14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

ON THE INFLUENCE OF DETECTION TESTS ON DETERMINISTIC PARAMETERS ESTIMATION.

(1) *Eric Chaumette,* (2) *Francois Vincent,* (3) *Jerome Galy and* (4) *Pascal Larzabal*

(1) Thales Naval France, Bagneux, France - eric.chaumette@fr.thalesgroup.com
(2) ENSICA, Toulouse, France - vincent@ensica.fr
(3) LIRM, Montpellier, France - galy@lirmm.fr
(4) SATIE/Ecole Normale Superieure, Cachan, France - Pascal.LARZABAL@satie.ens-cachan.fr

ABSTRACT

In non-linear estimation problems three distinct regions of operation can be observed [1][2]. In the asymptotic region, the Mean Square Error (MSE) of Maximum Likelihood Estimators (MLE) is small and, in many cases, close to the Cramer-Rao bound (CRB) [3]. In the a priory performance region where the number of independent snapshots and/or the SNR are very low, the MSE is close to that obtained from the prior knowledge about the problem. Between these two extremes, there is an additional transition region where MSE of estimators deteriorates with respect to CRB. The present paper provides exemples of improvement of MSE prediction by CRB, not only in the transition region but also in the a priori region, resulting from introduction of a detection step, which proves that this refinement in MSE lower bounds derivation is worth investigating.

1. INTRODUCTION

Lower bounds on the minimum mean square error (MSE) in estimating a set of parameters from noisy observations provide the best performance of any estimators in terms of the MSE. Originally they were introduced to investigate fundamental limits of a parameter estimation problem or to assess the relative performance of a specific estimator. They also have been widely used since as a mean to assess the exact MSE of MLE for problems where it is difficult to evaluate. They can be divided in two families [4]. The first family treats the set of parameters as an unknown deterministic quantity, and provides bounds on the MSE in estimating any selected values of the parameters ("locally" best estimators). The second family assumes that the parameters are random variables with known *a priory* distributions. In this paper, we will focus on the first family, i.e. deterministic parameters estimation. Historically the first MSE lower bound for deterministic parameters to be derived was the CRB [3][4], which has been the most widely used since. Its popularity is largely due to its simplicity of calculation, the fact that in many cases it can be achieved asymptotically (high SNR [5] and/or large number of snapshots [3][4]) by MLE, and last but not least, its noticeable property of being the lowest bound on the MSE of unbiased estimators, since it derives from the weakest formulation of unbiasedness at the vicinity of any selected value of the parameters [6]. This initial characterization of locally unbiased estimators has been improved first by Bhattacharyya's works [4][6][7] which refined the characterization of local unbiasedness, and significantly generalized by Barankin works [6], who established the general form of the greatest lower bound of any absolute moment of an unbiased estimator. In the particular case of MSE, his work allows the derivation of the highest lower bound on MSE (BB) since it takes into account the strongest formulation of unbiasedness, that is to say unbiasedness over an interval of parameter values including the selected value. Unfortunately the BB is generally incomputable [2]. Numerous works ([1][8][9] and additional references in [1] and [10]) devoted to the computing and placing of bounds on MSE have shown that the CRB and the BB can be regarded as key representative of two general classes of bounds, respectively the Small-Error bounds and the Large-Error bounds. Indeed, in non-linear estimation problems three distinct regions of operation can be observed. In the asymptotic region, the MSE is small and, in many cases, close to the Small-Error bounds. In the *a priory* 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 an additional ambiguity region, also called the transition region. In this region, the MSE of MLEs deteriorates rapidly with respect to Small-Error bounds and generally exhibits a threshold behavior corresponding to a "performance breakdown" [11] highlighted by Large-Error bounds. However in nearly all fields of science and engineering, a wide variety of processing requires a binary detection step designed to decide if a signal is present or not in noise. Intuitively, such a detection step is expected to improve the lower bounds tightness by selecting instances with relatively high signal energy - sufficient to exceed the detection threshold - and disregarding instances belonging to the *a priory* region that deteriorate the MSE. Additionally, as a detection step restricts the set of observations available for parameter estimation, any accurate MSE lower bound should take this statistical conditioning into account. In the first part we briefly recapitulate main theoretical results on the characterization of the joint detection and estimation problem for deterministic parameters introduced in [10] and extended since in [12]. Then, the second part aims at completing initial results presented in [10] by highlighting the influence of the type of the detection test on estimation performance.

2. CONDITIONAL LOWER BOUNDS

2.1 On lower bounds and norm minimization

Let $\mathbf{x}$ be the random observations vector and Ω be the observation space. Denote by $f_\theta(\mathbf{x})$ the probability density function (p.d.f.) of observations depending on an unknown deterministic real parameter θ. Let F_Ω be the real vector space of square integrable functions over Ω. A fundamental property of the MSE of a particular estimator $\widehat{g(\theta_0)}(\mathbf{x}) \in F_\Omega$ of $g(\theta_0)$, where θ_0 is a selected value of the parameter θ and $g(\theta)$ is a real function of real variable θ, is that it is a norm associated with a particular scalar product $\langle \mid \rangle_\theta$:

$$\begin{aligned} MSE_{\theta_0}\left[\widehat{g(\theta_0)}\right] &= \left\|\widehat{g(\theta_0)}(\mathbf{x}) - g(\theta_0)\right\|^2_{\theta_0} \quad (1)\\ \langle g(\mathbf{x}) \mid h(\mathbf{x})\rangle_{\theta_0} &= E_{\theta_0}\left[g(\mathbf{x})\, h(\mathbf{x})\right] \\ &= \int g(\mathbf{x})\, h(\mathbf{x})\, f_{\theta_0}(\mathbf{x})\, d\mathbf{x}. \end{aligned}$$

In the search for a lower bound on the MSE, this property allows the use of two equivalent [12] fundamental results: the generalisation of the Cauchy-Schwartz inequality to Gram matrices (generally referred to as the "covariance inequality" [1]) and the minimization of a norm under linear constraints [2][13]. Nevertheless, we shall prefer the "norm minimization" form as its use provides a better understanding of the hypotheses associated with the different lower bounds on the MSE [2][10][13].
Then, let $\mathbb{U}$ be an Euclidean vector space of any dimension (finite or infinite) on the body of real numbers $\mathbb{R}$ which has a scalar product $\langle \mid \rangle$. Let $\mathbf{c}_{[1,K]} = (\mathbf{c}_1, \dots, \mathbf{c}_K)$ be a free family of K vectors of $\mathbb{U}$ and $\mathbf{v} = (v_1, \dots, v_K)^T$ a vector of $\mathbb{R}^K$. The problem of the minimization of $\|\mathbf{u}\|^2$ under the K linear constraints $\langle \mathbf{u} \mid \mathbf{c}_k\rangle = v_k$ then has the solution:

$$\begin{aligned} \min\left\{\|\mathbf{u}\|^2\right\} &= \mathbf{v}^T \mathbf{G}_\mathbf{c}^{-1}\mathbf{v} \quad \text{for } \mathbf{u}_{opt} = \sum_{k=1}^{K} \alpha_k \mathbf{c}_k \quad (2)\\ (\alpha_1, \dots, \alpha_K)^T &= \boldsymbol{\alpha} = \mathbf{G}_\mathbf{c}^{-1}\mathbf{v}, \quad (\mathbf{G}_\mathbf{c})_{k',k} = \langle \mathbf{c}_k \mid \mathbf{c}_{k'}\rangle \end{aligned}$$

The above result (2) can be generalized to linear combinations of a family of N vectors $\mathbf{u}_{[1,N]} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N)$ where the minimization problem becomes [12] - $\boldsymbol{\lambda} \in \mathbb{R}^N, \boldsymbol{\lambda} \neq \mathbf{0}$ - :

$$\min\left\{\left\|\sum_{n=1}^{N} \lambda_n \mathbf{u}_n\right\|^2 = \boldsymbol{\lambda}^T \mathbf{G}_\mathbf{u} \boldsymbol{\lambda}\right\} \text{ under } \langle \mathbf{u}_n \mid \mathbf{c}_k\rangle = \mathbf{V}_{k,n} \quad (3)$$

- $(\mathbf{G}_\mathbf{u})_{n',n} = \langle \mathbf{u}_n \mid \mathbf{u}_{n'}\rangle$ - and leads to the matrix inequality:

$$\boldsymbol{\lambda}^T \mathbf{G}_\mathbf{u} \boldsymbol{\lambda} \geq \boldsymbol{\lambda}^T \left(\mathbf{V}^T \mathbf{G}_\mathbf{c}^{-1} \mathbf{V}\right) \boldsymbol{\lambda} \Leftrightarrow \mathbf{G}_\mathbf{u} \geq \mathbf{V}^T \mathbf{G}_\mathbf{c}^{-1} \mathbf{V} \quad (4)$$

2.2 Example of lower bound derivation: CRB

As an example of lower bound derivation using this remarkable approach, we propose a novel derivation of CRB in the multiple parameters case [12]. We consider now the case where the p.d.f. of the observations $f_{\boldsymbol{\theta}}(\mathbf{x})$ depends on a vector of K parameters $\boldsymbol{\theta} = (\theta_1, \dots, \theta_K)$ belonging to $\mathbb{R}^K$. Let $\boldsymbol{\theta}_0$ be a particular value of $\boldsymbol{\theta}$, and $\widehat{\mathbf{g}(\boldsymbol{\theta}_0)}(\mathbf{x})$ an estimator of $\mathbf{g}(\boldsymbol{\theta}_0)$ vector of N real functions of $\boldsymbol{\theta}$. In the multiple parameters context, it seems quite natural to consider that $\widehat{\mathbf{g}_n(\boldsymbol{\theta}_0)}(\mathbf{x})$ is a "locally unbiased" estimator of $\mathbf{g}_n(\boldsymbol{\theta}_0)$ if:

$$\begin{aligned} E_{\boldsymbol{\theta}_0 + d\boldsymbol{\theta}}\left[\widehat{\mathbf{g}_n(\boldsymbol{\theta}_0)}(\mathbf{x})\right] &= \mathbf{g}_n(\boldsymbol{\theta}_0 + d\boldsymbol{\theta}) + o(\|d\boldsymbol{\theta}\|) \quad (5)\\ &= \mathbf{g}_n(\boldsymbol{\theta}_0) + \frac{\partial \mathbf{g}_n(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}^T} d\boldsymbol{\theta} + o(\|d\boldsymbol{\theta}\|) \end{aligned}$$

where $\frac{\partial}{\partial \boldsymbol{\theta}} = \left(\frac{\partial}{\partial \theta_1}, \dots, \frac{\partial}{\partial \theta_K}\right)^T$, $\frac{\partial}{\partial \boldsymbol{\theta}^T} = \left(\frac{\partial}{\partial \boldsymbol{\theta}}\right)^T$ and $\frac{\partial h(\theta_0, \mathbf{x})}{\partial \theta} = \left.\frac{\partial h(\theta, \mathbf{x})}{\partial \theta}\right|_{\theta_0}$, which means that, up to the first order and in the neighborhood of $\boldsymbol{\theta}_0$, $\widehat{\mathbf{g}_n(\boldsymbol{\theta}_0)}(\mathbf{x})$ remains an unbiased estimator of $\mathbf{g}_n(\boldsymbol{\theta})$ independently of a - small - variation of the parameters. Considering as well that in the neighborhood of $\boldsymbol{\theta}_0$:

$$f_{\boldsymbol{\theta}_0 + d\boldsymbol{\theta}}(\mathbf{x}) = f_{\boldsymbol{\theta}_0}(\mathbf{x}) + \frac{\partial f_{\boldsymbol{\theta}_0}(\mathbf{x})}{\partial \boldsymbol{\theta}^T} d\boldsymbol{\theta} + o(\|d\boldsymbol{\theta}\|),$$

the requested locally unbiased property (5) is satisfied for all components of $\widehat{\mathbf{g}(\boldsymbol{\theta}_0)}(\mathbf{x})$ if the following linear constraints are verified:

$$\left\{\begin{aligned} E_{\boldsymbol{\theta}_0}\left[\widehat{\mathbf{g}(\boldsymbol{\theta}_0)}(\mathbf{x}) - \mathbf{g}(\boldsymbol{\theta}_0)\right] &= \mathbf{0} \\ E_{\boldsymbol{\theta}_0}\left[\left(\widehat{\mathbf{g}(\boldsymbol{\theta}_0)}(\mathbf{x}) - \mathbf{g}(\boldsymbol{\theta}_0)\right) \frac{\partial \ln f_{\boldsymbol{\theta}_0}(\mathbf{x})}{\partial \boldsymbol{\theta}}^T\right] &= \frac{\partial \mathbf{g}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}^T} \end{aligned}\right. \quad (6)$$

Then by defining:

$$\mathbf{u}_{[1,N]} = \widehat{\mathbf{g}(\boldsymbol{\theta}_0)}(\mathbf{x}) - \mathbf{g}(\boldsymbol{\theta}_0), \quad \mathbf{c}_{[1,K]} = \begin{pmatrix} 1 \\ \frac{\partial \ln f_{\boldsymbol{\theta}_0}(\mathbf{x})}{\partial \boldsymbol{\theta}} \end{pmatrix}$$

and considering scalar product (1), inequality (3) can be applied for $\mathbf{V} = \begin{bmatrix} \mathbf{0}^T \\ \frac{\partial \mathbf{g}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}^T} \end{bmatrix}^T$ and leads to:

$$E_{\boldsymbol{\theta}_0}\left[\left(\widehat{\mathbf{g}(\boldsymbol{\theta}_0)}(\mathbf{x}) - \mathbf{g}(\boldsymbol{\theta}_0)\right)\left(\widehat{\mathbf{g}(\boldsymbol{\theta}_0)}(\mathbf{x}) - \mathbf{g}(\boldsymbol{\theta}_0)\right)^T\right] \geq \left[\frac{\partial \mathbf{g}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}^T}\right] \mathbf{F}(\boldsymbol{\theta}_0)^{-1} \left[\frac{\partial \mathbf{g}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}^T}\right]^T \quad (7)$$

where:

$$\mathbf{F}(\boldsymbol{\theta}_0) = E_{\boldsymbol{\theta}_0}\left[\left(\frac{\partial \ln f_{\boldsymbol{\theta}_0}(\mathbf{x})}{\partial \boldsymbol{\theta}}\right)\left(\frac{\partial \ln f_{\boldsymbol{\theta}_0}(\mathbf{x})}{\partial \boldsymbol{\theta}}\right)^T\right] \quad (8)$$

is the well known Fisher Information Matrix (FIM).

As previously mentionned and higlighted hereabove, the principal merit of the "norm minimization" form is to raise explicitly in the first place the problem of the formulation of pertinent constraints ((5) and (6)), which then determine the value of the lower bound on the MSE.

2.3 Extension to Conditional Lower Bounds

More generally, all known bounds on the MSE – Cramer-Rao, Bhattacharya, Barankin, Hammersley-Chapman-Robbins and Abel bounds – are different solutions of the same generalized norm minimization problem (3) under sets of appropriate linear constraints (possibly infinite but countable) [2][10][12][13]. If the observations set is restricted to a subset D of Ω, for example by a detection step, then the p.d.f. of observations $f_\theta(\mathbf{x})$ becomes a conditional p.d.f. $f_\theta(\mathbf{x} \mid D) = \frac{f_\theta(\mathbf{x})}{P_D(\theta)}$ and scalar product (1) becomes:

$$\begin{aligned} \langle g(\mathbf{x}) \mid h(\mathbf{x}) \rangle_{\theta_0 \mid D} &= E_{\theta_0}\left[g(\mathbf{x})\, h(\mathbf{x}) \mid D\right] \\ &= \int_D \left[g(\mathbf{x})\, h(\mathbf{x})\right] f_{\theta_0}(\mathbf{x} \mid D)\, d\mathbf{x} \end{aligned} \tag{9}$$

where $\int_D f_\theta(\mathbf{x})\, d\mathbf{x} = P(D) = P_D(\theta)$ is the probability of conditioning event D. It is obvious that if subset D does not depend on parameter θ, scalar product definitions (1) and (9) are of the same form. Consequently, whatever bound is considered its conditional formulation will be obtained by substituting D and $f_\theta(\mathbf{x} \mid D)$ for Ω and $f_\theta(\mathbf{x})$ in the various expressions. For example, one can derive the two useful following expressions of the Conditional Fisher Information Matrix (CFIM) [10]:

$$\mathbf{F}(\boldsymbol{\theta} \mid D)_{k,l} = E_{\boldsymbol{\theta}}\left[\frac{\partial \ln f_{\boldsymbol{\theta}}(\mathbf{x})}{\partial \theta_k} \frac{\partial \ln f_{\boldsymbol{\theta}}(\mathbf{x})}{\partial \theta_l} \mid D\right] - \frac{\partial \ln P_D(\boldsymbol{\theta})}{\partial \theta_k} \frac{\partial \ln P_D(\boldsymbol{\theta})}{\partial \theta_l}$$

$$\mathbf{F}(\boldsymbol{\theta} \mid D)_{k,l} = -E_{\boldsymbol{\theta}}\left[\frac{\partial^2 \ln f_{\boldsymbol{\theta}}(\mathbf{x})}{\partial \theta_k \partial \theta_l} \mid D\right] + \frac{\partial^2 \ln P_D(\boldsymbol{\theta})}{\partial \theta_k \partial \theta_l}$$

that encompass usual unconditional FIM expressions (8).

Contrarily, if subset D does depend on parameter θ, generalization of lower bounds based on derivatives constraints - such as Cramer-Rao, Bhattacharya or Abel bounds [1][2][10] - is not an elementary exercise since it involves calculation of integral derivatives with respect to its domain. Although it is certainly an interesting mathematical problem, this case is of little interest for actual applications where realizable detection test - defining the conditioning set D - can not depend on the unknown value θ.

3. ON THE INFLUENCE OF DETECTION TESTS ON ESTIMATION PERFORMANCE

This section deals with estimation of the direction of arrival (DOA) of a signal source by means of a 2 sensors array to illustrate the main theoretical points ensuing from conditioning: conditional bound expressions, "conditional" efficiency, importance of bias at low SNR. Why this application? Firstly, because this technique is one of the oldest and most widely used high-resolution techniques, even nowadays, in most operational tracking systems [14]. Secondly, as in nearly all fields of science and engineering, its processing requires a detection step. Last but not least, a complete statistical prediction can be computed analytically for a Rayleigh signal source, including probability of detection, bias, MSE of MLE and CRB, conditionned by two different detection tests.

3.1 DOA estimation with a 2 sensors array

Assume that a signal source situated at an angle θ (deviation angle from array boresight) is received on a 2 sensors (Σ and Δ) array in the presence of a circular, zero mean, white (both temporally and spatially), complex Gaussian thermal noise. A common model of the observation equation dedicated to this problem - after Hilbert Filtering - is the following receiver signal vector:

$$\begin{aligned} \begin{bmatrix} \Sigma(t) \\ \Delta(t) \end{bmatrix} &= \alpha(t) \begin{bmatrix} g_\Sigma \\ g_\Delta \end{bmatrix} + \begin{bmatrix} n_\Sigma(t) \\ n_\Delta(t) \end{bmatrix} \\ &= \beta(t)\, \mathbf{x} + \mathbf{n}(t) \end{aligned} \tag{10}$$

where $\beta(t) = \alpha(t)\, g_\Sigma$, $\mathbf{x} = (1, r)^T$, $r = \frac{g_\Delta}{g_\Sigma}$, $n_\Sigma(t)$ and $n_\Delta(t)$ represent Gaussian receiver noise, g_Σ and g_Δ represent the one-way complex sensor voltage pattern at angle θ and $\alpha(t)$ represents the complex amplitude of the source (including power budget equation, signal processing gains).

In the particular case of a 2 sensors array, the angular information is contained in the ratio $r(\theta) = \frac{g_\Delta(\theta)}{g_\Sigma(\theta)}$, provided the function $\theta \longrightarrow r(\theta)$ is invertible. In actual 2 sensors arrays beamwidth/resolution constraint generally prevents this assumption from being verified for any θ in $\left[\frac{-\pi}{2}, \frac{\pi}{2}\right]$. Nevertheless with appropriate sensors patterns - uniform sum excitation for Σ and linear odd difference excitation for Δ - collocated and in phase [14], $r(\theta) = r_x(\theta)$ is real and the property can hold for θ belonging to Σ main beam, i.e. between the first pattern nulls. Such 2 sensors array are generally called monopulse antennas where $r_x(\theta)$ is the monopulse ratio and $\theta = r_x^{-1}\left(\frac{g_\Delta}{g_\Sigma}\right)$ is the deviation angle function.

If a linear relation $r_x = k\theta$ is assumed - which is true at the vicinity of boresight [14] - then statistical prediction of $\widehat{\theta} = \frac{\widehat{r}_x}{k}$ can be easily derived from statistical prediction of $\widehat{r}_x = \mathrm{Re}\{\widehat{r}\}$. It is the reason why in open literature the deviation angle function is generally reduced to a linear function characterized by a Monopulse Slope and most DOA statistical performance analysis are related to $\widehat{r}_x$. We will consider this approximation in the present paper.

3.2 Statistical prediction of monopulse ratio MLE

In the following, we focus on the case $r_x = r = 0$, which corresponds to a source signal located along the main axis of a monopulse antenna ($\theta = 0$). This is a reference

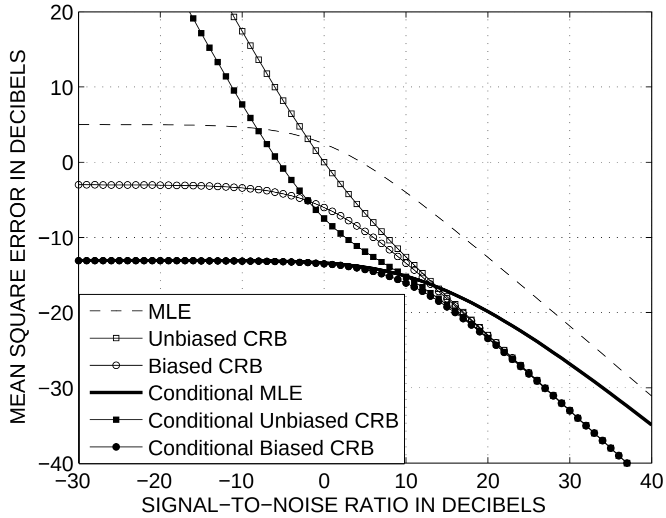

Figure 1: MSE of MLE and CRB of $\widehat{r}_x = \mathrm{Re}\left\{\frac{\Delta}{\Sigma}\right\}$ conditioned by LRT versus SNR

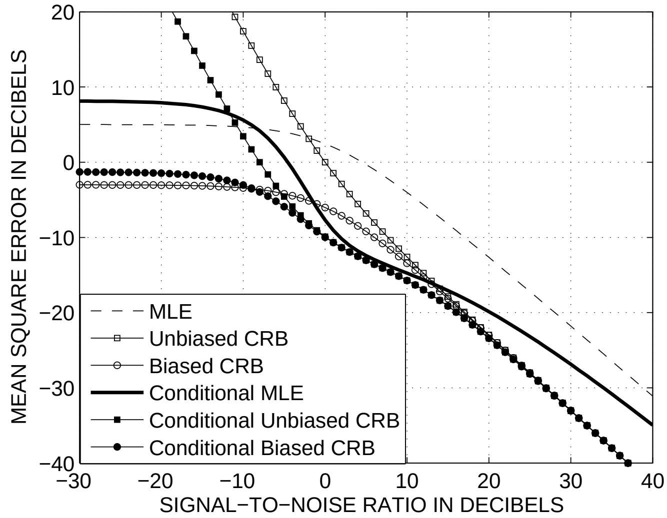

Figure 2: MSE of MLE and CRB of $\widehat{r}_x = \mathrm{Re}\left\{\frac{\Delta}{\Sigma}\right\}$ conditioned by GLRT versus SNR

case in the study of the performance of such receiving system [14], since it corresponds to the peak received energy. Moreover, for this particular DOA and a single snapshot, the two following detection tests:

$$|\Sigma|^2 \underset{H_0}{\overset{H_1}{\gtrless}} T \quad \text{and} \quad |\Delta|^2 + |\Sigma|^2 \underset{H_0}{\overset{H_1}{\gtrless}} T$$

correspond respectively to the Likelihood Ratio Test (LRT, Neyman-Pearson Test) and the Generalized LRT [15], where T is the detection threshold. The monopulse ratio MLE is then defined by $\widehat{r}_x = \mathrm{Re}\left\{\frac{\Delta}{\Sigma}\right\}$. The two joint detection and estimation problems have been anaytically characterized in terms of:

- Probability of Detection (P_D) and Probability of False Alarm (P_{FA}),
- conditional mean ($E\left[\widehat{r}_x \mid D\right]$) and variance ($Var\left[\widehat{r}_x \mid D\right]$),
- conditional CRB ($\mathbf{F}\left(\boldsymbol{\theta} \mid D\right)$),

conditioned by the event $D = \left\{|\Sigma|^2 \geq T\right\}$ of the LRT or the event $D = \left\{|\Delta|^2 + |\Sigma|^2 \geq T\right\}$ of the GLRT [10][12][15]. Therefore all results introduced in the next section rely on exact analytical formulas.

3.3 Results

Figures (1) and (2) display the values of the MSE of MLE of r_x and its related CRBs for biased - $E\left[\widehat{r}_x \mid D\right]$ is known - and unbiased estimates as a function of SNR on sum channel (Σ) for $P_{FA} = 10^{-4}$ (classical value in track mode for a radar). In figures legend, "MLE", "Unbiased CRB" and "Biased CRB" respectively stands for MLE, CRB for unbiased estimators and CRB including the known bias, computed for $P_{FA} = 0.999$ (negligible detection step), whereas "Conditional ..." take into account the detection test, which is the the LRT in figure (1), and the GLRT in figure (2).

First of all, let us mention that in the problem at hand, the p.d.f. of $\widehat{r}_x$ without conditionning follows a Student distribution with mean value 0 and a smoothly increasing variance [10] as the SNR decreases. Therefore, we are in the particular case where the transition region is smooth when the detection threshold effect is negligible (see 'MLE' curve in figures (1) and (2)), which is not the most general case in non-linear estimation problems.

Nevertheless some general considerations can still be drawn from this particular case. Intuitively, the detection step is expected to modify MSE behavior mainly in the transition region where it plays a crucial role in selecting instances with relatively high signal energy - sufficient to exceed the detection threshold - and disregarding instances mainly consisting of noise that deteriorate the MSE. This effect is confirmed by both figures (1) and (2) where the introduction of a detection step has decreased MSE values in the transition region. Additionnaly both figures highlights the necessity of using conditional form of a given lower bound to be able to keep at least its lower bound property. This is particularly well illustrated on figure (2) in the SNR region $[-10, 10]$ dB where the Conditional Unbiased CRB is obviously the only meaningful expression of the unbiased bound.

On the other hand, a surprising result is the tightness of:

- the Conditional Unbiased CRB in most of the transition region in case of the GLRT (figure (2)),
- the Conditional Biased CRB in all the *a priori* region in case of the LRT (figure (1)).

Indeed, a particular property of $\widehat{r}_x$ is to be non efficient when SNR tends to infinity, which originates from its Student p.d.f. (see 'MSE' curves in both figures (1) and (2)). This result reveals that conditioning of the ob-

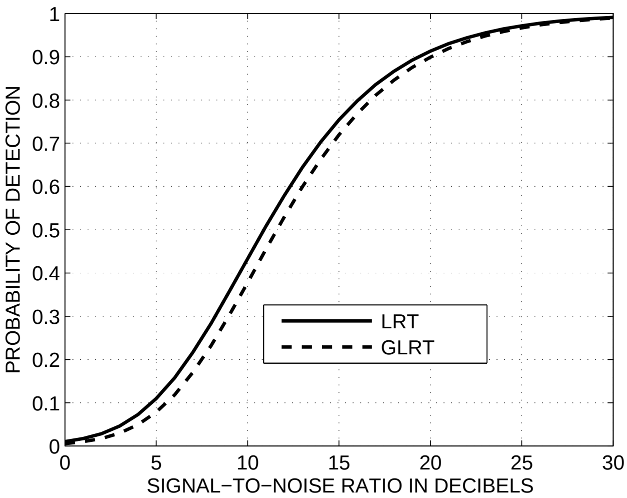

Figure 3: Probability of Detection versus SNR

servations by an event D may significantly modify the conditions required to attain the CRB and thus to obtain an efficient estimator, biased or not.
Additionally, it also shows that there is a limit to the pertinence of the information delivered by the CRB for unbiased estimates at very low SNR, even when conditioning is included, since threshold detection increase has almost no effect on the tightness of the Unbiased CRB in the *a priory* region, whatever the test. The main reason is that a locally unbiased estimator of source signal parameters generally does not exist asymptotically as the SNR decreases to 0. To overcome this limitation, one can resort to biased CRB. It is an attractive theoretical refinement if analytical expression of the bias is available as it is shown in both figure (1) and (2) where introduction of bias ("Conditional Biased CRB" curve) has restored CRB property in all regions of operation. Unfortunately the bias depends on the specific estimator and furthermore is hardly ever known in practice. This pessimistic observation must however be balanced against practical considerations. For example, it is doubtful whether this limit raises a genuine practical problem in the GLRT case, since it appears in an SNR region ($SNR < -10dB$) where the source signal is simply considered to be absent from an operational point of view $\left(P_D \leq 10^{-3}\right)$.

Last but not least, the most noticeable result is that the nature of the observations selection performed by each detection test is of first importance: whereas they lead to comparable P_D (see figure (3)) they induce a very different MSE behaviour as their selection effect increases, resulting in a completely opposite effect on the MSE value in the *a priori* region (decrease with LRT and increase with GLRT).

4. CONCLUSION

Despite they have been derived for CRB and for a particular case of MLE behavior (smooth transition region), the results introduced in the present paper shows that the problem of lower bound tightness at low SNR ($P_D < 1$) must be revisited for practical application involving a binary detection test. Indeed they clearly highlight that, in a joint detection and estimation problem, the nature of the detection step plays a crucial role in estimation performance preventing from drawing relevant forecasts from the unconditionned study case. Additionnaly, improvement of Unbiased CRB behavior in the transition region resulting from conditioning by the GLRT shows that this refinement in MSE lower bounds derivation for unbiased estimators is worth investigating, including Large-Error bound representatives.

REFERENCES

[1] J.S. Abel, "A bound on mean-square-estimate error", IEEE Trans., IT vol 39, pp. 1675-1680, 1993

[2] P. Forster, P. Larzabal, "On lower Bounds For Deterministic Parameters Estimation", ICASSP Conf. 2002, p1137-1140

[3] H. Cramer, "Mathematical Methods of Statistics", Princeton, Princeton University Press 1946

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

[5] A. Renaux, P. Forster, E. Chaumette, P. Larzabal, "On the High SNR CML Estimator Full Statistical Characterization", to appear in IEEE Trans. on SP

[6] E.W. Barankin, "Locally best unbiased estimates", Ann. Math. Stat., vol. 20, pp. 477-501, 1949

[7] A. Bhattacharyya,"On some analogues of the amount of information and their use in statistical estimation", Sankya, vol. 8 , pp. 1-14, 1946

[8] D.G. Chapman, H. Robbins, "Minimum variance estimation without regularity assumptions", Annals of Math. Stat, vol. 22, pp. 581-586, 1951

[9] R. McAulay, L.P. Seidman, "A useful form of the Barankin Lower Bound and its application to PPM threshold analysis", IEEE Trans, IT vol 15, pp. 273-279, 1969

[10] E. Chaumette, P. Larzabal, P. Forster, "On the Influence of a Detection Step on Lower Bounds for Deterministic Parameters Estimation", IEEE Trans., SP vol 53, pp. 4080-4090, 2005

[11] D.C. Rife, R.R. Boorstyn,"Single tone parameter estimation from discrete-time observations", IEEE Trans., IT vol 20, pp. 591-598, Sept. 1974

[12] E. Chaumette, "Contribution to the performance characterization of the joint Detection-Estimation problem", PhD Thesis, Ecole Normale Supérieure de Cachan, France, Dec. 2004

[13] F.E. Glave, "A new look at the Barankin Lower Bound", IEEE Trans., IT vol 18, pp. 349-356, May 1972

[14] N. Levanon, "Radar Principles", Wiley-Interscience, Wiley 1988

[15] E. Chaumette, P. Larzabal, "Optimal Detection Theory Applied to Monopulse Antenna", ICASSP Conf. 2004, Vol II p. 1065-1068