

On lower bounds for non-standard deterministic estimation

Nabil Kbayer, Jérôme Galy, Eric Chaumette, François Vincent, Alexandre Renaux, Pascal Larzabal

▶ **To cite this version:**

Nabil Kbayer, Jérôme Galy, Eric Chaumette, François Vincent, Alexandre Renaux, et al.. On lower bounds for non-standard deterministic estimation. IEEE Transactions on Signal Processing, 2017, 65 (6), pp.1538-1553. 10.1109/TSP.2016.2645538 . hal-01525498

HAL Id: hal-01525498

https://centralesupelec.hal.science/hal-01525498

Submitted on 21 May 2017

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.

On Lower Bounds for Non-Standard Deterministic Estimation

Nabil Kbayer, Jérôme Galy, Eric Chaumette, François Vincent, Alexandre Renaux and Pascal Larzabal

Abstract

We consider deterministic parameter estimation and the situation 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. Unfortunately, in the general case, this marginalization is mathematically intractable, which prevents from using the known standard deterministic lower bounds (LBs) on the mean-squared-error (MSE). Actually the general case can be tackled by embedding the initial observation space in an hybrid one where any standard LB can be transformed into a modified one fitted to non-standard deterministic estimation, at the expense of tightness however. Furthermore, these modified LBs (MLBs) appears to include the sub-matrix of hybrid lower bounds which is a LB for the deterministic parameters. Moreover, since in non-standard estimation, maximum likelihood estimators (MLEs) can be no longer derived, suboptimal non-standard MLEs (NSMLEs) are proposed as being a substitute. We show that any standard LB on the MSE of MLEs has a non-standard version lower bounding the MSE of NSMLEs. We provide an analysis of the relative performance of the NSMLEs, as well as a comparison with the modified LBs for a large class of estimation problems. Last, the general approach introduced is exemplified, among other things, with a new look at the well known Gaussian complex observation models.

Index Terms

Deterministic parameter estimation, estimation error lower bound, maximum likelihood estimation.

I. Introduction

As introduced in [1, p53], a model of the general deterministic estimation problem has the following four components: 1) a parameter space Θ_d, 2) an observation space $\mathcal{X}$, 3) a probabilistic mapping from parameter vector space Θ_d to observation space $\mathcal{X}$, that is the probability law that governs the effect of a parameter vector value $\boldsymbol{\theta}$[1] on the observation $\mathbf{x}$ and, 4) an estimation rule, that is the mapping of the observation space $\mathcal{X}$ into vector parameter estimates $\widehat{\boldsymbol{\theta}}(\mathbf{x})$. Actually, in many estimation problems [1][2][3][4], the probabilistic mapping 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) $\boldsymbol{\theta} \rightarrow \boldsymbol{\theta}_r \sim p(\boldsymbol{\theta}_r; \boldsymbol{\theta})$, ii) $(\boldsymbol{\theta}, \boldsymbol{\theta}_r) \rightarrow \mathbf{x} \sim p(\mathbf{x}|\boldsymbol{\theta}_r; \boldsymbol{\theta})$, and leading to a compound probability distribution:

$$p(\mathbf{x}, \boldsymbol{\theta}_r; \boldsymbol{\theta}) = p(\mathbf{x}|\boldsymbol{\theta}_r; \boldsymbol{\theta})\, p(\boldsymbol{\theta}_r; \boldsymbol{\theta}), \tag{1a}$$

$$p(\mathbf{x}; \boldsymbol{\theta}) = \int_{\Theta_r} p(\mathbf{x}, \boldsymbol{\theta}_r; \boldsymbol{\theta})\, d\boldsymbol{\theta}_r, \tag{1b}$$

Copyright (c) 2017 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org. DOI 10.1109/TSP.2016.2645538

This work appeared in part in the proceedings of IEEE workshop on Sensor Array and Multichannel Signal Processing, Rio de Janeiro, Brazil, 2016

Nabil Kbayer is with University of Toulouse, Tésa/Isae-Supaéro, 7 Boulevard de la Gare, 31500 Toulouse, France. Email: nabil.kbayer@isae.fr

Jérome Gally is with Université de Montpellier 2/LIRMM, 161 rue Ada 34392 Montpellier Cedex 5, France. Email: galy@lirmm.fr

Eric Chaumette and François Vincent are with University of Toulouse-ISAE, DEOS, 10 Avenue Edouard Belin, 31055 Toulouse, France. Email: eric.chaumette@isae.fr, francois.vincent@isae.fr.

Alexandre Renaux is with Universite Paris-Sud/LSS 3, Rue Joliot-Curie, 91192 Gif-sur-Yvette, France. Email: renaux@lss.supelec.fr

Pascal Larzabal is with Université Paris-Sud/SATIE, 61 av. du President Wilson, 94235 Cachan cedex, France. Email: pascal.larzabal@satie.ens-cachan.fr

This work has been partially supported by the iCODE institute, research project of the IDEX Paris-Saclay, by the DGA/DGCIS, by the DGA/MRIS (2015.60.0090.00.470.75.01), by the CNES, and by the ANR MAGELLAN.

[1]Throughout the present paper, scalars, vectors and matrices are represented, respectively, by italic (as in a or A), bold lowercase (as in $\mathbf{a}$) and bold uppercase (as in $\mathbf{A}$) characters.

where $p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})$ is the conditional probability density function (p.d.f.) of $\mathbf{x}$ given $\boldsymbol{\theta}_r$, and $p(\boldsymbol{\theta}_r;\boldsymbol{\theta})$ is the prior p.d.f., parameterized by $\boldsymbol{\theta}$. Therefore, deterministic estimation problems can be divided into two subsets: the subset of "standard" deterministic estimation problems for which a closed-form expression of $p(\mathbf{x};\boldsymbol{\theta})$ is available, and the subset of "non-standard" deterministic estimation problems for which only an integral form of $p(\mathbf{x};\boldsymbol{\theta})$ (1b) is available.

In any estimation problem, having explicit expressions for lower bounds on estimators performance is desired, since it can give insight on the inherent limitations of the problem, as well as being a reference to the optimality of the various estimators. For a long time, the open literature on deterministic lower bounds (LBs) on the mean squared error (MSE) has remained focused on standard estimation [5]-[21]. One noteworthy point is that all the LBs introduced in standard estimation derive, directly [22][23][24] or indirectly [18][25], from the seminal work of Barankin [11], who established the general form of the greatest lower bound on any sth absolute central moment ($s > 1$) of a uniformly unbiased estimator with respect to $p(\mathbf{x};\boldsymbol{\theta})$, generalizing the earlier works of Fisher [5], Frechet [6], Darmois [7], Cramér [8], Rao [9] and Battacharayya [10]. It is likely that the first LBs derived in the non-standard case appeared in [2], [3] and [4], to characterize three estimation problems which differ from one another mainly in the status of the additional random vector $\boldsymbol{\theta}_r$. In [2], authors derived a LB for estimators of $\boldsymbol{\theta}$ uniformly unbiased for all values of $\boldsymbol{\theta}_r$, that is uniformly unbiased with respect to $p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})$. In such a setting, $\boldsymbol{\theta}_r$ is treated as a nuisance parameter vector, i.e., it is not a parameter vector of interest, but the estimate of the interesting parameter vector $\boldsymbol{\theta}$ is nevertheless dependent on it. The LB obtained, aka the Miller and Chang bound (MCB), can be regarded as the first LB on the MLE of $\boldsymbol{\theta}$ deriving from $p(\mathbf{x}|\boldsymbol{\theta}_r;\boldsymbol{\theta})$ in place of $p(\mathbf{x};\boldsymbol{\theta})$, hereinafter referred to as a non-standard MLE (NSMLE) of $\boldsymbol{\theta}$. In [3], both $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_r$ are parameter vectors of interest and the LB derived is the first hybrid LB (HLB), that is a LB for a hybrid parameter vector consisting of a mixture of deterministic and random parameters. In [4], $\boldsymbol{\theta}_r$ is an unwanted parameter vector introduced by a two steps probabilistic mechanism (1a-1b) which estimation is not of interest, and the LB derived is the first LB in non-standard estimation for unbiased estimates in the Barankin sense, hereinafter referred to as a modified LB (MLB). Specifically, [2], [3] and [4] introduced three variants of the standard deterministic Cramér-Rao bound (CRB) [5][6][7][8][9] suitable for non-standard estimation, namely the MCB, the hybrid CRB (HCRB) and the modified CRB (MCRB)[2]. Later on, Reuven and Messer [29] put forward the so-called hybrid Barankin bound (HBB), that is a Barankin-type LB on the MSE of a hybrid parameter vector. It is noteworthy that the deterministic part[3] of the HBB provides a LB on the deterministic parameters able to handle the MSE threshold phenomena [17][19][20][22][23][24], and that one limiting form yields the HCRB [3]. This seminal work fitted to the hybrid estimation framework has naturally contributed to disseminate the concept of a framework where the random variables $(\boldsymbol{\theta}_r)$ marginalized out (1b) are regarded as unknown nuisance random parameters to be estimated. This is unambiguously recalled in abstracts and introductions of subsequent papers dealing with non-standard estimation such as [30]-[42].

As a first contribution sketchily disclosed in [43], a unified framework for the derivation of LBs on the MSE for unbiased estimates in non-standard estimation is more fully developed. The proposed framework exploits the fact that the MSE is the square of a norm defined on the vector space of square integrable functions $\widehat{\boldsymbol{\theta}}(\mathbf{x})$ and, therefore, all known standard LBs on the MSE can be formulated as the solution of a norm minimization problem under linear constraints [20][23][25][44]. This formulation of LBs not only provides a straightforward understanding of the hypotheses associated with the different LBs [20][23][44] (briefly overviewed in Section II), but also allows to obtain a unique formulation of each LB in terms of a unique set of linear constraints. Interestingly enough, the "norm minimization" approach simply exploits the general form of the marginalization formula (1b), that is $p(\mathbf{x};\boldsymbol{\theta})$ is simply the result of marginalizing over the intermediate random variables $\boldsymbol{\theta}_r$, without any reference to extraneous or nuisance random parameters. As an immediate consequence, which goes beyond mere semantic features, $\boldsymbol{\theta}_r$ is neither required nor expected to be estimated when searching for LBs for unbiased estimates of $\boldsymbol{\theta}$ in non-standard estimation. Indeed, the proposed rationale brings to light that the lack of a closed-form for the marginal p.d.f. $p(\mathbf{x};\boldsymbol{\theta})$ simply compels to embed the initial norm minimization problem on observation space $\mathcal{X}$ into the same norm minimization problem but on $\mathcal{X} \times \Theta_r$, leading to the derivation of LBs for unbiased estimator $\widehat{\boldsymbol{\theta}}(\mathbf{x},\boldsymbol{\theta}_r)$ of $\boldsymbol{\theta}$ in place of unbiased estimators $\widehat{\boldsymbol{\theta}}(\mathbf{x})$ of $\boldsymbol{\theta}$. This embedding mechanism allows to show that any

[2]For the sake of completeness: i) CRBs for synchronization parameter estimation derived earlier in [26][27] are in fact MCRBs, ii) [28] introduced earlier a tighter version of the MCRB [28, (20)] but with sufficient conditions [28, (21)] unnecessary restrictive.

[3]In the following, the deterministic part of a HLB denotes the HLB's sub-matrix which is a LB on the MSE of the determinisitic parameter vector.

LB derived in standard estimation has an analog form fitted to non-standard estimation obtained by substituting $p(\mathbf{x}, \boldsymbol{\theta}_r; \boldsymbol{\theta})$ for $p(\mathbf{x}; \boldsymbol{\theta})$ in the LB definition (see Section III), encompassing the historical LBs such as the MCRB [4][30] or the modified Bhattacharyya bound (MBaB) [34]. Therefore, it seems appropriate to keep on referring to these LBs for non-standard estimation as modified LBs (MLBs) as proposed initially in [4].
Further more, the proposed framework naturally allows to obtain tighter MLBs (see Subsections III-B1 and III-B2) by adding constraints compatible with unbiasedeness, in order to restrict the class of viable estimators $\widehat{\boldsymbol{\theta}}(\mathbf{x}, \boldsymbol{\theta}_r)$ and therefore to increase the minimum norm obtained. It is exemplified with two general subsets of additional constraints. The first subset yields the deterministic part of the class of HLBs based on linear transformation on the centered likelihood-ratio (CLR) function [45], which includes the HCRB [3][30][35], the HBB [29] and their combination via a compression matrix (the CCLRB) [45]. The second subset, combined with the first one, yields the deterministic part of an even more general class of HLBs obtained from linear transformation of a mixture of the McAulay-Seidman bound (MSB) and Bayesian lower bounds of the Weiss-Weinstein family, including the so-called hybrid McAulay-Seidman-Weiss-Weinstein bound (HMSSWB) [46]. It is noteworthy that the proposed unified framework allows to reformulate all known LBs on the MSE for unbiased estimates in non-standard estimation as a MLBs without any regularity condition on the (nuisance) random vector estimates, since $\boldsymbol{\theta}_r$ is neither required nor expected to be estimated. Furthermore, the proposed framework naturally incorporates possible tightness comparison between two MLBs. Indeed, if the subset of linear constraints associated with a MLB is included into the subset of linear constraints associated with another MLB, then the latter one is tighter. Thus, the proposed unified framework is a useful tool to look for the best possible trade-off between tightness, regularity conditions and computational cost, in the choice of a MLB for a given non-standard estimation problem. Last, the proposed framework not only proves straightforwardly the looseness of any MLB (including therefore the deterministic part of any HLB) in comparison with the standard LB, but also provides a very general "closeness condition" on $p(\mathbf{x}, \boldsymbol{\theta}_r; \boldsymbol{\theta})$ in order to obtain a MLB equal to the standard LB (see Subsection III-C1).

Since in non-standard estimation, maximum likelihood estimators (MLEs) can be no longer derived, "non-standard" MLEs (NSMLEs), i.e. joint MLEs of $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_r$ w.r.t. $p(\mathbf{x}|\boldsymbol{\theta}_r; \boldsymbol{\theta})$, are proposed as being a substitute. The idea underlying this proposal is that the closed-form of $p(\mathbf{x}|\boldsymbol{\theta}_r; \boldsymbol{\theta})$ is known in many estimation problems [47] and therefore the NSMLEs take advantage not only of asymptotic optimality and unbiasedeness w.r.t. $p(\mathbf{x}|\boldsymbol{\theta}_r; \boldsymbol{\theta})$ of MLEs, but also of the extensive open literature on MLE closed-form expressions or approximations [47]. These key features clearly make the "non-standard" maximum likelihood estimation more attractive than joint maximum a posteriori-maximum likelihood estimation (JMAPMLE), known to be biased and inconsistent in general [48], or than the Expectation-Maximization algorithm [49], impractical to implement in many non-standard problems of interest. Furthermore, as exemplified in [2], they are applications where the property of an estimator unbiased for all values of the random parameter $\boldsymbol{\theta}_r$ is desirable, even if a closed-form expression of $p(\mathbf{x}; \boldsymbol{\theta})$ is available. Last, as mentioned in [50, p. 6, 12], it appears that the NSMLEs are also the asymptotic forms of JMAPMLEs for a class of hybrid estimation problems when the number of independent observations tends to infinity [51]. To the best of our knowledge, NSMLEs have not received full attention so far, whereas they may be the only sensible estimators available in many estimation problems Therefore, as a second contribution we introduce a detailed study of the asymptotic properties of NSMLEs (see Section IV). Firstly, we show that NSMLEs are asymptotically (when the number of independent observations tends to infinity) suboptimal and that any standard LB on the MSE of MLEs has a non-standard version lower bounding the MSE of NSMLEs. Secondly, we exhibit a large class of estimation problems for which a comparison between non-standard LBs and modified LBs is possible.

Last, the general approach introduced is exemplified with existing application examples and a new look at the well known Gaussian complex observation models (see Section V-B).
The results introduced in the following are also of interest if a closed-form of $p(\mathbf{x}; \boldsymbol{\theta})$ does exist but the resulting expression is intractable to derive LBs and/or MLEs.
For the sake of simplicity, unless otherwise stated, we will focus on the estimation of a single unknown real deterministic parameter θ, although the results are easily extended to the estimation of multiple functions of multiple parameters [23][24] (see Sections III-A1 and V-B for examples).

A. Notations and Assumptions

In the following, the n-th row and m-th column element of the matrix $\mathbf{A}$ will be denoted by $A_{n,m}$ or $(\mathbf{A})_{n,m}$. The n-th coordinate of the column vector $\mathbf{a}$ will be denoted by a_n or $(\mathbf{a})_n$. The scalar/matrix/vector transpose,

conjugate and transpose conjugate are respectively indicated by the superscripts ^{T}, * and ^{H}. $[\mathbf{A},\mathbf{B}]$ denotes the matrix resulting from the horizontal concatenation of matrices $\mathbf{A}$ and $\mathbf{B}$. $\left(\mathbf{a}^T,\mathbf{b}^T\right)$ denotes the row vector resulting from the horizontal concatenation of row vectors $\mathbf{a}^T$ and $\mathbf{b}^T$. $\mathbf{I}_M$ is the identity matrix of order M. $\mathbf{1}_M$ is a M-dimensional vector with all components equal to one. If $\mathbf{A}$ is a square matrix, $|\mathbf{A}|$ denotes its determinant. For two matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A}\succeq\mathbf{B}$ means that $\mathbf{A}-\mathbf{B}$ is positive semi-definite (Löwner ordering). Furthermore, unless otherwise stated:

• $\mathbf{x}$ denotes a M-dimensional complex random observation vector belonging to the observation space $\mathcal{X}\subset\mathbb{C}^M$, θ denotes a real deterministic parameter belonging to the parameter space $\Theta_d\subset\mathbb{R}$, and $\boldsymbol{\theta}_r$ denotes a P_r-dimensional real random vector belonging to a subset Θ_r of $\mathbb{R}^{P_r}$.

• $p(\mathbf{x};\theta)$, $p(\mathbf{x},\boldsymbol{\theta}_r;\theta)$, $p(\mathbf{x}|\boldsymbol{\theta}_r;\theta)$, $p(\boldsymbol{\theta}_r;\theta)$ and $p(\boldsymbol{\theta}_r|\mathbf{x};\theta)$ denote, respectively, the marginal p.d.f. of $\mathbf{x}$, the joint p.d.f. of $\mathbf{x}$ and $\boldsymbol{\theta}_r$, the likelihood function of $\boldsymbol{\theta}_r$, the prior p.d.f. of $\boldsymbol{\theta}_r$ and the posterior p.d.f. of $\boldsymbol{\theta}_r$, parameterized by θ.

• $\mathcal{S}_{\mathcal{X}}$, $\mathcal{S}_{\mathcal{X},\Theta_r}$, $\mathcal{S}_{\mathcal{X}|\boldsymbol{\theta}_r}$, $\mathcal{S}_{\Theta_r}$ and $\mathcal{S}_{\Theta_r|\mathbf{x}}$ denote, respectively, the support of $p(\mathbf{x};\theta)$, $p(\mathbf{x},\boldsymbol{\theta}_r;\theta)$, $p(\mathbf{x}|\boldsymbol{\theta}_r;\theta)$, $p(\boldsymbol{\theta}_r;\theta)$ and $p(\boldsymbol{\theta}_r|\mathbf{x};\theta)$, i.e., $\mathcal{S}_{\mathcal{X}}=\left\{\mathbf{x}\in\mathbb{C}^M \mid p(\mathbf{x};\theta)>0\right\}$, $\mathcal{S}_{\mathcal{X},\Theta_r}=\left\{(\mathbf{x},\boldsymbol{\theta}_r)\in\mathbb{C}^M\times\mathbb{R}^{P_r} \mid p(\mathbf{x},\boldsymbol{\theta}_r;\theta)>0\right\}$, $\mathcal{S}_{\mathcal{X}|\boldsymbol{\theta}_r}=\left\{\mathbf{x}\in\mathbb{C}^M \mid p(\mathbf{x}|\boldsymbol{\theta}_r;\theta)>0\right\}$, $\mathcal{S}_{\Theta_r}=\left\{\boldsymbol{\theta}_r\in\mathbb{R}^{P_r} \mid p(\boldsymbol{\theta}_r;\theta)>0\right\}$,
and $\mathcal{S}_{\Theta_r|\mathbf{x}}=\left\{\boldsymbol{\theta}_r\in\mathbb{R}^{P_r} \mid p(\boldsymbol{\theta}_r|\mathbf{x};\theta)>0\right\}$[4].

• $E_{\mathbf{x};\theta}[\mathbf{g}(\mathbf{x})]$, $E_{\boldsymbol{\theta}_r;\theta}[\mathbf{g}(\boldsymbol{\theta}_r)]$ and $E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}[\mathbf{g}(\mathbf{x},\boldsymbol{\theta}_r)]$ denote, respectively, the statistical expectation of the vector of functions $\mathbf{g}(\ )$ with respect to $\mathbf{x}$, to $\boldsymbol{\theta}_r$, to $\mathbf{x}$ and $\boldsymbol{\theta}_r$, parameterized by θ, and satisfy:

$$E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}[\mathbf{g}(\mathbf{x},\boldsymbol{\theta}_r)]=E_{\mathbf{x};\theta}\left[E_{\boldsymbol{\theta}_r|\mathbf{x};\theta}[\mathbf{g}(\mathbf{x},\boldsymbol{\theta}_r)]\right]=E_{\boldsymbol{\theta}_r;\theta}\left[E_{\mathbf{x}|\boldsymbol{\theta}_r;\theta}[\mathbf{g}(\mathbf{x},\boldsymbol{\theta}_r)]\right]. \tag{2}$$

• $\mathcal{L}_2(\mathcal{S}_{\mathcal{X}})$, $\mathcal{L}_2(\mathcal{S}_{\mathcal{X},\Theta_r})$ and $\mathcal{L}_2(\mathcal{S}_{\mathcal{X}|\boldsymbol{\theta}_r})$ denote, respectively, the real Euclidean space of square integrable real-valued functions w.r.t. $p(\mathbf{x};\theta)$, $p(\mathbf{x},\boldsymbol{\theta}_r;\theta)$ and $p(\mathbf{x}|\boldsymbol{\theta}_r;\theta)$,

• $1_A(\boldsymbol{\theta}_r)$ denote the indicator function of subset A of $\mathbb{R}^{P_r}$.

II. An overview of lower bounds for standard estimation

A. On Lower Bounds and Norm Minimization

In the search for a LB on the MSE of unbiased estimators, two fundamental properties of the problem at hand, introduced by Barankin [11], must be noticed. The first property is that the MSE of a particular estimator $\widehat{\theta^0}\in\mathcal{L}_2(\mathcal{S}_{\mathcal{X}})$ of θ^0, i.e., $\widehat{\theta^0}\triangleq\widehat{\theta^0}(\mathbf{x})$, where θ^0 is a selected value of the parameter θ, is the square of a norm $\|\ \|_\theta^2$ associated with a particular scalar product $\langle\ |\ \rangle_\theta$:

$$MSE_{\theta^0}\left[\widehat{\theta^0}\right]=\left\|\widehat{\theta^0}(\mathbf{x})-\theta^0\right\|_{\theta^0}^2, \tag{3a}$$

$$\langle g(\mathbf{x})\mid h(\mathbf{x})\rangle_\theta=E_{\mathbf{x};\theta}[g(\mathbf{x})\,h(\mathbf{x})]. \tag{3b}$$

This property allows the use of two equivalent fundamental results: the generalization of the Cauchy-Schwartz inequality to Gram matrices (generally referred to as the "covariance inequality" [21][22]) and the minimization of a norm under linear constraints [20][23][25][44]. Nevertheless, we shall prefer the "norm minimization" form as its use:

• provides a straightforward understanding of the hypotheses associated with the different LBs on the MSE expressed as a set of linear constraints,

• allows to resort to the same rationale for the derivation of LBs whatever the observation space considered,

• allows to easily reveal LBs inequalities and tightness conditions without the complex derivations (based on the use of the covariance inequality) introduced by previous works [30][29][52].

The second property is that an unbiased estimator $\widehat{\theta^0}\in\mathcal{L}_2(\mathcal{S}_{\mathcal{X}})$ of θ^0 should be uniformly unbiased:

$$\forall\theta\in\Theta_d: E_{\mathbf{x};\theta}\left[\widehat{\theta^0}(\mathbf{x})\right]=\int_{\mathcal{S}_{\mathcal{X}}}\widehat{\theta^0}(\mathbf{x})\,p(\mathbf{x};\theta)\,d\mathbf{x}=\theta. \tag{4a}$$

[4]The Barankin approach [11] (see Section II) and its extensions to non-standard estimation (see Sections III and IV) require that all the supports be independent of θ.

If $\mathcal{S}_{\mathcal{X}}$, i.e., the support of $p(\mathbf{x};\theta)$, does not depend on θ, then (4a) can be recasted as:

$$\forall\theta\in\Theta_d : E_{\mathbf{x};\theta^0}\left[\left(\widehat{\theta^0}(\mathbf{x})-\theta^0\right)\upsilon_{\theta^0}(\mathbf{x};\theta)\right]=\theta-\theta^0, \tag{4b}$$

where $\upsilon_{\theta^0}(\mathbf{x};\theta)=p(\mathbf{x};\theta)/p\left(\mathbf{x};\theta^0\right)$ denotes the Likelihood Ratio (LR). As a consequence, the locally-best (at θ^0) unbiased estimator in $\mathcal{L}_2(\mathcal{S}_{\mathcal{X}})$ is the solution of a norm minimization under linear constraints:

$$\min_{\widehat{\theta^0}\in\mathcal{L}_2(\mathcal{S}_{\mathcal{X}})}\left\{\left\|\widehat{\theta^0}(\mathbf{x})-\theta^0\right\|_{\theta^0}^2\right\}\ \text{under}\ \left\langle\widehat{\theta^0}(\mathbf{x})-\theta^0\mid\upsilon_{\theta^0}(\mathbf{x};\theta)\right\rangle_{\theta^0}=\theta-\theta^0,\forall\theta\in\Theta_d. \tag{5}$$

Unfortunately, as recalled hereinafter, if Θ_d contains a non empty interval of $\mathbb{R}$, then the norm minimization problem (5) leads to an integral equation (9a) with no analytical solution in general. Therefore, since the seminal work of Barankin [11], many studies quoted in [22][23][24][53] have been dedicated to the derivation of "computable" LBs approximating the MSE of the locally-best unbiased estimator, which defines the Barankin bound (BB). All these approximations derive from sets of discrete or integral linear transform of the "Barankin" constraint (4b) and can be easily obtained (see next Section) using the following well known norm minimization lemma [54]. Let $\mathbb{U}$ be an Euclidean vector space on the field of real numbers $\mathbb{R}$ which has a scalar product $\langle\mid\rangle$. Let $(\mathbf{c}_1,\ldots,\mathbf{c}_K)$ be a family of K linearly independent vectors of $\mathbb{U}$ and $\mathbf{v}\in\mathbb{R}^K$. The problem of the minimization of $\|\mathbf{u}\|^2$ under the K linear constraints $\langle\mathbf{u}\mid\mathbf{c}_k\rangle=v_k$, $k\in[1,K]$ then has the solution:

$$\min\left\{\|\mathbf{u}\|^2\right\}=\|\mathbf{u}_{opt}\|^2=\mathbf{v}^T\mathbf{G}^{-1}\mathbf{v}, \tag{6}$$

$$\mathbf{u}_{opt}=\sum_{k=1}^{K}\alpha_k\mathbf{c}_k,\ \boldsymbol{\alpha}=\mathbf{G}^{-1}\mathbf{v},\ \mathbf{G}_{k',k}=\langle\mathbf{c}_k\mid\mathbf{c}_{k'}\rangle.$$

B. Lower Bounds via linear transformations of the McAulay-Seidman bound

The McAulay-Seidman bound (MSB) is the BB approximation obtained from a discretization of the Barankin unbiasedness constraint (4b). Let $\boldsymbol{\theta}^N=\left(\theta^1,\ldots,\theta^N\right)^T\in\Theta_d^N$ be a vector of N selected values of the parameter θ (aka test points), $\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x};\boldsymbol{\theta}^N\right)=\left(\upsilon_{\theta^0}\left(\mathbf{x};\theta^1\right),\ldots,\upsilon_{\theta^0}\left(\mathbf{x};\theta^N\right)\right)^T$ be the vector of LRs associated to $\boldsymbol{\theta}^N$, $\xi(\theta)=\theta-\theta^0$ and $\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)=\left(\xi\left(\theta^1\right),\ldots,\xi\left(\theta^N\right)\right)^T$. Then, any unbiased estimator $\widehat{\theta^0}\in\mathcal{L}_2(\mathcal{S}_{\mathcal{X}})$ verifying (4b) must comply with the following subset of N linear constraints:

$$E_{\mathbf{x};\theta^0}\left[\left(\widehat{\theta^0}(\mathbf{x})-\theta^0\right)\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x};\boldsymbol{\theta}^N\right)\right]=\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right), \tag{7a}$$

yielding, via the norm minimization lemma (6), the MSB [17]:

$$MSE_{\theta^0}\left[\widehat{\theta^0}\right]\geq\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T\mathbf{R}_{\boldsymbol{\upsilon}_{\theta^0}}^{-1}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right),\quad\left(\mathbf{R}_{\boldsymbol{\upsilon}_{\theta^0}}\right)_{n,m}=E_{\mathbf{x};\theta^0}\left[\upsilon_{\theta^0}(\mathbf{x};\theta^m)\,\upsilon_{\theta^0}(\mathbf{x};\theta^n)\right], \tag{7b}$$

which is a generalization of the Hammersley-Chapman-Robbins bound (HaCRB) previously introduced in [12] and [13] for 2 test points ($N=2$). Obviously, any given set of K ($K\leq N$) linear transformations of (7a):

$$E_{\mathbf{x};\theta^0}\left[\left(\widehat{\theta^0}(\mathbf{x})-\theta^0\right)\mathbf{H}_K^T\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x};\boldsymbol{\theta}^N\right)\right]=\mathbf{H}_K^T\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right),\quad\mathbf{H}_K=[\mathbf{h}_1,\ldots,\mathbf{h}_K],\ \mathbf{h}_k\in\mathbb{R}^N,\ 1\leq k\leq K, \tag{8a}$$

where $\mathbf{H}_K$ has a full rank, provides, via the norm minimization lemma (6), another LB on the MSE:

$$MSE_{\theta^0}\left[\widehat{\theta^0}\right]\geq\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T\mathbf{R}_{\mathbf{H}_K}^{\dagger}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right),\quad\mathbf{R}_{\mathbf{H}_K}^{\dagger}=\mathbf{H}_K\left(\mathbf{H}_K^T\mathbf{R}_{\boldsymbol{\upsilon}_{\theta^0}}\mathbf{H}_K\right)^{-1}\mathbf{H}_K^T. \tag{8b}$$

It is worth noting that, for a given vector of test points $\boldsymbol{\theta}^N$, the LB (8b) reaches its maximum if, and only if, the matrix $\mathbf{H}_K$ is invertible ($K=N$) [17][55, Lemma 3], which represents a bijective transformation of the set of constraints associated with the MSB (7a). Thus:

$$MSE_{\theta^0}\left[\widehat{\theta^0}\right]\geq\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T\mathbf{R}_{\boldsymbol{\upsilon}_{\theta^0}}^{-1}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)\geq\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T\mathbf{R}_{\mathbf{H}_K}^{\dagger}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right).$$

The BB [11, Theorem 4] is obtained by taking the supremum of (8b) over all the existing degrees of freedom $\left(N,\boldsymbol{\theta}^N,K,\mathbf{H}_K\right)$. All known LBs on the MSE deriving from the BB can be obtained with appropriate instantiations

of (8b), that is with appropriate linear transformations of the MSB[5] (7b). For example, under mild regularity conditions on $p(\mathbf{x};\theta)$, the CRB is the limiting form of the HaCRB, that is the MSB where $N=2$, $\boldsymbol{\theta}^2=\left(\theta^0,\theta^0+d\theta\right)^T$ and $d\theta\to 0$ [11][12][13][17][55]. More generally, appropriate linear transformations of the MSB (8a-8b) for finite values of N and K lead to the Fraser-Gutman bound (FGB) [14], the Bhattacharyya bound (BaB) [10], the McAulay-Hofstetter bound (MHB), the Glave bound (GlB) [20], and the Abel bound (AbB) [22]. Furthermore, the class of LBs introduced lately in [24] can also be obtained as linear transformations of the MSB (8a-8b) in the limiting case where $N,K\to\infty$. It suffices to define each $\mathbf{h}_k$ as a vector of samples of a parametric function $h(\tau,\theta)$, $\tau\in\Lambda\subset\mathbb{R}$, integrable over Θ_d, $\forall\tau\in\Lambda$, i.e., $\mathbf{h}_k^T=\left(h\left(\tau_k,\theta^1\right),\ldots,h\left(\tau_k,\theta^N\right)\right)$, $1\leq k\leq K$. In such a setting, one obtains the integral form of (8b) (see [25, Section 2] for details) released in [24, (34-36)]:

$$MSE_{\theta^0}\left[\widehat{\theta^0}\right]\geq TTB^h_{\theta^0}=\int\limits_{\Lambda}\Gamma^h_{\theta^0}(\tau)\,\beta^h_{\theta^0}(\tau)\,d\tau, \tag{9a}$$

where $\Gamma^h_{\theta^0}(\tau)=\int\limits_{\Theta_d}h(\tau,\theta)\left(\theta-\theta^0\right)d\theta$, and $\beta^h_{\theta^0}(\tau)$ is the solution of the following integral equation:

$$\Gamma^h_{\theta^0}\left(\tau'\right)=\int\limits_{\Lambda}K^h_{\theta^0}\left(\tau',\tau\right)\beta^h_{\theta^0}(\tau)\,d\tau, \tag{9b}$$

$$K^h_{\theta^0}\left(\tau,\tau'\right)=\iint\limits_{\Theta_d^2}h(\tau,\theta)\,R_{\upsilon_{\theta^0}}\left(\theta,\theta'\right)h\left(\tau',\theta'\right)d\theta d\theta', \tag{9c}$$

$$R_{\upsilon_{\theta^0}}\left(\theta,\theta'\right)=E_{\mathbf{x};\theta^0}\left[\frac{p(\mathbf{x};\theta)}{p\left(\mathbf{x};\theta^0\right)}\frac{p\left(\mathbf{x};\theta'\right)}{p\left(\mathbf{x};\theta^0\right)}\right]. \tag{9d}$$

Note that if $h(\tau,\theta)=\delta(\tau-\theta)$ (limiting case of $\mathbf{H}_N=\mathbf{I}_N$ where $N=K\to\infty$) then $K^h_{\theta^0}(\tau,\tau')=R_{\upsilon_{\theta^0}}(\tau,\tau')$ and (9a) becomes the expression of the BB [23, (10)][44, (6-7)]. As mentioned above, in most practical cases, it is impossible to find an analytical solution of (9b) to obtain an explicit form of the $TTB^h_{\theta^0}$ (9a), which somewhat limits its interest. Nevertheless, as highlighted in [24], this formalism allows to use discrete or integral linear transforms of the LR, possibly non-invertible, possibly optimized for a set of p.d.f. (such as the Fourier transform) in order to get a tight approximation of the BB.

III. Modified lower bounds for non-standard estimation

In the previous Section II-B, we have pointed out that in standard estimation, the computability of the MSB (7b) is the cornerstone to generate the class of LBs on the MSE of uniformly unbiased estimate deriving from Barankin's work [11]. Therefore it seems sensible to check whether or not the MSB is computable in non-standard estimation.

A. A new look at modified lower bounds

If $\mathcal{S}_{\mathcal{X},\Theta_r}$, i.e., the support of $p(\mathbf{x},\boldsymbol{\theta}_r;\theta)$, is independent of θ, then:

$$E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[g(\mathbf{x},\boldsymbol{\theta}_r)\right]=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[g(\mathbf{x},\boldsymbol{\theta}_r)\,\upsilon_{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r;\theta)\right],\quad \upsilon_{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r;\theta)=\frac{p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right)}. \tag{10}$$

Therefore, for any unbiased estimator $\widehat{\theta^0}\in\mathcal{L}_2(\mathcal{S}_{\mathcal{X}})$, (7a) can be reformulated as, $\forall n\in[1,N]$:

$$\begin{aligned}\theta^n-\theta^0 &= E_{\mathbf{x};\theta^0}\left[\left(\widehat{\theta^0}(\mathbf{x})-\theta^0\right)\upsilon_{\theta^0}(\mathbf{x};\theta^n)\right]\\ &= E_{\mathbf{x};\theta^n}\left[\widehat{\theta^0}(\mathbf{x})-\theta^0\right]\\ &= E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^n}\left[\widehat{\theta^0}(\mathbf{x})-\theta^0\right]\\ &= E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}(\mathbf{x})-\theta^0\right)\upsilon_{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r;\theta^n)\right],\end{aligned}$$

[5]Since there is a one-to-one correspondence between a LB and a set of linear constraints, in the following, a linear transformation of a given LB actually refers to the LB obtained from a linear transformation of the corresponding set of linear constraints.

that is in vector form:

$$\begin{aligned}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right) &= E_{\mathbf{x};\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x};\boldsymbol{\theta}^N\right)\right]\\ &= E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}^N\right)\right],\end{aligned} \tag{11}$$

where $\left(\upsilon_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^1\right),\ldots,\upsilon_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^N\right)\right)=\boldsymbol{\upsilon}_{\theta^0}^T\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}^N\right)$. Additionally, since $\widehat{\theta^0}\in\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$, then:

$$E_{\mathbf{x};\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)^2\right]=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)^2\right]. \tag{12}$$

Therefore:

$$\min_{\widehat{\theta^0}\in\mathcal{L}_2(\mathcal{S}_{\mathcal{X}})}\left\{E_{\mathbf{x};\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)^2\right]\right\}\ \text{under}\ \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)=E_{\mathbf{x};\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x};\boldsymbol{\theta}^N\right)\right], \tag{13a}$$

is equivalent to:

$$\min_{\widehat{\theta^0}\in\mathcal{L}_2(\mathcal{S}_{\mathcal{X}})}\left\{E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)^2\right]\right\}\ \text{under}\ \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}^N\right)\right]. \tag{13b}$$

Note that the equivalence between (13a) and (13b) holds only if $\widehat{\theta^0}\in\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$. Unfortunately, since $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$ is a subspace of $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$, the solution of (13b) cannot be given by the minimum norm lemma (6) in general, since the lemma provides a solution in $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$, that is the solution of:

$$\min_{\widehat{\theta^0}\in\mathcal{L}_2(\mathcal{S}_{\mathcal{X},\Theta_r})}\left\{E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta^0\right)^2\right]\right\}\ \text{under}\ \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta^0\right)\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}^N\right)\right], \tag{13c}$$

yielding the following modified MSB:

$$MSE_{\theta^0}\left[\widehat{\theta^0}\right]\geq\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T\mathbf{R}_{\boldsymbol{\upsilon}_{\theta^0}}^{-1}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right),\quad\left(\mathbf{R}_{\boldsymbol{\upsilon}_{\theta^0}}\right)_{n,m}=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\upsilon_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^m\right)\upsilon_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^n\right)\right], \tag{14}$$

in the sense that it is a LB for unbiased estimates belonging to $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$. One noteworthy point is that the modified MSB (14) is obtained from the MSB (7b) by substituting $E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\ \right]$ for $E_{\mathbf{x};\theta^0}\left[\ \right]$ and $\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}^N\right)$ for $\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x};\boldsymbol{\theta}^N\right)$. More generally, since (13a) and (13c) share a similar formulation, reasoning by analogy, one can state that any approximation of the BB deriving from linear transformations of the set of constraints associated with the MSB (8a-8b), has an analog formulation in non-standard estimation obtained by substituting $E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\ \right]$ for $E_{\mathbf{x};\theta^0}\left[\ \right]$ and $\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}^N\right)$ for $\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x};\boldsymbol{\theta}^N\right)$. Actually, this is obtained by substituting $p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)$ for $p\left(\mathbf{x};\theta\right)$ in any approximation of the BB. This result holds whatever the prior p.d.f. depends or does not depend on the deterministic parameters. In the end, we have simply embedded the search of the locally-best unbiased estimator initially performed in the vector space $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$ (5) into a larger vector space containing $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$, namely $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$, where the search of the locally-best unbiased estimator is formulated as:

$$\min_{\widehat{\theta^0}\in\mathcal{L}_2(\mathcal{S}_{\mathcal{X},\Theta_r})}\left\{E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta^0\right)^2\right]\right\}\ \text{under}\ E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta^0\right]=\theta-\theta^0,\forall\theta\in\Theta_d. \tag{15}$$

Indeed, if $\widehat{\theta^0}\in\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)\subset\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$, then (15) reduces to (5). From this perspective, it seems appropriate to refer to these LBs for unbiased estimates belonging to $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$ as modified LBs (MLBs) as it has been proposed initially in [4] and [30] for the modified CRB. Since (13a) and (13b) are equivalent and $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)\subset\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$, it follows naturally that the modified form of a LB is looser (lower or equal) than the standard form of the LB. This highlights the trade-off associated with MLBs in non-standard estimation: computability at the possible expense of tightness. However it is possible to increase the tightness of MLBs by adding constraints in order to restrict the class of viable estimators $\widehat{\theta^0}\in\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$ and therefore to increase the minimum norm obtained from (13c) as shown hereinafter in Section III-B.

1) Old and new modified lower bounds:

In the light of the above, the MCRB in vector parameter estimation $\boldsymbol{\theta}$ is obtained directly from the CRB:

$$\mathbf{CRB}_{\boldsymbol{\theta}} = E_{\mathbf{x};\boldsymbol{\theta}}\left[\frac{\partial \ln p\left(\mathbf{x};\boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}}\frac{\partial \ln p\left(\mathbf{x};\boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}^T}\right]^{-1} \rightarrow \mathbf{MCRB}_{\boldsymbol{\theta}} = E_{\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}}\left[\frac{\partial \ln p\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}}\frac{\partial \ln p\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}^T}\right]^{-1}, \quad (16)$$

and one can assert that $\mathbf{CRB}_{\boldsymbol{\theta}} \geq \mathbf{MCRB}_{\boldsymbol{\theta}}$, without having to invoke neither the Jensen's inequality [4] nor to prove specific matrix inequality [30, (4)]. Furthermore, the MCRB expression (16) is still valid if the prior depends on $\boldsymbol{\theta}$, which extends the historical results provided in [4] and [30] under the restrictive assumption of a prior independent of $\boldsymbol{\theta}$. In the same way, the MBaB of order K is obtained from the BaB [10][44, (19)]:

$$\left|\begin{array}{l} BaB_\theta = \mathbf{e}_1^T E_{\mathbf{x};\theta}\left[\boldsymbol{\varrho}\left(\mathbf{x};\theta\right)\boldsymbol{\varrho}\left(\mathbf{x};\theta\right)^T\right]^{-1}\mathbf{e}_1, \\ \boldsymbol{\varrho}\left(\mathbf{x};\theta\right)^T = \frac{1}{p(\mathbf{x};\theta)}\left(\frac{\partial p(\mathbf{x};\theta)}{\partial \theta},\ldots,\frac{\partial^K p(\mathbf{x};\theta)}{\partial^K \theta}\right) \end{array}\right| \rightarrow \left|\begin{array}{l} MBaB_\theta = \mathbf{e}_1^T E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\boldsymbol{\varrho}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)\boldsymbol{\varrho}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)^T\right]^{-1}\mathbf{e}_1, \\ \boldsymbol{\varrho}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)^T = \frac{1}{p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}\left(\frac{\partial p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{\partial \theta},\ldots,\frac{\partial^K p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{\partial^K \theta}\right), \end{array}\right.$$

where $\mathbf{e}_1^T = (1,0,\ldots,0)$. Therefore, with the proposed approach, we not only extend the result introduced in [34, (4)] under the restrictive assumption of a prior independent of θ, but we can also assert that $BaB_\theta \geq MBaB_\theta$, which has not been proven in [34].

As with the CRB and the BaB, the modified form of all remaining BB approximations released in the open literature, namely the FGB [14], the MHB [19], the GlB [20], the AbB [22], and the CRFB [24, (101-102)], can be easily obtained with the proposed framework. For instance, the modified form of the general class of LBs (9a-9d) proposed in [24, (34-36)] is obtained simply by updating the definition of $R_{\upsilon_{\theta^0}}\left(\theta,\theta'\right)$ (9d) as follows:

$$R_{\upsilon_{\theta^0}}\left(\theta,\theta'\right) = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\frac{p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)}{p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right)}\frac{p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta'\right)}{p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right)}\right], \quad (17)$$

and one can also assert that $TTB^h_{\theta^0} \geq MTTB^h_{\theta^0}$.

B. A general class of tighter modified lower bounds and its relationship with hybrid lower bounds

As mentioned above, it is possible to increase the tightness of MLBs by adding constraints in order to restrict the class of viable estimators $\widehat{\theta^0} \in \mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$ and therefore to increase the minimum norm obtained from (13c). However, such additional constraints must keep on defining a subset of $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$ including the set of unbiased estimates belonging to $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$, as shown in the following with two general subsets of additional constraints. The first subset can be related to historical works on hybrid LBs [3][29] but addressed in a different way. The second subset is a generalization of [46, (8)] reformulated according to the proposed framework.

1) A first class of tighter modified lower bounds:

Since:

$$\begin{aligned} p\left(\mathbf{x};\theta\right) &= \int\limits_{\mathcal{S}_{\Theta_r|\mathbf{x}}} p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)d\boldsymbol{\theta}_r \\ &= \int\limits_{\mathbb{R}^{P_r}} p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)1_{\mathcal{S}_{\Theta_r|\mathbf{x}}}\left(\boldsymbol{\theta}_r\right)d\boldsymbol{\theta}_r, \end{aligned}$$

then, after change of variables $\boldsymbol{\theta}_r = \boldsymbol{\theta}'_r + \mathbf{h}_r$ and renaming $\boldsymbol{\theta}'_r$ as $\boldsymbol{\theta}_r$:

$$p\left(\mathbf{x};\theta\right) = \int\limits_{\mathbb{R}^{P_r}} p\left(\mathbf{x},\boldsymbol{\theta}_r+\mathbf{h}_r;\theta\right)1_{\mathcal{S}_{\Theta_r|\mathbf{x}}}\left(\boldsymbol{\theta}_r+\mathbf{h}_r\right)d\boldsymbol{\theta}_r.$$

Therefore for any $\mathbf{h}_r$ such that:

$$1_{\mathcal{S}_{\Theta_r|\mathbf{x}}}\left(\boldsymbol{\theta}_r+\mathbf{h}_r\right) = 1_{\mathcal{S}_{\Theta_r|\mathbf{x}}}\left(\boldsymbol{\theta}_r\right),\ \forall\boldsymbol{\theta}_r \in \mathbb{R}^{P_r}, \quad (18a)$$

then:

$$p\left(\mathbf{x};\theta\right)=\int\limits_{\mathcal{S}_{\Theta_r|\mathbf{x}}}p\left(\mathbf{x},\boldsymbol{\theta}_r+\mathbf{h}_r;\theta\right)d\boldsymbol{\theta}_r, \tag{18b}$$

and for any unbiased estimator $\widehat{\theta^0}\in\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$, (7a) can be reformulated as, $\forall n\in[1,N]$

$$\begin{aligned}\theta^n-\theta^0 &= E_{\mathbf{x};\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\upsilon_{\theta^0}\left(\mathbf{x};\theta^n\right)\right]\\ &= \int\limits_{\mathcal{S}_{\mathcal{X}}}\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)p\left(\mathbf{x};\theta^n\right)d\mathbf{x}\\ &= \int\limits_{\mathcal{S}_{\mathcal{X}}}\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\int\limits_{\mathcal{S}_{\Theta_r|\mathbf{x}}}p\left(\mathbf{x},\boldsymbol{\theta}_r+\mathbf{h}_r;\theta^n\right)d\boldsymbol{\theta}_r d\mathbf{x}\\ &= E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\upsilon_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r+\mathbf{h}_r;\theta^n\right)\right],\end{aligned}$$

that is in vector form:

$$\begin{aligned}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right) &= E_{\mathbf{x};\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x};\boldsymbol{\theta}^N\right)\right]\\ &= E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r+\mathbf{h}_r;\boldsymbol{\theta}^N\right)\right].\end{aligned} \tag{19}$$

The identity (19) means that for any $\widehat{\theta^0}\in\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$, the two subsets of N constraints are equivalent system of linear equations yielding the same vector subspace of $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$: $span\left(\upsilon_{\theta^0}\left(\mathbf{x};\theta^1\right),\ldots,\upsilon_{\theta^0}\left(\mathbf{x};\theta^N\right)\right)$. Therefore, for any $\widehat{\theta^0}\in\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$, any set of $N\times K$ constraints of the form:

$$\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r+\mathbf{h}_r^k;\boldsymbol{\theta}^N\right)\right], \tag{20}$$

where $\left\{\mathbf{h}_r^1,\ldots,\mathbf{h}_r^K\right\}$ satisfy (18a), is equivalent to the set of N constraints (11). Fortunately this result does not hold a priori for all $\widehat{\theta^0}\in\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$ where the $N\times K$ constraints (20) are expected to be linearly independent (not necessarily true in the general case). As mentioned above, the main effect of adding constraints is to restrict the class of viable estimators $\widehat{\theta^0}\in\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$ and therefore to increase the minimum norm obtained from (6):

$$\min_{\widehat{\theta^0}\in\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)}\left\{E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta^0\right)^2\right]\right\}\text{ under }\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta^0\right)\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r+\mathbf{h}_r^k;\boldsymbol{\theta}^N\right)\right], \tag{21}$$

$1\leq k\leq K$, which remains smaller (or equal) than the minimum norm obtained for $\widehat{\theta^0}\in\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$ given by (13a). This LB ordering was previously introduced in [29, (29)], but only in the restricted case where the prior does not depend on the deterministic parameter and $\mathcal{S}_{\Theta_r|\mathbf{x}}=\mathbb{R}^{P_r}$, at the expense of a not straightforward derivation (see Subsections III.C and III.D in [29]). Note that the regularity condition (18a) only imposes on $1_{\mathcal{S}_{\Theta_r|\mathbf{x}}}\left(\boldsymbol{\theta}_r\right)$, $\mathbf{x}\in\mathcal{S}_{\mathcal{X}}$, to be of the following form:

$$1_{\mathcal{S}_{\Theta_r|\mathbf{x}}}\left(\boldsymbol{\theta}_r\right)=\left|\begin{array}{l}0\text{ if }\displaystyle\sum_{\mathbf{h}_r\in\mathcal{F}_{\mathbf{x}}}\left(\sum_{l\in\mathbb{Z}}1_{\mathcal{S}^0_{\Theta_r|\mathbf{x}}}\left(\boldsymbol{\theta}_r+l\mathbf{h}_r\right)\right)=0,\\ 1,\text{ otherwise,}\end{array}\right. \tag{22}$$

where $\mathcal{F}_{\mathbf{x}}$ and $\mathcal{S}^0_{\Theta_r|\mathbf{x}}$ are subsets of $\mathbb{R}^{P_r}$, what means that the complementary of $\mathcal{S}_{\Theta_r|\mathbf{x}}$ is the union (possibly uncountable) of periodic subsets of $\mathbb{R}^{P_r}$.

2) A second class of tighter modified lower bounds:

Let us recall that any real-valued function $\psi\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)$ defined on $\mathcal{S}_{\mathcal{X},\Theta_r}$ satisfying

$$\int\limits_{\mathcal{S}_{\Theta_r|\mathbf{x}}} \psi\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right) p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right) d\boldsymbol{\theta}_r = 0, \tag{23}$$

is a Bayesian LB-generating functions [56]. A well known example is, for $\gamma \in]0,1[$:

$$\psi_{\gamma}^{\mathbf{h}_r}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right) = \left(p\left(\mathbf{x},\boldsymbol{\theta}_r+\mathbf{h}_r;\theta\right)/p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)\right)^{\gamma} - \left(p\left(\mathbf{x},\boldsymbol{\theta}_r-\mathbf{h}_r;\theta\right)/p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)\right)^{1-\gamma}, \tag{24}$$

if $\left(\mathbf{x},\boldsymbol{\theta}_r\right) \in \mathcal{S}_{\mathcal{X},\Theta_r}$, and $\psi_{\gamma}^{\mathbf{h}_r}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right) = 0$ otherwise, yielding the Bayesian Weiss-Weinstein bound (BWWB). Let $\boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)$ be a vector of L linearly independent functions satisfying (23). Then $\forall g\left(.\right) \in \mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$:

$$E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[g\left(\mathbf{x}\right)\boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right)\right] = \mathbf{0}, \tag{25a}$$

which means that the subspace $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$ is orthogonal to $span\left\{\psi_1\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right),\ldots,\psi_L\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right)\right\}$ in $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$. Therefore, since (13a) can be reformulated as (13b), it is straightforward that (13a) is equivalent to:

$$\min_{\widehat{\theta^0}\in\mathcal{L}_2(\mathcal{S}_{\mathcal{X}})} \left\{E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)^2\right]\right\} \text{ under } \left| \begin{array}{r} \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right) = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x};\boldsymbol{\theta}^N\right)\right] \\ \mathbf{0} = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right)\right] \end{array} \right. . \tag{25b}$$

In other words, the addition of the set of L constraints $E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right)\right] = \mathbf{0}$ to any linear transformation of (7a) does not change the associated LB (8b) computed for $\widehat{\theta^0} \in \mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$. Fortunately, once again, this result does not hold a priori for all $\widehat{\theta^0} \in \mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$. Indeed, provided that $\boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)$ is chosen such that $E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}^N\right)\boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right)^T\right] \neq \mathbf{0}$ [55, Lemma 2], one can increase the minimum norm obtained from (13c) by computing:

$$\min_{\widehat{\theta^0}\in\mathcal{L}_2(\mathcal{S}_{\mathcal{X},\Theta_r})} \left\{E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta^0\right)^2\right]\right\} \text{ under } \left| \begin{array}{r} \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right) = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta^0\right)\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}^N\right)\right] \\ \mathbf{0} = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta^0\right)\boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right)\right] \end{array} \right. , \tag{26}$$

which remains smaller (or equal) than the minimum norm obtained for $\widehat{\theta^0} \in \mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$ given by (13a). First note that it is in general not possible to compare (21) with (26) since they derive from different subset of constraints. Second, (26) can be used with joint p.d.f. $p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)$ which does not satisfy the regularity condition (22) since functions (24) are essentially free of regularity conditions [56].

3) A general class of tighter modified lower bounds and its relationship with hybrid lower bounds:

The tightest modified LBs are obtained by combination of constraints (21) and (26) as the solution of:

$$\min_{\widehat{\theta^0}\in\mathcal{L}_2(\mathcal{S}_{\mathcal{X},\Theta_r})} \left\{E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta^0\right)^2\right]\right\} \text{ under } \left| \begin{array}{r} \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right) = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta^0\right)\boldsymbol{\upsilon}_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r+\mathbf{h}_r^k;\boldsymbol{\theta}^N\right)\right] \\ \mathbf{0} = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta^0\right)\boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right)\right], \end{array} \right. \tag{27}$$

$1 \leq k \leq K$, where $\boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right)$ satisfies (23).

Firstly, if we restrict (27) to (21), that is no function $\psi\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right)$ (23) is involved, then the solution of (21)(27) given by the minimum norm lemma (6) yields the deterministic part of the HLBs obtained as discrete forms [45, (30)] of linear transformations on the CLR function introduced in [45]. Following from similar argument given in Section II-B or in [46, Section III], one obtains the deterministic part of the HLB integral form proposed in [45], as linear transformations of (21) in the limiting case where $N, K \to \infty$. Note that in [45] two restrictive regularity conditions are assumed: i) $\mathcal{S}_{\Theta_r|\mathbf{x}} = \mathbb{R}^{P_r}$, ii) the prior does not depend on θ, which are relaxed with the proposed framework: HLBs obtained via linear transformations on the CLR function are still valid if the prior depends on θ as long as $\mathcal{S}_{\Theta_r|\mathbf{x}}$ satisfies (22), which includes $\mathbb{R}^{P_r}$. In contrast, according to (22) such bounds do not exist if $\mathcal{S}_{\Theta_r|\mathbf{x}}$ is a connected set of $\mathbb{R}^{P_r}$ and in most cases, if $\mathcal{S}_{\Theta_r|\mathbf{x}}$ is a disconnected subset of $\mathbb{R}^{P_r}$ (not stated in [45]). Last,

since the modified LB obtained from (21) is lower than or equal to the standard LB from (13a), one can assert that the deterministic part of any HLB obtained via the linear transformation on the CLR is looser (or equal) than the corresponding standard LB (not proven in [45]).

Secondly, if $\mathcal{S}_{\Theta_r|\mathbf{x}}$ does not satisfy (22), e.g., if $\mathcal{S}_{\Theta_r|\mathbf{x}}$ is an interval, then (21) can no longer be used to increase the minimum norm obtained from (13c). One solution is therefore to restrict (27) to (26) and, following from similar argument given in Section II-B, to resort to some of its possible integral forms obtained as the limiting cases where $N, K \to \infty$, where L has a finite value:

$$\min_{\widehat{\theta^0} \in \mathcal{L}_2(\mathcal{S}_{\mathcal{X},\Theta_r})} \left\{ E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0} \left[\left(\widehat{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r) - \theta^0 \right)^2 \right] \right\} \text{ under } \left| \begin{array}{r} \Gamma^h_{\theta^0}(\tau) = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0} \left[\left(\widehat{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r) - \theta^0 \right) \eta^h_{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r;\tau) \right] \\ \mathbf{0} = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0} \left[\left(\widehat{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r) - \theta^0 \right) \boldsymbol{\psi}(\mathbf{x},\boldsymbol{\theta}_r;\theta^0) \right] \end{array} \right., \tag{28a}$$

where $\eta^h_{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r;\tau) = \int\limits_{\Theta_d} h(\tau,\theta)\, \upsilon_{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r;\theta)\, d\theta$ and $\Gamma^h_{\theta^0}(\tau) = \int\limits_{\Theta_d} h(\tau,\theta)\left(\theta - \theta^0\right) d\theta$, or where $L \to \infty$, if we choose (24):

$$\min_{\widehat{\theta^0} \in \mathcal{L}_2(\mathcal{S}_{\mathcal{X},\Theta_r})} \left\{ E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0} \left[\left(\widehat{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r) - \theta^0 \right)^2 \right] \right\} \text{ under } \left| \begin{array}{r} \Gamma^h_{\theta^0}(\tau) = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0} \left[\left(\widehat{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r) - \theta^0 \right) \eta^h_{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r;\tau) \right] \\ \mathbf{0} = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0} \left[\left(\widehat{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r) - \theta^0 \right) \kappa(\mathbf{x},\boldsymbol{\theta}_r;\theta^0) \right] \end{array} \right., \tag{28b}$$

where $\kappa(\mathbf{x},\boldsymbol{\theta}_r;\theta^0) = \iint\limits_{\gamma,\mathbf{h}_r} \psi^{\mathbf{h}_r}_{\gamma}(\mathbf{x},\boldsymbol{\theta}_r;\theta^0)\, d\mathbf{h}_r d\gamma$. To the best of our knowledge, (28b) defines a new class of MLBs, whereas (28a) is a particular case of the deterministic part of HLBs introduced in [46].

Thirdly, if we restrict ourselves to the case where $\mathcal{S}_{\Theta_r|\mathbf{x}} = \mathbb{R}^{P_r}$, then (27) and its possible integral forms provide an extended class of MLBs which can be regarded as the deterministic part of an extended class of HLBs which includes all known HLBs introduced so far as shown in [45] and [46].

It is noteworthy that the proposed unified framework allows to reformulate all known LBs on the MSE for unbiased estimates in non-standard estimation as a MLBs without any regularity condition on the (nuisance) random vector estimates, since $\boldsymbol{\theta}_r$ is neither required nor expected to be estimated. Moreover, since any modified LB obtained from (27) is lower than or equal to its standard form (13a), one can assert that the deterministic part of any HLB is looser (or equal) than the corresponding standard LB, which is a new general result. Last, the deterministic part of any HLB is a valid MLBs whatever the prior depends on or does not depend on the deterministic parameter θ, which is another new general result.

4) Old and new tighter modified lower bounds:

A typical example is the case of the CRB. A tighter MCRB obtained from (21) for $N = 2$, $K = P_r$, where $\boldsymbol{\theta}^2 = \left(\theta^0 + d\theta, \theta^0\right)$ and $\mathbf{h}^k_r = \mathbf{u}_k h^k_r$, $1 \leq k \leq P_r$, leading to the following subset of constraints:

$$\mathbf{v} = d\theta \begin{pmatrix} 0 \\ \mathbf{e}_1 \end{pmatrix} = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0} \left[\left(\widehat{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r) - \theta^0 \right) \mathbf{c}_{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r) \right],$$

$$\mathbf{c}^T_{\theta^0}(\mathbf{x},\boldsymbol{\theta}_r) = \left(\upsilon_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right), \upsilon_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0 + d\theta\right), \upsilon_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r + \mathbf{u}_1 h^1_r;\theta^0\right), \ldots, \upsilon_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r + \mathbf{u}_{P_r} h^{P_r}_r;\theta^0\right) \right)$$

where $\mathbf{e}_1 = (1,0,\ldots,0)^T$ and $\mathbf{u}_k$ is the kth column of the identity matrix $\mathbf{I}_{P_r}$. By letting $\left(d\theta, h^1_r, \ldots, h^{P_r}_r\right)$ be infinitesimally small, which imposes that (22) reduces to: $\forall \mathbf{x} \in \mathcal{S}_{\mathcal{X}}$, $\mathcal{S}_{\Theta_r|\mathbf{x}} = \mathbb{R}^{P_r}$, the LB obtained from (6) is:

$$\overline{MCRB}_{\theta^0} = \mathbf{e}^T_1 \mathbf{F}\left(\theta^0\right)^{-1} \mathbf{e}_1, \quad \mathbf{F}(\theta) = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta} \left[\frac{\partial \ln p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{\partial\left(\theta, \boldsymbol{\theta}^T_r\right)^T} \frac{\partial \ln p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{\partial\left(\theta, \boldsymbol{\theta}^T_r\right)} \right]. \tag{29}$$

Since $\mathbf{F}(\theta) = \begin{bmatrix} f_\theta(\theta) & \mathbf{f}^T_{\boldsymbol{\theta}_r,\theta}(\theta) \\ \mathbf{f}_{\boldsymbol{\theta}_r,\theta}(\theta) & \mathbf{F}_{\boldsymbol{\theta}_r}(\theta) \end{bmatrix}$, therefore:

$$\overline{MCRB}_{\theta^0} = \frac{1}{f_\theta\left(\theta^0\right) - \mathbf{f}^T_{\boldsymbol{\theta}_r,\theta}\left(\theta^0\right) \mathbf{F}^{-1}_{\boldsymbol{\theta}_r}\left(\theta^0\right) \mathbf{f}_{\theta,\boldsymbol{\theta}_r}\left(\theta^0\right)} \geq \frac{1}{f_\theta\left(\theta^0\right)} = MCRB_{\theta^0}. \tag{30}$$

Actually (29) is also the deterministic part of the HCRB [3] and a similar derivation was proposed in [29], but under the unnecessary restrictive assumption of a prior independent of θ, as in [3] which introduced the HCRB as an

extension of the Bayesian CRB proposed in [1]. This condition was relaxed in [28, (20)] with sufficient conditions [28, (21)] unnecessary restrictive and which have been an impediment to the dissemination of their result. However if $\mathcal{S}_{\Theta_r|\mathbf{x}}$ is an interval of $\mathbb{R}^{P_r}$, then the tighter MCRB (29) cannot be derived any longer. Fortunately, as shown with the proposed rationale, an alternative tighter MCRB can be derived from (26). Indeed, for $N=2$, where $\boldsymbol{\theta}^2=\left(\theta^0+d\theta,\theta^0\right)$, the following subset of constraints:

$$\mathbf{v}=d\theta\begin{pmatrix}0\\ \mathbf{e}_1\end{pmatrix}=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\left(\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta^0\right)\mathbf{c}_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)\right],$$
$$\mathbf{c}_{\theta^0}^T\left(\mathbf{x},\boldsymbol{\theta}_r\right)=\left(\upsilon_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right),\upsilon_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0+d\theta\right),\boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^0\right)^T\right)$$

yields, as a limiting form where $d\theta\rightarrow 0$, via Lemma (6):

$$\overline{MCRB}_{\theta^0}^a=\mathbf{e}_1^TE_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\begin{pmatrix}\frac{\partial\ln p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{\partial\theta}\\ \boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)\end{pmatrix}\begin{pmatrix}\frac{\partial\ln p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{\partial\theta}\\ \boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)\end{pmatrix}^T\right]^{-1}\mathbf{e}_1,$$

that is:

$$\overline{MCRB}_{\theta^0}^a=\frac{1}{f_\theta\left(\theta^0\right)-\mathbf{f}\left(\theta^0\right)\mathbf{F}\left(\theta^0\right)^{-1}\mathbf{f}\left(\theta^0\right)}\geq\frac{1}{f_\theta\left(\theta^0\right)}=MCRB_{\theta^0},$$

where $\mathbf{F}\left(\theta\right)=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)\boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)^T\right]$, $\mathbf{f}\left(\theta\right)=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\boldsymbol{\psi}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)\frac{\partial\ln p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{\partial\theta}\right]$ and $f_\theta\left(\theta\right)=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\frac{\partial\ln p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{\partial\theta}\right)^2\right]$. In the same way, one could easily proposed an alternative to the tighter MBaB of order K deriving from (21), suitable to estimation problems for which $\mathcal{S}_{\Theta_r|\mathbf{x}}$ does not satisfy (22). Or for the modified form of any known standard LB.

C. On closeness, tightness, regularity conditions and implementation of modified lower bounds

1) On the closeness of MLBs to LBs:

Actually, a "closeness condition" required to obtain a modified LB equal to the standard LB is quite simple to express: it is necessary, and sufficient, that the estimator solution of the norm minimization under linear constraints (13c)(21)(26)(27) belongs to $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$, that is according to (6):

$$\widehat{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)_{opt}-\theta^0=\sum_{k=1}^K\alpha_k\left(\mathbf{c}_{\theta^0}\left(\mathbf{x},\boldsymbol{\theta}_r\right)\right)_k\in\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right),\tag{31}$$

a closeness condition fulfilled by a class of joint p.d.f. $p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)$ which depends on the vector of constraint functions chosen. For example, if we consider the $\overline{MCRB}_{\theta^0}$ (29) then the tightness condition is:

$$\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)_{opt}-\theta=\frac{\partial\ln p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)}{\partial\left(\theta,\boldsymbol{\theta}_r^T\right)}\boldsymbol{\alpha}\left(\theta\right)=\widehat{\theta}\left(\mathbf{x}\right)_{opt}-\theta.\tag{32}$$

Since $\mathbf{e}_1^T\mathbf{F}\left(\theta\right)^{-1}=\overline{MCRB}_\theta\left(1,-\mathbf{F}_{\boldsymbol{\theta}_r}^{-1}\mathbf{f}_{\boldsymbol{\theta}_r,\theta}\left(\theta\right)\right)$, therefore (32) is equivalent to:

$$\frac{\partial\ln p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)}{\partial\theta}-\mathbf{f}_{\theta,\boldsymbol{\theta}_r}\left(\theta\right)\mathbf{F}_{\boldsymbol{\theta}_r}^{-1}\left(\theta\right)\frac{\partial\ln p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)}{\partial\boldsymbol{\theta}_r}=\frac{\widehat{\theta}\left(\mathbf{x}\right)_{opt}-\theta}{\overline{MCRB}_\theta},$$

leading to the necessary, and sufficient, condition:

$$\frac{\partial\ln p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)}{\partial\theta\partial\boldsymbol{\theta}_r^T}=\mathbf{f}_{\theta,\boldsymbol{\theta}_r}\left(\theta\right)\mathbf{F}_{\boldsymbol{\theta}_r}^{-1}\left(\theta\right)\frac{\partial\ln p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)}{\partial\boldsymbol{\theta}_r\partial\boldsymbol{\theta}_r^T},\tag{33}$$

which has been introduced in [52, (34)] at the expense of a quite complex proof.

2) On tightness, regularity conditions and implementation of MLBs:

As mentioned above, the trade-off associated with MLBs is computability at the possible expense of tightness. Indeed, a key feature of the simplest form of the MLBs deriving from the MMSB (14), is to be essentially free of regularity conditions both on the joint p.d.f. $p(\mathbf{x}, \boldsymbol{\theta}_r; \theta)$ w.r.t the random parameters $\boldsymbol{\theta}_r$, and on the support $\mathcal{S}_{\Theta_r|\mathbf{x}}$. This feature still holds for the tighter MLBs obtained with Bayesian LB-generating functions (26)(28a-28b). In contrast, none of the existing HLBs, which are all obtained via linear transformations on the CLR function [45], can be used if $\mathcal{S}_{\Theta_r|\mathbf{x}}$ does not satisfy (22), that is, for instance, if $\mathcal{S}_{\Theta_r|\mathbf{x}}$ is a connected set of $\mathbb{R}^{P_r}$ and in most cases, if $\mathcal{S}_{\Theta_r|\mathbf{x}}$ is a disconnected subset of $\mathbb{R}^{P_r}$. Off course, any time an existing HLB can be derived, its deterministic part provides a tighter LB than the corresponding MLB deriving from (14), however at the expense of an increased computational cost (see (29)). Thus, the proposed unified framework is a useful tool to look for the best possible trade-off between tightness, regularity conditions and computational cost, in the choice of a MLB for a given non-standard estimation problem. In that perspective, non-standard estimation can take advantage of the works on computable approximations of the BB in standard estimation [22][23][24], which have shown that the CRB and the BB can be regarded as key representatives of two general classes of bounds, respectively the Small-Error bounds and the Large-Error bounds. Indeed, it is now well known that the Small-Error bounds, such as the CRB, are optimistic bounds in a non-linear estimation problem where the outliers effect generally appears [57][58][59]. This outliers effect leads to a characteristic threshold behavior of estimators MSE which exhibits a "performance breakdown" highlighted by Large-Error bounds [17]. Furthermore, it has been underlined that under the norm minimization approach, the Small Error bounds derive from linear constraints expressed at the true value θ^0 only, whereas the Large-Error bounds derive from linear constraints expressed at vectors of test points $\boldsymbol{\theta}^{N+1}$ including the true value [17][20][22][23][24]. The tightness of a given Large Error bound is at the expense of some computational cost; indeed as its tightness depends on the used vector of test points [24][60], it generally incorporates the search of an optimum over a set of vectors of test points. As a consequence, the final practical form proposed by each author is an attempt to optimize the trade-off between tightness and computational cost. For example, in [17] the main goal was to reduce the complexity of use of the BB by substituting the simplified form (7b) for the initial form (8b). In [19] and generalized in [20] and [22], the rationale is to combine a Small Error bound (CRB [19][20] or BaB [22]) with a Large Error bound (MSB [17]) in order to obtain a bound which accounts for both local and large errors and is able to handle the threshold phenomena. Indeed the use of derivatives is also helpful to decrease the computational burden since it allows to resort to smaller sets of tests point vectors to achieve similar tightness [22][24], however at the expense of the existence of the derivatives (although this condition is mild and generally satisfied). Last, the norm minimization approach naturally incorporates possible tightness comparison between two MLBs. Indeed, if the subset of linear constraints associated with a MLB is included into the subset of linear constraints associated with another MLB, then the latter one is tighter. To wrap up, when looking at a MLB, the following questions should be answered: i) is the strong regularity condition (22) satisfied?, e.g., is $\mathcal{S}_{\Theta_r|\mathbf{x}} \triangleq \mathbb{R}^{P_r}$?, ii) is the joint p.d.f. $p(\mathbf{x}, \boldsymbol{\theta}_r; \boldsymbol{\theta})$ differentiable w.r.t the random parameters? iii) which regions of operation of estimators, among the asymptotic region, the threshold region and the a priori region [58], are of interest?, iv) are analytic forms of $E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\boldsymbol{\psi}\left(\mathbf{x}, \boldsymbol{\theta}_r; \theta^0\right) \boldsymbol{\upsilon}_{\theta^0}^T\left(\mathbf{x}, \boldsymbol{\theta}_r; \boldsymbol{\theta}^N\right)\right]$ and $E_{\mathbf{x},\boldsymbol{\theta}_r;\theta^0}\left[\boldsymbol{\psi}\left(\mathbf{x}, \boldsymbol{\theta}_r; \theta^0\right) \boldsymbol{\psi}\left(\mathbf{x}, \boldsymbol{\theta}_r; \theta^0\right)^T\right]$ available? For instance, if the asymptotic and threshold regions are of interest, and if the answers to i), ii), iv) are positive, then at the expense of non negligible computational burden, the tightest MLBs will be probably obtained by deriving from (27) a combination of the tight version of the modified GlB (or of the modified CRFB) with Bayesian LB-generating functions (23), since the MLB obtained will incorporate most of the meaningful constraints available. If the asymptotic and threshold regions are of interest, and if the answers to i), ii), iv) are negative, then at the expense of a non negligible computational burden, the tightest MLBs is the MMSB (14).

IV. Non-standard maximum likelihood estimator for deterministic estimation

Let us recall that the widespread use of MLEs in deterministic estimation originates from the fact that, under reasonably general conditions on the observation model [8][61], the MLEs are asymptotically uniformly unbiased, Gaussian distributed and efficient when the number of independent observations tends to infinity. Additionally, if the observation model is Gaussian complex circular, some additional asymptotic regions of operation yielding uniformly unbiased Gaussian and efficient MLEs have also been identified at finite number of independent observations

[62][63][64][65][66]. If a closed-form of $p(\mathbf{x};\theta)$ does not exist or if a closed-form of $p(\mathbf{x};\theta)$ does exist but the resulting expression is intractable to derive the standard MLE of θ:

$$\widehat{\theta}_{ML}(\mathbf{x}) = \arg\max_{\theta\in\Theta_d}\{p(\mathbf{x};\theta)\}, \tag{34a}$$

a sensible solution in the search of a realizable estimator based on the ML principle is to look for:

$$\left(\widehat{\underline{\boldsymbol{\theta}_r}}(\mathbf{x}), \underline{\widehat{\theta}}(\mathbf{x})\right) = \arg\max_{\theta\in\Theta_d,\boldsymbol{\theta}_r\in\mathcal{S}_{\Theta_r|\mathbf{x}}}\{p(\mathbf{x}|\boldsymbol{\theta}_r;\theta)\}. \tag{34b}$$

In the following $\underline{\widehat{\theta}}(\mathbf{x})$ and $\widehat{\underline{\boldsymbol{\theta}_r}}(\mathbf{x})$ (34b) are referred to as "non-standard" MLEs (NSMLEs). The underlying idea is that, since in many estimation problems [1][2][3][4] $p(\mathbf{x},\boldsymbol{\theta}_r;\theta)$ is a compound probability distribution, i.e., $p(\mathbf{x},\boldsymbol{\theta}_r;\theta) = p(\mathbf{x}|\boldsymbol{\theta}_r;\theta)\,p(\boldsymbol{\theta}_r;\theta)$, the closed-form of $p(\mathbf{x}|\boldsymbol{\theta}_r;\theta)$ is known and the NSMLEs (34b) take advantage not only of the aforementioned properties, and in particular of the asymptotic uniform unbiasedeness w.r.t. $p(\mathbf{x}|\boldsymbol{\theta}_r;\theta)$, but also of the extensive open literature on MLE closed-form expressions or approximations [47]. These key features clearly make the "non-standard" maximum likelihood estimation more attractive than the two known following alternative approaches. The first alternative approach consists in deriving the joint maximum a posteriori-maximum likelihood estimate (JMAPMLE) of the hybrid parameter vector $\left(\boldsymbol{\theta}_r^T,\theta\right)$:

$$\left(\widehat{\boldsymbol{\theta}_r}_J(\mathbf{x}), \widehat{\theta}_J(\mathbf{x})\right) = \arg\max_{\theta\in\Theta_d,\boldsymbol{\theta}_r\in\mathcal{S}_{\Theta_r|\mathbf{x}}}\{p(\mathbf{x},\boldsymbol{\theta}_r;\theta)\}, \tag{35}$$

but suffers from a major drawback: the JMAPMLE is biased and inconsistent whatever the number of independent observations [48], except for a class of hybrid estimation problems yielding (34b) when the number of independent observations tends to infinity [50, p. 6, 12]. One point worthy of note is that the JMAPMLE may outperform the MLE (34a) in terms of MSE, especially with short data records, where MLE is indeed disarmed of its asymptotic optimality [48]. However the biasedness of the JMAPMLE prevents from the comparison of its MSE with deterministic LBs. Indeed, if any known bias can be taken into account in deterministic LBs formulation [11], the bias depends on the specific estimator and, furthermore, is hardly ever known in practice. The second alternative approach consists in resorting to the expectation-maximization (EM) algorithm [49]. In the general case the EM algorithm converges to a stationary point of $\ln p(\mathbf{x};\theta)$. The stationary point need not, however, be a local maximum. Indeed, if it is shown that, under suitable regularity conditions [67], it converges to the MLE (34a), it is also shown [68] that it is possible for the algorithm to converge to local minima or saddle points in unusual cases. Moreover, in non-standard estimation, the EM algorithm consists in the following iterative procedure:

$$\theta_{n+1} = \arg\max_{\theta\in\Theta_d}\left\{E_{\boldsymbol{\theta}_r|\mathbf{x};\theta_n}\left[\ln p(\mathbf{x},\boldsymbol{\theta}_r;\theta)\right]\right\}, \tag{36}$$

which is unlikely to be of practical use in many estimation problems of interest where $p(\boldsymbol{\theta}_r;\theta)$ is not a conjugate prior for the likelihood function $p(\mathbf{x}|\boldsymbol{\theta}_r;\theta)$ and $p(\boldsymbol{\theta}_r|\mathbf{x};\theta)$ is not computable. Last, in any case where the EM algorithm converge to the MLE (34a), its MSE is lower bounded by the MLBs.

A. Performance comparison

For that purpose, let us denote $\boldsymbol{\phi} = \left(\theta,\boldsymbol{\theta}_r^T\right)^T \in \Theta_d\times\mathbb{R}^{P_r}$, $p(\mathbf{x}|\boldsymbol{\phi}) \triangleq p(\mathbf{x}|\boldsymbol{\theta}_r;\theta)$ and $E_{\mathbf{x}|\boldsymbol{\phi}}[\,] \triangleq E_{\mathbf{x}|\boldsymbol{\theta}_r;\theta}[\,]$. Then any estimator $\widehat{\boldsymbol{\phi}}^T = \left(\widehat{\theta},\widehat{\boldsymbol{\theta}}_r^T\right) \in \mathcal{L}_2(\mathcal{S}_{\mathcal{X},\Theta_r})$, i.e., $\widehat{\boldsymbol{\phi}} \triangleq \widehat{\boldsymbol{\phi}}(\mathbf{x},\boldsymbol{\theta}_r)$, of a selected vector value $\boldsymbol{\phi}$[6] uniformly strict-sense unbiased [52], i.e., w.r.t. $p(\mathbf{x}|\boldsymbol{\phi})$, must comply with:

$$\forall\boldsymbol{\phi}' = \begin{pmatrix}\theta'\\ \boldsymbol{\theta}_r'\end{pmatrix} \in \Theta_d\times\mathbb{R}^{P_r} : E_{\mathbf{x}|\boldsymbol{\phi}'}\left[\widehat{\boldsymbol{\phi}}\right] = \boldsymbol{\phi}', \tag{37}$$

which implies that:

$$\forall\theta'\in\Theta_d : E_{\mathbf{x},\boldsymbol{\theta}_r';\theta'}\left[\widehat{\boldsymbol{\phi}}\right] = E_{\boldsymbol{\theta}_r';\theta'}\left[\boldsymbol{\phi}'\right] = \begin{pmatrix}\theta'\\ E_{\boldsymbol{\theta}_r';\theta'}\left[\boldsymbol{\theta}_r'\right]\end{pmatrix}, \tag{38}$$

[6]In this section, for sake of legibility, $\boldsymbol{\psi}$ denotes either the vector of unknown parameters or a selected vector value $\boldsymbol{\psi} \triangleq \boldsymbol{\psi}^0 = \left(\theta^0,\left(\boldsymbol{\theta}_r^0\right)^T\right)^T$.

that is $\widehat{\boldsymbol{\phi}} \in \mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$ is a uniformly wide-sense unbiased[7] [52] estimate of $\mathbf{g}\left(\theta\right)^T = \left(\theta, E_{\boldsymbol{\theta}_r;\theta}\left[\boldsymbol{\theta}_r^T\right]\right)$, i.e., w.r.t. $p\left(\mathbf{x}, \boldsymbol{\theta}_r; \theta\right)$. As the reciprocal is not true:

$$\forall \theta' \in \Theta_d : E_{\mathbf{x},\boldsymbol{\theta}_r';\theta'}\left[\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}'\right] = \mathbf{0} \;\not\Rightarrow\; \forall \boldsymbol{\phi}' \in \Theta_d \times \mathbb{R}^{P_r} : E_{\mathbf{x}|\boldsymbol{\phi}'}\left[\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}'\right] = \mathbf{0},$$

then $\mathcal{U}_S\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right) = \left\{\widehat{\boldsymbol{\phi}} \in \mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right) \text{ verifying (37)}\right\} \subset \mathcal{U}_W\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right) = \left\{\widehat{\boldsymbol{\phi}} \in \mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right) \text{ verifying (38)}\right\}$[8]. Let $\mathcal{U}_S\left(\mathcal{S}_{\mathcal{X}}\right)$ and $\mathcal{U}_W\left(\mathcal{S}_{\mathcal{X}}\right)$ denote the restriction to $\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X}}\right)$ of $\mathcal{U}_S\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$ and $\mathcal{U}_W\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$. First, $\forall \widehat{\boldsymbol{\phi}} \in \mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$:

$$E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}\left(\theta\right)\right)\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}\left(\theta\right)\right)^T\right] = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\boldsymbol{\phi}} - E_{\mathbf{x}|\boldsymbol{\phi}}\left[\widehat{\boldsymbol{\phi}}\right]\right)\left(\widehat{\boldsymbol{\phi}} - E_{\mathbf{x}|\boldsymbol{\phi}}\left[\widehat{\boldsymbol{\phi}}\right]\right)^T\right] + E_{\boldsymbol{\theta}_r;\theta}\left[\left(E_{\mathbf{x}|\boldsymbol{\phi}}\left[\widehat{\boldsymbol{\phi}}\right] - \mathbf{g}\left(\theta\right)\right)\left(E_{\mathbf{x}|\boldsymbol{\phi}}\left[\widehat{\boldsymbol{\phi}}\right] - \mathbf{g}\left(\theta\right)\right)^T\right]. \quad (39a)$$

Therefore, if $\widehat{\boldsymbol{\phi}} \in \mathcal{U}_S\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$:

$$E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}\left(\theta\right)\right)\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}\left(\theta\right)\right)^T\right] = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)^T\right] + \mathbf{C}_\theta\left(\boldsymbol{\phi}\right), \quad (39b)$$

where:

$$\mathbf{C}_\theta\left(\boldsymbol{\phi}\right) = E_{\boldsymbol{\theta}_r;\theta}\left[\left(\boldsymbol{\phi} - E_{\boldsymbol{\theta}_r;\theta}\left[\boldsymbol{\phi}\right]\right)\left(\boldsymbol{\phi} - E_{\boldsymbol{\theta}_r;\theta}\left[\boldsymbol{\phi}\right]\right)^T\right] = \begin{bmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{C}_\theta\left(\boldsymbol{\theta}_r\right) \end{bmatrix}.$$

Second, as $\mathcal{U}_S\left(\mathcal{S}_{\mathcal{X}}\right) \subset \mathcal{U}_W\left(\mathcal{S}_{\mathcal{X}}\right)$ and $\mathcal{U}_S\left(\mathcal{S}_{\mathcal{X}}\right) \subset \mathcal{U}_S\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$ and, finally:

$$\min_{\widehat{\boldsymbol{\phi}} \in \mathcal{U}_W\left(\mathcal{S}_{\mathcal{X}}\right)} \left\{E_{\mathbf{x}|\theta}\left[\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}\left(\theta\right)\right)\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}\left(\theta\right)\right)^T\right]\right\} \leq \min_{\widehat{\boldsymbol{\phi}} \in \mathcal{U}_S\left(\mathcal{S}_{\mathcal{X}}\right)} \left\{E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)^T\right]\right\} + \mathbf{C}_\theta\left(\boldsymbol{\phi}\right), \quad (40a)$$

and, in particular:

$$\min_{\widehat{\theta} \in \mathcal{U}_W\left(\mathcal{S}_{\mathcal{X}}\right)} \left\{E_{\mathbf{x}|\theta}\left[\left(\widehat{\theta} - \theta\right)^2\right]\right\} \leq \min_{\widehat{\theta} \in \mathcal{U}_S\left(\mathcal{S}_{\mathcal{X}}\right)} \left\{E_{\mathbf{x}|\theta}\left[\left(\widehat{\theta} - \theta\right)^2\right]\right\}. \quad (40b)$$

If we consider an asymptotic region of operation [8][61][62][63][64][65][66] for both $\widehat{\theta}_{ML}\left(\mathbf{x}\right)$ and $\widehat{\underline{\theta}}\left(\mathbf{x}\right)$, then $\widehat{\theta}_{ML}\left(\mathbf{x}\right)$ is wide-sense unbiased, i.e., $\widehat{\theta}_{ML} \in \mathcal{U}_W\left(\mathcal{S}_{\mathcal{X}}\right)$, $\widehat{\underline{\theta}}\left(\mathbf{x}\right)$ is strict-sense unbiased, i.e., $\widehat{\underline{\theta}} \in \mathcal{U}_S\left(\mathcal{S}_{\mathcal{X}}\right)$, and (40b) holds for $\widehat{\theta}_{ML}$ and $\widehat{\underline{\theta}}$. Thus, the NSMLEs of θ is in general an asymptotically suboptimal estimator of θ (in the MSE sense) in comparison with the MLE of θ within the set of unbiased estimates in the Barankin sense (4a). Therefore, from a theoretical as well as a practical point of view, it is of interest to investigate on a possible quantification of the suboptimality of the NSMLE, which can be obtained in some extent by LBs derivation and comparison.

B. Non-standard lower bounds

For any $\widehat{\boldsymbol{\phi}} \in \mathcal{U}_S\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$, let $\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}\right) = E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)^T\right]$ denotes its covariance matrix w.r.t. $p\left(\mathbf{x}|\boldsymbol{\phi}\right)$. Then by noticing that, $\forall \widehat{\boldsymbol{\phi}} \in \mathcal{U}_S\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$:

$$E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)^T\right] = E_{\boldsymbol{\theta}_r;\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}\right)\right], \quad (41)$$

one can derive LBs on the MSE of NSMLEs as follows. Firstly, the rationale outlined in Sections II-B and III-A is generalizable to vector parameter [17][23], that is any LB on $\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}\right)$, $\widehat{\boldsymbol{\phi}} \in \mathcal{U}_S\left(\mathcal{S}_{\mathcal{X}}\right)$, can be expressed as linear transformations of the ad hoc form of the MSB, that is in the present case, w.r.t. to $p\left(\mathbf{x}|\boldsymbol{\phi}\right)$ and for strict-sense unbiased estimates (37) satisfying:

$$E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)\boldsymbol{\upsilon}_{\boldsymbol{\phi}}^T\left(\boldsymbol{\Phi}^N\right)\right] = \boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right), \quad (42a)$$

[7]Regarding the deterministic parameter θ, uniform wide-sense unbiasedeness is another name for unbiasedeness in the Barankin sense (4a).

[8]In most cases, the inclusion is strict leading to strict inequalities (40a-45a)

where $\boldsymbol{\Phi}^N = \left[\begin{array}{ccc} \boldsymbol{\phi}^1 & \ldots & \boldsymbol{\phi}^N \end{array} \right]$, $\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right) = \left[\begin{array}{ccc} \boldsymbol{\phi}^1 - \boldsymbol{\phi} & \ldots & \boldsymbol{\phi}^N - \boldsymbol{\phi} \end{array} \right]$, $\boldsymbol{\upsilon}_{\boldsymbol{\phi}}\left(\boldsymbol{\Phi}^N\right) \triangleq \boldsymbol{\upsilon}_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\Phi}^N\right) = \left(\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}^1\right),\ldots,\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}^N\right)\right)^T$ and $\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}'\right) = p\left(\mathbf{x}|\boldsymbol{\phi}'\right)/p\left(\mathbf{x}|\boldsymbol{\phi}\right)$. By resorting to the generalization of (6) to a vector of estimators [23, Lemma 1], the solution of:

$$\min_{\widehat{\boldsymbol{\phi}}\in\mathcal{U}_S(\mathcal{S}_{\mathcal{X}})} \left\{\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}\right)\right\} \text{ under } E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\boldsymbol{\phi}}-\boldsymbol{\phi}\right)\boldsymbol{\upsilon}_{\boldsymbol{\phi}}^T\left(\boldsymbol{\Phi}^N\right)\right] = \boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right), \tag{42b}$$

is given by:

$$\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{MSB}\right) = \boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^N\right)\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)^T, \quad \widehat{\boldsymbol{\phi}}_{MSB} = \boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^N\right)\boldsymbol{\upsilon}_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\Phi}^N\right), \tag{42c}$$

where $\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}\left(\boldsymbol{\Phi}^N\right) = E_{\mathbf{x}|\boldsymbol{\phi}}\left[\boldsymbol{\upsilon}_{\boldsymbol{\phi}}\left(\boldsymbol{\Phi}^N\right)\boldsymbol{\upsilon}_{\boldsymbol{\phi}}^T\left(\boldsymbol{\Phi}^N\right)\right]$. Therefore:

$$\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{MSB}\right) \leq \min_{\widehat{\boldsymbol{\phi}}\in\mathcal{U}_S(\mathcal{S}_{\mathcal{X}})} \left\{\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}\right)\right\}, \tag{43a}$$

leading to (41):

$$E_{\boldsymbol{\theta}_r;\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{MSB}\right)\right] \leq \min_{\widehat{\boldsymbol{\phi}}\in\mathcal{U}_S(\mathcal{S}_{\mathcal{X}})} \left\{E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\boldsymbol{\phi}}-\boldsymbol{\phi}\right)\left(\widehat{\boldsymbol{\phi}}-\boldsymbol{\phi}\right)^T\right]\right\}. \tag{43b}$$

In any asymptotic region of operation of NSMLEs, since NSMLEs belong to $\mathcal{U}_S\left(\mathcal{S}_{\mathcal{X}}\right)$, then $E_{\boldsymbol{\theta}_r;\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{MSB}\right)\right]$ is a LB on the covariance matrix of NSMLEs. Therefore it seems sensible to refer to $E_{\boldsymbol{\theta}_r;\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{MSB}\right)\right]$ as a non-standard MSB (NSMSB) to make the difference with the modified MSB (14). Indeed, the NSMSB is a LB for $\widehat{\boldsymbol{\phi}} \in \mathcal{U}_S\left(\mathcal{S}_{\mathcal{X}}\right)$, i.e., strict-sense unbiased estimates, whereas the MMSB is a LB for $\widehat{\boldsymbol{\phi}} \in \mathcal{U}_W\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$, i.e., wide-sense unbiased estimates (see Section III-A). In the same vein, any Barankin bound approximation (BBA) on the MSE of MLEs resulting from a linear transformation of the MSB (7b), has a non-standard version, referred to as NSBBA or as NSLB hereinafter, and defined as

$$\mathbf{NSBBA} = E_{\boldsymbol{\theta}_r;\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{BBA}\right)\right], \tag{44}$$

where $\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{BBA}\right)$ is the LB resulting from the same linear transformation of (42c). As well as the NSMSB, any NSBBA lower bounds the MSE of NSMLEs in any asymptotic region of operation. Note that in general, the NSLBs cannot be arranged in closed form due to the presence of the statistical expectation. They however can be evaluated by numerical integration or Monte Carlo simulation [33]. It is then worth noticing that an equivalent form of (40a) is:

$$\min_{\widehat{\boldsymbol{\phi}}\in\mathcal{U}_W(\mathcal{S}_{\mathcal{X}})} \left\{E_{\mathbf{x}|\theta}\left[\left(\widehat{\boldsymbol{\phi}}-\mathbf{g}\left(\theta\right)\right)\left(\widehat{\boldsymbol{\phi}}-\mathbf{g}\left(\theta\right)\right)^T\right]\right\} \leq \min_{\widehat{\boldsymbol{\phi}}\in\mathcal{U}_S(\mathcal{S}_{\mathcal{X}})} \left\{E_{\boldsymbol{\theta}_r;\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}\right)\right]\right\} + \mathbf{C}_{\theta}\left(\boldsymbol{\phi}\right). \tag{45a}$$

Unfortunately, $\widehat{\boldsymbol{\phi}}_{BBA} \notin \mathcal{U}_S\left(\mathcal{S}_{\mathcal{X}}\right)$ and $\widehat{\boldsymbol{\phi}}_{BBA} \notin \mathcal{U}_W\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$ in general, therefore no general result can be drawn from (45a) on the ordering between $\mathbf{NSBBA} + \mathbf{C}_{\theta}\left(\boldsymbol{\phi}\right)$ and $\min_{\widehat{\boldsymbol{\phi}}\in\mathcal{U}_W(\mathcal{S}_{\mathcal{X}})} \left\{E_{\mathbf{x}|\theta}\left[\left(\widehat{\boldsymbol{\phi}}-\mathbf{g}\left(\theta\right)\right)\left(\widehat{\boldsymbol{\phi}}-\mathbf{g}\left(\theta\right)\right)^T\right]\right\}$, or any BBA computed on $\mathcal{U}_W\left(\mathcal{S}_{\mathcal{X}}\right)$.

C. Lower bounds comparison

If in general we cannot compare directly the performance of the NSMLEs and the MLEs, however since:

$$\min_{\widehat{\boldsymbol{\phi}}\in\mathcal{U}_W(\mathcal{S}_{\mathcal{X},\Theta_r})} \left\{E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\boldsymbol{\phi}}-\mathbf{g}\left(\theta\right)\right)\left(\widehat{\boldsymbol{\phi}}-\mathbf{g}\left(\theta\right)\right)^T\right]\right\} \leq \min_{\widehat{\boldsymbol{\phi}}\in\mathcal{U}_S(\mathcal{S}_{\mathcal{X}})} \left\{E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\boldsymbol{\phi}}-\boldsymbol{\phi}\right)\left(\widehat{\boldsymbol{\phi}}-\boldsymbol{\phi}\right)^T\right]\right\} + \mathbf{C}_{\theta}\left(\boldsymbol{\phi}\right),$$

we should be able to compare their associated LBs. Actually, some comparisons are possible but for a restricted class of non-standard estimation problems, as shown in the following. Using the rationale outlined in Section III-A, one can state that all MLBs on $\widehat{\boldsymbol{\phi}} \in \mathcal{U}_W\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$ derive from sets of discrete or integral linear transform of:

$$\forall n \in [1,N],\ \mathbf{g}\left(\theta^n\right) - \mathbf{g}\left(\theta\right) = E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\boldsymbol{\phi}}\left(\mathbf{x},\boldsymbol{\theta}_r\right) - \mathbf{g}\left(\theta\right)\right)\upsilon_{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^n\right)\right], \tag{46}$$

which can be rewritten as:

$$\forall n \in [1,N]\,,\ \mathbf{g}\left(\theta^n\right)-\mathbf{g}\left(\theta\right)=E_{\boldsymbol{\theta}_r;\theta^n}\left[\boldsymbol{\phi}\right]-\mathbf{g}\left(\theta\right)+E_{\boldsymbol{\theta}_r;\theta}\left[E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\boldsymbol{\phi}}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\boldsymbol{\phi}\right)\upsilon_\theta\left(\mathbf{x},\boldsymbol{\theta}_r;\theta^n\right)\right]\right],$$

and yields the general form of the MMSB for unbiased estimates of functions of θ:

$$\mathbf{MMSB}=\boldsymbol{\Xi}\left(\boldsymbol{\theta}^N\right)\mathbf{R}_{\boldsymbol{\upsilon}_\theta}^{-1}\left(\boldsymbol{\theta}^N\right)\boldsymbol{\Xi}\left(\boldsymbol{\theta}^N\right)^T, \tag{47}$$

where $\boldsymbol{\Xi}\left(\boldsymbol{\theta}^N\right)=\left[\ \mathbf{g}\left(\theta^1\right)-\mathbf{g}\left(\theta\right)\quad\ldots\quad\mathbf{g}\left(\theta^N\right)-\mathbf{g}\left(\theta\right)\ \right]$ and $\mathbf{R}_{\boldsymbol{\upsilon}_\theta}\left(\boldsymbol{\theta}^N\right)=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\boldsymbol{\upsilon}_\theta\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}^N\right)\boldsymbol{\upsilon}_\theta^T\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}^N\right)\right]$. If $p\left(\boldsymbol{\theta}_r;\theta\right)$ does not depend on θ, i.e. $p\left(\boldsymbol{\theta}_r;\theta\right)=p\left(\boldsymbol{\theta}_r\right)$, then:

$$E_{\boldsymbol{\theta}_r;\theta'}\left[\boldsymbol{\phi}\right]=\mathbf{g}\left(\theta\right),\ \upsilon_\theta\left(\mathbf{x},\boldsymbol{\theta}_r;\theta'\right)=\frac{p\left(\mathbf{x}|\boldsymbol{\theta}_r;\theta'\right)}{p\left(\mathbf{x}|\boldsymbol{\theta}_r;\theta\right)},$$

and (46) becomes:

$$\forall n \in [1,N]\,,\ \mathbf{g}\left(\theta^n\right)-\mathbf{g}\left(\theta\right)=E_{\boldsymbol{\theta}_r;\theta}\left[E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\boldsymbol{\phi}}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\boldsymbol{\phi}\right)\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}^n\right)\right]\right], \tag{48}$$

where $\boldsymbol{\phi}^n=\left(\theta^n,\boldsymbol{\theta}_r^T\right)^T$, that is any $\widehat{\boldsymbol{\phi}}\in\mathcal{L}_2\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$ satisfying (42a) also satisfies (46). Therefore, if $p\left(\boldsymbol{\theta}_r;\theta\right)$ does not depend on θ, $\widehat{\boldsymbol{\phi}}_{MSB}$ in (42c) satisfies (46) and the following inequality:

$$\underbrace{\boldsymbol{\Xi}\left(\boldsymbol{\theta}^N\right)\mathbf{R}_{\boldsymbol{\upsilon}_\theta}^{-1}\left(\boldsymbol{\theta}^N\right)\boldsymbol{\Xi}\left(\boldsymbol{\theta}^N\right)^T}_{\mathbf{MMSB}}\ \leq\ \underbrace{E_{\boldsymbol{\theta}_r;\theta}\left[\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^N\right)\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)^T\right]}_{\mathbf{NSMSB}}+\mathbf{C}_\theta\left(\boldsymbol{\phi}\right), \tag{49a}$$

where $\boldsymbol{\Phi}^N=\begin{bmatrix}\left(\boldsymbol{\theta}^N\right)^T\\ \boldsymbol{\theta}_r\mathbf{1}_N^T\end{bmatrix}$, $\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)=\boldsymbol{\Xi}\left(\boldsymbol{\theta}^N\right)=\begin{bmatrix}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T\\ \mathbf{0}\end{bmatrix}$ and $\boldsymbol{\upsilon}_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\Phi}^N\right)=\boldsymbol{\upsilon}_\theta\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}^N\right)$. In particular, regarding the estimation of θ, since $\theta=\left(1,\mathbf{0}^T\right)\boldsymbol{\phi}$, one obtains:

$$\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T\mathbf{R}_{\boldsymbol{\upsilon}_\theta}^{-1}\left(\boldsymbol{\theta}^N\right)\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)\leq E_{\boldsymbol{\theta}_r;\theta}\left[\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^N\right)\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)\right]. \tag{49b}$$

Interestingly enough, it is straightforward to extend (49a-49b) by introducing tighter NSLBs. It suffices to note that the addition of any subset of K constraints:

$$\forall k\in[N+1,N+K]\,,\boldsymbol{\phi}^k-\boldsymbol{\phi}=E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\boldsymbol{\phi}}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\boldsymbol{\phi}\right)\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}^k\right)\right],\boldsymbol{\phi}^k=\left(\theta^k,\left(\boldsymbol{\theta}_r^k\right)^T\right)^T,$$

to (48) restricts the class of viable estimators $\widehat{\boldsymbol{\phi}}\in\mathcal{U}_W\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$ and therefore increases the associated NSBBA (42c), leading to:

$$\begin{aligned}\boldsymbol{\Xi}\left(\boldsymbol{\theta}^N\right)\mathbf{R}_{\boldsymbol{\upsilon}_\theta}^{-1}\left(\boldsymbol{\theta}^N\right)\boldsymbol{\Xi}\left(\boldsymbol{\theta}^N\right)^T&\leq E_{\boldsymbol{\theta}_r;\theta}\left[\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^N\right)\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)^T\right]+\mathbf{C}_\theta\left(\boldsymbol{\phi}\right)\\&\leq E_{\boldsymbol{\theta}_r;\theta}\left[\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^{N+K}\right)\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^{N+K}\right)\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^{N+K}\right)^T\right]+\mathbf{C}_\theta\left(\boldsymbol{\phi}\right),\end{aligned} \tag{50a}$$

and, regarding the estimation of θ, to:

$$\begin{aligned}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T\mathbf{R}_{\boldsymbol{\upsilon}_\theta}^{-1}\left(\boldsymbol{\theta}^N\right)\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)&\leq E_{\boldsymbol{\theta}_r;\theta}\left[\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^N\right)\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)\right]\\&\leq E_{\boldsymbol{\theta}_r;\theta}\left[\boldsymbol{\xi}\left(\boldsymbol{\theta}^{N+K}\right)^T\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^{N+K}\right)\boldsymbol{\xi}\left(\boldsymbol{\theta}^{N+K}\right)\right],\end{aligned} \tag{50b}$$

where $\boldsymbol{\Phi}^{N+K}=\left[\boldsymbol{\Phi}^N,\begin{bmatrix}\theta^{N+1}&\ldots&\theta^{N+K}\\ \boldsymbol{\theta}_r^{N+1}&\ldots&\boldsymbol{\theta}_r^{N+K}\end{bmatrix}\right]$, $\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^{N+K}\right)=\left[\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right),\begin{bmatrix}\xi\left(\theta^{N+1}\right)&\ldots&\xi\left(\theta^{N+K}\right)\\ \boldsymbol{\theta}_r^{N+1}-\boldsymbol{\theta}_r&\ldots&\boldsymbol{\theta}_r^{N+K}-\boldsymbol{\theta}_r\end{bmatrix}\right]$, and $\boldsymbol{\xi}\left(\boldsymbol{\theta}^{N+K}\right)=\left(\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T,\xi\left(\theta^{N+1}\right),\ldots,\xi\left(\theta^{N+K}\right)\right)^T$. Then one can take advantage of the use of the numerous (standard) BBAs derived for parameter vector [23][24], however, at a cost of numerical integration or Monte Carlo simulation to evaluate their statistical expectation.

1) Old and new non-standard lower bounds:

A typical example is the NSCRB obtained for $N=2$, where $\boldsymbol{\theta}^2=(\theta,\theta+d\theta)$ leading to the following subset of constraints:

$$\begin{pmatrix}0\\ d\theta\end{pmatrix}=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\begin{pmatrix}1_{\mathcal{S}_{\mathcal{X},\Theta_r}}\left(\mathbf{x},\boldsymbol{\theta}_r\right)\\ \upsilon_\theta\left(\mathbf{x},\boldsymbol{\theta}_r;\theta+d\theta\right)\end{pmatrix}\right], \tag{51a}$$

which is equivalent to [55, Lemma 3]:

$$\begin{pmatrix}0\\ 1\end{pmatrix}=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\begin{pmatrix}1_{\mathcal{S}_{\mathcal{X},\Theta_r}}\left(\mathbf{x},\boldsymbol{\theta}_r\right)\\ \frac{\upsilon_\theta(\mathbf{x},\boldsymbol{\theta}_r;\theta+d\theta)-1}{d\theta}\end{pmatrix}\right], \tag{51b}$$

and can be reduced to [55, Lemma 2]:

$$1=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\frac{p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta+d\theta\right)-p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)}{d\theta p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)}\right], \tag{51c}$$

since $E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[1_{\mathcal{S}_{\mathcal{X},\Theta_r}}\left(\mathbf{x},\boldsymbol{\theta}_r\right)\left(\upsilon_\theta\left(\mathbf{x},\boldsymbol{\theta}_r;\theta+d\theta\right)-1\right)\right]=0$. Then by letting $d\theta$ be infinitesimally small, (49b) becomes [4, (5)]:

$$MCRB_\theta\triangleq E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\frac{\partial\ln p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right)}{\partial\theta}\right)^2\right]^{-1}\leq NSCRB_\theta\triangleq E_{\boldsymbol{\theta}_r;\theta}\left[E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\frac{\partial\ln p\left(\mathbf{x}|\boldsymbol{\phi}\right)}{\partial\theta}\right)^2\right]^{-1}\right], \tag{52}$$

where the $NSCRB_\theta$ is the MCB [2, (7)]. Following the rationale introduced in [14], a straightforward extension of (52) is obtained for $\boldsymbol{\theta}^N=\left(\theta^1,\ldots,\theta^N\right)^T$, $\theta^n=\theta+(n-1)\,d\theta$, $1\leq n\leq N$. Indeed the set of N associated constraints:

$$d\theta\mathbf{w}_N=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\boldsymbol{\upsilon}_\theta\left(\mathbf{x},\boldsymbol{\theta}_r;\boldsymbol{\theta}^N\right)\right], \tag{53a}$$

where $\mathbf{w}_N^T=(0,\ldots,N-1)$, by letting $d\theta$ be infinitesimally small, becomes equivalent to [14][55, Lemma 3]:

$$\mathbf{v}'=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\mathbf{b}'_\theta\left(\mathbf{x},\boldsymbol{\theta}_r\right)\right], \tag{53b}$$

where $\mathbf{v}'=(0,1,0,\ldots,0)^T$ and $\mathbf{b}'_\theta\left(\mathbf{x},\boldsymbol{\theta}_r\right)=\frac{1}{p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}\left(p\left(\mathbf{x},\boldsymbol{\theta}_r;\theta\right),\frac{\partial p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{\partial\theta},\ldots,\frac{\partial^{N-1}p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{\partial^{N-1}\theta}\right)^T$. Since $v'_1=0$ and $E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\mathbf{b}'_\theta\right)_1\left(\mathbf{x},\boldsymbol{\theta}_r\right)\left(\mathbf{b}'_\theta\right)_n\left(\mathbf{x},\boldsymbol{\theta}_r\right)\right]=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\frac{\partial^n p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{\partial^n\theta}\right]=0$, $2\leq n\leq N-1$, (53b) is actually equivalent to [55, Lemma 2]:

$$\mathbf{e}_1=E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\mathbf{b}_\theta\left(\mathbf{x},\boldsymbol{\theta}_r\right)\right], \tag{53c}$$

where $\mathbf{b}_\theta\left(\mathbf{x},\boldsymbol{\theta}_r\right)=\frac{1}{p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}\left(\frac{\partial p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{\partial\theta},\ldots,\frac{\partial^{N-1}p(\mathbf{x},\boldsymbol{\theta}_r;\theta)}{\partial^{N-1}\theta}\right)^T$, and (49b) becomes an inequality between the Battacharayya bounds (BaBs) [10] of order $N-1$:

$$MBaB_\theta\triangleq\mathbf{e}_1^TE_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\mathbf{b}_\theta\left(\mathbf{x},\boldsymbol{\theta}_r\right)\mathbf{b}_\theta^T\left(\mathbf{x},\boldsymbol{\theta}_r\right)\right]^{-1}\mathbf{e}_1\leq NSBaB_\theta\triangleq E_{\boldsymbol{\theta}_r;\theta}\left[\mathbf{e}_1^TE_{\mathbf{x}|\boldsymbol{\phi}}\left[\boldsymbol{\beta}\left(\mathbf{x};\boldsymbol{\phi}\right)\boldsymbol{\beta}^T\left(\mathbf{x};\boldsymbol{\phi}\right)\right]^{-1}\mathbf{e}_1\right], \tag{53d}$$

where $\boldsymbol{\beta}\left(\mathbf{x};\boldsymbol{\phi}\right)=\frac{1}{p(\mathbf{x}|\boldsymbol{\phi})}\left(\frac{\partial p(\mathbf{x}|\boldsymbol{\phi})}{\partial\theta},\ldots,\frac{\partial^{N-1}p(\mathbf{x}|\boldsymbol{\phi})}{\partial^{N-1}\theta}\right)^T=\mathbf{b}_\theta\left(\mathbf{x},\boldsymbol{\theta}_r\right)$. Therefore, with the proposed approach, we not only extend the result introduced in [34, (11)] under the restrictive assumption of a prior independent of θ, but we can also assert that $MBaB_\theta\leq NSBaB_\theta$ if the prior does not depend on θ, which has not been proven in [34].
As with the CRB and the BaB, (49b) also allows to derive inequalities between modified and non-standard forms of all remaining BB approximations released in the open literature, namely the FGB [14], the MHB [19], the GlB [20], the AbB [22], and the CRFB [24, (101-102)].
Furthermore, an example of a tighter NSLB can be easily derived from the usual NSCRB (52). Indeed by adding to (51a) the following $K=P_r$ constraints:

$$\mathbf{0}=E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\Phi}^K\right)\right],$$

where $\boldsymbol{\phi}^k = \begin{pmatrix} \theta \\ \boldsymbol{\theta}_r + \mathbf{u}_k h_r^k \end{pmatrix}$ and $\mathbf{u}_k$ is the kth column of the identity matrix $\mathbf{I}_{P_r}$, one obtains the following equivalent set of constraints [55, Lemma 3+Lemma 2]:

$$\mathbf{e}_1 = E_{\mathbf{x}|\boldsymbol{\phi}} \left[\left(\widehat{\theta}\left(\mathbf{x}, \boldsymbol{\theta}_r\right) - \theta \right) \mathbf{c}\left(\mathbf{x}; \boldsymbol{\Phi}^{K+1}\right) \right],$$
$$\mathbf{c}^T\left(\mathbf{x}; \boldsymbol{\Phi}^{K+1}\right) = \left(\tfrac{p(\mathbf{x}|\boldsymbol{\theta}_r;\theta+d\theta)}{p(\mathbf{x}|\boldsymbol{\theta}_r;\theta)d\theta} - \tfrac{1}{d\theta}, \tfrac{p(\mathbf{x}|\boldsymbol{\theta}_r+\mathbf{u}_1 h_r^1,\theta)}{p(\mathbf{x}|\boldsymbol{\theta}_r;\theta)h_r^1} - \tfrac{1}{h_r^1}, \ldots, \tfrac{p(\mathbf{x}|\boldsymbol{\theta}_r+\mathbf{u}_K h_r^K,\theta)}{p(\mathbf{x}|\boldsymbol{\theta}_r;\theta)h_r^K} - \tfrac{1}{h_r^K} \right).$$

By letting $\left(d\theta, h_r^1, \ldots, h_r^{P_r}\right)$ be infinitesimally small, then $\mathbf{c}\left(\mathbf{x}; \boldsymbol{\Phi}^{K+1}\right) \to \frac{\partial \ln p(\mathbf{x}|\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}$ and (50b) becomes [33, (24)]:

$$MCRB_\theta \triangleq E_{\mathbf{x},\boldsymbol{\theta}_r;\theta}\left[\left(\frac{\partial \ln p\left(\mathbf{x}, \boldsymbol{\theta}_r; \theta\right)}{\partial \theta} \right)^2 \right]^{-1} \leq NSCRB_\theta \triangleq E_{\boldsymbol{\theta}_r;\theta}\left[E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\frac{\partial \ln p\left(\mathbf{x}|\boldsymbol{\phi}\right)}{\partial \theta} \right)^2 \right]^{-1} \right]$$
$$\leq \overline{NSCRB}_\theta = E_{\boldsymbol{\theta}_r;\theta}\left[\mathbf{e}_1^T E_{\mathbf{x}|\boldsymbol{\phi}}\left[\frac{\partial \ln p\left(\mathbf{x}|\boldsymbol{\phi}\right)}{\partial \boldsymbol{\phi}} \frac{\partial \ln p\left(\mathbf{x}|\boldsymbol{\phi}\right)}{\partial \boldsymbol{\phi}^T} \right]^{-1} \mathbf{e}_1 \right]. \quad (54)$$

In [33] $\overline{NSCRB}_\theta$ was introduced under the restrictive assumption of a prior independent of θ, which can be relaxed as shown with the proposed framework. Furthermore, the above example illustrate that the tightest form of any NSLB is obtained when the set of unbiasedness constraints are expressed for $\boldsymbol{\phi}$ as in (50b) and not only for θ as in (49b).

D. Non-standard lower bounds (continued)

Any of the NSLBs mentioned in the previous section can be derived in the general case where $p\left(\boldsymbol{\theta}_r; \theta\right)$ does depend on θ, except that no general inequalities between MBBA and NSBBA can any longer be exhibited. Interestingly enough, since any existing standard LB can be obtained from (42c) as $\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{BBA}\right)$ (a multiple parameters version of (8b)), it has a non-standard counterpart $E_{\boldsymbol{\theta}_r;\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{BBA}\right)\right]$, which includes the FGB [14], the MHB [19], the GlB [20], the AbB [22], and the CRFB [24, (101-102)]. Last, let us recall that in general $\widehat{\boldsymbol{\phi}}_{BBA} \in \mathcal{U}_S\left(\mathcal{S}_{\mathcal{X},\Theta_r}\right)$, therefore the associated NSLB cannot be compared a priori neither with the MSE of $\widehat{\theta}_{ML} \in \mathcal{U}_W\left(\mathcal{S}_{\mathcal{X}}\right)$ nor with any of its LBs (computed with $p\left(\mathbf{x}; \theta\right)$). In particular, NSLBs are not in general neither upper bounds on the MSE of $\widehat{\theta}_{ML}$ nor on any of its LBs.

V. Application examples

A. Application examples in the open literature

Since the preliminary works [2], [3] and [4], non-standard estimation has given rise to a growing interest [26]-[42] as compound probability distributions arise in many applications such as telecom or radar and hybrid parameter vectors occur in miscellaneous estimation problem.

1) Application examples in telecom:

The MCRB has been often used in linearly modulated signals for synchronization problems involving the estimation of carrier-frequency offset, carrier phase, and timing epoch [4][26][27] where the calculation of the CRB is infeasible, while application of the MCRB leads to useful expressions with moderate analytical effort. In [30] the MCRB is computed for linearly modulated signals corrupted by correlated impulsive noise, wherein data symbols and noise power are regarded as nuisance parameters. In [32] linearly modulated signals are observed on a frequency-flat time-selective fading channel affected by additive white Gaussian noise. High signal-to-noise ratio asymptote of the CRB and MCRBs are derived and compared, for the joint estimation of all those channel parameters that impact signal detection, namely, carrier phase, carrier frequency offset, frequency rate of change, signal amplitude, fading power, and Gaussian noise power. [33] considers the problem of estimating the carrier frequency offset affecting a linearly modulated waveform received through a Rician fading channel. The relevant CRB, HCRB, MCRB and NSCRBs are calculated and the relative merits of these bounds are discussed, both in terms of their tightness and ease of calculation. [36] addresses the CRB and the MCRB for the joint estimation of the carrier frequency offset, the carrier phase and the noise and signal powers of binary phase-shift keying (BPSK), minimum shift keying (MSK), and quaternary phase-shift keying (QPSK) modulated signals corrupted by additive

white circular Gaussian noise. [40] focuses on the MCRBs related to estimating the parameters of a noisy hybrid modulated signal which combines the pseudo-random binary codes (PRBC) phase modulation and linear frequency modulation (LFM). The parameters to be estimated of the PRBC-LFM signal include signal amplitude, carrier phase, carrier frequency, chirp rate and symbol width.

2) Application examples in radar:

MCRBs and NSLBs have also been investigated in radar applications. [3] derives HCRB on source location accuracies achievable with far-field sources whose bearings are random and impinging on a two-dimensional array of sensors whose locations are not known precisely. [2] derived the NSCRB for estimators of the arrival time of a radar echo of random amplitude, unbiased for all values of the signal amplitude. In [29] a similar problem is studied for real valued radar echo via the HBB which deterministic part is shown to handle the MSE threshold phenomena. [31] addresses the derivation of the HCRB for the parameters of a signal composed of a mixture of spherically invariant random processes. The case of signal composed of a mixture of K-distributed clutter, Gaussian clutter, and thermal noise belongs to this set, and it is regarded as a realistic radar scenario. In [37][38] the problem of registration errors involved in the grid-locking problem for successful multisensor target tracking is addressed, i.e., attitude, measurement, and position biases. Linear least squares and non-linear expectation-maximisation estimators of these bias terms are derived and their statistical performances compared to the HCRB as a function of sensor locations, sensors number, and accuracy of sensor measurements. [39] computes the MCRB for the target parameter (delay, Doppler) estimation error using universal mobile telecommunications system (UMTS) signals as illuminators of opportunity for passive multistatic radar systems. In [41] the MCRB on the joint estimation of target location and velocity for a non-coherent passive MIMO radar, employing the OFDMbased L-band digital aeronautical communication system type 1 signals as signals of opportunity, is investigated. [42] considers the problem of estimating the parameters of a low-rank compound-Gaussian process in white Gaussian noise. This situation typically arises in radar applications where clutter is relevantly modeled as compound-Gaussian with a rankdeficient covariance matrix of the speckle. First, assuming the textures are deterministic, the CRB of the parameters describing the covariance matrix is derived, which enables one to assess the impact of the time-varying textures on the estimation performance. Then, considering the textures as random, HCRB, MCB, MCRB are derived and compared.

B. A new look at Gaussian observation models

In the framework of modern array processing [47], the observation vector $\mathbf{x}$ generally consists of a bandpass signal which is the output of an Hilbert filtering [1][69, §13], i.e., a complex circular vector $\mathbf{x} \sim C\mathcal{N}\left(\mathbf{m}_{\mathbf{x}}, \mathbf{C}_{\mathbf{x}}\right)$ with p.d.f.:

$$p_{\mathcal{CN}}\left(\mathbf{x}; \mathbf{m}_{\mathbf{x}}, \mathbf{C}_{\mathbf{x}}\right) = \pi^{-M} \left|\mathbf{C}_{\mathbf{x}}^{-1}\right| e^{-(\mathbf{x}-\mathbf{m}_{\mathbf{x}})^{H} \mathbf{C}_{\mathbf{x}}^{-1} (\mathbf{x}-\mathbf{m}_{\mathbf{x}})}.$$

Mostly two different signal models are considered: the deterministic (conditional) signal model and the stochastic (unconditional) signal model [70]. The discussed signal models are Gaussian and the parameters are connected either with the expectation value in the deterministic case or with the covariance matrix in the stochastic one. A simple and well known instantiation is the observation model formed from L independent snapshots of a linear superposition of an individual signal of interest and noise:

$$\mathbf{x}_l = \mathbf{s}\left(\tau\right) a_l + \mathbf{n}_l, 1 \leq l \leq L, \tag{55}$$

where $a_1, \ldots, a_L$ are the complex amplitudes of the signal, $\mathbf{s}\left(\ \right)$ is a vector of M parametric functions depending on a single deterministic parameter τ, $\mathbf{n}_l \sim \mathcal{CN}\left(\mathbf{0}, \sigma_{\mathbf{n}}^2 \mathbf{I}_M\right)$, $1 \leq l \leq L$, are independent and identically distributed (i.i.d.) Gaussian complex circular noises independent of the signal of interest. Additionally if the L components of $\mathbf{a} = \left(a_1, \ldots, a_L\right)^T$ are i.i.d. zero mean Gaussian complex circular random variables with variance σ_a^2, i.e. $\mathbf{a} \sim C\mathcal{N}\left(\mathbf{0}, \sigma_a^2 \mathbf{I}_L\right)$, then (55) is an unconditional signal model characterized by a p.d.f. $p\left(\mathbf{x}|\boldsymbol{\theta}\right)$, $\boldsymbol{\theta} = \left(\tau, \sigma_a^2, \sigma_{\mathbf{n}}^2\right)^T$,

given by [70]:

$$p\left(\mathbf{X}|\boldsymbol{\theta}\right)=\int p\left(\mathbf{X}|\mathbf{a},\tau,\sigma_{\mathbf{n}}^{2}\right)p\left(\mathbf{a}|\sigma_{a}^{2}\right)d\mathbf{a}=\frac{e^{-Ltr\left(\mathbf{C}_{\mathbf{x}}^{-1}\widehat{\mathbf{R}}_{\mathbf{x}}\right)}}{\left(\pi^{M}|\mathbf{C}_{\mathbf{x}}|\right)^{L}},$$

$$\mathbf{C}_{\mathbf{x}}=\sigma_{a}^{2}\mathbf{s}\left(\tau\right)\mathbf{s}\left(\tau\right)^{H}+\sigma_{\mathbf{n}}^{2}\mathbf{I}_{M},\ \widehat{\mathbf{R}}_{\mathbf{x}}=\frac{1}{L}\sum_{l=1}^{L}\mathbf{x}_{l}\mathbf{x}_{l}^{H},$$

$$p\left(\mathbf{X}|\mathbf{a},\tau,\sigma_{\mathbf{n}}^{2}\right)=\frac{e^{-\frac{L}{\sigma_{\mathbf{n}}^{2}}tr\left(\widehat{\mathbf{C}}_{\mathbf{n}}\right)}}{\left(\pi\sigma_{\mathbf{n}}^{2}\right)^{ML}},\ p\left(\mathbf{a}|\sigma_{a}^{2}\right)=e^{-\frac{\mathbf{a}^{H}\mathbf{a}}{\sigma_{a}^{2}}}\left(\pi\sigma_{a}^{2}\right)^{-L},$$

$$\widehat{\mathbf{C}}_{\mathbf{n}}=\frac{1}{L}\sum_{l=1}^{L}\left(\mathbf{x}_{l}-\mathbf{s}\left(\tau\right)a_{l}\right)\left(\mathbf{x}_{l}-\mathbf{s}\left(\tau\right)a_{l}\right)^{H}.$$

Then the MLE (34a) of τ, aka the unconditional MLE (UMLE), is obtained by minimization of the concentrated criterion [64]:

$$\widehat{\tau}=\arg\min_{\tau}\left\{\left|\widehat{\sigma_{a}^{2}}\mathbf{s}\left(\tau\right)\mathbf{s}^{H}\left(\tau\right)+\widehat{\sigma_{\mathbf{n}}^{2}}\mathbf{I}_{M}\right|\right\},\ \widehat{\sigma_{a}^{2}}=\frac{1}{\left\|\mathbf{s}\left(\tau\right)\right\|^{2}}\mathbf{s}^{H}\left(\tau\right)\left(\widehat{\mathbf{R}}_{\mathbf{x}}-\widehat{\sigma_{\mathbf{n}}^{2}}\mathbf{I}_{M}\right)\mathbf{s}\left(\tau\right),$$

$$\widehat{\sigma_{\mathbf{n}}^{2}}=\frac{1}{M-1}tr\left(\Pi_{\mathbf{s}(\tau)}^{\perp}\widehat{\mathbf{R}}_{\mathbf{x}}\right),\ \Pi_{\mathbf{a}}^{\perp}=\mathbf{I}_{M}-\frac{\mathbf{a}\mathbf{a}^{H}}{\left\|\mathbf{a}\right\|^{2}},$$

and the associated CRB, aka the unconditional CRB (UCRB), is [64, (4.64)][71]:

$$UCRB_{\tau}=\frac{\sigma_{\mathbf{n}}^{2}}{2h\left(\tau\right)L\sigma_{a}^{2}\frac{SNR}{SNR+1}},\tag{56}$$

where $SNR=\sigma_{a}^{2}/\sigma_{\mathbf{n}}^{2}\left\|\mathbf{s}\left(\tau\right)\right\|^{2}$ is the signal-to-noise ratio computed at the output of the single source matched filter [64], and $h\left(\tau\right)=\frac{\partial\mathbf{s}(\tau)}{\partial\tau}^{H}\Pi_{\mathbf{s}(\tau)}^{\perp}\frac{\partial\mathbf{s}(\tau)}{\partial\tau}$. The NSMLE (34b) of τ is actually the conditional MLE (CMLE) obtained by minimization of the concentrated criterion [64] $\underline{\widehat{\tau}}=\arg\min_{\tau}\left\{tr\left(\Pi_{\mathbf{s}(\tau)}^{\perp}\widehat{\mathbf{R}}_{\mathbf{x}}\right)\right\}$ and the associated NSCRB is $NSCRB_{\tau}=E_{\mathbf{a}|\sigma_{a}^{2}}\left[CCRB_{\tau}\left(\mathbf{a}\right)\right]$, where $CCRB_{\tau}\left(\mathbf{a}\right)=\sigma_{\mathbf{n}}^{2}/\left(2h\left(\tau\right)\left\|\mathbf{a}\right\|^{2}\right)$ denotes the conditional CRB associated to the CMLE [64, (4.68)]. First, it has been shown [64, (4.74)], in the case of a vector of unknown parameters $\boldsymbol{\tau}$, that asymptotically where $L\rightarrow\infty$:

$$\mathbf{C}_{\boldsymbol{\theta}}\left(\underline{\widehat{\boldsymbol{\tau}}}\right)\geq\mathbf{C}_{\boldsymbol{\theta}}\left(\widehat{\boldsymbol{\tau}}\right)=\mathbf{UCRB}_{\boldsymbol{\tau}}\geq\mathbf{CCRB}_{\boldsymbol{\tau}},\tag{57}$$

which illustrates that the act of resorting to the NSMLE (here the CMLE) is in general an asymptotic suboptimal choice in the MSE sense. However, in the case of single unknown parameter τ, (57) becomes:

$$C_{\boldsymbol{\theta}}\left(\underline{\widehat{\tau}}\right)=C_{\boldsymbol{\theta}}\left(\widehat{\tau}\right)=UCRB_{\tau}\geq CCRB_{\tau}=\frac{\sigma_{\mathbf{n}}^{2}}{2Lh\left(\tau\right)\sigma_{a}^{2}},$$

where $C_{\boldsymbol{\theta}}\left(\underline{\widehat{\tau}}\right)=CCRB_{\tau}\left(1+CCRB_{\tau}\frac{2Lh(\tau)}{\left\|\mathbf{s}(\tau)\right\|^{2}}\right)$ [64, (4.73)]. Second, since $\left\|\mathbf{a}\right\|^{2}/\sigma_{a}^{2}\sim\chi_{2L}^{2}$, i.e. a chi-squared random variable with 2T degrees of freedom, then [69]:

$$NSCRB_{\tau}=\left|\begin{array}{cc}\frac{\sigma_{\mathbf{n}}^{2}}{2L\sigma_{a}^{2}h(\tau)}\frac{L}{L-1} & \text{if } L\geq2\\ \infty & \text{if } L=1\end{array}\right..\tag{58}$$

Therefore, if $L\geq2$:

$$\frac{NSCRB_{\tau}}{UCRB_{\tau}}=\frac{NSCRB_{\tau}}{C_{\boldsymbol{\theta}}\left(\widehat{\tau}\right)}=\frac{L}{L-1}\frac{SNR}{SNR+1},$$

which illustrates the fact that NSLB are not in general neither upper bounds on the MSE of MLEs nor on any of its LBs. Last, the MCRB (29) can be computed from [55]:

$$\mathbf{F}\left(\boldsymbol{\theta}\right)=E_{\mathbf{a}|\sigma_{a}^{2}}\left[-\frac{\partial\ln p\left(\mathbf{a}|\sigma_{a}^{2}\right)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}^{H}}\right]+E_{\mathbf{a}|\sigma_{a}^{2}}\left[\mathbf{F}_{\boldsymbol{\phi}}\left(\mathbf{a}\right)\right],\quad\mathbf{F}_{\boldsymbol{\phi}}\left(\mathbf{a}\right)=E_{\mathbf{x}|\boldsymbol{\theta}_{r};\theta}\left[-\frac{\partial\ln p\left(\mathbf{X}|\mathbf{a},\tau,\sigma_{\mathbf{n}}^{2}\right)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}^{H}}\right],$$

where $\boldsymbol{\phi} = \left(\sigma_{\mathbf{n}}^2, \sigma_a^2, \mathbf{a}, \mathbf{a}^*, \tau\right)^T$,

$$E_{\mathbf{a}|\sigma_a^2}\left[-\frac{\partial \ln p\left(\mathbf{a}|\sigma_a^2\right)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}^H}\right] = \frac{1}{\sigma_a^2}\begin{bmatrix} 0 & 0 & \mathbf{0}^T & \mathbf{0}^T & 0 \\ 0 & \frac{L}{\sigma_a^2} & \mathbf{0}^T & \mathbf{0}^T & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_L & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_L & \mathbf{0} \\ 0 & 0 & \mathbf{0}^T & \mathbf{0}^T & 0 \end{bmatrix},$$

and [72]:

$$\mathbf{F}_{\boldsymbol{\phi}}\left(\mathbf{a}\right) = \begin{bmatrix} \frac{ML}{\left(\sigma_{\mathbf{n}}^2\right)^2} & 0 & \mathbf{0}^T & \mathbf{0}^T & 0 \\ 0 & 0 & \mathbf{0}^T & \mathbf{0}^T & 0 \\ \mathbf{0} & \mathbf{0} & \frac{\|\mathbf{s}(\tau)\|^2\mathbf{I}_L}{\sigma_{\mathbf{n}}^2} & \mathbf{0} & \frac{\rho_{\mathbf{s}}\mathbf{a}^*}{\sigma_{\mathbf{n}}^2} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{\|\mathbf{s}(\tau)\|^2\mathbf{I}_L}{\sigma_{\mathbf{n}}^2} & \frac{\rho_{\mathbf{s}}^*\mathbf{a}}{\sigma_{\mathbf{n}}^2} \\ 0 & 0 & \frac{\rho_{\mathbf{s}}^*\mathbf{a}^T}{\sigma_{\mathbf{n}}^2} & \frac{\rho_{\mathbf{s}}\mathbf{a}^H}{\sigma_{\mathbf{n}}^2} & \frac{2\varrho_{\mathbf{s}}\|\mathbf{a}\|^2}{\sigma_{\mathbf{n}}^2} \end{bmatrix},$$

where $\rho_{\mathbf{s}} = \frac{\partial \mathbf{s}(\tau)}{\partial \tau}^H \mathbf{s}\left(\tau\right)$ and $\varrho_{\mathbf{s}} = \left\|\frac{\partial \mathbf{s}(\tau)}{\partial \tau}\right\|^2$. Therefore $\overline{MCRB}_\tau = \sigma_{\mathbf{n}}^2/\left(2L\sigma_a^2\varrho_{\mathbf{s}}\right)$ and:

$$\frac{UCRB_\tau}{\overline{MCRB}_\tau} = \frac{\varrho_{\mathbf{s}}}{h\left(\tau\right)}\frac{SNR}{SNR+1}. \tag{59}$$

In the single tone estimation case for which $\mathbf{s}\left(\tau\right) = \left(1, e^{j2\pi\tau}, \ldots, e^{j2\pi(M-1)\tau}\right)^T$, then $\varrho_{\mathbf{s}}/h\left(\tau\right) = 4 - \frac{6}{M+1}$ and:

$$2\left(1+\frac{1}{SNR}\right) \leq \frac{UCRB_\tau}{\overline{MCRB}_\tau} \leq 4\left(1+\frac{1}{SNR}\right),$$

which can be an acceptable optimistic approximation of the $UCRB_\tau$ according to the application under consideration. An additional comparison of the relative looseness of the modified Hammersley-Chapman-Robbins bound for the unconditional signal model can be found in [29].

VI. Conclusion

In the present paper, we have addressed deterministic parameter estimation and the situation where a closed-form of the conditional p.d.f. does not exist or where a closed-form does exist but the resulting expression is intractable to derive either LBs or MLEs. We have provided a unified framework allowing to extend the previous theoretical works released on that problem [2][3][4][29][30][34][35][45][46][52]. First, in terms of intrinsic LBs by showing that any standard LB can be transformed into a modified one fitted to non-standard deterministic estimation, at the expense of tightness however. Second, in terms of relative LBs, i.e. dedicated to characterize the asymptotically suboptimal NSMLEs, by showing that any standard LB has a non-standard version lower bounding the MSE of NSMLEs. Last, for a broader perspective, let us mention that authors in [73] address compound probability distribution the other way round, and derive CR-type LBs on the MSE of estimators of random parameter subject to unknown deterministic nuisance parameters.

References

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

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

[3] 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

[4] 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

[5] R.A. Fisher, "On an absolute criterion for fitting frequency curves", Mess. of Math., 41: 155-160, 1912

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

[7] G. Darmois, "Sur les lois limites de la dispersion de certaines estimations", Rev. Int. Stat, 13: 9-15, 1945

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

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

[10] A. Bhattacharyya, "On some analogues of the amount of information and their use in statistical estimation," Sankhya Indian J. Stat, 8: 1-14, 1946

[11] E.W. Barankin, "Locally best unbiased estimates", Ann. Math. Stat., 20(4): 477-501, 1949.
[12] J.M. Hammersley, "Estimating Restricted Parameters", Journal of Roy. Stat. Soc., 12(2): 192-240, 1950
[13] D.G. Chapman, H. Robbins, "Minimum variance estimation without regularity assumptions", Ann. Math. Stat., 22: 581-586, 1951
[14] D.A.S. Fraser and I. Guttman, "Bhattacharyya bounds without regularity assumptions," Ann. Math. Stat., 23(4): 629-632, 1952
[15] J. Kiefer, "On minimum variance estimators", Ann. Math. Stat., 23(4): 627–629, 1952.
[16] W. James, and C. Stein, "Estimation with Quadratic Loss", Proc. Fourth Berkeley Symp. on Math. Statist. and Prob., 1:361-379, 1961
[17] 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
[18] W. R. Blischke, A. J. Truelove and P. B. Mundle, "On non-regular estimation.I.Variance bounds for estimators of location parameters", Journal of American Stat. Assoc., 64(327): 1056-1072, 1969.
[19] R. McAulay, E.M. Hofstetter, " Barankin Bounds on parameter estimation", IEEE Trans. on IT, 17(6): 669-676, 1971
[20] F.E. Glave, "A new look at the Barankin Lower Bound", IEEE Trans. on IT, 18(3): 349-356, 1972
[21] C. R. Blyth, "Necessary and sufficient conditions for inequalities of Cramér-Rao type," Ann. Math. Stat., 2(3): 464-473, 1974
[22] J.S. Abel, "A bound on mean-square-estimate error", IEEE Trans. on IT, 39(5): 1675-1680, 1993
[23] 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
[24] 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
[25] E. Chaumette, A. Renaux, P. Larzabal, "Lower bounds on the mean square error derived from mixture of linear and non-linear transformations of the unbiasness definition", in Proc. IEEE ICASSP, 2009
[26] M. Moeneclaey, "A simple lower bound on the linearized performance of practical symbol synchronizers," IEEE Trans. on Commun., 31(9): 1029-1032, 1983
[27] M. Moeneclaey, "A fundamental lower bound on the performance of practical joint carrier and bit synchronizers," IEEE Trans. on Commun., 32(9): 1007-1012, 1984
[28] H. Messer and Y. Adar, "New lower bounds on frequency estimation of a multitone random signal in noise," Signal Process., 18: 413-424, 1989
[29] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of a hybrid parameter vector," IEEE Trans. on IT, 43(3): 1084-1093, 1997
[30] F. Gini, R. Reggiannini, and U. Mengali, "The modified CRB in vector parameter estimation," IEEE Trans. on Commun., 46(1): 52-60, 1998
[31] F. Gini, "A radar application of a modified CRLB: Parameter estimation in non-Gaussian clutter," IEEE Trans. on SP, 46(7): 1945-1953, 1998
[32] M. Moeneclaey, "On the true and the modified Cramér-Rao bounds for the estimation of a scalar parameter in the presence of nuisance parameters," IEEE Trans. on Commun., 46(11): 1536-1544, 1998
[33] 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
[34] F. Lu and J. V. Krogmeier, "Modified Bhattacharyya bounds and their application to timing estimation", in Proc of IEEE WCNC, 2002
[35] S. Bay, B. Geller, A. Renaux, J.-P. Barbot, and J.-M. Brossier, "On the hybrid Cramér-Rao bound and its application to dynamical phase estimation," IEEE SP Letters, 15: 453-456, 2008.
[36] J. P. Delmas, "Closed-Form Expressions of the Exact Cramer–Rao Bound for Parameter Estimation of BPSK, MSK, or QPSK Waveforms", IEEE SP Letters, 15: 405-408, 2008
[37] S. Fortunati, A. Farina, F. Gini, M. S. Greco, A. Graziano, S. Giompapa, "Least Squares Estimation and Cramér-Rao Type Lower Bounds for Relative sensor Registration Process", IEEE Trans. on SP, 59(3): 1075-1087, 2011.
[38] S. Fortunati, F. Gini, A. Farina, A. Graziano, M. S. Greco, S. Giompapa, "On the application of the expectation-maximisation algorithm to the relative sensor registration problem", IET RSN, 7(2): 191-203, 2013
[39] S. Gogineni, M. Rangaswamy, B. D. Rigling, and A. Nehorai, "Cramér-Rao Bounds for UMTS-Based Passive Multistatic Radar", IEEE Trans. on SP, 62(1): 95-106, 2014
[40] W. Pei, D. Dongping, Q. Z. T. Bin, "Modified Cramer-Rao Bounds for Parameter Estimation of Hybrid Modulated Signal Combining PRBC and LFM", in Proc of IEEE ICCSE, 2014
[41] A. Filip, D. Shutin, "Cramér-Rao bounds for L-band digital aeronautical communication system type 1 based passive multiple-input multiple-output", IET RSN, 10(2): 348-358, 2016
[42] O. Besson, "Bounds for a Mixture of Low-Rank Compound-Gaussian and White Gaussian Noises", IEEE Trans. on SP, 64(21): 5723-5732, 2016
[43] J. Galy, Eric Chaumette, F. Vincent, A. Renaux and P. Larzabal, "Lower bounds for non-standard deterministic estimation", in Proc. of IEEE Workshop on SAM, 2016
[44] P. Forster, P. Larzabal, "On lower Bounds For Deterministic Parameter Estimation", in Proc. of IEEE ICASSP 2002
[45] K. Todros and J. Tabrikian, "Hybrid lower bound via compression of the sampled CLR function," in Proc. of IEEE Workshop on SSP, 2009
[46] C. Ren, J. Galy, E. Chaumette, F. Vincent, P. Larzabal, A. Renaux, "Hybrid Barankin-Weiss-Weinstein Bounds", IEEE SP Letters, 22(11): 2064-2068, 2015
[47] H.L. Van Trees, *Optimum Array Processing*, New-York, Wiley-Interscience, 2002
[48] A. Yeredor, "The Joint MAP-ML criterion and its relation to ML and to Extended Least-Squares," IEEE Trans. on SP, 48(12): 3484-3492, 2000
[49] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm", Journal of the Royal Statistical Society, 39(1): 1-38, 1977

[50] H. L. Van Trees and K. L. Bell, Eds., *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, Wiley/IEEE Press, 2007
[51] M. Stein, S. Bar, J. A. Nosseky, and J. Tabrikian, "Performance analysis for pilot-based 1-bit channel estimation with unknown quantization threshold", in Proc of IEEE ICASSP, 2016
[52] 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
[53] K. Todros and J. Tabrikian, "Uniformly Best Biased Estimators in Non-Bayesian Parameter Estimation", IEEE Trans. on IT, 57(11): 7635-7647, 2011
[54] D. G. Luenberger, *Optimization by vector space methods,* John Wiley & Sons, 1969.
[55] 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
[56] A. J.Weiss and E.Weinstein, "A general class of lower bounds in parameter estimation", IEEE Trans. on IT, 34(2): 338-342, 1988
[57] D.C. Rife, R.R. Boorstyn, "Single tone parameter estimation from discrete-time observations", IEEE Trans. on IT, 20(5): 591-598, 1974
[58] F. Athley, "Threshold region performance of Maximum Likelihood direction of arrival estimators", IEEE Trans. on SP, 53(4): 1359-1373, 2005
[59] C. D. Richmond, "Capon Algorithm Mean Squared Error Threshold SNR Prediction and Probability of Resolution", IEEE Trans. on SP, 53(8): 2748-2764, 2005
[60] J. Tabrikian, J.L. Krolik, "Barankin bounds for source localization in an uncertain ocean environnment", IEEE Trans. on SP, 47(11): 2917-2927, 1999
[61] E. L. Lehmann and G. Casella, *Theory of Point Estimation (2nd ed.)*. Springer, 1998
[62] P. Stoica, A. Nehorai, "Performances study of conditional and unconditional direction of arrival estimation", IEEE Trans. on SP, 38(10): 1783-1795, 1990
[63] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting", IEEE Trans. on SP, 39(5): 1110-1121, 1991
[64] S. Haykin, J. Litva, and T. J. Shepherd, Eds., *Radar Array Processing,* Chapter 4, pp. 99-151, Springer-Verlag, 1993
[65] J. Li and R. T. Compton, "Maximum likelihood angle estimation for signals with known waveforms", IEEE Trans. on SP, 41(9): 2850-2862, 1993
[66] A. Renaux, P. Forster, E. Chaumette, and P. Larzabal, "On the high snr cml estimator full statistical characterization", IEEE Trans. on SP, 54(12): 4840-4843, 2006
[67] C. F. J. Wu, "On the convergence properties of the EM algorithm", Ann. Statist., 11(1): 95-103, 1983
[68] G. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, John Wiley & Sons, 1996.
[69] S.M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Prentice-Hall, 1998
[70] P. Stoica and A. Nehorai, "Performances study of conditional and unconditional direction of arrival estimation," IEEE Trans. on SP, 38(10): 1783-1795, 1990.
[71] P. Stoica, E. G. Larsson, and A. B. Gershman, "The Stochastic CRB for Array Processing: A Textbook Derivation", IEEE SP Letters, 8(5): 148-151, 2001
[72] S. F. Yau and Y. Bresler, "A Compact Cramér-Rao bound Expression for Parametric Estimation of Superimposed Signals", IEEE Trans. on SP, 40(5): 1226-1230, 1992
[73] S. Bar and J. Tabrikian, "Bayesian Estimation in the Presence of Deterministic Nuisance Parameters-Part I: Performance Bounds," IEEE Trans. on SP, 63(24): 6632-6646, 2015.

Nabil Kbayer is a PhD student at the French Institute of Aeronautics and Space (ISAE-SUPAERO), University of Toulouse, France. He received an Engineer degree in signal processing from the French Institute of Aeronautics and Space (ISAE-ENSICA). His research interests include signal processing for navigation and positioning in challenging environments, multipath and NLOS mitigation and constructive use for cognitive navigation.

Jérôme Galy was born in 1970 at Toulouse (France). He studied Electronics and Signal Processing at ENAC high school (Toulouse, France), where he obtained an Engineer degree in 1993. He received the PhD in Signal Processing from ENSICA high school (Toulouse, France) in 1998. He has been since an Assistant Professor in electronics at University of Montpellier II (France) and a researcher at the LIRMM (Laboratoire d'Informatique, de Robotique et de Micro-Electronique de Montpellier). Main domains of interest are related to signal processing and its applications such as radar or microelectronics (Functional Electrical Stimulation).

Eric Chaumette was born in 1965 at Chartres (France). He studied Electronics and Signal Processing both at ENAC (Toulouse, France) where he obtained an Engineer degree in 1989, and at Toulouse University where he obtained a M.Sc. degree in Signal Processing in 1989. From 1990 to 2007, he was with Thales in various radar studies departments. From 2007 to 2013, he was with the Electromagnetic and Radar Division of the French Aerospace Lab (ONERA), Palaiseau, France, as a research engineer. Simultaneously, from 2000 to 2014, he was a research associate at laboratory SATIE, CNRS, Ecole Normale Supérieure de Cachan, France, where he received the PhD degree in 2004 and the "habilitation à diriger les recherches" in 2014. He is now with the Department of Electronics, Optronics and Signal of ISAE (Institut Supérieur de l'Aéronautique et de l'Espace). Main domains of interest are related to radar scene modelling, detection and estimation theory and navigation.



Fig. 1. Nabil Kbayer



Fig. 2. Jérôme Galy

Francois Vincent received the Engineer degree in electronics and signal processing from ENSEEIHT, University of Toulouse, France, the PhD. Degree in signal processing from University of Toulouse and the ”habilitation à diriger des recherches” from INPT, University of Toulouse, respectively in 1995, 1999 and 2009. From 1999 to 2001, he was a research engineer with Siemens, Toulouse. He joined ISAE-Supaéro, Toulouse, France in 2001 and he is currently a Professor. His research interests include radar and navigation signal processing.

Alexandre Renaux received the Ph.D. degrees in electrical engineering from Ecole Normale Superieure de Cachan, Cachan, France, in 2006. From 2006 to 2007, he was a Postdoctoral Research Associate at the Department of Electrical and Systems Engineering , Washington University in St. Louis, St. Louis, MO, USA. Since 2007, he has been an Assistant Professor (and currently Associate Professor since 2011) at the Physics Department of University Paris-Sud, Orsay, France. His research interests are in the field of estimation theory and lower bounds on the mean square error applied to statistical signal processing.

Pascal Larzabal (M’93) was born in the Basque country in the south of France in 1962. He entered the Ecole Normale Supérieure of Cachan (France), in 1985 where he received the Agrégation in Electrical Engineering in 1988. He received the PhD in 1992 and the ”habilitation à diriger les recherches” in 1998. He is now Professor of Electrical Engineering at University of ParisSud 11, France. He teaches electronic, signal processing, control and mathematiques. From 1998 to 2003 he was the director of Electrical Engineering IUP of the University of ParisSud. From March 2003 to March 2007 he was at the head of electrical engineering department in the Institut Universitaire de Technologie of Cachan. Since January 2007 he is the director of the laboratory SATIE in Paris. His research concerns estimation in array processing and spectral analysis for wavefront identification, radars, communications, tomography and medical imaging. His recent works concern modelling error, geographical positioning and radio astronomy



Fig. 3. Eric Chaumette



Fig. 4. François Vincent



Fig. 5. Alexandre Renaux



Fig. 6. Pascal Larzabal