Weiss-Weinstein bound for change-point estimation

Lucien Bacharach, Alexandre Renaux, Mohammed Nabil El Korso, Eric Chaumette

Weiss-Weinstein bound for change-point estimation

Lucien Bacharach*, Alexandre Renaux*, Mohammed Nabil El Korso† and Eric Chaumette‡
*Université Paris-Sud/L2S, France
†Université Paris-Ouest/LEME, France
‡Toulouse-ISAE/DEOS, France

***Abstract*—We compute the Weiss-Weinstein bound in the context of change-point estimation in a multivariate time series whatever the considered distribution of the data as well the prior. Closed-form expressions are then given in the case of Gaussian observations with change of mean and variance and in the case of parameter change in a Poisson distribution. The proposed bound is shown to be tighter than the previous bounds which were originally derived in the deterministic context and provides a better approximation of the maximum *a posteriori* estimator global mean square error.**

I. Introduction

Change-point detection and estimation in a multivariate time series is an important topic due to the wide range of underlying applications. One can cite speech processing, fault detection, medical imaging or radar [1]. Particularly, in radar signal processing, such problem arise in SAR image edge detection [2], clutter map segmentation [3], meteorological calibration [4], detection of target signals in clutter [5], etc.

A plethora of algorithms have been proposed in the literature to address this problem, but their statistical performances are generally conducted only by way of Monte-Carlo simulations. Indeed, few results are available concerning the consistency, the bias or the mean square error (MSE) of change point estimators and are only given in an asymptotic context [6]–[8].

Concerning the MSE, the problem comes from the fact that the classical Cramér-Rao bound cannot be computed due to the discrete nature of the parameters. Some authors have tried to overcome this difficulty by using approximations (see e.g. [9], [10]). A more natural way to compute a lower bound on the MSE in this context is to use other bounds than the Cramér-Rao bound which require less regularity assumptions (mainly the existence of the derivative of the likelihood function). To the best of our knowledge, this has been done first in [11] where the Chapman-Robbins bound has been studied. Then, this result has been extended in [12] to the case of multiple change point estimation by using the McAulay-Seidman bound. In both the aforementioned papers, it has been shown that the obtained bounds was quite optimistic with respect to the maximum likelihood estimator empirical MSE.

In this paper, we propose to analyze the problem in the Bayesian context. As well as in the deterministic context, the Bayesian Cramér-Rao bound cannot be computed due to the discrete nature of the parameters and the Bayesian Barankin bound (a Bayesian version of the Chapman-Robbins and McAuly-Seidman bounds) neither due to some regularity assumptions about the used parameter *a priori* support. We propose to study the Weiss-Weinstein bound which is known to be one of the tightest bound among the Bayesian bounds. We give a general form of this bound whatever the considered distribution of the data (and the prior) and we then apply it in the case of Gaussian observations with change of mean and variance and in the case of parameter change in a Poisson distribution. Finally, throughout simulation results, the proposed bound is shown to be tighter than the previous bounds which were originally derived in the deterministic context.

This work has been partially supported by the iCODE institute, research project of the IDEX Paris-Saclay, by the MAGELLAN project (ANR-14-CE23-0004-01), by the Display-Mastodons project of CNRS and by DGA MRIS.

II. Problem setup and background

We consider the general case of N independent observation vectors $\mathbf{Y} = [\mathbf{y}_1,\ \mathbf{y}_2,\ \ldots,\ \mathbf{y}_N] \in \Omega \subset \mathbb{R}^{M\times N}$, which can be obtained, for example, from a multiple sensor system. In the context of a single change-point estimation, these observations are modeled as follows [11], [12]:

$$\begin{cases} \mathbf{y}_i \sim p_1\left(\mathbf{y}_i;\boldsymbol{\eta}_1\right) & \text{for } i = 1,\ldots,\theta, \\ \mathbf{y}_i \sim p_2\left(\mathbf{y}_i;\boldsymbol{\eta}_2\right) & \text{for } i = \theta+1,\ldots,N, \end{cases} \tag{1}$$

where M is the size of the sample vector (*e.g.* the number of sensors), and p_j is a probability density function with parameters $\boldsymbol{\eta}_j \in \mathbb{R}^L$, $j \in \{1,2\}$. Note that if $M = 1$, the problem is reduced to the estimation of a single change point in a time series context. We assume that the probability density functions p_j, $j \in \{1,2\}$, belong to a common distribution. The unknown parameter of interest is the change-point location denoted $\theta \in \Theta = \mathbb{Z}$. The observations before and after the change point are assumed to be stationary. The change point θ is assumed to be a discrete random variable over the parameter space Θ with a known prior distribution $\Pr\left(\theta = i\right)$, $i \in \mathbb{Z}$.

Besides, we recall that the Weiss-Weinstein bound (WWB) for a parameter θ, denoted WWB, satisfies the following inequality w.r.t. the mean square error of any Bayesian estimator $\hat{\theta}\left(\mathbf{Y}\right)$ [13]

$$\mathbb{E}_{\mathbf{Y},\theta}\left\{\left(\hat{\theta}\left(\mathbf{Y}\right)-\theta\right)^2\right\} \geq WWB, \tag{2}$$

in which

$$\mathbb{E}_{\mathbf{Y},\theta}\left\{\left(\hat{\theta}\left(\mathbf{Y}\right)-\theta\right)^2\right\} = \sum_{i=1}^{N-1}\Pr\left(\theta=i\right)\int_{\Omega}\left(\hat{\theta}\left(\mathbf{Y}\right)-i\right)^2 p\left(\mathbf{Y}\mid\theta=i\right)d\mathbf{Y}, \tag{3}$$

and

$$WWB = \sup_{h,s} \frac{h^2 \mathbb{E}^2_{\mathbf{Y},\theta}\left\{\left(\frac{p_{\mathbf{Y},\theta}(\mathbf{Y},i+h)}{p_{\mathbf{Y},\theta}(\mathbf{Y},i)}\right)^s\right\}}{\mathbb{E}_{\mathbf{Y},\theta}\left\{\left[\left(\frac{p_{\mathbf{Y},\theta}(\mathbf{Y},i+h)}{p_{\mathbf{Y},\theta}(\mathbf{Y},i)}\right)^s - \left(\frac{p_{\mathbf{Y},\theta}(\mathbf{Y},i-h)}{p_{\mathbf{Y},\theta}(\mathbf{Y},i)}\right)^{1-s}\right]^2\right\}}, \tag{4}$$

where $p_{\mathbf{Y},\theta}(\mathbf{Y},i) \triangleq p(\mathbf{Y} \mid \theta = i)\Pr(\theta = i)$ for $i \in \Theta$, and where $i+h$ is the so-called test-point such that $i+h \in \Theta$, $h \neq 0$ and $0 < s < 1$. We notice that we have to optimize the bound over the test point and over parameter s. Nevertheless, it has been shown that choosing $s = 1/2$ is generally the best choice in terms of tightness (see *e.g.* [13], [14]). Consequently, we will only optimize the bound over h. We first simplify the aforementioned expression in the context of a discrete parameter, then we derive a closed-form expression for change point estimation in the following.

The expectation of the numerator of (4) can then be simplified as follows:

$$\mathbb{E}_{\mathbf{Y},\theta}\left\{\sqrt{\frac{p_{\mathbf{Y},\theta}(\mathbf{Y},i+h)}{p_{\mathbf{Y},\theta}(\mathbf{Y},i)}}\right\} = \sum_{i\in\Theta} \mathcal{B}(i+h,i)\sqrt{\Pr(\theta = i+h)\Pr(\theta = i)}, \tag{5}$$

in which

$$\mathcal{B}(i,j) \triangleq \int_\Omega \sqrt{p(\mathbf{Y} \mid \theta = i)\, p(\mathbf{Y} \mid \theta = j)}\, d\mathbf{Y}, \tag{6}$$

denotes the Bhattacharyya distance between $p(\mathbf{Y} \mid \theta = i)$ and $p(\mathbf{Y} \mid \theta = j)$ [15].

Concerning the denominator of (4), one has:

$$\begin{aligned}\mathbb{E}_{\mathbf{Y},\theta}&\left[\left(\sqrt{\frac{p_{\mathbf{Y},\theta}(\mathbf{Y},i+h)}{p_{\mathbf{Y},\theta}(\mathbf{Y},i)}} - \sqrt{\frac{p_{\mathbf{Y},\theta}(\mathbf{Y},i-h)}{p_{\mathbf{Y},\theta}(\mathbf{Y},i)}}\right)^2\right] \\ &= \mathbb{E}_{\mathbf{Y},\theta}\left[\frac{p_{\mathbf{Y},\theta}(\mathbf{Y},i+h)}{p_{\mathbf{Y},\theta}(\mathbf{Y},i)}\right] + \mathbb{E}_{\mathbf{Y},\theta}\left[\frac{p_{\mathbf{Y},\theta}(\mathbf{Y},i-h)}{p_{\mathbf{Y},\theta}(\mathbf{Y},i)}\right] \\ &\quad - 2\mathbb{E}_{\mathbf{Y},\theta}\left[\frac{\sqrt{p_{\mathbf{Y},\theta}(\mathbf{Y},i+h)\, p_{\mathbf{Y},\theta}(\mathbf{Y},i-h)}}{p_{\mathbf{Y},\theta}(\mathbf{Y},i)}\right]. \end{aligned} \tag{7}$$

The three terms of the right hand side of (7) can be calculated as follows:

$$\begin{aligned}\mathbb{E}_{\mathbf{Y},\theta}&\left[\frac{p_{\mathbf{Y},\theta}(\mathbf{Y},i+h)}{p_{\mathbf{Y},\theta}(\mathbf{Y},i)}\right] \\ &= \sum_{i\in\Theta}\Pr(\theta = i+h)\int_\Omega p(\mathbf{Y} \mid \theta = i+h)\, d\mathbf{Y} \\ &= \sum_{i=1}^{N-1}\Pr(\theta = i+h) = 1 - \frac{|h|}{N-1}, \end{aligned} \tag{8}$$

given the constraint $1 \leq i+h \leq N-1$. In the same way, the second term is given by:

$$\mathbb{E}_{\mathbf{Y},\theta}\left[\frac{p_{\mathbf{Y},\theta}(\mathbf{Y},i-h)}{p_{\mathbf{Y},\theta}(\mathbf{Y},i)}\right] = 1 - \frac{|h|}{N-1}, \tag{9}$$

Finally, the third term can be written:

$$\begin{aligned}&\mathbb{E}_{\mathbf{Y},\theta}\left[\frac{\sqrt{p_{\mathbf{Y},\theta}(\mathbf{Y},i+h)\, p_{\mathbf{Y},\theta}(\mathbf{Y},i-h)}}{p_{\mathbf{Y},\theta}(\mathbf{Y},i)}\right] \\ &= \sum_{i\in\Theta}\mathcal{B}(i+h,i-h)\sqrt{\Pr(\theta = i+h)\Pr(\theta = i-h)}, \end{aligned} \tag{10}$$

Consequently, from (4), (7)-(10), one obtains the WWB expression as:

$$WWB = \sup_h \frac{N(h)}{D(h)} \tag{11}$$

where

$$N(h) = h^2\left(\sum_{i\in\Theta}\mathcal{B}(i+h,i)\sqrt{\Pr(\theta = i+h)\Pr(\theta = i)}\right)^2, \tag{12}$$

and

$$\begin{aligned}D(h) = 2\Bigg(1 - \frac{|h|}{N-1} - \\ \sum_{i\in\Theta}\mathcal{B}(i+h,i-h)\sqrt{\Pr(\theta = i+h)\Pr(\theta = i-h)}\Bigg). \end{aligned} \tag{13}$$

Given the above expression of the WWB, its final closed-form expression will be obtained by deriving $\mathcal{B}(i+h,i)$ and $\mathcal{B}(i+h,i-h)$ which depend on the probability density functions p_1 and p_2. Note that $\mathcal{B}(i+h,i-h)$ can be simplified by using a change of variable $i-h=j$, keeping in mind that $j \in \Theta$:

$$\begin{aligned}\int_\Omega \sqrt{p(\mathbf{Y} \mid \theta = i+h)\, p(\mathbf{Y} \mid \theta = i-h)}\, d\mathbf{Y} = \\ \int_\Omega \sqrt{p(\mathbf{Y} \mid \theta = j+2h)\, p(\mathbf{Y} \mid \theta = j)}\, d\mathbf{Y}, \end{aligned} \tag{14}$$

meaning that finding a closed-form expression of $\mathcal{B}(i+h,i)$ is enough since we can easily deduce $\mathcal{B}(i+h,i-h)$ by replacing h with $2h$.

III. Proposed bound for change point estimation

First, let us calculate the closed-form expression of $\mathcal{B}(i+h,i)$. By using the observation model of (1) and the fact that the observations are assumed to be independent, one can write

$$\begin{aligned}&\mathcal{B}(i+h,i) = \\ &\int_\Omega \prod_{i=1}^{\theta+h}\sqrt{p_1(\mathbf{y}_i)}\prod_{i=\theta+h+1}^{N}\sqrt{p_2(\mathbf{y}_i)}\prod_{i=1}^{\theta}\sqrt{p_1(\mathbf{y}_i)}\prod_{i=\theta+1}^{N}\sqrt{p_2(\mathbf{y}_i)}\, d\mathbf{Y}. \end{aligned} \tag{15}$$

We investigate both cases $h > 0$ and $h < 0$. For $h > 0$:

$$\begin{aligned}&\mathcal{B}(i+h,i) \\ &= \int_\Omega \prod_{i=1}^{\theta} p_1(\mathbf{y}_i)\prod_{i=\theta+1}^{\theta+h}\sqrt{p_1(\mathbf{y}_i)}\prod_{i=\theta+h+1}^{N} p_2(\mathbf{y}_i)\prod_{i=\theta+1}^{\theta+h}\sqrt{p_2(\mathbf{y}_i)}\, d\mathbf{Y} \\ &= \int_{\Omega'}\prod_{i=\theta+1}^{\theta+h}\sqrt{p_1(\mathbf{y}_i)}\sqrt{p_2(\mathbf{y}_i)}\, d\mathbf{y}_{\theta+1}\cdots d\mathbf{y}_{\theta+h} \\ &= \left(\int_{\Omega'}\sqrt{p_1(\mathbf{y})\, p_2(\mathbf{y})}\, d\mathbf{y}\right)^h \triangleq \rho_{1,2}^h \end{aligned} \tag{16}$$

where $\Omega' \subset \mathbb{R}^M$ denotes the observation space w.r.t. $\mathbf{y}_i$, $i \in \{1,\ldots,N\}$ (*i.e.* ${\Omega'}^N \triangleq \Omega' \times \ldots \times \Omega' = \Omega$). In the same way, for $h < 0$:

$$\begin{aligned}
&\mathcal{B}(i+h,i) \\
&= \int_{\Omega'} \prod_{i=\theta+h+1}^{\theta} \sqrt{p_1(\mathbf{y}_i)}\sqrt{p_2(\mathbf{y}_i)} d\mathbf{y}_{\theta+h+1} \cdots d\mathbf{y}_\theta \\
&= \left(\int_{\Omega'} \sqrt{p_1(\mathbf{y})\, p_2(\mathbf{y})} d\mathbf{y} \right)^{-h} = \rho_{1,2}^{-h}. \qquad (17)
\end{aligned}$$

Consequenlty, for any h we obtain:

$$\mathcal{B}(i+h,i) = \left(\int_{\Omega} \sqrt{p_1(\mathbf{y})\, p_2(\mathbf{y})} d\mathbf{y} \right)^{|h|} = \rho_{1,2}^{|h|}, \qquad (18)$$

and

$$\mathcal{B}(i+h,i-h) = \left(\int_{\Omega} \sqrt{p_1(\mathbf{y})\, p_2(\mathbf{y})} d\mathbf{y} \right)^{2|h|} = \rho_{1,2}^{2|h|}. \qquad (19)$$

Since $\mathcal{B}(i+h,i)$ and $\mathcal{B}(i+h,i-h)$ given above do not depend on i, the general WWB expression is now given by:

$$WWB = \sup_h \frac{h^2 \rho_{1,2}^{2|h|} \left(\sum_{i\in\Theta} \sqrt{\Pr(\theta = i+h)\Pr(\theta = i)} \right)^2}{2\left(1 - \rho_{1,2}^{2|h|} \sum_{i\in\Theta} \sqrt{\Pr(\theta = i+h)\Pr(\theta = i-h)} \right)}. \qquad (20)$$

The previous bound holds for any probablity density function $p_1(\mathbf{y})$ and $p_2(\mathbf{y})$ and any prior distribution on θ. A natural distribution in the change point estimation problem is the discrete uniform distribution, *i.e.*,

$$\Pr(\theta = i) = \begin{cases} \frac{1}{N-1}, & \forall i \in \{1,\ldots,N-1\}, \\ 0, & \forall i \in \mathbb{Z} \setminus \{1,\ldots,N-1\}. \end{cases} \qquad (21)$$

In this case, it is easy to evaluate the remaining parts of the WWB. We have

$$\begin{aligned}
&\sum_{i\in\Theta} \sqrt{\Pr(\theta = i+h)\Pr(\theta = i)} \\
&\quad = \frac{1}{\sqrt{N-1}} \sum_{i=1}^{N-1} \sqrt{\Pr(\theta = i+h)} = 1 - \frac{|h|}{N-1}. \qquad (22)
\end{aligned}$$

Similarly,

$$\sum_{i\in\Theta} \sqrt{\Pr(\theta = i+h)\Pr(\theta = i-h)} = 1 - \frac{2|h|}{N-1}. \qquad (23)$$

Consequently, the WWB for the estimation of one change point using a discrete uniform prior distribution is given by

$$WWB = \sup_h \frac{h^2 \left(1 - \frac{|h|}{N-1}\right)^2 \rho_{1,2}^{2|h|}}{2\left(1 - \frac{|h|}{N-1} - \left(1 - \frac{2|h|}{N-1}\right)\rho_{1,2}^{2|h|}\right)}. \qquad (24)$$

The only term that remains to be calculated in (24) is $\rho_{1,2}$, which will be done in the next section in the particular cases of Gaussian and Poisson distributed series.

IV. Gaussian and Poisson cases

A. Gaussian distributions

We now assume that the observations $\mathbf{y}_i$ follow a Gaussian distribution before and after the change point θ. The means and covariance matrices before and after θ are denoted $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$, $\sigma_1^2\mathbf{I}$, and $\sigma_2^2\mathbf{I}$, respectively. It is then possible to obtain the term $\rho_{1,2}$ defined in (16):

$$\begin{aligned}
\rho_{1,2} &= \frac{1}{(2\pi\sigma_1\sigma_2)^{\frac{M}{2}}} \int_{\Omega'} \exp\left\{ -\frac{\|\mathbf{y}-\boldsymbol{\mu}_1\|^2}{4\sigma_1^2} - \frac{\|\mathbf{y}-\boldsymbol{\mu}_2\|^2}{4\sigma_2^2} \right\} d\mathbf{y} \\
&= \frac{1}{(2\pi\sigma_1\sigma_2)^{\frac{M}{2}}} \left(4\pi \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2} \right)^{\frac{M}{2}} \exp\left\{ -\frac{\|\boldsymbol{\mu}_2-\boldsymbol{\mu}_1\|^2}{4(\sigma_1^2+\sigma_2^2)} \right\} \qquad (25)
\end{aligned}$$

due to the identification of a M-dimensional Gaussian density with mean $\frac{\sigma_1^2\boldsymbol{\mu}_2+\sigma_2^2\boldsymbol{\mu}_1}{\sigma_1^2+\sigma_2^2}$ and covariance matrix $\frac{2\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}\mathbf{I}$. In the case of a single time series ($M = 1$), this expression can be simplified and written as a function of two signal-to-noise ratios[1] defined in [11] as $\mathrm{snr}_{1,2}^m = \frac{(\mu_1-\mu_2)^2}{\sigma_1^2}$, and $\mathrm{snr}_{1,2}^v = \frac{\sigma_2^2}{\sigma_1^2}$.

The expression of $\rho_{1,2}$ in (25) then becomes:

$$\rho_{1,2} = \sqrt{\frac{2\sqrt{\mathrm{snr}_{1,2}^v}}{1+\mathrm{snr}_{1,2}^v}} \exp\left\{ -\frac{\mathrm{snr}_{1,2}^m}{4\left(1+\mathrm{snr}_{1,2}^v\right)} \right\}, \qquad (26)$$

In the following, we treat the two cases of mean jump and variance jump separately.

1) Gaussian mean change: In this case, the variance before and after the change point remains the same: $\sigma_1^2 = \sigma_2^2$ *i.e.* $\mathrm{snr}_{1,2}^v = 1$. As a consequence, (26) can be simplified as follows:

$$\rho_{1,2} = \exp\left\{ -\frac{\mathrm{snr}_{1,2}^m}{4} \right\}. \qquad (27)$$

The associated Weiss-Weinstein bound is then obtained by plugging (27) into (24).

2) Gaussian variance change: In that case, the mean remains the same before and after the change point, *i.e.*, $\mu_1 = \mu_2$, which implies $\mathrm{snr}_{1,2}^m = 0$. The expression of $\rho_{1,2}$ then reduces to:

$$\rho_{1,2} = \sqrt{\frac{2\sqrt{\mathrm{snr}_{1,2}^v}}{1+\mathrm{snr}_{1,2}^v}}, \qquad (28)$$

and as in the previous case, the Weiss-Weinstein bound is obtained thanks to (24).

B. Poisson distributions

Another case of interest in signal processing occurs with Poisson distributed observations. We denote λ_1 and λ_2 as the parameters of the distribution before and after the change-point respectively. The corresponding $\rho_{1,2}$ term in (16) can

[1]These two quantities are denoted as signal-to-noise ratios (or amount of change in the statistical literature) since they clearly make up a measure of the change-point optimal estimation performance. For example, in the Gaussian case of a mean jump, the estimation of θ is expected to be as precise as the difference between the two means before and after the jump is large. Of course, a similar argument can be made in the case of a variance jump.

then be obtained by taking into account the discrete nature of the distribution[2]:

$$\rho_{1,2} = \sum_{k=0}^{+\infty} \sqrt{\frac{\lambda_1^k}{k!}\mathrm{e}^{-\lambda_1}\frac{\lambda_2^k}{k!}\mathrm{e}^{-\lambda_2}} = \mathrm{e}^{-\frac{\left(\sqrt{\lambda_1}-\sqrt{\lambda_2}\right)^2}{2}}. \qquad (29)$$

Finally, as in the Gaussian case, the WWB expression is obtained by substituting (29) into (24). We use a similar definition of the signal-to-noise ratio as in [11], *i.e.*, $\mathrm{snr}_{1,2} = \frac{(\lambda_1-\lambda_2)^2}{\lambda_1^2}$.

V. Simulation results and conclusion

We present in this section the simulation results in the cases of a Gaussian mean change (Fig. 1) and of a Poisson parameter change (Fig. 2). Due to the lack of space, we do not present the result for the case of a Gaussian variance change, for which the same behavior is observed. On these figures, the empirical global mean square error (GMSE) of the maximum *a posteriori* (MAP) estimator is obtained by averaging the square errors over 1024 Monte-Carlo simulations. We compare them with the proposed theoretical WWB, as well as the averaged deterministic Barankin bound (BRB) adapted from [11], so that the comparison with the WWB makes sense. We recall that due to some regularity assumptions about the used parameter *a priori* support, the original Bayesian Barankin bound (and the Bayesian Cramér-Rao bound) cannot be computed for this problem.

We notice that, as expected, the proposed bound is in good agreement with the MAP estimator GMSE.

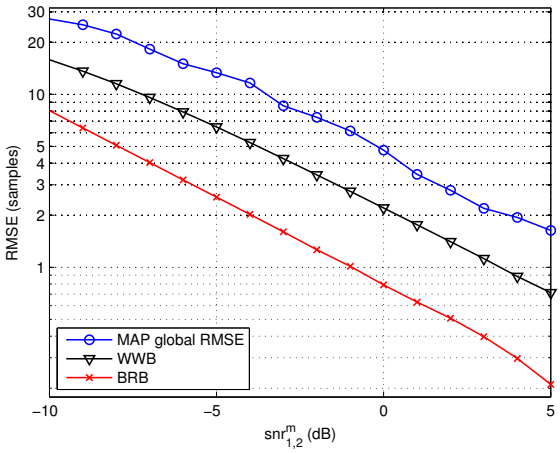

Fig. 1. MSE (circles), Barankin (crosses) bound and Weiss-Weinstein (triangles) bound in the Gaussian case for a mean change, in the same conditions as in [11] : $N = 128$, $\sigma^2 = 1$, $\mu_1 = 1$.

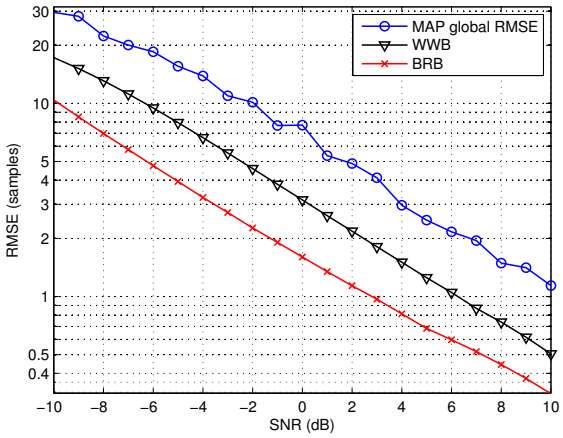

Fig. 2. MSE (circles), Barankin (crosses) bound and Weiss-Weinstein (triangles) bound in the case of Poisson distributions, in the same conditions as in [11] : $N = 128$, $\lambda_1 = 1$.

References

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

[2] J. Tourneret, M. Doisy, and M. Lavielle, "Bayesian off-line detection of multiple change-points corrupted by multiplicative noise: application to SAR image edge detection," *ELSEVIER Signal Processing*, vol. 83, no. 9, pp. 1871–1887, Sep. 2003.

[3] N. Brunel and F. Barbaresco, "Doppler and Polarimetric Statistical Segmentation for Radar Clutter map based on Pairwise Markov Chains," in *Proc. of IEEE RADAR*, Oct. 2007.

[4] A. Jann, "Reconciling the sequential probability ratio test with calibration," *Quarterly Journal of the Hungarian Meteorological Service*, vol. 113, no. 3, pp. 233–243, Jul. 2009.

[5] N. Nechval, K. Nechval, and E. Vasermanis, "Detection of target signals in clutter using change point statistics," in *Proc. SPIE 4541, Image and Signal Processing for Remote Sensing VII, 111*, Jan. 2002.

[6] S. B. Fotopoulos, S. K. Jandhyala, and E. Kapalova, "Exact asymptotic distribution of change-point MLE for change in the mean of Gaussian sequences," *The Annals of Applied Statistics*, vol. 4, no. 2, pp. 1081–1104, Nov. 2010.

[7] Y. Wu, "Bias of estimator of change-point detected by a CUSUM procedure," *Annals of the Institute of Statistical Mathematics*, vol. 56, no. 1, pp. 127–142, 2004.

[8] D. V. Hinkley, "Inference about the change-point in a sequence of random variables," *Biometrika*, vol. 57, no. 1, pp. 1–18, 1970.

[9] A. Bartov and H. Messer, "Analysis of inherent limitations in localizing step-like singularities in a continuous signal," in *Proc. IEEE SP Int. Symp. Time-Frequency Time-Scale Analysis*, Paris, France, Jun. 1996, pp. 21–24.

[10] A. M. Reza and M. Doroodchi, "Cramér-Rao lower bound on locations of sudden changes in a steplike signal," *IEEE Transactions on Signal Processing*, vol. 44, no. 10, pp. 2551–2556, Oct. 1996.

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

[12] P. S. La Rosa, A. Renaux, A. Nehorai, and C. H. Muravchik, "Barankin-type lower bound on multiple change-point estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5534–5549, Nov. 2010.

[13] E. Weinstein and A. J. Weiss, "A general class of lower bounds in parameter estimation," *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 338–342, Mar. 1988.

[14] D. T. Vu, A. Renaux, R. Boyer, and S. Marcos, "Some results on the Weiss-Weinstein bound for conditional and unconditional signal models in array processing," *ELSEVIER Signal Processing*, vol. 95, no. 2, pp. 126–148, Feb. 2014.

[15] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bulletin Calcutta Mathematical Society*, vol. 35, pp. 99–109, 1943.

[2]Notice that (16) is given for probability density functions. In the case of mass functions, it becomes $\rho_{1,2} = \sum_k \sqrt{\Pr_1(y=k)\Pr_2(y=k)}$.