A ZIV-ZAKAI TYPE BOUND FOR HYBRID PARAMETER ESTIMATION

Chengfang Ren(1), Jerome Galy(2), Eric Chaumette(3,4) Pascal Larzabal(4) and Alexandre Renaux(1)

(1) Universite Paris-Sud/LSS
3, Rue Joliot-Curie, 91192 Gif-sur-Yvette, France. Emails: cren@lss.supelec.fr, renaux@lss.supelec.fr
(2) Universite de Montpellier 2/LIRMM
161 rue Ada 34392 Montpellier Cedex 5, France. Email: galy@lirmm.fr
(3) ONERA/The French Aerospace Lab
Chemin de la Huniere Palaiseau Cedex 91123, France. Email: eric.chaumette@onera.fr
(4) Ecole Normale Superieure de Cachan/SATIE
61 av. du President Wilson, 94235 Cachan cedex, France. Email: pascal.larzabal@satie.ens-cachan.fr

ABSTRACT
In statistical signal processing, hybrid parameter estimation refers to the case where the parameters vector to estimate contains both non-random and random parameters. In this communication, we propose a new hybrid lower bound which, for the first time, includes the Ziv-Zakai bound well known for its tightness in the Bayesian context (random parameters only). For the general case of parameterized mean model with Gaussian noise, closed-form expressions of the proposed bound are provided.

Index Terms — Parameter estimation, Ziv-Zakai bounds, hybrid bounds, SNR threshold

1. INTRODUCTION
While Bayesian or non-Bayesian estimation techniques are now widely used in statistical signal processing, the technique called hybrid estimation has been developed more recently and suffers from a relative lack of results. Hybrid parameters refer to the case where the parameters vector to estimate contains both non-random and random parameters with a priori known probability density functions (p.d.f.). Such framework is useful for several signal processing applications, but it is not just the simple concatenation of Bayesian and non-Bayesian techniques. Indeed, new estimator has to be derived and one cannot use the Maximum Likelihood estimator for the non-Bayesian part and the Maximum A Posteriori estimator for the Bayesian part since the parameters can have a dependence. In the same way, performance analysis methods of such hybrid estimators has to be modified accordingly.

Signal processing community generally use the Hybrid Cramér-Rao Bound (HCRB) [1] for which some asymptotic achievability results [2] are known. The HCRB, as well as the classical CRB, is known to be simple to obtain for various problems (see Part III of [3]) but suffers from some drawbacks. The main one is its only asymptotic tightness in terms of number of samples or Signal-to-Noise Ratio (SNR) leading to the incapability of predicting the so-called threshold effect (i.e. large errors) on estimator mean square error in non-linear estimation problems. In order to fill this lack, other hybrid lower bounds have already been proposed, e.g. the Hybrid Barankin Bound (HBB) [4] or the Hybrid Barankin/Weiss-Weinstein bound (HBWWB) [5]. In each case, the key idea is to combine, in a tricky way, some lower bounds already known in the Bayesian and non-Bayesian framework. So far this combination has been done by resorting to the covariance inequality principle which seems to be the cornerstone to establish such hybrid lower bounds [5][6].

However, if we focus our attention on the Bayesian context (random parameter only), two families of bounds are known in the literature: the Weiss-Weinstein family and the Ziv-Zakai family. If the Weiss-Weinstein family is based on the covariance inequality principle, leading to a natural extension to the hybrid context [5], the Ziv-Zakai family [7] is based on a binary hypothesis testing problem not a priori linked to the covariance inequality principle. A first consequence is the necessity to resort to simulation in order to compare the tightness of these two families. A second one is the non-exhibition of an hybrid lower bound using one of the Ziv-Zakai family bounds. This paper aims to fill this lack since the Ziv-Zakai bounds are known to be very tight in the Bayesian context (see e.g. [8][9][10]).

First, by adapting an idea suggested in [11, p. 38], we propose an inequality between the hybrid MSE of a general class of estimators and a quantity, closely related to the Ziv-Zakai bound. We prove that this quantity is independent of the estimation scheme and is, consequently a new lower bound. Moreover, any lower bound on the MSE is a useful bound for signal processing problems if one are able to obtain closed-form expression for a large set of estimation problems. Therefore, in the second part of this paper we derive closed-form expressions of the proposed bound for the general case of Gaussian observation model with parameterized mean. This model is widely met in signal processing problems such that: spectral analysis [12], array processing [13], digital communications [14], etc. Finally, a comparison with the Maximum A Posteriori / Maximum Likelihood Estimator (MAPMLE) and existing bounds is given in a frequency estimation problem.
2. RELATION TO PRIOR WORK

In the Bayesian context, the Weiss-Weinstein bound [15] and the (extended) Ziv-Zakai bound [7] are both known to be tight while they come from two distinct theories with no a priori relationship. In the hybrid context, a bound including the Weiss-Weinstein bound for the random part has already been proposed [5]. The purpose of the present paper is to provide a hybrid bound including the Ziv-Zakai bound for the random part.

3. THE PROPOSED BOUND

Consider an observation space $\Omega$ of points $X$ and let $\theta = (\theta_d, \theta_r)^T$ denotes the hybrid parameter vector to estimate where $\theta_d \in \Pi_d \subseteq \mathbb{R}$ is an unknown deterministic parameter and where $\theta_r \in \Pi_r \subseteq \mathbb{R}$ is an unknown random parameter characterized by a prior p.d.f. which is assumed to be independent of $\theta_d$. In other words, $f(\theta_d, \theta_r) = f(\theta_d)$. Let $f(X, \theta) = f(X, \theta_d, \theta_r)$ denote the joint PDF of $X$ and $\theta_r$ parameterized by $\theta_d$. For any estimators $\hat{\theta} = (\hat{\theta}_d, \hat{\theta}_r)^T$ and for any $h_d$ and $h_r$ such that $\theta_d + h_d \in \Pi_d$ and $\theta_r + h_r \in \Pi_r$, satisfying the following assumptions:

1) $\forall \theta_d \in \Pi_d, f(X, \theta_d, \theta_r) = 0 \Rightarrow f(X, \theta_d + h_d; \theta_r) = 0$.
2) $\forall \theta_r \in \Pi_r, E_{\theta_d}[\theta_d | \theta_r] = \theta_d, E_{\theta_d, \theta_r + h_d}[\theta_d] = \theta_d + h_d$.
3) $\forall \theta_d \in \Pi_d, E_{\theta_d, \theta_r}[\theta_r - \theta_d] = 0$ and $E_{\theta_d, \theta_r + h_r, \theta_d}[\theta_r - (\theta_d + h_r)] = 0$.

Then, estimators MSE is bounded by (The proof is given in Appendix)

$$E_{\theta}[\hat{\theta} - \theta)^T (\hat{\theta} - \theta)] \geq CV^{-1}C^T, \quad (1)$$

where $A \succeq B$ means $A - B$ is a positive semidefinite matrix. Each element of matrix $V$ is given by

$$\begin{cases} \{V\}_{1,1} = \mu(h_1) - 1, \\ \{V\}_{2,2} = \beta(h_2, h_2) + \beta(-h_2, -h_2) - 2\beta(h_2, -h_2), \\ \{V\}_{1,2} = \{V\}_{2,1} = \alpha(h_1, h_2) - \alpha(h_1, -h_2). \end{cases} \quad (2)$$

The matrix $C$ is given by

$$C = \begin{pmatrix} h_{1d} & 0 \\ h_{1r} & h_{2r} & 0 & \alpha(0, h_1) \end{pmatrix} \quad (5)$$

Let us set $h_1 = (h_{1d} h_{1r})^T$ and $h_2 = (0 h_{2r})^T$ which are the so-called test points. Finally, $\mu(\cdot), \alpha(\cdot, \cdot)$ and $\beta(\cdot, \cdot)$ are defined by

$$\mu(h) = E_{X, \theta}[f^2(X, \theta + h)], \quad (6)$$

$$\begin{cases} \alpha(h_1, h_2) = E_{X, \theta}[f(X, \theta + h_1) f(X, \theta + h_2) - f(X, \theta + h_1) f(X, \theta - h_2)], \\ \beta(h_1, h_2) = E_{X, \theta}[f(X, \theta + h_1) f(X, \theta + h_2) - f(X, \theta - h_1) f(X, \theta + h_2)]. \end{cases} \quad (7)$$

We start by the expression of $\beta(\cdot, \cdot)$ since it can be directly obtained from the literature.

4.1. Expression of $\beta(\cdot, \cdot)$

For all $\delta_1 = (0 \delta_1)^T$ and $\delta_2 = (0 \delta_2)^T$ such that $\delta_1 \neq \delta_2$, $\beta(\delta_1, \delta_2)$ is similar to (2) in [17]

$$\beta(\delta_1, \delta_2) = E_{\theta_r} (I_{\delta_1} + I_{\delta_2} + I_{\delta_3} + I_{\delta_4}) \quad (11)$$

with

$$I_{\delta_1} = F_{\mathcal{N}(0, \frac{\sigma_r^2}{2}\mathbf{r})} (-\frac{\sigma_r^2}{2} \delta_1^2 - \frac{\sigma_r^2}{2} \delta_2^2),$$

$$I_{\delta_2} = e^{-\frac{\delta_1^2 + 2\sigma_r \delta_1}{\sigma_r^2}} F_{\mathcal{N}(m_{\delta_2}, \frac{\sigma_r^2}{2}\mathbf{r})} (-\frac{\sigma_r^2}{2} m_{\delta_2}^2 - \frac{\sigma_r^2}{2} \delta_2^2),$$

$$I_{\delta_3} = e^{-\frac{\delta_1^2 + 2\sigma_r \delta_1}{\sigma_r^2}} F_{\mathcal{N}(m_{\delta_3}, \frac{\sigma_r^2}{2}\mathbf{r})} (\sigma_r^2 m_{\delta_3}^2 - \frac{\sigma_r^2}{2} \delta_2^2),$$

$$I_{\delta_4} = e^{-\frac{\delta_1^2 + 2\sigma_r \delta_1}{\sigma_r^2}} F_{\mathcal{N}(m_{\delta_4}, \frac{\sigma_r^2}{2}\mathbf{r})} (\sigma_r^2 m_{\delta_4}^2 - \frac{\sigma_r^2}{2} \delta_2^2).$$

We define $F_{\mathcal{N}(m, \Sigma)}(a)$ as the value at the point $a$ of a normal cumulative distribution function parameterized by mean $m$ and covariance matrix $\Sigma$, $d(\delta_1) = g(\theta_d, \theta_r + \delta_1) - g(\theta_d, \theta_r)$, $b(\delta_1) = \frac{1}{\sqrt{2\pi}} \lVert d(\delta_1) \rVert^2 + \frac{\delta_1^2 + 2\sigma_r \delta_1}{2\sigma_r^2}$. The main challenge to compute the proposed bound is to obtain $\alpha(\cdot, \cdot)$ and $\beta(\cdot, \cdot)$ since the expression of $\mu(h)$ is already known in [16].
\[ m_{\beta_2} = - \text{Re} \left( \bar{d}^H (\delta_1) d (\delta_2) \right) \| d (\delta_2) \|^2 \], \\
m_{\beta_3} = \left( \| d (\delta_1) \|^2 - \text{Re} \left( \bar{d}^H (\delta_1) d (\delta_2) \right) \right), \\
m_{\beta_4} = \left( \| d (\delta_2) \|^2 - \text{Re} \left( \bar{d}^H (\delta_1) d (\delta_2) \right) \right), \\
\Gamma = \left( \| d (\delta_1) \|^2 - \text{Re} \left( \bar{d}^H (\delta_1) d (\delta_2) \right) \right) \| d (\delta_2) \|^2, \\
\Gamma' = \left( \| d (\delta_1) \|^2 - \text{Re} \left( \bar{d}^H (\delta_1) d (\delta_2) \right) \right) \| d (\delta_2) \|^2. \\
\]

Note that if \( \delta_1 = \delta_2 \), \( \Gamma \) and \( \Gamma' \) become singular. However, an analytic expression exists in [17] where (11) is modified with the quantities

\[ I_{\beta_1} = F_{N(0, \sigma_n^2 | d(\delta_1)|^2)} \left( - \frac{\sigma_n^2 b (\delta_1)}{2} \right), I_{\beta_2} = I_{\beta_3} = 0 \] (16)

and

\[ I_{\beta_4} = e^{\frac{2 |d(\delta_1)|^2}{\sigma_n^2} \! + \! 2 \! \frac{v_{\theta_1} + 2 v_{\theta_2}}{v_{\theta_2}}} \times \! F_{N(2 |d(\delta_1)|^2, \frac{\sigma_n^2 |d(\delta_1)|^2}{2})} \left( \frac{\sigma_n^2 b (\delta_1)}{2} \right). \] (17)

### 4.2. Expression of \( \alpha \) (\ldots)

The definition of \( \alpha (h_1, h_2) \) is given by (7). The main idea is to split integration domains in which the min operator can be substituted by \( \int_{\Omega_{h_1} h_2} \) or 1. Thus, we can split \( \alpha (h_1, h_2) \) into two parts, for all \( h_1 = (h_{1d}, h_{1r})^T \) and \( h_2 = (0, h_{2r})^T \)

\[ \alpha (h_1, h_2) = E_{\theta_1} \left( I_{\alpha_1} + I_{\alpha_2} \right) \] (18)

with

\[ I_{\alpha_1} = \int_{V_1} f (x, \theta + h_1) f (x, \theta) \, dx, \] (19)

and

\[ I_{\alpha_2} = \int_{V_2} f (x, \theta + h_1) f (x, \theta + h_2) f (x, \theta) \, dx, \] (20)

where

\[ V_1 = \{ x \in \Omega \mid \frac{f (x, \theta + h_2)}{f (x, \theta)} \geq 1 \}, \] (21)

and

\[ V_2 = \{ x \in \Omega \mid \frac{f (x, \theta + h_2)}{f (x, \theta)} < 1 \}. \] (22)

Some calculus similar to [17] lead to the following closed-form expressions:

\[ I_{\alpha_1} = e^{- \frac{2 \sigma_n^2 (\bar{h}_{1d} h_{1r} + h_{1r} h_{2r})}{2 \sigma_n^2}} \! \times \! F_{N(m_{\alpha_1}, \frac{\sigma_n^2 |d(\delta_2)|^2}{2})} \left( \frac{-\sigma_n^2 b (h_{2r})}{2} \right), \] (23)

and

\[ I_{\alpha_2} = e^{- \frac{2 \sigma_n^2 (\bar{h}_{1d} h_{1r} + h_{1r} h_{2r})}{2 \sigma_n^2}} \! \times \! F_{N(m_{\alpha_2}, \frac{\sigma_n^2 |d(\delta_2)|^2}{2})} \left( \frac{\sigma_n^2 b (h_{2r})}{2} \right), \] (24)

where the means are \( m_{\alpha_1} = - \text{Re} \left( \bar{d}^H (h_{1d}, h_{1r}) d (h_{2r}) \right) \) and \( m_{\alpha_2} = \text{Re} \left( \bar{d}^H (h_{1d}, h_{1r}) d (h_{2r}) \right) + \| d (h_{2r}) \|^2 \), and where \( \tilde{d} (h_{1d}, h_{1r}) = g (\theta_d + h_{1d}, \theta_r + h_{1r}) - g (\theta_d, \theta_r) \).

### 5. SIMULATION

To compare the proposed bound with other, one use the same observation model as in [5] (frequency estimation). Consequently, \( g (\theta_d, \theta_r) = \theta_d b (\theta_r) \) where \( \theta_d \) is the amplitude and \( b (\theta_r) = 1 + e^{i \theta_r} + e^{2i \theta_r} + \ldots + e^{(P-1) i \theta_r} \) is a normalised cisoid with angular frequency \( \theta_r \). The scenario is the following: \( P = 32 \), \( \theta_d = 1 \) and \( \sigma_n^2 = \frac{1}{2} \). From [18], the HCRB is \( 2 \times 2 \) diagonal matrix with entries \( \{ \text{HCRB} \}_{1,1} = \frac{\sigma_n^2}{2 \theta_d} \) and \( \{ \text{HCRB} \}_{2,2} = \left( \frac{2 \sigma_n^2}{\pi} \frac{P (P + 1)}{2 (P + 1)} - P^2 \right) + \frac{1}{\sigma_n^2} \). The HBB, the HBWWB which is given in [5] and the proposed bound are computed with \( h_1 \in [-1, 1] \times 4 \) where the sampling interval for the first component is \( \delta h_{1d} = 0.01 \) and \( h_2 \in [0, 4] \times \left[ -1, \frac{1}{2} \right] \) where the sampling interval for the second component is \( \delta h_{2r} = \frac{1}{8} \). Last, the MAPMLE is obtained by searching the best candidate \( s \in [0, 2] \) and \( \theta_r \in \left[ -\frac{1}{2}, \frac{1}{2} \right] \) maximizing the joint p.d.f. \( f (x, \theta_r, \theta_d, \theta_1, \theta_2) \). The empirical MSE of the MAPMLE is assessed with 1000 Monte-Carlo trials. We only plot on the figure (1) the HCRB, the HBB,

![Fig. 1. Comparison of MSE hybrid lower bounds versus SNR](image-url)
7. APPENDIX

If $\hat{\theta}$ is an estimator of $\theta$, from [20, p. 124], under some mild regularity assumptions and for any real-valued vector $v$ with finite second order moment (1) holds. The general expression of matrices involved in (1) are $C = E_X, \theta \left[ (\hat{\theta} - \theta) v^T \right]$ and $V = E_X, \theta \left[ vv^T \right]$. Note that (1) does not currently lead to a lower bound on the MSE since $C$ depends on $\hat{\theta}$. However, let us set $\nu = \left( v_d, v_r \right)^T$ where
\[
v_d = \begin{cases} \frac{f(X, \theta + h_1)}{f(X, \theta)}, & \text{if } \theta \in \left( \theta : f(X, \theta) > 0, X \in \Omega \right) \\ 0, & \text{else} \end{cases}
\]
and
\[
v_r = \min \left( \frac{f(X, \theta + h_2)}{f(X, \theta)}, 1 \right) - \min \left( \frac{f(X, \theta - h_2)}{f(X, \theta)}, 1 \right) \text{ for all } h_1 = (h_{1d}, h_{1r})^T \text{ and } h_2 = (0, h_{2r})^T.\]
By the definition of $V = E_X, \theta \left[ vv^T \right]$, this choice of $v_d$ and $v_r$ leads to the matrix $V$ given in (2), (3) and (4) without major mathematical difficulties. In order to provide a new hybrid lower bound independent of estimation scheme, we have to prove that $C$ does not depend on $\hat{\theta}$. It has already been proved that $\{C\}_{1,1}$ and $\{C\}_{2,1}$ do not depend on $\hat{\theta}$ in [5] (see Appendix) under the aforementioned assumptions 2) and 3).

Before calculating $\{C\}_{1,2}$ and $\{C\}_{2,2}$, we give a preliminary result: for any real-valued function $l(X, \theta_d)$ defined on $\Omega \times \Pi_d$ and for any $h = (0, h_r)^T$ where $h_r \in \Pi_r$, one has
\[
\int_{\Pi_r} l(X, \theta_d) \left( \min \left( \frac{f(X, \theta + h)}{f(X, \theta)}, 1 \right) - \min \left( \frac{f(X, \theta - h)}{f(X, \theta)}, 1 \right) \right) f(X, \theta) d\theta_r = \int_{\Pi_r} l(X, \theta_d) \left( \min \left( \frac{f(X, \theta + h)}{f(X, \theta)}, 1 \right) - \min \left( \frac{f(X, \theta - h)}{f(X, \theta)}, 1 \right) \right) f(X, \theta) d\theta_r. \tag{25}
\]
Note that
\[
\int_{\Pi_r} \left( \min \left( \frac{f(X, \theta + h)}{f(X, \theta)}, 1 \right) - \min \left( \frac{f(X, \theta - h)}{f(X, \theta)}, 1 \right) \right) d\theta_r = \left\{ \begin{array}{c} \int_{\Pi_r} \min \left( \frac{f(X, \theta + h)}{f(X, \theta)}, 1 \right) d\theta_r - \int_{\Pi_r} \min \left( f(X, \theta + h), f(X, \theta) \right) d\theta_r, \\ - \int_{\Pi_r} \min \left( f(X, \theta - h), f(X, \theta) \right) d\theta_r \end{array} \right. \tag{26}
\]
Let us study the first integral. By substituting $\theta' = \theta + h_r$, the integration domain is still $\Pi_r$ by assumption 1 and then,  
\[
\int_{\Pi_r} \min \left( f(X, \theta + h), f(X, \theta) \right) d\theta_r = \int_{\Pi_r} \min \left( f(X, \theta' + h_r), f(X, \theta + h_r) \right) d\theta'_r. \tag{27}
\]
Thus, using (27) into (25), one obtains
\[
\int_{\Pi_r} l(X, \theta_d) \left( \min \left( \frac{f(X, \theta + h)}{f(X, \theta)}, 1 \right) - \min \left( \frac{f(X, \theta - h)}{f(X, \theta)}, 1 \right) \right) f(X, \theta) d\theta_r = 0 \text{ a.e. } X \in \Omega \text{ and for every } \theta_d \in \Pi_d. \tag{28}
\]
Remarks:
- This result is similar to condition (1) in [15] with the slight difference that the joint PDF depends on $\theta_d$.
- If we chose $h = (h_{1d}, h_{1r})$ with $h_{1d} \neq 0$ in this preliminary result, then (28) would depend on $X$. Consequently, we would find that $\{C\}_{1,2}$ and $\{C\}_{2,2}$ would depend on $\hat{\theta}$. Therefore we use $h_2 = (0, h_{2r})^T$.

Now, concerning $\{C\}_{1,2}$, one has
\[
\{C\}_{1,2} = E_{X, \theta} \left[ (\hat{\theta}_d - \theta_d) v_r \right] = \int_{\Pi_r} (\hat{\theta}_d - \theta_d) \left( \min \left( \frac{f(X, \theta + h_1)}{f(X, \theta)}, 1 \right) - \min \left( \frac{f(X, \theta - h_1)}{f(X, \theta)}, 1 \right) \right) f(X, \theta) d\theta, dX = 0,
\]
and for every $l(X, \theta_d) = \hat{\theta}_d - \theta_d$.

Finally, concerning $\{C\}_{2,2}$, one has
\[
\{C\}_{2,2} = E_{X, \theta} \left[ (\hat{\theta}_r - \theta_r) v_r \right] = \int_{\Pi_r} (\hat{\theta}_r - \theta_r) \left( \min \left( \frac{f(X, \theta + h_2)}{f(X, \theta)}, 1 \right) - \min \left( \frac{f(X, \theta - h_2)}{f(X, \theta)}, 1 \right) \right) f(X, \theta) d\theta, dX = 0.
\]

By substitution $\theta'_r = \theta_r - h_{2r}$, the integration domain for $\theta'_r$ is still $\Pi_r$ by assumption 1 and we have
\[
\int_{\Pi_r} \theta'_r \min \left( f(X, \theta - h_2), f(X, \theta) \right) d\theta'_r = \int_{\Pi_r} \theta'_r \min \left( f(X, \theta'_r + h_{2r}), f(X, \theta'_r) \right) d\theta'_r \tag{30}
\]
Thus, plugging (31) into (29), one has
\[
\{C\}_{2,2} = h_{2r} E_{X, \theta} \left[ \min \left( \frac{f(X, \theta + h_2)}{f(X, \theta)}, 1 \right) \right]. \tag{32}
\]
Consequently, one has proved that the proposed choice of function $v$ leads to a matrix $C$ which does not depend on $\hat{\theta}$.

8. REFERENCES
