A TIGHTER BAYESIAN CRAMÉR-RAO BOUND

Lucien Bacharach[†], *Carsten Fritsche*[⋆], *Umut Orguner*[‡], *Eric Chaumette*[†]

[†] University of Toulouse/Isae-Supaero, Toulouse, France (lucien.bacharach@isae.fr, eric.chaumette@isae.fr)

[⋆] Linköping University - Dept. of Electrical Engineering, Linköping, Sweden (carsten@isy.liu.se)

[‡] Middle East Technical University, Dept. of Electrical Engineering, Ankara, Turkey (umut@metu.edu.tr)

ABSTRACT

It has been shown lately that any "standard" Bayesian lower bound (BLB) on the mean squared error (MSE) of the Weiss-Weinstein family (WWF) admits a "tighter" form which upper bounds the "standard" form. Applied to the Bayesian Cramér-Rao bound (BCRB), this result suggests to redefine the concept of efficient estimator relatively to the tighter form of the BCRB, an update supported by a noteworthy example. This paper lays the foundation to revisit some Bayesian estimation problems where the BCRB is not tight in the asymptotic region.

***Index Terms*—** Mean Squared Error, Bayesian Lower Bounds, Bayesian Cramér-Rao bound, Minimum Mean Squared Error.

I. INTRODUCTION

Extracting parameter estimates from noisy observations of an underlying signal is a common problem in fields such as signal processing, communications, system identification, control and economics. The problem has generated a myriad of different parameter estimation algorithms, see for instance [1] [2]. This paper is concerned with establishing fundamental lower bounds (LBs) on how well these algorithms may be expected to perform in the mean squared error (MSE) sense. Indeed, minimal performance bounds allow for calculation of the best performance that may be achieved in the MSE sense and are, hence, a very useful system analysis tool. Specifically, they allow a system designer to probe how parameter estimation accuracy is influenced by various system design decisions. This analysis is free from the mechanistic details of a particular parameter estimation algorithm. There are two main categories of lower bounds [2]. Those that evaluate the "locally best" behaviour of the estimator and those that consider the "globally best" performance. In the first case, the parameters being estimated are considered to be deterministic, whereas the second category considers the parameters as random variables with an *a priori* probability. This paper is concerned with the second category of bounds concerning random parameters, which are named Bayesian lower bounds (BLBs).

Historically, the Bayesian Cramér-Rao bound (BCRB), was the first BLB to be derived [3] [4] [5], and is still the most commonly used BLB, which is largely due to its simplicity of calculation. Nevertheless, it is now well known that the BCRB is an optimistic bound in a non-linear estimation problem where the outliers effect generally appears, leading to a quick increase of the MSE: this is the so-called threshold effect which is not predicted by the BCRB. Since the knowledge of the particular value for which the threshold effect appears is a key feature allowing to define estimators optimal operating area, tightness [6] [7] [9] [10] is a prominent quality looked for a BLB in non-linear estimation problems. This has led to a large body of research based, so far, on two main families, i) the Ziv-Zakai family (ZZF) resulting from the conversion of an estimation bounding problem into one bounding binary hypothesis testing [6] [7] [8] and, ii) the Weiss-Weinstein family (WWF), derived from a covariance inequality principle [3]- [5], [9]- [27], lately used also for the derivation of Bayesian cyclic bounds for periodic parameter estimation [28]- [30]. In each family, some bounds, generally called "large-error" bounds (in contrast with "small-error" bounds such as the BCRB), can predict the threshold effect [2]. Interestingly enough, authors in [31] have lately shown that any "standard" BLBs of the WWF admits a "tighter" form which upper bounds the "standard" form. Indeed, if we consider the simple case of estimating a single real-valued random parameter[1] $\theta \in \Theta \subset \mathbb{R}$ from a N-dimensional real-valued random observation vector $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^N$, then $\boldsymbol{\phi}(\mathbf{x},\theta) = (\phi_1(\mathbf{x},\theta), \ldots, \phi_L(\mathbf{x},\theta))^T$ where $E_{\theta|\mathbf{x}}[\phi_l(\mathbf{x},\theta)] = 0$ for a.e. $\mathbf{x} \in \mathcal{X}$, $1 \leq l \leq L$, are BLB-generating functions of the WWF [20, (6)] [12, (1)] [23] and:

$$
\begin{aligned}
&E_{\mathbf{x},\theta}\left[\left(\theta - E_{\theta|\mathbf{x}}[\theta]\right)^2\right] \geq \\
&E_{\mathbf{x}}\left[\begin{array}{l} E_{\theta|\mathbf{x}}[\theta\boldsymbol{\phi}(\mathbf{x},\theta)]^T E_{\theta|\mathbf{x}}\left[\boldsymbol{\phi}(\mathbf{x},\theta)\boldsymbol{\phi}(\mathbf{x},\theta)^T\right]^{-1} \\ \times E_{\theta|\mathbf{x}}[\theta\boldsymbol{\phi}(\mathbf{x},\theta)] \end{array}\right] \geq \\
&\qquad E_{\mathbf{x},\theta}[\theta\boldsymbol{\phi}(\mathbf{x},\theta)]^T E_{\mathbf{x},\theta}\left[\boldsymbol{\phi}(\mathbf{x},\theta)\boldsymbol{\phi}(\mathbf{x},\theta)^T\right]^{-1} \\
&\qquad \times E_{\mathbf{x},\theta}[\theta\boldsymbol{\phi}(\mathbf{x},\theta)]
\end{aligned} \tag{1}
$$

where $E_{\theta|\mathbf{x}}[\theta]$ is the minimum MSE (MMSE) estimator of θ and the right-hand side of the above inequality is the "standard" form of a BLB of the WWF. Moreover in [31] a "closeness condition" in order to obtain a "standard" form equal to the "tighter" form of any BLB of the WWF has been derived. Interestingly enough, this condition is unlikely to be met in most of estimation problems where the "tighter" form is a strict upper bound on the "standard" form, which may explain why the "standard" BLBs of the WWF are not always as tight as expected.

To highlight this recent theoretical result, we tackle the problem of estimating the variance σ^2, i.e. $\theta \triangleq \sigma^2$, of a zero-mean Gaussian random variable where the a priori pdf of σ^2 follows a beta distribution [2, p7]. In this problem, the BCRBs is not tight in the large sample regime and the maximum a posteriori (MAP) estimate seems not to be efficient [2, p11]. We say "seems" because we show that the tighter BCRB (TBCRB) derived from (1) converges in the large sample regime towards the MSE of the MAP, and towards the MMSE as well. It is a sound evidence that the existence of the TBCRB suggests to introduce an updated definition of efficiency and an updated class of efficient estimators. For didactic purposes,

This work has been partially supported by the DGA/MRIS (2018.60.0072.00.470.75.01) and by the Excellence Center at Linköping and Lund in Information Technology (ELLIIT).

[1]Throughout the present paper scalars, vectors and matrices are represented, respectively, by italic, bold lowercase and bold uppercase characters. The matrix/vector transpose is indicated by a superscript ^{T}. $tr(\mathbf{A})$ denotes the trace of matrix $\mathbf{A}$. "s.t." stands for "subject to", "w.r.t." stands for "with respect to" and "a.e." stands for "almost every". "i.i.d." stands for independent and identically distributed.

we introduce an alternative and simpler derivation of the TBCRB than the one provided in [31] to establish (1) valid for any BLB of the WWF.

From a more general perspective, one of the merits of the paper is to show with a noteworthy example that the general class of BLBs tighter than the WWF proposed in [31], referred to as the tighter WWF (TWWF) of BLBs hereinafter, are of interest from a practical point of view. Indeed it is clear that the main drawback of the BLBs of the TWWF (1) is that they are unlikely to be arranged in closed form, in general, due to the presence of the statistical expectation; they can however have a closed form for the TBCRB or be evaluated by numerical integration or Monte Carlo simulation.

II. TIGHTER BAYESIAN CRAMÉR-RAO BOUNDS AND AN UPDATED DEFINITION OF EFFICIENCY

In this section, for didactic purposes, we consider a completely different approach than the one previously used in [31], which allows to introduce an alternative and simpler derivation of the TBCRB associated to the estimation of θ. In return, the proposed derivation can not be extended to the problem of estimating any $g(\theta)$. If $g(\theta) \neq \theta$, then (1) must be used directly.
First, from the definition of the mean, for a.e. $\mathbf{x} \in \mathcal{X}$:

$$E_{\theta|\mathbf{x}}\left[\left(\theta - \widehat{\theta}(\mathbf{x})\right)^2\right] \geq E_{\theta|\mathbf{x}}\left[\left(\theta - E_{\theta|\mathbf{x}}(\theta)\right)^2\right].$$

Second, from the covariance inequality [2] [32], $\forall \phi(\mathbf{x},\theta)$ s.t. $E_{\theta|\mathbf{x}}[\phi(\mathbf{x},\theta)] < \infty$ for a.e. $\mathbf{x} \in \mathcal{X}$, then for a.e. $\mathbf{x} \in \mathcal{X}$:

$$E_{\theta|\mathbf{x}}\left[\left(\theta - E_{\theta|\mathbf{x}}(\theta)\right)^2\right] \geq \frac{E_{\theta|\mathbf{x}}\left[\left(\theta - E_{\theta|\mathbf{x}}(\theta)\right)\phi(\mathbf{x},\theta)\right]^2}{E_{\theta|\mathbf{x}}\left[\phi(\mathbf{x},\theta)^2\right]}.$$

Hence for a.e. $\mathbf{x} \in \mathcal{X}$:

$$E_{\theta|\mathbf{x}}\left[\left(\theta - \widehat{\theta}(\mathbf{x})\right)^2\right] \geq \frac{E_{\theta|\mathbf{x}}\left[\left(\theta - E_{\theta|\mathbf{x}}(\theta)\right)\phi(\mathbf{x},\theta)\right]^2}{E_{\theta|\mathbf{x}}\left[\phi(\mathbf{x},\theta)^2\right]}.$$

Moreover, if $\phi(\mathbf{x},\theta)$ is a BLB-generating function of the WWF, i.e. $E_{\theta|\mathbf{x}}[\phi(\mathbf{x},\theta)] = 0$ for a.e. $\mathbf{x} \in \mathcal{X}$, therefore for a.e. $\mathbf{x} \in \mathcal{X}$:

$$E_{\theta|\mathbf{x}}\left[\left(\theta - \widehat{\theta}(\mathbf{x})\right)^2\right] \geq \frac{E_{\theta|\mathbf{x}}\left[\theta\phi(\mathbf{x},\theta)\right]^2}{E_{\theta|\mathbf{x}}\left[\phi(\mathbf{x},\theta)^2\right]},$$

which leads to:

$$E_{\mathbf{x},\theta}\left[\left(\theta - \widehat{\theta}(\mathbf{x})\right)^2\right] \geq E_{\mathbf{x}}\left[\frac{E_{\theta|\mathbf{x}}\left[\theta\phi(\mathbf{x},\theta)\right]^2}{E_{\theta|\mathbf{x}}\left[\phi(\mathbf{x},\theta)^2\right]}\right]. \quad (2)$$

In the particular case where $\phi(\mathbf{x},\theta) = \partial \ln p(\theta|\mathbf{x})/\partial\theta$, under some mild regularity conditions (see [24, Section III]), $E_{\theta|\mathbf{x}}[\theta\phi(\mathbf{x},\theta)] = 1$ and (2) leads to the TBCRB:

$$E_{\mathbf{x},\theta}\left[\left(\theta - \widehat{\theta}(\mathbf{x})\right)^2\right] \geq E_{\mathbf{x}}\left[E_{\theta|\mathbf{x}}\left[\left(\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta}\right)^2\right]^{-1}\right]. \quad (3a)$$

Indeed, by the Jensen's inequality [32], one obtains that:

$$E_{\mathbf{x},\theta}\left[\left(\theta - \widehat{\theta}(\mathbf{x})\right)^2\right] \geq E_{\mathbf{x}}\left[E_{\theta|\mathbf{x}}\left[\left(\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta}\right)^2\right]^{-1}\right] \geq E_{\mathbf{x},\theta}\left[\left(\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta}\right)^2\right]^{-1}, \quad (3b)$$

where the right-hand side of the above inequality is the "standard" BCRB, more well-known as:

$$\mathrm{BCRB} = E_{\mathbf{x},\theta}\left[\left(\frac{\partial \ln p(\mathbf{x},\theta)}{\partial\theta}\right)^2\right]^{-1}. \quad (3c)$$

II-A. An updated class of efficient estimators

The covariance inequality (3a) becomes an equality [32] if the a posteriori pdf $p(\theta|\mathbf{x})$ satisfies $\partial \ln p(\theta|\mathbf{x})/\partial\theta = k(\mathbf{x})\left(\widehat{\theta}(\mathbf{x}) - \theta\right)$ for a.e. $\mathbf{x} \in \mathcal{X}$, that is:

$$\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta} = E_{\theta|\mathbf{x}}\left[\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta}^2\right]\left(\widehat{\theta}(\mathbf{x}) - \theta\right). \quad (4a)$$

Let $\partial^2\mu(\theta,\mathbf{x})/\partial^2\theta = E_{\theta|\mathbf{x}}\left[\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta}^2\right] = 1/\sigma^2_{\theta|\mathbf{x}}$. Then (4a) yields:

$$\ln p(\theta|\mathbf{x}) = \frac{\partial\mu(\theta,\mathbf{x})}{\partial\theta}\left(\widehat{\theta}(\mathbf{x}) - \theta\right) + \mu(\theta,\mathbf{x}), \quad (4b)$$

where $\partial\mu(\theta,\mathbf{x})/\partial\theta = \theta/\sigma^2_{\theta|\mathbf{x}} + u(x)$ and $\mu(\theta,\mathbf{x}) = \theta^2/\sigma^2_{\theta|\mathbf{x}}/2 + u(x)\theta + h(\mathbf{x})$. Finally, after some straightforward calculus, it appears that $p(\theta|\mathbf{x})$ solution of (4a) is of the form:

$$p(\theta|\mathbf{x}) = \frac{e^{-\frac{1}{2\sigma^2_{\theta|\mathbf{x}}}\left(\theta - \widehat{\theta}(\mathbf{x})\right)^2}}{\sqrt{2\pi\sigma^2_{\theta|\mathbf{x}}}}, \; \sigma^2_{\theta|\mathbf{x}} = \frac{1}{E_{\theta|\mathbf{x}}\left[\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta}^2\right]}, \quad (4c)$$

which means that the a posteriori probability density of θ must be Gaussian for all $\mathbf{x}$ in order for (3a) to become an equality.
It is an update of the result initially released by Shutzenberger [3] and Van-Trees [5, p73] based on the "standard" covariance inequality:

$$E_{\mathbf{x},\theta}\left[\left(\theta - \widehat{\theta}(\mathbf{x})\right)^2\right] \geq E_{\mathbf{x},\theta}\left[\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta}^2\right]^{-1},$$

which becomes an equality if $p(\theta|\mathbf{x})$ satisfies for a.e. $\mathbf{x} \in \mathcal{X}$:

$$\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta} = k\left(\widehat{\theta}(\mathbf{x}) - \theta\right), \; k = E_{\mathbf{x},\theta}\left[\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta}^2\right],$$

yielding the restricted case where $\sigma^2_{\theta|\mathbf{x}}$ does not depend on $\mathbf{x}$:

$$\sigma^2_{\theta|\mathbf{x}} = E_{\theta|\mathbf{x}}\left[\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta}^2\right]^{-1} = E_{\mathbf{x},\theta}\left[\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta}^2\right]^{-1} = \sigma^2_\theta.$$

In that case $\mathrm{TBCRB} = \mathrm{BCRB}$. If $\sigma^2_{\theta|\mathbf{x}}$ is not constant and does depend on $\mathbf{x}$ then $\mathrm{TBCRB} > \mathrm{BCRB}$. Thus the usual BCRB can not be reached in most of the estimation problems.
As a consequence, it seems appropriate to update the definition of an efficient estimate as follows:

Definition 1: $\widehat{\theta}(\mathbf{x})$ is an efficient estimate of θ if its MSE reaches TBCRB *(3a).*

Furthermore, the updated class of efficient estimators is given by (4c).

III. A NOTEWORTHY EXAMPLE

As in introduced in [2, p7], we wish to estimate the variance of a zero-mean Gaussian random variable. The observation vector $\mathbf{x}$ consists of N i.i.d. Gaussian samples: $\mathbf{x} \sim \mathcal{N}\left(0, \sigma^2\mathbf{I}\right)$. The parameter of interest is $\theta \triangleq \sigma^2$. The conditional pdf is:

$$p(\mathbf{x}|\theta) = (2\pi)^{-\frac{N}{2}}\theta^{-\frac{N}{2}}e^{-\frac{1}{\theta}\frac{\mathbf{x}^T\mathbf{x}}{2}}, \; \mathbf{x} \in \mathbb{R}^N. \quad (5a)$$

We assume the a priori pdf follows a beta distribution of the form:

$$p(\theta) = \beta(a,a)^{-1}\theta^{a-1}(1-\theta)^{a-1}, \; 0 \leq \theta \leq 1, \quad (5b)$$

where: $\beta(a,a) = \int_0^1 \theta^{a-1}(1-\theta)^{a-1}d\theta = \Gamma(a)^2/\Gamma(2a)$, and $\Gamma(.)$ is the gamma function, $\Gamma(a) = \int_0^\infty v^{a-1}e^{-v}dv$. This prior distribution is symmetric with mean $\mu_\theta = \frac{1}{2}$ and variance $\sigma^2_\theta = \frac{1}{4(2a+1)}$. When $a = 1$, the pdf is uniform; as the parameter a

increases, the pdf becomes narrower and, finally, as $a \to \infty$, we approach the known θ case. Combined with (5a), (5b) yields the following joint pdf: $0 \leq \theta \leq 1$, $\mathbf{x} \in \mathbb{R}^N$,

$$p(\mathbf{x}, \theta) = (2\pi)^{-\frac{N}{2}} \beta(a,a)^{-1} \theta^{a-\frac{N}{2}-1} (1-\theta)^{a-1} e^{-\frac{1}{\theta}\frac{\mathbf{x}^T\mathbf{x}}{2}}. \tag{5c}$$

III-A. Known results: BCRB and MAP

From (joint pdf):

$$-\frac{\partial^2 \ln p(\theta|\mathbf{x})}{\partial^2\theta} = -\frac{\partial^2 \ln p(\mathbf{x},\theta)}{\partial^2\theta} = \left(a-1-\frac{N}{2}\right)\theta^{-2} + \theta^{-3}\mathbf{x}^T\mathbf{x} + (a-1)(1-\theta)^{-2}. \tag{6}$$

Therefore the "standard" Bayesian Fisher Information Matrix (BFIM) can be easily computed as:

$$\begin{aligned} F &= E_{\mathbf{x},\theta}\left[\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta}^2\right] = E_{\mathbf{x},\theta}\left[-\frac{\partial^2 \ln p(\theta|\mathbf{x})}{\partial^2\theta}\right] \\ &= E_\theta\left[E_{\mathbf{x}|\theta}\left[\frac{a-1-\frac{N}{2}}{\theta^2} + \frac{\mathbf{x}^T\mathbf{x}}{\theta^3} + \frac{a-1}{(1-\theta)^2}\right]\right] \end{aligned}$$

where $E_{\mathbf{x}|\theta}\left[\mathbf{x}^T\mathbf{x}\right] = tr\left(E_{\mathbf{x}|\theta}\left[\mathbf{x}\mathbf{x}^T\right]\right) = N\theta$. Thus:

$$F = \left(a-1+\frac{N}{2}\right) E_\theta\left[\theta^{-2}\right] + (a-1) E_\theta\left[(1-\theta)^{-2}\right],$$

where [2, p10]:

$$E_\theta\left[\theta^{-2}\right] = \frac{\beta(a-2,a)}{\beta(a,a)}, E_\theta\left[(1-\theta)^{-2}\right] = \frac{\beta(a,a-2)}{\beta(a,a)},$$

leading to:

$$BCRB = \frac{1}{F},\ F = (N+4(a-1))\frac{\Gamma(a-2)\Gamma(2a)}{2\Gamma(2a-2)\Gamma(a)}. \tag{7a}$$

Last, if $a > 2$, then (7a) reduces to:

$$BCRB = \frac{1}{F},\ F = (N+4(a-1))\frac{2a-1}{a-2}, \quad a > 2. \tag{7b}$$

The MAP estimator,

$$\widehat{\theta}_{MAP}(\mathbf{x}) = \arg\max_{0\leq\theta\leq1}\{p(\theta|\mathbf{x})\} = \arg\max_{0\leq\theta\leq1}\{p(\mathbf{x},\theta)\},$$

is given by [2, p9]:

$$\begin{aligned} N &\neq 4(a-1) : \widehat{\theta}_{MAP}(\mathbf{x}) = \frac{-\delta-\sqrt{\delta^2-4\alpha\gamma}}{2\alpha} && (8a) \\ N &= 4(a-1) : \widehat{\theta}_{MAP}(\mathbf{x}) = \frac{\gamma}{\frac{1}{2}+\gamma} && (8b) \end{aligned}$$

where: $\alpha = 1 - \frac{4(a-1)}{N}$, $\delta = -\left(1-\frac{2(a-1)}{N}+\gamma\right)$, $\gamma = \frac{\mathbf{x}^T\mathbf{x}}{N}$.

III-B. New results: $p(\mathbf{x})$, $p(\theta|\mathbf{x})$, $E_{\theta|\mathbf{x}}[\theta]$, MMSE

From [33, 3.471(2.)], $\forall\mu, \beta \in \mathbb{C} \mid \mathrm{Re}\{\mu\} > 0, \mathrm{Re}\{\beta\} > 0$:

$$\int_0^1 \theta^{\nu-1}(1-\theta)^{\mu-1} e^{-\frac{\beta}{\theta}} d\theta = \beta^{\frac{\nu-1}{2}} e^{-\frac{\beta}{2}} \Gamma(\mu) W_{\frac{1-2\mu-\nu}{2},\frac{\nu}{2}}(\beta), \tag{9}$$

where $W_{\kappa,\mu}(z)$ is a Whittaker function [33, §9.22-9.23]. By applying (9) to:

$$p(\mathbf{x}) = (2\pi)^{-\frac{N}{2}} \beta(a,a)^{-1} \int_0^1 \theta^{\left(a-\frac{N}{2}\right)-1} (1-\theta)^{a-1} e^{-\frac{1}{\theta}\frac{\mathbf{x}^T\mathbf{x}}{2}} d\theta,$$

one obtains the following closed-form expressions:

$$p(\mathbf{x}) = \frac{\Gamma(a)\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)^{\frac{a-\frac{N}{2}-1}{2}} e^{-\frac{\mathbf{x}^T\mathbf{x}}{4}}}{\sqrt{2\pi}^N \beta(a,a)} W_{\frac{\frac{N}{2}-3a+1}{2},\frac{a-\frac{N}{2}}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right), \tag{10a}$$

$$p(\theta|\mathbf{x}) = \frac{\theta^{a-\frac{N}{2}-1}(1-\theta)^{a-1} e^{-\frac{1}{\theta}\frac{\mathbf{x}^T\mathbf{x}}{2}}}{\Gamma(a)\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)^{\frac{a-\frac{N}{2}-1}{2}} e^{-\frac{\mathbf{x}^T\mathbf{x}}{4}} W_{\frac{\frac{N}{2}-3a+1}{2},\frac{a-\frac{N}{2}}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)} \tag{10b}$$

Similarly, the minimum MSE (MMSE) estimator is given by:

$$\begin{aligned} E_{\theta|\mathbf{x}}[\theta] &= \int_0^1 \theta p(\theta|\mathbf{x}) d\theta = \frac{1}{p(\mathbf{x})}\int_0^1 \theta p(\mathbf{x},\theta) d\theta \\ &= \frac{\int_0^1 \theta^{\left(a-\frac{N}{2}+1\right)-1}(1-\theta)^{a-1} e^{-\frac{1}{\theta}\frac{\mathbf{x}^T\mathbf{x}}{2}} d\theta}{p(\mathbf{x})\sqrt{2\pi}^N \beta(a,a)} \\ &= \sqrt{\frac{\mathbf{x}^T\mathbf{x}}{2}} \frac{W_{\frac{\frac{N}{2}-3a}{2},\frac{a-\frac{N}{2}+1}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)}{W_{\frac{\frac{N}{2}-3a+1}{2},\frac{a-\frac{N}{2}}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)} \end{aligned} \tag{11}$$

and the MMSE can be computed as:

$$MMSE = E_{\mathbf{x},\theta}\left[\left(E_{\theta|\mathbf{x}}[\theta]-\theta\right)^2\right] = E_{\mathbf{x}}\left[E_{\theta|\mathbf{x}}\left[\theta^2\right] - E_{\theta|\mathbf{x}}[\theta]^2\right] \tag{12a}$$

where:

$$\begin{aligned} E_{\theta|\mathbf{x}}\left[\theta^2\right] &= \int_0^1 \theta^2 p(\theta|\mathbf{x}) d\theta = \frac{1}{p(\mathbf{x})}\int_0^1 \theta^2 p(\mathbf{x},\theta) d\theta \\ &= \frac{\int_0^1 \theta^{\left(a-\frac{N}{2}+2\right)-1}(1-\theta)^{a-1} e^{-\frac{1}{\theta}\frac{\mathbf{x}^T\mathbf{x}}{2}} d\theta}{p(\mathbf{x})\sqrt{2\pi}^N \beta(a,a)} \\ &= \frac{\mathbf{x}^T\mathbf{x}}{2} \frac{W_{\frac{\frac{N}{2}-3a-1}{2},\frac{a-\frac{N}{2}+2}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)}{W_{\frac{\frac{N}{2}-3a+1}{2},\frac{a-\frac{N}{2}}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)} \end{aligned} \tag{12b}$$

Thus:

$$MMSE = E_{\mathbf{x}}\left[\frac{\mathbf{x}^T\mathbf{x}}{2}\left(\frac{W_{\frac{\frac{N}{2}-3a-1}{2},\frac{a-\frac{N}{2}+2}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)}{W_{\frac{\frac{N}{2}-3a+1}{2},\frac{a-\frac{N}{2}}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)} - \left(\frac{W_{\frac{\frac{N}{2}-3a}{2},\frac{a-\frac{N}{2}+1}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)}{W_{\frac{\frac{N}{2}-3a+1}{2},\frac{a-\frac{N}{2}}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)}\right)^2\right)\right]. \tag{12c}$$

Last, by considering the change of variable $t = \frac{\mathbf{x}^T\mathbf{x}}{2}$, one obtains:

$$MMSE = \frac{\Gamma(2a)}{\Gamma(a)\Gamma\left(\frac{N}{2}\right)}\int_0^\infty \left(W_{\frac{\frac{N}{2}-3a-1}{2},\frac{a-\frac{N}{2}+2}{2}}(t) - \frac{\left(W_{\frac{\frac{N}{2}-3a}{2},\frac{a-\frac{N}{2}+1}{2}}(t)\right)^2}{W_{\frac{\frac{N}{2}-3a+1}{2},\frac{a-\frac{N}{2}}{2}}(t)}\right) t^{\frac{a+\frac{N}{2}-1}{2}} e^{-\frac{1}{2}t} dt. \tag{12d}$$

III-C. New results (cont): TBCRB

Since the tighter BCRB is defined by (3a):

$$\text{TBCRB} = E_{\mathbf{x}}\left[E_{\theta|\mathbf{x}}\left[\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta}^2\right]^{-1}\right],$$

$$E_{\theta|\mathbf{x}}\left[\frac{\partial \ln p(\theta|\mathbf{x})}{\partial\theta}^2\right] = -E_{\theta|\mathbf{x}}\left[\frac{\partial^2 \ln p(\theta|\mathbf{x})}{\partial^2\theta}\right],$$

we consider first the computation of the posterior BFIM $F_{\mathbf{x}} = E_{\theta|\mathbf{x}}\left[-\frac{\partial^2 \ln p(\theta|\mathbf{x})}{\partial^2\theta}\right]$, which, from (6), can be expressed as:

$$F_{\mathbf{x}} = \left(a - 1 - \frac{N}{2}\right) E_{\theta|\mathbf{x}}\left[\theta^{-2}\right] + \mathbf{x}^T\mathbf{x} E_{\theta|\mathbf{x}}\left[\theta^{-3}\right] + (a-1) E_{\theta|\mathbf{x}}\left[(1-\theta)^{-2}\right], \quad (13)$$

where, by using a similar approach as in (11) and (12b):

$$E_{\theta|\mathbf{x}}\left[\theta^{-2}\right] = \frac{\int_0^1 \theta^{\left(a-\frac{N}{2}-2\right)-1}(1-\theta)^{a-1} e^{-\frac{1}{\theta}\frac{\mathbf{x}^T\mathbf{x}}{2}} d\theta}{p(\mathbf{x})\sqrt{2\pi}^N \beta(a,a)} = \left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)^{-1} \frac{W_{\frac{\frac{N}{2}-3a+3}{2}, \frac{a-\frac{N}{2}-2}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)}{W_{\frac{\frac{N}{2}-3a+1}{2}, \frac{a-\frac{N}{2}}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)},$$

$$E_{\theta|\mathbf{x}}\left[\theta^{-3}\right] = \frac{\int_0^1 \theta^{\left(a-\frac{N}{2}-3\right)-1}(1-\theta)^{a-1} e^{-\frac{1}{\theta}\frac{\mathbf{x}^T\mathbf{x}}{2}} d\theta}{p(\mathbf{x})\sqrt{2\pi}^N \beta(a,a)} = \left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)^{\frac{-3}{2}} \frac{W_{\frac{\frac{N}{2}-3a+4}{2}, \frac{a-\frac{N}{2}-3}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)}{W_{\frac{\frac{N}{2}-3a+1}{2}, \frac{a-\frac{N}{2}}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)},$$

$$E_{\theta|\mathbf{x}}\left[(1-\theta)^{-2}\right] = \frac{\int_0^1 \theta^{\left(a-\frac{N}{2}\right)-1}(1-\theta)^{(a-2)-1} e^{-\frac{1}{\theta}\frac{\mathbf{x}^T\mathbf{x}}{2}} d\theta}{p(\mathbf{x})\sqrt{2\pi}^N \beta(a,a)} = \frac{\Gamma(a-2)}{\Gamma(a)} \frac{W_{\frac{\frac{N}{2}-3a+5}{2}, \frac{a-\frac{N}{2}}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)}{W_{\frac{\frac{N}{2}-3a+1}{2}, \frac{a-\frac{N}{2}}{2}}\left(\frac{\mathbf{x}^T\mathbf{x}}{2}\right)}.$$

Second, as $\text{TBCRB} = E_{\mathbf{x}}\left[\frac{1}{F_{\mathbf{x}}}\right]$, it is judicious to consider once again the change of variable $t = \frac{\mathbf{x}^T\mathbf{x}}{2}$, which leads to:

$$\text{TBCRB} = \frac{\Gamma(2a)}{\Gamma(a)\Gamma\left(\frac{N}{2}\right)} \times \int_0^\infty \frac{\left(W_{\frac{\frac{N}{2}-3a+1}{2}, \frac{a-\frac{N}{2}}{2}}(t)\right)^2 t^{\frac{a+\frac{N}{2}-1}{2}} e^{-\frac{1}{2}t}}{\left(\begin{array}{l}\left(a-1-\frac{N}{2}\right) W_{\frac{\frac{N}{2}-3a+3}{2}, \frac{a-\frac{N}{2}-2}{2}}(t) + \\ 2\sqrt{t} W_{\frac{\frac{N}{2}-3a+4}{2}, \frac{a-\frac{N}{2}-3}{2}}(t) + \\ (a-1)\frac{\Gamma(a-2)}{\Gamma(a)} t W_{\frac{\frac{N}{2}-3a+5}{2}, \frac{a-\frac{N}{2}}{2}}(t)\end{array}\right)} dt \quad (14)$$

III-D. Simulation results

To illustrate the known and new results, we ran a Monte-Carlo simulation for various value of N, i.e. $N \in \{2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048\}$, and a shape parameter $a = 3$ of the symmetric beta prior density (5b). In figure 1 we show the results from $20,000$ trials. The root MSE (RMSE) of the MAP estimate (8a) and of the MMSE estimate (11) averaged over trials are plotted as a function of N, as well as the square-roots of the bounds, i.e. the BCRB (7a) and the TBCRB (14), and the square-root of the theoretical value of the MMSE (12d). Note that the TBCRB (14) and the theoretical value of the MMSE (12d) have been computed with Mathematica, whereas the MAP estimate (8a) and the MMSE estimate (11) have been simulated with Matlab, which can not compute the Whittaker function $W_{\kappa,\mu}(z)$ when $N \geq 110$. As a consequence the comparison of the theoretical and empirical values of the MMSE is only available for $N \leq 64$, where the two values exhibit a perfect match, which validates (12d). Figures 1 clearly shows that the MMSE estimate (as well as the MAP estimate) is efficient in the large sample regime, provided that the updated definition of efficiency is used, that is relatively to the TBCRB (14) and not relatively to the BCRB (7a). Moreover, as expected, as N decreases and the prior information dominates the observation, both the MAP and the MMSE estimates converge to the a priori variance ($\sigma_\theta = 0.5/\sqrt{7}$), and the TBCRB converges to the BCRB.

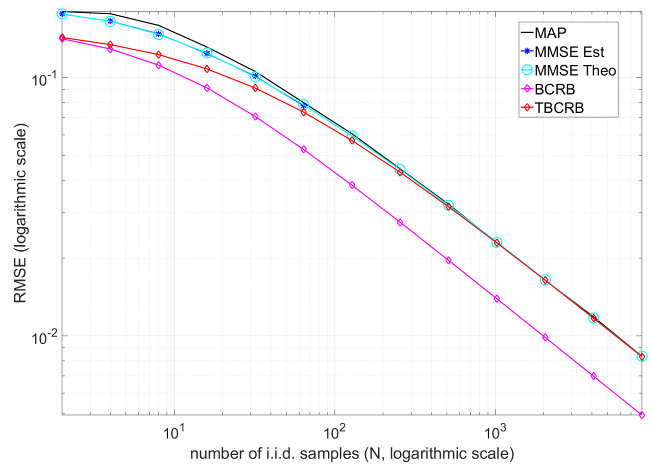

Fig. 1. RMSE, BCRB and TBCRB in estimating $\theta = \sigma^2$ versus N. $a = 3$.

IV. CONCLUSION

In [31] it has been shown that any "standard" BLBs of the WWF admits a "tighter" form which upper bounds the "standard" form (1). However the potential gain in tightness offered by the TWWF has not been quantified so far. By providing an illustrative example of the potential gain in tightness (of course, depending on the estimation problem), this paper lays the foundation to revisit some Bayesian estimation problems where the "standard" BCRB is not tight in the asymptotic region.

V. REFERENCES

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

[2] 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.

[3] M. P Shutzenberger," A generalization of the Frechet-Cramér inequality in the case of Bayes estimation," Bulletin of the AMS, 63(142), 1957

[4] J. J. Gart, "An extension of the Cramér-Rao inequality", Ann. Math. Stat., 30(2): 367-380, 1959

[5] H.L. Van Trees, *Detection, Estimation and Modulation Theory, Part 1*, New York, Wiley, 1968
[6] J. Ziv and M. Zakai, ”Some lower bounds on signal parameter estimation,” IEEE Trans. on IT, 15:386-391, 1969
[7] S. Bellini and G. Tartara, ”Bounds on error in signal parameter estimation,” IEEE Trans. on Com., 22:340-342, 1974
[8] K. L. Bell, Y. Steinberg, Y. Ephraim, and H. L. Van Trees, ”Extended Ziv-Zakaï lower bound for vector parameter estimation”, IEEE Trans. on IT, 43(2): 624-637, 1997
[9] B. Z. Bobrovsky and M. Zakai, ”A lower bound on the estimation error for certain diffusion processes,” IEEE Trans. on IT, 22:45-52, 1976.
[10] A. J. Weiss and E. Weinstein, ”A lower bound on the mean square error in random parameter estimation”, IEEE Trans. on IT, 31(5): 680-682, 1985.
[11] B.Z. Bobrovsky, E. Mayer-Wolf and M. Zakai, ”Some Classes of Global Cramér-Rao Bounds”, The Annals of Statistics, 15(4): 1421-1438, 1987
[12] E. Weinstein and A.J. Weiss, ”A general class of lower bounds in parameter estimation”, IEEE Trans. on IT, 34(2): 338-342, 1988
[13] M. Ghosh, ”Cramér-Rao bounds for posterior variances”, Statistics & Probability Letters, 17(3): 173-178, 1993
[14] 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.
[15] P. Tichavsky, C. H. Muravchik, and A. Nehorai, ”Posterior Cramér-Rao bounds for discrete-time nonlinear filtering”, IEEE Trans. on SP, 46(5): 1386-1396, 1998
[16] S. Reece and D. Nicholson, ”Tighter alternatives to the Cramér- Rao lower bound for discrete-time filtering”, in Proc. of Int. Conf. on Information Fusion, 1: 1-6, 2005
[17] A. Renaux, P. Forster, P. Larzabal, and C. Richmond ”The bayesian Abel bound on the mean square error”, in Proc. of IEEE ICASSP, 2006
[18] I. Rapoport and Y. Oshman, ”Weiss-Weinstein lower bounds for Markovian systems. Part 1 & 2” IEEE Trans. on SP, 55(5): 2016-2042, 2007
[19] A. Renaux, P. Forster, P. Larzabal, C. Richmond, and A. Nehorai, ”A fresh look at the bayesian bounds of the Weiss-Weinstein family”, IEEE Trans. on SP, Volume: 56, Issue: 11, pp. 5334-5352, Nov. 2008
[20] K. Todros and J. Tabrikian, ”General classes of performance lower bounds for parameter estimation - part II: Bayesian bounds”, IEEE Transactions on Information Theory, vol. 56, no. 10, pp. 5064–5082, Oct. 2010.
[21] F. Xaver, P. Gerstoft, G. Matz, and C. F. Mecklenbräuker, ”Analytic Sequential Weiss-Weinstein Bounds”, IEEE Trans. on SP, 61(20): 5049-5062, 2013
[22] D. T. Vu, A. Renaux , R. Boyer and S. Marcos, ”Some results on the Weiss-Weinstein bound for conditional and unconditional signal models in array processing”, Elsevier Signal Processing , 95: 126-148, 2014
[23] A. Weinstein, E. Weinstein, ”Inequalities for the Bayes Risk”, arXiv:1401.5187, 2014
[24] E. Chaumette, A. Renaux, and M. N. El Korso, ”A class of Weiss-Weinstein bounds and its relationship with the Bobrovsky-Mayer-Wolf-Zakai bounds”, IEEE Trans. on IT, 63(4): 1538-1553, 2017
[25] L. Bacharach, A. Renaux, M. N. El Korso, and E. Chaumette, ”Weiss-Weinstein bound on multiple change-points estimation”, IEEE Trans. on SP, May 2017
[26] C. Fritsche and F. Gustafsson, ”Bayesian Bhattacharyya Bound for discrete-time filtering revisited”, in Proc. IEEE Workshop on CAMSAP, 2017
[27] C. Fritsche, U. Orguner, F. Gustafsson, ”Bobrovsky-Zakai Bound for Filtering, Prediction and Smoothing of Nonlinear Dynamic Systems”, in Proc. ISIF FUSION, 2018
[28] E. Nitzan, T. Routtenberg, and J. Tabrikian, ”A New Class of Bayesian Cyclic Bounds for Periodic Parameter Estimation”, IEEE Trans. on SP, 64(1): 229-243, 2016
[29] E. Nitzan, T. Routtenberg, and J. Tabrikian, ”Bobrovsky-Zakai type bound for periodic stochastic filtering”, IEEE SP Letter, 25(10):1460-1464, 2018
[30] E. Nitzan, T. Routtenberg, and J. Tabrikian , ”Multivariate Bayesian Cramér-Rao-Type Bound for Stochastic Filtering Involving Periodic States”, in Proc. ISIF FUSION, 2018
[31] E. Chaumette and C. Fritsche, ”A General Class of Bayesian Lower Bounds Tighter than the Weiss-Weinstein Family”, in Proc. ISIF FUSION, 2018
[32] E. L. Lehmann and G. Casella, *Theory of Point Estimation (2nd ed.)*. Springer, 2011
[33] I.S. Gradshteyn and I.M. Ryzhik, Table of Integrals, Series, and Products (8th Ed), 2015