tilde{\phi}_1} & \frac{\partial \mathbf{g}_1}{\partial \theta_1} & \frac{\partial \mathbf{g}_1}{\partial \tilde{\phi}_2} & \frac{\partial \mathbf{g}_1}{\partial \theta_2} \end{bmatrix};$$

note that, due to the non-linear constraints at hand, we have $\frac{\partial \mathbf{g}_1}{\partial \tilde{\phi}_1} = \mathbf{0}$ but $\frac{\partial \mathbf{g}_1}{\partial \tilde{\phi}_2} \neq \mathbf{0}$.

Derivation of $\mathbf{g}_1(\theta_1, \theta_2)$ w.r.t. the parameters gives:

$$\frac{\partial \mathbf{g}_1}{\partial \tilde{\phi}_1} = \mathbf{0}, \quad \frac{\partial \mathbf{g}_1}{\partial \theta_1} = -\mathbf{Z}_1,$$

$$\frac{\partial \mathbf{g}_1}{\partial \tilde{\phi}_2} = -\text{Diag}\{\mathbf{Z}_2, \mathbf{Z}_3\} \mathbf{M}_2^\top, \quad \frac{\partial \mathbf{g}_1}{\partial \theta_2} = \mathbf{I}_{KMN}.$$

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

$$\mathbf{U} = \begin{bmatrix} \mathbf{I}_{(I+J+I_H+J_H)N-4N} \\ [\mathbf{0} \quad \mathbf{Z}_1 \quad \text{Diag}\{\mathbf{Z}_2, \mathbf{Z}_3\} \mathbf{M}_2^\top] \end{bmatrix}.$$

We thus obtain the Blind CCRB (**Blind-CCRB**):

$$\text{Blind-CCRB}(\tilde{\omega}) = \mathbf{U} \left[\mathbf{U}^\top \mathbf{F} \mathbf{U} \right]^{-1} \mathbf{U}^\top. \quad (42)$$

C. Performance bounds for fully-coupled HSR

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

1) *Scenario 1 – linear constraints*: In the first scenario, the most straightforward approach is to compute the reparameterized CRB using the fully-coupled model (6)–(7).

We consider the random real Gaussian distributed dataset \mathcal{Y} such that $\mathcal{Y} \sim \mathcal{N}(\mu(\tilde{\phi}_2, \theta_1), \Sigma)$, with

$$\mathcal{Y} = [\text{vec}\{\mathcal{Y}_1\}^\top \quad \text{vec}\{\mathcal{Y}_2\}^\top]^\top, \quad \Sigma = \text{Diag}\{\Sigma_1, \Sigma_2\},$$

$$\begin{aligned} \mu(\tilde{\phi}_2, \theta_1) &= \begin{bmatrix} \text{vec}\{[\mathbf{P}\mathbf{A}_2, \mathbf{Q}\mathbf{B}_2, \mathbf{C}_1]\} \\ \text{vec}\{[\mathbf{A}_2, \mathbf{B}_2, \mathbf{RC}_1]\} \end{bmatrix} \\ &= \underbrace{\begin{bmatrix} \mathbf{I}_K \boxtimes \mathbf{Q} \boxtimes \mathbf{P} \\ \mathbf{R} \boxtimes \mathbf{I}_{IJ} \end{bmatrix}}_{\tilde{\mathbf{P}}} \text{vec}\{[\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_1]\}. \end{aligned}$$

Note that the matrix $\tilde{\mathbf{P}}$ is constant, which means that we only have to compute the derivatives of $\text{vec}\{[\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_1]\}$ w.r.t. $(\tilde{\phi}_2, \theta_1)$. Those can be obtained from relationships between tensor unfoldings as in (38)–(40). As a result, we

can compute the reparameterized FIM for $(\tilde{\phi}_2, \theta_1)$ (denoted as $F_c(\tilde{\phi}_2, \theta_1)$) from the Slepian-Bangs formula as

$$F_c(\tilde{\phi}_2, \theta_1) = M_3 \begin{bmatrix} S_1^\top \\ S_2^\top \\ S_3^\top \end{bmatrix} \tilde{P}^\top \Sigma^{-1} \tilde{P} \begin{bmatrix} S_1 & S_2 & S_3 \end{bmatrix} M_3^\top. \quad (43)$$

The reparameterized CRB for $(\tilde{\phi}_2, \theta_1)$ can be then computed as $\text{CRB}_c(\tilde{\phi}_2, \theta_1) = F_c^{-1}(\tilde{\phi}_2, \theta_1)$. We can also compute the reparameterized CRB for the parameter x as

$$\text{CRB}_c(x) = \begin{bmatrix} \frac{\partial \mathbf{g}_3}{\partial \phi_2} & \frac{\partial \mathbf{g}_3}{\partial \theta_1} \end{bmatrix} \text{CRB}_c(\tilde{\phi}_2, \theta_1) \begin{bmatrix} \frac{\partial \mathbf{g}_3}{\partial \phi_2} & \frac{\partial \mathbf{g}_3}{\partial \theta_1} \end{bmatrix}^\top.$$

Equivalently, we can compute the CCRB from $F(\tilde{\omega})$ with

$$\mathbf{G} = \begin{bmatrix} \frac{\partial \mathbf{g}_2}{\partial \phi_1} & \frac{\partial \mathbf{g}_2}{\partial \theta_1} & \frac{\partial \mathbf{g}_2}{\partial \phi_2} & \frac{\partial \mathbf{g}_2}{\partial \theta_1} \\ \frac{\partial \mathbf{g}_1}{\partial \phi_1} & \frac{\partial \mathbf{g}_1}{\partial \theta_1} & \frac{\partial \mathbf{g}_1}{\partial \phi_2} & \frac{\partial \mathbf{g}_1}{\partial \theta_1} \end{bmatrix}. \quad (44)$$

Here, due to the linear constraints, we have $\frac{\partial \mathbf{g}_2}{\partial \theta_1} = \frac{\partial \mathbf{g}_2}{\partial \phi_2} = \mathbf{0}$ and $\frac{\partial \mathbf{g}_1}{\partial \phi_1} = \frac{\partial \mathbf{g}_1}{\partial \phi_2} = \mathbf{0}$. For other derivatives, we have

$$\begin{aligned} \frac{\partial \mathbf{g}_2}{\partial \phi_2} &= - \begin{bmatrix} \mathbf{I}_N \boxtimes \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \boxtimes \mathbf{Q} \end{bmatrix} M_2^\top, \\ \frac{\partial \mathbf{g}_2}{\partial \phi_1} &= \mathbf{I}_{(I_H+J_H)N}, \quad \frac{\partial \mathbf{g}_1}{\partial \theta_1} = -\mathbf{I}_N \boxtimes \mathbf{R}, \quad \frac{\partial \mathbf{g}_1}{\partial \theta_2} = \mathbf{I}_{KN}. \end{aligned}$$

The CCRB can then be computed using eq. (24).

2) *Scenario 2 – non-linear constraints*: In this subsection, the non-linear constraints in (32) yield to different bounds. In (44), we now have $\frac{\partial \mathbf{g}_1}{\partial \phi_2} \neq \mathbf{0}$ and

$$\begin{aligned} \frac{\partial \mathbf{g}_1}{\partial \theta_1} &= -\mathbf{Z}_1, \quad \frac{\partial \mathbf{g}_1}{\partial \theta_2} = \mathbf{I}_{KN}, \\ \frac{\partial \mathbf{g}_1}{\partial \phi_1} &= \mathbf{0}, \quad \frac{\partial \mathbf{g}_1}{\partial \phi_2} = -M_1 \text{Diag}\{\mathbf{Z}_2, \mathbf{Z}_3\} M_2^\top, \\ \frac{\partial \mathbf{g}_2}{\partial \phi_1} &= \mathbf{I}_{(I_H+J_H-2)N}, \quad \frac{\partial \mathbf{g}_2}{\partial \phi_2} = -M_1 \text{Diag}\{\mathbf{Z}_4, \mathbf{Z}_5\} M_2^\top. \end{aligned}$$

The matrices \mathbf{Z}_4 and \mathbf{Z}_5 are given in Appendix A, and the CCRB for $\tilde{\omega}$ is computed using eq. (24).

We can also consider the reparameterized CRB: we assume that $\mathcal{Y} \sim \mathcal{N}(\mu(\tilde{\phi}_2, \theta_1), \Sigma)$, with

$$\mathcal{Y} = [\text{vec}\{\mathcal{Y}_1\}^\top \quad \text{vec}\{\mathcal{Y}_2\}^\top]^\top, \quad \Sigma = \text{Diag}\{\Sigma_1, \Sigma_2\},$$

$$\mu(\tilde{\phi}_2, \theta_1) = \begin{bmatrix} \text{vec}\{\mathbf{P}\mathbf{A}_2\mathbf{D}_\alpha^{-1}, \mathbf{Q}\mathbf{B}_2\mathbf{D}_\beta^{-1}, \mathbf{C}_1\} \\ \text{vec}\{\mathbf{A}_2, \mathbf{B}_2, \mathbf{R}\mathbf{C}_1(\mathbf{D}_\alpha\mathbf{D}_\beta)^{-1}\} \end{bmatrix}.$$

The Jacobian of $\mu(\tilde{\phi}_2, \theta_1)$ is the matrix

$$\begin{bmatrix} \frac{\partial \mu}{\partial \phi_2} & \frac{\partial \mu}{\partial \theta_1} \end{bmatrix} = M_1 \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \mathbf{X}_3 \\ \mathbf{X}_5 & \mathbf{X}_6 & \mathbf{X}_4 \end{bmatrix} M_2^\top.$$

The matrices \mathbf{X}_i ($i = 1, \dots, 6$) are given in Appendix B.

VI. 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 [43]. The code is implemented in MATLAB and available online at https://github.com/cprevost4/CCRB_Software.

A. Simulations setup

The entries of the true CP factors $\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_1$ were generated once as i.i.d. real standard Gaussian variables, and the first rows of $\mathbf{A}_2, \mathbf{B}_2$ were set to ones. The true CP factors $\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_2$ were constructed according to the parameter constraints for each scenario.

In all experiments, the spatial degradation matrices \mathbf{P} and \mathbf{Q} are generated following Wald's protocol [27] with a Gaussian filter of length q and a downsampling ratio d . We also assume that $\mathbf{P} = \mathbf{Q}$. The spectral degradation matrix \mathbf{R} is a selection-and-averaging matrix that selects the common spectral bands of the SRI and MSI. We refer to 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 $\text{SNR}_i = 10 \log_{10} (\|\mathcal{Y}_i\|_F^2 / \|\mathcal{E}_i\|_F^2)$, ($i = 1, 2$). We fix SNR_2 to 20dB while SNR_1 varies from 5 to 60dB, unless otherwise specified.

The model parameters are retrieved using MLE. For estimation in the uncoupled case, we use ALS [28] 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]. For all algorithms, at most 5000 iterations are performed. To speed up the convergence of the coupled algorithms, the CP factors obtained by uncoupled ALS are used as initialization. The permutation ambiguities in the estimated factors are corrected by searching for the best column permutation of \mathbf{C}_2 with fixed \mathbf{C}_1 and applying that same permutation to \mathbf{A}_2 and \mathbf{B}_2 . This step is performed by merely maximizing the correlation between the estimated and true CP factors; but it could be performed optimally using the Hungarian algorithm [44].

We evaluate the total MSE on the parameters by averaging the squared errors through 500 noise realizations. For each realization, the best out of 10 initializations is picked. In the following figures, we plot our results for the parameters

$$\tilde{\psi} = \begin{bmatrix} \tilde{\phi}_2 \\ \theta_1 \end{bmatrix}, \quad \tilde{\xi} = \begin{bmatrix} \tilde{\phi}_1 \\ \theta_2 \end{bmatrix},$$

which correspond respectively to the CP factors of the SRI and the degraded factors.

B. Equivalence between CCRB and reparameterized CRB

In this subsection, we illustrate the results of [17], [31] regarding the equivalence between the CCRB in section III-D and reparameterized CRB in section III-E for the parameter $\tilde{\omega}$ with the models introduced above. For this first set of experiments, we first consider $I = J = 18$, $I_H = J_H = 6$, $K = 16$ and $K_M = 8$, and $N = 3$. In Figure 1, we show the total CCRB and reparameterized CRB for $\tilde{\omega}$ for the fully coupled model on a semi-log scale, for scenario 1 (linear constraints). In Figure 2, we consider scenario 2 and plot the total uncoupled CRB, CCRB and reparameterized CRB for $\tilde{\omega}$ in the fully-coupled and blind cases.

For both fully-coupled and blind-HSR, the CCRB and the reparameterized CRB for $\tilde{\omega}$ are numerically equivalent. Moreover, the Blind-CCRB is above the fully-coupled CCRB.

C. Choice of the rank

In this subsection, we investigate the choice of the tensor rank for the CPD of \mathcal{X} . We suppose that we wish to recover

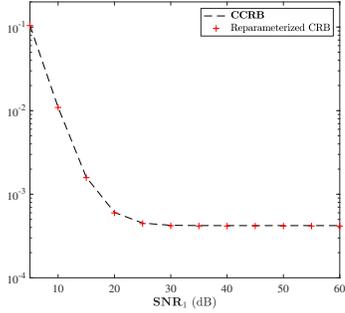


Fig. 1. Total bounds of $\tilde{\omega}$ with linear constraints

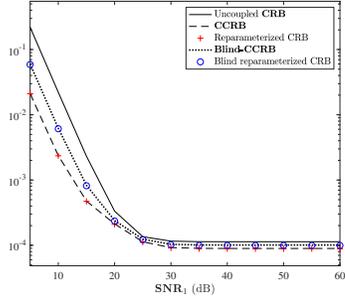


Fig. 2. Total bounds for $\tilde{\omega}$ with non-linear constraints

the CP factors of the SRI, namely $\mathbf{A}_{2,\text{th}}$, $\mathbf{B}_{2,\text{th}}$ and $\mathbf{C}_{1,\text{th}}$ with a rank $N_{\text{th}} = 3$. First, we fixed the “true” $\mathbf{A}_{2,\text{th}}$, $\mathbf{B}_{2,\text{th}}$ and $\mathbf{C}_{1,\text{th}}$ with i.i.d. real standard Gaussian entries. The unknown entries of the true factors were grouped in the parameter $\tilde{\psi}_{\text{th}}$:

$$\tilde{\psi}_{\text{th}} = \begin{bmatrix} \text{vec}\{\mathbf{A}_{2,\text{th}}\}_{2:I,:} \\ \text{vec}\{\mathbf{B}_{2,\text{th}}\}_{2:J,:} \\ \text{vec}\{\mathbf{C}_{1,\text{th}}\} \end{bmatrix}.$$

We generate the CP model with the dimensions in section VI-B and ranks N varying from 3 to 16. Thus, the first N_{th} columns of the true factors \mathbf{A}_2 , \mathbf{B}_2 , \mathbf{C}_1 are $\mathbf{A}_{2,\text{th}}$, $\mathbf{B}_{2,\text{th}}$ and $\mathbf{C}_{1,\text{th}}$, respectively. The remaining columns of these factors have i.i.d. real standard Gaussian entries. The first columns of \mathbf{A}_2 , \mathbf{B}_2 are also set to ones. The factors \mathbf{A}_1 , \mathbf{B}_1 , \mathbf{C}_2 are constructed according to model (6)–(7), that corresponds to the first scenario. The low-resolution tensors are constructed from these augmented CP factors. The CCRB is averaged over 100 realizations of the factors \mathbf{A}_i , \mathbf{B}_i , \mathbf{C}_i ($i = 1, 2$).

For $N \in \{3, \dots, 16\}$, we compute $\text{CCRB}(\tilde{\psi}_{\text{th}})$. In Figure 3, we plot the averaged CCRB as a function of SNR_1 and N .

We can see that, for all N , the CCRB decreases when SNR_1 decreases. Moreover, for all considered SNRs, the value of the CCRB increases when N increases; the best theoretical performance is obtained for $N = N_{\text{th}} = 3$.

However, one should note that in the HSR framework, the observed tensors are unlikely to be low-rank tensors. Indeed, the coupled CP formulation only performs a low-rank approximation of the data; in practice, choosing a high N allows for better modeling power [8].

D. Performance analysis and performance of the estimator

In this subsection, we evaluate the total MSE on the parameters provided by STEREO and compare it to the CCRB.

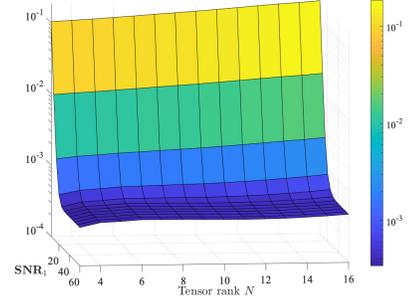


Fig. 3. $\text{CCRB}(\tilde{\psi}_{\text{th}})$ as a function of SNR_1 (dB) and N

We keep the same dimensions as in section VI-B. Whenever possible, we also compare the total MSE given by ALS and Blind-STEREO to the uncoupled CRB and Blind-CCRB, respectively. In Figures 4 and 5, we show the total performance bounds and the MSE for the parameters $\tilde{\omega}$ and \mathbf{x} respectively, on a semi-log scale.

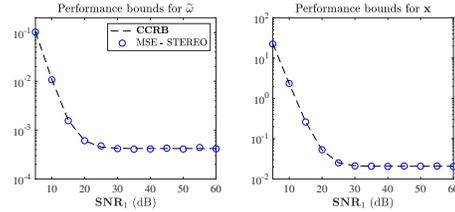


Fig. 4. Total bounds and MSE for $\tilde{\omega}$ (left) and \mathbf{x} (right): scenario 1

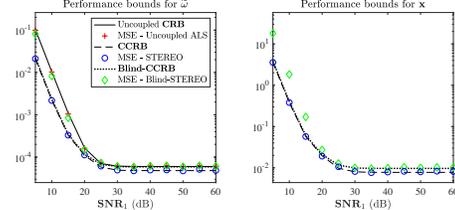


Fig. 5. Total bounds and MSE for $\tilde{\omega}$ (left) and \mathbf{x} (right): scenario 2

We can see that the total MSE given by STEREO follows the CCRB in both scenarios. When the scaling ambiguities are corrected on both tensors, the total MSE given by ALS reaches the uncoupled CRB. For the Blind-HSR problem, the total MSE given by Blind-STEREO is close to the uncoupled CRB for $\text{SNR}_1 < 20\text{dB}$ and follows the Blind-CCRB for $\text{SNR}_1 \geq 20\text{dB}$. Thus, in both scenarios, the estimators asymptotically reach their corresponding bounds.

In the next experiment, we study the convergence of STEREO with respect to the two estimation scenarios. For the first scenario, we generate the model according to (6)–(7), that corresponds to the first scenario. For scenario 2, we generate model (32) with non-linear constraints between the parameters. For each model, we compute the CCRBs in section V-C1 and section V-C2 for the parameter $\tilde{\psi}$. We also run STEREO according to each model.

We can see that, in both cases, the MSE given by STEREO follows the CCRB that corresponds to the model. On the left, the CCRB for scenario 1 is lower than that for scenario 2, while it is the opposite on the right. This figure shows that STEREO always reaches the CCRB, provided that the right model is employed.

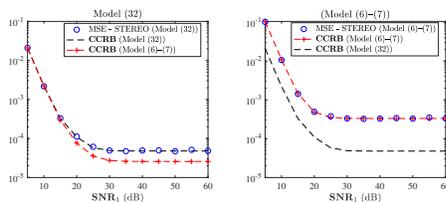


Fig. 6. Total CCRB and MSE provided by STEREO for $\tilde{\psi}$

E. Performance of STEREO without identifiability of \mathcal{Y}_1

In this subsection, we study a case where generic uniqueness of the HSI is not guaranteed, but the condition for unique noiseless recovery by STEREO is still satisfied. In particular, we investigate the performance of STEREO. Contrary to section VI-D, where the HSI and MSI 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 \mathcal{Y}_1 and \mathcal{Y}_2 . Thus we consider an adaptive 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. For this experiment, we generate the model as in section VI-B and $\text{SNR}_2 = 40\text{dB}$. We calculate $\text{CCRB}(\tilde{\omega})$ and compare it to the total MSE given by STEREO. We consider $\lambda = 1 \cdot 10^7$, $\lambda = 1$ and $\lambda = 1 \cdot 10^{-4}$. They correspond to the “true” regularization parameters for $\text{SNR}_1 = 5\text{dB}$, $\text{SNR}_1 = \text{SNR}_2$ and $\text{SNR}_1 = 60\text{dB}$, respectively. We plot the results on a semi-log scale in fig. 7.

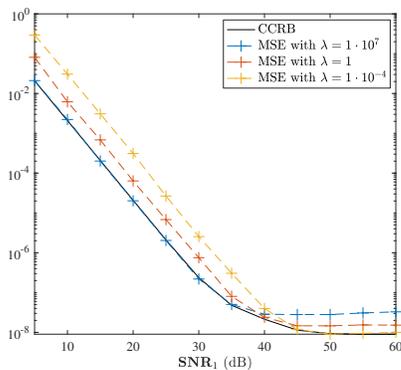


Fig. 7. $\text{CCRB}(\tilde{\omega})$ and total MSE for different λ

For $\lambda = 1 \cdot 10^7$, we notice that the total MSE reaches the CCRB for $\text{SNR}_1 \leq \text{SNR}_2$ even if λ is larger than the “true” λ . For higher SNR, the MSE is almost constant. For $\lambda = 1$, we can see that the MSE is above the CCRB for each noise level except $\text{SNR}_1 = \text{SNR}_2$. Finally, for $\lambda = 1 \cdot 10^{-4}$, while the MSE is above the CCRB for $\text{SNR}_1 \leq \text{SNR}_2$, it does reach the bound for higher SNR. Figure 7 shows that taking a small λ is likely to improve the performance of STEREO for high SNR, which is exactly what we are aiming at.

Thus we consider an adaptive choice of λ . That is, for each noise level, we successively run several iterations of STEREO with a decreasing $\lambda_a = \frac{c}{\sigma_2^2}$, possibly different from the “true” λ . Here, c is a constant that takes 5 decreasing values in $\{1, \dots, \sigma_1^2\}$, evenly distributed on a logarithmic scale. The value $\sigma_1^2 = 1$ corresponds to $\text{SNR}_1 = 0\text{dB}$; hence in our

experiments, we always have $\sigma_1^2 < 1$. We refer to this setup as “Adaptative λ ” and describe this procedure in Algorithm 1.

Algorithm 1: Adaptive procedure for STEREO

Input: $A_{2,0}, B_{2,0}, C_{1,0}, C_{2,0}; \lambda_a$

repeat

1. Run a few iterations of STEREO with λ_a ;
2. Decrease the value of c ;
3. Update $\lambda_a = \frac{c}{\sigma_2^2}$;

until $\lambda_a = \frac{\sigma_1^2}{\sigma_2^2}$;

To study the interest of the above procedure, we take $I_H = J_H = 4, I = J = 16, K_M = 10$ and $K = 20$. We also use $d = 4$ and $q = 3$. We consider $\text{SNR}_2 = 40\text{dB}$ while SNR_1 varies between 5dB and 60dB. For these dimensions, the generic uniqueness of \mathcal{Y}_1 is proved for $N \leq 9$ [25, Theorem 1.1], while condition (11) on unique recovery of the SRI by STEREO in the noiseless case gives $N \leq 16$. We address scenario 2 only, and tensor ranks $N = 10, N = 12$, and $N = 14$. For such ranks, uncoupled ALS performed on the HSI is not guaranteed to converge. To circumvent this limitation, in this subsection, we initialize STEREO as in [8]:

$$\begin{cases} A_{2,0}, B_{2,0}, C_{2,0} = \text{CPD}_N(\mathcal{Y}_2), \\ C_{1,0}^\top = (QB_{2,0} \odot PA_{2,0})^\dagger Y_2^{(3)}, \end{cases} \quad (45)$$

where the operation CPD_N returns estimated CP factors⁴ with rank N . In fact, initialization (45) implicitly considers that $\lambda = \infty$. For this reason, we expect STEREO not to converge when $\frac{\sigma_1^2}{\sigma_2^2}$ is low, that is, $\text{SNR}_1 \geq \text{SNR}_2$.

We run STEREO with $\lambda = \frac{\sigma_1^2}{\sigma_2^2}$ as well as the procedure with adaptive λ with one noise realization. We plot on a semi-log scale the total CCRB and MSE for $\tilde{\omega}$ as a function of SNR_1 .

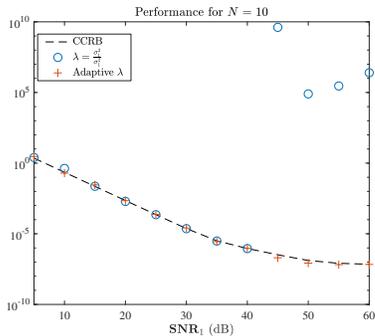
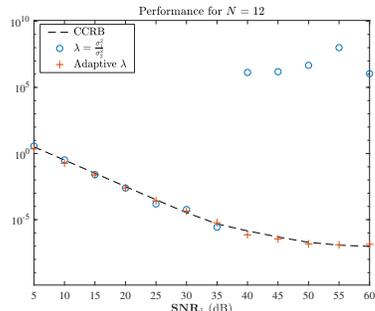
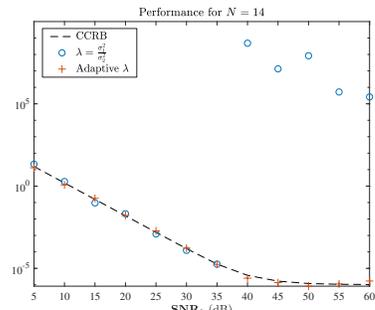
First, in Figures 8–10, we can see that for $\text{SNR}_1 \geq \text{SNR}_2$, STEREO with $\lambda = \frac{\sigma_1^2}{\sigma_2^2}$ does not converge indeed. Our guess is that the performance of the algorithm degrades when N is very large, especially when it becomes larger than (some of) the dimensions of the tensors.

However, running the procedure in Algorithm 1 does correct this behaviour: in all figures, the total MSE provided by STEREO reaches the CCRB in this setting. It should be mentioned that, for $\text{SNR}_1 < \text{SNR}_2$, the adaptive procedure is not needed for the MSE to reach the CCRB, since the “true” lambda is very large. Initialization (45) with $\lambda = \infty$ already provides a good estimation of the parameter $\tilde{\omega}$ in that case.

F. Influence of the spatial degradation matrices

In this subsection, we study the impact of the spatial degradation matrices on the CCRB. In practice, P and Q can be tuned by adjusting the filter size q and the downsampling ratio d : changing these parameters result in different acquisition schemes. On the first hand, when the downsampling ratio d varies while the filter size q is constant, the HSI possesses fixed spatial resolution but only contains a portion of the pixels in the SRI. On the other hand, when q varies while d is fixed,

⁴In practice, this operation is performed using TensorLab.

Fig. 8. $\text{CCRB}(\tilde{\omega})$ and total MSE, $N = 10$ Fig. 9. $\text{CCRB}(\tilde{\omega})$ and total MSE, $N = 12$ Fig. 10. $\text{CCRB}(\tilde{\omega})$ and total MSE, $N = 14$

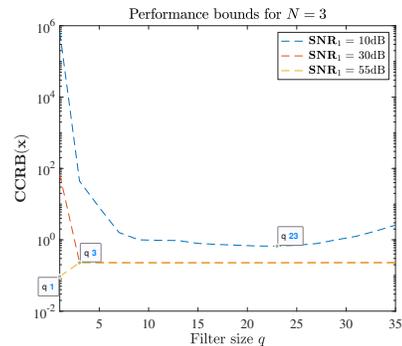
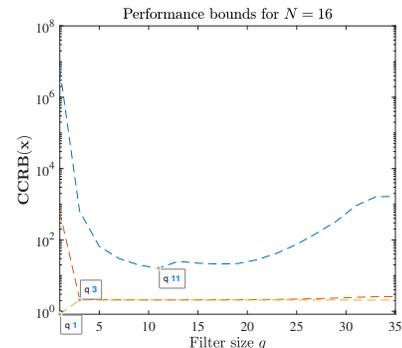
the HSI can be seen as a blurred SRI, as it contains a given spatial portion of the SRI, but with different pixel resolutions.

In this subsection, we investigate the influence of d and q on the reconstruction of the SRI (*i.e.*, we compute the CCRB for the parameter \mathbf{x}). We use the following dimensions: $I = J = 36$, $I_H = J_H = 6$, $K = 16$ and $K_M = 8$. Since we are only interested in the fully-coupled HSR problem, we resort to scenario 1 for model generation and bound derivation.

1) *Influence of the filter size:* First, we fix the downsampling ratio $d = 6$ and compute the CCRB for \mathbf{x} as a function of q for SNR_1 equal to 10dB, 30dB, and 55dB. We consider odd values of $q \in \{1, \dots, 36\}$. We recall that the expression for the Gaussian filter is available in Appendix C. We run the simulations for tensor ranks $N = 3$ and $N = 16$. Our results are displayed in Figures 11 and 12.

In both figures, the value of the CCRB decreases when the SNR increases. For moderate to high SNR (30dB and 55dB), the CCRB is almost constant. However, for low SNR (10dB), we can see that the CCRB increases from a certain q .

For $N = 3$, the filter size for which the CCRB is lowest becomes smaller when SNR_1 increases; for instance, for

Fig. 11. $\text{CCRB}(\mathbf{x})$ as a function of q for various SNR_1 and $N = 3$ Fig. 12. $\text{CCRB}(\mathbf{x})$ as a function of q for various SNR_1 and $N = 16$

$\text{SNR}_1 = 55\text{dB}$, the optimal filter size is $q = 1$. This result means that, from a certain noise level on \mathcal{Y}_1 , the presence of Gaussian blurring improves the theoretical performance of the model, which boils down to smoothing the images spatially.

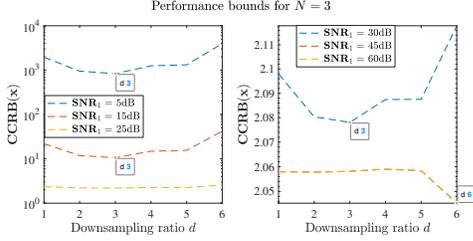
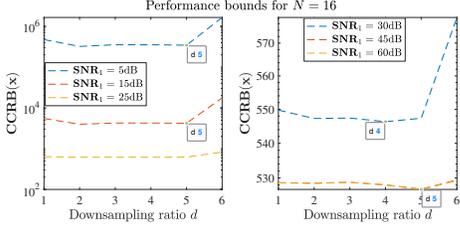
For $N = 16$, we can draw the same conclusions about the evolution of the optimal filter size, but we notice that, for $q \geq 13$ and low SNR, the CCRB increases much more than for $N = 3$. The optimal value of q in this case is $q = 11$.

2) *Influence of the downsampling ratio:* We now fix the filter size $q = 3$ and compute the CCRB for \mathbf{x} as a function of $d \in \{1, \dots, 6\}$ (to ensure that \mathbf{P} and \mathbf{Q} are full rank) and SNR_1 . Here, to depict the variations occurring at low-medium noise level, we separate the results in two subplots. We run the simulations for $N = 3$ and $N = 16$. The CCRB for \mathbf{x} is shown on Figures 13 and 14.

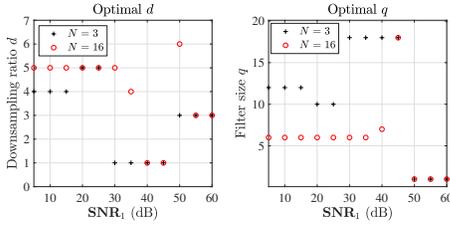
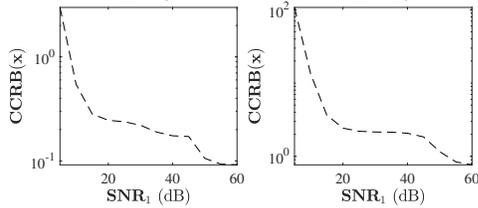
In Figure 13, for medium-high noise level, the value of d that gives the lowest CCRB is $d = q = 3$. For $d > 3$, the CCRB increases. However, we notice that when SNR_1 decreases, the CCRB varies significantly less. For SNR_1 equal to 45dB and 60dB, the lowest CCRB is obtained for $d = 6$.

In Figure 14, the value $d = 5$ provides the lowest bound regardless of the SNR. However, it must be noticed that the value $d \geq 5$ corresponds to the special case $d \geq I_H$. This means that more observations are available from the HSI, resulting in a higher CCRB, as illustrated in section VI-C.

3) *Optimal values of q and d :* Finally, we investigate the values of q and d for which the CCRB is lowest. We compute $\text{CCRB}(\mathbf{x})$ for $d \in \{1, \dots, 6\}$ and odd values of $q \in \{1, \dots, 36\}$. SNR_2 is fixed to 20dB, while SNR_1 varies from 5 to 60dB. We consider tensor ranks $N = 3$ and $N = 16$. For each SNR, we store the pair (q, d) for which the CCRB is the lowest: we plot our results in fig. 15. In fig. 16, we also

Fig. 13. $\text{CCRB}(\mathbf{x})$ as a function of d for $N = 3$ Fig. 14. $\text{CCRB}(\mathbf{x})$ as a function of d for $N = 16$

plot $\text{CCRB}(\mathbf{x})$ on a semi-log scale as a function of SNR_1 in dB.

Fig. 15. Optimal d and q as a function of SNR_1 (dB)Fig. 16. Optimal $\text{CCRB}(\mathbf{x})$ as a function of SNR_1 (dB)

In fig. 15 at low SNR, the optimal values of d are $d = 4$ and $d = 5$ for $N = 3$, and $d = 5$ for $N = 16$. This value decreases to $d = 1$ then increases again to $d = 3$ for $\text{SNR}_1 > 50$ dB and both ranks. The optimal filter size is $q = 1$ at $\text{SNR}_1 \geq 50$ for both ranks. For lower SNR, this value increases; while almost constant and equal to $q = 6$ for $N = 16$, it fluctuates more for $N = 3$. The large discrepancies in optimal d and q occurring at high SNR indicate that the choice of these parameters has little impact on the CCRB. Indeed, at high SNR, the cost function minimized by STEREO is very flat: thus a local minimum can be reached easily regardless of d and q .

In terms of CCRB, we notice a first elbow in the curves of fig. 16 at $\text{SNR}_1 = \text{SNR}_2$, then a second one for $\text{SNR}_1 = 45$ dB. This noise level corresponds to the lowest d and highest q in fig. 15 for both ranks.

VII. CONCLUSION

In this paper, we provided a full derivation of the CCRB for the HSR problem under various sets of constraints. In

our simulations, we assessed the performance of the MLE estimators STEREO and Blind-STEREO in optimal estimation conditions (*i.e.*, when the rank reduction allows for good estimation). Moreover, we illustrated the usefulness of the CCRB in terms of design of the hyperspectral measurement system. Finally, we displayed some cases where the tensor rank does not allow for correct estimation of the parameters by STEREO, and proposed an adaptive procedure to avoid non-converging cases.

APPENDIX A DERIVATIVES FOR CCRB

We give the expression of the matrices \mathbf{Z}_i ($i = 1, \dots, 5$) for the CCRB with non-linear constraints in sections V-B and V-C:

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

APPENDIX B DERIVATIVES FOR REPARAMETERIZED CRB

Here, we give the expression of the matrices \mathbf{X}_i ($i = 1, \dots, 6$) used for the reparameterized CRB in section V-C. The matrices $\mathbf{J}_{21}^{(i)}$, $\mathbf{J}_{31}^{(i)}$ are permutation matrices linking the second and third unfoldings to the first unfoldings of \mathcal{Y}_i .

$$\begin{aligned} \mathbf{X}_1 &= \left[(\mathbf{I}_N \odot (\mathbf{C}_1 \odot \mathbf{Q} \mathbf{B}_2 \mathbf{D}_\beta^{-1})) \mathbf{D}_\alpha^{-1} \boxtimes \mathbf{P} \right] \\ &\quad \left[\mathbf{I}_{IN} - (\mathbf{D}_\alpha^{-1} \boxtimes \mathbf{I}_I) (\mathbf{I}_N \odot \mathbf{A}_2) (\mathbf{I}_N \boxtimes \mathbf{P}_{1,:}) \right], \\ \mathbf{X}_2 &= (\mathbf{I}_N \boxtimes \mathbf{J}_{21}^{(1)}) \left[(\mathbf{I}_N \odot (\mathbf{C}_1 \odot \mathbf{P} \mathbf{A}_2 \mathbf{D}_\alpha^{-1})) \mathbf{D}_\beta^{-1} \boxtimes \mathbf{Q} \right] \\ &\quad \left[\mathbf{I}_{JN} - (\mathbf{D}_\beta^{-1} \boxtimes \mathbf{I}_J) (\mathbf{I}_N \odot \mathbf{B}_2) (\mathbf{I}_N \boxtimes \mathbf{Q}_{1,:}) \right], \\ \mathbf{X}_3 &= (\mathbf{I}_N \boxtimes \mathbf{J}_{31}^{(1)}) \\ &\quad \left[(\mathbf{I}_N \odot (\mathbf{Q} \mathbf{B}_2 \mathbf{D}_\beta^{-1} \odot \mathbf{P} \mathbf{A}_2 \alpha^{-1})) \boxtimes \mathbf{I}_K \right], \\ \mathbf{X}_4 &= (\mathbf{I}_N \boxtimes \mathbf{J}_{31}^{(2)}) \left[(\mathbf{I}_N \odot (\mathbf{B}_2 \odot \mathbf{A}_2)) (\mathbf{D}_\alpha \mathbf{D}_\beta)^{-1} \boxtimes \mathbf{R} \right], \\ \mathbf{X}_5 &= \left[(\mathbf{I}_N \odot (\mathbf{R} \mathbf{C}_1 \mathbf{D}_\beta^{-1} \odot \mathbf{B}_2)) \mathbf{D}_\alpha^{-1} \boxtimes \mathbf{I}_I \right] \\ &\quad \left[\mathbf{I}_{IN} - (\mathbf{D}_\alpha^{-1} \boxtimes \mathbf{I}_I) (\mathbf{I}_N \odot \mathbf{A}_2) (\mathbf{I}_N \boxtimes \mathbf{P}_{1,:}) \right], \\ \mathbf{X}_6 &= (\mathbf{I}_N \boxtimes \mathbf{J}_{21}^{(2)}) \left[(\mathbf{I}_N \odot (\mathbf{R} \mathbf{C}_1 \mathbf{D}_\alpha^{-1} \odot \mathbf{A}_2)) \mathbf{D}_\beta^{-1} \boxtimes \mathbf{I}_J \right] \\ &\quad \left[\mathbf{I}_{JN} - (\mathbf{D}_\beta^{-1} \boxtimes \mathbf{I}_J) (\mathbf{I}_N \odot \mathbf{B}_2) (\mathbf{I}_N \boxtimes \mathbf{Q}_{1,:}) \right]. \end{aligned}$$

APPENDIX C DEGRADATION MATRICES

Here, we explain in details how the degradation matrices are constructed (we consider that $\mathbf{P} = \mathbf{Q}$). As in [8], \mathbf{P} is constructed as $\mathbf{P} = \mathbf{S}_1 \mathbf{T}_1$, where \mathbf{T}_1 is a blurring Toeplitz matrix and \mathbf{S}_1 is a downsampling matrix. The blurring matrix is constructed from a Gaussian blurring kernel $\phi \in \mathbb{R}^{q \times 1}$ with a standard deviation $\sigma = \frac{q \sqrt{2 \log 2}}{4}$. For $m \in \{1, \dots, q\}$ and $m' = m - \lfloor \frac{q}{2} \rfloor$, we have

$$\phi(m) = \exp\left(\frac{-m'^2}{2\sigma^2}\right).$$

Thus, $\mathbf{T}_1 \in \mathbb{R}^{I \times I}$ can be seen as

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

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

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

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

REFERENCES

- [1] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," *IEEE Trans. Geosci. Remote Sens.*, vol. 5, no. 2, pp. 29–56, 2017.
- [2] G.A. Shaw and H.-h. K. Burke, "Spectral imaging for remote sensing," *Lincoln laboratory journal*, vol. 14, no. 1, pp. 3–28, 2003.
- [3] D. Lahat, T. Adali, and C. Jutten, "Multimodal data fusion: an overview of methods, challenges and prospects," *Proc. IEEE*, vol. 103, no. 9, pp. 1449–1477, 2015.
- [4] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled Nonnegative Matrix Factorization Unmixing for Hyperspectral and Multispectral Data Fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, 2012.
- [5] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Fast fusion of multi-band images based on solving a Sylvester equation," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4109–4121, 2015.
- [6] M. Simoes, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 3373–3388, 2015.
- [7] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J.-Y. Tourneret, "Multiband image fusion based on spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7236–7249, 2016.
- [8] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W.-K. Ma, "Hyperspectral Super-Resolution: A Coupled Tensor Factorization Approach," *IEEE Trans. Signal Process.*, vol. 66, no. 24, pp. 6503–6517, 2018.
- [9] S. Sahnoun and P. Comon, "Joint source estimation and localization," *IEEE Trans. Signal Process.*, vol. 63, no. 10, pp. 2485–2595, 2015.
- [10] X. Liu and N.D. Sidiropoulos, "Cramér-Rao lower bounds for low-rank decomposition of multidimensional arrays," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 2074–2086, 2001.
- [11] M. Boizard, R. Boyer, G. Favier, J.E. Cohen, and P. Comon, "Performance estimation for tensor CP decomposition with structured factors," in *Proc. ICASSP*, 2015.
- [12] R. Cabral Farias, J.E. Cohen, and P. Comon, "Exploring multimodal data fusion through joint decompositions with flexible couplings," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4830–4844, 2016.
- [13] C. Ren, R. Cabral Farias, P.-O. Amblard, and P. Comon, "Performance bounds for coupled models," in *Proc. 2016 IEEE SAM*, 2016, event-place: Rio de Janeiro, Brazil.
- [14] J.D. Gorman and A.O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1285–1301, 1990.
- [15] C. Prévost, K. Usevich, M. Haardt, P. Comon, and D. Brie, "Performance bounds for coupled CP model in the framework of hyperspectral super-resolution," in *Proc. IEEE CAMSAP*, 2019.
- [16] C. Prévost, E. Chaumette, K. Usevich, D. Brie, and P. Comon, "On Cramér-Rao lower bounds with random equality constraints," in *Proc. IEEE ICASSP*, 2020, pp. 5355–5359.
- [17] T. Menni, E. Chaumette, P. Larzabal, and J. P. Barbot, "New results on Deterministic Cramér-Rao bounds for real and complex parameters," *IEEE Trans. on SP*, vol. 60, no. 3, pp. 1032–1049, 2012.
- [18] T. Menni, J. Galy, E. Chaumette, and P. Larzabal, "Versatility of Constrained CRB for System Analysis and Design," *IEEE Trans. on AES*, vol. 50, no. 3, pp. 1841–1863, 2014.
- [19] P. Comon, "Tensors: A brief introduction," *IEEE Signal Process. Mag.*, vol. 31, no. 3, pp. 44–53, 2014.
- [20] T.G. Kolda and B.W. Bader, "Tensor Decompositions and Applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.
- [21] J.B. Kruskal, "Three-way arrays: Rank and uniqueness for trilinear decompositions," *Lin. Alg. Appl.*, vol. 18, no. 2, pp. 95–138, 1977.
- [22] R.A. Harshman and M.E. Lundy, "Data preprocessing and the extended PARAFAC model," *Research Methods for Multi-mode Data Analysis*, H. G. Law, C. W. Snyder, J. Hattie, and R. K. McDonald, eds., Praeger Publishers, New York, pp. 216–284, 1984.
- [23] J.B. Kruskal, "Rank, decomposition, and uniqueness for 3-way and n-way arrays," *Multway Data Analysis*, R. Coppi and S. Bolasco, eds., North-Holland, Amsterdam, vol. 18, no. 2, pp. 7–18, 1989.
- [24] L. Chiantini and G. Ottaviani, "On generic identifiability of 3-tensors of small rank," *SIAM Journal on Matrix Analysis and Applications*, vol. 33, no. 3, pp. 1018–1037, 2012.
- [25] L. Chiantini, G. Ottaviani, and N. Vannieuwenhoven, "An algorithm for generic and low-rank specific identifiability of complex tensors," *SIAM Journal on Matrix Analysis and Applications*, vol. 35, no. 4, pp. 1265–1287, 2014.
- [26] P. Mouroulis, R.O. Green, and T.G. Chrien, "Design of pushbroom imaging spectrometers for optimum recovery of spectroscopic and spatial information," *Applied Optics*, vol. 39, no. 13, pp. 2210–2220, 2000.
- [27] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogrammetric Eng. and Remote Sens.*, vol. 63, no. 6, pp. 691–699, 1997.
- [28] R. Bro, *Multi-way analysis in the food industry: Models, algorithms, and applications*, PhD Thesis, University of Amsterdam, The Netherlands, 1988.
- [29] C. Jauffret, "Observability and Fisher information matrix in nonlinear regression," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 43, no. 2, pp. 756–759, 2007.
- [30] T. Soderstrom, "Comments on 'order assumption and singularity of information matrix for pulse transfer function models'," *IEEE Trans. on Autom. Control*, vol. 20, no. 3, pp. 445–447, 1975.
- [31] E.L. Lehmann and G. Casella, *Theory of Point Estimation (2nd ed.)*, Springer, 1998.
- [32] D. Slepian, "Estimation of signal parameters in the presence of noise," *Trans. IRE Professional Group Inf. Theory*, vol. 3, no. 3, pp. 68–69, 1954.
- [33] Y. Qi, P. Comon, and L.-H. Lim, "Semialgebraic Geometry of Non-negative Tensor Rank," *SIAM J. Matrix Anal. Appl.*, vol. 37, no. 4, pp. 1556–1580, 2016.
- [34] P. Comon, Y. Qi, and K. Usevich, "Identifiability of an x-rank decomposition of polynomial maps," *SIAM Journal on Applied Algebra and Geometry*, vol. 1, no. 1, pp. 388–414, 2017.
- [35] V. Strassen, "Rank and optimal computation of generic tensors," *Linear Algebra and its Applications*, vol. 52–53, pp. 645–685, 1983.
- [36] S. Friedland, "On the generic and typical ranks of 3-tensors," *Linear algebra and its applications*, vol. 436, no. 3, pp. 478–497, 2012.
- [37] C. R. Rao, "Information and accuracy attainable in the estimation of statistical parameters," *Bull. Calcutta Math. Soc.*, vol. 37, pp. 81–91, 1945.
- [38] P. Stoica and T. L. Marzetta, "Parameter estimation problems with singular information matrices," *IEEE Trans. on SP*, vol. 49, no. 1, pp. 87–90, Jan. 2001.
- [39] T. L. Marzetta, "A simple derivation of the constrained multiple parameter Cramér-Rao bound," *IEEE Trans. on SP*, vol. 41, no. 6, pp. 2247–2249, June 1993.
- [40] P. Stoica and B. C. Ng, "On the Cramér-Rao bound under parametric constraints," *IEEE SP Letters*, vol. 5, no. 7, pp. 177–179, July 1998.
- [41] T. J. Moore, B. M. Sadler, and R. J. Kozick, "Maximum-Likelihood Estimation, the Cramér-Rao Bound, and the Method of Scoring With Parameter Constraints," *IEEE Trans. on SP*, vol. 56, no. 3, pp. 895–908, 2008.
- [42] G.A.F. Seber, *Matrix Handbook for Statisticians*, Wiley Series in Probability and Statistics, 2008.
- [43] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, *Tensorlab 3.0*, Available online, 2016.
- [44] J. Munkres, "Algorithms for the assignment and transportation problems," *Journal of the society for industrial and applied mathematics*, vol. 5, no. 1, pp. 32–38, 1957.