A Useful Form of the Abel Bound and Its Application to Estimator Threshold Prediction

Alexandre Renaux, L. Najjar-Atallah, Philippe Forster, Pascal Larzabal

A Useful Form of the Abel Bound and Its Application to Estimator Threshold Prediction

Alexandre Renaux, *Student Member, IEEE*, Leïla Najjar-Atallah, Philippe Forster, *Member, IEEE*, and Pascal Larzabal, *Member, IEEE*

Abstract—This correspondence investigates the Abel bound in order to predict the estimators mean square error (mse) threshold effect. A tractable and computationally efficient form of this bound is derived. This form combines the Chapman–Robbins and the Cramér–Rao bounds. This bound is applied to a data-aided carrier frequency estimation problem for which a closed-form expression is provided. An indicator of the signal-to-noise ratio threshold is proposed. A comparison with recent results on the Barankin bound (Chapman–Robbins version) shows the superiority of the Abel-bound version to predict the mse threshold without increasing the computational complexity.

Index Terms—Abel bound (AB), carrier synchronization, minimal bounds on mean square error (mse), performance analysis, threshold prediction.

I. INTRODUCTION

Analysis of estimators performance in terms of mean square error (mse) is of interest in many fields such as RADAR, digital communications, etc. For this purpose, the lower bounds on the mse give the fundamental limit that an estimator can expect to achieve. The most famous lower bound is the Cramér–Rao Bound (CRB) [1] due to its relative simplicity. But, when the scenario is critical [low signal-to-noise ratio (SNR) and/or low number of observations], the CRB is too optimistic. Indeed, when the observation time and/or the SNR decrease, we observe a performance breakdown of the estimator mse due to the appearance of outliers [2], [3], which are not handled by the CRB. Nevertheless, the knowledge of the SNR value or of the number of observations for which this breakdown occurs (the so-called threshold) is fundamental to characterize the performance of an estimator. A tighter bound than the CRB is the Barankin bound [4], which is the greatest lower bound on the variance of any unbiased estimator. Unfortunately, most estimators exhibit an unknown bias when the SNR decreases. Consequently, the Barankin bound and, more generally, deterministic bounds are not fully appropriate for investigating the ultimate performances for all SNR values. However, the Barankin bound is well known to exhibit the threshold phenomena and has been used for it. But it requires a multidimensional optimization over a set of r test points (with $r \to \infty$) leading to a huge computational cost. A useful alternative is given by the Chapman–Robbins bound (ChRB) [5] for which the set of test points is reduced to one ($r = 1$). It has already been used to predict the threshold, for example, in spectral analysis [6], in time-delay estimation [7], [8], or in direction of arrival estimation [9].

This paper focuses on the Abel bound (AB) [10] in order to predict the SNR threshold. This bound, denoted $AB_{m,r}$, combines the

Barankin bound with r test points and the mth-order Bhattacharyya bound [11]. As the Barankin bound, a nonlinear optimization of the AB over the set of r test points is needed, requiring approximately the same huge computational cost. This paper presents a simplified version of the AB with $m = r = 1$. This bound, $AB_{1,1}$, is a mixture between the CRB and the ChRB and provides a better prediction of the threshold than the ChRB does. The computational complexity of $AB_{1,1}$ is discussed and some low complexity variants are provided.

In order to illustrate the interest of the proposed bound, a data-aided (DA) synchronization application is presented. In this framework, for which bounds such as CRB or ChRB have been previously applied in [12]–[14], a closed-form expression of $AB_{1,1}$ is provided. A simple indicator of the SNR threshold is deduced from the optimal test point position.

The notational convention adopted is as follows: italic indicates a scalar quantity, as in a or A; lower case boldface indicates a vector quantity, as in $\mathbf{a}$; upper case boldface indicates a matrix quantity, as in $\mathbf{A}$. $\Re\{A\}$ is the real part of A. The complex conjugation of a quantity is indicated by a superscript $*$ as in A^*. The matrix transpose is indicated by a superscript T as in $\mathbf{A}^T$, and the complex conjugate plus matrix transpose is indicated by a superscript H as in $\mathbf{A}^H = (\mathbf{A}^T)^*$. The nth row and mth column element of the matrix $\mathbf{A}$ will be denoted by $\{\mathbf{A}\}_{n,m}$. $\odot$ denotes the Hadamard product (element by element product). $E_{\boldsymbol{\theta}}[\cdot]$ denotes the expectation operator with respect to the probability density function $p(\mathbf{x}; \boldsymbol{\theta})$. $\|\cdot\|$ denotes the norm and $|\cdot|$ the absolute value.

II. ABEL BOUND

A. Background on the AB

In [10], Abel derives an original bound which is a mixture of the Bhattacharyya and the Barankin bounds. Let us set $\mathbf{x}$ a random observations vector. Let $\mathbf{T}(\mathbf{x})$ an estimator of a vector function $\mathbf{f}(\boldsymbol{\theta})$ depending on a deterministic parameters vector $\boldsymbol{\theta}$. Let $\boldsymbol{\beta}(\boldsymbol{\theta})$ the bias of $\mathbf{T}(\mathbf{x})$ such that $E_{\boldsymbol{\theta}}[\mathbf{T}(\mathbf{x})] = \mathbf{f}(\boldsymbol{\theta}) + \boldsymbol{\beta}(\boldsymbol{\theta}) \overset{def}{=} \boldsymbol{\tau}(\boldsymbol{\theta})$. Let $\text{mse}\{\mathbf{T}(\mathbf{x})\} \overset{def}{=} E_{\boldsymbol{\theta}}[(\mathbf{T} - \mathbf{f}(\boldsymbol{\theta}))(\mathbf{T} - \mathbf{f}(\boldsymbol{\theta}))^T]$. Let $\boldsymbol{\eta}(\mathbf{x})$ a given matrix-valued function.

The key point of Abel's work is the following covariance inequality [10] (6) based on an inequality presented in [15, p. 123] ([10, (5)])

$$\text{mse}\{\mathbf{T}(\mathbf{x})\} \geq \boldsymbol{\beta}(\boldsymbol{\theta})\boldsymbol{\beta}^T(\boldsymbol{\theta}) + E_{\boldsymbol{\theta}}\left[(\mathbf{T} - E_{\boldsymbol{\theta}}[\mathbf{T}])\,\boldsymbol{\eta}^T\right]\left(E_{\boldsymbol{\theta}}[\boldsymbol{\eta}\boldsymbol{\eta}^T]\right)^{-1}$$
$$\times E_{\boldsymbol{\theta}}\left[\boldsymbol{\eta}\,(\mathbf{T} - E_{\boldsymbol{\theta}}[\mathbf{T}])^T\right]. \quad (1)$$

By choosing, in an appropriate manner, the matrix $\boldsymbol{\eta}(\mathbf{x})$ in the covariance inequality (1), Abel has unified some classical lower bounds.

More precisely, by choosing

$$\boldsymbol{\eta}_1(\mathbf{x}) = \frac{1}{p(\mathbf{x}; \boldsymbol{\theta})}\left[\frac{\partial p(\mathbf{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \quad \frac{\partial^2 p(\mathbf{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\otimes 2}} \quad \cdots \quad \frac{\partial^m p(\mathbf{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\otimes m}}\right]^T \quad (2)$$

where $\partial^s / \partial \boldsymbol{\theta}^{\otimes s}$ is the vector of all partial derivatives $\partial^s / \partial \theta_{i_1}\partial\theta_{i_2}\cdots\partial\theta_{i_s}$, the right-hand-side of (1) becomes the mth-order Bhattacharyya bound.

On the other hand, by choosing

$$\boldsymbol{\eta}_2(\mathbf{x}) = \left[\frac{p(\mathbf{x}; \boldsymbol{\theta}_1)}{p(\mathbf{x}; \boldsymbol{\theta})} - 1 \quad \frac{p(\mathbf{x}; \boldsymbol{\theta}_2)}{p(\mathbf{x}; \boldsymbol{\theta})} - 1 \quad \cdots \quad \frac{p(\mathbf{x}; \boldsymbol{\theta}_r)}{p(\mathbf{x}; \boldsymbol{\theta})} - 1\right]^T \quad (3)$$

where $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_r$ are a set of vectors (called test vectors) such that $\boldsymbol{\theta}_i \in \Theta$ the parameter space, the right-hand-side of (1) becomes the rth-order Barankin bound (r test points).

The AB is obtained by mixing the mth-order Bhattacharyya bound and the rth-order Barankin bound. The Bhattacharyya bound accounts for the estimates near the true parameter value and the Barankin bound

accounts for the estimates over the full range of parameters values. Therefore, Abel's idea is to combine the two bounds in order to obtain a bound which accounts for both local and large errors. Consequently, the AB represents a generalization of these two bounds. Indeed, by choosing $\boldsymbol{\eta}(\mathbf{x})$ as a concatenation of $\boldsymbol{\eta}_1(\mathbf{x})$ and $\boldsymbol{\eta}_2(\mathbf{x})$ such that $\boldsymbol{\eta}(\mathbf{x}) = [\boldsymbol{\eta}_1^T(\mathbf{x}) \ \boldsymbol{\eta}_2^T(\mathbf{x})]^T$, the covariance inequality (1) leads to the AB

$$\mathbf{AB}_{m,r}(\boldsymbol{\theta}_0) = \mathbf{BB}_m + \left[\boldsymbol{\alpha}\mathbf{K}_m^{-1}\mathbf{L} - \boldsymbol{\Phi}_{\{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_r\}}\right]$$
$$\times \left[\tilde{\mathbf{J}}_{\{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_r\}} - \mathbf{L}^T\mathbf{K}_m^{-1}\mathbf{L}\right]^{-1} \left[\boldsymbol{\alpha}\mathbf{K}_m^{-1}\mathbf{L} - \boldsymbol{\Phi}_{\{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_r\}}\right]^T \quad (4)$$

where $\{\boldsymbol{\theta}_i \text{ for } i = 1,\ldots,r\}$ is a set of parameter vectors for which the bias is specified and

$$\begin{cases} \boldsymbol{\alpha} = \left[\frac{\partial \boldsymbol{\tau}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \quad \frac{\partial^2 \boldsymbol{\tau}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{T\otimes 2}} \quad \cdots \quad \frac{\partial^m \boldsymbol{\tau}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{T\otimes m}}\right] \\ \{\mathbf{K}_m\}_{i,j} = E_{\boldsymbol{\theta}}\left[\frac{1}{p^2(\mathbf{x};\boldsymbol{\theta})}\frac{\partial^i p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\otimes i}}\frac{\partial^j p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{T\otimes j}}\right] \\ \quad \text{(Bhattacharyya information matrix)} \\ \mathbf{BB}_m = \boldsymbol{\beta}(\boldsymbol{\theta})\boldsymbol{\beta}^T(\boldsymbol{\theta}) + \boldsymbol{\alpha}\mathbf{K}_m^{-1}\boldsymbol{\alpha}^T \quad \text{(Bhattacharyya bound)} \\ \boldsymbol{\Phi}_{\{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_r\}} = [\boldsymbol{\tau}(\boldsymbol{\theta}_1) - \boldsymbol{\tau}(\boldsymbol{\theta}) \quad \boldsymbol{\tau}(\boldsymbol{\theta}_2) - \boldsymbol{\tau}(\boldsymbol{\theta}) \quad, \ldots, \\ \quad\quad\quad\quad\quad \boldsymbol{\tau}(\boldsymbol{\theta}_r) - \boldsymbol{\tau}(\boldsymbol{\theta})] \\ \left\{\tilde{\mathbf{J}}_{\{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_r\}}\right\}_{i,j} = E_{\theta_i}\left[\frac{p(\mathbf{x};\boldsymbol{\theta}_j)}{p(\mathbf{x};\boldsymbol{\theta})}\right] - 1 \\ \quad \text{(Barankin information matrix)} \\ \{\mathbf{L}\}_{i,j} = E_{\theta_j}\left[\frac{1}{p(\mathbf{x};\boldsymbol{\theta})}\frac{\partial^i p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\otimes i}}\right]. \end{cases} \quad (5)$$

Let us note that $\boldsymbol{\Phi}_{\{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_r\}}\tilde{\mathbf{J}}_{\{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_r\}}^{-1}\boldsymbol{\Phi}_{\{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_r\}}^T$ is the Barankin bound. Then, the AB only depends on the Bhattacharyya bound, on the Barankin bound and on an hybrid matrix $\mathbf{L}$.

B. Useful Form of the AB: $AB_{1,1}$

The sequel will focus on the estimation of a single parameter θ_0: $f(\theta_0) = \theta_0$. Moreover, the bias is assumed to be zero: $\beta(\theta_0) = 0$. Consequently, $\boldsymbol{\alpha} = [1 \ 0 \ \cdots \ 0]$. As the Barankin bound, a nonlinear optimization of the AB over the set of r test points is needed, requiring approximately the same huge computational cost. This paper presents a simplified version of the AB with $m = r = 1$.

Let us now detail the $AB_{1,1}$ expression. By setting for the single test point $\theta_1 = \theta_0 + h$, we have in (5)

$$\begin{cases} \alpha = 1 \\ BB_1 = K_1^{-1} = \text{CRB}(\theta_0) \\ \Phi_{\{\theta_1\}} = h \\ \tilde{J}_{\{\theta_1\}} = \int_{\Omega} \frac{p^2(\mathbf{x};\theta_1)}{p(\mathbf{x};\theta_0)}d\mathbf{x} - 1 \\ L(h,\theta_0) = \int_{\Omega} \frac{\partial \ln p(\mathbf{x};\theta)}{\partial \theta}\bigg|_{\theta=\theta_0} p(\mathbf{x};\theta_1)d\mathbf{x} \end{cases} \quad (6)$$

where $\text{CRB}(\theta_0)$ is the CRB, $p(\mathbf{x};\theta)$ is the likelihood of the observations, and Ω is the observations space. Let us denote by $\text{ChRB}(h,\theta_0)$ the ChRB without maximization over h. The last term $\tilde{J}_{\{\theta_1\}}$ can be connected straightforwardly to $\text{ChRB}(h,\theta_0)$ by

$$\tilde{J}_{\{\theta_1\}} = \frac{h^2}{\text{ChRB}(h,\theta_0)}. \quad (7)$$

By using (6) in (4), one obtains the $AB_{1,1}$

$$AB_{1,1}(h,\theta_0) = \text{CRB}(\theta_0) + \frac{(\text{CRB}(\theta_0)L(h,\theta_0) - h)^2}{\frac{h^2}{\text{ChRB}(h,\theta_0)} - L^2(h,\theta_0)\text{CRB}(\theta_0)}$$
$$= \frac{\text{CRB}^{-1}(\theta_0) + \text{ChRB}^{-1}(h,\theta_0) - 2\Psi(h,\theta_0)}{\text{CRB}^{-1}(\theta_0)\text{ChRB}^{-1}(h,\theta_0) - \Psi^2(h,\theta_0)} \quad (8)$$

where $\Psi(h,\theta_0) = L(h,\theta_0)/h$. This bound only depends on the CRB, on the ChRB without maximization over h, and on an additive term $\Psi(h,\theta_0)$. Then, in an estimation problem, if the CRB and the ChRB are available, the evaluation of the AB requires only the computation of $\Psi(h,\theta_0)$. For example, let us now focus on the widely spread Gaussian observation model with parameterized mean

$$\mathbf{x} = \mathbf{m}(\theta_0) + \mathbf{n} \quad (9)$$

where $\mathbf{x}$ is the $N \times 1$ complex observation vector, θ is the real unknown deterministic parameter of interest with true value θ_0, $\mathbf{m}$ is an $N \times 1$ complex deterministic vector depending (generally nonlinear) of θ, and $\mathbf{n}$ is the $N \times 1$ additive noise vector. The noise is assumed to be complex, circular, Gaussian with zero mean and with covariance matrix $\sigma^2\mathbf{I}_M$. For this model, which is widely used in the literature (spectral analysis, array processing, wireless communications, etc.), the form of $\Psi(h,\theta_0)$ is shown in Appendix I to be

$$\Psi(h,\theta_0) = \frac{2}{h\sigma^2}\Re\left\{(\mathbf{m}(\theta_0 + h) - \mathbf{m}(\theta_0))^H \frac{\partial \mathbf{m}(\theta)}{\partial \theta}\bigg|_{\theta=\theta_0}\right\}. \quad (10)$$

In the original paper, Abel gives some ways to choose the test points location. In our case, since the bound is greatly simplified, h will be directly chosen to maximize (8)

$$AB_{1,1}(\theta_0) = \sup_h \frac{\text{CRB}^{-1}(\theta_0) + \text{ChRB}^{-1}(h,\theta_0) - 2\Psi(h,\theta_0)}{\text{CRB}^{-1}(\theta_0)\text{ChRB}^{-1}(h,\theta_0) - \Psi^2(h,\theta_0)}. \quad (11)$$

Let us now compare the $AB_{1,1}$, the ChRB, and the CRB. For that purpose, the CRB can be interpreted as the minimum variance of an estimator which has both null bias and null bias derivative at the true value of the parameter θ_0 [16], [17]. In the same way, the ChRB is the supremum over h of the minimum variance of an estimator which has both null bias at θ_0 and $\theta_1 = \theta_0 + h$. Finally, the $AB_{1,1}$ is the supremum over h of the minimum variance of an estimator which has the three following constraints: a null bias at θ_0, a null bias derivative at θ_0, and a null bias at $\theta_1 = \theta_0 + h$. From the increase of constraints, it follows that:

$$AB_{1,1}(\theta_0) \geq \text{ChRB}(\theta_0) \geq \text{CRB}(\theta_0). \quad (12)$$

Consequently, the threshold of $AB_{1,1}(\theta_0)$ will be at least the same as the ChRB threshold.

III. APPLICATION TO THE DA FREQUENCY ESTIMATION PROBLEM

We will now apply the proposed bound to a DA frequency estimation problem, and we will give a way to obtain a threshold predictor based on this bound.

A. Observations Model

Consider a linearly modulated signal, obtained by applying a known data symbol sequence $\mathbf{a} = [a_0,\ldots,a_{N-1}]^T$ taken from a unit energy constellation to a square-root Nyquist transmit filter. The signal is transmitted over an additive white Gaussian noise channel. The output signal is sampled at the symbol rate which yields the observations

$$r_k = \rho a_k e^{j(k\theta_0 + \phi)} + n_k, \qquad k = 0,\ldots,N-1 \quad (13)$$

where $\{n_k\}$ is a sequence of i.i.d, circular $(E[\mathbf{n}\mathbf{n}^T] = \mathbf{0})$, zero mean complex Gaussian noise variable with variance 1, and ρ^2 denotes the SNR which is assumed to be known. The unknown deterministic parameter θ with true value $\theta_0 \in]-\pi,\pi]$ corresponds to the carrier angular pulsation offset. ϕ is the carrier phase and is assumed to be known

and compensated for (i.e., $\phi = 0$). Based on the data model (13), the likelihood of r_k is

$$p(r_k; \theta) = \frac{1}{\pi} e^{-\|r_k - \rho a_k e^{jk\theta}\|^2}. \tag{14}$$

Then, the likelihood of the observation vector $\mathbf{r} = [r_0 \cdots r_{N-1}]^T$ is given by

$$p(\mathbf{r}; \theta) = \prod_{k=0}^{N-1} p(r_k; \theta) = \frac{1}{\pi^N} e^{-\mathbf{r}^H \mathbf{r} - \rho^2 \mathbf{a}^H \mathbf{a} + 2\rho \mathrm{Re}(\mathbf{r}^H(\mathbf{a} \odot \mathbf{b}))} \tag{15}$$

where $\mathbf{b} = [1, e^{j\theta}, \ldots, e^{j(N-1)\theta}]^T$. In the sequel, we give closed form expressions of CRB, ChRB, and AB.

B. Bounds for the Frequency Estimation Problem

1) CRB and ChRB: For this problem, the CRB and the ChRB are given by [14]

$$\mathrm{CRB}(\theta_0) = \frac{1}{2\rho^2 \sum_{k=0}^{N-1} \|a_k\|^2 k^2} \tag{16}$$

and

$$\mathrm{ChRB}(\theta_0) = \sup_{0 \le h \le \pi} \mathrm{ChRB}(h, \theta_0)$$
$$= \sup_{0 \le h \le \pi} \frac{h^2}{e^{4\rho^2 \sum_{k=0}^{N-1} \|a_k\|^2 (1-\cos(kh))} - 1}. \tag{17}$$

As presented in [14], the ChRB gives an approximation of the ML threshold. As we will see in simulations, this approximation will be improved by using the AB.

2) AB: Here, the CRB and ChRB are available, and the AB only requires the evaluation of $\Psi(h, \theta_0)$. For our synchronization problem, a closed form of $\Psi(h, \theta_0)$ is (see Appendix II)

$$\Psi(h, \theta_0) = \frac{2\rho^2}{h} \sum_{k=0}^{N-1} \|a_k\|^2 k \sin(kh). \tag{18}$$

Using (16)–(18), in (11) leads to (19) shown at the bottom of the page.

It is interesting to note that the CRB, the ChRB, and the AB do not depend on θ_0 in this problem.

The tightest AB is given by an optimization over the test point h on the parameter support. Here, since $\mathrm{AB}_{1,1}(-h, \theta_0) = \mathrm{AB}_{1,1}(h, \theta_0)$, this optimization reduces to one over $[0, \pi]$. For a given SNR, let us set h_{opt} the value for which $\mathrm{AB}_{1,1}(\theta_0) = \mathrm{AB}_{1,1}(h_{\mathrm{opt}}, \theta_0)$. As we see in Fig. 1, for a different number of observations and a BPSK modulation, h_{opt} versus SNR has a very specific behavior. Indeed, when the AB starts to differ from the CRB, h_{opt} jumps from a value in the neighborhood of 0 to a value in the neighborhood of π. For lower bounds obtained by an optimization over test points, as the Barankin bound or the Weiss–Weinstein bound, it is known that this behavior is related to the ambiguity function of the problem. Indeed, the value of the test
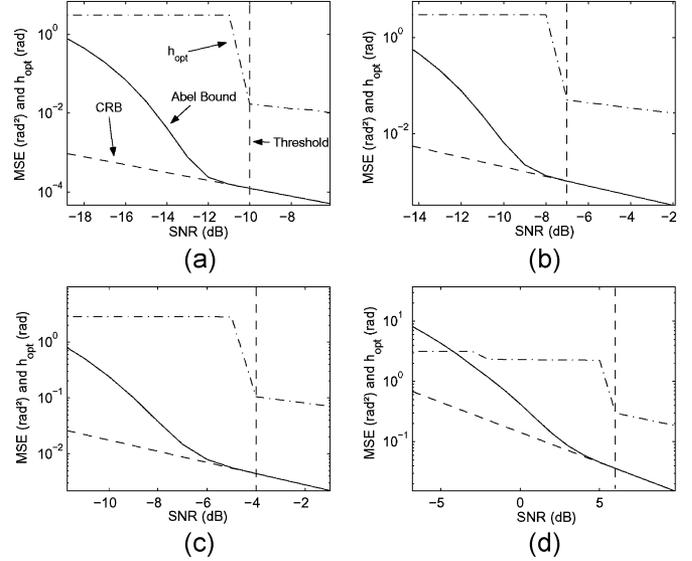

Fig. 1. Behavior comparison between the CRB (dashed line), the AB (solid line), and h_{opt} (dashed–dotted line). The y-axis scale represents: the mse (in radians squared) for the CRB, the mse (in radians squared) for the AB, and the pulsation value (in radians) for h_{opt}. (a) $N = 50$ observations. (b) $N = 20$ observations. (c) $N = 10$ observations. (d) $N = 4$ observations.

points which maximizes the bound are those for which the ambiguity function takes local maxima [9], [18]–[21].

Here, a first-order Taylor expansion of the pdf $p(\mathbf{r}; \theta + h)$ around $h = 0$ leads to

$$p(\mathbf{r}; \theta + h) = p(\mathbf{r}; \theta) + h \frac{\partial p(\mathbf{r}; \theta)}{\partial \theta} + o(h). \tag{20}$$

By using this expression in $\mathrm{ChRB}(h, \theta_0)$ and $\Psi(h, \theta_0)$, we obtain

$$\begin{cases} \lim_{h \to 0} \mathrm{ChRB}(h, \theta_0) = \mathrm{CRB}(\theta_0) \\ \lim_{h \to 0} \Psi(h, \theta_0) = \mathrm{CRB}^{-1}(\theta_0). \end{cases} \tag{21}$$

By reporting (21) in (8)

$$\lim_{h \to 0} \mathrm{AB}_{1,1}(h, \theta_0) = \mathrm{CRB}(\theta_0). \tag{22}$$

Consequently, at high SNR, where the AB achieves the CRB, $h \to 0$ and the estimates move within the main lobe of the ambiguity function.

In the threshold area, the value of h, which maximizes the right-hand-side of (19), is located on the last side lobe of the ambiguity function. This behavior is due to the fact that, since the side lobes of the ambiguity function have almost the same level, they are sensitive in the same manner to the noise. Then, due to the term h^2 in the numerator of (19), this is the furthest ambiguity which contributes more to the bound.

Finally, when the SNR is weak, the noise contribution is more important than the ambiguity function and the optimum value of h becomes

$$\mathrm{AB}_{1,1}(\theta_0) = \sup_{0 \le h \le \pi} \frac{2h^2 \rho^2 \sum_{k=0}^{N-1} \|a_k\|^2 k^2 + e^{4\rho^2 \sum_{k=0}^{N-1} \|a_k\|^2 (1-\cos(kh))} - 4h\rho^2 \sum_{k=0}^{N-1} \|a_k\|^2 k \sin(kh) - 1}{\left(2\rho^2 \sum_{k=0}^{N-1} \|a_k\|^2 k^2\right)\left(e^{4\rho^2 \sum_{k=0}^{N-1} \|a_k\|^2 (1-\cos(kh))} - 1\right) - 4\rho^4 \left(\sum_{k=0}^{N-1} \|a_k\|^2 k \sin(kh)\right)^2} \tag{19}$$
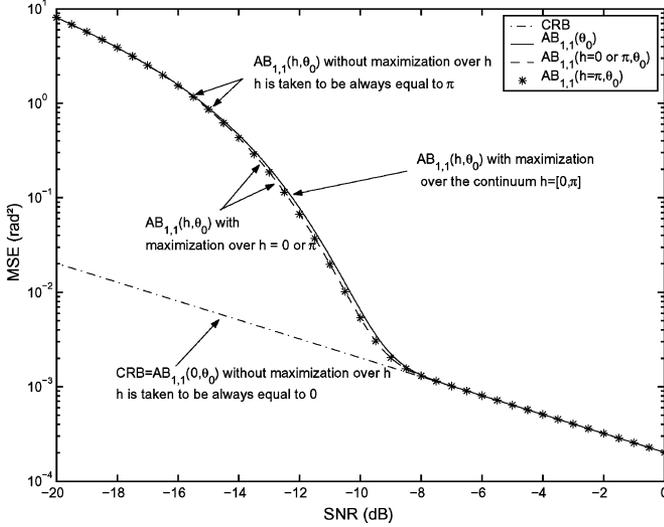
Fig. 2. Comparison of different variants of the AB for BPSK modulation with $N = 20$ observations.

Fig. 3. Maximum likelihood estimator mse, CRB, and AB versus SNR. $N = 10$ observations.

the limit of the parameter support π. These facts are particularly visible in Fig. 1(d), since for $N = 4$ observations, the only side lobe local maxima is not close to π.

C. SNR Threshold Predictor

Starting from (19), this section gives a method for predicting the threshold. This predictor is based on the monitoring of the value of h, which maximizes $AB_{1,1}(h, \theta_0)$.

Therefore, in the light of the optimal test point behavior described herein, we will define the SNR threshold predictor $\hat{\rho}_{\text{thres}}^2$ as

$$\hat{\rho}_{\text{thres}}^2 = \arg \max_{\rho^2} \left| \frac{\partial h_{\text{opt}}}{\partial \rho^2} \right|. \tag{23}$$

$\hat{\rho}_{\text{thres}}^2$ can be computed very easily from the AB. As expected, we note in Fig. 1 that the SNR threshold increases when the number of samples decreases.

D. Towards a Reduction of the AB Computational Cost

Since a closed form of h_{opt} is not available, the bound proposed in (19) requires an optimization which increases the computational cost. Nevertheless, the behavior of h_{opt} observed herein allows us to reduce the AB to an optimization only over two values: $AB_{1,1}(\theta_0) \approx \sup_{h=\{0,\pi\}} AB_{1,1}(h, \theta_0)$. Obviously, in this case, the obtained bound is weaker than the bound of (19), but the computational cost is largely reduced since the optimization is taken over only two points and not over a continuum of test points.

We can still reduce the computational cost by noting that the CRB part of the AB directly monitors the small values of h. This leads to: $AB_{1,1}(\theta_0) \approx AB_{1,1}(\pi, \theta_0)$. In this case, $\Psi(\pi, \theta_0) = 0$ and the AB is given by

$$AB_{1,1}(\pi, \theta_0) = CRB(\theta_0) + ChRB(\pi, \theta_0)$$

$$= \frac{1}{2\rho^2 \sum_{k=0}^{N-1} \|a_k\|^2 k^2}$$

$$+ \frac{\pi^2}{\frac{4\rho^2 \sum_{k=0}^{N-1} \|a_k\|^2 (1-(-1)^k)}{e} - 1}. \tag{24}$$

Fig. 2 superimposes $AB_{1,1}(\theta_0)$, $AB_{1,1}(0, \theta_0)$ (which is the CRB), the bound given by $\sup_{h=\{0,\pi\}} AB_{1,1}(h, \theta_0)$ and $AB_{1,1}(\pi, \theta_0)$ for $N = 20$ observations and a BPSK modulation. The last two bounds are very close and slightly weaker than the best bound $AB_{1,1}(\theta_0)$. Note that the computational cost of $AB_{1,1}(\pi, \theta_0)$ is the same as for the CRB.

IV. SIMULATION

The aim of this section is to examine the relevance of the derived bound for predicting the threshold in the DA frequency estimation problem. For this, the mse of the maximum likelihood algorithm (ML) is reported. Based on the likelihood given by (15), maximizing the observation likelihood is equivalent to searching for

$$\hat{\theta} = \arg \max_{\theta} \text{Re} \left(\sum_{k=0}^{N-1} r_k^* a_k e^{jk\theta} \right). \tag{25}$$

The simulation is performed with a BPSK pilot sequence which contains $N = 10$ symbols.

Fig. 3 superimposes the mse of ML (25) evaluated over 1000 Monte Carlo trials, the CRB (16), and the AB (19). This figure shows the threshold behavior of the ML estimator when the SNR decreases. A comparison with recent results on the Barankin bound (Chapman–Robbins version) [14] is also given. For the ChRB and the AB, the SNR threshold predictor is obtained by (23). The suggested AB version provides a better tool to predict the threshold without increasing the computation complexity. The predicted threshold value provides a good approximation of the effective SNR at which the ML estimator experiences the threshold behavior.

V. CONCLUSION

In this correspondence, the AB on the mse has been investigated. Under some assumptions, we simplified the original bound to obtain a bound which gives a better threshold prediction than the ChRB for slightly higher complexity. We also have shown that the AB can be expressed as a function of the CRB, the ChRB, and an additive term $\Psi(h, \theta_0)$. A closed form expression for $\Psi(h, \theta_0)$ in a widely met Gaussian observation model has been obtained. These general results have been applied to a digital communication frequency estimation problem. In this context, a theoretical closed form of the AB has been provided. Finally, as a by-product, a new SNR threshold indicator has been derived.

APPENDIX I
$\Psi(h, \theta_0)$ FOR A GAUSSIAN PARAMETERIZED MEAN OBSERVATION MODEL

The likelihood of the observations is

$$p(\mathbf{x}; \theta) = \frac{1}{(\pi\sigma)^N} e^{-\frac{1}{\sigma^2}\|\mathbf{x}-\mathbf{m}(\theta)\|^2}. \tag{26}$$

Then, $\Psi(h, \theta_0)$ can be rewritten

$$\Psi(h, \theta_0) = \frac{1}{h}\int_\Omega \left.\frac{\partial \ln p(\mathbf{x}; \theta)}{\partial \theta}\right|_{\theta=\theta_0} p(\mathbf{x}; \theta_1)\, d\mathbf{x}$$

$$= -\frac{1}{h\sigma^2} E_{\theta_1}\left[\left.\frac{\partial \|\mathbf{x}-\mathbf{m}(\theta)\|^2}{\partial \theta}\right|_{\theta=\theta_0}\right]. \tag{27}$$

Since

$$\frac{\partial \|\mathbf{x}-\mathbf{m}(\theta)\|^2}{\partial \theta} = -2\Re\left\{(\mathbf{x}-\mathbf{m}(\theta))^H \frac{\partial \mathbf{m}(\theta)}{\partial \theta}\right\} \tag{28}$$

then

$$\Psi(h, \theta_0) = \frac{2}{h\sigma^2}\Re\left\{E_{\theta_1}\left[(\mathbf{x}-\mathbf{m}(\theta_0))^H \left.\frac{\partial \mathbf{m}(\theta)}{\partial \theta}\right|_{\theta=\theta_0}\right]\right\}$$

$$= \frac{2}{h\sigma^2}\Re\left\{E_{\theta_1}\left[(\mathbf{x}-\mathbf{m}(\theta_0))^H\right]\left.\frac{\partial \mathbf{m}(\theta)}{\partial \theta}\right|_{\theta=\theta_0}\right\}$$

$$= \frac{2}{h\sigma^2}\Re\left\{(\mathbf{m}(\theta_0+h)-\mathbf{m}(\theta_0))^H \left.\frac{\partial \mathbf{m}(\theta)}{\partial \theta}\right|_{\theta=\theta_0}\right\}. \tag{29}$$

APPENDIX II
$\Psi(h, \theta_0)$ EVALUATION FOR DA FREQUENCY ESTIMATION PROBLEM

$\Psi(h, \theta_0)$ is given by (29). Here, $\sigma^2 = 1$ and $\mathbf{m}(\theta) = \rho(\mathbf{a}\odot\mathbf{b}(\theta))$. Then

$$\Psi(h, \theta_0) = \frac{2}{h}\Re\left\{(\rho(\mathbf{a}\odot\mathbf{b}(\theta_0+h))-\rho(\mathbf{a}\odot\mathbf{b}(\theta_0)))^H\right.$$

$$\left.\times \left.\frac{\partial\rho(\mathbf{a}\odot\mathbf{b}(\theta))}{\partial \theta}\right|_{\theta=\theta_0}\right\}$$

$$= \frac{2\rho^2}{h}\Re\left\{(\mathbf{a}\odot(\mathbf{b}(\theta_0+h)-\mathbf{b}(\theta_0)))^H\right.$$

$$\left.\times \left(\mathbf{a}\odot \left.\frac{\partial(\mathbf{b}(\theta))}{\partial \theta}\right|_{\theta=\theta_0}\right)\right\}$$

$$= \frac{2\rho^2}{h}\Re\left\{\sum_{k=0}^{N-1}\|a_k\|^2\left(e^{-jk(\theta_0+h)}-e^{-jk\theta_0}\right)\right.$$

$$\left.\times \left.\frac{\partial e^{jk\theta}}{\partial \theta}\right|_{\theta=\theta_0}\right\}$$

$$= \frac{2\rho^2}{h}\sum_{k=0}^{N-1}\|a_k\|^2 k\,\Re\left\{j\left(e^{-jkh}-1\right)\right\}$$

$$= \frac{2\rho^2}{h}\sum_{k=0}^{N-1}\|a_k\|^2 k\,\sin(hk). \tag{30}$$

REFERENCES

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

[2] H. L. VanTrees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968, vol. 1.

[3] D. C. Rife and R. R. Boorstyn, "Single tone parameter estimation from discrete time observations," *IEEE Trans. Inf. Theory*, vol. 20, no. 5, pp. 591–598, Sep. 1974.

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

[5] D. G. Chapman and H. Robbins, "Minimum variance estimation without regularity assumptions," *Ann. Math. Stat.*, vol. 22, pp. 581–586, 1951.

[6] L. Knockaert, "The Barankin bound and threshold behavior in frequency estimation," *IEEE Trans. Signal Process.*, vol. 45, no. 9, pp. 2398–2401, Sep. 1997.

[7] R. J. McAulay and L. P. Seidman, "A useful form of the Barankin lower bound and its application to PPM threshold analysis," *IEEE Trans. Inf. Theory*, vol. 15, no. 2, pp. 273–279, Mar. 1969.

[8] A. Zeira and P. Schultheiss, "Realizable lower bounds for time delay estimation: Threshold phenomena," *IEEE Trans. Signal Process.*, vol. 42, no. 5, pp. 1001–1007, May 1994.

[9] I. Reuven and H. Messer, "The use of the Barankin bound for determining the threshold SNR in estimating the bearing of a source in the presence of another," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 1995, pp. 1645–1648.

[10] J. S. Abel, "A bound on mean square estimate error," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1675–1680, Sep. 1993.

[11] A. Bhattacharrya, "On some analogues of the amount of information and their use in statistical estimation," *Sankhya Indian J. Stat.*, vol. 8, pp. 1–14, 1946.

[12] Y. Jiang, F. Sun, and J. Baras, "On the true Cramer Rao lower bound for the DA joint estimation of carrier phase and timing offsets," in *Proc. IEEE Int. Conf. Commun.*, 2000, pp. 331–335.

[13] ——, "On the performance limits of data aided synchronization," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 191–203, Jan. 2003.

[14] L. Atallah, J. P. Barbot, and P. Larzabal, "SNR threshold indicator in data aided frequency synchronization," *IEEE Signal Process. Lett.*, vol. 11, no. 8, pp. 652–654, Aug. 2004.

[15] E. L. Lehmann, *Theory of Point Estimation*. New York: Wiley, 1983.

[16] F. E. Glave, "A new look at the Barankin lower bound," *IEEE Trans. Inf. Theory*, vol. 18, no. 3, pp. 349–356, May 1972.

[17] P. Forster and P. Larzabal, "On lower bounds for deterministic parameter estimation," presented at the IEEE Int. Conf. Acoust., Speech, Signal Process., Orlando, FL, 2002.

[18] W. Xu, A. B. Baggeroer, and C. D. Richmond, "Bayesian bounds for matched-field parameter estimation," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3293–3305, Dec. 2004.

[19] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of a hybrid parameter vector," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1084–1093, May 1997.

[20] J. Tabrikian and J. Krolik, "Barankin bounds for source localization in an uncertain ocean environment," *IEEE Trans. Signal Process.*, vol. 47, no. 11, pp. 2917–2927, Nov. 1999.

[21] W. Xu, "Performances bounds on matched-field methods for source localization and estimation of ocean environmental parameters," Ph.D. dissertation, Dept. Ocean Eng., Massachusetts Inst. Technol., Cambridge, MA, Jun. 2001.