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 (CRB) 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 existing forms of the constrained CRB. 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 \( X \subset \mathbb{R}^M \), 3) a probabilistic mapping from parameter vector space \( \Theta_d \) to observation space \( X \), that is the probability law \( p(x; \theta) \) that governs the effect of a parameter vector value \( \theta \in \Theta_d \) on the observation \( x \in X \) and, 4) an estimation rule, that is the mapping of the observation space \( X \) into vector parameter estimates \( \hat{\theta} \xrightarrow{\Delta} \theta(x) \).

If a closed-form expression of \( p(x; \theta) \) 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 \( \theta \) 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 \( \Theta_d \) 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 steps probabilistic mechanism involving an additional random vector \( \Theta_r \), \( \theta_r \in \Theta_r \subset \mathbb{R}^{F_r} \), that is i) \( \theta \rightarrow \Theta_r \), ii) \( \Theta_r \rightarrow \Theta \), leading to a compound probability distribution: \( p(x; \theta; \theta_r) = p(x|\theta_r; \theta) p(\theta_r; \Theta_r) \). In this setting, some equality constraints on the unknown deterministic parameter vector \( \theta \) may depend on the random parameter vector \( \theta_r \), leading to equality constraints, a case which can not be tackled with the existing forms (standard or hybrid [22]) 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 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 \( \hat{\theta} \) 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) [2]. [10]). This formulation of LBs not only provides a straightforward understanding of the hypotheses associated with the different LBs [2], [10], but 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 linear
constrained variables involves the introduction 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,

$$E_{\theta + \delta \theta} \left[ \hat{\theta} - \theta + \delta \theta + \alpha (|\delta \theta|) \right] = 0 \quad \text{(1a)}$$

where $\alpha (\cdot)$ denotes the smallest oh notation, which means that, up to the first order and in the neighborhood of $\theta$, $\hat{\theta}$ remains an unbiased estimator of $\theta$ 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 holds:

$$E_{x, \theta} \left[ \hat{\theta} - \theta \right] = 0, \quad E_{x, \theta} \left[ (\hat{\theta} - \theta) \frac{\partial \ln p(x; \theta)}{\partial \theta} \right] = \mathbf{I}, \quad \text{(1b)}$$

must be satisfied 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 [23] §7.1) under LCs. Let $\mathbb{U}$ be an Euclidean vector space on the field of real numbers $\mathbb{R}$ which has a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathcal{C} = \{ x_1, \ldots, x_N \}$ be a family of $K$ linearly independent vectors and $U = \{ u_1, \ldots, u_N \}$ a family of $P$ vectors. Then

$$V^T G (V) ^{-1} V = \min_{U} \{ G(U) \} \quad \text{under} \quad \langle u_p, c_k \rangle = V_{k,p}, \quad \text{(2)}$$

where $G(W)$ denotes the Gram matrix associated to the family of $N$ vectors $W = \{ w_1, \ldots, w_N \}$ defined as $G_{n,m}(W) = \langle w_n, w_m \rangle$, $1 \leq n, m \leq N$. Indeed by defining $U = \hat{\theta} - \theta$ and $C = (1, \partial \ln p(x; \theta) / \partial \theta^T)$, and by considering the scalar product $\langle f(x), g(x) \rangle = E_{x, \theta} \left[ f(x) g(x) \right]$, lemma [2] can be applied for $V = \mathbf{0} \mathbf{1}$ [18] and leads to

$$E_{x, \theta} \left[ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right] \geq \text{CRB} (\theta) = F (\theta)^{-1} \quad \text{(3a)}$$

$$F (\theta) = E_{x, \theta} \left[ \frac{\partial \ln p(x; \theta)}{\partial \theta} \frac{\partial \ln p(x; \theta)}{\partial \theta} \right] \quad \text{(3b)}$$

where $F (\theta)$ 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

$$E_{x, \theta} \left[ (\hat{\theta} - \theta) \frac{\partial \ln p(x; \theta)}{\partial \theta} \right] = \mathbf{I}. \quad \text{(4)}$$

### II-B. 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 $\theta_r$, $\theta_r \in \Theta_r \subset \mathbb{R}^{P_r}$, that is i) $\theta \rightarrow \theta_r \sim p(\theta_r)$, ii) $(\theta, \theta_r) \rightarrow x \sim p(x|\theta, \theta_r)$, and leading to a conditional probability distribution:

$$p(x; \theta, \theta_r) = p(x|\theta_r) p(\theta_r), \quad \text{(5a)}$$

$$p(x; \theta) = \int_{\theta_r} p(x|\theta_r) d\theta_r, \quad \text{(5b)}$$

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

$$E_{x, \theta} \left[ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right] = E_{\theta_r, \theta} \left[ E_{x, \theta, \theta_r} \left[ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right] \right] \quad \text{(6)}$$

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

$$f_{\theta_r}(\theta) = 0, \quad f_{\theta_r}(\theta) \in \mathbb{R}^K, \quad 1 \leq K \leq P - 1, \quad \text{(7)}$$

where the matrix $\partial f_{\theta_r}(\theta) / \partial \theta^T \in \mathbb{R}^{K \times P}$ has full row rank ($K$), which defines $K$ random equality constraints.

### II-C. CRBs with Random Equality Constraints

Since the subset of $\Theta_\theta$ defined by the $K$ equality constraints [7], i.e. $\Theta_\theta = \Theta_\theta \cap \Theta$, is conditioned on the value of $\theta_r$, it seems sensible to first look for a CR-like bound conditioned on $\theta_r$, taking into account both local unbiasedness and equality constraints [7]. Conditionally to $\theta_r$, that is with respect to $p(x|\theta_r; \theta)$, local unbiasedness regarding the parameter vector $\theta$ reads

$$E_{x, \theta, \theta_r + \delta \theta} \left[ \hat{\theta} \right] = \theta + \delta \theta + o_{\theta_r}, \quad \text{vec} |\delta \theta|$$

and leads (similarly to (1a) and (2)) to the LCs

$$E_{x, \theta, \theta_r} \left[ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right] \frac{\partial \ln p(x|\theta_r; \theta)}{\partial \theta^T} \right] d\theta = \text{Id}_\theta \quad \text{(8)}$$

Moreover, if $\theta$ and $\theta + \delta \theta$ are constrained to belong to $\Theta_\theta$, thus, with some manipulation [17], when $||\delta \theta|| \rightarrow 0$,

$$\left\{ \begin{array}{l} f_{\theta_r}(\theta) = 0 \\ \partial f_{\theta_r}(\theta) / \partial \theta^T = 0 \end{array} \Rightarrow \left\{ \begin{array}{l} 0 = f_{\theta_r}(\theta) \\ d\theta = U_{\theta_r}(\theta) d\lambda \end{array} \right. \right.$$
defined by $K$ non-redundant equality constraints depending on a random parameter vector $\theta$, and, according to [6], its MSE matrix is lower bounded by the following randomly constrained CRB (RCCRB)

$$E_{x,\theta} \left[ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right] \geq \text{RCCRB} (\theta),$$

$$\text{RCCRB} (\theta) = E_{\theta,\theta} \left[ \text{CCRB}_{\theta,\theta} (\theta) \right].$$

(13)

II-D. Further considerations

Firstly, if no random constraints are taken into account, then $U_{\theta,\theta} (\theta) = \mathbf{I}$ and $\text{CCRB}_{\theta,\theta} (\theta) = \text{CRB}_{\theta,\theta} (\theta)$ which coincides with the tighter non-standard CRB (NSCRB) introduced in [24] and lately generalized in [2] (54). Moreover, the LCs [11] become equivalent to

$$E_{x,\theta} \left[ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right] = \sigma^2 \mathbf{I} + o \left( \sigma^4 \right),$$

(14)

where $\sigma^2$ is, w.r.t. $p(x|\theta; \theta)$ 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 [17] where $\theta_i$ is replaced with its NSMLE $\hat{\theta}_i$, leads to a constrained NSMLE asymptotically efficient with respect to $\text{CCRB}_{\theta,\theta} (\theta)$ and hence to $\text{RCCRB} (\theta)$. A conjecture left for future research. Secondly, in general,

$$\text{CCRB}_{\theta,\theta} (\theta) > U_{\theta,\theta} (\theta) \left( U_{\theta,\theta} (\theta) F_{\theta,\theta} (\theta) U_{\theta,\theta} (\theta) \right)^{-1} U_{\theta,\theta}^T (\theta),$$

(15)

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

$$p(x|\theta; \theta) = p(x; \theta) \Rightarrow \text{CCRB}_{\theta,\theta}^{-1} (\theta) = F_{\theta,\theta} (\theta) = F (\theta),$$

where $F (\theta)$ is the standard FIM [56], leading to

$$\text{RCCRB} (\theta) = E_{\theta,\theta} \left[ U_{\theta,\theta} (\theta) \left( U_{\theta,\theta}^T (\theta) F (\theta) U_{\theta,\theta} (\theta) \right)^{-1} U_{\theta,\theta}^T (\theta) \right],$$

(16)

which reduces to the standard CCRB

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

if the $K$ equality constraints [17] are non random.

III. APPLICATION TO THE HSR PROBLEM

Recently, tensor-based methods, using the inherent 3D nature of the HSI, were proposed [30, 31] for solving the HSR problem. In [30], 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 STEREO (Super-resolution TEStor REConstruction) 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 [35]. This can make the material identification difficult, since it is then unfeasible to acquire identical spectral signatures for the same material at different pixel locations [34].

In [27, 28], derivation of the CRB for uncoupled CP models have been provided. In [32], it was proposed to explore the CRB for complex tensors and partially coupled CP decompositions with (possibly) non-linear couplings.

In this section, we mainly follow [36, 37] in what concerns the tensor notation. We use the symbol $\otimes$ for the Kronecker and Khatri-Rao product, respectively. 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, $[A \bullet_p M_{1:p}] = \sum_i A_{ij} M_{ik}$. For the matrices $A \in \mathbb{R}^{I \times J}$, $B \in \mathbb{R}^{J \times K}$, $C \in \mathbb{R}^{K \times N}$, we will use a shorthand notation for the CPD:

$$[A, B, C] = [Z_N; A, B, C],$$

where $Z_N \in \mathbb{R}^{N \times N \times N}$ is a diagonal tensor of ones. For a tensor $Y \in \mathbb{R}^{I \times J \times K}$, its first unfolding is denoted by $Y^{(1)} \in \mathbb{R}^{K \times I \times J}$. The notation $\| \cdot \|_r$ stands for the Frobenius norm.

III-A. CP-based degradation model

We consider two tensors $Y_1 \in \mathbb{R}^{I \times J \times K}$ and $Y_2 \in \mathbb{R}^{J \times K \times M}$, denoting respectively an HSI and a MSI cube. While $I, J, H$ and $J, K, M$ denote the size of the images in the spatial dimensions, $K$ and $M$ stand for the size of the data cubes in the spectral dimensions. The spectral resolution of MSI is lower ($K_M \ll K$), while its spatial resolution is higher ($I > I_H, J > J_H$). The acquired MSI and HSI usually represent the same target, and $Y_1$ and $Y_2$ are viewed as two degraded versions of a single super-resolution image (SRI) $Y \in \mathbb{R}^{I \times J \times K}$. As in [30], we adopt the following degradation model (contraction of SRI):

$$Y_1 = Y \bullet [P \circ Q] + \varepsilon_1,$n
$$Y_2 = Y \bullet [A \circ R] + \varepsilon_2,$n

(16)

subject to $A_1 = PA_2, B_1 = QB_2,$ and $C_2 = \alpha R C_1,$

$$Y_1 = \left[ A_1, B_1, C_1 \right] + \varepsilon_1,$n
$$Y_2 = \left[ A_2, B_2, C_2 \right] + \varepsilon_2,$n

(17)

where $A_1 \in \mathbb{R}^{I \times J \times K}, B_1 \in \mathbb{R}^{J \times K \times N}, C_1 \in \mathbb{R}^{K \times N}, A_2 \in \mathbb{R}^{I \times J \times K}, B_2 \in \mathbb{R}^{J \times K \times N}, C_2 \in \mathbb{R}^{K \times N}$ are the factor matrices of the CPD. Thus, the SRI admits a CPD such that $Y = [A_2, B_2, C_1]$. We wish to estimate the factor matrices of the CPD of $Y_1$ and $Y_2$. Thus, we define the model parameters

$$\psi_1 = \text{vec}\{C_1\} \in \mathbb{R}^{KN}, \phi_1 = \text{vec}\{A_1\} \in \mathbb{R}^{(I_H + J_H)N},$$n
$$\psi_2 = \text{vec}\{C_2\} \in \mathbb{R}^{KN}, \phi_2 = \text{vec}\{A_2\} \in \mathbb{R}^{(I \times J)N},$$

(18)

corresponding to the vectorization of the factor matrices for each tensor. The choice to group $A_1$ and $B_1$ factors, separate from the $C_1$ ($i = 1, 2$) factor matrices is motivated by the fact that spatial and spectral degradations never occur in the same tensor, according

1 We suppose that the spatial degradation for the HSI is separable.
to model \([16]\). Thus, \(\psi_2\) and \(\phi_2\) can be seen as degraded versions of \(\psi_1\) and \(\phi_1\) by the spectral and spatial degradation matrices, respectively. In the problem at hand, \(\Theta^f = (\psi_1^f, \phi_1^f, \psi_2^f, \phi_2^f)\) is the vector of unknown deterministic parameters to be estimated and \(\alpha\) is a random parameter. Since \(\mathcal{E}_1\) and \(\mathcal{E}_2\) are i.i.d., the HSI and MSI are distributed according to

\[
\begin{align*}
    f_{\psi_1, \phi_1, \psi_2, \phi_2} &= (2\pi\sigma_2^2)^{-\frac{1}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{\|Y_1 - [A_1, B_1, C_1]\|_F^2}{2\sigma_1^2}} \\
    f_{\psi_2, \phi_2} &= (2\pi\sigma_2^2)^{-\frac{1}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{\|Y_2 - [A_2, B_2, C_2]\|_F^2}{2\sigma_2^2}}.
\end{align*}
\]

\[\tag{18}\]

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 \([40]\), we need to fix the first rows of \(A_i, B_i\) \((i = 1, 2)\) to known values; here, we normalize the factors by setting the first rows to ones. As a result, we define the parameters \(\phi_1\) and \(\phi_2\), only composed of the unknown entries of \(\phi_1\) \((\text{resp.} \phi_2)\). The matrices \(M_1 \in \mathbb{R}^{[(1+J-H)/2]N \times (1+J+H)N}\) and \(M_2 \in \mathbb{R}^{[(1+J-I)/2]N \times (1+J+I)N}\) are masks obtained by removing the \(2N\) entries corresponding to known entries of \(\phi_1\) \((\text{resp.} \phi_2)\). The CP model consists of the unknown entries of \(\phi_1\) \(\phi_2\), such that \(\phi_1 = M_1 \hat{\phi}_1\) and \(\phi_2 = M_2 \hat{\phi}_2\). Thus, we can define the deterministic parameter \(\bar{\theta} = (\psi_1^0, \phi_1, \psi_2, \phi_2)\). We consider the uncoupled CP model

\[\mathcal{X} = \{\text{vec}(Y_1), \text{vec}(Y_2)\} \sim \mathcal{N}(\mu(\bar{\theta}), \Sigma), \mu(\bar{\theta}) = \text{vec}([A_1, B_1, C_1]), \text{vec}([A_2, B_2, C_2])\]

and \(\Sigma = \text{Diag}(\sigma_1^2 I, \sigma_2^2 I)\). The FIM on \(\bar{\theta}\) in the uncoupled case is given by the Slepian-Bangs formula \([39]\),

\[F(\bar{\theta}) = \begin{bmatrix} \frac{\partial \mu(\bar{\theta})}{\partial \bar{\theta}} \end{bmatrix} \Sigma^{-1} \begin{bmatrix} \frac{\partial \mu(\bar{\theta})}{\partial \bar{\theta}} \end{bmatrix}^T.\]

For a full derivation of the FIM, see \([32]\) Section IV. Following model \([17]\), we can define, conditionally to the value of \(\alpha\), the equality constraints on \(\alpha\) as

\[f_\alpha(\bar{\theta}) = \begin{cases} |\alpha| (I \otimes R) \psi_1 - \psi_2, \\
\quad \phi_1 - M_1 \begin{bmatrix} I \otimes P & 0 \\ 0 & I \otimes Q \end{bmatrix} M_2^T \phi_2. \end{cases}\]

\[\tag{19}\]

Taking the derivative of \(f_\alpha(\bar{\theta})\) and setting it equal to zero yields

\[\frac{\partial f_\alpha(\bar{\theta})}{\partial \bar{\theta}} = \begin{bmatrix} |\alpha| (I \otimes R) & 0 & -I & 0 \\ 0 & I & 0 & -P \end{bmatrix}.\]

The expression for RCCR\(B\) is obtained from \(F(\bar{\theta})\) and any basis \(U_\alpha\) of ker \(\left(\frac{\partial f_\alpha(\bar{\theta})}{\partial \bar{\theta}}\right)^T\) by plugging them into \([15]\). In this case, the standard MLE \(\bar{\theta} = \arg \max_{\bar{\theta}} \{p(X; \bar{\theta})\}\) is asymptotically locally strict-sense unbiased and verifies \([15]\).

III-C. Simulations

In this section, we simulate the performance of the estimation of \(\bar{\theta}\) under additive Gaussian noise in the uncoupled case and in the constrained case with random equality constraints. We compare the MSE of the estimators to the bounds mentioned above. For basic tensor operations, we use TensorLab 3.0 \([41]\). The model parameters are retrieved using MLE. In the uncoupled case, an ALS algorithm \([42]\) 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 \([30]\), with a regularization parameter \(\lambda = \sigma_1^2 / \sigma_2^2\). The factor matrices estimated by the uncoupled ALS algorithm are used as initialization for the randomly constrained case to speed up the convergence of the algorithm. For the CP model to be identifiable, the scaling and permutation ambiguities are corrected by setting the first rows of the factors \(A_1, B_1, A_2, B_2\) to ones and searching for the best permutation of \(C_2\) with fixed \(C_1\). 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 Gaussian variables, and first rows of \(A_1, A_2, B_1, B_2\) are set to 1. The tensor rank for both HSI and MSI is \(N = 3\). The degradation matrices \(P, Q, R\) 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 \(\sigma_1^2\) such that it varies from 2 to \(2 \cdot 10^{-3}\). The SNR of \(Y_1\) varies from 5 to 60dB, while the SNR of \(Y_2\) is fixed to 20dB. For each value of \(\sigma_1^2\), we first compute \(\text{CRB}_\alpha(\bar{\theta})\) and \(\text{CRRB}_\alpha(\bar{\theta})\) by averaging over 200 realizations of \(\alpha\). We then compute \(\text{RCCR}\_\alpha(\bar{\theta})\) by averaging the conditional CCRB for all values of \(\sigma_1^2\). We also evaluate the total MSE on the parameters for each value of \(\sigma_1^2\) by averaging the squared errors over 500 noise realizations.

In Fig. 1.a, we show the limitations of the conditional CCRB. We plot the MSE given by uncoupled ALS and STEREO, as well as the uncoupled CRB. We also display \(\text{CRRB}_\alpha(\bar{\theta})\) for two values of \(\sigma_1^2 = 2\) and \(2 \cdot 10^{-3}\). Since the illumination coefficient \(\alpha\) mostly impacts the estimation of \(\psi_2\), we choose to display its performance bounds separately and group \(\psi_1\), \(\phi_1\), and \(\phi_2\) together. We can see that the MSE given by the uncoupled ALS algorithm reaches the uncoupled CRB for all parameters. For \(\psi_2\), the MSE given by STEREO does not fit any of the two conditional CCRB depicted. In particular, it is slightly below \(\text{CRRB}_\alpha(\psi_2)\) for \(\sigma_1^2 = 2\) and above it for \(\sigma_1^2 = 2 \cdot 10^{-3}\). For the other parameters, while the MSE of STEREO fits the conditional CCRB for \(\sigma_1^2 = 2\), it is slightly below the second curve for a low SNR. In Fig. 1.b, we plot the MSE, the uncoupled CRB and the RCCRB. Here, we can see that the MSE on \(\psi_2\) given by STEREO follows the RCCRB with a small gap for a SNR on \(Y_1\) 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


