



**HAL**  
open science

## Performance bounds for coupled CP model in the framework of hyperspectral super-resolution

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

### ► To cite this version:

Clémence Prevost, Konstantin Usevich, Pierre Comon, Martin Haardt, David Brie. Performance bounds for coupled CP model in the framework of hyperspectral super-resolution. CAMSAP 2019 - IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, Dec 2019, Le Gosier, Guadeloupe, France. <10.1109/CAMSAP45676.2019.9022504>. <hal-02303132v2>

**HAL Id: hal-02303132**

**<https://hal.science/hal-02303132v2>**

Submitted on 11 Oct 2019

**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.



HAL Authorization

# PERFORMANCE BOUNDS FOR COUPLED CP MODEL IN THE FRAMEWORK OF HYPERSPECTRAL SUPER-RESOLUTION

C. Prévost<sup>1</sup>, K. Usevich<sup>1</sup>, P. Comon<sup>2</sup>, M. Haardt<sup>3</sup>, D. Brie<sup>1</sup>

<sup>1</sup> CRAN, Université de Lorraine, CNRS, Vandoeuvre-lès-Nancy, France

<sup>2</sup> GIPSA-Lab, Univ. Grenoble Alpes, CNRS, Saint-Martin d’Hères, France

<sup>3</sup> Ilmenau University of Technology, Ilmenau, Germany

clemence.prevost@univ-lorraine.fr, konstantin.usevich@univ-lorraine.fr

pierre.comon@gipsa-lab.grenoble-inp.fr, Martin.Haardt@tu-ilmenau.de, david.brie@univ-lorraine.fr

## ABSTRACT

We derive the constrained Cramér-Rao bounds for a coupled CP model with linear constraints applied to the hyperspectral super-resolution problem. For this problem, we consider two tensors representing low-resolution hyperspectral and multispectral images. In a practical measurement setup, white Gaussian noise sequences are added to each tensor with different variances. The coupling constraints are expressed between the factor matrices of the canonical polyadic model for each tensor. We show that the estimator given by the coupled alternating least squares algorithm achieves the bounds for given Signal-to-Noise Ratios, but requires the knowledge of the ratio of variances of the additive Gaussian noise sequences on each tensor.

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

## 1. 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, 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 good spatial and spectral resolutions. This problem lies in the framework of multimodal data fusion [2] between heterogeneous datasets. In this case, the datasets have different parametric models with shared variables, linked through possibly non-linear deterministic or statistical relations.

In [3], a tensor-based method was proposed for solving the HSR problem. Assuming that the super-resolution image itself admits a low-rank canonical polyadic (CP) decomposition (CPD), the problem is reformulated as a coupled CP approximation. An alternating least squares (ALS) algorithm called STEREO (Super-resolution TENSOR REconstruction) was also proposed.

In [4, 5, 6], performance bounds for uncoupled CP models have been provided. In [7], a Bayesian framework was pro-

posed for flexible coupling models, and hybrid Cramér-Rao bounds were derived. In [8], constrained Cramér-Rao bounds (CCRB) [11] for partially coupled complex tensors and possibly non-linear couplings were explored. In this paper, we derive the CCRB for the HSR problem, where all variables are coupled through linear relations.

This paper is organized as follows: in Section 2, we present a general coupled model. In particular, we express the HSR model as a parametric coupled CP model. In Section 3, we derive the expressions for the Fisher information matrix (FIM) in the uncoupled case, and extend the work in [8] by formulating the CCRB for a fully coupled model. Section 4 presents our experiments on performance bounds and mean squared errors on the parameters for coupled and uncoupled ALS.

**Notation.** In this paper we mainly follow [9] in what concerns the tensor notation (see also [10]). We use the symbol  $\otimes$  for the Kronecker product and  $\odot$  for the Khatri-Rao product. We use  $\text{vec}\{\cdot\}$  for the standard column-major vectorization of a tensor or a matrix. The operation  $\bullet_p$  denotes contraction on the  $p$ th index of a tensor; for instance,  $[\mathcal{A} \bullet_1 \mathcal{M}]_{ijk} = \sum_{\ell} \mathcal{A}_{\ell jk} M_{i\ell}$ . For the matrices  $\mathbf{A} \in \mathbb{R}^{I \times N}$ ,  $\mathbf{B} \in \mathbb{R}^{J \times N}$ ,  $\mathbf{C} \in \mathbb{R}^{K \times N}$ , we will use a shorthand notation for the CPD.

$$[\mathbf{A}, \mathbf{B}, \mathbf{C}] = [\mathcal{I}_N; \mathbf{A}, \mathbf{B}, \mathbf{C}]$$

where  $\mathcal{I}_N \in \mathbb{R}^{N \times N \times N}$  is a diagonal tensor of ones. For the matrices  $\mathbf{A}, \mathbf{B}$ , the notation  $\mathbf{A} \succeq \mathbf{B}$  means that  $\mathbf{A} - \mathbf{B}$  is a positive semidefinite (PSD) matrix. By  $\partial_{\theta_i \theta_j} \triangleq \frac{\partial^2}{\partial \theta_i \partial \theta_j}$ , we denote the second order derivative operator w.r.t. parameters  $\theta_i$  and  $\theta_j^T$  arranged in a Hessian matrix.

## 2. COUPLED DATASET MODEL

### 2.1. General coupled model

We consider a general coupled model for two real datasets  $\mathbf{x}_1 \in \Omega_1 \subseteq \mathbb{R}^{n_1}$  and  $\mathbf{x}_2 \in \Omega_2 \subseteq \mathbb{R}^{n_2}$ :

$$\begin{cases} \mathbf{x}_1 \sim f_{\mathbf{x}_1; \theta_1, \phi_1} \text{ and } \mathbf{x}_2 \sim f_{\mathbf{x}_2; \theta_2, \phi_2} \\ g(\theta_1, \theta_2) = 0 \text{ and } h(\phi_1, \phi_2) = 0. \end{cases} \quad (1)$$

The probability density functions (PDF)  $f_{\mathbf{x}_1; \boldsymbol{\theta}_1, \boldsymbol{\phi}_1}$  and  $f_{\mathbf{x}_2; \boldsymbol{\theta}_2, \boldsymbol{\phi}_2}$  are parametrized by the unknown deterministic real parameter vectors  $(\boldsymbol{\theta}_1, \boldsymbol{\phi}_1) \in \Theta_1 \times \Phi_1 \subseteq \mathbb{R}^{m_1} \times \mathbb{R}^{p_1}$  and  $(\boldsymbol{\theta}_2, \boldsymbol{\phi}_2) \in \Theta_2 \times \Phi_2 \subseteq \mathbb{R}^{m_2} \times \mathbb{R}^{p_2}$ .

Unlike [8], this model is referred to as *fully coupled* since all parameters are coupled through the relations  $g(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = 0$  and  $h(\boldsymbol{\phi}_1, \boldsymbol{\phi}_2) = 0$ . We assume that:

- (i)  $g(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ ,  $h(\boldsymbol{\phi}_1, \boldsymbol{\phi}_2)$  are non-redundant deterministic functions differentiable for all  $(\boldsymbol{\theta}_1, \boldsymbol{\phi}_1) \in \Theta_1 \times \Phi_1$  and  $(\boldsymbol{\theta}_2, \boldsymbol{\phi}_2) \in \Theta_2 \times \Phi_2$ , respectively;
- (ii) the PDF  $f_{\mathbf{x}_1; \boldsymbol{\theta}_1, \boldsymbol{\phi}_1}$  and  $f_{\mathbf{x}_2; \boldsymbol{\theta}_2, \boldsymbol{\phi}_2}$  are differentiable w.r.t.  $(\boldsymbol{\theta}_1, \boldsymbol{\phi}_1)$  (resp.  $(\boldsymbol{\theta}_2, \boldsymbol{\phi}_2)$ ), and that their respective supports do not depend on these parameters;
- (iii) the variables  $\mathbf{x}_1$  and  $\mathbf{x}_2$  are statistically independent.

In a multimodal data fusion scenario, we also assume that  $f_{\mathbf{x}_1; \boldsymbol{\theta}_1, \boldsymbol{\phi}_1}$  and  $f_{\mathbf{x}_2; \boldsymbol{\theta}_2, \boldsymbol{\phi}_2}$  can be different even if the model parameters are the same.

In [11, Lemma 2], some regularity conditions were proposed, under which the maximum-likelihood estimator (MLE) for the constrained parameters is asymptotically optimal in the mean squared error (MSE) sense and achieves the CCRB. Such an estimator is called an *efficient constrained estimator*, provided that it exists [11]. Thus, we wish to analyze the performance of model (1) for the estimation of  $(\boldsymbol{\theta}_1, \boldsymbol{\phi}_1)$  and  $(\boldsymbol{\theta}_2, \boldsymbol{\phi}_2)$ .

## 2.2. HSR degradation model

We consider an HSI cube  $\mathcal{Y}_1 \in \mathbb{R}^{I_H \times J_H \times K}$  and a MSI cube  $\mathcal{Y}_2 \in \mathbb{R}^{I \times J \times K_M}$  acquired from existing sensors (for instance, LANDSAT or QuickBird [12]). While  $I$ ,  $I_H$ ,  $J$  and  $J_H$  denote the size of the images in the spatial dimensions,  $K$  and  $K_M$  stand for the size of the data cubes in the spectral dimensions. The spectral resolution of MSI is lower than that of HSI ( $K_M \ll K$ ), while its spatial resolution is higher ( $I > I_H$ ,  $J > J_H$ ). The acquired MSI and HSI usually represent the same target, and  $\mathcal{Y}_1$  and  $\mathcal{Y}_2$  are viewed as two degraded versions of a single super-resolution image (SRI) data cube  $\mathcal{Y} \in \mathbb{R}^{I \times J \times K}$ . The hyperspectral data fusion problem [1] consists in recovering the SRI  $\mathcal{Y}$  from  $\mathcal{Y}_1$  and  $\mathcal{Y}_2$ .

In this paper, as in [3], we adopt the following degradation model, that can be compactly written as contraction of SRI:

$$\begin{cases} \mathcal{Y}_1 &= \mathcal{Y} \bullet_1 \mathbf{P} \bullet_2 \mathbf{Q} + \boldsymbol{\mathcal{E}}_1, \\ \mathcal{Y}_2 &= \mathcal{Y} \bullet_3 \mathbf{R} + \boldsymbol{\mathcal{E}}_2, \end{cases} \quad (2)$$

where  $\mathbf{P} \in \mathbb{R}^{I_H \times I}$ ,  $\mathbf{Q} \in \mathbb{R}^{J_H \times J}$  and  $\mathbf{R} \in \mathbb{R}^{K_M \times K}$  are known degradation matrices<sup>2</sup>. In a realistic setup, the Wald's protocol [13] is commonly used for spatial degradation matrices  $\mathbf{P}$  and  $\mathbf{Q}$ , and the spectral degradation matrix  $\mathbf{R}$  is a selection-averaging matrix. In this paper, we consider the case where the first rows of  $\mathbf{P}$ ,  $\mathbf{Q}$  and  $\mathbf{R}$  are equal to the

<sup>1</sup> $n_i$ ,  $m_i$  and  $p_i$  denote the size of the vectors  $\mathbf{x}_i$ ,  $\boldsymbol{\theta}_i$  and  $\boldsymbol{\phi}_i$ , respectively.

<sup>2</sup>We suppose that the spatial degradation for the HSI is separable.

first rows of the identity matrix. The entries of the noise terms  $\boldsymbol{\mathcal{E}}_1 \sim \mathcal{N}(0, \sigma_1^2 \mathbf{I})$ ,  $\boldsymbol{\mathcal{E}}_2 \sim \mathcal{N}(0, \sigma_2^2 \mathbf{I})$  are independent and identically distributed (i.i.d.) real Gaussian variables with zero mean and variance  $\sigma_1^2 \mathbf{I}$  and  $\sigma_2^2 \mathbf{I}$ , respectively.

## 2.3. CP-based data fusion

In [3] it was proposed to model the SRI as a tensor of low tensor-rank  $N$ . Under the assumption that the SRI admits a low-rank CPD, the degradation model (2) becomes

$$\begin{cases} \mathcal{Y}_1 &= \llbracket \mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1 \rrbracket + \boldsymbol{\mathcal{E}}_1, \\ \mathcal{Y}_2 &= \llbracket \mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_2 \rrbracket + \boldsymbol{\mathcal{E}}_2, \end{cases} \quad (3)$$

subject to  $\mathbf{A}_1 = \mathbf{P} \mathbf{A}_2$ ,  $\mathbf{B}_1 = \mathbf{Q} \mathbf{B}_2$ , and  $\mathbf{C}_2 = \mathbf{R} \mathbf{C}_1$ ,

where  $\mathbf{A}_1 \in \mathbb{R}^{I_H \times N}$ ,  $\mathbf{B}_1 \in \mathbb{R}^{J_H \times N}$ ,  $\mathbf{C}_1 \in \mathbb{R}^{K \times N}$ ,  $\mathbf{A}_2 \in \mathbb{R}^{I \times N}$ ,  $\mathbf{B}_2 \in \mathbb{R}^{J \times N}$ ,  $\mathbf{C}_2 \in \mathbb{R}^{K_M \times N}$  are the factor matrices of the CPD. Thus, the SRI admits a CPD such that  $\mathcal{Y} = \llbracket \mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_1 \rrbracket$ .

Following [8], we define the model parameters

$$\begin{aligned} \boldsymbol{\theta}_1 &= \text{vec}\{\mathbf{C}_1\} \in \mathbb{R}^{KN \times 1}, & \boldsymbol{\phi}_1 &= \begin{bmatrix} \text{vec}\{\mathbf{A}_1\} \\ \text{vec}\{\mathbf{B}_1\} \end{bmatrix} \in \mathbb{R}^{(I_H + J_H)N \times 1}, \\ \boldsymbol{\theta}_2 &= \text{vec}\{\mathbf{C}_2\} \in \mathbb{R}^{K_M N \times 1}, & \boldsymbol{\phi}_2 &= \begin{bmatrix} \text{vec}\{\mathbf{A}_2\} \\ \text{vec}\{\mathbf{B}_2\} \end{bmatrix} \in \mathbb{R}^{(I + J)N \times 1}, \end{aligned}$$

corresponding to the vectorization of the factor matrices for each tensor. The choice to group  $\mathbf{A}_i$  and  $\mathbf{B}_i$  factors, separate from the  $\mathbf{C}_i$  ( $i = 1, 2$ ) factor matrices is motivated by the fact that spatial and spectral degradations never occur in the same tensor, according to the model (2). Thus,  $\boldsymbol{\theta}_2$  and  $\boldsymbol{\phi}_1$  can be seen as degraded versions of  $\boldsymbol{\theta}_1$  and  $\boldsymbol{\phi}_2$  by the spectral and spatial degradation matrices, respectively.

Since  $\boldsymbol{\mathcal{E}}_1$  and  $\boldsymbol{\mathcal{E}}_2$  are i.i.d., the HSI and MSI are distributed according to

$$\begin{cases} f_{\mathcal{Y}_1; \boldsymbol{\theta}_1, \boldsymbol{\phi}_1} &= \left( \frac{1}{\sqrt{2\pi\sigma_1^2}} \right)^{I_H J_H K} \\ &\exp \left( -\frac{1}{2\sigma_1^2} \|\mathcal{Y}_1 - \llbracket \mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1 \rrbracket\|_F^2 \right), \\ f_{\mathcal{Y}_2; \boldsymbol{\theta}_2, \boldsymbol{\phi}_2} &= \left( \frac{1}{\sqrt{2\pi\sigma_2^2}} \right)^{I J K_M} \\ &\exp \left( -\frac{1}{2\sigma_2^2} \|\mathcal{Y}_2 - \llbracket \mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_2 \rrbracket\|_F^2 \right), \end{cases} \quad (4)$$

where  $\|\cdot\|_F$  denotes the Frobenius norm. The criterion minimized by STEREO is

$$\|\mathcal{Y}_1 - \llbracket \mathbf{P} \mathbf{A}_2, \mathbf{Q} \mathbf{B}_2, \mathbf{C}_1 \rrbracket\|_F^2 + \lambda \|\mathcal{Y}_2 - \llbracket \mathbf{A}_2, \mathbf{B}_2, \mathbf{R} \mathbf{C}_1 \rrbracket\|_F^2, \\ \text{subject to } \mathbf{A}_1 = \mathbf{P} \mathbf{A}_2, \mathbf{B}_1 = \mathbf{Q} \mathbf{B}_2, \text{ and } \mathbf{C}_2 = \mathbf{R} \mathbf{C}_1.$$

This criterion is proportional to the log-likelihood criterion, provided that  $\lambda = \frac{\sigma_1^2}{\sigma_2^2}$ .

### 3. PERFORMANCE ANALYSIS

In this section, we derive the FIM for the uncoupled and coupled cases of model (3). In a HSR framework, we wish to estimate the parameters of the SRI model, *i.e.*  $\theta_1$  and  $\phi_2$ . Here, we also provide performance analysis for the degraded parameters  $\theta_2$  and  $\phi_1$ . We only focus on the case where the model is identifiable so that the Fisher information matrix (FIM)  $\mathbf{F}$  is non-singular and  $\text{CRB} = \mathbf{F}^{-1}$ .

#### 3.1. Uncoupled CP model

In this subsection, we consider the uncoupled CP model

$$\begin{cases} \text{vec}\{\mathcal{Y}_1\} \sim \mathcal{N}(\mu_1(\theta_1, \phi_1), \sigma_1^2 \mathbf{I}), \\ \text{vec}\{\mathcal{Y}_2\} \sim \mathcal{N}(\mu_2(\theta_2, \phi_2), \sigma_2^2 \mathbf{I}), \end{cases} \quad (5)$$

with  $\mu_1(\theta_1, \phi_1) = \text{vec}\{[\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1]\}$  and  $\mu_2(\theta_2, \phi_2) = \text{vec}\{[\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_2]\}$ .

Here, we only derive the CRB for the HSI  $\mathcal{Y}_1$ , but the expression for the MSI can be obtained similarly. In fact, to solve the scaling indeterminacy of the CPD [14], we need to fix the first rows of  $\mathbf{A}_1$ ,  $\mathbf{B}_1$  to known values; here we normalize the factors by setting the first rows to ones. As a result, we define the parameter  $\tilde{\phi}_1$ , only composed of the unknown entries of  $\phi_1$ . The matrix  $\mathbf{M}_1 \in \mathbb{R}^{(I_H+J_H-2)N \times (I_H+J_H)N}$  is a mask obtained by removing the  $2N$  entries corresponding to known parameters of  $\phi_1$  such that  $\tilde{\phi}_1 = \mathbf{M}_1 \phi_1$ . Thus, we want to analyze the performance of the uncoupled model by deriving the CRB on  $\theta_1$  and  $\tilde{\phi}_1$ .

The Fisher information matrix on the uncoupled parameters is given by:

$$\mathbf{F}(\theta_1, \tilde{\phi}_1) = \begin{bmatrix} \mathbf{D}_{\theta_1 \theta_1} & \mathbf{D}_{\theta_1 \tilde{\phi}_1} \\ \mathbf{D}_{\tilde{\phi}_1 \theta_1} & \mathbf{D}_{\tilde{\phi}_1 \tilde{\phi}_1} \end{bmatrix},$$

where  $\mathbf{D}_{\theta_i \theta_j} = -\mathbb{E}[\partial_{\theta_i \theta_j} \mathcal{L}_1]$  and  $\mathcal{L}_1 = \ln f_{\mathcal{Y}_1; \theta_1, \phi_1}(\mathcal{Y}_1)$  is the log-likelihood function of  $\mathcal{Y}_1$ .

Basic operations on tensor unfoldings yield:

$$\begin{aligned} \text{vec}\{[\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1]\} &= \underbrace{[(\mathbf{C}_1 \odot \mathbf{B}_1) \boxtimes \mathbf{I}_{I_H}]}_{\mathbf{S}_{A_1}} \text{vec}\{\mathbf{A}_1\}, \\ &= \underbrace{\mathbf{J}_{21}^{(1)} [(\mathbf{C}_1 \odot \mathbf{A}_1) \boxtimes \mathbf{I}_{J_H}]}_{\mathbf{S}_{B_1}} \text{vec}\{\mathbf{B}_1\}, \\ &= \underbrace{\mathbf{J}_{31}^{(1)} [(\mathbf{B}_1 \odot \mathbf{A}_1) \boxtimes \mathbf{I}_K]}_{\mathbf{S}_{C_1}} \text{vec}\{\mathbf{C}_1\}, \end{aligned}$$

where  $\mathbf{J}_{21}^{(1)}, \mathbf{J}_{31}^{(1)} \in \mathbb{R}^{I_H J_H K \times I_H J_H K}$  are permutation matrices mapping the entries of the second (resp. third) unfolding of  $\text{vec}\{[\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1]\}$  to those of the first unfolding.

The elements of  $\mathbf{F}(\theta_1, \tilde{\phi}_1)$  can thus be written in a com-

compact form due to the Slepian-Bangs formula [15]:

$$\begin{cases} \mathbf{D}_{\theta_1 \theta_1} &= \frac{1}{\sigma_1^2} \mathbf{S}_{C_1}^\top \mathbf{S}_{C_1}, \\ \mathbf{D}_{\theta_1 \tilde{\phi}_1} &= \frac{1}{\sigma_1^2} \mathbf{S}_{C_1}^\top \begin{bmatrix} \mathbf{S}_{A_1} & \mathbf{S}_{B_1} \end{bmatrix} \mathbf{M}_1^\top, \\ \mathbf{D}_{\tilde{\phi}_1 \tilde{\phi}_1} &= \frac{1}{\sigma_1^2} \mathbf{M}_1 \begin{bmatrix} \mathbf{S}_{A_1}^\top \\ \mathbf{S}_{B_1}^\top \end{bmatrix} \begin{bmatrix} \mathbf{S}_{A_1} & \mathbf{S}_{B_1} \end{bmatrix} \mathbf{M}_1^\top. \end{cases}$$

The CRB can then be computed by block inversion of the Fisher information matrix. For a full derivation of the CRB for the uncoupled CP models, please refer to [8].

#### 3.2. Coupled CP model

We now consider a joint estimation of the vector parameters  $\theta_1, \theta_2, \tilde{\phi}_1$  and  $\tilde{\phi}_2$  for the coupled CP model. According to [11, 8], a simple expression of the CCRB can be obtained such that

$$\text{CCRB} = \mathbf{F}^{-1} - \mathbf{F}^{-1} \mathbf{J}^\top \left( \mathbf{J} \mathbf{F}^{-1} \mathbf{J}^\top \right)^{-1} \mathbf{J} \mathbf{F}^{-1} \succeq \mathbf{0}, \quad (6)$$

where  $\mathbf{F} \triangleq \text{Diag}\{\mathbf{F}(\theta_1, \tilde{\phi}_1), \mathbf{F}(\theta_2, \tilde{\phi}_2)\}$ . Equation (6) means that the CCRB is always lower than the CRB.

The matrix  $\mathbf{J} \triangleq \mathbf{J}(\theta_1, \tilde{\phi}_1, \theta_2, \tilde{\phi}_2)$  is the Jacobian of the constraints in (3). They can be expressed with respect to the model parameters as

$$\begin{cases} g(\theta_1, \theta_2) &= (\mathbf{I}_N \boxtimes \mathbf{R}) \theta_1 - \mathbf{I}_{KMN} \theta_2 = 0, \\ h(\tilde{\phi}_1, \tilde{\phi}_2) &= \mathbf{I}_{(I_H+J_H-2)N} \tilde{\phi}_1 \\ &\quad - \underbrace{\mathbf{M}_1 \begin{bmatrix} \mathbf{I}_N \boxtimes \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \boxtimes \mathbf{Q} \end{bmatrix} \mathbf{M}_2^\top}_{\mathcal{P}} \tilde{\phi}_2 = 0, \end{cases} \quad (7)$$

where  $\mathbf{M}_2$  is a mask matrix obtained similarly to  $\mathbf{M}_1$  such that  $\tilde{\phi}_2 = \mathbf{M}_2 \phi_2$ .

Thus, a general expression for  $\mathbf{J}$  is

$$\mathbf{J} = \begin{bmatrix} \frac{\partial g(\theta_1, \theta_2)}{\partial \theta_1} & \mathbf{0} & \frac{\partial g(\theta_1, \theta_2)}{\partial \theta_2} & \mathbf{0} \\ \mathbf{0} & \frac{\partial h(\tilde{\phi}_1, \tilde{\phi}_2)}{\partial \tilde{\phi}_1} & \mathbf{0} & \frac{\partial h(\tilde{\phi}_1, \tilde{\phi}_2)}{\partial \tilde{\phi}_2} \end{bmatrix}.$$

In this particular case, we have

$$\mathbf{J} = \begin{bmatrix} \mathbf{I}_N \boxtimes \mathbf{R} & \mathbf{0} & -\mathbf{I}_{KMN} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{(I_H+J_H-2)N} & \mathbf{0} & -\mathbf{M}_1 \mathcal{P} \mathbf{M}_2^\top \end{bmatrix}.$$

From the expression of  $\mathbf{J}$ , it is easy to see that the sets of parameters  $(\theta_1, \theta_2)$  and  $(\phi_1, \phi_2)$  are uncoupled.

### 4. SIMULATIONS

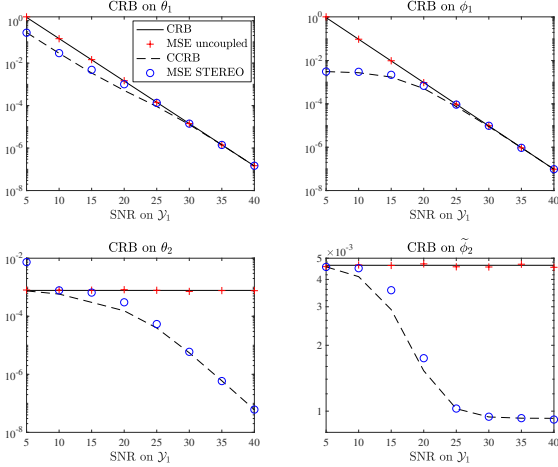
In this section, we simulate the performance of the coupled CP model under additive Gaussian noise for the constrained and unconstrained cases and compare it to the bounds presented in Section 3. For basic tensor operations we used TensorLab 3.0 [16].

The model parameters are retrieved using MLE. In the uncoupled case, an Alternating Least Squares (ALS) algorithm [17] is used, with random initialization for the factor matrices. For the coupled case, the algorithm proposed in [3] is used. It consists in a coupled ALS algorithm based on model (2). To speed up the convergence of the coupled algorithm, the factor matrices obtained by uncoupled ALS are used as initialization.

For the CP model to be identifiable, the scaling and permutation ambiguities are corrected by setting the first rows of the factors  $\mathbf{A}_1, \mathbf{B}_1, \mathbf{A}_2, \mathbf{B}_2$  to ones and searching for the best permutation of  $\mathbf{C}_2$  with fixed  $\mathbf{C}_1$ .

**Simulation settings.** We consider that  $I = J = K = 15$ ,  $I_H = J_H = 10$ , and  $K_M = 9$ . All CP factors are generated randomly according to i.i.d. real standard Gaussian variables, and first rows of  $\mathbf{A}_1, \mathbf{A}_2, \mathbf{B}_1, \mathbf{B}_2$  are set to 1. The tensor rank for both HSI and MSI is  $N = 3$ . The degradation matrices  $\mathbf{P}, \mathbf{Q}, \mathbf{R}$  are generated from identity matrices by keeping only the first  $I_H, J_H, K_M$  rows, respectively; this ensures that the matrices are full rank and normalized as in Subsection 2.2, which ensures that the coupling constraints in (7) are linear.

We evaluate the total MSE on the parameters by averaging the squared errors through 500 noise realizations. For each realization, the best of 5 initializations is picked. The SNR of  $\mathcal{Y}_2$  is fixed to 20dB while that of  $\mathcal{Y}_1$  varies from 5 to 40dB. For both uncoupled ALS and STEREO, 500 iterations at most are performed.

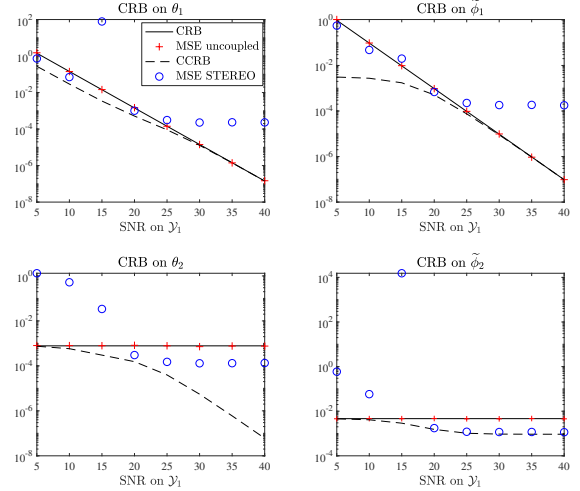


**Fig. 1.** Total MSE and performance bounds for the estimation of real coupled CP models on a semi-log scale with  $\lambda = \frac{\sigma_1^2}{\sigma_2^2}$

In Figure 1, we show the MSE and the performance bounds on a semi-logarithmic scale when the noise level is taken into account. By this, we mean that the regularization parameter  $\lambda$  is known [18, 19] and set to  $\lambda = \frac{\sigma_1^2}{\sigma_2^2}$ . We can see that the CCRB is indeed lower than the CRB. Moreover, the MSE for uncoupled ALS follows the CRB. Here, the MSE for STEREO follows closely the CCRB with a small gap when

the SNR for  $\mathcal{Y}_1$  is between 10dB and 25dB. In particular, the outlier point on the MSE for  $\theta_2$  for 5dB SNR may depict a case where the algorithm does not converge to a global minimum. For high SNR, the MSE for STEREO reaches the CCRB.

Those results show that the coupled CP algorithm proposed in [3] reaches the CCRB when the SNR of each tensor  $\mathcal{Y}_1, \mathcal{Y}_2$  is considered, and thus, the variances of the additive noises are needed. In other words, if STEREO performs noise variance normalization, then it is the MLE.



**Fig. 2.** Total MSE and performance bounds for the estimation of real coupled CP models on a semi-log scale with  $\lambda = 1$

In Figure 2, we show similar plots but in the case where the additive noise levels are not taken into account in the estimation of the parameters, *i.e.*  $\lambda = 1$ . In this case, the MSE for STEREO does not follow the CCRB. In particular, the outlier points on the MSE of  $\theta_1$  and  $\phi_2$  for 15dB SNR on  $\mathcal{Y}_1$  may be explained by the fact that the algorithm does not converge to a global minimum. For a SNR higher or equal to 20dB, the MSE is almost constant on a semi-logarithmic scale, and does not reach the CCRB.

## 5. CONCLUSION

We have proposed a formulation for constrained Cramér-Rao bounds for fully coupled CP model with linear constraints. In particular, we gave an expression of the CCRB for the HSR problem using an approximate low-rank CPD. Our experiments have shown that the considered coupled CP algorithm reaches the performance bounds provided that the noise variances ratio  $\lambda$  between the HSI and MSI is known, and is thus the MLE. If the noise variances are unknown, then the algorithm does not reach the bounds. Thus, the choice of this regularization parameter remains open: data-driven procedures can be envisaged [20]. Another approach would be to design methods that are robust to a misspecification of  $\lambda$  [21].

## 6. 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] 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.
- [3] 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.
- [4] S. Sahnoun and P. Comon, "Joint source estimation and localization," *IEEE Trans. Signal Process.*, vol. 63, no. 10, pp. 2485–2595, 2015.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] C. Ren, R. Cabral Farias, P.-O. Amblard, and P. Comon, "Performance bounds for coupled models," in *Proc. 2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2016.
- [9] P. Comon, "Tensors: A brief introduction," *IEEE Signal Process. Mag.*, vol. 31, no. 3, pp. 44–53, 2014.
- [10] T.G. Kolda and B.W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.
- [11] 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.
- [12] "Quickbird satellite," 2015, Available online.
- [13] 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.
- [14] P. Comon, "Tensor decompositions, state of the art and applications," in *Mathematics in Signal Processing V*, J. G. McWhirter and I. K. Proudler, Eds., pp. 1–24. Clarendon Press, Oxford, UK, 2002, available at arxiv:0905.0454.
- [15] 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.
- [16] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, "Tensorlab 3.0," Mar. 2016, Available online.
- [17] R. Bro, "Multi-way analysis in the food industry: Models, algorithms, and applications," Ph.D. dissertation, University of Amsterdam, The Netherlands, 1988.
- [18] J.E. Cohen, R. Cabral Farias, and P. Comon, "Joint tensor compression for coupled canonical polyadic decompositions," in *Proc. 24th European Signal Processing Conference (EUSIPCO 2016)*, 2016.
- [19] K. Naskovska and M. Haardt, "Extension of the semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalization to the efficient calculation of coupled CP decompositions," in *Proc. of 50th Asilomar Conf. on Signals, Systems, and Computers*, 2016.
- [20] P.C. Hansen, "The l-curve and its use in the numerical treatment of inverse problems," 1999.
- [21] F. Roemer and M. Haardt, "A semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalizations (SECSI)," *Signal Processing*, vol. 93, no. 9, pp. 2722–2738, 2013.