Performance Bounds for Change Point and Time Delay Estimation

Chengfang Ren*, Alexandre Renaux†, Sylvain Azarian*
*SONDRA, CentraleSupélec, 91192 Gif-sur-Yvette, France
†LSS, Université Paris-Sud, 91192 Gif-sur-Yvette, France

Abstract—This paper investigates performance bounds for joint estimation of signal change point and time delay estimation between two receivers. This problem is formulated as an estimation of the beginning of a well known message and the time shift of its arrival on two receivers from a noisy source. This scenario could be viewed as an extension of the classical time delay estimation [1], [2] with an additional change on the transmitted message. A theoretical derivation of the Barankin bound and a simplified version of this bound are proposed in order to predict the Maximum Likelihood Estimator (MLE) behavior. Simulations illustrate the validity of both bounds but it is pointed out that the MLE seems asymptotically efficient for the normalized time shift estimation contrary to the estimation of message’s starting time.

Index Terms—Performance bounds, Change point estimation, Time delay estimation

I. INTRODUCTION

Time delay estimation between two spatially separated sensors is fundamental for localization and tracking problems in radar, acoustics and ultrasonic applications [2], [3]. It turns out that this problem could also be interesting for physically separated sensors synchronization [4]. Previous studies on this problem always assume that the shape of reference signal did not change over time. But this hypothesis might not be verified for an opportunity source and only a portion of the reference signal can be exploited. To solve this problem, our method will be based on retrospective change point estimation with an additional parameter (time delay) to estimate. Change point estimation/detection problems have been widely investigated in the literature e.g. [5] for a general introduction and [6] for the specific offline change point estimation. In terms of performance analysis, the first asymptotic results concerning the MLE for very specific problems have been proposed in [7] and recently reanalyzed in [8] and [9]. The term “asymptotic” generally refers to the case where a large number of observations are collected before and after the change and then the asymptotic behavior of MLE in terms of probability density function (pdf) can be derived for most of the time. However, from a practical point of view, these results are currently hard to be useful. This is why the signal processing community has tried to bypass this drawback by focusing on lower bound on the Mean Square Error (MSE) and not directly on the pdf of the MLE.

In the context of change point estimation, the discrete nature of parameters to estimate implies that the regularity conditions are not satisfied in order to apply the classical Cramér-Rao bound [10], [11] similarly to the case of cyclic parameters estimation [12] or when the support of data pdf depends on the parameters to estimate [13]. Then, other bounds requiring less regularity conditions have been proposed to overcome this difficulty. The Chapman-Robbins bound has been derived in [14] in the context of one change point and extended to the multiple change point problem in [15]. In the Bayesian context, the Weiss-Weinstein bound has been studied in [16]–[18]. In this paper, we stay in the non-Bayesian context to study the Barankin (or McAulay-Seidman bound) [19], [20] for a change point estimation problem when (contrary to previous works) two sets of non synchronized data are available. Consequently, one has to additionally estimate the time delay between both receivers. We first start by presenting the mathematical framework of our analysis. Then, we derive the Barankin bound and propose also a simplified hybrid bound which is found to be surprisingly simple to compute. Finally, some simulation results are presented in order to show the interest of the proposed bounds.

II. MODEL SETUP

The application context of our paper is trying to synchronize multiple receivers on a well known message $m(t)$ (sinus, chirp, etc...) sent by a well located opportunity source $e(t)$ (could be a radio station, communication channel, etc...). For readability, we restrict our analysis to two receivers denoted $r_1$ and $r_2$. The geometry of both emitter and receivers are known, so we can easily compensate the radial distance difference between $e$ to $r_1$ and $e$ to $r_2$ by introducing a time delay on the received signals. Nevertheless, in order to simplify the analytic expression of the proposed bound, these two sensors are assumed to be equidistantly located from the opportunity source. The main goal here is to take advantage of this well known message $m(t)$ in order to synchronize our receivers. Therefore, the synchronization tasks consist in three steps: detect $m(t)$ then estimate exactly the starting time of this message. This time will be denoted by $\tau$ in the following. And finally, use this message to synchronize $r_1$ and $r_2$ by estimating the normalized time shift denoted by $\Delta$ between the received signals which could be viewed as the time delay between receivers. Since our goal is trying to establish performance bounds on the estimation of $\tau$ and $\Delta$, we will skip the detection part. Mathematically, we can formulate the estimation problem as following:
Assume that we have already detected a single message \( m(t) \) received by \( r_1 \) and \( r_2 \) between the time interval \([0; NT_e]\) where \( T_e \) is the common sampling interval of \( r_1 \) and \( r_2 \). We only assume phase jump between receivers’ clock and no frequency drift, otherwise it will be useless to synchronize them. Additionally, the sampling interval should respect the Nyquist-Shannon sampling theorem, i.e. \( 2BT_e < 1 \), where \( B \) is the bandwidth of \( m(t) \). This condition is sufficient for not loosing information on \( m(t) \) when sampling. Therefore, we can model the received signal as

\[
    r_1[k] = s(kT_e) + n_1[k], \quad k = 1, \ldots, N
\]

for the first receiver and

\[
    r_2[k] = s(kT_e + \Delta T_e) + n_2[k], \quad k = 1, \ldots, N
\]

for the second receiver where \( s(t) \) is the noiseless signal of form

\[
    s(t) = \begin{cases} 
    0, & \text{if } t \leq \tau \\
    m(t), & \text{if } t > \tau, 
    \end{cases}
\]

\( N \) is the total number of observations, \( \Delta T_e \) represents the fact that these two receivers are not synchronized \( (0 < \Delta < 1) \) and the receivers noise are represented by \( n_i[k], \ i = 1, 2, \) assumed to be independent and identically distributed (i.i.d.).

This model is similar to the classical time delay estimation problem proposed in [1]–[4] but with an additional parameter on the reference signal change. Indeed, the shape of the reference signal is assumed to be not static. Therefore the receivers should be synchronized on the portion of reference signal that provides useful information on the time shift between receivers signal. Additionally, the time delay considered is on the receivers sampling time and not on the (noiseless version of) received signal itself, this is why we can restrict the estimation of \( \Delta \) between 0 and 1.

Consequently, the pdfs of both receivers signal can be formulated as:

\[
    r_i[k] \sim \begin{cases} 
    f_1(r_i[k]) & \text{if } k \leq \nu \\
    f_2(r_i[k]) & \text{if } k > \nu 
    \end{cases}
\]

with \( \nu = \left\lfloor \frac{\tau}{T_e} \right\rfloor \) (\( \lfloor . \rfloor \) denotes the floor function). In order to simplify notations, we will work with normalized frequency which is equivalent to take the sampling period \( T_e = 1 \). One can also regroup equations (1) and (2) into a single formula \( r(t) = s(t) + n(t) \) such that \( r(k) = r_1[k] \) and \( r(k + \Delta) = r_2[k] \) with \( n(t) \) a strict stationary signal with the same pdf than the pdf of \( n_1[k] \) and \( n_2[k] \) (\( \overset{d}{=} \) means equality in distribution).

An illustration of the aforementioned problem is shown in figure (1) where the message \( m(t) \) is sinusoidal with a signal change at \( \tau = 0.00015 \) and a sampling shift of \( \Delta = 0.5 \) between both receivers. The goal will be to retrieve parameters \( \tau \) and \( \Delta \) from collected data \( r_1[k] \) and \( r_2[k] \).

Consequently, the received signals are modeled by:

\[
    r_1[k] \sim \begin{cases} 
    f_1(r(k)) & \text{if } k = 1, \ldots, \nu \\
    f_2(r(k)) & \text{if } k = \nu + 1, \ldots, N 
    \end{cases}
\]

\[
    r_2[k] \sim \begin{cases} 
    f_1(r(k + \Delta)) & \text{if } k = 1, \ldots, \nu \\
    f_2(r(k + \Delta)) & \text{if } k = \nu + 1, \ldots, N 
    \end{cases}
\]

III. PROPOSED BOUNDS

A. Relation to prior work

On the one hand, performance bounds have recently been developed for change point estimation [14], [15]. However, this bound did not take into account the estimation of an additional continuous parameter (time delay between receivers). On the other hand, time delay estimation with a change on the reference signal has not been investigated. Therefore, we establish in this paper a lower bound on the MSE in order to predict the accuracy of both parameters: time when signal pattern changes called “change point” and time delay between receivers. Since the time delay is estimated over a compact interval, we treat here an hybrid estimation problem. Consequently, the proposed bound differs from existing lower bounds in the literature.

B. General bound

The Barankin bound is the tightest lower bound on the variance of any unbiased estimator. In this paper, the bound is denoted \( BB(\theta) \) and is given by [20, Eqn. 2]:

\[
    BB(\theta) = \sup_{h_1, h_2} [h_1 h_2] D^{-1}(\theta, h)[h_1 h_2]^T
\]
with $h_1 = [h_\nu 0]^T$ and $h_2 = [0 h_\Delta]^T$. The so-called test-points $h_\nu$ and $h_\Delta$ are chosen such that $\nu + h_\nu \in \{2, \ldots, T - 1\}$ and $\Delta + h_\Delta \in [0; 1]$. The elements of the matrix $D(\theta, h)$ are defined as follows:

$$D_{1, i}(\theta, h) = \mathbb{E} \left[ \left( \frac{f(r; \theta + h_1)}{f(r; \theta)} \right)^2 \right] - 1, \ i = 1, 2 \quad (8)$$

and

$$D_{1, 2}(\theta, h) = D_{2, 1}(\theta, h) = \mathbb{E} \left[ \frac{f(r; \theta + h_1) f(r; \theta + h_2)}{f^2(r; \theta)} \right] - 1. \quad (9)$$

Since all component of observations vector $r$ are independent, thus its pdf is given by:

$$f(r; \theta) = \prod_{k=1}^\nu f_1(r(k)) f_2(r(k + \Delta)) \times \prod_{k=\nu+1}^N f_2(r(k)) f_2(r(k + \Delta)) \quad (10)$$

After some cumbersome calculus, one obtains

$$D_{1, 2}(\theta, h) + 1 = \prod_{k=\nu+1}^{\nu+h_\nu} \int \frac{f_1^2(r(k + \Delta))}{f_2^2(r(k + \Delta))} f_2(r(k + \Delta + h_\Delta)) dr(k + \Delta) \quad (12)$$

Let us set

$$D_{1, 1}(\theta, h) = \mathbb{E} \left[ \frac{f^2(r; \theta + h_1)}{f(r; \theta)} \right] - 1, \ \text{it is almost the previous one with } h_2 = h_1, \ \text{therefore one has}$$

$$D_{1, 1}(\theta, h) + 1 = \prod_{k=\nu+1}^{\nu+h_\nu} \int \frac{f_1^2(r(k + \Delta))}{f_2^2(r(k + \Delta))} f_2(r(k + \Delta + h_\Delta)) dr(k + \Delta) \quad (13)$$

Finally, $D_{2, 2}(\theta, h) = \mathbb{E} \left[ \left( \frac{f(r; \theta + h_1)}{f(r; \theta)} \right)^2 \right] - 1$ is obtained similar by previous calculus

$$D_{2, 2}(\theta, h) + 1 = \prod_{k=1}^\nu \int \frac{f_1^2(r(k + \Delta + h_\Delta))}{f_1^2(r(k + \Delta))} dr(k + \Delta) \times \prod_{k=\nu+1}^N \int \frac{f_2^2(r(k + \Delta + h_\Delta))}{f_2^2(r(k + \Delta))} dr(k + \Delta). \quad (14)$$

This expression can be applied to any problem by specifying the probability density functions $f_1$ and $f_2$. We next study the Gaussian case.

**C. Gaussian case**

Let us assume that $n_i[k], i = 1, 2$ are complex circular Gaussian with zero mean and variance $\sigma^2$. Then the pdf $f_1(r(k))$ and $f_2(r(k))$ are

$$f_1(r(k)) = \frac{1}{\pi \sigma^2} e^{-\frac{1}{2\sigma^2} ||r(k)||^2}, \ f_2(r(k)) = f_1(r(k) - m(k)) \quad (15)$$

By plugging the above expressions in $D_{1, 1}(\theta, h)$, one obtains

$$D_{1, 1}(\theta, h) + 1 = \prod_{k=\nu+1}^{\nu+h_\nu} \frac{1}{\pi \sigma^2} \int \left( e^{-\frac{1}{2\sigma^2} (||r(k)||^2 + ||r(k + \Delta)||^2)} \right) e^{\frac{1}{\sigma^2} (||r(k) - m(k)||^2 + ||r(k + \Delta) - m(k + \Delta)||^2)} dr(k) dr(k + \Delta). \quad (16)$$

By noticing that $\forall t$

$$\int \frac{1}{\pi \sigma^2} e^{-\frac{1}{2\sigma^2} ||r(t)||^2 + \frac{1}{\sigma^2} ||r(t) - m(t)||^2} dr(t) = e^{2 \frac{m(t)^2}{\sigma^2}},$$

one can obtain a simple expression for $D_{1, 1}(\theta, h)$

$$D_{1, 1}(\theta, h) = e^{\frac{1}{\sigma^2} \sum_{k=\nu+1}^{\nu+h_\nu} (||m(k)||^2 + ||m(k + \Delta)||^2)} - 1. \quad (17)$$

Let us set

$$||r(t)||^2 + 2 \text{Re}(r^*(t) (m(t) - m(t + h_\Delta))) \quad (18)$$

$$= ||r(t) + m(t)||^2 - ||m(t) - m(t + h_\Delta)||^2,$$

then

$$\int \frac{f_1(r(t))}{f_2(r(t))} f_2(r(t + h_\Delta)) dr(t) = \frac{1}{\pi \sigma^2} (||m(t) - m(t + h_\Delta)||^2 - ||m(t + h_\Delta)||^2),$$

and finally

$$D_{1, 1}(\theta, h) + 1 = \frac{1}{\pi \sigma^2} \sum_{k=\nu+1}^{\nu+h_\nu} \left( ||m(k + \Delta) - m(k + \Delta + h_\Delta)||^2 \right) \quad (19)$$

The last expression of $D_{2, 2}(\theta, h)$ is obtained by noticing that

$$\int \frac{f_1^2(r(t + h_\Delta))}{f_1^2(r(t))} dr(t) = 1,$$

and that

$$\int \frac{f_2^2(r(t + h_\Delta))}{f_2^2(r(t))} dr(t) = \frac{1}{\pi \sigma^2} \int \left( e^{-\frac{1}{2\sigma^2} (||r(t)||^2 + ||m(t)||^2)} \times e^{-\frac{1}{\sigma^2} (||r(t) - m(t)||^2 + ||m(t + h_\Delta)||^2)} \right) dr(t). \quad (20)$$
By letting
\[
\| r(t) \|^2 + 2 \Re (r^*(t)(m(t) - 2m(t + h_\Delta))) = \| r(t) + m(t) - 2m(t + h_\Delta) \|^2 - \| m(t) - 2m(t + h_\Delta) \|^2,
\]
one obtains
\[
\int \frac{f_2(r(t) + h_\Delta)}{f_2(r(t))} dr(t) = e^{\frac{1}{\nu} \left( \| m(t) - 2m(t + h_\Delta) \|^2 - 2 \| m(t + h_\Delta) \| \| m(t) \| \right)},
\]
and finally
\[
D_{2,2}(\theta, \mathbf{h}) = \frac{1}{\nu} \sum_{k=\nu+1}^{\nu+\nu|p|} \| m(k + \Delta) - m(k + \Delta + h_\Delta) \|^2
- \| m(k + \Delta + h_\Delta) \|^2 + \| m(k + \Delta) \|^2
= \sum_{k=\nu+1}^{\nu+\nu|p|} - 2h_\Delta \Re \left( m(k + \Delta) \frac{\partial m(k + \Delta)}{\partial \Delta} \right) + o(h_\Delta),
\]
and
\[
\sum_{k=\nu+1}^{\nu+\nu|p|} \| m(k + \Delta) - 2m(k + \Delta + h_\Delta) \|^2 - 2 \| m(k + \Delta + h_\Delta) \|^2 + \| m(k + \Delta) \|^2
= \sum_{k=\nu+1}^{\nu+\nu|p|} 2h_\Delta \left( \frac{\partial m(k + \Delta)}{\partial \Delta} \right)^2 + o(h_\Delta^2).
\]

By using these expressions into \( D(\theta, \mathbf{h}) \), one obtains a "hybrid" Cramér-Rao-Barankin bound, denoted \( \text{BB}(\nu, \Delta) \), which is very simple. This bound will be given in the next Section for a particular function \( m(t) \).

**IV. SIMULATIONS**

Let’s consider two receivers for which their receiving data respectively follow the model given in equations (1) and (2). This model assumes that receivers are equidistant from the opportunity source. Otherwise, a preprocessing step is necessary to compensate the time delay introduced by the radial distance gap from the opportunity source to each receiver.

The sent message is of sinus form given by \( m[k] \equiv m(kT_c) = e^{j2\pi f_0 k}e^{j2\pi f_0 k} \) where \( f_0 = fT_c \) is the normalized frequency. In order to estimate message arrival and the normalized time shift, \( \nu \Delta \), we perform the MLE given by
\[
\hat{\theta} = \arg \max_{\nu, \Delta} f(\mathbf{r}; \nu, \Delta)
\]
where \( \mathbf{r} \) is given in equation (6) with
\[
f(\mathbf{r}; \nu, \Delta) = \prod_{k=1}^{\nu} f_1(r(k)) f_2(r(k + \Delta)) \prod_{l=\nu+1}^{N} f_2(r(l)) f_2(r(l + \Delta)).
\]

Additionally, if we assume that both receiver noise follow a i.i.d. zero mean Gaussian distribution \( i.e. n_1[k] = n_2[k] \sim \mathcal{N}(0, \sigma^2) \), then, after some calculus, the ML criterion is given by the following expression:
\[
\hat{\nu}, \hat{\Delta} = \arg \max_{\nu, \Delta} \nu + \sum_{k=\nu+1}^{\nu+\nu|p|} \Re \left( e^{-j2\pi f_0 k} (r(k) + r(k + \Delta)e^{-j2\pi f_0 \Delta}) \right).
\]

Since parameter \( \nu \) is discrete and \( \Delta \) belongs to \([0; 1]\), the implementation of the MLE is performed by maximizing the above criterion over a grid search \([1; N]\) \times [0; 1] with a step \( \delta_\Delta = 10^{-3} \) for parameter \( \Delta \).

![Fig. 2. Comparison of MLE MSE, BB and BB for estimation of \( \Delta \). The xaxis is the amplitude of the MSE expressed in the square unit of the normalized time.](image-url)
the formula developed in Section (III-D), our simplified bound is given by:

\[
\text{BB}(\nu, \Delta) = \begin{bmatrix} \sup_{h_\nu} \frac{h_\nu^2 \sigma^2}{e^{\frac{h_\nu^2 \sigma^2}{2} - 1}} & 0 \\ 0 & \frac{\sigma^2}{8\pi^2 f_0^2 (N-\nu)} \end{bmatrix}.
\]  

We compare the MSE of the MLE to the Barankin bound and its simplified form for estimation of \(\Delta\) and \(\nu\) in the following setup: \(f_0 = 0.1, N = 45, \nu = 25\) and \(\Delta = 0.4\). The Barankin bound is the maximization of the criterion in (7) over a grid search with steps \(\delta h_\nu = 1\) and \(\delta h_\Delta = 10^{-3}\) and the simplified bound is obtained with the same step for \(\nu\) but does not depend on \(\Delta\). The MSE of the MLE is estimated with 1000 Monte-Carlo trials. In the figure (2), the Barankin bound and its simplified form give the same result since the estimation problem w.r.t. \(\Delta\) is linear. The performance of MLE estimate for \(\Delta\) is well predicted by both bounds. In figure (3), the MSE of the MLE for \(\nu\) is lower bounded by both Barankin bound and its simplified form. The simplified form gives almost the same performance than the Barankin with a lower computational cost. However, the prediction is not accurate when the SNR becomes higher than 15dB.

This difference is probably explained by the fact that \(\nu\) is a discrete parameter. Indeed, it is proved in [22] that there is no asymptotic efficiency of the MLE when estimating discrete parameters from an i.i.d. sequence. Here we are in the more constrained case where the observations are independent but not identically distributed and consequently, there is no reason that the MLE becomes efficient.

V. CONCLUSION

In this paper, we established the Barankin bound and a simplified version of this bound for a synchronization problem where an opportunity source is used. The MLE’s accuracy for the normalized time shift is well predicted by our proposed bounds. Therefore, the MLE seems asymptotically efficient for estimation of \(\Delta\). However, the prediction of MLEs MSE still need to be improved for parameter \(\nu\) because of its asymptotic inefficiency for discrete parameter estimation.

REFERENCES