

# LOWER BOUNDS ON THE MEAN SQUARE ERROR DERIVED FROM MIXTURE OF LINEAR AND NON-LINEAR TRANSFORMATIONS OF THE UNBIASNESS DEFINITION

Eric Chaumette, Alexandre Renaux, Pascal Larzabal

► **To cite this version:**

Eric Chaumette, Alexandre Renaux, Pascal Larzabal. LOWER BOUNDS ON THE MEAN SQUARE ERROR DERIVED FROM MIXTURE OF LINEAR AND NON-LINEAR TRANSFORMATIONS OF THE UNBIASNESS DEFINITION. International Conference on Acoustics, Speech, and Signal Processing, ICASSP-09, 2009, Taipei, Taiwan. 2009. <inria-00444810>

**HAL Id: inria-00444810**

**<https://hal.inria.fr/inria-00444810>**

Submitted on 7 Jan 2010

**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.

# LOWER BOUNDS ON THE MEAN SQUARE ERROR DERIVED FROM MIXTURE OF LINEAR AND NON-LINEAR TRANSFORMATIONS OF THE UNBIASEDNESS DEFINITION

Eric Chaumette<sup>(1)</sup>, Alexandre Renaux<sup>(2)</sup> and Pascal Larzabal<sup>(3)</sup>

<sup>(1)</sup> ONERA - DEMR/TSI, The French Aerospace Lab, Chemin de la Hunière, F-91120 Palaiseau, France

<sup>(2)</sup> Université Paris-Sud 11, L2S, Supelec, 3 rue Joliot Curie, F-91190 Gif-Sur-Yvette, France

<sup>(3)</sup> SATIE, ENS Cachan, CNRS, UniverSud, 61 av President Wilson, F-94230 Cachan, France

## ABSTRACT

It is well known that in non-linear estimation problems the ML estimator exhibits a threshold effect, i.e. a rapid deterioration of estimation accuracy below a certain SNR or number of snapshots. This effect is caused by outliers and is not captured by standard tools such as the Cramér-Rao bound (CRB). The search of the SNR threshold value can be achieved with the help of approximations of the Barankin bound (BB) proposed by many authors. These approximations result from a linear transformation (discrete or integral) of the uniform unbiasedness constraint introduced by Barankin. Nevertheless, non-linear transformations can be used as well for some class of p.d.f. including the Gaussian case. The benefit is their combination with existing linear transformation to get tighter lower bounds improving the SNR threshold prediction.

**Index Terms**— Parameter estimation, mean-square-error bounds, SNR threshold

## 1. INTRODUCTION

Minimal performance bounds allow for calculation of the best performance that may be achieved, in the Mean Square Error (MSE) sense, when estimating a set of model parameters from noisy observations. Historically the first MSE lower bound for deterministic parameters to be derived was the Cramér-Rao Bound (CRB) [8], which has been the most widely used since. Its popularity is largely due to its simplicity of calculation leading to closed form expressions useful for system analysis and design. Additionally, the CRB can be achieved asymptotically (high SNR [6] and/or large number of snapshots [8]) by Maximum Likelihood Estimators (MLE), and last but not least, it is the lowest bound on the MSE of unbiased estimators, since it derives from a local formulation of unbiasedness in the vicinity of the true parameters [3]. This initial characterization of locally unbiased estimators has been improved first by Bhattacharyya's works [8] which refined the characterization of local unbiasedness, and significantly generalized by Barankin works [1], who established the general form of the greatest lower bound on MSE (BB) taking into account a uniform unbiasedness definition (eq. (1)). Unfortunately the BB is the solution of an integral equation

---

This work has been partly funded by the European Network of excellence NEWCOM++ under the number 216715

This project is partially funded by both the Région Ile-de-France and the Digeo Research Park

with a generally incomputable analytic solution (eq. (8)). Therefore, since then, numerous works detailed in [3][7] have been devoted to deriving computable approximations of the BB and 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. These works have also shown that in non-linear estimation problems three distinct regions of operation can be observed. In the asymptotic region, i.e. at a high number of independent snapshots and/or at high SNR, the MSE is small and, in many cases, close to the Small-Error bounds. In the *a priori* performance region where the number of independent snapshots and/or the SNR are very low, the observations provide little information and the MSE is close to that obtained from the prior knowledge about the problem. Between these two extremes, there is a transition region where MSE of MLEs usually deteriorates rapidly with respect to existing MSE lower bounds (Large or Small) and exhibits a threshold behaviour, which corresponds to a "performance breakdown" of the estimators due to the appearance of outliers. 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. On the other hand, large-Error bounds suffer from their computational cost. Indeed, each BB approximation request the search of an optimum over a set of test points and their tightness depends on the chosen set of test points.

And tightness is the matter, since a more accurate knowledge of the BB allows a better prediction of the SNR threshold value.

Therefore, at least two strategies can be adopted. The first one consists in the search for the tightest and computable approximation of the BB, for a given set of test points [3]. The second one consists in exploiting a particular property of a given class of p.d.f. in order to design a lower bound able to reveal the set of test points optimizing its tightness [7]. All these approximations of the BB result from a linear transformation (discrete or integral) of the uniform unbiasedness constraint introduced by Barankin (eq. (1)). Indeed, they are different solutions of the same norm minimization problem under sets of appropriate linear constraints (eq. (4)) and derive from a simple formula (eq. (5)).

Another possible strategy to derive new BB approximations, possibly tighter, is to resort to non-linear transformations of the uniform unbiasedness constraint (eq. (1)). It seems that, since the original idea came from non regular-estimation (class of p.d.f. for which the CRB is the trivial bound 0) [2], its application to regular-estimation (class of p.d.f. for which the CRB is not the trivial bound 0) has been completely overlooked, and has sunk into oblivion.

Therefore the aim of this paper is to bring non-linear transformations out of oblivion. We show that the rationale introduced in [2]

is applicable to Gaussian p.d.f., which is one of the most important regular p.d.f. in signal processing, with application the to single tone threshold analysis. Moreover, the generalization of this rationale allows to introduce a more general class of possible transformations (eq. (12)) of the uniform unbiasedness constraint (eq. (1)), i.e. the mixture of integral linear and non-linear transformations, opening new directions in the search of computable tighter BB approximations improving the SNR threshold prediction.

Last, but not least, the next Section provides an unified simple framework for the derivation of any Barankin bound approximations based on linear transformations, including the last ones introduced in [7].

## 2. BARANKIN BOUND APPROXIMATIONS BY LINEAR TRANSFORMATIONS

For the sake of simplicity we will focus on the estimation of a single real function  $g(\theta)$  of a single unknown real deterministic parameter  $\theta$ . In the following, unless otherwise stated,  $\mathbf{x}$  denotes the random observation vector of dimension  $M$ ,  $\Omega$  the observations space, and  $p(\mathbf{x};\theta)$  the probability density function (p.d.f.) of  $\mathbf{x}$  depending on  $\theta \in \Theta$ , where  $\Theta$  denotes the parameter space. Let  $L^2(\Omega)$  be the real Hilbert space of square integrable functions over  $\Omega$ .

In the search for a lower bound on the MSE of unbiased estimators, two fundamental properties of the problem at hand, introduced by Barankin [1], must be noticed. The first property is that the MSE of a particular estimator  $\widehat{g(\theta^0)}(\mathbf{x}) \in L^2(\Omega)$  of  $g(\theta^0)$ , where  $\theta^0$  is a selected value of the parameter  $\theta$ , is a norm associated with a particular scalar product  $\langle \cdot | \cdot \rangle_{\theta^0}$ :

$$\begin{aligned} MSE_{\theta^0} \left[ \widehat{g(\theta^0)} \right] &= \left\| \widehat{g(\theta^0)}(\mathbf{x}) - g(\theta^0) \right\|_{\theta^0}^2, \\ \langle g(\mathbf{x}) | h(\mathbf{x}) \rangle_{\theta^0} &= E_{\theta^0} [g(\mathbf{x})^* h(\mathbf{x})]. \end{aligned}$$

The second property is that an unbiased estimator  $\widehat{g(\theta^0)}(\mathbf{x})$  of  $g(\theta^0)$  should be uniformly unbiased, i.e. for all possible values of the unknown parameter  $\theta \in \Theta$  it must verify:

$$E_{\theta} \left[ \widehat{g(\theta^0)}(\mathbf{x}) \right] = g(\theta) = E_{\theta^0} \left[ \widehat{g(\theta^0)}(\mathbf{x}) \nu(\mathbf{x};\theta) \right], \quad (1)$$

where  $\nu(\mathbf{x};\theta) = \frac{p(\mathbf{x};\theta)}{p(\mathbf{x};\theta^0)}$  denotes the Likelihood Ratio (LR). As a consequence, the locally-best (at  $\theta^0$ ) unbiased estimator is the solution of a norm minimization under linear constraints

$$\min \left\{ MSE_{\theta^0} \left[ \widehat{g(\theta^0)} \right] \right\} \text{ under } E_{\theta^0} \left[ \widehat{g(\theta^0)}(\mathbf{x}) \nu(\mathbf{x};\theta) \right] = g(\theta),$$

solution that can be obtained by using the norm minimization lemma

$$\begin{aligned} \min \left\{ \mathbf{u}^H \mathbf{u} \text{ under } \mathbf{c}_k^H \mathbf{u} = v_k, 1 \leq k \leq K \right\} &= \mathbf{v}^H \mathbf{G}^{-1} \mathbf{v} \\ \mathbf{u}_{opt} &= \sum_{k=1}^K \alpha_k \mathbf{c}_k, \quad \boldsymbol{\alpha} = \mathbf{G}^{-1} \mathbf{v}, \quad \mathbf{G}_{n,k} = \mathbf{c}_n^H \mathbf{c}_k \end{aligned} \quad (2)$$

Unfortunately, as shown hereinafter, if  $\Theta$  contains a continuous subset of  $\mathbb{R}$ , then the norm minimization under a set of an infinite number of linear constraints (1) leads to an integral equation (8) with no analytical solution in general. Therefore, since the original work of Barankin [1], many studies [3, and references therein][7] have been dedicated to the derivation of "computable" lower bounds approximating the MSE of the locally-best unbiased estimator (BB). All these approximations derive from sets of discrete or integral linear transform of the "Barankin" constraint (1), and accordingly of the LR, and can be obtained using the following simple rationale.

Let  $\boldsymbol{\theta}^N = (\theta^1, \dots, \theta^N)^T \in \Theta^N$  be a vector of  $N$  test points,  $\nu(\mathbf{x};\boldsymbol{\theta}^N) = (\nu(\mathbf{x};\theta^1), \dots, \nu(\mathbf{x};\theta^N))^T$  be the vector of LR associated to  $\boldsymbol{\theta}^N$ ,  $\xi(\theta) = g(\theta) - g(\theta^0)$  and  $\boldsymbol{\xi}(\boldsymbol{\theta}^N) = (\xi(\theta^1), \dots, \xi(\theta^N))^T$ . Any unbiased estimator  $\widehat{g(\theta^0)}(\mathbf{x})$  verifying (1) must comply with

$$E_{\theta^0} \left[ \left( \widehat{g(\theta^0)}(\mathbf{x}) - g(\theta^0) \right) \nu(\mathbf{x};\boldsymbol{\theta}^N) \right] = \boldsymbol{\xi}(\boldsymbol{\theta}^N), \quad (3)$$

and with any subsequent linear transformation of (3). Therefore, any given set of  $K$  ( $K \leq N$ ) independent linear transformations of (3):

$$E_{\theta^0} \left[ \left( \widehat{g(\theta^0)}(\mathbf{x}) - g(\theta^0) \right) \mathbf{h}_k^T \nu(\mathbf{x};\boldsymbol{\theta}^N) \right] = \mathbf{h}_k^T \boldsymbol{\xi}(\boldsymbol{\theta}^N), \quad (4)$$

$\mathbf{h}_k \in \mathbb{R}^N$ ,  $k \in [1, K]$ , provides with a lower bound on the MSE (2):

$$MSE_{\theta^0} \left[ \widehat{g(\theta^0)} \right] \geq \boldsymbol{\xi}(\boldsymbol{\theta}^N)^T \widetilde{\mathbf{G}}_{\mathbf{H}_K} \boldsymbol{\xi}(\boldsymbol{\theta}^N), \quad (5)$$

where  $\widetilde{\mathbf{G}}_{\mathbf{H}_K} = \mathbf{H}_K (\mathbf{H}_K^T \mathbf{R}_\nu \mathbf{H}_K)^{-1} \mathbf{H}_K^T$ ,  $\mathbf{H}_K = [\mathbf{h}_1 \dots \mathbf{h}_K]$  and  $(\mathbf{R}_\nu)_{n,m} = E_{\theta^0} [\nu(\mathbf{x};\theta^n) \nu(\mathbf{x};\theta^m)]$ . The BB is obtained by taking the supremum of (5) over all the existing degrees of freedom ( $N, \boldsymbol{\theta}^N, K, \mathbf{H}_K$ ). Moreover, for a given vector of test points  $\boldsymbol{\theta}^N$ , the lower bound (5) reaches its maximum iff the matrix  $\mathbf{H}_K$  is invertible ( $K = N$ ) [5], which represents a bijective transformation of the set of the  $N$  initial constraints (3):

$$MSE_{\theta^0} \left[ \widehat{g(\theta^0)} \right] \geq \boldsymbol{\xi}(\boldsymbol{\theta}^N)^T \widetilde{\mathbf{G}}_{\mathbf{I}_N} \boldsymbol{\xi}(\boldsymbol{\theta}^N) \geq \boldsymbol{\xi}(\boldsymbol{\theta}^N)^T \widetilde{\mathbf{G}}_{\mathbf{H}_K} \boldsymbol{\xi}(\boldsymbol{\theta}^N),$$

where  $\mathbf{I}_N$  is the identity matrix with dimension  $N$ . All known bounds on the MSE deriving from the Barankin Bound is a particular implementation of (5), including the most general formalism introduced lately in [7]. Indeed, the limit of (4) where  $N \rightarrow \infty$  and  $\boldsymbol{\theta}^N$  uniformly samples  $\Theta$  leads to the linear integral constraint:

$$E_{\theta^0} \left[ \left( \widehat{g(\theta^0)}(\mathbf{x}) - g(\theta^0) \right) \eta(\mathbf{x}, \tau) \right] = \Gamma_h(\tau), \quad (6)$$

$$\eta(\mathbf{x}, \tau) = \int_{\Theta} h(\tau, \theta) \nu(\mathbf{x};\theta) d\theta, \quad \Gamma_h(\tau) = \int_{\Theta} h(\tau, \theta) \xi(\theta) d\theta,$$

where each  $\mathbf{h}_k = (h(\tau_k, \theta^1), \dots, h(\tau_k, \theta^N))^T$  is the vector of samples of a parametric function  $h(\tau, \theta)$ ,  $\tau \in \Lambda \subset \mathbb{R}$ , integrable over  $\Theta$ ,  $\forall \tau \in \Lambda$ . Then, for any subset of  $K$  values of  $\tau$ ,  $\{\tau_k\}_{k \in [1, K]}$ , the subset of the associated  $K$  linear integral constraints (6) leads to the following lower bound (2):

$$\lim_{K \rightarrow \infty} \left\{ \begin{aligned} MSE_{\theta^0} \left[ \widehat{g(\theta^0)}(\mathbf{x}) \right] &\geq MSE_{\theta^0} \left[ \widehat{g(\theta^0)}_{lmvu}(\mathbf{x}) \right] \\ MSE_{\theta^0} \left[ \widehat{g(\theta^0)}_{lmvu}(\mathbf{x}) \right] &= \boldsymbol{\Gamma}_h^T \mathbf{R}_\eta^{-1} \boldsymbol{\Gamma}_h = \boldsymbol{\Gamma}_h^T \left( \frac{\boldsymbol{\alpha}}{\lambda} \right) \\ \widehat{g(\theta^0)}_{lmvu}(\mathbf{x}) - g(\theta^0) &= \sum_{k=1}^K \frac{\alpha_k}{\lambda} \eta(\mathbf{x}, \tau_k) \\ \mathbf{R}_\eta \left( \frac{\boldsymbol{\alpha}}{\lambda} \right) &= \boldsymbol{\Gamma}_h \end{aligned} \right. \quad (7)$$

where  $(\mathbf{R}_\eta)_{k,k'} = E_{\theta^0} [\eta(\mathbf{x}, \tau_k) \eta(\mathbf{x}, \tau_{k'})]$  and  $(\boldsymbol{\Gamma}_h)_k = \Gamma(\tau_k)$ . Therefore, when  $K \rightarrow \infty$  and the set  $\{\tau_k\}_{k \in [1, K]}$  uniformly samples  $\Lambda$ , by setting  $\frac{1}{\lambda} = d\tau = \tau_{k+1} - \tau_k$ ,  $\boldsymbol{\beta} = \frac{\boldsymbol{\alpha}}{\lambda}$ , the integral form of the above lower bound appears straightforwardly:

$$\left\{ \begin{aligned} MSE_{\theta^0} \left[ \widehat{g(\theta^0)}_{lmvu}(\mathbf{x}) \right] &= \int_{\Lambda} \Gamma_h(\tau) \beta(\tau) d\tau \\ \widehat{g(\theta^0)}_{lmvu}(\mathbf{x}) - g(\theta^0) &= \int_{\Lambda} \eta(\mathbf{x}, \tau) \beta(\tau) d\tau \\ \int_{\Lambda} K_h(\tau', \tau) \beta(\tau) d\tau &= \Gamma_h(\tau') \end{aligned} \right. \quad (8)$$

$$\begin{aligned}
K_h(\tau, \tau') &= E_{\theta^0} [\eta(\mathbf{x}, \tau) \eta(\mathbf{x}, \tau')] \\
&= \int_{\Theta} \int_{\Theta} h(\tau, \theta) R_{\nu}(\theta, \theta') h(\tau', \theta') d\theta d\theta', \\
R_{\nu}(\theta, \theta') &= E_{\theta^0} \left[ \frac{p(\mathbf{x}; \theta)}{p(\mathbf{x}; \theta^0)} \frac{p(\mathbf{x}; \theta')}{p(\mathbf{x}; \theta^0)} \right] = \int_{\Omega} \frac{p(\mathbf{x}; \theta) p(\mathbf{x}; \theta')}{p(\mathbf{x}; \theta^0)^2} d\mathbf{x},
\end{aligned}$$

which is exactly the main result introduced in [7] and is a generalization of the Kiefer Bound [4] ( $K = 2$ ). Note that if  $h(\tau, \theta) = \delta(\tau - \theta)$  (limit case of  $\mathbf{H}_N = \mathbf{I}_N$  where  $N = K \rightarrow \infty$ ) then  $K_h(\tau, \tau') = R_{\nu}(\tau, \tau')$  and (8) becomes the simplest expression of the exact Barankin Bound [3, (10)]. As mentioned above, in most practical cases, it is impossible to find either the limit of (7) or an analytical solution of (8) to obtain an explicit form of the exact Barankin Bound on the MSE, which somewhat limits its interest. Nevertheless this formalism allows to use discrete (4) or integral (6) linear transforms of the LR, possibly non-invertible, possibly optimized for a set of p.d.f. (such as the Fourier transform in [7]) in order to get a tight approximation of the BB.

### 3. BARANKIN BOUND APPROXIMATIONS BY NON-LINEAR TRANSFORMATIONS

At the opposite, the use of a non-linear transformation of the unbiasedness definition (1) of type

$$E_{\theta^0} \left[ \widehat{g(\theta^0)}(\mathbf{x}) t(\nu(\mathbf{x}; \theta)) \right] = h(g(\theta)) \quad (9)$$

is more obscure since it seems a difficult mathematical task to compute the bias transformation function  $h(\cdot)$  as a function of the LR transformation function  $t(\cdot)$  and of the LR. Nevertheless there is a class of estimation problems where non-linear transformations of the LR can be used to derive new lower bounds on the MSE. It is the class of estimation problems characterized by a p.d.f.  $p(\mathbf{x}; \theta)$  for which there exists at least one real valued function  $t(\cdot)$  such that, the transformation of p.d.f.  $p(\mathbf{x}; \theta)$  by  $t(\cdot)$  is still - up to a normalization constant, w.r.t.  $\mathbf{x}$ ,  $k(\theta, t)$  - a p.d.f. of the form  $p(\mathbf{x}; \cdot)$  but parameterized by a modified parameter value  $\gamma$ , function of the initial parameter  $\theta$  and of the transformation  $t(\cdot)$ :

$$t(p(\mathbf{x}; \theta)) = k(\theta, t) p(\mathbf{x}; \gamma(\theta, t)), \quad k(\theta, t) = \int_{\Omega} t(p(\mathbf{x}; \theta)) d\mathbf{x} \quad (10)$$

Then an unbiased estimator verifying (1) verifies as well,  $\forall \theta \in \Theta$ :

$$\begin{aligned}
\int_{\Omega} \widehat{g(\theta^0)}(\mathbf{x}) t(p(\mathbf{x}; \theta)) d\mathbf{x} &= k(\theta, t) \int_{\Omega} \widehat{g(\theta^0)}(\mathbf{x}) p(\mathbf{x}; \gamma(\theta, t)) d\mathbf{x} \\
&= k(\theta, t) g(\gamma(\theta, t))
\end{aligned}$$

what implies,  $\forall \theta \in \Theta$

$$\begin{aligned}
E_{\theta^0} \left[ \left( \widehat{g(\theta^0)}(\mathbf{x}) - g(\theta^0) \right) \frac{t(p(\mathbf{x}; \theta))}{p(\mathbf{x}; \theta^0)} \right] &= \\
&= k(\theta, t) [g(\gamma(\theta, t)) - g(\theta^0)].
\end{aligned}$$

In the most general case, if there exists a set of functions  $t_{\theta}(\cdot)$  verifying (10), then any unbiased estimator also verifies,  $\forall \theta \in \Theta$ :

$$\begin{aligned}
E_{\theta^0} \left[ \left( \widehat{g(\theta^0)}(\mathbf{x}) - g(\theta^0) \right) \frac{t_{\theta}(p(\mathbf{x}; \theta))}{p(\mathbf{x}; \theta^0)} \right] &= \\
&= k(\theta, t_{\theta}) [g(\gamma(\theta, t_{\theta})) - g(\theta^0)].
\end{aligned}$$

Therefore, if we update the definition of  $\nu(\mathbf{x}; \theta)$  and  $\xi(\theta)$  in (6) according to

$$\nu(\mathbf{x}; \theta) = \frac{t_{\theta}(p(\mathbf{x}; \theta))}{p(\mathbf{x}; \theta^0)}, \quad \xi(\theta) = k(\theta, t_{\theta}) [g(\gamma(\theta, t_{\theta})) - g(\theta^0)], \quad (11)$$

all the results released in the previous Section still hold, the linear integral transformation becoming a mixture of linear and non-linear integral transformations:

$$\begin{aligned}
\eta(\mathbf{x}, \tau) &= \int_{\Theta} h(\tau, \theta) \frac{t_{\theta}(p(\mathbf{x}; \theta))}{p(\mathbf{x}; \theta^0)} d\theta, \\
\Gamma_h(\tau) &= \int_{\Theta} h(\tau, \theta) k(\theta, t_{\theta}) [g(\gamma(\theta, t_{\theta})) - g(\theta^0)] d\theta.
\end{aligned} \quad (12)$$

The proposed rationale is a generalization of [2] where the authors has extended the Hammersley-Chapman-Robbins bound (HCRB) and the Bhattacharyya bound (BaB) for a particular non-linear transformation  $t_q(y) = y^q$ ,  $q \in ]0, 1[$ , to overcome a non-regular estimation problem: the estimation of the parameters of a Pearson Type III p.d.f.. Indeed for such a p.d.f., the usual bounds based on linear transformations such as the Cramér-Rao bound (CRB), the HCRB and the BaB yield the trivial bound 0.

At first sight, the proposed rationale does not seem appealing, since a non-linear transformation of type (9) or (10) is unlikely to exist whatever the form of the p.d.f., although the linear transformation of the LR (6) is always possible, however possibly yielding the trivial bound 0. It is probably the reason why the application of the proposed rationale to regular estimation problem has been completely overlooked, even by authors of [2], although it is applicable to Gaussian p.d.f..

#### 3.1. Application to the Gaussian observation model

We focus on  $M$ -dimensional complex circular Gaussian p.d.f.:

$$p(\mathbf{x}; \theta) = p(\mathbf{x}; \mathbf{m}(\theta), \mathbf{C}(\theta)) = \frac{e^{-(\mathbf{x}-\mathbf{m}(\theta))^H \mathbf{C}(\theta)^{-1} (\mathbf{x}-\mathbf{m}(\theta))}}{\pi^M |\mathbf{C}(\theta)|}$$

Then it is worth noticing that the transformation  $t_q(y) = y^q$  can be applied to the observation model resulting from a mixture of deterministic and stochastic signals in presence of Gaussian interference. Indeed, in this case  $\mathbf{m}(\theta) = \mathbf{m}(\varepsilon)$ ,  $\mathbf{C}(\theta) = \Psi(\zeta) \mathbf{C}_s \Psi(\zeta)^H + \mathbf{C}_n$ ,  $\theta = [\varepsilon^T, \zeta^T, \text{vec}(\mathbf{C}_s)^T, \text{vec}(\mathbf{C}_n)^T]^T$  and:

$$\begin{aligned}
t_q(p(\mathbf{x}; \theta)) &= k(\theta, q) p(\mathbf{x}; \gamma(\theta, q)) \\
k(\theta, q) &= \frac{\pi^{M(1-q)}}{q^{qM}} \left| \frac{\mathbf{C}(\theta)}{q} \right|^{1-q} \\
\gamma(\theta, q) &= \left[ \varepsilon^T, \zeta^T, \frac{\text{vec}(\mathbf{C}_s)^T}{q}, \frac{\text{vec}(\mathbf{C}_n)^T}{q} \right]^T
\end{aligned} \quad (13)$$

#### 3.2. Single tone threshold analysis

A reference problem in threshold analysis is the estimation of a single tone  $\eta \in ]-0.5, 0.5[$  for a deterministic observation model:

$$\begin{aligned}
\mathbf{x} &= \alpha \psi(\eta) + \mathbf{n}, \quad \psi(\eta) = [1, \dots, e^{j(M-1)2\pi\eta}]^T \\
p(\mathbf{x}; \theta) &= \frac{e^{-\|\mathbf{x} - \alpha \psi(\eta)\|^2}}{\pi^M}
\end{aligned} \quad (14)$$

In this problem,  $\boldsymbol{\theta} = [\eta, a, \sigma_n^2 = 1]$ ,  $\mathbf{m}(\boldsymbol{\theta}) = \alpha\psi(\eta)$ ,  $\mathbf{C}(\boldsymbol{\theta}) = \mathbf{I}$ ,  $a^2$  is the known SNR ( $a > 0$ ) and  $\mathbf{n}$  is a complex circular Gaussian noise, with zero mean and a known covariance matrix  $\mathbf{C}_n = \mathbf{I}$ . As a consequence:  $k(\boldsymbol{\theta}, q) = \frac{\pi^{M(1-q)}}{q^M}$ ,  $\boldsymbol{\gamma}(\boldsymbol{\theta}, q) = (\eta, a, \frac{1}{q})$  and  $g(\boldsymbol{\gamma}(\boldsymbol{\theta}, q)) = g(\boldsymbol{\theta}) = \eta$ .

Let us now consider the generalization of the HCRB, the simplest approximation of the BB (5) based on 2 test-points  $\boldsymbol{\eta}^2 = (\eta^0 + h, \eta^0 - h)^T$  where  $\mathbf{H}_2 = \mathbf{I}$ , obtained by using the non-linear transformation mentioned above (13). Application of (5) and (11) where  $\nu(\mathbf{x}; \boldsymbol{\theta}) = \frac{p(\mathbf{x}; \boldsymbol{\theta})^q}{p(\mathbf{x}; \boldsymbol{\theta}^0)}$  and  $\xi(\eta) = \frac{\pi^{M(1-q)}}{q^M} (\eta - \eta^0)$  leads to the following lower bound  $HCRB_q$ :

$$MSE_{\eta^0}[\hat{\eta}^0] \geq HCRB_q = \boldsymbol{\xi}(\boldsymbol{\eta}^2)^T \mathbf{R}^{-1} \boldsymbol{\xi}(\boldsymbol{\eta}^2)$$

$$\boldsymbol{\xi}(\boldsymbol{\eta}^2) = \frac{\pi^{M(1-q)}}{q^M} \begin{bmatrix} h \\ -h \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} R_{1,1} & R_{1,2} \\ R_{1,2} & R_{2,2} \end{bmatrix}$$

$$R_{1,1} = E_{\boldsymbol{\theta}^0} \left[ \frac{p(\mathbf{x}; \boldsymbol{\theta}^0 + h)^{2q}}{p(\mathbf{x}; \boldsymbol{\theta}^0)^2} \right] = \alpha e^{\frac{2qa^2}{2q-1}} \|\psi(\eta^0 + h) - \psi(\eta^0)\|^2$$

$$R_{1,2} = E_{\boldsymbol{\theta}^0} \left[ \frac{p(\mathbf{x}; \boldsymbol{\theta}^0 + h)^q p(\mathbf{x}; \boldsymbol{\theta}^0 - h)^q}{p(\mathbf{x}; \boldsymbol{\theta}^0)^2} \right]$$

$$= \alpha \frac{e^{\frac{2qa^2}{2q-1}} \left\| \frac{\psi(\eta^0 + h) + \psi(\eta^0 - h)}{2} - \psi(\eta^0) \right\|^2}{e^{2qa^2} \left\| \frac{\psi(\eta^0 + h) - \psi(\eta^0 - h)}{2} \right\|^2}$$

$$R_{2,2} = E_{\boldsymbol{\theta}^0} \left[ \frac{p(\mathbf{x}; \boldsymbol{\theta}^0 - h)^{2q}}{p(\mathbf{x}; \boldsymbol{\theta}^0)^2} \right] = \alpha e^{\frac{2qa^2}{2q-1}} \|\psi(\eta^0 - h) - \psi(\eta^0)\|^2$$

where  $\alpha = \frac{\pi^{2M(1-q)}}{(2q-1)^M}$ . The classical  $HCRB$ , i.e. the  $HCRB$  obtained from a linear transformation of the LR, is the particular case of its generalized form  $HCRB_q$ , obtained from a non-linear transformation of the LR, where  $q = 1$ .

The benefits on tightness of the introduction of a non-linear transformation is highlighted on figure (1) where the bounds

$$HCRB = \sup_h \{HCRB_{q=1}\} \text{ and } HCRB_{NL} = \sup_{h,q} \{HCRB_q\}$$

( $0 \leq h \leq 0.5$ ,  $\frac{1}{2} < q < 2$ ) are displayed and compared with both the CRB and the MSE of the MLE estimator for  $M = 10$  and  $\eta^0 = 0$ .

### 3.3. Results and Perspectives

First, the proof of the gain in tightness of the  $HCRB$  incorporating the non-linear transformation  $t_q(y) = y^q$  allows to state that:

in the case of the Gaussian p.d.f. family described in Section 3.1 (including both deterministic and stochastic observation models), there exists at least one non-linear transformation improving the tightness of any lower bound deriving from a linear transformation of the unbiasedness constraint (1).

Indeed, all these lower bounds derives from expression (5), which is a generalization of the  $HCRB$  to  $N$  test-points.

Second, an immediate improvement of tightness of all existing lower bounds can be obtained by using the integral form  $t_\theta(y) = y^{q(\theta)}$  of the discrete non-linear transformation  $t_\theta(y) = y^q$ , for various function  $q(\theta)$ .

Last, there are probably other non-linear transformations applicable to the Gaussian (or other) p.d.f which should increase the tightness of existing bounds ....

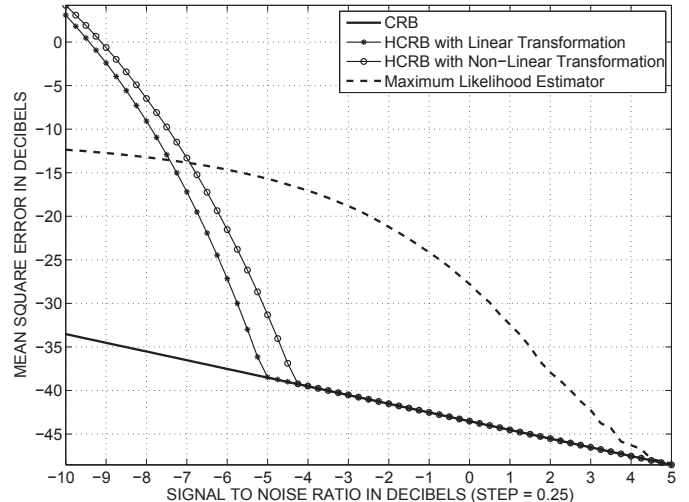


Fig. 1. Comparison of MSE lower bounds versus SNR

## 4. CONCLUSION

In this paper, we have shown that it is worth looking for non-linear transformations of the unbiasedness definition to be combined with linear transformations, in order to derive tighter Barankin bound approximations to improve the threshold prediction. Indeed such non-linear transformations exist for the Gaussian p.d.f.. Additionally, we provide an unified simple framework for the derivation of any Barankin bound approximations based on mixture of linear and non-linear transformations.

## 5. REFERENCES

- [1] E.W. Barankin, "Locally best unbiased estimates", Ann. Math. Stat., vol. 20, no. 4, pp. 477-501, 1949.
- [2] 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., vol. 64, no. 327, pp. 1056-1072, Sep. 1969.
- [3] 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, vol. 56, no. 11, pp. 5319-5333, Nov. 2008
- [4] J. Kiefer, "On minimum variance estimators", Ann. Math. Stat., vol. 23, no. 4, pp. 627-629, 1952.
- [5] 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, vol. 15, no. 2, pp. 273-279, Mar. 1969.
- [6] A. Renaux, P. Forster, E. Chaumette, and P. Larzabal, "On the high SNR CML estimator full statistical characterization," IEEE Trans. on SP, vol. 54, no. 12, pp. 4840-4843, Dec. 2006.
- [7] K. Todros and J. Tabrikian, "A new lower bound on the mean-square error of unbiased estimators", in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., pp. 3913-3916, 2008
- [8] H.L. Van Trees, *Detection, Estimation and Modulation Theory, Part 1*. New York: Wiley, 1968.