Estimation accuracy of non-standard maximum likelihood estimators

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

ESTIMATION ACCURACY OF NON-STANDARD MAXIMUM LIKELIHOOD ESTIMATORS

N. Kbayer[(1)], *J. Galy*[(2)], *E. Chaumette*[(3)], *F. Vincent*[(3)], *A. Renaux*[(4)] *and P. Larzabal*[(5)]

[(1)] Tésa/Isae-Supaero, Université de Toulouse, 7 Boulevard de la Gare, Toulouse, France (nabil.kbayer@isae.fr)
[(2)] Université de Montpellier 2/LIRMM, 161 rue Ada, Montpellier, France (galy@lirmm.fr)
[(3)] Isae-Supaero, Université de Toulouse, 10 av. Edouard Belin, Toulouse, France ([eric.chaumette, francois.vincent]@isae.fr)
[(4)] Université Paris-Sud/LSS, 3 Rue Joliot-Curie, Gif-sur-Yvette, France (renaux@lss.supelec.fr)
[(5)] Université Paris-Sud/SATIE, 61 av. du President Wilson, Cachan, France (pascal.larzabal@satie.ens-cachan.fr)

ABSTRACT

In many deterministic estimation problems, the probability density function (p.d.f.) parameterized by unknown deterministic parameters results from the marginalization of a joint p.d.f. depending on additional random variables. Unfortunately, this marginalization is often mathematically intractable, which prevents from using standard maximum likelihood estimators (MLEs) or any standard lower bound on their mean squared error (MSE). To circumvent this problem, the use of joint MLEs of deterministic and random parameters are proposed as being a substitute. It is shown that, regarding the deterministic parameters: 1) the joint MLEs provide generally suboptimal estimates in any asymptotic regions of operation yielding unbiased efficient estimates, 2) any representative of the two general classes of lower bounds, respectively the Small-Error bounds and the Large-Error bounds, has a "non-standard" version lower bounding the MSE of the deterministic parameters estimate.

Index Terms— Deterministic parameter estimation, maximum likelihood estimators, estimation error lower bounds

1. INTRODUCTION

As introduced in [1, p53], a model of the general deterministic estimation problem has the following four components: 1) a parameter space Θ consisting of a set of parameter vectors $\boldsymbol{\theta}$, $\Theta \subset \mathbb{R}^P$, 2) an observation space Ω consisting of a set of observation vectors $\mathbf{x}$, $\Omega \subset \mathbb{C}^M$, 3) a probabilistic mapping from Θ to Ω, that is the probability law that governs the effect of a vector parameters value $\boldsymbol{\theta}$ on the observation $\mathbf{x}$ and, 4) an estimation rule $\widehat{\boldsymbol{\theta}}(\mathbf{x})$, that is the mapping of Ω into estimates. Actually, in many estimation problems, the probabilistic mapping results from a two steps probabilistic mechanism, leading to a probability density function (p.d.f.) of the form:

$$p(\mathbf{x}|\boldsymbol{\theta}) = \int p(\mathbf{x}|\boldsymbol{\theta}_r, \boldsymbol{\theta})\, p(\boldsymbol{\theta}_r|\boldsymbol{\theta})\, d\boldsymbol{\theta}_r,\ \boldsymbol{\theta}_r \in \mathbb{R}^{P_r}, \quad (1)$$

where $\boldsymbol{\theta}_r$ is a random vector, and $p(\mathbf{x}|\boldsymbol{\theta}_r, \boldsymbol{\theta})$ and $p(\boldsymbol{\theta}_r|\boldsymbol{\theta})$ are known. Classical examples are the reception of M samples from a signal source either by a radar, a sonar or a telecom system in the presence of Gaussian or spherically invariant thermal noise [2]: $\mathbf{x} = \mathbf{s}(\boldsymbol{\theta}_\mathbf{s})\mathbf{a} + \mathbf{n}(\boldsymbol{\theta}_\mathbf{n})$, $\boldsymbol{\theta}^T = (\boldsymbol{\theta}_\mathbf{s}^T, \boldsymbol{\theta}_\mathbf{n}^T)$, $\boldsymbol{\theta}_r \triangleq \mathbf{a}$. In the case of an active radar [1][3], $\boldsymbol{\theta}_r \triangleq \mathbf{a}$ are the complex amplitudes of the received signals backscattered by P_r targets which may fluctuate according to a given (or measured) p.d.f. such as the Swerling laws [4]. Therefore, as recalled in [5], deterministic estimation problems can be divided into two subsets: the subset of "standard" estimation problems for which a closed-form expression of $p(\mathbf{x}|\boldsymbol{\theta})$ is available, and the subset of "non-standard" estimation problems for which only an integral form of $p(\mathbf{x}|\boldsymbol{\theta})$ (1) is available. Since in non-standard estimation, maximum likelihood estimators (MLEs) (2a) can be no longer derived, the use of joint MLEs of $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_r$ (2b) are proposed as being a substitute:

$$\widehat{\boldsymbol{\theta}}_{ML}(\mathbf{x}) = \arg\max_{\boldsymbol{\theta}\in\Theta} \{p(\mathbf{x}|\boldsymbol{\theta})\}, \quad (2a)$$

$$\left\{\underline{\widehat{\boldsymbol{\theta}}}(\mathbf{x}), \underline{\widehat{\boldsymbol{\theta}_r}}(\mathbf{x})\right\} = \arg\max_{\boldsymbol{\theta}\in\Theta, \boldsymbol{\theta}_r\in\Pi_{\boldsymbol{\theta}_r}} \{p(\mathbf{x}|\boldsymbol{\theta}_r, \boldsymbol{\theta})\}, \quad (2b)$$

where $\Pi_{\boldsymbol{\theta}_r}$ is the support of $p(\boldsymbol{\theta}_r|\mathbf{x}, \boldsymbol{\theta})$. It is a sensible solution in the search for a realizable estimator of $\boldsymbol{\theta}$. Indeed, the widespread use of MLEs originates from the fact that, under reasonably general conditions on the observation model [7][8], the MLEs are, in the limit of large sample support, asymptotically unbiased, Gaussian distributed and efficient. If the observation model is Gaussian, some additional asymptotic regions of operation yielding unbiased Gaussian and efficient MLEs have also been identified at finite sample support [9]-[13]. Historically, the open literature on the estimation accuracy of MLEs in terms of mean squared error (MSE), including the associated lower bounds (LBs), has remained focused on standard deterministic estimation (2a) [7][14]-[33]. It is the reason why $\underline{\widehat{\boldsymbol{\theta}}}(\mathbf{x})$ and $\underline{\widehat{\boldsymbol{\theta}_r}}(\mathbf{x})$ (2b) are referred to as "non-standard MLEs" (NSMLEs). Interestingly enough, despite its frequent occurrence in practical estimation problems, the study of NSMLEs has received little attention and the contributions have been limited to the derivation of two representatives of the Small-Error bounds, namely the Cramér-Rao bound (CRB) [6][34] and the Battacharayya bound (BaB) [35].

The aim of the present communication is to complete this initial characterization of estimation accuracy of NSMLEs. First, the intuitive idea [6][34][35] that NSMLEs are generally suboptimal estimates (in any asymptotic region of operation yielding unbiased efficient estimates) is rigorously established. Therefore, it is of interest to investigate the suboptimality of the NSMLEs, which can be, in some extent, quantified by lower bounds (LBs) derivation and comparison. Thus, as a second contribution, we show that any representative of the two general classes of LBs on the MSE, respectively the Small-Error bounds and the Large-Error bounds, has a non-standard version lower bounding the MSE of NSMLEs. Small-Error bounds are not able to handle the threshold phenomena, whereas it is revealed by Large-Error bounds that can be used to predict the threshold value. Last, some of the results introduced are exemplified by a new look at the well known Gaussian complex observation models.

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, by the CNES, and by the ANR MAGELLAN.

We focus on the scalar case, i.e. $\boldsymbol{\theta} \triangleq \theta$, although the results are easily extended to the estimation of a vector of parameters [30][31].

2. RELATION TO PRIOR WORK

Despite its frequent occurrence in estimation problems, the study of NSMLEs (2b) has received little attention and the contributions have been limited to the derivation of the appropriate CRB [6][34] and BaB [35]. Our contribution is two-fold: we show that 1) NSMLEs are generally suboptimal estimates in any asymptotic regions of operation yielding unbiased efficient estimates, 2) any standard Small-Error or Large-Error bound on the MSE has a non-standard version lower bounding the MSE of NSMLEs.

3. ON THE SUBOPTIMALITY OF NSMLES

In the present communication, the discussion is restricted to the case where the supports of $p(\mathbf{x}, \boldsymbol{\theta}_r|\theta)$ and $p(\boldsymbol{\theta}_r|\mathbf{x}, \theta)$ are independent of θ: $\Delta(\theta) = \left\{(\mathbf{x}, \boldsymbol{\theta}_r) \in \mathbb{C}^M \times \mathbb{R}^{P_r} \mid p(\mathbf{x}, \boldsymbol{\theta}_r|\theta) > 0\right\} \triangleq \Delta$ and $\Pi_{\boldsymbol{\theta}_r|\mathbf{x}}(\theta) = \left\{\boldsymbol{\theta}_r \in \mathbb{R}^{P_r} \mid p(\mathbf{x}, \boldsymbol{\theta}_r|\theta) > 0\right\} \triangleq \Pi_{\boldsymbol{\theta}_r|\mathbf{x}}$. Let $\mathcal{L}^2(\Omega)$, respectively $\mathcal{L}^2(\Delta)$, be the real Euclidean space of square integrable real-valued functions over the domain Ω, respectively Δ. Let us denote $\boldsymbol{\phi} = \left(\theta, \boldsymbol{\theta}_r^T\right)^T \in \Theta \times \mathbb{R}^{P_r}$, $p(\mathbf{x}|\boldsymbol{\phi}) \triangleq p(\mathbf{x}|\boldsymbol{\theta}_r, \theta)$ and $E_{\mathbf{x}|\boldsymbol{\phi}}[\,] \triangleq E_{\mathbf{x}|\boldsymbol{\theta}_r,\theta}[\,]$. Then any estimator $\widehat{\boldsymbol{\phi}} = \left(\widehat{\theta}, \widehat{\boldsymbol{\theta}}_r^T\right)^T \triangleq \widehat{\boldsymbol{\phi}}(\mathbf{x}, \boldsymbol{\theta}_r) \in \mathcal{L}^2(\Delta)$ of a selected vector value $\boldsymbol{\phi}$, uniformly unbiased for $p(\mathbf{x}|\boldsymbol{\phi})$, must comply with:

$$\forall \boldsymbol{\phi}' \in \Theta \times \mathbb{R}^{P_r} : E_{\mathbf{x}|\boldsymbol{\phi}'}\left[\widehat{\boldsymbol{\phi}}\right] = \boldsymbol{\phi}', \quad (3)$$

which implies that:

$$\forall \theta' \in \Theta : E^T_{\mathbf{x},\boldsymbol{\theta}'_r|\theta'}\left[\widehat{\boldsymbol{\phi}}\right] = E^T_{\boldsymbol{\theta}'_r|\theta'}\left[\boldsymbol{\phi}'\right] = \left(\theta', E^T_{\boldsymbol{\theta}'_r|\theta'}\left[\boldsymbol{\theta}'_r\right]\right), \quad (4)$$

that is $\widehat{\boldsymbol{\phi}}$ is an uniformly unbiased estimate of $\mathbf{g}(\theta)^T = \left(\theta, E^T_{\boldsymbol{\theta}_r|\theta}[\boldsymbol{\theta}_r]\right)$ for $p(\mathbf{x}, \boldsymbol{\theta}_r|\theta)$. As the reciprocal is not true:

$$\forall \theta' \in \Theta : E_{\mathbf{x},\boldsymbol{\theta}'_r|\theta'}\left[\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}'\right] = \mathbf{0} \nRightarrow \forall \boldsymbol{\phi}' \in \Theta \times \mathbb{R}^{P_r} : E_{\mathbf{x}|\boldsymbol{\phi}'}\left[\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}'\right] = \mathbf{0},$$

then $\mathcal{U}(\Delta) = \left\{\widehat{\boldsymbol{\phi}} \in \mathcal{L}^2(\Delta) \text{ verifying (3)}\right\} \subset \mathcal{V}(\Delta) = \left\{\widehat{\boldsymbol{\phi}} \in \mathcal{L}^2(\Delta) \text{ verifying (4)}\right\}$[1]. Let $\mathcal{U}(\Omega)$ and $\mathcal{V}(\Omega)$ be the restriction to $\mathcal{L}^2(\Omega)$ of $\mathcal{U}(\Delta)$ and $\mathcal{V}(\Delta)$. As $\forall \widehat{\boldsymbol{\phi}} \in \mathcal{L}^2(\Delta)$:

$$E_{\mathbf{x},\boldsymbol{\theta}_r|\theta}\left[\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}(\theta)\right)\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}(\theta)\right)^T\right] = E_{\mathbf{x},\boldsymbol{\theta}_r|\theta}\left[\left(\widehat{\boldsymbol{\phi}} - E_{\mathbf{x}|\boldsymbol{\phi}}\left[\widehat{\boldsymbol{\phi}}\right]\right)\left(\widehat{\boldsymbol{\phi}} - E_{\mathbf{x}|\boldsymbol{\phi}}\left[\widehat{\boldsymbol{\phi}}\right]\right)^T\right] + E_{\boldsymbol{\theta}_r|\theta}\left[\left(E_{\mathbf{x}|\boldsymbol{\phi}}\left[\widehat{\boldsymbol{\phi}}\right] - \mathbf{g}(\theta)\right)\left(E_{\mathbf{x}|\boldsymbol{\phi}}\left[\widehat{\boldsymbol{\phi}}\right] - \mathbf{g}(\theta)\right)^T\right], \quad (5a)$$

therefore, if $\widehat{\boldsymbol{\phi}} \in \mathcal{U}(\Delta)$:

$$E_{\mathbf{x},\boldsymbol{\theta}_r|\theta}\left[\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}(\theta)\right)\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}(\theta)\right)^T\right] = E_{\mathbf{x},\boldsymbol{\theta}_r|\theta}\left[\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)^T\right] + \mathbf{C}_\theta(\boldsymbol{\phi}) \quad (5b)$$

where $\mathbf{C}_\theta(\boldsymbol{\phi}) = \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{C}_\theta(\boldsymbol{\theta}_r) \end{bmatrix}$ and $\mathbf{C}_\theta(\boldsymbol{\theta}_r)$ is the covariance matrix of $\boldsymbol{\theta}_r$. Also, as $\mathcal{U}(\Omega) \subset \mathcal{V}(\Omega)$ and $\mathcal{U}(\Omega) \subset \mathcal{U}(\Delta)$, finally:

$$\min_{\widehat{\boldsymbol{\phi}} \in \mathcal{V}(\Omega)}\left\{E_{\mathbf{x}|\theta}\left[\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}(\theta)\right)\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}(\theta)\right)^T\right]\right\} \leq \min_{\widehat{\boldsymbol{\phi}} \in \mathcal{U}(\Omega)}\left\{E_{\mathbf{x},\boldsymbol{\theta}_r|\theta}\left[\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)^T\right]\right\} + \mathbf{C}_\theta(\boldsymbol{\phi}) \quad (6a)$$

and, in particular, as $\widehat{\theta} = \mathbf{e}_1^T \widehat{\boldsymbol{\phi}}$ where $\mathbf{e}_1 = (1, 0, \ldots, 0)^T$:

$$\min_{\widehat{\theta} \in \mathcal{V}(\Omega)}\left\{E_{\mathbf{x}|\theta}\left[\left(\widehat{\theta} - \theta\right)^2\right]\right\} \leq \min_{\widehat{\theta} \in \mathcal{U}(\Omega)}\left\{E_{\mathbf{x}|\theta}\left[\left(\widehat{\theta} - \theta\right)^2\right]\right\}. \quad (6b)$$

In any asymptotic regions of operation yielding unbiased efficient estimates, $\widehat{\theta}_{ML} \in \mathcal{V}(\Omega)$, $\underline{\widehat{\theta}} \in \mathcal{U}(\Omega)$ and both reach the minimum MSE. Thus, according to (6b), the NSMLEs of θ is generally an asymptotically suboptimal estimator of θ (in the MSE sense) in comparison with the MLE of θ. Therefore, from a theoretical as well as a practical viewpoint, it is of interest to investigate the suboptimality of the NSMLEs, which can be, in some extent, quantified by LBs derivation and comparison.

4. NON-STANDARD LOWER BOUNDS

It is worth noticing that an equivalent form of (6a) is:

$$\min_{\widehat{\boldsymbol{\phi}} \in \mathcal{V}(\Omega)}\left\{E_{\mathbf{x}|\theta}\left[\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}(\theta)\right)\left(\widehat{\boldsymbol{\phi}} - \mathbf{g}(\theta)\right)^T\right]\right\} - \mathbf{C}_\theta(\boldsymbol{\phi}) \leq \min_{\widehat{\boldsymbol{\phi}} \in \mathcal{U}(\Omega)}\left\{E_{\boldsymbol{\theta}_r|\theta}\left[E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)\left(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}\right)^T\right]\right]\right\}, \quad (6c)$$

since the latter form (6c) is the corner-stone to derive LBs on the MSE of NSMLEs. Indeed, since $\mathcal{U}(\Omega) \subset \mathcal{U}(\Delta)$, any LB on the MSE over $\mathcal{U}(\Delta)$ is a LB on the MSE over $\mathcal{U}(\Omega)$.

In its seminal work in standard estimation [19], Barankin has shown that the locally-best at θ uniformly unbiased estimator is the solution of a norm minimization under linear constraints [5, Section 2]:

$$\min_{\widehat{\theta} \in \mathcal{L}^2(\Omega)}\left\{\left\|\widehat{\theta}(\mathbf{x}) - \theta\right\|_\theta^2\right\} \text{ under } \left\langle \widehat{\theta}(\mathbf{x}) - \theta \mid \upsilon_\theta(\mathbf{x}; \theta') \right\rangle_\theta = \theta' - \theta, \forall \theta' \in \Theta, \quad (7)$$

where $\upsilon_\theta(\mathbf{x}; \theta') = \frac{p(\mathbf{x}|\theta')}{p(\mathbf{x}|\theta)}$ denotes the likelihood ratio (LR), and:

$$MSE_\theta\left[\widehat{\theta}\right] = \left\|\widehat{\theta}(\mathbf{x}) - \theta\right\|_\theta^2, \langle g(\mathbf{x}) \mid h(\mathbf{x}) \rangle_\theta = E_{\mathbf{x}|\theta}\left[g(\mathbf{x})\, h(\mathbf{x})\right].$$

Unfortunately, if Θ contains a continuous subset of $\mathbb{R}$, then the norm minimization (7) leads to an integral equation with no analytical solution in general. As a consequence, many studies quoted in [28]-[32] have been dedicated to the derivation of "computable" LBs approximating the MSE of the locally-best uniformly unbiased estimator, aka the Barankin bound (BB). All these approximations derive from sets of discrete or integral linear transform of the "Barankin" constraint (7):

$$E_{\mathbf{x}|\theta}\left[\left(\widehat{\theta}(\mathbf{x}) - \theta\right)\upsilon_\theta(\mathbf{x}; \theta')\right] = \theta' - \theta, \forall \theta' \in \Theta.$$

These results are readily generalizable to the parameters vector case [30][31], that is any Barankin bound approximation (BBA)

[1] In most cases, the inclusion is strict leading to strict inequalities (6a-6c)

on $\min_{\widehat{\boldsymbol{\phi}}\in\mathcal{U}(\Delta)}\left\{E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\boldsymbol{\phi}}-\boldsymbol{\phi}\right)\left(\widehat{\boldsymbol{\phi}}-\boldsymbol{\phi}\right)^T\right]\right\}$ can be derived from discrete or integral linear transforms of the set of constraints:

$$\forall n\in[1,N],E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\boldsymbol{\phi}}-\boldsymbol{\phi}\right)\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}^n\right)\right]=\boldsymbol{\phi}^n-\boldsymbol{\phi}, \tag{8a}$$

where $\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}^n\right)=\frac{p(\mathbf{x}|\boldsymbol{\phi}^n)}{p(\mathbf{x}|\boldsymbol{\phi})}$, that is as the solution of:

$$\begin{aligned}&\min_{\widehat{\boldsymbol{\phi}}\in\mathcal{U}(\Delta)}\left\{E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\boldsymbol{\phi}}-\boldsymbol{\phi}\right)\left(\widehat{\boldsymbol{\phi}}-\boldsymbol{\phi}\right)^T\right]\right\} \text{ under}\\ &\qquad E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\boldsymbol{\phi}}-\boldsymbol{\phi}\right)\boldsymbol{\upsilon}_{\boldsymbol{\phi}}^T\left(\boldsymbol{\Phi}^N\right)\right]=\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right),\end{aligned} \tag{8b}$$

where $\boldsymbol{\Phi}^N=\left[\boldsymbol{\phi}^1\ \dots\ \boldsymbol{\phi}^N\right]$, $\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)=\left[\boldsymbol{\phi}^1-\boldsymbol{\phi}\ \dots\ \boldsymbol{\phi}^N-\boldsymbol{\phi}\right]$, $\boldsymbol{\upsilon}_{\boldsymbol{\phi}}\left(\boldsymbol{\Phi}^N\right)\triangleq\boldsymbol{\upsilon}_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\Phi}^N\right)=\left(\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}^1\right),\dots,\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}^N\right)\right)^T$, which defines the following BBA [30, Lemma 1]:

$$\begin{aligned}&\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{BBA}\right)=\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^N\right)\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)^T,\\ &\qquad\widehat{\boldsymbol{\phi}}_{BBA}=\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^N\right)\boldsymbol{\upsilon}_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\Phi}^N\right),\end{aligned} \tag{8c}$$

where $\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}\left(\boldsymbol{\Phi}^N\right)=E_{\mathbf{x}|\boldsymbol{\phi}}\left[\boldsymbol{\upsilon}_{\boldsymbol{\phi}}\left(\boldsymbol{\Phi}^N\right)\boldsymbol{\upsilon}_{\boldsymbol{\phi}}^T\left(\boldsymbol{\Phi}^N\right)\right]$ and $\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{BBA}\right)=E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\boldsymbol{\phi}}_{BBA}-\boldsymbol{\phi}\right)\left(\widehat{\boldsymbol{\phi}}_{BBA}-\boldsymbol{\phi}\right)^T\right]$ is the covariance matrix of $\widehat{\boldsymbol{\phi}}_{BBA}$. Even if in general $\widehat{\boldsymbol{\phi}}_{BBA}\triangleq\widehat{\boldsymbol{\phi}}_{BBA}\left(\mathbf{x};\boldsymbol{\phi}\right)$ (8c) is a clairvoyant estimator and does not belong to $\mathcal{U}(\Omega)$, as:

$$\begin{aligned}&E_{\boldsymbol{\theta}_r|\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{BBA}\right)\right]\leq\\ &\qquad\min_{\widehat{\boldsymbol{\phi}}\in\mathcal{U}(\Omega)}\left\{E_{\mathbf{x},\boldsymbol{\theta}_r|\theta}\left[\left(\widehat{\boldsymbol{\phi}}-\boldsymbol{\phi}\right)\left(\widehat{\boldsymbol{\phi}}-\boldsymbol{\phi}\right)^T\right]\right\},\end{aligned} \tag{9}$$

$\mathcal{U}(\Omega)$ containing asymptotically the NSMLEs, it seems sensible to call $E_{\boldsymbol{\theta}_r|\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{BBA}\right)\right]$ a non-standard LB (NSLB) and to denote $\mathbf{NSLB}\triangleq E_{\boldsymbol{\theta}_r|\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{BBA}\right)\right]$ to make the difference with the modified LBs (MLB) which are LBs on $\min_{\widehat{\boldsymbol{\phi}}\in\mathcal{V}(\Delta)}\left\{E_{\mathbf{x},\boldsymbol{\theta}_r|\theta}\left[\left(\widehat{\boldsymbol{\phi}}-\mathbf{g}(\theta)\right)\left(\widehat{\boldsymbol{\phi}}-\mathbf{g}(\theta)\right)^T\right]\right\}$ [5]. In the same vein, $E_{\boldsymbol{\theta}_r|\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{BBA}\right)\right]$ can also be regarded as a non-standard BBA (NSBBA). 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.

Last, since $\mathcal{U}(\Delta)\not\subseteq\mathcal{V}(\Omega)$ and $\mathcal{V}(\Omega)\not\subseteq\mathcal{U}(\Delta)$, no general result can be drawn from (6c) on the ordering between $E_{\boldsymbol{\theta}_r|\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{BBA}\right)\right]+\mathbf{C}_{\theta}\left(\boldsymbol{\phi}\right)$ and $\min_{\widehat{\boldsymbol{\phi}}\in\mathcal{V}(\Omega)}\left\{E_{\mathbf{x}|\theta}\left[\left(\widehat{\boldsymbol{\phi}}-\mathbf{g}(\theta)\right)\left(\widehat{\boldsymbol{\phi}}-\mathbf{g}(\theta)\right)^T\right]\right\}$ or any BBA computed on $\mathcal{V}(\Omega)$.

4.1. Lower bound examples

A typical example is the CRB obtained for $N=2$, where $\boldsymbol{\phi}^1=\left(\theta,\boldsymbol{\theta}_r^T\right)^T$ and $\boldsymbol{\phi}^2=\left(\theta+d\theta,\boldsymbol{\theta}_r^T\right)^T$, leading to the following subset of constraints:

$$\begin{pmatrix}0\\ d\theta\end{pmatrix}=E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\begin{pmatrix}1\\ \upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}^2\right)\end{pmatrix}\right], \tag{10a}$$

which is equivalent to [33, Lemma 3]:

$$\begin{pmatrix}0\\ 1\end{pmatrix}=E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\begin{pmatrix}1\\ \frac{\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}^2\right)-1}{d\theta}\end{pmatrix}\right], \tag{10b}$$

and can be reduced to [33, Lemma 2]:

$$1=E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\frac{p\left(\mathbf{x}|\boldsymbol{\theta}_r,\theta+d\theta\right)-p\left(\mathbf{x}|\boldsymbol{\theta}_r,\theta\right)}{d\theta p\left(\mathbf{x}|\boldsymbol{\theta}_r,\theta\right)}\right], \tag{10c}$$

since $E_{\mathbf{x}|\boldsymbol{\phi}}\left[1\times\left(\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}^2\right)-1\right)\right]=0$. Then by letting $d\theta$ be infinitesimally small, (9) becomes:

$$NSCRB\triangleq E_{\boldsymbol{\theta}_r|\theta}\left[E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\frac{\partial\ln p\left(\mathbf{x}|\boldsymbol{\phi}\right)}{\partial\theta}\right)^2\right]^{-1}\right], \tag{11}$$

that is the Miller and Chang bound [6, (7)]. Following the rationale introduced in [22], a straightforward extension of (11) is obtained for $\boldsymbol{\Phi}^N=\left[\boldsymbol{\phi}^1\ \dots\ \boldsymbol{\phi}^N\right]$, $\boldsymbol{\phi}^n=\left(\theta+(n-1)\,d\theta,\boldsymbol{\theta}_r^T\right)^T$, $1\leq n\leq N$. Indeed the set of N associated constraints:

$$d\theta\left(0,\dots,N-1\right)^T=E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\boldsymbol{\upsilon}_{\boldsymbol{\phi}}\left(\boldsymbol{\Phi}^N\right)\right], \tag{12a}$$

by letting $d\theta$ be infinitesimally small, becomes equivalent to [22]:

$$\left(0,1,0,\dots,0\right)^T=E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\mathbf{b}'\left(\mathbf{x};\boldsymbol{\phi}\right)\right], \tag{12b}$$

where $\mathbf{b}'\left(\mathbf{x};\boldsymbol{\phi}\right)=\frac{1}{p(\mathbf{x}|\boldsymbol{\phi})}\left(p\left(\mathbf{x}|\boldsymbol{\phi}\right),\frac{\partial p(\mathbf{x}|\boldsymbol{\phi})}{\partial\theta},\dots,\frac{\partial^{N-1}p(\mathbf{x}|\boldsymbol{\phi})}{\partial^{N-1}\theta}\right)^T$. As $E_{\mathbf{x}|\boldsymbol{\phi}}\left[b_1'\left(\mathbf{x};\boldsymbol{\phi}\right)b_n'\left(\mathbf{x};\boldsymbol{\phi}\right)\right]=E_{\mathbf{x}|\boldsymbol{\phi}}\left[\frac{\partial^n p(\mathbf{x}|\boldsymbol{\phi})}{\partial^n\theta}\right]=0$, $2\leq n\leq N-1$, (12b) is actually equivalent to [33, Lemma 2]:

$$\mathbf{e}_1=E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\mathbf{b}\left(\mathbf{x};\boldsymbol{\phi}\right)\right], \tag{12c}$$

where $\mathbf{b}\left(\mathbf{x};\boldsymbol{\phi}\right)=\frac{1}{p(\mathbf{x}|\boldsymbol{\phi})}\left(\frac{\partial p(\mathbf{x}|\boldsymbol{\phi})}{\partial\theta},\dots,\frac{\partial^{N-1}p(\mathbf{x}|\boldsymbol{\phi})}{\partial^{N-1}\theta}\right)^T$, $\mathbf{e}_1=\left(1,0,\dots,0\right)^T$, and (9) leads to the BaBs [18] of order $N-1$ [35]:

$$NSBaB\triangleq E_{\boldsymbol{\theta}_r|\theta}\left[\mathbf{e}_1^T E_{\mathbf{x}|\boldsymbol{\phi}}\left[\mathbf{b}\left(\mathbf{x};\boldsymbol{\phi}\right)\mathbf{b}^T\left(\mathbf{x};\boldsymbol{\phi}\right)\right]^{-1}\mathbf{e}_1\right]. \tag{13}$$

Last, a key representative of the Large Error bounds is the Mcaulay-Seidman bound (MSB) [24] which is the practical form of the BB. Its non-standard form is : $NSMSB=E_{\boldsymbol{\theta}_r|\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{BBA}\right)\right]$.

4.2. Tighter non-standard lower bounds

Interestingly enough, it is quite simple to introduce tighter NSLBs. It suffices to note that the addition of any subset of K constraints:

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

to (8b) restricts the class of viable estimators $\widehat{\boldsymbol{\phi}}\left(\mathbf{x},\boldsymbol{\theta}_r\right)$ and therefore increases the associated NSBBA (8c), leading to:

$$\begin{aligned}&E_{\boldsymbol{\theta}_r|\theta}\left[\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^N\right)\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^N\right)^T\right]\leq\\ &\quad E_{\boldsymbol{\theta}_r|\theta}\left[\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^{N+K}\right)\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^{N+K}\right)\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^{N+K}\right)^T\right],\end{aligned} \tag{14a}$$

and, regarding the estimation of θ, to:

$$\begin{aligned}&E_{\boldsymbol{\theta}_r|\theta}\left[\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^N\right)\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)\right]\leq\\ &\quad E_{\boldsymbol{\theta}_r|\theta}\left[\boldsymbol{\xi}\left(\boldsymbol{\theta}^{N+K}\right)^T\mathbf{R}_{\boldsymbol{\upsilon}_{\boldsymbol{\phi}}}^{-1}\left(\boldsymbol{\Phi}^{N+K}\right)\boldsymbol{\xi}\left(\boldsymbol{\theta}^{N+K}\right)\right],\end{aligned} \tag{14b}$$

where $\boldsymbol{\xi}\left(\boldsymbol{\theta}^{L}\right)=\left(\theta^{1}-\theta,\ldots,\theta^{L}-\theta\right)^{T}$, $\boldsymbol{\Phi}^{L}=\left[\boldsymbol{\phi}^{1}\ \ldots\ \boldsymbol{\phi}^{L}\right]$, $\boldsymbol{\Xi}\left(\boldsymbol{\Phi}^{L}\right)=\left[\boldsymbol{\phi}^{1}-\boldsymbol{\phi}\ \ldots\ \boldsymbol{\phi}^{L}-\boldsymbol{\phi}\right]$ for $L\in\{N,N+K\}$.
A typical example is given by the NSCRB (11). Indeed by adding to (10a) the following $K=P_r$ constraints:

$$\mathbf{0}=E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\Phi}^{K}\right)\right],$$
$$\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\Phi}^{K}\right)=\left(\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}^{1}\right),\ldots,\upsilon_{\boldsymbol{\phi}}\left(\mathbf{x};\boldsymbol{\phi}^{K}\right)\right)^{T},$$

where $\boldsymbol{\phi}^{k}=\left(\theta,\left(\boldsymbol{\theta}_r+\mathbf{u}_k h_r^{k}\right)^{T}\right)^{T}$ and $\mathbf{u}_k$ is the kth column of the identity matrix $\mathbf{I}_{P_r}$, one obtains the following equivalent set of constraints [33, Lemma 3+Lemma 2]:

$$\mathbf{e}_1=E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\widehat{\theta}\left(\mathbf{x},\boldsymbol{\theta}_r\right)-\theta\right)\mathbf{c}\left(\mathbf{x};\boldsymbol{\Phi}^{K+1}\right)\right],$$
$$\mathbf{c}\left(\mathbf{x};\boldsymbol{\Phi}^{K+1}\right)=\left(\frac{1}{d\theta}\left(\frac{p(\mathbf{x}|\boldsymbol{\theta}_r,\theta+d\theta)}{p(\mathbf{x}|\boldsymbol{\theta}_r,\theta)}-1\right),\right.$$
$$\left.\frac{1}{h_1}\left(\frac{p(\mathbf{x}|\boldsymbol{\theta}_r+\mathbf{u}_1h_1,\theta)}{p(\mathbf{x}|\boldsymbol{\theta}_r,\theta)}-1\right),\ldots,\frac{1}{h_K}\left(\frac{p(\mathbf{x}|\boldsymbol{\theta}_r+\mathbf{u}_Kh_K,\theta)}{p(\mathbf{x}|\boldsymbol{\theta}_r,\theta)}-1\right)\right).$$

By letting $(d\theta,h_1,\ldots,h_{P_r})$ be infinitesimally small, then $\mathbf{c}\left(\mathbf{x};\boldsymbol{\Phi}^{K+1}\right)\rightarrow\frac{\partial\ln p(\mathbf{x}|\boldsymbol{\phi})}{\partial\boldsymbol{\phi}}$ and (14b) becomes [34, (24)]:

$$E_{\boldsymbol{\theta}_r|\theta}\left[E_{\mathbf{x}|\boldsymbol{\phi}}\left[\left(\frac{\partial\ln p\left(\mathbf{x}|\boldsymbol{\phi}\right)}{\partial\theta}\right)^{2}\right]^{-1}\right]\leq$$
$$E_{\boldsymbol{\theta}_r|\theta}\left[\mathbf{e}_1^{T}E_{\mathbf{x}|\boldsymbol{\phi}}\left[\frac{\partial\ln p\left(\mathbf{x}|\boldsymbol{\phi}\right)}{\partial\boldsymbol{\phi}}\frac{\partial\ln p\left(\mathbf{x}|\boldsymbol{\phi}\right)}{\partial\boldsymbol{\phi}^{T}}\right]^{-1}\mathbf{e}_1\right].\quad(15)$$

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 (14b), however, at an additional cost in numerical integration or Monte Carlo simulation to evaluate the additional statistical expectations.

4.3. Further considerations

Since any existing standard Small-Error [7][15]-[18] or Large-Error bound [19]-[26][28]-[31] on $\boldsymbol{\phi}$ can be obtained from (8c), it has a non-standard form obtained from $E_{\boldsymbol{\theta}_r|\theta}\left[\mathbf{C}_{\boldsymbol{\phi}}\left(\widehat{\boldsymbol{\phi}}_{BBA}\right)\right]$ [5]. Let us recall that in general $\widehat{\boldsymbol{\phi}}_{BBA}\triangleq\widehat{\boldsymbol{\phi}}_{BBA}\left(\mathbf{x};\boldsymbol{\phi}\right)\in\mathcal{U}\left(\Delta\right)$, therefore the associated NSLB can not be compared a priori neither with the MSE of $\widehat{\theta}_{ML}\in\mathcal{V}\left(\Omega\right)$ 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; comparisons are possible only on a "case-by-case basis". However if $p\left(\boldsymbol{\theta}_r|\theta\right)$ does not depend on θ, then one can derive inequalities between modified [5] and non-standard LBs (proofs are given in [36]). In the general case where $p\left(\boldsymbol{\theta}_r|\theta\right)$ does depend on θ, no general inequalities between modified and non-standard LBs can any longer be exhibited; comparisons are possible only on a "case-by-case basis".

5. A NEW LOOK AT GAUSSIAN OBSERVATION MODELS

In many practical problems of interest (radar, sonar, communication, ...), the complex M-dimensional observation vector $\mathbf{x}$ consists of a bandpass signal which is the output of an Hilbert filtering leading to a complex Gaussian circular vector $\mathbf{x}\sim\mathcal{CN}\left(\mathbf{m}_{\mathbf{x}},\mathbf{C}_{\mathbf{x}}\right)$ [1][37, §13][38]. Two particular signal models are mostly considered: the deterministic (conditional) signal model and the stochastic (unconditional) signal model [39]. In the deterministic case the unknown parameters are connected with the expectation value, whereas they are connected with the covariance matrix in the stochastic one. A simple and well known instantiation is:

$$\mathbf{x}_t=\mathbf{s}\left(\tau\right)a_t+\mathbf{n}_t,\quad 1\leq t\leq T,\quad(16)$$

where $a_1,\ldots,a_t$ 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, $\mathbf{n}_t\sim\mathcal{CN}\left(\mathbf{0},\sigma_{\mathbf{n}}^{2}\mathbf{I}_M\right)$, $1\leq t\leq T$, are independent and identically distributed (i.i.d.) Gaussian complex circular noises independent of the signal of interest. Additionally if $\mathbf{a}=\left(a_1,\ldots,a_T\right)^{T}\sim\mathcal{CN}\left(\mathbf{0},\sigma_a^{2}\mathbf{I}_T\right)$, then (16) is an unconditional signal model parameterized by $\boldsymbol{\theta}=\left(\tau,\sigma_a^{2},\sigma_{\mathbf{n}}^{2}\right)^{T}$, and the MLE (2a) of τ, aka the unconditional MLE (UMLE), is obtained by minimization of the concentrated criterion [11]: $\widehat{\tau}=\arg\min_{\tau}\left\{\left|\widehat{\sigma_a^{2}}\mathbf{s}\left(\tau\right)\mathbf{s}^{H}\left(\tau\right)+\widehat{\sigma_{\mathbf{n}}^{2}}\mathbf{I}_M\right|\right\}$. The associated CRB, aka the unconditional CRB (UCRB), is [11, (4.64)][40]:

$$UCRB_{\tau}=\sigma_{\mathbf{n}}^{2}\left(2h\left(\tau\right)T\sigma_a^{2}\frac{SNR}{SNR+1}\right)^{-1},\quad(17)$$
$$SNR=\frac{\sigma_a^{2}\left\|\mathbf{s}\left(\tau\right)\right\|^{2}}{\sigma_{\mathbf{n}}^{2}},\ h\left(\tau\right)=\frac{\partial\mathbf{s}\left(\tau\right)}{\partial\tau}^{H}\Pi_{\mathbf{s}(\tau)}^{\perp}\frac{\partial\mathbf{s}\left(\tau\right)}{\partial\tau},$$

where $\Pi_{\mathbf{a}}^{\perp}=\mathbf{I}_M-\mathbf{a}\mathbf{a}^{H}\left\|\mathbf{a}\right\|^{-2}$ and SNR is the signal-to-noise ratio computed at the output of the single source matched filter [11]. The NSMLE (2b) of τ is actually the conditional MLE (CMLE) obtained by minimization of the concentrated criterion [11]: $\widehat{\underline{\tau}}=\arg\min_{\tau}\left\{\sum_{t=1}^{T}\mathbf{x}_t^{H}\Pi_{\mathbf{s}(\tau)}^{\perp}\mathbf{x}_t\right\}$, and the associated NSCRB is:

$$NSCRB_{\tau}=E_{\mathbf{a}|\sigma_a^{2}}\left[CCRB_{\tau}\left(\mathbf{a}\right)\right],CCRB_{\tau}\left(\mathbf{a}\right)=\frac{\sigma_{\mathbf{n}}^{2}}{2h\left(\tau\right)\left\|\mathbf{a}\right\|^{2}},$$

where $CCRB_{\tau}$ denotes the conditional CRB associated to the CMLE [11, (4.68)]. First, it has been shown [11, (4.74)], in the case of a vector of unknown parameters $\boldsymbol{\tau}$, that asymptotically where $T\rightarrow\infty$:

$$\mathbf{C}_{\boldsymbol{\theta}}\left(\widehat{\underline{\boldsymbol{\tau}}}\right)\geq\mathbf{C}_{\boldsymbol{\theta}}\left(\widehat{\boldsymbol{\tau}}\right)=\mathbf{UCRB}_{\boldsymbol{\tau}}\geq\mathbf{CCRB}_{\boldsymbol{\tau}},\quad(18)$$

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, (18) becomes:

$$C_{\boldsymbol{\theta}}\left(\widehat{\underline{\tau}}\right)=C_{\boldsymbol{\theta}}\left(\widehat{\tau}\right)=UCRB_{\tau},$$

which highlights that in some particular cases the NSMLE may be asymptotically equivalent to the MLE in the MSE sense.
Second, since $\frac{\|\mathbf{a}\|^{2}}{\sigma_a^{2}}\sim\chi_{2T}^{2}$, i.e. a chi-squared random variable with 2T degrees of freedom, then [37]:

$$NSCRB_{\tau}=\left|\begin{array}{ll}\frac{\sigma_{\mathbf{n}}^{2}}{2T\sigma_a^{2}h(\tau)}\frac{T}{T-1} & \text{if } T\geq 2\\ \infty & \text{if } T=1\end{array}\right..\quad(19)$$

Therefore, if $T\geq 2$:

$$\frac{NSCRB_{\tau}}{UCRB_{\tau}}=\frac{NSCRB_{\tau}}{C_{\boldsymbol{\theta}}\left(\widehat{\tau}\right)}=\frac{T}{T-1}\frac{SNR}{SNR+1}\quad(20)$$

which illustrates the facts that NSLB are not in general neither upper bounds on the MSE of MLEs nor on any of its LBs.

6. REFERENCES

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

[2] F. Gini, R. Reggiannini, and U. Mengali, “The modified Cramér-Rao bound in vector parameter estimation,” IEEE Trans. Commun., 46(1): 52-60, 1998

[3] 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

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

[5] J. Galy, E. Chaumette, F. Vincent, A. Renaux and P. Larzabal, ”Lower bounds for non-standard deterministic estimation”, in Proc. of IEEE SAM 2016

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

[7] H. Cramér, *Mathematical Methods of Statistics*. Princeton, NJ: Princeton Univ. Press, 1946

[8] E. L. Lehmann and G. Casella, *Theory of Point Estimation (2nd ed.)*. Springer, 1998

[9] P. Stoica, A. Nehorai, ”Performances study of conditional and unconditional direction of arrival estimation”, IEEE Trans. on SP, 38(10): 1783-1795, 1990

[10] M. Viberg and B. Ottersten, ”Sensor array processing based on subspace fitting”, IEEE Trans. on SP, 39(5): 1110-1121, 1991

[11] B. Ottersten, M. Viberg, P. Stoica and A. Nehorai, ”Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing”, in *Radar Array Processing*, S. Haykin, J. Litva, and T. J. Shepherd, Eds. Berlin, Germany: Springer-Verlag, ch. 4, pp. 99-151, 1993

[12] J. Li and R. T. Compton, ”Maximum likelihood angle estimation for signals with known waveforms”, IEEE Trans. on SP, 41(9): 2850-2862, 1993

[13] 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

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

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

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

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

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

[19] E.W. Barankin, ”Locally best unbiased estimates”, Ann. Math. Stat., 20(4): 477-501, 1949.

[20] J.M. Hammersley, ”Estimating Restricted Parameters”, Journal of Roy. Stat. Soc., 12(2): 192-240, 1950

[21] D.G. Chapman, H. Robbins, ”Minimum variance estimation without regularity assumptions”, Ann. Math. Stat., 22: 581-586, 1951

[22] D.A.S. Fraser and I. Guttman, ”Bhattacharyya bounds without regularity assumptions,” Ann. Math. Stat., 23(4): 629-632, 1952

[23] J. Kiefer, ”On minimum variance estimators”, Ann. Math. Stat., 23(4): 627–629, 1952.

[24] 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

[25] R. McAulay, E.M. Hofstetter, ” Barankin Bounds on parameter estimation”, IEEE Trans. on IT, 17(6): 669-676, 1971

[26] F.E. Glave, ”A new look at the Barankin Lower Bound”, IEEE Trans. on IT, 18(3): 349-356, 1972

[27] C. R. Blyth, ”Necessary and sufficient conditions for inequalities of Cramér-Rao type,” Ann. Math. Stat., 2(3): 464-473, 1974

[28] J.S. Abel, ”A bound on mean-square-estimate error”, IEEE Trans. on IT, 39(5): 1675-1680, 1993

[29] H. L. Van Trees and K. L. Bell, Eds., *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. New-York, NY, USA: Wiley/IEEE Press, 2007.

[30] E. Chaumette, J. Galy, A. Quinlan, P. Larzabal, ”A New Barankin Bound Approximation for the Prediction of the Threshold Region Performance of Maximum-Likelihood Estimators”, IEEE Trans. on SP, 56(11): 5319-5333, 2008

[31] 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

[32] K. Todros and J. Tabrikian, ” Uniformly Best Biased Estimators in Non-Bayesian Parameter Estimation”, IEEE Trans. on IT, 57(11): 7635-7647, 2011

[33] 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

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

[35] F. Lu and J. V. Krogmeier, ”Modified Bhattacharyya bounds and their application to timing estimation”, in Proc of IEEE Wireless Communications and Networking Conference, 2002

[36] J. Galy, N. Kbayer, E. Chaumette, F. Vincent, A. Renaux and P. Larzabal, ”On Lower Bounds for Non-Standard Deterministic Estimation”, submitted to IEEE Trans. on SP

[37] S.M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998

[38] B. Picinbono, ”On circularity,” IEEE Trans. on SP, 42(12): 3473-3482, 1994

[39] P. Stoica and A. Nehorai, ”Performances study of conditional and unconditional direction of arrival estimation,” IEEE Trans. on SP, 38(10): 1783-1795, 1990

[40] 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