On Cramér-Rao lower bounds with random equality constraints

Clémence Prévost, Eric Chaumette, Konstantin Usevich, David Brie, Pierre Comon

ON CRAMÉR-RAO LOWER BOUNDS WITH RANDOM EQUALITY CONSTRAINTS

C. Prévost[†]*, E. Chaumette*[⋆]*, K. Usevich*[†]*, D. Brie*[†]*, P. Comon*[‡]*,*

[†] CRAN, Univ. of Lorraine, CNRS, Vandoeuvre-lès-Nancy, France ([clemence.prevost, konstantin.usevich, david.brie]@univ-lorraine.fr)

[⋆] Univ. of Toulouse/Isae-Supaero, Toulouse, France (eric.chaumette@isae.fr)

[‡] GIPSA-Lab, Univ. Grenoble Alpes, CNRS, Saint-Martin d'Hères, France (pierre.comon@gipsa-lab.grenoble-inp.fr)

ABSTRACT

Numerous works have shown the versatility of deterministic constrained Cramér-Rao bound for estimation performance analysis and design of a system of measurements. Indeed, most of factors impacting the asymptotic estimation performance of the parameters of interest can be taken into account via equality constraints. In this communication, we introduce a new constrained Cramér-Rao-like bound for observations where the probability density function (p.d.f.) parameterized by unknown deterministic parameters results from the marginalization of a joint p.d.f. depending on random variables as well. In this setting, it is now possible to consider random equality constraints, *i.e.*, equality constraints on the unknown deterministic parameters depending on the random parameters, which can not be addressed with the usual constrained Cramér-Rao bound. The usefulness of the proposed bound is illustrated by way of a coupled canonical polyadic model with linear constraints applied to the hyperspectral super-resolution problem.

Index Terms— Mean Squared Error, Deterministic parameters, Constrained Cramér-Rao bound, Random equality constraints

I. INTRODUCTION

As introduced in [1, p53], a model of the general deterministic estimation problem has the following four components: 1) a parameter space $\Theta_d \subset \mathbb{R}^P$, 2) an observation space $\mathcal{X} \subset \mathbb{R}^M$, 3) a probabilistic mapping from parameter vector space Θ_d to observation space $\mathcal{X}$, that is the probability law $p(\mathbf{x};\boldsymbol{\theta})$ that governs the effect of a parameter vector value $\boldsymbol{\theta} \in \Theta_d$ on the observation $\mathbf{x} \in \mathcal{X}$ and, 4) an estimation rule, that is the mapping of the observation space $\mathcal{X}$ into vector parameter estimates $\widehat{\boldsymbol{\theta}} \triangleq \widehat{\boldsymbol{\theta}}(\mathbf{x})$. If a closed-form expression of $p(\mathbf{x};\boldsymbol{\theta})$ is available, the estimation problem at hand is so-called a "standard" deterministic estimation problem [2]. In this setting, minimal performance bounds on the mean square error (MSE) matrix of $\widehat{\boldsymbol{\theta}}$ allow for calculation of the best performance that can be achieved, when estimating parameters of a signal corrupted by noise. Historically the first MSE lower bound (LB) for deterministic parameters to be derived was the Cramér-Rao bound (CRB), which was introduced to investigate fundamental limits of a parameter estimation problem or to assess the relative performance of a specific estimator (efficiency) [3]–[5]. It has since become the most popular LB due to its simplicity of calculation for various problems (see [6, §8.4] and [7, Part III]) but suffers from some drawbacks. Indeed, CRBs are asymptotically tight only (high signal-to-noise ratio (SNR) and/or large number of snapshots) and cannot predict the so-called threshold (i.e. large errors) on estimator MSE in non-linear estimation problems [8]–[11]. Therefore, provided that one keeps in mind the CRB limitations, that is, to become an overly optimistic LB when the observation conditions degrade (low SNR and/or low number of snapshots), the CRB is still a LB of great interest for system analysis and design in the asymptotic region.

Moreover, in many applications, the definition, in part or totally, of the parameter space Θ_d results from deterministic (non random) equality constraints, as mentioned in the seminal paper [12]. Since then, numerous works [13]–[17] have been devoted to extend the results introduced in [12]: 1) by providing a general reparameterization inequality and the equivalence between parameterization change and equality constraints; 2) by studying the CRB modified by constraints either required by the model or required to solve identifiability issues; 3) by investigating the use of parameters constraints from a different perspective: the value of side (*a priori*) information on estimation performance. All these works have shown the versatility of constrained Cramér-Rao bound (CCRB) for estimation performance analysis and design of a system of measurement, as highlighted in [18].

However, in many estimation problems [1], [19]–[21], the probabilistic mapping mentioned above results from a two-step probabilistic mechanism involving an additional random vector $\boldsymbol{\theta}_r$, $\boldsymbol{\theta}_r \in \Theta_r \subset \mathbb{R}^{P_r}$, that is i) $\boldsymbol{\theta} \rightarrow \boldsymbol{\theta}_r$, ii) $(\boldsymbol{\theta}, \boldsymbol{\theta}_r) \rightarrow \mathbf{x}$, leading to a compound probability distribution: $p(\mathbf{x}, \boldsymbol{\theta}_r; \boldsymbol{\theta}) = p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})\, p(\boldsymbol{\theta}_r;\boldsymbol{\theta})$. In this setting, some equality constraints on the unknown deterministic parameter vector $\boldsymbol{\theta}$ may depend on the random parameter vector $\boldsymbol{\theta}_r$, leading to random equality constraints, a case which can not be tackled with the standard form of the CCRB. It is therefore the aim of this paper to derive a CR-like bound able to take into account random equality constraints, that is a randomly constrained CRB (RCCRB).

The usefulness of the RCCRB is illustrated by way of a coupled canonical polyadic (CP) model with linear constraints applied to the hyperspectral super-resolution problem (HSR). This problem consists in fusing a multispectral data cube (MSI), which has a good spatial resolution but few spectral bands, and a hyperspectral data cube (HSI), whose spatial resolution is lower than that of the MSI. The aim of the HSR problem 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 [25] between heterogeneous datasets.

II. CRBS WITH RANDOM EQUALITY CONSTRAINTS

II-A. Background on standard CRBs

In standard deterministic estimation problems [2], the MSE matrix of $\widehat{\boldsymbol{\theta}}$ is a Gram matrix (general form of the square of a norm) [17] defined on the vector space of square integrable functions and, therefore, all known standard LBs on the MSE can be formulated as the solution of a norm minimization problem under linear constraints (LCs) [9], [10]. This formulation of LBs does not only provides a straightforward understanding of the hypotheses associated with the different LBs [9], [10], but it also allows to obtain a unique formulation of each LB in terms of a unique set of linear constraints. When the lower bound is the CRB, the set of

This work has been partially supported by the DGA/MRIS (2018.60.0072.00.470.75.01) and by the Agence Nationale de la Recherche under grants OPTIFIN (ANR-15-CE10-0007) and LeaFleT (ANR-19-CE23-0021).

linear constraints involved reduces to a set of derivative constraints [17]. Indeed, the CRB is the lowest bound on the MSE of unbiased estimators, since it is derived from the *weakest* formulation of unbiasedness, *i.e.* local unbiasedness,

$$E_{\mathbf{x};\boldsymbol{\theta}+d\boldsymbol{\theta}}\left[\widehat{\boldsymbol{\theta}}\right]=\boldsymbol{\theta}+d\boldsymbol{\theta}+\mathbf{o}\left(\|d\boldsymbol{\theta}\|\right), \tag{1a}$$

where $\mathbf{o}\left(.\right)$ stands for the small oh notation, which means that, up to the first order and in the neighborhood of $\boldsymbol{\theta}$, $\widehat{\boldsymbol{\theta}}$ remains an unbiased estimator of $\boldsymbol{\theta}$ independently of a - small - variation of the parameters. Interestingly, (1a) can be rewritten in terms of Taylor expansion of each side, and the uniqueness of Taylor expansion imposes that the following LCs

$$E_{\mathbf{x};\boldsymbol{\theta}}\left[\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right]=\mathbf{0},\ E_{\mathbf{x};\boldsymbol{\theta}}\left[(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})\frac{\partial\ln p\left(\mathbf{x};\boldsymbol{\theta}\right)}{\partial\boldsymbol{\theta}^{T}}\right]=\mathbf{I}, \tag{1b}$$

must be statisfied by any locally unbiased estimator. Then the CRB is easily obtained by using the following well known lemma on the minimization of a Gram matrix (with respect to the Löwner ordering [22, §7.7]) under LCs. Let $\mathbb{U}$ be an Euclidean vector space on the field of real numbers $\mathbb{R}$ which has a scalar product $\left\langle\ \cdot\ \middle|\ \cdot\ \right\rangle$. Let $\mathcal{C}=(\mathbf{c}_1,\ldots,\mathbf{c}_K)$ be a family of K linearly independent vectors and $\mathcal{U}=(\mathbf{u}_1,\ldots,\mathbf{u}_P)$ a family of P vectors. Then

$$\mathbf{V}^{T}\mathbf{G}(\mathcal{C})^{-1}\mathbf{V}=\min_{\mathcal{U}}\left\{\mathbf{G}(\mathcal{U})\right\}\ \text{s.t.}\ \left\langle\mathbf{u}_p\mid\mathbf{c}_k\right\rangle=\mathbf{V}_{k,p}, \tag{2}$$

where $\mathbf{G}(\mathcal{W})$ denotes the Gram matrix associated to the family of N vectors $\mathcal{W}=(\mathbf{w}_1,\ldots,\mathbf{w}_N)$ defined as $G_{n,n'}(\mathcal{W})=\left\langle\mathbf{w}_{n'}\mid\mathbf{w}_n\right\rangle$, $1\leq n,n'\leq N$. Indeed by defining $\mathcal{U}=\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}$ and $\mathcal{C}=(1,\frac{\partial\ln p(\mathbf{x};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^{T}})$, and by considering the scalar product $\left\langle f\left(\mathbf{x}\right)\mid g\left(\mathbf{x}\right)\right\rangle=E_{\mathbf{x};\boldsymbol{\theta}}\left[f\left(\mathbf{x}\right)g\left(\mathbf{x}\right)\right]$, lemma (2) can be applied with $\mathbf{V}=\left[\begin{array}{cc}\mathbf{0} & \mathbf{I}\end{array}\right]$ (1b) and leads to

$$E_{\mathbf{x};\boldsymbol{\theta}}\left[(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})^{T}\right]\geq\mathbf{CRB}\left(\boldsymbol{\theta}\right)=\mathbf{F}\left(\boldsymbol{\theta}\right)^{-1}, \tag{3a}$$

$$\mathbf{F}\left(\boldsymbol{\theta}\right)=E_{\mathbf{x};\boldsymbol{\theta}}\left[\frac{\partial\ln p\left(\mathbf{x};\boldsymbol{\theta}\right)}{\partial\boldsymbol{\theta}}\frac{\partial\ln p\left(\mathbf{x};\boldsymbol{\theta}\right)}{\partial\boldsymbol{\theta}^{T}}\right], \tag{3b}$$

where $\mathbf{F}\left(\boldsymbol{\theta}\right)$ is the Fisher information matrix (FIM). Last, it has been shown in [17] that the CRB (3a) is also obtained if (1b) is reduced to

$$E_{\mathbf{x};\boldsymbol{\theta}}\left[(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})\frac{\partial\ln p\left(\mathbf{x};\boldsymbol{\theta}\right)}{\partial\boldsymbol{\theta}^{T}}\right]=\mathbf{I}. \tag{4}$$

II-B. Random Equality Constraints

Actually, in many estimation problems [1], [19]–[21], the probabilistic mapping mentioned above results from a two steps probabilistic mechanism involving an additional random vector $\boldsymbol{\theta}_r$, $\boldsymbol{\theta}_r\in\Theta_r\subset\mathbb{R}^{P_r}$, that is i) $\boldsymbol{\theta}\rightarrow\boldsymbol{\theta}_r\sim p\left(\boldsymbol{\theta}_r;\boldsymbol{\theta}\right)$, ii) $(\boldsymbol{\theta},\boldsymbol{\theta}_r)\rightarrow\mathbf{x}\sim p\left(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}\right)$, and leading to a compound probability distribution:

$$p\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}\right)=p\left(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}\right)p\left(\boldsymbol{\theta}_r;\boldsymbol{\theta}\right), \tag{5a}$$

$$p\left(\mathbf{x};\boldsymbol{\theta}\right)=\int_{\Theta_r}p\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}\right)d\boldsymbol{\theta}_r, \tag{5b}$$

where $p\left(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}\right)$ is the conditional p.d.f. of $\mathbf{x}$ given $\boldsymbol{\theta}_r$, and $p\left(\boldsymbol{\theta}_r;\boldsymbol{\theta}\right)$ is the prior p.d.f., parameterized by $\boldsymbol{\theta}$. If only an integral form of $p\left(\mathbf{x};\boldsymbol{\theta}\right)$ (5b) is available, the estimation problem at hand is so-called a "non-standard" estimation problem [2]. In this setting,

$$E_{\mathbf{x};\boldsymbol{\theta}}\left[(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})^{T}\right]=E_{\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[E_{\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})^{T}\right]\right] \tag{6}$$

which allows to consider the addition of K non redundant equality constraints on the unknown deterministic parameter vector $\boldsymbol{\theta}$ depending on the random parameter vector $\boldsymbol{\theta}_r$, that is

$$\mathbf{f}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)=\mathbf{0},\ \mathbf{f}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)\in\mathbb{R}^{K},\ 1\leq K\leq P-1, \tag{7}$$

where the matrix $\frac{\partial\mathbf{f}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^{T}}\in\mathbb{R}^{K\times P}$ has full row rank (K), which defines K random equality constraints.

II-C. CRBs with Random Equality Constraints

Since the set of K equality constraints (7) $\mathcal{C}_{\boldsymbol{\theta}_r}\subset\Theta_d$ is conditioned on the value of $\boldsymbol{\theta}_r$, it seems sensible to first look for a CR-like bound conditioned on $\boldsymbol{\theta}_r$, taking into account both local unbiasedness and equality constraints (7). Conditionally to $\boldsymbol{\theta}_r$, that is with respect to $p\left(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}\right)$, local unbiasedness regarding the parameter vector $\boldsymbol{\theta}$ reads

$$E_{\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}+d\boldsymbol{\theta}}\left[\widehat{\boldsymbol{\theta}}\right]=\boldsymbol{\theta}+d\boldsymbol{\theta}+\mathbf{o}_{\boldsymbol{\theta}_r}\left(\|d\boldsymbol{\theta}\|\right),$$

and leads (similarly to (1a) and (4)) to the LCs

$$E_{\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[\left(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\frac{\partial\ln p\left(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}\right)}{\partial\boldsymbol{\theta}^{T}}\right]d\boldsymbol{\theta}=\mathbf{I}d\boldsymbol{\theta}. \tag{8}$$

Moreover, if $\boldsymbol{\theta}$ and $\boldsymbol{\theta}+d\boldsymbol{\theta}$ are constrained to belong to $\mathcal{C}_{\boldsymbol{\theta}_r}$, thus, with some manipulation [17], when $\|d\boldsymbol{\theta}\|\rightarrow 0$,

$$\left\{\begin{array}{c}\mathbf{f}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)=\mathbf{0}\\ \frac{\partial\mathbf{f}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^{T}}d\boldsymbol{\theta}=\mathbf{0}\end{array}\right.\Leftrightarrow\left\{\begin{array}{l}\mathbf{0}=\mathbf{f}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)\\ d\boldsymbol{\theta}=\mathbf{U}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)d\boldsymbol{\lambda}\end{array}\right.$$

where $\mathbf{U}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)\in\mathbb{R}^{P\times(P-K)}$ is a basis of $\ker\left(\frac{\partial\mathbf{f}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^{T}}\right)$ and $d\boldsymbol{\lambda}\in\mathbb{R}^{P-K}$. Therefore, conditionally to $\boldsymbol{\theta}_r$, a locally unbiased estimate of $\boldsymbol{\theta}$ is now required to be locally unbiased only on $\mathcal{C}_{\boldsymbol{\theta}_r}$, what means that LCs (8) must be satisfied only when $d\boldsymbol{\theta}=\mathbf{U}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)d\boldsymbol{\lambda}$ where $\|d\boldsymbol{\lambda}\|\rightarrow 0$, which yields the LCs

$$E_{\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})\left(\mathbf{U}_{\boldsymbol{\theta}_r}^{T}\left(\boldsymbol{\theta}\right)\frac{\partial\ln p\left(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}\right)}{\partial\boldsymbol{\theta}}\right)^{T}\right]=\mathbf{U}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right). \tag{9}$$

Additionally, another desirable property is that

$$E_{\mathbf{x}|\boldsymbol{\theta}_r+d\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[\widehat{\boldsymbol{\theta}}\right]=\boldsymbol{\theta}+\mathbf{o}\left(\|d\boldsymbol{\theta}_r\|\right),\ \forall\boldsymbol{\theta}\in\mathcal{C}_{\boldsymbol{\theta}_r}, \tag{10a}$$

which means that, up to the first order and in the neighborhood of $\boldsymbol{\theta}_r$, $\widehat{\boldsymbol{\theta}}$ remains an unbiased estimator of $\boldsymbol{\theta}\in\mathcal{C}_{\boldsymbol{\theta}_r}$ independently of a - small - variation of the parameter vector $\boldsymbol{\theta}_r$. Once again, (10a) can be rewritten in terms of the following LCs

$$E_{\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})\frac{\partial\ln p\left(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}\right)}{\partial\boldsymbol{\theta}_r^{T}}\right]=\mathbf{0}. \tag{10b}$$

Finally, conditionally to $\boldsymbol{\theta}_r$, a constrained CR-like bound fitted to the problem at hand is the lower bound associated with the LCs

$$\left\{\begin{array}{c}E_{\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})\left(\mathbf{U}_{\boldsymbol{\theta}_r}^{T}\left(\boldsymbol{\theta}\right)\frac{\partial\ln p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)^{T}\right]=\mathbf{U}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)\\ E_{\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})\frac{\partial\ln p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}_r^{T}}\right]=\mathbf{0}\end{array}\right. \tag{11}$$

that is, according to (2),

$$\mathbf{CCRB}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)= \tag{12a}$$

$$\mathbf{U}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)\left(\mathbf{U}_{\boldsymbol{\theta}_r}^{T}\left(\boldsymbol{\theta}\right)\mathbf{CRB}_{\boldsymbol{\theta}_r}^{-1}\left(\boldsymbol{\theta}\right)\mathbf{U}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)\right)^{-1}\mathbf{U}_{\boldsymbol{\theta}_r}^{T}\left(\boldsymbol{\theta}\right),$$

$$\mathbf{CRB}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)= \tag{12b}$$

$$\left(\mathbf{F}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)-\mathbf{F}_{\boldsymbol{\theta}_r}^{T}\left(\boldsymbol{\theta}_r,\boldsymbol{\theta}\right)\mathbf{F}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}_r\right)^{-1}\mathbf{F}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}_r,\boldsymbol{\theta}\right)\right)^{-1},$$

$$\mathbf{F}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}\right)=E_{\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[\frac{\partial\ln p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\frac{\partial\ln p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^{T}}\right],$$

$$\mathbf{F}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}_r\right)=E_{\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[\frac{\partial\ln p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}_r}\frac{\partial\ln p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}_r^{T}}\right],$$

$$\mathbf{F}_{\boldsymbol{\theta}_r}\left(\boldsymbol{\theta}_r,\boldsymbol{\theta}\right)=E_{\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[\frac{\partial\ln p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}_r}\frac{\partial\ln p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^{T}}\right].$$

Finally, if $\widehat{\boldsymbol{\theta}} \triangleq \widehat{\boldsymbol{\theta}}(\mathbf{x})$ is, conditionally to $\boldsymbol{\theta}_r$, a locally unbiased estimated belonging to a subset $\mathcal{C}_{\boldsymbol{\theta}_r}$ of the parameter space defined by K non redundant equality constraints depending on a random parameter vector $\boldsymbol{\theta}_r$, then, according to (6), its MSE matrix is lower bounded by the following randomly constrained CRB (RCCRB)

$$E_{\mathbf{x}|\boldsymbol{\theta}}\left[(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})^T\right] \geq \mathbf{RCCRB}(\boldsymbol{\theta}),$$
$$\mathbf{RCCRB}(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[\mathbf{CCRB}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta})\right]. \quad (13)$$

II-D. Further considerations

Firstly, if no random constraints are taken into account, then $\mathbf{U}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta}) = \mathbf{I}$ and $\mathbf{CCRB}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta}) = \mathbf{CRB}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta})$ which coincides with the tighter Non-Standard CRB ($\overline{\mathbf{NSCRB}}(\boldsymbol{\theta})$) introduced in [23] and lately generalized in [2, (54)]. Moreover, the LCs (11) becomes equivalent to

$$E_{\mathbf{x}|\boldsymbol{\theta}_r+d\boldsymbol{\theta}_r;\boldsymbol{\theta}+d\boldsymbol{\theta}}\left[\widehat{\boldsymbol{\theta}}\right] = \boldsymbol{\theta} + d\boldsymbol{\theta} + \mathbf{o}(\|(d\boldsymbol{\theta}; d\boldsymbol{\theta}_r)\|),$$

which is the definition of a locally strict-sense unbiased estimator [24]. This is sensible, since, as shown in [2, §IV], Non-Standard CRB are LBs on the "non-standard" MLEs (NSMLEs) defined as

$$(\widehat{\underline{\boldsymbol{\theta}}_r}, \underline{\widehat{\boldsymbol{\theta}}}) = \arg \max_{\boldsymbol{\theta}\in\Theta_d, \boldsymbol{\theta}_r\in\Theta_r} \{p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})\}, \quad (14)$$

where $\underline{\widehat{\boldsymbol{\theta}}}$ is, w.r.t. $p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})$ and under reasonably general conditions, asymptotically uniformly strict-sense unbiased, Gaussian distributed and efficient when the number of independent observations tends to infinity. Therefore, it seems likely that the method of scoring with parameter constraints [16] applied to random equality constraints (7) where $\boldsymbol{\theta}_r$ is replaced with its NSMLE $\widehat{\underline{\boldsymbol{\theta}_r}}$, leads to a constrained NSMLE asymptotically efficient with respect to $\mathbf{CCRB}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta})$ and hence to $\mathbf{RCCRB}(\boldsymbol{\theta})$. A conjecture left for future research. Secondly, in general,

$$\mathbf{CCRB}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta}) > \mathbf{U}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta})\left(\mathbf{U}_{\boldsymbol{\theta}_r}^T(\boldsymbol{\theta})\mathbf{F}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta})\mathbf{U}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta})\right)^{-1}\mathbf{U}_{\boldsymbol{\theta}_r}^T(\boldsymbol{\theta})$$

which means that the RCCRB proposed (13) is tighter than the expectation of the standard CCRB parameterized by $\boldsymbol{\theta}_r$ (see (15)). However, in the case where $p(\mathbf{x}, \boldsymbol{\theta}_r;\boldsymbol{\theta}) = p(\mathbf{x};\boldsymbol{\theta})\,p(\boldsymbol{\theta}_r;\boldsymbol{\theta})$, then

$$p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta}) = p(\mathbf{x};\boldsymbol{\theta}) \Rightarrow \mathbf{CRB}_{\boldsymbol{\theta}_r}^{-1}(\boldsymbol{\theta}) = \mathbf{F}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta}) = \mathbf{F}(\boldsymbol{\theta}),$$

where $\mathbf{F}(\boldsymbol{\theta})$ is the standard FIM (3b), leading to

$$\mathbf{RCCRB}(\boldsymbol{\theta}) =$$
$$E_{\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[\mathbf{U}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta})\left(\mathbf{U}_{\boldsymbol{\theta}_r}^T(\boldsymbol{\theta})\mathbf{F}(\boldsymbol{\theta})\mathbf{U}_{\boldsymbol{\theta}_r}(\boldsymbol{\theta})\right)^{-1}\mathbf{U}_{\boldsymbol{\theta}_r}^T(\boldsymbol{\theta})\right], \quad (15)$$

which reduces to the standard CCRB

$$\mathbf{CCRB}(\boldsymbol{\theta}) = \mathbf{U}(\boldsymbol{\theta})\left(\mathbf{U}^T(\boldsymbol{\theta})\mathbf{F}(\boldsymbol{\theta})\mathbf{U}(\boldsymbol{\theta})\right)^{-1}\mathbf{U}^T(\boldsymbol{\theta}),$$

if the K equality constraints (7) are non random.

III. APPLICATION TO THE HSR PROBLEM

Recently, tensor-based methods, using the inherent 3D nature of the HSI, were proposed for solving the HSR problem [29], [30]. In [29], the problem is reformulated as a coupled CP approximation, assuming that the SRI itself admits a low-rank canonical polyadic decomposition (CPD). An alternating least squares (ALS) algorithm called Super-resolution TEnsor REconstruction (STEREO) was also proposed. In hyperspectral imaging, each pixel of a data cube can be seen as the spectral signature of the material present at this specific location. Spectral signatures are of prime importance in material identification or object detection in hyperspectral remote sensing. In the HSR problem, if the acquisitions of the MSI and HSI are not performed simultaneously, it may happen that the illumination conditions vary [34]. In [26]–[28], derivation of the CRB for uncoupled CP models have been provided. In [31], it was proposed to explore the CCRB for complex tensors and partially coupled CP decompositions with (possibly) non-linear couplings. In this section we mainly follow [35], [36] in what concerns the tensor notation. We use the symbol $\boxtimes$ and $\odot$ for the Kronecker and Khatri-Rao product, respectively. We use $\mathrm{vec}\{\cdot\}$ for the standard column-major vectorization of a tensor or a matrix. The operation $\bullet_p$ denotes contraction on the pth index of a tensor; for instance, $[\boldsymbol{\mathcal{A}} \bullet_1 \boldsymbol{M}]_{ijk} = \sum_\ell \mathcal{A}_{\ell jk} M_{i\ell}$. For the matrices $\boldsymbol{A} \in \mathbb{R}^{I\times N}$, $\boldsymbol{B} \in \mathbb{R}^{J\times N}$, $\boldsymbol{C} \in \mathbb{R}^{K\times N}$, we will use $[\![\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}]\!]$ as a shorthand notation for the CPD. For a tensor $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{I\times J\times K}$, its first unfolding is denoted by $\boldsymbol{Y}^{(1)} \in \mathbb{R}^{JK\times I}$. The notation $\|\cdot\|_F$ stands for the Frobenius norm.

III-A. CP-based degradation model

We consider two tensors $\boldsymbol{\mathcal{Y}}_1 \in \mathbb{R}^{I_H\times J_H\times K}$ and $\boldsymbol{\mathcal{Y}}_2 \in \mathbb{R}^{I\times J\times K_M}$, denoting respectively an HSI and a MSI cube. 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 the 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 $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ are viewed as two degraded versions of a single super-resolution image (SRI) $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{I\times J\times K}$. As in [29], we adopt the following degradation model as contraction of the SRI:

$$\begin{cases} \boldsymbol{\mathcal{Y}}_1 &= \boldsymbol{\mathcal{Y}} \bullet_1 \boldsymbol{P} \bullet_2 \boldsymbol{Q} + \boldsymbol{\mathcal{E}}_1, \\ \boldsymbol{\mathcal{Y}}_2 &= |\alpha|\, \boldsymbol{\mathcal{Y}} \bullet_3 \boldsymbol{R} + \boldsymbol{\mathcal{E}}_2, \end{cases} \quad (16)$$

where $\boldsymbol{P} \in \mathbb{R}^{I_H\times I}$, $\boldsymbol{Q} \in \mathbb{R}^{J_H\times J}$ and $\boldsymbol{R} \in \mathbb{R}^{K_M\times K}$ are known degradation matrices[1]. The scalar parameter α is random real variable following a given probability distribution, and stands for the illumination coefficient between the HSI and MSI. We consider the case where the first rows of $\boldsymbol{P}$, $\boldsymbol{Q}$ and $\boldsymbol{R}$ are equal to the first rows of the identity matrix. The entries of the noise terms $\boldsymbol{\mathcal{E}}_1 \sim \boldsymbol{\mathcal{N}}(0, \sigma_1^2\boldsymbol{I})$, $\boldsymbol{\mathcal{E}}_2 \sim \boldsymbol{\mathcal{N}}(0, \sigma_2^2\boldsymbol{I})$ are independent and identically distributed (i.i.d.) real Gaussian variables with zero mean and variance $\sigma_1^2\boldsymbol{I}$ and $\sigma_2^2\boldsymbol{I}$, respectively. Under the assumption that the SRI admits a low rank-N CPD, the degradation model (16) becomes

$$\begin{cases} \boldsymbol{\mathcal{Y}}_1 &= [\![\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1]\!] + \boldsymbol{\mathcal{E}}_1, \\ \boldsymbol{\mathcal{Y}}_2 &= [\![\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_2]\!] + \boldsymbol{\mathcal{E}}_2, \end{cases} \quad (17)$$
$$\text{subject to } \boldsymbol{A}_1 = \boldsymbol{P}\boldsymbol{A}_2, \boldsymbol{B}_1 = \boldsymbol{Q}\boldsymbol{B}_2, \text{ and } \boldsymbol{C}_2 = |\alpha|\,\boldsymbol{R}\boldsymbol{C}_1,$$

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

We wish to estimate the factor matrices of the CPD of $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$. Thus, we define the model parameters

$$\boldsymbol{\psi}_1 = \mathrm{vec}\{\boldsymbol{C}_1\} \in \mathbb{R}^{KN}, \qquad \boldsymbol{\phi}_1 = \begin{bmatrix}\mathrm{vec}\{\boldsymbol{A}_1\}\\ \mathrm{vec}\{\boldsymbol{B}_1\}\end{bmatrix} \in \mathbb{R}^{(I_H+J_H)N},$$
$$\boldsymbol{\psi}_2 = \mathrm{vec}\{\boldsymbol{C}_2\} \in \mathbb{R}^{K_M N}, \quad \boldsymbol{\phi}_2 = \begin{bmatrix}\mathrm{vec}\{\boldsymbol{A}_2\}\\ \mathrm{vec}\{\boldsymbol{B}_2\}\end{bmatrix} \in \mathbb{R}^{(I+J)N},$$

corresponding to the vectorization of the factor matrices for each tensor. The choice to group $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ factors, separate from the $\boldsymbol{C}_i$ ($i = 1, 2$) factor matrices is motivated by the fact that spatial and spectral degradations never occur in the same tensor, according to model (16). Thus, $\boldsymbol{\psi}_2$ and $\boldsymbol{\phi}_1$ can be seen as degraded versions of $\boldsymbol{\psi}_1$ and $\boldsymbol{\phi}_2$ by the spectral and spatial degradation matrices, respectively. In the problem at hand, $\boldsymbol{\theta} = \begin{bmatrix}\boldsymbol{\psi}_1^\mathsf{T} & \boldsymbol{\phi}_1^\mathsf{T} & \boldsymbol{\psi}_2^\mathsf{T} & \boldsymbol{\phi}_2^\mathsf{T}\end{bmatrix}^\mathsf{T}$ is the vector of unknown deterministic parameters to be estimated

[1]We suppose that the spatial degradation for the HSI is separable.

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

$$\begin{cases} f_{\boldsymbol{\mathcal{Y}}_1;\boldsymbol{\theta}_1,\boldsymbol{\phi}_1} &= \left(2\pi\sigma_1^2\right)^{\frac{-I_H J_H K}{2}} e^{-\frac{\|\boldsymbol{\mathcal{Y}}_1 - [\![\boldsymbol{A}_1,\boldsymbol{B}_1,\boldsymbol{C}_1]\!]\|_F^2}{2\sigma_1^2}} \\ f_{\boldsymbol{\mathcal{Y}}_2;\boldsymbol{\theta}_2,\boldsymbol{\phi}_2} &= \left(2\pi\sigma_2^2\right)^{\frac{-IJK_M}{2}} e^{-\frac{\|\boldsymbol{\mathcal{Y}}_2 - [\![\boldsymbol{A}_2,\boldsymbol{B}_2,\boldsymbol{C}_2]\!]\|_F^2}{2\sigma_2^2}} \end{cases} \tag{18}$$

III-B. Performance analysis

Here, we focus on the case where the CP models are both identifiable. In fact, to solve the scaling indeterminacy of the CPD [39], we need to fix the first rows of $\boldsymbol{A}_i$, $\boldsymbol{B}_i$ $(i = 1, 2)$ to known values; here, we normalize the factors by setting the first rows to ones. As a result, we define the parameters $\widetilde{\boldsymbol{\phi}}_1$ and $\widetilde{\boldsymbol{\phi}}_2$, only composed of the unknown entries of $\boldsymbol{\phi}_1$ (resp. $\boldsymbol{\phi}_2$). The matrices $\boldsymbol{M}_1 \in \mathbb{R}^{(I_H+J_H-2)N\times(I_H+J_H)N}$ and $\boldsymbol{M}_2 \in \mathbb{R}^{(I+J-2)N\times(I+J)N}$ are masks obtained by removing the $2N$ entries corresponding to known entries of $\boldsymbol{\phi}_1$ (resp. $\boldsymbol{\phi}_2$) such that $\widetilde{\boldsymbol{\phi}}_1 = \boldsymbol{M}_1\boldsymbol{\phi}_1$ and $\widetilde{\boldsymbol{\phi}}_2 = \boldsymbol{M}_2\boldsymbol{\phi}_2$. Thus, we can define the deterministic parameter $\widetilde{\boldsymbol{\theta}} = \begin{bmatrix}\boldsymbol{\psi}_1^\mathsf{T} & \widetilde{\boldsymbol{\phi}}_1^\mathsf{T} & \boldsymbol{\psi}_2^\mathsf{T} & \widetilde{\boldsymbol{\phi}}_2^\mathsf{T}\end{bmatrix}^\mathsf{T}$. We consider the uncoupled CP model

$$\boldsymbol{\mathcal{X}} = \begin{bmatrix}\mathrm{vec}\{\boldsymbol{\mathcal{Y}}_1\} \\ \mathrm{vec}\{\boldsymbol{\mathcal{Y}}_2\}\end{bmatrix} \sim \boldsymbol{\mathcal{N}}(\mu(\widetilde{\boldsymbol{\theta}}), \boldsymbol{\Sigma}), \mu(\widetilde{\boldsymbol{\theta}}) = \begin{bmatrix}\mathrm{vec}\{[\![\boldsymbol{A}_1,\boldsymbol{B}_1,\boldsymbol{C}_1]\!]\} \\ \mathrm{vec}\{[\![\boldsymbol{A}_2,\boldsymbol{B}_2,\boldsymbol{C}_2]\!]\}\end{bmatrix}$$

and $\boldsymbol{\Sigma} = \mathrm{Diag}\{\sigma_1^2\boldsymbol{I}, \sigma_2^2\boldsymbol{I}\}$. The FIM on $\widetilde{\boldsymbol{\theta}}$ in the uncoupled case is given by the Slepian-Bangs formula [38]:

$$\boldsymbol{F}(\widetilde{\boldsymbol{\theta}}) = \left[\frac{\partial\mu(\widetilde{\boldsymbol{\theta}})}{\partial\widetilde{\boldsymbol{\theta}}^\mathsf{T}}\right]^\mathsf{T} \boldsymbol{\Sigma}^{-1} \left[\frac{\partial\mu(\widetilde{\boldsymbol{\theta}})}{\partial\widetilde{\boldsymbol{\theta}}^\mathsf{T}}\right].$$

For a full derivation of the FIM, see [31, Section IV]. Following model (17), we can define the equality constraints on $\widetilde{\boldsymbol{\theta}}$ as

$$\mathbf{f}_\alpha(\widetilde{\boldsymbol{\theta}}) = \begin{cases} |\alpha|\,(\boldsymbol{I}\boxtimes\boldsymbol{R})\boldsymbol{\psi}_1 - \boldsymbol{\psi}_2, \\ \widetilde{\boldsymbol{\phi}}_1 - \underbrace{\boldsymbol{M}_1\begin{bmatrix}\boldsymbol{I}\boxtimes\boldsymbol{P} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}\boxtimes\boldsymbol{Q}\end{bmatrix}\boldsymbol{M}_2^\mathsf{T}}_{\boldsymbol{\mathcal{P}}}\widetilde{\boldsymbol{\phi}}_2. \end{cases} \tag{19}$$

Taking the derivative of $\mathbf{f}_\alpha(\widetilde{\boldsymbol{\theta}})$ and setting it equal to zero yields

$$\frac{\partial\mathbf{f}_\alpha(\widetilde{\boldsymbol{\theta}})}{\partial\widetilde{\boldsymbol{\theta}}^T} = \begin{bmatrix}|\alpha|\,(\boldsymbol{I}\boxtimes\boldsymbol{R}) & \boldsymbol{0} & -\boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} & \boldsymbol{0} & -\boldsymbol{\mathcal{P}}\end{bmatrix}.$$

The expression for $\mathbf{RCCRB}(\widetilde{\boldsymbol{\theta}})$ can then simply be obtained from $\boldsymbol{F}(\widetilde{\boldsymbol{\theta}})$ and any basis $\boldsymbol{U}_\alpha$ of $\ker\left(\frac{\partial\mathbf{f}_\alpha(\widetilde{\boldsymbol{\theta}})}{\partial\widetilde{\boldsymbol{\theta}}^T}\right)$ by plugging them into (15). In this case, the standard MLE $\widehat{\widetilde{\boldsymbol{\theta}}} = \arg\max_{\widetilde{\boldsymbol{\theta}}}\left\{p\left(\boldsymbol{\mathcal{X}};\widetilde{\boldsymbol{\theta}}\right)\right\}$ is asymptotically locally strict-sense unbiased and verifies (13).

III-C. Simulations

In this section, we simulate the performance of the estimation of $\boldsymbol{\theta}$ under additive Gaussian noise in the uncoupled case and in the constrained case with random equality constraints. We consider that $\alpha \sim \boldsymbol{\mathcal{N}}(0, \sigma_\alpha^2)$ is a real Gaussian variable. We compare the MSE of the estimators to the bounds presented in Section III. For basic tensor operations, we use TensorLab 3.0 [40]. The model parameters are retrieved using MLE. In the uncoupled case, an ALS algorithm [41] is used. The factor matrices are initialized randomly. In the case where the factor matrices are linked through random equality constraints, we use STEREO, the algorithm proposed in [29], with a regularization parameter $\lambda = \frac{\sigma_1^2}{\sigma_2^2}$. The factor matrices estimated by the uncoupled ALS algorithm are used as initialization for the randomly constrained case to speed up the convergence of STEREO. For the CP model to be identifiable, the scaling and permutation ambiguities are corrected by setting the first rows of the factors $\boldsymbol{A}_1$, $\boldsymbol{B}_1$, $\boldsymbol{A}_2$, $\boldsymbol{B}_2$ to ones and searching for the best permutation of $\boldsymbol{C}_2$ with fixed $\boldsymbol{C}_1$. We consider that $I = J = K = 15$, $I_H = J_H = 10$, and $K_M = 9$. The elements of all CP factors are realizations of i.i.d. real standard Gaussian variables, and first rows of $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{B}_1, \boldsymbol{B}_2$ are set to 1. The CP factor matrices are fixed for all Monte-Carlo runs. The tensor rank for both HSI and MSI is $N = 3$. The degradation matrices $\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{R}$ 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 that the coupling constraints in (19) are linear. We consider different values for σ_α^2 such that it varies from 2 to $2\cdot10^{-3}$. The SNR of $\boldsymbol{\mathcal{Y}}_1$ varies from 5 to 60dB, while the SNR of $\boldsymbol{\mathcal{Y}}_2$ is fixed to 20dB. For each value of σ_α^2, we first compute $\mathbf{CRB}_\alpha(\widetilde{\boldsymbol{\theta}})$ and $\mathbf{CCRB}_\alpha(\widetilde{\boldsymbol{\theta}})$ by averaging over 200 realizations of α. We then compute $\mathbf{RCCRB}(\widetilde{\boldsymbol{\theta}})$ by averaging the conditional CCRB for all values of σ_α^2. We also evaluate the total MSE on the parameters for each value of σ_α^2 by averaging the squared errors over 500 noise realizations. In Fig. 1.a, we show the limitations of the deterministic conditional CCRB. We plot the MSE given by uncoupled ALS and STEREO, as well as the uncoupled CRB on a semi-logarithmic scale. We also display $\mathbf{CCRB}_\alpha(\widetilde{\boldsymbol{\theta}})$ for two values of σ_α^2: 2 and $2\cdot10^{-3}$. Since the illumination coefficient α mostly impacts the estimation of $\boldsymbol{\psi}_2$, we choose to display its performance bounds separately and group $\boldsymbol{\psi}_1$, $\widetilde{\boldsymbol{\phi}}_1$ and $\widetilde{\boldsymbol{\phi}}_2$ together. We can see that the

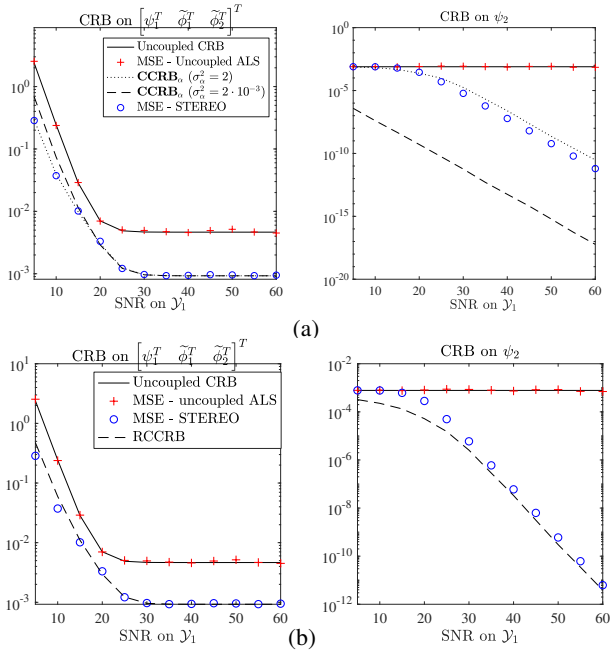

Fig. 1. Total MSE, CRB, conditional CCRB (1a) and RCCRB (1b) for the estimation of real coupled CP models on a semi-log scale

MSE given by the uncoupled ALS algorithm reaches the uncoupled CRB for all parameters. For $\boldsymbol{\psi}_2$, the MSE given by STEREO does not fit any of the two conditional CCRB depicted. In particular, it is slightly below $\mathbf{CCRB}_\alpha(\boldsymbol{\psi}_2)$ for $\sigma_\alpha^2 = 2$ and above the conditional CCRB for $\sigma_\alpha^2 = 2\cdot10^{-3}$. For the other parameters, while the MSE of STEREO fits the conditional CCRB for $\sigma_\alpha^2 = 2$, it is slightly below the second CCRB curve for a low SNR. This plot justifies the introduction of the proposed bound. In Fig. 1.b, we plot the MSE, the uncoupled CRB and the RCCRB on a semi-log scale. Here, we can see that the MSE on $\boldsymbol{\psi}_2$ given by STEREO follows the RCCRB with a small gap for a SNR on $\boldsymbol{\mathcal{Y}}_1$ between 5 and 15dB, and reached the bounds for high SNR. For the other parameters, the MSE on STEREO reaches the proposed bound. These results illustrate the usefulness of the RCCRB in this case.

IV. REFERENCES

[1] H.L. Van Trees, "Detection, Estimation and Modulation Theory, Part 1", New York, Wiley, 1968

[2] N. Kbayer, J. Galy, E. Chaumette, F. Vincent, A. Renaux and P. Larzabal, "On Lower Bounds for Non-Standard Deterministic Estimation", *IEEE Trans. on SP*, 65(6): 1538–1553, 2017

[3] M. Fréchet, "Sur l'extension de certaines évaluations statistiques au cas de petits échantillons", *Rev. Int. Stat*, 11: 182–205, 1943

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

[5] C.R. Rao, "Information and accuracy attainable in the estimation of statistical parameters", *Bull. Calcutta Math. Soc.*, 37: 81–91, 1945

[6] H.L. Van Trees, "Optimum Array Processing", New-York, Wiley-Interscience, 2002

[7] H. L. Van Trees and K. L. Bell, Eds., *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, Wiley/IEEE Press, 2007

[8] R. McAulay and L.P. Seidman, "A useful form of the Barankin lower bound and its application to PPM threshold analysis", *IEEE Trans. on IT*, 15(2): 273–279, 1969

[9] F.E. Glave, "A new look at the Barankin Lower Bound", *IEEE Trans. on IT*, 18(3): 349–356, 1972

[10] E. Chaumette, J. Galy, A. Quinlan, P. Larzabal, "A New Barankin Bound Approximation for the Prediction of the Threshold Region Performance of MLEs", *IEEE Trans. on SP*, 56(11): 5319–5333, 2008

[11] K. Todros and J. Tabrikian, "General Classes of Performance Lower Bounds for Parameter Estimation-Part I: Non-Bayesian Bounds for Unbiased Estimators", *IEEE Trans. on IT*, 56(10): 5064–5082, 2010

[12] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Trans. on IT*, vol. 26, no. 6, pp.1285–1301, Nov. 1990

[13] 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, Jun. 1993.

[14] P. Stoica and T. L. Marzetta, "Parameter estimation problems with singular information matrices", *IEEE Trans. on SP*, vol. 49, no1, pp. 87–90, Jan. 2001

[15] P. Stoica and B. C. Ng, " On the Cramér-Rao bound under parametric constraints", *IEEE SP Letters*, vol. 5, no. 7, pp. 177–179, Jul. 1998

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

[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*, 60(3): 1032–1049, 2012

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

[19] R. W. Miller and C. B. Chang, "A modified Cramér-Rao bound and its applications," *IEEE Trans. on IT*, 24(3): 398–400, 1978

[20] Y. Rockah and P. Schultheiss, "Array shape calibration using sources in unknown locations-part I: Far-field sources," *IEEE Trans. on ASSP*, 35(3): 286–299, 1987

[21] A. N. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramér-Rao bound and its application to synchronization problems," *IEEE Trans. on Commun.*, 42(2/3/4): 1391–1399, 1994

[22] R. A. Horn, C. R. Johnson, "Matrix Analysis (2nd Ed)", Cambridge University Press, 2013.

[23] F. Gini and R. Reggiannini, "On the Use of Cramér-Rao-Like Bounds in the Presence of Random Nuisance Parameters", *IEEE Trans. on Commun.*, 48(12): 2120–2126, 2000

[24] Y. Noam and H. Messer, "Notes on the tightness of the hybrid Cramér- Rao lower bound," *IEEE Trans. on SP*, 57(6): 2074–2084, 2009

[25] 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.

[26] S. Sahnoun and P. Comon, "Joint source estimation and localization", *IEEE Trans. Signal Process.*, vol. 63, no. 10, pp. 2485–2595, 2015.

[27] X. Liu and N.D. Sidiropoulos, "Cramér-Rao lower bounds for low-rank decomposition of multidimensionnal arrays", *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 2074–2086, 2001.

[28] 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*.

[29] 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.

[30] C. Prevost, K. Usevich, P. Comon, and D. Brie, "Coupled tensor low-rank multilinear approximation for hyperspectral super-resolution", in *Proc. ICASSP, 2019*.

[31] C. Ren, R. Cabral Farias, P-O. Amblard, and P. Comon, "Performance bounds for coupled models", in *Proc. 2016 IEEE SAM*, 2016.

[32] C. Prevost, K. Usevich, M. Haardt, P. Comon, and D. Brie, "Performance bounds for coupled CP model in the framework of hyperspectral super-resolution", in *Proc. CAMSAP, 2019*, available online at https://hal.archives-ouvertes.fr/hal-02303132.

[33] D. Manolakis, D. Marden and G.A. Shawn "Hyperspectral image processing for automatic target detection applications", *Lincoln laboratory journal*, 14(1), 79–116, 2003.

[34] A. Camacho, C.V. Correa and H. Arguello, "An analysis of spectral variability in hyperspectral imagery: a case study of stressed oil palm detection in Colombia", *International Journal of Remote Sensing*, 1–21, 2019.

[35] P. Comon, 'Tensors: A brief introduction", *IEEE Signal Process. Mag.*, vol. 31, no. 3, pp. 44–53, 2014.

[36] T.G. Kolda and B.W. Bader, "Tensor decompositions and applications", *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.

[37] 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.

[38] 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.

[39] 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.

[40] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, "Tensorlab 3.0," Mar. 2016, Available online.

[41] R. Bro, "Multi-way analysis in the food industry: Models, algorithms, and applications," *Ph.D. dissertation*, University of Amsterdam, The Netherlands, 1988.