DESIGNING SAR IMAGES CHANGE-POINT ESTIMATION STRATEGIES USING AN MSE LOWER BOUND

A. Mian[,†], L. Bacharach[‡], G. Ginolhac[*,†], A. Renaux[‡], M. N. El Korso[§] and J.-P. Ovarlez[*,¶]*

[*] CentraleSupélec/SONDRA, Plateau du Moulon, 3 rue Joliot-Curie, F-91190 Gif-sur-Yvette, France
[†] LISTIC, Université de Savoie Mont-Blanc, F-74944, Annecy le Vieux, France
[‡] L2S, Université Paris-Saclay, F-91190, Gif-sur-Yvette, France
[§] LEME, Université Paris-Nanterre, IUT Ville d'Avray, LEME-EA4416, France
[¶] DEMR, ONERA, Université Paris-Saclay, F-91123, Palaiseau, France

ABSTRACT

A growing problem in the remote sensing community concerns the estimation of change-points in a time series of Synthetic Aperture Radar (SAR) images. Although the methodologies of change-point estimation have already been investigated in the literature, there are, to the best of our knowledge, no study on the expected performance for the estimation of change-points in a Wishart distributed time series. This is mainly due to the fact that few results exist on change-point estimation performance in the mathematical literature: the classical central limit theorem does not apply and the classical Cramer-Rao Bound does not exist due to the discrete nature of the parameters. To fill this gap, this paper proposes to use a lower-bound on the Mean Square Error (MSE) with fewer regularity conditions. To this end, recent works on hybrid Cramer-Rao/Weiss-Weinstein bound have been adapted to the specific SAR problematic of interest. Since estimation strategies usually rely on a set of parameters which have to be set by the user, we show how the proposed lower bound allows performing an appropriate tuning. Moreover, the proposed bound is computationally efficient which enables an extensive analysis without a high computational cost.

Index Terms— Synthetic Aperture Radar; Wishart distribution; Hybrid Bounds; Cramer-Rao Bound; Weiss-Weinstein Bound; Change-point estimation;

1. INTRODUCTION

In the past years, interest has been shown in the analysis of Synthetic Aperture Radar (SAR) image time series for monitoring the evolution of large areas over time. Change-point detection and estimation in this context is useful for analysing the huge amount of data spanning large areas. Usually, when considering complex multivariate SAR images (such as polarimetric ones), the data is modelled by using a Gaussian distribution. The covariance parameter is then considered as a descriptor in order to assess the changes [1, 2]. Thus, recent works [3, 4] have considered the problem of change-point estimation in a time series of covariance matrices. In an adaptive context, these matrices are estimated through the well known Sample Covariance Matrix (SCM) estimator which is known to be Wishart distributed. This Wishart model has been used in various works pertaining to SAR applications such as change detection [1, 5], speckle reduction [6] or classification [7, 8].

The estimation of covariance matrices is done for a pixel using the surrounding spatial neighbourhood. The number of samples N used, referred as the number of looks, requires to be set by the user as a trade-off between the resolution obtained and performances of estimation. Moreover, this estimation step must be done by taking into account the dimension of the dataset p. Indeed, techniques exploiting time-frequency analysis [9] allow controlling the size of each pixel. Subsequently, the number of looks as well as the size of the vectors determine the performances of any given method of estimation and have to be chosen adequately.

In this paper, we consider this problem by using the Mean Square Error (MSE) of the estimated change-points as a measure of performance and propose a lower bound on its expected value. The only available theoretical results on change-point estimation concern specific problems such as monovariate Gaussian time series with an asymptotic assumption of a large number of data before and after the change [10, 11]. In order to overcome this difficulty, the signal processing community generally focusses on MSE analysis with Monte Carlo simulations, which are computationally expensive, or the Cramer-Rao Bound (CRB). However, in the context of change-point estimation, a part of the unknown parameter vector lies on a discrete space violating the regularity condition of the CRB. Fortunately, other lower bounds with weaker regularity conditions exist. Thus, we consider a hybrid lower bound between the CRB and the Weiss-Weinstein Bound (WWB) [12] to obtain a good trade-off between tightness and regularity conditions as suggested in [13].

2. RELATION TO PRIOR WORKS

Lower-bound for change-point estimation has been firstly considered in [14, 15] where only the Barankin Bound (BB) was considered. Then, the extension to a tighter Bayesian bound (i.e. the WWB) has been proposed in [16] where a general semi-closed-form expression focussing on change-points without specifying the distribution of the data, is given. This work has been extended by including the possibility of unknown additional parameters in [13]. Our aim is to adapt this semi-closed-form expression to the specific aforementioned SAR problem. To this end, we have to deal with the Wishart distribution which, to the best of our knowledge, was not included in any previous works.

This adaptation leads to a time-efficient lower-bound on the expected performances of any estimator of change-points in the context of SAR applications. The usefulness of this bound for the tuning of the parameters is then considered from a practical standpoint.

3. DATA MODEL

We consider a time series of T independent, Wishart-distributed matrices $\{\mathbf{X}_t \in \mathbb{S}_{\mathbb{H}}^p \,|\, t \in [\![1, T]\!]\}$, where $\mathbb{S}_{\mathbb{H}}^p$ is the set of positive definite Hermitian matrices, subjected to a single change-point t_C. This scenario corresponds to a case where the user has the knowledge that there is an abrupt change and wants to know its precise localisation. This kind of problem can arise, for example, when looking for the time at which some flood has occurred or a region where a forest has been cut. The case of multiple change-points could also be considered, but requires more complex derivations that will be let for further investigation. The studied scenario is then written as follows:

$$\begin{cases} \mathbf{X}_t \sim \mathbb{CW}(p, N, \boldsymbol{\Sigma}_0) & \text{for } t = 1, \ldots, t_C \\ \mathbf{X}_t \sim \mathbb{CW}(p, N, \boldsymbol{\Sigma}_1) & \text{for } t = t_C + 1, \ldots, T \end{cases}, \quad (1)$$

where $\mathbb{CW}$ is the central complex Wishart distribution with Probability Distribution Function (PDF):

$$f_{\mathbf{X}_t; \boldsymbol{\Sigma}}(\mathbf{X}_t; \boldsymbol{\Sigma}) = \frac{|\mathbf{X}_t|^{N-p}}{\Gamma_p(N) |\boldsymbol{\Sigma}|^N} \operatorname{etr}\left(\boldsymbol{\Sigma}^{-1} \mathbf{X}_t\right), \quad (2)$$

where, $\Gamma_p(N) = \pi^{p(p-1)/2} \prod_{j=1}^p \Gamma(N-j+1)$, $\Gamma(.)$ is the Gamma function and $\operatorname{etr}(.)$ is the exponential trace function.

In order to deal with complex parameters, we define the mapping from complex to real-case as a function $[\bullet]_{\mathbb{CR}} : \mathbb{C}^m \to \mathbb{R}^{2m}$, whose definition is: $[\mathbf{z}]_{\mathbb{CR}} = [\Re(\mathbf{z})^{\mathrm{T}}, \Im(\mathbf{z})^{\mathrm{T}}]^{\mathrm{T}}$.

We also define:

$$\mathbf{x} = \left[[\operatorname{vech}(\mathbf{X}_1)]_{\mathbb{CR}}^{\mathrm{T}}, \ldots, [\operatorname{vech}(\mathbf{X}_T)]_{\mathbb{CR}}^{\mathrm{T}} \right]^{\mathrm{T}} \in \mathbb{R}^{Tp^2},$$

$$\boldsymbol{\sigma} = \left[[\operatorname{vech}(\boldsymbol{\Sigma}_0)]_{\mathbb{CR}}^{\mathrm{T}}, [\operatorname{vech}(\boldsymbol{\Sigma}_1)]_{\mathbb{CR}}^{\mathrm{T}} \right]^{\mathrm{T}} \in \mathbb{R}^{2p^2},$$

where $\operatorname{vech}(\bullet)$ is the vectorisation operator with the upper triangular portion excluded.

The problem considered in this work is about the estimation of the unknown parameter $\boldsymbol{\theta} = \left[\boldsymbol{\sigma}^{\mathrm{T}}, t_C \right]^{\mathrm{T}}$, compromised of the covariance parameters as well as the change-point. We define $M = 2p^2 + 1$, the number of total unknown parameters.

Note that, in this paper, the change-point t_C is assumed to be random and the covariance matrices are assumed to be deterministic. This is due to the fact that the tightest lower bound (with good regularity conditions) is the deterministic BB which has already been shown to exhibit too optimistic results in the context of change-point estimation [14, 15] contrary to the WWB. Concerning the covariance matrices, it is well known that the CRB is an efficient approximation [17]. This is why we consider such a hybrid bound which allows us to manage a trade-off between tightness and regularity conditions.

Finally, the distribution of the observations in this model denoted $f_{\mathbf{x}, t_C; \boldsymbol{\sigma}}(\mathbf{x}, t_C; \boldsymbol{\sigma}) = f_{\mathbf{x}|t_C; \boldsymbol{\sigma}}(\mathbf{x}|t_C; \boldsymbol{\sigma}) f_{t_C}(t_C)$ where:

$$f_{\mathbf{x}|t_C; \boldsymbol{\sigma}}(\mathbf{x}|t_C; \boldsymbol{\sigma}) = \prod_{t=1}^{t_C} f_{\mathbf{X}_t; \boldsymbol{\Sigma}_0}(\mathbf{X}_t; \boldsymbol{\Sigma}_0) \prod_{t=t_C+1}^{T} f_{\mathbf{X}_t; \boldsymbol{\Sigma}_1}(\mathbf{X}_t; \boldsymbol{\Sigma}_1). \quad (3)$$

4. STATEMENT OF THE RESULT

The aim is to obtain a lower bound on the performances of any estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta} \in \boldsymbol{\Theta} = \mathbb{R}^{2p^2} \times \mathbb{N}$ using the hybrid CR/WWB. Note that, this does not consist only in the concatenation of the CRB of the unknown covariances $\boldsymbol{\sigma}$ and the WWB of the unknown change-point t_C but in obtaining a matrix that considers also the coupling between the parameters. To this end, we consider the global MSE defined as follows:

$$\mathbf{MSE}(\hat{\boldsymbol{\theta}}) = \mathbb{E}_{\mathbf{x}, t_C; \boldsymbol{\sigma}} \left\{ (\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta})^{\mathrm{T}} \right\}, \quad (4)$$

in which, $\mathbb{E}_{\mathbf{x}, t_C; \boldsymbol{\sigma}}$ denotes the expectation w.r.t the distribution $f_{\mathbf{x}, t_C; \boldsymbol{\sigma}}(\mathbf{x}, t_C; \boldsymbol{\sigma})$.

4.1. Covariance inequality

It has recently been shown in the literature [18] that eq. (4) can be bounded using the covariance inequality. In the context of our problem, we have the following proposition:

Proposition 4.1. Covariance Inequality
Let $\{\Psi_k(\mathbf{x}, \boldsymbol{\theta}) \,|\, k \in [\![1, M]\!]\}$ *be real-valued function set defined on* $\mathbb{R}^{Tp^2} \times \boldsymbol{\Theta}$ *such that the following integral exists and satisfies for almost every (a.e.)* $\mathbf{x} \in \mathbb{R}^{Tp^2}$, $\forall k \in [\![1, M]\!]$, $\int_{\boldsymbol{\Theta}} \Psi_k(\mathbf{x}, \boldsymbol{\theta}) f(\mathbf{x}, \boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} = 0$. *Then, using the definition of* $\mathbf{MSE}$ *at eq. (4), the following inequality holds:*

$$\mathbf{MSE}(\hat{\boldsymbol{\theta}}) \succeq \mathbf{V}\mathbf{P}^{-1}\mathbf{V}^{\mathrm{T}}, \quad (5)$$

where $\mathbf{A} \succeq \mathbf{B}$ *means that* $\mathbf{A} - \mathbf{B}$ *is positive semi-definite,* $\mathbf{V}$ *is a* $M \times M$ *matrix whose elements are given by*

$$(\mathbf{V})_{k,l} = \mathbb{E}_{\mathbf{x}, t_C; \boldsymbol{\sigma}} \left\{ \left((\hat{\boldsymbol{\theta}}(\mathbf{x}))_k - (\boldsymbol{\theta})_k \right) \Psi_l(\mathbf{x}, \boldsymbol{\theta}) \right\}, \quad (6)$$

$(\bullet)_{k,l}$ *is the k-th line and l-th column of a matrix, and* $\mathbf{P}$ *is a* $M \times M$ *matrix whose elements are given by*

$$(\mathbf{P})_{k,l} = \mathbb{E}_{\mathbf{x}, t_C; \boldsymbol{\sigma}} \{\Psi_k(\mathbf{x}, \boldsymbol{\theta}) \Psi_l(\mathbf{x}, \boldsymbol{\theta})\}. \quad (7)$$

To obtain a hybrid CR/WWB, Ψ_k are chosen as follows:

$$\Psi_k(\mathbf{x}, \boldsymbol{\theta}) = \begin{cases} \Psi_k^{\mathrm{CRB}}(\mathbf{x}, \boldsymbol{\theta}) & \text{for } k = 1, \ldots M - 1 \\ \Psi^{\mathrm{WWB}}(\mathbf{x}, \boldsymbol{\theta}) & \text{for } k = M \end{cases}, \quad (8)$$

where

$$\Psi_k^{\mathrm{CRB}}(\mathbf{x}, \boldsymbol{\theta}) = \begin{cases} \dfrac{\partial \ln f_{\mathbf{x}, t_C; \boldsymbol{\sigma}}(\mathbf{x}, t_C; \boldsymbol{\sigma})}{\partial(\boldsymbol{\sigma})_k} & \text{if } \boldsymbol{\theta} \in \boldsymbol{\Theta}' \\ 0 & \text{if } \boldsymbol{\theta} \notin \boldsymbol{\Theta}' \end{cases} \quad (9)$$

and

$$\Psi^{\mathrm{WWB}}(\mathbf{x}, \boldsymbol{\theta}) = $$
$$\begin{cases} \dfrac{f_{\mathbf{x}, t_C; \boldsymbol{\sigma}}^s(\mathbf{x}, t_C + h; \boldsymbol{\sigma})}{f_{\mathbf{x}, t_C; \boldsymbol{\sigma}}^s(\mathbf{x}, t_C; \boldsymbol{\sigma})} - \dfrac{f_{\mathbf{x}, t_C; \boldsymbol{\sigma}}^{1-s}(\mathbf{x}, t_C - h; \boldsymbol{\sigma})}{f_{\mathbf{x}, t_C; \boldsymbol{\sigma}}^{1-s}(\mathbf{x}, t_C; \boldsymbol{\sigma})} & \text{if } \boldsymbol{\theta} \in \boldsymbol{\Theta}' \\ 0 & \text{if } \boldsymbol{\theta} \notin \boldsymbol{\Theta}' \end{cases} \quad (10)$$

for $s \in]0, 1[$, h is such that $t_C + h \in [\![1, T - 1]\!]$ and $\boldsymbol{\Theta}' = \{\boldsymbol{\theta} \in \boldsymbol{\Theta} \,|\, f(\mathbf{x}, \boldsymbol{\theta}) > 0 \, a.e. \, \mathbf{x} \in \mathbb{R}^{Tp^2}\}$.

Note that, any value for the terms h and s will lead to a lower bound on the MSE. However, they must be chosen cautiously in order to obtain the tightest bound. To obtain the tightest bound on the change-point, we have to compute:

$$\mathbf{HCRWWB} = \sup_{h, s} \quad \mathbf{V}\mathbf{P}^{-1}\mathbf{V}^{\mathrm{T}}. \quad (11)$$

4.2. Hybrid bound for the change-point model

In the context of change-point estimation, the right-hand side of the inequality at eq. (5) can be obtained by using the semi closed-form expression provided in [13]:

$$\mathbf{V} = \begin{bmatrix} -\mathbf{I}_{2p^2} & \mathbf{0}_{2p^2,1} \\ \mathbf{0}_{1,2p^2} & v_{22} \end{bmatrix} \text{ and } \mathbf{P} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{12}{}^{\mathrm{T}} & P_{22} \end{bmatrix}, \qquad (12)$$

where the block-matrices are defined as follows:

- $\mathbf{P}_{11} = T/2 \operatorname{diag}(\mathbf{F}(\boldsymbol{\Sigma}_0), \mathbf{F}(\boldsymbol{\Sigma}_1))$, where $\mathbf{F}(\boldsymbol{\Sigma}_0)$ (resp. $\mathbf{F}(\boldsymbol{\Sigma}_1)$) is the Fisher information matrix with regards to $\boldsymbol{\Sigma}_0$ (resp. $\boldsymbol{\Sigma}_1$).

- $P_{22} = u(h)\left(\rho^{|h|}\left(\epsilon_h(2s)\right) + \rho^{|h|}\left(\epsilon_h(2s-1)\right)\right)$
  $-2u(2h)\rho^{|h|}\left(\epsilon_h(s)\right)$, where

$$u(h) \triangleq \begin{cases} (T-1-|h|)/(T-1) & \text{if } |h| < T-1 \\ 0 & \text{otherwise} \end{cases},$$

$$\epsilon_h(s) = \begin{cases} s & \text{if } h > 0 \\ 1-s & \text{if } h < 0 \end{cases} \text{ and }$$

$$\rho(s) \triangleq \int_{\mathbb{S}_{\mathbb{H}}^p} f_{\mathbf{X}_t;\boldsymbol{\Sigma}_0}^s(\mathbf{X}_t;\boldsymbol{\Sigma}_0) f_{\mathbf{X}_t;\boldsymbol{\Sigma}_1}^{1-s}(\mathbf{X}_t;\boldsymbol{\Sigma}_1) \, \mathrm{d}\mathbf{X}_t.$$

$$(13)$$

- $\mathbf{P}_{12} = \left[\mathbf{p}^{\mathrm{T}}, \mathbf{q}^{\mathrm{T}}\right]^{\mathrm{T}}$, where the elements of vectors $\mathbf{p}$ and $\mathbf{q}$ are given by:

$$(\mathbf{p})_\ell = -hu(h)\rho^{|h|-1}\left(\epsilon_h(s)\right)\phi_{\sigma_0,\ell}\left(\epsilon_h(s)\right),$$

$$(\mathbf{q})_\ell = hu(h)\rho^{|h|-1}\left(\epsilon_h(s)\right)\phi_{\sigma_1,\ell}\left(\epsilon_h(s)\right),$$

and given $j \in \{0,1\}$, $\ell \in [\![1, p^2]\!]$, $s \in]0,1[$:

$$\phi_{\sigma_j,\ell}(s) \triangleq \int_{\mathbb{S}_{\mathbb{H}}^p} \left.\frac{\partial \ln f_{\mathbf{X}_t;\boldsymbol{\Sigma}}(\mathbf{X}_t;\boldsymbol{\Sigma})}{\partial\left([\operatorname{vech}(\boldsymbol{\Sigma})]_{\mathbb{CR}}\right)_\ell}\right|_{\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_j} \times \qquad (14)$$
$$f_{\mathbf{X}_t;\boldsymbol{\Sigma}_0}^s(\mathbf{X}_t;\boldsymbol{\Sigma}_0) f_{\mathbf{X}_t;\boldsymbol{\Sigma}_1}^{1-s}(\mathbf{X}_t;\boldsymbol{\Sigma}_1) \, \mathrm{d}\mathbf{X}_t.$$

- $v_{22} = hu(h)\rho^{|h|}(\epsilon_h(s))$.

4.3. Derivation of $\mathbf{F}(\boldsymbol{\Sigma})$, $\rho(s)$ and $\phi_{\sigma_j,\ell}(s)$

In order to compute the bound, we finally need the closed-form expressions of $\mathbf{F}(\boldsymbol{\Sigma})$, $\rho(s)$ and $\phi_{\sigma_j,\ell}(s)$. We have:

Proposition 4.2. *The closed-form expression of $\rho(s)$ is given by:*

$$\rho(s) = \frac{\left|s\boldsymbol{\Sigma}_0^{-1} + (1-s)\boldsymbol{\Sigma}_1^{-1}\right|^{-N}}{|\boldsymbol{\Sigma}_0|^{sN}|\boldsymbol{\Sigma}_1|^{(1-s)N}}. \qquad (15)$$

Proof. Derived in [19]. The result can be obtained by doing the substitution $\mathbf{Y} = \mathbf{A}\mathbf{X}\mathbf{A}$, with $\mathbf{A} = \left(s\boldsymbol{\Sigma}_0^{-1} + (1-s)\boldsymbol{\Sigma}_1^{-1}\right)^{\frac{1}{2}}$, in the integral. $\qquad \square$

Proposition 4.3 (FIM of the Covariance for a $\mathbb{CW}$ distribution). *For $\boldsymbol{\Sigma} \in \mathbb{S}_{\mathbb{H}}^p$, we have:*

$$\mathbf{F}(\boldsymbol{\Sigma}) = f_{\mathbb{CR}}\left(N\mathbf{D}_p{}^{\mathrm{T}}(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1})\mathbf{D}_p\right), \qquad (16)$$

where $\mathbf{D}_p$ is the duplication matrix defined for any matrix $\mathbf{X} \in \mathbb{C}^{p \times p}$, by

$$\mathbf{D}_p\operatorname{vech}(\mathbf{X}) = \operatorname{vec}(\mathbf{X}),$$

and $f_{\mathbb{CR}} : \mathbb{S}_{\mathbb{H}}^{p^2/2} \to \mathbb{S}_{++}^{p^2}$, where $\mathbb{S}_{++}^{p^2}$ is the set of real symmetric matrices, is defined as

$$f_{\mathbb{CR}}(\boldsymbol{\Sigma}) = \frac{1}{2}\begin{bmatrix} \Re(\boldsymbol{\Sigma}) & -\Im(\boldsymbol{\Sigma}) \\ \Im(\boldsymbol{\Sigma}) & \Re(\boldsymbol{\Sigma}) \end{bmatrix}.$$

Proof. The result comes by noticing that the score function w.r.t. the covariance matrix of a Wishart distribution is the same as the score function of a Gaussian distribution, and by an appropriate change of variable ($\mathbf{X}_t = \sum_{i=1}^N \mathbf{z}_i\mathbf{z}_i^{\mathrm{H}}$ where $\{\mathbf{z}_i\}_{i=1,\ldots N} \sim \mathbb{CN}(\mathbf{0}_p, \boldsymbol{\Sigma})$, and $\mathbb{CN}$ is the complex Normal distribution). The mapping $f_{\mathbb{CR}}$ allows obtaining its form in the real parametrisation used in this paper. $\qquad \square$

Proposition 4.4. *The different terms of $\phi_{\sigma_j,\ell}(s)$ for $\ell \in [\![1, p^2]\!]$, $j \in \{0,1\}$ are given by $\phi_{\sigma_j,\ell}(s) = \left([\operatorname{vech}(\boldsymbol{\Phi}_j(s))]_{\mathbb{CR}}\right)_\ell$, where $\boldsymbol{\Phi}_j(s)$ is a $p \times p$ matrix given by:*

$$\boldsymbol{\Phi}_j(s) = N\rho(s)\boldsymbol{\Sigma}_j^{-1}\left(s\boldsymbol{\Sigma}_0^{-1} + (1-s)\boldsymbol{\Sigma}_1^{-1}\right)^{-1}\boldsymbol{\Sigma}_j^{-1} \qquad (17)$$
$$- N\rho(s)\boldsymbol{\Sigma}_j^{-1}.$$

Proof. See Appendix. $\qquad \square$

5. SIMULATIONS

In order to validate the bound derived in this paper, Wishart time series subjected to a change-point as described in eq. (1) have been generated. t_C is generated using a uniform random prior and the covariance matrices have been chosen as Toeplitz matrices of the form: $(\boldsymbol{\Sigma}_{k=0,1})_{i,j} = \alpha_k^{|i-j|}$.

Two estimators of the change-point have been considered:

- The Maximum A Posteriori (MAP) estimator which has the knowledge of the covariance matrices before and after the change:
$$\hat{t}_C = \operatorname*{argmax}_{t_C \in [\![1,T-1]\!]} f_{\mathbf{x},t_C}(\mathbf{x}, t_C). \qquad (18)$$

- The following hybrid estimator, derived from the Maximum A Posteriori/Maximum Likelihood [20], which estimates the covariance matrices before and after the change as well as the change point:
$$\hat{t}_C = \operatorname*{argmax}_{t_C \in [\![1,T-1]\!]} f_{\mathbf{x},t_C;\hat{\boldsymbol{\sigma}}}(\mathbf{x}, t_C; \hat{\boldsymbol{\sigma}}), \qquad (19)$$

where $\hat{\boldsymbol{\sigma}} = \left[[\operatorname{vech}(\hat{\boldsymbol{\Sigma}}_0)]_{\mathbb{CR}}^{\mathrm{T}}, [\operatorname{vec}(\hat{\boldsymbol{\Sigma}}_1)]_{\mathbb{CR}}^{\mathrm{T}}\right]^{\mathrm{T}}$ with:

$$\hat{\boldsymbol{\Sigma}}_0 = \frac{1}{t_C N}\sum_{t=1}^{t_C}\mathbf{X}_t \text{ and } \hat{\boldsymbol{\Sigma}}_1 = \frac{1}{(T-t_C)N}\sum_{t=t_C+1}^{T}\mathbf{X}_t.$$

First, let us consider Figure 1: the MSE of the estimated change-point for both estimators as well as the proposed bound has been plotted for several values of N. We first observe that the MAP performs better than the hybrid estimator, that is expected since it assumes the knowledge of the covariance matrices. The derived bound behaves similarly to the hybrid estimator: the drop in the error for high N is well described by the bound. Although the gap between

the bound and the hybrid estimator is rather high at low N, it decreases significantly for higher values. Since our aim is to design algorithms having good estimation performance, the region of interest is the one where the performance of the hybrid estimator is lower than those for 1 sample. In this region, the bound accurately describes the expected performance and can thus be used as an approximation.
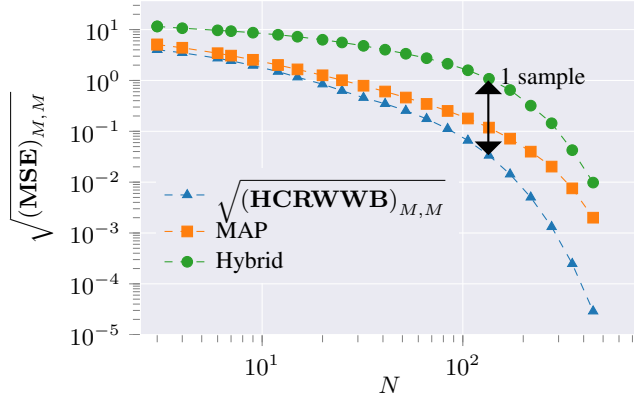

Fig. 1. MSE on the change-point for $p = 3$, $T = 50$, $\alpha_0 = 0.1$, $\alpha_1 = 0.3$. The estimators curves have been computed with 10^6 Monte Carlo trials.

The benefit of the bound is well described by Table 1 where the time-consumption[1] needed to compute the points of Figure 1 are given. Since the bounds need at least a thousand times less computational cost, it is preferable in a design context where many values for the parameters have to be tested.

Table 1. Time-consumption in seconds.

N	HCRWWB	MAP	Hybrid estimator
10	**0.17**	305	310
10^2	**0.17**	510	568
10^3	**0.17**	1462	1476

In order to illustrate the usability of the bound in a design context, the bound has been computed for an extensive set of parameters (p, N) in Figure 2. Since for $p < N$, the empirical covariance matrices cannot be inverted, the bound is not usable that explains the lack of performance in this region. We can observe the drop in the error for any value of p at some point when increasing N. The dashed line which corresponds to the region at which the bound is accurate enough, allows to limit a region where, for a given situation with fixed p, an appropriate value for N can be chosen to guarantee good estimation performance.

6. CONCLUSION

We have considered the derivation of a lower bound on the MSE of a change-point in the context of a SAR image time series. Simulations have shown a good behaviour for the proposed bound and a much lower computational cost with regards to a Monte-Carlo simulation.

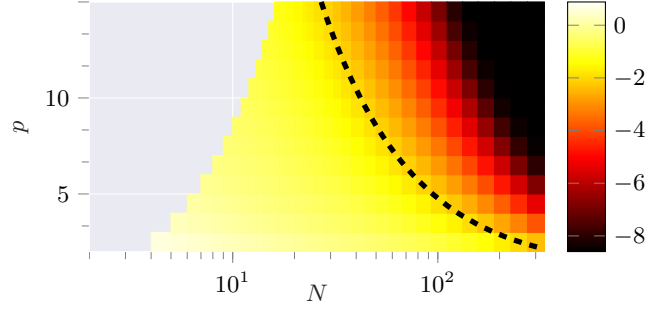[1]The simulations were done on two Intel(R) Xeon(R) CPU E5-2670 v3 @ 2.30GHz processors.

Fig. 2. Evolution of $\log_{10} \sqrt{(\textbf{HCRWWB})_{M,M}}$ for several parameters p and N, $T = 100$, $\alpha_0 = 0.1$ and $\alpha_1 = 0.3$. The dashed line corresponds to the region where $\sqrt{(\textbf{HCRWWB})_{M,M}} = 10^{-2}$.

Then a practical application of tuning has illustrated the usability of the bound for extensive analysis in a design context.

A. PROOF OF PROPOSITION 4.4.

Considering the complex nature of the matrices, we use the Wirtinger derivative [21] in our derivations. The result is then obtained using the complex to real mapping. Then we remark that since the derivative in eq. (14) does not involve the variable of integration and the fact that the right-hand side of the integral is made of scalar terms, it is possible to compute the following expression:

$$\boldsymbol{\Phi}_j(s) = \int_{\mathbb{S}_{\mathbb{H}}^p} \left.\frac{\partial \ln f_{\mathbf{X}_t;\boldsymbol{\Sigma}}(\mathbf{X}_t;\boldsymbol{\Sigma})}{\partial \boldsymbol{\Sigma}}\right|_{\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_j} \times$$
$$f_{\mathbf{X}_t;\boldsymbol{\Sigma}_0}^s(\mathbf{X}_t;\boldsymbol{\Sigma}_0) f_{\mathbf{X}_t;\boldsymbol{\Sigma}_1}^{1-s}(\mathbf{X}_t;\boldsymbol{\Sigma}_1)\, \mathrm{d}\mathbf{X}_t .$$

and we obtain our result by taking $\phi_{\sigma_j,\ell}(s) = ([\text{vech}\,(\boldsymbol{\Phi}_j(s))]_{\mathbb{CR}})_\ell$.

The Wirtinger derivative leads us to:

$$\frac{\partial \ln f_{\mathbf{X}_t;\boldsymbol{\Sigma}'}(\mathbf{X}_t;\boldsymbol{\Sigma}')}{\partial \boldsymbol{\Sigma}'} = -N\boldsymbol{\Sigma}'^{-1} + \left(\boldsymbol{\Sigma}'^{-1}\mathbf{X}_t\boldsymbol{\Sigma}'^{-1}\right)^{\mathrm{H}}.$$

Thus,

$$\boldsymbol{\Phi}_j(s) = -N\boldsymbol{\Sigma}_j^{-1}\rho(s) + \boldsymbol{\Sigma}_j^{-1}\left(\int_{\mathbb{S}_{\mathbb{H}}^p} \mathbf{X}_t \frac{|\mathbf{X}_t|^{N-p}}{\Gamma_p(N)|\boldsymbol{\Sigma}_0|^{sN}|\boldsymbol{\Sigma}_1|^{(1-s)N}}\right.$$
$$\times \text{etr}\left.\left(-(s\boldsymbol{\Sigma}_0^{-1} + (1-s)\boldsymbol{\Sigma}_1^{-1})\mathbf{X}_t\right) \mathrm{d}\mathbf{X}_t\right) \boldsymbol{\Sigma}_j^{-1},$$

$$= -N\boldsymbol{\Sigma}_j^{-1}\rho(s) + \boldsymbol{\Sigma}_j^{-1}\rho(s) \times$$
$$\left(\int_{\mathbb{S}_{\mathbb{H}}^p} \mathbf{X}_t \frac{|\mathbf{X}_t|^{N-p}}{\Gamma_p(N)\left|\left(s\boldsymbol{\Sigma}_0^{-1} + (1-s)\boldsymbol{\Sigma}_1^{-1}\right)^{-1}\right|^N} \times \right.$$
$$\text{etr}\left.\left(-(s\boldsymbol{\Sigma}_0^{-1} + (1-s)\boldsymbol{\Sigma}_1^{-1})\mathbf{X}_t\right) \mathrm{d}\mathbf{X}_t\right) \boldsymbol{\Sigma}_j^{-1},$$

$$= -N\boldsymbol{\Sigma}_j^{-1}\rho(s) + \boldsymbol{\Sigma}_j^{-1}\rho(s)\mathbb{E}_{\mathbf{X};\mathbf{A}}\{\mathbf{X}_t\}\boldsymbol{\Sigma}_j^{-1},$$

where $\mathbf{A} = \left(s\boldsymbol{\Sigma}_0^{-1} + (1-s)\boldsymbol{\Sigma}_1^{-1}\right)^{-1}$. Finally using the expectation of a Wishart distribution, we obtain the result.

B. REFERENCES

[1] K. Conradsen, A. A. Nielsen, J. Schou, and H. Skriver, "A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 1, pp. 4–19, Jan 2003.

[2] L. M. Novak, "Change detection for multi-polarization multi-pass sar," in *Algorithms for Synthetic Aperture Radar Imagery XII*. International Society for Optics and Photonics, 2005, vol. 5808, pp. 234–247.

[3] K. Conradsen, A. A. Nielsen, and H. Skriver, "Determining the points of change in time series of polarimetric SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 3007–3024, May 2016.

[4] A. Mian, J.-P. Ovarlez, G. Ginolhac, and A. M. Atto, "Robust detection and estimation of Change-Points in a time series of multivariate images," in *2018 26th European Signal Processing Conference (EUSIPCO) (EUSIPCO 2018)*, Roma, Italy, Sept. 2018.

[5] A. D. C. Nascimento, A. C. Frery, and R. J. Cintra, "Detecting changes in fully polarimetric SAR imagery with statistical information theory," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–13, 2018.

[6] A. Lopes and F. Sery, "Optimal speckle reduction for the product model in multilook polarimetric SAR imagery and the Wishart distribution," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 3, pp. 632–647, May 1997.

[7] J.-S. Lee, M. R. Grunes, T. L. Ainsworth, L.-J. Du, D. L. Schuler, and S. R. Cloude, "Unsupervised classification using polarimetric decomposition and the complex Wishart classifier," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 5, pp. 2249–2258, Sept 1999.

[8] M. Dabboor, M. J. Collins, V. Karathanassi, and A. Braun, "An unsupervised classification approach for polarimetric SAR data based on the Chernoff distance for complex Wishart distribution," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 7, pp. 4200–4213, July 2013.

[9] A. Mian, J.-P. Ovarlez, G. Ginolhac, and A. Atto, "Multivariate change detection on high resolution monovariate sar image using linear time-frequency analysis," in *2017 25th European Signal Processing Conference (EUSIPCO)*, Aug 2017, pp. 1942–1946.

[10] S. Fotopoulos and V. Jandhyala, "Maximum likelihood estimation of a change-point for exponentially distributed random variables," *Statistics & Probability Letters*, vol. 51, no. 4, pp. 423 – 429, 2001.

[11] S. B. Fotopoulos, V. K. Jandhyala, and E. Khapalova, "Exact asymptotic distribution of change-point MLE for change in the mean of gaussian sequences," *Ann. Appl. Stat.*, vol. 4, no. 2, pp. 1081–1104, 06 2010.

[12] 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, March 1988.

[13] L. Bacharach, M. N. El Korso, A. Renaux, and J.-Y. Tourneret, "A hybrid lower bound for parameter estimation of signals with multiple change-points," *IEEE Transactions on Signal Processing, submitted*, 2018.

[14] A. Ferrari and J.-Y. Tourneret, "Barankin lower bound for change-points in independent sequences," in *IEEE Workshop on Statistical Signal Processing, 2003*, Sept 2003, pp. 557–560.

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

[16] L. Bacharach, A. Renaux, M. N. El Korso, and E. Chaumette, "Weiss–Weinstein bound on multiple change-points estimation," *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2686–2700, May 2017.

[17] S. M. Kay, "Fundamentals of statistical signal processing: estimation theory.," 2010.

[18] C. Ren, J. Galy, E. Chaumette, P. Larzabal, and A. Renaux, "Hybrid Barankin–Weiss–Weinstein bounds," *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 2064–2068, Nov 2015.

[19] A. C. Frery, A. D. C. Nascimento, and R. J. Cintra, "Analytic expressions for stochastic distances between relaxed complex Wishart distributions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 2, pp. 1213–1226, Feb 2014.

[20] A. Yeredor, "The joint MAP-ML criterion and its relation to ML and to extended least-squares," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3484–3492, Dec 2000.

[21] W. Wirtinger, "Zur formalen theorie der funktionen von mehr komplexen veränderlichen," *Mathematische Annalen*, vol. 97, pp. 357–376, 1927.