BARANKIN BOUND FOR MULTIPLE CHANGE-POINT ESTIMATION

Patricio La Rosa, Alexandre Renaux, Arye Nehorai

BARANKIN BOUND FOR MULTIPLE CHANGE-POINT ESTIMATION

Patricio S. La Rosa, Alexandre Renaux, and Arye Nehorai

Department of Electrical and Systems Engineering,
Washington University in St. Louis, St. Louis, MO 63130, USA.

ABSTRACT

We derive the Barankin bound on the mean-squared error for multiple change-point estimation of an independent measurement sequence. We first derive a general form of this bound and give the structure of the so-called Barankin information matrix (BIM). We show that the BIM for the change-point parameters has a tri-diagonal structure which means that one change-point estimation depends on its neighboring change points. Using this result, we propose a computationally efficient inversion algorithm of the BIM. As an illustration, we analyze the case of changes in the mean vector of a Gaussian distribution.

Index Terms— Multiple change-point estimation, performance analysis, Barankin lower bounds on the mean-squared error.

1. INTRODUCTION

Estimation of changes in time series is an important research area with several applications, *e.g.*, speech processing, medical imaging, and econometrics. The literature concerning estimation algorithms for change-point estimation (see, *e.g.*, [1]) is abundant. However, less work has been done concerning the performance of such algorithms in terms of mean-square error (MSE).

We derive the Barankin bound (BB) [2] on the MSE for multiple change-point estimation for an independent measurement sequence. In contrast to the Cramér-Rao bound (CRB) [3], the BB is computable despite the discrete nature of the change-point parameter and regularity assumptions on the likelihood of the observations [4]. However, the BB requires the use of free parameters called test points, and, in order to obtain the best (*e.g.*, the tightest) bound, a nonlinear maximization over these test points has to be performed.

To the best of our knowledge, a particular case of the BB (the so-called Hammersley-Chapman-Robbins bound, [4, 5]) has already been studied only in the context of one change-point estimation in the foundational communication of Ferrari and Tourneret [6]. Here we extend the results in [6] to multiple change points. Particularly, we show that the estimation of one change point is corrupted by its neighboring change points and we give the details of the computation for the two change-point case which allows us to propose an efficient inversion algorithm of the Barankin information matrix. We apply our bound to the case of changes in the parameters of Gaussian observations and present one numerical example for comparing our bound to the performance achieved by the maximum likelihood estimator (MLE).

This work was supported by the Department of Defense under the Air Force Office of Scientific Research MURI Grant FA9550-05-0443 and AFOSR Grant FA 9550-05-1-0018.

2. PROBLEM FORMULATION

2.1. Observation model

We consider the general case of N independent vector observations $\mathbf{X} = [\mathbf{x}_1,\ \mathbf{x}_2,\ \ldots,\ \mathbf{x}_N]$, $\mathbf{X} \in \mathbb{R}^{M\times N}$, which can be obtained, for example, by a multiple sensor system and are modelled as follows:

$$\begin{cases} \mathbf{x}_i \sim p_1\left(\mathbf{x}_i;\boldsymbol{\eta}_1\right) \text{ for } i = 1,\ \ldots,\ t_1, \\ \mathbf{x}_i \sim p_2\left(\mathbf{x}_i;\boldsymbol{\eta}_2\right) \text{ for } i = t_1+1,\ \ldots,\ t_2, \\ \vdots \\ \mathbf{x}_i \sim p_{q+1}\left(\mathbf{x}_i;\boldsymbol{\eta}_{q+1}\right) \text{ for } i = t_q+1,\ \ldots,\ N. \end{cases} \tag{1}$$

where M is the size of the sample vector (*e.g.*, the number of sensors), q is the number of change-points, and p_j is a probability density function (or mass function for discrete random variables) with parameters $\boldsymbol{\eta}_j \in \mathbb{R}^L$. In other words,

$$\begin{aligned} \mathbf{x}_i \quad \sim \quad & p_j\left(\mathbf{x}_i;\boldsymbol{\eta}_j\right) \text{ for } i = t_{j-1}+1,\ \ldots,\ t_j, \\ & \text{with } j = 1,\ \ldots,\ q+1, \end{aligned} \tag{2}$$

where we define $t_0 = 0$ and $t_{q+1} = N$. Note that if $M = 1$, the problem is reduced to the estimation of changes in a time series. We assume that all probability density functions p_j belong to a common distribution. The unknown parameters of interest are the change-point locations $\{t_1, t_2, \ldots, t_q\}$ with $\{t_k \in \mathbb{N} - \{0\},\ k = 1,\ \ldots,\ q\}$. The observations between two consecutive change points are assumed to be stationary. Consequently, the $q \times 1$ vector of all unknown parameters for this model is

$$\mathbf{t} = [t_1, t_2, \ldots, t_q]\,. \tag{3}$$

Note that, since we are focused on the change-point estimation, we assume that the parameters $\boldsymbol{\eta}_j$ are known. The resulting bound will still be useful if these parameters are unknown, but more optimistic.

2.2. Barankin Bound

The P-order BB of a vector $\boldsymbol{\theta}_0 \in \mathbb{R}^K$, denoted by $\mathbf{BB}_P(\boldsymbol{\theta}_0)$, is given as follow (see [7] and [8, 9] for more details):

$$\mathbf{Cov}(\widehat{\boldsymbol{\theta}}) \geq \mathbf{BB}_P(\boldsymbol{\theta}_0) = \mathbf{H}(\boldsymbol{\Phi} - \mathbf{1}_{P\times P})^{-1}\mathbf{H}^T, \tag{4}$$

where $\mathbf{Cov}(\widehat{\boldsymbol{\theta}})$ is the covariance matrix of an unbiased estimator $\widehat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta}_0$. The notation $\mathbf{A} \geq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is a semi-positive definite matrix. The matrix $\mathbf{H}$ is a function of the set $\{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_P\}$, so-called "test points", left to the user, and it is given by:

$$\mathbf{H} = [\boldsymbol{\theta}_0 - \boldsymbol{\theta}_1,\ \ldots,\ \boldsymbol{\theta}_0 - \boldsymbol{\theta}_P]\,. \tag{5}$$

Here we define $\mathbf{h}_i = \boldsymbol{\theta}_0 - \boldsymbol{\theta}_i$ such that the matrix $\mathbf{H} \in \mathbb{R}^{K\times P}$ becomes $\mathbf{H} = [\mathbf{h}_1,\ \ldots,\ \mathbf{h}_P]$. Moreover, note that $\boldsymbol{\theta}_0 + \mathbf{h}_j \in \Theta$. In the following, for simplicity, we use the term "test point" for the vectors $\mathbf{h}_i$. Finally, $\boldsymbol{\Phi}$ is a $\mathbb{R}^{P\times P}$ matrix whose elements $[\boldsymbol{\Phi}]_{kl}$ are given by:

$$[\boldsymbol{\Phi}]_{kl} = \mathbb{E}[L(\mathbf{X}, \boldsymbol{\theta}_0, \mathbf{h}_k)L(\mathbf{X}, \boldsymbol{\theta}_0, \mathbf{h}_l)], \tag{6}$$

where $L(\mathbf{X}, \boldsymbol{\theta}_0, \mathbf{h}_j)$ is defined as follows:

$$L(\mathbf{X}, \boldsymbol{\theta}_0, \mathbf{h}_j) = \frac{p(\mathbf{X}; \boldsymbol{\theta}_0 + \mathbf{h}_j)}{p(\mathbf{X}; \boldsymbol{\theta}_0)}, \tag{7}$$

where $p(\mathbf{X}; \boldsymbol{\varphi})$ is the likelihood of the observations with parameter vector $\boldsymbol{\varphi}$. Note that the matrix $\boldsymbol{\Phi} - \mathbf{1}_{P\times P}$ is sometimes referred to as the Barankin information matrix (BIM) [10].

As already stated, test points are left to the user, since any set of test points in $\mathbf{BB}_P(\boldsymbol{\theta}_0)$ satisfies the inequality (4). Thus, the tightest BB, denoted by $\mathbf{BB}(\boldsymbol{\theta}_0)$, is given as follow:

$$\mathbf{BB}(\boldsymbol{\theta}_0) = \lim_{P\longrightarrow\infty} \sup_{\{\mathbf{h}_1,\ldots,\mathbf{h}_P\}} \mathbf{BB}_P(\boldsymbol{\theta}_0). \tag{8}$$

The solution of the above problem is computationally costly, since the limit on P implies that an infinite number of test points per parameter needs to be considered and a nonlinear maximization over the test points has to be performed. In the following we use a simplified version of the BB. In particular, for the parameters vector given in (3), *i.e.,* $\boldsymbol{\theta}_0 = \mathbf{t}$**,** we consider the classical assumption of one test point per parameter ($P = K = q$). Then, the structure of the matrix $\mathbf{H}$ is as follows:

$$\mathbf{H} \quad = \quad Diag\left([\alpha_1,\ \ldots,\ \alpha_q]^T\right), \tag{9}$$

where the vector $[\alpha_1,\ \ldots,\ \alpha_q]^T$ represents the set of test points associated to the parameters $\mathbf{t} = [t_1,\ t_2,\ \ldots,\ t_q]^T$. Note that $\alpha_j \neq 0$ is defined such that $t_j + \alpha_j$ ranges over all possible values of t_j, for $j = 1, \ldots, q$. Thus, $\alpha_j \in \{\mathbb{Z} \cap [t_{j-1} - t_j + 1, t_{j+1} - t_j - 1] - \{0\}\}$.

The matrix, $\boldsymbol{\Phi} - \mathbf{1}_{q\times q}$, corresponds to the BIM for change-point locations $\mathbf{t}$, denoted here by $\mathbf{BIM_t}$. In the following we will derive the elements of $\mathbf{BIM_t}$.

3. BARANKIN BOUND FOR MULTIPLE CHANGE-POINT ESTIMATION

To compute the BB for the change point localization parameters, we first need to compute $\mathbf{BIM_t}$, which depends on the matrix $\boldsymbol{\Phi}$. From Equations (6) and (7), the elements of $[\boldsymbol{\Phi}]_{kl}$, for $k, l = 1, \ldots, q$ are given by. Hence

$$[\boldsymbol{\Phi}]_{kl} = \int_\Omega \frac{p\left(\mathbf{X}; \boldsymbol{\theta}_0 + \mathbf{h}_k\right) p\left(\mathbf{X}; \boldsymbol{\theta}_0 + \mathbf{h}_l\right)}{p\left(\mathbf{X}; \boldsymbol{\theta}_0\right)} d\mathbf{X}. \tag{10}$$

where $p\left(\mathbf{X}; \mathbf{t}\right)$ is given by

$$p\left(\mathbf{X}; \mathbf{t}\right) = \prod_{i=1}^{t_1} p_1(\mathbf{x}_i; \boldsymbol{\eta}_1) \prod_{i=t_1+1}^{t_2} p_2(\mathbf{x}_i; \boldsymbol{\eta}_2) \cdots \prod_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i; \boldsymbol{\eta}_{q+1}), \tag{11}$$

and $p\left(\mathbf{X}; \mathbf{t} + \mathbf{h}_k\right)$ is given by

$$p\left(\mathbf{X}; \mathbf{t} + \mathbf{h}_k\right) \quad = \quad \prod_{i=1}^{t_1} p_1(\mathbf{x}_i; \boldsymbol{\eta}_1) \cdots \prod_{i=t_{k-1}+1}^{t_k+\alpha_k} p_k(\mathbf{x}_i; \boldsymbol{\eta}_k) \cdots \prod_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i; \boldsymbol{\eta}_{q+1}), \tag{12}$$

and where $p\left(\mathbf{X}; \mathbf{t} + \mathbf{h}_l\right)$ is same as Equation (12) ($k = l$).

In order to study $\boldsymbol{\Phi}$, we analyze its diagonal and non-diagonal elements separately.

3.1. Diagonal elements of $\boldsymbol{\Phi}$

Let $k = l$ in Equation (10). To simplify the analysis we consider the cases $\alpha_k > 0$ and $\alpha_k < 0$, obtaining the following expression:

$$[\boldsymbol{\Phi}]_{kk} = \begin{cases} \left(\int_\Omega \frac{p_k^2(\mathbf{x};\boldsymbol{\eta}_k)}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} d\mathbf{x}\right)^{\alpha_k}, & \text{if } \alpha_k > 0, \\ \left(\int_\Omega \frac{p_{k+1}^2(\mathbf{x};\boldsymbol{\eta}_{k+1})}{p_k(\mathbf{x};\boldsymbol{\eta}_k)} d\mathbf{x}\right)^{-\alpha_k}, & \text{if } \alpha_k < 0. \end{cases} \tag{13}$$

Remark: in the case of one change-point, *e.g.,* $l = k = 1$, Equation (13) is reduced to the result of Ferrari and Tournaret (see Equations (5) and (6) in [6]).

3.2. Non-diagonal elements of $\boldsymbol{\Phi}$

Following the same idea as for the diagonal elements, $[\boldsymbol{\Phi}]_{kl}$ for $k \neq l$ can be simplified by analyzing the four possible combinations of test-point ranges, namely,

$$\begin{cases} \alpha_k > 0 \text{ and } \alpha_l > 0, \\ \alpha_k < 0 \text{ and } \alpha_l < 0, \\ \alpha_k < 0 \text{ and } \alpha_l > 0, \\ \alpha_k > 0 \text{ and } \alpha_l < 0. \end{cases} \tag{14}$$

For the last case, *e.g.* $\alpha_k > 0$ and $\alpha_l < 0$, two subcases have to be analyzed: (i) $t_k + \alpha_k < t_l + \alpha_l$ and (ii) $t_k + \alpha_k > t_l + \alpha_l$. Note that since $k < l$, $t_k < t_l$ and since $\alpha_j \in \{\mathbb{Z} \cap [t_{j-1} - t_j + 1, t_{j+1} - t_j - 1] - \{0\}\}$, the subcase $t_k + \alpha_k > t_l + \alpha_l$ can appear only when $l = k+1$, or, in other words, when we are analyzing two neighboring change points. We will refer to this as the *overlapping* case. For the first three cases and subcase (i), Equation (10) becomes, after some calculus effort,

$$[\boldsymbol{\Phi}]_{kl} = 1. \tag{15}$$

For subcase (ii), keeping in mind that $\alpha_k > 0$ and $\alpha_{k+1} < 0$, Equation (10) becomes

$$[\boldsymbol{\Phi}]_{kl} = \begin{cases} \left(\int_\Omega \frac{p_k(\mathbf{x};\boldsymbol{\eta}_k) p_{k+2}(\mathbf{x};\boldsymbol{\eta}_{k+2})}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} d\mathbf{x}\right)^{\beta_{k,k+1}}, & \text{for } l = k+1, \\ 1, & \text{for } |k-l| > 1, \end{cases} \tag{16}$$

where $\beta_{k,k+1} = (t_k + \alpha_k) - (t_{k+1} + \alpha_{k+1})$.

3.3. Barankin information matrix $\boldsymbol{\Phi} - \mathbf{1}_{q\times q}$

Using Equations (13), (15), and (16), it is clear that $\mathbf{BIM_t}$ has at least a tri-diagonal structure:

$$\mathbf{BIM_t} = \begin{bmatrix} A_1 & B_1 & 0 & \cdots & 0 \\ B_1 & A_2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & A_{q-1} & B_{q-1} \\ 0 & \cdots & 0 & B_{q-1} & A_q \end{bmatrix}, \tag{17}$$

where

$$\begin{aligned} A_k \quad &= \quad [\boldsymbol{\Phi}]_{kk} - 1 \\ &= \quad \begin{cases} \left(\int_\Omega \frac{p_k^2(\mathbf{x};\boldsymbol{\eta}_k)}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} d\mathbf{x}\right)^{\alpha_k} - 1 & \text{if } \alpha_k > 0, \\ \left(\int_\Omega \frac{p_{k+1}^2(\mathbf{x};\boldsymbol{\eta}_{k+1})}{p_k(\mathbf{x};\boldsymbol{\eta}_k)} d\mathbf{x}\right)^{-\alpha_k} - 1 & \text{if } \alpha_k < 0, \end{cases} \end{aligned} \tag{18}$$

and

$$B_k = [\mathbf{\Phi}]_{kk+1} - 1 \quad (19)$$
$$= \begin{cases} 0, & \text{if } t_k + \alpha_k < t_{k+1} + \alpha_{k+1}, \\ \left(\int_\Omega \frac{p_k(\mathbf{x};\boldsymbol{\eta}_k) p_{k+2}(\mathbf{x};\boldsymbol{\eta}_{k+2})}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} d\mathbf{x} \right)^{\beta_{k,k+1}} - 1, \\ & \text{if } t_k + \alpha_k > t_{k+1} + \alpha_{k+1}. \end{cases}$$

Note also that the diagonal elements of $\mathbf{BIM_t}$ can be computed numerically in one step (*e.g.*, $\forall \alpha_k \gtrless 0$) as follows:

$$A_k = \left(\int_\Omega \left(\frac{p_k(\mathbf{x};\boldsymbol{\eta}_k)}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} \right)^{\epsilon_k} p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1}) d\mathbf{x} \right)^{|\alpha_k|} - 1, \quad (20)$$

where $\epsilon_k = \frac{1}{2}\left(3\frac{\alpha_k}{|\alpha_k|} + 1\right)$.

3.4. Barankin Bound computation

The next step of our analysis is to compute the BB for $\mathbf{t}$, $\mathbf{BB_t}$, given by

$$\mathbf{BB_t} = \mathbf{H} \left(\mathbf{BIM_t}\right)^{-1} \mathbf{H}, \quad (21)$$

where $\mathbf{H}$ is given by (9).

Two computational issues have to be addressed concerning the BB in general: the inversion of $\mathbf{BIM_t}$ and the multidimensional optimization of the the resulting bound over the test points.

Regarding the inversion of $\mathbf{BIM_t}$, for a given set of test points, it is clear that $t_k + \alpha_k > t_{k+1} + \alpha_{k+1} \Longrightarrow t_{k+1} + \alpha_{k+1} < t_{k+2} + \alpha_{k+2}$, since $\alpha_j \in \{\mathbb{Z} \cap [t_{j-1} - t_j + 1, t_{j+1} - t_j - 1] - \{0\}\}$. In other words, $\forall k$, if $B_k \neq 0$, then $B_{k+1} = B_{k-1} = 0$; therefore, $\mathbf{BIM_t}$ is block diagonal and the maximum size of one block is 2×2 leading to a straightforward inversion. Since the problem is reduced to finding, at worst, the inverse of several 2×2 matrices with the same structure, we propose to build the matrix $[\mathbf{BB_t}]$ using an "iterative" algorithm described in Fig. 1. It begins from the first change-point location parameter and proceeds inquiring on the existence of overlapping with the next segment. The matrix $\mathbf{\Gamma}$ is the general form of the BB for two neighboring change points obtained by inverting a block of size 2×2 in the case of overlapping, and is given as follows,

$$\mathbf{\Gamma} = \quad (22)$$
$$\frac{1}{\delta} \begin{bmatrix} \alpha_k^2 \left(\Delta^{|\alpha_{k+1}|}_{(k+2)(k+2)(k+1)} \right) & \alpha_k \alpha_{k+1} \left(-\Delta^{\beta_{k,k+1}}_{k(k+2)(k+1)} \right) \\ \alpha_k \alpha_{k+1} \left(-\Delta^{\beta_{k,k+1}}_{k(k+2)(k+1)} \right) & \alpha_{k+1}^2 \left(\Delta^{\alpha_k}_{kk(k+1)} \right) \end{bmatrix} \quad (23)$$

where

$$\delta = \Delta^{\alpha_k}_{kk(k+1)} \Delta^{|\alpha_{k+1}|}_{(k+2)(k+2)(k+1)} - \left(\Delta^{\beta_{k,k+1}}_{k(k+2)(k+1)} \right)^2, \quad (24)$$

and

$$\Delta^d_{ijk} = \left(\int_\Omega \frac{p_i(\mathbf{x};\boldsymbol{\eta}_i) p_j(\mathbf{x};\boldsymbol{\eta}_j)}{p_k(\mathbf{x};\boldsymbol{\eta}_k)} d\mathbf{x} \right)^d - 1. \quad (25)$$

On the other hand, when we have no overlapping, the scalar $\gamma_k = \alpha_k^2 / (A_k - 1)$ has to be computed, with A_k given by Equation (20).

Regarding the multidimensional optimization over the test points, in the continuous parameter case, the maximization is performed over a fine grid in order to find the optimum value test-point values. Here, the parameters are discrete, so the grid is already defined leading to an easier computation.

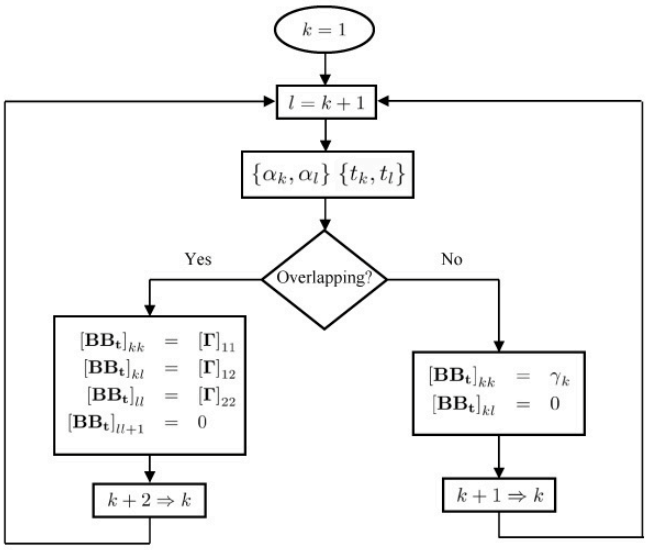

Fig. 1. Algorithm to compute the Barankin bound.

4. CHANGE IN THE MEAN OF A GAUSSIAN DISTRIBUTION

In this section, as an example, we study the well known linear Gaussian model with parameters in the mean. In particular, we apply the proposed bound to the problem of changes in the mean parameters. Let us assume that the vector of observations $\mathbf{x}_i \in \mathbb{R}^M$, for $i = 1, \ldots, N$, is modelled as follows:

$$\mathbf{x}_i = \mathbf{f}(\boldsymbol{\nu}_j) + \mathbf{n}_i, \quad (26)$$

where, $\mathbf{f}(\cdot)$ is a vector of known functions, $\boldsymbol{\nu}_j \in \mathbb{R}^F$ is an known parameter vector with $F \leq L$, $\mathbf{n}_i$ is a zero-mean Gaussian random vector with known covariance matrix $\mathbf{\Sigma}$. Then $\mathbf{x}_i$ are distributed as $\mathcal{N}\left(\mathbf{f}(\boldsymbol{\nu}_j), \mathbf{\Sigma}\right)$. Note that we restrict our analysis to the set of parameter vectors $\{\boldsymbol{\nu}_j\}$ such that the functions in $\mathbf{f}(\boldsymbol{\nu}_j)$ are injective. Below, we compute the elements of $\mathbf{BIM_t}$ different from zero:

(i) For $\alpha_k > 0$ and $\alpha_k < 0$, A_k is given as follows:

$$A_k = \exp\left\{ |\alpha_k| \left(\mathbf{f}(\boldsymbol{\nu}_k) - \mathbf{f}(\boldsymbol{\nu}_{k+1})\right)^T \mathbf{\Sigma}^{-1} \left(\mathbf{f}(\boldsymbol{\nu}_k) - \mathbf{f}(\boldsymbol{\nu}_{k+1})\right) \right\} - 1.$$

(ii) For $t_k + \alpha_k > t_{k+1} + \alpha_{k+1}$, B_k is given as follows:

$$B_k = \exp\left\{ \frac{\beta_{k,k+1}}{2} \left(\left(\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\boldsymbol{\nu}_k)\right) \mathbf{\Sigma}^{-1} \left(\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\boldsymbol{\nu}_k)\right)^T \right.\right.$$
$$+ \left(\mathbf{f}(\boldsymbol{\nu}_{k+2}) - \mathbf{f}(\boldsymbol{\nu}_{k+1})\right) \mathbf{\Sigma}^{-1} \left(\mathbf{f}(\boldsymbol{\nu}_{k+2}) - \mathbf{f}(\boldsymbol{\nu}_{k+1})\right)^T$$
$$\left.\left. - \left(\mathbf{f}(\boldsymbol{\nu}_k) - \mathbf{f}(\boldsymbol{\nu}_{k+2})\right) \mathbf{\Sigma}^{-1} \left(\mathbf{f}(\boldsymbol{\nu}_k) - \mathbf{f}(\boldsymbol{\nu}_{k+2})\right)^T \right) \right\} - 1.$$

As an illustration, we compare the MSE between the values of the change-point locations and their maximum likelihood estimations with our bounds. In particular, we analyze the case of multiple changes in the mean of a Gaussian distribution with fixed variance σ^2. We consider the scenario of time series ($M = 1$) with 3 change points in the mean values of a Gaussian distribution with common

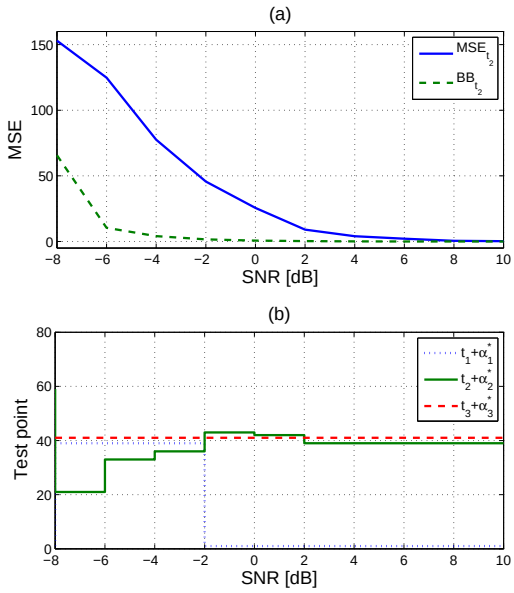

Fig. 2. Performance analysis for change point t_2: (a) MSE as a function of SNR using the maximum likelihood estimator and its corresponding Barankin bound. (b) Test points that maximize the BB for change point t_2 as a function of SNR.

variance. The locations of the change points are set to $t_1 = 20$, $t_2 = 40$, $t_3 = 60$; and the number of samples is $N = 80$; thus each segment has the same number of samples.

We define the signal-to-noise ratio (SNR) for the k^{th} change point as follows:

$$SNR_k = \left(\sigma^2\right)^{-1} \left(f(\boldsymbol{\nu}_{k+1}) - f(\boldsymbol{\nu}_k)\right)^2, \qquad (27)$$

and without loss of generality, we choose $f(\boldsymbol{\nu}_j) = \nu_j$ and $\sigma^2 = 1$. The means ν_j in each segment are set such that $SNR_1 = SNR_2 = SNR_3 = SNR$. We compute the maximum likelihood estimation of change-point locations assuming known the total number of changes and the distribution parameters in each segment, namely, the mean and variance. We illustrate the average MSE performance of the MLE for 1000 Monte Carlo experiments and study the performance as a function of SNR. In Figure 2(a), we illustrate the MSE performance of the maximum likelihood estimator for change point t_2, assuming knowledge of the means and variance. In the same figure, we illustrate the Barankin bound for the MSE of t_2 obtained such that the optimal test points $\{\alpha_1^*, \alpha_2^*, \alpha_3^*\}$ maximize the elements of the BB matrix (21) associated with t_2 only. As expected, it can be seen that the MSE of the MLE estimator approaches the BB as the SNR increases. In Figure 2(b) we illustrate the test point that maximizes the BB associated with t_2 . It can be seen that for SNR values before -2 dB the test points for t_1 and t_2 are overlapped, which implies that the Barankin information matrix is block diagonal and the corresponding bound for t_2 is a function of $[\boldsymbol{\Phi}]_{11,}$ $[\boldsymbol{\Phi}]_{12,}$ and $[\boldsymbol{\Phi}]_{22}$. For values above 2 dB the bound depends only on $[\boldsymbol{\Phi}]_{22,}$. Note that the test point approaches the change point values as SNR increases; *e.g.*, α_2 goes to 1 as SNR increases.

5. CONCLUSIONS

In this paper, we computed the Barankin bound on the mean-squared error (MSE) for multiple change-point estimation. The BIM structure revealed that the estimation of one change point is naturally perturbed by its two neighbors. Moreover, using this structure we proposed a computationally efficient inversion algorithm for the BIM.

6. REFERENCES

[1] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes, Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, Apr. 1993.

[2] E. W. Barankin, "Locally best unbiased estimates," *The Annals of Mathematical Statistics*, vol. 20, no. 4, pp. 477–501, Dec. 1949.

[3] H. Cramér, *Mathematical Methods of Statistics*, ser. Princeton Mathematics. New-York: Princeton University Press, Sept. 1946, vol. 9.

[4] D. G. Chapman and H. Robbins, "Minimum variance estimation without regularity assumptions," *The Annals of Mathematical Statistics*, vol. 22, no. 4, pp. 581–586, Dec. 1951.

[5] J. M. Hammersley, "On estimating restricted parameters," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 12, no. 2, pp. 192–240, 1950.

[6] A. Ferrari and J.-Y. Tourneret, "Barankin lower bound for change-points in independent sequences," in *Proc. of IEEE Workshop on Statistical Signal Processing (SSP)*, St. Louis, MO, USA, Sept. 2003, pp. 557–560.

[7] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of a hybrid parameter vector," *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 1084–1093, May 1997.

[8] R. J. McAulay and L. P. Seidman, "A useful form of the Barankin lower bound and its application to PPM threshold analysis," *IEEE Transactions on Information Theory*, vol. 15, pp. 273–279, Mar. 1969.

[9] R. J. McAulay and E. M. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Transactions on Information Theory*, vol. 17, pp. 669–676, Nov. 1971.

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